Performance Evaluation of LTE eSRVCC with Limited Access Transfers

Yi-Bing Lin, *Fellow, IEEE*, Ren-Huang Liou
Department of Computer Science
National Chiao Tung University
{liny, rhliou}@cs.nctu.edu.tw

Yingrong Coral Sung
Center for Information and Communications Technology
National Chiao Tung University
ycsung@cs.nctu.edu.tw

Pai-Chun Cheng
Department of Computer Science
National Chiao Tung University
pccheng@cs.nctu.edu.tw

Abstract

Long Term Evolution (LTE) evolved from *Universal Mobile Telecommunications System* (UMTS), which utilizes *IP Multimedia Core Network Subsystem* (IMS) to provide voice service. In most existing commercial operations, the LTE networks provide zonal coverage as compared with UMTS networks which provide full service coverage. A *User Equipment* (UE) can initiate or receive an IMS call in either LTE or UMTS. The UE uses LTE whenever it is available. If LTE is out of service, then the UE is transferred to UMTS. To support access transfer between LTE and UMTS during an IMS call, *3rd Generation Partnership Project* (3GPP) proposed *Enhanced Single Radio Voice Call Continuity* (eSRVCC). If the UE frequently moves back and forth between LTE and UMTS during an IMS call, it may incur large access transfer traffic. To resolve this issue, we propose the limited access transfer algorithm that limits the number of access transfers in an eSRVCC call (referred to as the transfer limit) to reduce the transfer traffic. An analytic model is proposed to investigate the

performance of the limited access transfer algorithm. Our study indicates that the selection of the transfer limit is not a trivial issue, and an appropriate transfer limit effectively reduces the access transfer traffic to enhance the LTE call control performance.

Index Terms: Access Transfer, Enhanced Single Radio Voice Call Continuity (eSRVCC), IP Multimedia Core Network Subsystem (IMS), Long Term Evolution (LTE)

1 Introduction

Long Term Evolution (LTE) [1] is an all-IP mobile broadband communication standard evolved from *Universal Mobile Telecommunications System* (UMTS) [2], which utilizes *IP Multimedia Core Network Subsystem* (IMS) to provide voice and multimedia services [3, 4]. In most existing commercial operations, the UMTS networks provide full service coverage. On the other hand, in the current deployment, the LTE networks provide zonal coverage as compared with UMTS. Therefore, when the LTE network is available, a *User Equipment* (UE) attaches to the LTE network to make an IMS call. When the LTE network is not available during an IMS call, the UE switches to the UMTS network to continue the call. This process is called *access transfer*. *3rd Generation Partnership Project* (3GPP) proposes *Enhanced Single Radio Voice Call Continuity* (eSRVCC) to transfer an IMS call between LTE and UMTS [5]. We note that IMS is an excellent core network infrastructure for integrating heterogeneous networks; for example, UMTS and WiFi integration [6]. Therefore, it is natural that 3GPP selects IMS for LTE and UMTS integration on access transfer.

Figure 1 illustrates a simplified LTE and UMTS network architecture for eSRVCC. In this architecture, *User Equipment* 1 (UE 1; Figure 1 (1)) accesses the LTE service through an *Evolved Node B* (eNB; Figure 1 (2)). UE 1 can also access the UMTS service through a Node

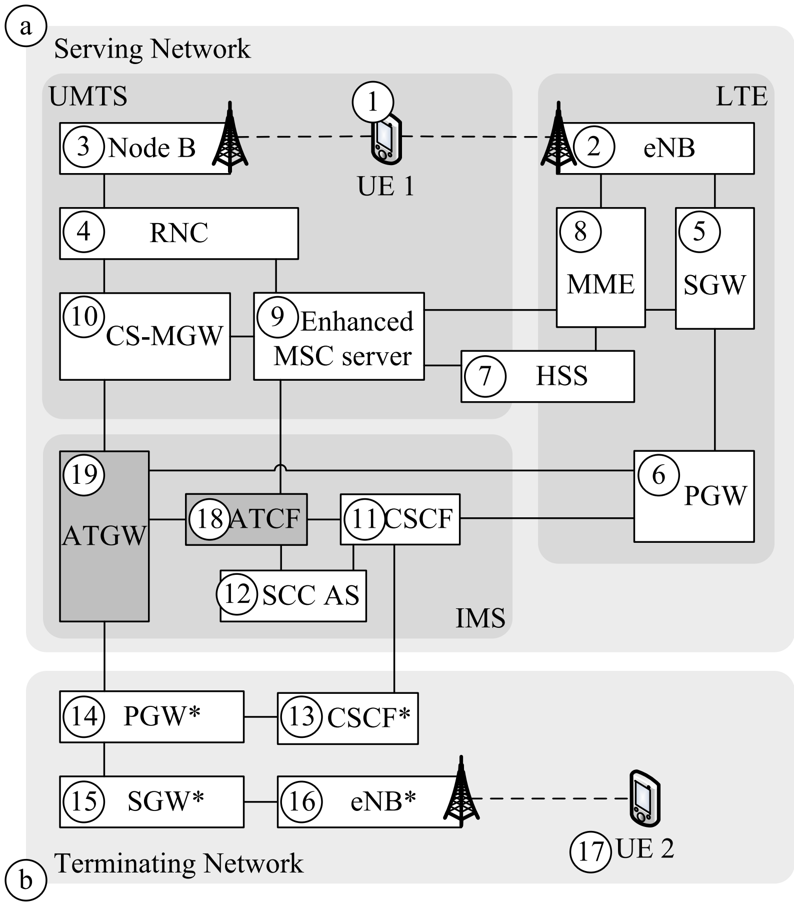


Figure 1: A Simplified LTE and UMTS Network Architecture for eSRVCC

B (Figure 1 (3)) and the *Radio Network Controller* (RNC; Figure 1 (4)). In the LTE core network, the *Serving Gateway* (SGW; Figure 1 (5)) is responsible for packet routing. The *Packet Data Network Gateway* (PGW; Figure 1 (6)) connects the SGW with external packet data networks (e.g., the Terminating Network in Figure 1 (b)). The *Home Subscriber Server* (HSS; Figure 1 (7)) is the master database containing all user subscription and location information. In particular, the HSS indicates if a user is offered the eSRVCC service. The *Mobility Management Entity* (MME; Figure 1 (8)) is responsible for mobility management through interaction with the HSS.

In the UMTS core network, the *Enhanced Mobile Switching Center server* (Enhanced MSC server; Figure 1 (9)) is an MSC server which is responsible for call control and is enhanced for eSRVCC support. The Enhanced MSC server communicates with the MME to allocate radio and gateway resources when performing the access transfer procedure. The *Circuit-Switched Media Gateway* (CS-MGW; Figure 1 (10)) is responsible for call and session connections. Specifically, it supports media conversion, bearer control, and payload processing. The CS-MGW performs media conversion between CS bearer channels and *Packet-Switched* (PS) media streams to support voice continuity of an IMS call. The IMS consists of various control functions and application servers to offer IP multimedia services through *Session Initiation Protocol* (SIP) [7,8]. The *Call Session Control Function* (CSCF; Figure 1 (11)) provides session management, service control, SIP proxy, and registrar functionalities in the IMS network, which consists of three functions. The *Interrogating-CSCF* (I-CSCF) is the external contact point that hides the internal configuration of an IMS network. The *Proxy-CSCF* (P-CSCF) contains limited address translation functions to forward SIP requests initiated by or destined to a UE. The *Serving-CSCF* (S-CSCF) supports the signaling interactions with the UE for registration, call setup, and supplementary service control. If a user has subscribed to the eSRVCC service, the S-CSCF always forwards SIP requests of a user to the *Service Centralization and Continuity Application Server* (SCC AS; Figure 1 (12)). The eSRVCC service introduces two new elements to the serving IMS network: *Access Transfer Control Function* (ATCF; Figure 1 (18)) and *Access Transfer Gateway* (ATGW; Figure 1 (19)). These two elements serve as the anchor points for control signals and media bearers, respectively. The ATGW partitions the media path of an IMS call between UE 1 and UE 2 (Figure 1 (17)) into two legs; i.e., UE 1's access leg and UE 2's access leg.

When UE 1 attaches to the LTE network and makes an IMS call to UE 2, UE 1's access leg is routed through (1)-(2)-(5)-(6)-(19) and UE 2's access leg is routed through (19)-(14)-(15)-(16)-(17). When UE 1 is switched from LTE to UMTS during an IMS call, the media path must be updated. In the eSRVCC *access transfer procedure*, UE 1's access leg is switched to (1)-(3)-(4)-(10)-(19). Since the voice bearer is always anchored at the ATGW, an access transfer does not affect UE 2's access leg; i.e., this leg remains as (19)-(14)-(15)-(16)-(17).

As mentioned before, the coverage of the LTE network is a subset of that of the UMTS network. If a UE frequently moves back and forth between LTE and UMTS during an IMS call, it may incur large access transfer traffic. To resolve this issue, we propose the limited access transfer algorithm that limits the maximum number of access transfers in a call (referred to as the *transfer limit*). This paper studies how to appropriately select the transfer limit to reduce the transfer traffic. The paper is organized as follows. Section 2 describes the eSRVCC access transfer procedure, and proposes the limited access transfer algorithm to reduce the transfer traffic. Section 3 proposes an analytic model to study the performance of an eSRVCC call with the transfer limit. Section 4 investigates the performance of eSRVCC by numerical examples, and conclusions are given in Section 5.

2 Limited eSRVCC Access Transfer Algorithm

The eSRVCC call setup and call release procedures are defined in 3GPP [3, 9], and the details are not given in this paper. This section describes the eSRVCC access transfer procedure. Then we propose the limited access transfer algorithm to reduce the transfer traffic caused by frequent access transfers.

2.1 eSRVCC Access Transfer

Suppose that UE 1 resides in LTE, and an IMS call is established between UE 1 and UE 2, which is anchored at the ATGW (the media path is (1)-(2)-(5)-(6)-(19)-(14)-(15)-(16)-(17) in Figure 1). Suppose that UE 1 moves from LTE to UMTS, and the serving eNB decides to transfer the call. Figure 2 illustrates the message flow for eSRVCC access transfer from LTE to UMTS with the following steps [10]:

Step 1. The eNB sends the Handover Required message to the MME. The MME initiates

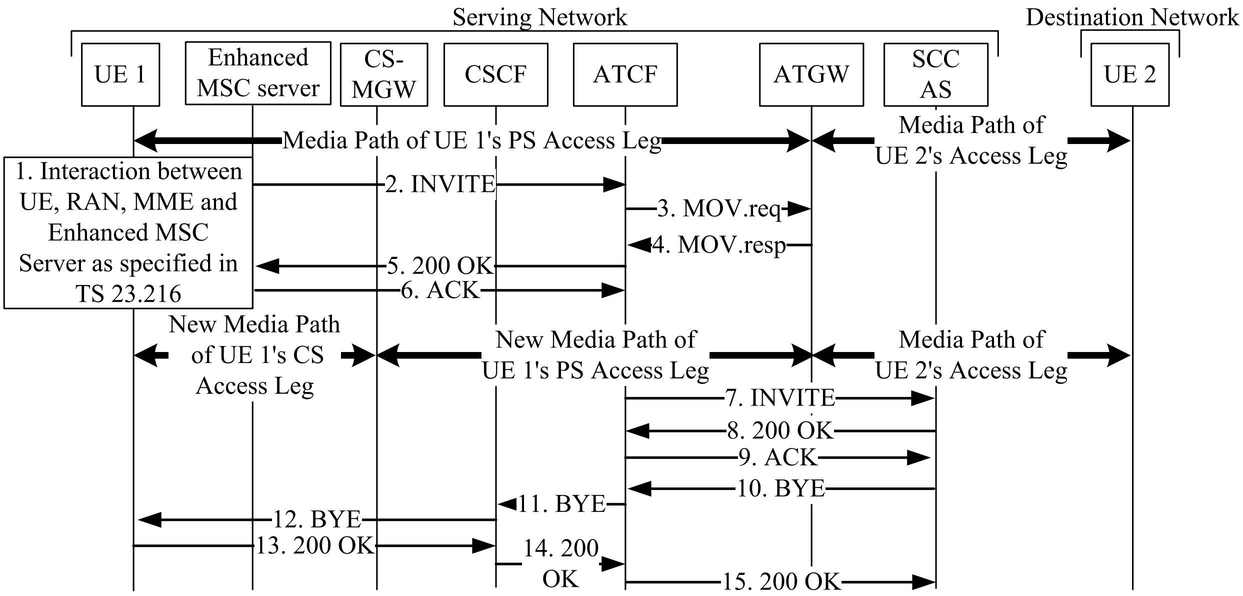


Figure 2: eSRVCC Access Transfer from LTE to UMTS

the PS-CS (LTE to UMTS) handover procedure by sending a PS to CS Request message to the Enhanced MSC server. The Enhanced MSC server allocates necessary radio resources and the CS-MGW resources for UE 1's access transfer. If the resources are allocated successfully, the Enhanced MSC server sends a PS to CS Response message back to the MME. Then the MME sends a Handover Command message to UE 1 to request UE 1 to handover from LTE to UMTS.

Step 2. Based on the *Session Transfer Number - Single Radio* (STN-SR; the ATCF address) contained in the PS to CS Request message, the Enhanced MSC server sends a SIP INVITE message to the ATCF. In this message, the *Session Description Protocol* (SDP) [11] is provided by the CS-MGW.

Steps 3 and 4. The ATCF correlates the SIP INVITE request to UE 1's access leg and UE 2's access leg by using the *Correlation Mobile Station International Integrated Services Digital Network number* (C-MSISDN) specified in the SIP INVITE message. Based on the SDP of the SIP INVITE message, the ATCF exchanges H.248 MOV.req and MOV.resp message pair with the ATGW to switch UE 1's access leg of the call from the PS domain (LTE) to the CS domain (UMTS).

Steps 5 and 6. After UE 1's access leg of the call was successfully modified, the ATCF sends a SIP 200 OK message to the Enhanced MSC server. Then the Enhanced MSC server sends a SIP ACK back to the ATCF. When UE 1 attaches to UMTS, the eSRVCC access transfer is executed without changing UE 2's access leg of the call.

Steps 7-9. After the access transfer, the ATCF sends a SIP INVITE message to inform the SCC AS that the access transfer has taken place.

Steps 10-15. The SCC AS sends a SIP BYE message to UE 1 to terminate the old access leg of the call.

After the eSRVCC access transfer, the media path is (1)-(3)-(4)-(10)-(19)-(14)-(15)-(16)-(17) in Figure 1. The scenario where UE 1 moves from UMTS to LTE is similar to that from LTE to UMTS, and the reader is referred to [9].

2.2 Limited Access Transfer Algorithm

To avoid frequent access transfers in a call, we propose the limited access transfer algorithm as follows. The access transfers of a UE can be classified into two types:

- **Type U-L:** access transfer from UMTS to LTE
- **Type L-U:** access transfer from LTE to UMTS

Define N as the U-L transfer limit; i.e., the maximum number of the U-L transfers that are allowed to be performed in a call. The limited access transfer algorithm reduces the transfer traffic caused by frequent access transfers with the following steps:

Step 1. When an eSRVCC call is initiated in UMTS or is transferred to UMTS for the first time, the Enhanced MSC server initiates the U-L transfer counter $N^* = 0$. This counter is kept in the Enhanced MSC server.

Step 2. If (a) the UE is in UMTS, (b) the radio quality of LTE is better than that of UMTS, and (c) $N^* < N$, then the network performs eSRVCC access transfer from UMTS to LTE and increments N^* by one. Otherwise, if the UE is in LTE and the radio quality of UMTS is better than that of LTE, then eSRVCC access transfer procedure is always performed to transfer the call from LTE to UMTS.

3 Analytic Model

This section proposes an analytic model for the limited access transfer algorithm. Figure 3 illustrates a timing diagram for call arrivals and access transfers. As mentioned in Section 2.2, the access transfers are classified into two types: the U-L transfers (that occur at t_1, t_4, t_6, and t_9 in Figure 3) and the L-U transfers (that occur at t_3, t_5, and t_7 in Figure 3). The transfer limit N is the maximum number of the U-L transfers that are allowed to be performed in a call. At the end of an eSRVCC call, the value of the counter N^* is the number of the U-L and L-U access transfers performed in a call. In Figure 3, $N^* = 5$ when the call is released. It is clear that $N^* \leq 2N + 1$ is enforced. Let $t_{L,i} = t_3 - t_1$ (also $t_5 - t_4$ and $t_7 - t_6$) and $t_{U,i} = t_4 - t_3$ (also $t_6 - t_5$ and $t_9 - t_7$) be the LTE and the UMTS residence times between the ith and the $(i+1)$th U-L transfers in a call, respectively. We assume that $t_{L,i}$ and $t_{U,i}$ are independent and identically distributed random variables with the density functions $f_L(\cdot)$ and $f_U(\cdot)$, the means $1/\eta_L$ and $1/\eta_U$, the variances V_L and V_U, and the Laplace transforms $f_L^*(s)$ and $f_U^*(s)$, respectively. Let the incoming calls to the UE be a Poisson process and the call holding time $t_c = t_8 - t_2$ be an exponential random variable with the mean $1/\mu$. Although the realistic call holding time distribution may not be the exponential distribution,

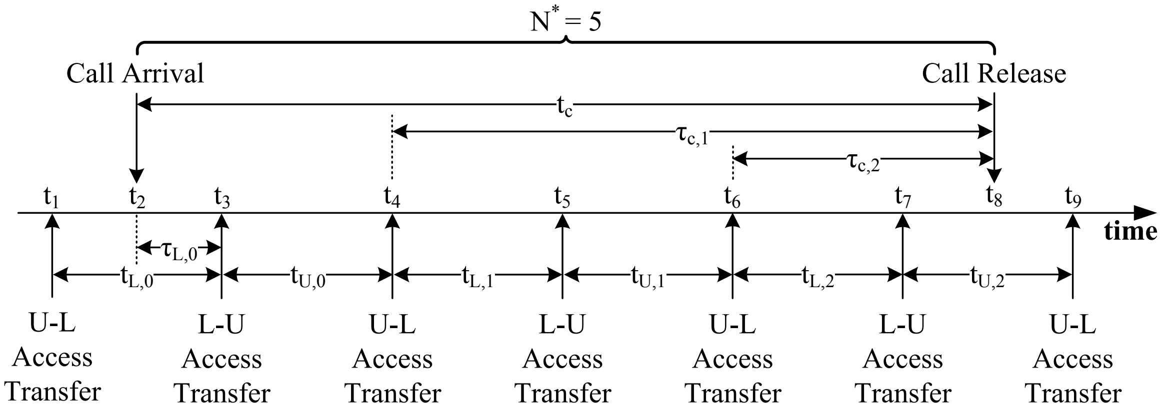


Figure 3: Timing Diagram for Call Arrivals and Access Transfers

the exponential distribution does provide the mean value analysis for a primary study on the trends of the call holding time impact. Also, this paper uses the analytic results based on the exponential assumption to validate the simulation experiments (to be elaborated in Appendix A), and the validated simulation model can then be extended to study realistic traffic distributions measured from the commercial operation. We define an *LTE-initiated call* as a call arriving at LTE (e.g., the call arriving at t_2 in Figure 3), and a *UMTS-initiated call* as a call arriving at UMTS. Since the call arrivals are a Poisson process, time point t_2 is a random observer of the period $[t_1, t_3]$. The interval $\tau_{L,0} = t_3 - t_2$ is the residual life of $t_{L,0}$ with the Laplace transform $r_L^*(s)$. Similarly, if a call arrives at UMTS, the residual life of $t_{U,0}$ is $\tau_{U,0}$ with the Laplace transform $r_U^*(s)$. From [12], $r_L^*(s)$ and $r_U^*(s)$ can be expressed as

$$r_L^*(s) = \left(\frac{\eta_L}{s}\right)[1 - f_L^*(s)] \text{ and } r_U^*(s) = \left(\frac{\eta_U}{s}\right)[1 - f_U^*(s)] \tag{1}$$

Let $\tau_{c,i} = t_8 - t_4$ (also $t_8 - t_6$) be the interval between when the ith U-L transfer occurs and when the call is released. From the memoryless property of the exponential distribution, $\tau_{c,i}$ has the same exponential distribution as t_c does. The variables used to describe the time distributions are summarized below.

- $1/\mu = E[t_c]$: the mean call holding time t_c

- $1/\eta_L = E[t_{L,i}]$: the mean LTE residence time $t_{L,i}$
- $1/\eta_U = E[t_{U,i}]$: the mean UMTS residence time $t_{U,i}$
- V_L: the variance for the $t_{L,i}$ distribution
- V_U: the variance for the $t_{U,i}$ distribution
- $f_L(\cdot)$: the density function for the $t_{L,i}$ distribution
- $f_U(\cdot)$: the density function for the $t_{U,i}$ distribution
- $f_L^*(s)$: the Laplace transform for the $t_{L,i}$ distribution
- $f_U^*(s)$: the Laplace transform for the $t_{U,i}$ distribution
- $r_L^*(s)$: the Laplace transform for the $\tau_{L,0}$ distribution
- $r_U^*(s)$: the Laplace transform for the $\tau_{U,0}$ distribution

We investigate the performance of the limited access transfer algorithm by considering the following two output measures:

- $E[N^*|N = n]$: the expected number of the U-L and the L-U access transfers performed in a call under the condition that the transfer limit $N = n$ (i.e., n is the maximum number of the U-L transfers that are allowed to be performed in a call). In this output measure, N^* is the counter value (Step 1 in Section 2.2) when the call is released. A smaller $E[N^*|N = n]$ value means fewer access transfers during a call, i.e., lower network cost.
- θ: the percentage of the time that the UE resides in LTE during a call. A larger θ value means higher LTE utilization. Note that a mobile operator attempts to utilize LTE as much as possible, and anticipates a large θ value for a call.

Our goal is to increase θ at reasonable low $E[N^*|N=n]$ cost. We note that $E[N^*|N=n]$ and θ are conflicting goals. Therefore, we need to select an appropriate N value to balance against these two output measures. To derive $E[N^*|N=n]$ and θ, we need extra parameters described as follows:

- P_L: the probability that a call arrives at LTE
- P_U: the probability that a call arrives at UMTS
- $E[N_{L,U-L}|N=n]$: the expected number of the U-L transfers performed in an LTE-initiated call under the condition that $N=n$
- $E[N_{L,L-U}|N=n]$: the expected number of the L-U transfers performed in an LTE-initiated call under the condition that $N=n$
- $E[N_{U,U-L}|N=n]$: the expected number of the U-L transfers performed in a UMTS-initiated call under the condition that $N=n$
- $E[N_{U,L-U}|N=n]$: the expected number of the L-U transfers performed in a UMTS-initiated call under the condition that $N=n$

By using the above parameters, the expected number $E[N^*|N=n]$ can be expressed as

$$\begin{aligned} E[N^*|N=n] &= P_L\{E[N_{L,U-L}|N=n]+E[N_{L,L-U}|N=n]\} \\ &\quad +P_U\{E[N_{U,U-L}|N=n]+E[N_{U,L-U}|N=n]\} \end{aligned} \tag{2}$$

In (2), a call is initiated either in LTE with probability P_L or in UMTS with probability P_U. An LTE-initiated call has the expected number $E[N_{L,U-L}|N=n]$ of U-L transfers and the expected number $E[N_{L,L-U}|N=n]$ of L-U transfers. On the other hand, a UMTS-initiated call has the expected number $E[N_{U,U-L}|N=n]+E[N_{U,L-U}|N=n]$ of transfers.

Since a sequence of $t_{L,i}$ and $t_{U,i}$ forms an alternating renewal process [12], P_L and P_U in (2) can be computed as

$$P_L = \frac{E[t_{L,i}]}{E[t_{L,i}] + E[t_{U,i}]} = \frac{\eta_U}{\eta_L + \eta_U} \text{ and } P_U = \frac{E[t_{U,i}]}{E[t_{L,i}] + E[t_{U,i}]} = \frac{\eta_L}{\eta_L + \eta_U} \quad (3)$$

The expected number $E[N_{L,U-L}|N = n]$ is derived as follows. Let random variable $T_{L,i}$ be the interval between when a call arrives and when the $(i+1)$th U-L transfer occurs. Then

$$T_{L,i} = \begin{cases} \tau_{L,0} + t_{U,0}, & \text{for } i = 0 \\ \tau_{L,0} + t_{U,0} + \sum_{j=1}^{i} (t_{L,i} + t_{U,i}), & \text{for } i \geq 1 \end{cases} \quad (4)$$

with the density function $f_{T,i}(\cdot)$ and the Laplace transform

$$f^*_{T,i}(s) = r^*_L(s) f^*_U(s) \left[f^*_L(s) f^*_U(s) \right]^i \quad (5)$$

Let $\Pr[N_{L,U-L} \leq i]$ be the probability that the number of the U-L transfers performed in the LTE-initiated call is less than or equal to i. From (4) and (5), it is clear that $\Pr[N_{L,U-L} \leq i] = \Pr[T_{L,i} > t_c]$, which is derived as

$$\begin{aligned} \Pr[T_{L,i} > t_c] &= \int_{T_{L,i}=0}^{\infty} f_{T,i}(T_{L,i}) \int_{t_c=0}^{T_{L,i}} \mu e^{-\mu t_c} dt_c dT_{L,i} \\ &= \int_{T_{L,i}=0}^{\infty} f_{T,i}(T_{L,i})(1 - e^{-\mu T_{L,i}}) dT_{L,i} \\ &= 1 - r^*_L(\mu) f^*_U(\mu) \left[f^*_L(\mu) f^*_U(\mu) \right]^i \end{aligned} \quad (6)$$

If $T_{L,0} > t_c$, the UE does not perform any U-L transfer, and $N_{L,U-L} = 0$. From (6),

$$\Pr[T_{L,0} > t_c] = \Pr[N_{L,U-L} = 0] = 1 - r^*_L(\mu) f^*_U(\mu) \quad (7)$$

For $i \geq 1$, from (6), $\Pr[N_{L,U-L} = i]$ is derived as

$$\begin{aligned} \Pr[N_{L,U-L} = i] &= \Pr[N_{L,U-L} \leq i] - \Pr[N_{L,U-L} \leq i-1] \\ &= r_L^*(\mu) f_U^*(\mu) \left[f_L^*(\mu) f_U^*(\mu) \right]^{i-1} \left[1 - f_L^*(\mu) f_U^*(\mu) \right] \end{aligned} \tag{8}$$

From (1), (7), and (8),

$$\begin{aligned} E[N_{L,U-L}|N = n] &= \sum_{i=0}^{n-1} i \Pr[N_{L,U-L} = i] + n \Pr[t_c > T_{L,n-1}] \\ &= \sum_{i=0}^{n-1} i \Pr[N_{L,U-L} = i] + n \left\{ 1 - \sum_{i=0}^{n-1} \Pr[N_{L,U-L} = i] \right\} \\ &= \left(\frac{\eta_L}{\mu} \right) [1 - f_L^*(\mu)] f_U^*(\mu) \left\{ \frac{1 - [f_L^*(\mu) f_U^*(\mu)]^n}{1 - f_L^*(\mu) f_U^*(\mu)} \right\} \end{aligned} \tag{9}$$

Since $N_{L,U-L} \leq n$ is constrained in the limited access transfer algorithm, exactly n U-L transfers are performed if $t_c > T_{L,n-1}$. Therefore, in (9), $\Pr[t_c > T_{L,n-1}]$ is the probability that a UE performs exactly n U-L transfers in a call. Similar to the derivation for $E[N_{L,U-L}|N = n]$, we have

$$E[N_{L,L-U}|N = n] = \left(\frac{\eta_L}{\mu} \right) [1 - f_L^*(\mu)] \left\{ \frac{1 - [f_L^*(\mu) f_U^*(\mu)]^{n+1}}{1 - f_L^*(\mu) f_U^*(\mu)} \right\} \tag{10}$$

$$E[N_{U,U-L}|N = n] = \left(\frac{\eta_U}{\mu} \right) [1 - f_U^*(\mu)] \left\{ \frac{1 - [f_L^*(\mu) f_U^*(\mu)]^n}{1 - f_L^*(\mu) f_U^*(\mu)} \right\} \tag{11}$$

and

$$E[N_{U,L-U}|N = n] = \left(\frac{\eta_U}{\mu} \right) [1 - f_U^*(\mu)] f_L^*(\mu) \left\{ \frac{1 - [f_L^*(\mu) f_U^*(\mu)]^n}{1 - f_L^*(\mu) f_U^*(\mu)} \right\} \tag{12}$$

From (3) and (9)-(12), (2) is re-written as

$$E[N^*|N=n] \;=\; \frac{\eta_L\eta_U\left\{2-[f_L^*(\mu)+1]\,[f_L^*(\mu)f_U^*(\mu)]^n\right\}}{(\eta_L+\eta_U)\mu} \tag{13}$$

For $n\rightarrow\infty$ (i.e., there is no constraint on the number of the U-L transfers), $E[N^*|N\rightarrow\infty]$ is derived as follows. In (13), $\lim_{n\rightarrow\infty}[f_L^*(\mu)f_U^*(\mu)]^n=0$ because $f_L^*(\mu)<1$ and $f_U^*(\mu)<1$. Therefore, $E[N^*|N\rightarrow\infty]$ is computed as

$$\begin{aligned}\lim_{n\rightarrow\infty}E[N^*|N=n] &= \left[\frac{\eta_L\eta_U}{(\eta_L+\eta_U)\mu}\right]\left\{2-[f_L^*(\mu)+1]\lim_{n\rightarrow\infty}[f_L^*(\mu)f_U^*(\mu)]^n\right\}\\ &= \frac{2\eta_L\eta_U}{(\eta_L+\eta_U)\mu}\end{aligned} \tag{14}$$

Intuitively, from the alternating renewal process, the rate of access transfers during a call is $2/(E[t_{L,i}]+E[t_{U,i}])$, and therefore $E[N^*|N\rightarrow\infty]$ can be expressed as

$$\lim_{n\rightarrow\infty}E[N^*|N=n] \;=\; \frac{2E[t_c]}{E[t_{L,i}]+E[t_{U,i}]}=\frac{2\eta_L\eta_U}{(\eta_L+\eta_U)\mu} \tag{15}$$

which is the same as (14).

The portion θ is derived as follows. Let $E[T_L^*|N=n]$ be the expected call holding time that the UE resides in LTE under the condition that $N=n$. Then θ can be expressed as

$$\theta=\frac{E[T_L^*|N=n]}{E[t_c]} \tag{16}$$

Similar to (2), $E[T_L^*|N=n]$ is expressed as:

$$E[T_L^*|N=n]=P_LE[T_{L,L}^*|N=n]+P_UE[T_{L,U}^*|N=n] \tag{17}$$

where $E[T_{L,L}^*|N=n]$ and $E[T_{L,U}^*|N=n]$ are the expected times that the UE resides in LTE for an LTE-initiated call and a UMTS-initiated call, respectively. Assuming that the UE performs i U-L transfers in a call, $E[T^*_{L,L}|N = n]$ can be partitioned into three parts:

Part 1 (the LTE call holding time before the 1st U-L transfer): The expected call holding time in $\tau_{L,0}$ is $E[min(t_c, \tau_{L,0})]$. In Figure 3, $[t_2, t_3]$ is the call holding time of this part.

Part 2 (the LTE call holding time between the 1st and the ith U-L transfers): If the UE performs i U-L transfers in a call for $i \geq 2$, then for $0 < j < i$, the expected call holding time in $t_{L,j}$ is $E[t_{L,j}|\tau_{c,j} > t_{L,j} + t_{U,j}]$. Note that there are $i-1$ $t_{L,j}$ intervals in a call, and the expected summation of these intervals is $\sum_{j=1}^{i-1} E[t_{L,j}|\tau_{c,j} > t_{L,j} + t_{U,j}]$. In Figure 3, $i = 2$, $j = 1$, and $[t_4, t_5]$ is the call holding time of this part.

Part 3 (the LTE call holding time after the ith U-L transfer): There are two possibilities:

Case A (for $0 < i < n$): If the UE performs i U-L transfers in a call for $0 < i < n$, the expected call holding time in $t_{L,i}$ is $E[min(\tau_{c,i}, t_{L,i})|\tau_{c,i} < t_{L,i} + t_{U,i}]$. In Figure 3, if $i = 2$ and $n \geq 3$, $[t_6, t_7]$ is the call holding time of this part.

Case B (for $0 < i = n$): If the UE performed i U-L transfers in a call for $0 < i = n$, the expected call holding time in $t_{L,i}$ is $E[min(\tau_{c,i}, t_{L,i})]$. In Figure 3, if $n = i = 2$, $[t_6, t_7]$ is the call holding time of this part.

Based on the above description, for $n = 0$, the UE does not perform any U-L transfer (i.e., we only need to consider Part 1), and therefore $E[T^*_{L,L}|N = 0]$ is expressed as

$$E[T^*_{L,L}|N = 0] = E[min(t_c, \tau_{L,0})] \tag{18}$$

For $n \geq 1$,

$$\begin{aligned} E[T^*_{L,L}|N=n] &= E[min(t_c, \tau_{L,0})] + \sum_{i=1}^{n-1} \Pr[N_{L,U-L}=i] \\ &\quad \times \left\{ \sum_{j=1}^{i-1} E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}] + E[min(\tau_{c,i}, t_{L,i})|\tau_{c,i} < t_{L,i}+t_{U,i}] \right\} \\ &\quad + \Pr[t_c > T_{L,n-1}] \left\{ \sum_{j=1}^{n-1} E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}] + E[min(\tau_{c,i}, t_{L,i})] \right\} \end{aligned} \quad (19)$$

In (19), the LTE call holding time before the 1st U-L transfer is $E[min(t_c, \tau_{L,0})]$ (Part 1). When the UE performs i U-L transfers for $0 < i < n$, the LTE call holding time after the 1st U-L transfer is $\sum_{j=1}^{i-1} E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}]$ (Part 2) plus $E[min(\tau_{c,i}, t_{L,i})|\tau_{c,i} < t_{L,i}+t_{U,i}]$ (Part 3 (A)) with the probability $\Pr[N_{L,U-L}=i]$. Similarly, when the UE performs i U-L transfers for $0 < i = n$, the LTE call holding time after the 1st U-L transfer is $\sum_{j=1}^{n-1} E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}]$ (Part 2) plus $E[min(\tau_{c,i}, t_{L,i})]$ (Part 3 (B)) with the probability $\Pr[t_c > T_{L,n-1}]$ (see the explanation of (9) for $\Pr[t_c > T_{L,n-1}]$).

From (1) and the derivation in [16], $E[min(t_c, \tau_{L,0})]$ and $E[min(\tau_{c,i}, t_{L,i})]$ in (18) and (19) are derived as

$$E[min(t_c, \tau_{L,0})] = \frac{1}{\mu} - \frac{\eta_L[1-f^*_L(\mu)]}{\mu^2} \text{ and } E[min(\tau_{c,i}, t_{L,i})] = \frac{1-f^*_L(\mu)}{\mu} \quad (20)$$

From the conditional expectation, $E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}]$ in (19) is expressed as

$$E[t_{L,j}|\tau_{c,j} > t_{L,j}+t_{U,j}] = \frac{E[t_{L,j}\&(\tau_{c,j} > t_{L,j}+t_{U,j})]}{\Pr[\tau_{c,j} > t_{L,j}+t_{U,j}]} \quad (21)$$

where

$$
\begin{aligned}
E[t_{L,j}\&(\tau_{c,j} > t_{L,j}+t_{U,j})] &= \int\limits_{t_{L,j}=0}^{\infty} t_{L,j} f_L(t_{L,j}) \int\limits_{t_{U,j}=0}^{\infty} f_U(t_{U,j}) \int\limits_{\tau_{c,j}=t_{L,j}+t_{U,j}}^{\infty} \mu e^{-\mu\tau_{c,j}} d\tau_{c,j} dt_{U,j} dt_{L,j} \\
&= -f_U^*(\mu)\left[\frac{df_L^*(s)}{ds}\right]\Big|_{s=\mu} \qquad (22)
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr[\tau_{c,j} > t_{L,j}+t_{U,j}] &= \int\limits_{t_{L,j}=0}^{\infty} f_L(t_{L,j}) \int\limits_{t_{U,j}=0}^{\infty} f_U(t_{U,j}) \int\limits_{\tau_{c,j}=t_{L,j}+t_{U,j}}^{\infty} \mu e^{-\mu\tau_{c,j}} d\tau_{c,j} dt_{U,j} dt_{L,j} \\
&= f_L^*(\mu) f_U^*(\mu) \qquad (23)
\end{aligned}
$$

From (23), $E[min(\tau_{c,i}, t_{L,i})|\tau_{c,i} < t_{L,i}+t_{U,i}]$ in (19) is derived as

$$
\begin{aligned}
E[min(\tau_{c,i}, t_{L,i})|\tau_{c,i} < t_{L,i}+t_{U,i}] &= \frac{E[min(\tau_{c,i}, t_{L,i})\&(\tau_{c,i} < t_{L,i}+t_{U,i})]}{\Pr[\tau_{c,i} < t_{L,i}+t_{U,i}]} \\
&= \frac{E[\tau_{c,i}\&(\tau_{c,i} < t_{L,i})] + E[t_{L,i}\&(t_{L,i} < \tau_{c,i} < t_{L,i}+t_{U,i})]}{1 - f_L^*(\mu) f_U^*(\mu)} \\
&\qquad (24)
\end{aligned}
$$

where

$$
\begin{aligned}
E[\tau_{c,i}\&(\tau_{c,i} < t_{L,i})] &= \int\limits_{t_{L,i}=0}^{\infty} f_L(t_{L,i}) \int\limits_{\tau_{c,i}=0}^{t_{L,i}} \tau_{c,i}\mu e^{-\mu\tau_{c,i}} d\tau_{c,i} dt_{L,i} \\
&= \left(\frac{1}{\mu}\right)[1 - f_L^*(\mu)] + \left[\frac{df_L^*(s)}{ds}\right]\Big|_{s=\mu} \qquad (25)
\end{aligned}
$$

and

$$
\begin{aligned}
E[t_{L,i}\&(t_{L,i} < \tau_{c,i} < t_{L,i}+t_{U,i})] &= \int\limits_{t_{L,i}=0}^{\infty} t_{L,i} f_L(t_{L,i}) \int\limits_{t_{U,i}=0}^{\infty} f_U(t_{U,i}) \int\limits_{\tau_{c,i}=t_{L,i}}^{t_{L,i}+t_{U,i}} \mu e^{-\mu\tau_{c,i}} d\tau_{c,i} dt_{U,i} dt_{L,i} \\
&= [f_U^*(\mu) - 1]\left[\frac{df_L^*(s)}{ds}\right]\Big|_{s=\mu} \qquad (26)
\end{aligned}
$$

From (20)-(26), both (18) and (19) are re-written as

$$E[T^*_{L,L}|N=n] = \frac{1}{\mu} - \left\{\frac{\eta_L[1-f^*_L(\mu)]}{\mu^2}\right\}\left\{1 - \frac{f^*_U(\mu)[1-f^*_L(\mu)]\{1-[f^*_L(\mu)f^*_U(\mu)]^n\}}{[1-f^*_L(\mu)f^*_U(\mu)]}\right\} \quad (27)$$

Similar to the derivation for $E[T^*_{L,L}|N=n]$, $E[T^*_{L,U}|N=n]$ is expressed as

$$E[T^*_{L,U}|N=n] = \frac{\eta_U[1-f^*_L(\mu)][1-f^*_U(\mu)]\{1-[f^*_L(\mu)f^*_U(\mu)]^n\}}{\mu^2[1-f^*_L(\mu)f^*_U(\mu)]} \quad (28)$$

From (3), (17), (27), and (28), (16) is re-written as

$$\theta = \frac{\eta_U}{\eta_L+\eta_U} - \frac{\eta_L\eta_U[1-f^*_L(\mu)][f^*_L(\mu)f^*_U(\mu)]^n}{\mu(\eta_L+\eta_U)} \quad (29)$$

If we do not limit the number of access transfers performed in a call, then $n \to \infty$ and $\lim_{n\to\infty}[f^*_L(\mu)f^*_U(\mu)]^n = 0$. From (29), we have

$$\lim_{n\to\infty}\theta = \frac{\eta_U}{\eta_L+\eta_U} - \left\{\frac{\eta_L\eta_U[1-f^*_L(\mu)]}{\mu(\eta_L+\eta_U)}\right\}\left\{\lim_{n\to\infty}[f^*_L(\mu)f^*_U(\mu)]^n\right\} = \frac{\eta_U}{\eta_L+\eta_U} \quad (30)$$

The intuition behind (30) is the following. If a call can be unlimitedly switched between LTE and UMTS, then the portion of the call in LTE is the probability that the UE stays in LTE; i.e., $\lim_{n\to\infty}\theta = E[t_{L,i}]/(E[t_{L,i}]+E[t_{U,i}]) = \eta_U/(\eta_L+\eta_U)$.

Assume that both $t_{L,i}$ and $t_{U,i}$ are Gamma distributed random variables. Then

$$f^*_L(s) = \left(\frac{1}{V_L\eta_L s+1}\right)^{\frac{1}{V_L\eta_L^2}} \text{ and } f^*_U(s) = \left(\frac{1}{V_U\eta_U s+1}\right)^{\frac{1}{V_U\eta_U^2}} \quad (31)$$

We consider the Gamma distribution because this distribution was widely used to model the UE movement in many studies [13–15].

We obtain $E[N^*|N = n]$ and θ by substituting (31) into (13) and (29). Equations (13), (14), (29), and (30) are used to validate against the discrete event simulation model in Appendix A, which shows that the discrepancies between the analytic and simulation results are within 1%. Therefore, the analytic analysis and simulation results are consistent.

4 Numerical Examples

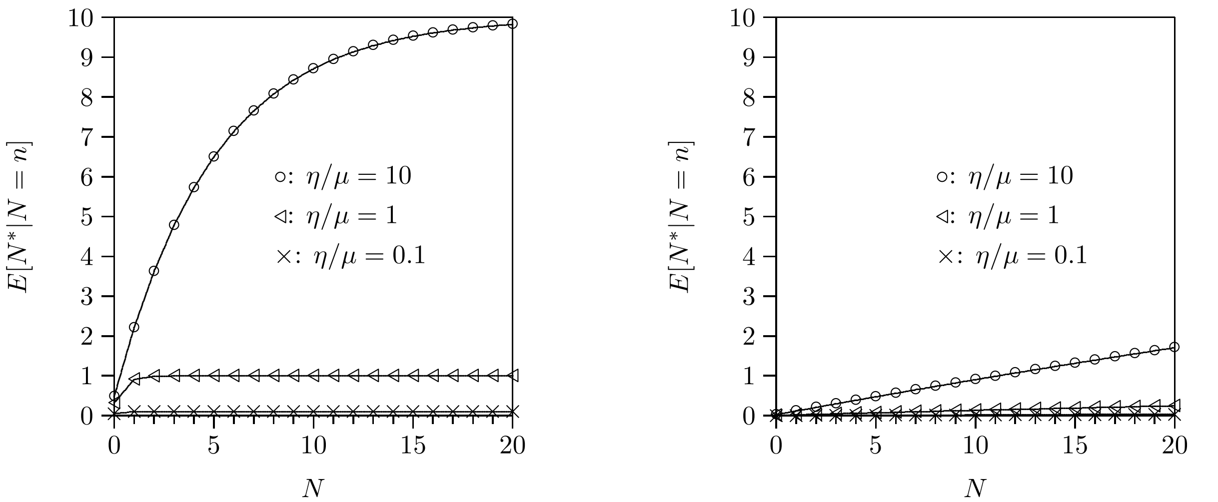
This section investigates the performance of the limited access transfer algorithm with the limit N. To simplify our discussion without loss of generality, we assume that $t_{L,i}$ and $t_{U,i}$ have the same distribution (i.e., $\eta_L = \eta_U = \eta$ and $V_L = V_U = V$). For other η_L/η_U and V_L/V_U values, the results are similar and are omitted. We first point out two facts:

Fact 1. When V increases, there are more short $t_{L,i}$ ($t_{U,i}$) periods and long $t_{L,i}$ ($t_{U,i}$) periods.

Fact 2. When the number of the U-L transfers reaches the limit N, the UE is forced to stay in UMTS after the call is transferred to UMTS.

Effects of η/μ and N on $E[N^*|N = n]$: Figure 4 shows that $E[N^*|N = n]$ increases as η/μ increases. It is trivial that when η/μ increases (i.e., the expected call holding time becomes longer), the access transfers are more likely to occur during a call, and thus larger $E[N^*|N = n]$ is expected. This figure also indicates that a large N incurs large $E[N^*|N = n]$ because the UE is allowed to perform more access transfers during a call. The effect of N becomes more significant when η/μ is large.

Effects of V on $E[N^*|N = n]$: Figure 5 (a) shows that when N is small (e.g., $N \leq 2$), $E[N^*|N = n]$ decreases as V increases. When V increases, due to Fact 1, there are more short and long $t_{L,i}$ ($t_{U,i}$) periods. From the property of the residual life [12], a call is more likely to fall in a long $t_{L,i}$ ($t_{U,i}$), and it is less likely that the UE makes access



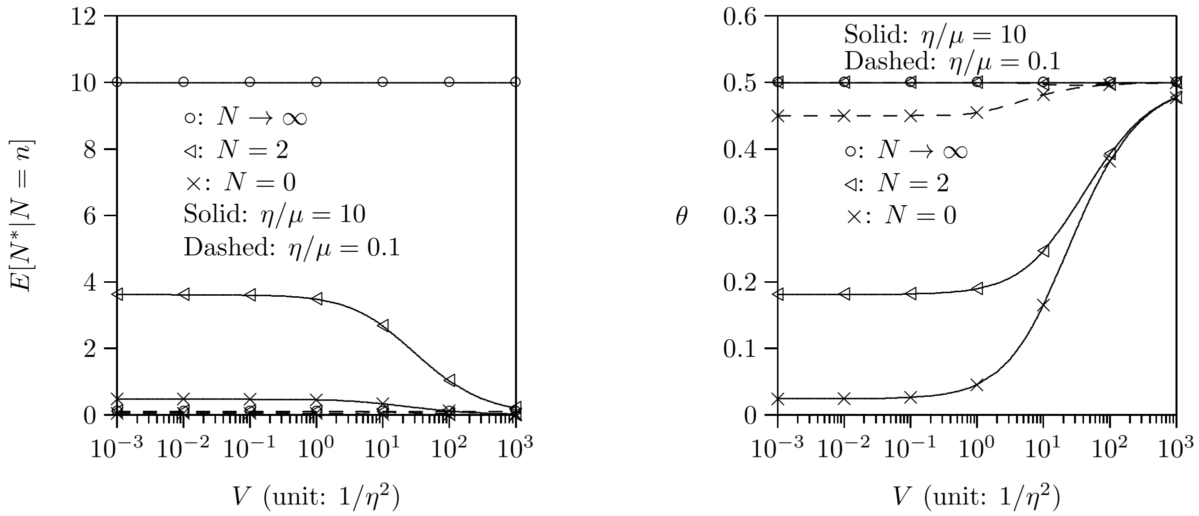
(a) $V = 0.001/\eta^2$ (b) $V = 1000/\eta^2$

Figure 4: The $E[N^*|N = n]$ Curves

transfer before the call is released. Also, for the call arriving at short $t_{L,i}$ ($t_{U,i}$), the number of access transfers during a call is constrained by N. Therefore, $E[N^*|N = n]$ decreases as V increases. When N is large (e.g., $N \to \infty$), $E[N^*|N = n]$ is not affected by V (see (14)). When $V \leq 1/\eta^2$, $E[N^*|N = n]$ is not sensitive to the change of V.

Effects of η/μ and N on θ: Since we assume that $\eta_L/\eta_U = 1$, the maximum θ is 1/2 (see equation (30)). Figure 6 shows that θ decreases as η/μ increases. When η/μ increases, the number of the U-L transfers is more likely to reach the limit N, and the UE is forced to stay in UMTS (see Fact 2). Therefore, θ decreases as η/μ increases. This figure also indicates that θ increases as N increases. Increasing N has the same effect as decreasing η/μ due to Fact 2. Thus, larger θ is expected.

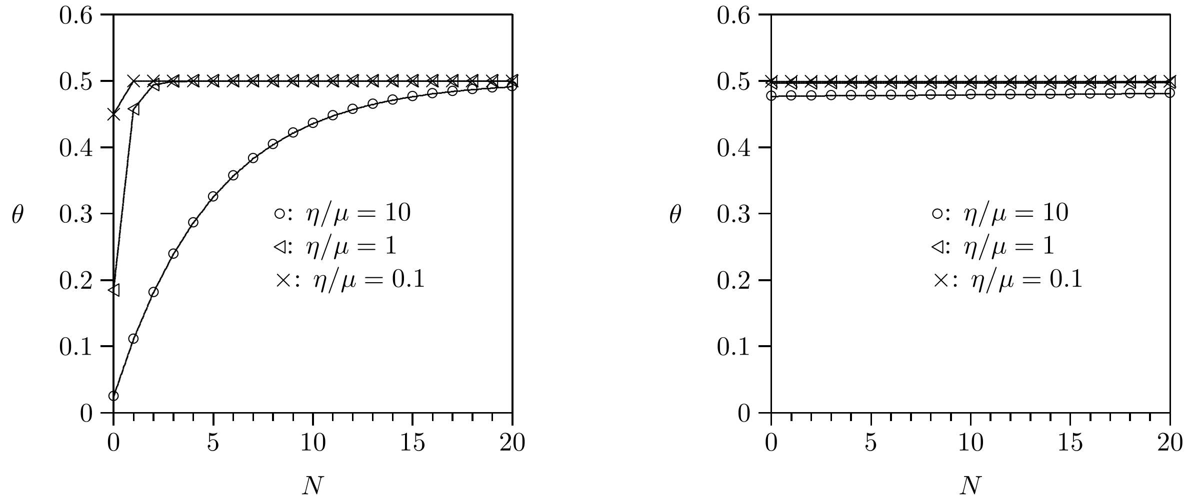
Effects of V on θ: Figure 5 (b) shows that when N is small (e.g., $N \leq 2$), θ increases as V increases. This phenomenon is explained as follows. When V increases, due to Fact 1, much longer $t_{L,i}$ ($t_{U,i}$) periods are observed. Since the calls are more likely to arrive at long $t_{L,i}$ ($t_{U,i}$) periods and are unlikely to be transferred, the number of the U-L transfers seldom reaches N and larger θ is observed. When N is large (e.g., $N \to \infty$), θ is not affected by V (see equation (30)).



(a) Effects of V on $E[N^*|N = n]$

(b) Effects of V on θ

Figure 5: Effects of V on $E[N^*|N = n]$ and θ



(a) $V = 0.001/\eta^2$

(b) $V = 1000/\eta^2$

Figure 6: The θ Curves

Based on the above discussion, when V and η/μ are small (i.e., $V \leq 1/\eta^2$ and $\eta/\mu \leq 1$), if $N \geq 3$, $E[N^*|N = n]$ is reasonably small and does not increase as N increases, and large θ is always observed (see the $\triangleleft$ and $\times$ curves in Figures 4 (a) and 6 (a)). Therefore, it is appropriate to select $N \geq 3$ to improve θ without significantly increasing $E[N^*|N = n]$.

When V is small and η/μ is large, θ and $E[N^*|N = n]$ significantly increase as N increases for $N \leq 10$, and these output measures are less sensitive to the change of N for $N \geq 15$ (see the $\circ$ curves in Figures 4 (a) and 6 (a)). In this case, the selection of N depends on the operation strategy of a mobile network, and is determined in network planning of the mobile operator. For example, if the mobile operator decides that $0.4 \leq \theta \leq 0.45$ is acceptable, then an N value ranges from 8 to 11 should be selected to satisfy this condition.

When V is large (i.e., $V > 1/\eta^2$), $E[N^*|N = n]$ increases as N increases (see Figure 4 (b)), and θ significantly increases as V increases (see the solid curves in Figure 5 (b)). A smaller N value should be selected to reduce $E[N^*|N = n]$ without seriously degrading θ (e.g., $1 \leq N \leq 8$). When V is extremely large (i.e., $V = 1000/\eta^2$), θ is always large in spite of the change of N (see Figure 6 (b)). In this case, $N = 0$ is selected to minimize $E[N^*|N = n]$ without significantly reducing θ.

5 Conclusions

This paper proposes using threshold N to limit the number of the access transfers for eSRVCC. An analytic model was developed to study the performance of the limited access transfer algorithm by measuring the expected number $E[N^*|N = n]$ of the access transfers in a call and the percentage θ of time that the UE resides in LTE during a call. Our study indicated the following results:

- When V is small (e.g., $V \leq 1/\eta^2$, which implies that the UE movement pattern is regular), $8 \leq N \leq 11$ are appropriate.

- When V is large (e.g., $1/\eta^2 < V < 1000/\eta^2$, which implies that the UE movement patter is irregular), $1 \leq N \leq 8$ are appropriate. When V is extremely large (e.g., $V = 1000/\eta^2$), $N = 0$ should be selected (i.e., the mobile operator should not allow the UE to switch from UMTS to LTE during a call).

The second conclusion is important. Most operators attempt to enforce access transfers for eSRVCC. Our study clearly indicated that if the user behavior is very irregular, access transfer should not be exercised. On the other hand, when the user behavior is regular, the N value needs to be carefully selected following our guidelines.

As a final remark, instead of setting the N values for the individual users, the telecommunications operators may consider the same N value for all users in the same geographic area for simplicity because the users in the same area may have similar call and mobility behavior (e.g., the vehicles on the highways). The operators may change the N values on weekdays, weekends, and holidays because the call and mobility behavior may change on different days. In the future, we will also study other approaches (e.g., timer-based scheme) to reduce access transfer traffic and enhance the LTE call control performance.

A Simulation Model

This appendix describes the discrete event simulation model for the LTE eSRVCC with limited access transfers. Let $\tau_{L,0}$ and $\tau_{U,0}$ be the residual lives of the LTE residence time $t_{L,i}$ and the UMTS residence time $t_{U,i}$, respectively. In the simulation, there are two methods to generate the samples of the residual lives $\tau_{L,0}$ and $\tau_{U,0}$.

Method 1 actually simulates the inter-call arrival times and the LTE/UMTS residence times. The samples of the residual lives $\tau_{L,0}$ and $\tau_{U,0}$ are computed from the time difference between the call arrival and the subsequent access transfer.

Method 2 generates the samples of the residual lives $\tau_{L,0}$ and $\tau_{U,0}$ from the residual life random number generators (to be elaborated later) [17].

This appendix describes the simulation model based on Method 2. The Method 1 simulation is similar to the one developed in the supplementary document of [15], and the details are omitted. In the remainder of this appendix, we first introduce the theorem for the residual life random number generation [17]. Then we describe our simulation model.

Theorem 1: Let t be a Gamma random variable with the mean $E[t]$ and the variance V. Define t^* to be a Gamma random variable with the mean $E[t]+V/E[t]$ and the variance $V+(V/E[t])^2$. Let random variable u be uniformly distributed over the interval (0,1). Let τ be the residual life of t. Then the distribution of τ is the same as the distribution of $u \times t^*$.

In the simulation, suppose that the samples of the call holding time t_c are obtained from an exponential generator G_c with the mean $1/\mu$, and P_L is the probability that a call will arrive at the LTE domain. Let $G(E[t], V)$ be the Gamma random number generator with the mean $E[t]$ and the variance V. The $t_{L,i}$ and $t_{U,i}$ samples are obtained from $G(E[t_{L,i}], V_L)$ and $G(E[t_{U,i}], V_U)$, respectively. The samples of u are obtained from a Uniform generator G_u. The samples of t_L^* and t_U^* are obtained from $G(E[t_{L,i}] + V_L/E[t_{L,i}], V_L + (V_L/E[t_{L,i}])^2)$ and $G(E[t_{U,i}] + V_U/E[t_{U,i}], V_U + (V_U/E[t_{U,i}])^2)$, respectively. According to Theorem 1, the samples of $\tau_{L,0}$ and $\tau_{U,0}$ are obtained from the residual life generators G_L and G_U, which multiply the u samples (generated from G_u) by the t_L^* and t_U^* samples (generated from $G(E[t_{L,i}] + V_L/E[t_{L,i}], V_L + (V_L/E[t_{L,i}])^2)$ and $G(E[t_{U,i}] + V_U/E[t_{U,i}], V_U + (V_U/E[t_{U,i}])^2)$, respectively.

We conduct the replicated simulation experiments. In every replicated run, we simulate a call arrival and the access transfers in this call. In the simulation model, an event e has two attributes:

- The *type* attribute indicates the event type. There are three event types. A **U-L Access Transfer** event represents that the UE performs an access transfer from UMTS to LTE during a call. An **L-U Access Transfer** event represents that the UE performs an access transfer from LTE to UMTS during a call. A **Call Release** event represents that a call is released.
- The *ts* attribute is the timestamp when the event occurs.

Six variables are used in the simulation:

- n^*: number of U-L access transfers in a call (i.e., in a replicated run)
- n_s^*: total number of L-U and U-L access transfers investigated in the simulation (i.e., in all replicated runs)
- N_c: total number of replicated runs (i.e., simulated calls) in the simulation
- T_L: portion of the call holding times that the UE resides in LTE
- T_U: portion of the call holding times that the UE resides in UMTS
- *domain*: a flag that indicates the domain (LTE or UMTS) where the UE resides

From the above variables, we compute

$$E[N^*|N = n] = n_s^*/N_c \text{ and } \theta = T_L/(T_L + T_U)$$

In the simulation, a clock ck is maintained to indicate the simulation progress, which is the timestamp of the event being processed. All events are inserted into the event list, and removed from the event list when the event is processed. Initially, ck is set to 0 to represent the call arrival time (t_2 in Figure 3). When the call is released (t_8 in Figure 3), a replicated run is complete. The clock ck is reset to 0, and the event list is re-initiated for the next replicated run. One million simulation runs are executed to obtain stable results. Figure 7 illustrates the simulation flow chart with the following steps:

Step 1. Set n_s^*, N_c, T_L, and T_U to 0.

Step 2. Initialize the event list. Set n^* and the simulation clock ck to 0.

Step 3. The domain where the UE resides when a call arrives is determined as follows. Generate a Uniform random number u which is drawn from G_u. If $u < P_L$, it means that the UE resides in LTE when the call arrives, and Step 4 is executed. Otherwise (i.e., the UE resides in UMTS), Step 5 is executed.

Step 4. The UE resides in LTE when the call arrives. Set *domain* to LTE. The **Call Release** event e_1 and first **L-U Access Transfer** event e_2 are generated and inserted into the event list. For event e_1, $e_1.type$ is **Call Release** and $e_1.ts$ is set to ck plus t_c obtained from G_c. For event e_2, $e_2.type$ is **L-U Access Transfer** and $e_2.ts$ is set to ck plus $\tau_{L,0}$ obtained from G_L. Then Step 6 is executed.

Step 5. The UE resides in UMTS when the call arrives. Set *domain* to UMTS. The **Call Release** event e_1 and first **U-L Access Transfer** event e_3 are generated and inserted into the event list. For event e_1, $e_1.type$ is **Call Release** and $e_1.ts$ is set to ck plus t_c obtained from G_c. For event e_3, $e_3.type$ is **U-L Access Transfer** and $e_3.ts$ is set to ck plus $\tau_{U,0}$ obtained from G_U.

Step 6. The first event e in the event list is deleted and is processed based on its type in Step 7.

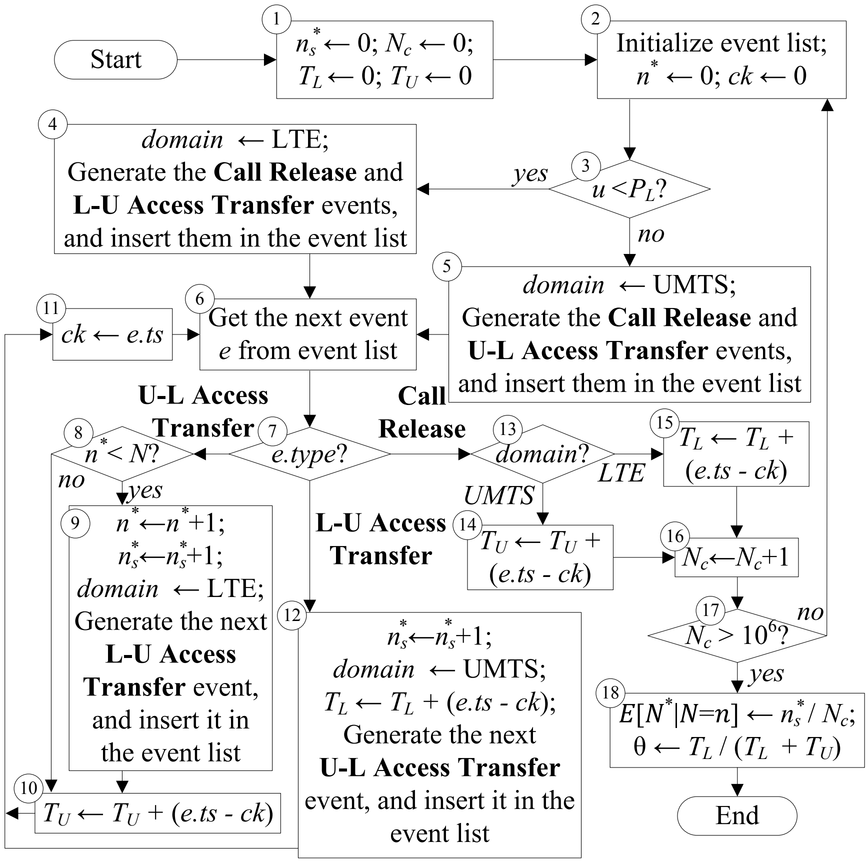


Figure 7: eSRVCC Simulation Flow Chart

Step 7. If *e.type* is **U-L Access Transfer**, then Step 8 is executed. If *e.type* is **L-U Access Transfer**, then Step 12 is executed. If *e.type* is **Call Release**, the simulation proceeds to Step 13.

Step 8 (U-L Access Transfer). If $n^* < N$, the network performs access transfer from UMTS to LTE, and Step 9 is executed. Otherwise, the transfer limit is reached, the UE has to remain in UMTS, and the simulation proceeds to Step 10.

Step 9. The UE executes the U-L Access Transfer procedure at UMTS. Increment both n^* and n_s^* by 1. Set *domain* to LTE. The next **L-U Access Transfer** event e_2 is generated and inserted into the event list, where $e_2.type$ is **L-U Access Transfer** and $e_2.ts$ is set to $e.ts$ plus $t_{L,i}$ obtained from $G(E[t_{L,i}], V_L)$.

Step 10. Calculate the time interval $e.ts - ck$ that the UE resides in UMTS. Increase T_U by this amount.

Step 11. Advance the simulation clock ck to $e.ts$, and proceed to Step 6.

Step 12 (L-U Access Transfer). The UE executes the L-U Access Transfer procedure at LTE. Increment n_s^* by 1. Set *domain* to UMTS. Calculate the time interval $e.ts - ck$ that the UE resides in LTE and increase T_L by this amount. The next **U-L Access Transfer** event e_3 is generated and inserted into the event list, where $e_3.type$ is **U-L Access Transfer** and $e_3.ts$ is set to $e.ts$ plus $t_{U,i}$ obtained from $G(E[t_{U,i}], V_U)$. The simulation proceeds to Step 11.

Step 13 (Call Release). If *domain* is *UMTS*, it means that the UE resides in UMTS when the call is released, and Step 14 is executed. Otherwise (i.e., the UE resides in LTE), the simulation proceeds to Step 15.

Step 14. Calculate the time interval $e.ts - ck$ that the UE resides in UMTS. Increase T_U by this amount. Then the simulation goes to Step 16.

Step 15. Calculate the time interval $e.ts - ck$ that the UE resides in LTE. Increase T_L by this amount.

Step 16. A call is released and N_c is incremented by 1.

Step 17. If one million of calls have been processed, then Step 18 is executed. Otherwise, the simulation proceeds to Step 2 for the next replicated run. In our experience, one million of **Call Release** events are enough to produce stable statistics, where the confidence intervals of the 99% confidence levels are within 3% of the mean values in most cases.

Step 18. The performance measures (i.e., $E[N^*|N = n]$ and θ) are computed, and the simulation is terminated.

References

[1] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access; Stage 2, Technical Specification 3GPP TS 23.401 version 12.0.0 (2013-03), 2013.

[2] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS); Service Description; Stage 2, Technical Specification 3GPP TS 23.060 version 12.0.0 (2013-03), 2013

[3] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2, Technical Specification 3GPP TS 23.228 version 12.0.0 (2013-03), 2013.

[4] Huei-Min Liao, Kai-Di Chang, Chi-Yuan Chen, Jiann-Liang Chen, Han-Chieh Chao, and Hua-Pei Chiang, "Emergency service model analysis to heterogeneous network environment," *2012 IEEE Globecom Workshops*, pp. 1027-1031, Dec 2012.

[5] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS) Service Continuity; Stage 2, Technical Specification 3GPP TS 23.237 version 12.0.0 (2013-03), 2013.

[6] Minghui Shi, Humphery Rutagemwa, Xuemin Shen, Jon W. Mark, and Aladdin Saleh, "A Service-Agent-Based Roaming Architecture for WLAN/Cellular Integrated Networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3168-3181, May 2007.

[7] Shun-Ren Yang and Wen-Tsuen Chen, "SIP multicast-based mobile quality-of-service support over heterogeneous IP multimedia subsystems," *IEEE Trans. Mobile Comput.*, vol. 7, no. 11, pp. 1297-1310, Nov 2008.

[8] Antonio Sanchez-Esguevillas, Belen Carro, Gonzalo Camarillo, Yi-Bing Lin, Miguel A. Garcia-Martin, and Lajos Hanzo, “IMS: the new generation of internet-protocol-based multimedia services,” *Proc. IEEE*, vol. 101, no. 8, pp. 1860-1881, August 2013.

[9] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia (IM) Core Network (CN) subsystem IP Multimedia Subsystem (IMS) service continuity; Stage 3, Technical Specification 3GPP TS 24.237 version 12.0.0 (2013-03), 2013.

[10] 3GPP, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Single Radio Voice Call Continuity (SRVCC); Stage 2, Technical Specification 3GPP TS 23.216 version 11.8.0 (2013-03), 2013.

[11] Phone Lin, Shan-Hung Wu, Chung-Min Chen, and Ching-Feng Liang, “Implementation and performance evaluation for a ubiquitous and unified multimedia messaging platform,” *Wireless Netw.*, vol. 15, no. 2, pp. 163-176, Feb. 2009.

[12] Leonard Kleinrock, *Queueing Systems, Vol. I, Theory*, Wiley, 1976.

[13] Ian F. Akyildiz, Joseph S. M. Ho, and Yi-Bing Lin, “Movement-based location update and selective paging for PCS networks,” *IEEE/ACM Trans. Netw.*, vol. 4, no. 4, pp. 629-638, Aug. 1996.

[14] Yi-Bing Lin and Yung-Chun Lin, “WiMAX location update for vehicle applications,” *Mobile Netw. and Appl.*, vol. 15, no. 1, pp. 148-159, Feb. 2010.

[15] Ren-Huang Liou, Yi-Bing Lin, and Shang-Chih Tsai, “An investigation on LTE mobility management,” *IEEE Trans. Mobile Comput.*, vol. 12, no. 2, pp. 166-176, Jan 2013.

[16] Yi-Bing Lin, “Performance modeling for mobile telephone networks,” *IEEE Network*, vol. 11, no. 6, pp. 63-68, Nov/Dec 1997.

[17] Hui-Nien Hung, Pei-Chun Lee, and Yi-Bing Lin, "Random number generation for excess life of mobile user residence time," *IEEE Trans. Veh. Technol.*, vol. 55, no. 3, pp. 1045-1050, May 2006.