INDIAN INSTITUTE OF MANAGEMENT CALCUTTA

WORKING PAPER SERIES

WPS No. 738/ November 2013

Managing Quality of Experience in 3G cellular network using priority queue

by

Samir K Sadhukhan
IIM Calcutta, Diamond Harbour Road, Joka P.O., Kolkata 700 104 India

Chayanika Bose
ETCE Dept. Jadavpur University Kolkata 700 032, India

Swarup Mandal
Wipro Technologies Kolkata, India

&

Debashis Saha
Professor, IIM Calcutta, Diamond Harbour Road, Joka P.O., Kolkata 700 104 India

Managing Quality of Experience in 3G cellular network using priority queue

Samir K Sadhukhan
IIM Calcutta, Diamond Harbour Road, Joka P.O., Kolkata 700 104 India
samir@iimcal.ac.in

Chayanika Bose
ETCE Dept. Jadavpur University Kolkata 700 032, India
chayanikab@ieee.org

Swarup Mandal
Wipro Technologies Kolkata, India
swarup.mandal@wipro.com

Debashis Saha
Professor, IIM Calcutta, Diamond Harbour Road, Joka P.O., Kolkata 700 104 India
ds@iimcal.ac.in

Abstract— **In a dual-homed 3G cellular network, signaling messages such as location update, handoff and paging arise due to the mobility of the users. These signaling messages are required to be processed efficiently to maintain or to improve the quality of service of the network. When a handoff fails, an ongoing call gets disconnected. A delay in paging increases time to establish a new call. The quality of signaling services have positive correlation with the users experience about the service. In this paper, we have considered priority processing to handoff messages over other signaling messages to reduce the delay in handoff processing in the network. We have analytically derived closed-form expressions for end-to-end delay of different type of signaling packets.**

Keywords- QoE, Signalling, queueing theory; cellular network; Handoff Management; 3G; dual homing;

I. INTRODUCTION

A 3G UMTS network is traditionally single homed as it has 'many-to-one' mapping from NodeBs to Radio Network Controllers (RNCs) in tier-1 and the same from RNCs to Mobile Switching Centres (MSCs)/Serving GPRS Support Nodes (SGSNs) in tier-2 [3]. But, in a mature network, over the years a certain percentage of users may exhibit a specific diurnal pattern [2][3] in their movement, and to address that change (known a priori) in mobility pattern, dual homing is shown to be a good option [3] as some inter-MSC handoffs can be converted to intra-MSC handoffs by virtue of dual homing of RNCs. This eliminates altogether some of the handoffs of the corresponding single-homed network. This contributes to quality of service (QoS) of the network. Further prioritizing the handoff processing over other signaling messages will reduce the delay of handoff processing. This will improve the ongoing call quality by accommodating more handoff processing by the network. Though there are several techniques to improve quality of experience (QoE) by identifying different types of signaling message prioritizing in a network, we will consider HO signaling message prioritization technique in this paper.

In this work, we have used the queuing analysis to calculate the signaling delay of a dual-homed network. We have suitably modified the model described in [1] to calculate the total delay for our dual-homed UMTS network.

The outline of this paper is as follows. Following introduction (Section I), section II presents timing diagram of signaling packets. Section III presents the signaling delay analysis for dual-homed network. Section IV concludes the paper.

II. TIMING DIAGRAM OF SIGNALING PACKETS AND ASSOCIATED DELAYS

We assume a simple dual-homed UMTS network that has one logical HLR. Let each MSC have one VLR only. Since the network is dual-homed, one RNC can be connected to at most two MSCs via two links. We represent MSC/P to denote the primary MSC of an RNC if the RNC is connected to MSC via primary link. A primary RNC of an MSC is connected via primary link. We also assume that each LA is a disjoint geographical area which constitutes collection of cells (i.e. NodeBs) under one RNC. When an MT crosses LA boundary, it generates an LU packet, by which the involved RNCs are informed of the location change of the MT [6].

We will consider processing of two types of handoff: *simple* handoff and *complex* handoff. In a simple handoff, only one MSC is involved in handling handoff process, whereas in complex handoff two MSCs are involved [3]. In a simple handoff MT changes its controlling RNC during a call without changing its controlling MSC. In case of complex handoff, both controlling RNC and controlling MSC of the MT changes. Thus when MT with an active call crosses the LA boundary, a handoff occurs. During handoff, an MT sends handoff signaling message to its controlling RNC. RNC sends relocation message to its MSC/P. At this point MSC decides the handoff type [5], and subsequent processing of the handoff takes place.

A signaling packet experiences transmission delay on wireless channel (both uplink and downlink), processing delays at different switches, such as NodeB, RNC and MSC, and processing delay at the wired network. Also there is some processing delay at MT.

To explain the various operations, we shall use the notations listed in Table 1 in the rest of the paper. The symbol '*' used in the Table I represents a place holder to indicated the type of signaling packets (e.g. LU, PG) to be used in its place.

TABLE I:LIST OF SYMBOLS USED FOR DELAY COMPONENT

$D^{i}_{TU}(*)$	Transmission delay on device uplink.
$D^{i}_{TD}(*)$	Transmission delay on device downlink
$D^{j}_{RU}(*)$	Processing delay on forward RNC-j
$D^{j}_{RD}(*)$	Processing delay on backward RNC-j
$D^{k}_{MU}(*)$	Processing delay on forward primary MSC-k
$D^{k}_{MD}(*)$	Processing delay on backward primary MSC-k
$D_{MP}(*)$	Processing delay for call setup, location update and handoff at the backend wired network
$D_{MT}(*)$	Transmission delay of call setup at MT.

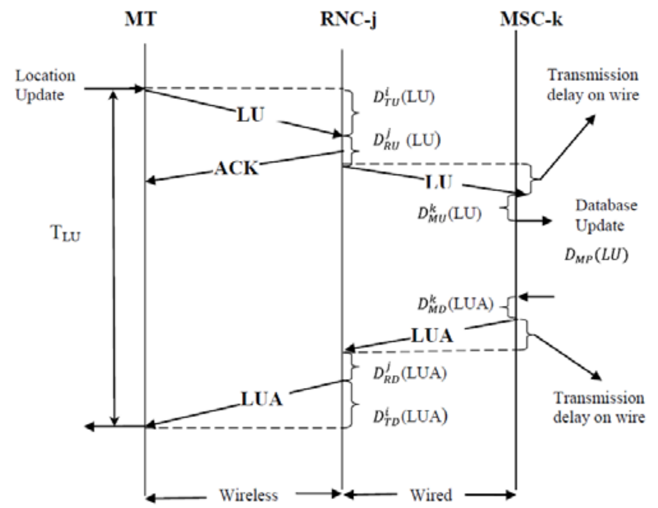

Fig. 1: Location Update and signaling packet delays in the network

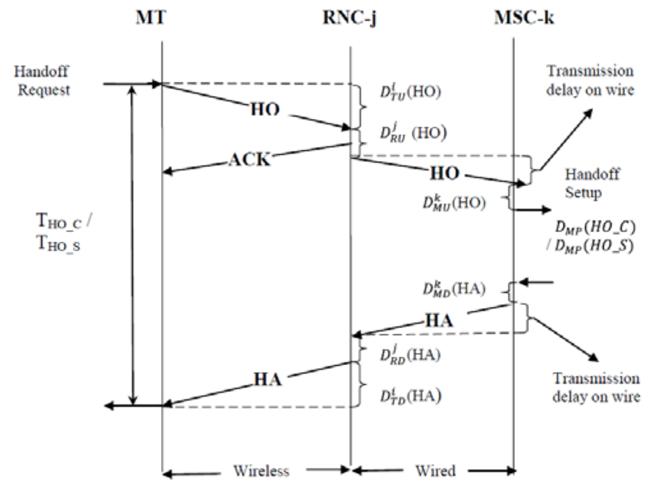

Fig. 2: Complex/Simple Handoff and signaling packet delays in the network

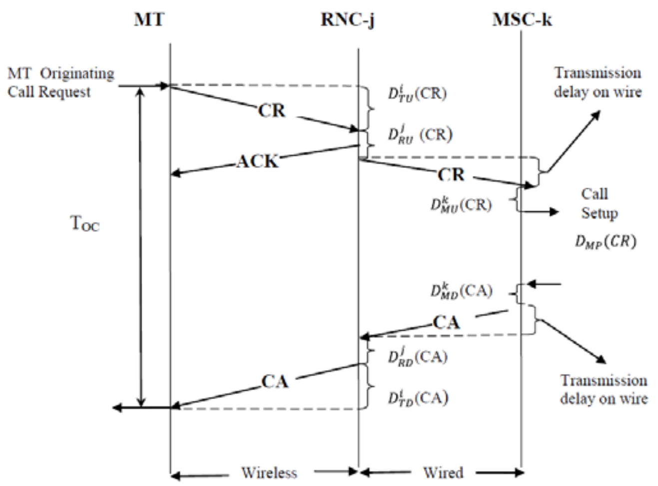

Fig. 3: MT Originating call and signaling packet delays in the network

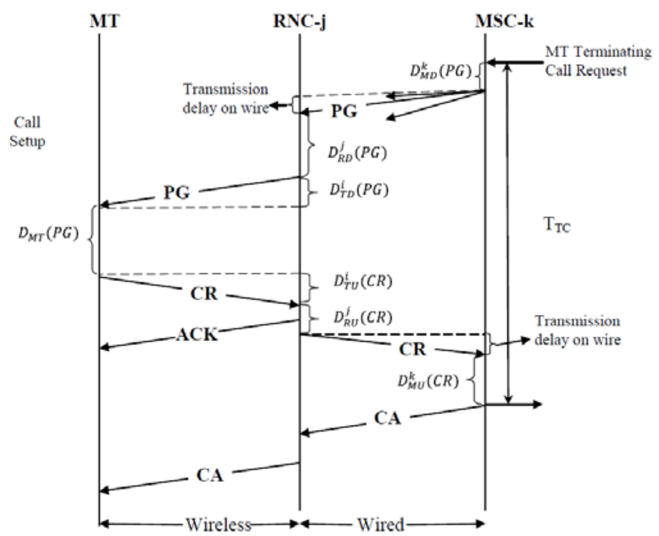

Fig. 4: MT Terminating Call and signaling packet delays in the network

We divide the MTs into two types: *busy* and *idle*. If an MT is making/receiving call, it is said to be busy; else idle. For an idle MT, two types of call arrival are possible. The first type is outgoing call initiated by MT by sending a CR (call request) packet to RNC. We denote this call as *MT-originating call*. The second type is incoming call initiated by another MT (say MT') and the call is targeted for the MT under consideration. Here, the MT will receive a PG packet (issued by MSC) and issue a CR packet in response to it. We denote this call as *MT-terminating call*. When an MSC receives an MT-terminating call request for the MT within its LA, PG packets are broadcast by MSC to RNCs in the LA and subsequently by RNCs to NodeBs to alert the MT of the incoming call. Upon receipt of the response from the destined MT, the call setup becomes possible. . Fig 1 and Fig 2 show the flow of signaling packets for LU and HO respectively whereas Fig 3 and Fig. 4 show the flow for MT-originating and MT-terminating call setup respectively.

The wireless channel is divided into the voice/data channel for user voice/data and the signaling channel for the signaling packets. In the wireless access network we have chosen Time division multiple access/Time division duplex (TDMA/TDD) [1] as a multiple access method. In TDMA/TDD protocol, a fixed length frame is divided into uplink and downlink dynamically. Both downlink and uplink consist of control slots and user data slot for each direction (Fig 5).

When the MT has a new signaling packet to transmit, it first sends the packet on the control slot, according to the slotted ALOHA protocol. If NodeB receives the signaling packet correctly, it returns the acknowledge packet during the next downlink. On the other hand, if packet collision occurs, the NodeB sends no response, so that MT will retry the transmission at the next frame. We will consider the delays caused by retransmissions to derive from the transmission delay on radio channel.

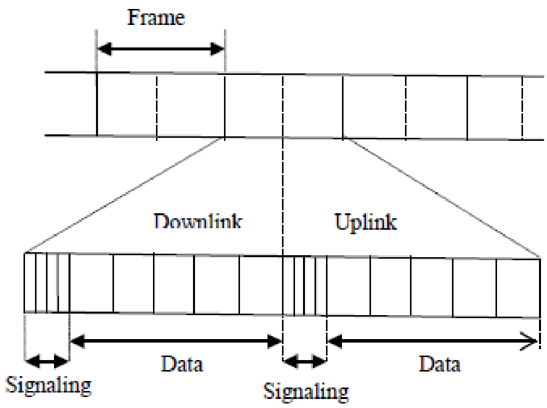

Fig. 5. Radio Frame structure of TDMA/TDD [1]

LU processing delay (T_{LU}): Fig. 1 shows the flow of LU request - MT→ RNC →MSC/P and the corresponding LUA (LU acknowledgement) is returned as MSC/P→ RNC → MT.

The different delay components in LU are shown in Fig. 1. The total delay for LU T_{LU} can be obtained as follows:

$$T_{LU} = D_{TU}^{i}(LU) + D_{RU}^{j}(LU) + D_{MU}^{k}(LU) + D_{MP}(LU) + D_{MD}^{k}(LUA) + D_{RD}^{j}(LUA) + D_{TD}^{i}(LUA)$$

Location search & MT Terminating Call setup processing delay (T_{TC}): Fig. 4 shows the flow of LS (location search) request for MT-terminating call -MSC/P→RNC→MT. T_{TC} is the time from the arrival of call request at MSC/P to the time when its response is sent to the wired network. For calls destined to some other MT, PG packet is broadcast from MSC/P to RNC in its LA, and sent to MTs via NodeB. Then the corresponding MT generates CR packet which is returned to the wired network.

The different delay components in MT-terminating call setup are shown in the Fig.4. The total delay can be obtained as follows:

$$T_{TC} = D_{MD}^{k}(PG) + D_{RD}^{j}(PG) + D_{TD}^{i}(PG) + D_{MT}(PG) + D_{TU}^{i}(CR) + D_{RU}^{j}(CR) + D_{MU}^{k}(CR)$$

MT Originating Call setup processing delay (T_{OC}): Fig. 3 shows the flow MT-originating call -MT →RNC→MSC/Pand the corresponding CA (call setup acknowledgement) is returned as MSC/P→ RNC → MT.

The different delay components in MT-originating call setup are shown in the Fig.3. The total delay can be obtained as follows:

$$T_{OC} = D_{TU}^{i}(CR) + D_{RU}^{j}(CR) + D_{MU}^{k}(CR) + D_{MP}(CR) + D_{MD}^{k}(CA) + D_{RD}^{j}(CA) + D_{TD}^{i}(CA)$$

Handoff processing delay ($T_{HO\text{-}S}/T_{HO\text{-}C}$): Fig. 2 shows the flow of HO request - MT→ RNC →MSC/P and the corresponding HA (HO acknowledgement) is returned as MSC/P→ RNC → MT. MSC/P will differentiate between simple handoff and complex handoff and necessary processing will take place at wired network.

The different processing delays for simple and complex handoff are shown in the Fig.2. The total delay can be obtained as follows:

$$T_{HO\text{-}S} = D^{i}_{TU}(HO) + D^{j}_{RU}(HO) + D^{k}_{MU}(HO) + D_{MP}(HO\_S) + D^{k}_{MD}(HA) + D^{j}_{RD}(HA) + D^{i}_{TD}(HA)$$
$$T_{HO\text{-}C} = D^{i}_{TU}(HO) + D^{j}_{RU}(HO) + D^{k}_{MU}(HO) + D_{MP}(HO\_C) + D^{k}_{MD}(HA) + D^{j}_{RD}(HA) + D^{i}_{TD}(HA)$$

III. QUEUEING NETWORK MODEL AND DELAY ANALYSIS

To analytically derive processing delays, we build a queuing network model as depicted in Fig.6 for dual-homed UMTS network. We define the following queues:

1. WCU (wireless channel uplink) queue: The uplink is contended by CR, LU, HO packets and additional CR packets in response to PG packets. The delays of those packets are determined by the approximate analysis which takes into account collision of the packets due to the multiple access over the uplink channel.

2. WCD (wireless channel downlink) queue: The packet delay at the downlink queue can be modeled by (M/D/1) queue processed in FIFO manner since the RNC can actively schedule the CA, LUA, HA, PG and ACK packets transmission on the wireless channel.

3. FRNC (forward RNC) queue: When RNC processes the LU, HO and CR packets, it returns acknowledgement (ACK) to the sending MT via the corresponding NodeB. At the same time RNC forwards the LU, HO and CR packet to primary MSC. The packets are processed in an FIFO manner and we have modeled FRNC queue as an M/M/1 queue. We have adopted priority processing of HO packets over LU and CR packets.

4. BRNC (backward RNC) queue: Signaling packets coming from MSC/P are processed in an FIFO manner at the RNC, modeled as M/M/1 queue. Since we have given priority processing for HO packets at the forward queue, we take priority processing for HA packets.

5. FMSC/P (forward MSC - primary) queue: MSC/P handles incoming packets from RNCs. Those packets are processed in an FIFO manner (M/M/1 queue) and forwarded to the wired network. Like in FRNC we take priority processing for HO packets.

6. BMSC/P (backward MSC - primary) queue: CA, HA and LUA packets are returned to RNC by MSC/P. PG packets coming from the wired network are also forwarded (broadcast) to the RNCs under this MSC. Those packets are processed in an FIFO manner and the queue is modeled as M/M/1 queue. Like in BRNC we take priority processing for HA packets.

7. CR, LU and HO_C and HO_S queues: We assume that CR, LU and HO_C and HO_S packets are processed separately in the wired network. CR, LU and HO_S and HO_C queues are used for calculating the delays of the corresponding packets at the wired network. We use infinite server queues for the scheduling discipline. We assume that separate queues are available to process simple and complex handoffs in the wired network.

Let us consider that there are a number of NodeBs, b number of RNCs and c number of MSCs in the dual-homed cellular network. Let I = $\{1,2,\ldots.a\}$ denote the set of NodeBs, J=$\{1,2,\ \ldots\ b\}$ denote the set of RNCs, and K=$\{1,2,\ldots c\}$ denote the set of MSCs. There can be at most one connection (one link) from NodeB to RNC and at most two connections (one primary link and one secondary link) from RNC to MSC. We will use a small letter to denote a member of the set represented by the corresponding capital letter; for example, iϵ I, j ϵ J, k ϵ K. A cell is a wireless coverage area under a NodeB. Let NodeB-i denote the i^{th} NodeB. Similarly for cell-i, RNC-j and MSC-k. Let us consider the following binary variables:

e_{ij} = 1, if NodeB-i is connected to RNC-j via a link,
0, otherwise
e'_{jk} = 1, if RNC-j is connected to MSC-k via a primary link,
0, otherwise

We assume the following:

1. Exponential distribution of the following random variables: (i) call duration in an MT (ii) MT's sojourn time in a cell (iii) processing time at MT (iv) processing time at RNC (v) processing time at MSC (vi) call setup processing time in the wired network (vii) location update and handoff processing time in the wired network.

2.Poisson distribution of random variables: (i) MT-originating call arrival (ii) MT-terminating call arrival.

3.The transmission delay at wired network is zero because it is negligibly small compare to other delays.

4.No processing delay at any NodeB.

5.A busy MT cannot generate CR packets.

6.Each random variable is independently and identically distributed (iid).

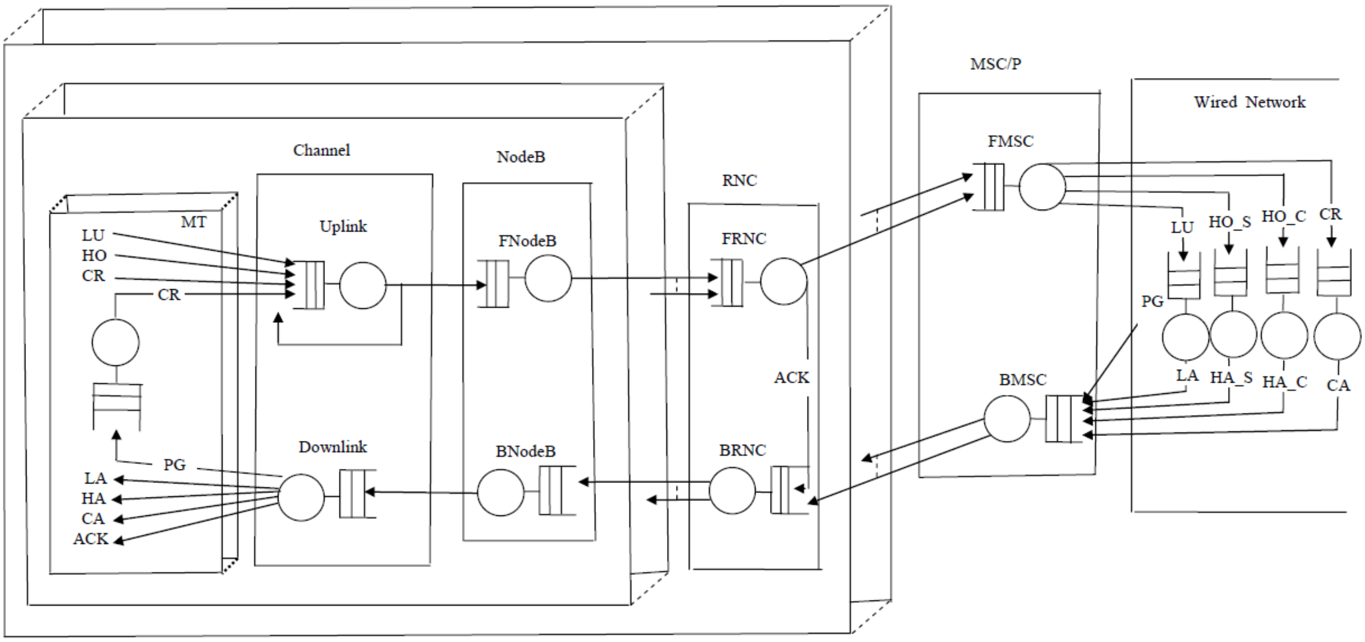

Figure 6: Queuing Network of dual-homed cellular network

TABLE II:LIST OF PARAMETERS OF THE NETWORK

Symbol	Description
N^i_{MT}	Number of MTs in cell-i
λ_T	Average MT-terminating call arrival rate in MT
λ_0	Average MT-originating call arrival rate in MT
$1/\mu$	Average call duration of an MT
$1/h$	Average MT sojourn time in a cell
$1/\mu_{MT}$	Average processing time in MT
$1/\mu_{RNC-j}$	Average processing time in RNC-j
$1/\mu_{MSC-k}$	Average processing time in MSC-k
$1/\mu_{CR}$	Average call setup time for an MT in wired network
$1/\mu_{LU}$	Average location update time in wired network
$1/\mu_{HO\_S}$	Average simple handoff processing time in wired network
$1/\mu_{HO\_C}$	Average complex handoff processing time in wired network
T	Frame length
n_u, n_d	Number of control slots in uplink and downlink respectively
α	Ratio of number of simple handoff packets to the total number of handoff packets in the network

Let us assume that the NodeB-i is under the RNC-j which is under the primary MSC-k. Before we go into the analysis of the delay from the queues adopted here, we first calculate mean arrival rates of CR, LU, PG, HO, CA, HA, LUA and ACK packets at the respective link (uplink and/or downlink) of cell-i.

Analysis of uplink transmission delay:

Let λ^i_{CR}be the CR packet arrival rate at uplink of the cell-i. CR packets are generated by the call request in both directions (MT-originating call and MT-terminating call). Moreover from our assumption, that a busy MT cannot generate CR packets, only idle MT can generate CR packets. Let p_{idle} be the probability that the MT has no active call. The average number of idle MTs of the cell-i is $p_{idle}*N^i_{MT}$. Then $\lambda^i_{CR} = (\lambda_o + \lambda_T)\, p_{idle}*N^i_{MT}$. To calculate p_{idle}, we assume each MT as M/M/1/1 queue system. Then $p_{idle} = \mu/(\mu + \lambda_o + \lambda_T)$

Let λ^i_{LU} be the arrival rate of the LU packets at the cell-i. We have $\lambda^i_{LU} = (p_{idle} * N^i_{MT}) * C^i * h$ where C^i is the probability that the cell boundary falls in the boundary of the LA [1]. When the MT is busy LU update request is piggybacked on the handoff request, so no extra channel request is needed. In this case LU update is assumed to be processed in parallel with handoff request.

When MT having active call moves into another location area, the MT generates HO packets. Therefore, the arrival rate of HO packets is given by $\lambda^i_{HO} = h * C^i * (1 - p_{idle}) * N^i_{MT}$.

So the total packet arrival rate is $(\lambda^i_{CR} + \lambda^i_{LU} + \lambda^i_{HO})$ at the uplink of cell-i.

Since Slotted ALOHA is considered for multiple access method at the uplink, collisions of the packets should be taken into consideration. Let the number of retransmissions until the packet is successfully admitted by FNodeB follows geometric distribution with parameter p^i_{col}, the probability of collision at i-th cell. Then the average number of retransmission before a successful transmission is given by $1/(1-p^i_{col})$. Thus the packet arrival rate including retransmission due to collision as

$$\lambda^i_{up} = (\lambda^i_{CR} + \lambda^i_{LU} + \lambda^i_{HO})/(1- p^i_{col}).$$

Assuming that the packets having arrived in the previous frame also compete for the control slots in the current frame, the packet collision probability, p^i_{col}, can be determined as

$$p^i_{col} = 1 - e^{-T\, \lambda^i_{up}/n_u}$$

Thus $p^i_{col} = 1 - e^{-T(\lambda^i_{CR}+\lambda^i_{LU}+\lambda^i_{HO})/(n_u*(1- p^i_{col}))}$ can be obtained iteratively.

Now after NodeB receives the control packet, it returns ACK in the downlink part of the next frame. Thus, the MT not receiving the ACK immediately, retransmits the packet in the next control slots. Thus the average number of MT's retransmission before a successful receive of ACK packet is $p^i_{col}/(1- p^i_{col})$.

We finally have the uplink delay

$$D^i_{TU}(*) = \frac{p^i_{col}}{(1-p^i_{col})} T + \frac{T}{2}$$

The second term of the above equation represents the average delay from packet arrival to the control slot at the next frame.

Analysis of downlink transmission delay:

When the MT-terminating call arrives at MSC, it broadcasts PG packets to all its primary RNCs which in turn broadcast to all its NodeBs. Since the PG packets are broadcast by MSC and replicated by all its primary RNCs and respective NodeBs, each cell in the location area will experience the same PG packet arrival rate as experienced by the corresponding MSC. Let λ^k_{PG} be the PG packet arrival rate at the primary MSC-k. Then

$$\lambda^k_{PG} = \lambda_T * p_{idle} * \sum_{j\epsilon J} \sum_{i\epsilon I} e'_{jk} e_{ij} N^i_{MT}$$

By our assumption, the NodeB-i is under the RNC-j which is under the primary MSC-k. Therefore, the arrival rate of PG packets at the RNC-j and at the cell-i is also λ^k_{PG} each.

Since the MSC receiving CR, HO and LU packets generates corresponding CA, HA and LUA packets, those arrival rates at the downlink of cell-i are obtained as

$$\lambda^i_{CA} = \lambda^i_{CR},\ \lambda^i_{HA} = \lambda^i_{HO} \text{ and } \lambda^i_{LUA} = \lambda^i_{LU}$$

Furthermore, RNC returns the ACK packet when it receives CR and LU packets correctly. The arrival rate of ACK packets at the downlink of cell-i is then given as

$$\lambda^i_{ACK} = \lambda^i_{CR} + \lambda^i_{LU} + \lambda^i_{HO}$$

We have the arrival rate of downlink at cell-i as follows:

$$\lambda^{ki}_{DOWN} = (\lambda^k_{PG} + \lambda^i_{CA} + \lambda^i_{HA} + \lambda^i_{LUA} + \lambda^i_{ACK})$$

We approximately model the downlink queue as an M/D/1 queue where transmission time of signaling packets is T/n_d. then we obtain the downlink delay at cell-i as follows

$$D^{i}_{TD}(*) = T^2\lambda^{i}_{DOWN}/(2n_d(n_d - T\lambda^{i}_{DOWN})) + \frac{T}{2}$$

Analysis of FRNC Queue: To calculate the delays due to different type of packets, we need to analyze FRNC queue at RNC-j. The arrival rates of CR, HO and LU packets at FRNC queue at RNC-j is given by $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{HO}$, $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{CR}$ and $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{LU}$ respectively. We have assumed the priority processing of HO packets over other signaling packets. We will derive the processing delay at FRNC at RNC-j by modeling it as a non-preemptive priority scheduling M/M/1 queuing system [4]. By assuming exponential distribution for processing time with means $1/\mu_{RNC\text{-}j}$ for all signaling packets, we can derive delays for HO, CR and LU packets as follows:

$$D^{j}_{RU}(HO) = 1/\mu_{RNC\text{-}j} + R/(1-\rho_1)$$

$$D^{j}_{RU}(CR) = D^{j}_{RU}(LU) = 1/\mu_{RNC\text{-}j} + R/\{(1-\rho_1)(1-\rho_1-\rho_2)\}$$

where $\rho_1 = (1/\mu_{RNC-j})\sum_{i\epsilon I} e_{ij}\lambda^{i}_{HO}$

$\rho_2 = (1/\mu_{RNC-j})\sum_{i\epsilon I} e_{ij}(\lambda^{i}_{CR}+\lambda^{i}_{LU})$

$R = (1/\mu_{RNC-j})^2\sum_{i\epsilon I} e_{ij}(\lambda^{i}_{HO}+\lambda^{i}_{CR}+\lambda^{i}_{LU})$

Analysis of BRNC Queue: We need to analyze BRNC queue at RNC-j. ACK packets should be considered in addition to CA, LUA, HA and PG packets. The arrival rates of CA, LUA, PG, HA and ACK packets at BRNC queue at RNC-j is given by λ^{k}_{PG}, $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{CA}$, $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{LUA}$, $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{HA}$, $\sum_{i\epsilon I} e_{ij}\lambda^{i}_{ACK}$ respectively. Since we have considered the priority for HO packets, we need to take the priority processing for HA packets too. We can derive the processing delay at BRNC queue at RNC-j by modeling it as a non-preemptive priority scheduling M/M/1 queuing system. By assuming exponential distribution for processing time with means $1/\mu_{RNC\text{-}j}$ for all signaling packets, we can derive delays for different types of packets as follows:

$$D^{j}_{RD}(HA) = 1/\mu_{RNC\text{-}j} + R/(1-\rho_1)$$

$$D^{j}_{RD}(CA) = D^{j}_{RD}(LUA) = D^{j}_{RD}(ACK) = D^{j}_{RD}(PG) = 1/\mu_{RNC\text{-}j} + R/\{(1-\rho_1)(1-\rho_1-\rho_2)\}$$

where $\rho_1 = (1/\mu_{RNC-j})\sum_{i\epsilon I} e_{ij}\lambda^{i}_{HA}$

$\rho_2 = (1/\mu_{RNC-j})(e'_{jk}\lambda^{k}_{PG} + \sum_{i\epsilon I} e_{ij}(\lambda^{i}_{CA}+\lambda^{i}_{LUA}+\lambda^{i}_{ACK}))$

$R = (1/\mu_{RNC-j})^2(e'_{jk}\lambda^{k}_{PG} + \sum_{i\epsilon I} e_{ij}(\lambda^{i}_{HA}+\lambda^{i}_{CA}+\lambda^{i}_{LUA}+\lambda^{i}_{ACK}))$

Analysis of FMSC Queue: When a RNC is connected to a MSC, it is said to be primary (secondary) RNC if the RNC is connected to the MSC via a primary (secondary) link. Thus a MSC can have both primary RNCs and secondary RNCs. The FMSC processes signaling packets from its primary RNCs. To calculate delays, we need to find the arrival rate of aggregate of each signaling packets from all primary RNCs under MSC-k. Since we have priority processing of HO packets, We will derive the processing delay at FMSC at MSC-k by modeling it as a non-preemptive priority scheduling M/M/1 queuing system . By assuming exponential distribution for processing time with means $1/\mu_{MSC\text{-}k}$ for all signaling packets, we can derive delays for HO, CR and LU packets as follows. Thus, for a given k, we have:

$$D^{k}_{MU}(HO) = 1/\mu_{MSC\text{-}k} + R/(1-\rho_1)$$

$$D^{k}_{MU}(LU) = D^{k}_{MU}(CR) = 1/\mu_{MSC\text{-}k} + R/\{(1-\rho_1)(1-\rho_1-\rho_2)\}$$

where $\rho_1 = (1/\mu_{MSC-k})\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{HO}e_{ij}e'_{jk}$

$\rho_2 = (1/\mu_{MSC-k})\sum_{j\epsilon J}\sum_{i\epsilon I}(\lambda^{i}_{LU}+\lambda^{i}_{CR})e_{ij}e'_{jk}$

$R = (1/\mu_{MSC-k})^2\sum_{j\epsilon J}\sum_{i\epsilon I}(\lambda^{i}_{LU}+\lambda^{i}_{CR}+\lambda^{i}_{HO})e_{ij}e'_{jk}$

Analysis of BMSC Queue: Note that at BMSC queue, the CA, LUA, HA and PG packets to primary RNCs are processed. We will derive the processing delay at BMSC at MSC-k by modeling it as a non-preemptive priority scheduling M/M/1 queuing system. By assuming exponential distribution for processing time with means $1/\mu_{MSC\text{-}k}$ for all signaling packets, we can derive delays for CA, HA, PG and LUA packets as follows. Thus, for a given k, we have:

$$D^{k}_{MD}(HA) = 1/\mu_{MSC\text{-}k} + R/(1-\rho_1)$$

$$D^{k}_{MD}(CA) = D^{k}_{MD}(LUA) = D^{k}_{MD}(PG) = 1/\mu_{MSC\text{-}k} + R/\{(1-\rho_1)(1-\rho_1-\rho_2)\}$$

where $\rho_1 = (1/\mu_{MSC-k})\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{HA}e_{ij}e'_{jk}$

$$\rho_2 = (1/\mu_{MSC-k})\,(\lambda^{k}_{PG} + \sum_{j\epsilon J}\sum_{i\epsilon I}(\lambda^{i}_{LUA} + \lambda^{i}_{CA}) e_{ij} e'_{jk})$$
$$R = (1/\mu_{MSC-k})^2(\lambda^{k}_{PG} + \sum_{j\epsilon J}\sum_{i\epsilon I}(\lambda^{i}_{LUA} + \lambda^{i}_{CA} + \lambda^{i}_{HA}) e_{ij} e'_{jk})$$

Analysis of the processing delay at the wired network: For handoff processing, whether a handoff is simple handoff or a complex handoff is determined at MSC [our ref.] and the simple handoff and complex handoff are processed in separate queues in the wired network. We assume there is no delay in scheduling of an arriving packet in the respective queues of the wired network. Assuming each queue is a M/M/1 queue with respective parameter, we can calculate the processing delays as follows. Thus, the processing delays at the wired network is given by

$$D_{MP}(LU) = 1/(\mu_{LU} - \sum_{k\epsilon K}\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{LU}\ e_{ij}\, e'_{jk})$$
$$D_{MP}(CR) = 1/(\mu_{CR} - \sum_{k\epsilon K}\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{CR}\ e_{ij}\, e'_{jk})$$
$$D_{MP}(HO\_S) = 1/(\mu_{HO\_S} - \sum_{k\epsilon K}\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{HO\_S}\ e_{ij}\, e'_{jk})$$
$$D_{MP}(HO\_C) = 1/(\mu_{HO\_C} - \sum_{k\epsilon K}\sum_{j\epsilon J}\sum_{i\epsilon I}\lambda^{i}_{HO\_C}\ e_{ij}\, e'_{jk})$$

where $\lambda^{i}_{HO\_S} = \alpha * \lambda^{i}_{HO}$ and $\lambda^{i}_{HO\_C} = (1-\alpha) * \lambda^{i}_{HO}$

Analysis of processing delay at MT:

$$D_{MT}(PG) = 1/(\mu_{MT} - \lambda_T)$$

Thus, the total delay of HO, LU, MT-originating call set up and MT-terminating call setup can be calculated analytically.

IV. CONCLUSION

In this paper, we have proposed a delay analysis of various signaling packets in dual-homed 3G UMTS cellular network. The signaling delays for LU, HO, MT-terminating call setup and MT-originating call setup are analytically derived using a queuing network model from a given dual-homed cellular network with parameters as shown in Table II. We have assumed that each cell in the network has different number of MTs and each LA has varying number of contiguous cells under it. The position of the cell in an LA is considered for calculating the delay of LU. We have considered priority processing of HO packets with non preemptive scheduling of signaling packets in our queuing model.

REFERENCES

[1] M. Sugano, D. S. Eom, M. Murata and H. Miyahara, "Performanance Analysis of Signalling Delay for Wireless Cellular Networks", Journal of Computing and Information Technology – CIT 12, 2004, 3, 251-262

[2] S K Sadhukhan, S Mandal, and D Saha, "A Heuristic Techneque for Solving Dual-homing Assignment Problem of 2.5G Cellular Networks", ICCTA, pp 66-71, 2007

[3] S K Sadhukhan, S Mandal, S R Biswas, P Bhaumik and D Saha, "Post-deployment Tuning of UMTS Cellular Networks through Dual-homing of RNCs", COMSNETS, 2009

[4] D. Bertsekas, R. Gallager, "Data Networks", Second Edition, Prentice-Hall, 1992

[5] S K Sadhukhan, U Tiwari, Vishal Agarwal, Vivek Agarwal, P Venkateswaran, "Mobility Management Protocol in Dual-homed UMTS networkks", Proc. of Third IEEE International Conference on Emerging Applications of Information Technology (EAIT), pp. 361-365, 2012.

[6] S. K. Sadhukhan, S. Mandal, K Orea, D. Saha, "A Novel Location Management technique for Dual-homed UMTS Cellular Networks", (poser paper): In Proc. IEEE COMSNET, Bangalore, India, January, 2012.