Delay Analysis of Wireless Cellular Networks for Better Qos

Satyasrikanth Palle, Shivashankar

Abstract: The message for call requests is created by mobile devices during a call which is then sent to a base station (BS). A BS processes the response of a call request and chooses to accept or deny the call. Signals such as location notifications, paging and switching due to user mobility take a significant share of the total traffic load within mobile cellular networks. Therefore, between signaling packets, the maximum allowable delays may differ. This time will be delayed because if the time is longer than the allowable pause. The quality of service is therefore reduced, which for service providers is not acceptable. In this paper, we propose an empirical model to determine an overall delay in the processing of wireless cell network signaling packets, which involves the delay in the radio channel and the wired component delay in processing. We are demonstrating the effectiveness of priority processing in reducing handoff delays. We also assess the delay between cells according to their positions in the area and their influence on processing delays by the number of nodes. In addition, we evaluate the difference in delay between cells depending on their position within the network area and how many stations influence time delayed processing.

Keywords: Quality of Service, Handoff delay, mobility, mobile cellular network, signaling.

I. INTRODUCTION

Nowadays Wireless Cellular Communication plays a dynamic role in transfer of information from one mobile node to another node considering the geographical region which is divided into smaller cell areas. All Base Stations (BS) are connected through a wireless connectivity via wireless channels. To a Base Station Controller (BSC) many base stations are connected which are then linked to a network substation. The Mobile Switching Center (MSC) network substation consists of the Register of Home Location (HLR) and Register Visitor Location (VLR). Update the routing data to authenticate the phone client. The Architecture of a Cellular network is shown in Figure 1. Through Channel access procedure the mobile station (MS) send a request for channel to start a communication session between the mobile nodes. Once the call request has been successfully answered, BS must communicate with its BSC and network subsystem. The processing of the request message is done by subsystem and will take decision to reject or accept the call.

Revised Manuscript Received on December, 30 2019.

Satyasrikanth Palle, Department of Electronics and Communication Engineering, Sri Venkateshwara College of Engineering, Bangalore, India
Shivashankar, Department of Electronics and Communication Engineering, Sri Venkateshwara College of Engineering, Bangalore, India

Figure 1: Architecture of Cellular Network

When call request messages are received, BS communicates effectively with its BSC and the network subsystem, and thus decides to accept or reject your application. If any idle wireless channel is present then BS accepts the call and allocates the channel for communication by utilizing the radio resources efficiently at BS by incorporating different call Admission control (CAC) schemes. The time to process a message with a call request varies with various CAC systems. For example, If the channel borrowing method does not have an unusual channel, the BS borrows from one of its neighboring cells and assigns it to the call demand. A BS is looking into its channel pool in a fixed guard channel scheme, for an idle channel. When the idle channel is identified, the channel is assigned to the demand call. These systems take no time to process a call request and are handled randomly. In terms of service providers’ time to provide impeccable service to high quality of services (QoS), it is important to analyze the processing delay of a call request and also the time to wait. If this time is longer than the permissible delay, the associated call will be blocked. In cellular wireless networks, a substantial proportion of the overall traffic load includes signal traffic such as location updates, paging, and handoff. These signal charges grow exponentially as mobile terminals increase and the size of radio cells decrease [1]. A notion of where the area of the cellular wireless network is initially divided into locations to reduce the amount of location updates data. That location area is comprised of several cells and a fundamental unit for mobile terminal location management to reduce the traffic load of locations modified by the larger location area shown in Figure 2. But there is a growing ambiguity as to the location of the mobile end terminal. This will raise the traffic load for paging. The load is also increased at the mobile switching center for each location. The scale of the location area therefore has a significant effect on packet delays in signaling [2].

DOI: 10.35940/ijeat.5517-9-9-2
We present an analytical method in this paper to assess the delay in complete interaction of wireless cellular network repositories. With this model, we are able to develop a system that meets the delay requirements of every processing type. The maximum processing delays for link configuration, update position and transfer are obtained from an empirical approach. In addition, to reduce the time for handoff and to test its efficacy by quantitative instances, we implement priority processing. The delay in communication should also be determined by the location of the cells in the region as the frequency of arrival of position update packages differs. Thus, according to the location of its position, we shall compare the delays in each cell. We will also discuss the impact on processing delays of a variety of terminals.

The rest of the paper is organized as follows. Section II describes the wireless network design and the transmission lag being considered. Section III demonstrates the analytical approach to measure each processing delays. In Section IV we discuss improving the quality of the priority handoff delay processing, the delay variations between cells and the effect of a number of terminals. At last, in Section V we conclude our paper.

II. LITERATURE REVIEW

Previous studies are overwhelmingly devoted to cellular network wired signaling traffic. Signaling load Signo 7SS7 in Pan-European GSM and in North American IS-41 have been tested by the researchers in [3]. Tipper et al. [4] talk about deficiencies at BS that increase and decrease call processing time. Chen et al. [5] addressed failure of hardware as well as failure of code. In order to mitigate the effects of BS faults, the network engineers introduce some fault tolerant techniques.

Varshney et al. [6] have suggested fault tolerance techniques for the recovery of BS faults. BS presumes that a call query message is handled following the recovery. In the event of failures and their recovery, it becomes necessary to measure the mean processing time for a message. The popular schemes of CACs include Haring et al., the fixed guard channel scheme [7], Li et al.’s proposed dynamic CAC scheme [8]. An elaborate CAC survey is being carried out by Cao et al. [9] and Ahmed [10].

The studies in [11] suggested an architecture using several millimeter wave micro BSs to supply user packets, where controls are addressed as a result of flexibility. The efficiency of networks is defined as the probability that a wireless network will satisfy the QoS of users [12]. The proportion of covered area within the spatial domain to the total service area is that where quality of service (QoS) can be satisfied [13]. The monitoring delayed on radio channels in cellular wireless networks was examined by Rubin et al. [14]. They looked at the lag of the forward signaling network from the base station to mobile terminals. The error was examined by them. They also addressed how the signaling lag relates to the size of the area.

For network availability analysis, the SINR model was proposed in [15] in which the likelihood of long blocklength short-lived error decoded in the SINR model was not considered. In addition, it was suggested a semi-analytical SINR method.

III. METHODOLOGY

3.1. System Model

The following shown in Figure 3 are the vital parts of Wireless Cellular Network they are:

- Mobile User (MU)
- Base Station (BS)
- Mobile Switching Center (MSC)

The area roofed by the wireless cellular network is separated into location regions. The neighboring cells and the devoted MSC are each in each region. The subsequent control must be performed under this arrangement.

- The location region generates a location update packet, whenever the MU crosses the frontier, by which the involved MSCs needs to update of the location change of the MU.
- Whenever a call appeal from the MU within in the location region is received by the MSC, paging packets are publicized by the concern BS in the location region to alert the MU for the incoming call. The call setup becomes possible once the acknowledgment of response from the intended MU. The channel is combination of signaling channels and data channels for signaling packets and user data. In the proposed work we are considering Time Division Multiple Access (TDMA) [16] as multiple access method, which gears the core physical structure. Several wireless ATM protocols have agreed TDMA as a most suitable mechanism for QoS guarantee. As shown in Figure 4, the frame length of TDMA protocol is divided into uplink and downlink form MU to BS dynamically, it is appropriate for environment where the request for both the direction changes. Both uplink and downlink consists of user data space for both directions.

Figure 3: Model of Wireless Communication Network
The slotted ALOHA protocol states that, when a signaling packet needs to be transmitted, the MU sends the packet on the control slot first. If the signaling packet is received correctly, the BS must send the message during the next downlink. If there is any collision, BS will not reply so that MU can transmit the frame again.

3.2. Flow and Processing Delay of Signaling Packet

The processing delays at BS and MSC, and at the wired network is due to the transmission delay of experienced by the signaling packets on both uplink and downlink. The following evaluates the signaling packets delay:

- **Originating Call Setup Delay ($T_{MD}$):** Let $T_{MD}$ signify the time duration from call appeal (CA) generated at MU to arrival of call acknowledgment (CAK) at MU as shown in Figure 5a. For calls intended for wired network. The flow of the CA packet is Uplink $\rightarrow$ BS $\rightarrow$ MSC. After processing at wired network the CAK is returned through MSC $\rightarrow$ BS $\rightarrow$ Downlink.

- **Terminating Call Setup Delay ($T_{W}$):** Let $T_{W}$ signify the time duration from the arrival of call appeal at MSC to the time when its reply is sent to the wired network, for calls intended for MU, from MSC to BS paging packets (PG) are broadcasted in its location region, and sent to MU via downlink. Then, the corresponding MU generates CA which will be reverted to wired network shown in Figure 5b.

- **Location Update Delay ($T_{L}$):** The flow of location update appeal $LUA$ is Uplink $\rightarrow$ BS $\rightarrow$ MSC, and corresponding location update acknowledgment $LUAK$ is returned through BS $\rightarrow$ Downlink as shown in Figure 5c.

- **Handoff Delay ($T_{HD}$):** Handoff appeal (HA) are sent through Uplink $\rightarrow$ BS $\rightarrow$ MSC. After request processed at the wired network, the corresponding handoff acknowledgment (HAK) is returned through MSC $\rightarrow$ BS $\rightarrow$ Downlink as shown in Figure 5d.
3.3. Model of Queuing Network

We are developing a queuing network model to analytically detect delays in signaling packets on mobile wireless networks to test the delays in signals. Then we describe every queue in turn.

1. **Uplink Queue**: CA, LUA and HA packets are included in the uplink. CA packets should also be considered in order to respond to PG packets. Such packets’ delays are calculated through a rough analyzes that take into consideration the packet collisions due to the TDMA multimetric access protocol.

2. **Forward Base (FB) Queue**: If the packet is read by BS, the approval (ACK) is returned to the MU sent through the corresponding BS. BS forwarded the packet simultaneously to MSC. The products are managed in a FIFO style.

3. **Forward Mobile (FM) Queue**: MSC manages all incoming BSs packages in the location region. Such packets will be sent to the Wired Network in a FIFO manner.

4. **CA, LUA and HA Queues**: The queues of CA, LUA and HA reflect delays in the respective wired network packets. Packets HA, LUA and CA are expected to be handled separately. The exact representation cannot be known and IS (Infinite Server) queues are treated as specified parameters for the scheduling discipline as well as service requirements.

5. **Backward Mobile (BM) Queue**: The MSC is returning CA, LUA and HA signage packets to BSs. The wired network PG packets are also sent via BS to the corresponding MU. These are treated in a FIFO way.

6. **Backward Base (BB) Queue**: The MSC signaling packets are handled in BS in a FIFO way.

7. **Downlink Queue**: The downlink queue packet delay may be modeled using FIFO queue because the BS Station may actively schedule the wireless transmission.

We present some observations and hypotheses to formulate the problems as follows:

- MU is spread across the network uniformly, so $N_{MU}$ mobile terminals are available for each cell.
- Arrivals of Mu originating and ending requests were accompanied by Poisson’s rate with $\lambda_{MU}, \psi_{MU}$ respectively
- Call time is distributed uniformly by average $1/\mu$ In contrast, before transfer into other cells, the period of MU in the cell remains exponential with a mean $1/h$. In general, the performance evaluation of cellular networks uses this approximation
- The handoff call and the new call can be treated in the same way by entering this approximation because of the property Poisson Arrivals See Time-Averages
- Time of processing in MT, BS, MSC, and wired component is distributed on an exponential basis.
  - $1/\mu_{MU}$ Mean duration of diagnosis at MU.
  - $1/\mu_{BS}$ Mean duration of diagnosis at BS.
  - $1/\mu_{MSC}$ Mean duration of diagnosis at MSC.
  - $1/\mu_{CA}$ Mean Call Setup time.
  - $1/\mu_{HA}$ Mean Handoff processing time.
  - $1/\mu_{LU}$ Mean Location update processing time.
- The delay in propagation of the wired network is considered negligible.
- CA packets are not created by the MU involved in the active call. Therefore, this possibility should be considered when taking into account the arrival of CR packets at the MSC.
In addition, we describe the delay for each component by using the following notations. The symbol $S$ is the type of signaling packets in these notations. However, we have included index $n$ for each cell to calculate the delays, as the frequency of arrival of location updates packets in each cell varies according to the location.

- $D_{RU}(S)$: Transmission delay on uplink.
- $D_{BD}(S)$: Transmission delay on Downlink
- $D_{BU}(S)$: Processing delay at FB
- $D_{DD}(S)$: Processing delay at BB
- $D_{MU}(S)$: Processing delay at FM
- $D_{MD}(S)$: Processing delay at BM
- $D_{NP}(S)$: Processing delay for call setup, location update and handoff at wired network
- $D_{NU}(S)$: Processing delay of call setup at a MU

### Uplink Transmission Delay Analysis:
We consider $\Psi_{CA}$ the uplink rate for the CA packet. The Call request in both directions produces CA packets. In addition, since active MU doesn't generate CA packets, only idle MUs create CA packets. Let $P_{bu}$ be the probability that the MU has the active calls. The average number of idle MUs in the cell is $(1 - P_{bu}) N_{MU}$. Then, the CA arrival rate is given as

$$\Psi_{CA} = (\Psi_{MW} + \Psi_{WM})(1 - P_{bu}) N_{MU}$$

By approximating each MU as M/M/1/1 queuing system $P_{bu}$ is derived

$$P_{bu} = \frac{1 + (\Psi_{MW} + \Psi_{WM})}{\mu}$$

The MU produces an HA packet when the MU with an active call moves into a different cell. The arrival rate of HA packages is therefore specified

$$\Psi_{HA} = h * P_{bu} * N_{NU}$$

On the other hand, it produces the LAU packet, if the idle MU moves into a different location. However, the rate of arrivals of LAU packets in cells varies depending on the location of the cell. Considering $b_n$ to represent small fraction of the length of the boundary of the $n$th cell when $n$th cell is located at the boundary of the location area.

If $n$th cell is not in the region, the MU will not generate the LAU packets in that cell. The $b_n$ coefficient can be calculated by the position of the $n$th unit. For instance, $n = 0$ is for centre, and $n = 1/2$ for other cells, for example. Thus, the arrival rate of LAU packets can be calculated in $n$th ($n = 1, 2, …, N_c$) cell in the field. We've have

$$\Psi_{LUA} = h * (1 - P_{bu}) * N_{MU} * b_n$$

The maximum arrival frequency for packets on the Uplink is given $\Psi_{CA} + \Psi_{HA} + \Psi_{LUA}$. The packet collision should be addressed since ALOHA is used for the Multiple Access Process. Let $P_{col}$ be the probability of a $n$th cell collision. If the mathematical distribution matches the number of retransmits, the average number of transmissions is $(1 - P_{col})$ until the packet is successfully accepted by FB. The packet arrival rate including retransmission is given as

$$\Psi_{n}^{UP} = \frac{1}{(1 - P_{col})} (\Psi_{CA} + \Psi_{HA} + \Psi_{LUA})$$

Let $L_f$ be the length of the frame, and $n_{id}$ is the number of the uplink control slots. By taking into account the packets arrived from the earlier frame for the control slots in the current frame, the probability of packet collision $P_{col}$ is determined as

$$P_{col}^n = 1 - e^{-L_f \Psi_{n}^{UP} / n_{id}}$$

BS will return a ACK in the downlink of the next frame once the control packet has been received. Thus the MU that does not receive the ACK transmits the packet directly into the next control slots. Then the uplink delay is eventually as follows:

$$D_{RU}^n = \frac{P_{col}^n L_f + L_f}{2}$$

### Downlink Transmission Delay Analysis:
When the request for the termination of the MU arrives at MSC, PG packets are sent to all BSs in the region. Therefore, the PG entry rate is received as

$$\Psi_{PG} = \Psi_{WN} (1 - P_{BU}) N_{MU} N_{C}$$

As the CA, HA and LA packets received by MSC generate ACK packages, the arrival rates are obtained as required.

$$\Psi_{CAC} = \Psi_{CA}, \Psi_{HCK} = \Psi_{HA}, \Psi_{LAK} = \Psi_{LA}$$

In fact, if BS receives CA, LA and HA correctly, the ACK packet is returned by BS. ACK packets are then given the arrival rate as

$$\Psi_{ACK} = \Psi_{CA} + \Psi_{HA} + \Psi_{LUA}$$

At $n$th cell we have the following downlink rate:

The number of downlink control slots be $n_{id}$. The downlink queue is approximately modeled as $M/D/1$ queue where

$$\Psi_{DOWN} = \Psi_{PG} + \Psi_{CA} + \Psi_{HA} + \Psi_{LUA} + \Psi_{ACK}$$

signaling packets are transmitted time $L_f / n_{id}$. Then at $n$-th cell we get the downlink delay:

$$…$$
\[ D_{RD}^N = \frac{L_f \psi_{down}^n}{2n_d(n_d - L_f \psi_{down}^n)} + \frac{L_f}{2} \]

**FB Queue Analysis**

The arrival rates of CA, HA, and LUA packets at the FB queue are given by \( \psi_{CA}^N, \psi_{HA}^N, \psi_{LUA}^N \) and \( \psi_{down}^n \), respectively. Priority scheduling can be initiated according to the type of signaling packet, to meet various delay requirements. If this is the case, the processing delay at FB can be interpreted by modeling it as a priority system for non-preemptive programming.

Take into account that higher-priority packets are HA than other signaling packets. By considering the exponential distribution for all signaling packets with same identical processing time is \( \frac{1}{\mu_{CA}} \), the packet delay of HA packets is given as

\[ D_{BU}^N(HA) = \frac{1}{\mu_{BS}} + \frac{R^N}{1 - \rho_1} \]

Lower priority CA and LUA packet delays are also suggested

\[ D_{BU}^N(CA) = D_{BU}^N(LUA) = \frac{1}{\mu_{BS}} + \frac{R^N}{(1 - \rho_1)(1 - \rho_1 - \rho_2^N)} \]

Where \( \rho_1 = \psi_{HA}/\mu_{BS} \),

\[ \rho_1 = \psi_{CA}^N + \psi_{LUA}^N \mu_{BS} \]

and

\[ R = \psi_{CA}^N + \psi_{HA}^N + \psi_{LUA}^N \mu_{BS} \]

**BB Queue Analysis**

In addition to CA, LA, HA and PG packets in the BB queue, ACK packets should be considered. The delay analysis can nevertheless be conducted in a manner similar to the FB queue analysis. Even if priority scheduling is added, it is valid.

**FM Queue Analysis**

The total arrival frequency for the Packet at the FM queue is \( N_c \psi_{CA}, N_c \psi_{HA}^N, \psi_{LUA}^N \) and \( \sum_{N=1}^{N_c} \psi_{LUA}^N \) for all BSs in the same region. Priority scheduling may also be implemented.

**BM Queue Analysis**

The BM packets come from CA, LUA, HA, and PG. Priority scheduling is also allowed.

**Processing Delay Analysis at the Wired Network**

Our assumption is that the average delay of transmission for each type of wired network signaling packet is based on the results of study of the M/M/1 queue. The wired network call setup time is given by

\[ D_{MP}^N(CA) = \frac{1/\mu_{CA}}{1 - N_c \psi_{CA}/\mu_{CA}} \]

**Processing Delay Analysis at Mobile User**

The frequency of entry of PG packets at MU is \( \psi_{WM}^N \), which means we get the average processing time at MU

\[ D_{MU}^N(PG) = \frac{1/\mu_{MU}}{1 - \psi_{WM}/\mu_{MU}} \]

**Derivation of Total Processing Delay**

On the basis of the evaluation of each element, the total average processing delay is eventually taken into account for each type of signaling packet.

- **MU Originating call setup Delay**

Simply summarizing the delay on each element, we can achieve the total average time to process the MT originating call configuration in the n th cell area as follows:

\[ T_{MW}^N = D_{MU}^N(CA) + D_{MB}^N(CA) + D_{MB}^N(CAK) + D_{RU}^N(CR) + D_{BU}^N(CR) + D_{BD}^N(CR) + D_{RD}^N(CA) \]

- **MU Terminating call setup Delay**

Similarly, for the MT termination call setup at the nth cell in the region, we have the total average processing period as follows:

\[ T_{MW}^N = D_{MD}^N(PG) + D_{MT}^N(PG) + D_{MB}^N(CAK) + D_{BD}^N(PG) + D_{BU}^N(PG) + D_{BD}^N(CAK) + D_{RD}^N(CAK) \]

- **Handoff Processing Delay**

Signaling handoffs are distributed through the same route as the call setup packets that begin with MU. Then we have

\[ T_{BA}^N = D_{MU}^N(HA) + D_{MP}^N(HA) + D_{MB}^N(HAK) \]

\[ + D_{RU}^N(HA) + D_{BU}^N(HA) + D_{BD}^N(HAK) + D_{RD}^N(HAK) \]
- Location update processing delay

Furthermore, by applying CR to LUA and CAK to LAK, we can achieve an average delay from Eq. 17

IV. RESULT & DISCUSSIONS

First, we show the case without priority processing. In this case, FIFO discipline processes all sorts of signaling messages on each variable. We show the total average time delay in the region below. For example, the overall average MU call setup time is calculated by

\[ T_{MW} = \frac{1}{N_c} \sum_{n=1}^{N_c} T_{MW}^n \]

The effect of the average processing delays is shown in figure 6a. The figure shows the originating call frequency on each cell to be calculated and increases the number of cells in each position region. As a result, the total source frequency in each region is increased by the size of the place zone. Nevertheless, for limited location regions, storage times are reduced as the figure shows. This is because the load of modified position packages can be reduced by larger location on radio and BS network. But the increase in PG load, which leads towards the bottleneck at MSC, is too wide a location region. It raises the delays in the package.

![Figure 6a: Without Priority Scheduling](image1)

Next, we investigated the effect of priority processing. Next, we carried out priority handling at base stations for the handoff and for the ACK package. That is, the priority over the FB queue was lower than the CA and LUA packets and the priority was higher than the CA and LUA packets in the BB queue and in the ACK packets. The priority analysis in BS was, however, almost unsuccessful, as shown in Figure 6b. The following example shows clearly the impact of priority processing on MSC in Figure 6c. The priority for HA Packets is lower than that of CA and LUA on FM queues, and the priority for HA Packets over CA and LUA Packets is higher. Due to increased delays in delivery of other items, the handoff time can be improved significantly. Nevertheless, if the position area in Figure 6c is small (Nc = 1 to 7), we can see that priority care is not as successful. The explanation is that uplink and downlink delays are prevalent in total processing delays in this area. The results showed that priority processing in MSC is successful because all packets in a position based on the MSC increased with the loading of MSC.

![Figure 6b: Higher Priority to Handoff Calls at Base Station](image2)

![Figure 6c: Higher Priority to Handoff Calls at MSC](image3)

The maximum delay, however, varies irregularly. The structure of the location area leads to the different lengths of the locality boundary. As the position load update in such a cell is strong, the average processing delays are high. On the other hand, the average processing delay is high in a normal structure (as in the case of Nc = 7). We note that it is important to take into account the location area structure in the system design, as a suitable structure exists in order to reduce the location traffic update load to a minimum.
Delay Analysis of Wireless Cellular Networks for Better Qos

V. CONCLUSION

In this paper, we have provided an empirical framework for the analysis of a maximum delay in the transmission of wireless networks signaling packets. We have extracted delays in call set-up, update of location and handoff packages by using our proposed model. In addition, the priority processing was implemented in the design to reduce the processed lag in handoff and its effectiveness was assessed by numerical instances. We also analyzed the delay difference between cells within the site to show how the site region structure influences the processing delays. Finally, we demonstrated the effect on processing delays of the number of terminals. We need to study the bottleneck under different system parameters on future research themes, in order to find out the aspect that dominates all processing delays. In the current paper, while handoff signal packets are simply given higher priorities than the other, it is important, on the basis of the practical delay needs of signaling packages, to examine the more appropriate priority assignment.

REFERENCES


AUTHORS PROFILE

Mr.Satyasrikanth Palle, M.E Assistant Professor, Department of Electronics & Communication Engineering, Sri Venkateshwara College of Engineering, Bangalore 562157

Dr.Shivashankar M.Tech. Ph.D Professor & Head, Department of Electronics & Communication Engineering, Sri Venkateshwara College of Engineering, Bangalore 562157