A Recursive Model for Controlling New Voice Call Congestion on GSM Network Using Priority Queue Discipline and Call Duration Control

*O. E. Oguike; **G. C. E. Mbah; **M. O. Oyesanya
*osondu.oguike@unn.edu.ng, **godwin.mbah@unn.edu.ng, **moses.oyesanya@unn.edu.ng

*Department of Computer Science
**Department of Mathematics
University of Nigeria, Nsukka, Enugu State, Nigeria.

Abstract - Voice call is the most widely subscribed service on the Global System for Mobile communication (GSM) network, as a result, the probability of congestion occurring on the network through voice call is very high. Whenever it occurs, there is a need to control it to provide better quality of service. The most popular scheme for controlling congestion is the call admission control, which defines the condition for rejecting and admitting new voice calls. The condition for rejecting and admitting new calls is based on the congestion parameter ‘New Voice Call Blocking Probability’, NVCBP, i.e. when the NVCBP is above a certain value the network begins to reject new voice calls, otherwise the new voice calls is admitted. The control scheme will help to control the NVCBP. However, the call admission control may lead to starvation, a situation where new voice call is denied access for a long period of time. Therefore, there is need for a new congestion control scheme that will give subscribers fair access into the network. This paper develops a recursive model to control voice call congestion using call duration control scheme, which depends on the NVCBP and subscriber priority level. Among the advantages of the proposed scheme is that it will eliminates starvation and gives value for money for high priority SIM cards, which are more expensive than low priority SIM cards and enjoy better quality of service than low priority SIM cards.

Keywords - new voice call blocking probability, congestion control, GSM network, congestion, recursive model, queuing model, call admission control, priority queue discipline.

I. INTRODUCTION

The GSM technology started in Europe in 1992 by a group called Group Special Mobile. Like a wide fire, it has spread to other parts of the globe. It is now known as Global System for Mobile communication, with the acronym GSM. Global System for Mobile communication, GSM is made up of a network, called GSM network [1]. The GSM network is made up of various functional entities in a hierarchical structure, which provide services to the GSM subscribers and GSM operators for monitoring and controlling the performance of the network [2]. Some of the characteristic features of GSM include the following: it is digital because its signal assumes discrete value at any time, it is a cellular network because it is based on the concept of cell as the smallest geographic area, which its coverage area is divided into. It is a wireless network because the functional entities of the network are connected without wire. It is a radio system because some of the functional entities of the network are connected by radio signal. It is also a mobile system because communication can be made while on motion [2, 3, 4, 5]. The main service that GSM provides to the user is voice related service. However, in the recent time, apart from voice calls, it can provide data communication services, like browsing, email and multimedia [2, 5]. The use of GSM technology as a means of communication offers the following advantages over the analogue mobile technology: clear and sharp sound quality, privacy of calls, handles more calls at a time; availability of roaming service in various countries, economy of acquisition, availability of data services etc [5].

These benefits and advantages have attracted millions of subscribers to the various services of GSM, thereby leading to a situation where many subscribers are competing for the use of limited communication resources, like channels. This problem situation is called congestion [5, 6].

The nodes of the GSM network are the base stations, each base station covers a geographic area, called cell [9, 13]. The base station consists of the base transceiver station, and the base station controller. The base transceiver station consists of radio antenna that connects to the subscriber’s mobile station, using the radio signal [8]. The base stations are also connected to other GSM network devices, called mobile switching centers [7]. The base stations are the grassroots communication devices because they serve as link between the subscribers and the GSM network. Each subscriber will be connected to the closest base station, based on the signal strength. The base stations contain the communication resources, like channels, which are used to setup and transmit calls between two mobile stations. Setting up calls between two mobile stations requires the use of an unused channel in the base station that is closest to the caller. The availability of unused channel at the base
station closest to the caller is one of the factors that determine if a call will be successfully setup or not.

II. PROBLEM STATEMENT AND THE QUEUING NETWORK

A. Problem Statement

Voice service is the oldest and the most widely subscribed service on the GSM network. All users of the GSM network subscribe to this service. Since the number of channels, which are used in the base station to setup calls is limited; therefore, when many subscribers try to setup calls in a particular base station at the same time, if there are insufficient channels, some of the calls will not be successfully setup because of unavailability of channels. This is called congestion or network overload. Congestion is therefore, one of the challenges of the GSM network and it is characterized by degradation in the quality of service, like high new voice call blocking probability, high voice call dropping probability etc. Congestion can be prevented by increasing the number of channels in the base station. It can also be controlled by allowing it to occur, whenever it occurs, the use of congestion control scheme, like call admission control can be used to bring it to an acceptable level. Though it can be argued that prevention is better than cure, but preventing congestion by increasing the number of channels in the base station may not be an ideal solution because the financial resources that will be used to increase the number of channels in the base station is limited.

Furthermore, the use of call admission control as congestion control mechanism means that some callers at various times and various locations can be starved from accessing the GSM network. Therefore, it is better to give every subscriber fair access to the network, by controlling the time duration those callers spend on their calls. This paper solves the problem of congestion, using this novel scheme, call duration control.

B. Queuing Network

The queuing system that will be used to represent the various priority levels of subscribers is the multiple queue, multiple channels/servers queuing system. Each queue will represent a priority level. As calls arrive at the base station, a queue allocator will assign the call to its priority queue. Because calls from any priority level can arrive at random, there is equal probability that an arriving call belongs to any priority level. The new calls in the various priority queues will compete for available channels, which a channel allocator will allocate to the new voice calls. Because the queue discipline in the queue network is a priority queue discipline, therefore the probability that a channel will be allocated to a new voice call in a particular priority queue will vary, it will depend on the priority level of that call. We assume that there are seven different priority levels, as specified in the eMLPP technical specification [12, 13]. Therefore, the probability that a new voice call with priority level \( p \) will be allocated to a channel is given in equation 1.1 as:

\[
\frac{p}{\sum_{j=1}^{7} p}
\]

(1.1)

Figure 1 illustrates the queuing network under consideration.

![Figure 1 Queue network under consideration](image)

Each queue in the queuing network can be described using this Kendall notation:

\[M|\mathcal{M}|k :: FCFS|c|\infty\]

where \( k = c \). The first \( M \) describes the arrival pattern of calls into the base station, which is Markovian, the second \( M \) describes the departure pattern of calls, which is Markovian, the letter \( k \) denotes the number of channels/servers in the base station, while the letter \( c \) denotes the number of new voice calls that can be in the base station. The queue discipline for each of the priority queues is first come first serve, while the infinity symbol denotes an infinite source from where calls can originate from. The condition, \( k = c \) is very important because in this queuing system, there is no queue buffer that will be used to hold calls that could not be allocated channels immediately on arrival, due to unavailability of channels. This means that if a new voice call arrives and there is no available channel, the call will be lost. The reason for this condition is because new voice calls require minimal delay during call setup;
therefore, there is no need for queue buffer during call setup.

III. REVIEW OF RELATED LITERATURE

Statistics from the study in [9] suggested that low performance of GSM network calls, which was characterized by high failure rates for both intra and inter network GSM calls, could be attributed to dearth of vital infrastructure, like efficient and constant power supply and transmission channels. The result from the study also indicated that the call failure rates was higher for inter network calls than for intra network calls. This result from the study is the basis for priority based calls, which will prioritize calls, especially at locations and time when congestion is likely to be high. As a result of using priority based queuing model in this study, the call failure rates for high priority calls will be reduced. It will also reduce the call blocking probabilities for high priority calls.

According to [10], congestion could be caused by lack of adequate infrastructures, inadequate channels. According to the authors, it could be managed by the use of priority based call setup mechanism.

Different algorithms that could be used to manage congestion were proposed by the authors of [11]. Among these algorithms is the use of priority allocation for setting up calls. The authors noted that different levels of priority should be integrated into the SIM cards, which would be determined by the nature of service that the owner of a particular SIM card would use it for. This means that people that make emergency medical calls or emergency security calls or very important personalities in the society, like president, prime ministers, kings and queens will register their SIM cards with high priority. Though this is part of the problems that this research paper aims at solving, the authors did not consider the mathematical queuing models that such priority based call setup method can use. However, they used the Erlang-B formula to model the blocking probability, and analyzed the performance of the Erlang-B model by choosing appropriate dimension of the network as specified in the parameters of the Erlang-B model with the aim of determining the call blocking probability.

Congestion has been identified in the literature as a common problem of the GSM network, which is as a result of limited network resources and increasing subscriber capacity [1, 5, 6]. Though different approaches can be used to handle the congestion problem, among the approaches is the congestion management approach. One of the techniques of congestion management approach is call admission control, which aims at determining the condition when new calls will be admitted into the network or rejected from accessing the network. In order to evaluate a particular call admission control scheme, modeling the blocking probability of GSM calls can be useful because the call blocking probability can be used to define the condition when new calls will be admitted into the network or rejected from accessing the network. During congestion, the blocking probability is expected to be high, while at non congestion period, it is expected to be low. Therefore, the call admission control can set the call admission/rejection condition to a particular value of the call blocking probability, so that whenever the call blocking probability reaches that value, any new call into the network will be rejected.

Seven priority levels were identified by [12, 13], as the number of priority levels on the GSM network. According to them, the first two priority levels were reserved for internal network use, while the remaining five priority levels were reserved for subscribers use. This GSM priority call service is part of the services of the enhanced Multi Level Precedence and Pre-emption (eMLPP) service of the second phase/generation of GSM technology. The provision of this service in phase-2 of the GSM technology means that every GSM subscriber has an associated level of priority. The service has two parts, which are precedence and pre-emption. Ongoing calls with lower priority will be preempted for higher priority calls whenever there is congestion at setup time, or at handover of higher priority call to a congested cell. This service forms the basis for developing mathematical models that is based on priority queuing system in this research paper.

It was remarked in [14] that most of the literature on congestion control focused on call admission control and network resource allocation. Call admission control, according to [14] specified the condition for admitting and rejecting new calls based on the availability of channels and the value of call blocking probability. Part of the problem that this research paper will solve will use priority based queuing model to model various performance metrics and compared them with non priority based queue discipline, like first come first serve discipline. The authors in [14] also proposed an algorithm that was based on ticket scheduling, which could be used to manage congestion.

Queuing approach was used in [15] to develop analytic/mathematical model for call completion in wireless mobile network. According to the authors, arrival of calls in GSM network had a random behavior with poisson distribution, while the overall service time was assumed to be exponentially distributed. A single server M/M/1 queuing model was used to model the queuing of calls that arrived when there was no available channel to allocate to the call. The authors assumed a non-priority based queue with a first in first out queuing discipline and steady states probabilities to develop mathematical model for the call completion rate. The performance of the model was evaluated against some network parameters, like queue size and traffic intensity. Though this research paper will use queuing approach to monitor the performance of GSM base stations, it will assume a different queuing model from the queuing model used in [15]. A parallel server queuing models will be assumed because different calls can be setup at the same time by a base station. A priority based approach will be
used in this paper as the queue discipline, therefore, a priority based, multiple queues, parallel server queuing system will be used in this research paper. The reason for the use of multiple queue, parallel server queuing system is because each queue will be used for a priority level, and more than one new voice call can be setup at the same time. Though the queue discipline is priority queue, we assume a non pre-emptive priority queue discipline. This assumption is necessary because new voice calls are very sensitive to time delay, and should not be pre-empted by higher priority calls. We also assume that arrival of new voice call occurs at random and termination of new voice calls occur at random. Since resources are scarce, we assume that the number of channels in the base station is finite. Since frequency/channel reuse technique can be used to increase the logical number of channels that a base station can use, we assume that the total number of channels that is available for use by a base station is the sum of the total physical channels and the total logical channels.

IV. DEVELOPMENT OF THE RECURSIVE MODELS

In this paper, we develop the following models:

A. Non Recursive Model for Erlang-B Blocking Probability

Suppose \( k \) is the number of channels in the base station and \( \lambda_j \) is the arrival rate when there are \( j \) calls in the base station, while \( \mu_j \) is the departure rate when there are \( j \) calls in the base station. The arrival and departure rates into the base station can be defined as:

\[
\lambda_j = \begin{cases} 
\lambda, & 0 \leq j \leq k - 1 \\
0, & \text{otherwise}
\end{cases}
\]

(1)

\[
\mu_j = \begin{cases} 
 j \mu, & 1 \leq j \leq k \\
0, & \text{otherwise}
\end{cases}
\]

(2)

A generalized model has been developed in [16], which is stated in equation (1.4.3) as the steady state probability that \( j \) calls are in the base station:

\[
P_j = \begin{cases} 
P_0, & j = 0 \\
\frac{\lambda^j}{j!} \frac{1}{\mu} P_0, & 1 \leq j \leq k
\end{cases}
\]

(3)

Equation (1.4.3) is the generalized model for the steady state probabilities. \( P_0 \) can be obtained as we sum all the probabilities and equate it to 1, i.e.

\[
\sum_{j=0}^{k} P_j = 1
\]

(4)

Using equations (1) and (2) in equation (3), we obtain the following:

\[
P_j = \begin{cases} 
P_0, & j = 0 \\
\frac{\lambda^j}{j!} \frac{1}{\mu} P_0, & 1 \leq j \leq k
\end{cases}
\]

(5)

Further simplification of equation (1.4.5), gives the following:

\[
P_j = \begin{cases} 
P_0, & j = 0 \\
\rho^j \frac{1}{j!} P_0, & 1 \leq j \leq k
\end{cases}
\]

(6)

Substituting the utilization factor or traffic intensity, which is given as \( \rho = \frac{\lambda}{\mu} \) in equation (6), we obtain the following:

\[
P_j = \begin{cases} 
P_0, & j = 0 \\
\rho^j \frac{1}{j!} P_0, & 1 \leq j \leq k
\end{cases}
\]

(7)

Using equation (7) in equation (4), we obtain the following:

\[
\sum_{j=0}^{k} P_j = \sum_{j=0}^{k} \frac{\rho^j}{j!} P_0 = 1
\]

(8)

Solving for \( P_0 \) in equation (8), we obtain the following:

\[
P_0 = \frac{1}{\sum_{j=0}^{k} \frac{\rho^j}{j!}}
\]

(9)

Using equation (9) in equation (7), we obtain the following:

\[
P_j = \begin{cases} 
1, & j = 0 \\
\frac{1}{\sum_{j=0}^{k} \frac{\rho^j}{j!}} \frac{\rho^j}{j!}, & 1 \leq j \leq k
\end{cases}
\]

(10)

From mathematical analysis, if the number of channels, \( k \) is very large, and the utilization factor, \( \rho \) is less than 1, the expression in equation (10) can be written as:

\[
\sum_{j=0}^{k} \frac{\rho^j}{j!} \approx e^{\rho}
\]

(11)

Therefore, using equation (11) in equation (10), we obtain the following:
The blocking probability is defined when \( j = k \), therefore, equation (12) can be written as:

\[
P_j = B(k, \rho) = \begin{cases} 
\frac{1}{e^\rho}, & j = 0 \\
\frac{\rho^j}{j! e^\rho}, & 1 \leq j \leq k
\end{cases}
\]  

Equation (13) is the Erlang-B loss function, which is the probability that new call will not be admitted into the base station because of unavailability of channels in the base station. It will be used to measure the level of new voice call congestion.

The problem with equation (13) is that it is an approximation; it is valid only when the approximation conditions hold. Therefore, we seek for an exact recursive solution of the Erlang-B formula.

### B. Recursive Model for the Erlang-B Blocking Probability

Two approaches will be used to develop the recursive model for the series in equation (11). We begin with the first approach, which will add the \( k \)th term recursively as shown below in equation (14).

\[
A(k) = \begin{cases} 
1, & k = 0 \\
\frac{\rho^k}{k!} + A(k - 1), & k \neq 0
\end{cases}
\]  

Therefore, using (14) in equation (13), we obtain the following:

\[
P_k = B(k, \rho) = \begin{cases} 
\frac{1}{A(k)}, & k = 0 \\
\frac{\rho^k}{k! A(k)}, & k \geq 1
\end{cases}
\]  

The second approach to obtaining the recursive model for the Erlang-B blocking probability will evaluate the series in equation (11) by determining the \( k \)th term of the series recursively, as shown below in equation (16).

\[
T(k) = \begin{cases} 
1, & k = 0 \\
\frac{\rho}{k} T(k - 1), & k \neq 0
\end{cases}
\]  

Therefore, we can rewrite equation (11) as:

\[
D(k) = \begin{cases} 
1, & k = 0 \\
T(k) + D(k - 1), & k \neq 0
\end{cases}
\]

Using equation (17) in equation (11), we obtain the following:

\[
D(k) = \sum_{j=0}^{k} \frac{\rho^j}{j!}
\]

Therefore, we write equation (1.4.13) as:

\[
P_k = B(k, \rho) = \begin{cases} 
\frac{1}{D(k)}, & k = 0 \\
\frac{\rho^k}{k! D(k)}, & k \geq 1
\end{cases}
\]

Equations (15) and (19) are two recursive models for the new voice call blocking probability for the \( p \)th priority queueing system, while equation (13) is the non-recursive model for the new voice call blocking probability for the \( p \)th priority queueing system.

Since Erlang-B new voice call blocking probability for the \( p \)th priority queue is identically independent distributed, therefore, the joint new voice call blocking probability, which is the probability that new voice call will be lost in all the priority queues can be obtained by using the multiplication law of probability. When we do that we obtain the following model:

\[
P_{\text{joint}}^{k} = \left\{ \begin{array}{ll}
\prod_{p=1}^{\gamma} \left( \frac{1}{D_p(k)} \right), & k = 0 \\
\prod_{p=1}^{\gamma} \left( \frac{\rho_p^k}{k! D_p(k)} \right), & k \geq 1
\end{array} \right.
\]

### C. Joint New Voice Call Duration

This is a novel model that will be used to estimate the duration of voice call, based on the new voice call blocking probability. By determining the duration of new voice call, the departure rate of new voice call will increase; as a result, the new voice call blocking probability will fall. In order to model the joint new call duration, we establish a novel relationship between the joint new voice call blocking probability and the joint new voice call duration. This relationship states that the joint new voice call blocking probability is inversely proportional to the joint new voice call duration. This relationship can be stated mathematically as:

\[
P_{\text{joint}}^{\text{int}} \alpha \frac{1}{C_{\text{joint}}^{\text{int}}}
\]

Removing the proportionality sign, we obtain the following:
where:

\[
P_{joint} = \frac{\sum_{j=1}^{n} \lambda_{admit,j}}{C_{joint}}
\]

\(22\)

Where:

\[
\sum_{j=1}^{n} \lambda_{admit,j}
\]

is the constant of proportionality,

and \(\lambda_{admit,j}\) is defined as the rate of arrival of new voice calls for the \(p^{th}\) priority queuing system that were admitted into the network because of availability of channels, which can be called, rate of admittance of new voice call for the \(p^{th}\) priority queue.

D. Rate of Admittance of New Voice Call

New calls that arrive can either be admitted into the network or be rejected (lost) on arrival, depending on the availability of channels. The rate of arrival of calls that are admitted into the network is an important parameter because it can be used to model other performance metrics. Therefore, we can say that the rate of arrival of calls is the sum of the rate of arrival of calls that are admitted into the network and the rate of arrival of calls that are lost. This can be written as follows:

\[
\lambda = \lambda_{admit} + \lambda_{lost}
\]

\(23\)

From equation (1.4.23), we obtain the following:

\[
\lambda_{admit} = \lambda - \lambda_{lost}
\]

\(24\)

The rate of arrival of calls that are lost can be obtained by multiplying the rate of arrival of call by the blocking probability (Erlang B) loss probability. Therefore, equation (24) can be expressed as:

\[
\lambda_{admit} = \lambda - \lambda(B(k, \rho))
\]

\(25\)

Simplifying equation (25), we obtain the following:

\[
\lambda_{admit} = \lambda(1 - B(k, \rho))
\]

\(26\)

V. RESULTS AND DISCUSSION OF THE RECURSIVE MODELS SIMULATION

Java programming language was used to simulate the recursive models as recursive Java methods. The variables that determine the joint call blocking probability, and the joint call durations are arrival rate, departure rate and the number of channels in the base station. Therefore, by varying the departure rate and keeping other variables constant, we can discover how the joint new voice call blocking probability and the joint new voice call duration will behave, for priority queue discipline and first come, first serve. A test program was written, and one of the variables was varied, while other variables were kept constant with the aim of obtaining the result of any of these independent variable, call blocking probability or call duration. Excel spreadsheet was used to analyze the results of the simulation.

Figure 2a shows the comparative analysis between priority queue discipline and first come first serve queue discipline, for joint new voice call blocking probability against arrival rate. It can be observed from the result that as the arrival rate of new voice calls increases, the joint new voice call blocking probability, for both priority discipline and first come first serve queue discipline increase. However, for any value of the arrival rate, the joint new voice call blocking probability for first come first serve queue discipline is always higher than the joint new voice call blocking probability for priority queue discipline. This comparison shows that priority queue discipline performs better than the first come first serve queue discipline. This result means that operators of GSM should watch for factors, like increase in the population of an area, which can lead to increase in the arrival rate of new voice call. The presence of such factor should make them to adopt the congestion prevention scheme or the congestion control scheme, which will help to reduce the new voice call blocking probability.

Figure 2a. Comparative Analysis Between Priority Discipline and FCFS for Joint New Voice Call Blocking Probability Against Arrival Rate
In figure 2b, as the arrival rate of new voice call increases, the joint new voice call duration decreases. This result means that when the call duration control scheme is used to control congestion, subscribers should watch for factors that can lead to increase in the arrival rate of new voice call, like increase in the population of an area, and time of the day. The presence and absence of the factors will guide subscribers when to make call, with the aim of obtaining long call duration.

Figure 2c shows how the joint new voice call blocking probability behaves at various hours of the day. It shows that as the hour of the day increases from the morning to afternoon hours, the joint new voice call blocking probability increases, but as the hour of the day increases from the afternoon hours to night hours, the joint new voice call blocking probability decreases. This result is consistent with real life situation. It means that subscribers can obtain better network performance i.e. low new voice call blocking probability in the early morning hours or late night hours.

Figure 2d shows the behavior of the new voice call blocking probability for the various priority levels. It can be seen that while all variables remain constant, as the priority level increases, the new voice call blocking probability falls. The result means that subscribers that desire good network performance, like low new voice call blocking probability should obtain high priority level SIM cards.

In figure 2e, we see how the joint new voice duration behaves at different hours of the day. It can be observed that as the hour of the day increases from the morning hours to the afternoon hours, the joint call duration decreases, but as the hour of the day increases from the afternoon hours to the night hours, the joint new voice call duration increases. This result will guide subscribers to know when to make calls with the aim of obtaining high call duration.
Figure 2f shows comparative analysis between priority discipline and FCFS queue discipline, for joint new voice call blocking probability against departure rate. The result shows that as the departure rate of calls increases, the joint new voice call blocking probability decreases, for both priority queue discipline and first come first serve queue discipline. This result can be used to explain the effect of setting time duration on subscriber’s calls. Setting time duration on subscriber’s call is expected to increase the departure rate of new voice calls, which will reduce the joint new voice call blocking probability. From the result, we observe that the priority queue discipline performs better than the first come first serve queue discipline; this is because for any value of the departure rate the joint new voice call blocking probability for priority queue discipline is always less than the joint new voice call blocking probability for first come first serve queue discipline.

The results of figures 2g and 2h are consistent with real life result.

VI. CONCLUSION

The paper has been able to develop recursive models for measuring congestion parameter of the GSM network, called new voice call blocking probability, using non pre-emptive queue discipline. Based on the value of the new voice call blocking probability, it develops another novel model, called new voice call duration, which can be used to control congestion on the GSM network. If the new voice call blocking probability is high, the new voice call duration will be low, leading to high departure rate, which will reduce the new voice call blocking probability, thereby controlling congestion.

REFERENCES


[9] Adeseko Ayeni, “Quality of Service Profile of the Nigerian GSM Networks”, Department of Electrical Engineering, University of Ilorin


[12] Digital cellular telecommunications system (Phase 2+); enhanced Multi-Level Precedence and Pre-emption service (eMLPP) - Stage 1 (GSM 02.67) ETSI European Telecommunications Standards Institute

[13] Digital cellular telecommunications system (Phase 2+); enhanced Multi-Level Precedence and Pre-emption service (eMLPP) - Stage 2 (GSM 03.67) ETSI European Telecommunications Standards Institute.

