

## Performance evaluation of Push-to-Talk Application over LTE Networks

Hisham Zorghani\*, Ivica Kostanic\*\*

\*Department of Electrical and Computer Engineering, Florida Institute of Technology, Florida, USA

\*\* Department of Electrical and Computer Engineering, Florida Institute of Technology, Florida, USA

Corresponding author: Hisham Zorghani

### ABSTRACT

Public safety has been using narrowband trunking systems for decades. These systems were designed with only voice calls in mind. Therefore, their data capabilities are limited. Long Term Evolution LTE has been selected as the platform for next generation public safety networks. It can support not only push-to-talk PTT but also push-to-multimedia. This paper applies queueing theory to evaluate the performance of PTT application over LTE networks. A proper queueing model based on PTT over LTE system is analyzed and implemented in SimEvents simulation. The results of the proposed mathematical model are compared with the simulation results for validation. Furthermore, since Erlang B formula is a general model, the simulation has been adopted to evaluate the system call blocking probability when a different type of calls such as *push-to-send-picture* is used in the system.

**Keywords**-Push-to-Talk PTT, next generation Public safety networks, call blocking probability

Date of Submission: 07-05-2018

Date of acceptance: 22-05-2018

### I. INTRODUCTION

A Land Mobile Radio (LMR) system is a narrow band Trunking Communication technology standard that is widely used in emergency services. This is due to its robustness in harsh environment and also its rich feature set that includes Push-to-Talk, Group Calling and Direct Communication mode. Even though the LMR standards have been evolved for many years, their data capability is still limited. [1], [2] Recently, migrating LMR systems into LTE system received much attention in academia, industry and standardization bodies. LTE technology standard has emerged as the broadband wireless communication technology that can transform Public Safety Communication (PSC) by providing broadband mission-critical voice and data [3]. Grade of Service (GoS) is a measure of traffic congestion in telephone networks that results in lost calls. According to public safety service requirements, call blocking probability must be very low. To mathematically investigate the performance of PTT over LTE network, a model of the system must be formulated. This may be done by using a branch of mathematics known as queueing theory [4]. PTT over LTE networks consists of many PTT users that can be modeled as customers who randomly arrive to a service station (eNodeB) to be served. Source population or number of subscribers (finite/infinite), probability distribution of the time between arrivals, the probability distribution of the time needed to serve

a customer, capacity of the queue (lossless/lossy), number of servers and queueing policy (e.g., first come first served) are attributes of the queueing model. PTT over LTE systems have characteristics that distinguish them from regular telephone networks such as one-to-many type of conversations. Therefore, Erlang-B formula does not adequately describe the actual PTT over LTE scenario and it cannot be used to calculate the call blocking probability [5]. In simple models, one could evaluate system performance by using only analytical means to arrive at close-form expression. However, in slightly more complicated models, close-expression may not be achievable [6].

This paper considers more realistic assumptions in the model process. It describes PTT over LTE by assuming a finite population size and exponential distributed call holding time with average of 3 seconds. Furthermore, with the plenty of bandwidth that LTE offers, other application, such as push-to-multimedia will be implemented in the next generation public safety networks. This different type of traffic has no longer exponential distribution in terms of call holding time. Since the assumption of Erlang B formula is so idealized, simulation can often be helpful. The advantage of simulation is its generality, it has no restrictions on the probability distributions involved. Also, with enough computer resources, very complicated models may be simulated, and performance measure of interest may be estimated to any desired

level of accuracy. Push-to-talk (voice) and push-to-send-picture (data) have been implemented in the simulation model to evaluate its performance in terms of call blocking probability. Finally, some implementations in channel management to the model have been introduced to improve its performance.

## II. SYSTEM DESCRIPTION

Fig. 1 shows a systematic representation of a queueing system. Source population consists of all users that are eligible for service in a given queueing system. In general, the most important property of the source population is its size. The arrival rate is one of the variables used to quantify the volume of generated traffic. The standard way to specify the arrival rate is through distribution of interarrival times. The server is a part of the queueing system capable of performing a service task. Service time is the period of time over which a server is allocated to an individual user. There are several different algorithms commonly referred to as the queueing discipline. The most common algorithm is First Come – First Serve FCFS which sometimes referred to as FIFO. Other examples of queueing discipline are Last In – First Out and priority queue.

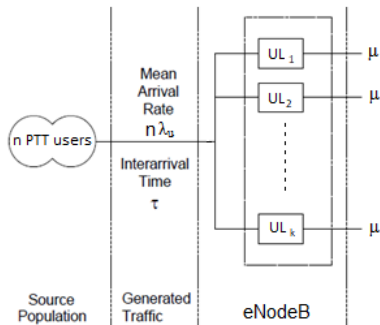


Fig.1 Systematic representation of a queueing system

2.1 Queueing model M/M/k/k/n (voice only)  
 To facilitate the analysis of queueing model, some popular symbols are defined in Table 1.

TABLE 1 Models parameters and their description

Symbol	Description
M	Poisson process arrivals
M	Exponent. dist. service time
k	Number of servers (uplinks)
k	System capacity
n	Population size
$\lambda_i$	$i^{th}$ system state arrival rate
$\lambda_u$	Arrival rate per PTT user
$\mu_i$	$i^{th}$ system state service rate
$\mu$	Service rate
$\rho_i$	$i^{th}$ system state offered

	traffic = $\lambda_i/\mu_i$
$P_i$	Probability of system state i

M/M/k/k/n model state transition diagram is illustrated in Fig. 2. Birth and Death process is used as a mathematical model for a queueing system. The framework of the process allows us to derive some results that describe the behavior of the queueing system.

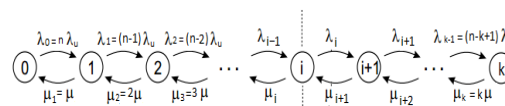


Fig. 2 M/M/k/k/n model state transition diagram  
 Where,

$$\lambda_i = (n - i)\lambda_u, \quad 0 \leq i < k \quad (1)$$

$$\mu_i = i\mu, \quad 0 < i \leq k \quad (2)$$

The equilibrium probabilities of the Birth-Death process can be used to develop the system states probabilities. In equilibrium, according to the transition diagram shown in Fig.2

$$n\lambda_u P_0 = \mu P_1 \quad (3)$$

$$(\lambda_{i-1} + \mu_{i+1})P_i = \mu_i P_{i-1} + \lambda_i P_{i+1} \quad (4)$$

$$(n - k + 1)\lambda_u P_{k-1} = k\mu P_k \quad (5)$$

From (1),

$$P_1 = (n\lambda_u/\mu)P_0 \quad (6)$$

Following the same approach,  $P_i$  maybe expressed as,

$$P_i = (n\lambda_u/\mu)((n - 1)\lambda_u/2\mu) \dots P_0 \quad (7)$$

$P_0$  probability is determined by using the normalization condition.

Let  $\lambda_u/\mu$  (offered traffic) be denoted as  $\rho$ . Then

$$P_i = \frac{\binom{n}{i} \rho^i}{\sum_{i=0}^k \binom{n}{i} \rho^i} \quad (8)$$

Where,

$$\binom{n}{i} = \frac{n!}{i!(n - i)!}$$

Because there are no waiting rooms in this model (lossy system) to hold calls that arrive when all k servers are busy, the blocking probability (BP) is the probability ( $P_k$ ) of having k PTT UEs in the system.

2.2 Model M/M/K/K (voice and data)

In this scenario, the model will be modified to address two classes of calls, voice and data. Call holding time for both classes follow exponential distribution and every transmission will occupy a single server. The source population  $n$  will be considered infinite to facilitate the mathematical analysis. Fig. 3 illustrates the two-dimensional state transition diagram of M/M/k/k lossy system [4]. In Fig. 3,  $\lambda_1$  and  $\lambda_2$  are arrival rates of calls of class 1 and class 2 respectively and  $\mu_1$  and  $\mu_2$  are service rates of calls of class 1 and class 2 respectively. The state of the system can be defined by a combination of the two integer numbers; denoted by  $(i, j)$ , which represent the number of calls of class 1 and the number of calls of class 2 respectively. The total number of system states is  $(k+1)(k+2)/2$ . The  $P_{i,j}$  is the probability that there are  $i$  calls of class 1 and  $j$  calls of class 2 in the system. The state space  $\Omega$  is finite

$$\Omega = \{(i, j): i + j \leq k, i \geq 0, j \geq 0\}$$

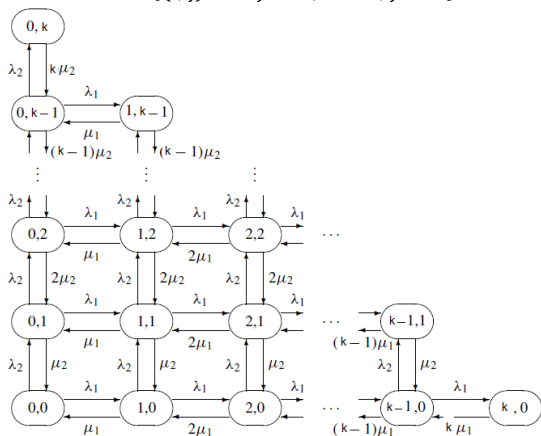


Fig. 3 State transition diagram for M/M/k/k lossy system with two call classes

According to the state transition diagram shown in Fig. 3, the balance equations

$$i\mu_1 P_{i,j} = \lambda_1 P_{i-1,j}, \quad 1 \leq i \leq k, 0 \leq j \leq k \quad (9)$$

$$j\mu_2 P_{i,j} = \lambda_2 P_{i,j-1}, \quad 0 \leq i \leq k, 1 \leq j \leq k \quad (10)$$

are satisfied by the product-form solution [4]

$$P_{i,j} = \frac{1}{G(k)} \frac{\rho_1^i \rho_2^j}{i! j!}, \quad (i, j) \in \Omega \quad (11)$$

Where,

$\rho_1 = \lambda_1/\mu_1$  is the offered traffic of class 1 (voice)

$\rho_2 = \lambda_2/\mu_2$  is the offered traffic of class 2 (data)

and  $G(k)$  is a constant that can be found by using normalization condition. The state probability becomes

$$P_{i,j} = \frac{\rho_1^i \rho_2^j}{\sum_{l=0}^k \frac{(\rho_1 + \rho_2)^l}{l!}} \quad (i, j) \in \Omega \quad (12)$$

2.3 Model M/G/K/K/n (voice and data)

In this scenario, the model of previous section is considered with finite source population. The second letter in Kendall notation  $G$  is suggested to refer to that Call Holding Time for voice and data calls follow General distributions. To analyze this model, we need to implement its two-dimensions state transition diagram. Fig. 4 shows state transition diagram for a system of 3 servers. Each PTT user generates voice calls with rate of  $\lambda_1$  and data calls with rate  $\lambda_2$ . Given such a queueing model, to evaluate its call blocking probability, one would like to use only analytical means, such as probability theory to arrive at closed-form expression. For simple models this is possible, but in more complicated models, a closed-form expression may not be obtainable. However, the results may still be found by solving system state transition diagram. [6]

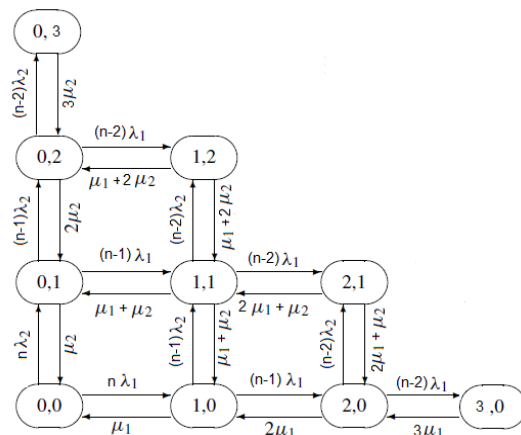


Fig. 4 State transition diagram of model M/G/3/3/n

Clearly, as the available uplinks in one cell site increases, the analytical or numerical calculation may become cumbersome or even impossible. In such situations, simulation can often be helpful. To get more accurate estimate of the quantity of interest, many different and independent samples should be applied to the model. The more samples are simulated, the more accurate the estimate of the quantity of interest that can be calculated by taking the average of all results of the simulation runs.

III. SIMULATION ASSUMPTIONS AND RESULTS

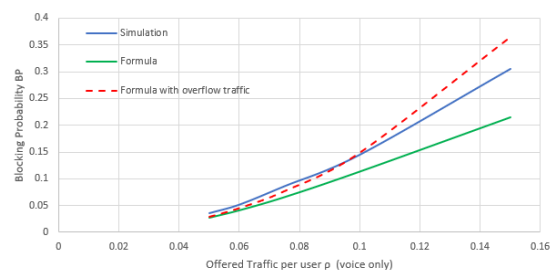
3.1 Simulation Assumptions

In PTT over LTE systems, each cell site has several predefined user and control radio channels. The number of available user channels in each cell determines the maximum number of simultaneous PTT calls that can be handled by that

specific cell. Each PTT call occupies one user channel. In tradition Trunking systems, there are two approaches of channel management (Transmission and Messaging Trunking). In Transmission Trunking, each radio transmission is treated as one call (the call ends and the channel is released when the user releases PTT button). However, in Messaging Trunking, the radio channel will be kept among one PTT group members for the whole conversation session. Transmission Trunking is 20-25% more efficient than Messaging Trunking. [7] In this study, Transmission Trunking behavior is adopted and implemented in SimEvents simulation. PTT call can be either a one-to-one call or a group call (one-to-many). In case of a group call, one channel must be available in each cell site that members of that specific group utilize during the time of the call. Otherwise, the call will be blocked. PTT user mobility and handover are also concerns. However, since PTT calls are relatively short (about 3.8sec), the chance of handover to happen is too low. In SimEvents simulation, one cell site with 3 user channels and 15 PTT users in 3 groups are implemented. Each PTT user generates a voice PTT calls of rate  $\lambda_1$  (Exponential CHT) and data PTT calls of rate  $\lambda_2$  (Uniform CHT). The interarrival time is exponential meaning arrival process is Poisson process. The average call holding time  $T_s=3$  seconds for voice and  $T_s=9$  seconds for data. The simulation run time is half hour(1800sec) executed five times with different initial seeds of each user at every run to ensure unrelated data. Moreover, during each simulation run, hundreds of results have been collected and averaged.

### 3.2 Simulation Results

Fig.4 shows a comparison between calculated call blocking probability BP of PTT over LTE system (voice callsonly) by using formula (8) and simulation results. It is notable that the formula underestimates the blocking probability of the system. However, by modeling the overflow traffic, and considering it in the model, it gives more accurate blocking probability. Overflow traffic is the traffic that blocked by the system calculated by using the blocking probability of formula 8. Therefore, the actual traffic of the system is the overflow traffic and the original traffic [5].

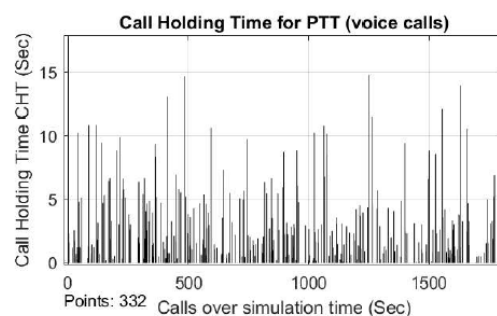


**Fig. 4** Call BP Comparison between simulation and mathematical model

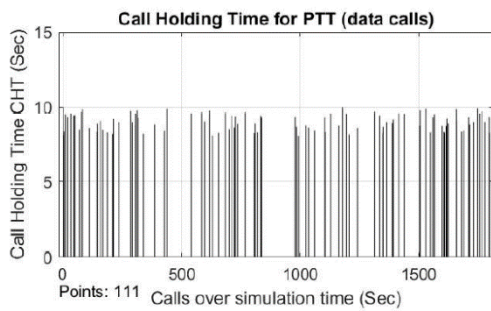
Table 2 shows some results of the simulation when users generate voice PTT calls with Exponential distributed CHT of average 3 seconds and data PTT calls with Uniform distributed CHT (min=8sec, max=10sec). Fig. 5 and Fig. 6 show the CHT of voice and data PTT calls generated by all user.

**TABLE 2** Results of the simulation when  $\rho = 0.075$

offered Traffic per user $\rho = 0.075$			
User Id	PTT voice	PTT data	Total per user
1	22	8	30
2	20	5	25
3	32	9	41
4	28	6	34
5	33	8	41
6	14	7	21
7	20	9	29
8	27	4	31
9	20	5	25
10	18	9	27
11	20	4	24
12	21	8	29
13	22	12	34
14	16	10	26
15	19	7	26
	332	111	443

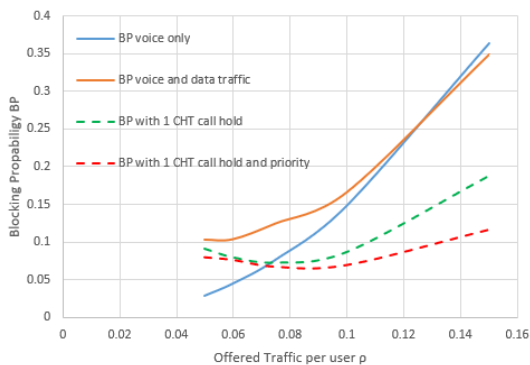


**Fig. 5** Call Holding Time CHT for voice PTT calls



**Fig. 6** Call Holding Time CHT for data PTT calls

Fig. 7 shows that with low traffic ( $\rho < 0.12$  Erlang) generated by PTT users, the system with PTT voice users only has better performance than PTT system with voice and data calls in terms of call blocking probability. For example, at ( $\rho = 0.05$  E) the system blocking probability is 4.8% for voice only and is 10% for voice and data. However, with high traffic load the performance is almost the same. The system is greatly enhanced by implementing holding calls for a period of one average CHT. Meaning that requests arrive to the system when all links are busy can be kept for a while before sending a busy tone. More enhancement to the model can be introduced by applying priority to voice PTT calls. For example, when generated traffic is 0.15 Erlang, call blocking probability of the system is 15.8% without giving priority to voice calls and 12% with voice priority.



**Fig. 7** Simulation results of Call BP in different scenarios

#### IV. CONCLUSION

In this paper, PTT over LTE system is modelled mathematically and simulated to evaluate its performance in terms of grade of service. The mathematical model is derived by using queueing theory with more appropriate assumptions that describes the system more accurately. The mathematical model results are compared with simulation results. In addition, SimEvents simulation is used for analysis of the system when different type of calls (voice and data) are deployed. Results show that by implementing holding requests for one average call holding time and prioritization of voice PTT calls, system call blocking probability is greatly enhanced.

#### REFERENCES

- [1]. A. Kumbhar, F. Koohifar, I. Guvenc, and B. Mueller, "A Survey on Legacy and Emerging Technologies for Public Safety Communications," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 97–124, 2017.
- [2]. T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 106–112, 2013.
- [3]. A. Raza, "LTE network strategy for Smart City Public Safety," 2016 *IEEE Int. Conf. Emerg. Technol. Innov. Bus. Pract. Transform. Soc.*, pp. 34–37, 2016.
- [4]. H. Takagi, "Derivation of Formulas by Queuing Theory," *Spectr. Requir. Plan. Wirel. Comm. Methodol. IMT-Advanced*, pp. 199–218, 2008.
- [5]. L. Ma, S. Wei, and W. Qiang, "A novel traffic analysis method for PoC over LTE based on retrial calling model," 2011 6th *Int. ICST Conf. Comm. Netw. China, CHINACOM 2011*, 2011.
- [6]. Pieter Tjerk de Boer, *analysis and efficient simulation of queueing models of telecommunication systems* (P.T. de Boer, Enschede, The Netherlands, 2000)
- [7]. Božidar Vujičić, *Modeling and characterization of traffic in a public safety wireless network*, master thesis, University of Montenegro, 1997

Hisham Zorghani "Performance evaluation of Push-to-Talk Application over LTE Networks" *International Journal of Engineering Research and Applications (IJERA)*, vol. 8, no.5, 2018, pp.88-92