



AN INTEGRATED APPROACH TO SOLVING RETRIAL QUEUE COMPLEXITY

C. L. Anioke^{1,*}, O. C. Nnamani² and P. A. Nwakamma³

^{1,2,3}DEPARTMENT OF ELECTRONIC ENGINEERING, UNIVERSITY OF NIGERIA NSUKKA, ENUGU STATE, NIGERIA

E-mail addresses: ¹ chidera.anioke@unn.edu.ng, ² obinna.nnamani@unn.edu.ng,

³ akachi.nwakamma@unn.edu.ng

ABSTRACT

Retrial queues have been proposed in literature as a means of managing lost traffic in wireless cellular networks. However, the retrial queue system has led to a more complex Markovic process in network analysis. This complexity is further worsened by the addition of handover prioritization. In this paper, a solution that can minimize this complexity has been studied. The solution involves an integration of computational and analytical modeling techniques. The work was simulated and tested in Matlab/Simulink with guard channels as a handover prioritization scheme. Further evaluations were carried out on the effect of blocking probability (PB) as a quality of service (QoS) parameter. Numerical results obtained from the integrated approach show that the retrial queue reduces the PB for all requests (new and handover). The result obtained from this paper simplifies the solution to the complexity found in retrial queue systems. Furthermore, it was also deduced that reduced retrial rate with corresponding increase in the retrial queue size improves the network quality without an increase in system complexity.

Key words: Markov, retrial, probability, network, traffic

1. INTRODUCTION

The wireless cellular system is affected by an exponential growth of subscribers. It is therefore imperative for the communication industry to provide better coverage and improved network capacity as well as an effective call/request control procedure to optimize network usage [1, 2]. As a result of the limited frequency spectrum and a limited number of channels per coverage area, the cellular concept was introduced to solve the problem of spectral congestion and user capacity [1]. However, this concept has over time faced challenges due to cell overlap and an increased traffic volume. This has resulted to an increase in the number of requests that are dropped or blocked while a communication is in progress or while trying to gain access to the network. Thus subscribers are unsatisfied with the quality of service (QoS) rendered by the network operator. Customer satisfaction can only be achieved if standard QoS is maintained at the network [1, 3, 4]. QoS comprise the speech quality as well as the availability and reliability of service within a given coverage area. However, with an increase in traffic intensity, QoS depends more on the uncommitted network resources and on the teletraffic engineering of the network. Standard QoS can be maintained if the

number of ongoing requests dropped while being handed over from one cell to another is minimized or insignificant. These requests are called handover requests. Therefore the network must have a very low handover call dropping probability. On the other hand, the incoming or new requests arriving at the network may be blocked if there are no free/common network resources available at their time of arrival. These requests are termed new requests. Therefore the new call blocking probability must also be made to be as small as possible.

1.1 The Retrial Queue

Retrial queues in wireless cellular networks have been studied for different telecommunication networks and computer systems. The retrial queue is a special type of queueing system that takes into account the retrial phenomenon whereby arriving requests that find all the channels in a cell occupied or busy tend to retry after a random interval of time, following a random order or the first-in-first-out (FIFO) order [6, 10, 20]. The retrial queue model is found in various practical situations such as in retail shopping where a customer who finds a long waiting line may wish to do something else to return at a later time with the hope that the

* Corresponding author, tel: +234 – 803 – 870 – 2995

queue dissolves [16]. As a matter of fact, the assumption that customers who leave the system as a result of inadequate channel resource are lost is just a first order approximation to a real situation. Usually such a customer returns to the system after a random time interval and tries to get service again [6, 16]. Retrial queues have been shown to improve network quality and customer satisfaction [7, 9, 10, 12]. The retrial queue unlike other classical queuing methods is designed to retry all requests whether new or handover that have been denied access to a channel resource. These requests are blocked or dropped leading to a waste of limited network resources. This means that such requests cannot be served. To minimize such wastes, the retrial queue allows requests in danger of being blocked to try to get access to the resource stream once again without leaving the network. Retrials could take place after a random time and for a given time interval. In this paper, only the first retrials have been accepted after which the request could be dropped if not served. The reason being that after the first retrial, subsequent retrials if allowed will cause congestion in the network since the same limited resources serve new requests.

1.2 Resource Reservation

A resource reservation scheme referred to as the guard channel (GC) scheme has also been added in the network design. GCs are used for prioritizing handover requests over new requests [10]. The GCs are resources that have been reserved for handover requests only. The need for GCs arises because handover requests are active requests. This means that they are already using up a channel resource. Active call drops lead to waste of limited resources. It is therefore in the best interest of the network operator and customers for dropped handover requests to be avoided to a great extent. In addition, new calls should be treated fairly and not to be entirely blocked from gaining access to the network as a result of the lack of free channels when the handover requests are absolutely comfortable.

2. RELATED WORKS

2.1 Standard Retrial Queue Models

In standard queueing models, a single queue for handover requests was proposed to queue blocked handover requests that should have been dropped [11, 17]. These requests are served according to some queueing discipline. Most literatures that are based on this theory may have an additional guard channel scheme or its variants like the Fractional Guard Channel (FGC) scheme, Limited Fractional Guard

Channel (LFGC) scheme and so on [15]. These schemes have one goal and that is to prioritize handover requests over new requests. Most authors believe that handover requests are the best determinants for the level of the network's QoS parameters. However, standard queueing models do not take the retrial phenomenon into account and therefore cannot be applied in solving a number of practically important problems such as fairness between receptions of new and handover requests [6, 11].

Various research methods combine the retrial queue model and maybe an additional prioritization technique for the handover calls such as an additional queue for the handover calls. A handover retrial mechanism and a guard channel scheme were analyzed in [7]. A recursive analysis was used to evaluate the system's advanced cell configuration. The recursive algorithm was used to obtain an analytical model from which the system performance measures were evaluated. Two models dealing with the call retrial phenomenon in a mobile communication system are presented in [7]. The first model considers a BS with a finite customer population and retrials. The second model focuses on the use of the Guard channel concept to prioritize the handover requests. A Markov chain modelling was proposed and a recursive method was developed to determine the state probabilities. The impact of customer retrial on the network's QoS was then evaluated. A new algorithm termed MOSEL (Modelling, Specification and Evaluation Language) was proposed in [18] to approximate and offer solution to the retrial queue problem and to evaluate the performance of the cellular network. The MOSEL tool was also used to formulate models and also calculate the system's performance measures. An approximate approach that employed probability generating functions and phase-merging techniques was employed for the performance evaluation of an M/M/2 retrial queue in which both servers are subject to active and idle breakdowns [6]. These techniques are used to analyze the limiting behaviour of the system. It has also been shown that algorithmic methods and approximate techniques are powerful tools for solving analytically problems that are difficult to manage [6, 8]. An approximate technique and the cut-off scheme were employed to calculate desired QoS metrics in an isolated cell under multi-parametric call admission control (CAC) in a cellular wireless network (CWN) in [5]. An analytical approach for the development of simple approximate CTMC models whose state spaces grow linearly with the number of calls that can be simultaneously made within a cell was proposed in

[19]. This approach was used to evaluate the call blocking probabilities with customer retrials due to blocked new calls or failed handovers. Analysis of the retrial queue system is a complex and often intractable method. Either of the approximate analytical or the computational approach [7, 11, 13] has been used in literature. However, none of these approaches has yielded precision and numerical stability required from such systems.

2.2 The Integrated Computational and Analytical Approach

This is the modeling approach adopted in this paper. The reason for this approach is that it aids in designing and analyzing a practical system that exhibits a real world scenario while reducing complexity in system analysis. The computational approach was done using the matlab/simulink block oriented simulation environment while the analytical approach was based on a two-dimensional Markov process and the phase-merging techniques. The two-dimensional Continuous-Time Markov Chain (CTMC) provides a very powerful and efficient means for the description and the analysis of a random process that is characterized by transition between states. Due to the transitions between and within levels of the two dimensional CTMC, deriving the steady-state probabilities in retrial queueing systems is challenging. Therefore, the phase-merging technique (PMA) aids in the analysis of the model complexity required for the adequate description of the real system [14, 20]. It is important to establish the best design criteria for improved network QoS standard while reducing system complexity.

The performance metrics to be considered are the blocking probability with retrial (PBr), blocking probability without retrial (PB), the handover call drop probability (PH), retrial queue length (K) and retrial rate (α).

3. MODEL DESIGN

3.1 The physical Model

The network considers both the new and handover calls; therefore an isolated node in an isolated cell is represented with two call types as shown in Figure 1. It is assumed that the cells in the adopted typical cellular network are identical. Therefore the analysis of an isolated cell can be extended to the entire network. Requests are modeled as a Markov-modulated Poisson process with exponentially distributed inter-arrival times given as λ_N and λ_H respectively. The call holding time defines the call departure rates, μ , which is also exponentially distributed. It is assumed that blocked new and handover calls will retry for a free channel the first time with a retrial probability θ or abandon the system and be lost with probability $1-\theta$ if blocked the second time. The admittance of more than two retrials for a single call in the system will increase the offered load on the network and this may lead to congestion [11]. Therefore a single retrial is allowed for each call in the retrial queue since a limited number of retrials will prevent the negative influence of retrials on new requests that are arriving at the network for their first attempt and also on handover calls crossing the cell boundaries. The time between retrials is exponentially distributed and the retrial rate is denoted by α . Among the channels in the system there are the free channels which accept new and handover requests, while the guard channels, g , are reserved solely for handover calls. Handover calls that were not served will retry for a free channel once. Such requests will be treated as a retried call and cannot access the guard channels. A queue with size K is provided to hold blocked calls for a first retrial. After a first retrial, if unsuccessful, the blocked calls will leave the system.

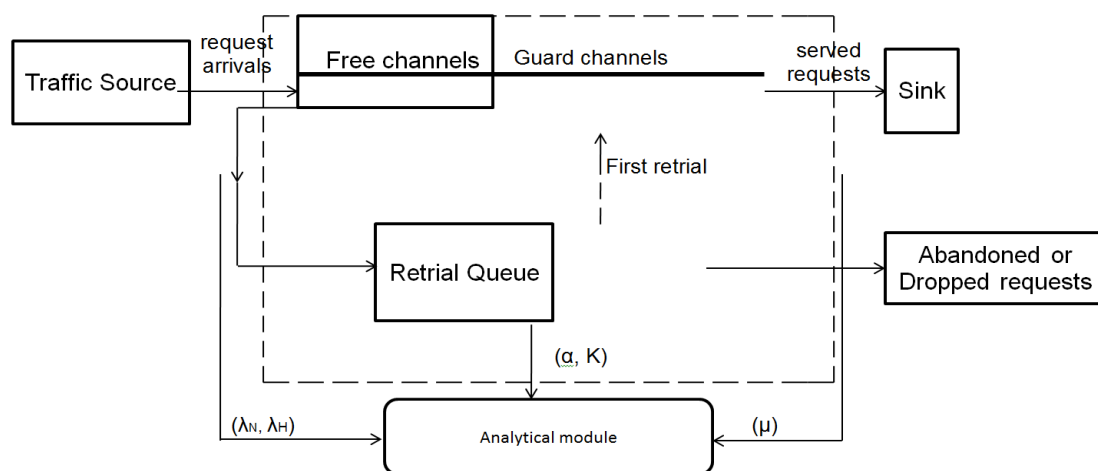


Figure 1: An integrated model of retrial queue for wireless cellular networks

3.2 The Analytical Model

The network is modeled analytically as a two dimensional continuous time Markov chain (CTMC) [9, 13]. The state of the system is described by a two-variable random process in continuous time, $\{X(t), Y(t) : t \geq 0\}$, where $X(t)$ is the number of occupied channels at time t and $Y(t)$ is the number of blocked new requests in the retrial queue. The state space of the CTMC is defined as: $S = \{i, j\}, i = 0, 1, 2, n - g, n - g + 1, \dots, n; j = 0, 1, 2, \dots, K$. $P(i, j)$ is the probability that the system is in a state (i, j) where $(i, j) \in S$.

The analysis of the Markov model is complex and will be simplified using the phase merging approximation (PMA) technique [9, 12, 13, 14]. Following the method of computation used in [9, 12, 13], the steady state probability can be determined by taking the product of the conditional probability distribution of the number of channels occupied at time t and the marginal probability of the different levels of the retrial queue. The approximate conditional probability $P_i = P_{i|j}$ can be expressed as [7, 9, 13]

$$P_i = P_{i|j} = \begin{cases} \frac{\rho^i}{i!} P(0) \text{ for } 1 \leq i \leq n - g \\ \frac{\rho^{n-g} \cdot \rho_H^{j-(n-g)}}{i!} P(0) \text{ for } n - g < i \leq n \end{cases} \quad (1)$$

The initial conditional probability $P(0)$ is also given as

$$P(0) = \left(\sum_{i=0}^{n-g} \frac{\rho^i}{i!} + \sum_{i=n-g+1}^n \frac{\rho^{n-g} \cdot \rho_H^{i-(n-g)}}{i!} \right)^{-1} \quad (2)$$

Consider each level of the retrial queue as a macrostate (an aggregate of all states in a level or a merged model) to form the overall state space of the merged model which is defined as $\tilde{S} \equiv j: j \geq 0$. The analysis of the overall state space of the merged model gives the approximate marginal probability distribution of the number of blocked new requests in the retrial queue. Transitions into the macrostates, ε_1 , and transitions out of the macrostates, ε_2 , can be expressed in equations (3) and (4);

$$\varepsilon_1 = \lambda_N \theta \sum_{i=n-g}^n P_i \quad (3)$$

$$\varepsilon_2 = j\alpha \left(\sum_{i=0}^{n-g-1} P_i + (1 - \theta_1) \sum_{i=n-g}^n P_i \right) \quad (4)$$

Therefore the total traffic intensity between the macrostates defined as q will be

$$q = \frac{\varepsilon_1}{\varepsilon_2} = \frac{\lambda_N \theta \sum_{i=n-g}^n P_i}{j\alpha (\sum_{i=0}^{n-g-1} P_i + (1 - \theta_1) \sum_{i=n-g}^n P_i)} \quad (5)$$

The steady-state marginal probability will be obtained by applying the Erlang formula to the merged model.

$$\varphi_j = \frac{q^j}{j!} \varphi_0 \quad (6)$$

φ_j is defined as the marginal probability for each level of the merged model and φ_0 is defined as the initial marginal probability for the merged model.

To obtain the initial marginal probability, all probabilities on the system of macrostates are summed to unity.

$$\sum_{j=0}^K \frac{q^j}{j!} \varphi_0 = 1 \quad (7)$$

$$\varphi_0 = \left(\sum_{j=0}^K \frac{q^j}{j!} \right)^{-1} \quad (8)$$

Substituting equation (8) into equation (6) the steady-state marginal probability will be

$$\varphi_j = \frac{\frac{q^j}{j!}}{\sum_{j=0}^K \frac{q^j}{j!}} \quad (9)$$

The approximate steady-state probability distribution $\{X(t), Y(t) : t \geq 0\}$ can be obtained by taking the product of the approximate conditional probability and the marginal probability.

From the steady state probability, the system's performance measures can be determined from equations (10) - (12)

$$PBr = \frac{\alpha(1 - \theta)}{\lambda_N \theta} \sum_{i=n-g+1}^n \sum_{j=1}^K j p(i, j) \quad (10)$$

$$PB = \sum_{i=n-g+1}^n \sum_{j=0}^K p(i, j) \quad (11)$$

$$PH = \sum_{j=0}^K p(n, j) \quad (12)$$

3.3 The Simulation Model

The simulation was carried out in Matlab/Simulink-an object oriented simulation software; with each aspect of the network in figure 1 designed using the simevents library. New and handover requests are seen as entities. Traffic sources generate entities at given intervals of time. Call arrivals are random and independent. This implies that calls/requests arrive to the network in a random and independent manner. The arrival rate λ , for generated entities is defined as

$$\lambda = \frac{\text{Total number of entities generated}}{\text{simulation time}} \quad (13)$$

The service rate for served entities can be defined as;

$$\mu = \frac{\text{Total number of entities served}}{\text{simulation time}} \quad (14)$$

Therefore the traffic intensity ρ can be defined as;

$$\rho = \frac{\lambda}{\mu} \tag{15}$$

The design represents an isolated cell, therefore calls arrive at the channels one at a time in order to avoid congestion in the cell. The channels are modeled using a multiple N-server. To design the guard channel which is reserved solely for handover calls, a threshold was set. Priority is given to handover requests by setting a threshold for channels that must not be exceeded by new requests and are reserved for handover requests. The number of entities currently in the multiple servers at simulation time *t* is used to ensure that only handover calls are carried above the set threshold. Served handover calls go to the sink while the dropped requests are sent to the retrial queue. The retrial queue is modeled to follow the FIFO queuing technique. Requests in the retrial queue are allowed to retry one at a time only for a first retrial. Such calls are lost if they are blocked the second time. Since the requests are made at different intervals of time, random inter-arrival (time difference between one request arriving

at the network and the next request arrival) and service times (time difference between one served request and the next) are chosen in the simulation. A flowchart of the simulation procedure is shown in Figure 5. Some parameters and their values are presented in table 1.

Table 1: Simulation parameters

Simulation parameters	Value
simulation time, <i>t</i>	300secs
inter-arrival times	0.01 to 1secs
service time	1 to 2 secs
Number of servers	10

Data collected from the simulation result are applied to the analytical model (section III). Figure 2 gives a flowchart of the simulation procedure and how priority was assigned to both the new and handover requests so as to achieve fairness.

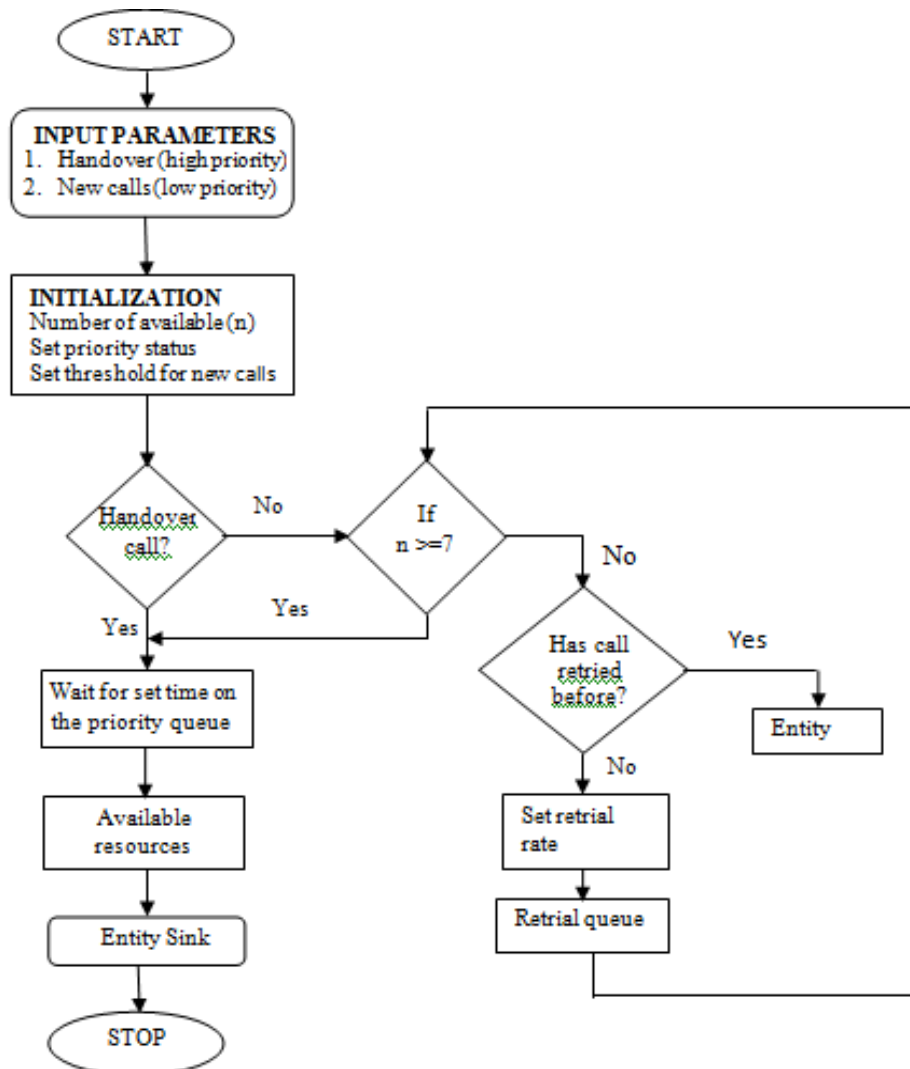


Figure 2: Flowchart of the simulation procedure

4. RESULT AND ANALYSIS

Figure 3 shows the graph of PB and PBr versus traffic intensity. The y-axis to the left reads values for PB while the y-axis to the right reads values for PBr. For a given traffic intensity, there is a reduction in the volume of calls blocked from gaining access into the network when the retrial queue is used. It can be deduced from the figure that the application of the reservation scheme and retrial queue is the best solution for improved QoS in the network. However, the proposed approach performed better particularly at low traffic intensity. This implies that rural and suburban deployments could benefit more from this system rather than a typical urban deployment where the traffic intensity is high. From Figure 4, it can be deduced that at low traffic intensities, the probability of dropped handover requests with and without retrial are almost equal. This is due to the fact that the guard channel prioritizes the reception of handover requests into the network over new requests. However, as traffic intensities increase, results obtained show that the PH tends to diverge. The PH with retrial is lower than that without retrial. This could be attributed to the fact that increases in traffic requests overwhelms the reserved channel resulting to an increase in the PH. The application of the retrial queue to the network minimizes handover requests at the risk of being dropped. Therefore the PH is minimized with both reservation and retrial.

There are other parameters that affect the PBr such as the retrial queue length and the retrial rate. From Figure 5, a graph of PBr versus K was plotted for different retrial rates (α) of 0.75, 0.5 and 0.25. The graph shows a sharp increase in the PBr below queue lengths of 4 and gradually becomes steady for queue lengths greater than 4. The response shows that a lesser α and more K gives an additional improvement in the PBr.

5. CONCLUSION AND RECOMMENDATION

In this paper, a solution to the problem of retrial queue complexity in wireless cellular networks has been proposed. Ease of analysis, precision and numerical stability are the defining factors that led to the use of the integrated computational and analytical approach. This approach aids in determining the parameters of a retrial queue and reservation system from a real world perspective.

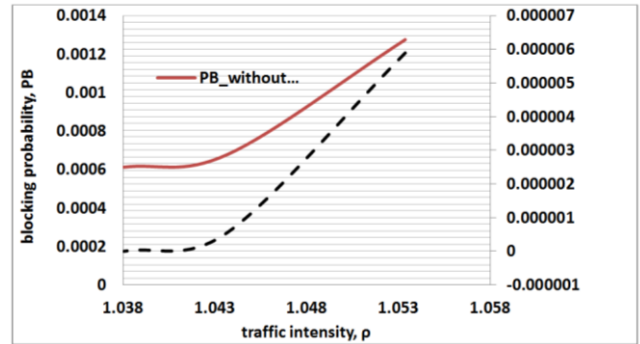


Figure 3: Blocking probability with and without retrial

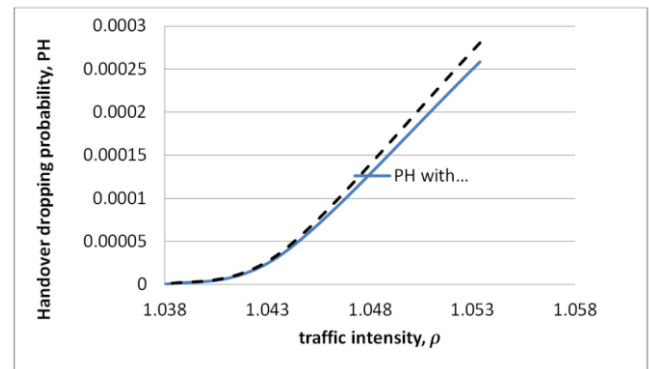


Figure 4: Handover dropping probability with retrial and without retrial

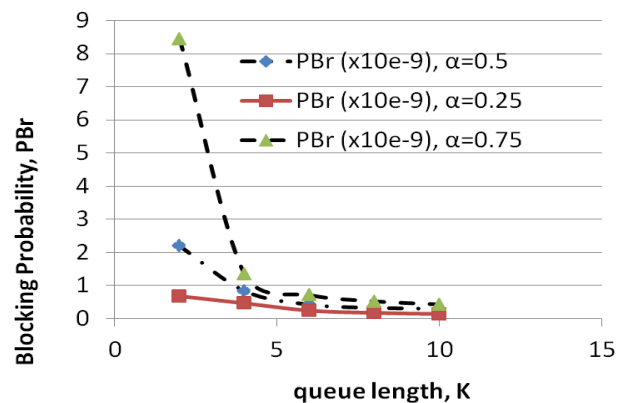


Figure 5: PBr vs K for different retrial rates

Designing systems that meet the practical system criteria are an added advantage. In addition to the retrial queue and guard channel scheme for improved blocking probability, it was also deduced that with a lower retrial rate and increased retrial queue size, the network will perform better in all traffic load conditions. Therefore the network operator is armed with an additional criteria needed to improve the network quality without an increase in the system complexity.

6. REFERENCES

- [1] Theodore S. Rappaport, *Wireless Communications Principles and Practice*. Prentice Hall, 2002
- [2] Michael Mandjes and Kurt Tutschku, *Efficient Call Handling Procedures in cellular mobile networks*. Institute of Computer Science University of Wurzburg, July 1996.
- [3] Sanjay Sharma, *Wireless and Cellular Communications*. S.K Kataria and Sons, January 2007.
- [4] G. Borgia, P. Camarda, A. D'Alconzo, A. De Biasi and M. Siviero, "Drop Call Probability in established Cellular Networks," *IEEE Vehicular Technology Conference*, vol. 5, June 2005.
- [5] A. Melikov and M. Fattakhova, "Algorithmic Approach to Qos Analysis in Multi-service Cellular Wireless Networks" *Journal of Transport and Telecommunication*, volume 10, 2009.
- [6] Jesus R. Artalejo and Antonio Gomez-Corral, *Retrial Queueing Systems: A Computational Approach*, Springerlink, 2008.
- [7] P. Trangia and M. Mandjes, "Modelling of Customer Retrial Phenomenon in Cellular Networks," *IEEE Journal on Selected Areas in Communication*, vol. 15, 1997.
- [8] J. M. Gimenez-Guzman et al, "Analysis of a cellular network with user redials and automatic handover retrials," *Proceeding of the 7th International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networking*, 2007.
- [9] Anioke C. L. and Ani C. I., "A reservation and retrial queue analysis for cellular networks", *International Journal of Engineering Research and Technology*, vol.3, February 2014.
- [10] Chidera I. Anioke, Obinna C. Nnamani and Cosmas I. Ani, "Call Drop Minimization Techniques for Handover Calls in Mobile Cellular Networks" *International Journal of Scientific & Engineering Research*, Volume 6, Issue 3, March 2015.
- [11] Tien Van Do, "A new computational algorithm for retrial queues to cellular mobile systems with guard channels", *Computers and Industrial Engineering, Elsevier*. September 2010.
- [12] N. Sherman and J. Kharoufeh, "An M/M/1 retrial queue with unreliable server", Department of Operational Sciences, Air Force Institute of Technology, USA.
- [13] Bong Dae Choi, Agassi Melikov and Amir Velibekov, "A Simple Numerical Approximation of Joint Probabilities of calls in service and calls in the Retrial Group in a Picocell," *Appl. Comput. Math*, vol. 7, no. 1, 2008.
- [14] Vladimir S. Korolyuk and Nikolaos Limnios, *Stochastic Systems in Merging Phase Space, Singapore*, World Scientific Publishing Co. Plc Ltd, 2005.
- [15] F.A. Cruz-Pérez and L. Ortigoza-Guerrero, *Fractional Resource Reservation in Mobile Cellular Systems*, Auebarch Publications, 2006.
- [16] I. Atencia, P. Moreno, "A Single-server Retrial Queue with General Retrial Times and Bernoulli Schedule," *Applied Mathematics and Computation*, Elsevier Inc., 2005.
- [17] J. Artalejo and G. Falin, "Standard and Retrial Queueing Systems: A Comparative analysis," *Revista Matematica Complutense*, Volume XV, 2002.
- [18] Janos Roszik, *Retrial queues and their application in the performance modelling of communication networks*, PhD Dissertation, Faculty of Informatics, University of Debrecen, 2006.
- [19] M. Ajmone Marsan, G. De Carolis, E. Leonardi, R. Lo Cigno and M. Meo, "Approximate Markovian Models of Cellular Mobile Telephone Networks with Customer Retrials," *IEEE International Conference on Communications*, June 2000.
- [20] Vladimir S. Korolyuk and Vladimir V. Korolyuk, *Stochastic Models of Systems*, Kluwer Academic Publishers, 1999.