International Journal of Mobile Communication & Networking.
ISSN 2231-1203 Volume 3, Number 1 (2012), pp. 1-15

http://www.ripublication.com

Handoff Management for Integrated Wireless Mobile Networks

Krishan K. Pandey[1], Yashu Vyas[2] and Ravi K. Pandey[3]

[1]COMES, University of Petroleum and Energy Studies (UPES), Deharadun-248007, Uttarakhand, India
E-mail: krisha.pandey@gmail.com, kkpandey@ddn.upes.ac.in
[2]Consultant, Deloitte USA, Moberly LN, Bentonville-72712, USA
[3]Government Engineering College, Ajmer, Rajasthan, India

Abstract

One of the key challenges in integrated wireless and mobile networks is to efficiently support multi-class services as each type of services has distinct characteristics and quality of service (QoS) requirements. Problems in handoff algorithm or its parameters may lead into call drops with a direct effect on user satisfaction. This is particularly critical for 3G systems, where high data rate users will be prime candidates for being dropped. Unnecessary handoff leads to degraded call quality and waste of capacity in signaling. Since the cell-size is constantly decreasing, therefore it is important to devise a handoff algorithm which identify users with different mobility and data rate characteristics and maximize the utilization of network infrastructure. A Multiple Queuing System for handoff in integrated real-time and non-real time service with priority reservation and preemptive priority handoff scheme is being analyzed which categorizes the service calls into four different types, namely, real time and non-real time service originating calls, and real-time and non-real-time handoff request calls and divide the channels among these four types of services according to their priorities. The system is modeled using a multidimensional Markov chain and a numerical analysis is presented to estimate blocking probabilities of originating calls, forced termination probability, and average transmission delay. Our results show that the predictions of the analytical model are in very good agreement with simulation results. Scheme significantly reduces the forced termination probability of real-time service calls. The probability of packet loss of non real-time transmission is shown to be negligibly small, as a non-real-time service handoff request in waiting can be transferred from the queue of the current base station to another one.

Keywords: Handoff, Quality of Service (QoS), Multidimensional Markov chain.

Introduction

In recent years there has been a phenomenal growth in the development and deployment of wireless services, evident from the proliferation of cellular data services and the emerging wireless multimedia applications. The next generation wireless and mobile networks are designed to support a true combination of real-time and non real-time services. The introduction of these novel applications like data delivery and real-time multimedia in 3G and 4G mobile cellular systems will not only significantly increase the traffic offered to wireless networks, but also create a heterogeneous traffic environment. Real-time service, such as voice transmission, has strict timing requirements. On the other hand, non real-time service, such as e-mail and ftp applications, can tolerate some delays. These types of services have distinct characteristics and performance requirements. Based on the different QoS's, efficient radio resource control and management are imperative to optimally support multiple types of service. Some of the radio resource management tasks performed by cellular systems include admission control, channel assignment, power control, and handoff[1]. An integrated radio resource management scheme can make necessary trade offs between the individual goals of these tasks to obtain better performance. Integrated radio resource management can increase system capacity within specified quality constraints. Due to the time and space varying nature of the cellular system, the radio resource management tasks need to be adaptive to factors such as interference, traffic, and propagation environment. Adaptive radio resource management tasks can reduce the initial cell planning and make re-planning easier, organized, and automatic. Some of the important objectives of resource management are global minimization of the interference level and handoffs and adaptation to varying traffic and interference scenarios.

Handover is the mechanism that transfers an ongoing call from one cell to another as a user moves through the coverage area of a cellular system. But with increasing demand for mobile computing services and limited available bandwidth, wireless networks increase the number of simultaneous users in the network systems by reducing the cell size. Next generation wireless networks are adopting micro/pico-cellular architecture. However, due to smaller cell size the number of mobile users crossing the cell boundaries is increasing; hence, the proliferation of handover calls. The way that handover calls are handled has a direct impact on the quality-of-service (QoS) provided to the mobile user (MU).

QoS provisioning is a major challenge for future integrated wireless mobile

[1]The terms handover and handoff are used interchangeably within this paper.

cellular networks as they provides complete multimedia services like video, voice and complete data streams along with internet. Providing QoS guarantees in wireless networks is more complicated than in wire line networks. The wireless link with its particular characteristics and the user mobility render QoS provisioning more difficult and complicated in wireless networks than in wire line networks. In determining the performance of these systems, many factors come into play. These factors depend heavily on the characteristics of the wireless channel such as signal fading, multipath distortion, limited bandwidth, rapidly changing propagation conditions, mutual interference of signals and vulnerability to eavesdropping and unauthorized access. The user mobility (handoff's) makes these effects even worse and guaranteeing the required QoS becomes more complicated. With the development of integrated wireless mobile systems, data has to be incorporated and its effects need to be taken into consideration. However, in order to meet future demands, a handoff strategy need to take different features of voice, video, and data services into account, i.e., the ideal handoff processes has to be service-dependent. For example, real-time service is very sensitive to interruptions. On the other hand, a transmission delay of non real-time service does not have much impact on the performance (delay insensitive). Therefore, a successful handoff without interruption is very important to real-time services, but not so critical for non real-time service.

In this paper we have analyzed the performance of a service-dependent preemptive and priority handoff scheme based on channel reservation for integrated real-time and non-real-time service wireless mobile networks. The analyzed handoff scheme provides priority reservation to voice handoff requests and dividing calls into four different classes: (i) Originating voice calls; (ii) Originating data calls; (iii) Voice handoff calls; and (iv) Data handoff calls. Depending upon the QoS parameters and traffic generated in the referenced cell few channels (not specific ones) are reserved for voice and data handoff requests and the voice and data handoff requests have their own finite queues.

Review literature

A good design of a handoff scheme requires that the blocking probability for new calls originated in a cell be minimized. However, from the user's point of view, handling of handoff request calls is more important, as forced termination of ongoing calls is considered much more disastrous than blocking of new calls. In addition, it should also reduce transmission delay of non real-time service calls and maximize channel utilization. Unfortunately, all requirements cannot be satisfied simultaneously and there are some tradeoffs between various parameters.

The study of handoff is not a new topic to the wireless communication and numerous methods are proposed for efficient handoff management. A simple way of giving priority to handoff requests is to reserve few channels, which is called a guard channel scheme. One of the earliest analytical frameworks for guard channel methods (GCM) was developed by Guerin [4]. Guerin proposed a novel approach, where a certain number of channels are used exclusively for handover calls and only queuing of originating (new) calls is investigated. This approach, not only minimizes the handover blocking probability for the handover calls, but also increases the total carried traffic in the network.

Since dropping of a call in progress is desirable than blocking a new call, various methods have been devised to prioritize handover calls over new calls. Hong and Rappaport [3] developed an analytic framework for GCM with Fist–in–First–out (FIFO) queuing of handover calls and no queuing of originating (new) calls. Results showed that the guard channel priority scheme with FIFO queuing of handover calls achieves smaller forced termination probability for handover calls compared to other schemes. Both the originating calls and handoff requests are allowed to be queued in [1]. Hong and Rappaport [5] proposed a priority oriented channel access scheme for cellular mobile radio telephone systems serving vehicular and portable radio telephone users where two–level priority reservation is provided. The calls generated are divided into three different classes: new vehicular calls, new portable calls and handoff calls. Priority is given to handoff calls over new calls and to vehicular calls over portable calls.

Cellular communication systems that support a mixture of platform types and queuing of handoff calls are considered in [6]. In pure loss systems, if there are no channels available at the target gateway, the handoff attempt fails and the call is forced to terminate. In delay systems however, handoff calls can be held in queue while the supporting mobile is within a transition region where acceptable performance can be provided by at least two different gateways (base stations). However in most of the studies discussed above, the research focuses on voice based cellular systems only and multiple type of services have not been taken into consideration.

With the development of integrated wireless mobile systems, non-real-time service has to be incorporated and its effect needs to be taken into account [7]. A handoff scheme for a non-real-time wireless network only has been studied in [8]. However, future wireless networks will be required to support multiple types of services simultaneously. In order to meet future demands, a handoff strategy needs to take different features of these services into account, i.e., the ideal handoff processes have to be service–dependent. For example, transmission of real-time service is very sensitive to interruptions. On the other hand, transmission delay of non-real-time service does not have any significant impact on the performance of their service, i.e., they are delay insensitive. Therefore, a successful handoff without interruption is very important for real time services, but not so critical for non real–time services. A two-dimensional model for integrated service cellular mobile systems has been proposed in [9], which assigns preemptive priority to real-time service calls. However, no distinction is made between originating and handoff requests. In [10], a handoff scheme for the integrated voice/data wireless network has been introduced, while only data service handoff requests are allowed to be queued.

Preemptive and Priority Reservation Handoff Scheme

In the analyzed handoff scheme, we classify the traffic into four different types of traffic as per QoS requirements and divides the channels into three major groups. Four different types of traffic are specified as follows: type 1 are real-time service calls originating in the cell (i.e., real–time originating calls); type 2 are real-time service calls in progress and handed off to the cell from other adjacent cells (i.e., real–time handoff calls); type 3 are non real time originating calls, and type 4 are non real time handoff calls. Three channel groups are called real-time channels (RCs) that are used by real time service users only, non real time channels (NCs) that are used by non real time service users only and the common handoff channels (CCs) that are reserved for both real time and non real time handoff calls.

System Model

Here we focus our attention on a single cell, which we call the reference cell. When a mobile user dials the number and contacts the base station of the reference cell, an originating call (type 1 or type 3) is generated in the reference cell. When a real-time service mobile user holding a channel enters the handoff area of the reference cell from a neighboring cell, a real-time service handoff request (type 2) is generated. There is no handoff area for non-real time service mobile users. Instead, we use the cell boundary, which is defined by the points where the received signal strength between two adjacent cells is equal. Therefore, when a non-real-time service mobile user holding a channel approaches the reference cell and crosses the cell boundary, a non-real-time service handoff request (type 4) is generated. The system model for the reference cell is shown in Fig 3.1.The total number of channels S of a reference cell are divided into three groups, namely,

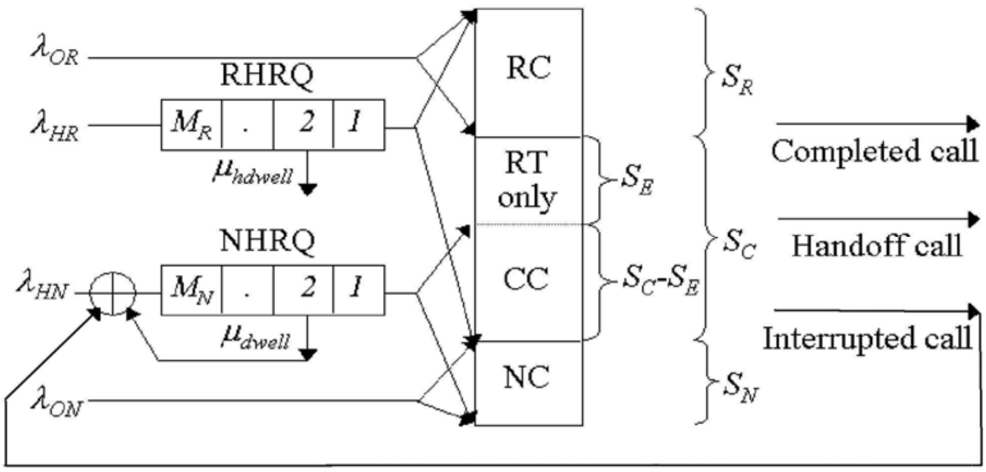

Figure 3.1: System Model for a reference cell.

1. Real-time service Channels (RC) group with capacity S_R ,
2. Common handoff Channels (CC) group with capacity S_C , and
3. Non-real-time service Channels (NC) group with capacity S_N ,

RC is reserved for real-time service calls only (type 1 and type 2), including both real-time originating calls (λ_{OR}) and handoff request calls (λ_{HR}). NC is reserved for non-real-time service calls (type 3 and type 4) only, including both non-real-time originating calls (λ_{ON}) and handoff request calls ($\lambda_{H\ N}$). As type 2 and type 4 traffic have more stringent QoS requirements than type 1 and type 2 CCs are reserved for the overflow of real-time and non-real-time service handoff requests from the first two groups of channels. Out of CC, a few channels are reserved exclusively for real-time service handoff requests only.

We use S_E to mark the predefined number of free channel reserved before non-real-time service handoff request calls could be served in CC. This reservation of channel depends on the QoS parameter and ratio of traffic intensity of real and non-real time calls. There are two queues in the reference cell, real-time service handoff request queue (RHRQ) and non-real-time service handoff request queue (NHRQ). RHRQ with finite capacity M_R serves the real-time service handoff request calls and NHRQ with finite capacity M_N serves the non-real-time ser- vice handoff request calls. The originating service calls do not have their own queues.

When an originating real-time service call (type 1) arrives, it can be served only if there are channels available in RC. Similarly, an originating non-real-time service call (type 3) can be served only if there are idle channels in NC. An originating real- time service call (or an originating non-real-time service call) is blocked if it finds no channels available in the RC (or NC). Fig 3.2 shows the flow diagram for handling originating calls in the handoff scheme proposed here.

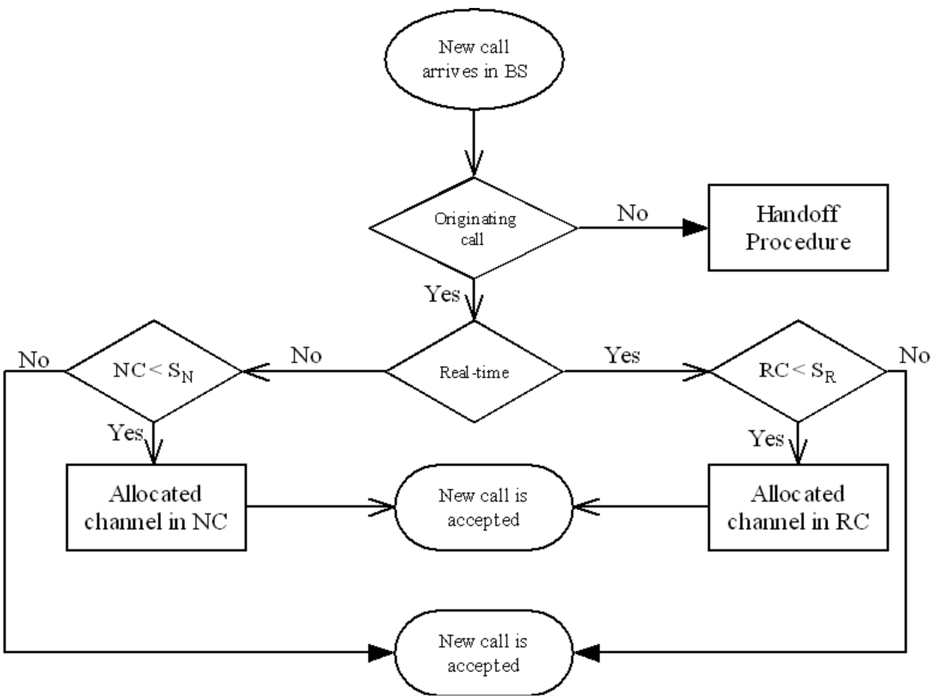

Figure 3.2: Flow diagram for handling originating calls.

Fig. 3.3 shows the flow diagram for handling handoff request calls. When a real-time service user holding a channel enters the handoff area of the reference cell, a handoff request (type 2) of real-time service is generated. The real time service handoff request call first checks whether there are channels available in RC on arrival. If RC is full, it checks whether there are channels available in CC. For the priority reservation handoff scheme, the real-time service handoff request call is put in RHRQ if both RC and CC are full. However, in the preemptive priority handoff scheme, the real-time service handoff request call has a higher priority and can be served even if there is no idle channel in both RC and CC. With preemptive procedure, the real-time service handoff request calls can get served by preempting one of the current non-real-time handoff request calls in CC if there is at least one ongoing non-real-time service call in CC and NHRQ that is not full. The interrupted non-realtime service call returns back to NHRQ and waits for an idle channel to be served based on the first-in-first- out rule. In this schemes, the real-time service hand off request is queued in RHRQ.

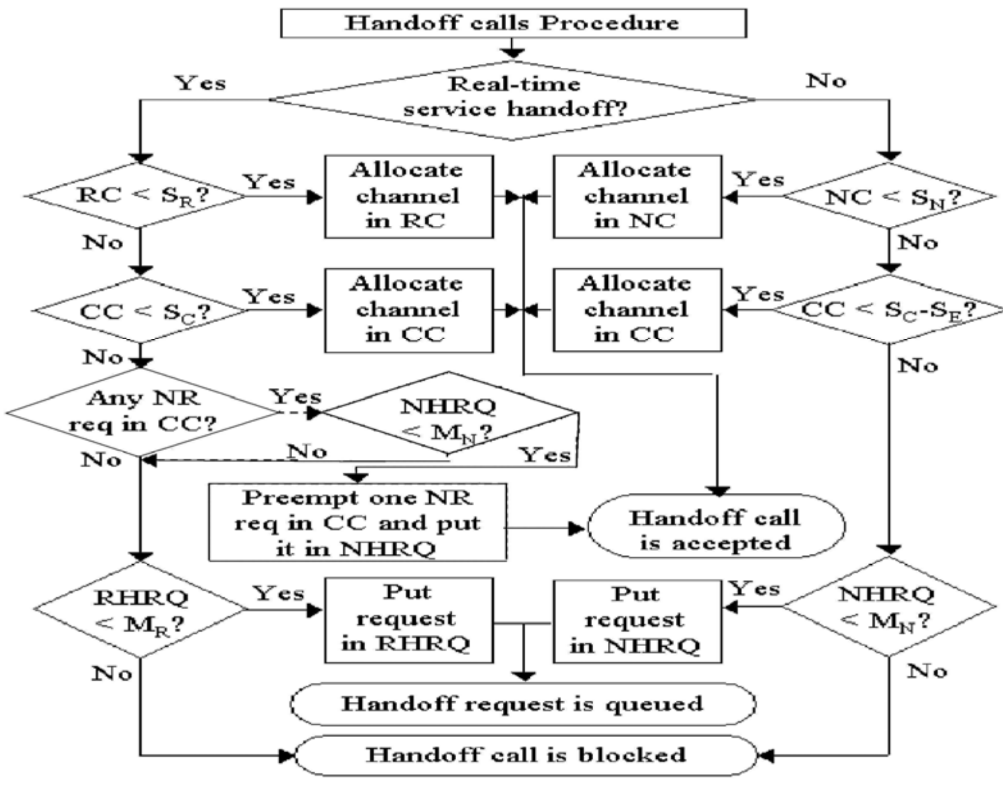

Figure 3.3: Flow diagram for handling handoff request calls.

The Iterative Algorithm

Beginning with an initial real time and non real time handoff arrival rate, λ_{HR} and λ_{HN}, we first derive the blocking probabilities of both real time and non-real time calls. To compute the probabilities, we use GTH algorithm [2] in solving the system of equations. We iterate this procedure till handoff arrival rates of both real time and non-real time calls converge and using these two quantities we estimates for the blocking probabilities of real time and non real time calls.

Algorithm

Input Parameters **:** Number of channels used for both the services real time and non real time S, S_R , S_C , S_E , S_N , buffer for real time calls M_R , buffer for non real-time M_N , real time call holding time $1/\mu_{C\ R}$, nonreal time call holding time $1/\mu_{C\ N}$, cell dwell time $1/\mu_{hdwell}$.

Output Measures **:** Blocking probabilities of new and handoff calls of both real time and nonreal time calls, forced termination probability of real time call, transmission delay in nonreal time service.

Step 1: Select initial values for λ_{HR} and λ_{HN} .

Step 2: Write down all transition balance equations and form the matrix like $\pi = \pi P$.

Step 3: Calculate the probabilities by solving system of equations by using GTH algorithm or SOR (Successive over Relaxation) method.

Step 4: Calculate the various performance metrics obtained by the equations given below;

$$E[C_R] = \sum_{i=0}^{S_R} \sum_{k=0}^{S_C-S_E} \sum_{j=0}^{S_C-k} (i+j) \sum_{l=0}^{S_N} P(i,j,k,l,0) + \sum_{k=0}^{S_C-S_E} (S_R+S_C-k) \sum_{j=S_C-k+1}^{S_C+M_R-k} \sum_{l=0}^{S_N} P(S_R,j,k,l,0)$$
$$+ \sum_{i=0}^{S_R} \sum_{k=0}^{S_C-S_E} \sum_{j=S_C-S_E-k}^{S_C-k} (i+j) \sum_{m=1}^{M_N} P(i,j,k,S_N,m)$$
$$+ \sum_{k=0}^{S_C-S_E} (S_R+S_C-k) \sum_{j=S_C-k+1}^{S_C+M_R-k} \sum_{m=1}^{M_N} P(S_R,j,k,S_N,m). \quad \ldots(A)$$

$$E[T_N] = \frac{N_h E[T_w](1-B_{HN})\lambda_H N}{(1-B_{HN})\lambda_H N + (1-B_{ON})\lambda_O N} \quad \ldots(B)$$

Step 5: Calculate the new handoff arrival rate of both real time and non-real call using equations and compare it with previous handoff arrival rate, if it does not converge, go to step 2 and use the new computed handoff arrival rates.

Numerical Results and Discussion

In this section, the performance of the analyzed scheme is evaluated both by simulation and analytical model. In the numerical analysis, we will consider the

GPRS environment. GPRS offers faster data transmission via GSM network within the range of 9.6 kbits to 115 kbits. This technology makes it possible for users to make telephone calls and transmits data at the same time. The main benefit of GPRS are that it reserves radio resources only when there is data to send and it reduces reliance on traditional circuit-switched network elements. With GPRS, an IP data transmission protocol is being introduced to GSM. IP is a data transmission protocol which is used in the internet. Before the introduction of GPRS, the radio capacity was used for calls and data transmission within the GSM network in a rather inefficient way. For data transmission the entire channel is occupied and is thus insufficiently used. With the GPRS technology, channel allocation is flexible where one to eight channels can be allocated to a user or one channel can be shared by several users, moreover, channel de-allocation is also supported in GPRS.

Our simulation program is written in C programming language and simulation was carried out by taking large number of calls generated in reference cell following a Poisson arrival process. The call holding time and cell dwell time for each type of service call is generated following a exponential distribution. A handoff will occur when residue call holding time exceed's the current cell dwell time. Assumptions for simulation;

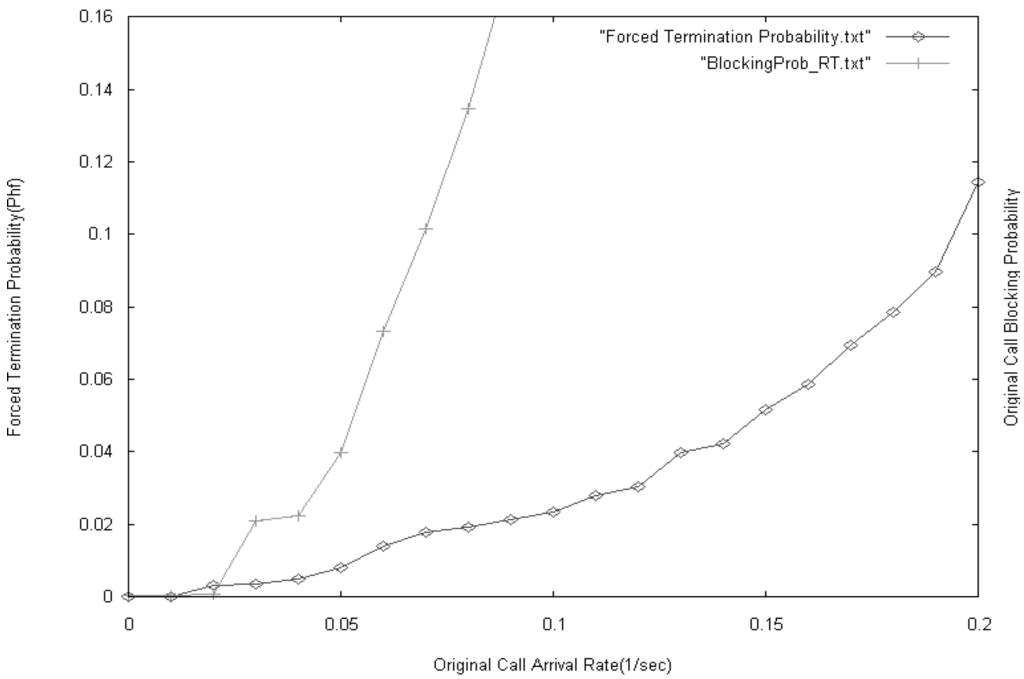

Figure 5.1: Real-time QoS parameters versus offered traffic.

.

- Real-time call holding time $E[T_{C\,R}]$ = 120 seconds,
- Non-Real time call holding time $E[T_{C\,N}]$ = 60 seconds,
- Number of channel in each cell

$S = S_R + S_C + S_N = 12,$
$S_R = 6, S_C = 3, S_N = 3, S_E = 1.$

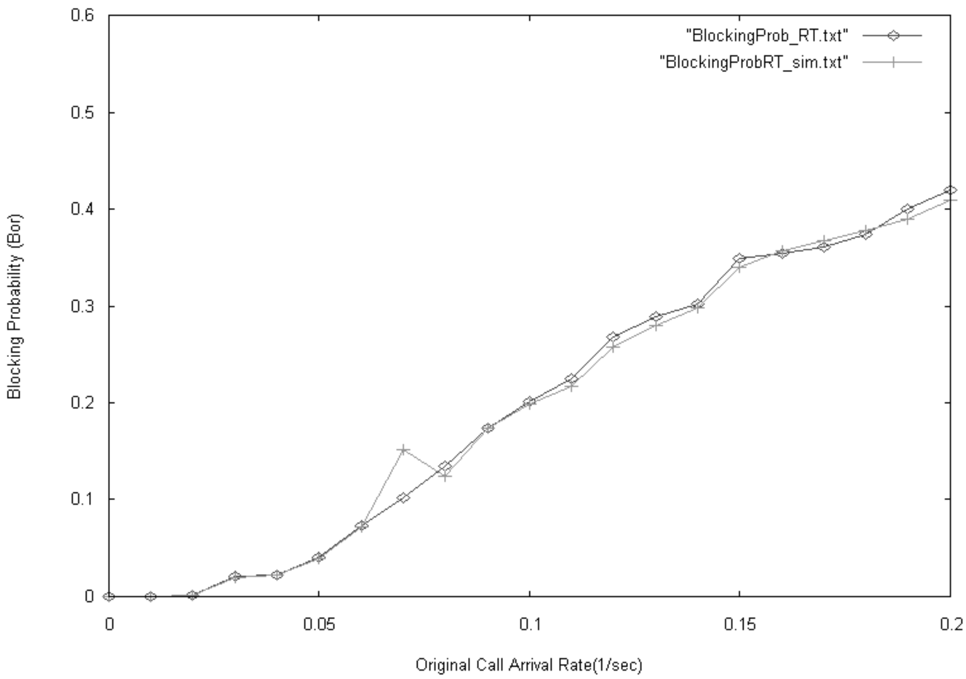

Figure 5.2: Blocking probability of Real-time service calls Versus offered traffic.

.

- Size of queuing buffer $M_R = 5$, $M_N = 50$.
- Average cell dwell time = 62.8 sec and Handoff area dwell time = 4 sec.

The ratio of originating real-time service call and non-real time service calls $\lambda_{OR} / \lambda_{ON}$ is set to 1.

Discussion

We compared several important QoS parameters, including blocking probability of originating calls for both real-time service and and non-real time service (B_{OR}) and (B_{ON}), forced termination probability of the real-time service calls P_{hf} and transmission delay of non real time service calls (T_N). From the various graphs obtained we can observe that simulation result , which assumes that the call holding time, the cell dwell time, and handoff area dwell time follows exponential distribution, and Analytical formula match each other and are consistent for all four QoS parameters. The difference between the result of simulation and analytical model for B_{OR} , B_{ON} , and P_{hf} is very small.

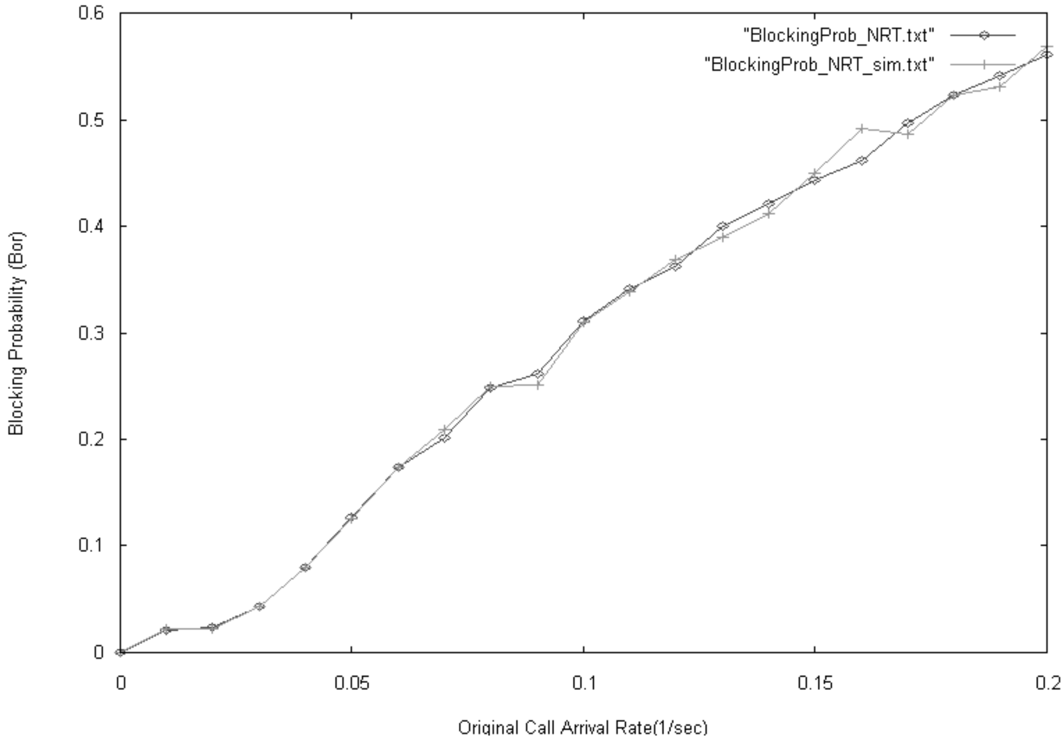

Figure 5.3: Blocking probability of Non-Real-time service calls versus offered traffic.

Figure 5.2 shows the comparative result of the analytical model and simulation model. It shows that as the call arrival rate increases the blocking probability for the real time call increases. Similarly figure 5.3 shows the comparative result of the analytical model and simulation results, blocking probability for non real time increases as the call arrival rate increases.

Figure 5.6 shows the comparative study of various QoS parameters, as the call arrival rate increases blocking probability of non real time increases very fast in comparison to real time service calls because in the analyzed scheme number of channels reserved for real time calls is more in comparison to non real time calls and there is no queue for originating calls. Rate of increase of forced termination probability of handoff request as the call arrival increases is less in comparison to blocking probabilities because real time handoff request has exclusively reserved channels for such calls, also it can borrow or share channels allocated for non real time service handoff request.

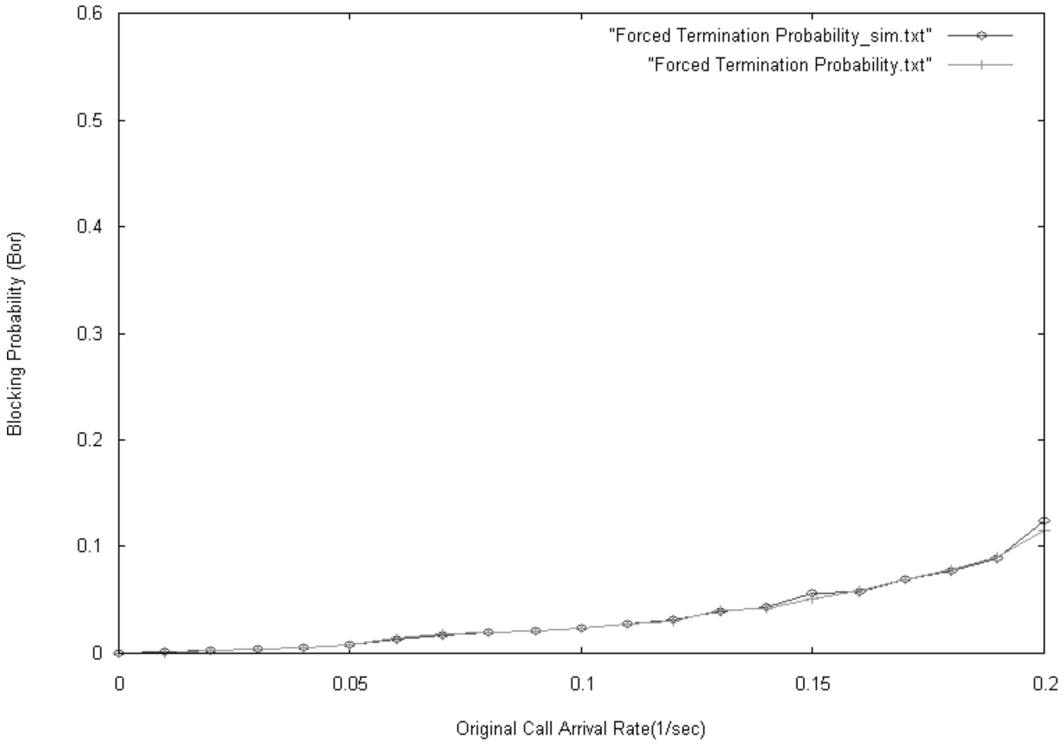

Figure 5.4: Forced Termination Probability of real-time Handoff request.

In addition to this, as the scheme is preemptive in nature, it means a non real handoff request can be preempted and placed in the queue and freed channel is allocated to real time handoff requests. Buffer size is also large for non real time handoff requests; this also is one of the reasons for lesser value of forced termination probability in comparison with blocking probabilities.

Figure 5.4 shows the comparative result of simulation and analytical model for forced termination probability of handoff request calls, there is almost complete agreement between both simulation and analytical model.

Figure 5.5 shows the transmission delay in the packets of non real time services. As the call arrival rate increase transmission delay increase because in that case non real time service calls are preempted for allocating those freed channels for real time calls.

Figure 5.7 shows the both QoS parameters for real time services i.e., B_{OR} and P_{hf} . Blocking probability increases drastically in comparison to forced transmission probability because handoff requests can be served in real time channels, its own allocated common channels and also by preempting the non-real time handoff requests served in the common shared channels, therefore handoff requests for real time services has maximum resources at its disposal.

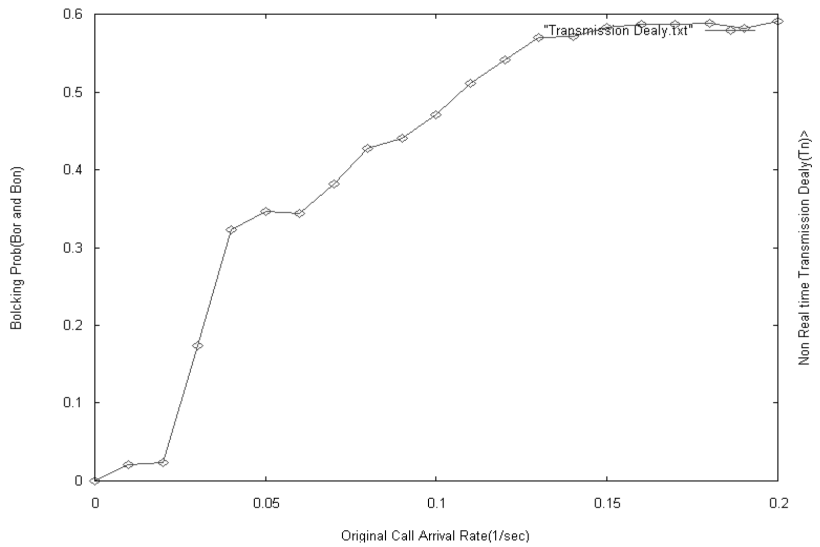

Figure 5.5: Transmission delay in Non Real time handoff request calls versus offered traffic

Since real- time service transmission is very sensitive to interruption, decreasing forced termination probability is more important than decreasing transmission delay of non-real time service. On the other hand blocking probability of originating non real time service calls in preemptive scheme is small. Because the non real time service handoff request can be transferred from the queue of one cell to another the forced termination probability for accepted non-real time service is negligibly small. The preemptive and priority reservation channel scheme have all the good properties of the previously available channel allocation schemes.

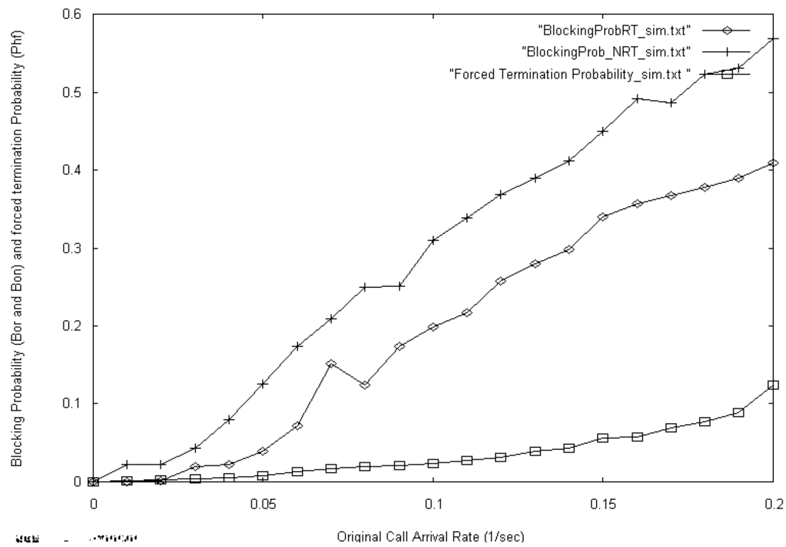

Figure 5.6: Comparative Result of Analytical model and Simulation model.

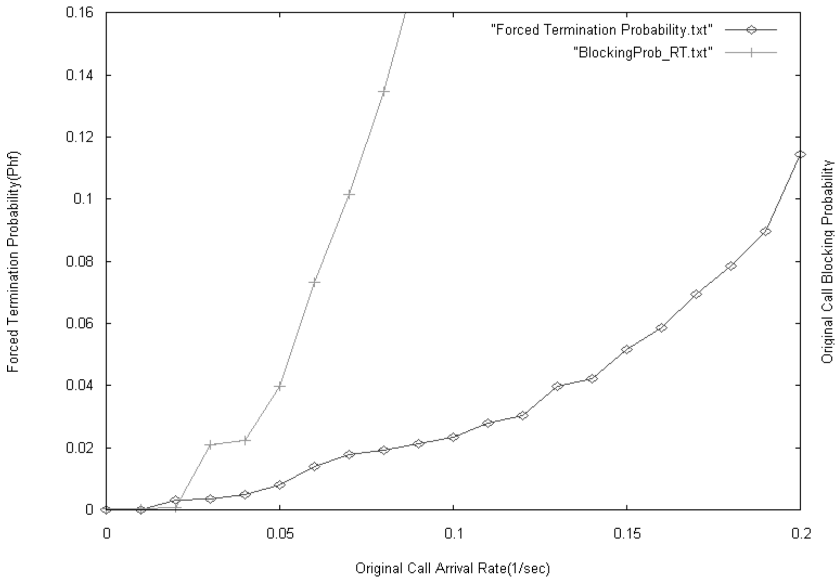

Figure 5.7: Real-time service QoS parameters versus offered traffic.

Conclusions and Future Scope

The design of a handoff scheme is an important consideration for the QoS in a wireless mobile network with integrated real-time and non-real-time services. A handoff scheme with priority reservation and preemptive priority procedure has been discussed and analyzed in the project. An analytical model for the system performance has been presented. Simulation with exponential distribution scenarios is obtained and results are observed to match with the analytical evaluations. Blocking probability of originating calls, forced termination probability of real-time service calls, and average transmission delay of non-real-time service have been evaluated. It is seen that forced termination probability of handoff request calls of real-time service mobile users can be decreased by this analyzed preemptive and priority reservation handoff schemes. Moreover, non-realtime service handoff requests do not fail, except for negligibly small blocking probability, as a non-real-time service handoff request can be effectively handled by transferring it from the queue of the reference cell to an adjacent cell. Through simulation we observed the following

- Providing buffers for non real time services traffic decreases the blocking probabilities of both, new and handover non real time calls.
- As the offered traffic increases transmission delay in the non real time services increases.
- Adopting the preemptive and priority based scheme for handover calls, decreases the blocking probability of real time services call.

Quality of Service (QoS) issue has been the main focus of network researchers and operators during the last decade. This is to provide the best service to the users with minimum network resources in order to maximize both the user satisfaction and the profit of the operators. Mobility research issue was considered separately from QoS research.

More recently, there is a research trial to combine QoS and mobility management in order to maintain QoS during the mobility of the users. Future wireless networks consists of a set of overlapping areas. Therefore, fast lossless handover between different network types (i.e., vertical handoff) will be crucial to be realization of seamless mobile multimedia networks. Future research is need to develop an analytic framework so that performance of low priority services can be enhanced without compromising high priority services.

Bibliography

[1] Qing-An Zeng, K. Mukumoto, and A. Fukuda : Performance Analysis of Mobile Cellular Radio System with Priority Reservation Handoff Procedures.Proc. IEEE VTC-94, vol.3, no.1, pp 1829-1833, June 1994. 2.5

[2] G. Latouche and V. Ramaswami : Introduction to Matrix Analytic Methods in Stochastic Modeling. Chapter 5, SIAM Journal. 4.3

[3] Daehyoung Hong and Stephen S. Rappaport : Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Non- prioritized Handoff Procedures. IEEE Trans. Veh. Tech., vol.VT-35, no.3, Aug 1986. 2.5

[4] R.A.Guerin : Queueing-Blocking System with Two Arrival Streams and Guard Channels. IEEE Trans. Commun., vo1. 36, pp. 153-163, Feb 1988. 2.5

[5] Daehyoung Hong and Stephen S. Rappaport : Priority Oriented Channels Access for Cellular Systems Serving Vehicular and Portable Radio Telephones. IEE Proc. I, CSV-136, no. 5, pp. 339-346, 1989. 2.5

[6] Stephen S. Rappaport and C. Purzynski : Multiple Call Handoff Problem with Queued Handoffs and Mixed Platform Types. IEE Proc. I, CSV-142, no. 1, pp. 31-39, 1995. 2.5

[7] Donald.J. Goodman: Trends in Cellular and Cordless Communication. IEEE Comm. Magazine, vol. 29, no. 6, pp. 31-40, June 1991. 2.5

[8] Qing-An Zeng and Dharma.P. Agrawal : Modeling and Efficient Handling of Handoffs in Integrated Wireless Mobile Networks. IEEE Trans. Vehicular Tech- nology, vol. 51, no. 6, Nov. 2002. 2.5

[9] Francis.N. Pavlidou: Two-Dimensional Traffic Models for Cellular Mobile Sys- tems. IEEE Trans. Comm., vol. 42, nos. 2/3/4, pp. 1505 - 1511, Feb./Mar./Apr. 1994. 2.5

[10] Qing-An Zeng and Dharma.P. Agrawal : Performance Analysis of a Handoff Scheme in Integrated Voice/Data Wireless Networks. Proc. IEEE VTC-2000, pp. 1986-1992, Sept. 2000. 2.5