Pre-emption based Call Admission Control with QoS and Dynamic Bandwidth Reservation for Cellular Networks

[1]Ankur Guha Roy, [2]Rajdeep Mandal, Iti Saha Misra
Electronics and Telecommunication Engineering Department
Jadavpur University, Kolkata, India
Email:itsankur45@gmail.com; justrajdeep@gmail.com
itimisra@cal.vsnl.net.in

Kalpana Saha (Roy)
Govt. College of Ceramic Technology, Kolkata, India
Email: klpnsh@yahoo.com

Abstract - **Call admission protocol (CAC) is a very important process in the provision of good Quality of Service (QoS) in cellular mobile networks. With micro/Pico cellular architectures that are now used to provide higher capacity, the cell size decreases with a drastic increase in the handoff rate. In this paper, we present modeling and simulation results to help in better understanding of the performance and efficiency of CAC in cellular networks. Handoff prioritization is a common characteristic, which is achieved through the threshold bandwidth reservation policy framework. Combined with this framework, we use pre-emptive call admission scheme and elastic bandwidth allocation for data calls in order to gain a near optimal QoS. In this paper, we also use a Genetic Algorithm (GA) based approach to optimize the fitness function, which we obtained by calculating the mean square error of predicted rejection values and the actual ones. The predicted values are calculated using a linear model, which relates the rejection ratios with different threshold values.**

Keywords-CAC, Preemtion, Dynamic Bandwidth reservation, GA

I. Introduction

Call admission protocol (CAC) make the decision whether to admit or reject the call. The decision is taken based on the requested bandwidth, network utilization, and the priority of the incoming call [1]. QoS provisioning in wireless networks is a challenging problem due to scarcity of wireless resources i.e. radio channels and due to the mobility of users. One of the most important connection-level QoS issues is how to reduce or control handoff drops due to lack of available resources in the new cell, since mobile users should be able to continue their ongoing connections. Since it is practically impossible to completely eliminate handoff drops the best one can do is to provide some form of probabilistic QoS to ensure rejection ratio to remain below a predefined value.

The future generation networks support different user with different requirements. Even the different customer may have different QoS requirements for the same type of calls, based on different cost per user. Today's cellular phones offer many other services such as text messaging, multimedia messages, web browsing, and video transmission. When a user initiates or receives a call, he may roam around the area covered by the networks. If the mobile user moves from one cell to another, with unfinished call, the network has to hand off the call from one cell to another at the cell boundary crossing without the users' awareness of handoff or without much degradation of service quality.

Call admission protocol plays an important role in the performance of the network. A good call admission protocol must decide whether to admit or reject the call requests. In order to do that, the protocol must take into account many factors and try to balance call rejection ratio. The maximum utilization is achieved by admitting any incoming request. In order to guarantee a certain call rejection ratio for high priority calls, we must leave some extra bandwidth in the network for high priority calls. Another factor is to be considered as the demand for wireless services continues to increases and the cell sizes are getting smaller. Small cell sizes causes shorter dwell time and more and more handoff during the life of call that adds another level of complexity since handoff calls should be treated differently than new calls.

A good call admission control (CAC) protocol should be fairly fast, it should reduce customer inconvenience and produce a good bandwidth utilization leading to increasing revenue for the

[1] Author at present a postgraduate student at IIT Madras, [2]Author at present a postgraduate student at IISc Bangalore

carrier. When a customer moves from a cell to another new cell, a request for call admission in the new cell be sent to the customer moved to, that call should be treated differently than a new call request in that new cell.
From customer's point of view it is much less desirable to drop a call in the middle of the connection because of the lack of bandwidth in the cell the customer is moving into rather to be denied admission at all [2]. All CAC protocols give priority to the handoff calls over new calls. Today's network consists of smaller cells with low power, so the calls are experiencing much more handoff compared with the high power larger cells of the previous years. A good CAC protocol must give priority to the handoff calls over new calls.

In this paper, we present a new call admission control protocol for wireless network that support dynamic bandwidth for data calls (both originating and handoff) also the handoff voice call can preempt handoff data call and originating data if bandwidth resource is unavailable. Our protocol takes into consideration of the remaining bandwidth and the threshold of each call in order to make a decision to accept or reject the call. In order to improve the total utilization, the data traffic is considered as Markov exponential process.

II. Previous Work

In the system where all calls have same bandwidth, and all calls have the same priority, First Come First Serve (FCFS) is used. FCFS produces a good utilization of communication medium under the above-mentioned conditions.
In [3], the author proposed the concepts of dividing bandwidth into different segments and group the call request into different categories, such that the call request in-group "j" can only be accepted if there is enough bandwidth in segments "j". This is similar to Multilevel Queuing architecture. The main problem in this concept is the waste of bandwidth when one type of call is missing.
In [4], the author proposed a call admission protocol to provide QoS guarantees for multimedia traffic in a heterogeneous Personal Communication System (PCS) network. Here two types of traffic are assumed i.e. real-time and non real-time-traffic .The bandwidth is divided into channels, each type of traffic may request a different number of channels. The channels are grouped into three groups, one for real time request, one for non real time request, and one is combined. The protocol depends on varying the boundaries between the three classes of channels to satisfy the requested QoS.
In [5], the authors proposed the call admission protocol for integrated voice and data service. Their protocol depends on the use of limited fractional guard channel policy.
In [6], a call admission protocol is introduced for wireless network that support call degradation. This protocol admits a new call based on their priority level and the required bandwidth. The protocol degrades the call in order to make bandwidth available for incoming calls.
The authors of [7] proposed a distributed algorithm for call admission in which information about the neighboring cell is taken into consideration in admitting any new call.

A simple but rather efficient algorithm for call admission is proposed in [8]. Here the authors proposed the use of a single buffer to hold the call request if there is not enough bandwidth and then admitted, or held in the buffer up to a maximum waiting time then dropped.
In [9], the authors assumed N different classes of customers each with a different arrival rate, service time and bandwidth requirements. A buffer of length N is used to store incoming call request. Here priority is introduced in the system. To add priority to the system they assume a threshold for every class. Here no-preemption is assumed and also bandwidth of each class is taken constant during the lifetime of the call.
In [10], the authors assumed a system with N types of calls and proposed the condition for acceptance of a call of class i and a non dynamic preemption scheme was suggested in which the future calls were needed to be known.
Our work mainly focuses on [10] and its further improvements. The preemption algorithm proposed in [10] was not physically realizable since it depends on knowledge of future calls of a particular time span. This would make the system non-causal. There was no provision for dynamic resource allocation; allocated call sizes were fixed during the runtime.
In our protocol we have considered the effect of pre-emption and making the bandwidth of data calls dynamic one. This is further extended by proposing an equation based on linear model for predicting a rejection ratio and its optimization.

III. Proposed Protocol

We assume the whole coverage area of the cellular system is divided into cells. There is some overlap area between cells called handoff area. It helps in a smooth handoff. New and handoff calls are admitted to each cell when users try to connect and request a specific bandwidth that depend on the application. The user is admitted, rejected or queued - waiting for another user to release some bandwidth.

We assume an interval of 100 time units. During this time the calls are generated following Markov Exponential Distribution. There are 4 classes of calls, originating data and voice and handoff data and voice. Each class has different arrival rate, service time, and different bandwidth requirements. To add priority to the system a threshold T_i is assumed for every call of class i. If the remaining bandwidth is more than T_i then the call of class i will be accepted. The priority of any call can be increased or decreased by decreasing or increasing its threshold value. We assume the bandwidth of data call is dynamic. The data call can take bandwidth from a range of values $[B_{min}, B_{max}]$

defined in [11], ideally each data call is provided maximum bandwidth i.e. B_{max} if available channels can accommodate, otherwise they will share equally the bandwidth left for them. But the allowed bandwidth should always be greater than or equal to B_{min}. We assume only the handoff voice calls can preempt ongoing and data calls. We also assume a buffer to store preempted calls. If data calls cannot get service at the time of its arrival then it is put up in a queue. This call is processed and admitted or dropped after certain duration.

TABLE-1: ALGORITHM OF THE PROPOSED PROTOCOL

```
when  a call of class i arrives

If  ((remaining BW>=Bi) && (remaining BW>=Ti)
      Accept the call
Else  if  ((remaining  BW<Bi)  &&  (remaining BW>=Ti) &&  ( no  request of class i call in the buffer ))
      B i = Bmin
If      ((remaining  BW>=Bmin  )  &&  (remaining BW>=Ti)
      accept the call
Else  if   ((remaining  BW<Bmin)  &&  (remaining BW>=Ti) && (no  request of class i call in the buffer ))
      preempt the class j call if possible
      store the preempted class j call in the buffer
     or accept the class i call and store it in the buffer
else
      reject the call

when a  call is completed add the BW of the call to the remaining BW
      check for call in the buffer
if ((remaining BW>=Bi)
       accept the call
```

A. Basic Protocol Algorithm

When a call request from class i with bandwidth B_i and threshold T_i arrives in the cell, if the remaining bandwidth of the cell is greater than B_i and T_i, the call is accepted. If the remaining bandwidth of the cell is less than B_i and greater than T_i for data calls the bandwidth of call is compressed and make equal to $B_i = B_{min}$. Then it is checked whether remaining bandwidth of cell is greater then B_{min}, if remaining bandwidth of cell is greater than B_{min} and T_i the call is accepted. Otherwise if the data call is kept in a waiting queue and processed at the start of every sec when it is either accepted (if resource is available) or rejected after certain duration. If the call is handoff voice call, and bandwidth is unavailable it preempts the ongoing data calls only after checking available preempted bandwidth. If there is enough bandwidth after preempting the new call to accept and the preempted data call is stored in the buffer. If bandwidth is unavailable then the voice call is dropped. When a call is completed the calls in the buffer are checked as first come first serve order. But priority is given to the call waiting in the buffer by threshold. If the remaining bandwidth of the cell is greater than the bandwidth of the priority call waiting in the buffer the call is accepted. This protocol can be described in an algorithmic form as shown in Table 1. Figure 2 is the flow chart of the proposed algorithm.

B. Pre-emption

In cellular system it is assumed that the handoff calls preempt the ongoing data calls. When there is a new voice handoff call, and ongoing data calls (both originating and handoff) in the cell the voice handoff call preempts the data calls. The preempted data calls are kept in a queue and processed at a later time. The performance metric used for Preemption is called transmission rate defined in [3]. The Pre-emption algorithm is shown in Table 2 and a flowchart is given in Figure 2.

where B_i= required bandwidth of call i.
T_i= assigned threshold value for call i
remBW_buffer = a variable to store the remaining BW value

TABLE 2: ALGORITHM OF THE PREEMPTION

```
          remBW_buffer = remaining BW
Step 1:
If  ((remaining BW< Bi) &&(remaining BW>=Ti)
          start preemption decision
if  voice handoff call  want to preempt the data call
          make remBW_buffer = remBW_buffer +
                                      B/W of  data  call
if  ((remBW_buffer > Bi) &&(Ti >Tj )
          the voice handoff call preempt the data
call   the voice call is accepted; the data call is
rejected and is stored in the buffer.
       Remaining BW= remBW_buffer
Else repeat from Step 1 for a certain interval.
```

C. Description of processes occurring in the cell

When a data call comes in the cell with bandwidth, and if the remaining bandwidth of the cell is greater than the call i is accepted, otherwise if the remaining bandwidth of the cell is less than then is made equal to and checked if it can be accepted otherwise the call is stored in the wait queue. When a voice handoff call arrives with bandwidth in the cell and the remaining bandwidth of the cell is less than then if there is ongoing data call present in the cell, first it is checked if the data calls if preempted, whether the handoff voice call can be accommodated. If it is seen that the call can be accommodated then the data calls are preempted and stored in a queue with proper information about the duration of time for which it was processed and the voice call is accepted. If it is found that even if the ongoing data calls are preempted the required bandwidth is unavailable then the voice call is dropped. The data calls, which were preempted and put up in the queue, are serviced at a later instant. The data calls kept in the wait queue and the preempted calls kept in the preemption queue are kept for a finite interval of time. If the calls cannot be accepted before that time then they are simply rejected after that time interval.

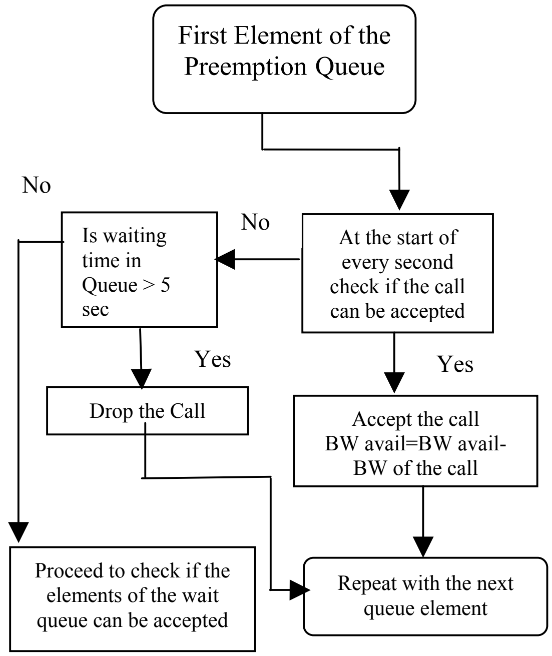

Figure 2. Flow diagram for processing in the cell

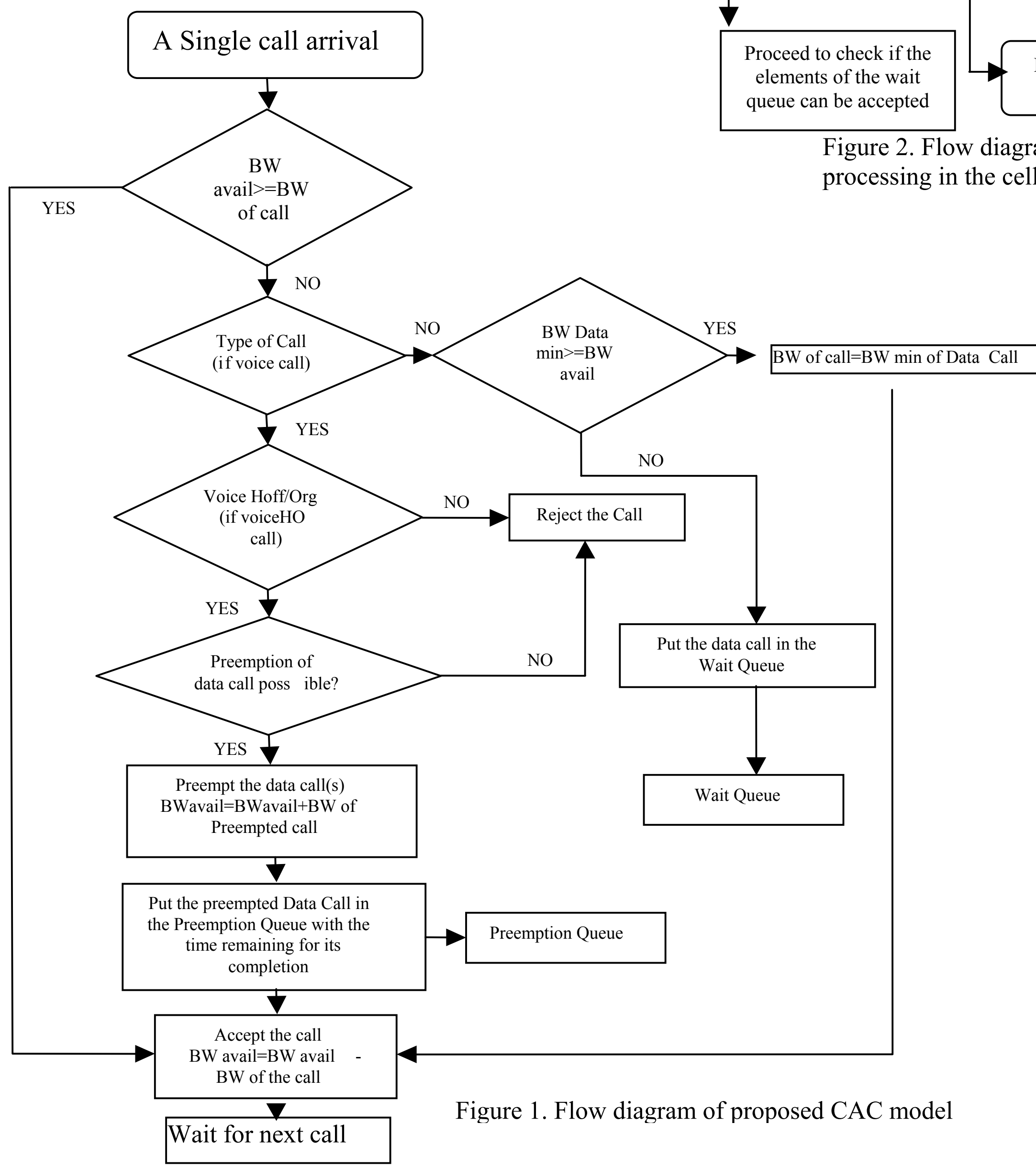

Figure 1. Flow diagram of proposed CAC model

IV. SIMULATION SETTING

The protocol given in [10] is simulated first for comparison purpose under MATLAB 7.0 environment.

In our simulation, the call arrival process is Poisson arrival process with average arrival rate (λ) 5,5,2,4 unit for 4 types of calls (viz. voice handoff, data handoff, voice originating, data originating respectively) i.e. we are operating with a system having large hand off call arrival rates and very small voice originating rate. It is to be noted that the choice of the system parameters depend on operating condition of the system and it has been chosen somewhat arbitrarily in this case for a specification of large hand off rate.

We consider the system with total bandwidth of 450 BBU (Basic Bandwidth Unit). The bandwidth requirement for voice calls is 20 BBU and 10 BBU for data calls in normal mode. Minimum allowed bandwidth for data calls is 5 BBU.

The Rejection Ratio (RR) is defined as,

$$\frac{\text{No. of calls rejected of particular type}}{\text{Total no. of calls rejected}}$$

We take the average rejection ratio as the average of last 20 values and then expressed in %. The simulation time is considered for 100 time units.

A. Determination of the rejection ratio by third order Polynomial

The RR is function of three threshold values for different types of calls and change of one threshold largely affect the RR of a particular type. If any one of the threshold parameter is varied in the range 1 to 80 (BW unit), in absence of the other threshold parameters (i.e. all the other threshold parameters are fixed at = 1) then the variation of RR for a given type of call can be approximated by a third order polynomial with the help of MATLAB tool.

$$RR_{\text{data hand-off}} = f_1(T_{idh}),$$

if T_{ido} and T_{ivo} are set to 1 to fix their effects on the system.

Similarly, $RR_{\text{data hand-off}} = g_1(T_{ido})$, if T_{idh} and T_{ivo} are set to 1 to fix their effects on the system.

By a set of simulations, we fit the following polynomials:

$$RR_{\text{data hand-off}} = -3.602*T_{idh}^3 \;\; -0.4897*T_{idh}^2 + 18.06*T_{idh} + 63.09 \quad (1)$$

$$RR_{\text{data hand-off}} = 2.394*T_{ido}^3 + 1.437*T_{ido}^2 - 14.77*T_{ido} + 31.14 \quad (2)$$

$$RR_{\text{data hand-off}} = .0001579*T_{ivo}^3 - 0.01891*T_{ivo}^2 + 0.2597*T_{ivo} + 46.79 \quad (3)$$

$$RR_{\text{data originating}} = 1.816*T_{idh}^3 + 0.4056*T_{idh}^2 - 9.767*T_{idh} + 18.26 \quad (4)$$

$$RR_{\text{data originating}} = -3.967*T_{ido}^3 - 2.005*T_{ido}^2 + 22.18*T_{ido} + 50.43 \quad (5)$$

$$RR_{\text{data originating}} = 1.058*T_{ivo}^3 - 1.044*T_{ivo}^2 - 4.777*T_{ivo} + 25.48 \quad (6)$$

$$RR_{\text{voice originating}} = 1.798*T_{idh}^3 - 0.1597*T_{idh}^2 - 8.127*T_{idh} + 16.6 \quad (7)$$

$$RR_{\text{voice originating}} = 1.692*T_{ido}^3 - 0.06822*T_{ido}^2 - 7.409*T_{ido} + 17.07 \quad (8)$$

$$RR_{\text{voice originating}} = -2.317*T_{ivo}^3 + 3.097*T_{ivo}^2 + 13.28*T_{ivo} + 29.01 \quad (9)$$

Since the algorithm ensures minimal rejection ratio of the voice hand off calls and simulation results show insignificant change of rejection ratio of voice HO calls with the variation of threshold parameters, we avoid fitting any polynomial for this type of call.

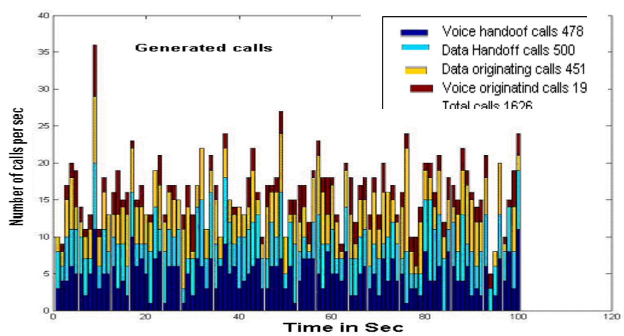

Figure 3. Number of different generated calls per sec using Poisson process

Figure 3 describes the call generated by Poisson process for different calls. Figures 4 and 5 are plotted to give the relative measures of the CAC protocol without preemption and with preemption respectively. It is evident that the number of voice hand of calls served at each interval increases drastically with preemption i.e. much emphasis is assigned on this type of calls.

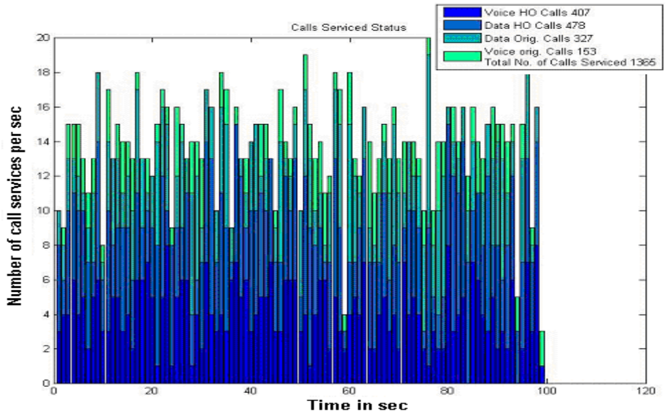

Figure 4. Calls serviced vs. time without preemption (total number of serviced call 1326)

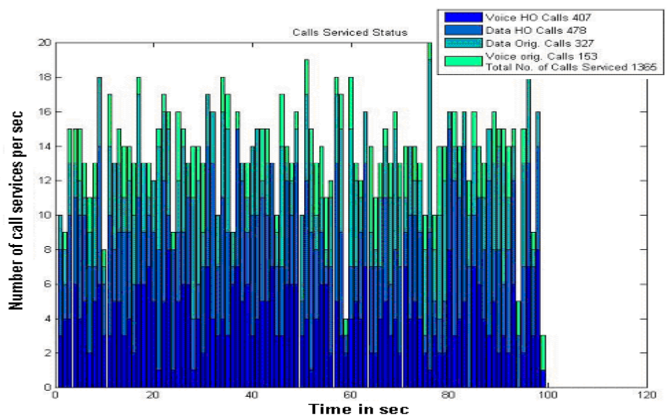

Figure 5. Calls serviced vs. time with preemption (Total number of calls services 1372)

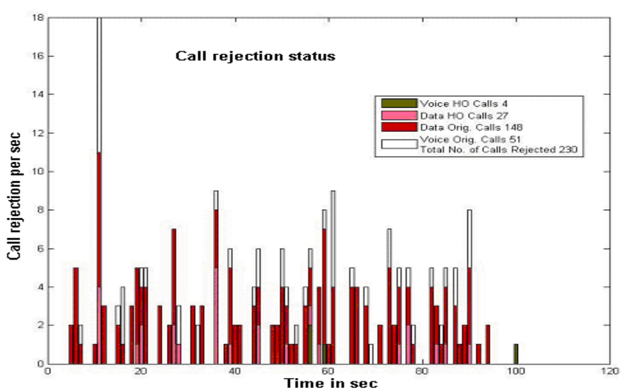

Figure 6. Calls rejected vs. time with preemption, voice call rejection status improves

The figures 6 and 7are plotted to give the relative measures of the CAC protocol without preemption and with preemption respectively in terms of call rejections per second. It is evident that the number of voice handoff calls rejected at each interval decreases considerably with preemption.

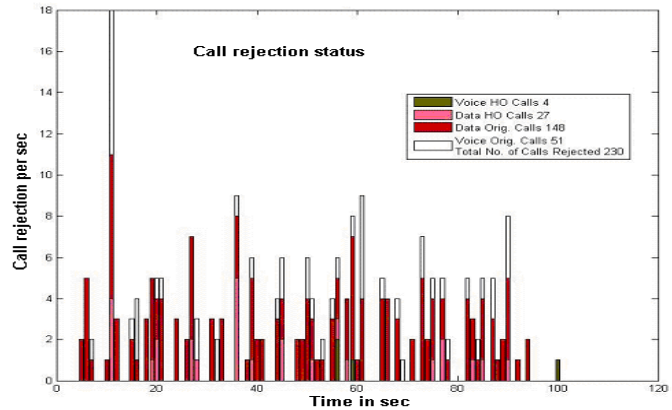

Figure 7. Calls rejected vs. time without preemption

The figures 8 and 9 are showing the BW usages of the CAC protocol without preemption and with preemption respectively. Since the system takes some time to stabilize, we have taken the average BW usage as the average of last 20 iterations. It is evident that average BW usage increases considerably with pre-emption.

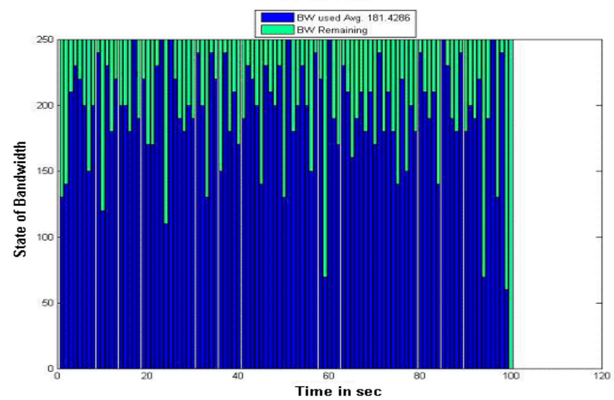

Figure 8. Bandwidth usage vs. time without preemption

The figures 10 and 11 show the change of current rejection ratio with time for the CAC protocol without preemption and with preemption respectively. It can be seen from the pictures that in both cases the system takes some time to stabilize. The rejection ratio of Voice Hand off call has decreased drastically with preemption.

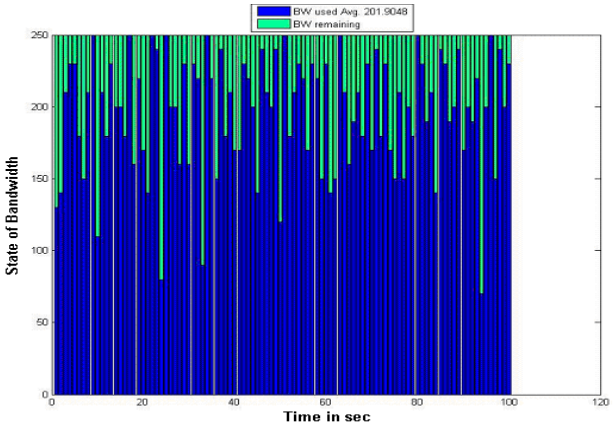

Figure 9. Bandwidth usage vs. time with preemption

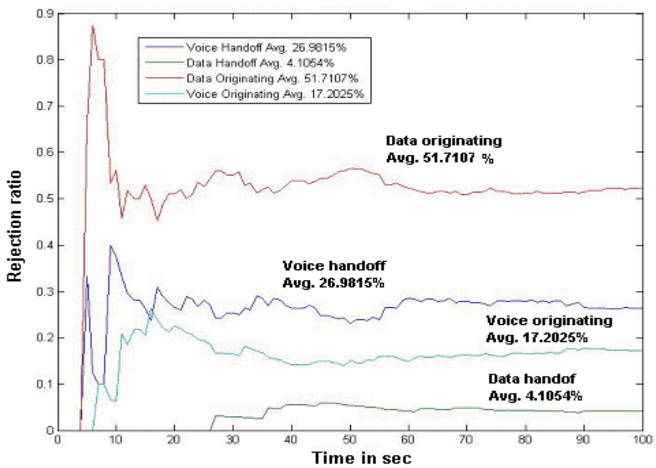

Figure 10. Rejection ratio vs. time without preemption

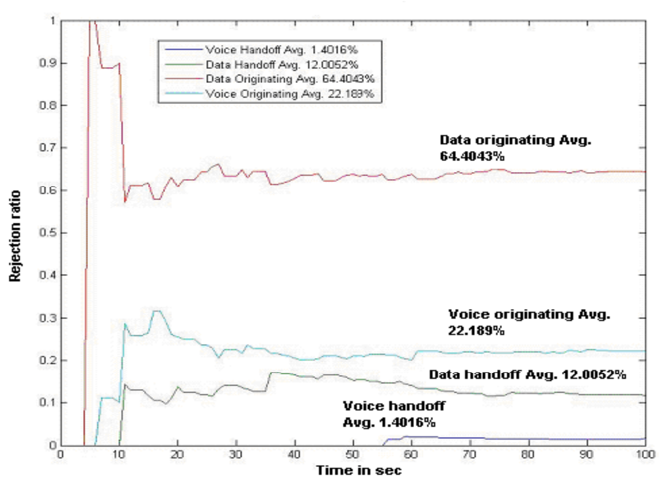

Figure11. Rejection ratio vs. time with preemption

B. Prediction of QoS parameters using linear Mathematical model:

A mathematical model is an abstract model that uses mathematical language to describe the behavior of a system. Mathematical models are used particularly in the natural sciences and engineering disciplines. Mathematical models are usually composed by variables, which are abstractions of quantities of interest in the described systems and operators that act on these variables, which can be algebraic operators, functions, differential operators, etc. If all the operators in a mathematical model present linearity, the resulting mathematical model is defined as linear. We assume that the RR is a linear combination of the third order polynomials obtained previously. By superposition principle, RR for i^{th} type of call can be modeled by:

$$RR_i = a_i \times f_i(T_1) + b_i \times g_i(T_2) + c_i \times h_i(T_3) \qquad (10)$$

We have to find the values of a_i, b_i, c_i (i=1,2,3) i.e. all total 9 coefficients. Optimization of a 9 variable function can be very efficiently performed by evolutionary computation technique. The most famous one of this kind is Genetic algorithm (GA) [12]. GA is a search technique used in computing to find true or approximate solutions to optimization and search problems. Thereby, it is used to compute the coefficient values. A fitness function is a particular type of objective function that quantifies the optimality of a solution. The fitness function is defined over the genetic representation and measures the quality of the represented solution. The fitness function is always problem dependent.

In this context, we define the fitness function as the mean square error value between the actual rejection ratio and our predicted rejection ratio.

Fitness function =

$$\sum_{i=1}^{3} (RR_{predicted,\, i} - RR_{actual,\, i})^2 \qquad (11)$$

Here, $RR_{predicted}$ is obtained by using the equations (1) to (9) with varying weight factors a_i, b_i & c_i.

We have used the MATLAB GA Toolbox and a_i, b_i, c_i are varied so that the predicted RR approaches actual RR.

RR_{actual} is obtained from the simulation of the actual system with different types of calls.

When fitness function is minimized, the predicted value will be the close approximation of the actual system value.

These optimized (a_i ,b_i ,c_i) values obtained from GA are chosen as the coefficient values and the simulation is run for different values of threshold parameters. It has been observed that the

predicted value of the RR closely follows the accurate value within a certain tolerance (±2% approx).
Therefore for a given set of threshold parameters, we can predict the rejection ratios almost accurately from the linear mathematical model.

C. Simulation Parameters:
Population Size =20, Crossover fraction =0.8, Generations =100, Stall generations = 50 (If there is no improvement in the best fitness value for the number of generations specified by Stall generations, the algorithm stops). Stall time limit =20 (If there is no improvement in the best fitness value for an interval of time in seconds specified by Stall time limit, the algorithm stops).

V. Conclusion

From the graphs it is evident that QoS of the voice handoff call undergoes a significant improvement since the rejection ratio is decreased. The curves show transient behavior initially but stabilize gradually. This is the reason behind selecting the average rejection ratio as the average of last 20 values. The threshold parameters can be calibrated properly for maintaining a certain QoS for all types of calls. The system could be made adaptive by threshold parameter optimization with changing call arrival rate. Thus the proposed CAC algorithm can handle required QoS in terms of rejection ratio by adjusting dynamic bandwidth reservation. Again, this paper shows the way to use GA based optimization tool for CAC problem.

References

[1] Sun J-Z, Sauvola J, Howie D. Features in future: 4G visions from a technical perspective. Proceedings of IEEE GLOBECOM’01, Vol. 6, Nov 2001, pp.3533-3537.

[2] C. H. Choi, M. I. Kim , and S.J. Kim, “Call Admission Control Using Moving Pattern of Mobile Users for Mobile Users for multimedia Networks,” in Proc. of the IEEE Conference on local Computer Networks , 431-440, 2002.

[3] Y.C Lai and S. F. Tsai “ A Fair Admission Control for Large bandwidth Multimedia Application ,” in Proc. of GLOBECOM, 312-322, November 2002.

[4] Y. R. Huang and J.M-.Ho, “Distributed call admission control for heterogeneous PCS network,” in IEEE Transactions on Computers 51(12), 1400-1409, December 2002.

[5] C. W. Leong, W. Zhuang, Y. Cheng and L. Wang, “Call Admission Control for integrated on/off voice and best-effort data services in mobile cellular communication ,” in IEEE Transactions on Communication, 52(4),778-790,May 2004

[6] M. Aboelaze, A. Elnaggar, and M. Musleh, “ A priority based call admission control protocol with call degradation for cellular networks,” in Proc. of The International Symposium on Wireless Communication Systems (ISWCS), Port Luis, Mauritius, 2004.

[7] Y. Iraqi, and R. Boutaba “An Adaptive Distributed Call Admission Control for QoS-Sensitive Wireless Mobile Networks” Proceedings of the 2000 IEEE Wireless Communication and Networks Conference WCNC. pp 449-453 Chicago, IL. Sept. 2000.

[8] Y.C. Lai and S. F. Tsai “ A Fair Admission Control for Large bandwidth Multimedia Application,” in Proc. of GLOBECOM, 312-322, November 2002.

[9] D. Levine, I. Akyidz, and M. Naghshineh, “ A Resource Estimation and Call Admission Algorithm for wireless Multimedia Networks Using the Shadow Cluster Concept”. in IEEE/ACM Transactions on Networking, 5(1), February 1997.

[10] M. A. Aboleaze, Fadi A. Aloul, “ A Call Admission Protocol for Wireless Cellular Multimedia Networks”, Proc. Int conf. WWC 2005.

[11] Li Zhong Li, Bin Li, Xi-Ren Cao, Bo Li, “ Call Level Performance Analysis For Multi-Services Wireless Networks”, in Proc. of IEEE conference on communication 1002-1007,2003

[12] Genetic Algorithms in Search, Optimization, and Machine Learning – David E. Goldberg