Journal of Mobile Multimedia, Vol. 3, No.2 (2007) 168-178

REOURCE MANAGEMENT AT CONNECTION LEVEL FOR MULTIMEDIA IN WIRELESS/MOBILE CELLULAR NETWORKS

M. SANABANI, S. SHAMALA, M. OTHMAN, and Z. ZUKARNAIN

University Putra Malaysia (UPM), Selangor, Malaysia
m.sanabani@gmail.com

Received October 25, 2006
Revised March 26, 2007

Quality of Service (QoS) guaranteed for multimedia services in next generation wireless/mobile cellular networks poses great challenges due to the scarce radio bandwidth. Therefore, the demand for effective and careful resource allocation is immensely needed. In this paper, we combine Call Admission Control (CAC) and Adaptive Resource Allocation (ARA) algorithm into a novel framework, which we call Connection-Level Resource Controller (CLRC) for provisioning connection-level QoS in Multimedia wireless/mobile cellular networks. Simulation results show that the framework is able to reduce Handoff Connection Dropping Probability (HCDP) for active users of real time services to zero level. Thus, it satisfies mobile users' needs resulting in a stable performance levels during heavy load periods. Furthermore, the framework provides a low New Connection Blocking Probability (NCBP), which is translated into high resource utilization. This is a highly desirable property from the service provider point of view.

Key words: Wireless cellular networks, resource management, quality of service, acceptance
Communicated by: I. Ibrahim and A. Hafid

1 Introduction

Future generations of wireless/mobiles cellular networks such as third-Generation (3G) and fourth-Generation (4G) are expected to support a varity of applications have diverse bandwidth requirement. This network will serve multiple classes of connections with each class having distinctively different Quality of Service (QoS) requirements. One side of the spectrum is real-time services requiring strong bandwidth guarantees as Real-Time Constant Bit Rate (RT-CBR). The other side of the spectrum is services which are adaptive in nature and can operate over a wide range of bandwidth as Real-Time Variable Bit Rate (RT-VBR) services and Non-Real-Time Unspecified Bit Rate (NRT-UBR) services.

In wireless/mobile cellular networks, a mobile user can freely roam within a network's coverage area, and may undergo a large number of handoff events during a typical session. When a user handoffs to a new cell, there may not be sufficient bandwidth to support his call. Under such a situation, there are two possibilities based on the nature of the service. In the case of a non-adaptive service, i.e., the bandwidth of a call is fixed throughout its life time, needing strict bandwidth guarantees, the call will be dropped. Whereas in the case of an adaptive service the call will not be dropped but will suffer bandwidth degradation.

Connection-level QoS of user in wireless/mobile cellular networks is usually quantitatively expressed in terms of probabilistic parameters such as New Connection Blocking Probability (NCBP) and Handoff Connection Dropping Probability (HCDP) [8]. A new call is initiated when a user requests a new call, while a handoff call occurs when an active user moves from one cell to another neighbouring cell. Thus, the NCBP is the probability of a new arriving call being rejected while the HCDP is the probability that an accepted call is terminated before the completion of its service, i.e., the probability that a handoff attempt fails. Providing multimedia services with QoS guarantees in wireless/mobile cellular networks presents great challenges due to (i) the limited bandwidth; and (ii) the high rate of handoff events as the next generation of wireless/mobile cellular networks will use micro/pico cellular architectures in order to provide higher capacity.

Recently, several Call Admission Control (CAC) algorithms and bandwidth adaptation algorithms have been proposed in wireless networks for QoS provisioning [1, 3, 4, 10, 11]. In [1, 4], it is assumed that all calls belong to a single class of adaptive multimedia traffic and receive varying bandwidth assignments from a discrete set of integer bandwidth values. An analytical model for wireless networks with adaptive resource allocation and a traffic restricting CAC is derived in [1]. In [4, 11], new QoS parameters for adaptive multimedia in wireless networks are introduced. Cell overload probability parameter is proposed in [11]. Multiple classes of adaptive multimedia services in cellular wireless networks have been introduced in the literature [3, 10] without considering the prioritization between new call arrivals and handoff calls traffic. However, in our framework we provide a reallocating policy that takes into account the separation between incoming traffic for each class and prioritizes handoff calls over new calls. A prioritization in the process of bandwidth adaptation among multiple classes of multimedia services is presented in [3] where the bandwidth of calls with lower priority is preferably adapted. However, in this approach the authors assume no handoff dropping which make their work impractical. This is not the case in our work since a handoff call can be dropped if it doesn't get the minimum bandwidth.

The main contribution of this paper is the introduction of a novel Adaptive Resource Allocation (ARA) algorithm for provisioning connection-level QoS in wireless/mobile cellular networks. The algorithm operates at the connection-level where the bandwidth of ongoing calls can be dynamically adjusted. Here, we combine CAC algorithm and ARA algorithm with micro-level into a new framework called Connection-Level Resource Controller (CLRC) that will always adapt for the new and handoff calls. Our specific objective in designing the CLRC framework is reducing the NCBP and the HCDP. This objective is essential from the point of view of the user. The service provider, on the other hand, is more interested in increasing revenues. For this purpose, we design ARA with a micro-level which allows the system to offer services whenever there is insufficient amount of bandwidth by intelligently adjusting bandwidth of ongoing calls which results in small values of the NCBP and the HCDP. Therefore, more calls are able to complete their services and as a result better utilization is obtained which translate into more revenue to the service provider. This framework supports multiple classes of fixed and adaptive multimedia services with diverse QoS requirements. The rest of this paper is organized as follows. Section 2 describes in detail the CLRC framework. Section 3 presents the system model. Section 4 reports the simulation results. Finally, conclusions and future works are presented in Section 5.

2 CLRC Framework

In this section, we describe the CLRC framework in detail (see Figure 1).

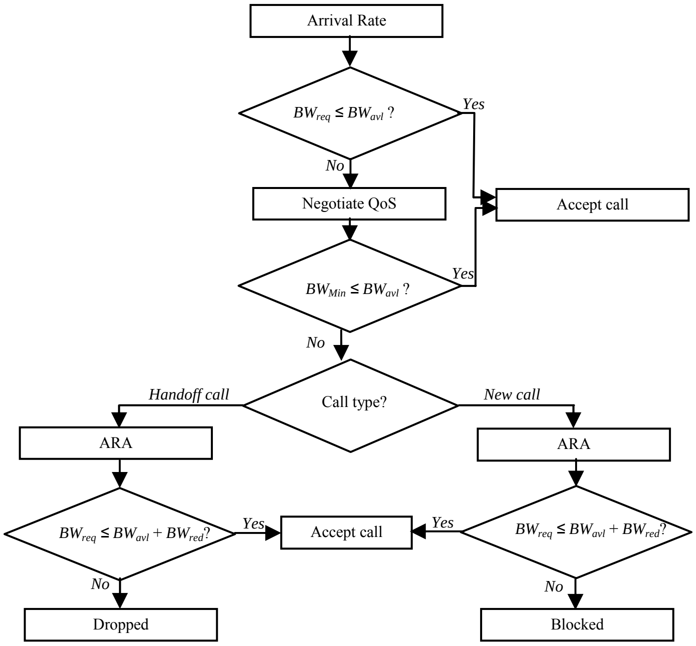

Figure 1: CLRC Framework

2.1 Reallocating Policy

Ideally, every call in a cell should allocate the maximum bandwidth (BW_{Max}) whenever possible. However, if the cell is over-loaded, some of the calls in the cell might receive a bandwidth lower than the requested bandwidth. When a new call or a handoff call arrives, some of the calls already in the cell are made to lower their bandwidth (the minimum bandwidth is BW_{Min}) to accommodate the newly arrived call. On the other hand, when a call completes or handoffs to other cells, some of the remaining calls in the cell may be provided an increase in their bandwidth (the maximum bandwidth is BW_{Max}).The proposed reallocating policy use different of above ideas to allocate, increase, and decrease bandwidth for the calls in a cell.

Under conditions of heavy traffic load, i.e. the sum of the requested bandwidth exceeds the unused bandwidth capacity so that not all the requests can be completely served, the role of bandwidth adaptation technique are essential. These algorithms are needed to reduce the requested or already connected call bandwidth allocation. In designing the algorithm, we assume that a service with degraded QoS is better than an outright rejection of service requests. The bandwidth reallocation (i.e. degradation or upgrading of resource allocation) module is deployed to reallocate the bandwidth capacity. The process of reallocating may involve either an upgrade or degrade of the bandwidth allocation based on the micro Acceptable Bandwidth Level (micro-ABL).

The micro-ABL is obtained by subtracting the maximum required bandwidth with the minimum required bandwidth. The difference is called the degradable range/spectrum. The degradable spectrum is further divided into N (i.e. where N = 1, 2 …n) levels, called micro-ABLs as shown in Figure 2. The concept of bandwidth allocation as a discrete component is applied into the structuring and derivation of the micro-ABLs. Subsequent to this theory, the bandwidth allocation for the micro-ABLs form the discrete set B= $\{BW_{Min}, BW_{Min+1}, BW_{Min+2},\ldots\ldots,BW_{Avg},\ldots\ldots,BW_{Max}\}$ where $BW_{Min+i} < BW_{Min+(i+1)}$, BW_{Min} is the minimum bound , BW_{Avg} is the average bound and BW_{Max} is the maximum bound for bandwidth allocation.

The computation and deployment of reallocating policy consumes an amount of time capable of jeopardizing the probability of a mobile to continue is connection. Thus, to avoid this time delay, the process of bandwidth reallocation is carried out in a distributed manner. Each Base Station (BS) does the computation process independent from other BSs. An important pre-requisite is to ensure that the computation process of bandwidth reallocation algorithm should be completed before the system does the real bandwidth reallocation of the ongoing connections in the network. The essence of the proposed CLRC is the cohesive working nature between the resource allocation/reallocation and the CAC.

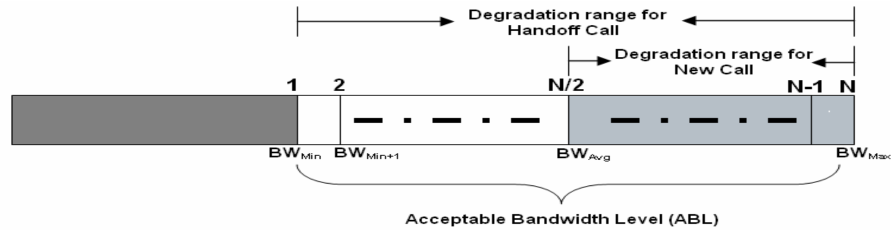

Figure 2: micro-ABLs range/spectrum

2.2 Call Admission Control

A CAC algorithm is a key factor that enables efficient system resource utilization while ensuring that connection-level QoS requirements are satisfied. CAC is always performed when a mobile initiates communication in a new cell, either through a new call or a handoff.

The CAC accepts an arrival call being new or handoff, the system attempts to allocate maximum bandwidth as bandwidth requested (BW_{req}) for this call. Thus, if the $BW_{req} \leq BW_{avl}$ (available bandwidth) for new or handoff calls, the arrival call will allocate a bandwidth requested. Otherwise, insufficient bandwidth a negotiate QoS process is invoked.

2.3 QoS Negotiate Process

This process negotiates the bandwidth of current arrival. The bandwidth of RT-VBR and NRT-UBR applications is adjusted from BW_{req} to the minimum bandwidth (BW_{Min}). Then CAC accepts the arrival call if the $BW_{Min} \leq BW_{avl}$ for new or handoff calls. That means that the arrival call will allocate a bandwidth not to be less than the minimum bandwidth. Otherwise, the system attempts to trigger ARA algorithm when QoS negotiation is ineffective based on type of call.

2.4 ARA Algorithm

The ARA performs two main procedures: degradation and upgrading. The degradation procedure is triggered when an accepted arriving call (new or handoff) arrives to an overloaded cell. The upgrading procedure is triggered when there is an outgoing handoff call or a call completion in the given cell.

2.4.1 Degrading Procedure

If the MT can not get at least the BW_{Min} because the $BW_{avl} < BW_{Min}$, a degradation procedure is triggered to degrade the bandwidth of some ongoing calls in the cell to attempt to allocate average bandwidth (BW_{Avg}) as follows; calls with the largest allocated bandwidth greater than BW_{Avg} and lower or equal priority to new arrival call are degraded to have lower bandwidth not less than BW_{Avg}. If the saved bandwidth (BW_{avl} + reduced bandwidth (BW_{red}) from ongoing connections) is larger than or equal to bandwidth requested, the new arrival call will be allocated a bandwidth requested. Otherwise further bandwidth degradation to accommodate the call cannot be performed. If the saved bandwidth is larger than or equal to BW_{Min}, the new arrival call will allocate a bandwidth not less than the BW_{Min}. If all above tests is not complied then blocking of new arrival call is done. Whereas, a handoff call; a degradation procedure is triggered to degrade the bandwidth of some ongoing calls in the cell to attempt to allocate BW_{Avg} as follows; calls with the largest allocated bandwidth greater than BW_{Min} and lower or equal priority to handoff call are degraded to have lower bandwidth not less than BW_{Min}. If the saved bandwidth is larger than or equal to the bandwidth requested, the handoff call will allocate the bandwidth requested. Otherwise, if the saved bandwidth is less than BW_{Avg} and no further bandwidth degradation to accommodate the call, the system will attempt to allocate a bandwidth at least BW_{Min} of the handoff call. If all above tests fail, then the dropping of the handoff call is done. In general, the degradation procedure follows these rules as follows:

- Only the ongoing calls in classes with lower or equal priority than the requested class priority, starting from the lowest priority class to equal priority class.
- Within the class, the calls receiving best service are degraded first. Subsequently, the calls are chosen according to ascending order of their micro-level of service.
- Only one micro-ABL of bandwidth can be degraded from ongoing call each time.
- The ongoing calls only used their average bandwidth and the minimum bandwidth cannot be degraded for new call request and handoff call, respectively.

2.4.2 Upgrading Procedure

As a call leaves the cell, whether outgoing handoff call or a call completion, the total available bandwidth increases. The system will invoke the bandwidth upgrading procedure to increase the bandwidth for one or more of the degraded calls to the bandwidth requested. The upgrading procedure starts from the higher priority and the most degraded calls in the cell. To achieve fairness among calls within the same class and among calls from different classes, the available bandwidth are divided into several shares and redistributed according to the following rules:

- The class with highest priority should be upgrade first.
- Within a class, the calls served worst should be upgraded first.
- The available bandwidth should be reallocated to as many as possible calls.

The upgrading procedure is terminated when there is no available bandwidth or every call in the cell has a bandwidth larger than or equal to bandwidth requested.

3 Performance Evaluation

In this section, we investigate the performance of the proposed framework that is assessed in its ability to provide adaptive QoS as compared to a non-adaptive scheme, adaptive scheme and Xiao scheme. The non-adaptive scheme is simulated assuming that a call must allocate its maximum bandwidth to be admitted and once accepted. However, its bandwidth cannot be changed throughout the lifetime. Then if such bandwidth is not available, the call is either blocked or dropped depending on whether the call is a new or handoff call. The selected non-adaptive scheme is from [9], a modified as adaptive scheme is from [7] that invoke QoS negotiate process without ARA algorithm when there is insufficient bandwidth for arrival (new or handoff) call and Xiao scheme is from [10, 11].

The performance metrics utilized encompasses two QoS parameters: NCBP and HCDP and bandwidth utilization. In most of the simulation results that follows, the performance measures are plotted as a function of the call arrival rate. The call arrival rate is the arrival rate of new calls measured as the average number of new call requests per second per cell. Before proceeding, we first describe the simulation model that is used in this paper.

3.1 Simulation Model

A discrete-event simulation model for a wireless cellular network environment in a two-dimensional (2-D) topology in which each cell has exactly six neighbouring cells is extensively developed using C++. The simulated area consists of 64 Omni-cells and has a uniform geographic distribution. It is assumed that a base station is located in the centre of a hexagonal. Fixed Channel Allocation (FCA) is assumed in the system.

3.1.1 Multimedia Services Model

In order to represent various multimedia applications these are assumed based on the call duration, bandwidth requirement and class of service. The classification strategy is adopted from [5, 6, 7]. The different application groups include real time (RT-CBR and RT-VBR) and data traffic sources non real time (NRT-UBR). The six applications are carefully chosen for a simulation; they are typical traffic seen in wireless networks and their respective parameter values are chosen from Table 1. The call duration is assumed to follow a geometric distribution between the minimum and maximum values and the bandwidth required values shown in the Table 1.The value closely represents realistic scenarios.

3.1.2 Traffic Model

A newly calls from the three groups (RT-CBR, RT-VBR and NRT-UBR) are generated with equal probability and appear anywhere in the cell. The traffic of the call requests are generated according to a Poisson process with rate λ_{new} (calls/second/cell) in each cell. The probability of a call belonging to traffic classes 1, 2...or 6 are 1/6, 1/6,1/3,1/9/1/9 and 1/9 respectively. Two types of calls (new and handoff calls) share the bandwidth of the cell. The parameters values used in the simulation shown in Table 1.

3.1.3 Mobility Model

Three parameters for mobility model are considered, as well as the initial position of a mobile, its direction and its speed. A newly generated call can appear anywhere in the cell with an equal probability. When a new call is initiated, a mobile is assigned a random initial position derived from a uniform probability distribution function over the cell area. To determine the cell coverage and handoff region threshold, the propagation model proposed by [2] is adopted where the Received Signal Strength (RSS). The mobile coordinate position can be used to determine the location of a mobile in the cell. When the base station detects that a mobile has entered the handoff region (by comparing its

RSS with certain threshold), the current base station sends a bandwidth request to the new base station (one of the neighbouring cells of current cell). That is as towards it the mobile is heading to in order to pre-allocate a bandwidth for expected handoff event. A mobile can travel in one of 8 directions with equal probability. As for handoff calls, the initial position is determined when the handoff event is done. A mobile is assigned a random direction upon entering a cell. A constant randomly selected speed is assigned to a mobile when it enters a cell either at call initiation or after handoff.

Table 1: Multimedia Traffic and system parameters used in simulation

Traffic Type-Class No.	**Type of Media**	**Bandwidth Requirements**		**Duration of call**		**Priority**
		Min.	**Max.**	**Min.**	**Max.**	
RT-CBR-1	Voice service and audio phone	30 Kbps	30 Kbps	1 min	10 min	6 *(Highest)*
RT-CBR-2	Video-phone and video-conference	256 Kbps	256 Kbps	1 min	30 min	5
RT-VBR-1	Interactive multimedia and video on demand	1 Mbps	6 Mbps	5 min	300 min	4
NRT-UBR-1	Email ,paging and fax	5 Kbps	20 Kbps	10 sec	2 min	3
NRT-UBR-2	Remote login and data on demand	64 Kbps	512 Kbps	30 sec	10 hour	2
NRT-UBR-3	File transfer and retrieval service	1 Mbps	10 Mbps	30 sec	20 min	1

Description	**Value**
Number of cells	64(8x8) Warp around
Cell radius	200m
Handoff region	10m
Max. bandwidth capacity of cell	30Mbps
Probing time interval	1 sec
The probability of moving direction	(0.5) Forward, and (0.5) Changing the direction
Direction of mobile	8 directions with 6 neighbouring
Speed of mobile	1 m/s – 10 m/s

4 Results and Discussions

The impact of the traffic nature is analyzed as it is needed by the indicator of the differentiated services. The scenario analyzed is the uniform traffic. The uniform traffic distribution in the simulations, where the traffic load is the same among all 64 cells is utilized. The effect of equal traffic loading in different cells can be shown in Figures 3-7.

Figures 3-4, depicts the NCBP and HCDP for real time traffic (RT-CBR and RT-VBR) versus the traffic load for proposed scheme and others schemes (non-adaptive, adaptive and Xiao scheme), respectively. Clearly, the NCBP performance for proposed scheme is better than that for others schemes due to the ARA algorithm employed in CLRC framework. The result indicates that the HCDP for proposed scheme and non-adaptive scheme is near zero under low traffic loads. This is trivially true by the nature of the algorithm. However, if the traffic is very heavy, HCDP is non-zero for others schemes, HCDP is significant even under low traffic loads. We observe that our proposed scheme outperforms than other scheme in terms of effective characterization of NCBP and HCDP for real time traffic.

Figures 5 depicts the NCBP and HCDP for non-real time traffic versus the traffic load for proposed scheme and others schemes (non-adaptive, adaptive and Xiao scheme), respectively. Clearly, the NCBP and HCDP performance for proposed scheme is better than that for non-adaptive scheme until system load λ= 3.5 and λ= 4.9, respectively. But in heavy load the performance for non adaptive scheme outperforms for proposed scheme as expect. This is because proposed tends to gradually reduce the operational bandwidth of the on-going non real-time connections for higher priority traffic class and therefore, preventing new non real-time calls from entering the system and cutting off handoff non real-time calls from continuing in the system as reflected in Figure 5. Also, the NCBP and HCDP performance for proposed scheme is better than that for adaptive scheme because our CLRC framework invokes the ARA algorithm in conjunction with the traffic-based CAC algorithm to keep the non real time NCBP and HCDP at a low rate. The Xiao scheme is better than proposed scheme in term of NCBP performance but in term of HCDP performance the proposed scheme is better. This is because the Xiao scheme doesn't consider the priority of handoff calls over new calls. Moreover, it also doesn't assign priority levels to the different types of traffic. Therefore, the real time traffic has no impact on non real time traffic.

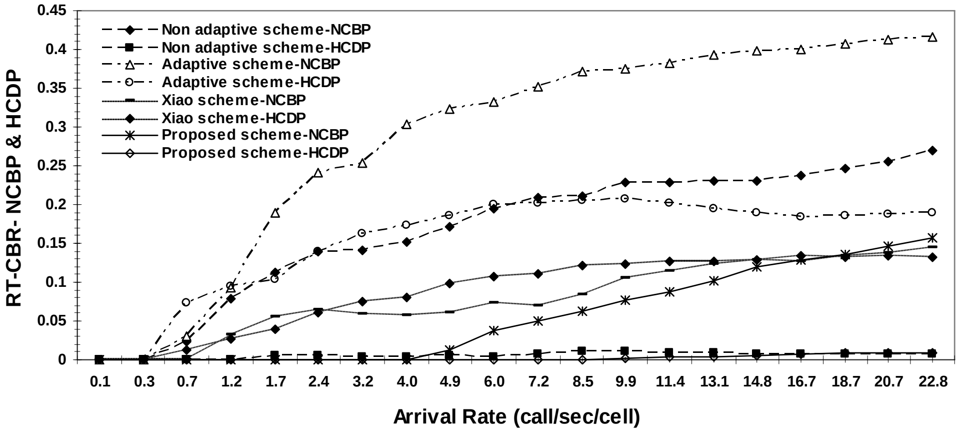

Figure 3: Real time (CBR) - NCBP and HCDP

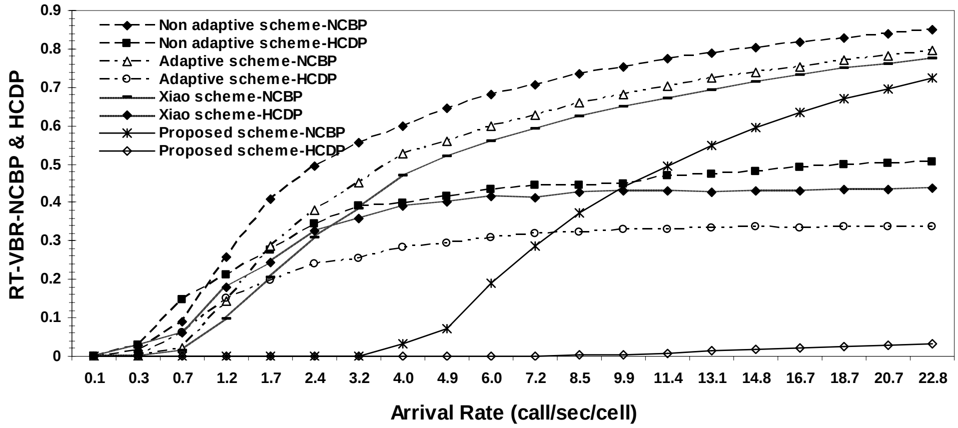

Figure 4: Real time (VBR) - NCBP and HCDP

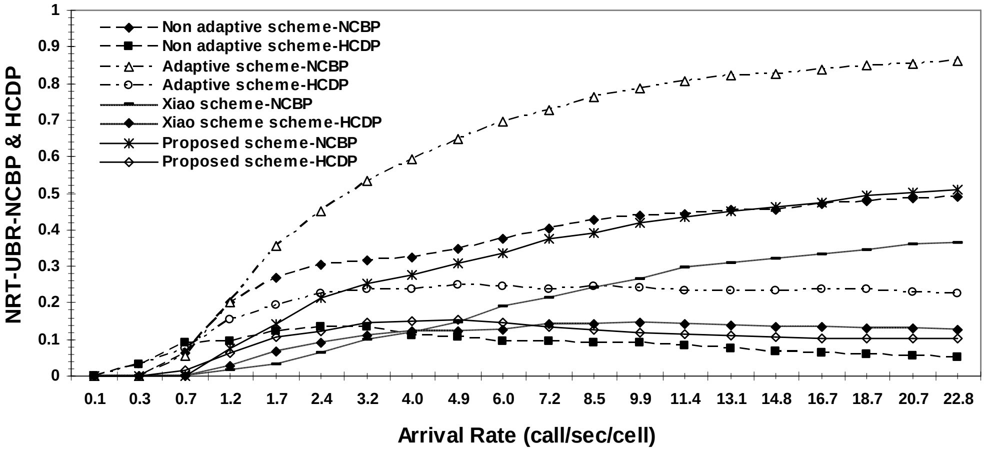

Figure 5: Non real time (UBR) - NCBP and HCDP

Figure 6 shows the performance comparison for all traffic services in terms of NCBP and HCDP. The proposed scheme outperforms the other schemes in terms of effective characterization of NCBP and HCDP. As the system becomes overloaded, our CLRC framework invokes the ARA algorithm in conjunction with the traffic-based CAC algorithm to keep the real-time NCBP and HCDP at a low rate. As the global traffic load increases, the average amount of available bandwidth decreases. Thus, new call requests are likely to be rejected and NCBP increases, but the HCDP of handoff calls quickly settles down due to higher priority then new calls.

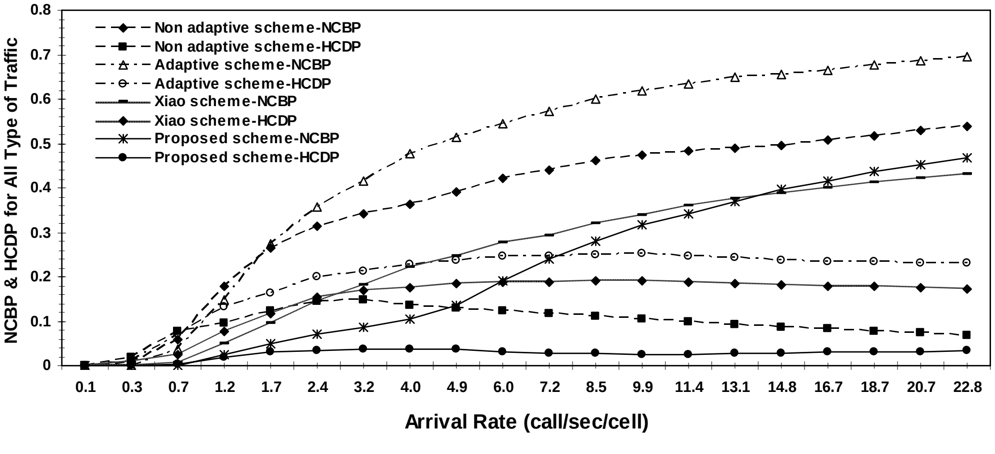

Figure 6: NCBP and HCDP versus the traffic load for all applications

Figure 7 shows the bandwidth utilization versus the traffic load for proposed scheme and other schemes. Clearly, the utilization for proposed scheme is better than that for other schemes. When the traffic load becomes higher, the advantage is more evident. The reason is due to the very adaptive nature of these services that allows the network to offer services whenever there is sufficient amount of resources by intelligently adjusting resource allocation. Also, micro-level offers more flexibility of bandwidth adjustment, thus resulting in better resource utilization.

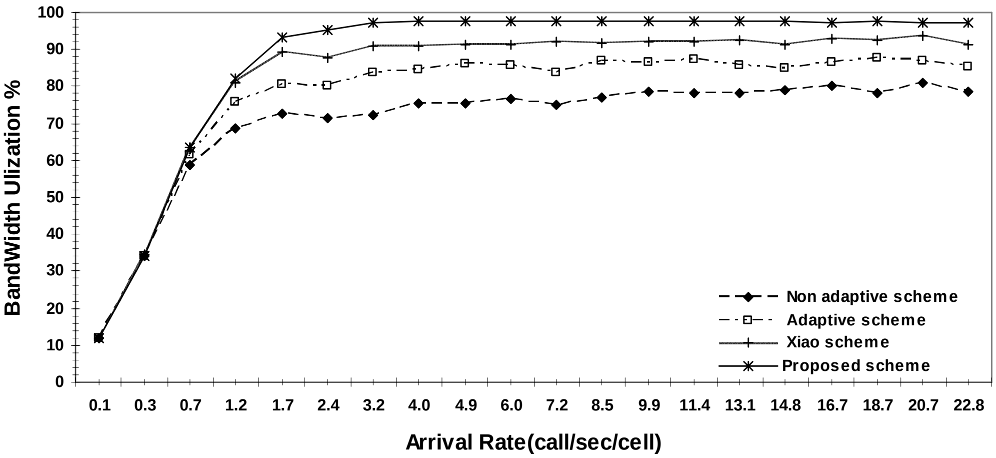

Figure 7: Bandwidth Utilization

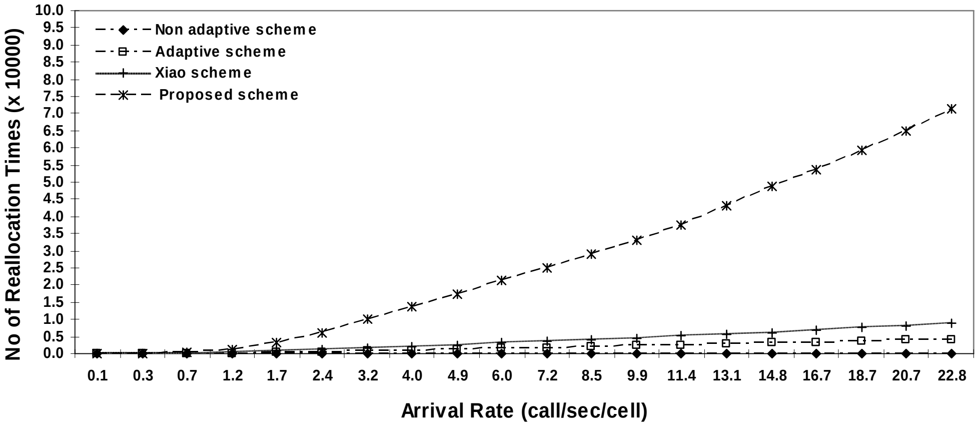

Figure 8: Number of Bandwidth Reallocation

Figure 8 shows the number of bandwidth reallocations for proposed scheme and other schemes. Non adaptive scheme performed zero number of reallocation because no adapted algorithm is performed while the adaptive scheme performed the lower than the other schemes because the adapted algorithm is limited to perform very crude reallocation. It is either a simple accept or reject with only one time bandwidth allocation, i.e. from BW_{Max} to BW_{Min}. For both Xiao and proposed schemes, they started off well with same number of reallocations when the load is low. From load λ=0.9 onwards, the proposed scheme started to reallocate more frequently than Xiao scheme as expected because the proposed scheme reallocate all the ongoing calls that have lower or equal priority for incoming call while Xiao scheme reallocate all ongoing calls that have the same priority only for incoming calls. The benefit which gets from the results of the proposed scheme may justify the cost of increased complexity that is represented in the number of reallocation times.

5. Conclusions and Future Work

In this paper, CLRC framework is proposed. It consists of two main components: CAC and ARA algorithms. The framework is designed for wireless/mobile cellular networks that support real-time multimedia services. The simulation results show an improvement and reduced values for the connection-level QoS parameters: HCDP and NCBP. The requirements of the mobile users are hence satisfied. Moreover, the results show that CLRC framework can utilize the bandwidth more efficiently which is highly desirable from the point of view of service providers.

The non-real time traffic suffered because highly priority of real-time traffic. That is why the NCBP and HCDP of non real time calls become higher in Figure 5. One way to deal with this problem is to introduce queue into the system. Since non-real-time calls usually may have delay tolerance to some extent, it is better to put them in a waiting queue rather the blocking or dropping them off. When resource is available, they can start or continue their connection.

References

1. Chou, C. and Shin, K., Analysis of combined adaptive bandwidth allocation and admission control in wireless networks. in Proceedings of IEEE INFOCOM 2002, **2** 676 –684.
2. Ebersman, H.G. and Tonguz, O. K., Handoff Ordering Using Signal Prediction Priority Queueing in Personal Communication Systems. IEEE Transactions Vehicular Technology 1999, **48** (1) 20-35.
3. Kim, S., Call admission control for prioritized adaptive multimedia services in wireless/mobile networks. in Proceedings of the IEEE Vehicular Technology Conference (VTC), (Tokyo-2000), 1536 –1540.
4. T. Kwon, Y. Choi, C. Bisdikian, and M. Naghshineh , QoS provisioning in wireless/mobile multimedia networks using an adaptive framework. Kluwer Wireless Networks, 2003, **9** (1) 51-59.
5. Malla, A., El-Kadi, M., Olariu, S. and Todorova, P., A fair resource allocation protocol for multimedia wireless networks. IEEE Transactions on Parallel and Distributed Systems, 2003, **14** (1) 63-71.
6. Ning L. and John B., Utility-maximization bandwidth adaptation for multi-class traffic QoS provisioning in wireless networks. in Proceedings of ACM workshop on Quality of service & security in wireless and mobile networks (Q2SWinet '05), (Montreal-2005), 136-143
7. T. Oliveria, C. J. B. Kim and T. Suda, An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks. IEEE Journal on Selected Areas in Communications, 1998, **16** (6) 858–874.
8. Rappaport, S.S., The multiple-call handoff problem in high-capacity cellular communications systems. IEEE Transactions Vehicular Technology, 1991, **40** (3) 546–557.
9. Soh, W.S. and Kim, H. S., Dynamic Bandwidth Reservation in Cellular Networks Using Road Topology Based Mobility Predictions. in Proceedings of IEEE INFOCOM (2004), 2766-2777.
10. Y. Xiao, C. Chen, and Y. Wang, Fair bandwidth allocation for multiclass of adaptive multimedia services in wireless/mobile Networks, in Proceedings of the IEEE Vehicular Technology Conference (VTC2001), 2081-2085.
11. Y. Xiao, C. L. P. Chen, and B. Wang, Bandwidth degradation QoS provisioning for adaptive multimedia in wireless/mobile networks, Computer Commu., 2002, **25** (13) 1153- 1161.