

A Priority Based Call Admission Control Protocol with Call Degradation for Cellular Networks

Mokhtar Aboelaze¹
Dept. of Computer Science and
Engineering
York University
Toronto, Canada

Ayman Elnaggar
Dept. of Electrical Engineering
Sultan Qaboos University
Oman

Maan Musleh
Dept. of Computer Engineering
The American University in Dubai
Dubai, UAE

Abstract—Call Admission Control (CAC) protocols play a very important role in the performance of wireless networks. In this paper, we present a call admission control protocol for cellular wireless networks. Our protocol depends on degrading the existing calls by reducing the bandwidth allocated to them in order to admit “important” calls. Our protocol assigns priorities for the incoming calls, and in the same time assigns priorities to the existing calls, both admitted calls are admitted according to their priorities and the existing calls are degraded according to their priorities. We show simulation results for the relation between network utilization, call blocking probability, and average assigned bandwidth during the life of the call.

Keywords—; call admission protocols; cellular networks; QoS; call degradation

I. INTRODUCTION

Wireless devices and connectivity through wireless networks are growing at an astonishing rate. Wireless networks are not only used for cellular phone applications, they are carrying different types of traffic, voice, video, and data [3]. They range in coverage from desktop area networks to national cellular networks. Users of such networks expect a specific Quality of Service QoS depending on their application and service contracts. The network must guarantee the QoS requested by each user, and in the same time maximize their revenue by maximizing the number of users (calls) admitted to the network.

Admission control policies play a crucial role in the performance of any network. When deciding to admit a call to the network there are many factors to be taken into consideration most of them are contradictory (network utilization, revenue, QoS, fairness, ...). The call admission control works in real-time, the algorithm used should be suitable for real-time implementation using the limited resources of the base station controller [15] There has been a lot of research in call admission that could be summarized as follows.

FCFS is probably the simplest of all call admission protocols, in this method if a call arrives and there is enough bandwidth to accommodate it, it is accepted, otherwise rejected. This method has shown to produce a good utilization of the bandwidth however it has been shown to be biased against calls with high bandwidth requirements. As a way of

introducing some priority in FCFS, bandwidth is divided into segments and call requests are grouped into different categories such that a call request from category i can only be admitted if there is enough bandwidth in segment i otherwise rejected. The main problem with this technique is the waste of bandwidth since we could have enough bandwidth in one segment but a call is rejected in another segment.

In [5], the authors proposed a general framework for bandwidth degradation and call admission in a multi-class traffic wireless network. Their objective is to model the changes in the revenue of the network due to admitting new users, in the same time they estimated the cost of degrading an ongoing call and considered it to be a negative revenue. The overall objective is to maximize the revenue. In their analysis they considered an exponential service time for the calls, and a Poisson arrival pattern. They did not differentiate between new calls and handoff calls, neither they considered the effect of handoff at all

In [17] The authors used the direction of the movement to predict the next cell and make an early reservation before the actual handoff occurs. They used both threshold distance and threshold time for early reservation. They assumed a real-time positioning technology in order to determine the position and the direction of movement of the user. They also used *channel borrowing* in the sense of migrating channels from cold cells to hot cells (within constraints) not in the sense of borrowing bandwidth from existing calls as we use in this work.

El-Kadi et al in [6] proposed a new call admission scheme. Their protocol depends on dividing the connections into real-time, and data connection. The real time connections are guaranteed at least a minimum B.W. where the data connections are assumed to tolerate a large delay and there is no guarantee minimum B.W. In their protocol, a handoff call is admitted if there is enough B.W. in the cell, or if the minimum required bandwidth for this call could be achieved by borrowing from other connections in the cell. However, in their protocol they did not support any priority schemes either for the admitted call, or the calls they borrow from.

In [9] the authors proposed a 2-level call admission scheme. In this protocol, The protocol is divided into 2 parts, Basic Call Admission Control (BCAC), and Advanced Call Admission Control (ACAC). The BCAC determine the admission based

¹ This work was done when the first author was visiting the American University in Dubai.

This work is supported by a grant from national Science and Engineering Research Council of Canada NSERC.

on the availability of B.W. The ACAC determine the call admission by utilizing the delay tolerance and priority queue algorithm, Depending on the type of the call blocked by the BCAC.

In [10] the authors proposed a novel but rather simple protocol in order to improve the fairness, they proposed the use of a single buffer in order to hold the call request if there is not enough bandwidth. If a call arrives and there is enough bandwidth, it is admitted right away. If there is not enough bandwidth, and the buffer is free, the call is admitted to the buffer and waits until there is enough bandwidth, then it is admitted and the buffer is cleared. If a call arrives and another call is waiting in the buffer, the new call is rejected. They proved that under a Poisson arrival and exponential service time, their protocol is optimally fair. However, one of the major drawback of their protocol is that it leads to a drop of utilization when there is a big difference in the bandwidth requirements of the different classes. The protocol also does not support priorities.

In [13], the author presented a study of the reliability of call admission control assuming an $M/M/c$ loss system at the call level and a bufferless link. The definition of reliability is the probability that the admitted call, or any existing call in the system, will receive less than its prescribed QoS. The author showed that many CAC proposed schemes are not acceptably reliable.

In [1] the author proposed a call admission protocols that can support differentiated fairness and maintain a good resource utilization for calls that require different and widely varied bandwidth. However the author assumed a fixed bandwidth for every call and no call degradation is allowed in order to admit new calls.

A new call admission protocol for voice and best-effort data is proposed in [11]. They proposed a 3-D birth death model that can effectively model both on/off voice traffic as well as best-effort data traffic. Their algorithm depends on searching the optimal admission region in order to maximize the resource utilization and in the same time not to violate the requested QoS of the individual users.

Admission control for integrated voice and data service using CDMA was proposed in [16]. In [12] a call admission protocol is proposed with the objective of guaranteeing the overall loss ratio of the system. Call admission protocols that takes into account the user velocity and the direction of movement in allocating bandwidth not only in the current cells but in the cell the user is moving into are studied in [7] and [8]. In [2], the authors proposed a QoS Provisioning for admission control for multimedia wireless networks.

The organization of our paper as follows, Section II states the mode we use in our simulation for both the network and the traffic. Section III explains our protocol. Section IV shows the protocol performance under different load and traffic assumptions, the paper ends with a conclusion and future work.

II. THE MODEL

In this section, we present the network model and the traffic model we used throughout our simulation.

A. Network Model

We assume a cellular network architecture where every cell is served by a base station. Base stations are connected together by using a wireless or wireline network. Users are roaming in the coverage area and when moving from a cell to another cell, handoff occurs. The call admission control protocol is responsible for deciding whether to admit or to block a new or a handoff request.

When the network is congested and the CAC decides to turn down an admission request, that is called a blocked call. However, if the admission request is coming from an active call in a neighboring cell that is moving into the cell's coverage area, that is called a dropped call. From the customer point of view, blocking a call is much more tolerable than dropping an active call. The admission control protocol must take this into consideration.

Although we did not study in this paper the handoff. However, this issue can be handled easily in our model as follows. Our protocol, as we will see in the next section, can support different priority levels for calls requesting admission. The handoff calls can be represented in our analysis as a high priority call. Our protocol does not require any information about the movements of the mobile terminals, it just treats handoff requests as a high priority which results in a preferential treatment for these calls.

We assume a Poisson arrival for calls. The service time for every call is exponentially distributed. Different classes of calls with different priorities, different arrival rate and service time, and different bandwidth requirements are considered. The call blocking ratio for each class and the resource utilization (network bandwidth) are the main parameters of interest.

B. Traffic model

We assume different types of traffic (audio, video or data). However for every type, we assume different classes or categories of users. For example consider a mobile terminal transmitting video, the quality of service offered to the terminal depends on the price paid for the service and can range over a wide range affecting the quality of the received video. The different types of traffic are characterized by different arrival rates, service rates, and bandwidth requirements.

Many applications especially audio and video transmission can support a variable bit rate. For example, MPEG-4 supports very low bit rate coding with bandwidth requirements of 5-64Kbps [3], while in Audio, using silence detection and sophisticated coding techniques [4] results in encoding that supports variable bit rate.

In this paper, we assume that every class has a maximum (requested) bandwidth, and a minimum bandwidth. The assumption is that, this class of traffic can supports degraded performance down to the minimum bandwidth. The network can, if needed, borrow some bandwidth from any user with a

condition of not violating its minimum bandwidth requirements. Constant bit rate sources are a special case where the minimum bandwidth is the same as the maximum bandwidth.

III. PROPOSED PROTOCOL

We assume n classes of users, such that users in class i require a maximum bandwidth of B_{maxi} , and a minimum bandwidth of B_{mini} . The difference between these two values is the degradation that a user in class i can tolerate (degradable bandwidth for class i). The degradable bandwidth ($B_{maxi} - B_{mini}$) for a call in class i is divided into λ_i segments, each segment is $(B_{maxi} - B_{mini}) / \lambda_i$. λ_i determine the granularity of the bandwidth borrowed from a connection in class i . It also determine the priority of class i connections as we will see in this section. A call is said to be in level s , if there are s segments borrowed from it. The network borrows segments from ongoing calls in a way such that the difference between the states of any two calls that did not exhaust all their degradable bandwidth is at most 1. A network is said to be in state s , if each node that did not exhaust its degradable bandwidth is either level s or $s-1$, and all the calls that have already exhausted their degradable bandwidth do have a number of segment in their degradable bandwidth equal to $s_x < s$. segments. Figure 1 shows the bandwidth requirements for 4 classes. Every class has its own B_{max} , B_{min} , and λ

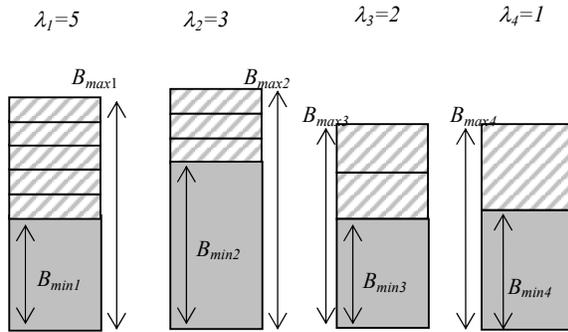


Figure 1. Bandwidth requirements for different types of traffic

Note that λ also determine the priority of class i as far as borrowing bandwidth is concerned. Since the network borrows bandwidth from users in such a way that the difference between any 2 levels in the network is at most 1, which means that is the network borrowed from class 1 a segment, it can not borrow another segment from class 1 unless it borrows a segment from all other calls in progress. That means the CAC protocol will borrow a segment from class 4, (The entire degradable bandwidth of class 4) before it borrows another segment from a call in class 1. Thus giving class 1 a higher priority than class 4 (for bandwidth degradation purpose).

A. Priority

The proposed protocol also supports priority in call admission. Every class has an admission priority p_i a call in class i can only be admitted if the network is in state s such that $s \leq p_i$

Our protocol has the advantage of having priority levels for different callers by treating them differently both in admitting them or not (based on the network state), and by deciding to borrow bandwidth from them. Low priority calls are not admitted if the network is beyond a specific state, also for high priority calls, we can increase the number of segments which leads to a smaller segment, and allowing borrowing is smaller quantities from high priority callers. That leads to a situation that if the network is saturated beyond a specific threshold, we can be sure that we will borrow from high priority calls only to admit another high priority call.

B. Handoff calls:

We opted not to deal with handoff calls any differently than new calls. In reality, handoff calls usually have higher priority than new calls. Customers consider getting a busy signal to be less of a nuisance than a call is terminated abruptly just because the customer moved to another cell. Thus handoff calls are usually given a higher priority in than new calls. However, in this protocol we do have priority levels that are assigned to each category, in this case, handoff calls are considered to have a higher priority than new calls, thus giving them a preferential treatment when it comes to accepting the call or not.

IV. SIMULATION RESULT

We have simulated our protocol and we report our results in this section, the important factors that we considered are the call blocking ratios for the different categories, the average bandwidth assigned to each call, and the utilization. However another factor we must take into consideration that is the number of changes in the bandwidth allocated to the call during its lifetime.

TABLE I. TRAFFIC CHARACTERISTICS

	BW requirements Required, min	Av. Duration	Type
1	30, 30Kbps	3 minutes	Voice
2	256, 256Kbps	5 minutes	Video conference
3	6, 1Mbps	10 minutes	Video on demand
4	20, 10Kbps	30 seconds	E-mail & Fax
5	512, 64Kbps	3 minutes	Data on demand
6	10, 1Mbps	2 minutes	File transfer

Table 1 shows the workload we used in our simulation. We used the workload that was presented in [14], and [6]. The mix of the different types of traffic will be reported for each experiment. Also the amount of degradable bandwidth (B_{min} , B_{max}), the priority levels, as well as the number of segments

will be stated for each experiment. Unless otherwise stated, the number of segments in the degradable bandwidth. λ_i is set to 10, and the probability p_i is 10 for all types.

In our simulation, we used the average bandwidth requirements in [14], but not the maximum. The reason for doing so, is the average and maximum do make sense in packet switching, but in circuit switching (what we are considering here), the bandwidth requested at the time of admission is the maximum, then the bandwidth may fluctuate down to the minimum during the life of the call.

One issue we have to consider is the mix of these different types of traffic. We could not find in the literature any model to predict the mix of traffic in wireless networks. We decided on using a mix that shares the cell bandwidth equally between the different types. The arrival rate for the different calls is adjusted such that the load on the network (number of calls) for the different types is such that the different types of traffic equally share the network bandwidth among them. For example, the arrival rate for type 4 (require 20Kpbs) is one and half times that of type 1 (require 30Kpbs).

Our system model is as follows: 30Mbps total bandwidth per cell. We modeled handoff calls as a different category of calls with the same traffic characteristics as the new calls in the same group but with a higher priority when needed. When we modeled the call blocking probability (and call drop probability for handoff calls) we reported it vs. utilization. The reason for reporting it vs. utilization instead of arrival rate is because arrival rate depends on the total cell bandwidth (an arrival rate of 0.1 calls per minute could be a very high load or a very low load depending on the total cell bandwidth). However reporting it vs. utilization indicates the behavior of the network as the traffic increases and the network becomes congested.

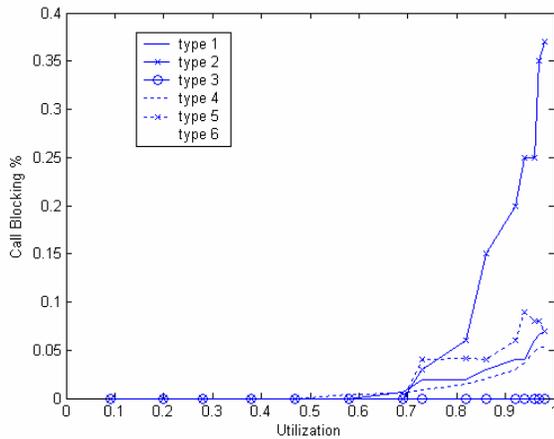


Figure 2. Utilization vs. rejection rate

The utilization in the Figures ranges from 0 to 0.1 (0-100%). However because of the ability of the call admission protocol to reduce the bandwidth of the existing calls in order to admit other calls, the actual load on the network goes beyond the 100% utilization. In our simulation, a 100%

utilization corresponds to a load that would require 150% utilization without the bandwidth degradation.

Figure 2 shows the call blocking probability vs. the network utilization. We can see from the graph that the call blocking probability is kept below 0.5% up to a load one and half times the utilization by using bandwidth degradation. Also the call blocking probability will be the highest for type 2 traffic that requires a constant bit rate of 256Kpbs with no degradation.

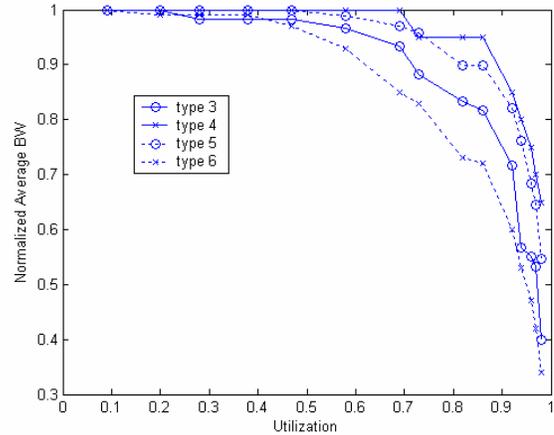


Figure 3. Utilization vs. normalized BW for the different traffic types

Figure 3 shows the normalized bandwidth for types 3 to 6 (types 1 and 2 require a constant bit rate with no degradation and the normalized bandwidth for them is always 1). Since different types of traffic require different bandwidth, normalized bandwidth (ratio between requested bandwidth and assigned one throughout the life of the call) is a good indication for customer satisfiability. Up to about 60% utilization every type is getting almost its required bandwidth. When the network starts to suffer congestion, bandwidth will be degraded for every type according to its tolerance to bandwidth degradation. Of course, type 6 with its huge bandwidth requirements (10Mbps) and high tolerance (down to 1Mbos) is the one who suffers the most.

Figure 4 shows how the protocol can be used to give a preferential treatment for handoff calls over new calls. In Figure 4, we simulated the network using the same setting of Figure 1. However, we considered that 30% of the calls in type 1 (audio calls) are handoff calls, the rest 70% are new calls. Both of these 2 types require the same bandwidth and the same tolerance for bandwidth degradation (0 degradation for type 1). The handoff calls are operating at a priority of 10 (highest priority in our setting), the rest of the calls are operating at priority of 7. When the network has a light load, both of new and handoff calls have the same very low call blocking rate. When the utilization increases, the call blocking rate for new calls approaches 1%, while the call blocking rate (in that case it is called call dropping rate since the call is moving from another cell) is kept below 0.1%. The rest of the traffic types performance is almost identical to Figure 1

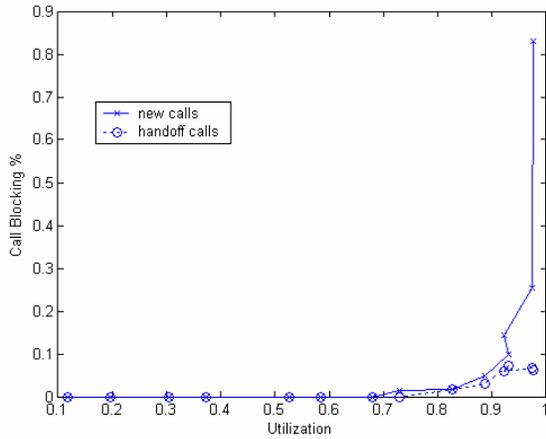


Figure 4. Call Blocking vs. Utilization for new and handoff calls

Figure 5 shows the effect of changing the priority on the call blocking rate. We choose type 3 traffic (it has the lowest call dropping priority in Figure 1). We varied type 3 priority from 1 to 10 and measured the call blocking rate. The Figure shows clearly that at priority of 1, the call blocking rate is 22%, with increasing the priority the call blocking rate decreases, at a priority of 6, the rate drops to 0. For this experiment the network utilization was kept constant at 85%.

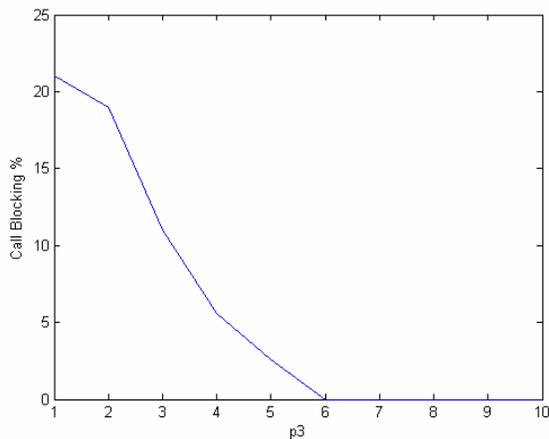


Figure 5. Call Blocking vs. priority for traffic type 3

CONCLUSION AND FUTURE WORK

In this paper we presented a call admission control protocol for cellular network. Our protocol depends on borrowing bandwidth from connections that can afford some performance degradation in order to admit new users to the network. Our protocol can assign priorities both in admitting traffic and in bandwidth degradation for different types of traffic. For future work, we would like to consider extending our protocol to guarantee QoS for both call blocking rate and the perceived quality of the call (number of BW changes).

REFERENCES

- [1] M. Aboelaze "A call admission protocol for cellular networks that supports differentiated fairness" Proceeding of the *IEEE 59th Vehicular Technology Conference VTC2004 Spring*, May 2004.
- [2] C. W. Ahn, and R. S. Ramakrishna "QoS provisioning dynamic connection-admission control for multimedia wireless networks using a Hopfield neural network" *IEEE Trans on Vehicular technology* vol 53, No. 1 Jan. 2004 pp 106-117.
- [3] C. H. Chia, and M. S. Beg "Realizing MPEG-4 video transmission over wireless bluetooth link via HCI" *IEEE Trans. on Consumer Electronic* vol 49, No. 4 Nov. 2003 pp 1028-1034
- [4] A. Crossman "A variable bit rate audio coder for videoconferencing" *Proceedings of the IEEE Workshop on Speech Coding for telecommunication* Oct. 13-15 1993 pp 29-30
- [5] S. Das, S. K. Sen, K. Basu, and H. Lin "A Framework for bandwidth begradation and ball bdmision control schemes for mobile multiclass traffic in next-generation wireless networks", *IEEE Journal on Selected Areas in Communications*, Vol. 21, No. 10, Dec. 2003 pp 1790-1802.
- [6] M. El-Kadi, S. Olariu, and H. Abdel-Wahab "A rate based borrowing scheme for qos provisioning in multimedia wireless networks", *IEEE Transactions on Para ller and Distributed Systems*, Vol. 13, No. 2 pp 156-166, Feb. 2002
- [7] M. M. Islam, M. Murshed, and L. S. Dooley "A direction based bandwidth reservation scheme for call admission control". *Proceedings of the International Conference on on Computer and Information technology*. (ICCI02), Dec. 2002. pp 345-349.
- [8] W. S. Jeon, D. G. Jeong. "Call admission control for mobile mulimedia communication with traffic asymmetry between uplink and downlink" *IEEE Trans, on Vehicular technology*, vol 50 No. 1 January 2001 pp 59-66
- [9] M. I. Kim, and S. J. Kim "A 2-level call admission control scheme using priority queue for decreasing new call blocking & handoff call dropping" *The 5th Semiannual Vehicular Technology Conference VTC2003-Spring* Volume 1 pp 459-461 April 22-25, 2003
- [10] Y. C. Lai, and Y. D. Lin "A fair admission control for large bandwidth multimedia applications" proceedings of *The 22nd International Conference on Distributed Computing Systems Workshops*, 2-5 July 2002. pp 317 -322
- [11] 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" *IEEE Trans, of Communication*, vol. 52, No. 5 May 2004. pp 778-790.
- [12] D. Mitra, M. I. Reiman, and J. Wang "Robust dynamic admission control for unified cell and call QoS in statistical multiplexers", *IEEE Journal on Select. Areas of Comm*. Vol 45, June 1998, pp 692-707.
- [13] S. Nananukul, "Efficiency and reliability of call admission control schemes using call-level behavior," *IEEE trans. on Communications*, Vol. 50, No. 10 Oct. 2002 pp. 1645-1651.
- [14] C. Oliveira, J. Kim, and T. Suda "An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks" *IEEE journal on Selected Areas in Communicatiobs*. Vol. 16, No. 6 August 1998, pp 858-874
- [15] Q. Ren, and G. Ramamurthy, "A real-time dynamic connection admission controller based on traffic modeling, measurement, and fuzzy logic control". *IEEE Journal on Select. Areas Comm*. Vol 18, Feb. 2000, pp 268-282.
- [16] C. Wu, Y. Tsai, and J. Chang "A quality-based birth-and-death queueing model for evaluating the performance of an integrated voice/data CDMA cellular system" *IEEE Trans. on Vehicular technology* vol. 48 Jan 1999, pp 83-89.
- [17] M. Wu, E. Wong, and J. J. Li "Performance evaluation of predictive handoff scheme with channel borrowing" Proceedings of *The 2003 IEEE International Performance, Computing, and Communications Conference* pp 531-536, 2003