PROVIDING DETERMINISTIC QUALITY OF SERVICE GUARANTEES IN MULTIMEDIA WIRELESS NETWORKS

Felix K. L. Lee and Mounir Hamdi [1]
Department of Computer Science
The Hong Kong University of Science & Technology
Clear Water Bay, Kowloon
Hong Kong

Abstract - One major challenge in designing future generation wireless networks is how to provide quality-of-service (QoS) guarantees to various multimedia applications. In this paper we consider deterministic quality-of-service guarantees. We propose a method for constructing a packet dropping mechanism that makes decisions as to when we need to drop packets while still not violating any QoS guarantees. This is achieved by employing: 1) An accurate traffic characterization of the VBR multimedia traffic streams; 2) A traffic regulator that can provide bounded packet loss; and 3) A traffic scheduler that can provide bounded packet delay. We performed a set of performance evaluation experiments. The results will demonstrate that our proposed QoS guarantee schemes can significantly support more connections than a system which do not allow any loss at the same required QoS.

1 Introduction

"Anytime, anywhere" communication, information access, and processing are much cherished in modern societies because of their ability to bring flexibility, freedom, and increased efficiency to individuals and organizations. Wireless communications, by providing ubiquitous and tetherless network connectivity to mobile users, are therefore bound to play a major role in the advancement of our society. Although initial proposals and implementations of wireless communications are generally focused on near-term voice and electronic messaging applications, it is recognized that future wireless communications will have to evolve towards supporting a wider range of applications, including voice, video, data, images, and connections to wired networks. This implies that future wireless networks must provide packet-based transport and bandwidth-upon-demand as well as support multimedia applications, thereby introducing a new set of challenges. One major challenge being how to provide quality-of-service (QoS) guarantees to various multimedia applications in a wireless environment.

Typical traffic in multimedia applications can be classified as either Constant-Bit-Rate (CBR) traffic or Variable-Bit-Rate (VBR) traffic. In particular, scheduling the transmission of VBR multimedia traffic streams in a wireless environment is very challenging and is still an open problem. In general, there are two ways to guarantee the QoS of VBR multimedia streams, either deterministically or statistically. The statistical models for characterizing bursty VBR traffic sources are either not powerful enough to capture the important burstiness and time correlation of realistic sources, or they are too complex for practical implementation for connection admission control (CAC) algorithms. The deterministic models would normally be more pessimistic. In particular, the throughput and the channel utilization of a deterministic bounded system are typically smaller than that of a statistical bounded system. However, their advantages are that they can provide deterministic bounds such that no admitted connection would violate its QoS guarantees.

Recently, CAC algorithms and medium access control (MAC) protocols have been proposed for wireless networks. Nearly all of the proposed algorithms only provide statistical, or soft, QoS guarantees. Even if a connection is admitted to the network, the network cannot fully guarantee that no QoS of any connection is violated. That is, the required or the desired QoS of some connections may still be violated in some cases.

In this paper we consider deterministic quality-of-service guarantees. We propose a method for constructing a packet dropping mechanism that makes decisions as to when we need to drop packets while still not violating any QoS guarantees. This is an extension of our earlier work [2], [6]. This mechanism is based on a mathematical framework that determines how many packets can be dropped while the required QoS can still be preserved. This is achieved by employing: 1) An accurate traffic characterization of the VBR multimedia traffic streams; 2) A traffic regulator that can provide bounded packet loss; and 3) A traffic scheduler that can provide bounded packet delay. The combination of traffic characterization, regulation, and scheduling can provide bounded loss ***and*** delay deterministically. This is a distinction from traditional determinitic QoS schemes where a 0% packet loss are always assumed with

---

[2] Email: hamdi@cs.ust.hk, Tel: (852) 2358 6984.

0-7803-5718-3/00/$10.00 ©2000 IEEE.

deterministically bounding the delay. The uniqueness of our scheme is that it can provide bounded loss ***and*** bounded delay deterministically.

We performed a set of performance evaluation experiments. The results will demonstrate that our proposed QoS guarantee schemes can significantly support more connections than a system which do not allow any loss at the same required QoS. As a result, the admitted connections are served with better hard QoS using our proposed scheme.

This paper is organized as follows. Section 2 reviews wireless MAC protocols and various traffic characterization methods. Section 3 and 4 introduce our proposed wireless access and traffic characterization and their performance evaluation for providing QoS guarantees. Finally, Section 5 concludes the paper.

2 Wireless MAC Protocols and Traffic Characterization

Although there are also a lot of related research on multimedia wireless MAC protocols and their call admission control algorithms, nearly all of the proposed algorithms only provide statistical QoS guarantees. As a result, these MAC protocols may not be able to satisfy connections which demand deterministic QoS guarantees.

In addition, there are also a lot of related research on both traffic characterization and packet scheduling, but much of this work cannot be directly applied to support traffic which requires a bounded QoS. For example [1], some approaches characterize the traffic sources using sophisticated stochastic models such as Markov-modulated, autoregressive, self-similar, TES, and S-BIND. These approaches are not capable for providing worst-case bounds on traffic arrivals and, therefore, cannot be used as the traffic characterization for deterministic bounded connections. Furthermore, the design of a real-time traffic policing mechanism for a stochastic traffic model is very difficult.

In this section we review some approaches which are suitable for use in networks with bounded QoS. We review a number of wireless MAC protocols. We then review a number of traffic characterization models and we discuss the advantages and the disadvantages of the different approaches.

Below, we summarize the characteristics of state-of-the-art protocols as far as serving multimedia applications is concerned. For more details, the reader is referred to [4], [5]. Some of the entries are based on subjective judgments rather than formal analysis.

Table 1: Summary of wireless MAC protocols.

	PRMA	**PRMA/DA**	**DSA++**	**MASCARA**	**BRMA**
Physical Layer	FDD	FDD	FDD	TDD	TDD
QoS Support	voice	voice, video, data	CBR, VBR, ABR	CBR, rt-VBR, nrt-VBR, ABR, UBR	CBR, VBR, ABR
Control overhead	Low	Medium	High	High	High
Access Technique	S-ALOHA	S-ALOHA	Split algorithm	Not defined yet	No
CAC	No	Yes	No	Yes	No

From the comparative results, it appears that the MAC protocols which use FDD can deal with the access contention procedures more quickly because they can send acknowledgment signals almost immediately after receiving the packets. However, TDD can be advantageous in some situations where frequencies are scare. In addition, when the traffic loading for the uplink and the downlink is unbalanced, TDD can use the bandwidth more efficiently by allocating more slots to one side than the other side.

Traffic characterization and policing in multimedia wireless networks is as important as the wireless access scheme.

Below, we summarize the relationship between the traffic models and their policing mechanisms in Table 2.

Table 2: Summary of traffic models

Traffic Model	**Policing Mechanisms**	**Traffic Constraint Function**
Peak-rate	Packet spacer	$A^*(t) = \left(\left\lfloor \frac{t}{X_{min}} \right\rfloor + 1\right) s^{max}$
(σ, ρ)	Jumping window	$A^*(t) = \sigma + \rho t$
$(\vec{\sigma}, \vec{\rho})$	Leaky bucket	$A_m{}^*(t) = \min\{\sigma_i + \rho_i t\}$
(r, T)	Multiple leacky buckets	$A^*(t) = \left(\left\lceil \frac{t}{T} \right\rceil + 1\right) rT$

Table 2: Summary of traffic models

Traffic Model	Policing Mechanisms	Traffic Constraint Function
D-BIND	Multiple moving widows	$A^*(t) = \frac{R_i I_i - R_{i-1} I_{i-1}}{I_i - I_{i-1}} + R_i I$

A study by Reibman and Berger [1] and another by Rathgeb [1] evaluate the accuracy with which the various traffic models can characterize VBR video. Rathgeb shows how the parameters of each traffic policing mechanism can be expressed in terms of parameters of the other mechanisms, enabling a direct comparison of the various mechanisms. The examples presented indicate that the leaky bucket is superior to both the windowing mechanisms for describing VBR video since neither the jumping windowing nor the moving windowing are capable of capturing the short-term burstiness.

Wrege *et al.* [3] showed that the $(\vec{\sigma}, \vec{\rho})$ traffic models which employ multiple leaky bucket mechanisms can accurately characterize VBR video.

Both Dittmann and Rathgeb showed that the moving window mechanism is significantly more difficult to implement than either the jumping window or leaky bucket mechanisms.

For these reasons, the networking community has focused primarily on characterizations that can be policed by the leaky bucket mechanisms, that is, the (σ, ρ) and $(\vec{\sigma}, \vec{\rho})$ traffic models.

3 Proposed MAC Protocol and Traffic Characterization

Before presenting our wireless acces protocol and our QoS guarantee mechanism, we need to accurately chatacterize multimedia traffic. Our proposed traffic model is based on the (σ, ρ)-model but with a small modification. Two more parameters added: maximum traffic rate ρ_{max} and minimum traffic rate ρ_{min}.

The (σ, ρ)-model assumes that all burstiness of traffic arrives at the same time, but in general, the burstiness do not come at the same time and it would normally come within a certain duration. Hence, our objective of adding the maximum traffic rate ρ_{max} is to characterize this fact. Our traffic model becomes similar to the special case of the (σ, ρ)-model with 2 pairs, (σ, ρ) and $(0, \rho_{max})$.

Most multimedia traffic is continuous,e.g., video and voice connections. Thus the connections have minimum rate. To characterize this fact, we add a minimum traffic rate ρ_{min}. The proposed traffic model has four parameter: burstiness σ, maximum traffic rate ρ_{max}, minimum traffic rate ρ_{min}, and average traffic rate ρ. The traffic constraint function is:

$$A^*(t) = \min(\sigma + \rho t, \rho_{max} t) \quad \text{for all } t \geq 0$$

In order to bound the packet loss, we employ a packet dropping mechanism. The uniqueness of our QoS mechanism is that we have the capability to bound the loss rate of packets besides bounding their delay. For this reason, we emply a packet dropping mechanism to select the packets to be dropped for each connection [2, 3].

Theorem: The optimal dropping of $L\%$ packets for the worst case of the $(\sigma, \rho, \rho_{max}, \rho_{min})$ traffic model is the same as the worst case of the $(\sigma', \rho', \rho'_{max}, \rho'_{min})$ traffic model where $\sigma' = \max(0, \sigma(1 - Lf))$, $\rho' = \rho(1 - L)$, $\rho'_{max} = \min(\rho(1 - L), (\rho_{max} - \rho)(1 - Lf) + \rho(1 - L))$, $\rho'_{min} = \rho_{min}$, and $f = \frac{1}{1 - \frac{\rho_{min}}{\rho}}$.

According to the Theorem, the traffic of a $(\sigma, \rho, \rho_{max}, \rho_{min})$-model with a loss rate $L\%$ can be reduced to $(\sigma', \rho', \rho'_{max}, \rho'_{min})$-model in the worst-case scenario of traffic arrival. It is obvious that the traffic of $(\sigma, \rho, \rho_{max}, \rho_{min})$-model can be reduced to $(\sigma', \rho', \rho'_{max}, \rho'_{min})$-model when traffic arrival is not in the worst case. We can use two leaky buckets to filter the traffic of the $(\sigma, \rho, \rho_{max}, \rho_{min})$-model and reduce it to $(\sigma', \rho', \rho'_{max}, \rho'_{min})$-model.

Having determined how to characterize traffic and the packet dropping acheme, we need to identify a wireless access scheme. Our proposed wireless MAC protocol is motivated by the Bandwidth Reservation Multiple Access (BRMA) [7]. Since the assignment of the mini-slots in BRMA is deterministic, both the request channels and the data channels is contention-free. This characteristic is suitable for us to provide deterministic QoS guarantee. Transmission time is divided into time slots which are further divided into mini-slots and data. Mini-slots are for sending the requests from terminals to the base station, while the data slots are for sending real data or packets to the base station.

In BRMA, each connection will be assigned one mini-slot in each frame. As a result, each connection do not need to content for the channel with other connections. But in our proposed protocol, we have to deal with VBR, CBR, ABR and UBR traffic where CBR and VBR connections need to have deterministic QoS guarantees. In our protocol, only CBR and VBR connections will be assigned one mini-slot in each frame, while ABR and UBR will content for the channel through a random access, e.g., Slotted ALOHA.

4 Performance Evaluation

To evaluate the performance of our proposed MAC protocol and QoS guarantee mechanisms, we performed a series of evaluation tests. These results complement our previous results [2]. We evaluate the performance of the protocol itself by varying different factors, e.g., traffic burstiness, average traffic rate, maximum traffic rate, minimum traffic rate, etc. The purpose of these evaluations is to see how these factors affect the performance of the proposed protocol.

We evaluate the performance of our proposed traffic characterization and wireless access by varying different factors such as the burstiness σ, the peak rate ρ_{max}, the minimum rate ρ_{min}, and the desired loss rate L. In the second test, we evaluate the performance using multiple classes of traffic. We plot an admission region for two classes of traffic.

Unless otherwise specified, all of the results in this section are generated with the parameters shown in Table 3.

Table 3: Traffic parameters for single type of traffic.

Channel Capacity	**45Mbps**
Cell Size	53Bytes
ρ	0.15Mbps
ρ_{max}	0.9Mbps
ρ_{min}	0.09Mbps
σ	100 cells
d	30ms
L	5%

First, we investigate the effect of the ratio of the minimum traffic rate to the average traffic rate on the channel utilization. As shown in Figure 8, we found that the channel utilization increases from 0.4 to 1 when the ratio of the minimum traffic rate to the average traffic rate increases from 0 to 1. A higher minimum traffic rate implies less variation of the traffic; thus the scheduling of traffic is easier and can be more efficient.

Next, we investigated the effect of the ratio of the maximum traffic rate to the average traffic rate on the channel ultilization. As shown in figure 9, we found that the channel utilization decreases from 1 to 0.2 when this ratio increases from 0 to 20. The channel utilization drops sharpely when the ratio is below 10, and it drops slowly when the ratio is above 10. In particular, a lower minimum traffic rate implies less variation of the traffic; thus the scheduling of these traffic are easier and can be more efficient.

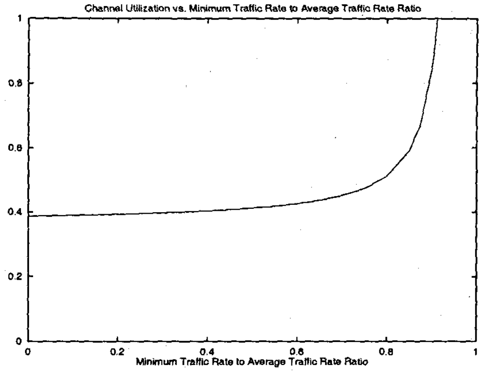

Figure 8: Utilization vs. ratio of minimum traffic rate to average traffic rate (ρ_{min} / ρ).

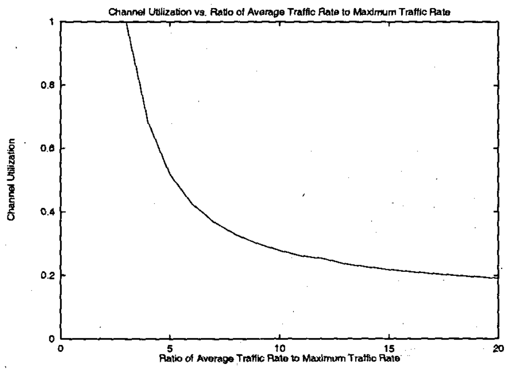

Figure 9: Utilization vs. ratio of maximum traffic rate to average traffic rate (ρ_{maxn} / ρ).

We also studied the effect of the traffic burstiness on the channel utilization as shown in figure 10. We found that the channel utilization decreases from 1 to 0.4 when the burstiness of the traffic increases from 0 to 120 cells. The result illustrates the fact that lower burstiness of the traffic implies less variation of the traffic; thus the scheduling of the traffic is easier and can be more efficient.

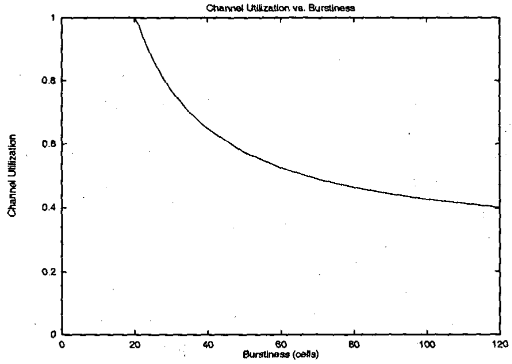

Figure 10: Utilization vs. burstiness of the traffic.

Figure 11 illustrates the effect of the packet loss rate on the channel utilization. The channel utilization increases from

0.368 to 0.425 when the desired bounded loss rate increases from 0% to 5%. By allowing 5% of packet loss for the connections, the channel utilization increases by 15%, thus 15% more connections can be admitted. According to the figure the channel utilization is directly proportional to the desired bounded loss rate.

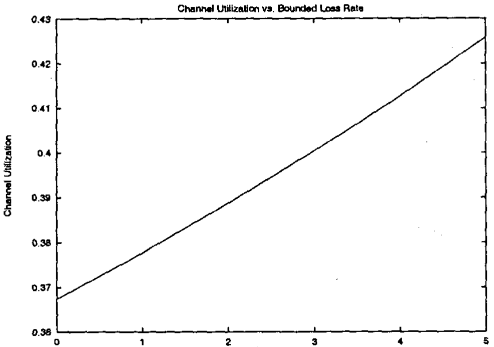

Figure 11: Utilization vs. desired loss rate.

Now, we assume that we have two classes of traffic and the channel capacity is 45Mbps. The other traffic parameters are shown in Table 2.

Table 4: Traffic parameters for multiple types of traffic.

Connection Class	1	2
ρ_i	0.15 Mbps	0.15 Mbps
σ_i	25 cells	250 cells
d_i	30ms	50ms
L_i	5%	5%

According to the Figure 12, the maximum number of class 1 connections that can be supported increases as the number of class 2 connections decreases, and vice versa. When the number of class 2 connections reduces to below a certain value, the maximum number of class 1 connections that can be supported stops to increase. The reason is that the resources saved by not serving class 2 connections are not enough to serve any more class 1 connections as the desired QoS, for instance delay, of class 1 connections is much higher than that of class 2 connections. The admission region increases as the ratio of minimum traffic rate to average traffic rate increases. This is due to the fact that the smaller ratio implies less variation of the traffic. When the ratio is 0.8, the area of the admission region of the proposed algorithm is around 40% larger than that of the algorithm which do not allow dropping.

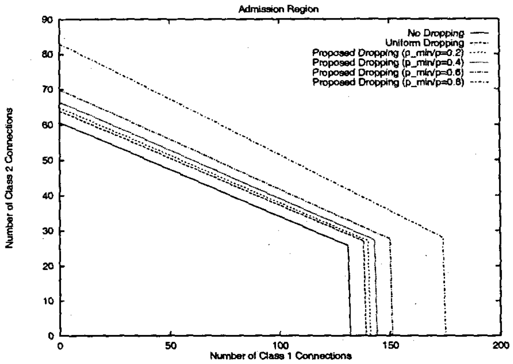

Figure 12: Admission region for two Different classes of traffic.

5 Conclusion

we proposed a QoS guarantees method for wireless networks which is based on a mathematical framework that accuratly characterizes multimedia traffic streams in conjunction with efficient scheduling and dropping algorithms. The uniqueness of this scheme is that it can provide bounded loss ***and*** bounded delay deterministically. By allowing certain bounded loss for the connections, more connections can be admitted into the network. We have performed a set of performance evaluation tests that demonstrated that our proposed algorithm can significantly support more connections than a system do not allow any loss.

References

[1] E. W. Knightly, ``Traffic models and admission control for integrated services networks," in *PhD thesis, University of California - Berkeley*, May 1996.

[2]. *F. K. L. Lee* and M. Hamdi, "Providing Deterministic Quality of Service Guarantee in Wireless Environment," *Proceedings of the 49th International Conference on Vehicular Technology (VTC '99)*, May 1999.

[3] J. Liebeherr, D. E. Wrege, and D. Ferrari, "Exact Admission Control in Networks with Bounded Delay Services," *IEEE/ ACM Trans. on Networking*, 1996.

[4] N. Passas, S. Paskalis, D. Vali, and L. Merakos, ``Quality-of-service-oriented medium access control for wireless ATM networks," in *IEEE Communications Magazine*, pp. 42-50, 1997.

[5] J. Sanchez, R. Martinez, and M. W. Marcellin, "A survey of MAC protocols proposed for wireless ATM," *IEEE Network*, pages 52--62, 1997.

[6] L. Wang, M. Hamdi, and K. B. Letaief, "Wireless channel access for multimedia personal communication," *Proceedings VTC'98*, 1998.

[7] Z. Zhang and T. Saadawi, ``A bandwidth reservation multiple access protocol for wireless ATM local networks," *Journal of Wireless Networks*, pp. 147-60, 1997.