

Adaptive Jitter Compensation on the Downlink of Future Mobile Multimedia Communications Systems

Glenn Platt and Jamil Y. Khan

School of Electrical Engineering & Computer Science

University of Newcastle

Callaghan, NSW, Australia

Email: glenn@ecemail.newcastle.edu.au, jkhan@ecemail.newcastle.edu.au

Abstract— As all-IP based mobile communications networks become a reality, it is important to understand how characteristics of these networks such as delay jitter can affect the overall performance of the system. We find that backbone network delay jitter can significantly influence the multiple access performance of an all-IP mobile communications system. To compensate for this, we introduce an adaptive jitter compensation technique that is linked to the system slot allocation and multiple access mechanisms. By linking these mechanisms, we are able to compensate for delay jitter, whilst also introducing novel resource allocation techniques for use with multimedia traffic. We detail the various techniques introduced, and show the performance enhancements possible when using such a scheme.

Keywords—MAC; delay jitter; wireless IP

I. INTRODUCTION

The amount of packet based traffic in mobile communications systems is continuing to grow, as new services are being introduced to current systems, and long-awaited third generation systems are reaching commercial realization. It is widely recognized that with the introduction of recent technologies such as the IP Multimedia Subsystem [1], Internet Packet (IP) based traffic will become more widespread and predominant in modern mobile communications systems. One of the main design philosophies of these systems is the seamless delivery of services, meaning that not only will future systems carry IP based user traffic, but that the radio access and core networks of such systems will be based on IP technologies.

Although IP technologies allow the introduction of very diverse and widespread network services, they are not without their disadvantages. Currently IP networks are effectively best-effort networks, which do not provide any means of stringently controlling end-to-end metrics such as delay, delay variation (jitter), or packet loss. We are interested in how these characteristics of IP networks can influence the performance of the resource allocation and multiple access strategies of a mobile communications system, an area that has received little attention thus far.

Consider for example the downlink of a mobile communications network, where packets being sent on the downlink arrive from a fixed IP-based network. Most mobile communications systems have a time-based multiple access

component, where system resources are controlled on a time basis, and a particular resource (be it, say, a time slot or CDMA code) is allocated to a user to use for a particular amount of time. Often users share access to the resource, where each user is granted access at regular, periodic intervals. The supervising multiple access control (MAC) entity of such systems pre-allocates transmission resources to users, based on the nature of the traffic to be transmitted. Essentially, a certain resource is scheduled to be available at a particular time for transporting a user's packet. Considering the downlink of such systems, if the scheduled packet is late or lost due to congestion or link failures in the fixed IP network, then the packet cannot be transmitted using the resource allocated to it, and this resource is wasted, with a subsequent effect on system capacity and/or multiplexing efficiency. Various studies by industry and the Mobile Internet Forum have highlighted such an issue [2], but the problem has received little further attention.

We previously considered such an issue in [3], where we introduced the concept of an adaptive buffering mechanism to try to minimize the loss in performance caused by delay jitter on the backbone network. In this paper we continue the investigation into the performance gains possible by such a jitter compensation technique. We show that by combining the proposed jitter compensation technique with the system MAC and slot allocation mechanisms, significant additional performance gains can be realized, especially regarding system utilization and capacity. We also introduce consideration of various classes of traffic in the slot allocation and jitter compensation schemes, and detail a novel way of providing service to traffic with low end-to-end delay requirements, without taking resources from higher-priority traffic.

II. PRIOR WORK IN DELAY AND JITTER COMPENSATION

In this document, when we refer to jitter control in mobile communications systems, we are referring to trying to maximize the multiple access performance, and thus the system capacity and utilization of system resources, by ensuring that the packet scheduled to be carried on a resource arrives in time for its transmission. That is, we wish to minimize the delay variation that makes a packet's arrival time unpredictable. If the packet is late, or experiences

excessive jitter, then it will arrive too late for transmission, and the resource scheduled for this packet is wasted.

The particular components of the mobile communications system of interest here are the interface node between the fixed IP network and the mobile network, the packet scheduler, access mechanism, base station and associated entities. For the remainder of the paper we refer to all of these entities collectively as the base station (BS). We assume that packets arrive at the BS from a fixed IP based network, and that the BS takes care of scheduling system resources for packet transmission. For the remainder of the paper we consider such resources to be timeslots, where one packet is transmitted per timeslot.

We have previously concluded that “destination-end” style jitter compensation techniques are the only applicable way of controlling delay jitter in the situation of interest [3]. Destination-end style jitter compensation techniques have received considerable interest for application in streaming and real-time transmissions on fixed networks. A number of such destination-end jitter compensation techniques are reviewed in [4] and [5]. Generally speaking, all the destination-end jitter compensation techniques reviewed use a buffering mechanism to attempt to restore the inter-packet temporal relationships created at the transmitter. The buffer is generally controlled by a scheduling mechanism that schedules when packets are removed from the buffer and delivered to their destination. These techniques all possess one fundamental trade-off- whilst the buffering mechanism improves the inter-packet temporal relationships, it also introduces additional end-to-end delay.

In [3] we presented such a destination-end style jitter compensation technique, designed to improve the multiple access performance of a mobile communications system experiencing jitter on the backbone network. In the following section, we review the controlling jitter compensation algorithm, and detail the significant performance enhancements that have been made to it.

III. JITTER COMPENSATION MECHANISM

In this work, we find that significant gains are possible if the jitter compensation mechanism is linked with other resource management mechanisms such as the slot allocation and MAC components of the system. Thus, in this section, we first review the jitter compensation algorithm, detailing its method of operation. We then cover the slot allocation and MAC strategies used in our system, and detail how these have been linked with the jitter compensation mechanism.

A. Jitter Compensation Algorithm

Considering the destination-end jitter compensation algorithms reviewed in [4] and [5], we have designed a destination-end based jitter compensation technique that adapts the size of the jitter buffer, and thus the additional delay, based on recent jitter measurements. As opposed to jitter compensation algorithms for streaming systems, which focus on the inter-packet temporal relationships, the main aim of our algorithm is to maximize the number of downlink slots that are occupied,

whilst continuing to respect the QoS requirements of the traffic to be carried.

Essentially, the jitter compensator monitors a partial range m of all previous packet delays, and sets the size of the jitter buffer based on the jitter level observed in this range of m received packets. The jitter buffer size is set based on the mean jitter value across the m measurements in the partial range. Setting the jitter buffer size based on the mean jitter value results in a satisfactory tradeoff between minimizing the number of too-late for transmission packets, whilst also keeping end-to-end delay at a minimum.

The size of the partial range can be tuned based on how conservative we wish to be. A small partial range means the jitter compensation algorithm will track rapidly changing jitter levels, yet may also attempt to cope with spurious, excessive jitter levels that are effectively noise, resulting in excessive buffer sizes and buffering delays. Too large a partial range means the jitter compensation algorithm will be too slow to track rapid jitter variations, resulting in poor choice of buffer size, and subsequent performance degradation.

Each data traffic stream operating on a different timeslot in the system has its own jitter buffer, and jitter buffer control mechanism. We adjust the size of the jitter buffer in between mobile station (MS) resource reservations. As a MS may not be assigned the same timeslot for consecutive reservations, the BS maintains a list of currently active MS's, and the appropriate jitter buffer size for each MS. Thus, when an active MS is assigned a timeslot, we are able to identify the appropriate jitter buffer size for this MS. If, after monitoring m jitter delays, the algorithm decides that the jitter buffer size needs to be changed, we update the buffer size entry for this MS, and at the next reservation, assign this new buffer size.

Note that the buffering technique does not operate on conversational class [6] (real-time) traffic, as previous results have shown that even very small buffer sizes result in excessive delays for real-time voice traffic [7]. Thus, the traffic types of interest to the buffering mechanism are streaming (audio/video), interactive (for example, web browsing), and background (for example sending/receiving email) class traffic. Although conversational class traffic is not buffered (and thus conversational class packets that arrive too late are simply dropped), in this latest work system utilization is still maximized, as with the integration of the jitter compensation and slot allocation mechanisms, voice traffic slots that are empty due to a late voice packet are used for transmission of other packets. Further details are available in the following section.

B. Slot Allocation and MAC Mechanism.

The MAC component of the model is based on reservation style MAC techniques such as PRMA++ [8]. Mobile stations contend for access to traffic timeslots using designated reservation slots. As per PRMA++, MS's generating conversational class traffic such as voice traffic hold a timeslot for the duration of their speech period or “talkspurt”.

We have introduced some advanced priority mechanisms designed to improve the downlink multiplexing efficiency of the system, by improving the access of non-conversational class mobile stations (referred to as “data terminals” for the remainder of this section), whilst still preserving the QoS given to conversational class traffic. The downlink slot allocation strategy operates adaptively, modifying its behavior based on system load. At low levels of system load, all classes of traffic can use system resources for as long as they require. As system load increases, to enable other data terminals access to the resources, data terminals are only allowed to reserve a “block” of resources, after which they relinquish the resources to another waiting terminal. We use a priority queuing mechanism to ensure conversational class mobile stations maintain precedence over data terminals. If system load increases to the point where conversational class stations are waiting for access, to maintain the QoS available to such stations at high load we introduce a “forced relinquish” mechanism, which ensures conversational class traffic can access the medium whenever it wishes.

In a time-constrained mobile communications system, large jitter compensation buffer sizes are impractical, due to the significant end-to-end delays they introduce. Thus, when using a reasonable buffer size, a number of packets will still arrive too late for transmission, even after the jitter buffering. In our system such packets are placed in a separate queue, named the “backlog queue”, where they are later transmitted. Given the jitter compensation mechanism previously introduced [3], a number of significant enhancements can be made to the system performance:

- The end-to-end delay measurements for all packets in the system can be improved if we deal with the “backlog” queue in an intelligent fashion. It is not appropriate for packets to sit in the backlog queue for a large amount of time, as the end-to-end delay experienced by these packets will be excessive. Nor is it appropriate for these backlogged packets to be transmitted using resources allocated for other traffic operating in the system, as multiplexing efficiency would suffer.
- If a data packet arrives too late for transmission, even after jitter buffering, then this packet’s slot will go wasted. Additionally, conversational class packets are not buffered at all, and thus a significant number of conversational class packet slots can be wasted due to backbone packet jitter. If a technique can be introduced that makes use of these otherwise wasted slots, system multiplexing efficiency can be further improved.
- One class of traffic in multimedia mobile communications systems, the background-class traffic [6], does not have strict end-to-end delay requirements, and can tolerate moderate end-to-end delay. It would be advantageous to use the otherwise wasted resources mentioned in the points above to transmit such traffic. Such a technique has the advantage of allowing background class traffic to be supported without decreasing the system capacity

available to delay-constrained traffic, whilst also utilizing what would be otherwise wasted resources.

Considering these points, we have linked the jitter compensation mechanism with the system slot allocation mechanism. The system operates a backlog queue for each downlink transmission slot. If a packet intended for transmission on a certain slot arrives too late for transmission, this packet is placed on the tail of the appropriate backlog queue for this slot (Note this only applies to non-conversational class traffic). When a transmission slot goes empty because the packet intended for this transmission slot has not arrived, the packet at the head of the backlog queue for this slot is transmitted in the now empty slot. When a particular mobile station relinquishes its transmission slot, packets remaining in the backlog queue for this slot are destroyed- they are no longer eligible for transmission, as a new mobile station will soon acquire this slot.

Meanwhile, considering the background class traffic, once a background class traffic source is granted admission to the system, this source is given a *virtual* traffic slot assignment, and the downlink destination MS is instructed to listen to this virtual slot number. The slot assignment is referred to as “virtual” as the background class traffic must share this slot assignment with conversational class traffic. Normally the assigned slot is occupied by conversational class traffic, for example voice packets. At such times, the background traffic class destination MS ignores the arriving voice packets, as they are intended for a conversational class MS. However, when the downlink slot goes empty because the conversational class packet did not arrive in time, then at this instant, the background class packet is removed from its queue, and transmitted on the empty slot. As the background class MS has been listening to this shared slot, it will detect the background class packet, and receive it correctly.

During low system load, when a number of downlink traffic slots remain empty, we maximize the transmission of background class traffic. As the destination background class mobile station continues to listen to its virtual slot allocation until all expected packets have been received, the destination mobile station will be listening to this traffic slot, and thus receive the packets correctly.

The performance of our three new mechanisms- the adaptive jitter compensation mechanism, the backlogged packet queue transmission scheme, and the background traffic class virtual channel assignment scheme has been studied using a simulation model, detailed in the following section.

IV. SYSTEM MODELING

We have integrated the jitter compensation, backlogged traffic and virtual channel assignment mechanisms into a complete MAC and resource allocation event-based model. We split the model description into two sub-sections- firstly we describe the general components of our simulation model, and then describe how we have approached modeling jitter on the downlink of the communications system.

A. General Model Description

The model contains three types of traffic sources and partner destination mobile stations- voice sources generating conversational class voice traffic, data sources generating interactive class data traffic, and background sources generating background class traffic. Voice traffic is generated using Brady's talkspurt model [9], interactive traffic is generated using an on-off generator designed to model the reception of HTML files, and background class traffic is generated by a fixed file generator producing small data files at various intervals (imitating the production of so-called "Multimedia Messaging Service" (MMS) files) [10]. The BS entity of the model controls both the slot allocation and jitter compensation mechanisms. The BS is connected to an IP-style backbone network, over which downlink user and control traffic arrives. Whilst traveling this backbone network, the traffic experiences varying end-to-end delays, corresponding to the packet delay jitter found in packet-based networks.

B. Modeling Delay Jitter on the Backbone Network

Realistically modeling the delay jitter in a mobile communications network has proven to be a difficult exercise. Whilst the measurement of jitter levels in packet networks is straightforward, and a number of authors have presented measured jitter levels in packet-based networks [11] [12] [13], using these results to realistically model the jitter expected in a mobile communications system backbone network is complicated. Firstly, the sheer size and heterogeneity of networks such as the Internet mean that large variations exist between the various studies looking at packet jitter on the Internet [13], and finding commonalities between these studies is difficult. Further, the extrapolation of results taken from the Internet, and applying these to the IP backbone network of a mobile communications system is not straightforward. An IP backbone network would generally be a closed network, composed of links of significantly differing bandwidths- whilst the branch links to the individual transceiver stations are often based on low throughput microwave transmission, the major trunk links are often high-speed optical links with low load. To the best of our knowledge, no measurements or predictions of the jitter levels experienced in such networks exist, and our thus our research into this issue is continuing.

To demonstrate the performance benefits of our proposed techniques, we examine their performance over a large range of network packet jitter, covering the entire variety of jitter levels measured from the Internet. Generally, individual packets from the same source experience approximately the same end-to-end delay as their neighbors, with the delay variation between packets (the jitter) having a very low mean. However, as measured on the Internet [13], groups of packets from the same source may experience significantly different levels of end-to-end delay between them, corresponding to the changing states of a packet network. We also add noise to the jitter simulation- "spikes" of delay, causing large jumps in the level of jitter measured through the simulation.

V. RESULTS

Simulation results are shown in the following figures, and basic simulation parameters are presented in table 1. Figure 1 compares the slot utilization benefits provided by the three proposed schemes. These results are taken at fairly significant levels of jitter, and as shown, slot utilization without some form of jitter compensation can suffer considerably. The slot utilization for the scheme without jitter compensation actually decreases as load increases- at low levels of load a significant

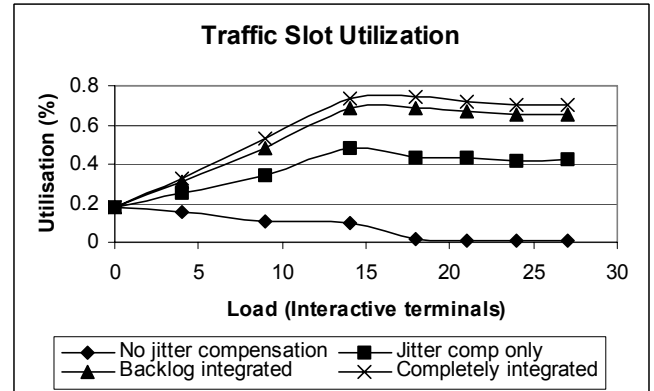


Figure 1. Traffic slot utilization vs. load.

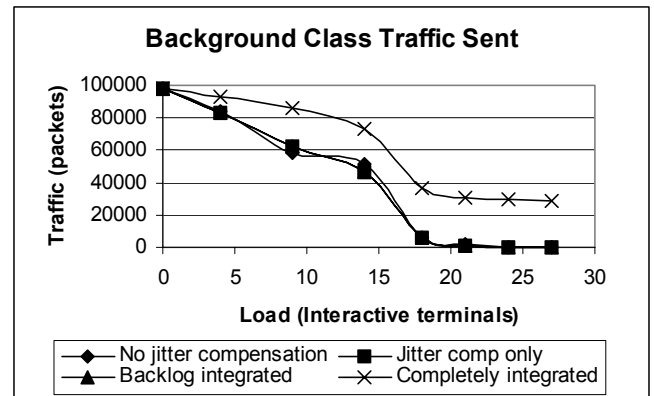


Figure 2. Background class traffic sent vs. load.

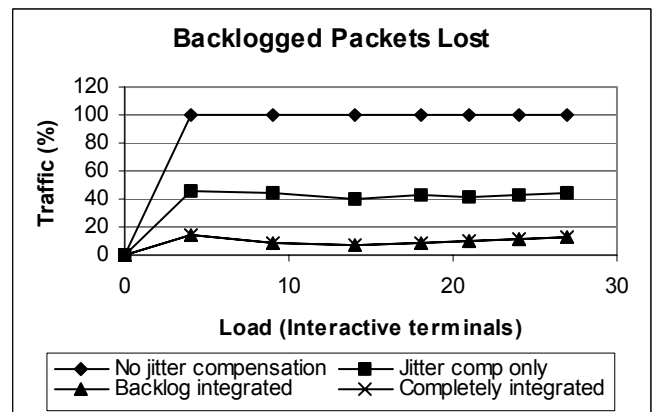


Figure 3. Backlogged packets lost vs. load.

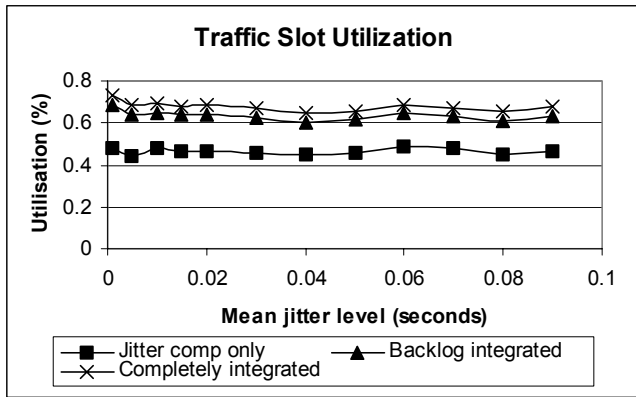


Figure 4. Traffic slot utilisation vs. jitter severity.

TABLE 1.

System Parameters	
Reservation slots: 4	Warmup time: 32s
Uplink slots: 21	Data block maximum reservation size: 200 frames
Downlink slots: 21	End-to-end delay mean max: 1s
Frame size: 10ms	Size of delay "group": 1000
Voice stations: 10	Max delay noise: 20s
Background stations: 21	Delay noise probability: 0.001
Data terminals: as per graphs	Jitter compensator m : 80 packets
Transmission speed: 2 Mbps	Jitter mean (figures 1-3): 0.001s
Simulation length: 5 min	

amount of background class traffic can be sent, yet as the load increases, less background traffic can be sent, the interactive and conversational class traffic that is sent suffers considerable jitter, and only small amounts of it arrive on time for transmission. Introducing our jitter compensation scheme significantly improves the slot utilization, and the subsequent inclusion of the backlog queue integration scheme, and following this the integration of the background traffic virtual channel assignment scheme, provide further utilization gains. As shown in figure 2, at high levels of load, if background traffic is only able to access unreserved slots, then very little background traffic is actually transmitted. However, with the integration of the background traffic scheme, background traffic is also transmitted in virtual allocation slots partnered with a conversational class terminal, and can still receive a reasonable amount of service.

Figure 3 shows the number of backlogged packets that are lost. With no jitter compensation scheme, any packet that does not arrive in time for its timeslot cannot be transmitted, and thus losses are high. Introduction of the adaptive jitter compensation scheme reduces the losses, but packets that are still too late for transmission after jitter buffering are destroyed, and so losses remain significant. Integrating the jitter compensation scheme with the slot allocation mechanism significantly improves this situation, as still too-late packets are placed on the backlog queue, and transmitted whenever an empty slot occurs. Note that packets are still lost with backlog queue integration- if a particular MS has packets sitting in the backlog queue when it relinquishes its resources, these are lost. We are currently looking at ways of further minimizing this loss. Figure 4 is included to show the performance of our proposed scheme under a large range of jitter severity. As

shown, the trends displayed in the previous graphs remain at all levels of jitter.

VI. CONCLUSIONS

IP backbone networks can significantly influence the performance of the multiple access and resource allocation protocols of mobile communications systems. Characteristics such as delay jitter can degrade resource utilization, and system capacity, and reduce the QoS available to system traffic.

We have presented an adaptive jitter compensation scheme, integrated with the system MAC and resource allocation mechanisms, that compensates for the negative effects of delay jitter on the fixed network, resulting in significantly improved system utilization. Additionally, we introduced a virtual channel assignment strategy that further improves system utilization, whilst allowing low priority traffic to be supported without degrading the QoS available to high priority mobile stations.

REFERENCES

- [1] 3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects, "IP based multimedia services framework", TR 22.941, v0.7.7, November 2001.
- [2] Mobile Wireless Internet Forum, "IP in the RAN as a transport option in 3rd generation mobile systems", technical report MTR-006, Release v2.0.2, June 2001.
- [3] G. Platt, J. Y. Khan, "Jitter compensation on the downlink of future mobile multimedia communications systems", *Proc. IEEE WCNC2003*, March 2003, in press.
- [4] N. Laoutaris, I. Stavrakakis, "Intrastream synchronisation for continuous media streams: A survey of playout schedulers", *IEEE Network*, May/June 2002, pp. 30-40.
- [5] R. Ramjee, J. Kurose, D. Towsley, H. Schulzerine, "Adaptive playout mechanisms for packetised audio applications in wide-area networks", *Proc. IEEE INFOCOM '94*, vol. 2, June 1994, pp. 680-688.
- [6] 3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects, "QoS concept and architecture", TS 23.107, v3.6.0, April 2001.
- [7] G. Platt, "A 'downlink aware' multiple access protocol for future multimedia communications: A simulation study", *Proc. OPNETWORK2002*, August 2002.
- [8] A. Arregui, J. Dunlop, "Stability analysis of the contention mechanism of the PRMA++ protocol", *Proceedings IEEE VTC '98*, vol. 3, May 1998, pp. 2124-2128.
- [9] P.T. Brady, "A statistical analysis of on-off patterns in 16 conversations", *Bell Syst. Tech. Journal*, vol 47, January 1968, pp. 73-91.
- [10] J. Sevanto, "Multimedia Messaging Service for GPRS and UMTS", *Proc. IEEE WCNC'99*, vol.3, September 1999, pp. 1422-1426.
- [11] V. Paxson, "End-to-end Internet packet dynamics", *Proc. ACM SIGCOMM 97*, September 1997.
- [12] D.C. Verma, H. Zhang, D. Ferrari, "Delay jitter control for real-time communication in a packet switching network", *Proc. TRICOMM 91*, April 1991.
- [13] D. Loguinov, H. Radha, "Large-scale experimental study of Internet performance using video traffic", *ACM SIGCOM Computer Communications Review*, vol. 32, issue 1, January 2002, pp. 7-19.