Technical University of Crete

Dept. of Electronics and Computer Engineering

Telecommunications Division

Information & Computer Networks Laboratory

Medium Access Control and Traffic Policing for Multimedia Integrated Access in High Capacity Wireless Channels

Πολυχρόνης Κουτσάκης

PhD Dissertation

Fall 2002

Στους γονείς μου

Ευχαριστίες

Ένα μεγάλο ευχαριστώ, κατ' αρχήν, στον επιβλέποντα καθηγητή μου Μιχάλη Πατεράκη, γι' αυτά τα πέντε ιδιαίτερα δημιουργικά χρόνια της συνεργασίας μας. Η συμμετοχή του σε κάθε βήμα αυτής της δουλειάς ήταν καθοριστική για την εξέλιξή της, και η αληθινή φιλία του με τιμάει πολύ.

Ένα δεύτερο, στον αγαπημένο μου φίλο (και, τώρα πια, κουμπάρο) Σπύρο Ψυχή, τόσο για την βοήθειά του στην δουλειά μου, από την οποία προέκυψαν κοινές δημοσιεύσεις, όσο και για την ψυχική στήριξή του στην προσπάθειά μου.

Ένα ακόμη ευχαριστώ σε όλους τους μεταπτυχιακούς και προπτυχιακούς φοιτητές των Εργαστηρίων Τηλεπικοινωνιών, Πληροφορίας και Δικτύων, που βοήθησαν πάντα σε ότι μεγάλο ή μικρό τους χρειάστηκα.

Να είστε όλοι καλά.

Πολυχρόνης Κουτσάκης

Χανιά, Νοέμβριος 2002

Abstract

In this work, we consider the design and performance evaluation of multiple access schemes that multiplex voice traffic at the talkspurt level to efficiently integrate voice, video and data traffic in third generation picocellular wireless networks. We also develop Call Admission Control and Traffic Policing Algorithms for transmitting multiple-layered videoconference movies over a wireless channel of high capacity. The design goals include maximizing the system capacity and satisfying quality of service requirements such as voice packet dropping probability, video packet dropping probability and voice, video and data packet access delays. These goals are complicated by the limited radio channel bandwidth and by the contradictory nature of voice, video and data traffic.

ΠΕΡΙΕΧΟΜΕΝΑ

Σελ.

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Introduction

The first part of this chapter contains a brief reference on wired and wireless networks and a review of the Asynchronous Transfer Mode (ATM) technique and of the combination of wireless networks with ATM. The second part of the chapter introduces our specific field of work and the third part focuses on the description of the contribution of our work to the field.

I. Background

1. Wired and Wireless Networks

The coexistence of wired and wireless communication networks is a hint that none of the two network types is able to service, on its own, all of the users needs. The basic advantage of wireless networks is the user's ability to move, even during the use of the network services. Clearly, this is something that wired networks can not offer. This ability is of great importance to users who, either because of the nature of their work or due to recreation reasons, are constantly in motion and simultaneously, need to communicate. On the other hand, wired networks offer reliable and high-speed communication, thus giving the users the ability to exploit a large variety of services. Especially after the introduction of fiber optics, the transmission speed and reliability have increased considerably, so that a wide variety of services can be offered (e.g., video, high quality image and sound applications). On the contrary, the use of the air as transmission medium in wireless networks leads to limited channel bandwidth, which in turn means limited transmission speed, as well as a high error rate, which is not acceptable by most of the applications. For this reason, the applications supported until today by the wireless cellular communications networks involve mostly voice and low rate data transmission (e.g., Short Message System).

However, recent technological developments in wireless communications, especially regarding the speed and quality of the transmission, have made possible the support of many more applications. These developments include:

- the capacity increase of wireless channels, mostly due to the improvement of the transmitter and receiver technologies
- the reduction of transmission power, which leads to greater independence for the wireless terminals
- the development of simpler and more easy-to-use equipment
- the reduction of the total cost for building and maintaining a wireless network [30].

Still, these technological achievements are not by themselves enough to efficiently support the new applications. The existing architectures and protocols have been designed to efficiently support the existing voice and low data rate applications, thus their use in new applications presents problems. For this reason, a new network design is required, both in the architecture and the protocol level. This way, the new applications will be adequately supported, and the network resources will be used more efficiently.

The ATM (Asynchronous Transfer Mode) technique can solve many problems towards this direction. This technique, which is briefly described in the next section, has the ability to adjust to different traffic conditions and QoS (Quality of Service) requirements, therefore it can support a large variety of applications. Still, the fact that the original design of ATM did not include the support of wireless networks, makes certain adjustments necessary.

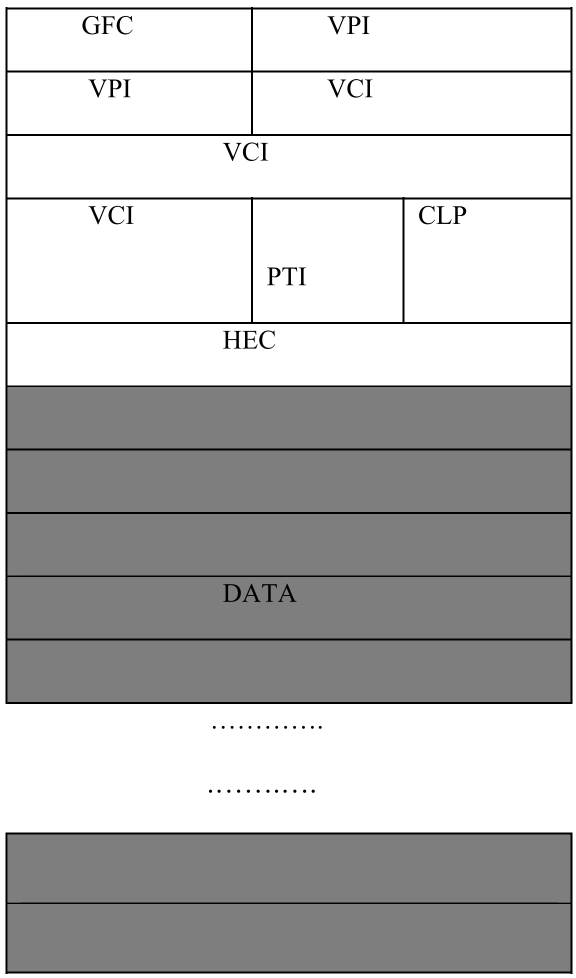
2. Asynchronous Transfer Mode

The ATM has been selected as the official transmission mode for BISDN (Broadband Integrated Services Digital Network). BISDN is a network architecture supporting a large applications' spectrum (voice, image, multimedia, etc.). The term ″asynchronous″ is not referring to the actual transmission, which in most cases is synchronous, but to the way the available bandwidth is reserved. The channel time is divided in fixed size time slots, which are dynamically reserved by the various network users, depending on their needs.

The ATM technique is defined with the help of a set of principles [31]:

- The information is transmitted with fixed length data units, which are called cells. The cells consist of a header and a data field. The cell structure is described below.
- ATM uses virtual circuits to transmit information.
- The main use of the cell header is the identification of the cells which belong to the same connection.
- The basic unit of an ATM network is the ATM switch. An ATM switch can be characterized as a network device with many ports. Its role is to switch the cells it receives from the input ports to the appropriate output ports.
- The cell identification headers are of local importance only. They are not absolute addresses and they are translated in each ATM switch.

The cell size is 53 bytes. Of these, 5 bytes consist the header, and the remaining 48 bytes are used for information transmission (payload). The cell structure is presented in the following figure, where each row represents one byte of the cell.



The use of the header fields is:

- **GFC (G**eneric **F**low **C**ontrol): (4 bits). It is used to regulate the information rate in an ATM network. Its full use is still an object of research.

- **VPI (V**irtual **P**ath **I**dentifier): (8 bits). It defines a set of virtual circuits, which are routed through the same path. VPI, together with VCI, determines the connection to which the cell belongs.
- **VCI (V**irtual **C**hannel **I**dentifier): (16 bits). It uniquely defines a connection between two end nodes.
- **PTI (P**ayload **T**ype **I**dentifier): (3 bits). It determines whether the data belong to a certain user or consist of network administration information.
- **CLP (C**ell **L**oss **P**riority): (1 bit). It determines whether the cell is of low priority and thus can be dropped in case of congestion somewhere in the network.
- **HEC (H**eader **E**rror **C**ontrol): (8 bits). It is used by the end node, for error control of the cell's header.

To achieve high quality of service, the connections are classified in five classes [32]:

1. **C**onstant **B**it **R**ate (**CBR)**. These are connections which demand a constant bandwidth for their entire duration. The applications which use this type of connection are usually voice or circuit emulation applications and have strict requirements on the cell transmission delay (maximum value and jitter).
2. **r**eal-**t**ime **V**ariable **B**it **R**ate (**rt-VBR)**. This type of connections is mainly used for real time applications (i.e., applications which demand very low delays). In these connections, the required transmission rate, and consequently the bandwidth demands, vary over time. Cells which fail to be transmitted within the maximum permissible delay are considered of

extremely low value for the corresponding applications, and are usually dropped. The videoconference connections used in our study are of this type.

3. **n**on-**r**eal-**t**ime **V**ariable **B**it **R**ate (**nrt-VBR)**. These are connections similar to the ones described above, but their delay requirements are loose (i.e., not strict). On the other hand, these connections have very small tolerance in packet losses, compared to the real-time VBR connections, and for this reason advanced control and error detection/ correction mechanisms are needed for these connections. Examples include internetworking (Web hosting or e-commerce), store and forward non-real-time video and audio, client-server data applications.
4. **A**vailable **B**it **R**ate (**ABR)**. The ABR connections have neither delay nor delay jitter requirements. Their only requirement from the network is that they are provided with bandwidth at least equal to a declared minimum value. Examples include a wide variety of applications wherein a connection is setup before data is transmitted, such as Internet services (FTP, Internet browsing).
5. **U**nspecified **B**it **R**ate (**UBR)**. These connections are used for non-real-time applications, with neither delay nor delay jitter requirements, and with no minimum bandwidth demand. Classic data computer communication applications, such as electronic mail, belong to this class. Since there is no minimum bandwidth demand, these connections have the lowest priority and use the bandwidth that is left unused by all the connections of the other classes.

As a multiplexing technique, ATM has significant advantages in comparison to the Synchronous Transfer Mode (STM), which uses static bandwidth reservation. The advantages mainly focus on the efficient use of bandwidth. According to STM, the channel time is organized in frames of equal duration, which in turn are divided into time slots. For each connection, one time slot is reserved in each frame, and the information is transmitted within this slot. An STM channel is defined by the location of its corresponding time slot within the channel frame. Clearly, STM is very efficient for CBR connections. However, its efficiency falls dramatically when it has to support VBR connections, since is has to reserve for each connection bandwidth equal to its maximum transmission rate.

In an ATM network, the support of VBR connections is facilitated via the use of **statistical multiplexing**. With statistical multiplexing, the sum of the maximum transmission rates of all the active connections can surpass the channel bandwidth, while the available bandwidth is dynamically apportioned depending on the (variable) connection bandwidth requirements. With this method, a better usage of the network is achieved, along with the ability to support more concurrent connections. Still, the danger exists that, when the sum of the connections transmission rates surpasses the channel capacity, congestion might occur. In this case, some packets are either dropped, or temporarily kept in storage buffers, which results in increased packet delays. Some of the dropped packets must be retransmitted, which further aggravates the network congestion.

In order to deal with congestion, there is the need of traffic control algorithms, which must be capable to easily adjust to the different QoS requirements of the various connections. For example, voice applications have strict requirements regarding the maximum cell transmission delay and delay jitter. On the other hand, data applications usually demand very low error probabilities and retransmission of dropped cells, while their delay requirements are either loose or non-existent.

The traffic control algorithm is divided in three parts:

1. Call Admission Control (CAC)
2. Usage Parameter Control (UPC)
3. Congestion Control

With CAC, the network decides whether it will accept a new connection or not. The decision can be based on:

i) the characteristics of the traffic that the connection intends to introduce into the network,

ii) the connection QoS requirements

iii) the current network status

The traffic characteristics and the QoS requirements are defined through a set of parameters which are declared by the user when applying for the establishment of a connection. The most common parameters are the mean and peak cell rate, the maximum burst size, the maximum cell delay tolerance and the maximum cell delay variation tolerance. Based on these parameters, the CAC algorithm decides whether it would accept the new connection or not. The algorithm should not be too strict, which would lead to low network utilization, nor too loose, thus creating congestion. If the connection is accepted, it is assumed that the network and the user agree on a traffic contract. Based on this contract, the network is bound to satisfy the user's QoS requirements, and the user is bound to operate according to its traffic description parameters [33].

However, in ATM networks bandwidth reservation is dynamic, so there is nothing that prevents a connection from violating the traffic contract and transmitting at a rate higher than the one stated. This could happen not only because of the user's bad intention, but also because of a miscalculation of the required bandwidth. Since the ATM cells corresponding to the additional rate might cause network congestion, a mechanism of usage parameter control is necessary. This mechanism controls the traffic which the user introduces into the network, and its aim is to protect the network and the other users from violations of the traffic contract, deliberate or not. A good usage parameter control mechanism should combine [34]:

- simplicity, so that it can be easily applied by the user
- responsiveness to traffic parameter violations
- tolerance, because of system inaccuracies.

Despite the use of call admission control and usage parameter control mechanisms, there still exists the possibility of congestion in some part of the network, because of temporary buffer overload. The aim of congestion control is to recognize the phenomenon and to implement mechanisms that will reduce its consequences. The congestion controls which have been proposed for use in ATM networks are preventive. The use of reactive congestion control methods is forbidden, mainly for two reasons:

i) the response of such methods to congestion is very slow, and this could be destructive to real-time applications, and

ii) the propagation delay is very large, compared to the very small cell transmission time. Thus the number of cells introduced in the network within a few propagation delays time after the congestion begins, can be very big.

Thus, even if the reactive congestion control mechanism reacts fast, congestion will further build up and continue for quite some time. For this reason, preventive methods have been suggested, in order to avoid the congestion problem [35-37].

3. Wireless Networks

The goal of wireless communication networks is to allow user access to the capabilities of the wired networks from everywhere. The development of digital coding techniques and the continuous increase in integrated circuits density have made possible the realization of second generation wireless digital systems. The digitization allows the use of Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) techniques, which present significant advantages in comparison to the Frequency Division Multiple Access (FDMA) technique, which is used in analog systems.

With the TDMA technique, the wireless channel is divided in time slots, which can be dynamically apportioned to each user, depending on the user current needs. This leads to better channel utilization. With the CDMA technique, one frequency can be simultaneously used by many users, and this is achieved because the different signals are separated by using different modulation codes. Other advantages of the TDMA and CDMA techniques are [38]:

- smoother coexistence and communication with the digital wired networks
- capability of voice and data integration
- capability of future channel capacity increase, with the development of voice coders
- lower required transmission power (thus bigger independence duration)
- smaller system complexity.

4. Combination of Wireless Networks and ATM Technology

The combination of wireless networks and ATM technology aims at combining the advantages of wireless communication and the users freedom of movement, with the capability of serving different traffic types and the guarantee of QoS that ATM provides. This combination is not easy, because ATM has been designed for environments very different than those of wireless networks. More specifically, ATM assumes:

- stable users
- ample and constant bandwidth, which can be dynamically apportioned, depending on the current connection needs
- full duplex and point-to-point transmission
- very good transmission quality, and for this reason the incorporated detection and error correction techniques are very limited
- small bandwidth aggravation from the physical layer (physical overhead).

On the contrary, in the wireless environment

- user movement is considered certain
- the wireless channel bandwidth is small and variable, depending on the channel quality
- the transmission is usually half-duplex, because of lack of available frequencies, and it is point-to-multipoint, so there is the need for a more specific routing
- the transmission quality is bad, and for this reason well developed detection and error correction techniques are required
- the aggravation from the physical layer is large, mainly due to the synchronization delay between the transmitter and the receiver [39].

From all the above, it is clear that it is necessary to design specific techniques for the wireless segments of an ATM path, in order to guarantee QoS comparable to that of wired ATM connections. There are three basic requirements which must be fulfilled in a wireless ATM system:

1) efficient support of all ATM classes in the wireless channel (this is achieved with the use of an efficient multiple access control protocol),
2) limitation of the consequences of the relatively high error rate (this is achieved with the use of an error control mechanism),
3) avoidance of large cell dropping during user handover (several proposals exist for an efficient procedure which could achieve this, e.g. [40-42]).

II. Our field of work

1. Significance of the efficient support of all ATM classes in the wireless channel

Our work focuses on the first of the three requirements of a wireless ATM system mentioned above. In this section we will analyze its significance.

One of the most important developments in information technology in recent years has been the emergence of multimedia communications. Multimedia data consists of a range of different types of information integrated together, including text, audio, images and video sequences. Multimedia communication involves the transmission of these data types over communication networks.

High-speed packet-switched network architectures have the ability to support the wide variety of multimedia services, the traffic streams of which will have widely varying traffic characteristics (bit-rate, performance requirements). The main goal of wireless communication is to allow the user access to the capabilities of the global packet-switched network at any time without regard to location or mobility. This will be achieved via major advances in wireless digital communications technology and due to the continuous proliferation of small, portable and inexpensive computing devices.

Of the different classes of multimedia data, digital video information presents the greatest challenge to designers of communication networks and systems. Real-time digital video transmission has a number of demanding requirements: it needs a high transmission bandwidth, it must be transmitted with minimal delay, and it cannot tolerate a high error rate. The reason for these more stringent requirements is that, compared to packet telephony, video services are more prone to cell loss. Because of the high bit rate involved, cell loss occurs more frequently for video services, e.g., for a cell loss rate of 1E-8 and a cell length of 32 bytes, the average period between two lost cells for a HDTV video service encoded at 140 Mbps is 3 min, while for 64Kbps telephony, under the same conditions, the average period is more than 4 days. If a cell is lost in the middle of a video frame, where the primary video frame is transmitted over multiple ATM cells, then the phase alignment of subsequent primary frames is lost for the rest of the video frame, and, depending on the synchronization scheme used, a cell loss may thus corrupt a large part of the frame. These kinds of errors can be compared to service interruptions, and should occur less than once a month [54].

2. Cellular Wireless Networks and their Evolution

Most of the current and future wireless networks are and will be based on the cellular concept [30]. Rather than using one base station (and high transmission power) to cover a large service area (e.g., a city), multiple base stations are used to partition the area into *cells*[1]. Since identical carrier radio frequencies can be used in cells that are far enough apart to eliminate interference, multiple mobile-base pairs can use the same frequency simultaneously (*frequency reuse*). A procedure known as *handover* is used to maintain continuous coverage over the service area[2]. It is interesting to observe that the cellular concept is a staple of the broadcast radio and television industries, thus the significance of the work in [30] is that frequency reuse is accomplished over much smaller geographic areas. The ability to maximize system capacity through frequency reuse, along with advances in microelectronics technology, make mobile telecommunication attractive to both the service provider and the consumer.

---

[1] The information unit named "cell" in Part I of the introduction is from now on going to be named "packet".
[2] Whenever the mobile crosses cell boundaries, it must change its frequency to correspond with that of the new base station.

First generation cellular systems have been designed for high power and high mobility (vehicular) voice users. They employ analog FM radio technology and frequency division multiplexing to provide channel access. The first generation North American cellular system, Advanced Mobile Phone Service (AMPS), was deployed in 1983. Other first generation systems based on different standards were used in European countries (e.g., UK, France and Germany each had their own) and Japan.

Second generation cellular systems exploit high power (0.5-5 W) digital radio technologies to increase system capacity, while continuing to focus primarily on vehicular voice users. The second generation Pan-European system, Global System for Mobile Communications (GSM), is based on TDMA technology [43]. GSM was deployed in 1993 and, as well as being implemented throughout Europe, has been adopted by many countries around the world. There are now over 100 members of the GSM Memorandum of Understanding (MoU) group, and around 70 GSM networks are operating in Europe, the Middle East, Africa, Asia Pacific and Australasia. What began as a pan-European technology has become the world's most widespread digital cellular standard.

Additionally, one will see reference to Digital Cellular System 1800 (DCS 1800) which is GSM operating at 1.8 GHz [44]. Several standards exist for second generation systems in North America: IS-54/IS-36 are based on TDMA [45]; and, IS-95 is based on CDMA, a direct sequenced spread spectrum technology [46,47]. IS-54 has been introduced into the largest cellular markets and IS-95 has been deployed in several parts of the USA. Finally, we note that

Japan has deployed its own second generation cellular systems, Personal Digital Cellular (PDC), based on TDMA [48].

As for the 2.5 generation cellular systems:

The European Telecommunication Standards Institute (ETSI) has developed a set of standards that define a General Packet Radio Service (GPRS). GPRS allows mobile devices to be connected via their IP addresses. It will allow digital transmissions to be undertaken at speeds up to 115kb/s per channel, using continuously enabled connections that will only read the packets that are addressed to them.

As services are introduced, GPRS will provide cost effective and efficient use of network resources for packet mode data applications that exhibit one or more of the following characteristics:

- intermittent, non-periodic (i.e. bursty) data transmissions in which the time between successive transmissions greatly exceeds the average transfer delay
- frequent transmissions of small volumes (<500 octets) of data
- infrequent transmission of larger volumes of data.

Within GPRS, two different bearer service types have been defined:Point-To-Point (PTP) and Point-To-Multipoint (PTM). Point-To-Point Connectionless Network Service (PTP-CLNS) provides a service in which single packets are transmitted from a service subscriber to a destination user. Point-To-Point Connection Oriented Network Service (PTP-CONS) provides a service in which multiple packets are sent between the service subscriber and a single destination user. PTM - Multicast (PTM-M) allows messages to be transmitted throughout geographical areas defined by the service requester. PTM - Group Call (PTM-G) allows messages to be transmitted to a specified receiver group. IP Multicast (IP-M) allows messages to be sent, using the Internet Protocol, between participants of an IP-M group.

GPRS packet-based service should cost users less than circuit-switched services since communication channels are being used on a shared-use, as-packets-are-needed basis, rather than dedicated only to one user at a time.

Another interesting development from ETSI is EDGE (Enhanced Data rates for Global Evolution). This will allow GSM operators to use existing GSM radio bands to offer wireless multimedia IP-based services and applications at speeds up to 384kbit/s. EDGE uses the same TDMA frame structure, logic channel and 200kHz carrier bandwidth as today's GSM networks, which allows existing cell plans to remain intact. EDGE is commercially available since 2001, and has been implemented in North America (North Carolina, South Carolina, coastal Georgia, eastern Tennessee and Puerto Rico). In Europe, the main focus remains on WCDMA (Wideband Code Division Multiple Access) technology. The trans-atlantic difference is caused by two factors: the difference in frequency distribution among operators, and the competition that GSM faces from CDMA 2000 in the US. In Europe, WCDMA has a new frequency spectrum that gives operators capacity. In the US frequencies are divided in a way that leaves no available spectrum for WCDMA. The implementation of EDGE in Europe is expected to lift GSM services (in Europe all operators use the GSM standard) so they work well in both WCDMA and GSM.

The first "2.5-generation" cellular phone system in the U.S. has been introduced in 2001 in the Seattle area, in the guise of an AT&T Wireless GPRS (General Packet Radio Service) system, and it's expected to expand. AT&T Wireless will provide unified messaging services, including cellular e-mail and a unified e-mail box utilizing text-to-speech technology, which enables e-mail to be sent as either text or voice and to be read as text or heard as voice, regardless of how it was sent, as well as voice mailboxes. Initially, the only phone able to work with this service, a Motorola 7382i Timeport, is limited to a maximum of 40 kilobits/second [59]. AT&T doesn't yet have roaming agreements in place with other carriers; that means that these phones' voice and data will go mute outside of their cities. But these limitations will surely pass, with the system eventually supporting a real-world data rate of about 100 kilobits/second. Then, about a year from now, AT&T plans to upgrade this system to EDGE technology, which should raise the peak data rate to a respectable 388 kilobits/second. This faster data service will initially cost $50 for 400 minutes of voice, and 1 megabyte of data.

Still, this decade is the Asian mobile decade. In 2000, while Europeans sent data over their cellular networks at 9.6 and North Americans at 14.4 kilobits per second, 64-kilobit-per-second networks were operating in Japan and Korea. Now while Europe is rolling out 2.5-generation GPRS service and North Americans hope to follow the European lead, the Japanese are already constructing the infrastructure for third-generation cellular service and plan to have the system in operation before the end of 2002. In Japan, i-mode is the new platform for mobile phone communications that has revolutionized the way nearly one-fifth of the people in Japan live and work. Introduced in February 1999, this remarkably convenient, new form of mobile service has attracted over 28 million subscribers. With i-mode, cellular phone users get easy access to more than 40,000 Internet sites, as well as specialized services such as e-mail, online shopping and banking, ticket reservations, and restaurant advice. Users can access sites from anywhere in Japan, and at unusually low rates, because their charges are based on the volume of data transmitted, not on the amount of time spent connected. NTT DoCoMo's i-mode network structure not only provides access to i-mode and i-mode-compatible content through the Internet, but also provides access through a dedicated leased-line circuit for added security. Finally, we must note that in China, the number of cellular subscribers jumped to 85 million at the end of 2000, almost double that of a year earlier, while growth rates were much lower in Europe and North America.

Ideally, using digital technology, third generation wireless networks will integrate the needs of pedestrian users (low power and low mobility) with the needs of the vehicular users, and provide ″seamless″ integration of various types of service (e.g., cordless phones, paging systems and wireless data networks) as well. Two other major efforts that are underway to develop mobile telecommunication standards for the emerging third generation systems are the universal mobile telecommunication system (UMTS) which is being developed as a Pan-European standard, and the future public land mobile telecommunications (FPLMTS) which is being developed as a global standard by the International Telecommunications Union (ITU) [49,50].

3. The need for an efficient MAC protocol

It is generally recognized that a hierarchical wireless network architecture is required to meet the needs of high and low mobility users, and to increase system capacity and coverage area.

Today's combination of cordless telephones and cellular networks is a simple case of a two-tiered solution.

In the two-tiered architecture proposed in [14], the low mobility user is served by microcells (on the order of 100m radius), while the high mobility user continues to be served by macrocells (minimum radius of approximately 1km). Handovers occur whenever a pedestrian crosses microcell boundaries and whenever a vehicular user crosses macrocell boundaries. To exploit the existing cellular structure, the microcells are embedded within the cellular grid. As shown in Figure 1.1, the macrocell base stations are star connected to a Mobile Switching Center (MSC) that provides access to the wired network infrastructure, and the clustered microcell base stations are connected to the macrocell base station over a high speed local area network to localize the network control associated with microcell handovers.

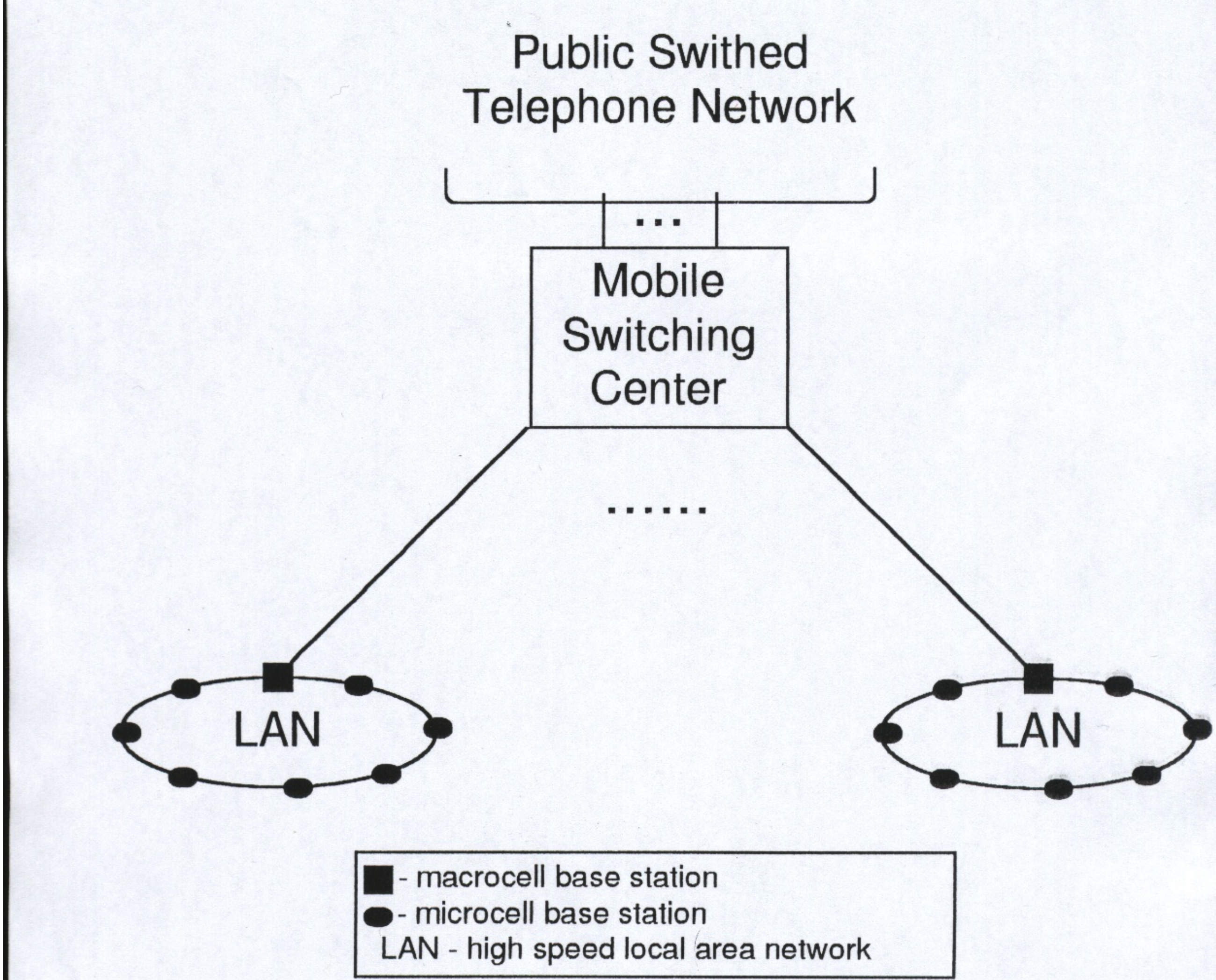


Figure 1.1 Illustration of the two tiered architecture proposed in [14].

In a third generation cellular system, the system capacity can be increased by:

a) using a cellular structure with a cell size as small as possible (microcells and picocells) to increase frequency reuse. Microcell (picocell) diameters are usually of the order of a few hundred (dozen) meters, therefore the round-trip propagation delay within a microcell is negligible (of the order of 1 μs or less).

b) using efficient multiple access control (MAC) protocols to exploit the variations in access and service required by disparate sources.

A well designed multiple access protocol will reduce system costs by maximizing system capacity, integrating different classes of traffic (e.g., voice, data and video, as opposed to today's picture, where wireless networks are essentially optimized for voice communications only), and satisfying the diverse and usually contradictory QoS requirements of each traffic class (such as voice packet dropping probability, voice packet access delay, video packet dropping probability and data message delay) whilst apportioning the limited radio channel bandwidth among them.

Most multiple access protocols are based on one of the known multiple access techniques, i.e., FDMA, CDMA or TDMA. In wireless ATM networks, the lack of available frequencies and the demand for dynamic bandwidth reservation, especially for VBR connections, make the use of FDMA inefficient. On the other hand, CDMA limits the maximum transmission rate of a connection to a relatively low value, which is a problem for broadband applications (i.e., requiring transmission rates >2 Mbps-WCDMA is being introduced to tackle this problem). Thus, most protocols in this field use a dynamic TDMA scheme [51], due to its capability to adjust to the needs of each connection, reserving more or less time slots in each case.

III. Our work

1. Previous work

In our previous work, which was conducted during an M.Sc.Thesis and led to journal publications [2, 18], two new medium access control (MAC) protocols for TDMA mobile wireless communications have been proposed and investigated. These protocols are based on reservation random access algorithms, and the mobile terminals make their slot reservations by transmitting packets in the *request* slots of equal-length time frames. The *information* slots of the time frames are allocated by the base station scheduler to the mobile terminals based on their requests and the quality of service (QoS) that each (voice, data or video) terminal demands. A summary of the work conducted and the main conclusions and contributions are as follows.

a) Via an extensive simulation study, we explored the performance of the first protocol when integrating voice and data traffic over two wireless channels, one of medium capacity (referring mostly to outdoor microcellular environments) and one of high capacity (referring to an indoor picocellular environment). Data message arrivals were assumed to occur according to a Poisson process and to vary in length according to a geometric distribution. We evaluated the voice packet dropping probability and access delay, as well as the data packet access and data message transmission delays for various voice and data load conditions. Two novel protocol ideas of ours (<u>the sharing of certain request slots among voice and data terminals with priority given to voice</u>, and <u>the use of a fully dynamic low-voice-load mechanism</u> in order to minimize unnecessary data message delays, reduce the voice packet dropping probability and hence increase the channel throughput by ″exploiting″ all the available information slots within each frame in the best possible way) were combined with two useful ideas which have been proposed in other MAC schemes [5,9], and this combination led to highly efficient voice sources multiplexing results along with most satisfactory voice and data performance and QoS requirements servicing, as shown from the comparison with other schemes in the field [18,7,10].

b) The performance of the second protocol was again explored via simulations, when integrating voice, video and data packet traffic over a wireless channel of high capacity. Both the video model and the actual video streams used in our study refer to videoconferencing traffic. Depending on the number of video users admitted in the system, the protocol varied: 1) the request bandwidth dedicated to resolving the voice users contention, and 2) the probability with which the base station grants information slots to successful voice requests, in order to preserve full priority for video traffic. We evaluated the voice packet dropping probability and mean access delay, as well as the video packet dropping probability, for various voice and video load conditions, and the data packet access and data message delays. Our scheme achieved high voice capacity and aggregate channel throughput results along with voice, video and data quality of service (QoS) requirements servicing, in all cases of traffic load, and provided better results than another efficient scheme which used the same traffic parameters [2,19].

2. The work in this Dissertation

In this Dissertation, we design a dynamic TDMA-based multiple access scheme which multiplexes voice traffic at the vocal activity (talkspurt) level to efficiently integrate voice

(Constant Bit Rate, CBR On/Off Traffic), video (Variable Bit Rate, VBR) and bursty data traffic in outdoor and indoor microcellular (picocellular) environments. We evaluate our scheme by testing its performance with various video sources, data traffic models and channel conditions. We also propose a Call Admission Control (CAC) and Traffic policing mechanism for transmitting multiple-layered videoconference streams over a wireless channel of high capacity. We focus both on MPEG-4 and H-263 coded videos, and evaluate the performance of the proposed mechanism.

Within the microcell (picocell), spatially dispersed source terminals share a radio channel that connects them to a fixed base station. The base station allocates channel resources, delivers feedback information and serves as an interface to the MSC. Since the base station is the sole transmitter on the downlink channel, it is in complete control of the downstream traffic, using TDMA to relay information to the users. Thus, we focus on the uplink (mobiles to base station) channel, where a MAC scheme is required in order to resolve the source terminals contention for channel access.

We assume that voice, video and data packet traffic is generated by mobile users who access the network with small, lightweight and low-power devices (i.e., in the category of low tier PCS [52]).

Speech alternates between periods of talk (talkspurts) and silence. Thus, voice terminals only require channel access during talkspurt and the time periods corresponding to silence gaps within a conversation can be used to transmit packets from other source terminals (i.e, multiplexing occurs at the talkspurt level). Both voice and video packet delay requirements are strict, because delays in both types of communication are annoying to a listener or viewer. Thus, voice and video packets must be delivered within specified maximum delays. Whenever the delay experienced by a voice or a video packet exceeds the corresponding maximum delay, the packet is dropped. Speech can withstand a small (1-2%) amount of dropped packets without suffering large quality degradation [4], at least one which can be perceived by humans. Video traffic is even less tolerant in the amount of dropped packets that it can withstand (0.01-0.02%) [19]. On the other hand, data applications are more tolerant of delays (delays of up to 200 ms are often acceptable), but 100% delivery of correct packets is often required (e.g. in the case of a file transfer) [12].

This work is organized in six chapters, which actually comprise of two parts. The first part, which includes Chapter 2 only, is an extension of our previous work conducted in the framework of the M.Sc. Thesis, as our scheme integrates voice, video and bursty data packet traffic and we compare our results with those in [2]. The second part, which includes Chapters 3-5, is a presentation of the work done based on our idea that, although our previous work showed remarkably high throughput results, its performance evaluation was based on the simulation of theoretical traffic models and not of actual traffic streams and on ideal channel conditions. Therefore, in this second part, our simulations considered the effect of channel errors and we used actual videoconferencing (MPEG-4 and H.263) streams and e-mail data models fitted to actual email data traffic, for video and data traffic, respectively.

Chapter 6 contains the main Conclusions of our work and its contribution to the field, as well as some ideas for future work.

Chapter 2

A Bandwidth Reservation Mechanism for Multimedia Integrated Access in TDMA-based High Capacity Wireless Networks

2.1. Introduction

This chapter presents the design and performance evaluation of our multiple access scheme which multiplexes voice traffic at the vocal activity (talkspurt) level to efficiently integrate voice (Constant Bit Rate, CBR On/Off Traffic), video (Variable Bit Rate, VBR) and bursty data traffic in high capacity wireless channel picocellular environments.

Within the picocell, spatially dispersed source terminals share a radio channel that connects them to a fixed base station. The base station allocates channel resources, delivers feedback information and serves as an interface to the mobile switching center (MSC). The MSC provides access to the fixed network infrastructure. We focus on the uplink (mobiles to base station) channel, where a MAC scheme is required in order to resolve the source terminals contention for channel access.

2.2. System Model

2.2.1 Channel Frame Structure

The uplink channel time is divided into time frames of equal length. The frame duration is selected such that a voice terminal in talkspurt generates exactly one packet per frame. As shown in Figure 2.1 (which presents an example of the channel frame structure), each frame consists of two *types* of slots. These are the *request* slots and the *information* slots.

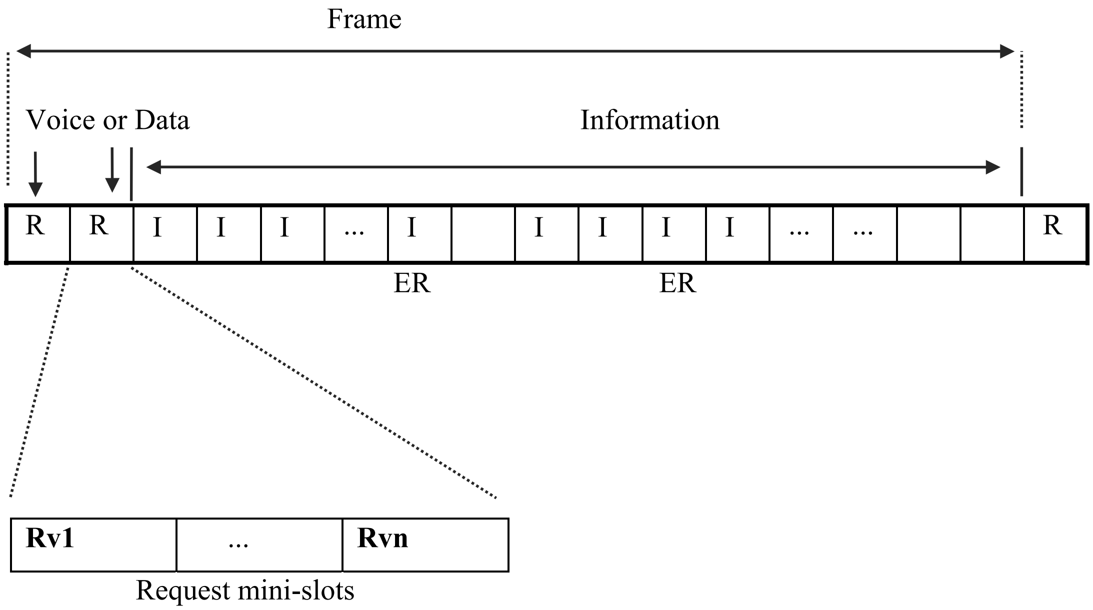


Figure 2.1. An example of a channel frame structure showing the request and information slots within a frame.

Each information slot accommodates exactly one, fixed length, packet that contains voice, video or data information and a header. As for the request slots, which are used by terminals to place their reservation requests to the Base Station (BS), *we introduce the idea* [17] *that the request slots can be shared by voice and data terminals (first by voice terminals and, after the end of voice contention, by data terminals)*, in order to optimize the request bandwidth. The request slots are subdivided into mini-slots and each mini-slot accommodates exactly one, fixed length, request packet. The request must include a source identifier. Since we assume that all of the voice source state transitions occur at the frame boundaries[3], we place all request intervals at the beginning of the frame, in order to minimize the voice packet access delay.

Voice and data terminals do not exhaust their attempts for a reservation within the request intervals. *Any other free, at the time, information slot of the frame can be temporarily used as an extra request slot (ER slot)* for voice or data terminals (with priority given to voice terminals by the base station). Each one of the ER slots is further divided into mini-slots, just like a standard request (R) slot. This approach has been introduced and implemented in [5,9].

By using more than one minislot per request slot, a more efficient usage of the available request bandwidth is possible. The number of minislots can not, of course, be increased at will, since a large number of minislots would mean that their duration would be very short and the time for the request packet to be transmitted to the BS and the BS to send an acknowledgement to the requesting terminal would not suffice. The concept of reserving a minimum bandwidth for terminals to make reservations helps to keep the voice access delay within relatively low limits and gives clearly better performance than the PRMA [6] and quite a few PRMA-like algorithms, such as DPRMA [19], where the absence of request slots leads to a continuously decreasing probability of finding available information slots as traffic load increases, and hence to greater access delays [2]. As shown from our results and also pointed out in [5], a request bandwidth of less than 3% of the total bandwidth is usually sufficient for high system performance (in the sense that is suffices for the requesting terminals to transmit their requests, while at the same time it does not consume a large portion of the bandwidth and leaves enough ″space″ for terminals with a reservation to transmit their information packets).

---

[3] The explanation for this assumption will be given in the next section.

No request slots are used for the video terminals, because of two reasons, which will be be analyzed in section 2.2.3.

2.2.2 Voice Traffic Model

Our primary voice traffic model assumptions are the following:

1) Voice terminals are equipped with a voice activity detector [6]. Voice sources follow an alternating pattern of talkspurts and silence periods (on and off), and the output of the voice activity detector is modeled by a two-state discrete time Markov chain.
2) All of the voice source transitions (e.g., talk to silence) occur at the frame boundaries. This assumption is reasonably accurate, taking into consideration that the duration of a frame is equal to 12 ms here, while the average duration of the talkspurt and silence periods exceeds 1 sec.
3) The voice delay limit is equal to 40 ms. [19]

2.2.3 Video Traffic Model

We adopt the same video traffic model with the one in [19]. This model is based upon work done by Heyman et al [15,55]. In this study of actual videoconferencing traffic, video frames (VFs) were found to be generated periodically and to contain a varying number of cells in each frame. The distribution of the number of cells per VF was found to be described by a gamma (or equivalently negative binomial) distribution. A Markov chain model can be constructed that demonstrates the transition from one state to the next. A ″state″ represents the number of video packets (cells) that a video frame contains. The transition matrix is computed as:

$$\boldsymbol{P} = \rho\, \boldsymbol{I} + (1\text{-}\rho)\, \mathbf{Q} \qquad (1)$$

where **I** is the identity matrix, ρ is the autocorrelation coefficient (0.98459 from [16]), and each row of the **Q** matrix is composed of the probabilities ($f_0, \ldots, f_M, \boldsymbol{F}_M$). The quantity fm has the negative binomial distribution and represents the probability that a video frame contains m cells. The value of M represents the peak cell rate and $\boldsymbol{F}_M = \Sigma_{k>K}\, fm$.

The statistics for video conferencing traffic that were obtained in [55] were the result of coding a video sequence with a modified version of the H.261 standard. The results showed a peak cell generation rate of 220 cells/VF (2.112 Mbps), an average generation rate of 104.8 cells/VF (1.006 Mbps), and a standard deviation of 29.7 cells/VF (0.285 Mbps). The cell size was taken equal to 48 bytes, which is equivalent to the ATM cell size. New VFs are assumed to arrive every 40 msecs (i.e., 25 VFs per second).

2.2.4 Data traffic models

Data traffic has low priority compared to voice traffic, and data messages are generated by a large unknown number of data terminals (theoretically infinite). Data messages vary in length according to a geometric distribution with parameter q and mean B=1/q. B is expressed in packets per message, and the data rate packet arrival is equal to λB packets/frame, where λ is the data message arrival rate.

We consider two different cases for the data message arrivals:

1) The aggregate data message arrivals are Poisson distributed with mean λ messages per frame, and

2) the interarrival times of the aggregate data message arrival process are assumed independent, and identically distributed according to a Pareto distribution with shape parameter α and location parameter k. We use the bursty Pareto model because it simulates well data traffic generated by interactive applications, the importance of which is growing and will be even more significant in future mobile systems [20,21]. The location parameter k corresponds to the minimum interarrival time between packets. It can be easily shown that if $\alpha<=2$, the distribution has infinite variance, while if $\alpha<=1$ the mean is infinite as well. Therefore, the Pareto distribution is heavy tailed with infinite mean and variance. This mathematical property accounts for the strong burstiness of the data message arrival process in this case.

Finally, an upper limit on the mean data message delay, equal to 200 ms, is assumed (the data message delay is defined as the time period from the instant a data terminal generates a message, until it completes the transmission of the last packet of its message in a reserved slot). Many types of data transmitted over wireless links today are "soft real-time", which means that the interaction between users is not as demanding as that of video and voice terminals, respectively, but still the data transmitted need to be delivered without errors and within a reasonable, short amount of time.

2.2.5 Actions of Mobile Terminals

Voice terminals with packets, and no reservation, contend for channel resources using a random access protocol to transmit their request packets only during the request intervals. The base station broadcasts a short binary feedback packet at the end of each mini-slot indicating only the presence or absence of a collision within the mini-slot (collision (C) versus non-collision (NC)).

Since the feedback packet is short (several bits) and the propagation delay within a picocell is negligible, we assume that the feedback information is immediately available to the terminals (i.e., before the next mini-slot). Upon successfully transmitting a request packet the terminal waits until the end of the corresponding request interval to learn of its reservation slot. If unsuccessful within the request intervals of the current frame, the terminal attempts again in the request intervals of the next frame (based on the transmission protocol described in 2.2.7). A terminal with a reservation transmits freely within its reserved slot. Generally, a terminal that fails to transmit a request tries again in successive frames until it succeeds. However, since voice packets that age beyond the voice delay limit are dropped, a voice terminal may stop transmitting requests without ever succeeding, because all of its packets have timed out and it has transitioned into silence.

The same reservation procedure is followed by the data terminals, right after the end of the voice contention period, which is indicated by two consecutive NC feedback packets. Since data traffic is the most tolerant of delays, data terminals have the lowest priority in acquiring the information slots they demand. Each data user is allowed to reserve just one slot per frame, although the average data message length is assumed quite larger in our scheme. This “restriction” is explained later in this section.

Video terminals, as already mentioned in section 2.1 and shown in Figure 1, do not have any request slots dedicated to them. This happens for two reasons:

1) Video sources “live” permanently in the system, they do not follow an ON-OFF state model like voice sources.

2) Video traffic follows a multi-state Markov model, in which however state transitions do not occur very often.

Thus, there is no need for granting request bandwidth to the video terminals, as it would be wasted in most cases. *Video terminals convey their requirements (in slots) to the base station by transmitting them within the header of the first packet of their current video stream, in the first idle channel information slot.*

2.2.6 Base Station Scheduling

To allocate channel resources, the BS maintains a dynamic table of the active terminals within the picocell. Upon successful receipt of a voice or data request packet, the BS provides an acknowledgment and queues the request. The BS allocates channel resources at the end of the corresponding request interval, and follows a different allocation policy for each of the three types of terminals.

Video terminals have absolute priority in acquiring the slots they demand. If a full allocation is possible, the BS then proceeds to allocate any still available information slots to the requesting voice terminals. Otherwise, if a full allocation is not possible, the BS grants to the video users as many of the slots they requested as possible (i.e., the BS makes a partial allocation). The BS keeps a record of any partial allocations so that the remaining requests can be accommodated whenever the necessary channel resources become available. In either allocation type case, the BS allocates the earliest available information slots to the video terminals, which, if needed, keep these slots in the following channel frames, until the next video frame (VF) arrives.

Voice terminals which have successfully transmitted their request packets do not acquire all the available (after the servicing of video terminals) information slots in the frame. This is a choice we make, in order to prevent a percentage of the voice terminals (the talkspurt of which has a mean duration of 1 sec, i.e. longer than 80 channel frames) from entering the system. If this happened, voice terminals would keep their dedicated slots for the whole duration of their talkspurt, and thus video terminals, which need many slots, would not find enough slots to transmit in, and the particularly strict video QoS requirements would be violated. *In order to implement this idea of preventing a percentage of voice terminals from immediately entering the system, the BS in our scheme allocates a slot to each requesting voice terminal with a probability* ***p**** *(the value of which is determined via simulation, for various video traffic loads).* The requests of voice terminals which ″ fail ″ to acquire a slot, based on the above BS slot allocation policy, remain queued. The same holds for the case where the resources needed to satisfy a voice request are unavailable. Voice terminals with queued requests must continuously monitor the base-to-mobile channel. Reserved slots are deallocated immediately, i.e., a voice terminal holding a reservation signals the BS upon the completion of its talkspurt, by setting a special bit in the header of its last talkspurt packet. Upon call transmission completion, or when an active voice terminal exits the picocell (handover) the BS will delete the table entry after some prescribed period of time. Within each priority class, the queuing discipline is assumed to be First Come First Served (FCFS).

The above comments on why we need to use **p***, stand for the explanation of our “defensive” slot allocation policy to data users, as well. As video traffic is quite bursty, an effort of allocating to a data terminal as many slots as it needs in one frame, in order to transmit its whole message, would again result in the deterioration of the video QoS.

Finally, in order to preserve the video QoS within the desired strict limits, *we enforce the following scheduling policy for the video terminals*:

As already mentioned, video terminals acquire from the BS the earliest available information slots within a frame and they can keep transmitting in these slots in the subsequent frames, until the arrival of the next VF. Taking into account the fact that new VFs arrive every 40 msecs and that the channel frame duration is 12ms in our study (i.e., a VF arrives every 3.33 channel frames), we understand that it is quite possible that certain slots dedicated to a video terminal may happen to be placed after the point of arrival (PA) of the next VF (see an example in Figure 2.2a, where two of the slots allocated to a video terminal are placed after PA and they are marked with bold characters). Thus, in order to avoid an unnecessary dropping of video packets, *the BS in our scheme performs a "reshuffling "*. In each frame where a new VF is expected, the BS first finds all the available slots in the frame which precede PA, and then replaces as many as possible of the slots dedicated to the video terminal after PA with an equal number of available slots before PA (see Figure 2.2b). This way, video packets are lost only if the available slots before PA are less than the dedicated video slots after PA, or if there aren't enough available slots to accommodate the VF of a video terminal in the first place.

PA

Res	A	Res	A	Res	Res	A	**Res**	A	A	A	**Res**	A	A

Figure 2.2a. Status of information slots at the beginning of a frame in which a new VF will arrive.

PA

Res	**Res**	Res	**Res**	Res	Res	A	A	A	A	A	A	A	A

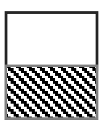
Figure 2.2b. "Reshuffling" done by the BS.

2.2.7 Transmission Protocols

Quite a few reservation random access algorithms have been proposed in the literature, for use by contending voice terminals to access a wireless TDMA channel (e.g., PRMA [6], Two-Cell Stack [12], Controlled Aloha [11], Three-Cell Stack [7]). In our study, we adopt the *two-cell stack* reservation random access algorithm, due to its operational simplicity, stability and relatively high throughput when compared to the PRMA (Aloha-based) ([8]) and PRMA-like algorithms, such as in [5,10,19]. In this protocol, the contending set of voice terminals is split probabilistically into two equiprobable subsets <u>at the beginning of a channel frame</u> (notice that the voice terminals split *before* the beginning of the voice contention. Our results for the algorithm without this initial split showed the voice capacity to be lower by several terminals). Only one of these subsets is transmitted in the first voice request minislot.

The operation of the collision resolution mechanism of the protocol can be visualized by a two-cell stack (as shown in Figure 2.3), where in a given voice request minislot the bottom cell contains the transmitting terminals, and the top cell contains the withholding terminals. If the transmitting set is found to contain more than one terminal (i.e., a collision occurs within the corresponding minislot), it is split probabilistically into two subsets (with equal probability), one of which remains in the transmitting cell while the other one joins the withholding cell. The end of the terminal contention for the request minislots is uniquely identified by the occurrence of two consecutive non-collisions. For more details on window type collision resolution algorithms the interested reader is referred to [11,53].

The actions of the data terminals, in order to acquire access to the channel resources, are similar to those of voice terminals. The *two-cell stack* blocked access collision resolution algorithm [13] is adopted for use by the data terminals in order to transmit their data request packets. This algorithm is of window type, with FCFS-like service.



r = 1, withhold

r = 0, transmit

Figure 2.3. Visualizing the two-cell stack algorithm.

2.3. System Parameters

We use computer simulations to study the performance of our MAC scheme. The simulations were conducted with the parameters contained in Table 2.1. Each simulation point is the result of an average of 10 independent runs (Monte-Carlo simulation), each simulating 305,000 frames (the first 5,000 of which are used as warmup period).

The channel rate is 9.045 Mbps (from [18,19]). The channel is error-free and without capture. The 12 ms of frame duration accommodate 256 slots, the size of which has been chosen equal to the size of an ATM cell. *As shown in Figure 2.4, the number of shared voice and data request slots is not fixed in the scheme. It depends on the number of video sources admitted into the system*[4]*, and it varies accordingly between 1 and 5 slots (see Table 2.2. where the values of the probability* ***p**** *are presented as well).*

Even for the case where 5 request slots are needed, this corresponds to a 1.95% request bandwidth only, which is well within the desired range (<3%) . We should note that:

1) In our design, we chose the number of minislots per request interval (4), to allow for guard time and synchronization overheads, for the transmission of a generic request packet (e.g., 40 bits long) that contains the source identifier, along with some data (e.g., priority, slots required, etc.), and for the propagation delay within the picocell (0.1μs for a picocell with radius 20m).
2) Because of assumption 3 of our voice traffic model, all voice request intervals are located at the beginning of each frame.
3) The maximum transmission delay for video packets is set to 40 msecs [19], with packets being dropped when this deadline is reached. That is, all video packets of a VF must be delivered before the next VF arrives. The maximum transmission delay for voice packets is also set to 40 msecs [19].
4) The allowed voice packet dropping probability is set to 0.01 [19], whereas the allowed video packet dropping probability is set to 0.0001 [19].

---

[4] The channel bandwidth consumed by each video source is large, and thus, when we examine cases with a small number of video sources, the system can accommodate a significantly larger number of voice sources (because each one of them consumes much less information bandwidth than a video source). In this case, more voice request slots are needed in order to allow voice sources to enter the system without significant dropping of voice packets.

Design Parameters	
Channel Rate (Mbps)	9.045
Speech Codec Rate (Kbps)	32 [9,12,18]
Frame Duration (ms)	12
Slots per Frame	256
Slot duration (μs)	46.875
Request slots per frame	1-5
Minislots per request slot	4
Packet size (bytes)	53 (5 header)
Voice delay limit (ms)	40
Mean talkspurt duration (s)	1.0
Mean silence duration (s)	1.35
Maximum voice dropping probability	0.01
Peak video bit generation rate (Mbps)	2.112
Average video bit generation rate (Mbps)	1.006
Standard deviation of video bit rate (Mbps)	0.285
Video delay limit (ms)	40
Maximum video dropping probability	0.0001
Maximum data message delay (ms)	200

Table 2.1. Experimental System Parameters

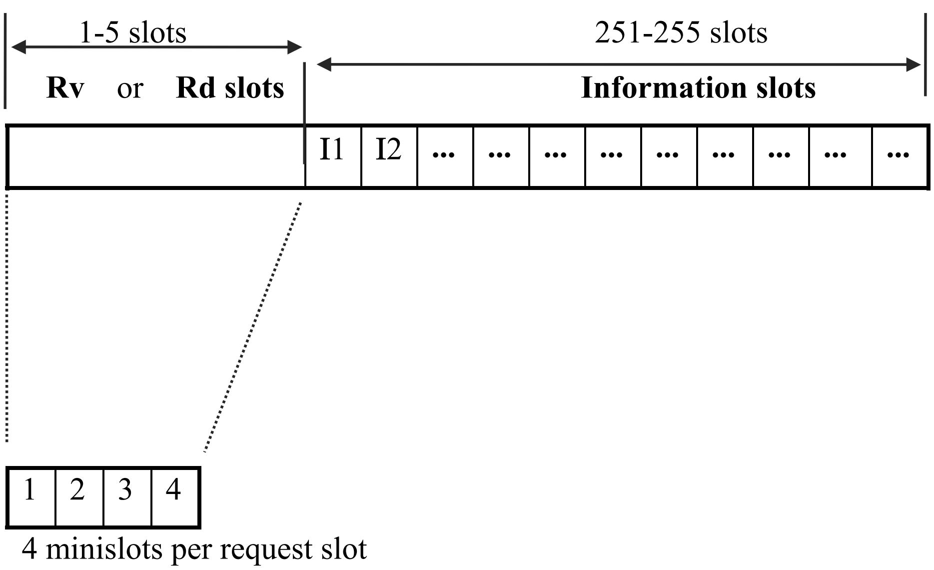


Figure 2.4. Frame structure for the 9.045 Mbps channel.

Number of Video users	Number of request slots	Probability p*
6	1	0.0072
5	1	0.03
4	2	0.06
3	2	0.085
2	3	0.128
1	4	0.18
0	5	1

Table 2.2. Adjustable request bandwidth depending on the number of video users.

2.4. Voice-Video-Data Integration Results and Discussion

The two parameters of the Pareto distribution are related through the following equation:

$$(\alpha-1)/\alpha = \lambda * k \quad (3),$$

which means that the ratio $(\alpha-1)/\alpha$ is equal to the ratio of the minimum data message interarrival time (k) to the mean data message interarrival time $(1/1/\lambda)$.

We have chosen to investigate the cases when the above ratio is equal to 0.2, 0.3, and 0.4 respectively. Thus, our simulations have been conducted for three different values of the shape parameter α, 1.25, 1.429 and 1.667, respectively. These three values are all within the interval (1,2), which means that the data message interarrival time Pareto distributions considered have finite mean and infinite variance. The closer the value of α is to 1, the burstier the distribution is.

Our results are presented in Tables 2.3-2.6, and Figures 2.5,2.6. Tables 2.3-2.6 present the results for the maximum voice capacity obtained for various values of the data message arrival rate λ, and for various numbers of video users (0,1,3, and 5, respectively), when all the QoS requirements for the three traffic types are satisfied. Figures 2.5 and 2.6 present the decrease (%)

in the maximum voice capacity suffered by the system, when the data message arrivals are characterized by Pareto distributed interarrival times, instead of exponentially distributed interarrival times (Poisson arrivals).

Our results lead us to the following observations:

1) Regarding the Poisson distribution, it is clear from the results in the Tables that the smoothest transitions for the maximum voice capacity, as λ increases, take place when no video users exist in the system. This is explained by the fact that video traffic is quite bursty, and thus, the number of voice terminals has to be drastically decreased as the data message arrival rate increases, in order to cope with the burstiness of the video traffic and still be able to preserve the QoS requirements for each traffic type.

2) The results confirm that, for the smaller values of the Pareto shape parameter α, i.e. for the "burstier" Pareto distribution, the decrease in the maximum voice capacity is the greatest compared to the results for the Poisson distribution. Still, we observe from our results in Tables 2.3,2.5, and 2.6 that as λ increases, and especially for $\lambda>2$, the differences in the maximum voice capacity achieved for different α values are smoother than the ones for smaller λ values. This means that the burstiness given by the smaller α values is overcome by the fact that, for $\lambda>2$, the data load becomes "heavy" for the system to accommodate. The only exception seems, by the results presented in Table 2.4, to be the case where 1 video user is admitted in the system. In this case, we observe that, as λ increases, the differences in the maximum voice capacity achieved by the system for the different α values increase as well. This can be explained by the fact that 1 video source alone behaves in a quite more bursty way than the superposition of 3 or 5 simultaneously active video sources does, and thus, the

most restraining parameter in the case of 1 video user and Pareto data message interarrival times is the video dropping probability.

3) From Figures 2.5 and 2.6 (and also from Tables 2.4 and 2.6), we observe, when 5 video users are admitted in the system, that the decrease in the maximum voice capacity, when using the Pareto data message interarrival distribution instead of the Poisson message arrival distribution, is very large compared to the decrease when 1 video user is admitted in the system. This is a general conclusion, which can be deducted from all four Tables: as the video load increases, the maximum voice capacity achieved with the Pareto data traffic model suffers a greater % decrease, for all the λ values, in comparison to the maximum voice capacity achieved when using the Poisson data traffic model. The very small allowed video dropping probability, as well as the burstiness of the Pareto data traffic model in comparison to the Poisson one, is responsible for this result.

4) As shown in Tables 2.3 and 2.5 (this also stands for the other two cases), our scheme achieves quite high throughput for all the cases investigated in this paper, i.e. not only in the case of the Poisson data traffic model, but also in the case of the very bursty Pareto data traffic model. The good results achieved by our scheme are a consequence of the combination of three factors: a) our voice slots allocation policy, b) our video slots scheduling policy, and c) the use of the unused information slots as extra request slots. These three factors enable the proposed scheme to ″exploit″ all the available information slots within each frame in the best possible way.

λ (mes./frame)	α=1.25		α=1.429		α=1.667		Poisson	
	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.
0.5	521	89.9	529	91.3	531	91.6	557	96.0
1	518	91.0	527	92.5	529	92.9	549	96.3
1.5	510	91.2	514	91.9	524	93.6	539	96.2
2	488	89.1	491	89.6	495	90.3	530	96.2
3	449	85.7	453	86.4	454	86.5	510	96.0
4	432	86.0	434	86.3	435	86.5	494	96.5

Table 2.3. Maximum Voice Capacity and Throughput when 0 video users are admitted in the system.

λ (mes./frame)	α=1.25	α=1.429	α=1.667	Poisson
0.5	409	413	414	433
1	392	396	397	410
1.5	378	381	384	400
2	366	371	373	385
3	336	347	354	362
4	310	322	327	328

Table 2.4. Maximum Voice Capacity when 1 video user is admitted in the system.

λ (mes./ frame)	α=1.25		α=1.429		α=1.667		Poisson	
	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.
0.5	221	75.7	232	77.6	241	79.1	264	82.7
1	201	74.0	214	76.1	222	77.5	244	81.0
1.5	185	72.8	195	74.5	203	75.9	229	80.0
2	170	71.9	178	73.3	188	74.9	215	79.3
3	158	73.1	165	74.2	174	75.7	192	78.7
4	138	72.9	144	73.9	147	74.4	173	78.7

Table 2.5. Maximum Voice Capacity and Throughput when 3 video users are admitted in the system.

λ (mes./frame)	α=1.25	α=1.429	α=1.667	Poisson
0.5	19	29	38	90
1	14	18	22	70
1.5	12	16	20	57
2	9	11	14	44
3	1	2	4	27
4	0	1	3	6

Table 2.6. Maximum Voice Capacity when 5 video users are admitted in the system.

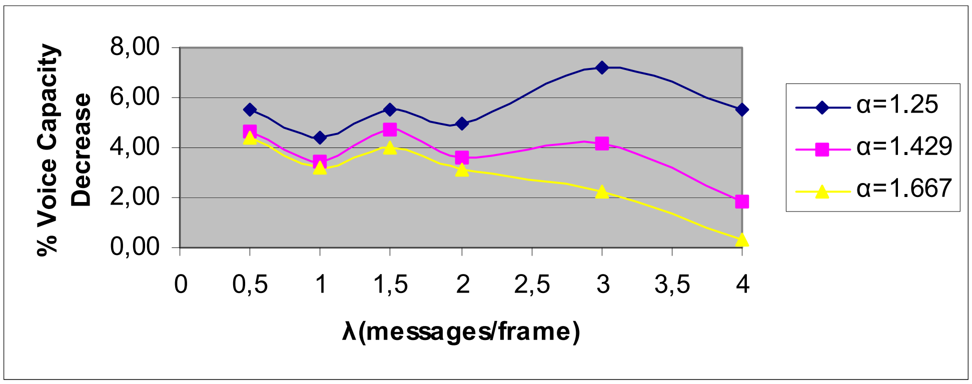


Figure 2.5. Percentage Decrease in the Maximum Voice Capacity when using the Pareto data traffic model, in comparison to the Poisson data traffic model (1 video user is admitted in the system).

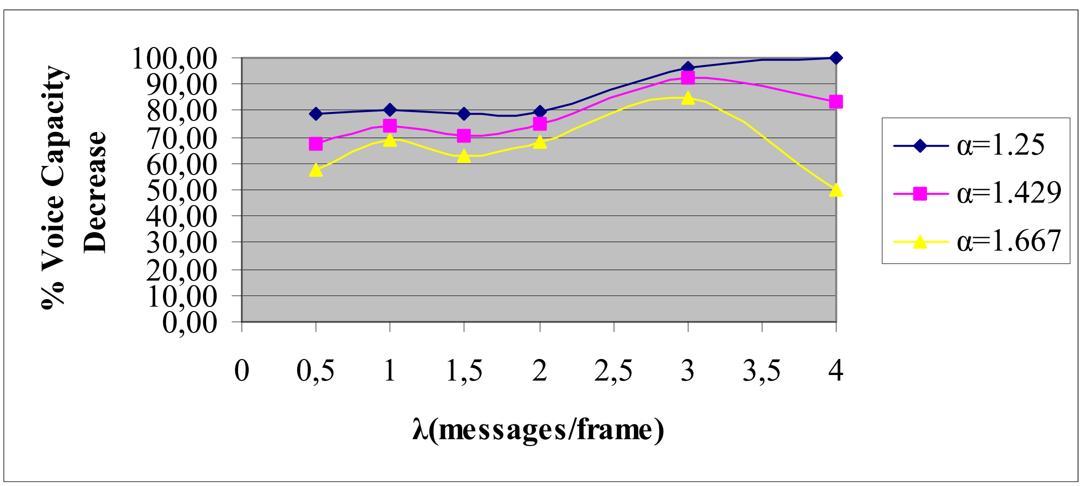


Figure 2.6. Percentage Decrease in the Maximum Voice Capacity when using the Pareto data traffic model, in comparison to the Poisson data traffic model (5 video users are admitted in the system).

2.5. Conclusions

In this chapter we have proposed and evaluated a new TDMA-reservation-based MAC scheme for integrating voice, video and data packet traffic in a high capacity picocellular environment. The scheme uses dynamic slot allocation by the base station, based on the required QoS of each packet traffic stream. Video traffic is offered absolute priority over voice and data traffic, due to its more stringent quality of service requirements.

The proposed scheme achieves high throughput when integrating all three traffic types, despite the very restraining video dropping probability limit.

Finally, the evaluation of the scheme's performance with two very different data message arrival models, shows the very large decrease in voice capacity when the data message interarrival times are Pareto distributed and thus the data arrival process is very bursty, in comparison to the case of Poisson data message arrivals.

Chapter 3

Integrating Voice, Video and Email Data Packet Traffic over Wireless TDMA Channels with Errors

3.1. Introduction

In this chapter, realistic channel conditions and data traffic models are introduced to our scheme. More specifically, we explore via an extensive simulation study the performance of our medium access control (MAC) protocol (which we introduced in the previous chapter) when integrating voice, video and actual e-mail data packet traffic over a wireless channel of high capacity. The two major differences from our study in the previous chapter are *the different data traffic model* used and *the introduction of errors* in the system.

As in the previous chapter, our protocol varies: a) the request bandwidth dedicated to resolving the voice users contention, and b) the probability with which the base station grants information slots to voice users, in order to preserve full priority for video traffic. We evaluate the voice and video packet dropping probabilities for various voice and video load conditions, and the average email data message delays. Our scheme achieves high aggregate channel throughput in all cases of traffic load, despite the fact that all transmissions are subject to error due to noise over the wireless channel.

3.2. Voice-Video-Data Integration

3.2.1 Channel Frame Structure

The channel frame structure is the same as the one presented in section 2.2.1.

3.2.2 Voice, Video and Data Traffic Models

The voice and video traffic models have been presented in sections 2.2.2 and 2.2.3, respectively. We adopt the data traffic model based on statistics collected on email usage from the Finish University and Research Network (FUNET) [3]. The probability distribution function f(x) for the length of the data messages of this model was found to be well approximated by the *Cauchy* (0.8,1) distribution. The packet inter-arrival time distribution for the FUNET model is exponential.

3.2.3. Actions of Voice, Video and Data Terminals, Base Station Scheduling, and Voice-Data Transmission Protocol

With the exception of the base station scheduling mechanism for data users, all the other above elements of our work are identical to the ones described in sections 2.2.5-2.2.7.

We study two cases of incorporating email data users into the system:

In both cases, each data user is allowed to reserve <u>just one slot per frame</u>.

In the first case studied, this slot <u>is guaranteed</u> until the completion of the email message transmission. The "defensive" choice of granting only one slot per frame to data users is easily explained by the fact that video traffic is quite bursty. Thus, the number of voice terminals has to be drastically decreased as the data message arrival rate increases, in order to cope with the burstiness of the video traffic mainly but also with the burstiness of the data traffic, and still be able to preserve the QoS requirements for each traffic type. This explains our ″defensive″ choice, as this would lead to a further decrease of the maximum voice capacity in order to preserve the video QoS requirements.

The second case is the one in which the <u>BS "preempts" data reservations</u> in order to service voice requests. Thus, whenever new voice requests are received and every slot within the frame is reserved, the BS attempts to service the voice requests by canceling the appropriate number of reservations belonging to data terminals (if any). When data reservations are canceled, the BS notifies the affected data terminal and places an appropriate request at the front of the data request queue.

Finally, in order to preserve the strict video QoS, we enforce again the "reshuffling" policy which was described in 2.2.6.

3.3 Channel Error Models

We use a two-state Markov model and an N-state Markov model to emulate the process of packet transmission errors (from [1]). In the two-state Markov model, the channel switches between a "good state" and a "bad state", s0 and s1 respectively: packets are transmitted correctly when the channel is in state s0, and errors occur when the channel is in state s1. The N-state Markov model (presented in Figure 3.1) comprises 6 states in the uplink channel, which is here under study. State $\mathbf{S_o}$ represents the "good state" and all other states represent the "bad states". When the channel is in state $\mathbf{S_o}$, it can either remain in this state, with probability 1-$\mathbf{p_0}$, or make the transition to state $\mathbf{S_1}$, with probability $\mathbf{p_0}$. When the channel is in state $\mathbf{S_n}$, n € [1,4], the transition of the channel state is either to the next higher state (with probability $\mathbf{p}_n$) or back to state $\mathbf{S}_o$ (with probability 1- $\mathbf{p}_n$). This means, that the channel does not remain in the same "bad state" for more than 1 slot. If the channel is in the last state ($\mathbf{S}_5$), it will always return to state $\mathbf{S}_o$. With this model, it is only possible to generate burst errors of at most length N-1. The transition probabilities of the two channel error models are presented in Table 3.1. The only difference between our models and the ones in [1] is that we have chosen a different value for the probability $\mathbf{p}_{good}$, i.e., the steady state probability that the channel is in the good state. In [1], this probability is equal to 0.9328, whereas in our models is larger, and equal to 0.99995. This is necessary, in order to be able to accommodate the type of video traffic we study, the QoS requirements of which are very strict.

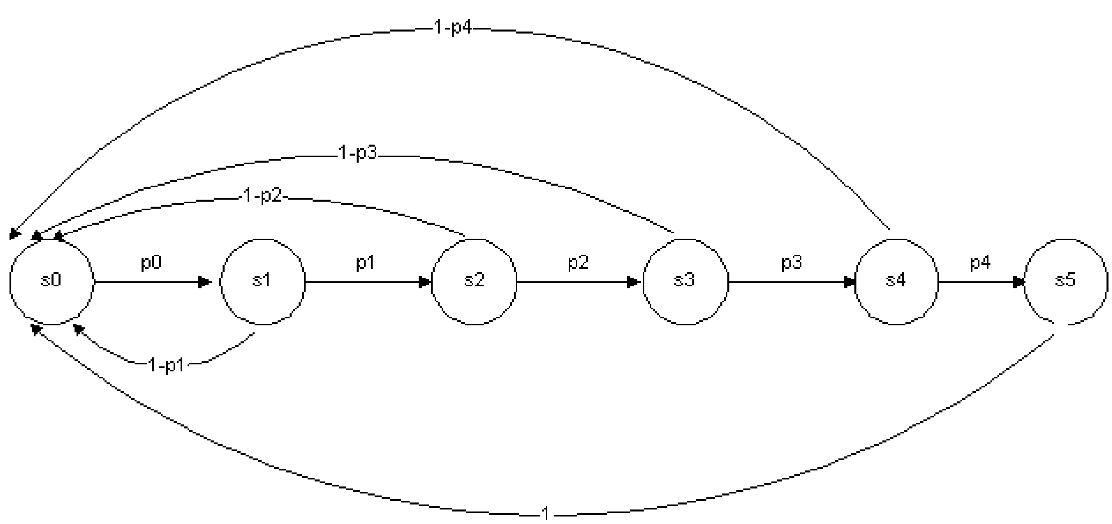


Figure 3.1. N-state Markov model.

3.4. System Parameters

The system parameters are the same as the ones presented in Chapter 2.

3.5. Results and Discussion

As in Chapter 2, each computer simulation point is the result of an average of 10 independent runs (Monte-Carlo simulation), each simulating 305,000 frames (the first 5,000 of which are used as warmup period).

Tables 3.2,3.3,3.4 and 3.5 present the results of our scheme (VVEDI, i.e., Voice, Video, and Email Data Integration) for different numbers of video users within the system. We present the maximum voice capacity and the average email data message delay for different email message arrival rates (λ messages/frame) and for both channel error models examined [58,60].

Our results show that, for the cases of both the error models, the data preemption mechanism helps significantly in the increase of the voice capacity. Also, for all data message arrival rates, we observe that in the presence of the N-state error model our scheme achieves slightly better results than the 2-state error model. This can be explained based on the observation that, although the two error models have the same probability of good-bad and bad-good state transitions, the N-state model is less bursty, due to the fact that this model can only generate error bursts of at most length N-1, i.e., 5 slots in this case.

Figure 3.2 presents the channel throughput (%) achieved by our scheme (number of slots used/frame, divided by the number of information slots in the frame) for the 2-state error model, for different numbers of video users and without preemption. Figure 3.3 presents the throughput achieved by our scheme for the N-state error model, for different numbers of video users, with preemption. Thus, these two figures present the "minimum" and "maximum" set of throughputs achieved by our scheme for all the data message arrival rates under study. As shown in both figures, even for the case of 5 video users, in which the system is heavily loaded with bursty and demanding sources and it has to cope with the transmission errors, the throughput achieved is quite high (steadily above 70% throughput).

The reasons for which VVEDI achieves such good throughput results are:

1) Our proposed video slot allocation mechanism is very dynamic, thus achieving higher bandwidth utilization.
2) The use of the probability **p*** for the allocation of slots to voice terminals enhances the support of the very demanding video traffic in the system.
3) With the proposed above mechanism and the use of ER slots, our scheme "exploits" the maximum amount of slots within the frame.
4) The data preemption policy proves itself to be very effective, as it imposes just a small extra delay on email data messages (of the order of 600-1000 ms, which is totally acceptable for email traffic), while at the same time it helps our system to increase its voice capacity significantly (i.e., around 3% in the case of 1 video user and around 6% in the case of 3 video users).

Table 3.6 includes results which have been extensively presented in [2]. In that work (the scheme called VVDI), we have integrated voice, video (the same models as in this work) and a different data model (from [7]): the aggregate message arrivals are Poisson distributed with mean λ messages per frame. Additionally, we assumed that the messages vary in length according to a geometric distribution with parameter q and mean B=1/q. B is expressed in packets per message, and the steady state data packet arrival rate is equal to λB packets/frame (B is equal to 8 in our study, i.e., an e-mail message in our present study has ten times more packets than a data message in VVDI). Also, an upper limit on the mean data message delay, equal to 200 ms, is assumed in VVDI, whereas in VVEDI no such limit has been set. Other significant differences between VVDI and VVEDI are, as already mentioned before, the absence of channel errors and the preemption policy in VVDI. We present the results in Table 3.6, in order to compare them with the results presented in Tables 3.7 and 3.8 (this is the reason why we have chosen the email data message arrival rate to be 10 times smaller than the data message arrival rate in VVDI, so that the average data packet load will be the same in all cases examined). It is shown that, in the case of no preemption for email data users, this scheme achieves a throughput about 7% less than VVDI with 2 video users and about 4% less than VVDI with 4 video users. In the case of preemption of email data users, this difference is even smaller (about 5% for 2 video users and 3% for 4 video users). Taking into account the facts of the presence of channel errors in VVEDI, and the much larger data message size, which is a "burden" for the channel and can cause the "loss" of information slots (in the sense that they could have been allocated to newly arrived video packets which need them much more urgently), it is once again shown that the performance of VVEDI is highly satisfactory.

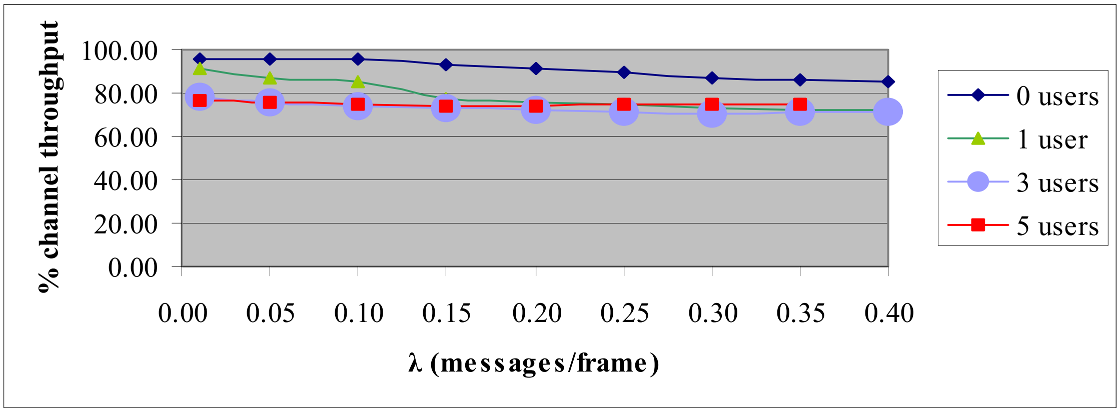


Figure 3.2. 2-state error model, no preemption.

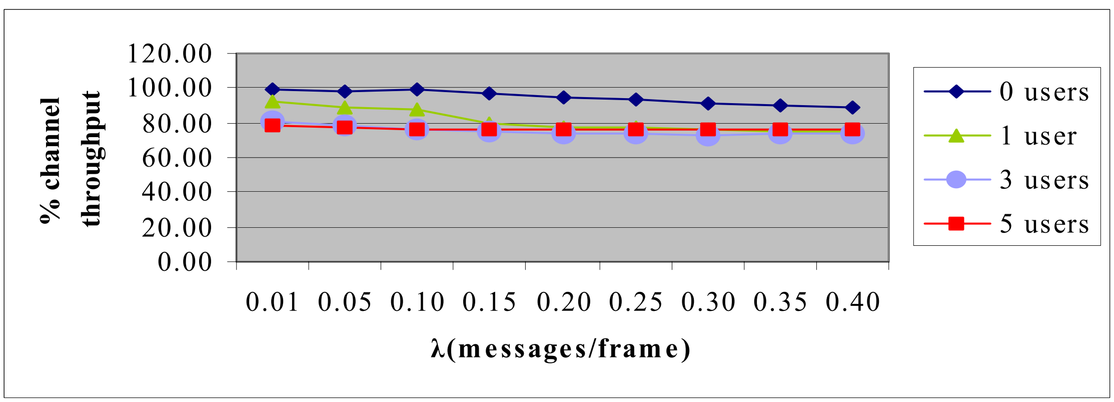


Figure 3.3. N-state error model, with preemption.

	Two-state	N-state
P0		0.0000446
P1		0.100324
P2		0.164083
P3		0.149606
P4		0.526316
P5		0.000000
Pr (Good state)	0.99995	0.99995
Pr (Good-Bad)	0.0000235	0.0000446
Pr (Bad-Good)	0.46945	0.8924

Table 3.1. Channel error models' parameters.

2-state error model					N-state error model			
	No preemption		Preemption		No preemption		Preemption	
λ	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)
0.01	560	1800.11	577	2789.77	560	1834.25	578	2869.28
0.05	552	1385.20	568	2247.47	553	1496.08	569	2326.41
0.1	543	1362.49	558	2013.51	545	1328.40	562	1912.56
0.15	520	1276.43	540	2009.45	522	1316.62	542	1999.47
0.2	498	1288.58	512	1986.35	500	1247.36	515	2017.59
0.25	481	1360.92	498	2156.70	483	1401.23	500	2096.91
0.3	452	1294.40	474	2221.39	453	1276.42	477	2203.17
0.35	439	1270.84	462	2089.96	443	1309.57	465	2109.73
0.4	428	1313.36	442	2127.52	430	1365.86	446	2193.26

Table 3.2 Data Message Delay and Maximum Voice Capacity for 0 video users and set data message arrival rate.

2-state error model					N-state error model			
	No preemption		Preemption		No preemption		Preemption	
λ	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)
0.01	460	1998.02	468	2307.69	461	1897.09	469	2411.13
0.05	430	1356.45	436	1943.35	432	1425.12	437	2000.67
0.1	406	1293.18	418	2007.11	409	1386.55	420	1987.36
0.15	352	1304.26	361	1966.25	356	1247.53	366	1943.25
0.2	331	1226.53	341	1860.96	334	1268.46	345	1966.64
0.25	318	1300.12	329	1932.29	321	1295.89	334	1997.39
0.3	300	1267.68	313	2091.16	304	1326.74	319	1962.81
0.35	286	1271.06	294	2065.38	293	1311.37	298	2102.52
0.4	275	1326.88	285	2001.55	281	1308.19	289	2096.31

Table 3.3 Data Message Delay and Maximum Voice Capacity for 1 video user and set data message arrival rate.

	2-state error model				N-state error model			
	No preemption		Preemption		No preemption		Preemption	
λ	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)
0.01	235	2002.47	247	2436.95	235	1946.36	248	2371.22
0.05	214	1370.18	227	1864.26	215	1488.61	229	2009.64
0.1	194	1402.36	206	1923.90	195	1438.04	208	1978.92
0.15	178	1439.45	189	1912.68	180	1377.93	191	1964.37
0.2	165	1415.36	176	1877.19	167	1364.82	178	1953.56
0.25	151	1329.94	161	1950.82	153	1325.53	164	2049.72
0.3	138	1356.27	148	1994.30	142	1417.71	152	1946.35
0.35	130	1296.54	139	2012.80	134	1386.80	143	1976.39
0.4	121	1423.83	131	1970.45	124	1401.68	135	2075.29

Table 3.4 Data Message Delay and Maximum Voice Capacity for 3 video users and set data message arrival rate.

2-state error model					N-state error model			
	No preemption		Preemption		No preemption		Preemption	
λ	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)	NV	EDmD(ms)
0.01	78	1909.64	88	2425.31	78	1863.75	89	2384.62
0.05	68	1503.21	76	2107.91	69	1587.29	78	2117.83
0.1	52	1388.48	60	2008.30	52	1493.37	62	2186.53
0.15	36	1401.29	44	2124.52	37	1426.01	47	2099.90
0.2	29	1432.30	38	2056.48	30	1508.42	40	2126.85
0.25	22	1464.42	30	1961.07	24	1453.47	32	2091.46
0.3	14	1381.36	21	2023.34	16	1398.29	22	2104.71
0.35	5	1376.28	13	2178.56	7	1441.72	15	2132.96
0.4	x		2	2155.87	x		4	2188.13

Table 3.5 Data Message Delay and Maximum Voice Capacity for 5 video users and set data message arrival rate.

λ (mes. /frame)	**Maximum Voice Capacity and Throughput (%)**			
	2 Video users		4 Video users	
0.1	377	88.6	201	83.2
0.5	350	85.3	185	82.1
1.0	330	83.5	164	80.1
1.5	314	82.4	150	79.4
2.0	300	81.6	136	78.6
2.5	287	81.3	124	78.2
3.0	277	80.9	115	78.2
3.5	265	80.5	104	78.0
4.0	258	80.9	94	77.9

Table 3.6 Maximum Voice Capacity and Channel Throughput for a set number of video users and set Poisson data message arrival rate.

λ (mes. /frame)	**Maximum Voice Capacity and Throughput (%)**							
	2 Video users				4 Video users			
	No preemption		Preemption		No preemption		Preemption	
0.01	322	80.3	336	82.7	165	78.7	176	80.6
0.05	299	77.6	314	80.2	140	75.7	150	77.4
0.1	278	75.7	294	78.4	128	75.3	137	76.8
0.15	264	74.9	281	77.8	113	74.3	121	75.7
0.2	254	74.8	270	77.5	104	74.4	112	75.8
0.25	239	73.8	253	76.2	98	75.0	106	76.3
0.3	227	73.4	241	75.8	88	74.9	97	76.4
0.35	218	73.5	228	75.2	75	74.3	83	75.6
0.4	208	73.4	217	74.9	63	73.8	69	74.8

Table 3.7 Maximum Voice Capacity and Channel Throughput for a set number of video users and set e-mail data message arrival rate, under the 2-state error model.

λ (mes. /frame)	Maximum Voice Capacity and Throughput (%)							
	2				4			
	No preemption		Preemption		No preemption		Preemption	
0.01	323	80.4	338	83.0	166	78.9	177	80.7
0.05	302	78.2	315	80.4	142	76.1	152	77.7
0.1	280	76.0	297	78.9	130	75.6	139	77.2
0.15	268	75.6	284	78.3	115	74.7	125	76.4
0.2	258	75.5	272	77.8	107	74.9	115	76.3
0.25	244	74.7	257	76.9	100	75.3	108	76.7
0.3	230	73.9	245	76.5	92	75.6	99	76.7
0.35	221	74.0	233	76.0	78	74.8	86	76.1
0.4	212	74.0	222	75.7	66	74.3	72	75.4

Table 3.8 Maximum Voice Capacity and Channel Throughput for a set number of video users and set e-mail data message arrival rate, under the N-state error model.

3.6. Conclusions

In this chapter we have evaluated the performance of our channel access control scheme for integrating voice, video and email data packet traffic in a high capacity picocellular environment with errors. Video traffic is offered absolute priority over voice and data traffic, due to its more stringent quality of service requirements, but in this study voice traffic can "preempt" data traffic, in order to preserve its priority over data traffic.

Via an extensive simulation study we demonstrate that the proposed scheme achieves high throughput when integrating all three traffic types, despite the very restraining video dropping probability limit, the burstiness of the data model used and the presence of errors in the packet transmission.

The results achieved by our scheme are a consequence of the combination of four factors: a)our voice slots allocation policy, b)our video slots scheduling policy, c) the preemption of data packets and d) the use of the unused information slots as extra request slots.

Chapter 4

Integrating Videoconference Traffic from MPEG-4 and H.263 Video Coders with Voice and E-mail Data Packet Traffic over Wireless Networks

4.1. Introduction

In this chapter, we extend our work to include actual videoconferencing data, along with the realistic e-mail data model fitted to actual email data which was incorporated in the system, in the work presented in the previous chapter. Our multiple access protocol multiplexes voice traffic at the vocal activity (talkspurt) level to efficiently integrate voice (Constant Bit Rate, CBR On/Off Traffic), MPEG-4 and H.263 video streams (Variable Bit Rate, VBR) and bursty e-mail data traffic in a high capacity picocellular system. The performance of the protocol integrating voice, MPEG-4 or H.263 video and e-mail data packet traffic over a wireless channel of high capacity is explored via simulation. Our scheme achieves high aggregate channel throughput in all cases of traffic load, while preserving the Quality of Service (QoS) requirements of each traffic type.

This is one of the first works, to the best of our knowledge, that considers the integration of MPEG-4 and H.263 streams with other types of packet traffic over wireless TDMA networks.

4.2. Multiple Traffic Type Integration

4.2.1 Channel Frame Structure, Voice and Data Traffic Models

The channel frame structure is the same as the one presented in section 2.2.1, the voice traffic model is presented in section 2.2.2 and the data traffic model in section 3.2.2.

4.2.2 MPEG-4 and H.263 Video Streams

The MPEG group initiated the new MPEG-4 standards in 1993 with the goal of developing algorithms and tools for high efficiency coding and representation of audio and video data to meet the challenges of video conferencing applications. The standards were initially restricted to low bitrate applications but were subsequently expanded to include a wider range of multimedia applications and bitrates. The most important addition to the standards was the ability to represent a scene as a set of audiovisual objects. The MPEG-4 standards differ from the MPEG-1 and MPEG-2 standards in that they are not optimized for a particular application but integrate the encoding, multiplexing, and presentation tools required to support a wide range of multimedia information and applications. In addition to providing efficient audio and video encoding, the MPEG-4 standards include such features as the ability to represent audio, video, images, graphics, text, etc. as separate objects, and the ability to multiplex and synchronize these objects to form scenes. Support is also included for error resilience over wireless links, the coding of arbitrary shaped video objects, and content-based interactivity such as the ability to randomly access and manipulate objects in a video scene [23].

In our study [56], we use the trace statistics of actual MPEG-4 streams from [24]. The video streams have been extracted and analyzed from a camera showing the events happening within an office. We have used the high quality version of the movie, which has a mean bit rate of 400 Kbps, a peak rate of 2 Mbps, and a standard deviation of 434 Kbps. New video frames arrive every 40 msecs. We have set the maximum transmission delay for video packets to 40 msecs, with packets being dropped when this deadline is reached. That is, all video packets of a VF must be delivered before the next VF arrives. The allowed video packet dropping probability is set to 0.0001 [19].

H.263 is a video standard that can be used for compressing the moving picture component of audio-visual services at low bit rates. It adopts the idea of PB frame, i.e., two pictures being coded as a unit. Thus a PB-frame consists of one P-picture which is predicted from the previous decoded P-picture and one B-picture which is predicted from both the previous decoded P-picture and the P-picture currently being decoded. The name B-picture was chosen because parts of B-pictures may be bidirectionally predicted from the past and future pictures. With this coding option, the picture rate can be increased considerably without increasing the bit rate much [24].

In our study, we use the trace statistics of actual H.263 streams from [24], and in particular the streams of the same movies that we study with MPEG-4 encoding. We have used four coding versions of the movie:

a) the unspecified target bit rate (VBR) coding version of the movie, which has a mean bit rate of 91 Kbps, a peak rate of 500 Kbps, and a standard deviation of 32.7 Kbps,
b) the high quality version, which has a mean bit rate of 256 Kbps, a peak rate of 1.4 Mbps and a standard deviation of the bit rate equal to 505 Kpbs.

c) the medium quality version, which has a mean bit rate of 64 Kbps, a peak rate of 320 Kbps and a standard deviation of the bit rate equal to 127 Kpbs.

d) The low quality version, which has a mean bit rate of 16 Kbps, a peak rate of 84 Kbps and a standard deviation of the bit rate equal to 46 Kbps.

The maximum transmission delay for the video packets of a video frame is equal to the time before the arrival of the next video frame[5], with packets being dropped when the deadline is reached. The allowed video packet dropping probability is set to 0.0001 [19,25].

4.2.3. Actions of Voice, Video and Data Terminals, Base Station Scheduling, and Voice-Data Transmission Protocol

With the exception of the base station scheduling mechanism for data users, all the other above elements of our work are identical to the ones described in sections 2.2.5-2.2.7.

In order to efficiently incorporate email data users into the system, the BS "preempts" data reservations in order to service voice requests (with the same mechanism used in Chapter 3). Thus, whenever new voice requests are received and every slot within the frame is reserved, the BS attempts to service the voice requests by canceling the appropriate number of reservations belonging to data terminals (if any). When data reservations are canceled, the BS notifies the affected data terminal and places an appropriate request at the front of the data request queue.

For video users, we use again our "reshuffling" policy (section 2.2.6).

Finally, we adopt again the *two-cell stack* reservation random access algorithm for voice users and the *two-cell stack* blocked access collision resolution algorithm for data users (section 2.2.7).

---

5 The interframe period in the H.263-encoded movies is not constant, as in MPEG-4 encoding. It is an integer multiple of 40 ms.

4.3. System Parameters

The channel rate is 9.045 Mbps (from [19]). The 12 ms of frame duration accommodate 256 slots. The number of request slots shared by voice and data users is fixed, and equal to 2 (Figure 4.1). It has been found, through simulation, that this number of request slots suffices for voice and data users in all the examined cases of video load (which include significantly high loads). We should also note that:

1. In our design, we chose the number of minislots per request interval (4), to allow for guard time and synchronization overheads, for the transmission of a generic request packet, and for the propagation delay within the picocell.
2. Because of assumption 2 of our voice traffic model, all voice request intervals are located at the beginning of each frame.
3. The average email data message length has been found (by simulation) to be 80 packets. We do not impose an upper limit on the average email data message delay, as this is a type of traffic that can withstand a delay of a number of seconds or even more. Thus, we simply evaluate the average email data message delay in our study.
4. The value of the *probability* ***p**** is chosen equal to 0.1 (10%). Other values of ***p**** have also been tried out through simulation, and it has been found that the chosen value gives very satisfactory results for all the examined cases of video load. Due to the great complexity of the scheme, the optimal ***p**** is extremely difficult to be obtained analytically.

4.4. Results and Discussion

Table 4.1 presents the simulation results of our scheme, when integrating voice and MPEG-4 video streams and satisfying the QoS requirements of both traffic types. As expected, the very bursty nature of the video streams leads to a significant decrease of the maximum voice capacity (and, consequently, of the channel throughput) as the number of video users in the system increases. It should be pointed out that the reason for which the channel throughput is smaller in the case of 1 video user than in the case of 2 video users in the system, is explained by the fact that 1 video stream is burstier than the superposition of multiple video streams. As the number of video terminals increases their traffic superposition is smoother, however the load increase is great and leads to the decrease of the channel throughput, because of the significantly smaller number of voice users who can be multiplexed.

Table 4.2 presents the simulation results of our scheme, when integrating all three traffic types: voice, MPEG-4 video streams, and email data. We present the voice capacity for different email message arrival rates (λ messages/frame) and the corresponding channel throughput. We examine the cases of λ being equal to 0.1, 0.3, and 0.5 messages/frame, (i.e., 256 Kbps, 768 Kbps and 1.28 Mbps respectively), and we observe from our results that, for a given number of video terminals, as λ increases, the channel throughput increases as well. This proves the efficiency of our data preemption mechanism, which allows the incorporation of larger data message arrival rates into the system without significant reduction of the voice capacity or violating the strict QoS requirements of video and voice traffic[6]. The reason for the reduction of the voice capacity despite the data preemption mechanism in favor of voice, is the fact that data

[6] The “x” symbol in the last row of the table means that this load can not be supported by the system.

users are not preempted in favor of video users as well, and thus less voice users can enter the system in order to preserve the strict QoS requirements of video traffic.

The results in both Tables 4.1 and 4.2 show that our scheme achieves a good channel throughput (over 60 and occasionally over 70 percent) when the number of video users is between 1-5, i.e. for low and medium video load. When the number of video users becomes larger, the throughput decreases below 60 percent, due to the very bursty nature of video traffic and its very stringent QoS requirements. If, however, the QoS requirements for video traffic were more “tolerant” (e.g., if we considered the acceptable upper limit of the video dropping probability to be 0.001 instead of 0.0001, which can be reasonable, given the nature of the movie (videoconferencing)), then the throughput of our scheme would increase considerably (by about 8-10%, as observed from our simulations).

Tables 4.3-4.6 present the simulation results of our scheme, when integrating voice and H.263 video streams and satisfying the QoS requirements of both traffic types. Again, as in the case of the MPEG-4 video streams, the bursty nature of the video streams leads to a significant decrease of the maximum voice capacity (and, consequently, of the channel throughput) as the number of video users in the system increases. Still, in this case, as we can see from the results in all four Tables, and as expected from the different traffic parameters of the H.263 video streams, a significantly larger number of video users can be supported by the system. Of course, this huge “gain” when comparing the number of H.263 movies that the system is able to accommodate with the respective number of MPEG-4 movies, is obtained at the cost of video image quality (which is quite better when using MPEG-4 encoding). By observing the last rows of each of the four Tables, we can see that the channel throughput, for the cases of 16 Kbps, 64 Kbps and unspecified mean bit rate movies, is quite high and between 61 and 85 percent. Still, there is a very large difference between the channel throughputs achieved for 16 Kbps and 64 Kbps target mean bit rate video coding, in comparison to the throughputs achieved for 256 Kbps target mean bit rate video coding. In the latter case, which is the closest (in the values of its mean, peak and deviation parameters) to the MPEG-4 movie encoding, the throughput is very low, especially in the case where more than three 256 Kbps-mean bit rate video users access the system. This result can be explained by both the facts: a) that the mean target bit rate of the users in this case is much higher, and b) that the peak-to-mean bit rate ratio is again (as in the cases of 16 Kbps and 64 Kbps coding) quite high, causing the peak to be 1.4 Mbps. Therefore, it is the burstiness of the video coding that hinders the system from reaching higher channel throughputs. However, the results for 256-Kbps video coding with less than 3 video users present in the system show that even for this video coding our scheme achieves quite high (and in certain cases, very high) throughputs while satisfying the very strict Quality of Service (QoS) requirements of each traffic type.

The results of Tables 4.3 and 4.6 are also presented in Figures 4.2 and 4.3, respectively, in order to visually present the cases of the lowest and the highest range in the throughput achieved by our scheme when the number of video users is increased or decreased in the system.

Tables 4.7, 4.8 and 4.9 present the simulation results of the scheme, when integrating all three traffic types: voice, H-263 video streams, and email data.

We present the voice capacity for different email message arrival rates (λ messages/frame) and the corresponding channel throughput. We examine the cases of λ being equal to 0.01, 0.05, 0.1 and 0.3 messages/frame, (i.e., 25.6 Kbps, 128 Kbps, 256 Kbps and 768 Kbps, respectively), and we observe from our results that, for a given number of video terminals, as λ increases, the channel throughput remains more or less stable. This proves again the efficiency of our data preemption mechanism, which allows the incorporation of larger data message arrival rates into the system without significant reduction of the voice capacity or violating the strict QoS requirements of video and voice traffic[7].

The results presented in all the Tables show that the examined scheme achieves good channel throughput (steadily over 60 and occasionally close to 80 percent) for all cases except, again, for the case of servicing more than 2 video users of 256-Kbps target mean bit rate coding. If, however, the QoS requirements for video traffic were more "tolerant" (e.g., if we considered the acceptable upper limit of the video dropping probability to be 0.001 instead of 0.0001, which would be reasonable, given the nature of the movie (videoconferencing)), then the throughput of our scheme would be considerably larger (by about 8-10%, as observed from our simulations).

The average email data message delay is, in all the examined cases, lower than 2 seconds, which is considered completely acceptable for email traffic. It does not increase linearly with the increase of the data message arrival rate, as would be expected for a given number of video terminals. This happens mainly because of the burstiness and the exceeding variance of the Cauchy distribution, which produces the length of the data messages of our model. Another reason for this is the fact the voice capacity decreases with the increase of the data message arrival rate.

A final comment that needs to be made is that we do not further "penalize" the data users, by using a second preemption mechanism which would also cancel data reservations in favor of video reservations, because: a) video users are already given the highest priority, and b) even in

[7] The "x" symbol in the last row of the table means that this load can not be supported by the system.

the case of abnormally high data message arrival rates (e.g., for λ=2, which corresponds to 160 data packets/frame, i.e. 6.25 Mbps of data), data users with newly generated messages and without slot reservations would acquire on average only 2 slots per frame (i.e.,1 slot/arriving message).

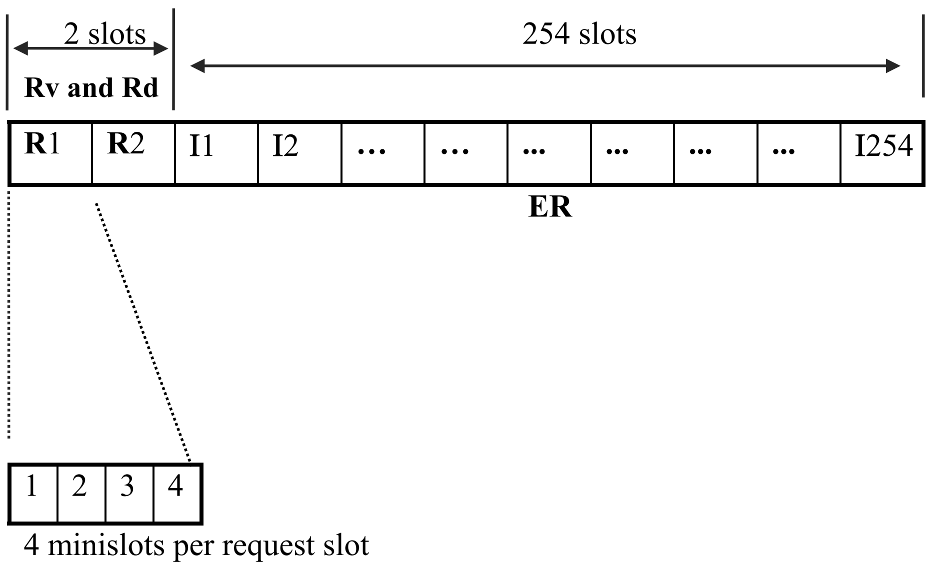


Figure 4.1. Frame structure for the 9.045 Mbps channel.

Number of Video Users	Voice Capacity (Maximum Number of Voice Terminals)	Channel Throughput (%)
9	0	43.9
8	77	51.9
7	124	54.8
6	190	60.9
5	238	64.0
4	279	65.9
3	336	70.5
2	378	72.6
1	402	71.7

Table 4.1. Voice Capacity and Channel Throughput for various MPEG-4 video loads.

Number of video users	λ (messages/frame)	Voice Capacity (Maximum Number of Voice Terminals)	Channel Throughput (%)
1	0.1	319	61.0
	0.3	288	62.1
	0.5	252	62.4
3	0.1	279	64.1
	0.3	248	65.2
	0.5	212	65.5
5	0.1	187	58.6
	0.3	155	59.6
	0.5	123	60.5
7	0.1	86	51.3
	0.3	47	51.4
	0.5	x	x

Table 4.2. Voice Capacity and Channel Throughput for various MPEG-4 video and data loads.

Video Users	59	50	45	40	35	30	25	20	15	10	5
Voice Capacity	0	60	96	130	164	198	234	269	307	347	386
Through-put (%)	65.51	65.63	66.03	66.21	66.32	66.50	67.04	67.31	68.14	69.23	70.21

Table 4.3. Voice Capacity and Channel Throughput for unspecified target mean bit-rate H.263 video.

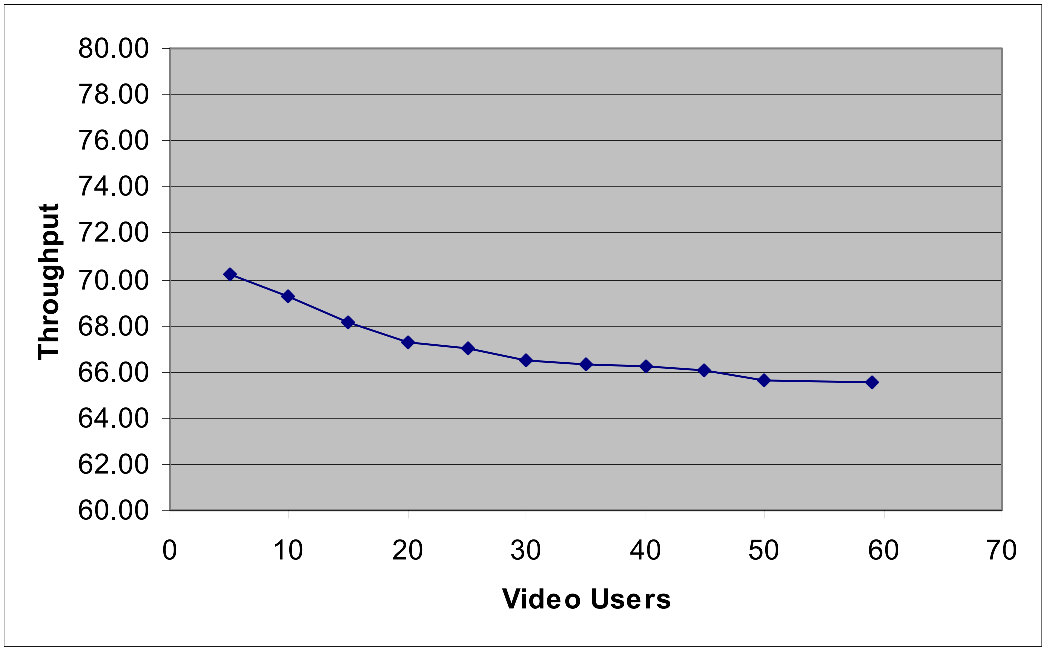


Figure 4.2. Throughput vs. number of video users for unspecified target mean bit-rate H.263 video.

Video Users	329	320	300	250	200	150	100	50	30	10
Voice Capacity	1	11	36	102	173	235	303	395	434	496
Through-put (%)	67.95	67.65	67.70	68.47	70.03	70.03	70.96	76.64	78.75	83.41

Table 4.4. Voice Capacity and Channel Throughput for 16-Kbps target mean bit-rate H.263 video.

Video Users	76	70	60	50	40	30	20	10	5	3
Voice Capacity	4	31	82	143	215	280	349	422	472	498
Through-put (%)	61.16	60.98	61.62	63.55	67.44	70.21	74.02	78.48	82.36	85.30

Table 4.5. Voice Capacity and Channel Throughput for 64-Kbps target mean bit-rate H.263 video.

Video Users	10	9	8	7	6	5	4	3	2	1
Voice Capacity	9	39	69	104	143	196	233	288	360	444
Through-put (%)	33.06	35.05	37.01	39.89	42.91	48.75	51.87	57.64	66.67	76.32

Table 4.6. Voice Capacity and Channel Throughput for 256-Kbps target mean bit-rate H.263 video.

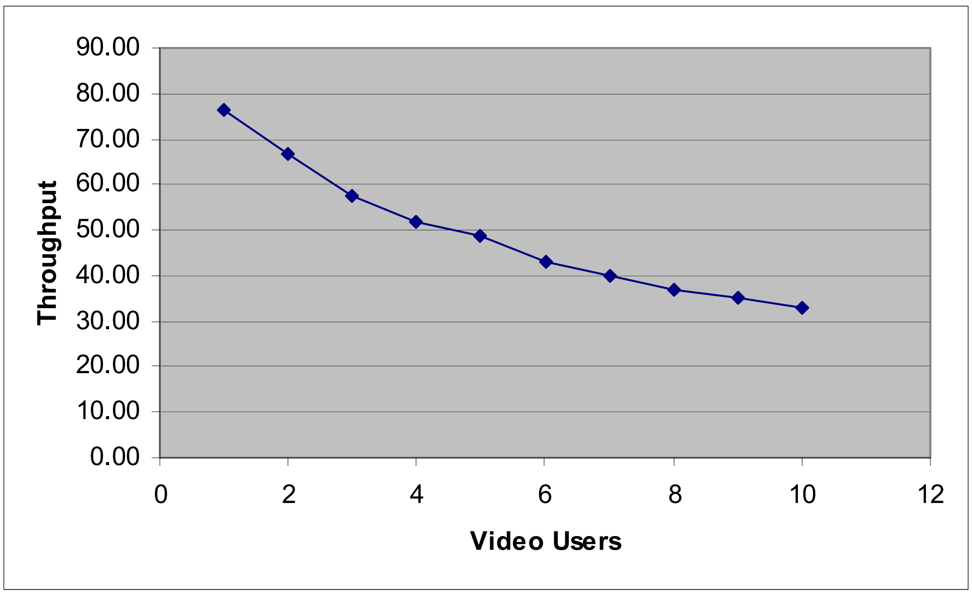


Figure 4.3. Throughput vs. number of video users for 256-Kbps target mean bit-rate H.263 video.

No. video users / Messages/ Frame		0.01	0.05	0.1	0.3
40	Cap.	417	410	404	376
	EMmD (ms)	1050.78	1080.82	804.69	1294.17
	Throug	78.56	78.14	78.05	77.93
80	Cap.	333	327	323	292
	EMmD (ms)	1472.81	1869.37	1013.01	963.86
	Throug. (%)	72.79	72.65	72.58	72.18
160	Cap.	215	209	206	172
	EMmD (ms)	879.18	967.73	940.56	1031.18
	Throug. (%)	69.49	69.40	69.20	69.19
240	Cap.	113	108	104	71
	EMmD (ms)	980.34	1027.42	969.56	1012.04
	Throug. (%)	68.67	68.41	68.35	68.24
329	Cap.	×	×	×	×
	EMmD (ms)	-	-	-	-
	Throug. (%)	-	-	-	-

Table 4.7. Voice-Video-Data Integration Results for 16-Kbps target mean bit rate video encoding.

Messages/ Frame No. video users		0.01	0.05	0.1	0.3
10	Cap.	422	399	386	360
	EMmD	756.04	1312.89	1786.43	1105.62
	Throug.	78.46	78.30	78.18	78.72
20	Cap	349	315	307	276
	EMmD	927.20	1373.47	2101.68	1204.69
	Throug.	74.14	74.03	73.93	73.84
40	Cap	212	177	165	138
	EMmD	1150.50	1621.38	1278.88	1255.33
	Throug.	67.76	66.96	66.57	66.33
60	Cap	81	77	73	41
	EMmD	1063.11	1277.93	1585.66	1267.73
	Throug.	61.66	61.59	61.56	61.57
76	Cap	3	×	×	×
	EMmD	980.25	-	-	-
	Throug.	61.18	-	-	-

Table 4.8. Voice-Video-Data Integration Results for 64-Kbps target mean bit rate video encoding.

Messages/ Frame No. video users		0.01	0.05	0.1	0.3
1	Cap	434	412	402	370
	EMmD	876.41	879.17	1017.85	1134.65
	Throug.	76.35	75.84	75.62	75.28
4	Cap	226	220	217	185
	EMmD	982.87	1034.69	936.05	1031.26
	Throug.	51.32	50.29	50.60	50.73
7	Cap	103	99	96	66
	EMmD	1478.11	664.75	848.12	974.53
	Throug.	39.77	39.46	39.41	39.32
10	Cap	8	4	×	×
	EMmD	1091.08	862,20	-	-
	Throug.	33.08	33,10	-	-

Table 4.9. Voice-Video-Data Integration Results for 256-Kbps target mean bit rate video encoding.

4.5. Conclusions

In this chapter, we have evaluated the performance of our medium access control (MAC) protocol for mobile wireless communications when integrating voice, MPEG-4 or H.263 video and e-mail data packet traffic over a wireless channel of high capacity. Our work presented in this Chapter is one of the first in the literature to study the integration of MPEG-4 and H.263 video streams with other packet traffic types. The proposed scheme is shown to achieve high aggregate channel throughput in all cases of traffic load, while preserving the Quality of Service (QoS) requirements of each traffic type.

Chapter 5

Call Admission Control and Traffic Policing Mechanisms for the Wireless Transmission of Layered Videoconference Traffic from MPEG-4 and H.263 Video Coders

5.1. Introduction

The use of actual video traces in the work presented in Chapter 4, led us to the idea to focus on the case where only video (MPEG-4 and H.263) traffic is present in the system, and develop appropriate Call Admission Control (CAC) and Traffic Policing. These mechanisms are essential in an ATM network, in order to ensure both the preservation of the QoS requirements of all the users and the efficient use of the network's resources.

Therefore, in this chapter we explore the performance of a CAC and Traffic Policing mechanism which we propose for transmitting multiple-layered videoconference movies over a wireless channel of high capacity. We consider both MPEG-4 and H.263 coded movies, and in the latter case our scheme achieves high aggregate channel throughput, while preserving the very strict Quality of Service (QoS) requirements of the video traffic. We consider MPEG-4 movies with two coding layers (high and low bit rate), and H.263 movies with three coding layers (high, medium and low bit rate).

5.2. System Model

In this study we investigate the case where video traffic is the only traffic in the system. Nevertheless, the frame duration is still chosen to be equal to the time a voice terminal needs to generate a new voice packet, as in the previous chapters. Assuming that the speech codec rate is 32 Kbps and that the packet length is equal to the size of an ATM cell, yields the frame duration of 12 ms.

5.2.1 MPEG-4 and H.263 Video Streams

We used the trace statistics of actual MPEG-4 streams from [24]. The video streams have been extracted and analyzed from a camera showing the events happening within an office. We have used two coding versions of the movie:

a) the high quality version, which has a mean bit rate of 400 Kbps, a peak rate of 2 Mbps, and a standard deviation of the bit rate equal to 434 Kbps, and
b) the low quality version, which has a mean bit rate of 90 Kbps, a peak rate of 1 Mbps, and a standard deviation of the bit rate equal to 261 Kbps.

New video frames arrive every 40 msecs. We have set the maximum transmission delay for video packets to 40 msecs, with packets being dropped when this deadline is reached. The allowed video packet dropping probability is set to 0.0001 [19,25].

For H.263 video, we use the trace statistics of actual H.263 streams from [24], and in particular the streams of the same movies that we study with MPEG-4 encoding. We have used three coding versions of the movie:

a) the high quality version, which has a mean bit rate of 256 Kbps, a peak rate of 1.4 Mbps and a standard deviation of the bit rate equal to 505 Kbps.

b) the medium quality version, which has a mean bit rate of 64 Kbps, a peak rate of 320 Kbps and a standard deviation of the bit rate equal to 127 Kbps.

c) the low quality version, which has a mean bit rate of 16 Kbps, a peak rate of 84 Kbps and a standard deviation of the bit rate equal to 46 Kbps.

In this case, the maximum transmission delay for the video packets of a video frame is equal to the time before the arrival of the next video frame[8], with packets being dropped when the deadline is reached. The allowed video packet dropping probability is again set to 0.0001 [19,25].

5.2.2. Actions of Video Terminals and Base Station Scheduling

The actions of the video terminals and the Base Station scheduling are the same as the ones presented in paragraphs 2.2.5 and 2.2.6.

5.3 Statistical Multiplexing

Along an ATM network, the bit streams of different sources are multiplexed for further transmission or switching. The resulting bit stream is the statistical superposition of the stochastic processes describing the individual source streams. This process is referred to as statistical multiplexing.

The ability to exploit the statistical multiplexing gains depends largely on the ability to control the statistical multiplexing process. At any time, the network must offer the specified cell loss figures, while exploiting the statistical multiplexing gains as far as possible. For an ATM network, however, no dynamic control is possible. Therefore, the control of the statistical multiplexing must be performed at call setup, where the originating application

[8] The interframe period in the H.263-encoded movies is not constant, as in MPEG-4 encoding. It is an integer multiple of 40 ms.

must provide the network with some parameters, describing the expected statistical behavior of the corresponding source (e.g., video source). Within the network, a path must then be selected whereupon this bit stream can be served. This selection must be made so that the cell loss performance of the previously admitted bit streams of other sources on this path is not degraded below specifications. Therefore, the network must keep track of the statistical load of all internal links. It is further necessary to monitor the admitted sources during their active periods, in order to protect the network from a single source exceeding its assigned quotas.

Possible statistical parameters provided by the originating source include the average bit rate, the bit rate variance and the peak bit rate. If a sufficient number of independent sources are multiplexed, the resulting distribution of the aggregate bit rate will approach a normal distribution with an average equal to the sum of the individual averages, and a variance equal to the sum of the individual variances. From these, the maximal allowable load on a link can be calculated for a given cell loss performance by reference to the cumulative normalized normal distribution [27].

5.4 Call Admission Control and Traffic Policing Mechanisms

In our study, we investigated an approach on the subject of Call Admission Control and Traffic Policing for both cases of movies (MPEG-4, H.263), which was a combination of two mechanisms:

1) For each new request from a video terminal to enter the system, the *equivalent bandwidth* of the superposition of the MPEG-4 (or H.263) movies was calculated. The calculation was based on the approximate formula presented in [27]. If the expected equivalent bandwidth, with the addition of the new video stream, was less than the channel rate (9.045 Mbps), then the new video terminal was allowed to enter the system. If, on the

other hand, the expected equivalent bandwidth surpassed the channel rate, then the new video terminal was asked to "decrease" its demands (move to a lower quality coding, if possible). If the problem persisted, then another video terminal of high quality coding was asked to "decrease" its demands. If, finally, all high quality coding video terminals which had entered the system decreased their demands and the equivalent bandwidth of the superposition still surpassed the channel rate, then the new request was rejected.

2) To avoid a situation where a video terminal would "pass" the CAC mechanism but would violate its declared mean and peak parameters and cause excessive video packet dropping for all admitted video terminals, a second mechanism was implemented in the system. Each minute, the system checked the video packet dropping, and if it exceeded the upper limit of 0.0001 in two consecutive checks (i.e., for two consecutive minutes), then the previous "demand decreasing" policy was implemented once more, starting from the movie which last entered the system. This policy will be referred to, as the "Traffic Policing mechanism".

5.5. Results and Discussion

In our experiments, we assume that each video request is with equal probability (50%) of high or low quality (MPEG-4 coding), and with equal probability (33.3%) of high, medium or low quality (H.263 coding).

Our simulation results have shown that the method presented in [27], for calculating the equivalent bandwidth of the superposition of the movies, greatly overestimates the actual bandwidth requirements. This "problem" is not restricted to just this method, but it has been encountered in other methods as well (e.g., the authors in [28] comment on their own method as overestimating the actual bandwidth). Although a somewhat conservative CAC mechanism is preferable to underestimating the bandwidth requirements of connections (which could result in network congestion), the use of the CAC policy in our mechanism proves to lead to very significant channel throughput deterioration. Therefore, we have examined [57] the same cases of MPEG-4 and H.263 video loads with the use only of the Traffic Policing mechanism and, as shown in Tables 5.1-5.2 and Figures 5.1-5.4, the improvement in channel throughput was substantial (more than 10% in many cases).

Tables 5.1 and 5.2 present the simulation results of our scheme, when integrating only two types of H.263 video streams: high quality coding streams with medium or low quality coding streams, respectively. Table 5.1 presents the simulation results when the CAC algorithm (at the "entrance" of the system) is used. Table 5.2 presents the results when the CAC algorithm is not in use, and only the Traffic Policing mechanism is implemented.

Similarly, Figures 5.1.1, 5.1.2, 5.2.1 and 5.2.2 present the results of our scheme when integrating all types of H.263 coding movies, with and without the use of the CAC algorithm respectively. In this case, as expected from our policy for permitting a new video terminal to enter the system, the steady state condition for the system consists of medium and low quality movies only, because all high quality movies are asked to decrease their demands for the system to accommodate as many more video users as possible.

Finally, Figures 5.3 and 5.4 present the results of our scheme when integrating the two types of MPEG-4 coding movies, with and without the use of the initial CAC algorithm respectively. It is evident from the Figures that the throughput achieved by our scheme is greater in the cases where high quality movies is the "dominant" type in the system.

A general comment that needs to be made about all our results is that the very bursty nature of the video streams is responsible for the system's inability to reach high channel throughput results.

Considering the results presented in Figures 5.2.1 and 5.2.2, it should be pointed out that the channel throughput is higher when one of two types is the dominant one in the system (i.e., when either the medium quality movies or the low quality movies mainly use the channel bandwidth). On the contrary, when both types of movies contribute more or less equally to the use of the channel bandwidth, the channel throughput decreases. This can be explained by the fact that the high burstinesses of the two types of movies, combined with their different traffic characteristics, lead to the decrease of the channel throughput. The same phenomenon is not encountered in the results shown in Table 5.2, where high quality movies are integrated with medium or low quality movies, because the burstinesses of the movies are balanced by the much higher mean bit rate of the 256-Kbps movies.

256K movies	64K movies	16K movies	Total number of movies	Throughput (%)
6	0	0	6	21.37
5	3	0	8	29.53
	0	19	24	30.86
4	7	0	11	31.28
	0	42	46	32.63
3	12	0	15	30.77
	0	73	76	35.62
2	20	0	22	33.95
	0	112	114	40.09
1	31	0	32	38.21
	0	165	166	46.92
0	57	0	57	47.05
	0	290	290	60.46

Table 5.1. H.263 video, with the use of the CAC algorithm. (case with high quality movies)

256K movies	64K movies	16K movies	Total number of movies	Throughput (%)
10	0	0	10	31.7
9	2	0	11	38.14
	0	15	24	39.69
8	4	0	12	37.7
	0	28	36	39.83
7	7	0	14	37.32
	0	46	53	41.23
6	12	0	18	37.83
	0	68	74	41.83
5	19	0	24	40.21
	0	97	102	44.18
4	28	0	32	43.59
	0	134	138	46.01
3	38	0	41	47.97
	0	182	185	54.92
2	50	0	52	54.5
	0	235	237	62.84
1	61	0	62	58.88
	0	275	276	67.13
0	76	0	76	60.54
	0	329	329	67.82

Table 5.2. H.263 video, without the use of the initial CAC algorithm. (case with high quality movies)

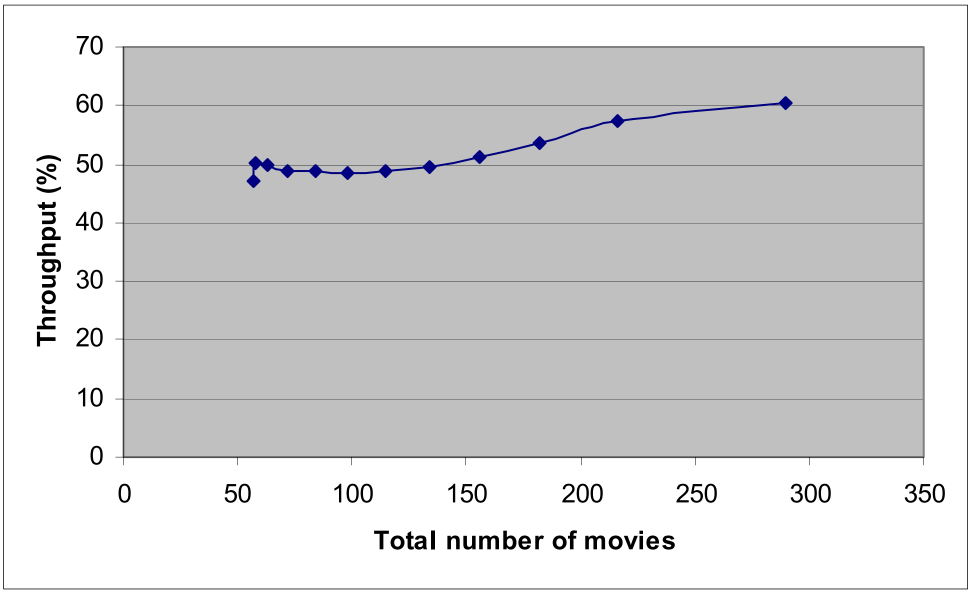


Figure 5.1.1. Throughput vs. total number of movies for H.263 video, with the use of the CAC algorithm.

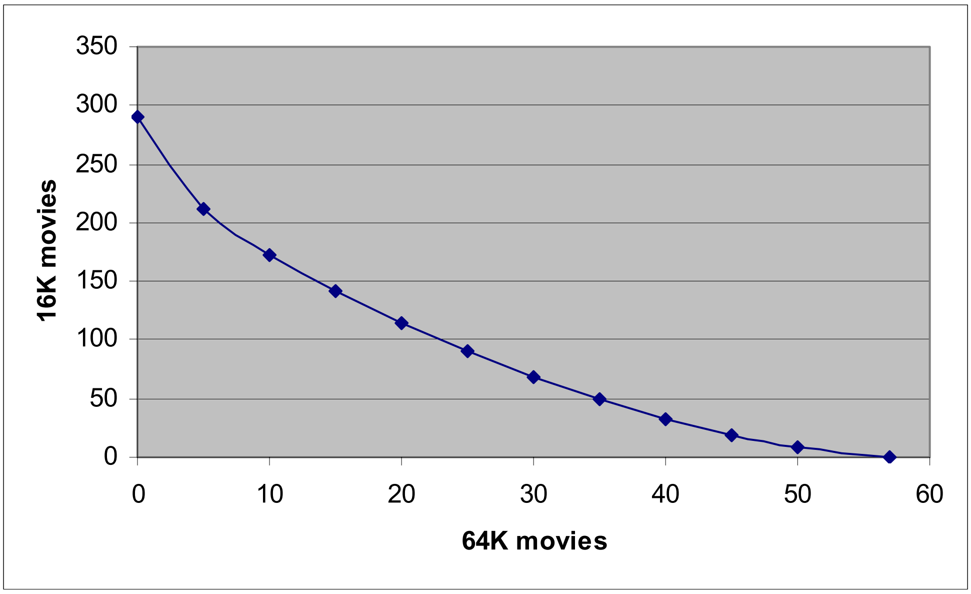


Figure 5.1.2. Relation of 64K and 16K movies for H.263 video, with the use of the CAC algorithm.

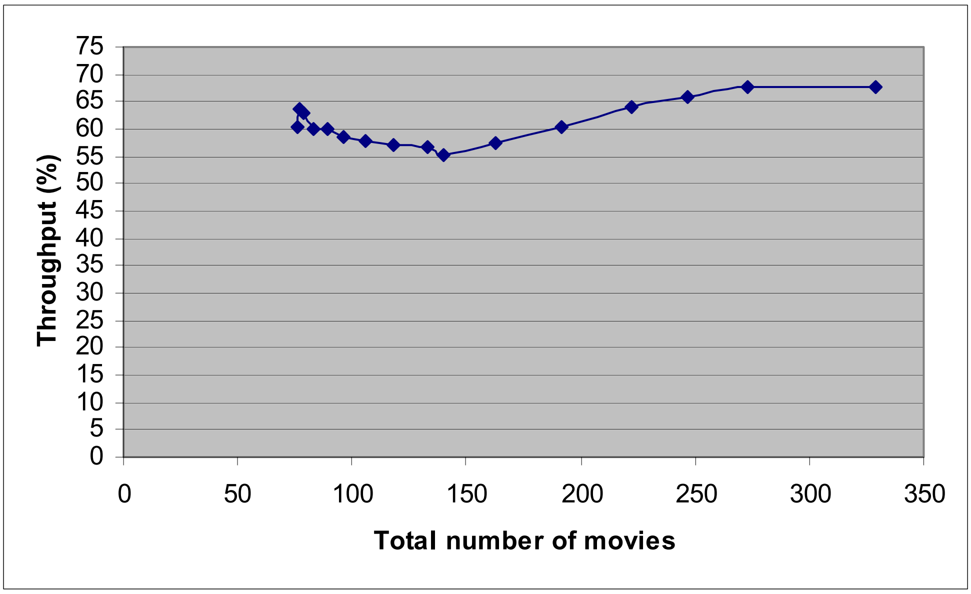


Figure 5.2.1. Throughput vs. total number of movies for H.263 video, without the use of the CAC algorithm.

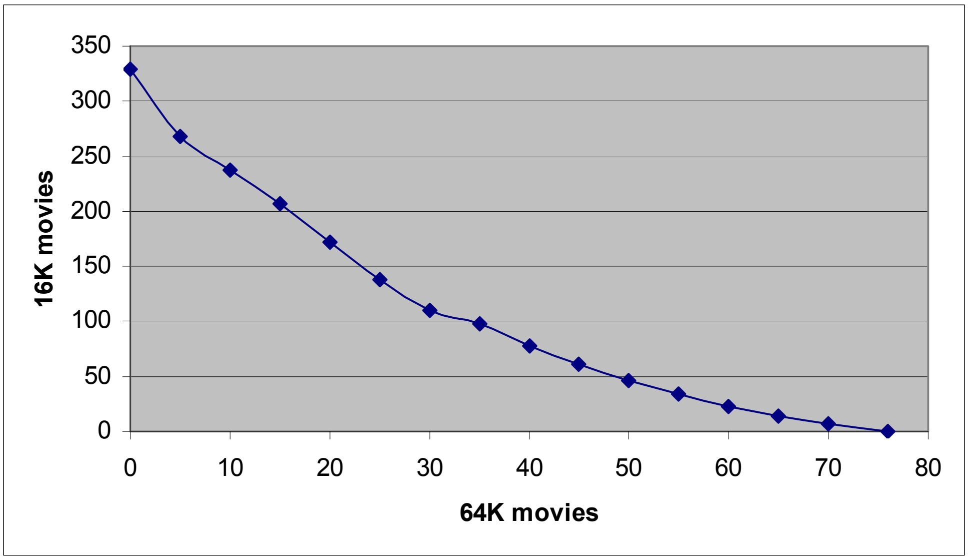


Figure 5.2.2. Relation of 64K and 16K movies for H.263 video, without the use of the CAC algorithm.

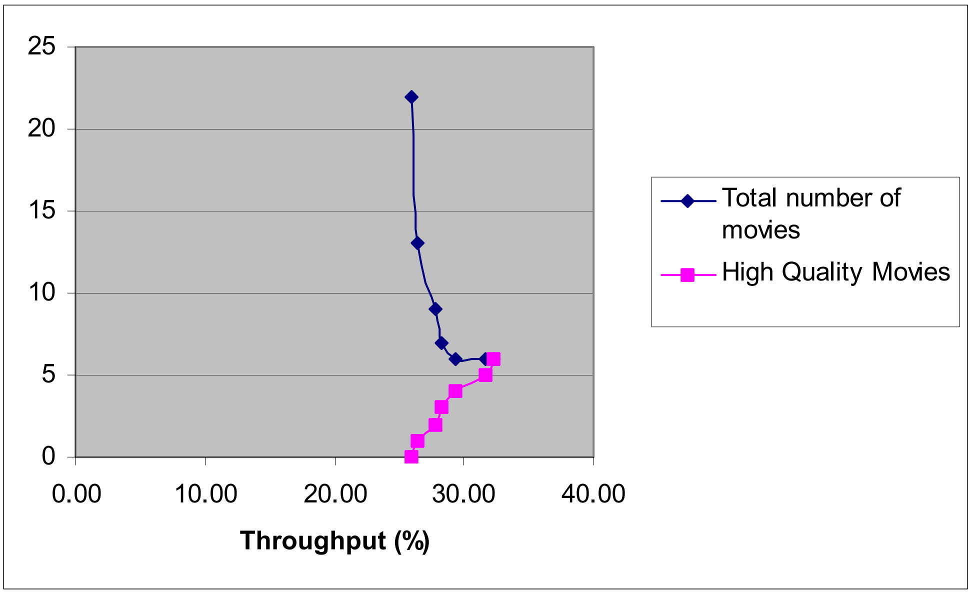


Figure 5.3. Mpeg-4 video, with the use of the CAC algorithm.

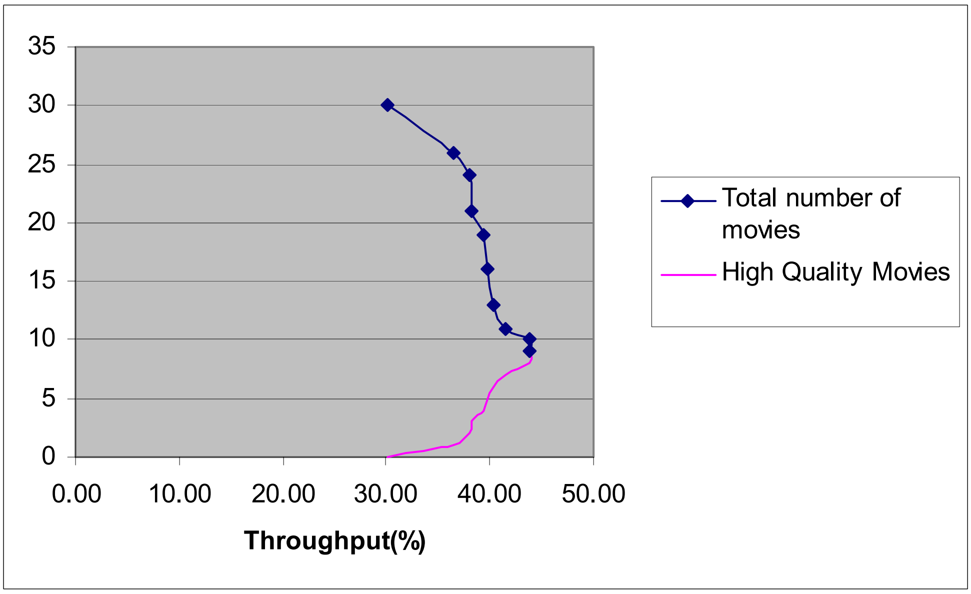


Figure 5.4. Mpeg-4 video, without the use of the CAC algorithm.

5.6. Conclusions

In this chapter, we have focused on the case where *only video* (MPEG-4 and H.263) traffic is present in the system. We have proposed and investigated the performance of a Call Admission Control and Traffic Policing mechanism for transmitting multiple-layered videoconference movies over a wireless channel of high capacity, depending on the user's needs and requests.

Our scheme, which is one of the first in the literature to study the integration of MPEG-4 and H.263 video streams, is shown to achieve high aggregate channel throughput in many cases of traffic load, while preserving the Quality of Service (QoS) requirements of each traffic type. The throughput is not high in all cases examined, because the throughput performance of the protocol depends heavily on the type of movies that demand access to the channel, some of which are very bursty in our study.

Chapter 6

Summary and Contribution of this Work-

Future Work

We have considered the design and performance evaluation of multiple access schemes that multiplex voice traffic at the talkspurt level to efficiently integrate voice, video and data traffic in third generation picocellular wireless networks. We have also developed Call Admission Control and Traffic Policing Algorithms for transmitting multiple-layered videoconference movies over a wireless channel of high capacity. The design goals included maximizing the system capacity and satisfying quality of service requirements such as voice packet dropping probability, video packet dropping probability and voice, video and data packet access delays. These goals were complicated by the limited radio channel bandwidth and by the contradictory nature of voice, video and data traffic. In this chapter, we briefly summarize our results, discuss the contribution of this work to the corresponding research field and present some ideas on how our work can be extended.

6.1 Summary

The first part of the Dissertation, Chapter 2, presents the design and performance evaluation of a new TDMA-reservation-based MAC scheme for integrating voice, video and data packet traffic in a high capacity picocellular environment. The scheme multiplexes voice traffic at the vocal activity (talkspurt) level to efficiently integrate voice (Constant Bit Rate, CBR On/Off Traffic), video (Variable Bit Rate, VBR) and bursty data traffic in high capacity wireless channel picocellular environments. We focused on the uplink (mobiles to base station) channel, where a MAC scheme is required in order to resolve the source terminals contention for channel access. The scheme uses dynamic slot allocation by the base station, based on the required QoS of each packet traffic. Video traffic is offered absolute priority over voice and data traffic, due to its more stringent quality of service requirements. The proposed scheme achieves high throughput when integrating all three traffic types despite the very restraining video dropping probability limit. Also, the evaluation of the scheme's performance with two very different data message arrival models, shows a very large decrease in voice capacity when the data message interarrival times are Pareto distributed and thus the data arrival process is very bursty, in comparison to the case of Poisson data message arrivals.

The second part of the Dissertation, Chapters 3-5, is a presentation of the work done based on our idea that, although our previous work showed remarkably high throughput results, its performance evaluation was based on the simulation of theoretical traffic models and ideal channel conditions, and not of actual traffic streams and channel conditions.

More specifically, in Chapter 3 we introduced realistic channel conditions (errors in the wireless transmission) and an e-mail data traffic model fitted to actual email data traffic

In Chapter 4, we further extended our work to include actual videoconferencing data (MPEG-4 and H.263 encoding).

In Chapter 5, we considered the case where only video users arrive in the system (we used again actual videoconferencing data), but they are subject to a Call Admission Control (CAC) and a Traffic Policing Mechanism, which we developed.

6.2 Contribution of this work

In this work, we have tackled issues that had not yet been addressed in the relevant literature of the field. More specifically, we have examined the potentials of voice, video and data traffic integration over a high capacity wireless channel according to a scheduling mechanism which has many novel aspects. The results derived from our simulations, both for the channel throughput and for the voice, video and data packet delays were very satisfactory, given the high burstiness of both the video and the data traffic.

The expansion of our research to consider the case where only MPEG-4 and H.263 video users arrive in the system, being subject to a Call Admission Control (CAC) and a Traffic Policing Mechanism which we developed, is also a novel contribution to the relevant literature. The results of our study show that, although the combination of our two mechanisms works well and helps significantly improve the channel throughput (in comparison to a scenario where no CAC and Traffic Policing existed), the throughput is not high in all cases examined. The throughput performance of the protocol depends heavily on the type of movies that demand access to the channel. When one type of movies is the dominant one in the system, the throughput is usually high, but when two or three types of movies contribute more or less equally to the use of the channel bandwidth, the channel throughput is lower.

We believe that the channel access and transmission scheduling mechanisms we have proposed, together with our CAC and Traffic Policing Mechanisms contribute to the state-of-the-art and can definitely help in future efforts to examine video and data traffic integration in high capacity wireless environments.

6.3 Future work

Regarding to possible future work, is should be pointed out that all the schemes presented in our work are flexible and can, with mild modifications, be extended to cover:

a) the integration of voice and data traffic with video traffic originating from actual movies (not videoconferencing), which have quite different traffic characteristics and QoS requirements from videoconferencing applications.
b) the investigation of different system assumptions, such as channel capture and significant noise errors in the channel (much more severe than the ones we have investigated in our study, due to the very strict video packet dropping probability).
c) the investigation of the system's performance when using different coding techniques for voice and video sources, and under less strict video QoS requirements.

REFERENCES

1. Chi-Yuan Hsu, Antonio Ortega, "Rate Control for Robust Video Transmission over Burst-Error Wireless Channels", *IEEE Journal on Selected Areas in Communications, Vol. 17, No.5, May 1999.*
2. P. Koutsakis and M. Paterakis, "On Multiple Traffic Type Integration over Wireless TDMA Channels with Adjustable Request Bandwidth", *International Journal of Wireless Information Networks,* Vol. 7, No.2, pp. 55-68, 2000.
3. Q. Pang, A. Bigloo et. al., "Service Scheduling for General Packet Radio Service Classes", *Proceedings of the IEEE Wireless Communications and Networking Conference 1999 (WCNC'99), New Orleans, USA*.
4. B. Mukherjee, " Integrated voice-data communication over high speed fiber optic networks ", *IEEE Computer Magazine*, Vol. 24, No. 2, pp. 49-58, 1991.
5. N. M. Mitrou, Th. D. Orinos and E. N. Protonotarios, " A Reservation Multiple Access Protocol for Microcellular Mobile-Communication Systems ", *IEEE Transactions on Vehicular Technology*, Vol. 39, No. 4, Nov. 1990, pp. 340-351.
6. S. Nanda, D. J. Goodman and U. Timor, " Performance of PRMA: A packet voice protocol for cellular systems ", *IEEE Transactions on Vehicular Technology*, Vol. 40, pp. 584-598, 1991.
7. A. C. Cleary and M. Paterakis, " An Investigation of Reservation Random Access Algorithms for Voice-Data Integration in Microcellular Wireless Environments", *Int. J. Wireless Info. Networks*, Vol. 2, No. 1, Jan. 1995, pp. 1-16.
8. A. C. Cleary and M. Paterakis, " Design and Performance Evaluation of an RRA Scheme for Voice-Data Channel Access in Outdoor Microcellular Wireless Environments ", *Mobile Networks and Applications (MONET) Journal, ACM and Baltzer Science Publishers*, Vol. 2, No. 1, 1997, pp. 31-43.
9. N. M. Mitrou, G. L. Lyberopoulos and A. D. Panagopoulou, " Voice and Data Integration in the Air-Interface of a Microcellular Mobile Communication System ", *IEEE Transactions on Vehicular Technology*, Vol. 42, No. 1, Feb. 1993, pp. 1-13.

10. W. C. Wong and D. J. Goodman, ″ A packet reservation multiple access protocol for integrated speech and data transmission ″, *IEE Proceedings-1*, Vol. 139, Dec. 1992, pp. 607-612.

11. D. Bertsekas and R. Gallager, ″ Data Networks ″, 2nd Ed., Prentice Hall Inc., 1992.

12. A. C. Cleary and M. Paterakis, ″On the Voice-Data Integration in Third Generation Wireless Access Communication Networks ″, *European Transactions on Telecommunications and Related Technologies*, Vol. 5, No. 1, Jan.-Feb. 1994, pp. 11-18.

13. M. Paterakis and P. Papantoni-Kazakos, ″ A Simple Window Random Access Algorithm with Advantageous Properties ″, *IEEE Transactions on Information Theory*, Vol. IT-35, No. 5, September 1989, pp. 1124-1130.

14. C. Li, L.J. Greenstein and R.D. Gitlin, "A Microcell/Macrocell Architecture for Low -and High-Mobility Users", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 6, August 1993, pp.885-891.

15. D. Heyman, A. Tabatabai, T. V. Lakshman, ″ Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks″, *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 2, No. 1, March 1992.

16. H. Heeke, ″ Statistical multiplexing gain for variable bit rate video codecs in ATM networks″, *Proc. of the 4th Int. Packet Video Workshop, Tokyo, Japan, March 1991.*

17. P. Koutsakis and M. Paterakis, ″Near-Optimal Voice-Data Integration over Third Generation Medium and High Capacity Wireless TDMA Channels″, *Proceedings of the IEEE Wireless Communications and Networking Conference 1999 (WCNC '99), New Orleans, USA.*

18. P. Koutsakis and M. Paterakis, ″*Highly Efficient Voice-Data Integration over Medium and High Capacity Wireless TDMA Channels* ″, ACM/Baltzer Wireless Networks Journal, Vol. 7, No.1, pp. 43-54, 2001.

19. D. A. Dyson and Z. J. Haas, "A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM", *ACM/Baltzer MONET Journal, 1999,* Vol. 4, No.2, pp. 87-99, 1999.

20. G. Anastassi, D. Grillo, L. Lenzini and E. Mingozzi, "A Bandwidth Reservation Protocol for Speech/Data Integration in TDMA-Based Advanced Mobile Systems", *Proc. of the 1996 IEEE Infocom Conf., pp.722-729.*

21. Z. Harpantidou and M. Paterakis, "Random Multiple Access of Broadcast Channels with Pareto Distributed Packet Interarrival Times", *IEEE Personal Communications, Vol.5, No.2, pp.48-55, April 1998.*

22. T. Suda and T. Bradley, " Packetized voice/data integrated transmission on a token passing ring local area network", *IEEE Transactions on Communications*, Vol. COM-37, No. 3, pp. 238-244, 1989.

23. http://www.kom.e-technik.tu-darmstadt.de/acmmm99/ep/gringeri/

24. http://www-tkn.ee.tu-berlin.de/research/trace/trace.html.

25. ITU-T Recommendation, H.263, 3/1996

26. P. Koutsakis, S. Psychis and M. Paterakis, “On the Integration of MPEG-4 Video Streams with Voice and E-mail Data Packet Traffic over Wireless Picocellular Networks”, *Proceedings of the 2001 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2001), San Diego, California*.

27. W. Verbiest and L. Pinoo, “A Variable Bit Rate Video Codec for Asynchronous Transfer Mode Networks”, *IEEE Journal on Selected Areas in Communications, Vol. 7, No. 5, June 1989.*

28. R. Guerin, H. Ahmadi, M. Naghsineh, “Equivalent Capacity and its Application to Bandwidth Allocation in High-Speed Networks”, *IEEE Journal on Selected Areas in Communications, Vol. 9, No. 7, September 1991.*

29. H. Armbruster, "The flexibility of ATM: Supporting Future Multimedia and Mobile Communications", *IEEE Communications Magazine*, Vol. 33, No. 2, Feb. 1995, pp. 76-84.

30. V. H. MacDonald, "The Cellular Concept", *Bell Sys. Tech. Journal*, Vol. 58, No. 1, pp. 15-41, January 1979.

31. T. M. Chen and S. S. Liu, "ATM Switching Systems", *Artech House, Inc.*, ISBN 0-89006-682-5, 1995.

32. P. A. Ramsdale, "Personal Communications in the UK-Implementation of PCN using DCS 1800", *International Journal of Wireless Information Networks*, Vol. 1, No. 1, pp. 29-36, January 1994.

33. W. Stallings, "ISDN and Broadband ISDN with Frame Relay and ATM", Prentice Hall, 1995.

34. N. Passas, "Multiple Access Control in Wireless ATM Networks", *Ph.D. Dissertation, Dept. of Informatics, University of Athens*, Dec.1997.

35. K. Sriram, "Methodologies for bandwidth allocation, transmission scheduling, and congestion avoidance in broadband ATM networks ", *Computer Networks and ISDN Systems*, Vol. 26, No. 1, pp. 43-59, 1993.

36. O.T.W. Yu and V.C.M. Leung, "Adaptive resource allocation for prioritized call admission over an ATM-based wireless PCN ", *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 7, pp. 1208-1225, September 1997.

37. B. L. Mark and G. Ramamurthy, "Joint source-channel control for real-time VBR over ATM via dynamic UPC renegotiation", *Proceedings of the IEEE Globecom '96*, Vol. 3, pp. 1726-1731, London, UK, November 1996.

38. J. E. Padgett et al., "Overview of Wireless Personal Communications", *IEEE Communications Magazine*, Vol. 33, No. 1, pp. 28-41, January 1995.

39. F. Bauchot, "MASCARA: A Wireless ATM MAC Protocol", *Proceedings of the Wireless ATM Workshop*, Helsinki, Finland, September 1996.

40. H. Mitts et al., "Lossless handover for wireless ATM", *ACM Mobile Networks and Applications Journal,* Vol. 1, No. 3, pp. 299-312, December 1996.

41. C.-K. Toh, "A hybrid handover protocol for local area wireless ATM networks", *ACM Mobile Networks and Applications Journal*, Vol. 1, No. 3, pp. 313-334, December 1996.

42. A. Kaloxylos et al., "Smart Buffering Technique for Lossless Hard Handover in Wireless ATM networks", *Proceedings of the IEEE International Conference on Communications '96*, pp. 240-244, Dallas, TX, June 1996.

43. M. Rahnema, "Overview of the GSM System and Protocol Architecture", *IEEE Communications Magazine*, pp. 92-100, April 1993.

44. P. A. Ramsdale, "Personal Communications in the UK-Implementation of PCN using DCS 1800", *International Journal of Wireless Information Networks*, Vol. 1, No. 1, pp. 29-36, January 1994.

45. TIA/EIA IS-54, "Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard", Telecommunications Industry Association, April 1992.

46. A. J. Viterbi, ″Spread Spectrum Communications- Myths and Realities″, *IEEE Communications Magazine*, pp. 11-18, May 1979.

47. R. Kohno, R. Meidan and L.B. Milstein, ″Spread Spectrum Access Methods for Wireless Communications″, *IEEE Communications Magazine*, pp. 58-67, January 1995.

48. K. Kinoshita, M. Kuramoto and N. Nakajima, ″Development of a TDMA Digital Cellular System Based on Japanese Standard″, *Proceedings of the 41st IEEE Vehicular Technology Conference (VTC '91)*, pp. 642-645.

49. C. C. Evci, ″RACE-UMTS for Third Generation Wireless Networks″, *Ann. Telecommun.*, Vol. 47, No. 7-8, pp. 267-281, July-August 1992.

50. R. Pandya, ″Emerging Mobile and Personal Communication Systems″, *IEEE Communications Magazine*, pp. 44-51, June 1995.

51. D. D. Falconer et al., ″Time Division Multiple Access Methods for Wireless Personal Communications″, *IEEE Communications Magazine*, Vol. 33, No. 1, pp. 50-57, January 1995.

52. D. C. Cox, ″Wireless Personal Communications: What is it? ″, *IEEE Pers. Commun.*, Vol. 2, No. 2, Apr. 1995, pp. 20-35.

53. R. Rom and M. Sidi, ″Multiple Access Protocols: Performance and Analysis″, *Springer Verlag*, 1990.

54. W. Verbiest, L. Pinoo and B. Voeten, "The Impact of the ATM Concept on Video Coding", *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988.

55. D. Heyman, T. V. Lakshman, A. Tabatabai and H. Heeke, ″ Modeling teleconference traffic from VBR video coders″, *Proceedings of the IEEE International Conference on Communications '94 (ICC'94),* Vol. 3, pp. 1744-1748, 1994.

56. P. Koutsakis, A. Parathyras and M.Paterakis, "Dynamic Bandwidth Reservation Scheduling for the Integrated Wireless Access of H-263 Videoconference Traffic with Voice and E-mail Packet Traffic", *Proceedings of the 2002 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002),* Lisbon, Portugal.

57. P. Koutsakis, M.Paterakis and S. Psychis, "Call Admission Control and Traffic Policing Mechanisms for the Wireless Transmission of Layered Videoconference Traffic from MPEG-4 and H.263 Video Coders", *Proceedings of the 2002 IEEE International*

Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002), Lisbon, Portugal.

58. P. Koutsakis and M.Paterakis, ″Integrating Voice, Video and E-mail Data Packet Traffic over Wireless TDMA Channels with Errors”, *International Journal of Wireless Information Networks*, Vol. 8, No.4, pp. 217-227, 2001.

59. *PCWorld.com*, July 19, 2001.

60. P. Koutsakis, M.Paterakis and S. Psychis, ″Integrating Voice, Video and Bursty Data Packet Traffic Over Burst-Error Wireless TDMA Channels with Adjustable Request Bandwidth”, *Proceedings of the 2001 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2001), San Diego, California*.