Technical University of Crete

Dept. of Electronic and Computer Engineering

Telecommunications Division

Information and Computer Networks Laboratory

On the Design of MAC Protocols and Transmission Scheduling Schemes for the Integration of Video Streams, Voice and Data Traffic over High Speed Packet Switched Wireless Networks

Spyros Psychis

M.Sc. Thesis

Summer 2003

Στην Ειρήνη, τους γονείς
και τα αδέρφια μου

Ευχαριστίες

Η εργασία αυτή θα ήταν αδύνατον να πραγματοποιηθεί χωρίς την υποστήριξη και την καθοδήγηση του Καθηγητή Μιχάλη Πατεράκη. Η εκτίμησή μου για εκείνον δεν περιορίζεται στο επιστημονικό και επαγγελματικό του έργο, αλλά επεκτείνεται στην ανθρώπινη πλευρά και τον χαρακτήρα του. Θεωρώντας την συνεργασία μας όλα αυτά τα χρόνια σαν εξαιρετική εύνοια της τύχης στο πρόσωπό μου, δεν μπορώ παρά να τον ευχαριστήσω ειλικρινά.

Ένα μεγάλο ευχαριστώ στον πολύτιμο φίλο και συνάδελφο Πολυχρόνη Κουτσάκη. Η βοήθειά του ήταν καθοριστική και συνεχής. Η κουμπαριά του ήταν η σημαντικότερη από τις κοινές δημοσιεύσεις μας και πραγματικά με τιμά κάθε στιγμή.

Ευχαριστώ την σύζυγό μου Ειρήνη Ρασούλη, για την ανεκτικότητα, την υποστήριξη και κυρίως την αγάπη της, από την οποία αντλώ συνεχώς ενέργεια. Σου δίνουν απίστευτη δύναμη τα σταθερά σημεία στη ζωή.

Ευχαριστώ τους γονείς και τα αδέρφια μου για την εμπιστοσύνη τους. Ελπίζω ότι τους κάνω περήφανους.

Τέλος ευχαριστώ τους συνεργάτες μου στο Κέντρο Ανάπτυξης και Διαχείρισης Δικτύου του Πολυτεχνείου, στο Εργαστήριο Πληροφορίας και Δικτύων και στο Ινστιτούτο Τηλεπικοινωνιακών Συστημάτων που πάντα πρόθυμα έδωσαν την βοήθειά τους.

Εύχομαι από την καρδιά μου να είστε όλοι σας πάντα καλά.

Σπύρος Ψυχής
Χανιά, Ιούλιος 2003

"The Only Real Problem is
Scaling… All the others inherit from that one.
If you can scale, everything else must be working."

Mike O'Dell - Chief Scientist UUNET, November 1998

On the Design of MAC Protocols and Transmission Scheduling Schemes for the Integration of Video Streams, Voice and Data Traffic over High Speed Packet Switched Wireless Networks

ABSTRACT

In this thesis we propose and evaluate mechanisms for the multiplexing and the integrated delivery of various traffic types such as Video, Data and Voice, over a Time Division Wireless Cellular High Speed Packet Switched Network. Two different scenarios are examined where the Video Traffic results from Videoconference sessions and High Quality Media playback, respectively. In both cases we focus on the uplink channel and we thoroughly investigate, by using actual MPEG-4 Video Streams, the system's performance under a variety of possible loads. Our goal is to achieve high aggregate channel throughput while preserving the Quality of Service requirements of each traffic type.

Chapter 1

Introduction and Background Review

Introduction

In this chapter we present a short review of the wireless networking platforms available today, an introductory review of the Wireless Networking concept, of the Asynchronous Transfer Mode and of the MPEG-4 video encoding. We also discuss the combination of Wireless Networks and the ATM technology. These introductory paragraphs are necessary since the reader of the Thesis must understand the environment in which the proposed mechanisms are designed and fitted to work into. Finally we present the reasons that lead us to introduce new Dynamic Time Division Media Access Protocols and the problems and challenges that we must face.

I. Background

a. Wireless Networking

During the last decade a booming technological evolution took place in the field of Telecommunication Systems. At the same time Wireless Communication dominated not only because it allows the users to move, even during the time they use the network services, but also because these network services are nowadays enhanced with all the features and capabilities that the wired networks offer. Thus, problems and issues that until now appeared only on wired networks now must be solved on wireless and vice versa. And of course the main concern in wireless communication is the same like in wired: scaling, in other words how to extend the broadband frontier to the end user. Therefore, the system capacity needs to grow as much as the demands of the various services grow. Wired Networks face the same problem but in a much smaller scale, since the usage of fiber optics in the network core together with the introduction of advanced modulation and multiplexing techniques, like Wavelength Division Multiplexing (WDM) and Dense Wavelength Division Multiplexing (DWDM), offer a reliable and high-speed communication. In addition, the usage of Digital Subscriber Line Techniques (xDSL), in the last mile of wired networks significantly increases the available bandwidth to the end user.

On the contrary, the use of the air as the 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 unacceptable by most of the network oriented applications. For this reason, the applications supported until recently by the wireless cellular communications networks involved mostly voice and low rate data transmission (e.g., the Short Message System - SMS).

b. What are 2G, 2.5G and 3G Wireless Networks?

i. 2G

The Wireless Industry is a fast moving market place with significant technological advances. The migration from second generation networks towards third and fourth generation systems present and create different business needs and opportunities.

Second generation protocols use digital encoding and include GSM, D-AMPS (TDMA) and CDMA (IS-95). 2G networks are in current use around the world. The protocols behind 2G networks support voice and some limited data communications, such as Fax and SMS, and most 2G protocols offer different levels of encryption, and security.

GSM or Global Systems for Mobile Communications is the most widely used digital mobile phone system and the de facto wireless telephone standard in Europe. It was originally defined as a pan-European open standard for a digital cellular telephone network to support voice, data, text messaging and cross-boarder roaming. GSM is now one of the world's main 2G digital wireless standards. Today's GSM platform is a hugely successful wireless technology and an unprecedented story of global achievement. In less than ten years since the first GSM network was commercially launched, it became the world's leading and fastest growing mobile standard, spanning over 190 countries.

Today, GSM technology is in use by more than one in ten of the world's population and it is estimated that at the end of 2002 there were 787 million GSM subscribers across 190 countries of the world. The growth of GSM continues unabated with more than 160 million new customers in the last 12 months. Since 1997, the number of GSM subscribers has increased by a staggering 10 fold. During late 2003 or early 2004, it is predicted that global GSM subscribers will smash through the one billion mark. [2]

The progress hasn't stopped there. Today's GSM platform is living, growing and evolving and already offers an expanded and feature-rich 'family' of voice and data enabling services. Second-generation GSM networks deliver high quality and secure mobile voice and data services (such as SMS/Text Messaging) with full roaming capabilities across the world.

ii. 2.5G

The move into the 2.5G world begun with the General Packet Radio Service (GPRS). GPRS is a radio technology for GSM networks that adds packet-switching protocols, shorter setup time for ISP connections, and the possibility to charge by the amount of data sent, rather than connection time. 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 supports flexible data transmission rates as well as continuous connection to the network. GPRS is the most significant step towards 3G.

*E*nhanced *D*ata Rates for *G*SM *E*volution *(EDGE)* will also be a significant contributor in 2.5G. EDGE supports data, multimedia services and applications at up to 384 Kbps. It was developed initially by Ericsson for mobile network operators who fail to win Universal Mobile Telephone System (UMTS) spectrum and will let them compete with 3G networks offering similar data services and applications, without having a 3G license. It uses the same TDMA frame structure, logic channel and 200 kHz carrier bandwidth as today's GSM networks, which allows existing cell plans to remain intact. It has been described as a software overlay for GSM/GPRS networks since in most cases all that is needed on the Base Station is a simple software upload. EDGE, introduces a new modulation technique along with improvements in the radio protocol that allow operators to use existing GSM frequency spectrums, 800, 900, 1800 and 1900 MHz, more effectively. It uses 16 Quadrature Amplitude Modulation (16QAM), rather than normal GSM Gaussian Modulation Shift Keying (GMSK). [30]. It is commercially available since 2001, and has been implemented in North America (North Carolina, South Carolina, coastal Georgia, eastern Tennessee and Puerto Rico). Telecom Italia Mobile (TIM) announced on June 2003 that will be the first major European operator to deploy EDGE. Although several other operators in the region will follow TIM's lead, either deploying EDGE to provide multimedia to sparsely populated areas or using it as an interim technology to fill the gap caused by delays in the deployment of 3G, the widespread deployment across Western Europe should not be expected. Having spent 85 billions euros on their licenses, most operators will prefer to stay committed to 3G.

In Europe, the main focus, towards the 3G, remains on Wideband Code Division Multiple Access (WCDMA) technology (for details see the next paragraph). 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.

iii. 3G

3G, or Third Generation protocols, is a new network technology being developed for mobile network operators. Using a new spectrum, it will support much higher data rates, measured in Mbits per second. Offering greater capacity and efficiency, it will offer high-speed data services to support bandwidth hungry multimedia applications like full-motion color video, videoconferencing and 'always on' Internet access.

The 3G standard that has been agreed for Europe, Japan, China and the rest of Asia is referred to as the Universal Mobile Telecommunications System (UMTS). The technology chosen to implement UMTS is Wideband Code Division Multiple Access (WCDMA) [1], [3]. With WCDMA, carrier bandwidth will jump from 200KHz up to 5MHz, allowing for the much greater capacity and data speeds (up to 2Mbps) inherent in a wideband system. New frequency allocations, terminals and base stations will be required for the changeover to WCDMA. [31]

UMTS is already a reality. Japan launched the world's first commercial WCDMA network in 2001 (DoCoMo’s i-Mode, owned by NTT), and in the fourth quarter of 2002 UK operator Hutchison 3G has opened its own network to customers under the "3" brand. Several other pilot and pre-commercial trials are operational in the Isle of Man, Monaco and other places in Europe.

c. Asynchronous Transfer Mode (ATM)

ATM is the world's most widely deployed backbone technology. This standards-based transport medium is widely used within the core--at the access and in the edge of telecommunications networks to send data, video and voice at ultra high speeds.

ATM is best known for its easy integration with other technologies and for its sophisticated management features that allow carriers to guarantee quality of service (QoS). These features are built into the different layers of ATM, giving the protocol an inherently robust set of controls.

Sometimes referred to as cell relay, ATM uses short, fixed-length packets called cells for transport. Information is divided among these cells, transmitted and then re-assembled at their final destination [4].

ATM is divided into layers (See Fig. 1.1). The physical layer is divided into two parts. The ATM physical medium sublayer is responsible for transmission of data over the physical medium, regardless of the type of medium used. The transmission convergence sublayer is responsible for the transmission and reception of frames over framed transmission facility, such as E3 or T3. ATM cells are packed into these frames and unpacked at the remote end. This sublayer also performs error detection/correction but only on the ATM header. This prevents the cells from being sent to the wrong destination.

The ATM Layers

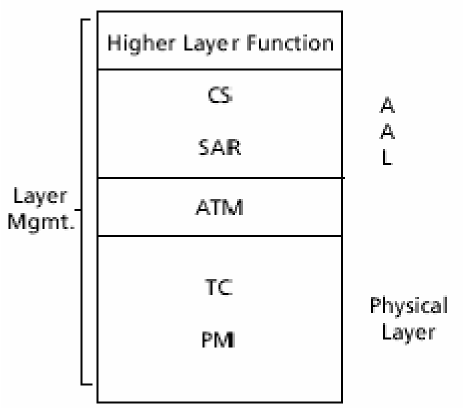

Figure 1.1

The ATM layer is responsible for multiplexing cells over the interface. At endpoints, the ATM layer generates and interprets cell headers (endpoints do not route cells).

The ATM Adaptation Layer (AAL) is used mostly by endpoints. It is divided into two sublayers: Segmentation and Reassembly (SAR) and the Convergence Sublayer (CS).

The SAR reconstructs data that has been segmented into different cells (reassembly). It is also responsible for segmenting data that cannot fit within a 48-byte payload of an ATM cell (segmentation).

The CS determines the class of service to be used for a transmission. This will depend on the bit rate (constant or variable bit rate), the type of data, and the type of service to be provided (connection-oriented or connectionless). The class of service assigned determines the QoS parameters necessary for the transmission [35].

The basic principles of ATM as put forward by CCITT in Recommendation I.150 are:

- ATM is considered as a specific packet oriented transfer mode based on fixed length cells of 53 bytes. Each cell consists of an information field and a header (48 and 5 bytes respectively), which is mainly used to determine the virtual channel and to perform the appropriate routing (Figure 1.2). Cell sequence integrity is preserved per virtual channel.

The use of the header fields is:

i. **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.

ii. VPI (Virtual **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.

iii. VCI (Virtual **C**hannel **I**dentifier): (16 bits). It uniquely defines a connection between two end nodes.

iv. PTI (Payload **T**ype **I**dentifier): (3 bits). It determines whether the data belong to a certain user or consist of network administration information.

v. CLP (Cell **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.

vi. HEC (Header **E**rror **C**ontrol): (8 bits). The end node, for error control of the cell's header, uses it.

The ATM Cell

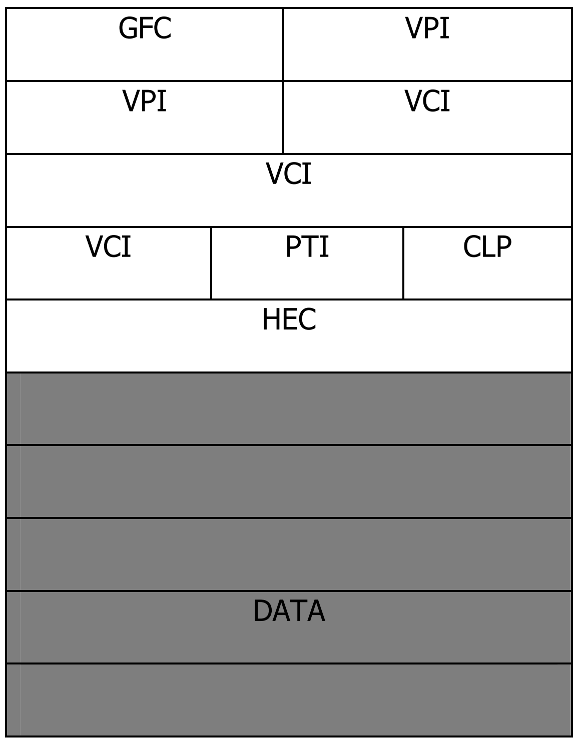

Figure 1.2

- ATM is connection-oriented. The header values are assigned to each section of a connection for the complete duration of the connection. Signaling and user information are carried on separate virtual channels.
- The information field of ATM cells is carried transparently through the network. No processing, like error control, is performed on it inside the network.
- All services (voice, video, data, etc.) can be transported via ATM, including connectionless services. To accommodate various services an adaptation function is provided to fit information of all services into ATM cells and to provide service specific functions (e.g. clock recovery, cell loss recovery, etc).

d. The Moving Picture Experts Group (MPEG) framework and the MPEG-4 standard

The MPEG standards are an evolving set of standards for video and audio compression and for multimedia delivery developed by the Moving Picture Experts Group.

MPEG-1 was designed for coding progressive video at a transmission rate of about 1.5 million bits per second. It was designed specifically for Video-CD and CD-i media. MPEG-1 audio layer-3 (MP3) has also evolved from early MPEG work.

MPEG-2 was designed for coding interlaced images at transmission rates above 4 million bits per second. MPEG-2 is used for digital TV broadcast and DVD. An MPEG-2 player can handle MPEG-1 data as well.

MPEG-1 and -2 define techniques for compressing digital video by factors varying from 25:1 to 50:1. The compression is achieved using five different compression techniques:

1. The use of a frequency-based transform called Discrete Cosine Transform (DCT).
2. Quantization, a technique for losing selective information (sometimes known as lossy compression) that can be acceptably lost from visual information.
3. Huffman coding, a technique of lossless compression that uses code tables based on statistics about the encoded data.
4. Motion compensated predictive coding, in which the differences in what has changed between an image and its preceding image are calculated and only these differences are encoded.
5. Bi-directional prediction, in which some images are predicted from the pictures immediately preceding and following the image.

The first three techniques are also used in JPEG file compression.

A proposed MPEG-3 standard, intended for High Definition TV (HDTV), was merged with the MPEG-2 standard when it became apparent that the MPEG-2 standard met the HDTV requirements.

MPEG-4 is a much more ambitious standard and addresses speech and video synthesis, fractal geometry, computer visualization, and an artificial intelligence (AI) approach to reconstructing images. It addresses a standard way for authors to create and define the media objects in a multimedia presentation, how these can be synchronized and related to each other in transmission, and how users are to be able to interact with the media objects.

It is an ISO standard that provides a standardized and worldwide accepted technological framework that enables the integration of the production, storage distribution and content access paradigms. It leverages existing digital video content by supporting the MPEG-1 and MPEG-2 coding standards. Furthermore, it enables richer development of new digital video applications and services. An MPEG-4 scene is composed of a number of audiovisual objects, which can be either static or time varying in nature. [8]

MPEG-4 Visual provides a natural video coding algorithm that is capable of operating from 5 kbps with a spatial resolution of QCIF (144x176 pixels). It is ITU-T H.263 compatible, in the sense that an MPEG-4 video decoder also correctly decodes an H.263 bit stream. The algorithm scales up to bit rates of some Mbit/s and optimization has been carried for ITU-R 601 resolution pictures (288x720@50Hz and 240x720@59.94 Hz). The recently developed Studio Profile can operate at over 1 Gbit/s. In addition it has a so-called Fine Granularity Scalability mode that allows transmission of the same video content at different bit rates from one coded version of the movie [9].

The audio platform of MPEG-4 (MPEG-4 Audio) provides a complete coverage of the bitrate range of 2 to 64 kbit/s. Good speech quality is obtained already at 2 kbit/s and transparent quality of monophonic music sampled at 48 kHz at 16 bits per sample is obtained at 64 kbit/s. Appealing quality is obtained at 16 kbit/s for stereo music.

MPEG-4 Systems provides the object composition technology referred to above. This is based on Virtual Reality Modeling (VRML), but extends the original VRML functionality by allowing the inclusion of streamed audio and video, natural objects, generalized URL, composition update and, most important, a very effective compression for VRML type information. This technology is called BIFS (Binary Format for Scene Description). As VRML only addresses 3D worlds, MPEG-4 Systems defines all 2D nodes that are needed by 2D-only applications according to the same organization of 3D. As done for MPEG-1 and 2, MPEG-4 Systems provides precise synchronization of audio and video objects. MPEG-4 Systems supports both push and pull delivery of content. MPEG-4 also has support for Object Content Identification (OCI), so that searches in data bases of MPEG-4 objects become possible.

The last component of MPEG-4 is referred to as the Delivery Multimedia Integration Framework (DMIF) and this provides three types of abstraction. The first is an abstraction from the transport protocol, which can be RTP/UDP over IP enabled networks, AAL5 over ATM networks or MPEG-2 Transport Stream (MPEG-2 TS) or others. The ability to identify delivery systems with different QoS, provides a very effective way to charge the cost of delivery depending on the QoS required by the application or by the user. The second is the abstraction of the application from the delivery type. Interactive (client-server), local or broadcast delivery is seen through a single interface. Furthermore, in the interactive case DMIF provides an abstraction from the signaling mechanisms of the delivery system.

II. 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 user's 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 [10].

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 that must be fulfilled in a wireless ATM system:

i. Efficient support of all ATM classes in the wireless channel (this is achieved with the use of an efficient multiple access control protocol).

ii. Limitation of the consequences of the relatively high error rate (this is achieved with the use of an error control mechanism).

iii. Avoidance of large cell dropping during user handover (several proposals exist for an efficient procedure which could achieve this, e.g. [11, 12, 13]).

III. Channel Access: why use TDMA?

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 [15], due to its capability to adjust to the needs of each connection, reserving more or less time slots in each case.

IV. Architect's Perspective

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.

Hierarchical Cellular System

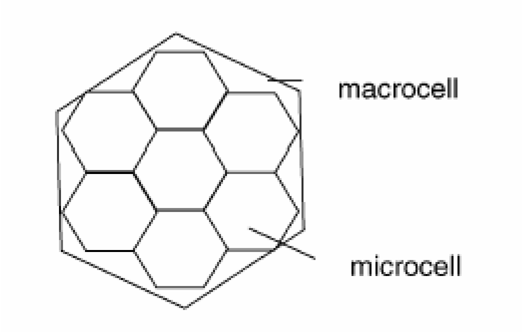

Figure 1.3

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) (Figure 1.3).

Two tiered Hierarchical Wireless Network Architecture

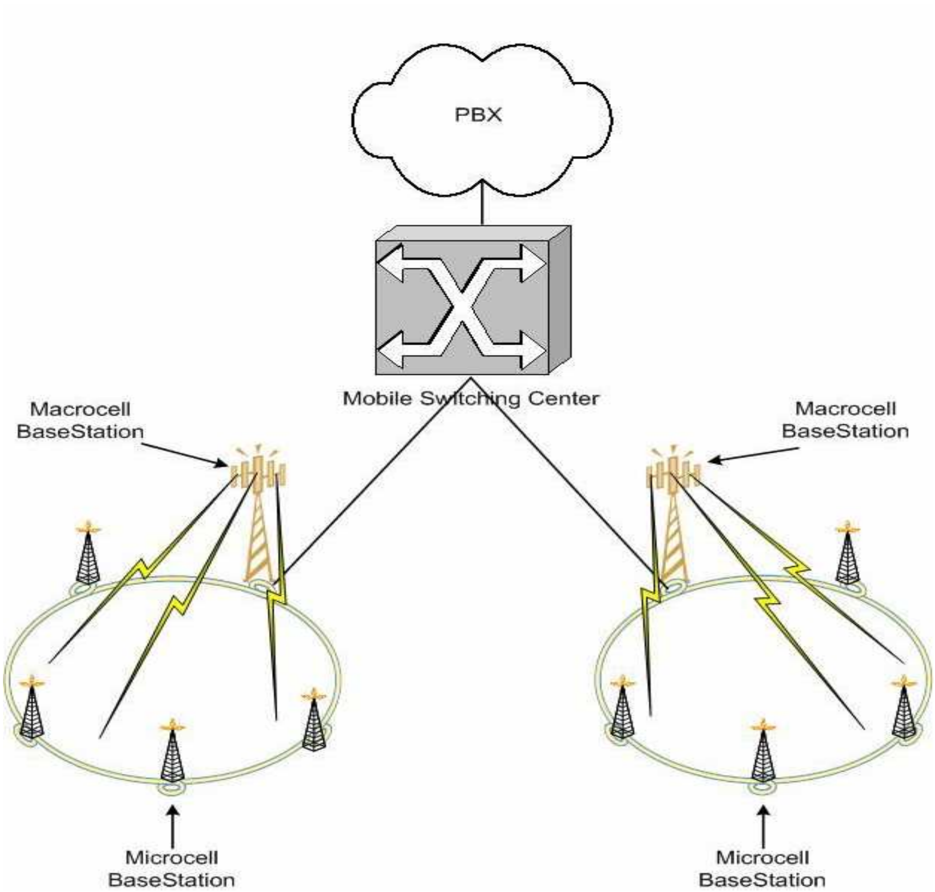

Figure 1.4

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 1.4, 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 (by wire or wireless) to localize the network control associated with microcell handovers.

V. Why we need efficient MAC Protocols?

The MAC protocol allows the mobile nodes to avoid contention while sharing a common channel for data transmission. It plays a key role in determining the efficiency of channel utilization in the network. Since the wireless bandwidth is scarce, there is a need for efficient MAC protocols for mobile networks.

System capacity can be increased by:

a. Using a cellular structure with a cell size as small as possible (microcells) in order to reduce transmitting power and to increase frequency reuse. Microcell diameters are usually of the order of a few hundred meters; therefore the round-trip propagation delay within a microcell is negligible (of the order of 1 μs or less). As we already mentioned in section IV modern wireless network architectures solve this problem efficiently.

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

MAC protocols must be simple and easy to implement in order to require the minimum processing power for their implementation.

Especially the delivery of high quality live video material is a service with very strict QoS requirements. Consequently, the design of MAC strategies that integrate this class of service with others, which often require contradictory QoS guaranties, is needed. These MAC strategies must also achieve the optimization of the usage of the system capacity.

VI. Our Work

The work presented in this Thesis focuses 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, and is divided in two parts.

In the first part, presented in Chapter 2, we extend previous approaches that have been made on the problem [16, 17]. We combine the tactics we followed at these works with physical medium characteristics that have been proposed in [18]. A MAC scheme that efficiently integrates voice (Constant Bit Rate, CBR, On/Off Traffic), videoconference traffic streams (Variable Bit Rate, VBR) and bursty data traffic in high capacity micro cellular environments is presented and investigated. Mobile Terminals are considered to be handy video telephones with integrated video camera, able to transmit low quality and low scene activity video to the uplink channel.

The contribution of this work, beyond the design of an efficient MAC scheme, extends to the base station scheduling scheme. Both of these schemes contribute to the achievement of the QoS required by the various traffic streams.

In the second part, which is presented in Chapter 3, the MAC scheme is modified in order to integrate only two types of traffic, (i.e., video and data), in a wireless channel of equally high capacity as the one introduced in the first part of the work. However, in this case, Mobile Terminals (MTs) are considered to be high performance devices with extended storage capabilities which can act like cache memories, streaming multimedia material to other MTs (in the same or in adjacent micro cells) in order to reduce client start-up latencies and improve the network performance.

During the last five years a lot of work has been done on various caching schemes that permit the fair and effective sharing of the network resources (for example see [19], [20], and [21] and the references therein). Most of them have been deployed especially for the retrieval and delivery of high quality stored multimedia content. And since the notion of broadband access is transferred to the wireless field, most of these caching and content placement architectures must be extended. As we already mentioned, in this case MTs can act like cache memories and even the whole micro cell can be faced as a cache farm from the rest of the network. Therefore, we envision high quality stored video streaming on the uplink channel (MTs to the BS) in order for these streams to be delivered for playback to their destinations.

The reason for not including voice traffic in the integration in this case is that the video traffic considered is very bursty and highly demanding in its bandwidth requirements, therefore allowing little room for any other type of time sensitive traffic to be serviced. The small amount of remaining bandwidth is used for the service of the data traffic.

As for the downlink channel we assume throughout the Thesis, that since the base station is the sole transmitter on that channel, it is in complete control of the downstream traffic, using TDMA to relay information to the users.

Chapter 2

Multiplexing Live Video Streams & Voice with Data over a High Capacity Packet Switched Wireless Network

Introduction

A multiple access scheme (MAC) that efficiently integrates voice (Constant Bit Rate, CBR, On/Off Traffic), video (Variable Bit Rate, VBR) and bursty data traffic in high capacity micro cellular environments is presented in this chapter. The delivery of high quality live video material is a service with very strict QoS requirements. Consequently the design of MAC strategies that integrate this class of service with others, which often require contradictory QoS guaranties, is needed. Within the microcell, spatially dispersed mobile terminals (MTs) share a radio channel that connects them to a fixed base station or a wireless hotspot. The base station allocates the channel resources, delivers scheduling and feedback information and serves as an interface to the mobile switching center (MSC), which provides access to the fixed network infrastructure and the Internet.

I. Channel Structure, Actions of Terminals and Base Station Scheduling.

The uplink channel time is divided into time frames of equal length. The frame duration is selected such that an active voice terminal (i.e. a terminal in talkspurt) generates exactly one packet per frame. Each frame consists of two *types* of intervals. These are the *request* intervals and the *information* intervals.

Within an information interval, each slot accommodates exactly one, fixed length, packet that contains voice, video or data information and a header.

The voice and data terminals use the request intervals in order to transmit their requests to the BS. Voice terminals are given priority to transmit their requests to the BS. When all the Voice Requests have been transmitted the Data Request transmission follows. The concept of reserving a minimum bandwidth for voice terminals to make channel reservations helps to keep the voice access delay within relatively low limits and gives clearly better performance than the PRMA [22] and quite a few PRMA-like algorithms, such as DPRMA [23], where the absence of request slots leads to a continuously decreasing probability of finding available information slots as traffic increases, and hence to greater access delays.

The request intervals consist from slots, which are subdivided into mini-slots, and each mini-slot accommodates exactly one, fixed length, request packet. By using more than one minislot per request slot, a more efficient usage of the available request bandwidth is possible. Since we assume that all of the voice source state transitions occur at the frame boundaries, we place all request intervals at the beginning of the frame, in order to minimize the voice packet access delay.

We adopt the idea 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 use of the request bandwidth. Furthermore, we assume that the number of request slots per channel frame is variable, depending on the number of the video terminals. [24]

Figure 4 gives an example of the channel frame structure showing the request and information slots within a frame. Notice that in this example each request slot consists of two minislots.

Frame Structure

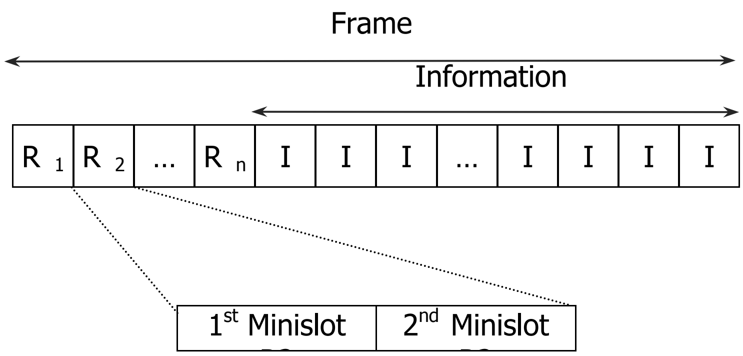

Figure 4

The video terminals envoy their slot requests to the base station by transmitting them within the header of the first packet of their current video frame. This is possible due to the nature of the video sources, which we assume here that they transmit low activity video content with a few and small changes in bit rate overtime. Thus there is no need to waste channel slots in order to facilitate the transmission of video request packets to the BS.

The BS allocates channel resources <u>at the end of the corresponding request interval</u>, and follows a different allocation policy for video terminals than that for voice terminals.

Video terminals have the highest 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. If this happened, voice terminals would keep their dedicated slots for the whole duration of their talkspurt, and thus video terminals, would not find enough slots to transmit in, and this could lead to a violation of the video QoS requirements.

Consequently, in our scheme the BS allocates a slot to each requesting voice terminal with a probability ***p*.*** When there are no video terminals in the system, ***p**** is set equal to 1.

Each data user can receive only one information slot per channel frame and can keep it just as long as it has data packets to transmit.

We do not use preemption of data reservations but still voice users are given priority both in slots and in allocation policy.

II. Voice Traffic

We assumed that the Voice Terminals are equipped with Voice Activity Detectors (VAD) [22, 32, 33]. These VADs can recognize the transitions of the human voice between silence and talkspurt and vice versa. Thus Voice Sources follow an alternating pattern of talkspurts and silence periods (on and off). During the talkpurt periods the Voice Sources generate traffic at a bitrate of 32 Kbps. The mean talkspurt duration is 1.0 secs and the mean silence duration is 1.35 secs. This assumption is directly related to the physical characteristics of the VAD (i.e. depends on how quickly the VAD recognizes the voice transitions). The upper delay limit that a voice packet can suffer is equal to 40 ms [23, 34]. Finally for all the voice source transitions (e.g., talk to silence) occur at the frame boundaries which 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.

III. Data Traffic

We adopt the data traffic model based on statistics collected on email usage from the Finish University and Research Network (FUNET) [25]. The model simulates email data traffic. The probability distribution function f(x) for the length of the data messages was found to be well approximated by the *Cauchy* (0.8,1) distribution. The message inter-arrival time distribution for the FUNET model is exponential. In our work, we assume an upper limit on the average data message delay equal to 2 seconds.

IV. Video Streams

In our study, we use the trace statistics of actual MPEG-4 streams from [26]. 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 the bit rate equal to 434 Kbps.

All the streams are encoded at 25 video frames per second, which means that a new video frame is generated every 40 msecs. In our study we have made the assumption that the maximum transmission delay for video packets is 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 [23].

V. Simulation Results and Discussion

An extensive simulation study was carried out. The simulator was developed using the GNU C programming language. The platform used for the simulation runs was a SUN ULTRA. Each run simulated one hour of actual network activity (300005 channel frames). All the results shown correspond to averages of 10 independent runs (Monte Carlo method).

Initially we simulated the system under all possible movie loads from 0 to 22 movies with no active data terminals. These runs helped us to specify the boundaries where the number of the request slots must change in order to accommodate the generated voice and data request packets. Additionally the value of the *probability* ***p**** was chosen equal to 0.1 (10%), since 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. Table 2.1 shows the results of these runs.

A careful reviewer of the results shown above will notice that the reduction of the Voice Terminals capacity as the number of Video Terminals increases does not appear in a completely linear manner. This is also shown in Figure 2.1 where the results are graphically presented. Each different shadow on the bars corresponds to a different value of R-slots.

VIDEO	VOICE	R-SLOTS	AVERAGE THROUGHPUT (%)
0	1258	30	94.57
1	1092	25	84.30
2	1062	25	84.25
3	1003	25	82.03
4	896	25	76.19
5	870	25	76.45
6	847	25	76.93
7	763	20	72.82
8	734	20	72.85
9	712	20	73.40
10	591	20	66.51
11	560	20	66.39
12	514	20	65.14
13	441	12	61.86
14	412	12	61.89
15	363	12	60.42
16	296	5	57.59
17	265	5	57.47
18	212	5	55.69
19	130	5	51.73
20	98	5	51.54
21	70	5	51.64
22	0	5	48.59

Table 2.1

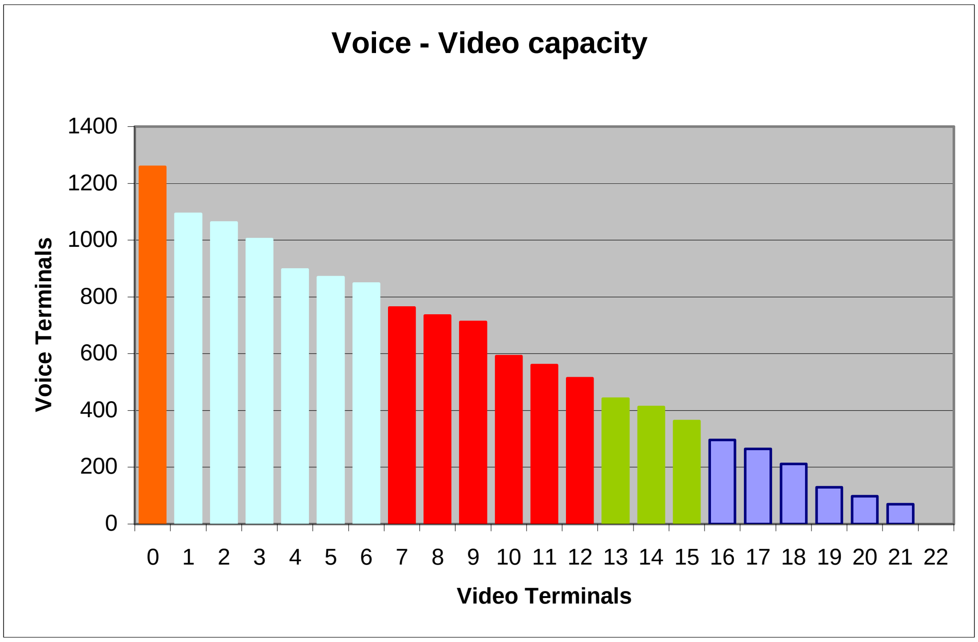

Figure 2.1

This non-linear reduction appears due to the relatively few changes in the values of R. The reason we do not change the number or R slots each time a new Video Terminal enters the system, is that such an approach would increase the system complexity since the Base Station would have to change the channel frame structure quite often. Therefore, we assume that the Base Station has a predefined small set of operating modes, which derive from the number of R slots reserved.

During the second set of experiments we simulated the system under Voice and Data traffic only. Since there is no Video in the system the value of R was set to 30 slots. We tried to accommodate into the system as many voice terminals as possible given that the data throughput should match the data load value. Table 2.2 shows the difference between the theoretical calculation of the Channel Voice capacity and the actual one.

The theoretical calculation of the voice capacity was made with the equation:

[(total number of information slots-estimated data throughput) / probability that a voice terminal is in talkspurt].

λ	**Voice Terms Capacity (Theoretical)**	**Voice Terms Capacity**	**Difference**
0,25	1212	1212	0
0,5	1165	1135	-30
0,75	1118	1078	-40
1	1071	1030	-41
1,25	1024	985	-39
1,5	978	938	-40
1,75	930	890	-40
2	883	845	-38

Table 2.2

We observe that the actual Voice Capacity is from 0 to 4.3 % smaller than the one predicted theoretically, i.e., the result achieved by our scheme is very close to the theoretical maximum. Our simulations also show that the data mean message delay is significantly lower than the assumed upper bound of 2 secs. The average data message delay does not exceed 1 sec, except in the case $\lambda=0.25$ where the number of voice terminals is quite high.

During the third set of experiments we simulated the system under certain mixtures of Voice, Video and Data traffic. Table 2.3 shows some representative results.

R-Slots	λ	Video Terms	Voice Terms	Average Throughput (%)
5	0,5	20	0	51,24
12	0,5	15	288	61,85
12	1,5	15	94	61,40
20	0,5	10	496	66,44
20	1,5	10	298	65,69
25	0,5	5	780	76,75
25	1,5	5	586	76,30
25	0,5	1	971	82,27
25	1,5	1	771	81,37
30	1,5	0	938	91,72

Table 2.3

We notice that, for the same number of Video Terminals, when λ increases by 1 (e.g. from 0.5 to 1.5, with 15 active Video Terminals), the system's voice capacity decreases by approximately 194 voice terminals. The reason for this is that a voice terminal is in talkspurt with probability 0.425 and each voice terminal in talkspurt requires one information slot per channel frame. Therefore, the voice capacity decrease of 194 terminals corresponds on average to approximately 83 slots, which is almost equal to the average increase of 80 slots needed by the data terminals because of the increase in the value of λ by one message per frame (i.e. on average 80 packets per frame). This means that our protocol achieves an almost stable channel throughput for a given number of video terminals, as shown in the last column of Table 3. The first and the last rows of the Table include the minimum (no voice, many bursty video sources are present) and the maximum (no video, many CBR voice sources are present) channel throughputs achieved by our system, respectively.

VI. Conclusions

In this chapter, we have proposed and investigated the performance of a mechanism for transmitting videoconference streams, voice and data over a high capacity wireless channel. We evaluated its performance through an extensive simulation study in which we used actual MPEG-4 stream traces together with data generated by a synthetic model fitted to measurements of data traffic from a National University and Research Network (FUNET). Our scheme is shown to achieve high aggregate channel throughput, expressed in terms of the maximum number of video and voice terminals with the maximum data load sustained by the system, in many cases of traffic load, while preserving the Quality of Service (QoS) requirements of each traffic type.

Chapter 3

On the Integration of MPEG-4 streams Pulled Out of High Performance Mobile Devices and Data Traffic over a Wireless Network

Introduction

In this chapter we propose and evaluate mechanisms for the multiplexing and the integrated delivery of Video and Data Traffic over a Wireless Cellular High Speed Packet Switched Network. We envision a system where the Mobile Terminals are considered to be high performance devices with extended storage capabilities which can act like cache memories streaming multimedia material. We focus on the uplink channel and we investigate the system's performance under a variety of possible loads which consist of actual MPEG-4 streams and Data Traffic.

I. Frame structure, BS scheduling and Terminal actions

The uplink channel time is divided into time frames of equal length. Each frame consists of two types of intervals. These are the request intervals and the information intervals. Each of these intervals is divided into a number of time slots.

Within an information interval, each slot accommodates exactly one, fixed length, packet that contains video or data information and a header. Each request slot is subdivided into two mini-slots and each mini-slot accommodates exactly one, fixed length, request packet. By using more than one mini slot per request slot, a more efficient usage of the available request bandwidth is possible, [16]. 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)). Upon successfully transmitting a request packet the terminal waits until the end of the corresponding request interval to learn of its reservation slot (or slots). If unsuccessful within the request intervals of the current frame, the terminal attempts again in the request intervals of the next frame. A terminal with a reservation transmits freely within its reserved slot [23].

We adopt the idea that video and data terminals can share the request slots (first by the video terminals, and after the video contention period, by the data terminals). The reason we require video terminals to contend in order to reserve information slots is that we are dealing with high quality stored video content which, in contrast to videoconference streams, generates highly bursty information. Video requirements change frequently over time and it is impossible to ensure that their demands will reach the BS successfully if we use techniques similar to the ones described in [16],[17]. The two-cell stack reservation random access algorithm [28] is used to resolve the collisions, among video request packets.

After the end of the video contention period, data terminals without a reservation and with message ready for transmission, begin their contention in order to reserve information slots. In the case that two or more data request packets collide, the two-cell stack random access algorithm [27] is used to resolve the collision, due to its operational simplicity, stability and high throughput.

The BS allocates channel resources at the end of the corresponding request interval, and follows a different allocation policy for video terminals than that for data terminals.

Video terminals have higher 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 data terminals. Since we assume here that data slot reservations can not be preempted in favor of video terminals waiting for transmission, we choose not to allocate more than one slot per frame to a requesting data terminal. Otherwise, if a full allocation is not possible, the BS grants to the video terminals as many of the slots they requested as possible (i.e., the BS makes a partial allocation). 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.

There are two reasons for adopting the partial allocation policy, described above, to the video users. The first is that in our system there are 10 available reservation slots (20 minislots) in each channel frame, which practically means that almost all of the possible collisions among video terminal request packets can be resolved within the contention period of one channel frame. The second reason is that the video terminals are transmitting high quality bursty video. The size of a video frame is usually of the order of a few kilobytes, which means that after segmentation it corresponds to a few hundred channel cells (of ATM size). Especially in MPEG-4 coding, the correlation factor between two adjacent video frames, is extremely high. Therefore, even in the case that a few tens of cells do not reach the destination; the decoder can successfully reconstruct the video frame. Additionally, MPEG-4 incorporates very capable error correction techniques and can therefore suffer relatively higher packet loss rates compared to other media standards without serious degradation of the playback quality. Consequently, we try to conform to the QoS requirement on packet dropping while spreading the packet losses in time. This spreading of losses can lead to an imperceptible blurring during video playback, which is expected to be acceptable when watching a movie, instead of concentrating the lost cells in time, and consequently receiving video “hiccups” and severe losses in the decoder.

When a video terminal wants to decrease its bit rate, it releases all the slots that were previously allocated to it and are no longer necessary. The BS realizes the change in the bit rate when the first currently unnecessary reserved slot is released by the video terminal (and is observed empty by the BS in the current frame), and consequently allocates the newly released slots to any other requesting terminals, after the end of the request interval of the next frame

When the bit rate of a video terminal increases then the terminal must enter the contention process in order to acquire the additional information slots it needs. In such case, the terminal releases all of its currently reserved information slots before entering the contention period. When it passes the contention period successfully, then the required information slots are reserved and the video terminal can freely transmit within them. The rationale behind this policy is fairness. We prefer to let each video terminal compete with other video terminals at an equal basis (since video has absolute priority), in order to “spread” the video packet losses uniformly.

II. Data Traffic

We adopt the same Data Traffic that was presented in Chapter 2 from FUNET [25].

III. Video Streams

In our study, we use the trace statistics of actual MPEG-4 streams from [26]. The video streams that we used, were carefully selected in order to cover a broad range of movie characteristics. Table 3.1 shows the movies we used together with their mean and peak bit rates.

Video Stream Characteristics

Movie Name	Mean Bit rate (Kbps)	Peak Bit rate (Kbps)
News	720	3400
Parking Lot Cam	790	2800
Silence of the Lambs	580	4400
The Simpsons	1300	8800
Soccer	1100	3600
Star Wars	280	1900

Table 3.1

Figures 3.1 through 3.6 show the histograms of the bit rate of each movie. Horizontal axis shows for each video frame the corresponding bit rate quantized at 1 Kbps sectors. Vertical axis shows the number of video frames that correspond to a specific bit rate. Histograms delineate the bit rate behavior for each movie; a spread distribution corresponds to a bursty stream. Notice that the axes are not under the same scale.

All the streams are encoded at 25 video frames per second, which means that a new video frame is generated every 40 msecs. Since the video streams are pulled out from Mobile Terminals that act like caches and not dedicated cache engines with optimization or smoothing functionalities, no packet shaping mechanisms were used. Video content is delivered to the network in the initial form that was encoded. In our study we have made the assumption that the maximum transmission delay for video packets is 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 from [23].

News Stream

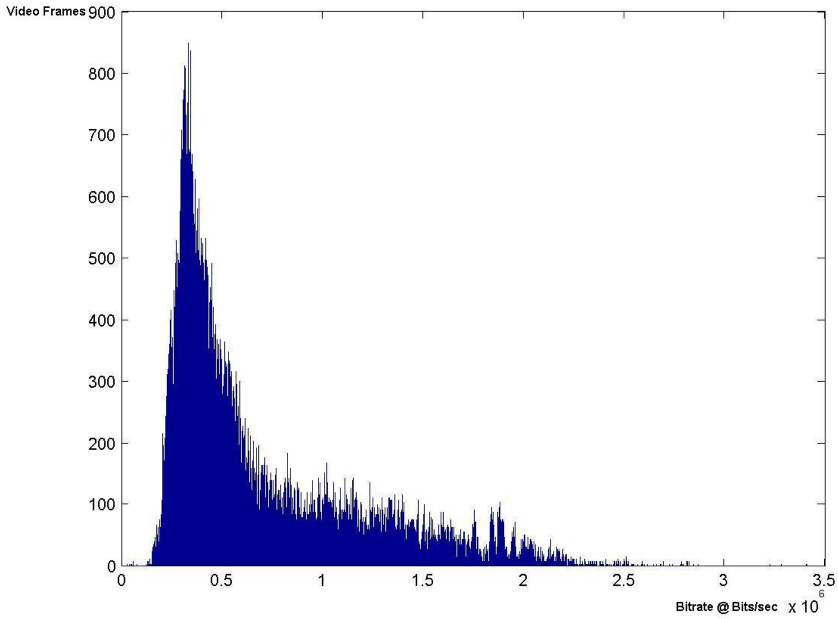

Figure 3.1

Parking Lot Stream

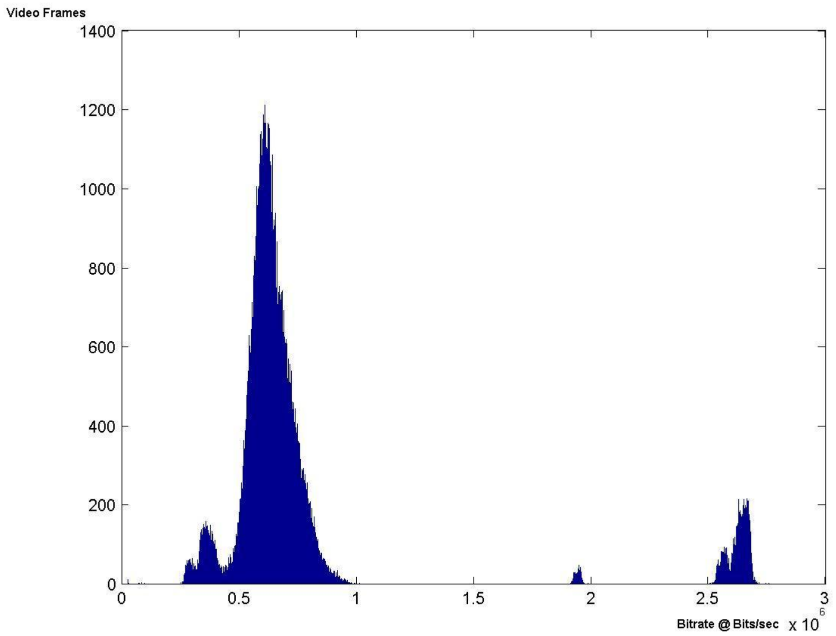

Figure 3.2

Silence of the Lambs Stream

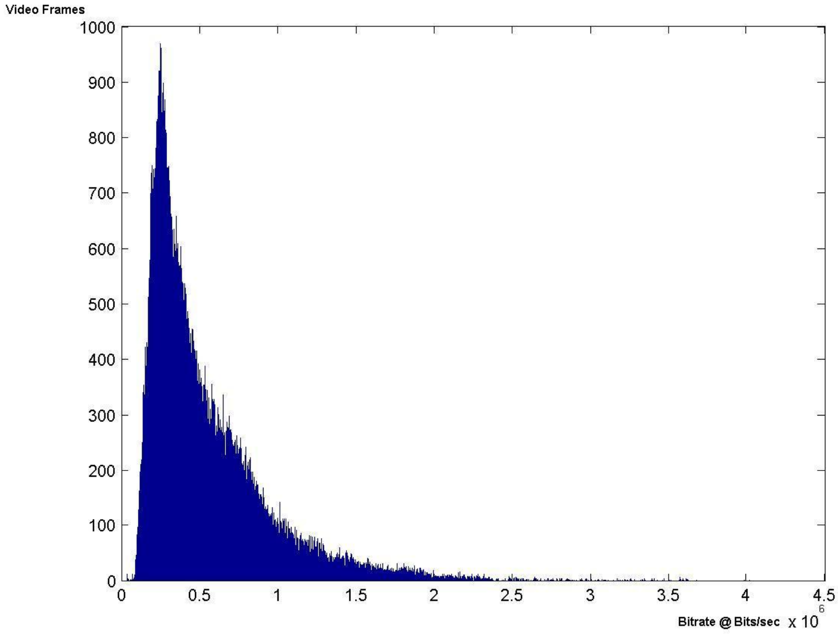

Figure 3.3

The Simpsons Stream

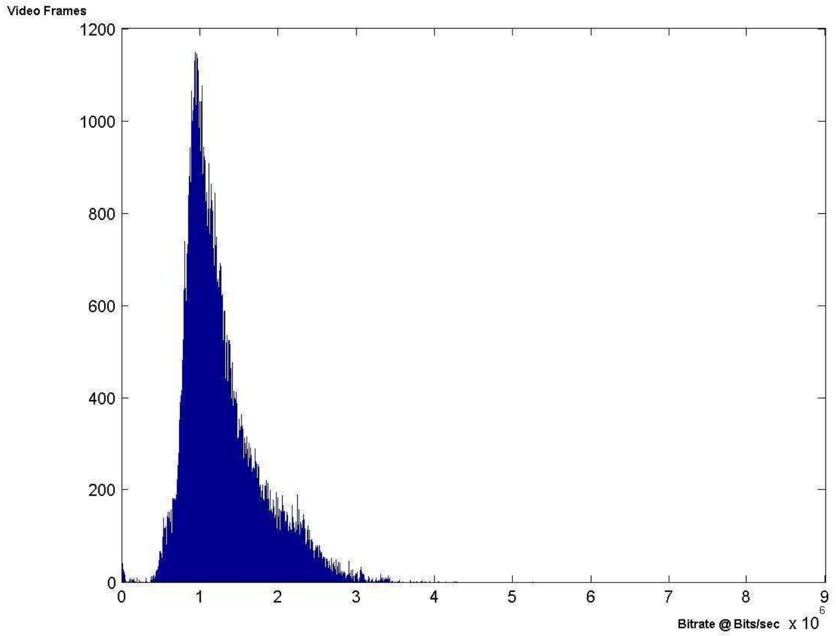

Figure 3.4

The Simpsons Stream

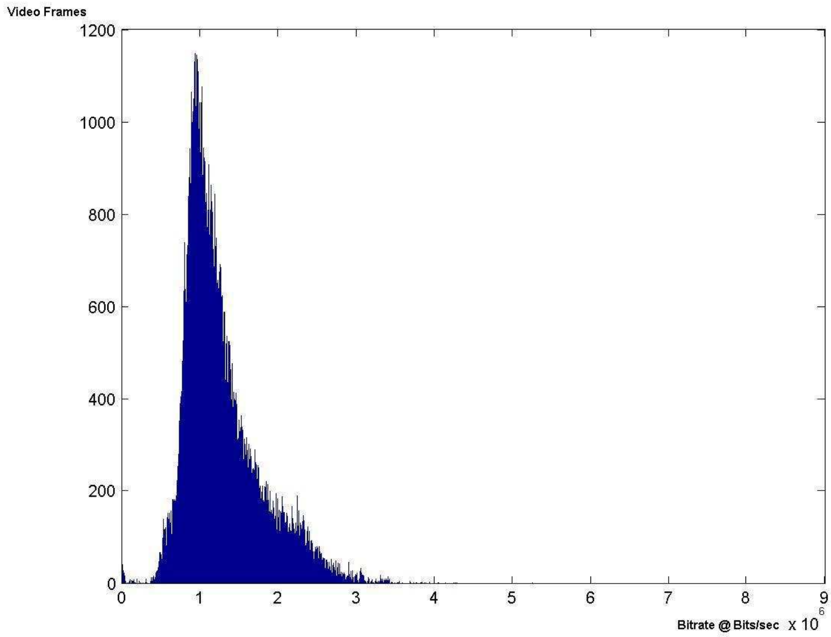

Figure 3.5

Soccer Stream

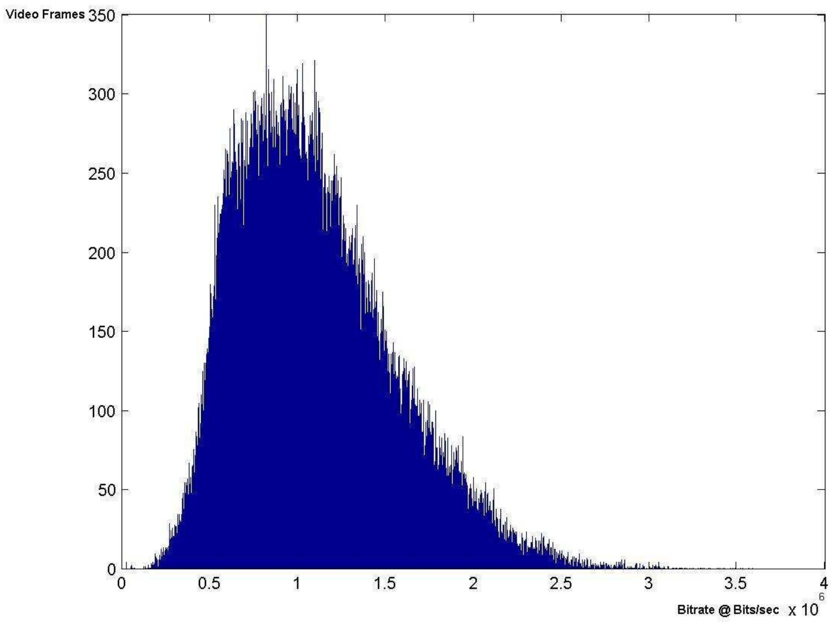

Figure 3.6

Star Wars Stream

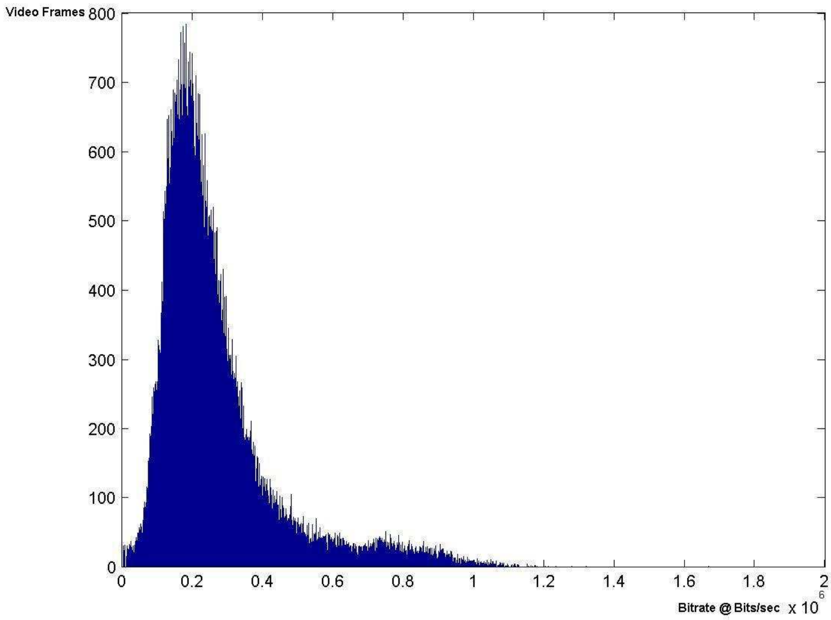

Figure 3.7

IV. System Parameters

The channel rate is 20Mbps (from [18]). As mentioned before, in this study we investigate the case where the system traffic includes video and data. Nevertheless, high capacity wireless channels will often be used for integrating voice, video and data traffic, and the frame duration is usually chosen to be equal to the time a voice terminal needs to generate a new voice packet [29].

Assuming a speech codec rate of 32 Kbps and that the information packet length is equal to the size of an ATM cell, yields the channel frame duration of 12 ms. The 12 ms of frame duration accommodate 566 slots (556 information slots plus 10 reservation slots). Consequently, the above channel's information payload transmission speed is slightly above 18 Mbps (because of the ATM headers encapsulation and the existence of reservation slots).

Finally, given that the packet length is equal to the size of an ATM cell, it can be seen that the average data message length is equal to 80 packets.

V. Simulation Results and Discussion

An extensive simulation study was carried out. The simulator was developed using the GNU C programming language. The platform used for the simulation runs was a SUN ULTRA. Each run simulated one hour of actual network activity (300005 channel frames). We simulated the system under all possible movie loads from 1 to 6 movies. For each load, all different combinations were examined (63 in total), from 5 times each (Monte Carlo method). The lack of similar work in the literature (i.e., proposed mechanisms for handling Video Streams of this quality, together with performance evaluation of such mechanisms) prevents us from making result comparisons.

Table 3.2 and Figure 3.1 show the results for each movie separately. "The Simpsons" turned out to be the most demanding stream in contrast with "Star Wars" which was the less demanding, although the highest channel throughput appears with the Soccer stream.

λ & Throughput results for Video

Movie		λ	Throughput
A	News	5.1	86.47266
B	Parking Lot	4.85	83.91547
C	Silence of the Lambs	4.78	79.1234
D	The Simpsons	3.1	68.67518
E	Soccer	5.05	92.85171
F	Star Wars	5.85	88.76075
Average		4.788333	83.29986

Table 3.2

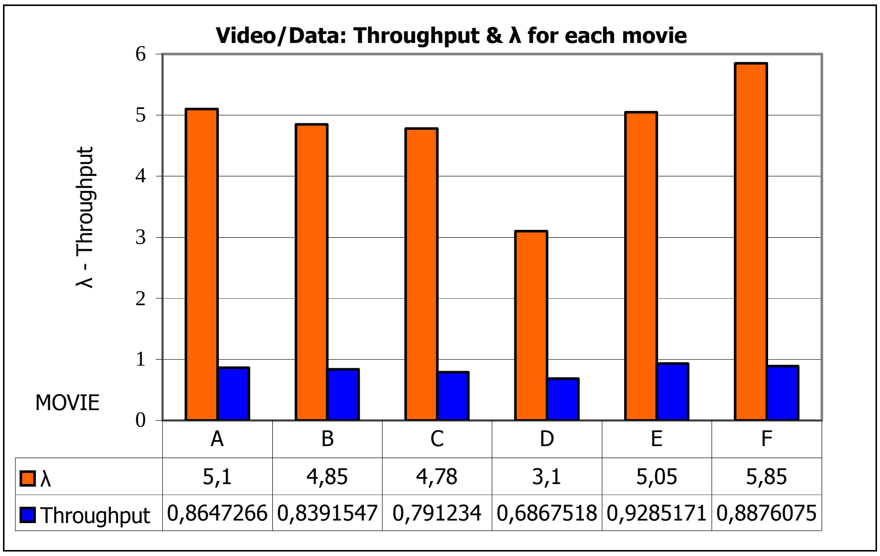

Figure 3.1

Figure 2 and Table 3 show the average results of the simulation runs under all possible movie loads. As the number of video terminals increases, the aggregate bit rate gets smoother, since the superposition of the movies is known from the literature to always be less bursty than a single movie, and the average bandwidth of the superposition is close to the sum of the mean bit rates of the movies. This is proven once more in our study, as the case of the system accommodating only one movie is shown to provide the smallest throughput than any other case of movie load.

Average λ & Throughput for Video/Data

System Load	Average λ	Average Throughput
1 Movie	4.788333	83.29986
2 Movies	3.8366	83.97877
3 Movies	2.96	85.90973
4 Movies	2.2625	91.91385
5 Movies	1.266667	97.34089
6 Movies	0.31	94.18723

Table 3.3

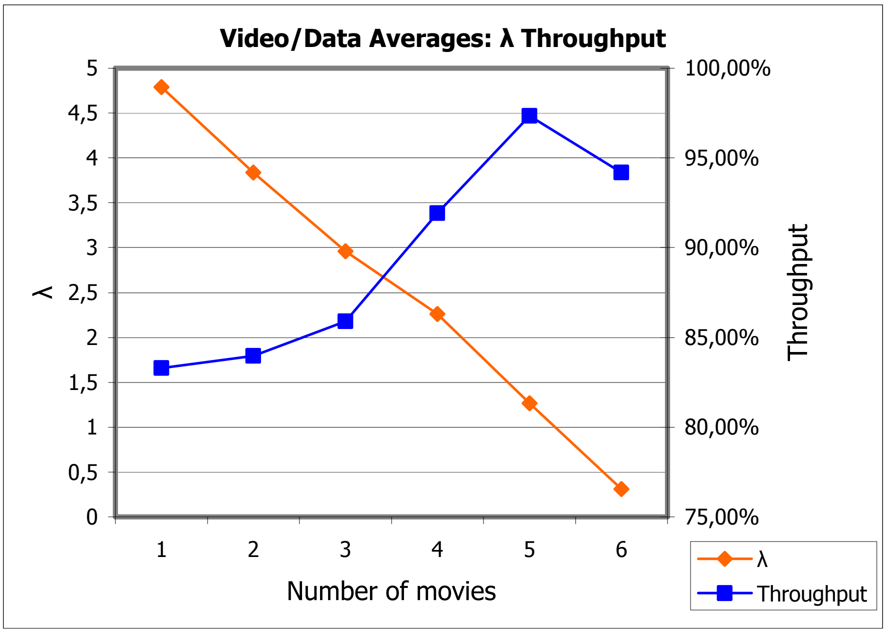

Figure 3.2

The λ parameter corresponds to the maximum data message arrival rate (expressed in messages per frame) sustained by the system (i.e. such that the video packet dropping probability <0.0001 and the average data message delay < 2 seconds).

Another comment that needs to be made is about the throughput of the system when there are 5 active video terminals. This throughput is remarkably high (97,3%), and the reason is that in this case the channel gets filled up with video packets but there is still some empty space for data packets, which almost ideally fill the channel.

On the contrary, when six movies are present in the system, the channel throughput decreases. The reason for this result can be explained if we sum up the peak bit rates of the six movies. The total peak rate reaches 25 Mbps, which is significantly higher than the total channel capacity. Of course, the multiplexing gain is also significant and it is rather rare that all of the 6 video terminals would transmit at their peak bit rates at the same time but still their superposition surpasses a few times (i.e., for a few video frames) the channel capacity of 20 Mbps. This forces the system to experience, in these video frames, severe video packet dropping, because of the very strict video dropping probability requirements, and therefore leads to the decrease in throughput in comparison with the 5 active video terminals scenario.

A final comment concerns the average data messages delays of our mechanism (the corresponding results are not shown here). In all cases examined in this study, the average data messages delays were found to be roughly constant and approximately equal to 960 msecs (80 packets per message * 12 msecs per frame).

This can be explained as follows. In the case of one Video Terminal being active in the system, the maximum data message arrival rate sustained by the system is on average 4.79 messages per frame and the 20 minislots available per frame for resolving the collisions among the contending data requests are enough. Most of the data request packet collisions are resolved within one frame. The corresponding data terminals receive slot reservations in the same frame with high probability, since the utilization of the frame is roughly equal to 83% and each Data Terminal can be allocated only one reserved slot per frame. Similarly, in the case of six active Video Terminals, the maximum data message arrival rate sustained, is equal to 0.31 messages per frame, and the 20 request minislots are sufficient for handling the new video request packets (at most six in each frame, if all Video terminals undergo a bit rate increase at the same time which happens rarely) and the new data message arrivals, which correspond to one new arrival per three frames on average.

VII. Conclusions

In this chapter we proposed a mechanism for transmitting high quality stored multimedia streams and data traffic, over a wireless channel of high capacity. We evaluated its performance through an extensive simulation study in which we used actual MPEG-4 stream traces together with data generated by a synthetic model fitted to measurements of data traffic from a National University and Research Network (FUNET). The results of our study show that the proposed mechanism achieves high aggregate channel throughput in all cases of traffic load, while preserving the Quality of Service (QoS) requirements of each traffic type.

Chapter 4

Summary and Contribution of this Work & Ideas for Future Work

I. Summary and Contribution

In this work we developed mechanisms for the integrated delivery of different traffic types that vary from High Quality Video Streams to Data Traffic over a Time Division Wireless Cellular High Speed Packet Switched Network. The aim is to design efficient MAC mechanisms in order to satisfy the diverse nature of the traffic types that the network must accommodate and the contradictory QoS requirements of each traffic type. In all cases we focused on the uplink channel (mobiles to the Base Station).

In Chapter 1, we presented the technological environment in which we must adapt and the restraints and obstacles which need to be considered in order to understand the problems and challenges that we must face. We reviewed technologies like ATM, and MPEG-4 in order for the reader of the Thesis to obtain an understanding of the system's framework and we discussed the need for new MAC schemes that adapt to the above technological platforms.

In the second Chapter we closely examined and extended approaches that have been made during our previous work (not reported in the Thesis). An enriched physical medium was introduced, on which videoconference traffic stream transmissions integrated with voice and data traffic transmissions was carried out. Actual videoconference traces were used in order to evaluate the performance of our schemes. Beyond the adaptation of the MAC scheme, our contribution extends to the Base Station scheduling mechanism. The combination of these two approaches contributes to the achievement of the desirable QoS metrics of the various traffic types.

In Chapter 3, the MAC scheme is properly modified in order to achieve the integration of two types of traffic (Video and Data). In this scenario however the Video traffic was generated by High Performance Mobile devices that were streaming multimedia content to other MT's over the wireless network infrastructure. More over those mobile devices are assumed to act like cache engines in order to achieve lower client startup latencies and reduce the traffic load from the Base Station of the wireless network to the outer network. No traffic shaping was used for reasons explained in that Chapter.

We believe that the channel access and transmission scheduling mechanisms we have designed and evaluated in this Thesis, contribute to the state-of-the-art and can help in future efforts towards the integrated transmission of video, voice and data traffic in high capacity wireless networks.

II. Ideas for Future Work

Given the flexibility that characterizes the protocols we have introduced, only mild modifications are needed in order to adapt them into the needs of various scenarios.

The performance evaluation of the proposed mechanisms when used over a wireless error prone channel is one of the first cases that must be examined in our future work.

Further investigation is also required in the case that we are using different media encoding techniques together with less strict QoS requirements for the time sensitive traffic types.

Finally, embedding the proposed mechanisms in a larger system framework is a matter deserving further attention and study. Such a case could be when trying to use the proposed MAC and BS scheduling schemes together with a content placement scheme throughout the network (e.g. assuming that cache farms are attached to the BS's or dispersed throughout the network) and with a mobility prediction mechanism, in order to achieve an integrated wireless environment in which the end user would be able to experience delivery of high quality multimedia content while on the move.

References

[1] Marconi Plc http://www.marconi.com

[2] GSM World http://www.gsmworld.com

[3] The UMTS Forum http://www.umts-forum.org

[4] The ATM Forum http://www.atmforum.com

[8] A. Basso, S. Varakliotis, R. Castagno*, "Transport of MPEG-4 over IP/RTP"*, in Proceedings of the PV Workshop, 2000.

[9] Leonardo Chiariglione,"*MPEG-4, why use it?"* Telecom Italia Lab–Italy, *http://leonardo.telecomitalialab.com/paper/mpeg-4/*

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

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

[12] 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.

[13] 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.

[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. 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.

[16] 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", in Proceedings of the IEEE Personal Indoor Mobile Radio Communications (PIMRC) 2001 Conference, San Diego, USA.

[17] 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", in Proceedings of the IEEE PIMRC 2001 Conference, San Diego, USA.

[18] N.Passas D.Skyrianoglou and L.Merakos, *"Traffic Scheduling in Wireless ATM Networks",* in Proceedings of the IEEE ATM'97 Workshop, Lisbon, Portugal, May 1997.

[19] C. Papadimitriou, S. Ramanathan, P.V. Rangan and S. SampathKumar, "Multimedia information caching for personalized video-on-demand", Comput. Commun., vol. 18, pp. 204-216, Mar. 1995.

[20] Y.Guo, S.Sen, and D.Towsley "Prefix Caching assisted Periodic Broadcast: Framework and Techniques to Support Streaming for Popular Videos", in Proceedings of the 2002 IEEE Intl. Conf. on Communications (ICC).

[21] C. Vassilakis, M. Paterakis and P. Triantafillou: "Video Placement and Configuration of Distributed Video Servers on Cable TV Networks", ACM Multimedia Systems Journal, Vol. 8, No. 2, 2000, pp. 92 - 104.

[22] 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.

[23] D.A. Dyson and Z.J. Haas, *"A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM"*, ACM / Kluwer Mobile Networks (MONET) Journal, 4(2), pp. 87-89, Jan 1999.

[24] 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.

[25] Q. Pang, A. Bigloo, et al., *"Service Scheduling for General Packet Radio Service Classes"*, in Proc. of the 1999 IEEE Wireless Communications and Networking Conference (WCNC '99), New Orleans, USA.

[26] http://peach.eas.asu.edu/index.html

[27] M. Paterakis and P. Papantoni – Kazakos, "A Simple Window Random Access Algorithm With Advantageous Properties", IEEE Trans. on Inform. Theory, Vol. IT-35, No. 5, September 1989, pp. 1124-1130.

[28] 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.

[29] P. Koutsakis, "Medium Access Control and Traffic Policing for Multimedia Traffic Integrated Access in High Capacity Wireless Channels", PhD Dissertation ECE Dept, Technical University of Crete, Chania 2002

[30] http://www.ericsson.com/edge/

[31] http://*www.qualcomm.com*

[32] D.C. Cox, "Wireless Personal Communications: What is it?", IEEE Personal Communications Vol. 2, Apr. 1995, pp. 20-35

[33] P.T. Brady, "A Technique for Investigating On - Off Patterns of Speech", Bell Sys. Tech Journal, Jan 1965.

[34] 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.

[35] Travis Russell, *"Telecommunications Protocols"*, 1997 McGraw-Hill Telecommunications, pages 372-386.