



TECHNICAL UNIVERSITY OF CRETE
ELECTRONIC & COMPUTER ENGINEERING DEPARTMENT

THESIS

***Scheduling for efficient telemedicine traffic
transmission over Next Generation Cellular
networks***

Submitted by: Kouskouli Theodora

Advisor: Koutsakis Polychronis

CHANIA, APRIL 2012

Abstract

The ability to adjust the allocated bandwidth of ongoing calls to cope with wireless network resource fluctuations is becoming increasingly important in cellular networks. In this thesis, we propose a Multiple Access Control (MAC) protocol, for the integrated transmission of regular and telemedicine traffic transmission over next generation cellular networks.

Telemedicine traffic transmission has gained in importance during the past few years. Due to the fact that it carries critical information regarding the patients' condition, the expedited and errorless transmission of multimedia telemedicine traffic is of fundamental importance. The prioritized or guaranteed transmission of telemedicine traffic, however, can lead to the violation of the Quality of Service (QoS) requirements of regular traffic users and to the loss of guaranteed bandwidth in cases when it is left unused, due to the infrequent nature of telemedicine traffic. To resolve these problems, we propose an adaptive bandwidth reservation scheme based on road map information and on user mobility and a fair scheduling scheme for telemedicine traffic transmission over cellular networks. The combination of the two schemes achieves high channel bandwidth utilization while offering full priority to telemedicine traffic over regular traffic.

More specifically, the joint transmission of voice, real-time video, electrocardiogram signals, and medical scans in a cellular multi-user environment is considered, taking into account that real-time delivery of medical videos, images etc. requires high level of quality of service as little tolerance to loss and delay is acceptable. How to keep the handoff packet dropping probability under a prespecified upper limit is a very important QoS issue in cellular networks because mobile users should be able to maintain

ongoing sessions even during their hand-off from one cell to another. Our protocol prioritizes hand-off over non-handoff users while managing to achieve fairness among users of the same type of traffic; three scheduling approaches are used in our work and we compare their efficiency in terms of fairness. Our simulations results demonstrate the advantages of our protocol for emergency care traffic transmission from high-speed moving ambulance vehicles to hospitals.

Acknowledgements

I would first like to thank my advisor, Assistant Professor Koutsakis Polychronis, for giving me knowledge, understanding and strength to complete this study. He has been always there through all the course of this study, so I thank him for his help, guidance and words of encouragement.

I would like to thank Professors Aggelos Bletsas and Michael Paterakis for their participation in the examination committee.

In addition, many thanks to Thomas Hall, Panagiotis Kotsarinis, Annalia Kouloktsi, Xristina Lionoudaki, Vaggelis Maratsolas, Thodwris Mathioudakis and Eleni Tseva for their important contribution to the completion of this thesis.

Special thanks also go to my family and my friends for their help and support. Their patience and encouragement gave me the necessary motivation to complete this study.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	v
Table of Contents	vi
List of Figures	ix
List of Tables	x
List of Abbreviations	xi
1 Introduction	1
1.1 Overview	1
1.2 Contribution of this work	3
1.3 Organization of the Thesis	5
2 Description of System Model	6
2.1 Regular Multimedia Traffic	6
2.2 Telemedicine Traffic	10
3 Multimedia Integration Multiple Access Control (MI-MAC)	13
3.1 Channel Frame Structure	13
3.2 Base Station Scheduling and Actions of Terminals	15
3.3 Two-Cell Stack algorithms	17
3.3.1 Two-cell stack reservation random access algorithm	18
3.3.2 Two-cell stack blocked access collision resolution algorithm	19
3.4 System state transitions	19
3.5 Channel Error Model	21

4	<u>The Proposed Scheduling Scheme</u>	24
4.1	<u>Introduction of the scheduling ideas</u>	24
4.2	<u>Scheduling Priorities and Contention Resolution</u>	28
4.3	<u>Fair Scheduling</u>	31
4.4	<u>FCFS – EDF – SJF Algorithms</u>	33
4.5	<u>Jain’s Fairness Index</u>	35
5	<u>Adaptive Bandwidth Reservation based on Mobility and Road Information</u>	36
5.1	<u>Overview</u>	36
5.2	<u>Network and Mobility Models</u>	37
5.2.1	<u>The Mobility Model in [42]</u>	38
5.3	<u>Adaptive Bandwidth Reservation Scheme</u>	41
6	<u>Results and Discussion</u>	46
6.1	<u>Simulation Setup</u>	46
6.2	<u>Results</u>	46
6.2.1	<u>FCFS results</u>	46
6.2.2	<u>FCFS results using a different map</u>	58
6.2.3	<u>FCFS-EDF-SJF results</u>	62
6.2.4	<u>Open cells results</u>	68
6.2.5	<u>Jain’s Fairness results</u>	70
7	<u>Conclusion and Future Work</u>	80
	<u>Bibliography</u>	81

LIST OF FIGURES

2.1	<u>The voice source activity model</u>	7
3.1	<u>Channel Frame Structure</u>	14
3.2	<u>State transition diagram for an active terminal</u>	20
3.3	<u>Channel Error Model</u>	22
5.1	<u>Road Map and cellular network model</u>	37
5.2	<u>Road Map and cellular network model</u>	39
6.1	<u>Effect of regular video traffic on X-Ray traffic</u>	48
6.2	<u>Percentage of regular video users who experience packet loss larger than 1%, versus the number of regular video users</u>	49
6.3	<u>Regular video packet dropping vs. the number of regular video users</u>	50
6.4	<u>Effect of regular voice traffic on telemedicine video traffic</u>	51
6.5	<u>Effect of regular voice traffic on telemedicine image traffic</u>	52
6.6	<u>Effect of telemedicine video traffic on regular video traffic</u>	53
6.7	<u>Percentage of telemedicine video users who experience packet loss larger than 0.01%, versus the number of telemedicine video</u>	55
6.8	<u>Telemedicine video packet dropping versus the number of telemedicine video users</u>	55
6.9	<u>Effect of telemedicine video traffic on X-ray traffic</u>	56
6.10	<u>Effect of telemedicine video traffic on telemedicine image traffic</u>	57
6.11	<u>Improvement on handoff voice packet dropping probability with the use of adaptive bandwidth reservation</u>	57
6.12	<u>Effect of regular video traffic on X-Ray traffic</u>	59
6.13	<u>Effect of regular voice traffic on telemedicine video traffic</u>	60
6.14	<u>Effect of telemedicine video traffic on regular video traffic</u>	61

6.15	<u>Telemedicine video packet dropping versus the number of telemedicine video users</u>	61
6.16	<u>Regular video packet dropping vs. the number of regular video users</u>	63
6.17	<u>Effect of regular voice traffic on telemedicine video traffic</u>	64
6.18	<u>Effect of telemedicine video traffic on regular video traffic</u>	65
6.19	<u>Telemedicine video packet dropping versus the number of telemedicine video users</u>	66
6.20	<u>Effect of telemedicine video traffic on X-ray traffic</u>	67
6.21	<u>Effect of telemedicine video traffic on regular video traffic</u>	68
6.22	<u>Telemedicine video packet dropping versus the number of telemedicine video users</u>	69

LIST OF TABLES

3.1	Channel Error Parameters	23
6.1:	9% traffic load	71
6.2:	15% traffic load	71
6.3:	43% traffic load	71
6.4:	51% traffic load	72
6.5:	58% traffic load	72
6.6:	79% traffic load	72
6.7:	83% traffic load	73
6.8:	88% traffic load	73
6.9:	92% traffic load	73
6.10:	100% traffic load	74
6.11:	9% traffic load	74
6.12:	15% traffic load	75
6.13:	43% traffic load	75
6.14:	51% traffic load	75
6.15:	79% traffic load	76
6.16:	83% traffic load	76
6.17:	88% traffic load	76
6.18:	92% traffic load	77
6.19:	100% traffic load	77
6.20:	60% traffic load	78
6.21:	80% traffic load	78
6.22:	110% traffic load	78

LIST OF ABBREVIATIONS

3G	Third Generation Communication Systems
4G	Fourth Generation Communication Systems
BS	Base Station
CBR	Constant Bit Rate
CRP	Collision Resolution Period
ECG	Electro-Cardiogram
EDF	Earliest Deadline First
FCFS	First Come First Served
GSM	Global System for Mobile communications
GPS	Global Positioning System
LAN	Local-area Network
MI-MAC	Multimedia Integrated - Medium Access Control Protocol
MSC	Mobile Switching Center
QoS	Quality of Services
SJF	Shortest Job First
SNR	Signal-to-noise ratio
TDMA-FDD	Time Division Multiple Access with Frequency Division Duplex
VAD	Voice Activity Detector
VBR	Variable Bit Rate
VF	Video Frame
WCDMA	Wideband Code Division Multiple Access

Chapter 1:Introduction

1.1 Overview

Telemedicine has been defined as using telecommunications to provide medical information and services [1], and is already being employed in many areas of healthcare, such as intensive neonatology, critical surgery, pharmacy, public health and patient education. Mobile networks have brought about new possibilities in the field of telemedicine due to the wide coverage provided by cellular networks and their capability of providing service to moving vehicles. In car accidents, pre-hospital teams provide on-scene initial assessment and resuscitation and transmit this information to a physician mainly via voice communication; therefore the physician can only make an assessment based on what is described and is unable to continuously monitor an injured victim through visual communication (e.g., video), while at the same time receiving data of major importance regarding the victim's vital signs [2, 3]. Given that the patients' arrival to the hospital can be delayed for a variety of reasons, mainly having to do with traffic congestion, it is clear, that ambulances need to be able to provide the patient with health services that would only have been possible, until recently, when the patient would be admitted to the hospital.

The ultimate goal for all telemedicine applications is to improve the well-being of patients and bring medical expertise fast and at low cost to people in need [4, 5]. Thus, in addition to ambulance vehicles, it is also of critical importance for the provision of health care services at understaffed areas like

rural health centers, ships, trains, airplanes, as well as home monitoring (e.g., middle-aged people in the future will have to be monitored at home as it will be impossible for everybody to stay at the hospital) [4,6]. Mobile healthcare (M-health, “mobile computing, medical sensor and communication technologies for healthcare” [8]) is a new paradigm that brings together the evolution of emerging wireless communications and network technologies with the concept of “connected healthcare” anytime and anywhere. Various M-health studies have been conducted within the last few years, on very significant aspects of public health [2, 3, 6, 8, 9, 10, 11, 12]. In many of these studies, the efficient use of the cellular network resources was of paramount importance for the correct and rapid transmission of all types of telemedicine traffic (video, audio and data), especially since only a limited uplink (and in some cases downlink) bandwidth could be devoted to the transmission of this very urgent and crucial, for the patients’ life, type of traffic. Also, some of the related research projects include the development of low-cost, low-power, multifunctional medical micro sensor nodes [47], a low power Medium Access Control protocol for wireless medical sensor networks [43], the transmission of medical and context-aware data from mobile patients to healthcare centers over heterogeneous wireless networks [48], QoS provisioning over the wireless channel between the Body Sensor Network (BSN) Gateway and the wireless Access Points (AP)[46] and methods for providing rural telemedicine with quality video transmission [45]. In addition, various applications have been introduced, e.g., the mobile tele-echography robotic application over WCDMA [7], the mobile Tele-Ultrasonography in M-health over 3G networks [8], the Mobile Teletrauma System presented in [2] using CDMA over 3G and a handheld device, called personal wireless hub (PWH), for each mobile

patient. PWH first gathers and aggregates the vital signs and context-aware data for various telemedicine applications. PWH transmits the aggregated data to the remote healthcare center over multiple wireless interfaces such as cellular, WLAN, and WiMAX.

As pointed out in the above-mentioned telemedicine research efforts, beyond 3G wireless networks can be a sufficient test bed for the development of efficient telemedicine traffic transmission mechanisms, due to the much higher channel rate they are expected to provide in comparison to previous generation's cellular networks. Next generation cellular networks will be able to provide voice, data and streamed multimedia to users on an “anytime, anywhere” basis. Although there is no formal definition for what 4G (Fourth-Generation Communications System) will be, its commonly assumed objective is that it will be a fully IP-based integrated system. This will be achieved after wired and wireless technologies converge and will be capable of providing very high data rates both indoors and outdoors, with premium quality and high security.

1.2 Contribution of this work

Mobile telemedicine is one of the most exciting technologies today for improving public health conditions. Many studies have been conducted in order to optimize aspects of mobile telemedicine [[2](#), [3](#), [6](#), [8](#), [9](#), [10](#), [11](#), [12](#)]. One common characteristic of these studies is that they focus solely on the transmission of telemedicine traffic over the cellular network, without taking into account the fact that regular traffic, which represents the bulk of traffic in the network, has

strict Quality of Service (QoS) requirements as well. Also, in many of these studies, despite the importance of the systems used, the accuracy in information transmission was low (e.g., [11]). Hence, our thesis aims at proposing a resource allocation mechanism which will enable the efficient integration of telemedicine traffic with regular traffic over next generation cellular networks, with highest priority offered to telemedicine traffic.

In previous work [13], MI-MAC (Multimedia Integration Multiple Access Control Protocol) was introduced, which was shown to achieve superior performance in comparison to other TDMA and WCDMA-based protocols of the literature when integrating various types of multimedia traffic (video, voice, email and web data) over next generation cellular networks. MI-MAC does not consider the transmission of urgent traffic, like telemedicine traffic. In MI-MAC, within each priority class the queuing discipline is assumed to be First Come First Served (FCFS). Hence the average performance evaluation metrics will give no insight on the QoS of each individual wireless subscriber; therefore, it could be the case that certain users have their QoS severely violated while others get exceptional QoS, which would give an acceptable average QoS over the total number of users. This approach is unfair to users who arrive later in the network and hence are placed at the bottom of the Base Station (BS) service queue; the problem is especially significant in the case of video users, where early arriving users may dominate the channel by being allocated large numbers of slots, allowing just a small number of resources to be available for users arriving later. For this reason, we introduce a fair bandwidth allocation scheme in MI-MAC, as well as a scheduling scheme in order to improve the protocol's performance and reserve the proper bandwidth among cells for handoff traffic, and especially, for handoff telemedicine traffic.

Individual priorities are also set among the various types of mobile telemedicine traffic, based on their current importance for medical health care. Our work extends the recent work in [42] by adapting the bandwidth reservation scheme, (based on road map information and on user mobility), which was proposed in that work, for a different cellular architecture and different scheduling parameters. We, also, implement two more queuing priority algorithms and we compare their fairness against FCFS, which was the only queuing priority algorithm used in [42]. Jain's fairness index is used in our work to conduct the comparison.

1.3 Organization of the Thesis

The rest of the thesis is organized as follows. Chapter 2 gives a basic description of the telemedicine and regular traffic types and models used in our work. In Chapter 3, we provide a brief overview of MI-MAC protocol which is used as the basis of our work. Chapter 4 presents our proposed improved and fair scheduling scheme which we add on MI-MAC. In Chapter 5 we discuss our proposal for an adaptive bandwidth reservation scheme, which enables the system to exploit its knowledge of the users' mobility patterns. Chapter 6 includes our simulation results and a discussion on them, and in Chapter 7 we present our conclusions and ideas for future work.

Chapter 2: Description of System Model

In our work we consider the integration of regular and telemedicine traffic transmission. We consider voice, video, e-mail and short message service (sms) as representatives of regular traffic, and electro-cardiograph (ECG), X-ray, video and high-resolution medical still images for telemedicine traffic, in order to study the practical scenario of many different types of users simultaneously attempting to access the network and hence aggravating the access for telemedicine users. We describe below the characteristics of each traffic type.

2.1. Regular Multimedia Traffic

Four types of “regular” multimedia traffic are considered in our thesis. MPEG-4 video-conference, voice, email and mobile text messages (sms), which are the most common traffic types in cellular networks.

- Voice: The speech codec rate is 32 kb/s, and voice terminals are equipped with a voice activity detector (VAD) [15]. 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 (Figure 2.1). The mean talkspurt duration is 1 s and the mean silence duration is 1.35 s. The talkspurt to silence transition probability is P_{ST} and the silence to talkspurt transition probability is P_{TS} . The talkspurt and silence periods are geometrically distributed with mean $1/P_{TS}$ and $1/P_{ST}$ frames, respectively. Therefore, at steady state, the probability

that a terminal is in talkspurt (speech activity), P_T , or silence, P_S , is obtained from the following equations:

$$P_T = P_{ST} / (P_{ST} + P_{TS}) \quad (2.1)$$

$$P_S = 1 - P_T \quad (2.2)$$

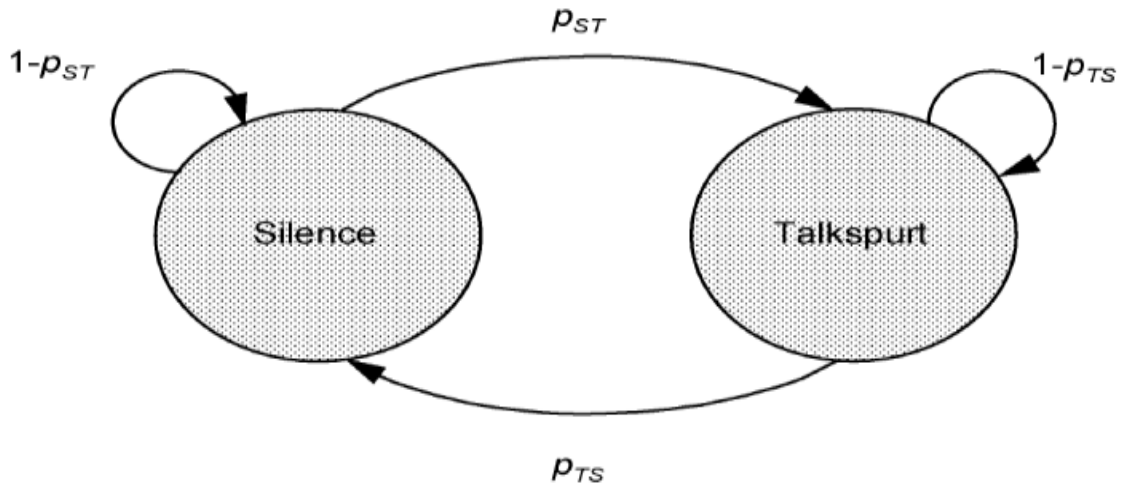


Figure 2.1: The voice source activity model

The number of active voice terminals N in the system is assumed to be constant over the period of interest. 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 s. Reserved slots are deallocated immediately. This implies that a voice terminal holding a reservation signals the BS upon the completion of its talkspurt (the same assumption is made for slots

reserved by data and video terminals). The allowed voice packet dropping probability is set to 1%, and the maximum transmission delay for voice packets is set to 40 ms.

- Email: We adopt the data traffic model based on statistics collected on e-mail usage from the Finnish University and Research Network (FUNET) [14]. The probability distribution function $f(x)$ for the length of the e-mail data messages of this model was found to be well approximated by the Cauchy (0.8, 1) distribution. The packet interarrival time distribution for the FUNET model is exponential, and the average e-mail data message length is 80 packets. A quite strict (considering the nature of this type of traffic) upper bound is set on the average e-mail transmission delay, equal to 5 s. The reason for this strict bound is that mobile users sending emails will be quite demanding in their QoS requirements, as they will expect service times similar to those of short message service traffic.
- SMS: Short Message Service (SMS) is a store-and-forward service that relies on a Short Message Service Center (SMSC). SMS messages are especially suitable for the transmission of small data bulks and for transmissions repeating in long time intervals (minutes to hours). The SMS payload is 140 bytes (including a header of 13 bytes) [16]. The message inter-arrival time distribution is considered exponential. Estimations of GSM networks' SMS transmission delays refer to delays of 2-30 s [17]. In this work, in order to test our system under the strictest QoS requirements, we set an upper bound of 2 s in SMS transmission delay.

- Regular MPEG-4 Video Streams: The MPEG 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 bit rate applications but were subsequently expanded to include a wider range of multimedia applications and bit rates. 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, coding arbitrarily shaped video objects, and content-based interactivity such as the ability to randomly access and manipulate objects in a video scene. In our study, we use the trace statistics of actual MPEG-4 streams from the publicly available library of frame size traces of long MPEG-4 and H.263 encoded videos provided by the Telecommunication Networks Group at the Technical University of Berlin [18]. The two video streams used in our study have been extracted and analyzed from a camera showing the events happening within an office and a camera showing a lecture, respectively. We have

used the high quality version of the videos: one has a mean bit rate of 400 kb/s, a peak rate of 2Mb/s, and a standard deviation of 434 kb/s, and the other one has a mean rate of 210 kbps, peak rate of 1.5 Mbps and standard derivation of 182 kbps. New video frames (VFs) arrive every 40 ms. We have set the maximum transmission delay for video packets to 40 ms, 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. This strict requirement is necessary, because streaming video requires bounded end-to-end delay so that packets arrive at the receiver in a timely fashion to correctly be decoded and displayed. If a video packet does not arrive on time, the play out process will pause, which is annoying to the human eye. The allowed video packet dropping probability is set to 1% [19], as the loss of regular video packets is not of equally critical importance as that of telemedicine video packets for which the maximum allowed video packet dropping probability is 0.01%, as it will be explained in Section 2.2. The two video traces are chosen with equal probability by regular video users. Video sources have exponentially distributed sessions with a mean duration of five minutes (this duration has been denoted by global trials as an expected one for users of another wireless cellular video application) [29].

2.2 Telemedicine Traffic

Four types of telemedicine traffic are considered in our work: Electro-Cardiograph (ECG), X-ray, video and high-resolution medical still images.

- Electro-Cardiograph (ECG): Similarly to [2, 12] which use data from the MIT-BIH arrhythmia database, we consider that ECG data is sampled at 360 Hz with 11 bits/sample precision. The arrival rate of ECG users is set to be λ_E user/frame following a Poisson distribution. The transmission of ECG traffic should be rapid and lossless, due to the critical nature of the data; additionally, we have set a strict upper bound of just 1 channel frame (12 ms) for the transmission delay of an ECG packet.

- X-Ray: We consider that a typical X-ray file size is 200 Kbytes [6] and that the aggregate X-Ray file arrivals are Poisson distributed with mean λ_X files/frame. The upper bound for the transmission delay of an X-Ray file, which again needs to be lossless, is set to 1 minute. A discussion on the strictness of this bound will be made in Chapter 4.

- Medical Images: Medical image files have sizes ranging from 15 to 20 Kbytes/image [2] and are Poisson distributed with mean λ_I files/frame. The upper bound for the transmission delay of a medical image is set to 5 seconds (this bound is much stricter than those used in [2, 6, 10, 12]), and the transmission needs to be lossless.

- Telemedicine Video: Since H.263 is the most widely used video encoding scheme for telemedicine video today, we use in our simulations real H.263 video-conference traces from [18] with mean bit rate of 91 Kbps, peak rate of 500 Kbps and standard deviation of 32.7 Kbps. The video frames arrive with constant rate (every 80 ms) with variable frame sizes. We have set the

maximum transmission delay for video packets to 80 ms, with packets being dropped when this deadline is reached; i.e., all packets of a video frame must be delivered before the next video frame arrives. Due to the need for very high-quality telemedicine video, the maximum allowed video packet dropping probability is set to 0.01%.

It needs to be mentioned that existing work in the field assumes much looser QoS requirements in order for the network to be able to meet them. For example, in [3] the upper delay bound for the transmission of an ECG packet is 1 second, and the upper bound on voice packet dropping is set to 3%.

Chapter 3: Multimedia Integration

Access Control (MI-MAC)

The Multimedia Integration Multiple Access Control (MI-MAC) protocol, introduced in [13] and based on Time Division Multiple Access with Frequency Division Duplex (TDMA-FDD), is one of the first works in the relevant literature for wireless picocellular networks that efficiently integrates voice (Constant Bit Rate, CBR, On/Off Traffic), bursty email, and sms traffic with either MPEG-4 or H.263 video streams (Variable Bit Rate, VBR) in high capacity picocellular systems with burst-error characteristics. The protocol was shown to be a good candidate for next generation cellular networks, as it outperformed (in simulation results and conceptually) other TDMA and Wideband Code Division Multiple Access (WCDMA)-based protocols when evaluated over a wireless channel with burst-error characteristics.

3.1 Channel Frame Structure

Our work focuses on the uplink (wireless terminals to Base Station) channel. The uplink channel time is divided into time frames of fixed length. The frame duration (12 ms accommodating 566 slots) is selected such that a voice terminal in talkspurt generates exactly one packet per frame. The packet size is

considered to be equal to 53 bytes, 48 of which contain information. This choice for the packet size was used in MI-MAC in order to ease the comparison with other protocols, and we have kept it, as the packet size does not influence the efficiency of our scheduling scheme. As shown in Fig. [3.1](#), which presents the channel frame structure, each frame consists of two types of intervals. These are the request interval and the information interval.

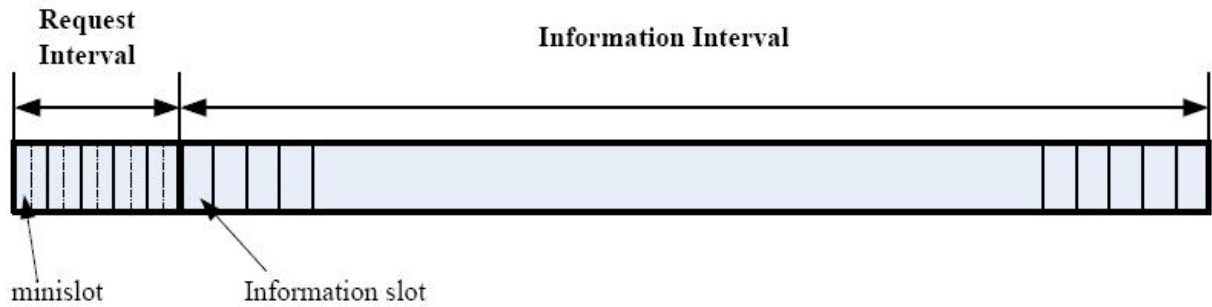


Figure 3.1: Channel Frame Structure

By using more than one minislot per request slot, a more efficient usage of the available request bandwidth (in which users contend for channel access) is possible. Our thesis considers a 20 Mbps channel as a conservative estimation, given that next generation cellular networks are planned to have transmission rates exceeding 20 Mbps. We chose the number of minislots per request slot to be equal to 2, to allow for guard time and synchronization overheads, for the transmission of a generic request packet and for the propagation delay within the picocell. Each minislot accommodates exactly one fixed-length request packet. Within an information interval, each slot accommodates exactly one fixed-length packet that contains voice, video, or data information and a header. Any free

information slot of the current channel frame can be temporarily used as an extra request (ER) slot to resolve the contention between requesting users. ER slots are again subdivided into two minislots. The function and operation of ER slots are exactly the same as those of the regular request slots.

3.2 Base Station Scheduling and Actions of Terminals

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 minislot, indicating only the presence or absence of a collision within the minislot [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.

To resolve contention among all requesting users, different priorities were assigned to different types of users. The four types of telemedicine traffic are transmitted first; their priority order will be explained in Section 4. The four types of regular traffic follow, with priority order: video, voice, email, and sms. The above prioritization by isolating each type of traffic and letting it contend only with traffic of the same type is feasible due to the use of the two-cell stack reservation random access algorithm (by video and voice terminals) and the two-

cell stack blocked access collision resolution algorithm [20] (by email and sms terminals) to resolve contention (this algorithm is of window type, with FCFS-like service, and will be discussed in Section 3.3). Apart from their operational simplicity, stability and relatively high throughput when compared to the PRMA [21] and PRMA-like algorithms [22], the stack algorithms have the additional advantage of offering a clear indication of when contention has ended (for two-cell stack this happens when two consecutive non-collision signals are transmitted by the BS in the downlink). To allocate channel resources, the Base Station maintains a dynamic table of the active terminals within the picocell. Upon successful receipt of a voice or data request packet, the Base Station provides an acknowledgment and queues the request. The BS allocates channel resources at the end of the corresponding request interval.

Specifically, for a video terminal, if a full allocation is possible, which means that the number of idle information slots is larger than the number of requested slots, the BS assigns this user all its requested slots. If a full allocation is not possible, the BS grants to the video user as many of the requested slots as possible (partial allocation). For email and sms users, the BS allocates one slot per frame for each user. Voice terminals that have successfully transmitted their request packets do not acquire all the available information slots in the frame. If this happened, voice terminals would keep their dedicated slots for the whole duration of their talkspurt (on average, more than 80 channel frames), and thus video terminals would not find enough slots to transmit in; hence, the particularly strict video QoS requirements would be violated. Consequently, the BS allocates a slot to each requesting voice terminal with a probability p^* , as in [13]. The probability p^* for the allocation of slots to voice users varied according to the

video load. In this thesis, a near-optimal value of p^* (0.09) has been found through extensive simulations, which works well for all video loads examined. 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 when the resources needed to satisfy a voice request are unavailable. With the use of p^* , what we are actually achieving is to increase voice packet dropping to its maximum allowed limit, in order to facilitate the transmission of the much burstier video traffic. Within each priority class, the queuing discipline is assumed to be FCFS.

In addition, the BS “preempts” email and sms reservations in order to service video and voice requests. Thus, whenever new video or voice requests are received and every slot within the frame is reserved, the BS attempts to service them by canceling the appropriate number of reservations belonging to data (email/sms) 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 corresponding request queue.

3.3 Two-Cell Stack algorithms

As mentioned in Section 3.2, the contention among terminals within each type of traffic is resolved by the 2-cell stack algorithm [20]. This section provides a brief discussion of the two-cell stack reservation random access algorithm (used by video and voice terminals) and the two-cell stack blocked access collision resolution algorithm (used by email and sms terminals).

3.3.1 Two-cell stack reservation random access algorithm

Each terminal uses a counter, r , as follows:

1. At the start of every request interval the contending terminals initialize their counter, r , to 0 or 1 with probability $1/2$.
2. Contending terminals with $r = 0$ transmit into the first request slot. With x being the feedback for that transmission, the transitions are as follows:
 - if $x = \text{non-collision}$:
 - if $r = 0$, the request packet was transmitted successfully.
 - if $r = 1$, then $r = 0$.
 - if $x = \text{collision}$:
 - if $r = 0$, then reinitialize r to 0 or 1 each with probability $1/2$.
 - if $r = 1$, then $r = 1$.
3. Repeat step 2, until either two consecutive feedbacks indicating non-collision occur or the interval ends.

The operation of this protocol can be depicted by a two-cell stack, where in a given request mini-slot the bottom cell contains the transmitting terminals (those with $r = 0$), and the top cell contains the withholding terminals (those with $r = 1$). The especially attractive feature of this algorithm is that two consecutive non-collisions indicate that the stack is empty.

3.3.2 Two-cell stack blocked access collision resolution algorithm

To transmit data request packets, the data terminals follow the two-cell stack random blocked access collision resolution algorithm during consecutive data request intervals, due to its operational simplicity, stability and relatively high throughput.

A blocked access mechanism is established by the following first time transmission rule for newly generated data messages. Terminals with new message arrivals may not transmit during a collision resolution period (CRP). A CRP is defined as the interval of time that begins with an initial collision (if any) and ends with the successful transmission of all data request packets involved in that collision (or, if no collision occurred, ends with that mini-slot). In the first mini-slot following a CRP, all of the terminals whose message arrived within a prescribed allocation interval, of maximum length D , transmit with probability one. Terminals involved in a collision follow rules 2 ~ 3 in Section 3.3.1 and the conclusion of the CRP is identified by two consecutive feedbacks indicating non-collision.

3.4 System state transitions

As shown in figure [3.2](#), an active terminal is described as being in one of four states: silent, contender, queued, or reserved. A silent terminal has no packet to transmit and does not require channel resources. Once the terminal has information to transmit, it enters the contender state and remains there until it either successfully transmits a request packet or drops all of its packets (in the

case of video and voice terminals). Since the requests are queued at the BS, the terminal enters the queued state and remains there until it either receives a reservation or exits talkspurt. After receiving a reservation, the terminal enters the reserved state and transmits one (or more, in the case of video terminals) packet(s) per frame into its slot(s) until it exhausts its packets and returns to the silent state.

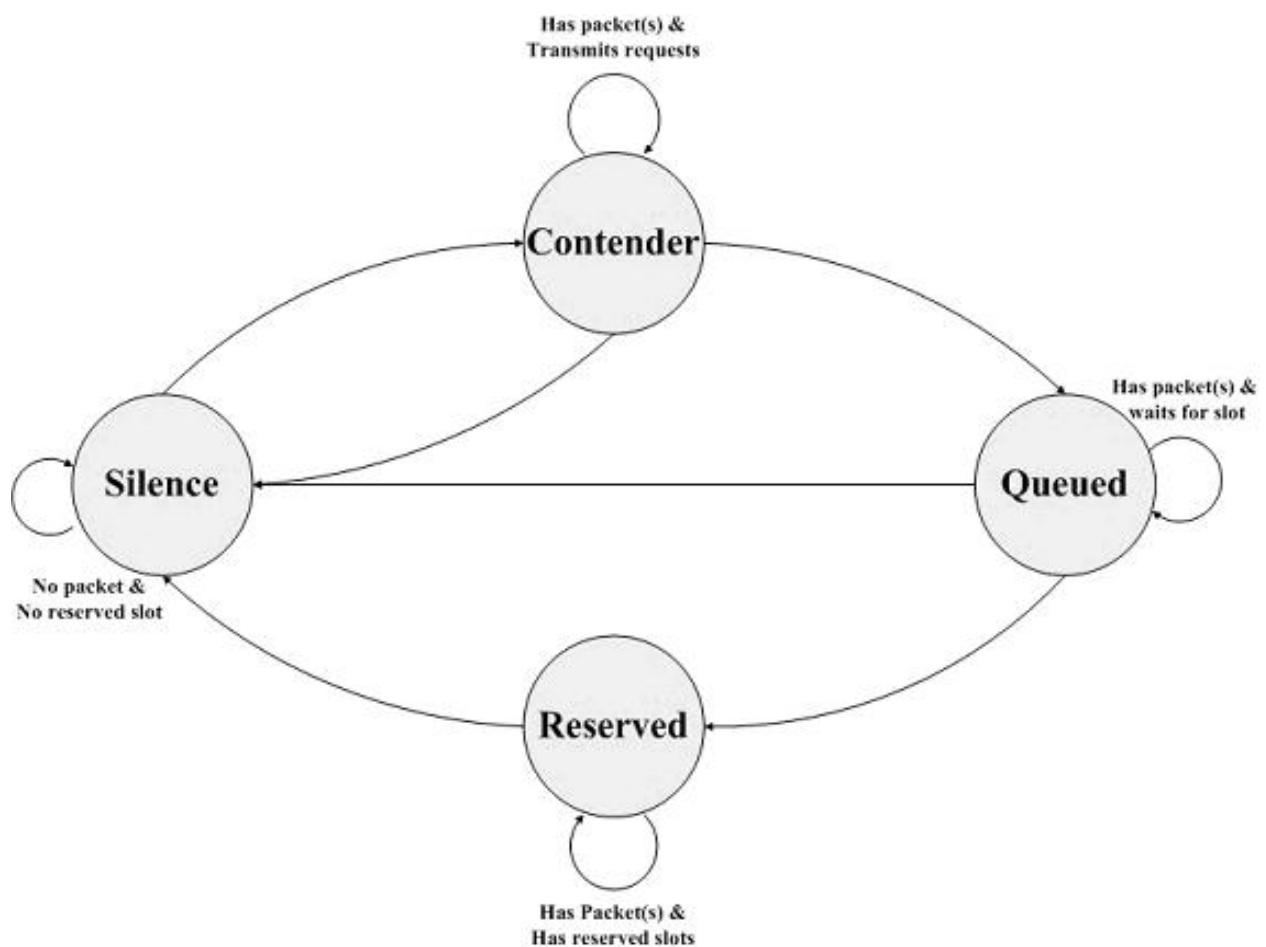


Figure 3.2: State transition diagram for an active terminal

3.5 Channel Error Model

The most widely adopted wireless channel error model in the literature is the Gilbert-Elliot model [23, 24]. The Gilbert-Elliot model is a two-state Markov model where the channel switches between a “good state” (always error-free) and a “bad state” (error-prone). However, many recent studies have shown that the Gilbert-Elliot model fails to predict performance measures depending on longer-term correlation of errors [25], minimizes channel capacity [26], and leads to a highly conservative allocation strategy [27].

A better choice for a more robust error model for wireless channels is the one we adopt in our study, and which was adopted in [13,28] . This model, with the use of the short and long error bursts, makes more accurate predictions of the long-term correlation of wireless channel errors than the Gilbert-Elliot model. The error model consists of a three-state discrete-time Markov chain, where one state is the “good state” (error-free) and the other two states are the “bad states”, the long bad and the short bad state, as we can see in Fig 3.3. A transmission is successful only if the channel is in the “good state” (G); otherwise, it fails. The difference between the long bad (LB) and short bad (SB) states is the time correlation of errors: LB corresponds to long bursts of errors, SB to short ones.

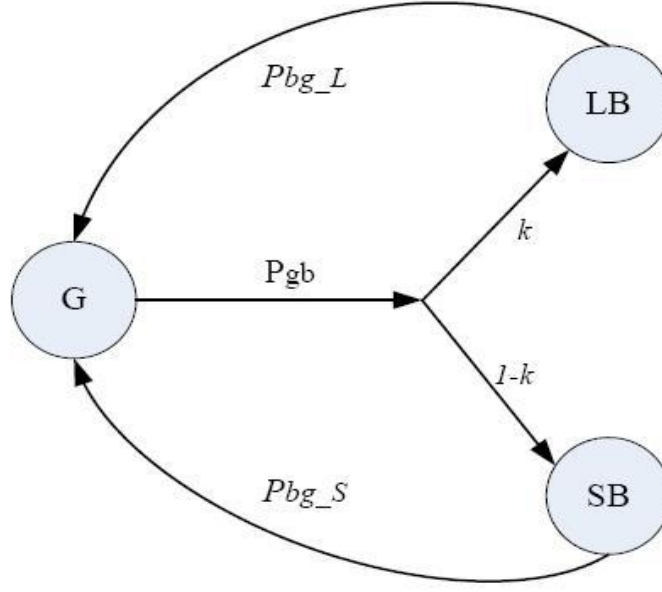


Figure 3.3: Channel Error Model

The parameters of the error model are presented in Table 3.1. The average number of error bursts, in slots, experienced when the states LB and SB are entered, are, respectively, given by $B_{LB} = 1/P_{bg_L}$ and $B_{SB} = 1/P_{bg_S}$, where P_{bg_S} is the transition probability from state SB to G, and P_{bg_L} is the transition probability from state LB to G. Similarly, the average number of consecutive error-free slots is given by $B_G = 1/P_{gb}$, where P_{gb} is the probability to leave state G. The parameter k is the probability that the Markov chain moves to state LB, given that it leaves state G; k also represents the probability that an error burst is long (i.e., the fraction of long bursts over the total number of error bursts). Similarly to [13], we have chosen the value of the probability P_{bad} , i.e., the steady-state probability that the channel is in a bad state, to be equal to $8 * 10^{-5}$; this value has been chosen in order to test an "almost worst" case scenario for our system, as the telemedicine

video packet dropping probability is set to 0.01% and, by choosing a value of bad state probability larger than the upper bound on telemedicine video packet dropping, the strict QoS requirement of telemedicine video users will certainly be violated. The values for P_{gb} and for the parameter k have been taken from [28], as well as the ratio between P_{bg_S} and P_{bg_L} . The value for P_{bg_L} is derived from the steady-state behavior of the Markov chain, for the bad state probability chosen.

$P_{good} = 0.99992$
$B_G = 1/P_{gb} = 65160 \text{ slots}$
$B_{SB} = 1/P_{bg_S} = 2.38 \text{ slots}$
$B_{LB} = 1/P_{bg_L} = 59.53 \text{ slots}$
$k = 0.5$

Table 3.1: Channel Error Parameters

Chapter 4: The Proposed

Scheduling Scheme

4.1 Introduction of the scheduling ideas

As explained in Chapter 3, we use the work on the MI-MAC protocol [13], as a basis for our scheduling scheme, and we significantly extend it to focus on the efficient scheduling for transmissions from mobile telemedicine users. Our work represents the next step in the relevant work of our research group in [42], with the differences explained in Section 1.2. Certain design limitations had been adopted in the study of [13] in order to facilitate the comparison of MI-MAC with other protocols in the literature:

1. Since the protocol was evaluated over one cell of the network, no traffic was considered to be arriving from other cells (handoff traffic).
2. Since video sources were assumed to “live” permanently in the system they did not have to contend for channel resources.
3. Since a picocellular wireless cellular architecture was assumed (picocell radius 10 ~ 50 m), the assumption was made that all users perceived the same uplink channel condition.

In order to evaluate the protocol's performance when integrating telemedicine traffic with “regular” traffic in a realistic wireless cellular network scenario, these assumptions need to be waived. More specifically, in our work the following respective additions/changes have been made to the previous wireless scenario which was studied in [\[13\]](#):

1. A portion of the traffic arriving in each cell is handoff traffic from the other cells in the network. Handoff traffic is treated with full priority, with the use of the adaptive bandwidth reservation scheme which will be presented in Chapter 5.
2. Video sources do not “live” permanently in the system, but have exponentially distributed sessions with a mean duration of five minutes [\[29\]](#). This “relieves” a burden from the information interval of the channel, but adds a significant burden to the request interval, which has to compensate for the increase in contention as video users attempt to regain channel access. Similarly to previous works in the literature (e.g., [\[30\]](#)), we have found that a small percentage of the bandwidth suffices to be used for requests. This percentage is 4.4% in our work (25 slots used for requests out of the 566 slots of the channel frame); this value has been found via extensive simulations to provide a good tradeoff between allowing sufficient bandwidth for terminals to transmit their requests and allowing a large enough number of slots for terminals with a reservation to transmit their information packets.
3. In reality, however small the picocell radius, the channel fading experienced by each user is different, since users are moving independently of each

other; therefore, in the present work fading per user channel is considered. As explained in Chapter 3, we adopt the robust three-state (good state, short bad state, long bad state) error model for wireless channels presented in [28] and we introduce the idea that the system should take advantage of the “problem” created when a regular video user experiences a “long bad” channel state (error burst) and is unable to transmit in its allocated uplink slots; this would normally lead to the dropping of the video packets scheduled to be transmitted in these slots, and consequently to higher average video packet dropping probability and the system's failure to satisfy the very strict QoS requirements of real-time videoconference traffic. This new proposed mechanism aims at allocating as many of these slots as possible to other video terminals awaiting for packet transmission, in order to decrease their transmission delay. Although conceptually simple, the above approach is not equally simple to implement. The quality of each user's channel can be indicated by the signal-to-noise ratio (SNR) function; as shown in [31], in a FDD system (such as ours, which is a TDMA-FDD one) using pilot symbols that are inserted in the downlink with a certain time-frequency pattern, the mobile terminals can effectively estimate their SNR function and send it to the BS, which can then make its scheduling decisions based on all the collected cross-layer information from the terminals. This process, however, introduces both errors and delays in the estimates. Due to the random nature of the channel, it is impossible for the BS to precisely determine the state of the channel. The best estimate a BS can provide is a probability distribution over the possible channel states

[32] , which is our assumption in this work, i.e., that the probabilities of the Markov-chain error model have been derived with the above procedure.

Still, the BS cannot know with certainty the type of channel state transition that takes place for a mobile terminal when it leaves the good state, i.e., if the terminal's channel has entered the SB state or LB state. Therefore, the BS can only make an estimation of each mobile video terminal's channel conditions, by monitoring the slots allocated to the terminal and checking whether the terminal is transmitting in them or not. If the total number of a terminal's failed transmissions within its allocated slots surpasses a given threshold, the BS in our scheme deduces that the terminal is in LB state, as the probability that it is in SB is very small given the high number of corrupted transmissions. Based on the channel error model it is easy to confirm by both analysis and simulation that the probability that a mobile terminal's channel is in SB when more than 6 slots have been wasted is 6.55%; hence we have set the threshold to be 6 subsequent transmission failures (choosing a higher threshold would result in a more accurate prediction of the channel condition, as the probability of a mistake in the prediction would be significantly lower; however, it would also lead to a higher number of lost slots while the BS is awaiting to make that prediction). When the BS determines that a mobile video terminal is in LB state, if that terminal has more reserved slots in the current channel frame, the BS deallocates these slots. Full priority for these slots is given to handoff telemedicine video terminals, followed by telemedicine video users originating from within the cell, then by hand-offed regular video users and finally by regular video users originating from

within the cell; the allocation of the abandoned slots within each priority type is FCFS.

When the channel of the mobile terminal to which the slots were originally allocated returns to the good state, the terminal needs to inform the BS of this change, if it still has packets to transmit. This is done by transmitting a request packet. The terminal has to follow this procedure also in the case of a wrong estimation by the BS (i.e., if it was in SB state despite the long error burst). Therefore, in the (unlikely but not improbable) case of a wrong estimation, this does not directly influence the throughput achieved by our protocol in heavy traffic loads (slots are simply allocated to other telemedicine video and regular video users) but it results in an unnecessary increase of contention.

4.2 Scheduling Priorities and Contention Resolution

When resolving the contention among all requesting users, the BS needs to service the telemedicine traffic first, due to its urgency. To achieve this objective, we need to guarantee highest priority to telemedicine traffic. The priority order used by the BS in our proposed scheme is the following: ECG, X-Ray, telemedicine image, telemedicine video. The choice of priorities has been made based on the importance that each of these traffic types currently has for medical care[[2,3,4,6,8,9,10,11,12](#)]. Highest priority is given to handoff telemedicine traffic, with telemedicine traffic originating from within the cell following in priority, in the same order (provided, of course, that the telemedicine users from within the cell have successfully transmitted their requests at the beginning of the frame request interval). Handoff regular traffic is transmitted next, with priority (video,

voice, email, sms), based on the strictness of the QoS requirements for each traffic type (video and voice have the same QoS requirements of less than 1% packet dropping, but video traffic is much burstier, therefore it is granted priority over voice). Regular traffic originating from within the cell is transmitted last, with the same priority order.

It needs to be mentioned again that, similarly to [\[13\]](#), we are able to ensure the priority of telemedicine traffic with the use of the two-cell stack protocol [\[20\]](#) for contention resolution. By exploiting the two-cell stack's advantage of clearly defining the end of contention among users of the same priority class, users of lower priority classes cannot affect the QoS of users of higher priority classes in our system. The use of the two-cell stack protocol also enables users who are moving from cell X to cell Y without having been able to access the channel in cell X, to transmit their request packets to the BS of cell Y with higher priority than new users originating from within cell Y; i.e., only when contention among the request packets of handoff telemedicine users (who were not yet transmitting in cell X) has ended, will regular users originating from within cell Y be able to transmit their request packets.

We employ the two-cell stack reservation random access algorithm for telemedicine video, regular video and voice terminals, and the two-cell stack blocked access collision resolution algorithm to resolve the contention of ECG, X-ray, medical image, email and sms terminals.

The major scheduling problem in serving telemedicine traffic is that not only does it have very strict QoS requirements, but it is also very bursty; hence it is necessary for it to be transmitted immediately when it arrives, but on the other hand the choice of constantly dedicating request bandwidth to it will often result

in the loss of that bandwidth. This problem can be solved with the use of a Call Admission Control (CAC) module at the entrance of the system; the BS is hence notified of which types of telemedicine traffic are active in the cell, and can decide how many request slots should be dedicated to telemedicine users. For example, in the extreme case when all four types of telemedicine traffic are present, both from handoffs and from traffic within the cell, at least 8 of the 25 request slots will be needed (i.e., one slot per type of handoff telemedicine traffic and one slot per type of telemedicine traffic from within the cell) to be dedicated to telemedicine users (the two-cell stack protocol needs a minimum of 2 minislots to resolve contention or to denote the absence of contention in a specific channel frame).

On the other hand, if more than 8 slots are needed to resolve the contention among users of each telemedicine traffic type, then contention will continue until all collisions have been resolved; only then will users of regular traffic (both handoff and from within the cell) get the opportunity to transmit their request packets. Since the case of 8 or more request slots being needed to resolve telemedicine traffic contention is quite infrequent, because of the nature of telemedicine traffic, it will be clear from our results that our scheme can satisfy the strict QoS requirements and the urgency of telemedicine traffic by devoting most of the time less than $8/566=1.4\%$ of the total bandwidth for telemedicine request packets.

4.3 Fair Scheduling

If users of the same type of traffic are served in a FCFS order once they are admitted into the network (as in [13] and in most relevant works on MAC protocols in the literature), the average performance evaluation metrics will give no insight on the QoS of each individual wireless subscriber; therefore, it could be the case that certain users have their QoS severely violated while others get exceptional QoS, which would give a seemingly acceptable average QoS over the total number of users. This approach is, however, unfair to users who arrive later in the network and hence are placed at the bottom of the BS service queue; the problem is especially significant in the case of telemedicine video and regular video users, where early arriving users may dominate the channel by being allocated large numbers of slots, allowing just a small number of resources to be available for users arriving later.

For this reason, we introduce the following Fair Scheduling scheme for telemedicine video and regular video users (the scheme is enforced separately among users of each of the two types of traffic, since telemedicine video users have higher priority). The BS allocates bandwidth by comparing the channel resources to the total requested bandwidth, currently, from all active video users. If the available bandwidth is larger than the total requested bandwidth, all users will be assigned as many slots as they have requested. If, however, the available bandwidth is smaller than the total requested bandwidth, then the available bandwidth will be shared among video users proportionally. More specifically, let M be the number of currently idle information slots in the frame and B_i the amount of bandwidth that will be assigned to video terminal i in every channel frame. B_i is given by:

$$B_i = M * (D_i / \sum_i D_i) \quad (4.1)$$

where D_i is the i^{th} user's requested bandwidth and $\sum_i D_i$ is the total bandwidth requested by all of the video terminals at that moment.

It is intuitively clear, and it will also be shown from our results in Section 6, that with the use of this formula the number of telemedicine and regular, respectively, video users whose QoS is violated significantly decreases. The above scheme does not need to be implemented on any of the other types of traffic considered in our work, besides video, since they are allocated only one slot per frame.

However, an additional scheduling policy is needed for X-Ray and medical image traffic, as the upper bounds for their transmission delays are equally strict with those for telemedicine video and ECG traffic. This strictness can be explained by the fact that X-Ray and medical image terminals are allocated only one slot per frame (close to 35 Kbps, similarly to regular voice, email and sms traffic), to allow for the significantly larger numbers of slots needed by telemedicine and regular video users. Therefore, a typical X-Ray file of 200 Kbytes needs 50 seconds to be transmitted (while the upper bound for its transmission delay is set to 1 minute) and an average-sized medical image file of 17.5 Kbytes needs 4.4 seconds to be transmitted (while the upper bound for its transmission delay is set to 5 seconds). The allocation of only one slot per frame to these types of traffic, although defensive enough to prevent cases where newly arriving telemedicine video traffic cannot find enough resources to transmit, is not the most efficient in terms of bandwidth utilization. On the other hand, if these types of traffic are constantly granted more than one slot per frame, this could lead to the existence of too few idle slots for the bursty telemedicine video users. Therefore, in order to maximize

system bandwidth utilization we use the following scheduling policy. After the end of the request interval at the beginning of each frame, the BS is aware of whether there are information slots during the current frame which will be left idle. These slots are allocated, only for the current frame, to X-Ray and medical image users who have already entered the network (with priority to X-Ray users), as additional slots to their guaranteed single slot per frame. Hence, the telemedicine traffic transmission is expedited and channel throughput is increased. This policy does not need to be extended to ECG traffic, as ECG users need only one slot per frame.

4.4 FCFS – EDF – SJF Algorithms

As mentioned earlier, in [13, 42] and in most of the relevant MAC protocols' literature, users of the same type of traffic are served in a FCFS order once they are admitted into the network. The First-Come-First-Served algorithm is the simplest scheduling algorithm. Processes are dispatched according to their arrival time on the ready queue. Jobs arriving are placed at the end of queue, the dispatcher selects the first job in the queue and this job runs to completion. The FCFS scheduling is fair in the human sense of fairness but it is unfair in the sense that long jobs make short jobs wait. FCFS is not useful in scheduling interactive users because it cannot guarantee a good response time.

The Earliest Deadline First (EDF) algorithm operates on the logic that, among the users of one traffic class (e.g. voice), the user with the nearest deadline to transmit (a packet, message or video frame) will be accommodated first, instead of the user that arrived first (i.e. transmitted a request packet

successfully in an earlier minislot). Earliest deadline first or “least time to go” is a dynamic scheduling algorithm used in real-time operating systems. It places processes in a priority queue. Whenever a scheduling event occurs (task finishes, new task released, etc.) the queue will be searched for the process closest to its deadline. EDF is an optimal scheduling algorithm, in the following sense: if a collection of independent jobs, each characterized by an arrival time, an execution requirement, and a deadline, can be scheduled (by any algorithm) such that all the jobs complete by their deadlines, then EDF will schedule this collection of jobs such that they all complete by their deadlines. However, when the system is overloaded, the set of processes that will miss deadlines is largely unpredictable (it will be a function of the exact deadlines and the time at which the overload occurs.) This is a considerable disadvantage to a real time systems designer. The algorithm is also difficult to implement in hardware.

Shortest Job First (SJF), also known as Shortest Job Next (SJN), is a scheduling policy that selects the waiting process with the smallest execution time to execute next, and this process runs to completion. Shortest job next is advantageous because of its simplicity and because it maximizes process throughput (in terms of the number of processes run to completion in a given amount of time). However, it has the potential for process starvation, for processes which will require a long time to complete if short processes are continuously added. We did not use SJF algorithm on voice users since we cannot predict a priori how long a conversation will last, while we know, e.g., the size of data messages or video frames.

4.5 Jain's Fairness Index

As explained in Section 4.3, aggregate performance evaluation metrics reveal very little, if any, information, regarding the QoS of each wireless subscriber. Therefore, we should examine the impact of FCFS, EDF, SJF not only on commonly used performance metrics, such as throughput and delay, but also on fairness. We used Jain's Fairness Index [\[41\]](#) which is defined as follows:

$$J(x_1, \dots, x_n) = \frac{(\sum_{k=1}^n x_k)^2}{n \sum_{k=1}^n x_k^2}, \quad (4.2)$$

where n is the number of users and x_k is the throughput of user k . This index is continuous and bounded between 0 and 1, with 1 denoting maximal fairness. It is also very intuitive. If a ratio y of the users are treated fairly and $(1 - y)$ are starved, then the resulting fairness index is y . We study fairness separately for telemedicine video users, regular video users and voice users. The throughput is measured in terms of the average number of allocated slots over 83 frames, which approximately corresponds to the average talkspurt duration (1 second); n is the number of users that were active during each 83-frame window (which is called a superframe).

Chapter 5: Adaptive Bandwidth Reservation

based on Mobility and Road Information

5.1 Overview

Within a picocell, spatially dispersed source terminals share a radio channel that connects them to a fixed base station (BS). The BS 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.

It is a common assumption in past studies that the dissatisfaction of a wireless cellular subscriber who experiences forced call termination while moving between picocells is higher than that of a subscriber who attempts to access the network for the first time and experiences call blocking [33,34]. For this reason, it is important that the system is able at any point in time to accommodate newly arriving handoff calls in any cell of the network. On the other hand, the policy of reserving a significant amount of bandwidth for possible handoff calls may lead to a portion of the bandwidth being left unused, due to small volumes of handoff traffic, while at the same time the remaining available resources for newly generated traffic from within the cell may not suffice. Two of the most common methods to handle handoff traffic are to calculate either sojourn time [35, 36] or handoff probability/rate [37, 38, 39, 40] based on some updated location information. Besides the basic location information gained through the Global

Positioning System (GPS), our mobility model uses road map information as a fundamental component of its function.

5.2 Network and Mobility Models

We consider an architecture of seven hexagonal cells which we place in two different topologies. In the first topology, we assume the “circular” case, where after leaving the last cell (G) a user enters cell (A) again. In the second topology, mobiles can get in or out of our system only via cell A and cell G, which are placed at the two edges of our network model [\[44\]](#). So, all cells are connected in a straight line, as shown in figure [5.1](#).

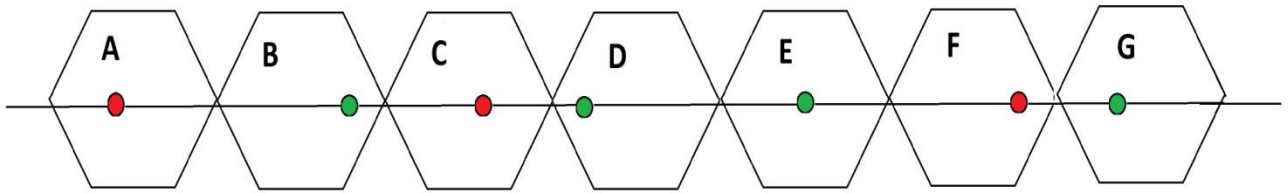


Figure 5.1: Road Map and cellular network model

In both topologies, each cell has two neighbors. The cell diameter is 300 meters. One road, which is modeled by a straight line, passes through cells and connects them. Each new call is generated with a probability of 50% to be moving on the road and 50% to be stationary. Moving users are assumed to be traveling only on the road. The initial location of a moving user on the road is a uniform random variable between zero and the length of that road. During their call, stationary callers remain stationary and mobile users travel at a constant speed.

Mobile users can travel in either of the two directions of a road with an equal probability, and with a speed chosen randomly in the range of [36, 90] Km/h. One traffic light is located randomly within each cell. A mobile user arriving at the traffic light of the cell might continue to go straight, or turn around with probabilities 0.9 and 0.1, respectively. If a mobile user chooses to go straight at the traffic light, it needs to stop there with probability 0.5 for a random time between 0 and 30 seconds due to a red traffic light, or else passes with probability 0.5. If the user chooses to turn around, it needs to stop there for a random time between 0 and 60 seconds due to the traffic signal [36, 37]. Each base station is loaded with the road map of its coverage area and its neighboring cells. Mobile stations report their position to the BS of their cell through a control channel. The position information includes the mobile user's exact location (cell), moving direction, and speed, and can be provided with an accuracy of 1m through GPS [36, 37, 40].

5.2.1 The Mobility Model in [42]

As explained earlier, this thesis significantly extends the work of [42], and in Section 6 we compare our results with those of that work. In [42], the following map was utilized in order to observe how efficient the adaptive bandwidth reservation is.

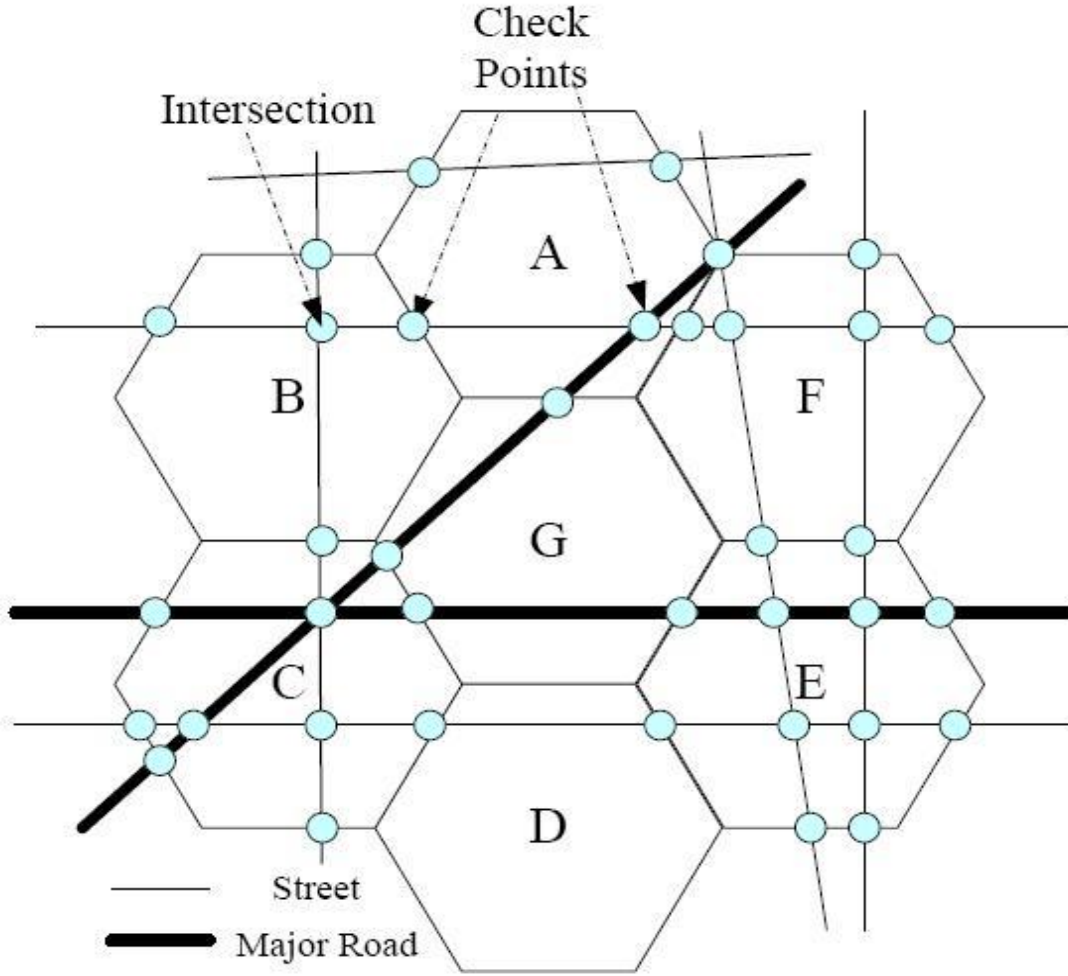


Figure 5.2: Road Map and cellular network model

In this map, each cell has six neighbors. The cell diameter is 300 meters. Roads are modeled by straight lines. Each road is assigned a weight (ω_j for road j), which represents the traffic volume. Each new call is generated with a probability of 50% to be moving on the road and 50% to be stationary. Moving users are assumed to be traveling only on the roads, and are placed on each road i with probability

$$\omega_i / \sum_{j=1}^N \omega_j \quad (5.1)$$

where N is the total number of roads.

The initial location of a moving user on a particular road is a uniform random variable between zero and the length of that road. During their call, stationary callers remain stationary and mobile users travel at a constant speed. Mobile users can travel in either of the two directions of a road with an equal probability, and with a speed chosen randomly in the range of [36, 90] Km/h. At the intersection of two roads, a mobile user might continue to go straight, or turn left, right, or around with probabilities 0.55, 0.2, 0.2 and 0.05, respectively. If a mobile user chooses to go straight or turn right at the intersection, it needs to stop there with probability 0.5 for a random time between 0 and 30 seconds due to a red traffic light. If the user chooses to turn left or around, it needs to stop there for a random time between 0 and 60 seconds due to the traffic signal.

More specifically, [42] sets the road intersections and cell boundaries to be the “check-points” of the system, as shown in Figure 5.2. Each cell boundary represents a unique check-point, while around each road intersection four check points are set, one in each possible direction that the user may choose after reaching the intersection. Each of the check points is assumed to be placed at a distance of 10m from the intersection, which is a distance that even the slowest moving vehicles (with a speed of 36 Km/h) will cover within 1s after passing the intersection (naturally, if there are less than four possible directions for a mobile to follow after reaching an intersection, the number of check points needed is smaller than four). Mobile stations only need to update their position information to the BS of their cell when they arrive at a check-point.

5.3 Adaptive Bandwidth Reservation Scheme

In previous studies [36,37,40], the bandwidth reservation schemes required mobile stations to report their location information to the BS every T seconds (1 second in [40], 10~45 seconds in [36] and 60 seconds in [37]). For this period of time if there was a probability based on the mobile's trajectory that it would move to a new cell, then a certain amount of bandwidth was reserved in all possible future cells that the mobile may move to. In [37], a prediction was made by the system based on each updated position information and the road map was randomly generated for each simulation. The common disadvantage of these approaches is that the length of the report period yields a tradeoff between the prediction accuracy and the computational load imposed on the system. If mobile stations report their position frequently, the computational load of processing this information and using it for bandwidth reservation will be too high; on the other hand, if mobile users' position reports are too infrequent, the prediction on the mobile's trajectory can be untrustworthy and lead to an unnecessary waste of the reserved bandwidth in neighboring cells. Because of this tradeoff, in [36,40], some additional dynamic mechanisms are used in order to adjust the amount of reserved bandwidth for users in neighboring cells, based on the quality of the system performance. In our study however, we propose the use of distance-based information reports in order to eliminate the problems introduced by the time-based location information reports of the mobile station to the BS.

More specifically, we adopt the idea that mobile stations only need to update their position information to the BS of their cell when they arrive at a traffic light. If the BS of the current cell of a mobile station predicts, based on the

station's location (at a traffic light) and speed that the station is going to move to another cell, it sends a notification to the BS of that cell, including the current bandwidth used by the station and the estimated arrival time at the next traffic light. Hence, the proper amount of bandwidth is reserved for the station. For telemedicine video and regular video terminals, the bandwidth that is reserved in the next cell is equal to the remaining bandwidth that the terminal will need to complete its transmission (this bandwidth is declared in each video user's initial request to the BS). For all other types of users, the bandwidth that is reserved in the next cell is equal to their current bandwidth, so that they will seamlessly continue its transmission. How to keep the probability of hand-off drops within a prespecified limit is a very important quality-of-service (QoS) issue in cellular networks because mobile users should be able to maintain ongoing sessions even during their hand-off from one cell to another. Our proposed approach not only guarantees the existence of adequate resources in the new cell for handoff users but also reduces the information update frequency. The prediction and adaptive reservation process executed by the BS can be summarized as follows, in algorithm 1.

Algorithm 1: Distance-based Adaptive Bandwidth Reservation Scheme

for Each new user in the system do

if The mobile station is not stationary **then**

When the mobile station arrives at a traffic light

 Update the position information

 Estimate the next cell boundary and calculate the arrival time at that

```
        Reserve the proper amount of bandwidth for the station in the
        next cell
    end if
end for
```

Our proposed adaptive bandwidth reservation scheme based on distance-based location updates has two major advantages:

- The position update duration is unique for every mobile user, based on their different initial position and speed. This is impossible to “capture” with the time-based location updates proposed [36, 37, 40]. Additionally, it creates less computational load than the time-based updates which, in most proposals in the literature, are synchronized and therefore require simultaneous information transmission from the terminals to the BS.
- An important parameter used to evaluate a bandwidth reservation scheme is bandwidth efficiency [37], which is calculated by $f = N_r/N_q$, where N_r is the reserved bandwidth and N_q is the actual bandwidth utilized by hand-offed users. The closer f is to 1, the higher is the efficiency achieved by the bandwidth reservation scheme, since this would mean that there is neither lack nor waste of bandwidth in the reservation procedure.

We argue that with our distance-based approach the bandwidth efficiency of our scheme can asymptotically reach 1. The reasons are:

- Our scheme guarantees that there is no waste of bandwidth. For all users ready to handoff, the exact amount of bandwidth they need is reserved in the new cell and this cell is known with precision, with the use of the traffic lights. The only case when bandwidth can be wasted, with the use of our adaptive bandwidth reservation approach, is when a mobile station which is predicted to move to another cell makes a stop before entering this cell (this case is not included in our adopted mobility model). Given the fact that the distance between traffic lights and cell boundaries is in all cases much smaller than the cell diameter of 300 meters, this case can be considered a rare exception.
- Our scheme also guarantees that there will be no lack of bandwidth for hand-offed users. As explained in Chapter 4, the bandwidth needed for all types of telemedicine and regular mobile stations which do not transmit video is close to 35 Kbps (just one slot per channel frame). Therefore, this bandwidth is very small compared to the total channel capacity of 20 Mbps and can generally be reserved in the next cell with the rare exception of cases when the channel is overloaded with traffic. In the case of hand-offed telemedicine users (transmitting any type of telemedicine traffic), if the needed bandwidth is not available, then email and web users are preempted in the new cell in order for the system to grant this bandwidth to the high-priority telemedicine traffic. When email and web reservations are canceled, the BS notifies the affected data terminals and places them at the front of the respective (email or web) queue of terminals awaiting

bandwidth allocation. The priorities among telemedicine traffic types in bandwidth reservation are set in the same way as the scheduling priorities which were discussed in Chapter 4.

Chapter 6: Results and Discussion

6.1 Simulation Setup

We use computer simulations to study the performance of our scheme. The simulator is written in C programming language. Each simulation point is the result of an average of 5 independent runs (Monte-Carlo simulation), each simulating 155000 frames (1/2 h of network operation), the first 5000 of which are used as warm-up period. In our results, we use different traffic “combinations” from all types of traffic considered in our work, in order to test the system's performance in a large variety of cases. In this way, we try to produce results representative of different practical cases, where one type of telemedicine traffic might be more dominant than others in any given moment. Each simulation point presents the average result over 5 combinations which were used to create a specific traffic load.

6.2 Results

6.2.1 FCFS results

We present below the first part of our results, in figures 6.1-6.11. These figures include results derived with the use of the FCFS (First Come First Served) algorithm.

For the first 3 figures we created 5 “scenarios” where telemedicine traffic corresponded to 5%-10% (1-2 Mbps) of total channel capacity and regular traffic consisted of video users only.

Figure [6.1](#) shows how the increase in the number of regular video users affects the QoS of X-Ray traffic. The maximum number of video users (62) corresponds to 95% of the total traffic being generated by regular traffic. The average delay for the start time of transmission of X-Ray files increases very slowly and does not become larger than 230 ms even for very high numbers of regular video users. This means that the upper bound of 1 min for the transmission delay of an X-ray file is met in all cases examined in this paper (<230 ms until the transmission starts, i.e., until the specific file acquires a slot in which it will be transmitted, plus 50 s for the actual transmission of the file). The Poisson-distributed X-ray file arrivals in our simulations have a rate λ_x (files/frame) that varies between 0.0022 and 0.0052. These rates are quite high, because for example, a rate of 0.0052 X-ray file transmission per 75 ms frame corresponds to an average transmission of four X-ray files per minute in the seven picocells shown in our map, which, in real life, would correspond to a massive crisis situation. As also shown in Fig. [6.1](#), even for zero video users, there is a delay of slightly less than nine channel frames (104 ms) in the start time of transmission of an X-ray file due to the existence of handoff traffic and other types of telemedicine and regular traffic in the cells. Regarding the QoS of the other types of telemedicine traffic (for the same range of regular video users), telemedicine image files are transmitted on average within less than 50 ms, ECG packets are transmitted, on average, within half a frame (6 ms) and the average packet dropping probability of telemedicine video ranges between 0.0035% and 0.0039%

as it is affected only by the wireless channel errors. All the aforementioned values are far below the acceptable upper bounds; hence, our adaptive bandwidth reservation and scheduling schemes are shown to provide the required QoS for all types of telemedicine traffic.

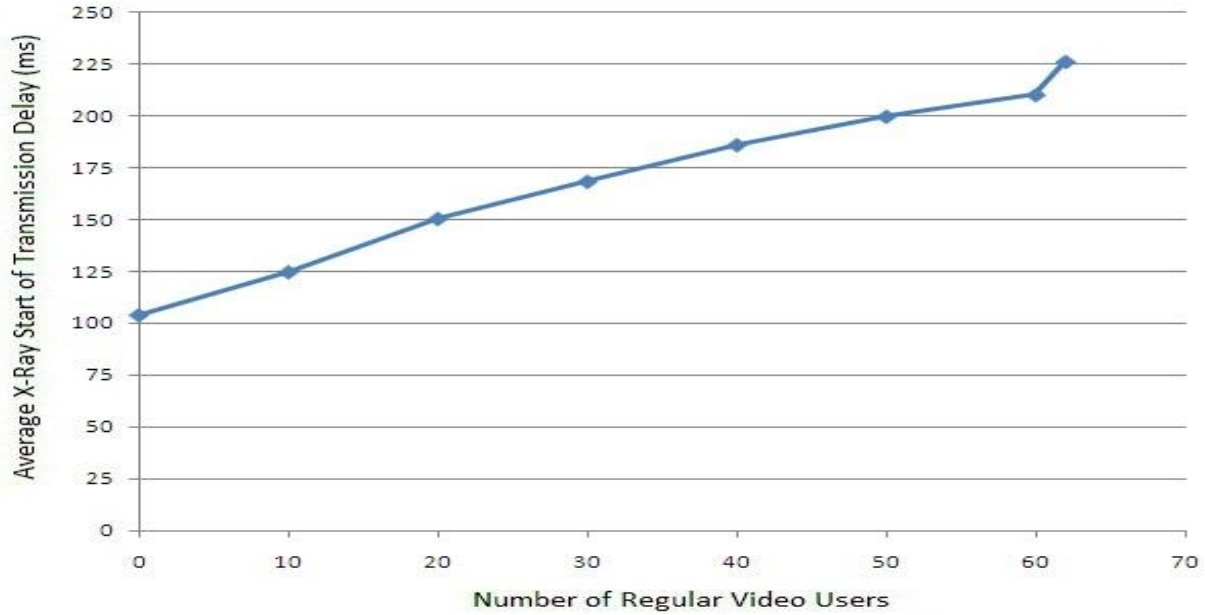


Figure 6.1: Effect of regular video traffic on X-Ray traffic

Figures [6.2](#) and [6.3](#) show that the increase in the number of regular video traffic clearly affects that type of traffic, because the video-packet-dropping probability of regular video users significantly rises above the 1% acceptable upper bound. This conclusion, combined with the results presented in Fig. [6.1](#), shows that the increase in the number of regular video traffic is clearly affecting only that type of traffic; therefore, our combined schemes succeed in offering absolute priority to telemedicine traffic. Figs. [6.2](#) and [6.3](#) also show that

the use of our fair scheduling scheme substantially helps improve both the average video-packet-dropping probability over all regular video users and the individual QoS of each regular video user. The improvement in the individual QoS, as shown in Fig. 6.2, is due to the allocation of available resources in each channel frame proportionally to each user's requested bandwidth. The result shown in Fig. 6.3 with regard to the improvement in the average video packet dropping is, again, due to the proportionate allocation; our scheme prevents the case where a user whose transmission deadline is not imminent may dominate the channel, hence not allowing users with imminent transmission deadlines to transmit.

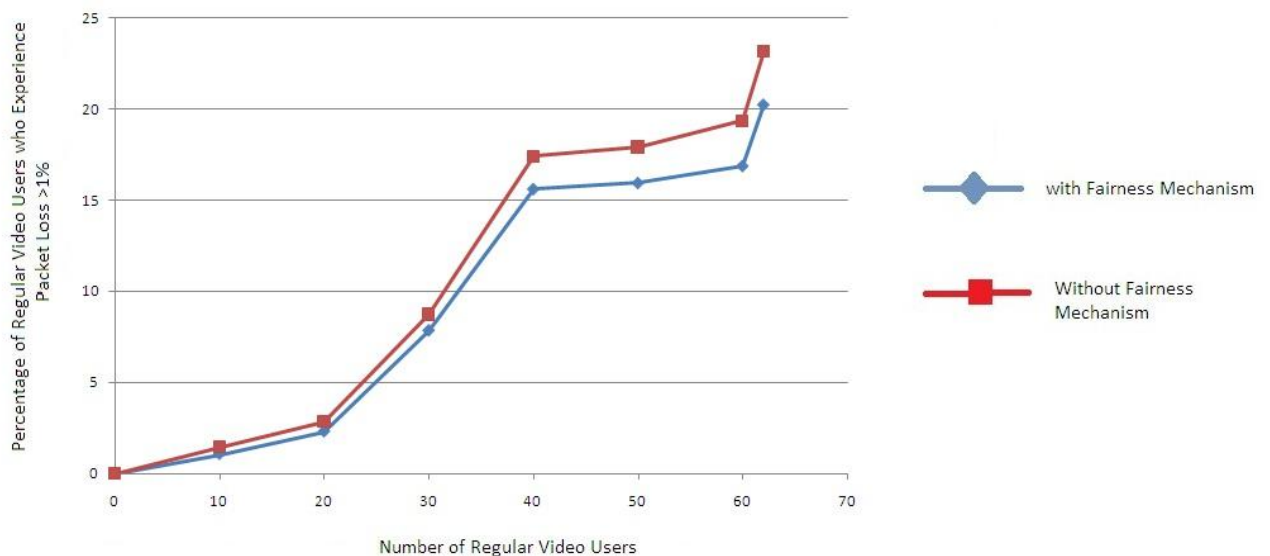


Figure 6.2: Percentage of regular video users who experience packet loss larger than 1%, versus the number of regular video users

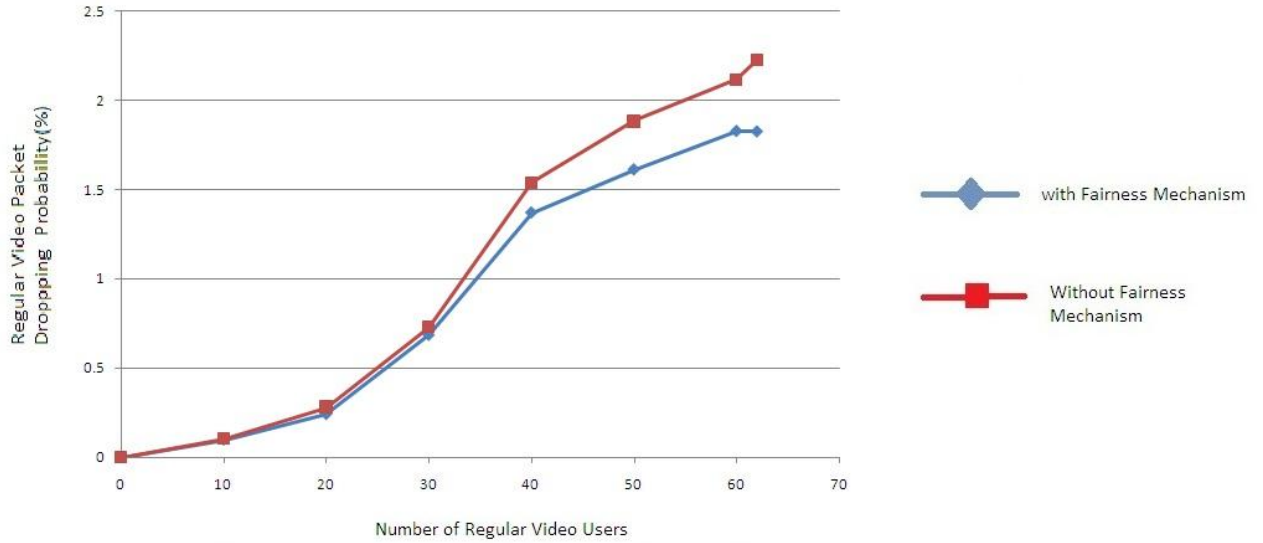


Figure 6.3: Regular video packet dropping versus the number of regular video users

For the results presented in figures [6.4-6.5](#) we created 5 “scenarios” where telemedicine traffic corresponded to 5%-10% (1-2 Mbps) of total channel capacity and regular traffic consisted of voice users only.

Figures [6.4](#) and [6.5](#) show the almost negligible effect on the QoS of telemedicine users that stems from the increase in the number of voice users. We present results on the telemedicine video packet dropping and telemedicine images transmission delay, and it is clear in both figures that only in the case of very high voice loads can there be deterioration in telemedicine traffic QoS. The reason, again, is that our combined scheduling and adaptive bandwidth reservation schemes guarantee full priority to all types of telemedicine traffic. The average telemedicine image traffic transmission delay does not exceed 60 ms, as shown in Fig. [6.5](#), even when the number of voice users exceeds 1397, which corresponds to 95% of the total channel capacity being utilized by voice users.

Similarly to the results presented in Fig. 6.1, even for 0 voice users, there is a small delay for the transmission of a telemedicine image file due to the existence of handoff traffic and other types of telemedicine and regular traffic in the cells. For the same channel utilization by voice users, we found that the average X-Ray traffic transmission delay does not exceed 140 ms. In addition, the telemedicine-video-dropping probability is shown in Fig. 6.4 to remain below the strict upper bound of 0.01% until the number of voice users exceeds 1000, which corresponds to 68% of the total channel capacity. Hence, only in the case of a very heavily loaded channel with voice traffic, does the telemedicine traffic experience some deterioration in its QoS.

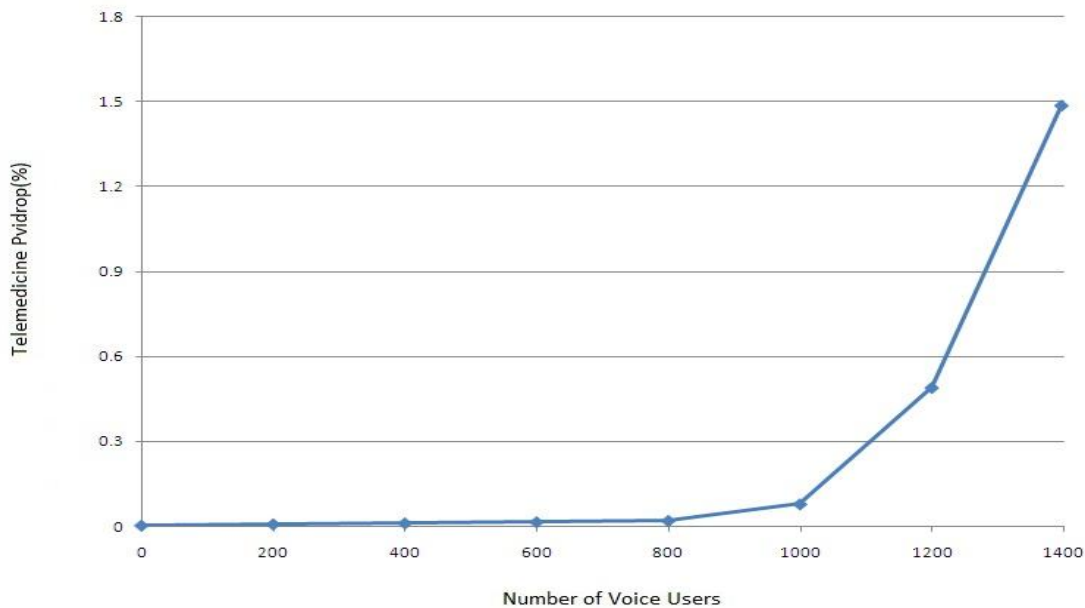


Figure 6.4: Effect of regular voice traffic on telemedicine video traffic

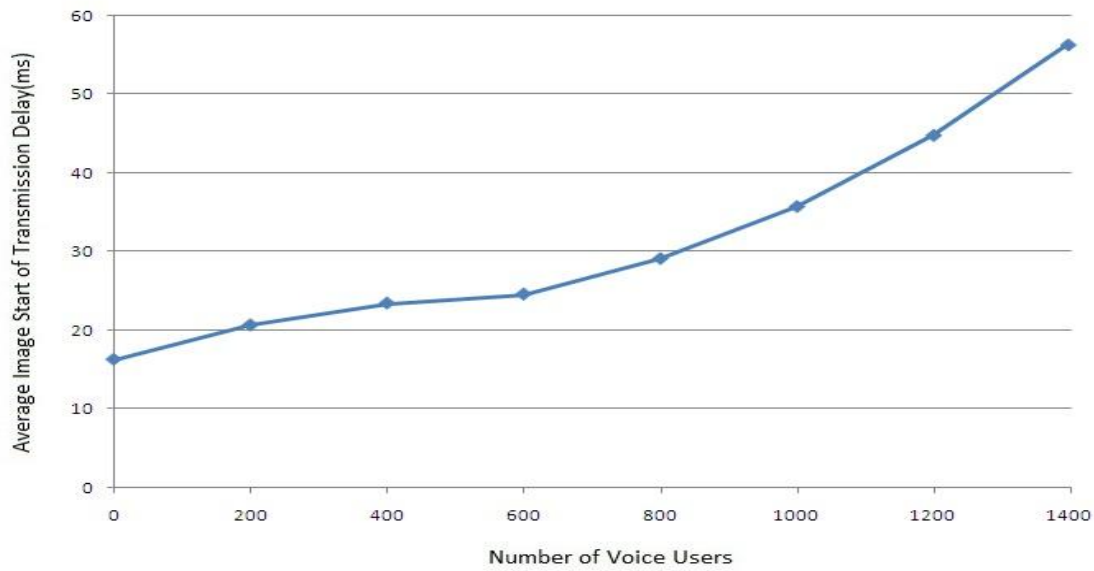


Figure 6.5: Effect of regular voice traffic on telemedicine image traffic

For the results presented in figures 6.6-6.8 we created 5 “scenarios” where telemedicine traffic corresponded to 5%-10% (1-2 Mbps) of total channel capacity and consisted only of telemedicine video users while the rest of the load consisted of all types of regular traffic.

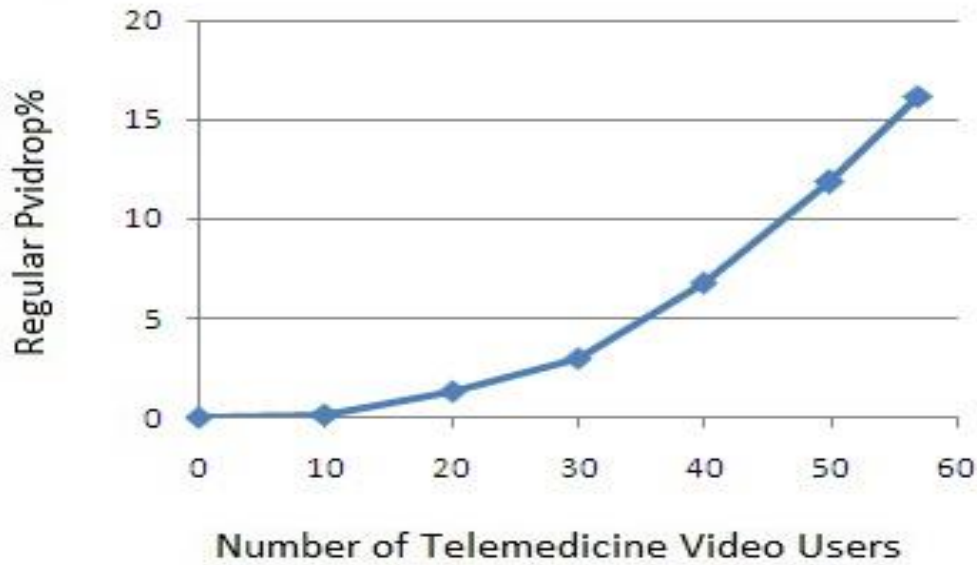


Figure 6.6: Effect of telemedicine video traffic on regular video traffic

Figure [6.6](#) shows the effect that the increase in telemedicine load has on regular video traffic, which is the most bursty of all regular traffic types. This increase results in a very significant increase in regular video packet dropping due to the absolute priority of telemedicine traffic. Even in this case, however, there need to be more than 18 telemedicine video users present in the seven picocells for the QoS requirement of a maximum of 1% regular video packet dropping to be violated.

Figs. [6.7](#) and [6.8](#) are similar in nature to Figs. [6.2](#) and [6.3](#), but in this case, we study the effect of the increase of the number of telemedicine video users to the telemedicine-video-packet-dropping probability. It is clear in the figures, again, that our fair scheduling scheme significantly improves both the average telemedicine-video-packet-dropping probability over all telemedicine video users and the individual QoS of each telemedicine video user. An interesting comparison can be made between Figs. [6.2](#) and [6.3](#) and Figs. [6.7](#) and [6.8](#). In Figs. [6.2](#) and [6.3](#), the QoS requirement for regular video packet dropping ($< 1\%$) is violated when a larger percentage of users experiences a packet loss $> 1\%$, compared with the percentage of telemedicine video users who experience a packet loss $> 0.01\%$ when the QoS requirement for telemedicine video packet dropping ($< 0.01\%$) is violated. This is the case both for the implementation without a fairness mechanism and for the implementation with the fairness mechanism that we propose. The reason for this result is that our schemes are designed to offer maximum priority to telemedicine traffic; therefore, when telemedicine video packet dropping increases, the increase is almost “uniform” for all telemedicine video users. On the contrary, regular video users may exhibit differences in their video packet loss (even with the use of the fairness mechanism), because they are of lower priority, and hence, some of these users may not transmit in a traffic overload situation.

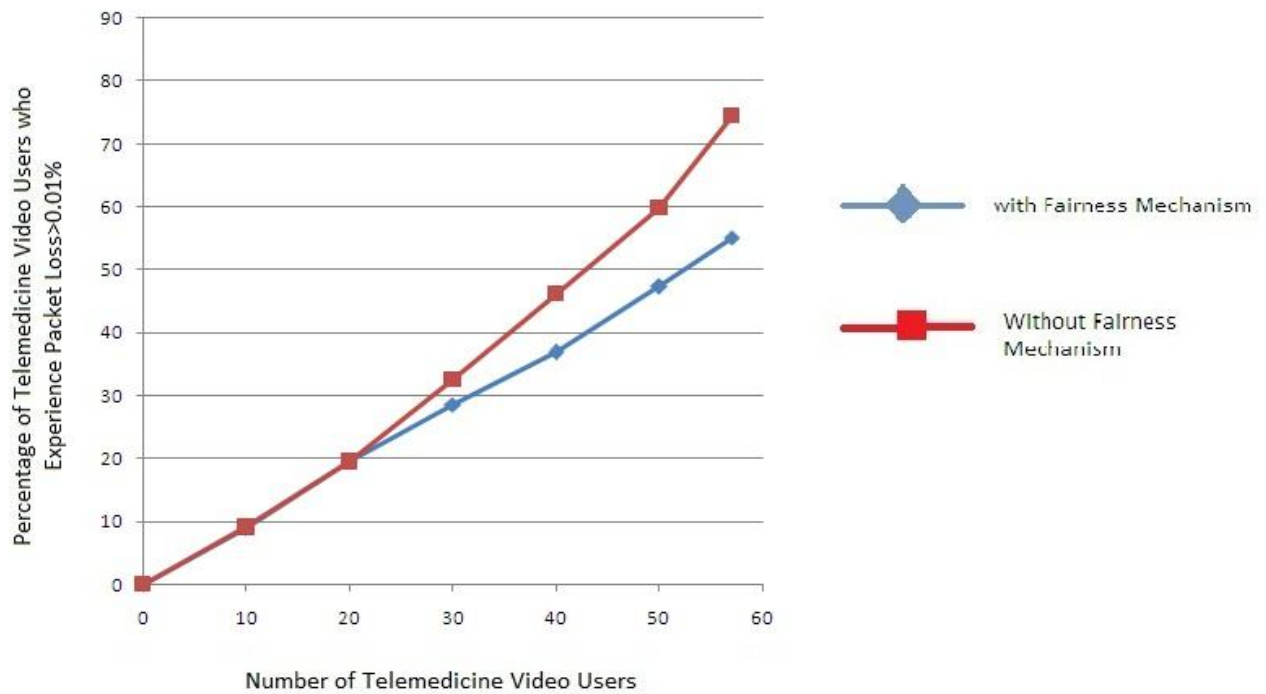


Figure 6.7: Percentage of telemedicine video users who experience packet loss larger than 0.01%, versus the number of telemedicine video users

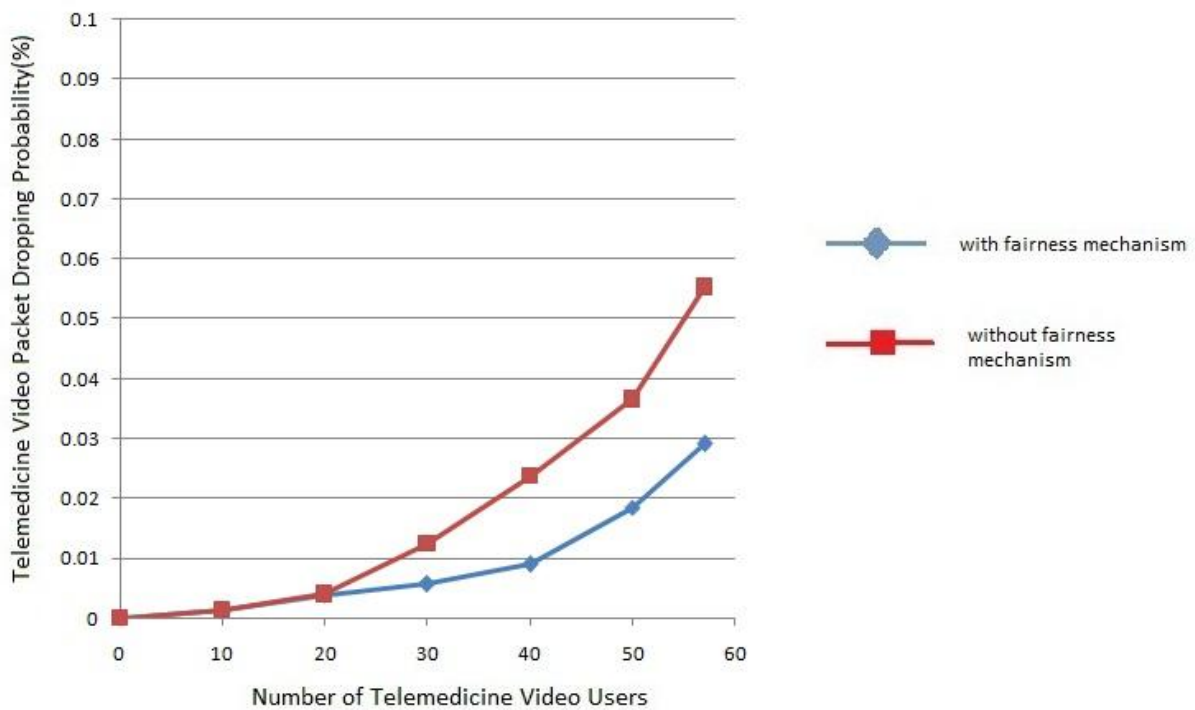


Figure 6.8: Telemedicine video packet dropping versus the number of telemedicine video users

For figures [6.9](#) and [6.10](#) we created 5 “scenarios” where telemedicine traffic corresponded to 5%-10% (1-2 Mbps) of total channel capacity and consisted of telemedicine video and X-Ray users(or tele-image users for figure 10,respectively) while all types of regular traffic are present in the system, with the exception of regular video traffic.

Figures [6.9](#) and [6.10](#) show that only under a medium-to-high number of telemedicine video users can there be an impact on the QoS of X-ray traffic and telemedicine image traffic. The reason is that both types of traffic have a higher priority in our scheme than telemedicine video.

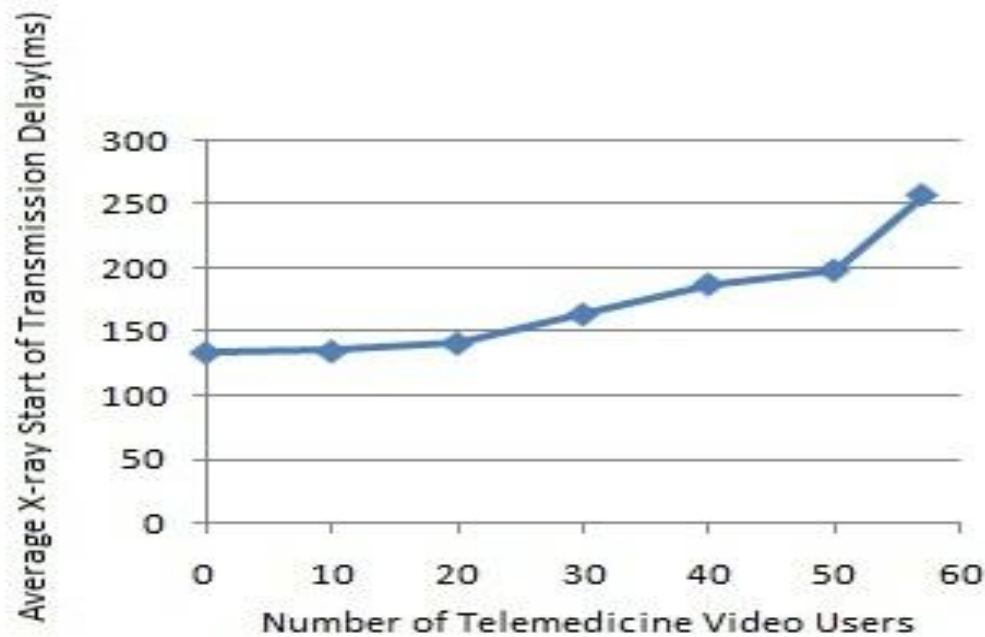


Figure 6.9: Effect of telemedicine video traffic on X-ray traffic

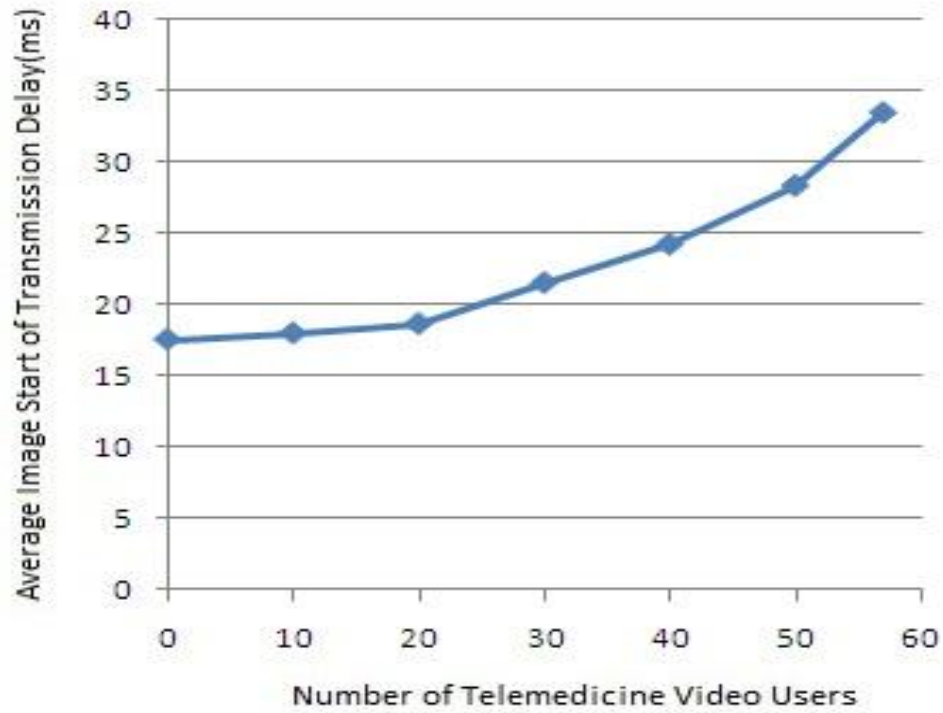


Figure 6.10: Effect of telemedicine video traffic on telemedicine image traffic

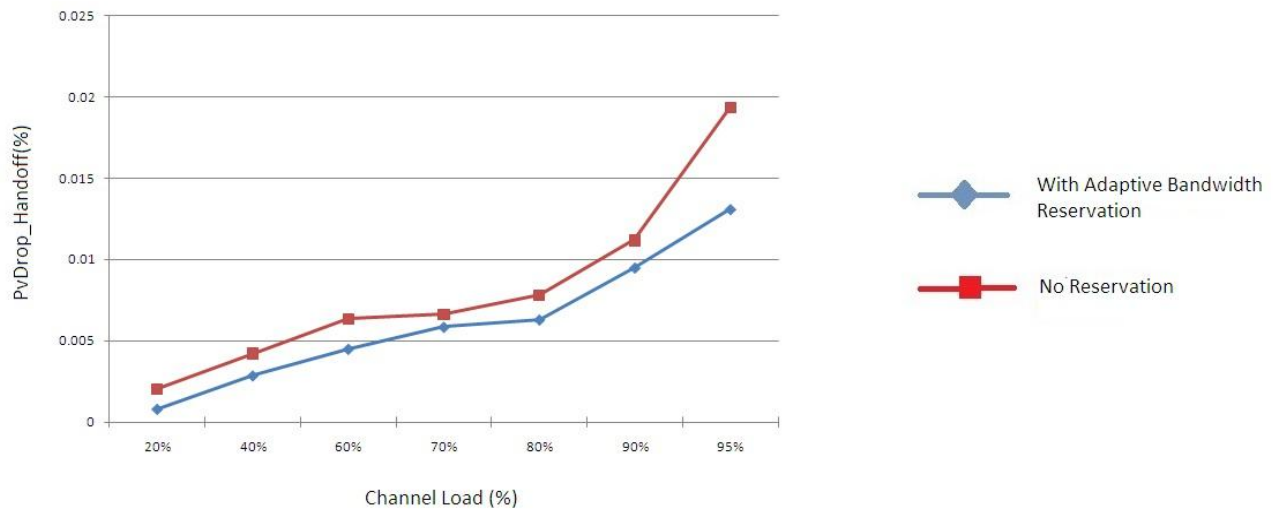


Figure 6.11: Improvement on handoff voice packet dropping probability with the use of adaptive bandwidth reservation

Figure [6.11](#) shows the significant improvement achieved by the use of our adaptive bandwidth reservation scheme on the QoS of the most widely used cellular application, i.e., voice. The voice-packet-dropping probability of handoff voice users is quite smaller, on average, compared with the case when the bandwidth reservation scheme is deactivated. The reason is that, by intelligently reserving bandwidth in adjacent cells with bandwidth efficiency f almost equal to 1, our scheme helps to significantly decrease contention for channel resources. Therefore, once again our proposal is shown to improve the QoS of regular traffic, while always offering highest priority to telemedicine users.

On the whole, our results for all other types and combinations of telemedicine and regular traffic confirm that telemedicine traffic is negligibly affected by an increase in regular traffic, whereas regular traffic is severely affected by increased loads of telemedicine traffic. Therefore, our combined adaptive bandwidth reservation and fair scheduling schemes guarantee the absolute priority of telemedicine traffic and can also preserve the required QoS of regular traffic under all the traffic scenarios used in our extensive simulation study.

6.2.2 FCFS results using a different map

In this section we present four Figures (6.12-6.15) which contain results from the work in [\[42\]](#). The respective Figures from our work are Figures [6.1](#), [6.4](#), [6.6](#) and [6.8](#). We proceed to compare the results from the two works. This

comparison concerns the case of FCFS scheduling only, as [42] did not use any other scheduling algorithm.

In Figure 6.12 the slope of the increase in X-Ray transmission delay is similar to that in Figure 6.1, but the maximum X-Ray transmission delay is almost double in [42] than in our work (460 vs. close to 230 ms).

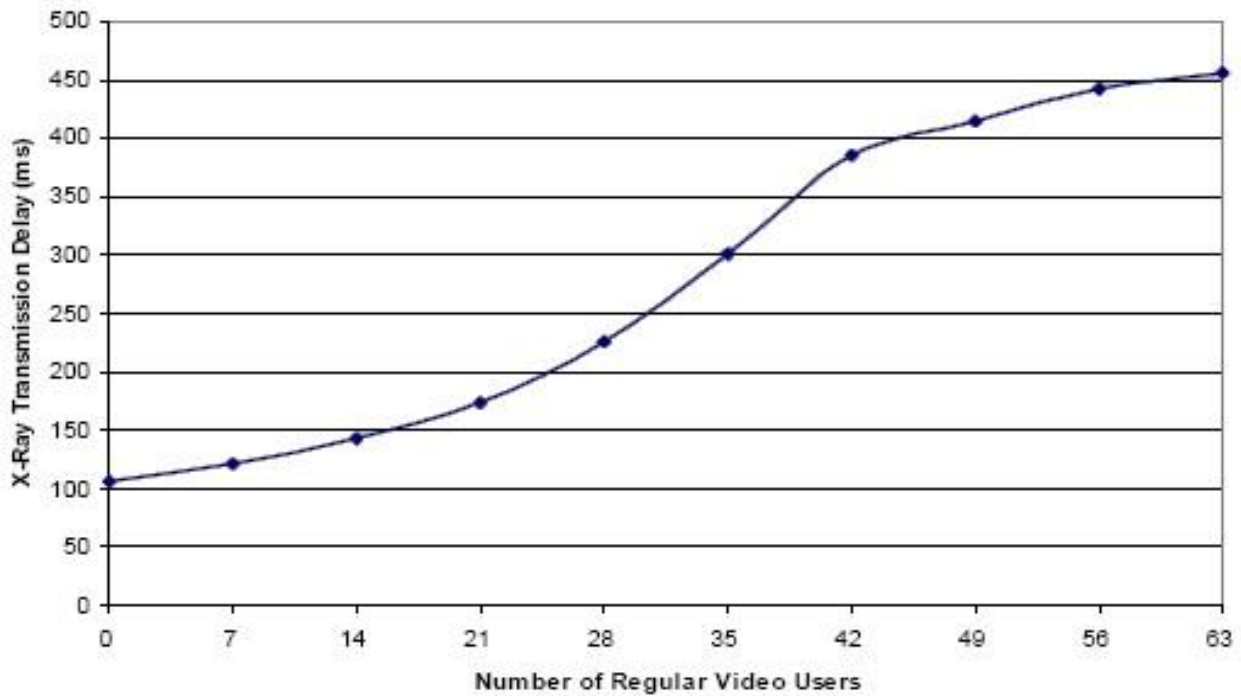


Figure 6.12: Effect of regular video traffic on X-Ray traffic

In Figure 6.13, similarly to the results presented in Figure 6.4, for a small number of voice terminals the telemedicine video packet dropping probability is shown to remain below the strict upper bound of 0.01%. However, for voice loads

higher than 80% of the total channel capacity we observe that the highest value of telemedicine video packet dropping probability reaches 2.8%, in contrast to 1.5%, which is the maximum value in Figure [6.4](#).

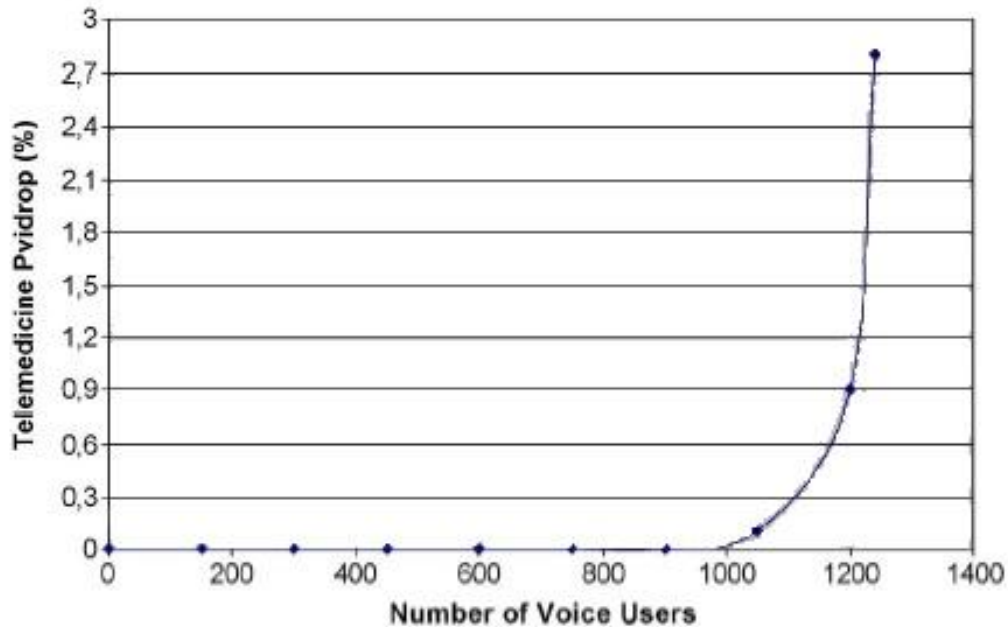


Figure 6.13: Effect of regular voice traffic on telemedicine video traffic

Figures [6.14](#) and [6.15](#) are once again similar in nature to figures [6.6](#) and [6.8](#). We study the effect that the increase in telemedicine video load has on regular and telemedicine video traffic. Using our road map (figure [5.1](#)) we notice that regular video packet dropping probability does not surpass 20% and telemedicine video packet dropping probability 0.06%, due to the absolute priority of telemedicine traffic. Figures [6.14](#) and [6.15](#) show that the corresponding values are 40% and 0.09% when the road map presented in Figure [5.2](#) is used with FCFS scheduling.

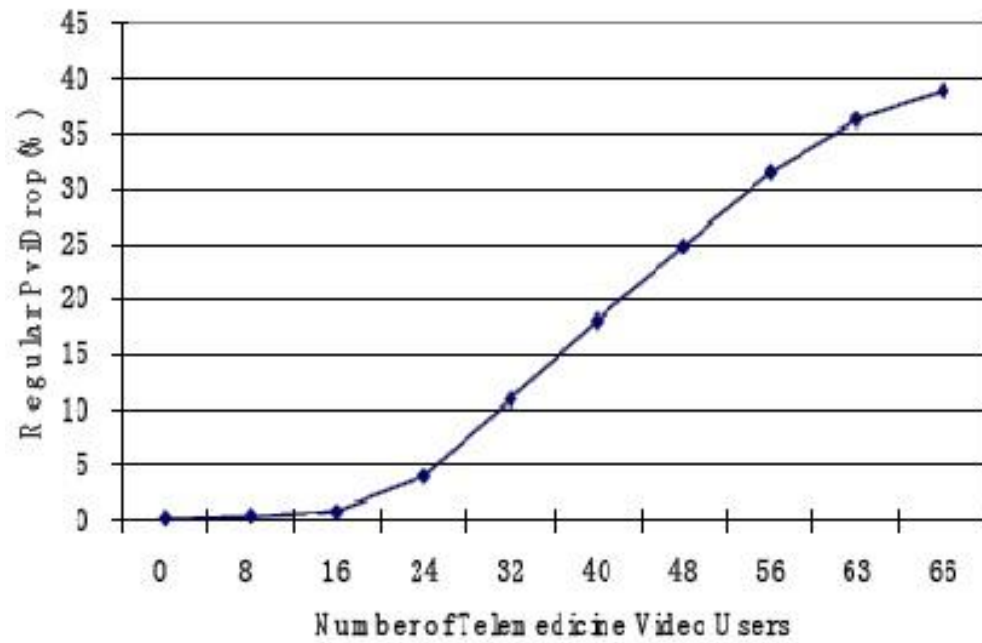


Figure 6.14: Effect of telemedicine video traffic on regular video traffic

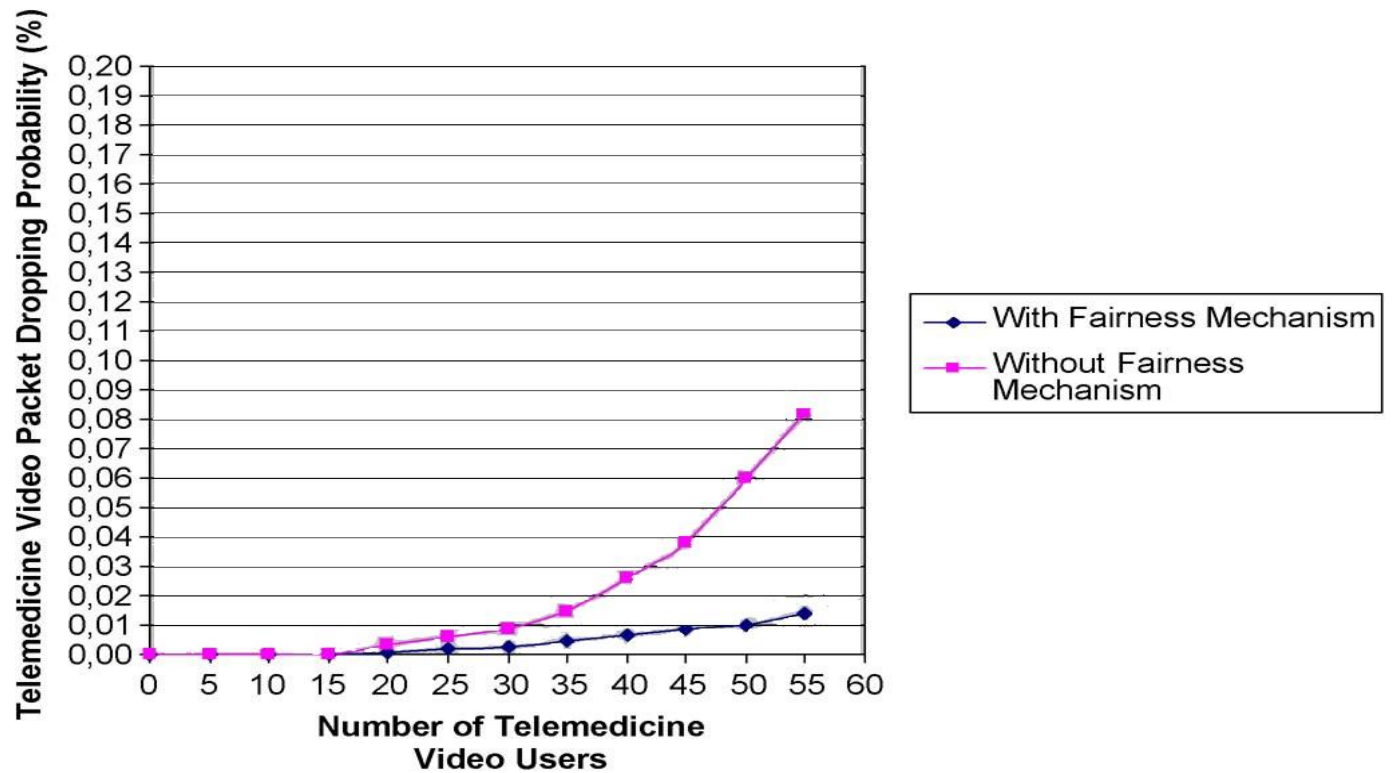


Figure 6.15: Telemedicine video packet dropping versus the number of telemedicine video users

The common conclusion of all the comparisons is that, when using FCFS, the results in [42] are very close in nature but worse in terms of the user QoS than those in our work. The reason is the more complicated road map used in [42], with every road having a different weight; in that model, there is a higher probability of traffic overload in a cell. Still, for the purposes of our work, which aims at comparing different scheduling algorithms in terms of the user QoS and the fairness they achieve, our simpler road model, taken from [44], is a better choice in order to reach important conclusions without increasing the computational complexity of the system.

6.2.3 FCFS-EDF-SJF results

Figures 6.16- 6.20 present results derived with FCFS (First Come First Served), EDF (Earliest Deadline First) and SJF (Shortest Job First) algorithms. For these results we used the same “scenarios” for channel load that we created for the corresponding figures in the previous section.

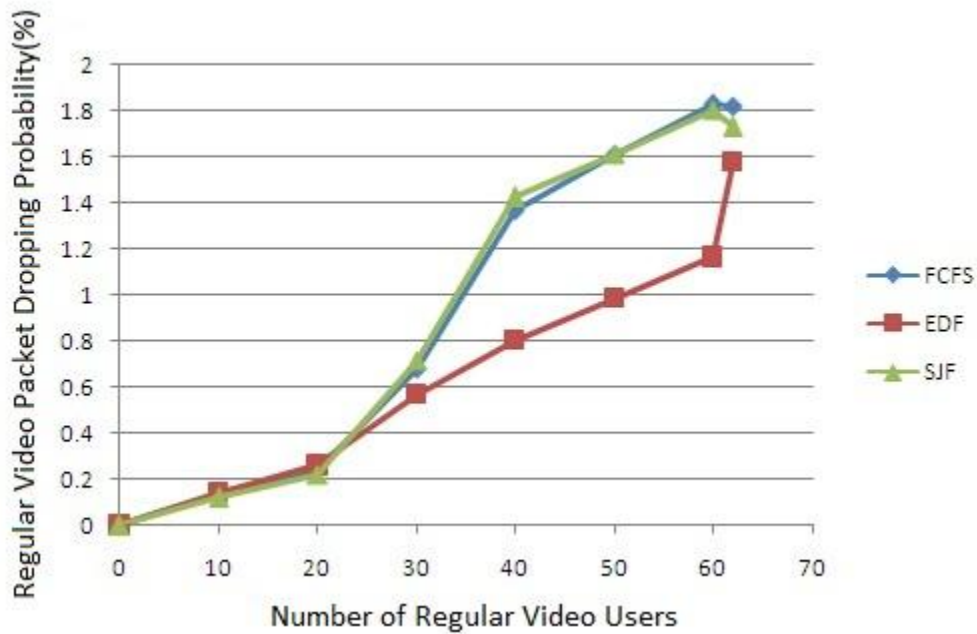


Figure 6.16: Regular video packet dropping versus the number of regular video users

Figure [6.16](#) shows that the increase in the number of regular video users clearly affects that type of traffic, because the video-packet-dropping probability of regular video users significantly rises above the 1% acceptable upper bound. Initially, the three algorithms have almost the same efficiency. However, when the number of users exceeds 20, the EDF algorithm turns out better video packet dropping probability results for regular video users, with FCFS being marginally better than SJF overall.

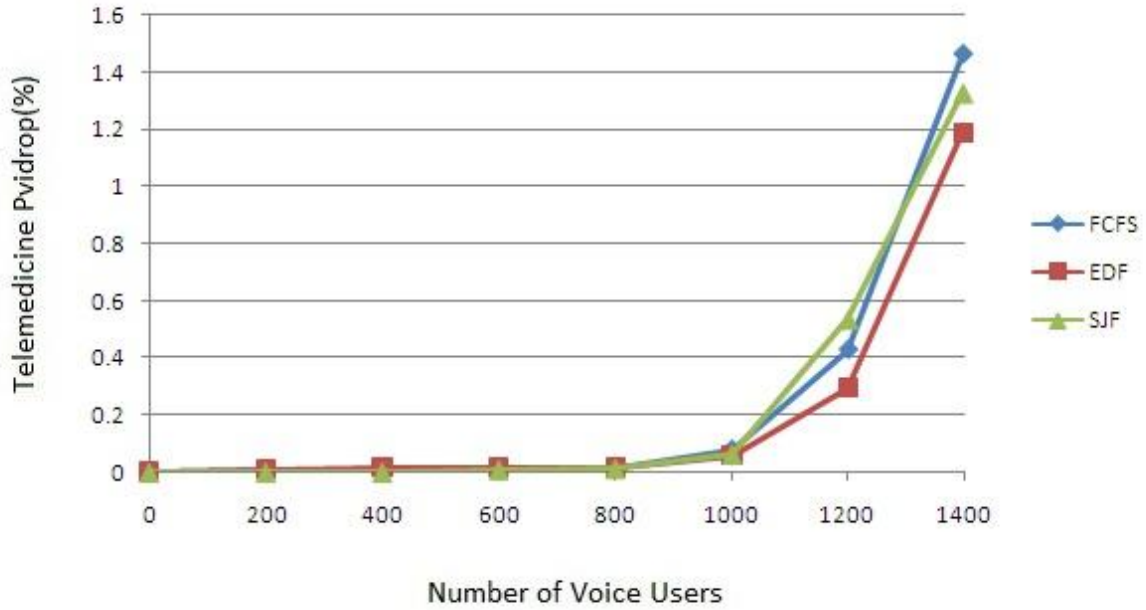


Figure 6.17: Effect of regular voice traffic on telemedicine video traffic

Figure [6.17](#) shows once again that an increase in the number of voice users does not influence telemedicine video packet dropping whatever the choice of algorithm. Only in the case of very high voice loads can there be deterioration in telemedicine traffic QoS. The reason, again, is that our combined scheduling and adaptive bandwidth reservation schemes guarantee full priority to all types of telemedicine traffic. For very high voice loads, EDF achieves the lowest telemedicine-video-dropping probability.

Figures [6.18](#) and [6.19](#) show that the three queuing priority algorithms produce almost identical results for low and medium telemedicine video loads, in

terms of regular and telemedicine video packet dropping. EDF excels once again in the case of high loads.

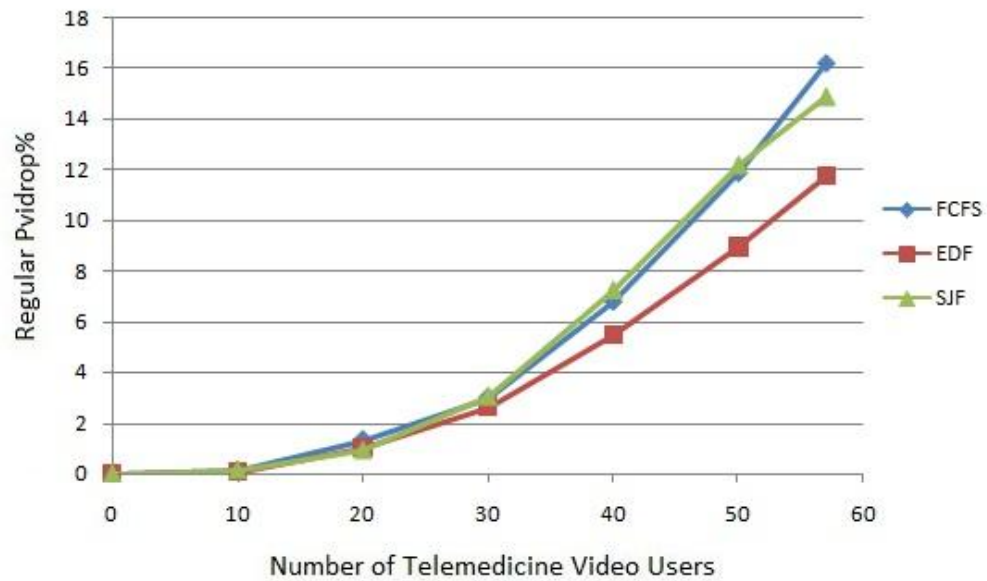


Figure 6.18: Effect of telemedicine video traffic on regular video traffic

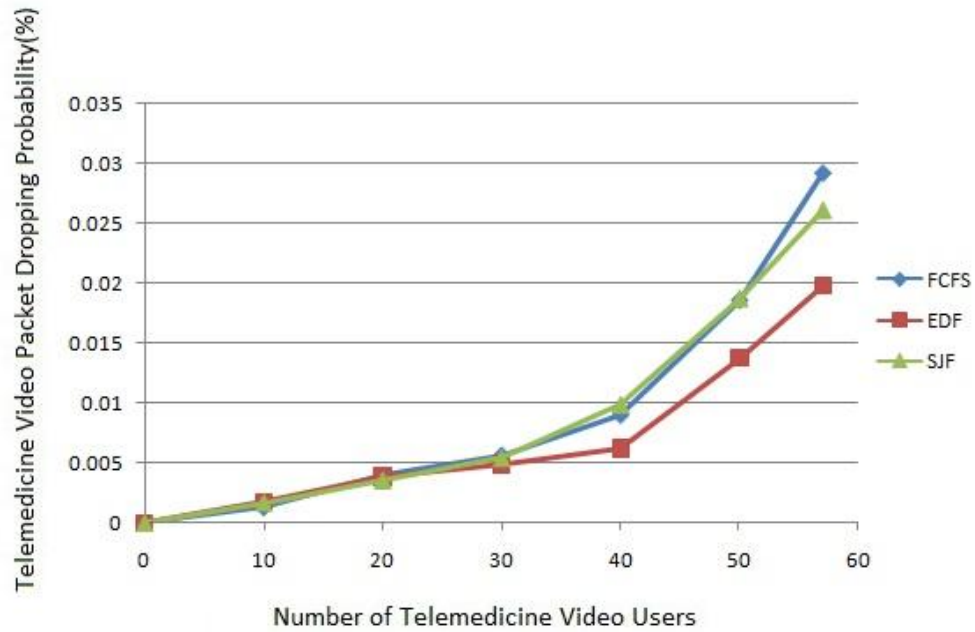


Figure 6.19: Telemedicine video packet dropping versus the number of telemedicine video users

Figure [6.20](#) shows that, regardless of the telemedicine video users' load, the three algorithms have comparable performance in terms of the X-Ray transmission delay.

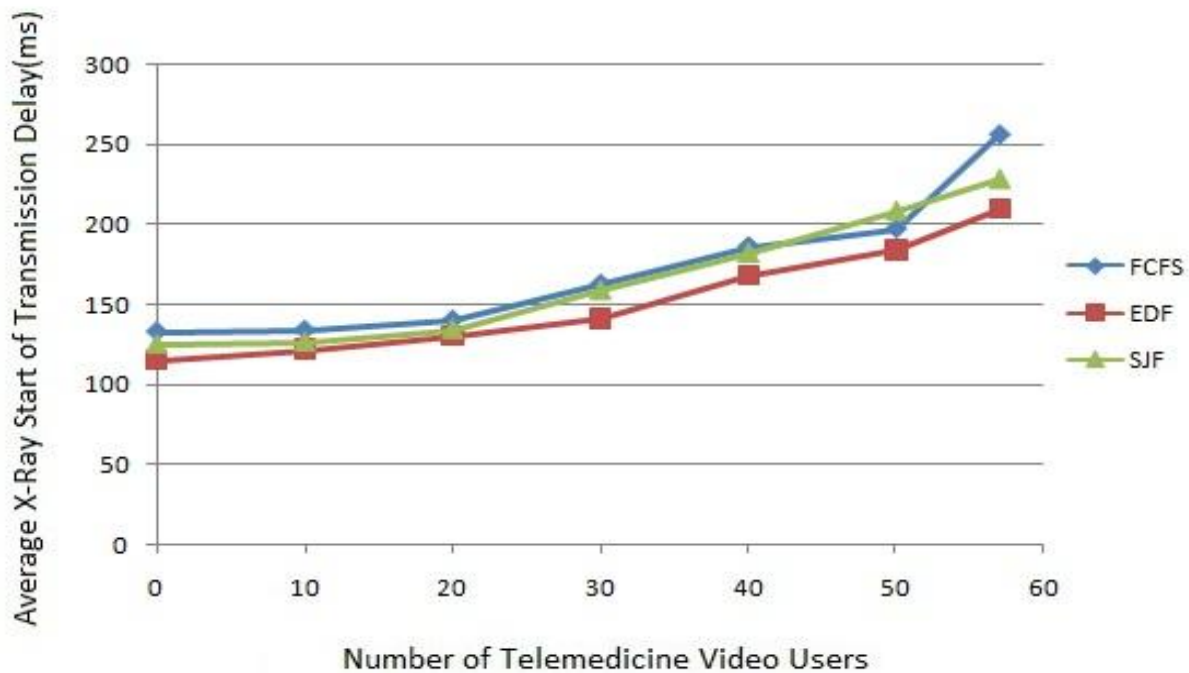


Figure 6.20: Effect of telemedicine video traffic on X-ray traffic

It is clear from all the figures presenting the results with the use of the three algorithms that the best results are achieved by EDF. The reason is that it accommodates users based on their deadline, hence it manages to satisfy their QoS requirements in time. The FCFS and SJF algorithms base their respective policies on the arrival time and the size of new information, thereby ignoring the urgency of users' needs. The more aggressive policy implemented by EDF, of course, has the disadvantage that it can lead to unfairness; users with later deadlines can experience longer delays than they do, e.g., with the use of FCFS. However, our results on Section 6.2.5 will show that this tradeoff is of almost negligible importance. Finally, we need to note that, although FCFS and SJF have comparable results for almost all traffic loads, in very high load conditions FCFS

performs always worse than SJF because it fails to offer any kind of priority based on the users' needs.

6.2.4 Open cells results

In this section we indicatively present results derived for the second road map that we created. All cells are connected in a straight line and mobiles can get in or out of our system only via cell A and cell G, which are placed at the two edges of the network.

The first figure illustrates the increase of regular video packet dropping probability and the second one the increase of telemedicine video packet dropping probability, as we increase the number of telemedicine video users. It is useful to compare and contrast figures [6.21](#) and [6.22](#) to the corresponding figures [6.18](#) and [6.19](#), where we used the first map.

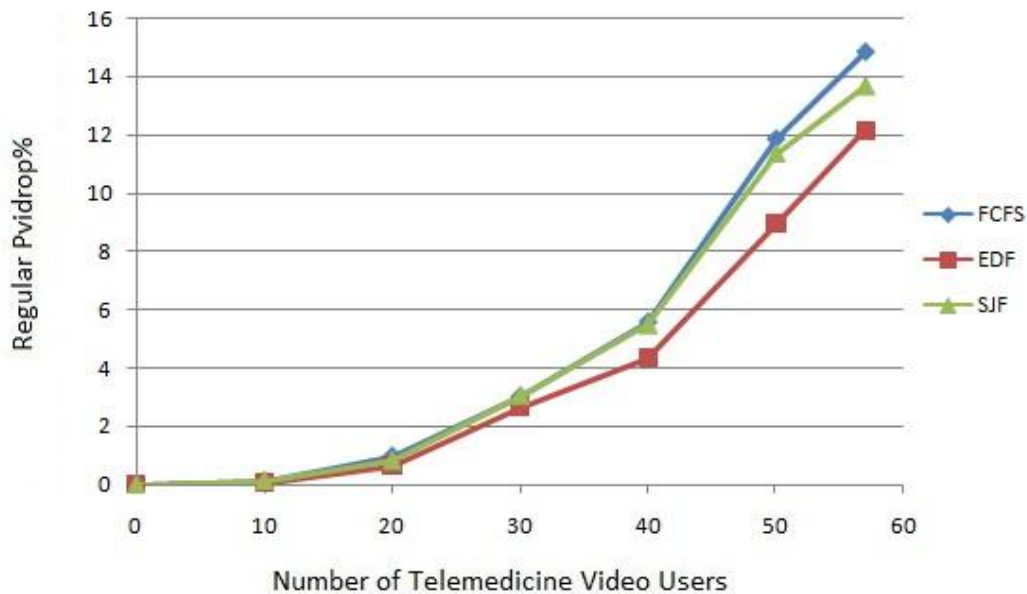


Figure 6.21: Effect of telemedicine video traffic on regular video traffic

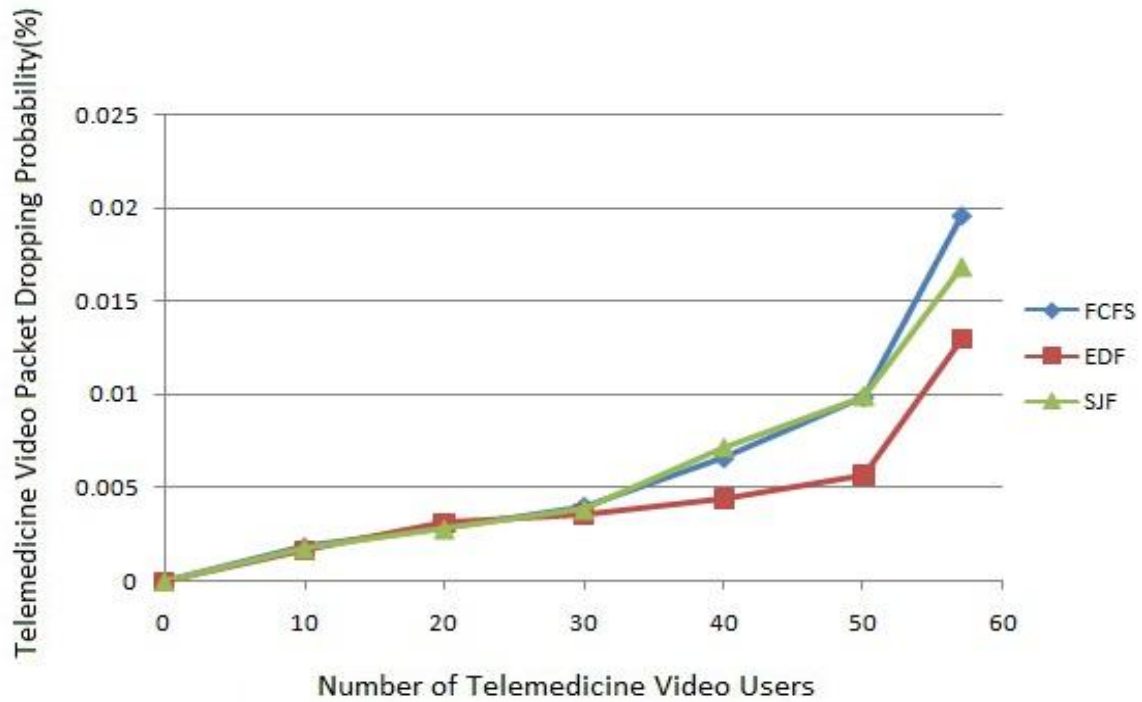


Figure 6.22: Telemedicine video packet dropping versus the number of telemedicine video users

Comparing figures [6.18](#) and [6.21](#), we note that they are similar in nature, but in figure [6.18](#) the maximum regular video packet dropping probability exceeded 16% with the use of the FCFS algorithm while in figure [6.21](#) it does not surpass 15%. The same 1% difference is shown for SJF and a 0.5% difference for EDF. The respective difference is much larger, as a percentage, when comparing the results in figures [6.19](#) and [6.22](#). Using the first map, telemedicine video packet dropping probability was 0.03% with FCFS, 0.026% with SJF and 0.02% with EDF, while with the second map the corresponding values are 0.02%, 0.017% and 0.0128%.

Our conclusion from this comparison is the following: when cells are “open” and users can go out of the system through cells A and G, regular video packet dropping probability and telemedicine video packet dropping probability are much lower. The reason is that when the seven cells are connected in a circular fashion, users never leave the system; hence, their aggregate number steadily grows, considering that new users keep arriving in the system. On the contrary, when users can leave the system via cells A and G, the aggregate users’ number fluctuates. This explains the worse QoS achieved by all algorithms in the case of the first map; this case actually represents a worst-case scenario for our system.

6.2.5 Jain’s Fairness results

Below, we have created 10 different combinations of all loads of users from very low (9%) to very high (100%) and we present the fairness results with the use of Jain’s fairness index for voice, video and telemedicine video users.

✓ Throughput-Based Fairness results

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9510	0.9513	-
Video Fairness	0.9978	0.9981	0.9976
Televideo Fairness	1.0000	1.0000	1.0000

Table 6.1:9% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9478	0.9474	-
Video Fairness	0.9979	0.9980	0.9977
Televideo Fairness	0.9971	0.9975	0.9970

Table 6.2:15% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9394	0.9382	-
Video Fairness	0.9919	0.9906	0.9903
Televideo Fairness	0.9968	0.9971	0.9967

Table 6.3:43% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9476	0.9470	-
Video Fairness	0.9868	0.9889	0.9859
Televideo Fairness	0.9963	0.9961	0.9965

Table 6.4:51% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9447	0.9497	-
Video Fairness	0.9866	0.9894	0.9857
Televideo Fairness	0.9957	0.9958	0.9958

Table 6.5:58% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9738	0.9638	-
Video Fairness	0.9862	0.9879	0.9840
Televideo Fairness	0.9948	0.9955	0.9933

Table 6.6:79% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9850	0.9884	-
Video Fairness	0.9862	0.9883	0.9853
Televideo Fairness	0.9893	0.9897	0.9888

Table 6.7:83% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9794	0.9879	-
Video Fairness	0.9869	0.9879	0.9845
Televideo Fairness	0.9937	0.9944	0.9932

Table 6.8:88% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9924	0.9899	-
Video Fairness	0.9889	0.9886	0.9860
Televideo Fairness	0.9933	0.9942	0.9928

Table 6.9:92% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9838	0.9925	-
Video Fairness	0.9876	0.9901	0.9866
Televideo Fairness	0.9920	0.9931	0.9902

Table 6.10:100% traffic load

As we mentioned in Section 4.4, we did not apply the SJF algorithm for voice users. The results presented in Tables 6.1-6.10 show that EDF once again outperforms FCFS and SJF in terms of fairness, something intuitively explained due to the nature of the algorithm. FCFS outperforms SJF, which creates some unfairness by always servicing first the users with the least packets to transmit, and hence leading larger transmissions to possible packet dropping. Still, the use of throughput as a fairness metric does not suffice; delays and packet dropping are of equal importance, therefore we used them in our study and we present the results below.

✓ *Delay-Based Fairness results:*

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9442	0.9430	-
Video Fairness	0.9973	0.9959	0.9692
Televideo Fairness	0.8529	0.8527	0.7726

Table 6.11:9% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9502	0.9462	-
Video Fairness	0.9948	0.9845	0.9552
Televideo Fairness	0.9928	0.9927	0.8637

Table 6.12:15% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9460	0.9393	-
Video Fairness	0.9884	0.9848	0.9522
Televideo Fairness	0.9916	0.9906	0.8318

Table 6.13:43% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9463	0.9436	-
Video Fairness	0.9887	0.9868	0.9439
Televideo Fairness	0.9894	0.9903	0.8437

Table 6.14:51% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9489	0.9465	-
Video Fairness	0.9824	0.9745	0.9299
Televideo Fairness	0.9900	0.9895	0.7590

Table 6.15:79% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9560	0.9452	-
Video Fairness	0.9859	0.9799	0.9233
Televideo Fairness	0.9886	0.9880	0.8291

Table 6.16:83% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9760	0.9628	-
Video Fairness	0.9804	0.9787	0.9268
Televideo Fairness	0.9883	0.9849	0.8518

Table 6.17:88% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9512	0.9438	-
Video Fairness	0.9857	0.9703	0.9248
Televideo Fairness	0.9886	0.9770	0.8625

Table 6.18:92% traffic load

User's type/Algorithm	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
Voice Fairness	0.9489	0.9471	-
Video Fairness	0.9653	0.9503	0.9225
Televideo Fairness	0.9869	0.9867	0.8451

Table 6.19:100% traffic load

As we observe from the above Tables, the FCFS algorithm is the fairest of the three as far as delay is concerned. SJF introduces the longest delay, since users with large numbers of packets to transmit are forced to wait for smaller transmissions to finish. EDF presents a slightly higher delay than FCFS, because by serving first users who have the shortest deadline, it incurs higher delays to users whose deadline is not imminent. Finally, we present three more cases, one with a medium load (60%), one with a high load (80%) and one corresponding to a traffic

overload (110% of the channel capacity) and we evaluate fairness in terms of delay and packet dropping.

✓ Delay and Packet Dropping-Based results:

	DELAY			P D R O P		
User's type/Algorithm	FCFS	EDF	SJF	FCFS	EDF	SJF
Voice Fairness	0.9461	0.9430	-	0.9383	0.9186	-
Video Fairness	0.9915	0.9910	0.9314	0.9868	0.9841	0.9193
Televideo Fairness	0.9926	0.9918	0.9455	0.9916	0.9912	0.8766

Table 6.20: 60% traffic load

	DELAY			P D R O P		
User's type/Algorithm	FCFS	EDF	SJF	FCFS	EDF	SJF
Voice Fairness	0.9369	0.9327	-	0.9138	0.9094	-
Video Fairness	0.9799	0.9753	0.9226	0.9338	0.9203	0.8248
Televideo Fairness	0.9906	0.9893	0.9317	0.9893	0.9824	0.8465

Table 6.21: 80% traffic load

	DELAY			P D R O P		
<i>User's type/Algorithm</i>	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>	<i>FCFS</i>	<i>EDF</i>	<i>SJF</i>
<i>Voice Fairness</i>	0.8977	0.8813	-	0.8948	0.8723	-
<i>Video Fairness</i>	0.8858	0.8552	0.7912	0.6937	0.6779	0.6219
<i>Televideo Fairness</i>	0.9727	0.9788	0.8864	0.9551	0.9501	0.8065

Table 6.22: 110% traffic load

For the results presented in Tables 6.20-6.22, we have kept constant the number of voice users and we increased the number of video and telemedicine video users. Hence, the 60% load in Table [6.20](#) consists of 40% voice, 17% video and 3% telemedicine video, the 80% load in Table [6.21](#) consists of 40% voice, 35% video and 5% telemedicine video and the 110% load in Table [6.22](#) consists of 40% voice, 60% video and 10% telemedicine video. These Tables clearly show that as we increase the channel load, all three algorithms' fairness decreases significantly, especially when considering packet dropping as a metric. The traffic type that is influenced the most is regular video traffic, as voice is less demanding in bandwidth, and telemedicine video users have absolute priority in scheduling, therefore they are minimally affected. It is clear, however, from all the results presented in Tables 6.1-6.22, that our scheme achieves excellent fairness results for all traffic types and for all traffic loads that do not exceed the maximum channel capacity.

Chapter 7: Conclusion and Future Work

This thesis has focused on the problem of scheduling integrated traffic transmissions from urgent types of traffic, like telemedicine, with regular wireless traffic over next generation cellular networks. We have extended a recent work [42] on a new MAC protocol by using three different scheduling algorithms over a simple network topology and by evaluating the algorithms' fairness based on a number of metrics.

Our results have clearly shown that the EDF algorithm excels over SJF and FCFS, which has been widely used in the literature as it is the intuitively simplest choice and is marginally fairer than EDF.

In future work, we intend to experiment with more scheduling algorithms from the literature and to propose an algorithm of our own, which will incorporate the advantages of EDF and will provide increased fairness in comparison to FCFS. In order to achieve this, we believe that we will need periodic bandwidth reallocation from the BS to the wireless users, based on efficient traffic modeling.

Bibliography

- [1] D. A. Perednia and A. Allen, "Telemedicine Technology and Clinical Applications," *the Journal of the American Medical Association (JAMA)*, vol. 273, no. 6, pp. 483-488, Feb 1995.

- [2] Y. Chu and A. Ganz, "A Mobile Teletrauma System Using 3G networks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 4, pp. 456-462, 2004.

- [3] J. R. Gallego, A. Hernandez-Solana, M. Canales, J. Lafuente, A. Valdovinos, and J. Fernandez-Navajas, "Performance Analysis of Multiplexed Medical Data Transmission for Mobile Emergency Care over the UMTS channel," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 13-22, 2005.

- [4] A. Bhargava, M. F. Khan, and A. Ghafoor, "QoS management in Multimedia Networking for Telemedicine Applications," in *Proceedings of the IEEE Workshop on Software Technologies for Future Embedded Systems*, pp. 39-42, 2003.

- [5] B. Tulu, S. Chatterjee, and S. Laxminarayan, "A Taxonomy of Telemedicine Efforts with Respect to Applications, Infrastructure, Delivery Tools, Type of Setting and Purpose," in *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS)*, 2005.

- [6] S.C. Voskarides, C. S. Pattichis, R. Istepanian, C. Michaelides, and C. N.Schizas, "Practical Evaluation of GPRS use in a Telemedicine System in Cyprus," in *Proceedings of the 4th IEEE International EMBS Special Topic Conference on Information Technology Applications in Biomedicine*, pp. 39-42, 2003.
- [7] W. Yao, R. Istepanian, A. Salem, H. Zisimopoulos, and P. Gosset, "Throughput Performance of a WCDMA system Supporting Tele-echography Services," in *Proceedings of 4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine*, pp. 310-313, April 2003.
- [8] S.Garawi, R. S. H. Istepanian, and M. A. Abu-Rgheff, "3G Wireless Communications for Mobile Robotic Tele-Ultrasonography Systems," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 91-96, 2006.
- [9] E. A. V. Navarro, J. F. N. J. R. Mas, and C. P. Alcega, "Performance of a 3G-Based Mobile Telemedicine System," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 1023-1027, 2006.
- [10] C. H. Salvador, M. P. Carrasco, M. A. G. de Mingo, A. M. Carrero, J. M. Montes, L. S. Martin, M. A. Caverio, I. F. Lozano, and J. L. Monteagudo, "Airmed-Cardio: A GSM and Internet Services-Based system for Out-of-Hospital Follow-up of Cardiac Patients," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 73-85, 2005.

- [11] S. Pavlopoulos, E. Kyriacou, A. Berler, S. Dembeyiotis, and D. Koutsouris, "A Novel Emergency Telemedicine System Based on Wireless Communication Technology -AMBULANCE," *IEEE Transactions on Information Technology in Biomedicine*, vol. 2, no. 4, pp. 261-267, 1998.
- [12] C.-Y. Huang and S.-G. Miaou, "Transmitting SPIHT Compressed ECG Data over a Next-Generation Mobile Telecardiology Testbed," in *Proceedings of the 23rd IEEE International Conference of the Engineering in Medicine and Biology Society*, 2001.
- [13] P. Koutsakis, S. Psychis, and M. Paterakis, "Integrated Wireless Access for Videoconference from MPEG-4 and H.263 video coders with Voice, Email and Web traffic," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 1863-1874, 2005.
- [14] Q. Pang, A. Bigloo, V. C. M. Leung, and C. Scholefield, "Service Scheduling for General Packet Radio Service Classes," in *Proceedings of IEEE Wireless Communications and Networking Conference*, 1999.
- [15] 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.
- [16] ETSI. Digital Cellular Telecommunications System (Phase 2+); Technical Realization of the Short Message Service (SMS); Point-to-Point (PP). (GSM 03.40)

[17] [Online]<http://www.heitec.sk/EN/xml/heikon.htm>

[18] F. H. P. Fitzek and M. Reisslein, "MPEG-4 and H.263 Video Traces for Network Performance Evaluation," *IEEE Network*, vol. 15, no. 6, pp. 40-54, 2001.

[19] C.-F. Tsai, C.-J. Tsang, F.-C. Ren, and C.-M. Yen, "Adaptive Radio Resource Allocation for Downlink OFDMA/SDMA systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2007.

[20] A. C. Cleary and M. Paterakis, "An investigation of Stack Based Algorithms for Voice Packet Transmission in Microcellular Wireless Environments," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 1995.

[21] 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, no. 3, pp. 584-598, 1991.

[22] D. Dyson and Z. J. Haas, "A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM," *Mobile Networks and Applications (MONET) Journal*, vol. 4, no. 2, pp. 87-99, 1999.

[23] E. N. Gilbert, "Capacity of a Burst-noise Channel," *Bell System Technical Journal*, vol. 39, pp. 1252-1265, 1960.

[24] E. O. Elliot, "Estimates of Error Rates for Codes on Burst-noise Channels," *Bell Labs Technical Journal*, vol. 42, pp. 1977-1997, 1963.

- [25] A. Willig, "A New Class of Packet- and Bit-level Models for Wireless Channels," in *Proceeding of the 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communication (PIMRC)*, 2002.
- [26] H. S. Wang and N. Moayeri, "Finite State Markov channel - A useful Model for Radio Communications Channels," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163-171, 1995.
- [27] M. Hassan, M. M. Krunz, and I. Matta, "Markov-based Channel Characterization for Tractable Performance Analysis in Wireless Packet Networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 821-831, 2004.
- [28] M. Bottigliengo, C. Casetti, C.-F. Chiasserini, and M. Meo, "Short-term Fairness for TCP flows in 802.11b WLANs," in *Proceedings of IEEE Infocom*, 2004.
- [29] E. C. Baig, "Will Consumers Tune in to a Tiny TV in their hand?", [Online]: http://www.usatoday.com/tech/wireless/2006-08-17-mobile-tv_x.htm, August 2006.
- [30] 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, pp. 1-13, 1993.

- [31] Y. Li, "Pilot-symbol Aided Channel Estimation for OFDM in Wireless Systems," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 4, pp. 1207-1215, 2000.
- [32] T. Holliday, A. Goldsmith, and P. Glynn, "Wireless Link Adaptation Policies: QoS for Deadline Constrained Traffic with Imperfect Channel Estimates," in *Proceedings of the International Conference on Communications (ICC)*, vol. 5, pp. 3366-3371, 2002.
- [33] F. R. Yu, V. W. S. Wong, and V. C. M. Leung, "A New QoS Provisioning Method for Adaptive Multimedia in Wireless Networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 3, pp. 1899-1909, 2008.
- [34] Q. Song and A. Jamalipour, "A Negotiation-Based Network Selection Scheme for Next-Generation Mobile systems," in *Proceedings of the IEEE Globecom*, 2006.
- [35] K. L. Yeung and S. Nanda, "Channel Management in Microcell/Macrocell Cellular Radio Systems," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 601-612, Nov 1996.
- [36] S. Choi and K. Shin, "Adaptive Bandwidth Reservation and Admission Control in QoS-sensitive Cellular Networks," *IEEE Transactions on Vehicular Technology*, vol. 13, pp. 882-897, Sep 2002.

- [37] D.-S. Lee and Y.-H. Hsueh, "Bandwidth-reservation Scheme based on Road Information for Next-generation Cellular Networks," *IEEE Transactions on Vehicular Technology*, vol. 53, pp. 243-252, Jan 2004.
- [38] R. Zander and J. M. Karlsson, "A Rate-Based Bandwidth Borrowing and Reservation Scheme for Cellular Networks," in *Proceedings of the 60th IEEE Vehicular Technology Conference*, vol. 2, pp. 1123-1128, Sep 2004.
- [39] N. Nasser, "Enhanced Blocking Probability in Adaptive Multimedia Wireless Networks," in *Proceedings of the 25th IEEE International Conference on Performance, Computing, and Communications*, 2006.
- [40] W. S. Soh and H. S. Kim, "A Predictive Bandwidth Reservation Scheme using Mobile Positioning and Road Topology Information," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, pp. 1078-1091, Oct 2006.
- [41] Raj Jain, "The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling," Wiley-Interscience, New York, NY, April 1991.
- [42] L. Qiao and P. Koutsakis, "Adaptive Bandwidth Reservation and Scheduling for Efficient Wireless Telemedicine Traffic Transmission", *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 2, 2011, pp. 632-643.

- [43] I. E. Lamprinos, A. Prentza, E. Sakka, D. Koutsouris ,“A Low Power Medium Access Control Protocol for Wireless Medical Sensor Networks” , Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, USA, September 1-5, 2004.
- [44] T. Kwon, Y. Choi, C. Bisdikian and M. Naghsineh, “QoS Provisioning in Wireless/Mobile Multimedia Networks Using an Adaptive Framework”, Wireless Networks, Vol. 9, no. 1, 2003, pp. 51-59.
- [45] P. Malindi, “Methods for providing rural telemedicine with quality video transmission”, Cape Peninsula University of Technology, CPUT Theses & Dissertations,paper 243, 1-1-2007.
- [46] C. Chigan and V. Oberoi, "QoS Provisioning in Sensor Enabled Telemedicine Networks", presented at IJHISI, 2007, pp.12-30.
- [47] F. Hu and Sunil Kumar, “QoS Considerations in Wireless Sensor Networks for Telemedicine”, in Proceedings of SPIE ITCOM Conference,Orlando,FL , 2003.
- [48] N. Challa, H. Çam and M. Sikri, “Secure and Efficient Data Transmission over Body Sensor and Wireless Networks”, EURASIP J. Wireless Comm. and Networking, March 2008.