

Master Thesis
Electrical Engineering
Thesis no: MEE10:44
June 2010



Impact of Packet Losses on the Quality of Video Streaming

JOHN Samson Mwela
&
OYEKANLU Emmanuel Adebomi

School of Computing
Blekinge Institute of Technology
Box 520
SE – 371 79Karlskrona
Sweden

Internet : www.bth.se/com
Phone : +46 457 38 50 00
Fax : + 46 457 271 25

This thesis is submitted to the School of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:

Author(s):

JOHN Samson Mwela

Blekinge Institute of Technology

E-mail: mwelasj@yahoo.com

OYEKANLU Emmanuel Adebomi

Blekinge Institute of Technology

E-mail: manuelbomi@yahoo.com

Supervisor

Tahir Nawaz Minhas

School of Computing

Examiner

Dr. Patrik Arlos, PhD

School of Computing

School of Computing

Blekinge Institute of Technology

Box 520

SE – 371 79 Karlskrona

Sweden

ABSTRACT

In this thesis, the impact of packet losses on the quality of received videos sent across a network that exhibit normal network perturbations such as jitters, delays, packet drops etc has been examined.

Dynamic behavior of a normal network has been simulated using Linux and the Network Emulator (NetEm). Peoples' perceptions on the quality of the received video were used in rating the qualities of several videos with differing speeds.

In accordance with ITU's guideline of using Mean Opinion Scores (MOS), the effects of packet drops were analyzed. Excel and Matlab were used as tools in analyzing the peoples' opinions which indicates the impacts that different loss rates has on the transmitted videos. Statistical methods used for evaluation of data are mean and variance.

We conclude that people have convergence of opinions when losses become extremely high on videos with highly variable scene changes.

Keywords: packets, CODEC, streaming, Mean Opinion Score.

ACKNOWLEDGMENTS

Praise be to God the Almighty for abundant blessings poured into my heart at different levels of my studies. I believe all the successes I obtained in my life are not by my own powers but they are graces from God. I am also thankful to my grandmother Cassiana Salama for her teachings on tolerance and endurance. I thank my priest Canon Julius Lugendo for his spiritual services especially at the time of discouragement. My children Kihomo and JP have braved a long difficult time without their father believing in what I pursue. Let this opens up doors of your educational successes. I thank you all.

Samson

Deep appreciation goes to Jehovah my God the Almighty for being with me all the days of my life. Jehovah is gently teaching me, leading me by the hands and directing my affairs. Through thick and thin, He has been a God that never gives up on his loyal ones. He has given me the grace to achieve, the courage to excel, the wisdom to progress and abundant understanding more than my contemporaries. I really appreciate his Godly steadfastness, love, humility and his support. One thing I have asked from Jehovah-it is what I have looked for, that I may dwell in the house of Jehovah all the days of my life, to behold his pleasantness and to look with appreciation upon his temple.

Emmanuel

We also hold in his esteem our indefatigable lecturer in person of Dr. Patrik Arlos from whom we have gained so much both in all the courses he taught us and in this thesis. Dr. Arlos gave us time latitude and intuition; he asked deeply incisive question that makes us discover hidden knowledges and values. Overall, we shall never forget all the valuable trainings we have received from him. Our Supervisor, Mr. Tahir Nawaz Minhaz has also given much in terms of time, attention and knowledge impartation. His inputs have proven to be quite invaluable over time. We deeply appreciate him. We pray that God Almighty continue to bless these ones and their families.

Emmanuel and Samson

CONTENTS

IMPACT OF PACKET LOSS ON THE QUALITY OF VIDEO STREAM TRANSMISSION...	I
ABSTRACT	II
ACKNOWLEDGMENTS.....	III
CONTENTS.....	IV
LIST OF FIGURES.....	V
LIST OF TABLES.....	VI
LIST OF ABBREVIATIONS.....	VII
1 INTRODUCTION.....	1
1.1 ROLES OF TRANSPORT PROTOCOLS ON VIDEO TRANSMISSION	1
1.2 RESEARCH QUESTIONS.....	2
1.3 AIMS AND OBJECTIVES	2
1.4 EXPECTED OUTCOMES.....	2
1.5 RESEARCH METHODOLOGY.....	3
1.6 THESIS OUTLINE.....	3
2 BACKGROUND AND RELATED WORK.....	4
2.1 COMPARISON TO THIS WORK	6
3 DESIGN AND IMPLEMENTATION:.....	7
3.1 REQUIREMENTS:.....	7
3.2 ARCHITECTURE	7
3.3 SYSTEM DESIGN	8
3.3.1 <i>Network Emulation and Traffic Control Overview</i>	9
3.3.2 <i>Client</i>	9
3.3.3 <i>Server</i>	9
3.3.4 <i>Evaluation and Selection of Simulation and Experimental Tools</i>	10
3.3.5 <i>Operation of VLC</i>	11
3.4 IMPLEMENTATION.....	12
3.4.1 <i>Server</i>	13
3.4.2 <i>NetEm's configuration</i>	13
3.4.3 <i>Measurement Point Setup</i>	14
3.4.4 <i>Consumer</i>	14
3.4.5 <i>Client</i>	14
3.5 TESTING OF THE SYSTEM.....	14
4 DATA COLLECTION AND ANALYSIS.....	15
4.1 MEAN AND VARIANCE	15
4.2 DATA COLLECTION	16
4.3 DATA ANALYSIS.....	16
5 RESULTS AND DISCUSSION.....	17
5.1 EFFECT OF PACKET SIZES ON VIDEO QUALITY	18
5.2 MEAN OPINION SCORE FOR DIFFERENT TYPES OF VIDEOS	19
5.3 AMOUNT OF PACKET LOSSES LEADING TO USERS' REJECTION OF VIDEOS	19
5.4 DIFFERENT PEOPLE'S OPINIONS ON THE SAME VIDEO	21
6 CONCLUSION.....	23
7 FUTURE WORK	24
8 APPENDICES.....	25
9 REFERENCE.....	31

LIST OF FIGURES

Figure 1: Model representation in block diagram	7
Figure 2: Client- server communications.....	8
Figure 3: TCP/IP reference model representing packet transmission on different layers.....	8
Figure 4: UDP message format [18].....	10
Figure 5: Ethernet frame format [19]	10
Figure 6: IP datagram format [20].....	11
Figure 7: VLC and its main features [22]	12
Figure 8: MOS for all scenarios	17
Figure 9: Graph of MOS for packet loss in the range of 0 to 1 percent.....	18
Figure 10: MOS results for packet transmission loss of 10 percent	19
Figure 11: MOS graph for “Foreman” received video. Transmission packet size is 512 bytes	20
Figure 12: MOS graph for "Football" video. Transmission packet size is 512 bytes.....	20
Figure 13: MOS graph for "News" received video. Transmission packet size is 1024 bytes	21
Figure 14: Variance of Users' Perception	22
Figure 15: MOS graph for “Foreman” received video. Transmission packet size is 1500 bytes.....	28
Figure 16: MOS graph for “Foreman” received video. Transmission packet size is 1024 bytes.....	28
Figure 17: MOS graph for "Football" video. Transmission packet size is 1500 bytes	29
Figure 18: MOS graph for "Football" video. Transmission packet size is 1024 bytes.....	29
Figure 19: MOS graph for "News" received video. Transmission packet size is 1500 bytes	30
Figure 20: MOS graph for "News" received video. Transmission packet size is 512 bytes..	30

LIST OF TABLES

Table 1: Video and audio codec parameters settings.....	13
Table 2: Summarized description of packet loss.....	15
Table 3: Mean Opinion Score	17

LIST OF ABBREVIATIONS

AAC	– Advanced Audio Coding
AVC	– Advanced Video Coding
CDA	– Canonical Discriminant Analysis
CLI	– Command Line Interface
CODEC	– Coder/Decoder
DAG	– Digital Acquisition and Generation
DVD	– Digital Versatile Disc
GOP	– Group of Pictures
ITU	– International Telecommunication Union
JPEG	– Joint Producer Expert Group
LAN	– Local Area Network
MPEG	– Moving Picture Experts Group
MP	– Measurement Point
MS	– Mixed Streaming
MOS	– Mean Opinion Score
MTU	– Maximum Transmission Unit
NetEm	– Network Emulator
QoE	– Quality of users' Experience
QoS	– Quality of Service
RLC	– Radio Link Control
TC	– Traffic Control
TCP/IP	– Transmission Control Protocol/Internet Protocol
UDP	– Universal Datagram Protocol
UMTS	– Universal Mobile telecommunication System
VLC	– VideoLAN Client
VOD	– Video on Demand
WCDMA	– Wideband Code Division Multiple Access

1 INTRODUCTION

The need for more reliable transmission of real time video streams over the Internet has recently become more urgent due to the applicability of video streaming to many areas of human endeavor. Areas like telemedicine, coordination of relieve efforts in disaster areas, videoconferencing and coverage of football tournaments etc demands real-time video communications.

The proliferation of mobile telecommunications technologies like 3G and 4G has been quite useful in facilitating real time video transmission. Even though communication technology in a broad perspective is enjoying much innovation, real-time data transmission still rely on using the transport layer user datagram protocol (UDP) which although has an advantage of providing a fast video transmission in one hand but on the other hand is infamous in terms of the video quality because of packets loss, delay, jitter and out of order packet delivery. These are some problems of the video streaming over the UDP that requires urgent solutions due to the fact that users' perception and the required responses are often times dictated by the received video quality.

The invention of video coding technologies like H.264/AVC (Moving Picture Experts Group- MPEG 4 Part 10) has impacted positively on the industry. It is capable of producing good quality video when compared with previous CODEC (Coder/Decoder) versions such as MPEG 2. It can consistently produce about 50% better compression rate than the MPEG 2 [1]. It has achieved the most profound compression efficiency for many applications including streaming, multimedia communication and video compression in real time manner. What more, H.264 has achieved this feat without increasing the design complexity, thus making it practical and affordable.

H.264 works by making use of the redundancies in the video that it is encoding. The redundancies in question can be spatial or temporal redundancies. Spatial redundancies are those that exist within the same picture while temporal redundancies are those that abound between the pictures. Temporal redundancies make use of a technology known as motion estimation whereby each subsequent frame that constitute the video are subdivided into grids or block formats and the entire blocks are scanned for matching textures. Once a matching texture between two blocks is located, H.264 can later reproduce the texture of one of the blocks using pointing vectors that can locate the matched reference block. If however motion estimation cannot locate matching references, H.264 can make use of the textures of nearby frames to predict or suggest the texture of the next frame. This is spatial redundancy. In this regard, it is only the referenced block and the difference between the prediction and the actual block texture that will be stored [2].

The goal of H.264 is not to recreate the original picture in its exact format but to use an optimal means to achieve a reduction in the data rate with a focus on achieving laudable visual quality. It is a 'lossy' compressor. The H.264 audio version is the Advanced Audio Coding (AAC) also known as MPEG 4 Part 3. Both of them are technically MPEG 4 CODECS although in technical parlances, they are referred to with their specific ITU-T Standard given names.

1.1 Roles of Transport Protocols on Video Transmission

As laudable as the achievement of H.264 is for multimedia messaging, the gains are besotted by the fact that real time video transmission in most cases takes place on an unreliable UDP framework. The Transmission Control Protocol (TCP) would have been the best bet for real time transmission but its inherent design forces it to retransmit any dropped packet or any packet that the receiver does not acknowledge as being received as well as assuring the orderly delivery of sequence of video packets.

For real time video transmission, usage of TCP will eventually lead to a clogged network and the purpose of real time streaming will be defeated. This leaves UDP as the next prime candidate for real time transmissions. It does not retransmit any dropped packet or any un-acknowledge packet. It works in the best effort mode.

As a result, if any real time packet is dropped, it now devolves on the decoder at the receiver's end to suggest or recreate such packet so that the video viewer will have an appreciable Quality of user Experience (QoE). The granularity of video output that makes it possible to detect errors is delimited by the viewers' QoE. The H.264 is reputed to be the preferred CODEC for video streaming for real time applications, but its mode of operation which makes it store vectors that acts as pointers in lieu of storing the entire compressed video makes it susceptible to the shortcomings of UDP. For example, if a lost packet contains whole picture plus a vector that points to important information that may lead to the recreation of important segment of a video part, the effect of the loss will multiply through the entire video stream and the user at the receivers' end will inadvertently suffer reduced QoE.

How does the impact of UDP affect the output of H.264 videos? How do the lossy nature of H.264 and the best effort behavior UDP affect the QoE of the user at the receiver's end? In a nutshell, what are the impacts of losses on the quality of video transmission? These questions and sundry ones are what this thesis work is all about. In the thesis such questions as the effect of loss on different MTU sizes will be considered. The percentage of loss on particular MTU sizes that may lead to total rejections of a received video will also be ascertained. The result of the thesis may be useful for researchers in designing better methods by which users may enjoy better QoE.

1.2 Research Questions

A more focused approach on carrying out of the thesis project is paramount among the key factors required to successfully obtain the intended research outcomes. Outcomes of the research need to answer some of the well formulated research questions that lead to developing hypothesis and research methodologies. In this thesis, our findings will be mostly based on answering the following research questions.

- What is the relationship between packet loss and video quality?
- Is the size of the lost packet having any impact on video frame?
- What is the percentage of packet loss that may lead to total rejection of received videos at the users end?
- Is it possible to identify particular video frame at network layer in packet form?

1.3 Aims and objectives

The thesis goal is to analyze the effect of quality metrics such as packet loss on faithful transmission and reception of video packet over an Internet Protocol (IP) network. This will include:

- Simulating a transmission network that implements all the necessary details synonymous with a normal network
- Analyze and investigate the relationship between the packet loss and video quality.

1.4 Expected outcomes

Based on prevailing network quality metrics/parameters, the thesis report contains results in form of graphs detailing the comparison of amount of packet loss relating to the different videos used and also the received video quality as a result of users/viewers QoE. The document will also contain recommendations on the tolerable threshold of packets loss for improvement of video streams transmission.

1.5 Research Methodology

A Network Emulator (NetEm) will be used to simulate a normal dynamic behavior of communications network. Packet size and percentages of packet dropped will be varied during transmissions. The received video quality will be determined by interviewing people who will be watching a Video LAN Client (VLC) player.

1.6 Thesis Outline

The remaining parts of the thesis are organized as follows: Chapter two is a review of related work. Chapter three will focus on design and implementation, Chapter four will contain data analysis and a review of system's performance. Chapter five will feature results and discussions on thesis experimental observations. Chapter six, seven, eight and nine provide conclusion, future works, appendices and references respectively.

2 BACKGROUND AND RELATED WORK

A typical H.264 encoder yields three types of frames after the coding process, the Intra (I-Frame), the Prediction (P-Frame) and the Bi-directional (B-Frames) respectively. While information from the encoded still image is contained in I frames, P frames are directional in nature and they are a product of the previous B and I frames. The preceding P and I frames also produces the B frames. Each subsequent video sequence is a series of repeating streams of these frames aptly referred to as Group of Pictures (GOP) [3].

At the receiver, the success or otherwise of decoding a compressed bit streams is a function of receiving the reference I frames since they transmit the vectors that produces information serving as pointers to the locations of frames with similar characteristics. The loss of a reference frame may have a ripple effect on the entire GOP, if the lost frame contains information about the properties and locations of other frames. This makes errors in the reference frames to have more debilitating effects than having errors in the derived frames. If possible, the reference frames should be accorded a higher level of protection than other frames in the GOPs [4]. Significant performance gains are achievable if the frames that contain the most important data are received.

In some cases, video traces are used to solve the problems of evaluation of the video quality. Since it is impractical to return the entire received GOPs to the sender for comparison purpose with the original GOPs, so video traces has proven to be quite invaluable. Also worthy of note is that mechanisms that are based on retransmission strategies are infeasible for streaming purposes on the Internet due to the fact that an added round trip may culminate into too much latency for a reasonable seamless flow.

Conventionally, a video frame is conceptually larger in size when compared with the Maximum Transfer Unit (MTU) for a UDP packet. It is allowable for designers working with 3G wireless networks to select MTUs ranging from small values around 512 bytes to large values of about 1500 bytes for IP packets. Although video frame sizes can be much larger or smaller than the MTU, with the consequence that the Ethernet frames will have to be resized when they get to the IP level if their sizes differs from that allowed by the MTU. If the video frame sizes are smaller, they are promptly funneled onto the Internet but they are parsed down into the MTU size if they are larger [5].

One of the major effects of utilizing UDP for video streaming purposes is that bit errors will lead to whole packets getting rejected as a result of the checksum failure of the packet. The decoder will as a result, copy the parameters defining the previous frames into the present frame location. The viewers of such video will notice some artifacts that cause the video quality of that particular frame to suffer degradation. If the frames that suffers these artifacts are part of the very precious I or P frame, then the error will have a multiplier effect and will be propagated downstream of the affected frames.

In a bid to minimize the impact of packet loss on video stream transmission, many research works are ongoing in the research community. For instance, a research work by Ron et al [5] analyzed viewers' perceived quality of a compressed video stream transmitted over a lossy IP network with a quality of service mechanism. The parameters of the encoding schemes used include the transmission bit rate, compression depth, frame size and the frame rate. The result of the work showed that it is feasible to classify and predict the subjective quality of the video stream from set parameters by using linear discriminant analysis known as Canonical Discriminant Analysis (CDA).

2.1 Related works

In [1], the impact of frame rate on a secured real time transmission of video over IP networks was examined. The most suitable frame rate in order to achieve better data rates and fewer frame and packet loss was determined. [6] Presents the concept of mixed reliability video streaming, i.e. Mixed Streaming (MS) which reduces the impact of video propagation errors in error prone wireless networks. MS delivers a video file using both reliable and best efforts connections. Results shows that MS reduces the impact of errors by making sure that error on reference frames are corrected.

Mackenzie [7] et al investigated several mapping schemes for a variety of video content to see how the quality of the decoded video is affected as the numbers of concurrent video connections is increased. It was shown that as the numbers of concurrent video approaches the network capacity, some mapping schemes shows a cliff edge drop in quality while others offers more acceptable gradual quality degradation. Ulrich et al [8] investigated the perceptual image quality metrics for real-time quality assessment of Motion JPEG2000 (MJ2) video streams over wireless channels. A number of metrics were evaluated and their recommendation forms a basis for this thesis project.

It is proposed in [9] that throughput histograms can be used as a Quality-of-Service indicator and also further research is recommended in order to obtain quantitative results such as thresholds for sending signals to users or applications. During the course of this thesis, work will be done to analyze the proposed metrics, investigate and propose some useful parameters which may include thresholds as proposed in [9].

[10] Proposed the adaptation of the number of video frames transmitted per seconds based on the current load of the cell phone network and the data size of the video frames. A real time artificial intelligence based decision making component was added to the proposed algorithm. This was found to greatly reduce the number of dropped frames. However, the work did not explore the ability of the algorithm to reduce or increase the data size of video frames by altering the quality of the frames via an additional video compression parameter. In [11], an attempt was made to find a suitable packet size in order to achieve better video quality. It was found that the smaller the packet size the worse the video transmission quality.

[12] Investigates the performance of video streaming using MPEG 4 CODECS over Universal Mobile Telecommunications System (UMTS) channels. The channel is dedicated with the normally varying channel conditions. The simulation models used to evaluate the streaming performance caters for both unacknowledged and acknowledged mode. It was eventually deduced that the acknowledged mode outperforms the unacknowledged mode. The optimum Radio Link Control (RLC) size that should be employed determines the video quality. This RLC size is found to be dependent on the channels bit rate and the size of the encoded video frame. The unacknowledged mode does not provide for error recovery, this makes its performance to be poor since the MPEG 4 encoder is aware of error propagation and erroneous data. In acknowledged mode, there is an increment in end to end delay of the video frame, this enables the acknowledge mode to be able to support block error rates of upwards of 40% before total deterioration.

An analysis of the effects of channel properties on the quality of received video streams was carried out in [13] and [14]. The quality of transmitted H.264 video over a UMTS dedicated channel was compared to the quality obtained in a memory-less channel. A live UMTS channel was employed to ascertain the correctness of the simulation. It was deduced that if the resulting IP packet probability of error remains unchanged, then the error correlation properties of the network or link layer model will not exhibit any direct effect on the video stream quality. Furthermore, it is deducible that proper modeling of the error characteristics of the link layer is a profound

determinant of the quality of the received video streams. Works embarked on in [15] and [16] are similar to this approach. The major difference however is that while [15] focuses on mobility, [16] focuses more on adaptation and error correction.

In [4], an investigation of the packet scheduling function of a satellite system providing a unidirectional point to multipoint services to mobile users is carried out. The satellite system may be considered as an overlay broadcast or multicast system with frequency division duplex features (FDD) that may complement the functions of conventional point to multipoint third generation (3G) systems. The downlink channel characteristic of the satellite is synonymous with those obtainable on Wideband Code Division Multiple Access (WCDMA) radio access scheme. Enormous challenges were encountered in determining the optimal packet rates in view of perturbations introduced by UDP and satisfying the physical channel utilization that most satellite broadband service providers will want.

In [17], several techniques that provides for error resilient transmission of video streaming in view of several network disturbances over the wireless mobile network were examined. The focus is on the H.264/AVC standard. Consideration was given to the impact of the exploitation of residual redundancies in the video streams received at the decoder. The considered packets may otherwise have been rejected, but the redundancy provides for error localization and impairment detection. Additional video packets were not required to reduce the distortions. Results indicate that performance may be considerably improved if the variable length codes were re-synchronized. This technique may be aided by side information that is 'out-of-the-stream'. Additionally, techniques that borders on error concealment were considered. It was concluded that no single error concealment strategy performs to expectation in all considered situation. If scene changes occur, then temporal interpolation is found wanting, but otherwise, it performs better than spatial interpolation. As such, a method is proposed that hybridized the best parts of these two methods. This culminates in a low complexity approach that adapts a hybrid of the spatial and temporal interpolations to dynamically evolving situations.

A cross layer design approach for that utilizes UMTS radio link information at the receiver's application layer also leads to improved quality at practically no cost. Often, link layer errors of the UMTS radio link are predictable. As such, when the intervals between low and high error probabilities are differentiated, then priority levels can be assigned to video packets based on the semantics of the expected radio link perturbations.

2.2 Comparison to this work

Most of the research works reviewed in the preceding sections evaluates the quality of service (QoS) parameters that affect performance of the received multimedia contents. Some of these works have anchored their findings on compression/decompression commonly known as coding and decoding; and their effect on the quality of multimedia contents.

Other findings were geared towards analyzing the performance during transmission of the respective multimedia contents. Some other works were biased towards an evaluation of measurement metrics including subjective and objective analysis. Almost all the papers reviewed in this thesis analyze such QoS parameters like packet drops, delay, jitter etc. In this thesis, an attempt will be made to find the threshold of packet loss that could lead to total rejection of multimedia content by users.

Some of the methods proposed and implemented in the reviewed paper will also be used as implementation tools in this thesis.

3 DESIGN AND IMPLEMENTATION:

3.1 Requirements:

The study is based on an evaluation of different amount of packets that are dropped/ lost during transmission and also studying different packet sizes and how they affect quality of received video. Basically, consideration is given to the fact that packet losses during transmission are experienced in both wired and wireless channels, though the effects on wireless channels is much more pronounced due to its dynamic behavior.

In this project, a channel is modeled by placing a computer equipped with NetEm between the transmitting and the receiving computers. Different amount of packet losses are configured prior to starting of packet transmission which are channeled through an intermediate NetEm traffic shaping computer, dropping certain amount of packets and finally delivering the remaining packets to the intended destination.

Determination of the amount of packets sent before and those received after the NetEm traffic shaping is accomplished by using a measurement point (MP) that has a DAG (Data Acquisition and Generation) Card with two ports, one for capturing information from the transmission wiretap and the other one from the receiving wiretap. The MP then compare the amount of sent and received packets whose difference provides the amount of packet lost.

Three scenarios are implemented. The first scenario deals with a slow motion picture like news videos that has a relatively fast moving background. The second scenario involves intermediate motion videos which includes a construction site in which the foreman elaborates something and video pictures are changing to different views of the building. The last but not least is the fast motion video picture featuring a football game. It is anticipated that motion of videos can also affect the quality of video after packet loss.

Opinions on the received video quality for all scenarios with different packet losses and sizes are collected from a different people, which in turn will be used for analyzing the Quality of users Experience on the received video.

3.2 Architecture

Real-time video streaming system implemented in this work comprises of three major components namely streaming server, communication channel and a streaming receiver. Block diagram for the components of the video communications system is shown in Figure 1.

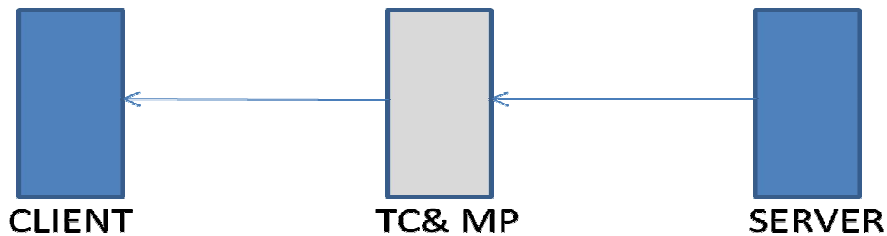


Figure 1: Model representation in block diagram

For the client to access video streaming resources from the server, it needs to initiate a request for the service which in this document is referred to as REQUEST command. The server then sends a REPLY for the requested video clips (figure 2). Between the server and client there is a traffic control and measurement point (TC & MP) system responsible for shaping traffics passing through it. One among the

activities performed by TC & MP system is to drop a certain amount of packets by using NetEm software.



Figure 2: Client- server communications

Shown in Figure 1 are the three major components of the client-server communication system for a video streaming. On the server side, video files are encoded and compressed by using a H.264 encoding technique. Before transmission, the size of each Ethernet frame is controlled at server's interface. As explained thoroughly in the succeeding design subsection, different packet sizes are set at the Data-link layer. TC & MP component is responsible for emulating the lossy transmission medium. Packets are randomly dropped based on the set packet drop values.

Receiver component is responsible for the reception and playback of the received video clips. Measurement Point (MP) is a computer system responsible for capturing the transmitted packets as well as the received ones and the difference of which is amount of packets dropped by NetEm.

3.3 System Design

Video streaming is a client- server communication system. Normally, videos run at the application layer of the OSI reference model (see Figure 3). Below the application layer, there exists Transport, Internet/ Network, Data Link and Physical layers that facilitate transmission of data from one computer or related system to another. For the client to access resources in the server, and the server to deliver resources to a client computer, there must be a socket that opens a port at the transport/ application layer's interface.

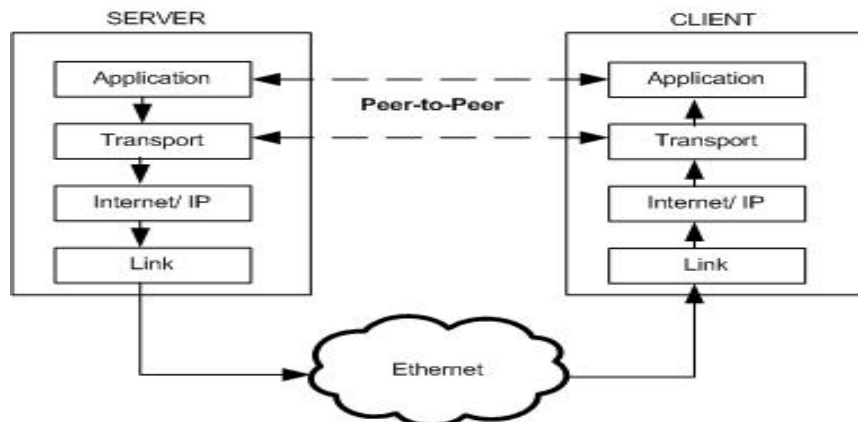


Figure 3: TCP/IP reference model representing packet transmission on different layers

In this design, a server is a computer that streams video whereas a client is a computer that access online video streams from server. Figure 3 gives an overview of the full client server communication system.

3.3.1 Network Emulation and Traffic Control Overview

NetEm operates with Linux and has variety of emulation functionalities. It is used in this thesis for emulating wide area network properties. Practically, a network has the inherent property of dropping, delaying, duplicating and reordering packets that are transmitted through it. To simulate this network behavior, NetEm is used.

For a NetEm to operate, a Linux command line tool known as traffic control abbreviated as “tc” is used. Some of the functionalities that NetEm provides include emulation of a wide area network delays, packet loss, packet corruption, packet duplication, and packet re-ordering. Packet loss function is a function used in this thesis project for dropping different amounts of packets basing on the configured parameters.

Traffic control is one of the powerful tools used for controlling transmission quality of service. It has three elements namely “classification”, “scheduling” and “queuing”. Queuing discipline in Linux combines both queuing and scheduling. The simplest and default queuing discipline in IP communication uses the first-in-first-out (FIFO) queuing discipline which operates in first-come first-served basis.

This makes it possible to emulate a wide area network by grouping packets into different classes with different priorities. In this project a default queuing discipline (qfifo) is used.

3.3.2 Client

A video stream application runs on the client computer connected with a server by an Ethernet Local Area Network (LAN). The client computer runs Ubuntu 9.10 Linux Operating System (OS). It is powered by a 1.6GHz Intel Atom N270 Processor, a DIMM Synchronous 533 MHz (1.9 ns), 1GB, 64 bits Single Channel Memory, the storage memory is an Intel 82801GBM/GHM (ICH7 Family) SATA AHCI Controller Hard Disk Drive, the Network Interface Card (NIC) is an Attansic L1e Gigabit Ethernet Adapter, 100MB/s. It also has an Integrated Intel Graphics Media Accelerator 950.

In order for a server to deliver packets to the client, the transport layer port should be opened first. Because of the personal firewall in Linux system, it is also necessary to control the incoming and outgoing traffics at the interface level. Implementation steps are described in the succeeding section.

3.3.3 Server

The Server is powered by Intel^(R) Core^(TM)2 CPU U7500 1.06GHz, 800MHz, 64 bits Core 2 Processor, a 1024MB DDR2 SDRAM Single Channel Memory, 120GB SATA Hard Drive is the storage memory, the multimedia audio controller is an 82801G (ICH7 Family) High Definition Audio Controller, 64 bits, 33MHz. The network interface is equipped with Intel PRO/Wireless 3945ABG [Golan], 32 bits, 33MHz (wireless=IEEE 802.11abg) and a Broadcom NetXtreme BCM5752 Gigabit Ethernet PCI Express, 1GB/s, 64 bits, 33MHz. It runs a Linux Ubuntu 9.10 Operating system.

A streaming server application runs on the server component. This component comprises of computer and Ethernet switch. The switch is used to enable connectivity between a streaming server and the NetEm TRAFFIC SHAPER and between the streaming server and MP (see Figure 3). The main reason for connecting streaming server to MP is to enable the later to capture traffic generated by streaming server before they reach the NetEm for packet dropping exercise.

This thesis evaluates the impact of different packet sizes on the quality of video stream reception. There is the need for control of packet sizes before transmission. This task may be implemented at different layers of the TCP/ IP (Transmission Control

Protocol/internet Protocol) reference model (Figure 3). If Transmission Control Protocol were used in this project there would be an option of configuring/ tuning the Maximum Segment Sizes (MSS) before delivering to network/ Internet layer for further encapsulation. Because of the real-time requirement of the video streaming, TCP is not used and instead the User Datagram Protocol (UDP) is used as a transport protocol (Figure 4).

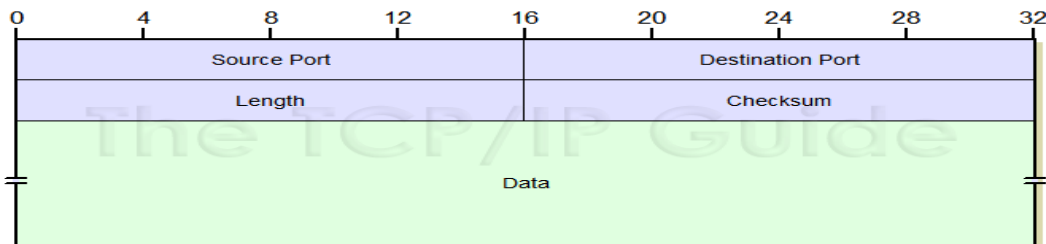


Figure 4: UDP message format [18]

It is also possible to control the packet size at network layer by tuning a network layer segment size. All the UDP datagram delivered to network layer are segmented into small chunks called packets. These packets (IP Datagram) are encapsulated at Link layer to form frames which are transmitted through a physical network. The transmitted size of frames depends on the size of the transmission medium such as MTU. For our case, an Ethernet (Figure 5) network is used and has an MTU of 1500 bytes. All the data arriving at the Link will also have to be segmented to a size that does not exceed the Ethernet MTU.

Any packet size bigger than Ethernet MTU will have to be segmented, conversely, any packet that is smaller than MTU will just be transmitted; this is equivalent to tuning packets sizes at Link layer directly other than on higher layers.

Starting Delimiter (1 byte)	Destination Address (6 bytes)	Source Address (6 bytes)	Length/ Type (2 bytes)	LLC header and Information field (46 - 1500 bytes)	Frame Check Sequence (4 bytes)
-----------------------------	-------------------------------	--------------------------	------------------------	--	--------------------------------

Figure 5: Ethernet frame format [19]

Destination IP Address field and Destination UDP port number are important fields for implementation of video streaming application. For Ethernet frame, an information field is useful for tuning frames to different packet sizes. In this thesis

3.3.4 Evaluation and Selection of Simulation and Experimental Tools

As appropriate, any meaningful project of this magnitude requires the provision of the needed experimental and simulation tools. Among the features that are relevant for implementation of this thesis work that VLC has, include real-time video streaming & transcoding, video playback, etc. Besides these features, VLC also has socket functionality that enables communications between two peer to peer or client- server computers. It was therefore decided to use VLC that has all the required features for undertaking of this thesis.

Tuning of packet sizes was found to be a challenge that faces VLC software. The software gives possibility of changing only bit rate and frame rate. However these parameters do not solve the packet size challenge experienced on VLC which is a paramount requirement for thesis evaluation. This challenge can be addressed by varying packet size at Link Layer of TCP/ IP reference models (Figure 3). Commands available in the Operating System solves the problem of tuning different packet sizes.

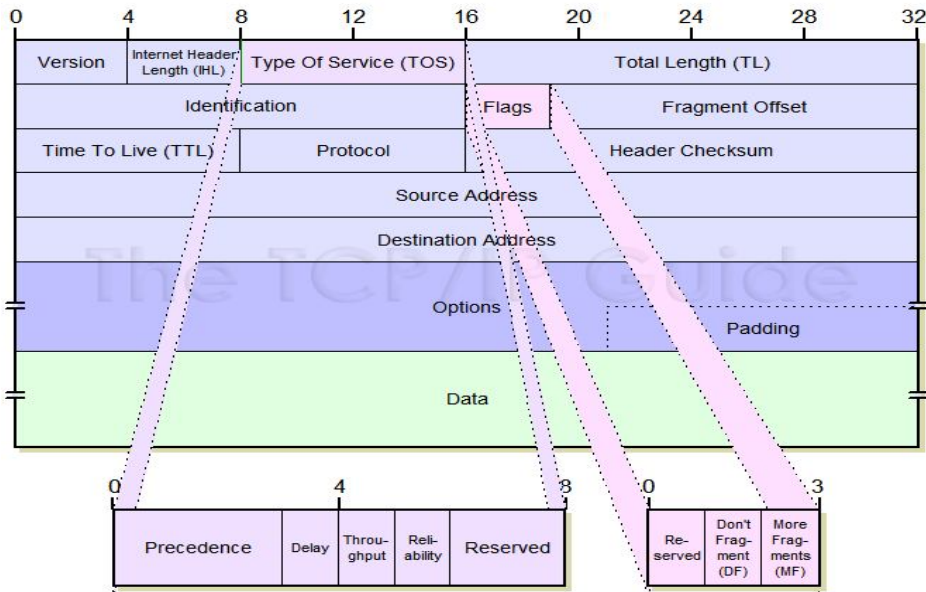


Figure 6: IP datagram format [20]

Linux Operating System is an open source OS and it has some useful commands like “ifconfig” on Command Line Interface (CLI) that can be used together with VLC. All the commands used in simulation are described in the succeeding subsection. Other important and necessary features that Linux provides include Traffic Control and Network Emulation, firewall configuration that allows or denies the incoming and outgoing packet at the interface are also available in Linux.

The UDP-sender and UDP-receiver is another simulation tool considered for use in this project. This is an open source program that can be installed in Linux machine for real-time client/ server multimedia communications. After thorough evaluation of this tool it was discovered that apart from transmission, the UDP-sender waits for acknowledgement from receiver for recently sent packets. If an acknowledgement is not received, the sender retransmit the unacknowledged packets until it get feedback. If the percentage of packet drops is high, it takes so long time for transmission of the entire whole file. This is in fact not suitable for real-time video streaming implemented in this project and therefore ignored.

Laboratory experiment is another activity performed in this thesis. The required equipments are traffic shaper, Measurement Points (MP) with DAG Cards, and consumer computers. Other equipments include Ethernet switch and cables. Connectivity of different components and equipments are shown in Figure 2. Having all these equipments in the laboratory facilitates implementation of real-time video streaming of which some packets are dropped randomly during transmission.

Two tools used for data analysis are Matlab and Microsoft excel. After data collection, Microsoft is used for recording and saved in a file which will later on exported into Matlab .mat files. A Matlab code is then developed to calculate the mean opinion values and variance values. All these parameters are used for analysis of data and determine the deviation of users’ opinions for each variable.

3.3.5 Operation of VLC

VLC is selected as the paramount tool for implementing video streaming in this project because it contains almost all the required functionalities. This subsection provides a brief overview of VLC operations and also briefly discusses its main components. Depicted in Figure 7 are three main components of VLC streaming

solution [22], namely “streamer”, “network” and “client”. Input to streamer are video files, DVD, acquisition cards with soft encoding, MPEG hardware encoding card, satellite and digital terrestrial TV. There are currently three types of streamers namely VideoLAN Client (VLC) streamer, VideoLAN service (VLS) streamer and VOD server.

The streamer used for this thesis is VLC and therefore VLS and VOD are not discussed. As seen in Figure 7, lots of inputs options are available for the VideoLAN solution, however in this project; the video file option was used and implemented. The main feature provided by VLC streamer is the transcoding technique that does codec of the streaming file into a format that can easily be read by intended clients. To enable the client to have access to VLC streamer, transport/ application port protocol is first opened and also the IP address of the intended destination is also configured. Intended destination in this case may be reached via a multicasting or a unicast arrangement. In this project, a unicast option was selected.

Prior to starting streaming, the client media player is first initiated to receive the files that are transmission by the streamer. A protocol port and IP address of the client are among the important parameter that should be configured for the client to communicate with a server.

VideoLAN Streaming Solution

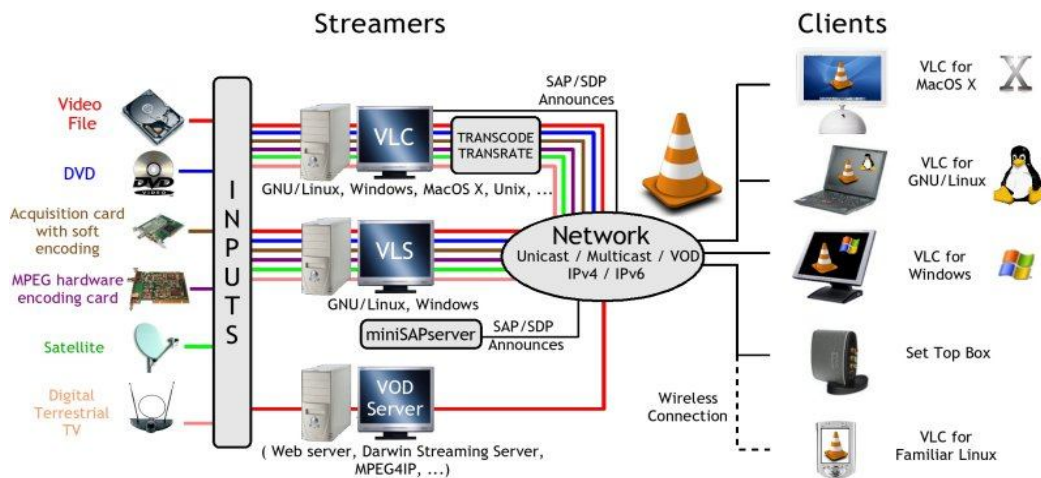


Figure 7: VLC and its main features [22]

3.4 Implementation

Prior to starting laboratory experiments, all the equipments were configured and tested. All computers were connected as shown in Figure 2 and tested by using a ping command. Order of configuration was as follows;

1. Firewall configuration on both client and the streaming server
2. Initiate a VLC player at the client computer
3. Configuring packet loss on the Traffic Shaper server
4. Setting a file name on the Consumer computer
5. Starting MP computer

3.4.1 Server

Experimental setup is as depicted in Figure 1 and Figure 3. Stream server's settings were accomplished at application, transport and Link layers of TCP/ IP model. Before transmission, various streaming parameters of VLC streaming server, like video coding, settings of frame rate and bit rate etc, were configured as shown in table 1.

Procedures used to initiate file streaming on the stream server are as follows

- Open VLC media player
- Choose streaming
- Insert file to be streamed
- Click a stream button to continue
- Choose the protocol used in streaming. In this thesis UDP
- Enter IP address and UDP port of the recipient/ destination host. In this thesis the address and were 10.0.1.189 and 1235 respectively
- Set transcoding parameters as shown in Table 1.

Table 1: Video and audio codec parameters settings

PARAMETER	VALUE
Video Codec	
Codec	MPEG-4
Bit rate (kbps)	1024
Frame rate (fps)	25
Resolution (pixel)	352x288
Aspect Ratio parameters	
Scale	1 (default)
Width	0 (default)
Height	0 (default)
Audio Codec	N/A

Packet size at the server was configured at Link layer by using the following `ifconfig` command.

```
ifconfig eth0 mtu packet "size"
```

Three different packet sizes used for this thesis were 1500 bytes, 1024 bytes and 512 bytes.

3.4.2 NetEm's configuration

The NetEm is an important machine which is responsible for dropping some packets passing across it based on the configured rules related to percentage of packet loss. Discussed in details on one of the preceding sections is a Linux command that emulates dynamic properties of a network. Various QoS performance parameters like delay, loss, etc can be configured by using a command. In this thesis only a packet loss parameter was set using the following commands

```
# tc qdisc add dev eth0 root loss 0.2%  
# tc qdisc change dev eth0 root loss 0.2%  
# tc qdisc del dev eth0 root
```

Percentage of packet losses used in this thesis are 0%, 0.1%, 0.3%, 0.5%, 0.7%, 1%, 3%, 5%, 7% and 10%. Assuming that the percentage of packet loss is set to 0.1%, it is equivalent to saying that for every one thousand packets transmitted from the

streaming server to the client, the shaper drops one packet. It should be noted that NetEm drops these packets randomly.

3.4.3 Measurement Point Setup

The MP machine has two DAG cards. One card is for capturing the transmitted packets before passing the NetEm traffic shaper and the second card captures packets after traffic shaping exercise. After completion of each video streaming session, the flushing command is issued in order to force the remaining packets out of the DAG buffer.

3.4.4 Consumer

All the data captured by MP are saved in the consumer server. File names should be set before starting of any video streaming exercise. This setting is important because without which an MP cannot know where to save data that it captures during transmission.

3.4.5 Client

Prior to starting the video streaming server, the client computer should be set ready to packets transmitted by the server. This is important because the video file is in the server and the client does not know a specific time when the server will start transmitting video frames. To avoid this problem, the following parameters should be configured first in the client computer.

To initiate playback at the client the following settings were done

- Open a VLC media player
- Choose network stream
- Select UDP from a list of transport protocols
- Enter IP- address of the client machine, i.e. 10.0.1.189
- Enter port number of the client machine, 1235
- Click play

3.5 Testing of the system

Before beginning laboratory exercise the system was tested to observe its performance and that it performs well according to the expectation. The reason behind testing of the system is trying to reduce errors and mistakes that might be caused by the improper settings and configuration of the system.

To avoid the problem just mentioned, the system was tested by using different types of data and the results were analyzed and compared with theories governing the problem at hand. Some of the parameters used for testing were connectivity, transmission of packets for different packet sizes tuned at Link layer, file streaming between the streamer and the client, testing of various MP, NetEm, and CONSUMER servers' connectivity and their communications.

NetEm was tested together with client and server. The NetEm parameters are set first followed by initiating the client and finally video streaming is started in the streaming server. Real-time monitoring is conducted at the client to observe some artifacts and deterioration of the received video. The consumer was tested to see if it stored the captured files and if it converts them into a readable and understandable format.

MP was also tested before starting experiment. For an MP to give the intended outputs, all the DAG cards should be functioning. If any or all of them are not working, the streaming may continue as usual but no data will be captured by the MP.

Our test on the system showed that the system works well according to expectations.

4 DATA COLLECTION AND ANALYSIS

Collection of video streaming data in this thesis is carried out into two phases. The first phase involves laboratory experimentation while the second phase involves collection of users' opinions. Laboratory experiment is carried-out using the system developed in this project as depicted in Figure 2 while video viewers' opinions are collected by using a questionnaire attached as appendix A. Starting of the second phase is dependent on the first phase because data obtained in phase one are used in the second one.

Two statistical methods used for analysis of data in this project are mean and variance/standard deviation of all the data for each variable. Matlab is used for computation of mean and variance of data for each study case. However, for ease of reference, mean and variance formula are derived in the succeeding subsection.

4.1 Mean and Variance

In this project, each study case known as vector has twenty seven elements. These are the people whose opinions were taken on the survey of this study. Assuming a vector element is represented by $x(n)$, for $n \in [1, 2, \dots, N]$, where, N equals 27. The mean of viewers' opinions become

$$\bar{x}_{ij} = \frac{1}{N} \sum_{n=1}^N x_{ij}(n) \quad (1)$$

Where, \bar{x}_{ij} is an MOS value for a j percents of the dropped packet for video type i . Conversely, $x_{ij}(n)$ means n^{th} user's opinions for a received video of type i which has lost j percents of packets. Three types of videos used in this study are described as follows.

- $i = 1$ represents "Foreman" video type
- $i = 2$ represents "Football" video type
- $i = 3$ represents "News" video type

There are ten variables of packet losses used in this study and are described in Table 2.

Table 2: Summarized description of packet loss

j (Index)	Packet loss (%)	Description
1	0.0	No packet is lost
2	0.1	For every a thousand packets one is lost
3	0.3	For every a thousand packets three of them are dropped
4	0.5	For every a thousand packets five of them are dropped
5	0.7	For every a thousand packets seven of them are dropped
6	1	For every a thousand packets ten of them are dropped
7	3	For every a thousand packets thirty of them are dropped
8	5	For every a thousand packets fifty of them are dropped
9	7	For every a thousand packets seventy of them are dropped
10	10	For every a thousand packets a hundred of them are dropped

Accordingly, the variance is computed as follows

$$\text{var}_{ij} = \frac{1}{N-1} \sum_{n=1}^N (x_{ij}(n) - \bar{x}_{ij})^2 \quad (2)$$

It is described in (2) that variance is computed by taking summation of a square of the difference between the n^{th} element of vector x viewed a video clip of type i which

has lost j percentage of packets and its respective mean divided by total amount of elements minus one.

4.2 Data Collection

Laboratory experiment was conducted according to the setup shown in figure 4. As described in chapter 1, this thesis' objective is about evaluation of the impact caused by packet losses and packet sizes on qualities of the transmitted video streaming clips. Authors have opted to use the subjective method commonly known as mean opinion score (MOS) to determine quality of the received streams proposed by proposed and authorized by ITU-T. There are also other methods like SCS (Switching between error Concealment and frame Skipping), that one can use to analyze qualities of the received video qualities as suggested in [21].

ITU-T being the highest authority/ Agency of the United Nations responsible for coordination of telecommunications industry has defined one of the most important parameter for the measurement of qualities of multimedia contents based on experience of users perception. Opinions of users based on their perception of the quality of the received multimedia is believed to enhance multimedia users' satisfaction on multimedia contents regardless of communication problem experienced during transmission on telecommunication network.

According to ITU-T standard, the Mean Opinion Score (MOS) is classified into five quality groups namely Excellent (5), Good (4), Fair (3), Poor (2), and Bad (1). Apparently, It may also be said that the excellent multimedia means the impairment of video quality is imperceptible, good means the impairment is perceptible but not annoying, fair quality means the multimedia's impairment is slightly annoying, poor means the received multimedia is annoying and bad means the received multimedia is very annoying.

4.3 Data Analysis

Statistical properties described in section 4.1 are used for evaluation and analysis of data. Mean opinion score obtained through questionnaire shown in appendix A is computed and are therefore used for drawing histograms and plots. Comparison between the MOS and packet drops, video speeds and packet sizes are used for data analysis.

5 RESULTS AND DISCUSSION

This research experiments was conducted at BTH in Karlskrona City, Sweden of which twenty seven people were invited to participate. Table 3 presents the viewers opinions on qualities of different video clips. Variables used in this thesis were “packet loss” and “packet sizes”. Also used are different videos, one of the video has fast moving and changing scenes. It is called Football. The scenes in the second video called Foreman does not change rapidly as that of Football, but its scenes varies and changes faster when compared with the third video called News. The first column on the table contains three types of abbreviations. FM1500 is an abbreviation for “foreman” video with transmission packet size of 1500 bytes while FT and NE represent “football” and “news” videos respectively. Packet sizes for FT and NE should be described in the same way as for the case of “foreman” videos.

Table 3: Mean Opinion Score

VIDEO	MOS for each Percentage of Packet Loss									
	0%	0.1%	0.3%	0.5%	0.7%	1%	3%	5%	7%	10%
TYPE										
FM1500	3.93	3.78	3.33	2.52	3.22	2.78	2.19	2.78	2.07	1.70
FM1024	3.96	3.67	2.81	2.37	3.19	2.19	1.96	1.67	1.30	1.44
FM512	4.04	3.33	3.78	2.22	2.74	2.04	1.85	1.78	1.74	1.26
FT1500	3.70	3.41	3.52	2.70	3.33	2.33	1.96	2.11	1.44	1.26
FT1024	3.85	3.52	3.30	2.11	2.15	2.15	1.37	1.52	1.44	1.19
FT512	3.52	2.59	2.44	2.07	2.00	1.78	1.26	1.22	1.41	1.15
NE1500	3.85	3.65	3.78	3.15	3.67	3.30	2.67	2.19	3.04	1.96
NE1024	3.85	3.63	3.33	2.85	3.26	2.96	2.19	2.56	2.19	1.56
NE512	3.70	3.52	2.93	2.59	3.41	2.48	2.04	2.26	1.44	1.30

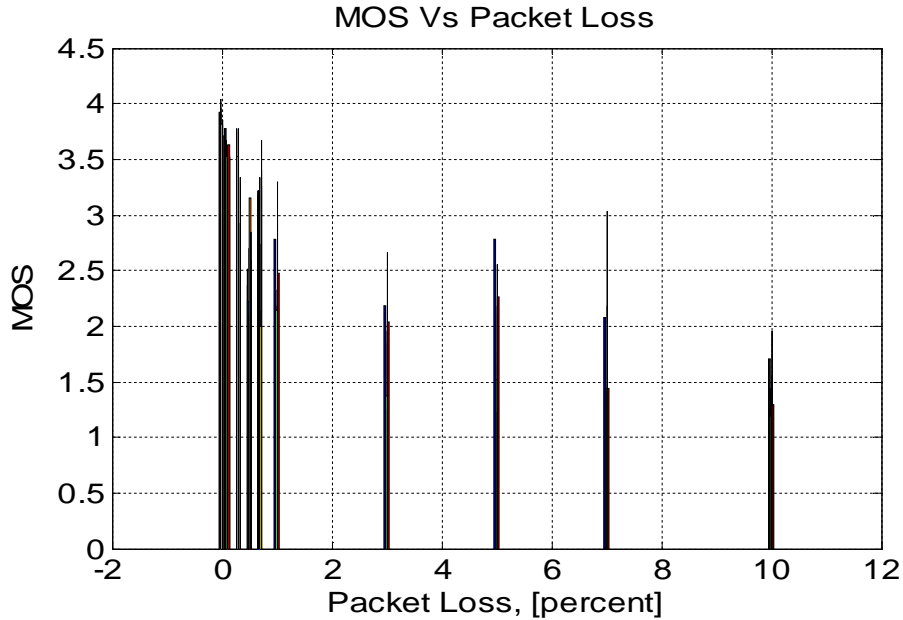


Figure 8: MOS for all scenarios

It can generally be observed in Table 3 that the smaller the amount of dropped packets the better is the MOS and thence the user’s perception on the received video

quality. This argument is strongly supported for instance with the comparison of results between the seven- percent of packet loss and those of zero percent of the lost packets.

General trend shows that as packet drops increases in the network, MOS value decays exponentially as observed in Figure 8. The graph shows different lines with colors that described different types of video which are not clearly observed. For analysis purpose, Figure 8 is divided into three parts. The first part cuts and zooms-in the portion between 0 and 1 value of packet loss axis, the resulting graph is presented in Figure 9, the second part zooms-in a packet loss axis value number 10 and the graph is presented in Figure 10 and the third part zooms-in packet-loss axis value number 7 with the graph being presented in Figure 15.

Our observation and discussion on these results reveals that packet loss may significantly affect quality of the received video. We analyze the effect of different packet sizes on video frames basing our analysis on user’s perception as well as the percentage of packet losses.

5.1 Effect of packet sizes on video quality

Figure 9 presents a zoomed-in portion of Figure 8. For each packet loss values, nine results are presented for different packet sizes and video types. Each group of histograms contains ten blocks of different colors. The first three represent three different packet sizes, 1500 bytes, 1024 bytes and 512 bytes for foreman video. The next three represents results opinions for football video and the last three represent news video.

Also, we need to clarify that the first three histograms with dark blue, blue and pale blue colors represents the received foreman videos which were transmitted with packet size of 1500 bytes, 1024 bytes and 512 bytes respectively. General trend as depicted in Figure 9 shows that the bigger the packet size, the higher the MOS value. This means that viewers have better quality of experience when the videos were transferred in larger packet sizes. The last three histograms represents the received news videos with Ethernet frame sizes of 1500 bytes, 1024 bytes and 512 byte while the a fourth to sixth histogram represent news with the same packet size as for foreman and football. MOS results show the same trend for different packet sizes of football and news videos as was the case for foreman video described above.

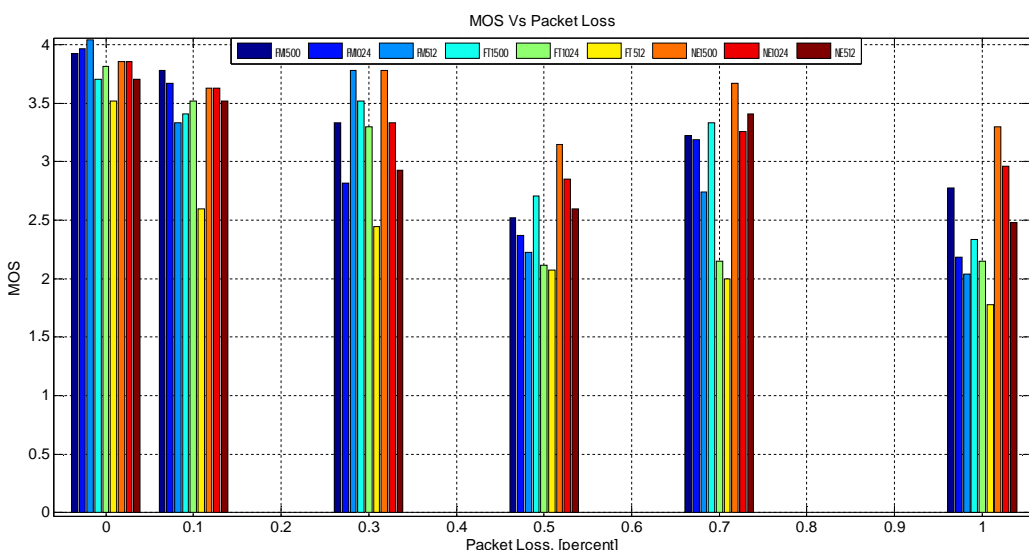


Figure 9: Graph of MOS for packet loss in the range of 0 to 1 percent

As a result of the relevance of packet sizes to video quality as revealed in Figure 10, settings of frame sizes during transmission can have great impact on video quality. We therefore recommended by the use of bigger packet sizes for improved reception of streaming video.

5.2 Mean Opinion Score for different types of videos

Three types of videos were streamed. Figure 10 is a zoomed-in portion of Figure 8 windowed at packet loss value of 10%. This portion is deliberately selected to describe the effect of various video scenes transmitted across the lossy channel basing on MOS results. It is observed from Figure 10 that users' perception on news video scored high MOS grade followed by foreman and finally football. It can generally be said that MOS values for all of the video ranges between poor and bad grades. In other words, the qualities of all videos were 'annoying' and 'very annoying'.

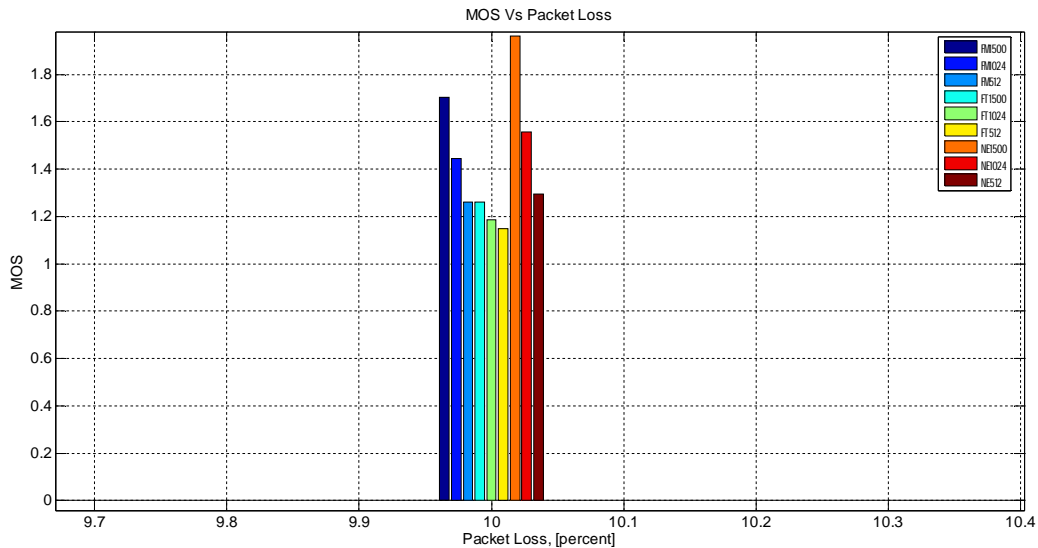


Figure 10: MOS results for packet transmission loss of 10 percent

5.3 Amount of packet losses leading to users' rejection of videos

Table 3 provides results for all video types received from the network that drops different amount of packets of different sizes. In this subsection, our discussion is based on the amount of packets that can be dropped by the network beyond which users' rejects them. According to ITU-T, and as discussed in 4.2, the MOS values are grouped into five quality groups of videos namely, imperceptible, perceptible but not annoying, slightly annoying, annoying and very annoying. In this discussion, a fair MOS value is considered as the value that people find the video is okay for them to continue watching.

It is observed from Table 3 that the received football video with the packet size of 1500 bytes, people seem to be quite comfortable with the video that has lost about 0.3% of packets. When the packet size is 1024 bytes, users' opinions on 'foreman video' shows that about 0.2% of packet losses seems to be okay and that the video is slightly annoying. For the case of football, results show a fair grade for a received video that has lost up to 1% of its packets.

Figure 11, Figure 12 and Figure 13 provide graphs that further describe the arguments discussed in this subsection.

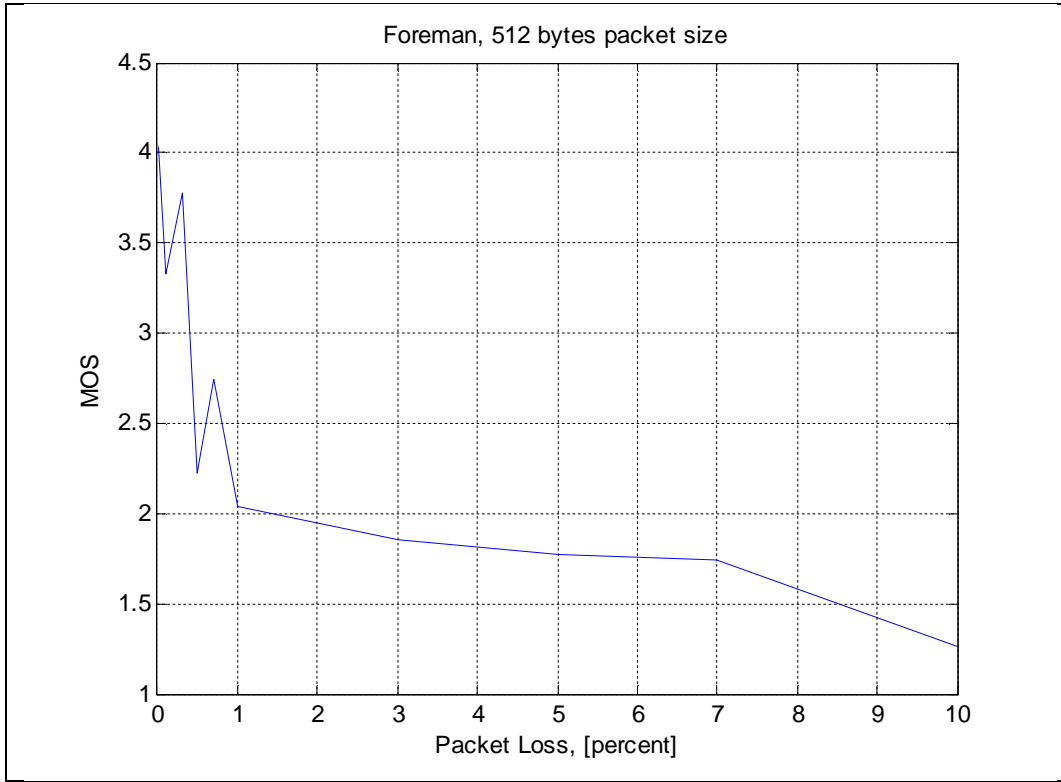


Figure 11: MOS graph for "Foreman" received video. Transmission packet size is 512 bytes

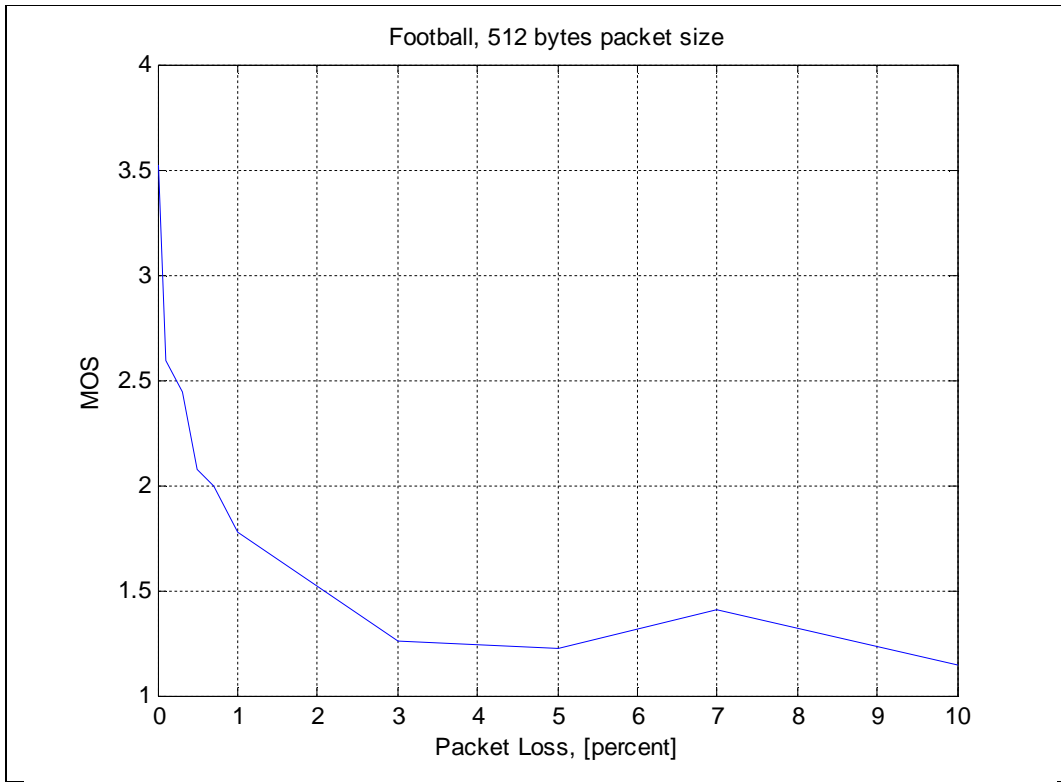


Figure 12: MOS graph for "Football" video. Transmission packet size is 512 bytes

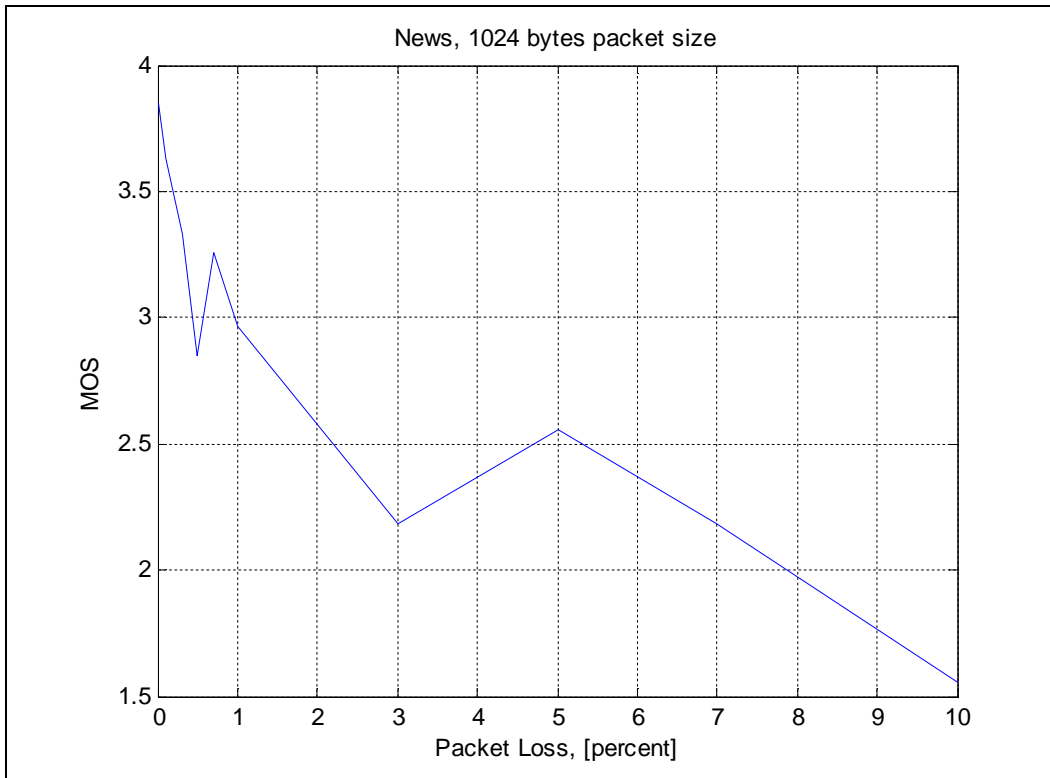


Figure 13: MOS graph for "News" received video. Transmission packet size is 1024 bytes

5.4 Different people's opinions on the same video

For all type of videos considered, there is observed a progressive decay of MOS values as the percentages of dropped packets increases. This lends credence to the fact that increase in packet loss will lead to increase in artifacts thus leading to poor users' perception and low ratings for any of the video types considered.

Consider Figure 14 that shows the variances of each of the videos. It is observable that at 10% of packet loss, the variance of viewers' opinions is low for football video. This means that more people agreed that the football video that has fast moving scenes has poorer video qualities. This fact is quite noticeable as there is convergence of opinion with the variance tending towards zero.

Also it is noticeable that as more packets are dropped, there is divergence of opinion about whether the quality of videos with slowly changing scenes such as news is bad. While some people are still comfortable with the video display, others on the other hand think that the video quality is very poor. The variance value of viewers' opinions for the videos with slowly changing scenes such as news is quite high signifying a lot of divergence in peoples' opinions, this may be attributable to the fact that more people are still comfortable with such videos as the scenes represented although changing, has in most cases stationary subjects (humans in this case)

Generally speaking, variances for 0% and 0.1% of packet losses are observed to be low for all videos and packet sizes. Variance is high for 0.3% and 1% of packet losses whereas a range between 3% and 10% of packet drops the variance value is low. These results could be translated as how difficult is for viewers to make decision on the videos they watch in the range between 0.3% and 1% of packets. But, low values of

variance for a range between 3% and 10% shows that viewers were easily able to decide on whether the video is bad or good.

It may further be observed that at low values of packet losses, variance for football and foreman video clips were higher than news video clips. However, as the percentage of packet losses increase the variance of opinions for news video overtook those of foreman and football.

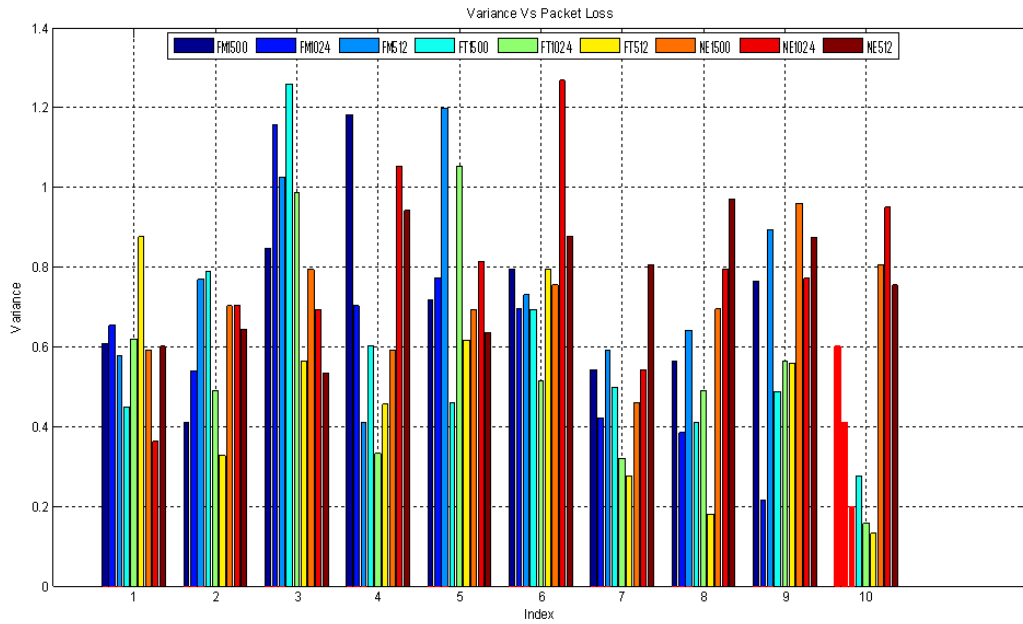


Figure 14: Variance of Users' Perception

6 CONCLUSION

In this thesis, we have observed the impact of packet losses on video stream transmission. Due to different packet sizes and different percentages of losses, there were differing opinions from different viewers that watches the received videos.

A normal UDP channel has been simulated and the losses have been introduced to the network using NetEm which is a tool that can emulate Wide Area Networks. Viewers have qualities of experiences which differ sharply from one another.

When amount of packet lost increases, result shows that MOS suffers an exponential decay. Many people were able to observe that the videos were poor as the percentages of losses peaks to the maximum allowable loss percentage for this thesis.

A highpoint of the achievement of this work is that it is observable from viewers' responses that the larger the packet sizes, the better the quality of videos received. This is attributable to the fact that if the videos are segmented into many smaller packets, there is high possibility that the video during transmission may suffer more losses of the very important I frames. The effect of the loss of these frames will ripple through the entire GOPs and the received video will have much more artifacts leading to poor MOS from viewers.

On the other hand, if the packets are larger coupled with the fact that NetEm looses packets randomly, more I frames have improved chances of surviving the best effort UDP transmission onward to the receivers' side.

Hence, it is better for designers to have it on mind that bigger packet sizes will lead to better quality of received video and that smaller frame sizes will lead to poorer video quality.

In a nutshell, the effect of packet size on video stream transmission dictates the sending of larger packets from the senders' side for viewers' at the receiver's side to have a much more enjoyable video quality experience.

The contribution of this thesis will aid in operations such as emergency relieve operations, telemedicine, videoconferencing and coverage of football tournaments.

7 FUTURE WORK

Future works along this line may include identification of packets as well as change of packet size at network layer. This may lead to the design of better CODEC that is more intelligent. The CODEC in question may be able to more intelligently reconstruct a deformed frame during transmission when the GOPs were received at the receivers end.

Also, this will lead to better protection of the vector carrying I frames. Since the loss of I frames oftentimes have multiplier effects on the received video frame, identification of these I frames will lead to better prioritization and protection of the frames during transmission. Losses accruing to them will be minimized and users will have a more qualitative experience of their multimedia videos.

The overall benefits of the above stated cannot be over-emphasized, UDP will be more appreciated as its inherent shortcomings will no longer lead to poor perception when packets are lost as the packets that may eventually be lost may not be the all important I frames.

8 APPENDICES

APPEINDIX A

NAME -

OCCUPATION - STUDENT

LECTURER

OTHERS (PLEASE SPECIFY)

GENDER -

AGE -

Please rate the video clip shown to you according to your perception of the video quality.

The highest score being 5 (Ex'lent) while the lowest is 1 (Bad). Overall distribution of points is as follows: EX'LENT (5), GOOD (4), FAIR (3), POOR (2), BAD (1)

FOREMAN (A)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOREMAN (B)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOREMAN (C)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOOTBALL (A)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOOTBALL (B)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOOTBALL (C)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

NEWS (A)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

NEWS (B)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

NEWS (C)

Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Video 8	Video 9	Video 10
Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent	Ex'lent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

FOOTBALL- MISCELLANEOUS

CASE 1			CASE 2			CASE 3			CASE 4		
Vid 1	Vid 2	Vid3	Vid1	Vid 2	Vid 3	Vid 1	Vid 2	Vid 3	Vid 1	Vid 2	Vid 3
Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent	Ex'ent
Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:	RT:

MANAGEMENT OF FILE NAMES

Video Name	Given questionnaire name	Description
	Video_file_name (letter)	There are three types of videos used for laboratory experiments namely Foreman, News and Football. The first nine questionnaire tables are named as either Foreman (A), Football (B), News (C), etc. Letter A inside the bracket

			represents packet size 1500 bytes, B represents 1024 bytes and C is for 512 bytes packet size.
Video_1.mpg	Video 1		Received video with packet loss of 0.7%
Video_2.mpg	Video 2		Received video with packet loss of 5%
Video_3.mpg	Video 3		Received video with packet loss of 0.1%
Video_4.mpg	Video 4		Received video with packet loss of 0%
Video_5.mpg	Video 5		Received video with packet loss of 0.3%
Video_6.mpg	Video 6		Received video with packet loss of 7%
Video_7.mpg	Video 7		Received video with packet loss of 10%
Video_8.mpg	Video 8		Received video with packet loss of 0.5%
Video_9.mpg	Video 9		Received video with packet loss of 3%
Video_10.mpg	Video 10		Received video with packet loss of 1%
F1_a.mpg	Vid 1	CASE 1	Received video with packet loss of 5%, 1500 bytes, football video file
F1_b.mpg	Vid 2		Received video with packet loss of 5%, 1024 bytes, football video file
F1_c.mpg	Vid 3		Received video with packet loss of 5%, 512 bytes, football video file
F2_a.mpg	Vid 1	CASE 2	Received video with packet loss of 0%, 1500 bytes, football video file
F2_b.mpg	Vid 2		Received video with packet loss of 0%, 1024 bytes, football video file
F2_c.mpg	Vid 3		Received video with packet loss of 0%, 512 bytes, football video file
F3_a.mpg	Vid 1	CASE 3	Received video with packet loss of 0.5%, 1500 bytes, football video file
F3_b.mpg	Vid 2		Received video with packet loss of 0.5%, 1024 bytes, football video file
F3_c.mpg	Vid 3		Received video with packet loss of 0.5%, 512 bytes, football video file
F4_a.mpg	Vid 1	CASE 4	Received video with packet loss of 10%, 1500 bytes, football video file
F1_b.mpg	Vid 2		Received video with packet loss of 10%, 1024 bytes, football video file
F1_c.mpg	Vid 3		Received video with packet loss of 10%, 512 bytes, football video file

APPENDIX B

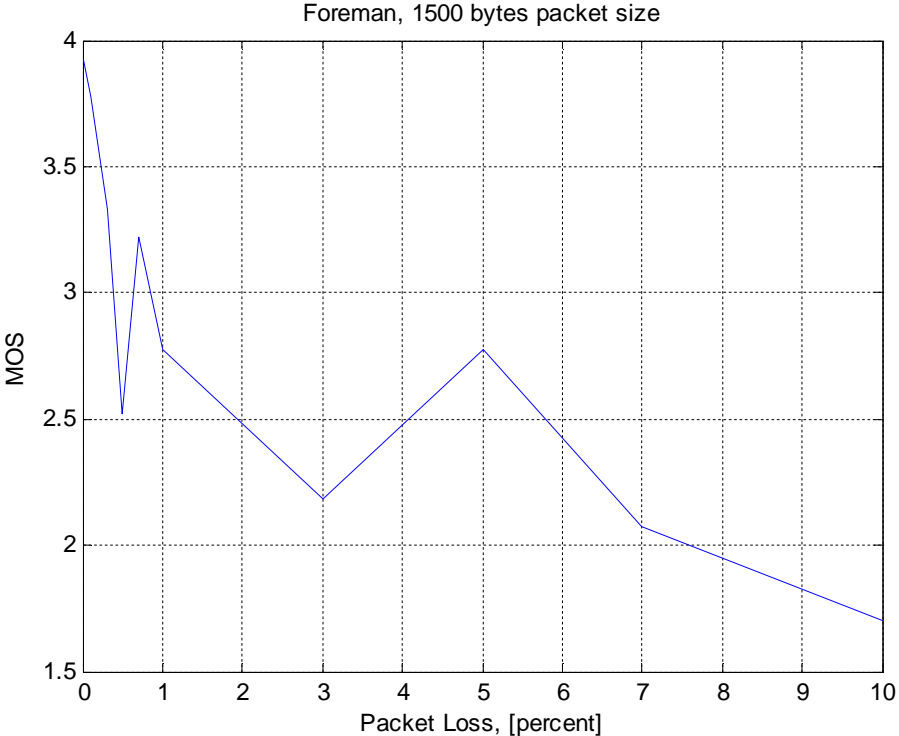


Figure 15: MOS graph for “Foreman” received video. Transmission packet size is 1500 bytes

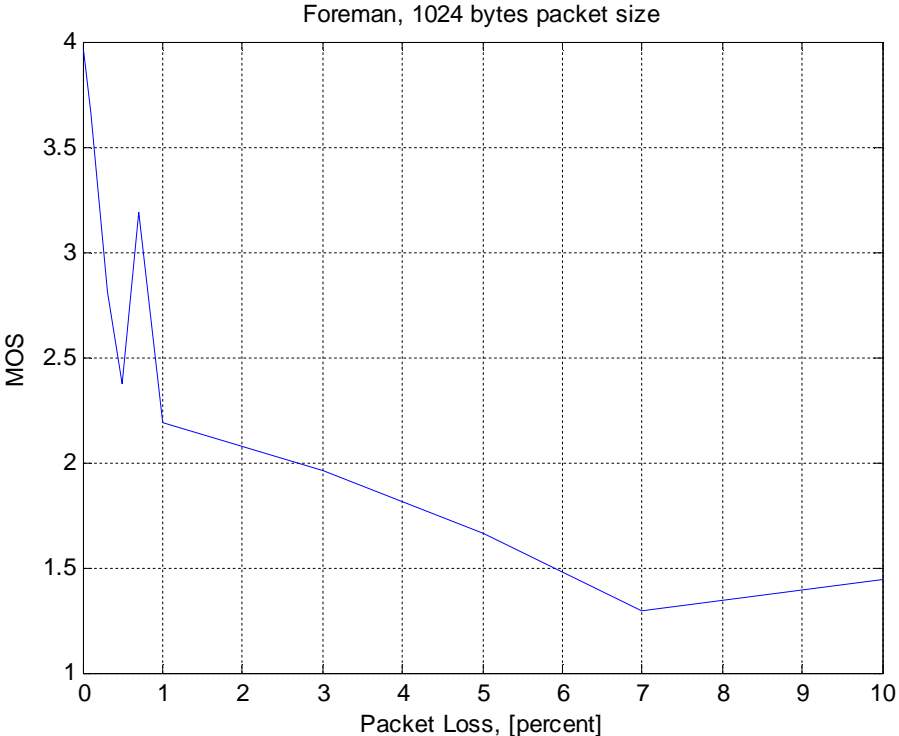


Figure 16: MOS graph for “Foreman” received video. Transmission packet size is 1024 bytes

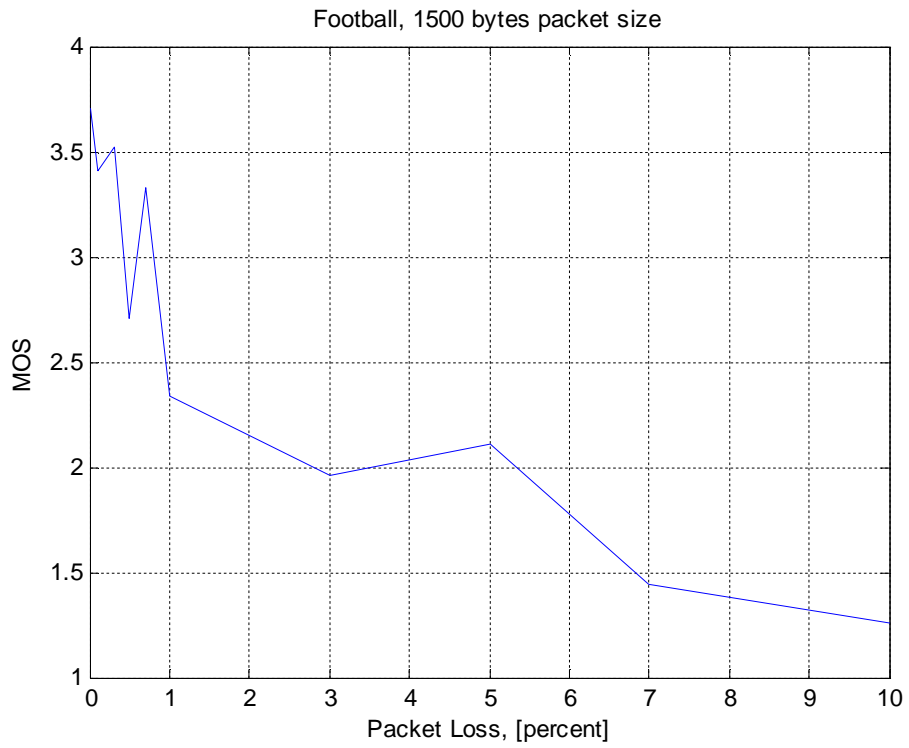


Figure 17: MOS graph for "Football" video. Transmission packet size is 1500 bytes

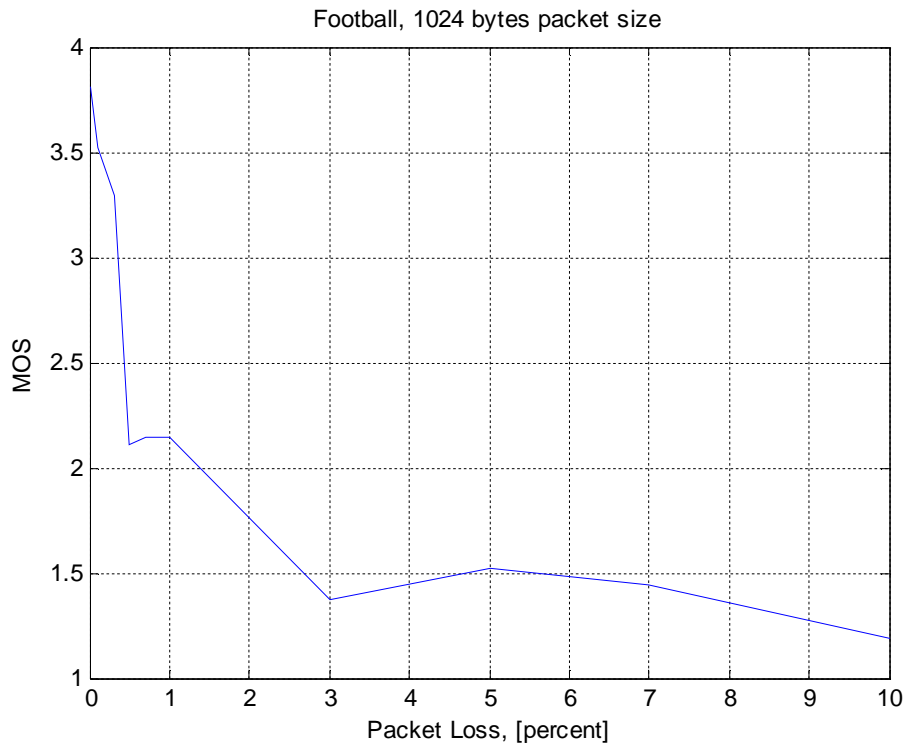


Figure 18: MOS graph for "Football" video. Transmission packet size is 1024 bytes

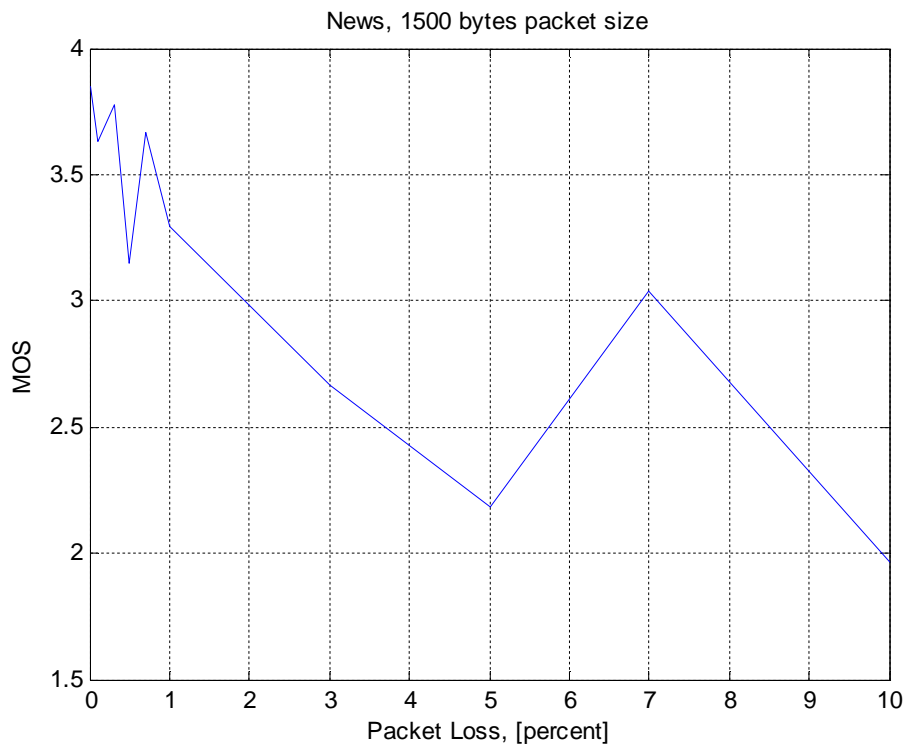


Figure 19: MOS graph for "News" received video. Transmission packet size is 1500 bytes

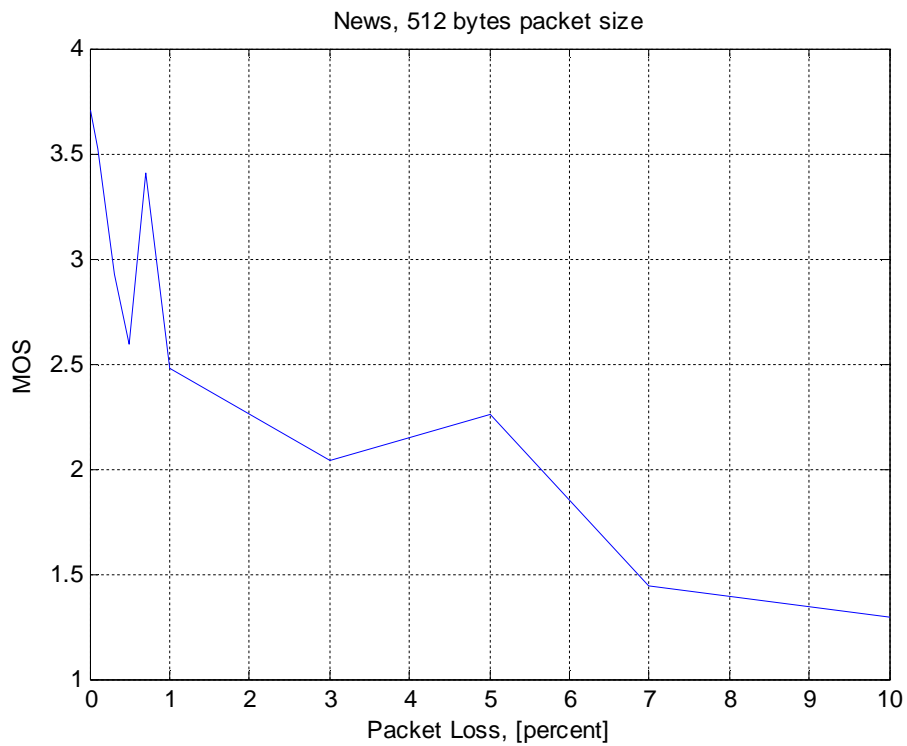


Figure 20: MOS graph for "News" received video. Transmission packet size is 512 bytes

9 REFERENCE

- [1] N. Khalifa and H. Elmahdy, "The Impact of Frame Rate on Securing Real Time Transmission of Video over IP Networks," in IEEE ICNM, pp. 57 – 63, March 2009.
- [2] H.264- The new standard for digital video quality and performance, available: www.divx.com/en/software/divx-plus/codec-pack/h264, accessed on 08-05-2010
- [3] M. Koul, "ANALYSIS of the Effects of Packet Loss and Delay Jitter on MPEG-4 Video Quality," Dept. of Electrical Engineering, The University of Texas at Arlington, Arlington, Texas 76019; 2008
- [4] M. Karaliopoulos, P. Henrio, K. Narenthiran, E. Angelou and B. Evans, "Packet scheduling for the delivery of multicast and broadcast services over S-UMTS," International Journal Of Satellite Communications And Networking; Int. J. Satell. Commun. Network. 2004; 22:503–532 (DOI: 10.1002/sat.797)
- [5] R. Shmueli, O. Hadar, R. Huber, M. Maltz and M. Huber, "Effects of an Encoding Scheme on Perceived Video Quality Transmitted Over Lossy Internet Protocol Networks", in *IEEE Trans. Broadcasting*, Vol. 54, pp. 628 – 640, Sept. 2008.
- [6] J. Daka and D. Makaroff, "Mixed Reliability Video Streaming in Wireless Networks," in IEEE 2008 International Symposium on a WoWMoM, pp. 1 – 8, Aug. 2008.
- [7] R. MacKenzie, D. Hands and T. O'Farrell, "QoS of Video Delivered over 802.11e WLANs," in IEEE International Conference on Communications, pp. 1 – 5, June 2009
- [8] E. Ulrich, T. Kusuma and H. Zepernick, "Perceptual Quality Assessment of Wireless Video Applications," Department of Radio Communications, Blekinge Institute of Technology, Karlskrona, Sweden.
- [9] M. Fiedler, K. Tutschku, P. Carlsson and A. Nilsson, "Identification of Performance Degradation in IP Networks Using Throughput Statistics," Department of Telecommunications & Signal Processing, Blekinge Institute of Technology, Karlskrona, Sweden.
- [10] V. Uti and R. Fox, "Improving Delivery Rate for Real Time Mobile Streaming Video," in proc. 21st IEEE International Conference on Tools with Artificial Intelligence, pp. 49- 53, 2009
- [11] C. Lin, C. Ke, C. Shien and N. Chilamkurti, "The Packet Loss Effect on MPEG Video Transmission in Wireless Networks," in Proceedings of the 20th International Conference on Advanced Information Networking and Applications (AINA'06), vol.1, pp. 565 – 572, 2006
- [12] A. Lo, G. Heijenk and I. Niemegeers, "Evaluation of MPEG 4 Video Streaming Over UMTS/WCDMA Dedicated Channels,"; in Proceedings of the 1st International Conference on Wireless Internet (WICON), pp. 182- 189, 2005

- [13] M. Lee and J. Kim, "Video Frame Rate Control for Non-Guaranteed Network Services with Explicit Rate Feedback," in proc. IEEE global Telecommunications conference, vol.1, pp. 293 – 297, 2000
- [14] T. Tsai and J. Leou, "A Rate Control Scheme for H.264 Video Transmission," in proc. IEEE International Conference on Multimedia and Expo (ICME), vol. 2, pp. 1359 – 1362, 2004.
- [15] H. Zheng, S. Wang and J. Copeland, "QoS Aware Mobile Video Communications," in proc. IEEE Military Communications Conference (MILCOM), vol. 2, pp. 1231 – 1235, 1999.
- [16] P. Nunes, L. Soares and F. Pereira, "Automatic and adaptive network-aware macroblock intra refresh for error-resilient H.264/AVC video coding," in 16th IEEE ICIP, pp. 3073 – 3076, Nov. 2009.
- [17] D. Nemethova, M. Rupp and M. Kieffer, "Error Resilient Transmission of Video Streaming over Wireless Mobile Networks," DISSERTATION Institut für Nachrichtentechnik und Hochfrequenztechnik Technische Universität Wien Österreich, Germany, 2007
- [18] UDP Message Format, on "The TCP/ IP guide, available: http://www.tcpipguide.com/free/t_UDPMessageFormat.htm, accessed on 8th May 2010.
- [19] IEEE 802.3/ Ethernet version 2 frame format. Ref. Ethernet frame format, available: <http://publib.boulder.ibm.com/infocenter/iserics/v5r3/index.jsp?topic=/rzajy/rzajyeternetframeformat.htm>, accessed on 8th May 2010
- [20] IP Datagram General Format, "on The TCP/ IP guide," available http://www.tcpipguide.com/free/t_IPDatagramGeneralFormat.htm, accessed on 8 May 2010
- [21] S. Tasaka and H. Yoshimi, "Enhancement of QoE in Audio–Video IP Transmission by Utilizing Tradeoff between Spatial and Temporal Quality for Video Packet Loss,"
- [22] Streaming, Muxers and Codecs, available: <http://www.videolan.org/doc/streaming-howto/en/ch01.html> , accessed on 6th March 2010.