UNDERSTANDING THE IMPACT OF NETWORK IMPAIRMENT OVER INTERNET-BASED TELEMEDICINE VIDEO TRAFFIC

Bengisu Tulu
Samir Chatterjee
Claremont Graduate University

Abstract
The Internet is designed to provide best effort service, which is not suitable for real-time applications. Therefore, carrying real-time applications over the Internet presents a number of challenges. While quality of service technologies can overcome some of these challanges, most medical facilities and patient homes do not utilize them since there is a high cost associated. Telemedicine and related technologies provide a promising opportunity to improve our healthcare reach to wider populations. The low cost and ubiquity of the Internet can help deliver healthcare to a wider population. However, the unreliable connection properties of packet-based systems hamper the quality of Internet-based telemedicine applications. This research is aimed to explore the delay, jitter, and drop cutoff points for Internet-based videoconferencing applications where video quality stays above the acceptable peak-signal-to-noise-ratio value (25dB). An experimental study is conducted on our testbed where behavior of the Internet is emulated on a Linux PC using NIST Net, a network emulation tool. Video file utilized during these experiments is an ophthalmology telemedicine session and it was obtained from the Regenstrief Institute for Health Care. Video quality measurements are computed using Video Quality Metric (VQM) software, which was developed by the Institute for Telecommunication Sciences (ITS). Our analysis and results indicate that jitter has a severe effect on the video quality whereas the effect of delay is almost invisible. Drop effects are similar to jitter with less impact on video quality. Results presented in this paper include the cutoff points for these variables as well.

1. Introduction

Telemedicine in particular can spread critical medical expertise across a region and around the globe. Even though telemedicine has evolved over the past 30 years, today most telemedicine implementations require expensive leased telecommunication circuits to provide secured reliable connections. The low cost and ubiquity of the Internet can help deliver healthcare to a wider population through smart Internet-based applications. However, the unreliable connection properties of packet-based systems and their vulnerability to various impairments that affect the physical, network, and application layers hamper the quality of Internet-based telemedicine applications. This research is aimed to explore the delay, jitter, and packet drop cutoff points for Internet-based videoconferencing applications over telemedicine where video quality stays above an acceptable objective measurement level.

Existing audio/video quality measures, both objective and subjective, were developed for the Public Switched Telephone Networks and the broadcast networks. The Internet has very different characteristics compared to these old traditional communication technologies. Therefore, existing quality measures are not necessarily well suited for audio/video applications over the Internet. Moreover, telemedicine has specific requirements for quality and these requirements need to be taken into account while evaluating the quality of Internet-based telemedicine audio/video applications. For example, the application area (such as ophthalmology) and application purpose (diagnosis versus consultation) have an important effect on the quality expectations of the users. Therefore, in this study we recognize the various dimensions of telemedicine reported in [1] and select the ophthalmology general eye exam as our application area for initial exploration.

2. Network Impairments Affecting Quality in Telemedicine over IP Networks

Since the Internet was not designed for real-time applications and only provides best effort service, carrying real-time applications over the Internet presents a number of challenges. These include lack of guarantee in terms of bandwidth, packet loss, delay, and jitter, all of which affect the quality of voice and video over the Internet as reported in various studies [2-4].

Packet loss – Unlike circuit-switched networks, in packet switched networks no physical end-to-end circuit is established [4]. Packets are transmitted from the source to the destination over the Internet with the help of routers. Arriving packets at a router are first queued and then transmitted one-by-one, usually with the first in first out (FIFO) policy. However, if the queue (buffer) of a router is already full when a packet arrives, then this packet is dropped and consequently, is not transmitted to its destination. Network congestion occurs when routers start dropping packets. The effects of packet loss on real-time multimedia applications are critical. During a voice conversation, human cognition can handle only a certain amount of packet loss. If too many packets are lost, the voice becomes incomprehensible. For video the effect of extensive packet loss is more acute. If packet loss happens, some parts of the video cannot be decoded and displayed. It is important to understand the effects of packet loss on the perceived quality of voice and video applications. Researchers have developed various techniques to overcome, or at least ease, the effects of packet loss on applications; some of these techniques are discussed in [4, 5].

Packet Delay – End-to-end packet delay is typically caused by a number of components [4]: (1) codec delay is the time it takes to convert analog data to digital and vice versa, (2) serialization delay is the time it takes to place a packet on the transmission line, and is determined by the speed of the line, (3) queuing delay occurs at the various switching and transmission points of the network, such as routers and gateways, where packets wait in the queue to be transmitted over the same outgoing link, and (4) propagation delay is the time required by signals to travel from one point to another, which is fixed as determined by the speed of light. The effects of large packet delay become even more severe for voice communications, as timing is an important characteristic of voice. This is especially true when an interactive conversation is being transmitted on the network; delay effects can turn the conversation into a half-duplex mode where one speaks and other listens and pauses to make sure the other is done. Echo is another unwanted effect of packet delay. Various techniques were also developed to overcome these problems over packet-based networks since in current circuit-switched networks the primary source of delay is propagation delay.

Packet Delay Variation (Jitter) – Packet delay variation refers to the variation or gaps between packet arrival times at the receiving buffer. This occurs due to the variability in queuing and propagation delays. To eliminate the effects of this variation, usually a playout buffer is used. The receiver holds the first packet in the buffer for a specific amount of time before playing it out. Therefore, a small jitter is tolerable but large fluctuation causes difficulty in decoding and playback and causes quality degradation. The effects of delay variation are similar to the effects of packet loss. Large variation in delay will result in some packets arriving long after the playout time scheduled for them based on the buffer size. The receiver will discard these packets since they are out of order.

3. Research design

The goal of this study is to understand the effects of impairments that occur in IP-based networks on video quality in low-cost telemedicine settings. Therefore, tools that are utilized in the experiments provide low-cost solutions on broadband Internet connections.

3.1 Experimental testbed

The problem under investigation is the measurement of degradation in quality of video after it is transmitted through the Internet where certain impairments occur. A testbed is designed to emulate the Internet traffic and control impairment parameters (delay, jitter, and drop) while transmitting an actual telemedicine multimedia session. Figure 1 illustrates the configuration of the testbed utilized in this study.

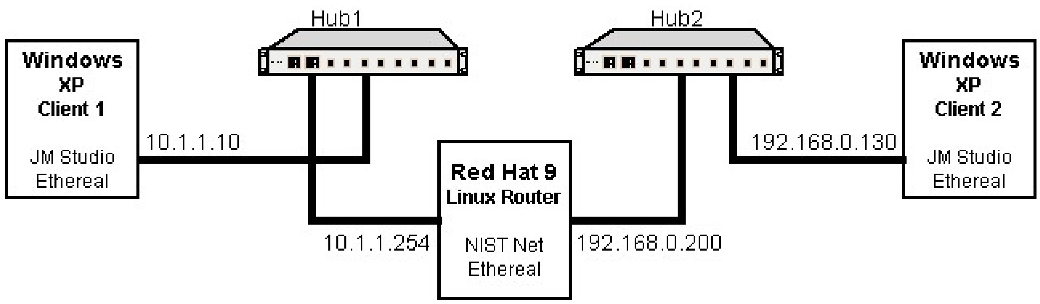

Figure 1. Experimental testbed

There are two laptops running Microsoft Windows XP, JM Studio for audio/video transfer using RTP, and Ethereal for network packet analysis. A Red Hat 9 Linux router running NIST Net is utilized to emulate Internet traffic. Ethereal program runs on both network interfaces to monitor the network traffic and control the accuracy of the emulation happening at the router level. A brief description of these applications is provided in Table 1.

Table 1. Applications running on the testbed

Name/ Type	Description
JM Studio [6] / Java based media player	The Java Media Framework API (JMF) enables audio, video and other time-based media to be added to applications and applets built on Java technology. JMStudio is an application developed based on JMF which can capture, play, record audio and video files. It can also receive and play RTP Media Streams .
Ethereal [7] / Network Analyzer	It is a free network protocol analyzer for Unix and Windows. It provides features to examine data from live network or from a capture file on disk.
NIST Net [8] / Network Emulation Package	It allows a single Linux box set up as a router to emulate a wide variety of network conditions, such as packet loss, duplication, delay and jitter, bandwidth limitations, and network congestion. This allows testing of network-adaptive protocols and applications in a lab setting.

The telemedicine video sequence selected for the experiments was a telediagnosis case in ophthalmology (this is an eye testing video sequence by a physician). It was obtained from the Regenstrief Institute for Health Care in mpeg video format. The high quality video file (which is originally 800 pixels x 720 lines) obtained in mpeg format is resized to 352 pixels x 288 lines (CIF size) based on the telehealth technology guidelines [9] where technology standards were first defined for teleophthalmology. These standards were categorized under different purposes of ophthalmology examinations. According to the purposes identified in these standards, the video used in our experiments presents an eye examination for external assessment. The technology guidelines for real-time external assessment indicate that CIF image size is suitable for this type of examinations.

3.2 Experiment Design

In experimental design each variable that affects the response variable and has several alternatives is called a factor. Quality of video depends upon several factors. The experiments in our study are designed to isolate the effects of each factor from those of others so that meaningful results can be obtained. Proper experimental design allows determining if a factor has a significant effect or if the observed difference is simply due to random variations caused by measurement errors and parameters that are not controlled [10].

In our study, general full factorial design was selected with three factors. In order to capture experimental errors, we had 2 repetitions of each experiment and calculated the experimental error. Table 2 presents the three factors – variables that affect the video quality – used in this experimental design and their levels – how they are manipulated. These levels are selected based on previous studies on the Internet backbone behavior [3] and preliminary experiments. Video codecs and codec related parameters are controlled during the experiments. Video codec selected for experiments is H.263 with CIF (352x288) video size. This experimental design required 5*5*4*2=200 experiments including the repetition factor. Each experiment generated one processed video file. Once objective measurements are completed, the amount of degradation based on PSNR was calculated.

Table 2. Factors and their values

Factors	Level 1	Level 2	Level 3	Level 4	Level 5
Delay(ms)	50	100	200	300	400
Delay Variation/ Jitter (ms)	0	2	5	10	25
Drop (%)	0	5	10	25	-

Response variable used in this study is the peak-signal-to-noise-ratio (PSNR) objective quality measure. The PSNR is usually reported in decibels (dB) [2]. An image with a PSNR of 25 dB or below is usually unacceptable. Between 25 dB and 30 dB, perceived quality usually improves and above 30 dB, images are often perceived as good as the original image. Markopoulou [12] notes that the PSNR is exclusively used as a quality measure, partly because of its mathematical traceability and partly because of the lack of better alternatives. It has also been noted [5] that the PSNR does not always correlate well with subjective measures.

The quality degradation in the video files is measured objectively using Video Quality Metric (VQM) Software [11], which was developed by the Institute for Telecommunication Sciences (ITS). It is designed for bench-top laboratory testing and is available for the PC and selected UNIX platforms. The tool implements video calibration algorithms (i.e., spatial registration, valid region estimation, system gain and level offset, temporal registration), root cause analysis algorithms (i.e., calibration problem detection, video artifact detection), and five video quality models (i.e., TV, Videoconferencing, General, Developer, PSNR). The videoconferencing and PSNR models are used during the experiments to calculate the video quality.

3.3 Experiment Protocol

Original mpeg video was first downgraded to CIF size and converted into avi format since the VQM tool used requires video sequences in avi format. Later this avi file was edited to create a video sequence of 14 seconds and black frames are included at the beginning and at the end of the file to eliminate any loss as a result of initialization or termination problems. This new file is called the *original sequence* from this point on.

In the experimental testbed illustrated in Figure 1, the original sequence was stored on *Client 1*. JMStudio was started on *Client 1* and *Transmit* option was used to send the original sequence using H.263 over RTP. On the receiving side, *Client 2* was running JMStudio with an RTP session on the port *Client 1* was transmitting video. Once the RTP session was established, *Export* option of JMStudio was used to store the transmitted file on *Client 2* in avi file format with YUV video color option. Meanwhile, NIST Net was set on the Linux Router to emulate a network based on different levels of factors in the experimental design. Traffic on all the network interfaces in the experimental testbed was analyzed using Ethereal during the transmission of the multimedia file.

4. Results

Initial analysis was conducted at the bivariate level as illustrated in Table 3. Results of correlations indicate that there is a negative relationship between video quality and jitter level and a positive relationship between video quality and the frames per second (fps) the video sequence is transmitted and recorded. An interesting finding was regarding the negative relationship between the delay level and the fps video sequence was recorded.

Table 3. Correlation between factors

	DELAY	JITTER	DROP	FPS	PSNR
DELAY	1				
JITTER	.000	1			
DROP	-.120	.000	1		
FPS	-.316(**)	-.249(*)	-.162	1	
PSNR	-.173	-.589(**)	-.222(*)	.479(**)	1

Correlation is significant at the 0.01 level (**), 0.05 level (*) (2-tailed).

Findings of the stepwise linear regression indicate that jitter has a significant effect on the video quality. Delay did not stay in the equation and hence has no effect on video quality degradation. Number of frames per second also has a significant effect on the quality as indicated by the regression results. Packet drop stayed in the equation, however its effects are not significant when compared to other factors. These results point out that the degradation in the video quality is mainly caused by the variation in delay and delay does not have any effect on this degradation. Insignificant effects of delay may be an effect of the experimental setup. The video is streamed from one end to another. Lack of interactivity may result in higher toleration of delay since the need for synchronizing two parties is not necessary.

Table 4. Stepwise Linear Regression Results

R	R^2	Adj. R^2	error
0.690	0.476	0.453	0.724

Factor	beta	t	sig.
Jitter	-0.481	-5.299	<.001
fps	0.316	3.469	0.001
Packet Drop	-0.187	-2.150	0.035

Based on the results of the experiments, following conditions should provide a PSNR value of 25dB or more which corresponds to acceptable video quality: (1) no delay variation/jitter, no packet drop, up to 400ms delay; (2) no delay, no delay variation/jitter, up to 5% packet drop.

In this experimental model, we had $2^3 - 1 = 7 effects$, where 3 are main effects (delay, jitter, drop), 3 two-factor interactions (delay-jitter, delay-drop, jitter-drop), 1 three-factor interaction (delay-jitter-drop). In order to understand the effects of individual factors and their interactions on the video quality, further analysis will be conducted using ANOVA and they will be reported during the presentation.

5. Conclusion

This research identified the effects of network impairment on video quality for a telemedicine video. Main contributions are as follows: (1) significant effects of jitter on video with high movement is identified, (2) certain cutoff points for the three values were provided to achieve reasonable video quality. Future research will use these findings in predicting video quality real time based on the three factors utilized during this study.

6. References

[1] 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 the Proceedings of 38th Annual Hawaii International Conference on System Sciences (HICSS- 38), Hawaii, USA, 2005.

[2] D. A. Rosenthal, "Analyses of selected variables effecting video streamed over IP," *International Journal of Network Management*, vol. 14, no. 3, pp. 193-211, 2004.

[3] A. P. Markopoulou, F. A. Tobagi, and M. J. Karam, "Assessing the quality of voice communications over internet backbones," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 747-760, 2003.

[4] M. Hassan, A. Nayandoro, and M. Atiquzzaman, "Internet telephony: services, technical challenges, and products," *IEEE Communications Magazine*, vol. 38, no. 4, pp. 96 - 103, 2000.

[5] S. Mohamed, "Automatic Evaluation of Real-Time Multimedia Quality: a Neural Network Approach." Rennes: University of Rennes I, 2003.

[6] 2005, "http://java.sun.com/products/java-media/jmf/2.1.1/jmstudio/jmstudio.html," *accessed on: June 15, 2005*.

[7] 2005, "http://www.ethereal.com/," *accessed on: June 15, 2005*.

[8] R. Blum, *Network Performance Open Source Tool Kit: Using Netperf, tcptrace, NIST Net, and SSFNet*. Indianapolis, Indiana: Wiley Publishing Inc., 2003.

[9] C. W. Flowers, "Technology Guidelines by Type of Clinical Service: Ophthalmology," in *Telehealth Technology Guidelines*, J. Tracy, Ed. http://telehealth.hrsa.gov/pubs/tech/techhome.htm: Office for the Advancement of Telehealth (OAT), Health Resources and Services Administration, accessed on May 24, 2005, 1999.

[10] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modelling*. New York, NY: John Wiley & Sons, Inc., 1991.

[11] S. Wolf and M. Pinson, "Video Quality Measurement Techniques," U.S. DEPARTMENT OF COMMERCE - National Telecommunication and Information Administration NTIA Report 02-392, June 2002.

[12] A. P. Markopoulou, "Assessing the Quality of Multimedia Communications Over Internet Backbone Networks," in *Department of Electrical Engineering*. Stanford, CA: Stanford University, 2002.