Audio Transmission over the Internet: Experiments and Observations*

Lopamudra Roychoudhuri, Ehab Al-Shaer, Hazem Hamed and Gregory B. Brewster
School of Computer Science, Telecommunications and Information Systems
DePaul University
Chicago, IL., U.S.A.
{lroychou,ehab,hhamed,brewster}@cs.depaul.edu

Abstract—**The performance of IP telephony systems is highly dependent on the audio codecs and their reaction to packet loss and instantaneous delays. Understanding the interaction between audio encoding and the dynamic behavior of the Internet is significant for designing adaptive audio transport mechanisms. For this purpose, we conducted a large-scale audio transmission experiment over the Internet in a 12-month period using various Internet sites. As a result of this experiment, we have made a number of new observations to assess the audio quality of G.711 and G.728 codecs under different loss and delay conditions. The paper also states a number of recommendations for implementing efficient adaptive FEC and playout mechanisms.**

Keywords- packet audio; Internet experiments; audio codecs; IP telephony; VoIP.

I. INTRODUCTION

The human ear is more sensitive to quality degradation than the human eye [9]. Hence it is crucially important to maintain the quality of the audio in a multimedia transmission. In the absence of guarantee of quality in the current Internet, one needs to get a feel for the network criteria, such as bandwidth, packet loss, delay and jitter present in the Internet in order to study the impact of these factors on the quality of the received audio. Also, since current codecs have a very diverse range in terms of degree of compression and underlying technologies, one can intuitively say that they should react differently under different network conditions.

For this purpose, we conducted a set of experiments on the Internet over a period of one year. We took two speech segments; encoded with two different encoders, and sent them to fourteen echo servers in various parts of the United States and the world. Round trip time (RTT), jitter, packet loss and out of order packets were calculated from the echoed audio data. Selected clips from the echoed audio data were tested for subjective audio quality analysis and to observe the effects of network condition on different encoders.

In the area of examining the effects of network performance on audio perception, we studied the performance of two different audio codecs: PCM Mu-law and G.728, using the data collected from our experiments. The codecs were compared in terms of how they were affected by different loss conditions of total packet loss, loss of one packet ('single loss') and losses of consecutive 2, 3, 4, and greater than 4 packets ('loss bursts'). We have observed that under similar loss conditions, the two codecs are affected very differently in terms of audio quality. Thus the quality of an audio transmission depends not only on the network conditions in general, but also on the effects of these conditions on the choice of codecs used. These observations thus provide valuable insight for designing mechanisms: (1) where codecs can be switched efficiently based on the changing loss conditions, such as total loss ratio and degrees of burstiness, and (2) where multiple encoders can be mixed in a certain ratio for bandwidth optimization.

The area of Internet measurement is not new. Some of the well-known work has been done by Paxson [16], who used TCP data for this purpose. Some researchers used UDP probes for similar work [1,2,4,12] that was not necessarily audio data, whereas in some other studies audio data compressed with a particular codec was used [3,5,14]. The distinction of our work from these lies in using more sites, different codecs and longer durations of experiments. But more importantly, our studies have revealed new areas of observations. For example, the Internet delay-related observations, which include cases of strong correlation of RTT variation and packet loss ratio and differences at network-level and application-level RTT, have been presented in [17]. This paper presents another subset, namely the observations regarding packet loss impacts on different audio codecs. With each observation, we have presented analysis and recommendations that are highly relevant to adaptive audio transmission. This work is the pre-cursor of designing a control mechanism for audio transmission that adapts to the network conditions and dynamically maintains the received audio quality.

We describe the experiments in section 2. Section 3 describes and explains our observations. We present the related work in section 4. We summarize the conclusions and enumerate our future work in section 5.

II. DESCRIPTION OF EXPERIMENTS

Tools. Two speech segments, each of 5 minutes, one using G.711 [6] PCM Mu-law and the other using G.728 [7] CELP coding were recorded using hardware audio compression by Osprey [19]. PCM Mu-law is non-linear logarithmic PCM encoder, of sampling rate 8000 samples/second, 8-bit sample coding, and bit-rate of 64 kbps. G.728 is a more compressed

*This material is, in part, based upon work supported by the National Science Foundation under Grant No. 0125068. Any opinions, findings, conclusions or recommendations are those of the authors and do not necessarily reflect the views of NSF.

TABLE I. EXPERIMENT SITES

Site,City,State/Country	Smooth RTT min/max(ms)
US	
a.psc.edu, Pittsburgh, PA	41.16/61.09
ftp.cse.buffalo.edu, Buffalo, NY.	36.05/129.82
caltech.edu, Pasadena, CA.	101.70
ftp.gatech.edu, Atlanta, GA.	61.03/85.03
beach.utmb.edu,Galveston, TX.	169.57
www.udallas.edu, Dallas, TX.	46.72
www.ucsf.edu, San Francisco, CA.	70.26
International	
ftp.rhrz.uni-bonn.de, Bonn, Germany	139.83
ftp.contrib.de, Germany	115.35
psy.uq.edu.au, Brisbane, Australia	234.42/363.13
isc.asueng.eun.eg, Cairo, Egypt	241.44/329.42
ecnet.ec, Ecuador	613.91
sunsite.ust.hk, Hong Kong, Hong Kong	318.27
www.ust.hk, Hong Kong, Hong Kong	259.31

audio encoding of indexing rate 1600 and bit-rate 16 kbps. The speech segments were clear and of excellent quality to begin with. A program was developed and used to read and encapsulate audio packets from the speech segment file, send the packets to a remote UDP echo server, and then receive and save the echoed data in log files for postmortem analysis. The packets were time-stamped and sequence-numbered so that various path criteria such as latency, jitter, packet loss (single and loss bursts), out of order packets, could be calculated. In addition, we used ping and traceroute during the experiment to record the path characteristics. The received audio clips were classified based on the path characteristics and their audio quality was measured subjectively using the Mean Opinion Score (MOS) [22].

Experiment Duration and Locations. We ran our experiment with fourteen UDP echo servers from various parts of the world and the U.S. (Table 1). In each experiment, both G.728 and Mu-law audio data were sent from DePaul University to each server every four hours for a week, a number of times between February 2001 and February 2002. Each transmission experiment lasted for 140 minutes and was repeated 48 to 54 times in a row. A total of 2010 data files and echo audio files were saved out of these experiments. Graphs were generated based on the aggregate data collected during this year. The smoothed RTT values were also calculated for each site in each run using the Exponential Weighted Moving Average with a smoothing factor of 0.9. In Table 1, the two smoothed RTT values for a particular site indicate values derived from different experiments. We observed that the routes, number of hops and smoothed RTT values varied considerably for some sites over one year.

III. OBSERVATIONS

This section presents the observations that we made after collecting and analyzing the audio file data and the traceroute data. These observations are divided into two categories: pertaining to the Internet in general, and pertaining to audio transmission in particular. With each observation, we have presented an analysis and a recommendation pertaining to audio transmission.

A. *Observations pertaining to the Internet Behavior*

In general, we have noticed that, compared to earlier studies, the Internet has become busier in terms of showing much less diurnal and weekly cycles. Number of hops, measuring distance in the networking world, is not an indicator of a path/link quality anymore – in other words, for a site to have a greater number of hops in the path does not really imply that the link is worse in terms of delay and congestion. But most of our notable observations have been in the area of packet loss/RTT variation correlation, and choices of RTT measurement techniques. From these observations we realize that RTT is an important factor, which needs to be considered in predicting packet loss and estimating FEC (Forward Error Correction [3,10]). A comprehensive report of this section can be found in [17]. In summary, we have observed that:

- RTT increase is an indicator of packet loss at times. However, this correlation is not linear.
- Considering end-to-end RTT alone may not be enough to predict packet loss and congestion. Per-hop RTT measurement is also necessary for this purpose.
- A high smoothed RTT value alone does not always reflect that the path/link is congested, especially for International sites. Similarly, number of hops alone is not an indicator of suitability of the link for VoIP.
- The network and application level RTT show more or less the same variation pattern in general. But on occasions, the network and application level RTT vary from each other considerably, signifying that the network-level RTT may not necessarily match the application level RTT.

B. *Packet Loss Impact on Audio Quality*

After collecting a number of audio echo files as large as 2010 samples from the experiments, we chose 36 representative files. The selections were done based on total percent loss and loss burst percentages. We used a sufficient number of human testers from various linguistic backgrounds, and did subjective quality Mean Opinion Scoring (MOS) by listening to the clips and judging their clarity and intelligibility. The following is the five-category rating scale of MOS: 5-Excellent, 4-Good, 3-Fair, 2-Poor, and 1-Bad. According to ITU standards, Mu-law scores 4.2 and G.728 scores 4 in the MOS scale [23]. Table 2 contains a subset of the tested files in terms of loss ratio percentage data and scores obtained, with the first entry under each category being the original file with perfect score. Fig. 1 is the graph showing the correlation of the MOS scores with the total loss ratio.

Observation 1: *Under high loss situations, G.728 performs more poorly than Mu-law in terms of speech intelligibility. But as the loss decreases, G.728 gradually improves in quality – under low loss conditions G.728 becomes better than Mu-law. However, for both Mu-law and G.728, the threshold of total loss ratio for acceptable speech quality is comparable.* From Fig. 1 we see that G.728 rates significantly lower than Mu-law consistently when the loss ratio is higher than 15%. But this

TABLE II. SUBJECTIVE AUDIO QUALITY MEASUREMENT FOR SELECTED AUDIO ECHO FILES

File	Loss Ratio %	Single loss ratio %	Loss burst ratio %	MOS
Mu-law				
1	0	0	0	4.20
2	2.6	93.85	6.15	3.60
3	8.16	81.86	18.14	2.80
4	13.4	78.21	21.79	2.42
5	15.48	71.58	28.42	2.20
6	16.04	53.37	46.63	2.54
7	21.6	59.44	40.56	2.06
8	30.44	44.94	55.06	1.24
G.728				
1	0	0	0	4.00
2	2.88	91.67	8.33	3.74
3	8.04	67.16	32.84	3.30
4	10.92	87.91	12.09	2.92
5	11.28	61.7	38.3	2.52
6	15.25	62.2	37.8	2.16
7	21.49	57.54	42.46	1.72
8	28.93	52.7	47.3	1.04

trend changes around 15% loss. Both the encoders score comparably (2.1-2.5) in the range of 11% to 15% loss ratio. Under 10% loss ratio, G.728 consistently performs better than Mu-law. This observation is pertinent when the packet sizes of Mu-law and G.728 are comparable.

One 960-byte packet of Mu-law contains 120ms (5mins/2500) of speech information, whereas one 720-byte packet of G.728 contains 360ms (5mins/833) of speech information (the 960-byte size of a Mu-law packet was selected based on Osprey recommendation to support real-time audio with minimum delay, whereas the 720-byte size of a G.728 packet was enforced by Osprey). According to [21], the duration of a phoneme, or the smallest unit of speech comprehension is 40-80ms. Hence a single packet loss affects G.728 by losing about 6 phonemes, whereas a single loss of a Mu-law packet loses about two phonemes. Hence in the case of a packet loss for G.728, a significant part of a word is lost, whereas, a packet loss in Mu-law loses a smaller number of phonemes, hence a word can be guessed, especially in the case of single loss. Under high loss situations, though Mu-law loses a greater number of packets (741 for Mu-law vs. 241 for G.728 in a 30% loss situation), it loses less audio information than G.728 at each occurrence of loss. Hence it maintains an overall better quality than G.728. G.728, on the other hand, becomes almost unintelligible with big chunks of audio data being lost, with MOS scores of 1.72 and 1.04 respectively in a 21.49% and 28.93% loss conditions, as opposed to Mu-law scoring 2.06 and 1.24 under similar loss conditions (Table 2).

Under lower loss conditions, G.728 has a smaller number of interruptions, since the number of packets lost is fewer (41 for G.728, 122 for Mu-law in a 4.9% loss situation). Each loss occurrence loses part or full word, but the rest of the speech is clear and uninterrupted. Mu-law, on the other hand, loses smaller number of phonemes at each loss occurrence, but it has about three times more loss occurrences, hence more interruptions, than G.728. Hence it seems that it becomes equally or more annoying than G.728 in overall terms of speech quality.

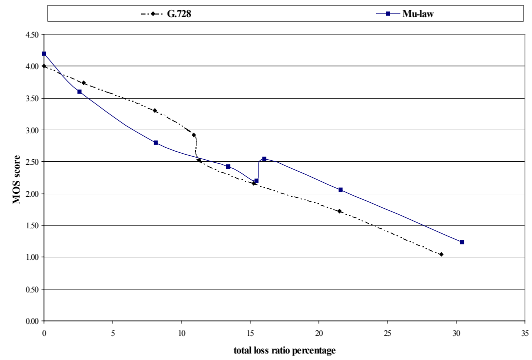

Figure 1. MOS Score vs. Total Loss Ratio

Since a G.728 packet contains three times more speech information than Mu-law, it was expected that G.728 would be more susceptible to packet loss than Mu-law in terms of intelligibility. Unintuitively, we found that both G.728 and Mu-law total loss ratio thresholds for acceptability (MOS score 3.5) were around 2.5%-3%. Since a G.728 packet contained about 6 phonemes, it appears that though part of a word is lost, under low loss conditions a human can guess, and hence understand the speech, even if 6 consecutive phonemes are lost.

This observation makes us realize that the following: (1) an adaptive mechanism needs to be aware of such loss thresholds that affect the behaviors of codecs, (2) under low loss conditions, a high bit-rate codec is replaceable by a low bit rate codec without any significant quality sacrifice, which provides an efficient way of saving bandwidth.

Observation 2: *Burst loss seems to be a more critical factor for G.728 than it is to Mu-law.* In Table 2, under G.728 we see that G.728 scores drastically lower for file 5 (score 2.52, loss burst ratio 38.3%) than file 4 (score 2.92, loss burst ratio 12.09%) even though their total loss ratios are comparable (11.28% vs. 10.92%). Whereas, for Mu-law, the score for file 5 is worse than file 6 (2.20 vs. 2.54), even though both total loss ratio and loss burst for file 6 are actually higher than file 5 (16.04% vs. 15.48%, 46.63% vs. 28.42%).

It was also noted that, for the lower scoring G.728 file 5, burst-of-2 ratio was 50.9% of the total loss burst, bursts-greater-than-4 ratio being 30.5%, whereas for higher scoring file 4, these figures were 72.7% and 0% respectively. Hence G.728 is not only susceptible to the total loss burst ratio, but also to bursts of higher degree. The same figures for higher scoring Mu-law file 6 were 59.9% and 13.9%, as opposed to 78.2% and 0% for lower-scoring file 5. This shows that the Mu-law score was not dependent on the degree of burstiness. The details of this observation can be found in [18].

A burst of two or three consecutive packets in Mu-law will lose about 4-6 phonemes (it appears that humans can tolerate a loss of that degree). But a burst of two to three consecutive packets in G.728 loses twelve to eighteen phonemes, which deteriorates the speech quality sharply. In other words, the MOS value of Mu-law is more determined by the total loss percentage, rather than the loss burst percentage. In contrast, the MOS value of G.728 is more critically dependent on the

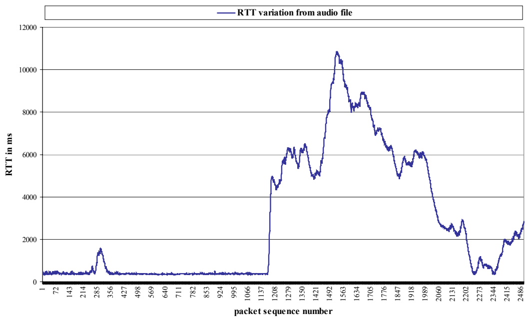

Figure 2. Packet loss vs. RTT – isc.asueng.eun.eg Day-8 12:00pm

loss burst percentage and degrees of burstiness.

Observation 3: *In the case of high loss, correcting the high percentage of random packet loss by FEC for Mu-law will consume higher bandwidth, compared to G.728. Similarly, in the case of low loss, there is a loss threshold under which the gain achieved by FEC techniques may not justify the overhead.* We have observed that sites have loss ratios as high as 25% to 30% at times. We also see that, out of total loss, single loss has a much higher percentage than loss bursts (~81% vs. 19% in the case of U.S. sites, 77% vs. 23% in case of International). Single packet loss is correctable by FEC methods, but FEC requires bandwidth overhead. For example, there is 33% overhead if 1 parity packet is generated per 2 original packets in parity-coding XOR technique [11]. Hence under high loss situations, when the link is already congested, FEC will cause more congestion when Mu-law is used, as it consumes 4 times more bandwidth than G.728. On the same note, adding FEC to a transmission on a low-loss link will be wasteful, since it will just add unnecessary traffic to the network without improving the audio quality significantly. For example, from the observations under G.728 (Table 2), at the loss ratio of 2.88%, the MOS value is 3.74 (above acceptable), and at the loss ratio of 8.04%, the MOS value is 3.30 (close to acceptable) – hence it can be safely assumed that for G.728, adding FEC to a low loss condition of 3% will not serve any worthwhile purpose, since the speech is already acceptable in spite of the loss.

Hence an efficient FEC scheme should be based not only on packet loss ratio, but also be based on the codec used and the percentage of the link capacity left, for efficient bandwidth use.

C. Delay Impact on Audio Quality

Delay and jitter are two of the most critical factors that affect the quality of audio transmission. In real-time, if the packets are delayed too long, they have to be discarded, which would appear as loss of packets. Similarly, in the presence of jitter, it is not possible for the receiving site to play out an audio packet as soon as it receives it, thus negatively affecting the quality of real-time voice communication.

Observation 4: *For almost all sites, the smoothed traceroute RTT value is less than the maximum permitted delay bounds for audio. But at some times, the per-packet RTT exceeds the allowable limits drastically. The per-packet jitter is also observed to exceed the allowable limits considerably.*

In ITU-T Recommendation G.114 [20], the acceptable delay is 150ms for most real-time user applications in traditional telephony using G.711. The M2E (Mouth to Ear) delay bounds (the time that elapses between the moment the talker utters and the moment the listener hears, that can be tolerated for voice data) are extended for various low bit rate compression schemes [20]. Provided one is aware of the impact of delay on the quality of user applications, the mouth to ear delay bound for G.711 is 400ms and that for G.728 is 324ms, beyond which the delay is not acceptable. The corresponding values for RTT are 800ms and 648ms respectively. Except for one site, all the smoothed RTT values fall well under these thresholds, although the smoothed RTT value differs considerably among the sites (Table 1).

In contrast, analyzing the data behind Fig. 2, we see that for Egypt on Day 8 12:00pm, the per-packet RTT exceeds the allowable values of 648ms and 800ms 47% of the times respectively, the RTT values going up as high as 10861ms. On another occasion for Egypt (Day 4 4:00am), the per packet RTT exceeds the allowable values 100% of times, RTT going up as high as 5832 ms.

In contrast to audio delay bounds, the audio jitter bound should be around 50ms [8] at most, so that the receiver can smoothly playback. We have observed that the RTT variation consistently exceeds the above limit, particularly for the International sites. For example, analyzing G.728 data for Australia on Day-1 4:00pm, we observed that the delay variation, i.e. the difference of the per-packet delay with that of the next packet, exceeds the limit of 50ms 36% of the times, whereas the delay exceeds the allowable limit only 16% of the times.

This reveals that jitter is a persistent problem on the Internet, and hence delay adjustment buffering and jitter smoothing algorithms need to be in place under all situations to counteract it. Delay, on the other hand, exceeds allowable limits as exceptions or bursts, only for some sites, as seen in the observations. This means that delay adjustment mechanisms need to operate in two modes – a *regular* mode with normal buffering and loss concealment, and an *exception* mode, which needs to be equipped to handle high delay conditions more robustly. There are numerous citations of adaptive delay adjustment algorithms [15], in which the authors present an adaptive delay adjustment algorithm that tracks the network delay of recently received packets and dynamically adjusts playout delay. This area needs to be explored in light of our findings.

IV. RELATED WORK

Comparing our experiment with those of Paxson [16], that work was done to measure the Internet in general. In '94 and '95, most of the Internet traffic was TCP. Our experiment with UDP audio data is more relevant with multimedia over the current Internet. In a similar work done in 1992, Bolot [2] examined one connection from France to the U.S. to analyze the end-to-end packet delay and loss behavior over the Internet. Maxemchuk and Lo [12], in order to examine loss and delay over the Internet, and enumerate factors affecting the quality of a connection, present experiments where they generated traffic by three sources sending data to a U.S. destination. Borella et.al. [4] analyzed Internet packet loss statistics for speech transmission using three different Internet paths in the U.S. Arai et.al. [1], in their attempt to measure packet-loss ratio, its time dependency and the frequency of bursts in the Internet, conducted experiments with probe UDP packets from four host sites to one destination in Japan. Bolot and Garcia [3], as part of the initial study for an adaptive audio control mechanism, did a set of experiments using PCM coder between France and the UK. Hagsand et.al. [5] present and examine observations in terms of loss, delay and jitter from a similar experiment done with four sites and with PCM-encoded voice recording. Mohamed et.al. [14], as part of presenting a neural network control mechanism, present and examine observations for four sites. Mena and Heidemann [13] analyze streaming Read Audio traffic emanating from a popular Internet audio service.

All the above experiments suffer from limited number of sites and limited number of encoders used in the experiments. Our experiment is comparable in time and duration to the above papers. In addition, it includes more sites from more varied geographical locations around the globe, as well as within the United States. This has given us a better understanding of the network conditions over current global Internet. We used two compression techniques instead of one, in order to get an understanding of how the above factors affect each of these schemes, and what parameters can be changed for each scheme in order to perform optimally. The observations listed in this paper have not been addressed previously.

V. CONCLUSIONS AND FUTURE WORK

This paper provides important insights into Internet in order to develop adaptive and robust transport protocols for audio transmission over the Internet. Following is a summary of the observations and recommendations out of our experiments.

- Various attributes of packet loss (such as high vs. low loss, single vs. burst packet loss) can have vastly different effects on different audio codecs. For example, under high loss conditions of 15% or above, G.728 performs much worse than Mu-law, whereas, the behavior reverses under lower loss conditions (below 10%). Adaptation techniques need to be aware of these thresholds to provide appropriate loss recovery (FEC) and efficient bandwidth utilization.

- Jitter is seen to be a more persistent problem in the Internet. Hence jitter smoothing algorithms are needed to be in place under all circumstances. On the other hand, delay exceeds allowable limits as exceptions or bursts. Delay adjustment algorithms thus need to operate in a *regular* mode, with normal buffering and loss concealment, when the transmission delay stays within the allowable limits, and an *exception* mode for high delay bursts, which should handle high delay conditions more robustly.

In the future, we plan to measure factors such as path asymmetry – by setting up a collaborative framework involving receivers at remote locations, which will allow one-way measurements, including other factors like per-hop RTT. We plan to conduct this experiment on a regular, periodic basis to efficiently track the changing Internet. In addition, we plan on using multicast in our transmissions. We will also explore Internet2 in terms of measurements of this kind.

REFERENCES

[1] M. Arai, A. Chiba, K. Iwasaki, “Measurement and Modeling of Burst Packet Losses in Internet End-to-End Communications,” *Pacific Rim Int. Symp. On Dependable Computing (PRDC)*, pp. 260-267, Dec. 1999.

[2] J. Bolot, “End-to-end packet delay and loss behavior in the Internet,” *ACM Sigcomm*, San Francisco CA, pp.289-298, September 1993.

[3] J. Bolot, A. Vega-Garcia, “Control Mechanisms for Packet Audio in the Internet,” *Proceedings IEEE Infocom*, San Francisco, CA, pp 232-239, April 1996.

[4] M. Borella, D. Swider, S. Uludag, G. Brewster, “Internet Packet Loss: Measurement and Implications for End-to-end QoS,” *Proceedings, International Conference on Parallel Processing,* Aug. 1998.

[5] O. Hagsand, K. Hanson & I. Marsh, “Measuring Internet Telephony Quality: Where are we today?” *Globecomm Telecommunications Conference* – Globecomm ’99.

[6] ITU-T Recommendation G.711 (11/88) - http://www.itu.int/itudoc/itu-t/rec/g/g700-799/g711.html

[7] ITU-T [G.728] Recommendation G.728 (09/92) - http://www.itu.int/itudoc/itu-t/rec/g/g700-799/G.728.html

[8] M. Karam, F. Tobagi, “Analysis of the Delay and Jitter of Voice Traffic Over the Internet,” *IEEE Infocom*, Anchorage, Alaska, April 2001.

[9] F. Kuo, W. Effelsberg, J.J. Garcia-Luna-Aceves, *Multimedia Communications: Protocols and Applications*, Prentice Hall, 1998, pp.68.

[10] S. Lin, D.J. Costello, *Error Correcting Coding, Fundamentals and Applications*, Englewood Cliffs, NJ, Prentice Hall, 1983.

[11] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman, V. Stemann, “Practical Loss-Resilient Codes,” *ACM Symposium on Theory of Computing (STOC)*, El Paso, TX, May 1997.

[12] N. F. Maxemchuk, S. Lo, “Measurement and Interpretation of Voice Traffic on the Internet,” *Proc. ICC*, Montreal, Canada, June 1997.

[13] A. Mena, J. Heidemann, “An Empirical Study of Real Audio Traffic,” *IEEE Infocom*, March 2000.

[14] S. Mohamed, F. Cervantes-Perez, H. Afifi, “Integrating Network Measurements and Speech Quality Subjective Scores for Control Purposes,” *IEEE Infocom* 2001.

[15] S. B. Moon, J. Kurose, D. Towsley, “Packet Audio Playout Delay Adjustment: Performance Bounds and Algorithms,” *ACM/Springer Multimedia Systems,* Vol 6, pp. 17-28, January 1998.

[16] V. Paxson, “End-to-End Internet Packet Dynamics,” *IEEE/ACM Transactions on Networking,* Vol 7, No. 3, June 1999, pp. 277-292.

[17] L. Roychoudhuri, E. Al-Shaer, H. Hamed, G.B. Brewster, “On Studying the Impact of the Internet Delays on Audio Transmission,” *IEEE IP Operations and Management (IPOM)*, Dallas, TX, Oct. 2002.

[18] L. Roychoudhuri, E. Al-Shaer, H. Hamed and G.B. Brewster, “Audio Transmission over the Internet: Experiments and Observations - A Technical Report,” *DePaul CTI PhD Symposium*, November 2001. http://www.mnlab.cs.depaul.edu/~lroychou/phd.htm

[19] http://www.viewcast.com/products/osprey/osprey200.html

[20] D. De Vleeschauer, J. Janssen, G. H. Petit, “Delay and Distortion Bounds for Packetized Voice Calls of Traditional PSTN Quality,” *IPTEL*, Berlin, 12-13 April, 2000.

[21] A. Watson, M. A. Sasse, “Measuring Perceived Quality of Speech and Video in Multimedia Conferencing Applications,” *ACM Multimedia* 1998 pp. 55-60.

[22] ITU-T Recommendation P.800 (08/96) - Methods for subjective determination of transmission quality.

[23] W.T.K. Wong, R.M. Mack, B.M.G. Cheetham, X.Q. Sun, “Low Rate Speech Coding for Telecommunications,” *BT Technol J* Vol 14, No.1 January 1996.