Simulation of redundant traffic encoding in VoIP systems

Samy El-Tawab [a], M. Nazih El-Derini [a], Hussein H. Aly [a] and Emad Aboelela [b]
[a] *Computers and Systems Eng. Dept., Faculty of Eng., Alexandria University, Alexandria, Egypt*
[b] *Computer and Info Science, Dept., of University of Massachusetts Dartmouth, North Dartmouth, MA 02780, USA*

The quality of Voice over Internet Protocol (VoIP) does not yet match the quality of a circuit-switched telephone network. One formidable problem is that the Internet was designed for data communications; consequently, packets suffer a long and variable delay that decreases voice quality. We proposed a multi-stream (dual) transmission of real-time voice over best-effort packet networks such as today's Internet, where multiple redundant descriptions of the voice stream are sent over independent network paths. The main target of the proposed system is to improve the quality of VoIP service over congested and slow communication links by using routing diversity while utilizing the efficiency of the Internet from the sender to the receiver and vice versa. We have used Network Simulator version 2 (ns-2) as a simulation tool that is a standard network Simulator used in large number of related research. The simulation uses two different paths to send the packets on rather than sending the packets on a single path to the receiver. Then, it tests the packet loss ratio, end-to-end delay and jitter. Our results have showed that sending a redundant copy on another path would improve quality of the VoIP system.

مازالت جودة نقل الصوت عبربروتوكول الانترنت ليست بجودة شبكة الهواتف المحلية. وتعتبر المشكلة الاساسية هى انه تم تصميم شبكة الانترنت لنقل البيانات، وبالتالى فإن الصوت يتعرض الى مشكلات التأخير والذى يقلل من جودة الصوت. نتقدم بفكرة لنقل الصوت فوق شبكات نقل البيانات مثل شبكة الانترنت، وذلك عن طريق إرسال نسخة من الصوت على مسارين مختلفين. الهدف الرئيسى هو تحسين جودة نقل الصوت عبر بروتوكول الانترنت فوق المسارات البطيئة وذلك باستخدام الارسال المتنوع مع الحفاظ على فعالية الانترنت من الكثير من الابحاث المتعلقة بالشبكات. المحاكى المستخدم يقوم بإرسال الصوت على مسارين مختلفين بدلا من مسار واحد، ويختبر نسبة الفاقد، التأخير والانحراف المعيارى للتأخير. اظهرت النتائج أن ارسال نسخة للصوت على مسار مختلف يحسن من نظظام نقل الصوت عبر بروتوكول الانترنت.

Keywords: Voice over IP, Packet loss, Latency, Jitter, Encoding, Multi-stream transmission

1. Introduction

Voice over Internet Protocol (VoIP) is the integration and convergence of voice and data networks, services, and applications. It enables voice traffic to be carried over an IP network (e.g. the global Internet). Thereby communication modes such as email, voice mail, fax, pager, real-time human speech, and multimedia videoconferencing can be merged into a single integrated system.

Because the internet is designed for data and is not dependent on fixed locations, VoIP can be used to reduce long distance voice communication cost, although at the expense of Quality of Service (QoS). Extensive research is going on to make packet switched networks more reliable for multimedia communication. As voice applications are delay sensitive, it is necessary to have a well-engineered end-to-end network to successfully use VoIP. Fine-tuning the network to adequately support VoIP services and deliver good QoS involves implementing a series of protocols and features. Traffic shaping considerations must also be taken into account to ensure the reliability of the voice connection [1].

Conventional techniques to reduce packet loss are retransmission and Forward Error Control (FEC) [2]. Retransmission is time-consuming making it unsuitable for real-time application. FEC techniques are designed to overcome the packet loss if the loss is below some threshold value. But if the packet loss is above the threshold value, then only small amount of data can be recovered [3]. In ref. [4], routing diversity along with FEC is proposed. It forwards the packets simultaneously over multiple redundant paths rather than choosing one optimal path. In ref. [3], two schemes to implement routing diversity are proposed. The first one, path diversity via

IP source routing, is not applicable as this feature is turned off on network nodes due to security reasons. The other approach, path diversity via relays, is useful to emulate route diversity behaviour but sending a stream through a series of known nodes can add more delay to the voice packets. So, it might not result in an accurate evaluation of path diversity technique.

The proposed system takes advantage of the largely uncorrelated statistical characteristics of loss and delay. As a result, the probability of a negative disturbance, such as packet erasure or increasing delay, impacting both channels at the same time will be small [5].

The goal of this proposed system is to improve the quality of VoIP service over congested and slow communication links by using routing diversity while utilizing the efficiency of the Internet. Routing diversity helps reducing the packet loss, by explicitly sending different subsets of a packet stream over different paths. It usually provides better performance than seeing the behaviour of any individual random path.

The rest of this paper is organized as follows: section 2 discusses the Components of VoIP, section 3 summarize the problems that faces the Quality of VoIP, section 4 presents our simulation model, section 6 presents the simulation results. Finally, section 7 gives conclusions and suggests some future work.

2. Components of VoIP

The Public Switch Telephone Network (PSTN) is the collection of all the switching and networking equipment that belongs to the carriers that are involved in providing telephone service. VoIP is being promoted to augment, if not eventually replace, the current PSTN infrastructure. The overall requirements of an IP telephony solution can be split into four categories: *signaling, coding, transport and gateway control.*

2.1. Signaling

Once a user dials a telephone number (or clicks a name hyperlinked to a telephone number), signaling is required to determine the status of the called party- available or busy – and to establish the call. There are multiple and complex levels of signaling that must take place in order to initiate and complete a call; their complexities escalate when VoIP users in packet networks communicate with PSTN subscribers [6].

2.2. Coding

A prerequisite for digital transmission systems is that the information to be transmitted can be converted into a sequence of pulse combinations, which are then transmitted practically without any noticeable distortion. Consequently, Analog information - such as human speech - must be converted into digital form. The accuracy of A/D conversion is crucial to the subscriber's perceived quality. The digit combination must be so detailed that the Analog speech can be reproduced without distortion or disturbances in the receiving equipment. At the same time, our ambition is to reduce the amount of digital information in order to better utilize the available network capacity. Coders are usually divided into two main classes: *waveform coders* and *voice coders (vocoders).* In addition, there are *hybrid coders* that combine the characteristics of the two main types. Waveform coding means that the amplitude variations of the Analog signal (the voice curve) are described by a number of measured values. These values are then pulse-coded and sent to the receiving end. The signal's Analog appearance is reproduced in the receiving equipment by means of the received measured values. The method makes it possible to obtain a very high level of voice quality, since the received voice curve is a true copy of the one transmitted.

The voice coder is a parametrical coder. Instead of transmitting a direct description of the voice curve, a number of transmitted parameters describe how the curve has been generated. Parametrical coding requires a defined model of how the voice curve is created. The quality will be averaged (the received speech sounds "synthetic") but, on the other hand, signals can be transmitted with a very low bit rate. A hybrid coder sends a number of parameters as well as a certain amount of waveform-coded information. This type of voice coder, which provides a reasonable compromise between voice quality and coding efficiency, is used in today's digital mobile telephone systems [7].

2.3. PCM and DPCM (differential PCM)

Pulse Code Modulation (PCM) may be chosen as an example of technologies for A/D conversion. PCM is a type of waveform coding and is standard for voice coding in the telephone network. The bit rate generated per call - 64 kbit/s - has been a decisive factor in switching and transmission design. In DPCM, A signal that has been sampled according to the sampling theorem shows a high degree of correlation between adjoining samples. In other words, two samples next to each other are relatively similar. This means that there is much to be gained - in terms of bandwidth - by coding the differences between adjoining samples instead of the absolute value of each sample. Fig. 1 shows that four bits can be used instead of eight. DPCM has one disadvantage, if the analog input signal varies too much between the samples, it cannot be represented by only four bits but will be cut off. [7]

2.4. Transport

Typical internet applications use TCP/IP, whereas VoIP uses RTP/UDP/IP. TCP/IP is a reliable connection-oriented network communications protocol suite. But it is not suitable for real-time communications, such as speech transmission, because the acknowledgement/retransmission feature would lead to excessive delays. UDP provides unreliable connectionless delivery service using IP to transport message between end points in an internet. RTP, used in conjunction with UDP, provides end-to-end network transport functions for applications transmitting real-time data, such as audio and video, over unicast and multicast network services [8].

Once Signaling and encoding occur, RTP and RTCP are utilized to transport the voice packets. Media streams are packetized according to a predefined format and placed in RTP packets. RTP provides delivery monitoring of its payload types through sequencing and time-stamping. RTCP offers insight on the performance and behavior of the stream, such as voice stream jitter. RTP and RTCP are designed to be independent of the signaling protocol, encoding schemes and network layers implemented. Applications typically run RTP on top of UDP to make use of its multiplexing and checksum services [6].

RTCP is based on the periodic transmission of control packets to all participants in the session, using the same distribution mechanism as the data packets [6].

RTCP performs the following functions:

- Provides feedback on the quality of the data distribution (primary function).
- Carries a persistent transport-layer identifier for an RTP source.
- Controls the rate in order for RTP to scale up to a large number of participants.

2.5. Gateway control

Gateways are responsible for converting packet-based audio formats into protocols understandable by PSTN systems.

3. Quality of VoIP

The basic routing philosophy on the internet is "best effort", which serves most users well enough but it is not adequate for the time-sensitive, continuous stream transmission required for VoIP. It is imperative for an implementation of VoIP to remain cognizant of quality. Quality encompasses many factors; the ones that will be examined here are QoS, packet loss, jitter and latency [6].

3.1. Quality of service

QoS refers to the ability of a network to provide better service to selected network traffic over various underlying technologies

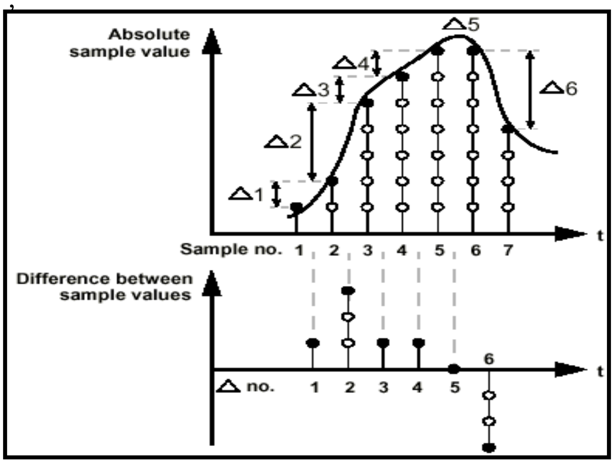


Fig. 1. Differential PCM (DPCM).

including IP-routed networks. QoS features are implemented in network routers to provide better and more predictable network service by:

- Supporting dedicated bandwidth
- Improving loss characteristics
- Avoiding and managing network congestion
- Shaping network traffic
- Setting traffic priorities across the network.

Real-time voice applications have different characteristics and requirements from those of traditional data applications. Because they are real-time based, voice applications tolerate minimal variation of delay affecting delivery of their voice packets. Voice traffic is also intolerant of packet loss, out-of-order packets, and jitter, all of which gravely degrade the quality of the voice transmission delivered to the recipient end user. To effectively transport voice traffic over IP, mechanisms are required that ensure reliable delivery of packets with low and controlled latency. [6]

3.2. Packet loss

UDP/IP networks cannot provide a guarantee that packets will be delivered at all, much less in order. Packets will be dropped under peak loads and during periods of congestion. Due to time sensitivity of voice transmission, the normal TCP-based retrans-mission schemes are not appropriate. Approaches used to compensate for packet loss include interpolation of speech by replaying the last packet, and sending of redundant information. Packet losses greater than ten percent are generally intolerable unless the encoding scheme implemented provide extraordinary robustness [6].

3.3. Jitter

Because IP networks cannot guarantee the delivery time of data packets (or their order), the data will arrive at very inconsistent rates. The variation in inter-packet arrival rate is the jitter, which is introduced by variable transmission delay over the network. Remov-ing jitter to allow an equable stream requires collecting packets and storing them long enough to permit the slowest packets to arrive in time to be played in the correct sequence. Each jitter buffer, is used to remove the packet delay variation that each packet is subjected to as it is transmitted through the network [6].

3.4. Latency

Latency is the time delay incurred in speech by the IP telephony system. One –way latency is the amount of time measured from the moment the speaker starts to talk until the listener actually hears the word. Round trip latency, of course, is the sum of the two one-way latency figures that compose the user's call. The lower the latency, the more natural interactive conversation becomes and the additional delay incurred by the VoIP system is less discernable. In a VoIP implementation that is primary used a cost-reduction or toll bypass application, studies suggest that users will tolerate one-way latency up to 20 ms. Furthermore; user perception of the link quality can be mapped in terms of one-way latency, as shown in fig. 2 [6].

4. Simulation model

The usual scenario is that, voice data is packetized, sent over a single path to the receiver, our simulation uses two different paths to send the packets on, it sends the

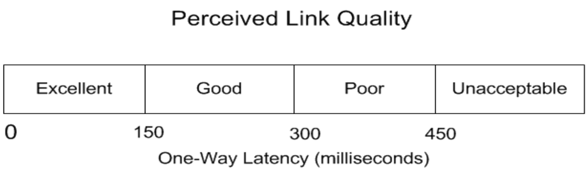


Fig. 2. Quality perception vs. latency.

same packet on the two paths and test the packet loss ratio, end-to-end delay and jitter.
In fig. 3 the setting of this model are selected according to the typical settings of a set of network configurations, Chain Configuration and Jumping Bottleneck Configuration [9], [11]. Our network contains. 4 routers, two on each path. We have two nodes connected to each other through two different paths; each path contains number of routers. We measure each of packet loss ratio, End-to-end delay and jitter in different cases. In each case, we will use different packet size (1Kbytes…64 K Bytes).All the link bandwidth varies between 128 Kbps to 1 Mbps [11].

Voice over IP (VoIP) is susceptible to network behaviours, referred to as delay and jitter, which can degrade the voice quality to the point of being unacceptable to the average user. Delay is the time taken from point-to-point in a network. Delay can be measured in either one-way or round-trip delay. To get a general measurement of one-way delay, measure round-trip delay and divide the result by two. Jitter is the variation in delay over time from point-to-point. If the delay of transmissions varies too widely in a VoIP call, the call quality is greatly degraded. The amount of jitter tolerable on the network is affected by the depth of the jitter buffer on the network equipment in the voice path. The more jitter buffer available, the more the network can reduce the effects of jitter.

Packet loss is losing packets along the data path, which severely degrades the voice application. Prior to deploying VoIP applications, it is important to assess the delay, jitter, and packet loss on the data network in order to determine if the voice applications work.

Fig. 4 shows the trace file format used in our simulator. It traces the packet number,

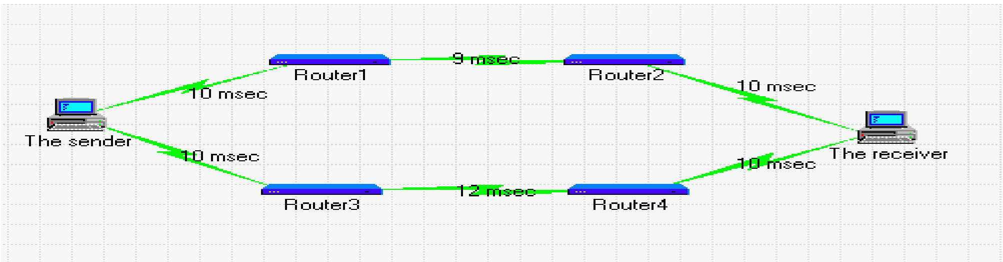
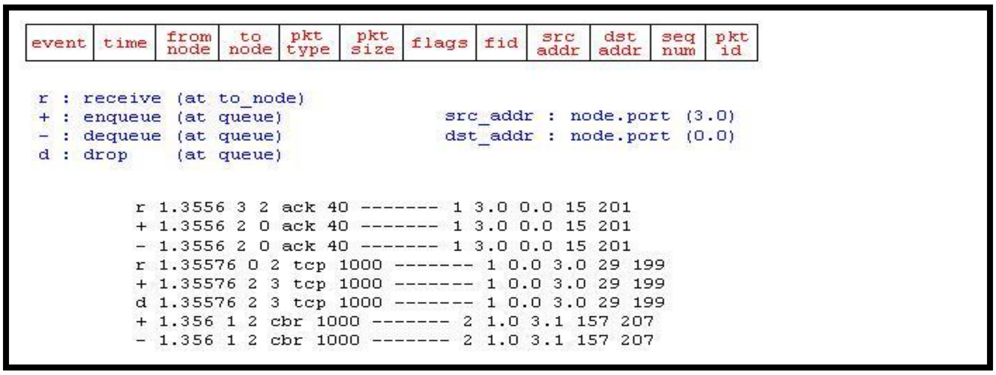


Fig. 3. The simulation model.



event	time	from node	to node	pkt type	pkt size	flags	fid	src addr	dst addr	seq num	pkt id

```
r : receive  (at to_node)
+ : enqueue  (at queue)          src_addr : node.port (3.0)
- : dequeue  (at queue)          dst_addr : node.port (0.0)
d : drop     (at queue)

r 1.3556 3 2 ack 40 ------- 1 3.0 0.0 15 201
+ 1.3556 2 0 ack 40 ------- 1 3.0 0.0 15 201
- 1.3556 2 0 ack 40 ------- 1 3.0 0.0 15 201
r 1.35576 0 2 tcp 1000 ------- 1 0.0 3.0 29 199
+ 1.35576 2 3 tcp 1000 ------- 1 0.0 3.0 29 199
d 1.35576 2 3 tcp 1000 ------- 1 0.0 3.0 29 199
+ 1.356 1 2 cbr 1000 ------- 2 1.0 3.1 157 207
- 1.356 1 2 cbr 1000 ------- 2 1.0 3.1 157 207
```

Fig. 4. The trace file format.

time sent, from node, to node, packet type, packet size, packet sequence and some flags.

5. Simulation results

We will measure the packet loss in Path1, Path2 and the total packet loss; we start by calculating the packet loss in Path1 using the following flowchart, fig. 5. Packet loss calculations in Path1 or Path2 Calculating the total packet loss in the network using the following flowchart.

We have calculated each of the packet loss, end-to-end delay and jitter in the case of same packet size (same coder) or different packet size (different coders) are sent using two paths to the receiver .We also handle the case that the receiver just store his data without using a playout buffer or time stamp for packet arrival and the case that the receiver uses a playout buffer to store, order and play the packets with time stamp for packet arrival.

In this case, we have used a copy of the packets to be sent on path 2 as a redundant copy of our packets sent on path 1. We also have used a playout buffer, the playout buffer gives a time stamp for the arrival of packets, if a packet is not received from any of the two paths until the time stamp ends, then it is supposed to be lost (time stamp) . Even if it is received later, it is dropped. The buffer also put the packets in the correct order then plays the packets when the buffer is full or the buffer has not receive any packets for a specific time. The following table shows the configuration of our simulator in this case.

The following graph shows the packet loss ratio with time, we conclude that the packet loss ratio in the new total received packets from both paths is less than any of the two paths or even their summation, also notice the increasing of the total when using a playout buffer.

The increasing in Packet loss when using the playout buffer returns to the existing of time stamp used in this case. The time stamp of 30 msec allows the packet loss ratio to increase rapidly as we can see in fig. 7. We changed the time stamp to different values and compared between the total with time stamp values of 30msec, 40msec and 50msec. The following graph shows a comparison between total packet losses with different time stamp values ranged from 30msec to 60 msec.

From the above graph, we notice that the packet loss ratio decreases when increasing the time stamp used in our simulator until it reaches the same packet loss ratio obtained when no time stamp is used. We can notice that in 30msec time stamp the packet loss ratio is higher than the one in 40msec. When reaching a specific value ≈ 60 msec, any increasing in the time stamp has no effect anymore.

The following graph shows a comparison between (Same Packet size - same encoding PCM, using playout buffer) and (Different Packet size - Different encoding PCM, DPCM - using playout buffer).

From fig. 8, we have noticed that there is a small difference between (Same packet size sent) and (Different packet size sent). This concluded that sending a copy from the packet is almost equal to sending the same packet encoded in different method.

From the pervious results, we notice that the time stamp used is a parameter that affects the performance of our model. VoIP typically tolerates delays up to 150ms before the quality of the call is unacceptable. Our model showed that increasing the time stamp decreases the amount of packet loss ratio.

5.1. End-to-end delay

We have calculated the *End-to-End delay* (the amount of time measured from the moment the speaker starts to talk until the listener actually hears the word), in our model, we calculate the end-to-end delay by the amount of time from the packet is sent until it is played). We have used a playout buffer with time stamp of 30msec.

From fig. 9, we have noticed that the end to end delay increases with time (packet id). In the beginning, the system is empty. The time taken from the sender to the receiver is the summation of delays in the links from the sender to the receiver. By time increases, the traffic increases and the end-to-end delay increases until it reaches a constant value. Also we have noticed that the total delay is the minimum between the delays of the two paths.

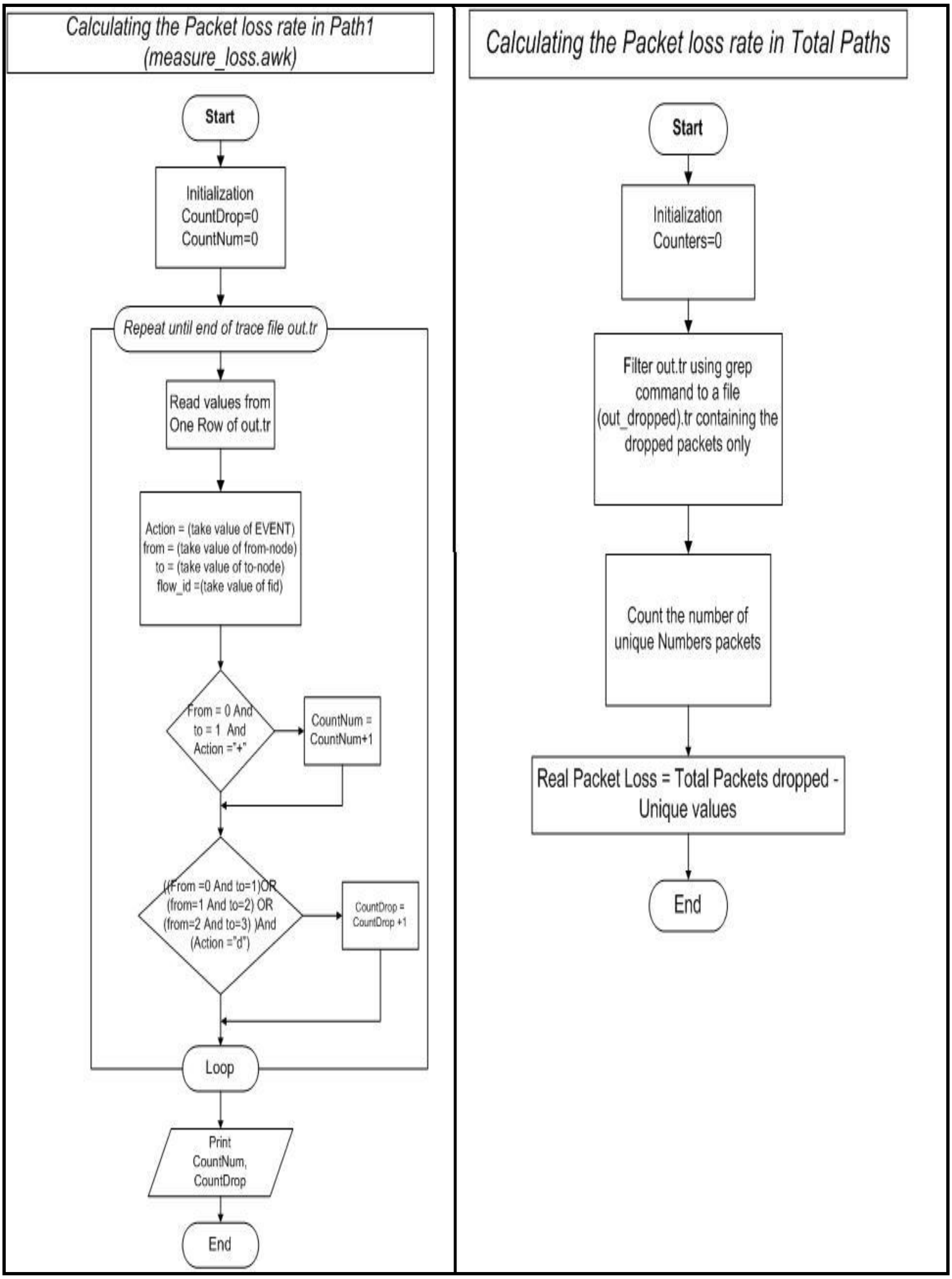


Fig. 5. Packet loss calculations in total paths.

Table 1
Configuration of our simulator

Sending	Redundant copy of the packet on the two paths
Buffer	Playout buffer (stores packets in order) and plays the content when buffer is full.
Buffer size	7000 Packets
Time stamp	Packet was not received until a specific time is considered lost.
Time stamp	30 msec

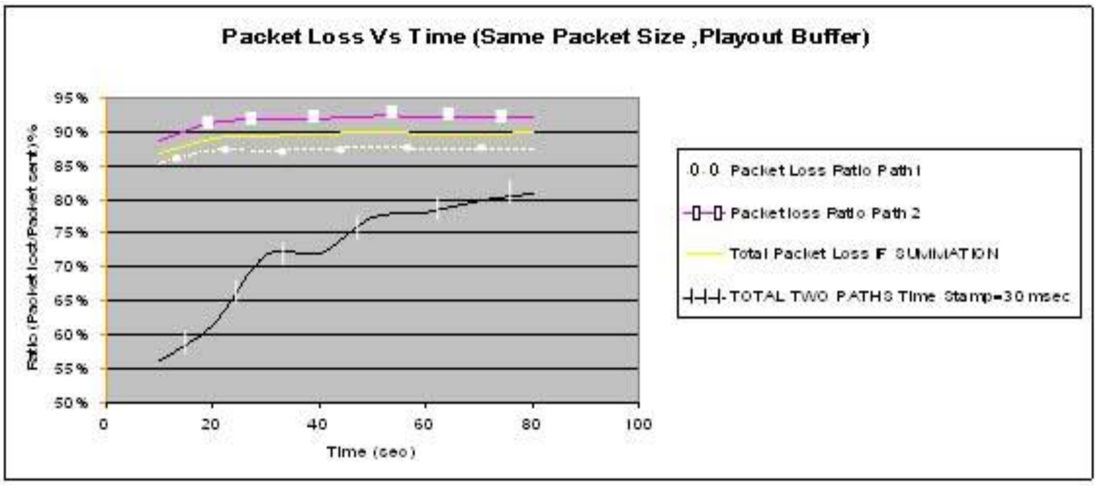


Fig. 6. packet loss case I.

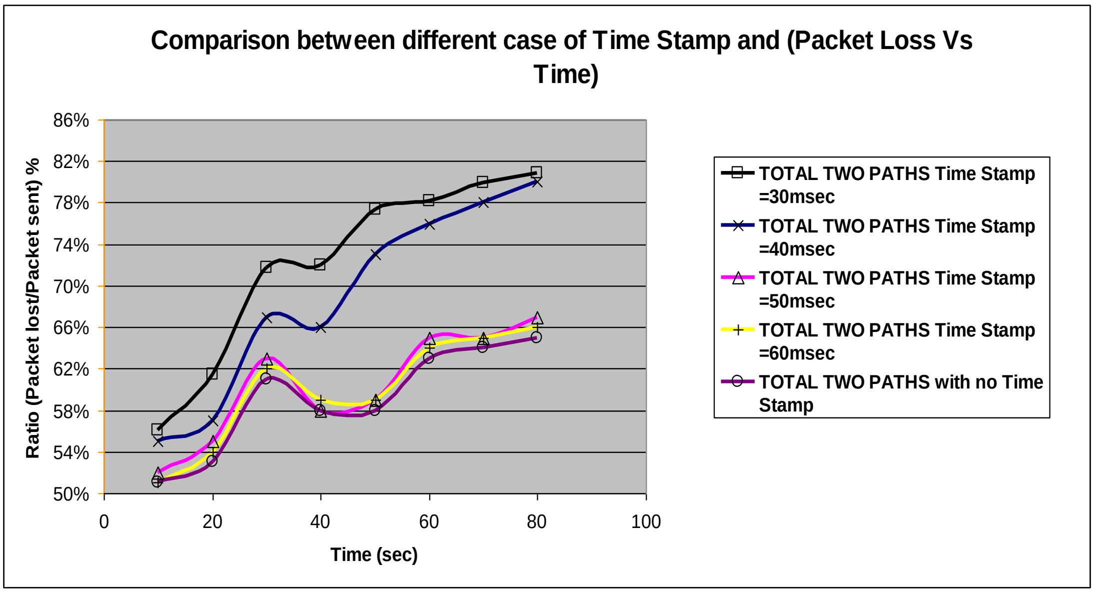


Fig. 7. Comparison between packet loss ratio using different time stamp values.

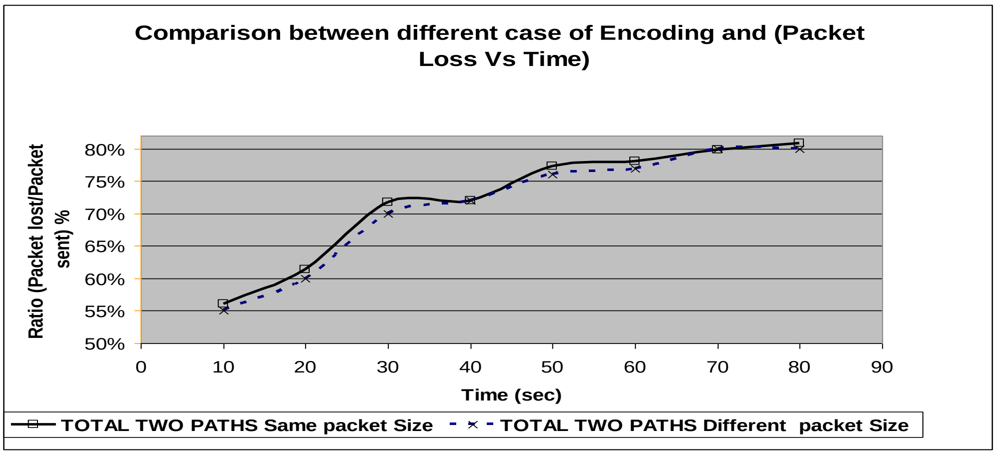


Fig. 8. Comparison between case I and II.

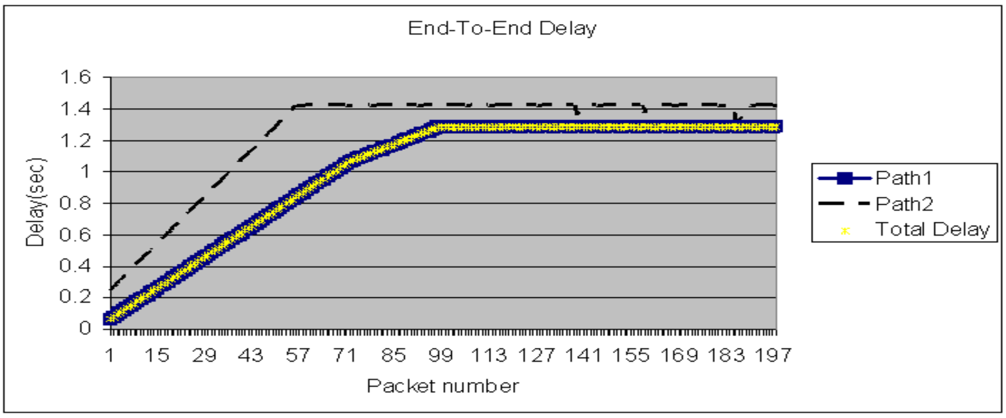


Fig. 9. End-to-end delay.

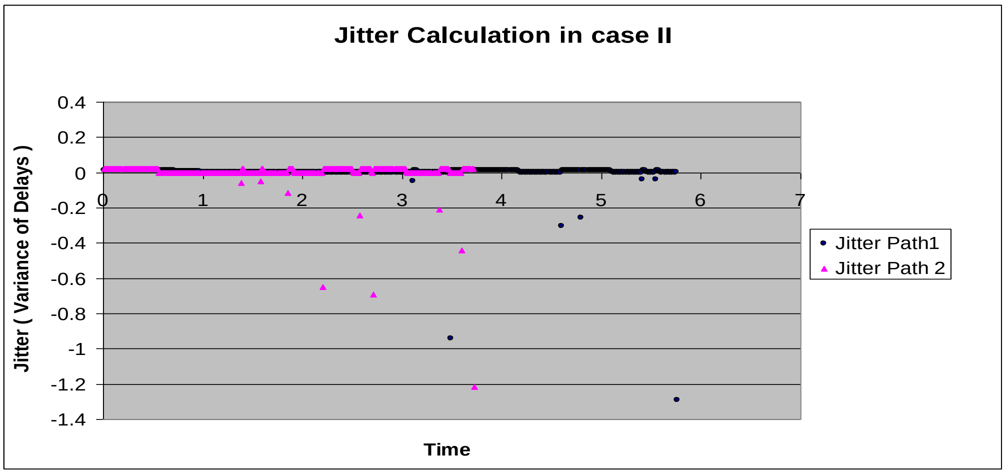


Fig.10. Jitter calculation.

Table 2
Description of some Coders [6]

Standard	Description	Bit rate	MOS
G.711	*Pulse code modulation* using eight bits per sample, sampling at 8000 Hz	64 kbps	4.3
G.723.1	Dual rate speech coder designed with low bit rate video telephony in mind. The G.723.1 coder needs a 7.5 ms lookahead and used one of these coding schemes: Multipulse Maximum Likelihood Quantization (MP-MLQ) Algebraic CELP (*ACELP*)	6.3 and 5.3 kbps respectively	4.1
G.726	Coder *using ADPCM.* Contains obsolete standards G.721 and G.723	16,24,32 and 40 kbps	2-4.3
G.728	Low Delay CELP (LD-CELP)	16 kbps	4.1

5.2. Jitter

In voice Over IP (VoIP), jitter is the variation in the time between packets arriving, caused by network congestion, timing drift, or route changes. A jitter buffer can be used to handle jitter [11].

RTCP jitter is measured in terms of packet to packet delay. If we consider the delay between two consecutive packets to be T_a and T_b, then the variation is represented as ABS (T_b-T_a). The mean of the packet to packet delay variation can be given by MPPDV = mean (ABS ($t_i - t_{i-1}$)). The MPPDV in this case represents the jitter levels in scenarios in which the packets arrive early and late in an alternate fashion.

Constant jitter: In this, the variation in delay is more or less constant.

Transient jitter: An unnatural incremental delay, sometimes only by single packets.

This transient jitter can explain the behaviour of some points in figure 10.

We used Jitter buffer of size 7000 packets is used to reduce jitter from the voice stream; however, in the process of reducing jitter, the buffers can increase delay and packet loss. Jitter buffers are either adaptive or fixed. We used a fixed buffer length. Adaptive jitter buffers can vary their size as per the amount of traffic. The impact of jitter can be measured on a VoIP service by using a jitter buffer emulator that can deduce the number of packets that will be discarded [10].

6. Conclusions and future work

The packet loss, end-to-end delay and jitter measurements can aid in the correct design and configuration of traffic prioritization, as well as buffering parameters in the data network equipment. Our results have showed that sending a redundant copy on another path would improve quality of the overall system. This has been shown by improving the packet loss Ratio, end-to-end delay and jitter.

For the future, our interest will be focused on using an adaptive jitter buffer; Jitter buffers can adjust automatically with the delay in traffic; this permits the data to be retained for maximum time before it has to be discarded. The jitter buffer is sensitive to the recent minimum delays and is aware of the maximum permissible delay. This helps it to adjust to any changes in delay. An increase in jitter levels or the presence of a discard event is a trigger for adaptive jitter buffers to react. In the presence of a discard event, the jitter buffer size is increased. For jitter events that happen close to one another, an adaptive jitter is preferred; however, for jitter that occurs over a period of time, as in a LAN, increasing the size of the jitter buffer may lead to delay. Jitter modelling should be such that IP network emulation can be carried out with the help of data obtained by using a time series model.

- Also choosing more effective coders, example for coders:

1. Waveform Coders:
   – PCM ➔ Pulse code Modulation
   – DPCM ➔ Differential Pulse code Modulation
   – APCM ➔ Adaptive Differential Pulse code Modulation
2. Voice Coders:
   – *LPC* ➔ *Linear Predictive Coding*
3. Hybrid Coders:

–RELP ➔ Residual Excited Linear Prediction
–CELP ➔ Codebook Excited linear Prediction
–MPE ➔ Multipulse Excited Coding
–RPE ➔ Regular Pulse Excited Coding

Acknowledgements

This paper is sponsored by the US-Egypt science and technology join fund in cooperation with the United States National Science Foundation (NSF) and the Egyptian Ministry of Scientific Research under project number INF8-001-009

References

[1] Cisco: Voice over IP Quick Start Guide. Text Part Number: 78-4936-01 (1998).
[2] J.-C. Bolot, S. Fosse-Parisis, and D. Towsley. "Adaptive FEC-based error Control for Internet Telephony", In Proceedings of IEEE INFOCOM '99, Vol. 3, pp. 1453–1460 (1999).
[3] G. John, Apostolopoulos, “Reliable Video Communication Over Lossy Packet Networks Using Multiple State Encoding and Path Diversity”, in Proc.Visual Commun. And Image Processing, Jan (2001).
[4] Tommi Koistinen, “Protocol Overview: RTP and RTCP” Nokia Telecommunications.
[5] Multi-Stream Voice Over IP Using Packet Path Diversity: By Y.J. Liang, E.G. Steinbach, and B. Girod. Proceedings IEEE Fourth Workshop on Multimedia Signal Processing, pp. 555-560 (2001).
[6] Overview of Voice Over IP: by Princy Mehta and Sanjay Udani, University of Pennsylvania (2001).
[7] Uyless Black, "Voice Over IP ", Prentice Hall, 2nd Edition, Published November, pp. 145-164, ISBN 0130652040 (2001).
[8] Voice Over Internet Protocol (VoIP): By BUR Goode, Proceedings of The IEEE, Vol. 90 (9), pp. 1495-1517, September (2002).
[9] Tamer M. El-Sayed , Mohamed N. El-Derini , Mohamed S. Selim, Magdi A. Ahmed, "Improved Feedback Consolidation Algorithm for Multicast ABR Connections in ATM Networks", the proceeding of the ISCA 14th International Conference on Computer Applications in Industry and Engineering (CAINE-2001) , Las Vegas, Nevada, USA, November 27-29, pp. 120-124 (2001).
[10] Andy Hagans and Jimmy Atkinson, "Voice over IP News Organization", September 11, 2005 http://www.voipnow.org/2005/09/causes_of_jitte.html,
[11] Jeewan Jyoti, Samy El-Tawab, M. Nazih El-Derini, Emad Aboelela, and Hussein Aly: “Improving Quality of Service for Voice-Over-IP Using Routing Diversity", the 23rd Biennial Symposium on Communications, Kingston, Ontario, Canada (2006).

Received May 22, 2006
Accepted September 30, 2006