Performance Enhancement of the Playout Buffer of VoIP based Social Networking Applications.

1. Introduction:

Online social network is a novel technology over the Internet that has become a buzzword for the exchange of information and connectivity. It is basically a web-based service that provides ways for users across the globe to interact in multiple ways, such as file sharing, blogging, discussion groups, Real Time Communication (RTC), communities building among people sharing common interests. According to Goth (2008) these networks derived their names from their structure which is made up of individuals (or organizations) referred to as "nodes" that are tied (connected) by one or more specific types of interdependency to share contents or to communicate.

A Social Networking Site (SNS), as explained by Goth (2008), is a channel or a platform that integrates and provides multiple communication tools and technologies that aim at delivering RTC. The real fascination is of exchange of information instantly or with negligible latency. RTC aims at having peer-to-peer communication that can include:

- Telephony in the conventional sense
- Mobile and cellular telephone
- Two-way or multi-way amateur radio
- IM (instant messaging)
- VoIP (Voice over IP, also called Internet Telephony)
- IRC (Internet Relay Chat) or other chatting modes
- Live videoconference communication
- Live teleconference communication
- Robotic telepresence.

Out of these VoIP is becoming tremendously popular because of the benefits offered over Public Switched Telephone Network (PSTN). Sankaran et al. (2010) recognized Internet Telephony as the technology to deliver voice over Internet Protocol (IP). They defined it as a group of methodologies, communication protocols, and transmission technologies to deliver voice communications and multimedia sessions over Internet Protocol (IP), which is the base protocol for communication over the Internet.

SNSs have given a new meaning to the communication over Internet. A SNS has multiple groups and communities interacting with each other. There is variety of proprietary and open source VoIP based clients available for voice based communication over these sites which help the netizens to stay connected around the world through Internet for free or at a very nominal price. Real-time Transport Protocol (RTP) is the main protocol for voice transmission in modern VoIP networks.

Sankaran et al. (2010) observed that VoIP networks that are also called as "packet-switched network" (Datagram based), on contrary to Public Switch Telephone Network (PSTN) which is a circuit-switched network, takes a peer-to-peer communication path between two end stations, separately routed through the network. A VoIP communication over social network is gaining more and more popularity and is likely to pick up more acceleration in near future due to its multiple advantages, like huge cost cutting, resource sharing and many more.

To setup a PC-to-PC VoIP communication one needs to have the following (Wu, 2008):

- VoIP application software called as "softphone software".
- Sound card.
- Internet access.

The basic steps involved in conducting VoIP communication are:

- The Analog to Digital Converter (ADC) converts analog voice to digital signals (bits).
- Encoding and compression of bits in a proper format for transmission by "codec" (coder-decoder)
- Insertion of voice packets in data packets using a real time protocol (typically RTP over UDP over IP).
- A signaling protocol (ITU − T H.323) is used to call users.
- At receiving end playout buffer receives and holds the packets and reorder them.
- Codec decode and decompress the packets and extract data.
- Digital to Analog Convertor (DAC) to convert digital signals to analog voice signals.
- Sound card (or phone) to play the voice.

All the above steps must be performed in a real time fashion because we cannot wait for too long for a vocal answer during VoIP communication.

According to the steps involved in VoIP communication, we can have a base architecture for VoIP process that is depicted in Figure1, given below:

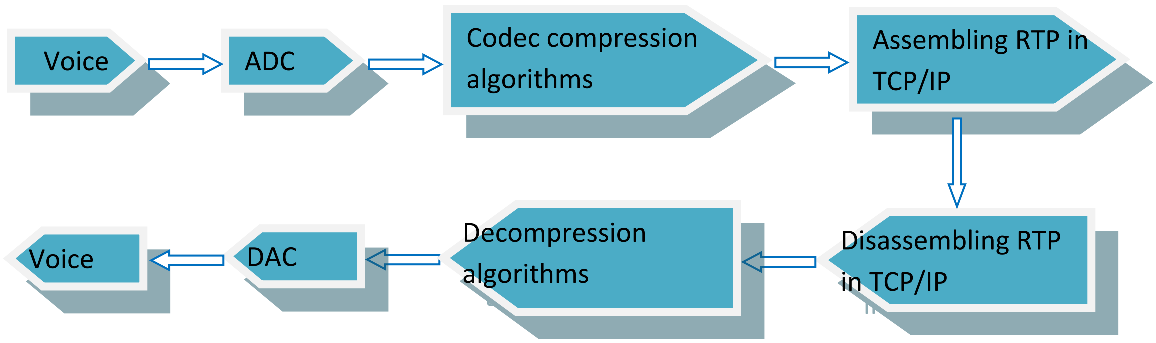


Figure 1. Base architecture of VoIP process

A circuit-switched network, such as the PSTN, provides a dedicated communication path between two end stations where as datagram-based packet-switched networks segment the original data into multiple packets, which are then separately routed through the network. By default, there is no dedicated path or bandwidth for datagram-based packet-switched networks.

The Internet was primarily developed for data networking and not voice; therefore, there are certain performance issues associated with VoIP implementation and the kind of quality provided by it. Malfait et al. (2008) pointed that due to low tolerance for latency (delay) in real-time communication, toll-quality voice transmission can be obtained on a packet-switched network only after the following issues have been resolved:

- **Jitter** – It is the variance in delay between packets. Excessive delay can cause for uneven, difficult-to-hear voice communication. End-to-end delay can be due to following reasons:
  - Delay incurred in encoding.
  - Due to packetization.
  - Due to the path distance from sender to the receiver (propagation, transmission, queuing).
  - Delay incurred in the Playout buffer.
  - Decoder

- **Packet Loss** – Packets may get lost either due to delay or due to congestion (dropped because of full queue) and transmission error in the wireless/wired interface.

- **Packet Sequence** – Due to the nature of voice communication, received packets need to be processed in the same order in which they were sent from the original source. The connectionless feature of IP network causes out-of-order packages.

- **Acoustic Echo** – Acoustic echo is the reflection of a sound signal. The power, or amplitude, of the acoustic echo and the amount of delay between the originating signal and the reflecting signal (the acoustic echo) determine whether the echo is detectible or bothersome to the persons who are talking.

As the load on social networks is unpredictable we have varying rate or number of packets arriving at receiving end during a VoIP communication. Due to this variation the packets may get lost. There are many factors that affect quality of voice and packet loss is one of them. As per Bolot (1993) packet loss may occur due to many reasons congestion of router and gateways, overloaded links, excessive collisions on a LAN, physical media errors, delay in packet transmission from sender to receiver, packet arriving too late at the receiver side, heavy loading (other than links), the variation in packet inter arrival time (jitter), loss of voice packets from sender to receiver, playout buffer overflow.

However, packet loss creates a real problem when the percentage of the lost packets exceeds a certain threshold (roughly 5% of the packets). Yeung et al.(2005) observed that packet losses occurring in large groups, called as “packet bursts”, severely effect packet loss rate. Thus, it is important to know both the percentage of lost packets, as well as whether these losses are grouped into packet bursts.

Playout buffer

Maximum packet loss occurs at senders’ or receivers’ buffer. Mazurczyk et al. (2010) discussed that whenever a router receives an audio stream for VoIP, it must compensate for the jitter (delay) that is encountered. The mechanism that handles this function is called as the “playout delay buffer” or simply “playout buffer”. The playout delay buffer must buffer these packets and then play them out in a steady stream to the Digital Signal Processors (DSPs) to be converted back to an analog audio stream. The playout delay buffer is also sometimes referred to as the “jitter buffer”.

Playout buffer algorithm

A playout buffer algorithm monitors the time-stamp and reception time of the i-th packet and adjusts the playout deadline. A good playout algorithm should be able of minimizing both buffering delay and late packet loss. Ramjee et al. (1994) asserted that these two conflicting goals have led to various playout algorithms. The playout buffer algorithms can be categorized into following types:

- Algorithms that perform continuous estimation of network delays and jitter to calculate playout deadlines. They are also called as “Reactive algorithms”.

- Algorithms that maintain a histogram of packet delays and choose the optimal playout delay from that histogram. Such algorithms are also called as "Histogram-based algorithms".
- Algorithms that monitor packet loss ratio or buffer occupancy and adjust the playout delay accordingly.
- Algorithms that aim in maximizing user satisfaction.

In conjunction with the above algorithms Ramjee et al. (1994) also described that the buffering techniques can be categorized into two types, i.e., "fixed" (uses fixed buffer time) and "adaptive" (adjust playout delay).

Adaptive technique and algorithms prove to be better than the fixed one, as they are capable of adjusting buffer dimension according to the network scenario. Such algorithms suite the best for our application's requirement, as the jitter buffer of a VoIP based social network application must be capable enough of resizing its buffer according to the condition of the network, so as to have least packet loss. Though we have many adaptive VoIP playout buffer dimensioning algorithms, their relevance for VoIP communication over social networks needs to be further explored.

2. Review of Literature:

VoIP implementations have many performance issues associated with them. Several factors (physical or software related) are there that affect the quality of voice. Packet loss is one such impairment that adversely affects voice quality. Mazurczyk et al (2010) analyzed experimental data results with respect to Real-time Transport Protocol (RTP) packet losses including physical losses and losses caused by jitter buffer (late packet drops and buffer overflows) and evaluated its feasibility for implementing RTP steganographic methods based on real VoIP traffic. The area of this research was just confined to steganographic methods only. Effect of packet loss on voice quality was also not taken into consideration.

Apart from performance issues, that are Quality of Service (QoS) and Quality of Experience (QoE), user satisfaction also accounts a lot for such time sensitive services. Narbutt & Murphy (2005) assessed the quality of VoIP transmission affected by playout buffer scheme. They justified that we cannot make choice of buffer algorithm and its parameters without considering its effect on user satisfaction. For this, a new method was proposed to evaluate playout buffer algorithm which provides a direct link to the perceived conversation speech quality by estimating user satisfaction from time varying transmission impairment. Though a new method was proposed by them for the evaluation of playout buffer but nothing was done for its improvement.

After analyzing the previous playout buffer algorithm techniques Narbutt & Murphy (2005) highlighted the fact that the standard playout buffer strategy uses an estimate (Exponentially Weighted Moving Average) of the mean and variance of network delay to set the playout deadline which is characterized by fixed, constant weighing factor. They replaced this constant weighing parameter by dynamic one, which was dynamically adjusted according to the observed delay variations. This algorithm helps in improving trade-off between buffering delay and late packet loss at the receiver end. The work has limitations that it could not provide the benefits of the suggested algorithm in voice quality improvement.

Tuning of end-to-end delay was done by Pinto & Christensen (1999) who developed a new adaptive "gap based" algorithm that could be tuned for both end-to-end delay and packet loss to satisfy a user desired tolerance. This research did not provide any kind of playout adjustment.

Boutremans & Boudec (2003) also developed a joint rate-error-playout buffer and Forward Error Correction (FEC) adjustment scheme for VoIP. This scheme provides better quality than the playout adjustments and FEC provided earlier, and is important because of a threshold effect when the end-to-end delay of interactive audio is around 150 ms. The perceived audio quality was presented as a function of both the end-to-end delay and the distortion of the voice signal. It uses a channel model for both loss and delay. They showed that it provides better quality than the adjustment schemes for playout buffer and FEC, that were previously published but in their work the user desired tolerance was not considered.

Sometimes these end-to-end delays stretch so much that they take the shape of "spikes", which are the result of long delays. Chang (2004) used an algorithm selector to select jitter buffer algorithms like linear regression and spike (long delays) detection algorithm. He also introduced a new linear regression algorithm with dynamic parameters to adapt to the fast changing wireless environment during VoIP. The real life implementation of this algorithm was not provided.

Wu (2008) analyzed the performance of system during VoIP in Ethernet LAN and WLAN scenario, as the number of VoIP clients and traffic distribution were changed. The simulation work done by him can help organizations, researchers and designers understand how well VoIP will perform on a local network prior to adopt it and design a network for its deployment. He did not show the impact of the changes in the number of VoIP clients on the quality of voice.

Several researchers have proposed refinements and proposals through their studies to reduce packet loss, adjust playout buffer dynamically according to the varying network load and parameters and voice quality enhancement; this motivated others to actually find out the real implementation of the principles provided by such researchers. Wu et al. (2009) made an empirical evaluation of playout buffer dimensioning in Skype, Google Talk and MSN Messenger, by using an objective Quality of Experience (QoE) metric and proved that none of them adjusted the buffer size for the network loss rate. However, the number of VoIP calls was fixed to 10 calls and packet loss rate was also not taken into account in this research.

Applications like Google Talk and Skype have strong support for VoIP over social network, which is the latest buzz word for 24x7 connectivity, but at the same time also suffers from poor voice quality. This opens a new genre for researches where lot of work can be done. A lot of research work has been done on refining playout buffer algorithms so as to improve upon packet loss and quality of VoIP communication but the actual implementation of these algorithms and the behaviour of jitter buffer for social network based VoIP applications is something where lot of research is to be done.

3. Motivation/Justification and Relevance:

Real-time voice traffic suffers through many different challenges that severely effect voice quality. Out of many, delay and packet loss are the ones which are much talked about. These challenges can be taken care of either at sender end or at receiver end. Receiver based methods are better than the sender based methods because they do not require changes to an encoding scheme or network infrastructure.

Voice packets are buffered and reordered at receivers' playout buffer that introduce inter-packet gap to remove jitter before sending it to receivers' codec. Wu (2008) stated that Playout buffer holds these packets by an amount of "buffer time" that compensates for the network delay variance without excessively delaying the playout. Packets arriving after the buffer time are considered to be lost. The buffer may stretch its buffer time to compensate for such losses which results in high quality at the cost of low conversational interactivity. Similarly a sudden burst of RTP streams may result in buffer overflow resulting in packet loss and low quality. Thus a good playout scheme should be capable enough to trade-off playout delay and packet loss rates in order to achieve best voice quality. Since a social network also suffers from this ever fluctuating load, there is a need to travel around to substantiate how well the playout buffers of VoIP based social networking applications are being adjusted.

Many researches claim at providing buffer algorithms that ultimately help in reducing packet loss rate but the actual implementation of such algorithms by real VoIP based applications needs to be investigated. Wu et al (2009) found that the three most popular VoIP applications do not consider the network loss rate in their buffer dimensioning algorithms. Since these applications also have strong support for social network, which is the future of these VoIP applications, we come up with a new scope of research for the testability of these applications in social network scenario.

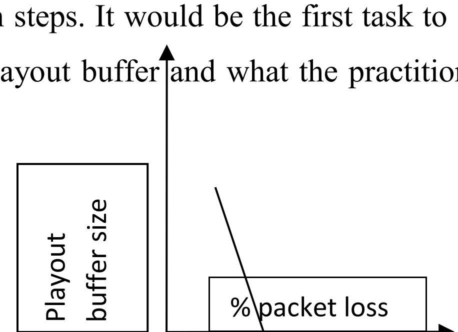
4. Objectives:

Since many researches are going on for fine tuning of the playout buffer, the research aims at proposing an effective playout buffer algorithm. The objectives of the proposed research are:

i. To analyze and evaluate the behavior of playout buffer as the number of VoIP calls change over VoIP based social network applications.
ii. Analyze packet loss rate and its correlation with playout buffer size.
iii. Evaluation of the quality of voice received at the receiver's end (through subjective or objective test score).
iv. Refining the existing playout buffer algorithm of the VoIP based social networking application or proposing a new algorithm so that we can have least packet loss, best voice quality and user interactivity, while changing the number of VoIP calls during communication over a social network.
v. Comparative result analysis of existing and refined/new algorithm(s) to verify proposed improvements.

5. Methodology:

The proposed research work would be implemented in certain steps. It would be the first task to find out whether a gap between what the researchers say about playout buffer and what the practitioners actually face exists or not.



In order to meet the desired research objectives there is a need to formulate certain strategies that would help us in efficiently meeting our goal. The complete plan of our research work can be summarized in following sequential steps:

- Intensive study of different research papers, journals and documents showing details of ongoing research in the field of playout buffer dimensioning.
- Identifying different parameters and other factors which influence the quality of VoIP calls.
- The proposed work will use an experimental network of components meeting our requirement.
- VoIP based social network application will be selected.
- To describe selection criteria for number of VoIP calls and their duration.
- A VoIP session would be initiated with single VoIP call and number of calls will be increased progressively.
- The actual playout buffer size evaluation (either by using some tool or some available method).
- An effective network traffic analysis tool will be used to evaluate the rate of packet loss.
- Evaluation of the effect of number of calls on application's playout buffer size and packet loss rate.
- Assessment of voice quality either through subjective or objective testing.
- Identifying limitation(s) and addressing practical problems of VoIP calls performance.
- Refinement of existing algorithm or proposing a new one.
- Implementation of new/refined algorithm.
- Reanalysis of playout buffer through new/refined algorithm for same set of parameters.
- Reassessment of voice quality.
- Comparative result analysis of previous and refined/new algorithm will verify the objective(s) achieved.
- Conclusion drawn on the basis of Comparative result analysis of previous and refined/new algorithm.
- Thesis will be prepared to record research activities and results of different phases in graphical or tabular format.

Implementation of the above methodology will result in an improved quality playout buffer algorithm for VoIP based social networking application.

6. Time Plan:

The research work is planned to be completed in following phases and in the approximated time span of 18 months. Every phase has roughly been divided in tenure of justified period in following way:

Table 1: Time Plan of proposed research work.

Months → **Activity ↓**	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**
Literature survey	**--------- (2months)**								
Performance evaluation and quality analysis of playout buffer		**----------- (2-3months)**							
Refinement			**------------- (3-4months)**						
Implementation				**--------- (1-2months)**					
Reanalysis and evaluation					**------------- (1-2months)**				
Comparison and results						**------- (1-2months)**			
Thesis writing							**---------------- (3-4months)**		

7. Tools and Techniques:

The research work will require certain tools to carry out the work successfully. Following is the list of few tools that may be required during the research depending on the requirement.

- VoIP based social network application (proprietary or open source) – e.g. Google talk or Skype.
- Network simulator to create artificial network traffic – e.g. NS2.
- Network traffic analyzing tools (proprietary or open source) – e.g. RADCOM's AudioPro.
- Comparison tools – e.g. MATLAB, SPSS, MS-Excel.
- Graphical tools – e.g. Gliffy, Lucid Chart.

The usage of above mentioned tools will purely depend on their inevitability in the course of research.

This research will also open new scope to analyze and evaluate the other time sensitive and quality bound applications over the social networks.

8. References :

[i] Bala, P., Krishna Prasad, G.Murali, G.Gurukesava Das, K.Siva Krishna Rao (2009), "Congestion Controlling for Streaming Media Through Buffer Management and Jitter Control", *International Journal of Computer Science and Network Security (IJCSNS)*, VOL.9 No.2, pp. 362-371

[ii] Berto, R.A. (2002), http://tldp.org/HOWTO/pdf/VoIP-HOWTO.pdf, vol. 1.7, August 7, 2002

[iii] Bolot, J.C. (1993) , "End-to-end Packet Delay and Loss Behaviour in the Internet", *Conference proceedings on Communications architectures, protocols and application*, *ACM* SIGCOMM 1993, NewYork, vol.23, pp.289-298

[iv] Boutremans, C., J.Y. Le Boudec (2003), "Adaptive Joint Playout Buffer and FEC Adjustement for Internet Telephony", *22nd Annual Joint Conference of the IEEE Computer and Communications IEEE Societies*, 30th March-3rd April 2003, vol.1, pp. 652 - 662

[v] Chang, Z. (2004), " VoIP Jitter Buffer Algorithm Improvement for Wireless Environment", *Master thesis project, DELFT University of Technology*, http://ce.et.tudelft.nl/publicationfiles/1279_536_myreport8.3.pdf

[vi] Goth, G. (2008), "Are Social Networking Sites Growing Up?", *IEEE Distributed Systems Online*, vol. 9, pp. 3

[vii] Malfait, L., Phil Gray, Martin J. Reed (2008) "Objective listening quality assessment of speech communication systems introducing continuously varying delay (Time-Warping): A Time alignment issue", *IEEE International Conference on Acoustics, Speech and Signal Processing*, ICASSP 31st March – 4th April 2008, Las Vegas, pp.4213-4216

[viii] Mazurczyk, W., Krzysztof Cabaj and Krzysztof Szczypiorski (2010), "What are suspicious VoIP delays?", *SpringerLink's book on Multimedia tools and Applications*, April 2010, http://www.springerlink.com/content/yg34t43l893q1253/

[ix] Narbutt, M., A. Kelly, P. Perry, L. Murphy (2005), "Adaptive VoIP playout scheduling: Assessing user satisfaction", *Internet Computing, IEEE Computer Society*, July-August 2005, vol.9, Issue:4, page(s): 28-34

[x] Narbutt, M., M L. Murphy (2004), "A new VoIP adaptive playout algorithm", *Telecommunications Quality of Services: The Business of Success, IEEE*, 2-3 March 2004, page(s):99-103

[xi] Pinto, J. and Kenneth J. Christensen (1999), "An algorithm for playout of packet voice based on adaptive adjustment of talkspurt silence periods", *Conference on Local Computer Networks*, Lowell, MA, 18-20th October 1999, Page(s): 224 - 231

[xii] Ramjee, R., J. Kurose, D. Towsley, and H. Schulzrinne (1994), "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks," 13th *Proceedings of IEEE INFOCOM, Networking for Global communications*, 12th – 16th June 1994, Toronto, Canada, vol. 2, pp. 680 - 688

[xiii] Sankaran, N.P., Rao Shruti Raghunath, Raut Mayuresh Sanjay (2010), "Replacing Conventional Telephones by VoIP: Their Major Constraints & Solutions", *International Journal of Computer Applications* 2010, Vol. 1 – No. 12, pp. 0975 – 0987

[xiv] Wu Di (2008), "Performance Studies of VoIP over Ethernet LANs", http://aut.researchgateway.ac.nz/bitstream/10292/677/4/DiW.pdf

[xv] Wu, Chen-Chi, Kuan-Ta chen, Yu-Chun Chang, Chin-Laung Lei (2009), "An Empirical Evaluation Of VOIP Playout Buffer Dimensioning In Skype, Google Talk, And MSN Messenger", *Proceedings of the 18th International workshop on Network and operating systems support for digital audio and video*, Taiwan Roc Proceeding NOSSDAV '09

[xvi] Yeung,K. Z., K.L. Li and V.O.K (2005), "On Bursty Packet Loss Model For TCP Performance Analysis", *Workshop on High Performance Switching and Routing*, HPSR 12-14th May 2005, Page(s): 292 - 296

Abstract:

Social networks have given a new meaning to the world of Internet. They are offering multiple tools and applications for people to make their conversation more interactive, and VoIP is one of those. Playout buffer of VoIP applications play an important role in maintaining a balance between jitter and Quality of Service (QoS) issues. Many Social Network applications are offering VoIP facility with them but are unable to cope up with the impairments associated with it. This paper offers an analysis

of how playout buffer of such applications can be fine tuned to get better user interactivity and voice quality.

1. Introduction

Voice over Internet Protocol (VoIP) helps in having voice communication over IP. Since Internet was primarily developed for the transfer of data and not for voice there are a few factors that sternly affect the quality of voice. VoIP implementation suffers from impairments like Delay, Jitter, Spiky delay, Packet loss, Acoustic echo and Packet sequence. Unlike circuit switched Public Switched Telephone Network (PSTN), VoIP is packet switched network based. Packets reach the destination at different time and in different order, therefore there is a need of some mechanism to hold the packets, reorder them and play them as a single stream. This is achieved by a buffer called as "playout buffer" or "jitter buffer". The time at which the packets are actually being played is called as "playout time". The most decisive role of this buffer is to maintain a trade-off between voice quality and user interactivity. A large sized playout buffer would definitely offer a good quality of voice but will compromise with user interactivity and vice versa, therefore a good playout buffer must keep a balance between voice quality and user interactivity. There are certain algorithms that help Playout buffer in achieving its objectives.

The playout buffer algorithms can be categorized as "fixed" and "adaptive". Fixed playout buffer algorithms are for fixed size buffer that doest adjust its size, while adaptive buffer algorithms help the playout buffer in adaptively adjust its size according to varying network conditions.

Buffer size adjustment is an important optimization problem that has had gathered much attention of researchers over the past few years. Many researches have been going on in improving upon buffer size adjustment. Theoretically different types of playout buffer algorithms have been proposed that help in fine-tuning the playout buffer but the real life applications' based implementation of these algorithms needs to be explored. There is a need to identify the best available playout algorithm(s) from social network perspective.

2.