TD(07)009

8th COST290 Management Committee and Technical Meeting, Malaga, Spain, February 15-16, 2007

Coding scheme impact on the IP-QoS network utilization and voice quality

Eduard-Cristian Popovici, Lecturer, Politehnica University of Bucharest,
eduard@elcom.pub.ro
Tatiana Radulescu, Professor, Politehnica University of Bucharest,
tatiana.radulescu@comm.pub.ro

Abstract

This article presents the impact of speech coding scheme on network utilization and voice quality for voice-over-IP (VoIP) applications over QoS-aware networks. We mention the coding scheme evaluated and give details of our network and application simulation framework. We use for simultation IT Guru OPNET's tool. Based on simulation results, different aspects of the coding schemes impact on network load and voice quality are revealed and some ways to exploit them for a better network utilization and QoS improvement are discussed.

Keywords

IP Network Utilization, QoS, QoS–IP Neyworks, Voice Coding Scheme, End-to-End Delay

Working Group 1

1. Introduction

Our work aim to reveal criteria for decisions concerning the speech coding scheme used in voice-over-IP (VoIP) applications over QoS-aware networks. We think that such criteria should be based on evaluations of impact of coding schemes on network utilization and voice quality impairments. For voice quality impairments evaluations we choosen the E-Model specified by ITU-T [2] [3] for objective evaluations of voice quality degradations, as an alternative to the subjective MOS score. Our performance evaluations concerned six speech coding schemes specified by ITU-T and ETSI (GSM).

The simulation framework used in our evaluations is IT Guru tool from OPNET. We given details of the simulated network and services. Based on simulation results, we analyses different aspects of coding schemes impact on network load and voice quality. We found ways to exploit this aspects so that the network load and voice quality can benefit.

In section 2 we give details about quality assesment objectives and metrics, including the E-Model and the R factor. In section 3 we anticipate some expected effects of coding schemes on network load and QoS. Section 4 is concerned with details about simulation network and performance evaluation. In section 5 we conclude with a discussion of criteria for decisions concerning the speech coding scheme used in VoIP applications over QoS-aware networks.

2. User perceived voice quality (QoS) assesment

2.1. User perceived QoS for voice-over-IP (VoIP) communications

Quality of Service (QoS) is collective effect of service performance which determine the degree of satisfaction of a user of the service [1] [7] [8]. Evaluation of voice quality is subjective because it’s a measure of the intelligibility and clarity of speech as perceived by the listener.

In VoIP environments, two elements emerge as the primary factors affecting voice inteligibility and conversation quality: clarity and delay. Clarity generically refers to a voice signal’s fidelity, clearness, lack of distortion, and intelligibility. In a VoIP environment, clarity problems are caused by delay, packet loss and uncontrolled jitter. Clarity is also significantly impacted by the codecs used on the voice channel.

End-to-end delay is the time it takes a voice signal to travel from talker to listener. When end-to-end delay reaches about 150 milliseconds, participants in a telephone conversation begin to notice its effects. Between 250 to 400 milliseconds, normal conversation is difficult. End-to-end delay above 400 milliseconds can make normal conversations impossible.

2.2. Metrics for VoIP service quality evaluation

Traditionally, the clarity of a voice signal or voice channel has been measured subjectively according to ITU-T Recommendation P.800 resulting in a mean opinion score (MOS). MOS values can range from 1 to 5 with 5 being the best possible score.

ITU defined in Recommendation G.107 an E-Model for estimating the voice quality of IP telephony. The output of the E-Model is a scalar Transmission Rating Factor called the R-Value (or simply “R”). It combines individual impairments that result from both the signal's properties and the characteristics of the transport medium into a single R rating, that ranges from 0 to 100. The R-rating is a linear combination of the perceived impairments converted to appropriate psycho-acoustic scales.

With appropriate modeling, R can be correlated to MOS with remarkable accuracy [2]. The relationship between R-ratings, speech quality and MOS values is shown in Table 1.

Table 1. Categories of voice transmission quality

R-value range	MOS range	Quality	User satisfaction	Voice QoS example (end-to-end delay)
90 - 100	4.34 - 5.00	Best	Very satisfied	< 100 ms
80 - 90	4.03 - 4.34	High	Satisfied	< 100 ms
70 - 80	3.60 - 4.03	Medium	Some users dissatisfied	< 150 ms
60 - 70	3.10 - 3.60	Low	Many users dissatisfied	< 400 ms
< 60	< 3.1	Poor	Nearly all uses dissatisfied	>400 ms

There are many factors that can impair voice quality [4], and each of these can be associated with an R-Degradation. The coding technique has a profound effect on the response to channel impairments. The overall R-Value is:

$$R = Ro - Is - Id - Ie - A$$

where:

Ro = basic Signal-to-noise ratio (SNR), in absence of any impairments
Is = impairments occurred simultaneously (connection loudness, quantisation distortion)
Id = impairment due to delay of voice signal
Ie = impairment due to compression techniques (loss due to very low bit-rate encoding) and loss
A = user tolerance of some degradation due to special situations (mobility, hard-to-reach locations)

3. Expected effects of coding scheme on QoS and network utilization

3.1. Voice compression standards used in evaluations

We used in our evaluations some of G.7xx suite of ITU-T standards for audio compression and de-compression, primarily used in telephony. G.711 represents logarithmic pulse-code modulation (PCM) with 8 bits samples for signals of voice frequencies, sampled at the rate of 8000 samples/second, on an 64 kbps channel. G.723.1 is the ITU-T standard for speech coding in PSTN videophones utilizing Algebraic Code-Excited Linear Prediction (ACELP) coding at 5.3 kbit/s and Multipulse Maximum Likelihood Quantization (MP-MLQ) at 6.3 kbit/s. G.726 is the Recommendation for speech coding at 40, 32, 24, and 16 kbit/s (variable bit rates) utilizing Adaptive Differential Pulse Code Modulation (ADPCM) (we use only 32 kbps as speed coding). G.278 is the ITU-T Recommendation for speech coding at 16 kbit/s utilizing Low-Delay Code-Excited Linear Prediction Coding (LD-CELP). G.729 is the ITU-T Recommendation for speech coding at 8 kbit/s utilizing Conjugate Structure Algebraic Code-Excited Linear Prediction Coding (CS-ACELP). Another standard used in our evaluations is ETSI GSM 06.10 Full-rate, the coding standard for first generation digital European Global System for Mobile communications (GSM) cellular system, utilizing Regular Pulse Excitation Long Term Prediction (RPE-LTP) coding at a net bit rate of 13 kbit/s.

Table 2 presents the information of interest for our evaluation, concerning all coding standards we mentioned [2]. We can observe higher degradation of voice quality for lower bit rates coding schemes.

Table 2. Coding scheme contribution to end-to-end delay and perceived QoS

Standard	Voice coding type	Coding rate (kbps)	Frame Size (ms)	Compression Impairment, Ie	Rating R
G.711	PCM	64	0,125	0	94,3
G.726	ADPCM	16		**50**	44,3
		24		**25**	69,3
		32		7	87,3
		40	0,125	2	92,3
G.728	LD-CELP	16	0,625	7	87,3
G.729	CS-ACELP	8	10	**10**	84,3
G.723.1	ACELP	5,3	30	**19**	75,3
	MP-MLQ	6,3		**15**	79,3
GSM-FR	RPE-LPT	13	20	**20**	74,3

3.2. QoS (end-to-end delay) for VoIP applications

One of the most important design considerations in implementing voice is minimizing one-way, end-to-end delay. Voice traffic is real-time traffic; if there is too long a delay in voice packet delivery, speech will be unrecognizable. Delay is inherent in voice-networking and is caused by a number of different factors. The International Telecommunication Union (ITU) considers network delay for voice applications in Recommendation G.114 (see Table 1).

End-to-end delay is the time it takes a voice signal to travel from talker to listener. This voice signal delay is the additive result of VoIP/IP network processing and packet transport. Delay affects the quality of a conversation without affecting the actual sound of the voice signal – delay does not introduce noise or distortion into the voice channel.

Generally, applying QoS does not affect fixed-network delay because fixed-network delay is a property of the medium. Upgrading to higher-speed media such as 10 Gigabit Ethernet and newer network hardware with lower encoding and decoding delays, depending on application, may result in lower fixed-network delay.

Packetization delay—Amount of time that it takes to segment, sample, and encode signals, process data, and turn the data into packets. Serialization delay—Amount of time that it takes to place the bits of a packet, encapsulated in a frame, onto the physical media. Propagation delay—Amount of time it takes to transmit the bits of a frame across the physical wire. Processing delay—Amount of time it takes for a network device to take the frame from an input interface, place it into a receive queue, and place it into the output queue of the output interface.

Table 3 contains estimations of propagation delay introduced by different transmission or processing systems [5]. We can observe critical values of delays for sattelite and mobile systems, thus imposing very high constraints on other end-to-end delay components, like coding delays.

Table 3. Estimated contribution to end-to-end delay for different systems

Transmission or processing system	Contribution to end-to-end delay
Cable systems	4-6 μs/km
Submarine optical fibre: - transmit terminal - receive terminal	13 ms 10 ms
Satellite system: - 400 km altitude - 14 000 km altitude - 36 000 km altitude	12 ms **110 ms** **260 ms**
PLMS (Public Land Mobile System)	**80-110 ms**

Variable-network delay—The congestion affect the overall latency of a packet in transit from source to destination. Applying QoS mechanisms does affect the variable-network delay. Queuing delay, the amount of time a packet resides in the output queue of an interface, has the main contribution to the variable-network delay.

Fixed-network delay is introduced by network when the electrical and optical signals travel the media en route to the receiver (see Table 4). For high networks which include many transit nodes for inter-region interconnections, network delay can be very large, so is necessary to use coding scheme with lower coding/decoding delay.

Table 4. Estimated QoS parameters (including network delay) for different global routes

	Network Delay	Network Packet Loss	Data Delivery Rate	Network Jitter
INTRA-REGION				
North America	55 ms	0.30%	99.90%	2 ms
Europe	45 ms	0.30%	99.90%	2 ms
Asia	105 ms	0.30%	99.90%	2 ms
South Pacific	70 ms	0.30%	99.90%	2 ms
INTER-REGION				
Europe to North America	95 ms	0.30%	99.90%	2 ms
Japan to North America	130 ms	0.30%	99.90%	2 ms
Singapore to North America	300 ms	0.30%	99.90%	2 ms
India to North America	350 ms	0.70%	99.70%	2 ms
South Pacific to North America	210 ms	0.30%	99.90%	2 ms
Latin America to North America	135 ms	0.70%	99.30%	2 ms

The voice coding schemes mentioned before could introduce high delays, from packetization and look-ahead processing. In ITU-T G.108 Recommendation [3], the minimum delay attributable to codec-related processing in IP-based systems with multiple frames per packet is calculated as:

$$(N + 1) \times \text{Frame size} + \text{Look-ahead delay}$$

where N is the number of frames in each packet, Frame size could be obtained from Table 2 and Look-ahead delay depends on coding scheme (5 ms for G.729 and 7.5 ms for G.723.1)

In our evaluations we used 1 to 3 frames per packet for G.729. The corresponding contribution of codec-related processing to end-to-end delay can reach 65 ms. For the same packetization interval, 30 ms, G.723.1 contribution of codec-related processing to end-to-end delay is 67.5 ms. One important effect of high compression delays is the reduction of the communication quality (Rating). It imposes high requirements on network delays, especially for satellite and mobile communications, as explained before. Therefore, when choosing a coding scheme, a trade-off should be made between the coding scheme contribution to the end-to-end delay (which affects the perceived voice quality) and the network load corresponding. To obtain relevant data for such a compromise we realize evaluations of how the traffic load depends on the coding scheme in use, to determine which scheme better use the network resources. We take into account the expected influence of the packet length to establish a more efficient trade-off voice quality-network load.

4. Simulation models and evaluation results

4.1. Network model description

The IT Guru OPNET's tool is a very powerful network simulator which provides a Virtual Network Environment that models the behavior of an entire network, including its routers, switches, protocols, servers, workstations, different kinds of links and individual applications [6].

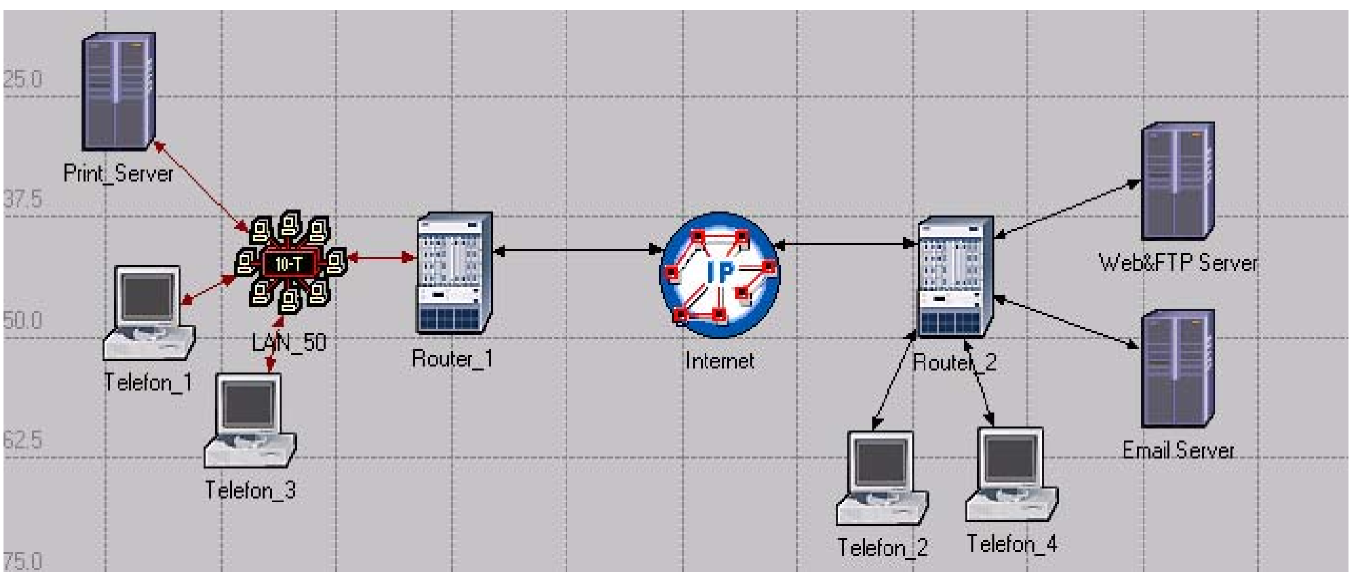


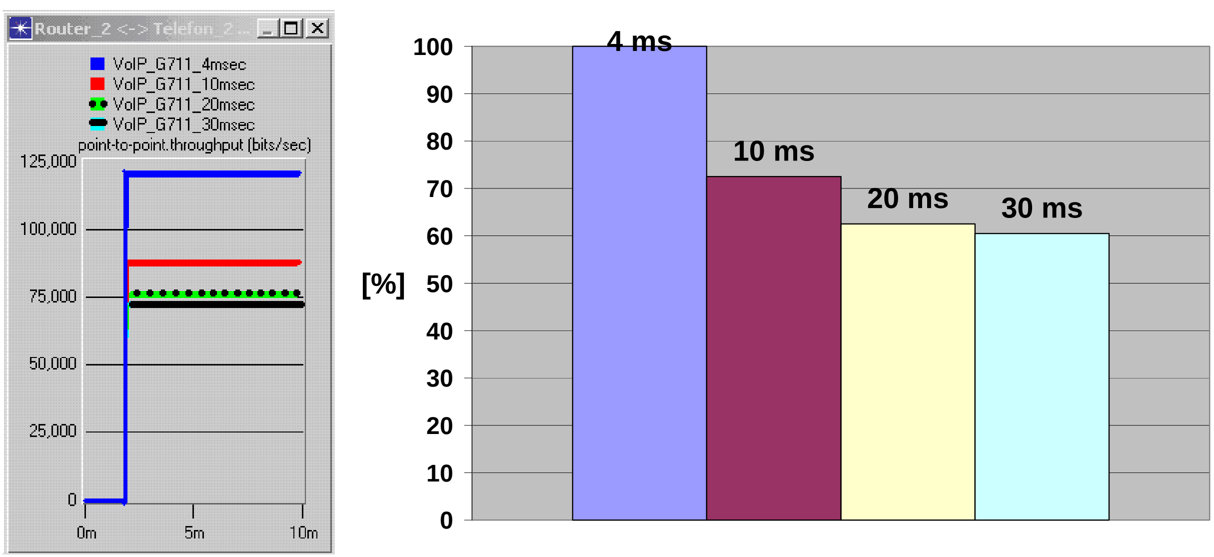
Figure 1. Network model in IT Guru OPNET

Our network model (see Figure 1) consists of two routers (*Router_1* and *Router_2*) connected via IP backbone (*Internet*), an 10BaseT access LAN which supports up to 50 workstation (*LAN_50*) and two types of traffic sources: data and voice. The data traffic include traffic from ftp, http, Email and Print applications. VoIP traffic is generated as two telephonic conversations: one between *Telefon_1* and *Telefon_2*, and another between *Telefon_3* and *Telefon_4*.

We configured the network model to assure the *Best Effort* type of service for data traffic which has no real time constraints, and an *Interactive Voice* type of service for VoIP traffic. The *Interactive Voice* profile includes Integrated Service as QoS mechanism. All network nodes involved in voice traffic delivery are configurated to support Integrated Services mechanism, by means of ***Classification, Queuing and Scheduling*** (CQS) and resource reservation by means of RSVP. The classification criteria is ToS (Type of Service). For voice we select ToS =6 which correspond at Interactive voice. The queueing scheme used in our evaluations is WFQ (Weighted Fair Queuing).

4.2. Simulation results and interpretation

In the beginning, we evaluated the impact of packetization interval on network utilization, throughput and QoS using end-to-end delay. The throughput was evaluated in two kinds of network segments: the 10BaseT Ethernet links and the point to point (*ppp*) links. We started with the G.711 voice coding scheme, which was presumably the worst of all. Figure 2 shows evaluation results for this coding scheme over *ppp* links. The results obtained emphasize an improvement of network utilization, because the throughput decreases for higher packetization interval (10 to 30 ms), and also little variations between throughputs for higher packetization interval.



(a) OPNET evaluation results **(b) Relative network utilization (reference value – 4ms packetization interval)**

Figure 2. Network utilization for G.711 coding scheme and different packetization intervals

Next, we tried to evaluate how the voice traffic generated in 10BaseT and point-to-point links, depends on the coding scheme and the packetization interval. We chose as packetization interval: 10, 20 and also 30 msec. Therefore, we repeated such performance evaluations for all coding schemes and packet sizes. There is a restriction related to packetization interval for GSM (20 msec) and G.723.1 (30 msec). Results for 10BaseT links are shown in Figure 3 and for *ppp* links in Figure 4.

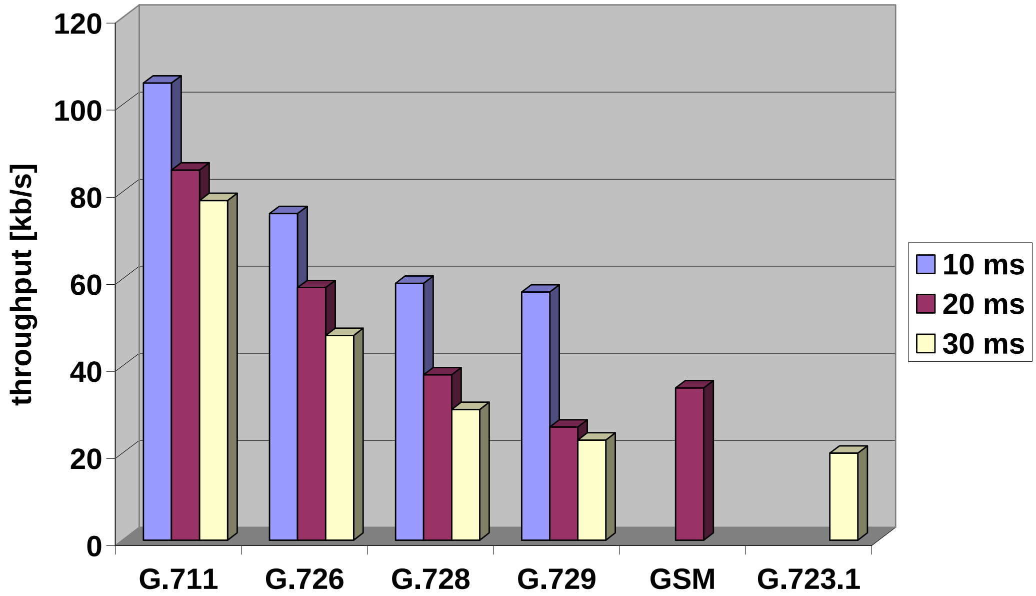


Figure 3. Network load (10BaseT links) for different coding scheme and packet sizes

As expected, the used voice encoders for high and middle bit rates, such as G.711 and G.726 involve high network loads. On the other hand, these schemes have little gain from increased packetization interval. The best performance we obtained for G.723.1 and G.729 coding schemes of ITU-T.

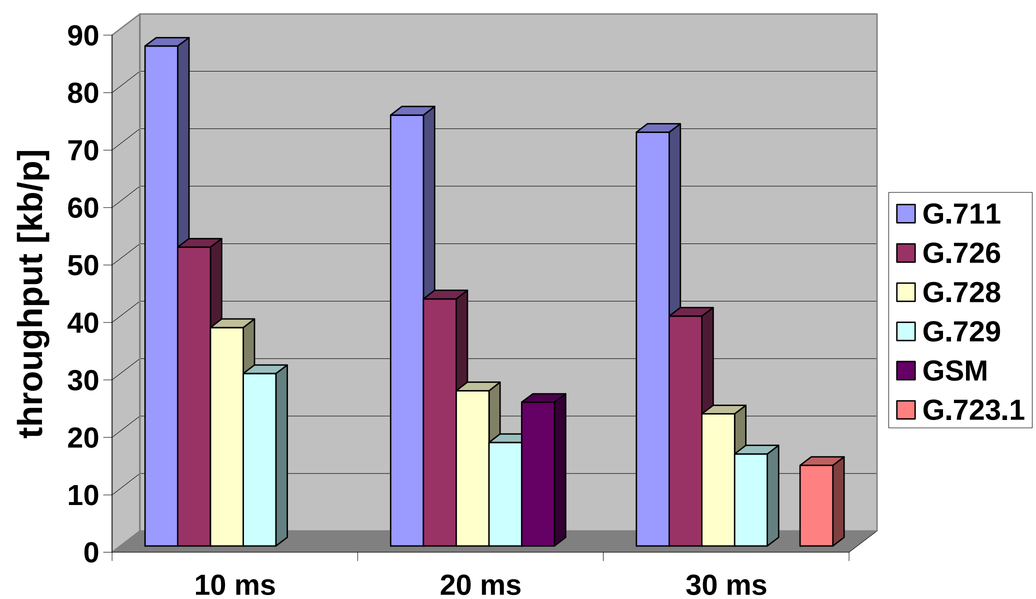


Figure 4. Network load (point-to-point links) for different coding scheme and packet sizes

Low bit rates coding schemes can lead to important decrease in network load utilization compared to G.711. This network load improvement we encountered at packetization interval bigger than 10 ms. Although, there are little differences in performance between 20 ms and 30 ms. As we already explained, 10 ms difference in end-to-end delay can have an important impact on overall voice quality, especially for sattelite and mobile systems. Therefore, a good trade-off between efficient network utilization and high voice quality can be obtained using a 20 ms packetization interval. Among the evaluated encoders, G.729 has the best performances at this packet size value.

5. Conclusion

In this work, we analyzed the impact of voice coding scheme, used in VoIP applications, on network bandwidth utilization and perceived quality of voice. We based our analysis on simulations results obtained with a network simulation framework, IT Guru from OPNET. The tool was configured to offer a QoS-aware framework for VoIP applications simultaneously with data applications like ftp, and http.

Our main interest in this evaluation was to find how much the voice packetization interval influence the network load (the used throughput), for each coding scheme. Therefore, we used in our simulations a board range of voice encoders: G.711, G.723.1, G.726, G.278 and G.729 from ITU-T and GSM from ETSI. We started with the worst of all from the point of view of network utilization, G.711 coding scheme, as reference. We repeated performance evaluations for all coding schemes and packet sizes from 4 ms to 30 ms. Packetization interval is restricted by standard at 20 ms for ETSI GSM and at 30 ms for ITU-T G.723.1.

Evaluation results shown not only high network loads for voice encoders at high bit rates, such as G.711 and G.726, but little gain from increased packetization interval too. The best performance, a network load five times lower than the reference one, we obtained for ITU-T G.723.1 and G.729 voice encoders. Such improvement of network bandwidth utilization load we encountered at packetization interval bigger than 10 ms.

In search for a good trade-off trade-off between efficient network utilization and high voice quality we observed little differences in performance, for the same coding scheme, when we increased the packetization interval from 20 to 30 ms. Thus, we concluded that from the evaluated coding schemes ITU-T G.729 coding scheme is the most appropriate compromise.

As future developments, we planned to study effects of different coding schemes on queuing delays, which could affect the overall voice quality especially in near congestion conditions. Another step foreword will be to evaluate the use of different IP-QoS mechanisms for congestion control.

References

[1] Grenville Armitage: "Quality of Service in IP Networks: Foundations for a Multi-Service Internet", Macmillan Technical Publishing, April 2000

[2] ITU-T Recommendation G.107: "The E Model, a computation model for use in transmission planning", International Telecommunication Union, 2005

[3] ITU-T Recommendation G.108: "Application of the E-model: A planning guide", International Telecommunication Union, 2001

[4] ITU-T Recommendation G.113: Transmission impairments due to speech processing, International Telecommunication Union, 2001

[5] ITU-T Recommendation G.114: One-way transmission time, International Telecommunication Union, 2003

[6] OPNET IT Guru Homepage, http://www.opnet.com/services/university/itguru_academic_edition.html

[7] Srinivas Vegesna: "IP Quality of Service", Cisco Press, 2003

[8] Tim Szigeti, Christina Hattingh: "End-to-End QoS Network Design: QoS in LANs, WANs and VPNs", Cisco Press, 2005