Impact of the Background Traffic on Speech Quality in VoIP

Peter Počta[1], Peter Kortiš[1], Peter Palúch[2], Martin Vaculík[1]

[1]*Dept. of Telecommunications, FEE, University of Žilina, Univerzitná 1, SK-01026, Slovakia*

{pocta, kortis, vaculik}@fel.uniza.sk

[2]*Dept. of InfoCom Networks, FMSI, University of Žilina, Univerzitná 1, SK-01026, Slovakia*

{paluch}@fri.uniza.sk

Abstract

This paper describes measurements of an impact of background traffic on speech quality in an environment of *IP* networks. The simulated background traffic consists of three types of current traffics in telecommunication networks such as data transfer service, multimedia streaming service and Web service. The background traffic was generated by means of the accomplished *D -ITG* traffic generator. The impact of these types of traffic and traffic load on speech quality using the test sequence and speech sequences is the aim of this paper. The assessment of speech quality is carried out by means of the accomplished *PESQ* algorithm. The proposal of a new method of improved detection of the critical conditions in telecommunication networks from the speech quality point of view is presented in this paper. Conclusion implies the next application of the method of improved detection of the critical conditions for the purpose of algorithms for link adaptation from the speech quality point of view in an environment of *IP* networks. The primary goal of these algorithms is improving speech quality in the *VoIP* connections, which are established in the competent link.

Keywords

Perceptual Evaluation of Speech Quality (*PESQ*), test sequence, speech sequence, intrusive measurement, Voice over Internet Protocol (*VoIP*), Voice Transmission Quality of Service (*VTQoS*), background traffic

1 Introduction

Voice over Internet Protocol (*VoIP*), the transmission of packetized voice over *IP* networks, has gained much attention in recent years. It is expected to carry more and more voice traffic for its cost-effective service. However, the current Internet, which was originally designed for data communications, provides *best-effort* service only, posing several technical challenges for real time *VoIP* applications. Speech quality is impaired by packet loss, delay and jitter. Assessment of perceived speech quality in the *IP* networks becomes an imperative task to manufacturers and well as service providers.

Speech quality is judged by human listeners and hence it is inherently subjective. The Mean Opinion Score (*MoS*) test, defined by *ITU-T P.800* [1] is widely accepted as a norm for speech quality assessment. However, such subjective test is expensive and time-consuming. It is impractical for frequent testing such as routine network monitoring.

Objective test methods have been developed in recent years. They can be classified into two categories: signal-based methods and parameter-based methods. Signal based methods use two signals as the input to the measurements, namely, a reference signal and the degraded signal, which is the output of the system under test. They identify the audible distortions based on the perceptual domain representation of two signals incorporating human auditory models. These methods include Perceptual Speech Quality Measure (*PSQM*), Measuring Normalizing System (*MNB*), Perceptual Analysis Measurement System (*PAMS*), and Perceptual Evaluation of Speech Quality (*PESQ*). Among them, *PSQM* and *PESQ* [2] were standardized by *ITU-T* as *P.861* and *P.862* respectively. Parameter-based methods predict the speech quality through a computation model instead of using real measurement. A typical model is the *E-model* as defined by *ITU-T G.107*. The *E-model* includes a set of parameters characterizing the end-to-end voice transmission as its input, and the output can be transformed into a *MoS* scale for prediction. The algorithm *PSQM* is based on comparison of the power spectrum of the corresponding sections of the reference and the degraded signals. The results of this algorithm more correlate with the results of listening tests, in comparison with *E-model*. At present, this algorithm is no longer used because of a raw time-alignment. Instead of it the algorithm *PESQ* is rather used. The algorithm *PESQ* facilitates with very fine time-alignment and one single interruption are also taken into account in the calculation of *MoS*. It is possible to use *PESQ* in mobile networks as well as in networks based on packet transmission. The disadvantages include impossibility to use it for codec with data rate lower than 4 kbps and higher calculation load what is caused by recursions in the algorithm.

Some works have been carried out on the effects of packet loss and jitter on speech quality. Particularly, [3], [4], [5] examined these effects in the *MoS* domain for certain packet loss rate and packet sizes. In [3], a formula based on the subjective *MoS* test was suggested, where linear *PCM*, and random packet loss were used, and the lost packet were replaced by silence. It models that *MoS* drops logarithmically with increasing packet loss rate or packet size. In [4], several common speech coders, and random packet loss were used without error concealment. The same formula as in [3] was used to fit *MoS* measured by *PAMS*. In [5], several packet loss rates, packet sizes and error concealment techniques for codec *ITU-T G.729* were examined. *MoS* is applied as an index for speech quality and is measured by *PESQ*. A formula, which quantifies these effects in the *Ie* domain, and finally incorporated into the current *E-model*, is then proposed. *MoS* can be predicted from the formula for the given network conditions without doing real measurements.

Here we focus on the impact of background traffic on speech quality of transmission sequences in the environment of *IP* networks. The background traffic was generated by means of *D-ITG* traffic generator [6]. The simulated background traffic consists of three types of current traffics. The current traffics are: data transfer service, multimedia streaming service and Web service. The increasing traffic load causes the increasing jitter and packet loss. In general, speech quality drops with increasing packet loss and jitter. The impact of these types of traffic and traffic load on speech quality is the aim of this paper. The speech quality is assessment by means of the accomplished *PESQ* algorithm. The proposal of a new method of improved detection of the critical conditions in telecommunication networks from the speech quality point of view is presented at the end of this paper.

The rest of the paper is organized as follows: Section 2 reviews the impact of background traffic on voice. Section 3 describes measurement scenario. Section 4 presents the measurement results. In Section 5 we propose a method of improved speech quality centered detection of the critical conditions in telecommunication networks. Section 6 concludes the paper and suggests some future studies.

2 Impact of background traffic on voice services

As a real-time application, *VoIP* is delay-sensitive but can tolerate a certain level of packet loss. Hence, delay and jitter are the main *QoS* measures. Each voice packet should be transmitted within delay bounds. Also, the jitter (i.e., variation of voice packet delay) should be carefully controlled as it may degrade speech quality more severely than delay. Traditionally, an appropriately designed playout buffer is an effective way to deal with jitter and make the voice understandable [7]. Therefore, impact of change of networks parameters (jitter, packet loss) on speech quality is the main goal of this paper. The change of network parameters is realized by means of background traffic. The investigation of these influences allows to design the methods for improving speech quality in *IP* networks.

The actual time period, in which packet appears in network node (hub, switch or router), depends of system load of the system. Given the network node model depicted in Figure 1, in which the output of every port is modelled by queuing system, it's primary goal is to hold the packets and transmit them when the output line is free. Consider that all packets are served periodically (synchronous output line) without priorities. Have a time between packet departures given by T and the packet service rate $\mu = 1/T$. The offered load should be defined as follows $\rho = \lambda / \mu$, where λ is the mean packet arrival rate.

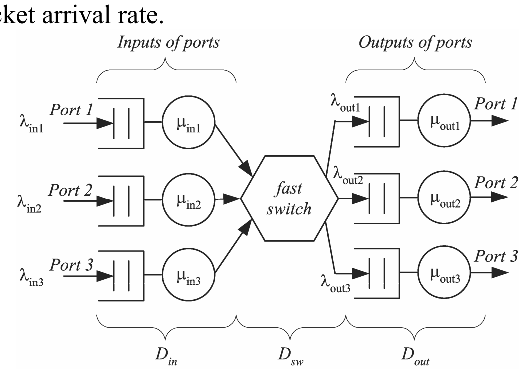

Fig. 1 Three port network node queuing model

The main system characteristics as are packet loss probability and mean packet delay are determined by the distribution of packets in the buffer [8, 9, 10]. Let N_i be the number of packets in the buffer at the end of i-th period and $A_i(N_i)$ denotes the number of packets that arrive during the i-th period, assuming that there are N_i packets in the system. The state of the system at the end of the next period is given by:

$$N_{i+1} = N_i - 1 + A_{i+1}(N_i) \ ; \quad N_i \geq 1 \tag{1}$$

If the system is empty at the beginning of period, at the end of the next period the system contains only messages that arrived during the period:

$$N_{i+1} = A_{i+1}(N_i) \ ; \quad N_i = 0 \tag{2}$$

Have a finite buffer that can hold maximum B packets. Let a_n is the probability of n packet arrivals in a period and P_i; $i = 0, 1, ..., B$ is the steady-state probability that there are i packets in the buffer at the beginning of the period. The probabilities are given by:

$$P_i = \begin{cases} P_0 a_i + \sum_{j=0}^{i} P_{j+1} a_{i-j} \ ; & 0 \leq i < B \\ P_0 \sum_{j=B}^{\infty} a_j + \sum_{j=0}^{B-1} P_{j+1} \sum_{k=B-j}^{\infty} a_k \ ; & i = B \end{cases} \tag{3}$$

In each of these equations, we can solve for P_{i+1}:

$$P_{i+1} = \frac{P_i - \sum_{k=1}^{i} P_k a_{i-k+1} - P_0 a_i}{a_0}; \quad 0 \leq i < B \tag{4}$$

The reason for this simplicity lies in the fact that the decrease over a period time can be at most one packet. The average number of packets in the buffer should be calculated from:

$$\bar{N}_P = \sum_{i=0}^{B} i P_i \tag{5}$$

An outgoing slot is empty only if the buffer is empty; accordingly, the rate on the output line is $(1-P_0)/T$ packets per second. Since the rate at which packets arrive is $\sum_{i=1}^{\infty} i a_i / T$ packets per second, the packet loss rate normalized to the input rate is

$$\bar{L} = 1 - \frac{1-P_0}{\sum_{i=1}^{\infty} i a_i} \tag{6}$$

Following Little's formula [10], the average delay of packet in seconds is

$$\bar{D}_P = \frac{\bar{N}_P}{\frac{1-P_0}{T}} = \frac{T \sum_{i=0}^{B} i P_i}{1-P_0} \tag{7}$$

The delay of packets in the system depends on the number of packets in the buffer that has stochastic character (characterized by probabilities P_i). Because of that the delay of individual packets is stochastic, what causes delay variation (jitter).

Assume that the arrival process is composed of the many sources. In this case, the arrival process should be estimated by Poisson process and therefore the probabilities a_n are given by formula:

$$a_n = \frac{(\lambda T)^n}{n!} e^{-\lambda T} \ ; n = 0, 1, 2, ... \tag{8}$$

The figures 2 and 3 show the average delay and packet loss for Poisson arrival process for various buffer lengths and serving time $T = 1 \mu s$.

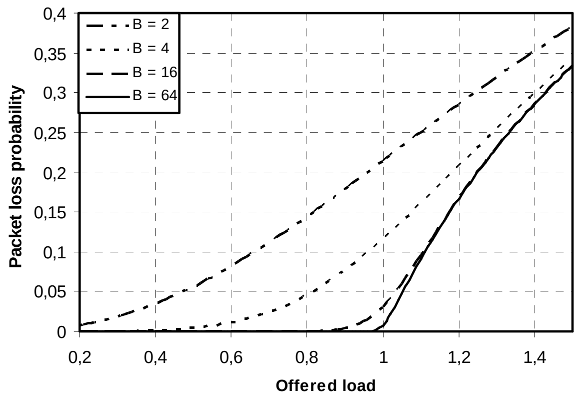

Fig. 2 Packet loss versus offered load

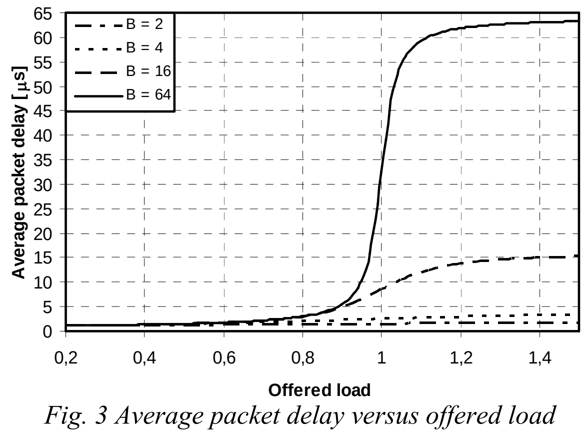

Fig. 3 Average packet delay versus offered load

The upper mentioned approach takes into the account only delay of the packet caused by queuing system. The voice quality is affected by the end-to-end packet delay that consists of the following main components [11]:

- Propagation delay - depends on the physical distance of communication path and communication medium.

- Transmission delay – the time it takes the network interface to send out the packet.
- Queuing delay – the time that a packet has to spend in the queues before it can be processed.
- Codec processing delay – including codecs's algorithm delay and lookahead delay.
- Packetization/depacketization delay – the time needed to build data packets at the sender as well as to strip off packet headers at the receiver.
- Playout buffer delay, the time waited at playout buffer at the receiver's terminal.

3 Measurements

3.1 Experimental setup

One-way *VoIP* session was established between two hosts (*VoIP* Sender and *VoIP* Receiver), via the *IP* network, in *IEEE 802.3i* 10Base-T Ethernet (Figure 4).

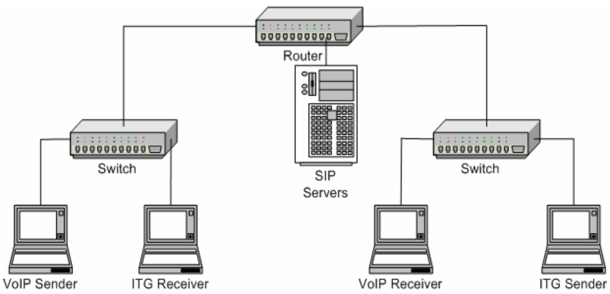

Fig.4 Measurement setup

Two stations (*ITG* Sender and *ITG* Receiver) equipped with the accomplished *D-ITG* traffic generator were used to generate and receive background traffic. *ITG* Sender generated the User Datagram Protocol (*UDP*) and Transmission Control Protocol (*TCP*) packets of length 1024 bytes. Background traffic is described in the chapter 3.3. Voice traffic was generated using *VoIP* clients. Session Initiation Protocol (*SIP*) is used for established *VoIP* connection. For the measurement we chose the *ITU-T G.729A* encoding scheme [12]. In the measurement, one frame were encapsulated into a packet in turn, it corresponding to a packet size of 10 ms. Adaptive jitter buffer, packet loss concealment and voice activity detector are implemented in the using *VoIP* clients.

The *ITU-T* recommendation *P.862.3* [13] recommends to use a sequence in duration in the range from 8 seconds to 30 seconds for the purpose of speech quality measurement. We decided to use the sequence in duration of 30 seconds for the needs of our measurements. The sequence in duration of 30 seconds enables to realize the precise measurement of speech quality. The duration of each measurement was defined 30 seconds. The measurements were performed for six different testing conditions. The sequences described in chapter 3.2 are utilized for transmission through the given *VoIP* connection. Finally, *PESQMoS* was measured by *PESQ* algorithm.

3.2 Description of the sequences

Two types of sequences were used for the purpose of the measurement. The first type of sequence is the test sequence that is composed from simple signals. Speech sequences are the second type of sequence. The speech sequences are composed from speech records.

3.2.1 Description of the test sequence

The test sequence consists of non-speech-like (fully artificial) signals. These signals are closer defined in *ITU-T* recommendation *P.501* [14] and the recommendation divides them into deterministic and random signals. Development of the method for more precise detection of the critical conditions in the telecommunication networks from the speech quality point of view, was our motivation for the work with simple signals. The duration of the test sequence is set to 30 seconds. The test sequence is composed of the following signals introduced and evaluated in [15]:

- Sinusoidal signal with frequencies 300, 800, 1000, 1700, 2400, 3000 Hz ,
- Square bipolar signal with frequencies 300, 400, 500, 600, 635, 670 Hz ,
- Gaussian white noise with $\mu = 0$ and $\delta =$ 0.0001; 0.001; 0.005; 0.01; 0.025; 0.05.

The principle of creation of the final test sequence is based on arranging the parts of the test sequence, which are shown in Figure 5 and Figure 6. The final test sequence consists of six sections. Each section consists of five parts. The arrangement shown in Figure 5 is used once and then the arrangement shown in Figure 6 is used four times to form the first section of final test sequence. The arrangement shown in Figure 6 is used five times to form the other sections of the final test sequence. The signals step-by-step have got the values defined above. That means, in the second section of the test sequence (from 5 sec. to 10 sec.), the signals have the following values: Square bipolar signal $f = 400$ Hz, Gaussian white noise $\delta = 0.001$ and Sinusoidal signal $f = 800$ Hz. The values of the signals in the first section of the test sequence (from 0 sec. to 5 sec.) are the same as those in Figure 5 and Figure 6. The test sequence was stored in 16-bit, 8000 Hz linear *PCM*.

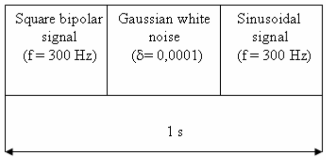

Fig.5 Initial part of test sequence (test seq)

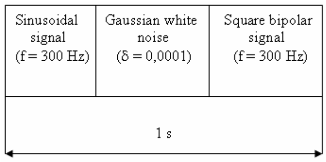

Fig.6 Second part of test sequence (test seq)

The choice of the test sequence for intrusive measurement of *VTQoS* is published in [17]. The optimization of the test sequence for *ITU-T G.729* encoding scheme is published in [19].

It was found that the test sequence composed from simple signals is more sensitive to the interferences in environment of *IP* network than a sequence composed from speech samples. Hence, such test sequence is more suitable for more precise detection of the qualitative changes in the *IP* networks. We compared these two types of sequences in [21].

3.2.2 Description of the speech sequences

The speech sequences selection should follow the criteria given by *ITU-T* recommendation *P.830* [16] and *ITU-T* recommendation *P.800* [1]. The speech sequences should include bursts separated by silence periods, and are normally of 1-3 seconds long. Also it should be active for 40-80% of their duration. The speech sequences are composed from speech records. In our experiments, these speech records come from a Slovak speech database. In each set, two female and two male speech utterances were used. The speech sequences, which are of 30 seconds long with 57 % average value of active speech interval, were stored in 16-bit, 8000 Hz linear *PCM*.

3.3 Background traffic

Background traffic has been generated by *D-ITG* traffic generator. The primary goal of background traffic is to simulate standard traffic that appears in *IP* network, which includes data transfer via Hypertext Transfer Protocol (*HTTP*) and File Transfer Protocol (*FTP*), multimedia streams for real-time applications. The simulated background traffic includes three types of communication:

- "Data transfer service", which includes *FTP* and other non specified services, is represented as information stream with constant bit rate based on *TCP*.
- "Multimedia streaming service" represents real-time multimedia applications and therefore is based on information stream with constant bit rate. The *UDP* is used in this case.
- "Web service" that is simulated as a sequence of separated data bursts with Poisson distribution of packet rate. The active period of burst is 400 ms and the bursts appear periodically every two seconds. *TCP* was used for the purpose of this service.

The measurements have been performed for six different testing conditions. The selected bit rates of three above mentioned types of communication and average traffic load of background traffic are described in Table 1 and Figure 7. The simulation of the multimedia streaming service was carried out from the point of view of the impact of traffic of this service on speech quality. Note *D-ITG* traffic generator doesn't allow the simulation of the multimedia streaming service using Real-time Transport Protocol (*RTP*) but the *RTP* based streaming service has the same impact on speech quality as the streaming using *UDP*. The aim of this measurement was to investigate how these types of traffic and traffic load effect speech quality.

Table 1 Performance evaluation of testing conditions

Testing condition	**Data transfer Service [Mb/s]**	**Streaming service [Mb/s]**	**Web service [Mb/s]**	**Average traffic load [%]**
0	0	0	0	0
1	2	2.5	0.5	50.5
2	2.25	2.82	0.56	59.8
3	2.5	3.14	0.61	65.7
4	2.75	3.45	0.68	70.8
5	3	3.76	0.74	77.2

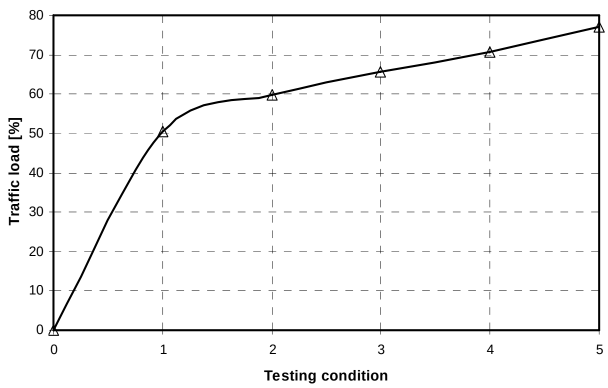

Fig.7Traffic load for given testing conditions

The traffic load was measured by means of *Wireshark* network analyzer [24].

3.4 Assessment of the speech quality

PESQMoS was evaluated by the *PESQ* metric [2], the most recent *ITU-T* standard for objective speech quality assessment. *PESQ* combines merits of *PAMS* and *PSQM99* (an updated version *PSQM*), and adds new methods for transfer function equalization and averaging distortions over time. It can be used in wider range of network conditions, and gives higher correlation with subjective tests and the other objective algorithms [2], [18]. Unlike to the conversational model, *PESQ* is a listening-only model; the degraded sample is time-aligned with the reference sample during pre-processing. The *PESQMoS* values do not reflect the effects of delay on speech quality.

4 Measurement results

The measurement was independently performed 10 times under the same testing conditions. The *PESQMoS* results were averaged out and the standard deviation was kept within 0.085 *PESQMoS* for the test sequence and 0.111 *PESQMoS* for the speech sequences.

Figure 10 shows the measurement results for the test sequence and speech sequences. The graphs represent the dependence of *PESQMoS* change on the testing conditions. The testing conditions represent a few types of network conditions. Each network condition is described by traffic load. The increasing traffic load causes jitter and also packet loss increase. In general, speech quality drops with increasing packet loss and jitter. Figure 7 shows the traffic load for given testing conditions. The transmission rates for given testing conditions are described in Table 1. The impact of background traffic on the jitter (delay variation) and packet loss in *VoIP* connection is shown in Figure 8 and Figure 9.

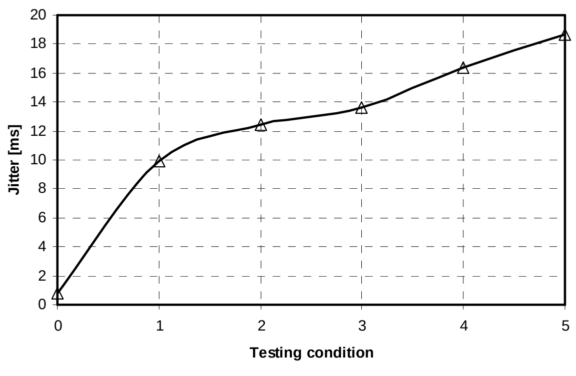

Fig.8 Impact of background traffic on average value of jitter in VoIP connection

The 1550 voice packets were approximately transmitted during one 30 seconds long *VoIP* connection. The average value of jitter ranged from 0.825 to 18.625 ms and the total packet loss ranged from 0.35 to 13.88 % for these measurements. The total packet loss consists of two components. The first component is lost packets and the second component is dropped packets.

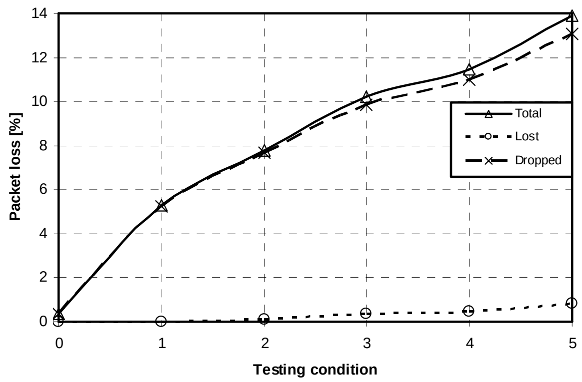

Fig.9 Impact of background traffic on packet loss in VoIP connection

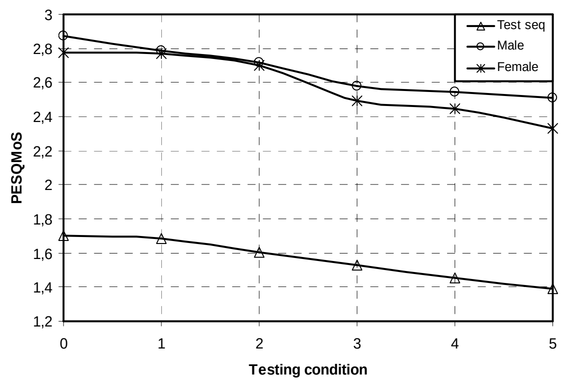

Fig.10 Impact of background traffic on speech quality of speech sequences and test sequence

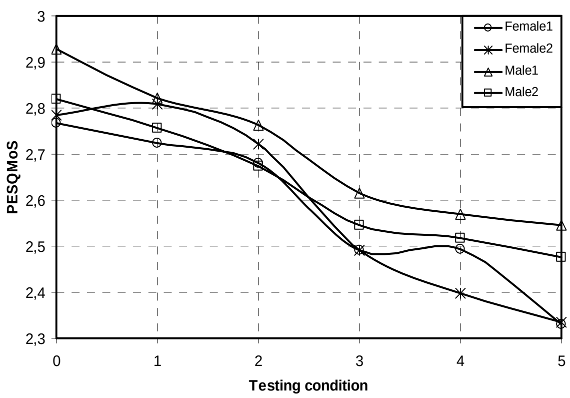

Fig.11 Impact of background traffic on speech quality of speech sequences

Figure 11 shows the measurement results for the speech sequences. We can see some differences in impact of background traffic on speech quality among the individual speech sequences in Figure 11. It is caused by a different arrangement of the speech sequences as well as by differences in duration of their active speech intervals among speakers. Figure 10 represents only average values separately for female and male speech sequences. As seen in Figure 10, the test sequence has smoother characteristic than the speech sequences have. From the speech quality point of view, the test sequence responds to network parameters changes more sensitively than the speech sequences do. It allows to carry out the detection of qualitative changes in telecommunication networks more precisely. We aim to use it for the development of a method for improved speech quality based detection of critical conditions in telecommunication networks.

5 Proposal of the method of improved detection of critical conditions in telecommunication networks from the speech quality point of view

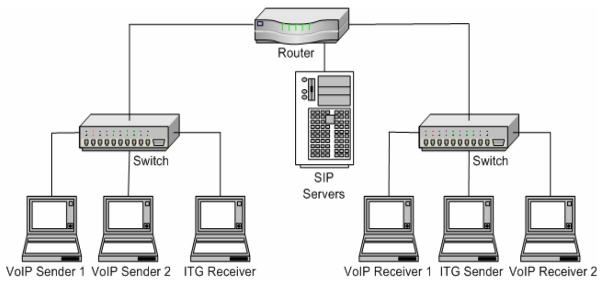

Fig.12 A typical network scenario in the environment of IP networks

Figure 12 represents a typical network scenario in the environment of *IP* networks. *VoIP* connection No.1 is used for the real *VoIP* transmission and *VoIP* connection No.2 is used for the purpose of improved detection of critical conditions in wireless telecommunication networks. This method is based on the intrusive measurement of Voice Transmission Quality of Service (*VTQoS*). The test sequence is transmitted through the *VoIP* connection No.2. *ITG* stations represent the data traffic stations. The proposal of the method of improved speech quality centered detection of critical conditions in telecommunication networks follows the experiment, which is described in this paper. Such improved detection allows to realize earlier response to the change of network parameters and it may avoid the impairment of speech quality. The principle of the improved detection of the critical conditions is based on the simple scheme. The measurement by using the test sequence is carried out in successive 30 seconds steps and the duration of the measurement is set to 30 seconds. The *PESQMoS* value is computed for each measurement. The *PESQMoS* is used as a threshold parameter for the critical conditions. The threshold value for critical conditions is set to 1.575 *PESQMoS* for the test sequence. This critical decision threshold has been derived from the results, which were obtained by the measurements described in this paper. As seen from the measurement results (Figure 10), the *PESQMoS* value 1.575 for the test sequence corresponds with about 63 percents value of traffic load. The *PESQMoS* values of all the speech sequences (Figure 11) are fast declined under the traffic load value of about 63 percents. The jitter and packet loss parameters (Figure 8, 9) are very fast increasing above 63 percents the traffic load value. Table 2 figures out the decision level for the relevant network conditions.

Table 2 Decision level for the relevant network conditions

PESQMoS value	The network condition
> 1.575	standard
< 1.575	critical

When the *PESQMoS* value for test sequence is higher than the threshold value, the network is situated in the standard network conditions from the speech quality point of view. It means that speech quality is kept within a tolerable range. The change of network parameters is not necessary. The critical network conditions are expected when the *PESQMoS* value for test sequence is below the threshold value. The impairment speech quality is expected and a change of network parameters is required. The acceptable speech quality can not be expected when a change of the network parameters is not done. The network parameter change represents prioritizing voice, congestion avoidance methods and etc.

6 Conclusion and future work

This paper investigated the impact of background traffic on speech quality in *VoIP* applications. The different traffic testing conditions were used for the purpose of the measurements. The each testing condition consists of the three types of current traffics, which exist in the telecommunication networks. The current traffics are: data transfer service, multimedia streaming service, Web service. The results show that test sequence responds to the change of the network parameters more sensitively than the speech sequences from the speech quality point of view. It allows to perform improved detection of qualitative changes in telecommunication networks. We propose the method for such speech quality centered detection of critical conditions in telecommunication networks. In the future we aim to expand this method to algorithms for link adaptation from the speech quality point of view in the environment of *IP* networks. The primary goal of these algorithms is improving speech quality in *VoIP* connections, which are established in the competent link. The first type of the algorithm will be based on the fragmentation of data packets. The fragmentation of the large datagrams to packets of a size small enough to allow to satisfy the delay requirements of the delay-sensitive traffic. Small delay-sensitive packets could be interleaved between fragments of the large datagram. The second type of the algorithm will be based on the Generic Traffic Shaping (*GTS*) [22] and Weighted Random Early Detection (*WRED*) [22] functions, too. The principle of this algorithm will be based on prioritizing voice by means of congestion avoidance methods (*WRED*, *GTS*). Future work will also focus on development and verification of these algorithms for such speech quality centered link adaptation.

7 References

[1] ITU-T Rec. P.800 "Methods for subjective determination of transmission quality" International Telecommunications Union, Geneva, 1996.

[2] ITU-T Rec. P.862 "Perceptual Evaluation of Speech Quality" International Telecommunications Union, Geneva, 2001.

[3] Yamamoto, L., Beerends, J. G.: Impact of network performance parameters on the end-to-end perceived speech quality, In *Proceedings of Expert ATM Traffic Symposium*, Mykonos, Greece, 1997.

[4] Duysburgh, B., Vanhastel, S., Vreese, B., Petrisor, C., Demeester, P.: On influence of best-effort network conditions on the perceived speech quality of VoIP connections, In *Proceedings of IEEE International Conference on Computer Communications and Networks*, Phoenix, USA, 2001, pp 334-339.

[5] Ding, L. Goubran, R. A.: Assessment of effects of packet loss on speech quality in VoIP, In *Proceedings of the 2nd IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications, 2003, HAVE 2003. Ottawa, Canada, September 2003, pp. 49-54.*

[6] Distributed Internet Traffic Generator, http://www.grid.unina.it/software/ITG/

[7] Wang, P., Jiang, H., Zhuang, W.: IEEE 802.11e Enhancement for voice service, In *IEEE Wireless Communications*, vol. 13, no. 1, pp. 30-35, Feb. 2006, ISSN 1536-1284.

[8] Vastola, K. S.: Performa Modelling and Analysis of Computer Communication Networks, [online], [2006-10-11], http://networks.ecse.rpi.edu/_vastola/pslinks/erf/node1.html.

[9] Rukovanský, I.: Hodnotenie výkonnosti počítačových systémov (in Slovak), Prague, SNTL, 1989, ISBN 80-03-00057-2.

[10] Hayes, J. F., Babu, T. V. J. G.: Modeling and Analysis of Telecommunication Networks, John Wiley & Sons, 2004, ISBN 0-471-34845-7.

[11] Sun, L.: Speech Quality Prediction for Voice over Internet Protocol Networks, Dissertation thesis submitted to the University of Plymouth, January 2004.

[12] ITU-T Rec. G.729 "Coding of speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Exited Linear prediction (CS-ACELP)" International Telecommunications Union, Geneva, 1996.

[13] ITU-T Rec. P.862.3 "Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2" International Telecommunications Union, Geneva, 2005.

[14] ITU-T Rec. P.501 "Test signals for use in telephonometry" International Telecommunications Union, Geneva, 2000.

[15] Počta, P., Vaculík, M.: Method of choice of test signals for automatic intrusive measurement VTQoS, In *Proceedings of conference MESAQIN 2005*, Prague (Czech Republic), 2005, ISBN 80-01-03262-0.

[16] ITU-T Rec. P.830 "Subjective Performance Assessment of digital telephone-band and wideband digital codecs " International Telecommunications Union, Geneva, 1996.

[17] Počta, P., Vaculík, M.: Determination of optimal test sequence for automatic intrusive measurement VTQoS on environment of fixed telecommunication network, In *Proceedings of conference RTT 2005*, Ostrava (Czech Republic), 2005, ISBN 80-248-0897-8.

[18] Rix, A.W., Beerends, J.G., Hollier, M.P., Hekstra, A.P.: Perceptual Evaluation of Speech Quality (PESQ) a new method for speech quality assessment of telephone network and codecs, In *Proceedings of IEEE international conference on Acoustics, speech and signal processing, Salt Lake City, USA, 2001, pp.749-752.*

[19] Počta, P., Vaculík, M.: Researching of coders influence on basic measurement signals used in optimal test sequence (in Slovak), In *AEEE Journal*, Žilina (Slovakia), No.3, Vol.5/2006, ISSN 1336-1376.

[20] Holub, J., Šmíd, R., Bachtík, M.: Child Listeners as the Test Subject – Comparison with Adults and P.862, In *Proceedings of conference MESAQIN 2003*, Prague (Czech Republic), 2003, ISBN 80-01-02822-4.

[21] Počta, P., Vaculík, M.: Comparison of test sequence for intrusive measurement of VTQoS with speech sequences in the environment of IP networks, In *Proceedings of conference MESAQIN 2006,* Prague (Czech Republic), 2006, ISBN 80-01-03503-4.

[22] Cisco IOS Quality of Service Solutions Configuration Guide, Release12.0, [online], [2007-01-23], http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/qos_c/index.htm

[23] Pennock, S.: Accuracy of the Perceptual Evaluation of Speech Quality (PESQ) algorithm, In *Proceedings of conference MESAQIN 2002*, Prague (Czech Republic), 2002, ISBN 80-01-02515-2.

[24] Wireshark network analyzer, http://www.wireshark.org/