ISSN 1392 – 1215 ELEKTRONIKA IR ELEKTROTECHNIKA. 2005. Nr. 7(63)

T 180 TELEKOMUNIKACIJŲ INŽINERIJA

Investigation the Ability of Objective Measures of the Perceptual Speech Quality in Mobile Networks

A. Kajackas, A. Anskaitis
Department of Telecommunications Engineering, Vilnius Gediminas Technical University,
Naugarduko str. 41, LT-03227 Vilnius, Lithuania; e-mails: algimatas.kajackas; aurimas.anskaitis@el.vtu.lt

Introduction

An important performance aspect of telephony networks is the perceived conversational speech quality of voice calls. The overall quality is dependent on many factors including signal level, echo, delay, signal-to-noise ratio, codec type. In wireless networks such as GSM, WCDMA, and voice-over-IP, there are also additional factors affecting speech quality such as packet loss.

This paper continues in [1, 2] initiated work. The purpose of the research is to show possibilities and ways to construct and deploy special module for the de facto perceived service quality evaluation. Demand for such module can be justified because quality perceived by individual users varies greatly for different users. Quality of Service (QoS) depends on user location, time, network load and so on. When user moves his communication conditions changes also changes his perceived quality level.

The main purpose of this work is to choose means and methodology for the evaluation of perceived voice quality under different non-stationary conditions. AMR-n codecs are used in voice quality analysis (3GPP TR 26.975 V5.0.0).

Initial conditions

In current systems voice is divided into length T time intervals (typically 10, 20, 30 ms), coded and transmitted. Transmitted packet is named frame. Some frames may arrive with errors to the receiver. It is impossible to decode such damaged frames. Therefore damaged frames are erased. Decoder uses some kind of substitution algorithm instead of normal decoding when it receives erased frame.

Particular mobile telecommunication conditions may determine many lost frames. Such flow of lost frames is the direct reason of voice quality degradation. There are many investigations in the field of voice quality evaluation under different frame loss conditions. However majority of investigations deals with relatively long speech segments and stationary loss conditions.

Till now performed research shows that sparse lost frame patterns worsens speech quality but not much. For example, the ETSI TIPHON project offers the following quality grades for IP telephony voice packet loss: < 0,5% for class 1 = gold, from 0,5% to 1% for class 2 = silver, and 1% to 2% for class 3 = bronze.

Voice quality is deteriorated much more when lost patterns are bursty. Parts of words or even complete words may be lost. In [3] it is shown that under poor communication conditions it is common that separate words or groups of words may be lost. Duration of such erasures is between few tens of milliseconds to few seconds.

The methods of voice quality evaluation and their evolution

The need to measure speech quality is a fundamental requirement in communications systems for technical, legal and commercial reasons. The measurements of speech quality in communication networks may be performed with intrusive and non-intrusive means. The measurements are intrusive since test calls are injected into the network to conduct the measurements. This way of measurement may be useful in evaluation of statistical network supplied quality of service. It is also required for testing of non-intrusive methods.

Problem of voice quality evaluation has become especially important when special purpose voice coding methods emerged. Vocal quality is an interaction between an acoustic signal and a listener. The acoustic signal itself does not possess quality, it evokes it in the listener. The first in ITU P.800 recommendation described method for subjective voice quality evaluation was MOS (Mean-Opinion-Score). The ITU recommended MOS test method for listening-only tests is the Absolute Category Rating (ACR). A MOS ACR value is normally obtained as an average opinion of quality based on asking people to grade the quality of speech signals on a five-point scale (5 Excellent, 4 Good, 3 Fair, 2 Poor, 1 Bad) under controlled conditions as set out in the ITU standard .

Commonly used are ACR and DCR (Degradation Category Rating) methods. The DCR method compares the system under test with a high quality fixed reference and the degradation (from "Inaudible" to "Very annoying") is rated on a five-point scale. The average of the opinion scores of subjects in DCR is called Degradation Mean Opinion Score (DMOS). This method is suitable when the impairment (especially digital) is small. It may therefore be particularly useful for evaluating similar digital speech processing algorithms. DCR method may serve as a means for system optimization once.

As subjective measure MOS can only be representative under test conditions in which they were recorded. Listening tests performed with other conditions than those used in the characterization phase may lead to a different set of MOS results. Regardless of that, the MOS method seems being very suitable and is used for the QoS evaluation experiments when analyzing new coding or transmission methods and devices and resulting voice quality impairments.

An objective measurement is made using measurement equipment and is repeatable given the same external conditions. Nowadays, there are quite a number of algorithms like Perceptual Speech Quality Measure (PSQM) and Perceptual Evaluation of Speech Quality (PESQ) that provide an objective MOS-equivalent score for a voice call. PSQM (ITU-T. Rec. P.861) was originally designed to evaluate codec quality. PSQM+ is an enhancement of PSQM to cover short duration temporal clipping as often seen in wireless communications.

PESQ (ITU-T. Rec. P.862) is an intended replacement of PSQM. It was developed by combining the two advanced speech quality measures PSQM+ and PAMS (*Perceptual Analysis Measurement System*). PESQ compares an original speech sample $x(t)$ with its transmitted and hence degraded version $y(t)$. After some preprocessing, such as level- and time-alignment, both the original and degraded speech signals are transformed into a psychoacoustic representation which models the properties of the human auditory system. The output of PESQ is a prediction of the perceived quality that would be given to $y(t)$ by subjects in a subjective listening test and directly produces an objective MOS ACR in the range $(-1..4.5)$.

PESQ provides significantly higher correlation with subjective opinion than the PSQM [4]. Even for the test cases for which PESQ has been designed it might not perform perfectly. In [5], the author compared the speech quality prediction of PESQ with human conducted subjective tests, covering test conditions with impairments due to coding distortions and packet losses. The difference between the PESQ MOS value and the MOS from subjective tests is below 0.25 for 70% and below 0.50 for 90% of all test conditions. The correlation between subjective and objective tests is about 0.93.

Generalizing voice quality evaluation methods the following diagram has been made (Fig .1.)

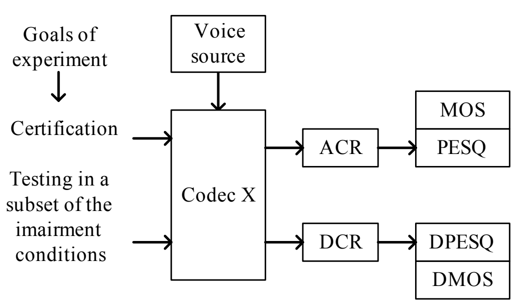

Fig. 1. Voice quality evaluation methods

The beginning of the diagram – voice source and object under investigation – codec. Two separate investigation purposes are depicted. The first goal is certification of new codec. This kind of research seeks to find the absolute quality rating Q_{ACR}. Second goal – testing of quality degradation in a subset of impairment conditions. The research seeks to find relations between impairments and perceived speech quality degradation. For this purpose DCR quality measure Q_{DCR} is better suited.

It is important to notice that perceived voice quality can be expressed as:

$$Q_{\mathrm{perc}} = Q_{\mathrm{ACR}} - Q_{\mathrm{DCR}} . \tag{1}$$

The perceived voice quality and quality degradation are related additively. Such scale is not convenient if the purpose is to relate quality and pricing. For pricing it is possible to use multiplicative form which evaluates different impairments.

All the objective methods of speech perceptual quality measurement that have been mentioned in this paper (PSQM, PESQ etc) were designed to characterize the effect of static impairments introduced by low-bit-rate vocoders. They compute an objective perceptual quality score by comparing the original transmitted input signal $x(t)$ with the received output signal $y(t)$. As such, it requires knowledge of the original voice signal that was transmitted. Therefore all mentioned methods can not be directly applied to evaluation of quality of real conversation. But these methods are required for the research purposes.

Recommended voice quality

For transmission planning purposes ITU-T (Rec. G.107) has created a E-Model. This model does not define a new way of quality evaluation. It only determines the methodology of evaluation of separate factors. The E-model predicts the subjective quality of a telephone call based on its transmission parameters. It combines the impairments caused by these transmission parameters into a rating R, which can be used to predict subjective user reactions. One exceptional feature of the E-Model is the assumption that the psychological effect of uncorrelated sources of impairments is additive. This assumption is based on empirical results in the field of psychophysical research.

The rating factor R is composed of:

$$R = R_0 - I_s - I_d - I_e + A , \tag{2}$$

where R_0 represents, in principle, the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise; I_s is the combination of all impairments, which occur simultaneously with the voice signal, such as those caused by quantization, by too loud connection and by too loud side tone; I_d represents the impairments caused by delay; I_e , the equipment impairment factor, represents impairments caused by the low bit rate codec; it also includes impairments due to packet loss of random nature; A- the advantage factor.

ITU in Rec. G.108 also set a connection between R values, speech transmission quality category and user satisfaction level (Table 1).

An estimated Mean Opinion Score (*MOS*) for the conversational situation in the scale 1-5 can be obtained from the *R*-factor using the formula [6]:

$$MOS = \begin{cases} 1, R < 0, \\ 1 + 0.035 \cdot R + \\ + R \cdot (R - 60) \cdot (100 - R) \cdot 7 \cdot 10^{-6}, R < 100, \\ 4.5, R > 100. \end{cases} \quad (3)$$

The graph plotted according to this formulae is depicted in Fig. 2. Also graph shows different user satisfaction levels and respective MOS values. For example R values between 50-60 are described as “Nearly all dissatisfied”. Correspondent MOS values are 2.575-3.1 and in MOS terminology are described as “Fair”.

In the rating formula (2) *A* is the advantage factor which allows for compensation of impairment factors when there are other advantages to the user from the access to a given network/service. Examples of defining the factor *A* are provided in ITU-T Rec. G.108. There it is recommended to set *A* = 0 for a conventional (wirebound) telephony service, *A* = 5 for a cellular network in a building, *A* = 10 for a typical mobile network (in a geographical area or moving in a vehicle) services and *A* = 20 for a satellite communication services provided in remote areas

Table 1. Definition of categories of speech transmission quality

Range of Rating	Category of Quality	User satisfaction
$90 \le R < 100$	Best	Very satisfied
$80 \le R < 90$	High	Satisfied
$70 \le R < 80$	Medium	Some users dissatisfied
$60 \le R < 70$	Low	Many users dissatisfied
$50 \le R < 60$	Poor	Nearl all users dissatisfied

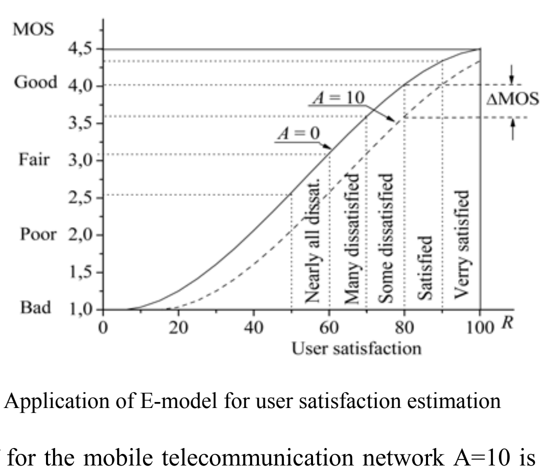

Fig. 2. Application of E-model for user satisfaction estimation

If for the mobile telecommunication network A=10 is chosen, according to E-model higher distortions are allowed (noise, delay, jitter, packet loss). This phenomenon is illustrated in fig. 2 by dashed line. According to this curve for the same satisfaction levels lower MOS values are assigned. For example “Nearly all dissatisfied” region is projected into MOS values between 2.064-2.575.

Table 1 sets officially recommended limits for different quality classes. Decrease in R value by 10 changes (worsens) Category of Quality. On the ground of these data it is logically to distinguish five degradation classes for iQoS purposes. These classes may be coincided with MOS DCR scale – from “Inaudible” to “Very annoying”. For distinct degradation levels different tarrifs should be applied.

Quality of AMR codecs

Generalizing the analysis of voice quality Table 2 was made. This table describes properties of AMR-n codecs in good conditions (clean speech and without errors). The fourth column of the same table gives PESQ values [5]. Difference between MOS ACR and PESQ never exceeds 0.25. The fifth column in table 2 named “User satisfaction” was estimated on the basis of formula (3) and Fig. 2.

It is necessary to notice that only two AMR codecs (AMR-7, AMR-6) are characterized as satisfying users. The other types of codecs falls to category “Some dissatisfied” or “Many dissatisfied”.

Table 2. Properties of AMR-*n* codecs

Codec mode	Rrate kbit/s	MOS [3]	PESQ [5]	User Satisfaction
AMR-7	12.2	4.06	3.92	Satisfied
AMR-6	10.2	4.06	3.84	
AMR-5	7.95	3.91	3.7	Some dissatisfied
AMR-4	7.4	3.83	3.65	
AMR-3	6.7	3.77	3.55	
AMR-2	5.9	3.72	3.48	
AMR-1	5.15	3.50	3.32	Many dissatisfied
AMR-0	4.75	3.50	3.3	

Any information transmission system should fulfil a list of requirements when communication conditions are good. This initial quality of service level may be called “nominal quality level”. Nominal quality levels may be different for different technologies. For example, nominal voice quality is higher in ISDN than in GSM system. In GSM networks it is logically to use AMR-7 voice quality as nominal.

Research of the impact of lost frames

Some works have been carried out on the effects of packet loss on speech quality. Cox and Perkins compared the impact of random and burst packet loss [6]. They found that for low packet loss rates a burst distribution gave a higher subjective quality than a non bursty distribution whereas for high packet loss rates the converse was true. At the lower 1 and 2 percent rates there is no significant effect of the distribution (random vs. bursty) variable, whereas with the high 5 percent rate opinion suffers more when packets are lost in bursts. Comprehensive research on lost packets influence on speech quality is given in Tiphon report. For the packet loss conditions, frames were removed from the speech samples randomly with a frequency determined by the test condition (e.g., 1% of the frames). Relation of erased frames with speech quality is described in ITU recommendations (G.113/Appendix I. 2002).

Ding and Goubran conducted a modelling in which impairment factor I_e grows logarithmically with increasing packet loss rate or packet size [7].

Conway [8] has proposed original methodology for the research of lost packets impact. This methodology is named Framed PESQ. A basic assumption in the proposed method is that the encoded digital speech signal is transmitted in a framed format. The erased frames are signalled to the decoder of the received encoded speech signal. An exact PESQ measurement of the speech quality is provided in a known reference signal.

Framed PESQ evaluates a known reference signal instead of transmitted speech. There is a doubt whether such method is suitable for iQoS.

The development of speech quality metrics relying only on GSM transmission parameters RxQual, FER has been presented in [9, 10]. The proposed speech quality measure is an empirical function which depends on RxQual, FER.

The main purpose of these works is to predict MOS scores using measured parameters such as RxQual, FER. Authors state that it is possible to predict the subjective speech quality very accurately. The graphs presented in the mentioned work are not very impressive in the sense that maximum deviation of predicted MOS scores from true MOS scores exceeds one.

In [9, 10] it is indirectly shown that FER and RxQual parameters contain enough information for voice quality prediction.

It is possible to construct many kinds of models which take as an input above mentioned parameters and give as an output predicted MOS score for conversation.

Impact of lost frames on the quality of single words

After the review of published works it is possible to make conclusion that there is not enough data to predict speech quality when communication conditions are near critical. Therefore we have made single word degradation research when 1, 2, 3, …, n frames are lost. Different lost patterns were used – from deterministic to purely random. Simulation system is depicted in Fig. 3. The main components of a system are AMR coder, AMR decoder, PESQ measurement algorithm and packet loss imitator.
AMR codec uses 20 ms lasting packets for speech transmission. Every word consists of some of these packets. Typical value is 20-40 packets per word. In our research we will simulate packet loss in every possible position in a word. After this loss we calculate PESQ score for original and damaged words. In this way we have got dependence of PESQ score on loss position for one packet loss in a word in every possible place. Also we have simulated consecutive loss of two, three and more packets starting from every possible place in a word. Graphs of these simulations are presented in Fig. 4

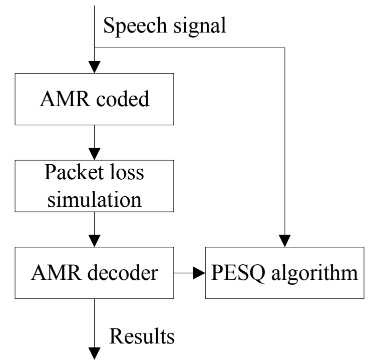

Fig. 3. Simulation scheme

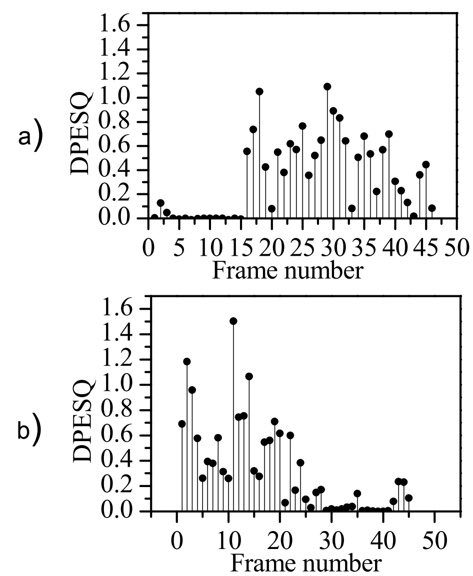

Fig. 4. . DPESQ as function of position of lost frame: a) word "mama", b) word "namas"

We will try to explain results which we get shortly. From graphs it is obvious that speech quality after lost packet substitution substantially depends on loss position. There are places in a word where packet can be lost without noticeable speech quality degradation. Erasure of the packet in some other location decreases speech quality much more. We can see the same tendency in every investigated word. Different words are sensitive for losses in particular positions.

As it can be seen from Fig. 4 many DPESQ values are higher then 0.5. It means that only one lost packet in a word may decrease voice quality so that user satisfaction changes to lower category.

When instead of one lost packet we erasure two or more consecutive packets in a word, speech quality degrades yet more. This is expected result of course. It is natural that when we have some consecutive packets erasures there is no point in a word where delta PESQ has small value.

Then during a mobile call, the iQoS expressed as $Q_k(t)$ would constantly change, reflecting the actual QoS received by that particular user. Evaluation of the received quality may be done over reasonably short time period, e.g. half- or one second. As humans may tell a word in a half- or one-second interval, impairment of quality over that period could be well noticeable.

Analysis of Fig. 4. shows there is not deterministic relation between location of erasure and speech quality. A set of DPESQ values looks like an array of random numbers.

For the purpose to obtain some statistically meaningful data the following experiment was executed. At first 50 different words were chosen. In every separate word k $(k = 1, 2, ..., 10)$ frames were erased randomly and independently. For every word 20 realizations of erasures were generated. After erasure DPESQ scores were calculated.

Conditional empirical cumulative distribution functions were calculated. $F(x \mid k) = P(DPESQ \le x)$, when k packets are lost in the word.

In the Fig. 5 there are presented functions $F(x \mid k)$, where $k = 1, 2, 3, 5, 10$. In the same figure unconditional cumulative distribution function is shown (labelled as unconditional) with premise that distribution of the number of frame errors is uniform.

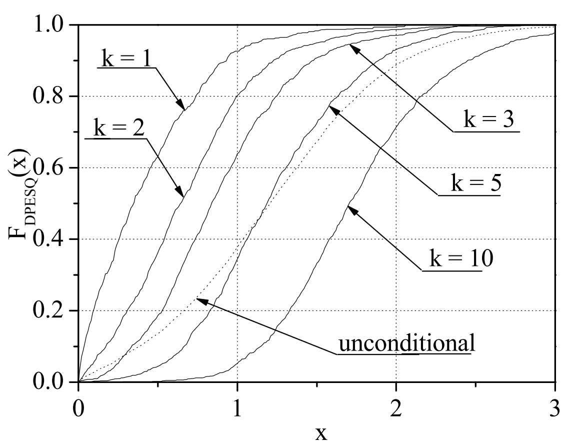

Fig 5. Empirical conditional cumulative distribution functions of quality degradation

Probability that DPESQ value will exceed x is equal to 1 - $F(x \mid k)$. If probabilities that $k = 1, 2, 3, ..., K$ frames will be erased are known and equal $p(k)$ then cumulative distribution function may be expressed as:

$$P(\text{DPESQ} > x) = \sum_{k=1}^{K} p_k (1 - F(x \mid k)) . \qquad (4)$$

From above formula it is clear that overall probability of quality degradation depends on the distribution of frame losses and on the conditional distributions of quality degradation values.

It can be seen that probability of DPESQ exceeding 1 is less than 0.1 when there is only one lost frame in a word. In such cases the meaning of word can be recognized easily. Situation becomes more complicated when number of errors increases. Sometimes word may become unrecognizible and it is the most serious impairment.

Let us notice that we have not a purpose to assign particular number to speech quality yet. We only construct some probabilistic quality degradation measure. Particular number may be obtained by averaging, for example.

The distributions were obtained analyzing single words. From them it is possible to draw some conclusions. Only some lost frames may impair speech quality hardly. On the contrary, sometimes many lost frames are not vital for word recognition.

Conclusions

Majority of quality evaluation methods have been created for the new codecs testing. All ITU recommended quality evaluation methods are based on assumption that communication conditions are stationary. This assumption is doubtless when impact of noise or disturbances is analyzed. Assumption about stationary conditions is extended to lost packets also. For example, ITU recommendation G.113 states that one percent of lost packets increases quality degradation factor I_e by 4. It means ITU relates quality with percent of lost packets, still it is obvious that lost packets in silent intervals do not impact quality at all.

The fact ITU recommendation G.108 allows tolerating some quality degradation for network planning purposes, should be seriously discussed. Major goal of such discussion is the relation between quality and pricing. People may tolerate services of lower quality but pricing should be related with quality.

There is a need to notice that ITU methodology about quality evaluation of voice signals is very important for general QoS evaluation. But in mobile networks communication conditions changes rapidly, perceived quality level constantly varies. PESQ can not adequately account for the voice quality degradation caused by the dynamic impairments such as random series of packet loss.

In mobile networks iQoS measurement can be analyzed as a random process. And it is not stationary. Analyzing this process it is necessary distinguish between quality of small voice segments and overall quality. On purpose to relate perceived quality and pricing in mobile communication conditions it is necessary to use appropriate algorithms which take into account these conditions. In the process of search of such algorithms initially we have chosen PESQ method. But it is necessary to check results of such analysis. The best and maybe the only checking method is subjective MOS tests.

This paper has investigated the effects of packet loss in single words. The results show that sometimes only one lost frame may impair speech quality hardly, still many lost frames sometimes has no impact on speech quality.

Acknowledgment. *We would like to thank Lithuanian State Science and Studies Fondation for partial support for this work.*

References

1. **Kajackas A., Batkauskas V., Medeišis A.** Individual QoS Rating for Voice Services in Cellular Networks // IEEE Communications Magazine. – June 2004. . – P. 88-93.
2. **Kajackas A., Anskaitis A., Guršnys D.** Individual Quality of Service concept in Next Generations Telecommunications networks.
3. **Kajackas A., Batkauskas V.** Impact of Electro-magnetic Interference on QoS in Cellular Networks // Electromagnetic Compatibility. . – Wroclaw, 2004. – P. 244-247.
4. **Rix A. W., Beerends J. G., Hollier M. P. and Hekstra A. P.** PESQ – the new ITU standard for end-to-end speech quality assessment. AES 109th Convention. – Los Angeles, 2000 Sep. . – P. 1 – 18.
5. **Pennock S.** Accuracy of the Perceptual Evaluation of Speech Quality (PESQ) algorithm // Proc. Workshop Measurement of Speech and Audio Quality in Networks MESAQIN. – Prague, Jan. 2002.
6. **Cox R., Perkins R.** Results of a Subjective Listening Test for G.711 with Frame Erasure Concealment. Committee contribution T1A1.7/99-016. – May, 1999.
7. **Ding L., Goubran R.** Assessment of Effects of Packet Loss on Speech Quality in VoIp // Proceedings. The 2nd IEEE Internatioal Workshop. – HAVE, 2003. – P. 49–54.
8. **Conway A. E.** Output-Based Method of Applying PESQ to Measure the Perceptual Quality of Framed Speech Signals. WCNC 2004 // IEEE Communications Society. – P. 2521–2526.
9. **Werner M., Kumps K., Tuisel U., Beerend J., Vary P.** Parameter-based Speech Quality Measures for GSM // Proc. IEEE Intl. Symposium Personal, Indoor and Mobile Radio Communications (PIMRC), Beijing, Sep. 2003. – P. 2611–2615.
10. **Werner M., Junge T., Vary P.**. Quality Control for AMR Speech Channels in GSM Networks**.** ICASSP 2004. – P.1076–1079.

Pateikta spaudai 2005 07 14

A. Kajackas, A. Anskaitis. Kalbos kokybės vertinimas paketinio perdavimo tinkluose // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2005. – Nr. 7(63). – P. 10–15.

Tęsiama iQoS modulio kūrimo galimybių analizė. Išsamiai išnagrinėta mokslinėje spaudoje sukaupta balso kokybės vertinimo patirtis. Atliktus išsamią skelbtų darbų analizę parodoma, kad paplitę balso kokybės vertinimo būdai tiesiogiai netinka balso kokybei realaus pokalbio metu vertinti, nes jie rėmėsi prielaida, kad balso iškraipymai susidaropagal stacionariesiems procesams būdingus dėsningumus. Pavienių žodžių iškraipymų tyrimai rodo, kad daugeliu atveju net vienas ištrintas paketas balso kokybę pažemina tiek, kad vartotojo pasitenkinimo kategorija pereina į gretimą žemesnę. Determinuoto ryšio tarp kokybės degradavimo ir ištrintų freimo vietos žodyje nepastebėta, todėl galima daryti prielaidą, kad tas ryšys yra stochastinis. Gauti empiriniai balso kokybės degradavimo tikimybių pasiskirstymai patvirtina šią prielaidą. Tęsiant tyrimus siekiama nustatyti sąsajas tarp per tam tikrą laikotarpį ištrintų freimų kiekio ir kokybės degradavimo lygio.Suformuluotas tam tikras sąrašas su iQoS įgyvendinimu susijusių uždavinių ir aptariami jų sprendimo būdai. Il. 5, bibl 10 (anglų kalba; santraukos lietuvių, anglų ir rusų k.).

A. Kajackas, A. Anskaitis. Investigation the Ability of Objective Measures of the Perceptual Speech Quality in Mobile Networks // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005. – No. 7(63). – P. 10–15.

There is continued the work of iQoS module creation. Possibilities to create special module which should measure de facto perceived speech quality (iQoS module) are investigated. The paper thoroughly reviews scientific publications about QoS measurement. After analysis of published works it is shown that currently prevailing methods of speech quality evaluation are not directly applicable for the evaluation of conversational speech quality because they make an assumption that distortions are stationary.

Research of single word distortions shows that in most cases even one lost frame worsens voice quality so that user satisfaction category changes to lower level. Deterministic relation between lost frame position in a word and quality degradation was not observed so it is possible to postulate that relation is stochastic in nature. Obtained empirical distributions of voice quality degradation confirm this assumption. The purpose of research continuation is to determine correlation between amount of erased frames and perceived quality level. The paper has formed a list of tasks which is necessary to solve in order to implement iQoS concept. Ill 5, bibl. 10 (in English; summaries in Lithuanian, English, Russian).

А. Каяцкас, А. Анскайтис. **Исследование возможностей объективного определения воспринимаемого качества речи в условиях мобильной связи // Электроника и электротехника. – Каунас: Технология, 2005. – № 7(63). – С. 10–15.**

Анализируются возможности создания специального модуля iQoS, назначение которого – оценивать качество фактически представленного сервиса. Анализ известных способов оценки качества речевых сигналов, выполненный по опубликованным материалам, показал, что все они работоспособны в условиях, когда искажения являются регулярными. По этой причине ни один из распространенных способов определения качества речевых сигналов непосредственно во время разговора не может быть применен. Исследование искажений отдельных слов показывают, что во многих словах даже один потерянный пакет качество речи переводит в более низкую категорию. Детерминированной взаимосвязи между потерянными пакетами и качеством речи не обнаружено. Остается предположение, что эта взаимосвязь – случайная. Продолжение исследований – поиск мер, устанавливающих взаимосвязь между воспринимаемым качеством речи и количеством потерянных пакетов. Сформулирован список задач, связанных с разработкой модулей. Ил. 5, библ. 10, (на английском языке; рефераты на литовском, английском и русском яз.).