ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2006. Nr. 4(68)
ELEKTRONIKA IR ELEKTROTECHNIKA

T 180 TELECOMMUNICATION ENGINEERING
TELEKOMUNIKACIJŲ INŽINERIJA

Investigation of Voice Frame Erasures in GSM

A. Kajackas, D. Gursnys
Department of Telecommunications Engineering, Vilnius Gediminas Technical University, Naugarduko str. 41, LT-03227 Vilnius, Lithuania; e-mail: algimatas.kajackas@el.vtu.lt, darius.gursnys@el.vtu.lt

Introduction

Packet loss is a key factor determining the quality of many multimedia applications. These applications experience degradation in quality with increasing packet loss and delay in the network. The impact of lost packets on voice quality has been investigated in many publications [1, 2]. Performed research shows that sparse lost packets only slightly worsen speech quality. Voice quality is deteriorated much more when lost packets are bursty. In such conditions parts of words or even complete words may be lost.

Performed researches show that GSM communication conditions are constantly varying. Lost packets are rare when user is in good communication conditions. Otherwise the number of lost packets is increasing and they become bursty. The probability of lost packets also depends on channel codes and consequently on parity bits. The experimental investigations of GSM frame-level error traces have been carried out at the reliable link layer (Radio Link Protocol – RLP). The models of lost packet series in GSM network were developed based on these RLP experimental results [3, 4].

In GSM channel coding of RLP packets is specified in [5] while channel coding of speech data is specified in [6]. The channel codes of RLP packets are not the same as channel codes of speech data. Speech data are protected with different number of parity bits compared to RLP data packets. Various voice codecs have different channel coding schemes. This is the main difference between voice and RLP packet encoding. Therefore the models of RLP lost packet series can't be applied for investigation of voice transmission systems.

In this paper we present a method for collection of lost speech frame traces in GSM system. The main idea of measurements is based on speech frame substitution (SFS) function which is implemented in many codecs. The basic substitution operation for lost frames is the repetition of the last good received frame. As a result, substituted and previous frames become highly correlated.

Voice codecs

The human speech consists of talkspurt and silence time intervals (Fig. 1) which are separated by the functional Voice Activity Detector (VAD) in voice encoder. GSM is a digital system, so speech signals, inherently analog, have to be digitized. There are several types of codecs used for voice coding: GSM 06.10 FR, GSM 06.20 HR, GSM 06.60 EFR, AMR-*n*.

In all GSM systems speech signal $v(t)$ is divided into $T = 20$ ms time intervals (voice frames)

$$v_i(t) = \begin{cases} v(t), \text{when } (i-1)T < t < iT, \\ 0, \quad \text{otherwhere.} \end{cases} \quad (1)$$

Each of voice frames is encoded $v_i(t) \Rightarrow c_i$.

The voice coder delivers to the channel encoder a sequence of data blocks c_i, where $i = 1, 2, \dots$ (Fig. 1).

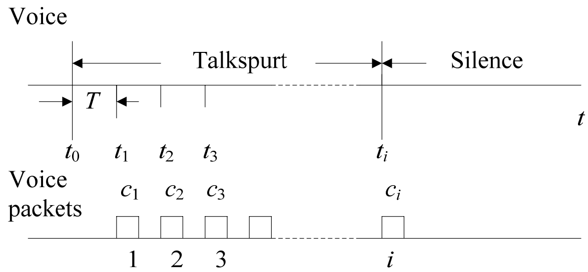

Fig. 1. Time diagram of voice encoding

Using different codecs on the same voice signal we get different data blocks c_i. Every pronounced word is represented by a particular set of data blocks $C = \{c_i\}_{i=1}^{I}$ which can be recomposed as vector

$$C = c_1 c_2 \dots c_I . \quad (2)$$

The number I of data blocks in the set depends on length of pronounced word and varies for different words. Voice signal contains many words, therefore each of them is coded with different vector C.

The voice talkspurt intervals are coded into data blocks $\{c_i\}$ and then sent to communication channel. In the silence intervals only particular silence descriptor (SID) frames are transmitted, therefore VAD disconnects the transfer of voice frames for this time period (this is called discontinuous transmission).

In GSM FR speech encoder one data block corresponds to one speech frame and contains 260 information bits, including 182 bits of class 1 (protected bits), and 78 bits of class 2 (not protected).

The class 1 bits are divided into the class 1a and class 1b. The most important class 1a bits are protected by a cyclic code (3 parity bits) and the convolutional code whereas the class 1b bits are protected by the convolutional code only [6]. Class 2 is without any protection.

When the channel decoder has decoded transmitted data block, the CRC is checked and if it is wrong, the whole speech frame is discarded. Then decoder uses speech frame substitution function SFS [7].

Marking of substituted frames

Normal decoding of lost speech frames would result in very unpleasant noise effects. In order to improve the subjective quality, lost speech frame $v_i^*(t)$ is substituted with either a repetition or an extrapolation of the previous good speech frame $v_{i-1}(t)$. This substitution is done so that it gradually will decrease the output level, resulting in silence at the output [7]

$$v_i^*(t) = \alpha v_{i-1}(t), \tag{3}$$

where $\alpha < 1$, i – frame number.

After the substitution, strongly correlated adjacent segments $v_{i-1}(t)$, $v_i^*(t)$ of the signal are created

$$R(v_{i-1}, v_i) = \frac{\left| \int_{(i-1)T}^{iT} v_{i-1}(t-T) v_i^*(t) dt \right|}{\sqrt{E_{i-1} \cdot E_i^*}} \to 1, \tag{4}$$

where

$$E_i = \int_{(i-1)T}^{iT} v_i^2(t) dt \tag{5}$$

is the energy of i-th signal segment.

According to eq. (3) the energy of substituted frame is lesser compared to the energy of last good frame. These two features of substituted frame - strong correlation and decreasing energy are used as metrics for detection of lost voice packets.

Experimental test-bed

The scheme of experimental test-bed, integrated in GSM network infrastructure, is showed in Fig. 2. It consists of stationary and mobile parts. The stationary part – desktop PC1 with acoustic system for communication status monitoring and mobile phone MS1 are connected to the PC1's sound card. This part is immovable during the measurements.

PC1 → MS1 → GSM network → MS2 → PC2

Fig. 2. Measurement scheme

The mobile part – laptop PC2, mobile phone MS2, connected to the PC2's sound card. PC1 is generating special test signal which simulates voice signal. The test signal is sent to microphone circuit of MS1. Then connection between MS1 and MS2 is established. Depending on the position of MS1 and MS2 in the GSM network, connection process involves one or two base transceiver stations and mobile switching center.

In order to investigate lost frames in particular radio channel, it is important to have one of the mobile phones (MS1) in exceptionally good communication conditions where all packets are transmitted without loss. The second mobile phone (MS2) should be located in the area of investigation.

The described test-bed measures lost frame traces in the radio channel between base transceiver station and MS2. The received test signal is passed to PC2. The special software analyzes this signal and creates the information file about good and erased (lost) packets.

The lost voice packet traces have to be used in order to evaluate voice quality and simulate codec behavior.

Test signal

The real speech signal has silence periods in which only SID frames are transmitted. From another point of view, VAD analyzes incoming signal and disconnects the transmission if this signal is not similar to the speech. In this case receiver generates comfort noise.

The proposed system can actually decide which decoded voice packets are correctly received and which are decoded after the substitution, only if it uses special test signal.

The spectrum of voice simulating signal $s(t)$ must be similar to the spectrum of the real speech. The duration of test signal period should be long enough, at least several 20 ms time intervals.

It is experimentally investigated, that VAD does not disconnect transmission when the test signal consists of sinusoid segments with different frequencies

$$s(t) = s(t, \Omega_i) = A \sin(\Omega_i t - \varphi_i), \tag{6}$$

where $(i-1)T < t < iT$, $i = 1, 2, 3, \ldots$.

The correlation between good received adjacent speech frames must be minimal to ensure correct detection of lost packets. The SFS function in the decoder after the substitution crates correlated voice signal which is the criteria for determining of the lost packet position in the trace. That is, when

$$R(s_{i-1}(t), s_i(t-T)) << 1. \tag{7}$$

Eq. (6) fulfils this condition when each pair of any sinusoid segments has following frequencies $\mathbf{\Omega}_i$ and $\mathbf{\Omega}_{i+1}$

$$|\Omega_{i+1} - \Omega_i| = \frac{2\pi}{T} k, \tag{8}$$

where k – any integer number.

The frame of test signal and MS frame are not timed. The frame position in test signal is shifted by $\tau = t_x - t_0$ compared to MS frame. In this case the test signal passes minimum correlation condition (7), when the frequencies of sinusoid segments Ω_i are multiples of $\Omega_0 = 2\pi / T$ and

$k >> 1$ in eq. (8). In the GSM system we have $F_0 = \Omega_0 / 2\pi = 50$Hz.

From codec description [8] we notice, that frequency response has less variation in the frequency range from 250 Hz to 1300 Hz. To avoid distortions of amplitude the frequency band of test signal must be limited to this range.

The test signal which corresponds to conditions mentioned above can be expressed:

$$s(t) = s_6(t) + s_1(t-T) + s_7(t-2T) + \\ + s_2(t-3T) + ... + s_5(t-10T) + s_{11}(t-11T), \quad (9)$$

where $s_i(t) = \begin{cases} \sin(\Omega_i \cdot t), & t \in [0,T], \\ 0, & t \notin [0,T], \end{cases}$

$f_i = 250 + 100 \cdot (i-1)$ Hz, i = 1, 2, …, 11, T = 20ms.

The test signal (9) has following features:

- The energy E_i of sinusoid segments is constant;
- The adjacent frames of the test signal have minimum correlation.

These features determine the analysis algorithm of received test signal. There are two successive analysis windows and each of them has duration of 20 ms. They move over received test signal and the correlation between segments of signals which are inside these windows is calculated:

1. If correlation is small, the frame was received without errors;
2. If correlation is strong, the frame was lost. We assume that this is the first lost frame;
3. If there are more consecutive lost frames, they can be detected by observing strong correlation and decreasing energy.

In Fig. 3 theoretically calculated normalized correlation is showed, when one frame at $t_1 = 0.04$ s, series of two frames at $t_2 = 0.22$ s and series of three frames at $t_3 = 0.42$ s were lost. At these time moments the correlation coefficient reaches one. The correlation between adjacent frames in other time periods is considerably smaller.

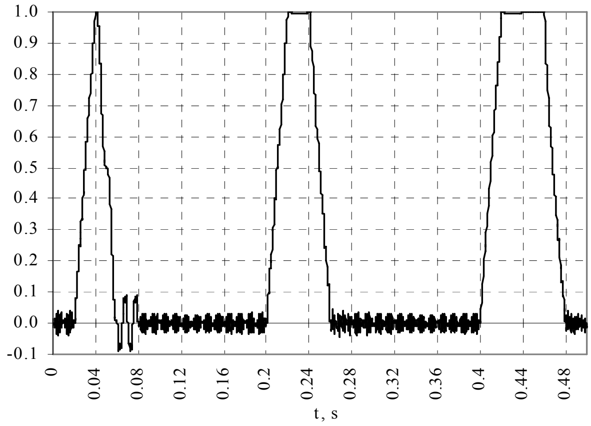

Fig. 3. Theoretically calculated standardized correlation

In Fig. 4 normalized correlation of test signal that was propagated through the real GSM system is shown. In the time interval where one frame was substituted, correlation coefficient reaches one as in theoretical calculation case.

The increased correlation in other time intervals (compared to the theoretical calculation) is a result of signal distortions in the real GSM speech codec and voice input-output channels.

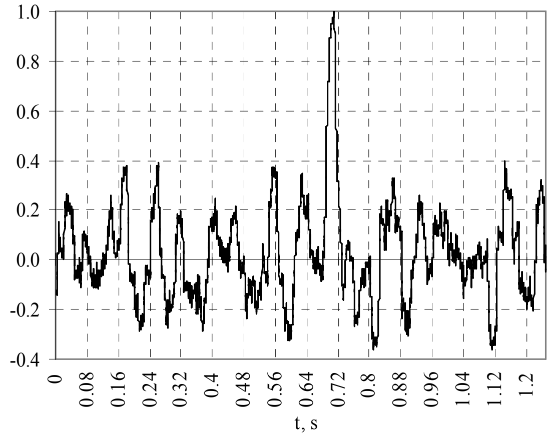

Fig. 4. Experimentally estimated standardized correlation

Experiments and results

Characteristics of the radio channel are varying and depend on noise, frequency hopping, interference with other transmitting stations, etc. Single frame can be lost at any time, but regular losses of frames (significant speech quality degradation) occur when the received signal level is below – 80 dBm. We are interested in the voice quality when end-user is in bad communication conditions and therefore, the place of experiment must be chosen respectively.

The measurements were performed in a wooded area to ensure low signal level. In Fig. 5 preliminary frame erasure trace is showed, which is collected by moving (at 40 km/h) from a forest (bad communication conditions) to a small town (good communication conditions).

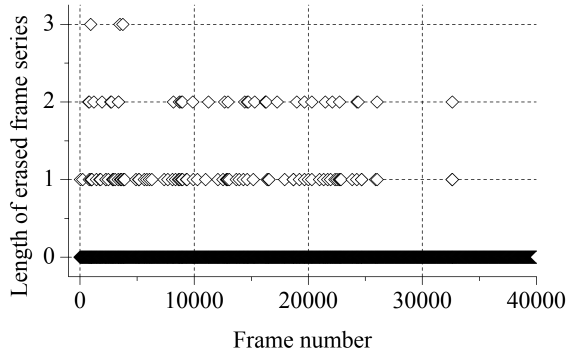

Fig. 5. Experimentally collected frame erasure trace

The trace is 40000 frames long and consists of 1, 2 and 3 successively erased frame series (null length means good received frame).

Conclusion and future work

In this paper, we have presented a framework for measurement and collection of voice frame erasure traces in GSM network. For measurements we created special test-bed system. The main idea of measurements is based on speech frame substitution function which is implemented in the codec. The proposed scheme of experimental test-bed, integrated in GSM network infrastructure, consists of stationary and mobile parts. The stationary part generates the special test signal. The measurements of lost frames are performed in mobile part using end-user equipment.

The collected traces of erased frames will be used for development of models of lost voice frame series and for estimation of individual voice quality.

The frame trace collection method was developed for GSM system, but it can be used for investigation of 3G and other communication systems too.

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

References

1. **Sun L., Wade G., Lines B., Ifeachor E.** Impact of Packet Loss Location on Perceived Speech Quality // Proceedings of 2nd IP-Telephony Workshop (IPTEL '01) Columbia University. New York. –P. 114–122, 2-3 April, 2001.
2. **Clark A. D.** Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality, IPTEL'2001. - New York, April, 2001. - P. 123-127.
3. **Ji P., Liu B. Towsley D., Ge Z., Kurose J.** Modeling frame-level errors in GSM wireless channels. GLOBECOM 2002. – Vol.3. – P. 2483–2487.
4. **Konrad A., Zhao B. Y., Joseph A. D., Ludwig R..** A Markov-based Channel Model Algorithm for Wireless Networks. Wireless Networks. – Vol. 9, No. 3. – P. 189–199, 2003/05.
5. ETSI GSM Technical Specification 04.22, GSM Radio Link Protocol for data and telematic services, Version 6.1, November 1998.
6. ETSI GSM Technical Specification 05.03. Channel coding, Version 8.9.0, April 2005.
7. ETSI GSM Technical Specification 06.11, Substitution and muting of lost frames for full rate speech channels, Version 8.0.1, November 2000.
8. ETSI GSM Technical Specification 06.10, Full rate speech; Part 2: Transcoding, Version 4.3.0, June 2001.

Presented for publication 2006 02 10

A. Kajackas, D. Gursnys. Investigation of Voice Frame Erasures in GSM // Electronics and Electrical Engineering. – Kaunas: Technologija, 2006. – No. 4(68). – P. 47-50.

The experimental investigation of individual voice quality is carried out in common GSM communication system. There are scientific publications in which methods of GSM frame-level error measurements are presented. These works involve data transfer channel, RLP packet traces. In this paper we take note, that in GSM cellular communication system voice and RLP packets are coded with different channel codes, voice packets have different calculation of parity bits. We have presented a framework for measurements and collection of voice frame erasure traces in GSM network. For measurements we constructed special test-bed system. The main idea of measurements is based on speech frame substitution function which is implemented in the codec. The scheme of experimental test-bed, integrated in GSM network infrastructure, consists of stationary and mobile parts. The stationary computer generates the special test signal. The measurements of lost frames are performed in mobile part by using end-user equipment. The collected traces of erasure frames will be used for development of models of lost voice frame series and for estimation of individual voice quality. The preliminary voice frame erasure trace analysis is presented. Ill. 5, bibl. 8 (in English; summaries in English, Russian and Lithuanian).

А. Каяцкас, Д. Гуршнис. Исследование потерянных в GSM сети потоков фреймов голоса // Электроника и электротехника. – Каунас: Технология, 2006. - № 4(68). – С. 47-50.

На примере распространенной GSM системы проводится экспериментальное исследование способов оценки индивидуального качества передачи голоса. В научной печати имеются работы, в которых описан способ определения ошибочных пакетов в GSM сети, представлены результаты измерений и на их основе разработанные модели. Однако упомянутые данные получены для канала передачи данных на основе RLP протокола. В данной работе обращается внимание на то, что в GSM системе пакеты голоса кодируются не теми канальными кодами как данные, иначе определяются контрольные биты. Предложен метод определения серий пакетов, стертых при передаче по GSM сети. Он основан на специфике восстановления стертых пакетов в декодере источника. Измерения производятся с применением оборудования пользователя, поэтому полученные данные пригодны для оценки индивидуального качества сервиса. Для экспериментов пришлось разработать специальную измерительную систему, создать специальный измерительный сигнал. Сигнал реального голоса из-за присущих интервалов молчания для измерений непригоден. Представлены данные проведенных измерений. Ил 5, библ. 8 (на английском языке; рефераты на английском, русском и литовском яз.).

A. Kajackas, D. Guršnys. GSM tinkle prarastų balso paketų srautų tyrimai // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2006. – Nr. 4(68). – P. 47-50.

Individualios balso kokybės eksperimentiniai tyrimai atliekami paplitusios GSM ryšio sistemos pagrindu. Mokslinėje spaudoje yra darbų, kuriuose aprašyti GSM paketų lygmens klaidų matavimo būdai, pateikti tyrimo rezultatai bei ištrintų paketų serijų modeliai. Minimi darbai remiasi duomenų perdavimo kanalu, RLP paketų srautais. Šiame darbe atkreiptas dėmesys, kad GSM korinio ryšio sistemoje balso paketai yra koduojami kitokiais kanalo kodais nei duomenys, kitaip skaičiuojami kontroliniai bitai. Pasiūlytas GSM sistemoje ištrintų balso paketų srautų matavimo metodas. Jis pagrįstas dekoderio prarastų balso paketų pakeitimo algoritmo ypatumais. Matavimai atliekami naudojant vartotojo įrangą, todėl gauti duomenys tiesiogiai tinka individualiai paslaugos kokybei vertinti. Matavimams teko sukurti specialią sistemą, specialų balsą imituojantį signalą. Įprastinis balso signalas dėl jame pasitaikančių tylos intervalų nėra tinkamas. Pateikti pirminių matavimų rezultatai. Il. 5, bibl. 8 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).