Implementation of a Steganographic Algorithm in an Internet VoIP Phone

Damian Bachmat and Zbigniew Piotrowski
Faculty of Electronics, Military University of Technology
Gen. S. Kaliskiego 2 Str., Warsaw 00-908, Poland

Abstract— The paper describes implementation of the steganographic method based on the least significant bits (LSB) in a program Internet phone VoIP. In the developed program Internet phone VoIP a library PJSIP has been used. Selected technology — LSB is characterized by easiness of implementation and high data rate and perceptual transparency of a signal with embedded steganographic sequence. In case of the LSB method a change in the least significant bits in a datagram takes place after quantizer of an analog-digital converter circuit. This solution enabled simple implementation of the embedding algorithm by modification of a source code of a codec G.711. The way of the method implementation protects data from errors since mechanisms of the error control and correction of the RTP protocol monitored correctness of sent data. One of the advantages of the method is the fact that for determined number of the least significant bits it does not cause an audible speech degradation and therefore, modified speech signal is perceptually transparent for a subscriber at the receiver side. The paper includes results of tests of developed steganographic phone, among others, subjective audible tests based on signal fidelity estimation standard — ITU-R BS 1116-1, tests of minimal, required time of transmission and integrity tests of steganographic sequence.

1. INTRODUCTION

Voice over IP (VoIP) or in other words Internet telephony is one of “IP world” services, which changes the picture of all telecommunication. It is a real time service enabling the user to do audio or audio-video connections within the packet network that is IP network. So-far there have been a lot of applications, also mobile ones, enabling to make use of this technology. A huge potential is in PJSIP libraries which are available as so-called open source for each user. The asset of making use from PJSIP libraries stack is the fact that it is possible to have access to the source codes of the component elements out of which the customer software VoIP is built. The access gives the possibility for modification of the library source codes so that they would be adapted to user’s requirements. As the Internet telephony popularity is growing in the telecommunication, there appear new concepts for using this technology for safety data transmission, and it becomes the target for applying different marking methods and steganographic techniques. VoIP technology is also widely used in the tele-informative elements in the military connection systems. Steganography makes use of many different features of VoIP technology for the purposes of embedding secret message in the carrier which in this case are IP packets. To hide data, the network steganography uses the network protocols and the relations between them. From the steganography nature results that it should not be visible for a third party user in any way, but it should not deteriorate the quality of services offered by VoIP. Moreover, one of the most difficult requirements to be met is the data resistance to degrading impacts. The packet losses or lossy encoding, frequently occurring in VoIP technology, have a degrading impact on the hidden data. Additionally, it is required from the electronics equipment (e.g., VoIP phones) to meet the electromagnetic compatibility prerequisite [1–9]. In the article there is presented the implementation of the technique modifying the least significant bit (LSB) in the programming Internet telephone based on stack of PJSIP libraries, algorithm of embedding and extraction of message was realised by the modification of encoder source codes and standard decoder G.711 with compounding according to the encoding curve A (A-law).

2. CHARACTERISTICS OF PROJECTED SOLUTION

The accepted solution assumes the embedding of hidden message in the least significant bit of the first eight sound samples in the signal frame. The embedding process occurs just after the encoder on the samples encoded according to the standard G.711. Each 120 sound samples go to the loading space of RTP protocol, where the header details are added. So obtained packet goes to further lower layers and is sent to the addressee with whom VoIP talk is conducted. The implemented algorithm takes the data form text file byte after byte and embeds each of them in the separate packet. The packets of RTP protocol transferring the speech samples with the length of 20 ms make the channel with the bit rate ca. 400 bps.

2.1. Embedding Process

The principle of speech samples transmission as well as all mechanisms supporting the error-free data recovery at the receiver side, indicate the favourable place for implementation of the message bit embedding algorithm in the data flow just after the encoder G.711. In the general overview this technique consists in the replacement of the least significant bit of the encoded samples of speech signal at the level of application layer, before placing the data to the loading space of RTP protocol. The encoder in its operation works on the data transcription from the input buffer — 16 bit samples, to output buffer — 8 bit encoded samples. For encoding each 16 bit sample there is called the function linear2alaw (pj_int_16_t*), which returns 8 bit samples with no mark. Algorithm of the encoder G.711 at the first step takes 16-bit samples, calls the encoding function and then transcripts the encoded values to output buffer. The implemented marking technique changes the sample value just before their transcription to output buffer, so as to leave the message bits on the least significant bit. The placing is realised by means of suitable operators of binary shifts as well as multiplication with proper mask, because the smallest data unit in C language is 1 byte. The embedding process occurs on the first 8 bits of each frame. Each byte of the message is per one frame with duration of 20 ms, which is given by the general download rate of the method equal to:

$$K_w = \frac{8\,\text{bit}}{20\,\text{ms}} = 400\,\text{b/s} \tag{1}$$

2.2. Extraction Process

At the receiver side the decoder source code was also modified properly to save the data from the last layer (LSB) bits of encoded samples just before their processing in the decoder. The values coming from the source buffer are multiplied by mask which separates the least significant bit from the whole value of the sample. This bit is then binary shifted to the right by the set number of places depending on the sample number, because n-th bit of undisclosed message is transmitted by n-th encoded sample. After writing down the least significant 8 bits from the first samples, and "locating" them in byte, it is stored to the file. The whole operation is repeated after next calling the decoder function. The decoder function uses the internal decompression function linear2alaw (pj_int_8_t*), which is called for each speech sample.

3. DESCRIPTION OF LSB METHOD AND ITS VARIATIONS

The modification of the least significant bit (LSB) is the simplest steganographic technique providing the large payload. In this technique the data is hidden in the least significant bit (bits) of the sound signal samples.

The influence of the value change of one bit in the sample on the signal quality is neglected. However, the change of the least significant bits causes the formation of some noise which is acceptable as long as its value is below the perceptive threshold. While increasing the number of changed bits we cause formation of the noise with higher power. If the noise value exceeds the threshold value, it becomes detectable by Human Auditory System. The use of bigger number of bits per sample for embedding the data increases the payload while decreasing the perceptive transparency at the same time.

4. IMPLEMENTATION OF THE PROGRAMMING VOIP INTERNET TELEPHONE

PJSIP is the stack of libraries written in C language which is characterised with small size and great capacity. This stack is distributed on GNU license of General Public License. In the packet available on the page [10] there are examples of VoIP client's scripts, by means of which after compilation it is possible to make connection within local network or by using Internet after log-in. In the application using the stack of PJSIP libraries datagrams/packets are made by means of proper modules. Hence, e.g., RTP datagram is made by the header concatenation as well as the data field in the module named RTP, and TCP packet is made by TCP module. In additional the latter is the connection-oriented protocol, so its module must provide additional functions for connecting and keeping the session. "Raw" data coming from audio stream (microphone port) is encoded according to the chosen speech encoder, and then it is introduced to the transport spaces of next protocols which add their own control information.

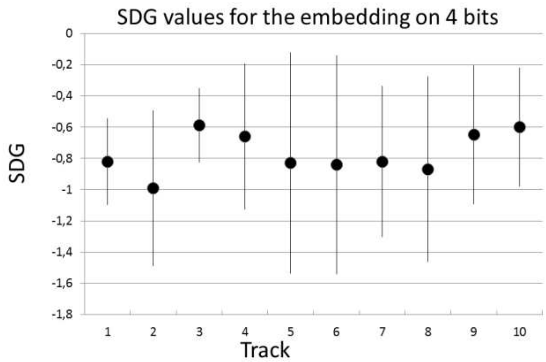

Figure 1: SDG values for 4 least significant bits used for the embedding.

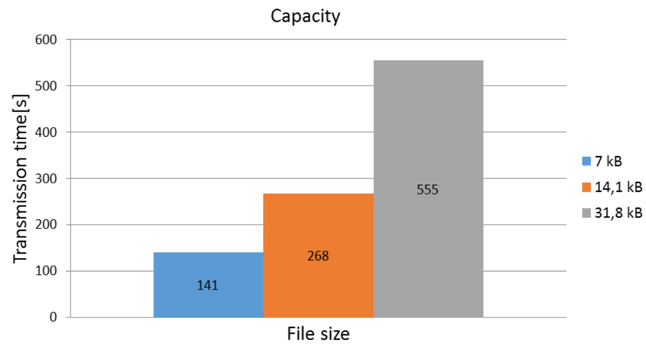

Figure 2: Embedding time depending on the file size.

5. RESULTS FROM EXPERIMENTS

In the first place of conducting the tests, the correctness of the algorithm operation was checked, and then some amendments were introduced. The whole test area was divided into three main areas:

- Subjective listening tests — ITU-R BS 1116-1.
- Tests of minimal required time for transmission.
- Watermark integrity tests.

Three independent tests were conducted, first of them while using the algorithm in MATLAB environment, the second and third one directly on the operating application, at the connection between two stations.

5.1. Tests of Minimal Required Time for Transmission

The capacity tests of the embedding algorithm were conducted in order to determine the minimum length of talk, which enables to send the file with set length. The algorithm theoretical download rate was defined as 400 bps. You should remember, however, about the mechanisms such as VAD, which may decrease this value during the talk. The dependence of the embedding capacity is strictly related to so-called *fill factor*. This factor defines silence duration (inactive frames) in relation to the total duration time of talk. The fill factor value equal to 0.2 during 60 seconds talk means that only in 48 seconds of talk you may embed the mark. The real rate of data hidden transmissions in the talk signal drops to the value of 320 b/s. The fill factor of typical phone call is ca. 0.4 [11].

All tests were conducted for the fill factor equal to 0.95, whereas the embedding process was realised on the least significant bit of the first eight samples encoded according to G.711.

5.2. Watermark Integrity Tests

During the conducted tests, there were no losses in the transmitted packets and therefore any falsification in the watermark structure. Information on statistical parameters of the talk conducted was presented via applications after finishing each talk.

6. CONCLUSIONS

The aim of the work was to implement the data embedding algorithm in the programming Internet telephone VoIP, operating based on the stack of PJSIP libraries. The aim was reached and the effectiveness tests of the developed algorithm were presented. The implementation of the algorithm in the programming telephone was preceded with the execution of marking program in Matlab environment, for earlier selected method. The work proves that there is possibility for changing bits of the data stream at the datagram level so that they could be used for transmitting the watermark information. The selected technique for marking is characterised with the implementation easiness, large download rate in comparison to the steganographic methods operating on the signal in bit form just after the quantizer of the analogue-digital converter system. The way of method implementation protects the data from errors occurring because the mechanisms of control and error correction of the individual layers are monitoring the data integrity. One of the method advantage is the fact that it does not cause large signal degradation of the transmitted speech, so it is transparent. The transmission of data — mark bits takes place with no noticeable changes in the talk quality by the users.

REFERENCES

1. Piotrowski, Z., L. Zagoździński, P. Gajewski, and L. Nowosielski, "Handset with hidden authorization function," *European DSP Education & Research Symposium EDERS 2008, Proceedings*, 201–205, Published by Texas Instruments, 2008, ISBN: 978-0-9552047-3-9.
2. Nowosielski, L., M. Wnuk, and C. Ziółkowski, "Interlaboratory tests in scope of measurement of radio disturbance," *European Microwave Week 2009, EuMW 2009, Science, Progress and Quality at Radiofrequencies, Conference Proceedings — 39th European Microwave Conference, EuMC 2009*, Article No. 5296286, 288–291, 2009.
3. Nowosielski, L., J. Patka, and M. Siłaczuk, "Modelling of electromagnetic wave propagation with the use of the ray-tracing method," *PIERS Proceedings*, 2701–2705, Guangzhou, China, Aug. 25–28, 2014.
4. Nowosielski, L. and C. Piotrowski, "Honeycomb ventilation grill shielding effectiveness measuring methodology," *PIERS Proceedings*, 2692–2696, Guangzhou, China, Aug. 25–28, 2014.
5. Nowosielski, L. and M. Wnuk, "Compromising emanations from USB 2 interface," *PIERS Proceedings*, 2666–2670, Guangzhou, China, Aug. 25–28, 2014.
6. Kaszuba, A., R. Checinski, M. Kryk, J. Lopatka, and L. Nowosielski, "Electromagnetically shielded real-time MANET testbed," *PIERS Proceedings*, 2706–2710, Guangzhou, China, Aug. 25–28, 2014.
7. Nowosielski, L., and J. patka, "Measurement of shielding effectiveness with the method using high power electromagnetic pulse generator," *PIERS Proceedings*, 2687–2691, Guangzhou, China, Aug. 25–28, 2014.
8. Gruszczynski, M., M. Wnuk, and L. Nowosielski, "Multisystem microstrip antenna for mobile communications," *IEEE Antennas and Propagation Society, AP-S International Symposium (Digest)*, Article No. 6349179, 2012.
9. Nowosielski, L., M. Wnuk, and C. Ziółowski, "Interlaboratory tests in scope of measurement of radio disturbance," *European Microwave Week 2009, EuMW 2009: Science, Progress and Quality at Radiofrequencies, Conference Proceedings — 39th European Microwave Conference, EuMC 2009*, Article No. 5296286, 288–291, 2012.
10. http://www.pjsip.org/.
11. Beritelli, F., S. Casale, and G. Ruggeri, "Performance evaluation and comparison of ITU-T/ETSI voice activity detectors," *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings, (ICASSP '01)*, 1425–1428, 2001.