Computer Communications 70 (2015) 95–108

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Improved adaptive partial-matching steganography for Voice over IP

Hui Tian [a,*], Jie Qin [a], Shuting Guo [a], Yongfeng Huang [a,b], Jin Liu [a], Tian Wang [a], Yonghong Chen [a], Yiqiao Cai [a]

[a] College of Computer Science and Technology, National Huaqiao University, Xiamen 361021, China
[b] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 28 July 2014
Revised 20 August 2015
Accepted 20 August 2015
Available online 29 August 2015

Keywords:
Steganography
Covert communication
Voice over IP
Partial matching
Matrix encoding

ABSTRACT

In Voice over IP scenarios, adaptive partial-matching steganography (APMS) is an effective approach to balance steganographic transparency and bandwidth, in which the strategies involved, however, have not achieved the optimal steganographic performance yet. Therefore, in this paper, we try to improve APMS from three aspects. First, when the current partial similarity value (PSV) falls between the two given substitution thresholds, the embedding process is guided with unequal probabilities increased with the PSVs instead of identical probabilities, which enhances the embedding chances of cover parts with larger PSVs. Second, for cover parts considered unsuitable for information hiding in the original strategies, the matrix encoding strategy is employed to conceal secret messages with no more than one bit changed, which improves cover usage rate while maintaining high embedding efficiency. Third, in the APMS strategy with two flag bits, one more encrypted form of the secret message is added into the comprehensive evaluation of partial matching, which makes the embedding process more efficient. The proposed strategies are evaluated with ITU-T G.711 (A-law) as the codec of the cover speech, and compared with the original ones and some existing approaches. The experimental results demonstrate that the improved strategies can strike excellent balances between the steganographic transparency and bandwidth, and significantly outperform both the former ones and other previous approaches. Furthermore, the results of detection tests show they have outstanding performance of steganalysis resistance.

1. Introduction

Steganography, which is an art and science of information hiding, has captured the imagination of researchers for long years [1,2]. Particularly, in the past few years, it has been developed from limited applications to extensive deployments, and the steganographic covers have been also extended from images to almost all multimedia. As a dynamic steganographic cover, Voice over Internet Protocol (VoIP) or named IP telephony, has attracted significant attention in recent years [3–5]. It possesses three main advantages over traditional storage media. First, the real-time nature of VoIP provides better security for secret messages by virtue of its instantaneity, because it does not give eavesdroppers sufficient amount of time to detect possible abnormity due to hidden messages. Second, VoIP can be considered a multi-dimensional carrier in that both the packet protocol headers and the payload data can be used to hide data. Third, the length of a VoIP conversation is dynamic and variable, so it can supply enough cover data according to the requirement of covert communication and also make VoIP-based steganography even harder to detect [5]. In virtue of these advanced characteristics, VoIP-based steganography provides an effective solution for secure covert communication [5].

Many researchers have carried out useful research on steganography over VoIP, and various approaches have been put forward. We would briefly introduce the related work in the next section. The interested reader could also refer to a recent exhaustive survey given by Mazurczyk [5] for more details. Among the existing techniques, Least-Significant-Bit (LSB) steganography is most popularly employed. Typically, Dittmann et al. [6], Wang and Wu [7] and Aoki [8] suggested different LSB approaches based on ITU-T G.711 speech. Su et al. [9] and Liu et al. [10] respectively found that the LSBs of fixed codebook parameters in ITU-T G.729a and G.729 have the best transparency for information hiding. Xu and Yang [11] analyzed the redundancy of each parameter of ITU-T G.723.1 codec and pointed out that the lowest five bits of the LSP VQ indices can be used to hide secret information.

Recently, more and more researchers have begun to pay attention to how to improve the steganographic transparency and security (undetectability) of LSB approaches. For example, Huang et al. [12] attempted to introduce LSB matching steganography to the covert

* Corresponding author. Tel.: +86 18959267996; fax: +86 592 6162556.
E-mail address: htian@hqu.edu.cn (H. Tian).

http://dx.doi.org/10.1016/j.comcom.2015.08.011
0140-3664/

communication based on VoIP to enhance the transparency; Wei et al. [13] proposed an adaptive VoIP steganography approach to improve the embedding security via three measures, namely dynamically choosing multiple bits based on the VoIP vector value, dynamically changing embedding intervals and modifying neighbor bits to offset the distortion; Miao and Huang [14] presented an adaptive steganography scheme to improve the security and transparency by choosing lower embedding bit rates in the flat blocks and higher embedding bit rates in the sharp blocks; Liu et al. [15] adopted least-significant-digits rather than LSBs to hide secret messages, which can increase around 30% of steganographic capacity, and induce less distortion given the same capacity.

In our previous work, we presented an adaptive partial-matching steganography (APMS) scheme, in which the notion of Partial Similarity Value (PSV) is introduced to evaluate the partial matching between covers and secret messages to be embedded. Moreover, we suggested two concrete strategies [16], respectively called APMS-1 and APMS-2 for the convenience of description in this paper. [1] APMS-1 evaluates the partial similarity between the covers and the encrypted secret messages. APMS-2, however, considers the partial similarities between not only covers and encrypted messages but also covers and original messages. By properly setting a low threshold and a high threshold of PSV, both APMS-1 and APMS-2 can adaptively balance steganographic transparency and bandwidth. Nevertheless, our recent observations through research and experiments suggest that the original strategies do not achieve the best performance for three reasons. First, when the calculated PSV is between the two substitution thresholds, whether the cover part is replaced depends on the current output of an m sequence, which means that the embedding probability of each cover part in this scenario is approximately 0.5. That is a very simple measure to strike a possible tradeoff between steganographic transparency and bandwidth without regard to the difference of the PSVs. Apparently, it is more reasonable to make the embedding probability of each cover part increased with the corresponding PSV. Second, when the calculated PSV is less than the low threshold, or between the two thresholds and the current output of the m sequence is "0", the corresponding cover part is not used to hiding any secret messages, which lowers the cover usage rate. Third, APMS-2 does not make full use of flag bits. The adopted two flag bits can indicate four states, but only three of them are employed. Thus, we can consider introducing one more transformation of the secret message part (e.g. another encrypted form) to participate in the evaluation of the partial matching, which may further improve steganographic performance.

Therefore, we are motivated to propose the improved APMS strategies, which can provide better steganographic performance. The main improvements include:

- When the calculated PSV falls between the two given substitution thresholds, unequal probabilities are adopted to guide the embedding process. The probabilities are uniformly distributed between 0 and 1, and increased with the PSV. Consequently, the larger the PSV is, the more embedding chance the corresponding cover part has.
- For the "unused" cover parts, the matrix encoding strategy (MES) [4,17] is introduced to embed secret messages with no more than one bit changed, which thereby improves cover usage rate while maintaining high embedding efficiency.
- To take full advantage of all four states represented by the flag bits in APMS-2, another encrypted form of secret messages using a new m sequence is added into the comprehensive evaluation of the partial matching, which makes the embedding process more efficient.

[1] The naming of the two approaches also implies the number of their adopted flag bits, namely, APMS-1 and APMS-2 respectively employ one flag bit and two flag bits.

The proposed strategies are evaluated with ITU-T G.711 (A-law) [18,19] as the codec of cover speech in StegVoIP [20,21] that is a prototypical covert communication system based on VoIP. The experimental results demonstrate that the improved strategies can strike excellent balances between steganographic transparency and bandwidth, and outperform the original ones.

The remainder of this paper is organized as follows. Section 2 reviews the related work regarding steganography based on VoIP. To make this paper self-contained, Section 3 first introduces the original APMS schemes for VoIP. The improved APMS strategies are further proposed in Section 4, which are followed by their evaluation and experimental results that are presented in Section 5. Finally, concluding remarks are made in Section 6.

2. Related work

Existing VoIP-based steganography methods evolve from two distinct origins, i.e., covert channels over various network protocols and steganography techniques based on traditional multimedia like image, audio and video, so they can be naturally classified into two major categories, namely, protocol steganography approaches that utilize the specific protocols of VoIP as the carriers and speech steganography approaches that employ the digital speech signal as the carrier.

Generally, protocol steganography has two principal techniques. One embeds the secret messages into redundant fields of protocol headers in subtle ways. The other encodes secret messages by varying packet rates, which is equivalent to modulating the inter-packet times. In the VoIP scenario, the former has been popularly implemented in IP [4,22,23], RTP/RTCP [22,24,25,26] and SIP with SDP [27]. In contrast, the application of the latter is comparatively rare, possibly considering the QoS (Quality of Service) of VoIP. However, Mazurczyk and Lubacz [28] suggested a steganography named LACK (Lost Audio Packets Steganography) for VoIP, which conceals secret messages into the payloads of some intentionally delayed packets. To increase the complexity of steganalysis, Hamdaqa and Tahvildari [29] further provided a reliability and fault tolerance mechanism for LACK based on a modified (k, n) threshold at the cost of a portion of the steganographic bandwidth. They are creative attempts, which may stimulate more steganography ideas using the modulation of inter-packet times.

Compared with the protocol steganography, speech steganography has attracted much more attention of the research community, which mainly includes the following approaches.

2.1. Quantization index modulation (QIM) steganography

The basic principle of the QIM steganography is to utilize the redundancy in quantization stage and embed secret messages by modifying the quantized values. It is first proposed by Chen and Wornell [30] for information hiding in images, and introduced to achieve steganography in low bit-rate speech streams by Xiao et al. [31]. Instead of random division of the codebook, Xiao et al. [31] proposed a complementary neighbor vertices (CNV) algorithm to reasonably divide the codebook based on the relations between codewords, and further presented an improved QIM approach, called CNV-QIM steganography. The results of experiments based on iLBC and G.723.1 codecs show that it only slightly decreases the speech quality, and provides a steganographic bandwidth of 100 bit/s. However, due to adopting a static codebook division, CNV-QIM steganography is insecure. Moreover, it is vulnerable to a state-of-the-art steganalysis presented by Li et al. [32] exploiting unbalanced and correlated characteristics of the codeword distribution. Therefore, we were motivated to present an improved QIM steganography (Sec-QIM) [33], which introduces a key-based codebook division strategy to improve security, and incorporates random position selection and matrix encoding strategy to resist the steganalysis at the expenses of the embedding