International Journal of Digital Crime and Forensics
Volume 11 • Issue 4 • October-December 2019

Steganalysis of AMR Based on Statistical Features of Pitch Delay

Yanpeng Wu, Xiamen Meiya Pico Information Co., Ltd., Xiamen, China
Huiji Zhang, Xiamen Meiya Pico Information Co., Ltd., Xiamen, China
Yi Sun, Xiamen Meiya Pico Information Co., Ltd., Xiamen, China
Minghui Chen, Xiamen Meiya Pico Information Co., Ltd., Xiamen, China

ABSTRACT

The calibrated matrix of the second-order difference of the pitch delay (C-MSDPD) feature has been proven to be effective in detecting steganography based on pitch delay. In this article, a new steganalysis scheme based on multiple statistical features of pitch delay is present. Analyzing the principle of the adaptive multi-rate (AMR) codec, the pitch delay values in the same frame is divided into groups, in each of which, a pitch delay has a closer correlation with the other ones. To depict the characteristic of the pitch delay, two new types of statistical features are adopted in this article. The new features and C-MSDPD feature are together employed to train a classifier based on support vector machine (SVM). The experimental result shows that, the proposed scheme outperforms the existing one at different embedding bit rates and with different speech lengths.

KEYWORDS

Adaptive Multi-Rate Speech, C-MSDPD, C-PDDPD, Markov Transition Probability, MDPD, Pitch Delay, Speech Steganalysis, Speech Steganography, Support Vector Machine

1. INTRODUCTION

Steganography is a security technique that utilizes digital files or network protocols to embed secret messages (Provos & Honeyman, 2003). Compared with traditional security technology, steganography has the advantage of concealment, which will make it undetectable for attackers. Accordingly, steganography can be applied to covert communication.

The research of steganography is mainly concentrated in images. Content-adaptive steganographic methods are the most secure schemes in recent years. Compare with traditional steganographic methods, content-adaptive steganographic methods can provide better security to resist the statistical detection. Filler, Judas and Fridrich (2010) developed a framework with Syndrome-Trellis Codes (STCs), which could be used for minimizing additive distortion between cover and stego images. There are many algorithms implemented by STCs, such as highly undetectable stego (HUGO) method (Bas, 2010), spatial-universal wavelet relative distortion (S-UNIWARD) method (Holub and Fridrich, 2013) et al. To enhance the security of covert communication, Sedighi, Cogranne and Fridrich (2016) proposed a method by using an estimated multivariate Gaussian cover image model to minimize the statistical detect ability. Content-adaptive image steganographic methods increase the difficulty of detection, but steganalysis technologies also make some progress in these years.

DOI: 10.4018/IJDCF.2019100105

Rich-model based steganalysis is the modern methods for stego images detection. Fridrich and Kodovsky (2012) first design a rich-model based steganalysis method for images steganography. In their method, high dimensional features and ensemble classifier are employed to enhance the detection accuracy. Then Goljan, Fridrich and Cogranne (2014) designed an extension of the spatial rich model for color images. To detect the content-adaptive image steganographic methods, Denemark, Boroumand and Fridrich (2016) design some high order features by the knowledge of the selection channel. Luo et al. (2016) analysis the character of STCs and designed a steganalysis method for HUGO steganography. The method can not only detect the stego images but also extract the secret messages. Recently, Liu, Yang and Kang (2017) proposed a steganalysis method combines convolutional neural network with rich-models and ensemble classifiers. Experimental results show that the method has better performance than the state-of-the-art one. However, due to the structure and character differences between the parameters of image and speech, it is hard to directly employ the steganalysis methods on image to achieve effective detection for speech steganography.

In recent years, with the development of mobile network and smart phone, Voice over IP (VoIP) has become widely employed by mobile communication such as network telephone or instant message. Compared with other carriers for covert communications, VoIP has obvious advantages, for example, its large volume for embedding data could provide high covert bandwidth, and its instantaneity could provide real-time communication environment. Therefore, there are many works have been done for the steganography based on VoIP. As a standard of speech compression, AMR is widely employed by 3G, 4G systems or VoIP in speech services. Due to its great performance on speech compression, AMR is adopted as the file format for many communication applications such as intent message or speech recorder on smart phones. Therefore, the steganography of AMR speech codec has attracted extensive attention in recent years.

In general, the steganography based on VoIP can be divided into two classes (Mazurczyk, 2013): The first one carries out information hiding by modifying the protocol of VoIP (Huang, Yuan, Chen & Xiao, 2011; Jiang, Tang, Zhang, Xiong & Yip, 2016; Mazurczyk & Lubacz, 2010), the other one embeds secret messages by manipulating the parameters of speech codec during or after encoding process. The steganography based on parameter modification is the most common approach for covert communication based on VoIP. Because of the redundancy of compress speech, slightly change of parameter would not affect the speech quality obviously. Wu and Yang (2006) found that fixed codebook indices are ideally suitable for embedding secret message. They proposed an approach of steganography based on Analysis-by-Synthesis (ABS) by modifying the fixed codebook index parameter in the course of encoding process. To enhance the security of steganography, Wu, Cao, and Li (2015) adopted matrix coding in the modification of fixed codebook index afterward. Geiser and Vary (2008) proposed a steganography method based on an alternative search strategy of the fixed codebook. The experimental results demonstrate that the method causes a negligible effect on the subjective quality of the coded speech with high speed of secrete messages transmission. Miao et al. (2012) also choose fixed codebook indices as the carrier to embed messages, and they used an embedded factor to control the embedding capacity. The method can embed message with both high capacity and low speech distortion by adjusting the factor during the process of speech encoding.

Besides the fixed codebook indices, Liner Prediction Coefficient (LPC) is another feasible domain for steganography. Xiao, Huang, and Tang (2008) utilized an algorithm called complementary neighbor vertices (CNV) to divide the codebook into two parts. Quantization index modulation (QIM) is employed to embed bits into LPC during codebook searching. To enhance the security of QIM, Tian, Liu and Li (2014) introduced a novel steganographic method based on random position selection and matrix encoding strategy. The experimental results show that the approach has greater steganalysis resistance than the Xiao's one (Xiao et al. 2008). Liu, Tian, Lu and Chen (2015) utilized the matrix embedding strategy to hide secret information during the linear predictive coding process. The method has better performance for resist steganalysis and lower speech distortion.