Fast Fourier Transform-based steganalysis of covert communications over streaming media

Fast Fourier Transform-Based Steganalysis of Covert Communications over Streaming Media

Jinghui Peng, Shanyu Tang, Jia Li

Abstract—Steganalysis seeks to detect the presence of secret data embedded in cover objects, and there is an imminent demand to detect hidden messages in streaming media. This paper shows how a new steganalysis algorithm based on Fast Fourier Transform (FFT) can be used to detect the existence of secret data embedded in streaming media. The proposed algorithm uses machine parameter characteristics and a network sniffer to determine whether the Internet traffic contains streaming channels. The detected streaming data is then transferred from the time domain to the frequency domain through FFT. The distributions of power spectra in the frequency domain between original VoIP streams and stego VoIP streams are compared in turn using t-test, achieving the p-value of 7.5686E-176 which is below the threshold. Results indicate that the proposed FFT-based steganalysis algorithm is effective in detecting the secret data embedded in VoIP streaming media.

J. Peng is with the School of Computing and Engineering, University of West London, London, UK (email: 21368391@student.uwl.ac.uk).

***Keywords*—**Steganalysis, Security, Fast Fourier Transform (FFT), Streaming media

I. INTRODUCTION

COVERT communication can be used to transmit confidential information on mobile telecommunications. There are three main ways to implement it: secure channel, encryption technology and information hiding. The secure channel is a private communications path established by the communicating parties, which is not accessible for others. It has high security but high cost and poor extensibility. Encryption technology largely depends on the length of the key used for encryption and decryption. As computer processing capabilities increase rapidly, it becomes less reliable to increase system security by increasing the key length. Digital steganography has drawn people's attention in the field of information hiding. Based on encryption technology, it embeds confidential information into seemingly innocuous transmissions, that is, the encrypted confidential information is "invisible", which is unlikely to be detected by attackers, thus reducing the probability of confidential information being attacked. From the perspective of information transmission security, steganography is one of the most advanced

S. Tang is with the School of Computing and Engineering, University of West London, London, UK (email: shanyu.tang@uwl.ac.uk).

J. Li is with the Freelance Consultancy, Merton, London, UK (email: lily.jjl@gmail.com).

information hiding technologies.

A new generation of mobile telecommunications has emerged as a result of advances in wireless communications and mobile terminal technology. The third generation of mobile telecommunications technology (referred to as 3G) adopts the IMT-2000 international standard, which offers voice service, higher data transfer rates (200 Kbps), broadband multimedia communications, and wireless access to the Internet. The fourth generation of mobile telecommunications (referred to as 4G), using OFDMA and MIMO (TD-LTE), is designed to provide high-speed (100 Mbps) data transmission services such as VoIP and IP Multimedia Subsystem (IMS), with a fibre broadband experience similar to a fixed-line network. With the advent of the era of 3G and 4G, streaming media such as VoIP has been widely used on the mobile Internet, providing a new dynamic cover object for information hiding, especially steganography.

Streaming media steganography has attracted the attention of information security experts all over the world. On the one hand, streaming media contain plenty of redundancy, which can be used to hide confidential information. Compared with image, audio, text and other multimedia files and network channels, streaming media are better cover objects. On the other hand, the widespread use of streaming media on mobile telecommunications networks has portrayed a variety of new mobile Internet services: mobile instant messaging, mobile TV, mobile content sharing, mobile E-reading, mobile social, mobile advertising and so on. Therefore, steganography in streaming media has broad application prospects in the field of mobile Internet.

Steganalysis, the countermeasure technology of covert communication in the field of information hiding, is developing rapidly with the strong demand of investigation into covert communication. Steganographic technology is very likely to be exploited by hostile agents, terrorists and evil forces. By hiding their secret information in streaming media, they intend to avoid content scrutinising and tracing, and use it to organise crimes and terrorist activities, and steal military and commercial information. It would endanger national and public security, and undermine social stability. Therefore, with the rapid development of mobile Internet, it is imminent to develop steganalysis technology for mobile networks, especially steganalysis of streaming media steganography. At present, research in streaming media steganalysis on the mobile Internet is at an early stage, and few preliminary results have been published.

This study was aimed to devise a new steganalysis method for streaming media that are ubiquitous on the mobile Internet and explore ways of universal and real-time detection and countermeasure of steganography in streaming media on the mobile Internet.

II. RELATED WORK

Properly speaking, scientific research in steganalysis (information hiding detection) has lapsed. Steganalysis is to detect the statistical differences between cover object and stego object, such as histogram analysis for LSB steganography, chi-square analysis, RS analysis, steganographic capacity analysis, as well as steganalysis for carrier content

(semantic information layer analysis). Steganalysis of multimedia carriers includes analysis of communications data between specific monitored objects, a large amount of abnormal traffic, abnormal behaviour and status of users, etc., which is called pragmatic information layer steganalysis. It makes a search easier by narrowing the scope of specific analysis [1].

In recent years, the crime of using steganography in streaming media is increasing gradually [2], so it is urgent to study steganalysis technology for streaming media. Kraetzer and Dittmann first proposed a VoIP steganalysis method based on voice feature vector Mel-Cepstrum [3]. Assuming that secret information is not fixedly and continuously embedded into the voice carrier from the beginning to the end of the call, it used χ^2 statistical analysis to analyse the Mel feature of the hidden channel. Later, their work was extended to VoIP steganalysis using Support Vector Machine (SVM) classifier [4].Liu et al. also studied speech steganalysis based on time derivative spectrum and Mel eigenvector [5] [6]. By calculating the distance between the speech signal and its denoising residue, Takahasi and Lee use SVM classifier to detect the hidden information [7]. This method could detect LSB, direct sequence spread spectrum (DSSS), frequency hopping spread spectrum (FHSS) and echo steganography. Garateguy et al. put forward a steganalysis method that relies on RTP packet classification and random projection matrix [8]. It detected the VoIP packets with hidden information through knowing normal data stream structure, and the disadvantage of this method is time-consuming. Arackaparambil et al. presented a statistical distribution-based steganalysis method for VoIP steganography [9], but the measurement range and the length of the detection

window seriously affect the success rate of detection. For low bit rate G.723.1 codec, Li and Huang suggested a voice stream steganalysis algorithm based on Complementary Neighbour Vertex - Quantization Index Modulation (CNV-QIM) [10], which used a distributed histogram, state transition model and time-consuming SVM classifier. Li et al. designed a mechanism for extracting the spread spectrum of hiding data from digital media by using multi-carrier/signature iterated generalised least squares (M-IGLS) [11], but it can only detect image hiding effectively. Janicki et al. studied the steganalysis method of VoIP transcoding information [12], but this method was only valid for G.711 codec. Ren et al. pointed out that data hiding in AMR codec would increase the probability of the same pulse position in the same track, and proposed a speech steganalysis algorithm based on SVM classifier [13], but this method was effective only when the embedding rate was more than 30%. Huang et al. proposed a steganalysis method for VoIP streaming media using a sliding window mechanism to improve RS algorithm [14]. Compared with RS algorithm, the detection time was reduced by 64%, but the detection experiment used G.711 codec and VoIP data streams stored in a database, so the computation was heavy. In addition, for G723.1 codec, they also presented a steganalysis method based on quadratic statistics and regression analysis [15], which can be used to analyse the compressed VoIP voice stream stored in the database and estimate the size of embedded secret information.

Research in streaming media steganography is in its infancy, and the technical approaches still follow image steganography. Most of the research results are limited to embedding and extraction algorithms for hiding data in streaming media. In the scenario

of mobile Internet, it should take into account the time-varying of steganographic capacity and the dynamic delay of, the mobile network, which is caused by the time-variance characteristics of streaming media in the process of steganography, so that a real-time covert communication over streaming media is hard to realise. Traditional information theory and existing steganographic theory cannot describe the time-variance characteristics of steganographic capacity and the dynamic delay characteristics of mobile Internet. Therefore, it is imminent to establish a new theoretical model and method for covert streaming media communications on the mobile Internet.

Steganalysis of streaming media steganography is far from mature. Some existing steganalysis methods need classifiers such as SVM and conventional statistical analysis, which are computationally intensive and time-consuming, thus not meeting the requirements of real-time detection of streaming media steganography on the mobile network. Other steganalysis methods are designed for a certain codec or rely on specific steganographic algorithms, which mean they are not universal and difficult to put into practical applications. Therefore, it is imperative to explore a new steganalysis method for streaming media on the mobile Internet.

To take on the challenges mentioned above, we seek new ways to provide a universal (independent of specific steganographic algorithm and codec) and real-time detection scheme for steganalysis of streaming media on the mobile Internet, through identifying code streams and examining speech cepstrum feature vectors of streaming media with FFT.

III. PROPOSED STEGANALYSIS ALGORITHM BASED ON FFT

Steganalysis is a technology used to detect whether steganography has used to hide secret data in cover objects. The purpose of steganalysis is to identify disguised cover objects containing secret data from ordinary cover objects with little or no knowledge of steganographic algorithms. The simplest steganalysis method is to compare the message or file to be tested with the known message or original file. The difference between the stego object containing secret data and the original payload can be examined through comparisons. Nowadays, steganalysis is becoming important for computer forensics, such as tracking and scanning suspicious documents/audio clips /video clips to effectively prevent criminal and terrorist activities, as well as stopping unauthorised data leakage to protect data privacy. Due to the broad application of streaming media on the mobile Internet, steganalysis of streaming media becomes more and more complex. To cope with this challenge, a novel steganalysis method that uses speech cepstrum feature vectors and Fast Fourier Transform (FFT) is devised in this paper to detect streaming media with messages embedded.

The process of detecting whether secret data is hidden in the cover object is usually modelled as a classification problem. The steganalysis algorithm receives the objects to be tested as input, and classify them into original cover objects and stego objects with data embedded. Steganalysis algorithms can be divided into two categories: Blind steganalysis and Targeted steganalysis. Blind steganalysis is a method that detects all

possible covert channels without knowing steganography algorithms or whether steganography exists. Targeted steganalysis is detection of a specific steganographic algorithm. In both cases, steganalysis is modelled as a classification problem, and techniques available include pattern recognition, classification and machine learning. Classification and machine learning require high computational power and are time-consuming, so the preferred method is pattern recognition. Because of the diversity of streaming media codec and steganographic algorithm, this research focuses on blind steganalysis of streaming media.

For mobile Internet, VoIP streaming media is the main reliable cover object for covert communication. Through close analysis of VoIP streaming media signals, it is known that the effective characteristic parameters of VoIP audio signals include pitch period, linear prediction parameter (LPC), line spectrum pair parameter (LSP), Cepstrum Coefficient, Perceptual Linear Predictive parameters (PLP) and WV spectrum parameters. The main difficulty in pitch detection is that it is impossible to accurately determine each pitch period, the beginning and ending positions in the voiced segment. The pitch period varies from person to person and has a wide range of variation. Linear predictive parameters are widely used in the fields of speech recognition, speech coding and speech synthesis. Lattice method proposed by Makhoul is particularly effective. Line spectrum pair parameter is another form of linear prediction parameter, which is closely related to the spectral characteristics of speech signals, and its quantisation and interpolation characteristics are superior to reflection coefficients. Perceptual Linear Predictive parameters are based on the characteristic parameters of the auditory model, which are

easily interfered by the frequency response of the communication channel. WV spectrum is a high-order, non-stationary signal spectrum, which involves complex time and frequency domain transformation, and is related to the choice of frame length. There are three kinds of cepstrum parameters for VoIP audio signals: LPCC parameters, MFCC parameters, and ASCC parameters. The preliminary idea proposed in this paper is to analyse the characteristic changes in three cepstrum parameters of VoIP streaming media packets before and after data embedding, use FFT calculation to process the cepstrum parameters by frequency transformation, and extract the characteristic threshold of the cepstrum parameters, which is then used to detect the existence of steganography in streaming media, so as to establish a general steganalysis technique for detecting stream media steganography.

In this section, we identify the changes in three cepstrum parameters (LPCC parameters, ASCC parameters and MFCC parameters) of VoIP audio signals caused by steganography in streaming media on the mobile Internet. FFT is used to calculate the characteristic threshold of cepstrum parameters of VoIP streaming media, and a general steganalysis system that uses the cepstrum FFT feature vector to detect stego VoIP streaming media packets with data embedded. The flow chat of the steganalysis system is shown in Figure 1.

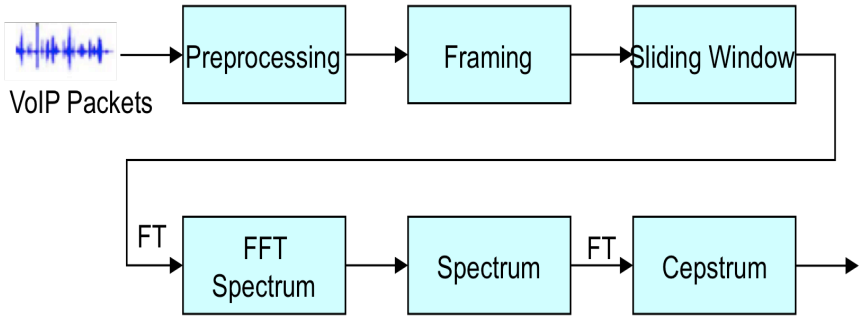

Fig.1 Steganalysis using speech Cepstrum FFT features

The Linear Predictive Cepstrum Coefficient (LPCC) is the representation of LPC in the cepstrum domain. LPCC assumes that an audio signal is an autoregressive signal with low computational complexity. It describes vowels well, but has poor ability to describe consonants and has poor anti-noise performance.

The Accent Sensitive Cepstrum Coefficient (ASCC) is the energy spectrum of audio signals calculated by the filter set method. Though a filter set consists of 16 filters, the filter banks are divided into different frequencies according to the accent sensitive scale. The ASCC coefficients are calculated using modified inverse discrete cosine transform.

Based on the auditory characteristics of the human ear, the Mel Frequency Cepstrum Coefficient (MFCC) first converts a speech spectrum into a non-linear spectrum based on Mel frequency standard, and then converts it into the cepstrum domain. MFCC has no premise assumptions, good speech recognition ability and anti-noise ability, and requires high computational complexity and accuracy. Preliminary research shows that MFCC has better performance than LPCC in some language recognition. The filter set is used to calculate the Mel cepstrum coefficients. Since the human perception of sound above 1000

Hz follows an approximate linear relationship on the logarithmic frequency coordinates, these filters are of equal width at the Mel frequency coordinates. The Mel cepstrum coefficient *Cmel* (*n*) is affected by various factors such as the number, shape, distribution and energy spectrum of the filters in the filter set. It can be obtained by modified inverse discrete cosine transform (IDCT):

$$C_{mel}(n) = \sum_{k=1}^{K} \theta(M_k) \cos\left(n(k-0.5)\frac{\pi}{K} \right) \qquad (n = 1,2,..., p) \tag{1}$$

In the formula, $\theta(M_k)$ denotes the output energy of the *k*th filter, and *p* is the order of MFCC parameters.

The Fourier Transform is one of the most widely used tools for spectrum data analysis, which was used to transform data from the time domain to the frequency domain (and vice versa). The Fast Fourier Transform (FFT) algorithm was devised as a means of reducing the number of computations involved in (and therefore the time required for) finding regular discrete Fourier transforms. The frequency information can be investigated by Fourier transformation of the autocorrelation function into the frequency domain to give power spectral density functions. Autocorrelation measurements are used to extract regularities or periodic deterministic data from noise. The autocorrelation function ($R(\tau)$) is defined as the product of the values of signal $x(t)$ at time t and $x(t+\tau)$ at some time ($t+\tau$) later, average over the observation time T:

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t+\tau)\, dt \qquad (2)$$

$$\tau = r \cdot \Delta t$$
$$r = 1, 2, 3, \ldots, m$$

According to the Weiner-Khintchine theorem [16], the power spectral density functions ($G(f)$) and the autocorrelation function ($R(\tau)$) are related by Fourier transformation:

$$G(f) = 2\int_{-\infty}^{\infty} R(\tau) \exp(-j\omega\tau)\, d\tau \qquad (3)$$

where j and ω are time parameters.

With the designed data processing system, 81920 points data in the time domain (intensity- time) were acquired for audio streaming signals (8 seconds), and used to determine the frequency by Fast Fourier Transform (FFT). The frequencies (f) of the signals were computed with the follow equation [16]:

$$f = \frac{n}{N\Delta t} \qquad n = 0, 1, \ldots N\text{-}1 \qquad (4)$$

where N is the number of data points collected from the signal (here equal to 81920), Δt the sampling period, and n is the data point number corresponding to the peaks of the power spectral density in the frequency domain after FFT. The power spectral densities were calculated by using equations (3) and (4).

IV. RESULTS AND DISCUSSION

In this study, we identify VoIP streaming media channel from mobile Internet communication through machine parameter characteristics, then use a network sniffer to analyse SIP data packets to determine the existence of VoIP channel; use FFT method to transfer VoIP audio data from the time domain to the frequency domain, and then compare the distribution of power spectra in the frequency domain between original VoIP streams and VoIP streams to be tested, determining whether the current streaming media packet contains any secret information; the next streaming media packet is detected in turn to realise real-time detection of covert channel in mobile interconnected streaming media.

The mute parameters and interrupt parameters in machine parameters are two unique characteristics of streaming media signals passing through VoIP channel. This study proposes to recognise VoIP channel by these two parameters. The signal level of natural speech does not exhibit sharp changes in the convex and concave peaks (i.e., mute), but after compression coding these peaks would be effected due to computational inaccuracies and packet switching. If a signal level has a steep or falling point, the signal contains a mute, and the presence of a mute indicates that it has passed through the VoIP channel. Similarly, the natural voice level signal does not appear to suddenly stop in the pitch period, and after compression coding and packet switching transmission, the abrupt

termination of voice occurs due to the influence of calculation error and packet loss. This feature can be used to recognise the existence of VoIP streaming media channel.

In addition, through a network sniffer and analyser (such as Wireshark, WildPacket original datagram capturer and analyser), VoIP media streams are analysed, protocol decoding is carried out, and SIP data packets are examined to identify the existence of VoIP streaming media channel. In the process of streaming media VoIP communication, online measurement of various voice quality parameters, such as PESQ, MOS-PSQM, PAMS, can also be used with network sniffers.

The main purpose of steganography is to convey secret information in seemingly ordinary channels without causing any suspicion. The security performance of a steganographic system is usually evaluated by statistical undetectability, which refers to the difficulty of reliably determining the secret information embedded in the cover object. In other words, a secure steganographic system means statistical undetectability.

In this study, a steganalysis method based on Fast Fourier Transform (FFT) is proposed to detect the covert channel of mobile streaming media in real time. FFT is a method of converting data from time domain to frequency domain. By comparing the distribution differences of two groups of audio streams in the frequency domain, we recognise whether they are different, and then determine whether the audio streams to be detected contains secret data.

In our experiments, steganalysis tests on VoIP communications were implemented on original VoIP streams and VoIP Streams to be detected. We detected whether the VoIP streams contain secret data by comparing the distribution of two groups of audio streaming data in the frequency domain. The experimental results are discussed in detail below.

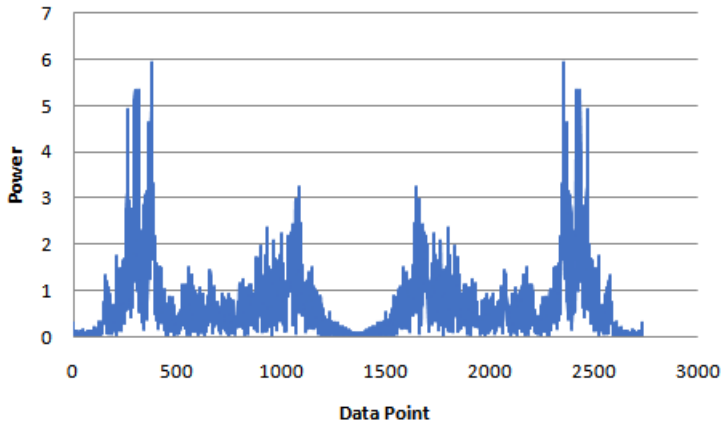

Fig. 2 Power spectra in frequency domain of the original VoIP streams

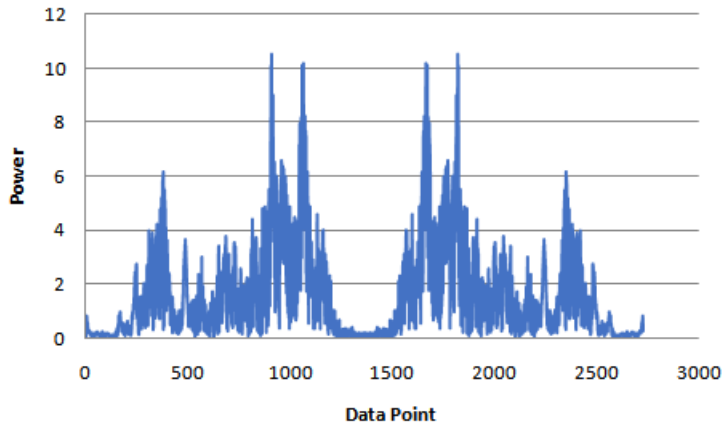

Fig. 3 Power spectra in frequency domain of the VoIP streams with secret data embedded

Figures 2 and 3 show the results of the audio samples of VoIP streams. The original data was obtained from the experiments published in [17]. We collected data points from the original VoIP streams and the VoIP streams to be detected, and then transferred them from the time domain to the frequency domain by the proposed FFT algorithm. As the figures show, their power spectra were similar, but the peaks changed significantly, which indicates the possibility of secret information embedded in the VoIP streams to be detected. We further tested the difference between them by t-test to determine whether the streams to be detected contain secret information.

The t-test is a statistical hypothesis test in which the test statistic follows a Student's t-distribution under the null hypothesis. It uses t-distribution theory to infer the probability of difference occurrence, thereby comparing whether the difference between the two means is significant.

A two-sample t-test is to test whether the difference between the means of two samples and the population represented by them is significant. This test is used when it can be assumed that the two distributions have the same variance. The t statistic used to test whether the means are different can be calculated as follows:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}(\frac{1}{n_1} + \frac{1}{n_2})}} \qquad (5)$$

where S_1^2 and S_2^2 are the variances of the samples, n_1 and n_2 are the sample sizes.

Usually, we used the default alpha of 0.05 as the threshold. When the calculated p-value is below the threshold, then there is significant difference between the two samples.

t-Test: Two-Sample Assuming Equal Variances

	Variable 1	Variable 2
Mean	0.761833868	1.615130783
Variance	0.6624058	2.801982919
Observations	4000	4000
Pooled Variance	1.73219436	
Hypothesized Mean Difference	0	
df	7998	
t Stat	-28.99458273	
P(T<=t) one-tail	3.7843E-176	
t Critical one-tail	1.645044168	
P(T<=t) two-tail	7.5686E-176	
t Critical two-tail	1.960260582	

Fig. 4 Results of T-test

Among the variables that appear in the results sheet (Figure 4), depending on our experiment, the most important result is the p-value, highlighted in the figure in the green cell. Since we chose the alpha value of 0.05, if the p-value is less than 0.05, then there is a statistically significant difference between the means of two trials. In this experiment, p = 7.5686E-176, which is < 0.05, so the difference is significant.

The experimental results show there is significant difference between the power spectra in the frequency domain of the original VoIP streams and the VoIP stream to be detected, indicating that the VoIP streams to be detected contain hidden data and our steganalysis algorithm is effective in detecting covert channels.

V. CONCLUSION

In the study, we proposed a FFT-based steganalysis algorithm for covert communications over streaming media on the mobile Internet, thereby realising the detection of secret data hidden in streaming media. The proposed steganalysis algorithm can identify VoIP streaming media channel from mobile internet communications, and determine whether the current streaming media packet contains secret information by comparing the power spectra generated using FFT. Our experimental results have shown that the proposed steganalysis algorithm is effective in detecting hidden message in streaming media. Further studies should investigate how to accurately acquire the media packets needed for detection when heavy packet lose occurs.

ACKNOWLEDGMENT

This work was supported in part by various industrial sponsors under Grant 28801.

REFERENCES

[1] X. Niu, Y. Yang (2006). Study on the Frame of Information Steganography and Steganalysis. Acta Electonica Sinica, 34(12A), 2421-2424.

[2] J. Lubacz, W. Mazurczyk, and K. Szczypiorski (2010). Voice over IP. IEEE Spectrum, 40-45.

[3] C. Kraetzer, and J. Dittmann (Jan. 2007). Mel-Cepstrum based steganalysis for VoIP-steganography. Proc. IS&T/SPIE Symposium on Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII, San Jose, USA, pp. 650-661.

[4] C. Kraetzer, and J. Dittmann (2007). Pros and cons of Mel-Cepstrum based audio steganalysis using SVM classification. Information Hiding, Springer Berlin Heidelberg, 359-377.

[5] Q. Liu, A. H. Sung, and M. Qiao (2009). Temporal derivative-based spectrum and mel-cepstrum audio steganalysis. IEEE Transactions on Information Forensics and Security, 4(3), 359-368.

[6] Q. Liu, A. H. Sung, and M. Qiao (2011). Derivative-based audio steganalysis. ACM Transactions on Multimedia Computing, Communications and Applications, 7(3), 1-19.

[7] T. Takahashi T, and W. Lee (2007). An assessment of VoIP covert channel threats. Proc. 3rd Int Conf Security and Privacy in Communication Networks, Nice, France, 371-380.

[8] G. Garateguy, G. Arce, and J. Pelaez (2011). Covert channel detection in VoIP streams. Proc. 45th Annual Conference on Information Sciences and Systems, Baltimore, Maryland, 1-6.

[9] C. Arackaparambil, G. Yan, S. Bratus, and A. Caglayan (2012). On tuning the knobs of distribution-based methods for detecting VoIP covert channels. Proc. Hawaii International Conference on System Sciences, Hawaii, 2431-2440.

[10] S. Li, H. Tao, and Y. Huang (2012). Detection of quantization index modulation steganography in G. 723.1 bit stream based on quantization index sequence analysis. Journal of Zhejiang University Science C, 13(8), 624-634.

[11] M. Li, M. K. Kulhandjian, D. A. Pados, S. N. Batalama, and M. J. Medley (2013). Extracting spread-spectrum hidden data from digital media. IEEE Transactions on Information Forensics and Security, 8(7) 1201-1210.

[12] A. Janicki, W. Mazurczyk, and K. Szczypiorski (2014). Steganalysis of transcoding steganography. Ann. Telecommun., vol. 69, 449-460.

[13] Y. Ren, T. Cai, M. Tang, and L. Wang (2015) AMR steganalysis based on the probability of same pulse position. IEEE Transactions on Information Forensics and Security, 10(9), 1801-1811.

[14] Y. F. Huang, S. Tang, and Y. Zhang (2011). Detection of covert Voice over Internet Protocol communications using sliding window-based steganalysis. IET Communications, vol. 5, iss. 7, 929-936.

[15] Y. Huang, S. Tang, C. Bao, and Y. J. Yip (2011). Steganalysis of compressed speech to detect covert Voice over Internet Protocol channels. IET Information Security, 5(1), 26-32.

[16] E.O. Brigham, (1988). The Fast Fourier Transform and Its Applications. Prentice-Hall Inc., NJ, USA.

[17] S. Tang, Y. Jiang, L. Zhang, Z. Zhou, Audio steganography with AES for real-time covert voice over Internet protocol communications, Science China Information Sciences. 57 (3)(2014) 1-14.