symmetry

MDPI

Article

Steganalysis of Inactive Voice-Over-IP Frames Based on Poker Test

Jie Liu [1], Hui Tian [1,*], Chin-Chen Chang [2], Tian Wang [1], Yonghong Chen [1] and Yiqiao Cai [1]

[1] College of Computer Science and Technology, National Huaqiao University, Xiamen 361021, China; liujiecs@hqu.edu.cn (J.L.); wangtian@hqu.edu.cn (T.W.); iamcyh@hqu.edu.cn (Yo.C.); caiyq@hqu.edu.cn (Yi.C.)

[2] Department of Information and Computer Science, Feng Chia University, Taichung 40724, Taiwan; alan3c@gmail.com

* Correspondence: htian@hqu.edu.cn; Tel.: +86-592-6162497

Received: 20 July 2018; Accepted: 7 August 2018; Published: 11 August 2018

Abstract: This paper concentrates on the detection of steganography in inactive frames of low bit rate audio streams in Voice over Internet Protocol (VoIP) scenarios. Both theoretical and experimental analyses demonstrate that the distribution of 0 and 1 in encoding parameter bits becomes symmetric after a steganographic process. Moreover, this symmetry affects the frequency of each subsequence of parameter bits, and accordingly changes the poker test statistical features of encoding parameter bits. Employing the poker test statistics of each type of encoding parameter bits as detection features, we present a steganalysis method based on a support vector machine. We evaluate the proposed method with a large quantity of speech samples encoded by G.723.1 and compare it with the entropy test. The experimental results show that the proposed method is effective, and largely outperforms the entropy test in any cases.

Keywords: steganalysis; Voice over Internet Protocol (VoIP); inactive frame; poker test; G.723.1

1. Introduction

Steganography is a technique of covert communication by embedding secret messages into seemingly innocent digital media such as audio [1–3], image [4–10] and video [11,12]. Like other security techniques, such as encryption [13,14], the misuse of steganography by lawbreakers will pose a threat to network security and public safety. To confront this challenge, its countermeasure, steganalysis, has received increasing attention. The aim of steganalysis is to detect, extract and destroy the secret messages embedded in digital media, where determining whether the suspicious media contain secret messages is the precondition of other operations [15].

In recent years, Voice over VoIP has emerged as a popular communication service over the Internet for its convenience and instantaneity. With the widespread application of VoIP, researchers have paid more and more attention to VoIP-based steganography. Compared with traditional carriers, there are many advantages for VoIP-based carriers, such as immediacy, high steganographic bandwidth and alterable steganographic length [15]. In general, VoIP-based steganography can be classified into the following two categories. One employs the relevant protocols of VoIP as carriers, for example, Mazurczyk and Szczypiorski [16] used the redundant data area in session initiation protocol (SIP) to embed secret messages and Forbes [17] created a covert channel by modifying the timestamp in real-time transport protocol (RTP). The other is to embed the secret messages in VoIP payload, which attracted more attention from the research community last decade [15] for its higher steganographic bandwidth. Low bit rate codecs are widely applied in VoIP for its high compression ratio; most of the steganographic algorithms in VoIP payload are conducted on them. For most low bit rate codecs, there are three feasible embedding domains, including fixed codebook (FCB) [18–20], liner prediction coefficients (LPC) [21–23] and adaptive codebook (ACB) [24–27]. For example, Geiser and Vary [18] presented a steganography by modifying the FCB search strategy to embed secret messages during the encoding process, and the embedding capacity can reach up to 35 bits per subframe with adaptive multi-rate (AMR) 12.2 kbit/s mode. Later, Miao et al. [19] proposed another steganography to limit the pulse positions in FCB to embedding secret messages and further introduced an embedding factor to control the embedding capacity. Liu et al. [22] introduced the genetic algorithm into Vector Quantization (VQ) division and replaced the quantization index set of LPC with secret messages, which has a better performance than random division of VQ. Huang et al. [25] proposed an embedding algorithm with high steganographic capacity, which was accomplished by adjusting the closed-loop pitch period range of a subframe according to secret message bits. Due to it being integrated into an encoding process, there is no delay when embedding and extraction. Janicki et al. [27] proposed a steganography algorithm based on approximating the F0 parameter of pitch in a speex codec, which can be applied without any steganographic cost. Recently, Huang et al. [28] improved the Voice Activity Detection (VAD) algorithm to keep the VAD result invariant after steganography and modified several types of parameter bits in inactive frames in G.723.1 with 6.3 kbit/s mode to embed the secret messages, whose steganographic bandwidth can reach up to 101 bits per frame. Lin [29] proposed another improved VAD algorithm to keep the VAD result unchanged after steganography and extended Huang's method to 5.3 kbit/s mode.

As for VoIP steganalysis, there is no universal detection method currently, but some effective steganalysis methods [30–37] have been proposed to detect steganographic algorithms, which modify specific encoding parameters. For example, Li et al. [30] pointed out that the codeword of LPC became asymmetrical after embedding secret messages and proposed a quantization codework correlation network model to detect LPC-based steganography, which has a good detection performance even in short sample length. Lin et al. [31] first introduced the recurrent neural network (RNN) into steganalysis and designed a two-layer network to detect the LPC-based steganography. The experimental results show that Lin's method [31] has better detection performance than Li's method [30] and can achieve a good detection accuracy when the sample length is only 0.1 s at the embedding rate of 100%. After analyzing the search rule of pitch delay, a Markov matrix of the second-order difference of pitch delay was presented as steganalysis features by Ren et al. [32] to detect steganography in ACB. According to the short-time stability of speech, Tian et al. [33] proposed a series of steganalysis features to completely describe the characteristics among the pulse positions in FCB, which included the Markov matrix of the pulse positions in the same subframe and the joint probability distributions of pulse positions among different subframes. Because the steganalysis feature vector had a high feature dimension, adaptive boost was applied to feature selection. The experimental results show that Tian's method [33] outperforms the state of the arts [34,35]. However, for Huang's [28] and Lin's [29] methods, several types of encoding parameter bits in inactive frames are modified and there is no effective detection method currently. To fill this gap, we analyze the effect of steganography on the encoding parameter bits and present a support vector machine based steganalysis method with a large number of speech samples encoded by G.723.1. Specifically, our contribution in this work can be summarized as follows: (1) we present a steganalysis method of inactive voice-over-IP frames based on poker test, which is the first work aiming to detect steganography of inactive VoIP frames; (2) we analyze the impacts on parameter bits induced by the steganographic process, and model the steganalysis feature using the poker test statistics of different parameters; and (3) we comprehensively evaluate the detection performance of the presented scheme by experiments and comparisons with the traditional entropy test [38]. The experimental results demonstrate that our scheme can effectively detect the steganography of inactive VoIP frames, and significantly outperforms the entropy test.

The rest of this paper is organized as follows. Section 2 introduces the improved VAD algorithm and the steganography in the inactive frames. In Section 3, theoretical analyses of the proposed features have been presented. The support vector-machine based steganalysis method is presented in Section 4. Experiments and performance analysis are given in Section 5. Finally, Section 6 concludes the paper.

2. Related Work

2.1. Improved VAD Algorithm

ITU G.723.1 [39] is a hybrid codec with two encoding modes: 5.3 kbit/s mode and 6.3 kbit/s mode and each frame is coded into various parameters. The length of each frame of both modes is 30 ms, and the speech bits of each frame are 160 bits with 5.3 kbit/s mode and 192 bits with 6.3 kbit/s mode. The bit allocation of encoding parameters of each frame is listed in Tables 1 and 2.

VAD is to determine whether the current frame is active or inactive by comparing the energy of the current frame with a threshold [39]. The inactive state of frames may be affected by steganography, so it is necessary to keep the VAD results of the sender and the receiver consistent. However, the original VAD algorithm is related to the previous frame of the current frame. If the inactive state of the previous frame is impacted by embedding secret messages, the inactive state of the current frame may be also affected when applied original VAD algorithm to detect inactive frames, which leads to an altered VAD result. To solve the problem, Huang et al. [28] improved the algorithm of autocorrelation coefficients, which made the autocorrelation coefficients of the current frame were independent with the previous frame. Then, the stateless coefficients were used to calculate residual energy, which was then compared with the threshold, so the VAD results kept invariable after steganography. Due to the silence compression function being optional for G.723.1 codec and there being some bits in the VoIP packet header to direct whether to use the silence compression function during encoding process, Lin [29] used the one free bit generated by disabling the silence compression function to mark the inactive frames embedded with secret messages. When decoding the VoIP streams at the receiver end, there is no need to run the VAD algorithm again and the secret messages are extracted from the inactive frames that have been marked.

Table 1. Bit allocation of G.723.1 codec with 6.3 kbit/s mode [39].

Parameters	Subframe 0	Subframe 1	Subframe 2	Subframe 3	Subtotal (bits)
Adaptive codebook lags (Olp/Aclg)	7	2	7	2	18
LPC indices (Lsf)	-	-	-	-	24
Grid index (Grid)	1	1	1	1	4
All the gains combined (Mamp)	12	12	12	12	48
Pulse positions (Ppos)	20	18	20	18	73
Pulse signs (Pamp)	6	5	5	5	22
Total	-	-	-	-	189

Table 2. Bit allocation of G.723.1 codec with 5.3 kbit/s mode [39].

Parameters	Subframe 0	Subframe 1	Subframe 2	Subframe 3	Subtotal (bits)
Adaptive codebook lags (Olp/Aclg)	7	2	7	2	18
LPC indices (Lsf)	-	-	-	-	24
rid index (Grid)	1	1	1	1	4
All the gains combined (Mamp)	12	12	12	12	48
Pulse positions (Ppos)	12	12	12	12	48
Pulse signs (Pamp)	4	4	4	4	16
Total	-	-	-	-	158

2.2. Steganography in Inactive Frame

Huang et al. [28] and Lin [29] selected the parameters to embed secret messages by evaluating the effects after being modified. There are 101 bits per frame for steganography with 6.3 kbit/s mode, 81 bits per frame for steganography with 5.3 kbit/s mode, and the suitable parameters for steganography are listed in Tables 3 and 4. The selected parameters are used to embed the secret messages with the following algorithm that involves three steps:

Step 1: Voice activity detection. Speech samples are divided into frames, and each frame is input into the VAD detector, where the inactive frames are marked with a tag.

Step 2: Encoding and embedding secret messages in inactive frames. All frames are encoded without applying silence compression function. If the frame has been marked in Step 1, suitable parameters of the frame will be embedded with secret messages.

Step 3: Encapsulation and send. All the frames are encapsulated in VoIP packets, which are transmitted over the Internet.

Table 3. Parameters of the inactive frame suitable for embedding secret messages with 6.3 kbit/s mode [28].

Parameter Name	**Lsf**	**Grid**	**H_Ppos**	**L_Ppos**	**Pamp**	**Total Bits**
Number of bits	2	4	13	60	22	101

Table 4. Parameters of the inactive frame suitable for embedding secret messages with 5.3 kbit/s mode [29].

Parameter Name	**Olp**	**Lsf**	**Gains**	**Grid**	**Pamp**	**Ppos**	**Total Bits**
Number of bits	2	3	8	4	16	48	81

3. Steganalysis Based on Poker Test

Poker Test [40] is a technique to determine whether a given bit sequence satisfies the characteristics of a truly random sequence. Let X be a bit sequence of length n and m be the length of a subsequence of X such that

$$\left\lfloor \frac{n}{m} \right\rfloor \geq 5 \cdot 2^m. \tag{1}$$

For a given bit sequence X, it can be divided into k non-overlapping subsequences each of length m, which can be written as

$$k = \left\lfloor \frac{n}{m} \right\rfloor. \tag{2}$$

Let F_i be the frequency of the i-th type of subsequence of length m, where $1 \leq i \leq 2^m$. The poker test statistic is defined as

$$V = \frac{2^m}{k}\left(\sum_{i=1}^{2^m} {F_i}^2\right) - k, \tag{3}$$

which approximately follows a χ^2 distribution with $2^m - 1$ degrees of freedom. In the proposed steganalysis method, each parameter suitable for steganography in all inactive frames can form a bit sequence. The bit sequence can be considered as consisting of a series of subsequences, which can be expressed as

$$X = \{S_1, S_2, ..., S_j\}, (1 \leq j \leq k), \tag{4}$$

where S_j is the subsequence of X of length m. Denote P as the set which contains all the 2^m types of subsequences. Then, the frequency of the i-th type of subsequence can be calculated by

$$F_i = \sum_{j=1}^{k} I(S_i = S_j), (S_i \in X, S_j \in P), \tag{5}$$

where $I(x)$ is expressed as

$$I(x) = \begin{cases} 1, x \text{ is true} \\ 0, \text{else} \end{cases}. \tag{6}$$

Let b_i be the i-th bit in X, the probabilities b_i = 1 and b_i = 0 are denoted as p (b_i = 1) and p (b_i = 0), respectively; denote the embedding rate as r, the probabilities for b_i = 1 and b_i = 0 after steganography are $p'(b_i$ = 1) and $p'(b_i$ = 0), which can be expressed as

$$p'(b_i = 1) = (1 - \frac{r}{2}) \cdot p(b_i = 1) + \frac{r}{2} \cdot p(b_i = 0), \tag{7}$$

$$p'(b_i = 0) = (1 - \frac{r}{2}) \cdot p(b_i = 0) + \frac{r}{2} \cdot p(b_i = 1). \tag{8}$$

By subtracting the above two equations, we can obtain

$$p'(b_i = 1) - p'(b_i = 0) = (1 - r)(p(b_i = 1) - p(b_i = 0)). \tag{9}$$

From Equation (9), it can be concluded that the distribution of 0 and 1 in X tends to be symmetric as the embedding rate increases. When the distribution of 0 and 1 becomes symmetric and the bit sequence is long enough, the values of F_i will be nearly equal. For example, let m = 2, there are four types of subsequences ({0,0}, {0,1}, {1,0}, {1,1}) and the probability of each subsequence is approximately equal to 0.25. Based on this, the values of F_i satisfy the following equations:

$$\sum_{i=1}^{2^m} F_i = k, \tag{10}$$

$$F_1 = F_2 = ... = F_i = ... = F_{2^m}, (1 \leq i \leq 2^m). \tag{11}$$

According to the Cauchy–Buniakowsky–Schwarz inequality [41], we can obtain:

$$\sum_{i=1}^{2^m} {F_i}^2 = \frac{k^2}{2^m}. \tag{12}$$

Putting Equation (12) into Equation (3), we can obtain:

$$\begin{aligned} V &= \tfrac{2^m}{k}(\sum_{i=1}^{2^m} {F_i}^2) - k \\ &= \tfrac{2^m}{k}(\tfrac{k^2}{2^m}) - k \\ &= k - k \\ &= 0 \end{aligned} . \tag{13}$$

However, because the length of X is limited, the value of V is generally non-zero. We can still reach the conclusion from the above analyses that the poker test statistic of X tends to decrease with the increase of the embedding rate. To verify the inference, we calculate the poker test statistics of each parameter suitable for steganography at different embedding rates (from 0.1 to 1.0). Table 5 shows the poker test statistics of each parameter suitable for steganography at different embedding rates with 6.3 kbit/s mode. Apparently, although the embedding secret messages have different effects on different parameter sequences, the poker test statistics tend to decrease with the increase of embedding rate. Therefore, the poker test statistics of each parameter sequence can be applied as steganalysis features.

Table 5. Poker test statistics with 6.3kbit/s mode at various embedding rates.

Parameters	**Embedding Rate**										
	0	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
Lsf	93.68	93.68	94.66	79.55	67.59	56.76	43.99	17.81	5.93	18.78	12.47
Grid	3.01	3.09	2.60	2.93	3.13	6.04	4.22	5.76	5.23	3.10	0.59
H_Ppos	4.12	5.23	4.28	6.55	3.18	3.00	3.31	1.30	5.23	4.71	0.39
L_Ppos	7.52	8.25	6.26	6.67	5.88	4.48	2.73	7.93	2.71	2.85	3.06
Pamp	201.45	203.87	182.85	161.63	119.60	131.89	84.86	74.00	30.68	4.37	2.99

4. SVM-Based Steganalysis Method

In this section, the steganalysis method based on the support vector machine (SVM) has been presented, and the proposed features in Section 3 are employed as the steganalysis features. The steganalysis method includes a training process and a detection process. Specifically, the training process is divided into three steps as follows:

Step 1: Sample preparation. Collect a great quantity of speech samples encoded by G.723.1 with both encoding modes and embed secret messages with the steganography in Section 2.2 at different embedding rates.

Step 2: Feature extraction. Extract the proposed features in Section 3, of which the extraction process is shown in Figure 1.

Step 3: Classifier training. Train the SVM classifier with the feature vector built in Step 2.

Similarly, the detection process contains two steps as follows:

Step 1: Feature extraction. Extract the proposed features from the samples to be detected.

Step 2: Decision-making. Input the features extracted in Step 1 into the trained SVM classifier to determine whether the samples to be detected contain secret messages according to the classification results.

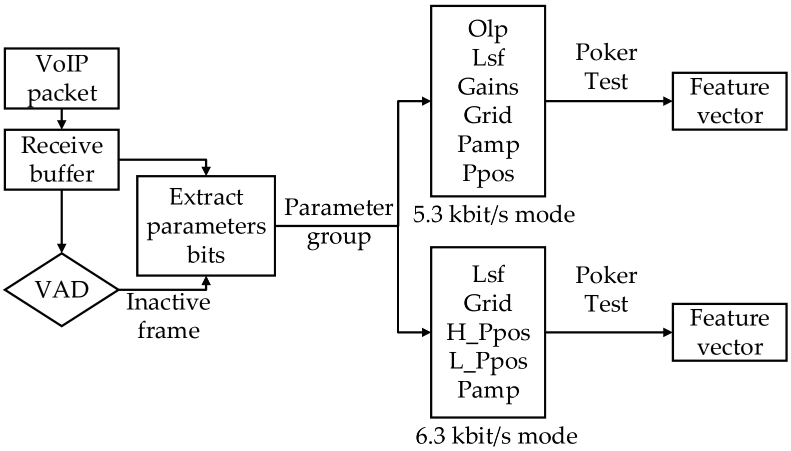


Figure 1. The extraction process of steganalysis features.

5. Experimental Result and Analysis

5.1. Experiment Setup and Performance Evaluation

In this paper, we gather a large number of speech samples with a length of 10 s (333 frames) from language-learning lessons to evaluate the performance of the proposed method without loss generality. Specifically, the experimental dataset consists of 2200 speech samples, which are 8000 Hz sampled and 16-bit quantized. All of these samples involve four types, namely, Chinese male speech samples, Chinese female speech samples, English male speech samples and English female speech samples. Note that, in the experiments, we only focus on the inactive frames in these speech samples. The distribution of inactive frames is shown in Figure 2, and secret messages are produced randomly. In this paper, the SVM with radial basis function (RBF) kernel is implemented based on LibSVM [42] in C-style, where $c = 1$ and $g = 1/1064$. Meanwhile, half samples are used to train the classifiers, and the other half are used to test the performance of the proposed method.

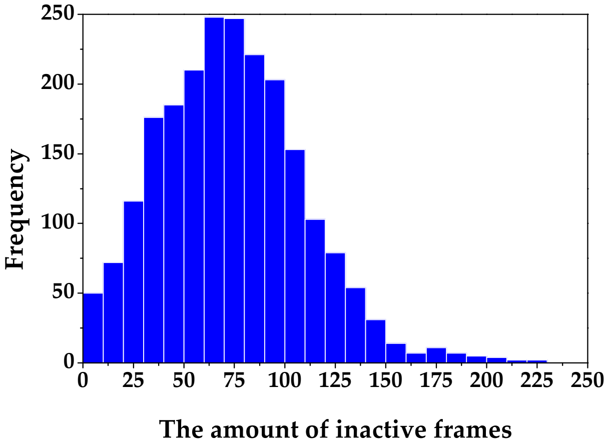


Figure 2. The distribution of inactive frames.

The poker test statistics of all suitable parameters for steganography are calculated with $m = 2$, which reaches the best detection performance in our experiments and satisfies Equation (1). Since there is no detection method to detect the targeted steganography, we compare our method with the entropy test [38], which has been already used to detect symmetry [43] and steganography [34,36]. In the entropy test, all of the parameters suitable for steganography can form a sequence; then, the entropies of the eight types of binary sequences ({0,0,0}, {0,0,1}, . . . {1,1,1}) are calculated respectively as steganalysis features. Employing the entropies of the eight types of binary sequences as steganalysis features reaches the best detection performance in our experiments.

The performances of both the steganalysis methods are evaluated by accuracy (ACC), false positive rate (FPR) and false negative rate (FNR). The accuracy is the percentage of the samples that are correctly classified in the total of test samples. ACC can be calculated by

$$\mathrm{ACC} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{TP}}, \tag{14}$$

where N_{TP} is the quantity of true positives, namely, the quantity of steganographic samples identified as steganographic samples; N_{TN} is the quantity of true negatives, namely, the quantity of cover samples identified as cover samples; N_{FP} is the quantity of false positives, namely, the quantity of cover samples identified as steganographic samples; N_{FN} is the quantity of false negatives, namely, the quantity of steganographic samples identified as cover samples. False positive rate (FPR) is the probability of false positives in the total number of negatives, which can be expressed as

$$\mathrm{FPR} = \frac{N_{FP}}{N_{TN} + N_{FP}}, \tag{15}$$

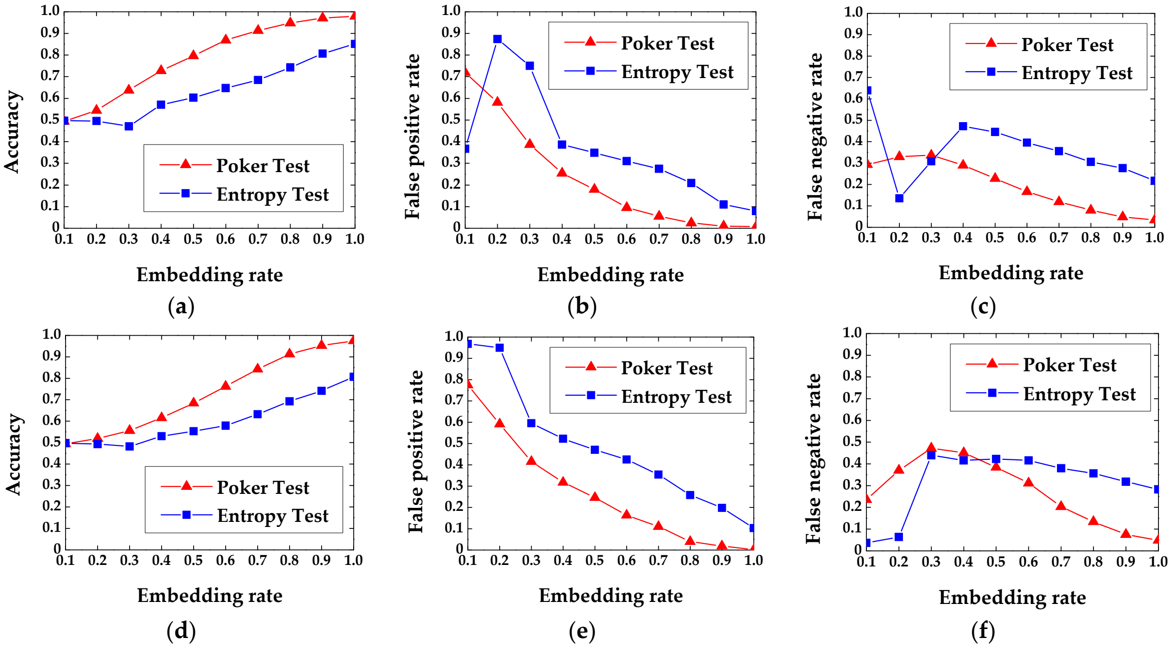
where the sum of N_{TN} and N_{FP} is the total number of negatives, that is, the total number of cover samples. False negative rate (FNR) is the probability of false negatives in the total number of positives, which can be expressed as

$$\text{FNR} = \frac{N_{FN}}{N_{TP} + N_{FN}}, \tag{16}$$

where the sum of N_{TP} and N_{FN} is the total number of positives, that is, the total number of steganographic samples.

5.2. Performance and Analysis

In our steganalysis experiments, 2200 ten-second samples are embedded with secret messages at different embedding rates (from 0.1 to 1.0), respectively. Figure 3 shows the experimental results of detection accuracy, FPR and FNR at different embedding rates.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**)

Figure 3. Accuracy (ACC), false positive rate (FPR) and false negative rate (FNR) at various embedding rates. (**a**) ACC with 5.3 kbit/s mode; (**b**) FPR with 5.3 kbit/s mode; (**c**) FNR with 5.3 kbit/s mode; (**d**) ACC with 6.3 kbit/s mode; (**e**) FPR with 6.3 kbit/s mode; (**f**) FNR with 6.3 kbit/s mode.

From these charts, we can reach the following conclusions: first, for both of the detection methods, the detection accuracy increases as the embedding rate increases, which means that the detection performance is positively correlated with an embedding rate of a given steganography. Furthermore, from Figure 3a,d, it can be observed that the difference of detection accuracy between entropy test and the poker test also increases as the embedding rate increases. Particularly, for entropy tests, the accuracy is nearly 85% with 5.3 kbit/s at the embedding rate of 100% while the poker test can reach the same level of accuracy at the embedding rate of 60%, which means that the proposed method has much better detection performance than an entropy test.

Second, FPR and FNR decrease according to the increase of the embedding rate and the proposed method has lower FPR and FNR than an entropy test. The abnormality at low embedding rates may be that the parameter bits are ordered in the initial state and when the embedding rate is low, few parameters are modified, which lead to the parameters bits slightly disordered; with the embedding rate increasing, more parameters are replaced by secret messages, which makes the parameter bits become ordered again.

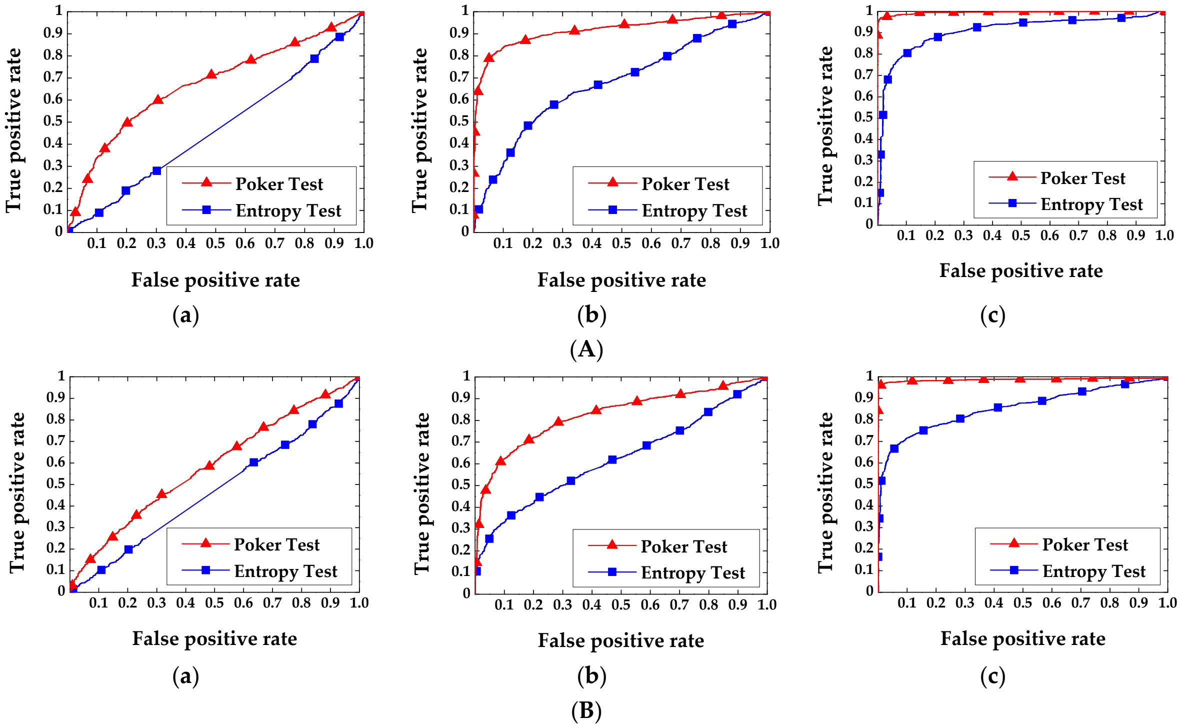
To further evaluate the performances of steganalysis methods, the receiver operating characteristic (ROC) curves at the embedding rates of 30%, 60%, 100% are drawn in Figure 4. To conduct ROC curves, True Positive Rate (TPR) and True Negative Rate (TNR) need to be calculated firstly. TPR is the proportion of true positives out of all positives, which is calculated by

$$\mathrm{TPR} = \frac{N_{TP}}{N_{TP} + N_{FN}}. \tag{17}$$

TNR is the proportion of true negatives out of all negatives, which is calculated by

$$\mathrm{TNR} = \frac{N_{TN}}{N_{TN} + N_{FP}} \tag{18}$$

The results reconfirm that the proposed method has better performance than the entropy test. Moreover, we can get another conclusion from Figure 4 that the proposed method with 5.3 kbit/s mode slightly outperforms that with 6.3 kbit/s mode.



(a) (b) (c)

(A)

(a) (b) (c)

(B)

Figure 4. The receiver operating characteristic (ROC) curves at various embedding rates with both encoding modes. (**a**) ROC at embedding rate of 30%; (**b**) ROC at embedding rate of 60%; (**c**) ROC at embedding rate of 100%; (**A**) the ROC with 5.3 kbit/s mode; (**a**) ROC at embedding rate of 30%; (**b**) ROC at embedding rate of 60%; (**c**) ROC at embedding rate of 100%; (**B**) the ROC with 6.3 kbit/s mode.

6. Conclusions

Steganography of inactive VoIP frames is a state-of-the-art information hiding method for VoIP, which lacks effective countermeasures. To fill this gap, we present a steganalysis method of inactive voice-over-IP frames based on a poker test. Specifically, we analyze the impacts on parameter bits induced by the steganographic process, model the steganalysis feature using the poker test statistics of different parameters, and finally present an SVM based steganalysis method to detect steganography of inactive VoIP frames. We evaluate the proposed method with a large number of speech samples encoded by G.723.1 and compare it with the traditional entropy test. The experimental results show that the proposed method is effective and achieves much better detection performance than the entropy test.

Author Contributions: All authors contributed the preparation of this paper. H.T. and J.L. designed the method, conducted the literature review, guided the experiments, analyzed the results and wrote the manuscript; J.L. designed and conducted the experiments; C.-C.C. supervised the research work and provided suggestions for improvement; T.W., Y.C. (Yonghong Chen) and Y.C. (Yiqiao Cai) analyzed the experimental results and provided revising suggestions.

Funding: This research is funded by the National Natural Science Foundation of China under Grant Nos. U1536115 and U1405254, the Natural Science Foundation of Fujian Province of China under Grant No. 2018J01093, the Program for New Century Excellent Talents in Fujian Province University under Grant No. MJK2016-23, the Program for Outstanding Youth Scientific and Technological Talents in Fujian Province University under Grant No. MJK2015-54, the Promotion Program for Young and Middle-aged Teachers in Science and Technology Research of Huaqiao University under Grant No. ZQN-PY115, Program for Science and Technology Innovation Teams and Leading Talents of Huaqiao University under Grant No. 2014KJTD13, the Opening Project of Shanghai Key Laboratory of Integrated Administration Technologies for Information Security under Grant No. AGK201710 and the Subsidized Project for Postgraduates' Innovative Fund in Scientific Research of Huaqiao University No. 17013083007.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Naidu, T.R.K.; Kumar, G.P.; Prasad, T.G. Overview of digital audio steganography techniques. *Int. J. Emerg. Technol. Eng.* **2016**, *3*, 62–66.
2. Joshi, R.; Venugopala, P.S. Improved security in audio steganography using packet forger at the third level. In Proceedings of the 2017 International Conference on Smart Technologies for Smart Nation (SMARTTECHCON), Bangalore, India, 17–19 August 2017; pp. 363–368.
3. Liu, H.; Liu, J.; Hu, R.; Yan, X.; Wan, S. Adaptive Audio Steganography Scheme Based on Wavelet Packet Energy. In Proceedings of the 2017 IEEE 3rd International Conference on Big Data Security on Cloud (BIGDATASECURITY), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), Beijing, China, 26–28 May 2017; pp. 26–31.
4. Hussain, M.; Wahab, A.W.A.; Javed, N.; Jung, K.H. Hybrid Data Hiding Scheme Using Right-Most Digit Replacement and Adaptive Least Significant Bit for Digital Images. *Symmetry* **2016**, *8*, 41. [CrossRef]
5. Rajendran, S.; Doraipandian, M. Chaotic Map Based Random Image Steganography Using LSB Technique. *Int. J. Netw. Secur.* **2017**, *19*, 593–598.
6. Kordov, K.; Stoyanov, B. Least Significant Bit Steganography using Hitzl-Zele Chaotic Map. *Int. J. Electron. Telecommun.* **2017**, *63*, 417–422. [CrossRef]
7. Stoyanov, B.P.; Zhelezov, S.K.; Kordov, K.M. Least significant bit image steganography bit image steganography algorithm based on chaotic rotation equations. *Mathematiques* **2016**, *69*, 845–850.
8. Aziz, M.; Tayarani-N, M.H.; Afsar, M. A cycling chaos-based cryptic-free algorithm for image steganography. *Nonlinear Dyn.* **2015**, *80*, 1271–1290. [CrossRef]
9. Almutairi, A. A Comparative Study on Steganography Digital Images: A Case Study of Scalable Vector Graphics (SVG) and Portable Network Graphics (PNG) Images Formats. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 170–175. [CrossRef]
10. Mondal, B.; Sinha, N.; Mandal, T. A Secure Image Encryption Algorithm Using LFSR and RC4 Key Stream Generator. In Proceedings of the 3rd International Conference on Advanced Computing, Networking and Informatics, Bhubaneswar, India, 23–25 June 2016; pp. 227–237.
11. Kar, N.; Mandal, K.; Bhattacharya, B. Improved chaos-based video steganography using DNA alphabets. *ICT Express* **2018**, *4*, 6–13. [CrossRef]
12. Dasgupta, K.; Mondal, J.K.; Dutta, P. Optimized video steganography using genetic algorithm (GA). *Procedia Technol.* **2013**, *10*, 131–137. [CrossRef]
13. Kordov, K.; Lachezar, B. Using Circle Map for Audio Encryption Algorithm. *Math. Softw. Eng.* **2017**, *2*, 183–189.
14. Bashir, Z.; Wątróbski, J.; Rashid, T.; Zafar, S.; Sałabun, W. Chaotic Dynamical State Variables Selection Procedure Based Image Encryption Scheme. *Symmetry* **2017**, *9*, 312. [CrossRef]

15. Mazurczyk, W. VoIP steganography and its Detection—A survey. *ACM Comput. Surv.* **2013**, *46*, 20. [CrossRef]
16. Mazurczyk, W.; Szczypiorski, K. Covert Channels in SIP for VoIP Signalling. *Commun. Comput. Inf. Sci.* **2008**, *12*, 65–72.
17. Forbes, C.R. A New Covert Channel over RTP. Master's Thesis, Rochester Institute of Technology, New York, NY, USA, 2009.
18. Geiser, B.; Vary, P. High rate data hiding in ACELP speech codecs. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 4005–4008.
19. Miao, H.; Huang, L.; Chen, Z.; Yang, W.; Al-hawbani, A. A new scheme for covert communication via 3G encoded speech. *Comput. Electr. Eng.* **2012**, *38*, 1490–1501. [CrossRef]
20. Zhijun, W.; Yongpeng, S. An implementation of speech steganography for iLBC by using fixed codebook. In Proceedings of the 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 14–17 October 2016; pp. 1970–1974.
21. Liu, P.; Li, S.; Wang, H. Steganography integrated into linear predictive coding for low bit-rate speech codec. *Multimed. Tools Appl.* **2017**, *76*, 2837–2859. [CrossRef]
22. Liu, P.; Li, S.; Wang, H. Steganography in vector quantization process of linear predictive coding for low-bit-rate speech codec. *Multimed. Syst.* **2017**, *23*, 485–497. [CrossRef]
23. Tian, H.; Liu, J.; Li, S. Improving security of quantization-index-modulation steganography in low bit-rate speech streams. *Multimed. Syst.* **2014**, *20*, 143–154. [CrossRef]
24. Nishimura, A. Data Hiding in Pitch Delay Data of the Adaptive Multi-Rate Narrow-band Speech Codec. In Proceedings of the 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kyoto, Japan, 12–14 September 2009; pp. 483–486.
25. Huang, Y.; Liu, C.; Tang, S.; Bai, S. Steganography integration into a low-bit rate speech codec. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1865–1875. [CrossRef]
26. Wu, Z.; Yang, W.; Yang, Y. ABS-based speech information hiding approach. *Electron. Lett.* **2003**, *39*, 1617–1619. [CrossRef]
27. Janicki, A. Pitch-based Steganography for speex voice codec. *Secur. Commun. Netw.* **2016**, *9*, 2923–2933. [CrossRef]
28. Huang, Y.F.; Tang, S.; Yuan, J. Steganography in inactive frames of VoIP streams encoded by source codec. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 296–306. [CrossRef]
29. Lin, R.S. High capacity information hiding scheme using VAD algorithm. In Proceedings of the 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Nantou, Taiwan, 27–29 May 2016; pp. 1–2.
30. Li, S.; Jia, Y.; Kuo, C.C.J. Steganalysis of QIM Steganography in Low-Bit-Rate Speech Signals. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**, *25*, 1011–1022. [CrossRef]
31. Lin, Z.; Huang, Y.; Wang, J. RNN-SM: Fast Steganalysis of VoIP Streams Using Recurrent Neural Network. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1854–1868. [CrossRef]
32. Ren, Y.; Yang, J.; Wang, J.; Wang, L. AMR Steganalysis Based on Second-Order Difference of Pitch Delay. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 1345–1357. [CrossRef]
33. Tian, H.; Wu, Y.; Chang, C.C.; Huang, Y.; Chen, Y.; Wang, T.; Cai, Y.; Liu, J. Steganalysis of adaptive multi-rate speech using statistical characteristics of pulse pairs. *Signal Process.* **2017**, *134*, 9–22. [CrossRef]
34. Miao, H.; Huang, L.; Shen, Y.; Lu, X.; Chen, Z. Steganalysis of compressed speech based on Markov and entropy. *Lect. Notes Comput. Sci.* **2013**, *8389*, 63–76.
35. Ren, Y.; Cai, T.; Tang, M.; Wang, L. AMR Steganalysis Based on the Probability of Same Pulse Position. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 1801–1811.
36. Malik, H.; Subbalakshmi, K.P.; Chandramouli, R. Nonparametric Steganalysis of QIM Steganography Using Approximate Entropy. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 418–431. [CrossRef]
37. Yang, W.; Tang, S.; Li, M.; Cheng, Y.; Zhou, Z. Steganalysis of Low Embedding Rates LSB Speech Based on Histogram Moments in Frequency Domain. *Chin. J. Electron.* **2017**, *26*, 1254–1260. [CrossRef]
38. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
39. Dual Rate Speech Coder FOR Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s. Available online: https://www.itu.int/net/itu-t/sigdb/speaudio/AudioForm-s.aspx?val=1117231 (accessed on 1 May 2018).

40. Menezes, A.J.; Vanstone, S.A.; Oorschot, P.C.V. *Handbook of Applied Cryptography*, 1st ed.; CRC Press Inc.: Boca Raton, FL, USA, 1996; ISBN 0-8493-8523-7.
41. Ibrahim, A.; Dragomir, S.S. A Survey on Cauchy–Bunyakovsky–Schwarz Inequality for Power Series. In *Analytic Number Theory, Approximation Theory, and Special Function*; Springer: New York, NY, USA, 2014; pp. 247–295.
42. Chang, C.C.; Lin, C.J. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [CrossRef]
43. Garrido, A. Symmetry and Asymmetry Level Measures. *Symmetry* **2010**, *2*, 707–721. [CrossRef]