MBNA Publishing House Constanta 2021

Proceedings of the International Scientific Conference SEA-CONF

SEA-CONF PAPER • OPEN ACCESS

State of the art for Machine Learning used in audio Steganalysis

To cite this article: M. PREDA, S. BÎRLEANU and C. RĂCUCIU, Proceedings of the International Scientific Conference SEA-CONF 2021, pg.203-206.

Available online at www.anmb.ro

ISSN: 2457-144X; ISSN-L: 2457-144X

doi: 10.21279/2457-144X-21-026

State of the art for Machine Learning used in audio steganalysis

M Preda, S Bîrleanu, C Răcuciu

Bucharest, Romania
mirela.preda@mta.ro

Abstract. Steganography represents a method of transmitting hidden information, applicable on a large scale of digital files (images, audio file, video, text). Steganalysis, the countermeasure of steganography, is purposed for detecting files with steganographic content. To identify these files, machine learning algorithms are used with the purpose of raising the efficiency and precision of the steganographic process. Currently, steganography and steganalysis techniques have raised an increased interest, especially regarding audio files. Based on the fact that these are used and known on a large scale, they represent an optimal solution for steganography. This study presents the State of the art for audio steganalysis instruments based on Machine Learning.

1. Introduction

Artificial intelligence has lately an increased presence in our lives. Machine learning is a part of artificial intelligence, which has two approach, depending on the labeling of the data: supervised learning and unsupervised learning. It can be used to detect hidden content in steganographic multimedia files. Most steganalysis tools have been identified for the image files but, there has been an increase of interest in audio files also.

2. Steganalysis methods

The purpose of steganalysis is to be able to build an optimal system to identify steganographic files, knowing the embedding method and the statistical model. In practice, this situation is not common, as this information missses and steganalysis systems needs to be developed using algorithms for feature extraction and machine learning [1].

The existing steganalysis methods based on machine learning are divided, according to the purpose, into three categories: for the residual calculation, for the extraction of the characteristics and for the classification in two classes (steganographic files or unaltered files) [2]. Increased attention must be paid to the selection of the most revealing and optimal features, because the accuracy of the detection depends on this aspect.

A steganalysis system is described in Figure 1.

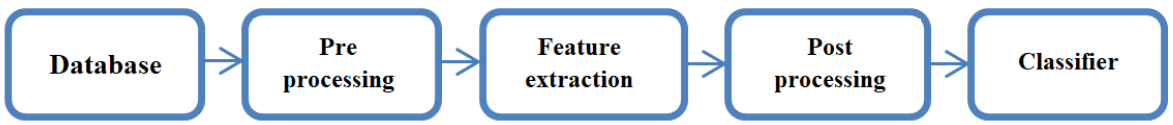

Figure 1. Stages of a steganalysis system

2.1 Based on CNN

CNN neural networks are the most representative models of deep learning and are used successfully in classifying images or audio files. However, for steganalysis the requirements are slightly different, because it is necessary to capture the artifacts introduced by steganography which are much weaker signals than the media content.

For steganographic algorithms in the field of entropy, conventional methods have difficulties in identifying the audio steganographic files, so in [3], CNN networks were used to steganalyze audio files in MP3 format.

Due to the wakness of the signal introduced by steganography comparing to the image content / audio content, applying CNN directly to the image / audio data may suffer due to the negative impact of the content, thus leading to a weak local minimum of the trained model. One such CNN network is RHFCN [4], made for steganalysis of MP3 files, which contains an HPF (High-Pass Filter) module, which highlights the traces left by the hidden content, so that the network becomes more efficient in detecting stego files. Another change for the optimization of the network is the replacement of fully connected (Fc) layers with convolutional layers.

In order to capture the artifacts left by the steganography, we proceeded similarly to the image analysis, where a high-pass filter is applied on the input image [5]. A steganalytical system is developed for audio files in which CNN networks are used. The proposed scheme is efficient for detecting steganographic models in the time domain, but which is inefficient for audio files in MP3 or AAC format.

The analysis of the LSB and STC algorithms (Syndrome-Trellis Code) was performed with a CNN-based steganalysis system [6], with an improved version of the convolutional layer, in order to bring an increased sensitivity of the detection of weak signals. Another change is the replacement of Rectified Linear Unit (ReLU) with Truncated Linear Unit (TLU) as the activation function.

2.2. Based on SVM

A high amount of documentation was found in the specific literature presenting steganalytic systems based on SVM.

In [7] is described a system that detects audio steganographic files, which were generated with StegHide software. The classification of the files in the two classes (clean and stego) was made using the SVM classifier, trained with five types of steganographic files, depending on the capacity (5%, 10%, 20%, 40% and 60%). The system can detect files that have steganographic content with a capacity of 5%. The smaller the amount of hidden content inserted, the harder it is to detect.

The first published study [8] that uses deep residual network to extract classification features from audio files and an SVM classifier that evaluates the performance of feature selection, has achieved remarkable results for AAC and MP3 audio files. The spectogram (representation of the frequency spectrum) with windows of different sizes is used to extract the universal characteristics introduced by the steganographic schemes. A better accuracy was noticed in detecting steganographic files for three main embedding domains concerning the compression parameters of AAC: Modified Discrete Cosine Transform coefficients (MDCT), the scale factor and Huffman coding.

In [9], derivate-based and wavelet-based methods have better results than the method presented by Kraetzer and Dittmann. These methods were tested on steganographic files made with the Hide4PGP, Invisible Secrets, Steghide and LSB steganographic tools achieved better results compared to a signal-based mel-cepstrum method. Compared to the two new methods, on average, the derivative-based solution is superior to the wavelet-based method. In this method, core-based SVM with radial base function (RBF) is used to detect and differentiate steganograms from unaltered signals.

Another steganalytical algorithm is described in [10], which identifies files generated with four steganographic methods, Direct Sequence Spread Spectrum (DSSS), Quantization Index Modulation (QIM), ECHO embedding (ECHO), and Least Significant Bit embedding (LSB). The efficiency of this

system is 85%, and the characteristics processed in this paper are in the frequency domain, based on Short Time Fourier Transform (STFT).

The universal steganalysis algorithm described in [11] is created also by using frequency domain features and two approaches are used. An ANOVA approach (analysis of variance) and an SFFS approach (sequential forward floating search method), each complemented by two classification methods. Each approach was combined with linear regression (LR) and supports vector machines (SVM). After the four combinations of ANOVALR, ANOVA-SVM, SFFS-LR and SFFS-SVM, it was found that the SFFS approach and the choice of an SVM classifier bring superior results.

However, in case of an universal steganalysis, in which the steganographic algorithm used is not known, it is not recommended to use an SVM classifier, because it can only detect the algorithms used in training sets, being a type of supervised learning. - AAST (AMSL Audio Steganalysis Toolset) [12].

2.3. Based on ANN

The artificial neural network (ANN) is a nonlinear dynamic system that has advantages such as self-adaptation, self-organization and self-learning.

The most common steganographic technique is LSB (least significant bit), and a type of steganalytical algorithm for this method is described in [13]. This type of algorithm detects audio files in MP3 format using machine learning, more precisely, a multilayer perceptron with three layers (an input, a hidden and an output layer). Each layer in the network is built on a group of neurons, and each neuron is the exit of the first layer and the entry of the next layer. The proposed system proved to be efficient for MP3 files smaller than 15MB, but it represents a punctual steganalysis solution, being applied only for one method.

In [14] wavelet features are used to identify audio files that have been incorporated with LSB, QIM (Quantization index method) and AM (addition method) steganographic algorithms. For file classification, an RBF network based on radial functions was chosen. This method reduces the size of the feature vector and simplifies the projection of the classifier, proving a better performance than the system [7], where an SVM classifier was used.

In [15] is described a steganalytical technique that has been tested on three types of audio steganography (StegHide, Hide4PGP and DWT-FFT), with better results especially in files made with Hide4PGP or speech files. This technique uses a complex classifier, called the Autoregressive Time Delay Neural Network (AR-TDNN), which had an efficiency rate of 82.73%. The AR component of the network offers the possibility of recognizing a sequence of previously learned sequences, thus making the classification decisions.

3. Conclusion

Machine learning algorithms are increasingly used and have obtained better performances compared to classical classification methods. Based on this State of the art, there is an increased performance of the use of artificial ANN networks, which better performed compared to SVM classifiers. A classifier that uses ANN networks could be a good candidate for developing an universal steganalytic tool. Even if CNN networks are used successfully in classifying images or audio files, in the case of steganalysis, the results are not the most accurate. With the advancement of steganalysis algorithms based on machine learning, the detection of traces left by steganographic algorithms is facile, but an important aspect remains selecting the best representative features for classification.

References

[1] Ghasemzadeh H, Kayvanrad M. Comprehensive review of audio steganalysis methods. *IET Signal Processing*. 2018;12(6):673-687.

[2] Lin Y, Wang R, Yan D, Dong L, Zhang X. Audio Steganalysis with Improved Convolutional Neural Network. *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, 2019.

[3] Wang Y, Yang K, Yi X, Zhao X, Xu Z. CNN-based Steganalysis of MP3 Steganography in the

Entropy Code Domain. *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, 2018.
[4] Wang Y, Yi X, Zhao X, Su A. RHFCN: Fully CNN-based Steganalysis of MP3 with Rich High-pass Filtering. *ICASSP - IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2019.
[5] Chen B, Luo W, Li H. Audio Steganalysis with Convolutional Neural Network. *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*, 2017.
[6] Lin Y, Wang R, Yan D, Dong L, Zhang X. Audio Steganalysis with Improved Convolutional Neural Network. *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, 2019.
[7] Xue-Min Ru, Hong-Juan Zhang, Xiao Huang. Steganalysis of audio: attacking the Steghide, *International Conference on Machine Learning and Cybernetics*, 2005.
[8] Ren, Yanzhen, D. Liu, Qiaochu Xiong, Jianming Fu and Lina Wang. Spec-ResNet: A General Audio Steganalysis scheme based on Deep Residual Network of Spectrogram. *ArXiv* abs/1901.06838, 2019.
[9] Qingzhong Liu, Sung A, Mengyu Qiao. Temporal Derivative-Based Spectrum and Mel-Cepstrum Audio Steganalysis. *IEEE Transactions on Information Forensics and Security*. 2009;4(3):359-368.
[10] Wei Y, Guo L, Wang Y, Wang C. A blind audio steganalysis based on feature fusion. Journal of Electronics (China). 2011;28(3):265-276.
[11] Özer, H., B. Sankur, N. Memon and I. Avcibas. Detection of audio covert channels using statistical footprints of hidden messages. *Digit. Signal Process.* 16 (2006): 389-401.
[12] Krätzer, Christian and J. Dittmann. Pros and Cons of Mel-cepstrum Based Audio Steganalysis Using SVM Classification. *Information Hiding* (2007).
[13] Alhaddad M, Alkinani M, Atoum M, Alarood A. Evolutionary Detection Accuracy of Secret Data in Audio Steganography for Securing 5G-Enabled Internet of Things. *Symmetry*. 2020;12(12):2071.
[14] Jian-Wen Fu, Yin-Cheng Qi, Jin-Sha Yuan. Wavelet domain audio steganalysis based on statistical moments and PCA. *International Conference on Wavelet Analysis and Pattern Recognition*, 2007.
[15] Rekik S, Selouani S-A, Guerchi D, Hamam H. An Autoregressive Time Delay Neural Network for speech steganalysis. *11th International Conference on Information Science, Signal Processing and their Applications (ISSPA),* 2012, http://dx.doi.org/10.1109/isspa.2012.6310612