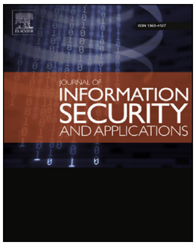
Available online at www.sciencedirect.com

ScienceDirect

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



Steganalysis based on steganography pattern discovery

Hedieh Sajedi *

Department of Computer Science, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran

ARTICLE INFO

Article history:
Available online 19 April 2016

Keywords:
Steganography
Evolutionary fuzzy rules
Steganalysis
Blind steganalysis

ABSTRACT

The goal of steganalysis algorithms is detection of stego images from clean images. Each steganography method based on its embedding mechanism puts a special pattern on the stego images. Finding this pattern in the images leads us to employ a classifier to be constructed specially for detecting stego images which are the results of a special steganography algorithm. In this paper, to have high detection accuracy, we propose an approach for Steganography Pattern Discovery (SPD). Our proposed approach employs an evolutionary method to extract the signature of stego images against clean images via fuzzy if–then rules. Based on the discovered knowledge, suitable trained models for steganalysis can be employed and stego images will be detected with high accuracy. Using SPD, we can predict the type of steganography method from a stego image. Employing SPD can enhance the approaches, which assume that a special steganography method is used. The effect of SPD before applying steganalysis methods has been investigated by some steganography and steganalysis techniques and it has been validated using some image databases. The results indicate that the pattern of a steganography method is extracted well and the type of steganography method used to make a stego image can be predicted with high accuracy.

1. Introduction

Steganography is the science of imperceptible communications. In cryptography, the attacker is able to identify, catch, and change the transmitted information (Kahn, 1996), nevertheless, steganography is used when we need to hide the existence of communicating. Steganography methods embed secret messages within visually innocent covers. Typical medias that can cover secret messages are image, video, and audio files (Munuera, 2007).

Invisible ink, covert channel, microdot, and spread-spectrum communication are some famous and ancient steganographic methods (Kahn, 1996; Norman, 1973). A famous classic steganographic model is the prisoners' problem. In this problem, Alice and Bob are in a jail and they plan to escape together (Simmons, 1984). The communications between them are monitored by Wendy, who is a warden. In this regard, they must hide the secret messages in another innocuous-looking means (cover object) to achieve the stego object. Afterward, the stego object is sent through the public channel. For more explanation about applications of steganography method, refer to Simmons (1984) and Westfeld and Pfitzmann (1999).

The fundamental requirement of steganographic systems is that the stego object should be perceptually indistinguishable to the degree that it does not raise suspicion. In other words, the hidden information should introduce only slight modification to the cover object (Wu and Shih, 2006).

Various image steganography methods have been proposed in the literature. Due to the great use of JPEG images,

* Department of Computer Science, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran. Tel.: 009821 61112915.
E-mail address: hhsajedi@ut.ac.ir.

http://dx.doi.org/10.1016/j.jisa.2016.04.001

embedding in Discrete Cosine Transform (DCT) domain is well known. Steganography methods like F5 (Westfeld, 2001), Model-based (MB) (Sallee, 2003), Perturbed Quantization (PQ) (Fridrich et al., 2004), and YASS (Solanki et al., 2007) embed secret messages in images by modifications of carefully chosen DCT coefficients. In addition, some methods have been proposed which embed messages in other transform domains, such as Contourlet transform (Sajedi and Jamzad, 2008). The method presented in Sajedi and Jamzad (2008) embeds secret messages in Contourlet coefficients of a cover image.

Adaptive steganography schemes like WOW (Holub and Fridrich, 2012) started with the advancement of coding schemes (Fridrich et al., 2014) capable of embedding messages while nearly optimally minimizing arbitrarily defined additive distortion functions. Since the capacity of steganography methods is limited based on the properties of images, and the goal of the steganography methods is to be undetectable, the research in Sajedi and Jamzad (2009a, 2010a, 2010c, 2010d) is about increasing the embedding capacity of steganography methods to provide the capability of embedding larger secret messages.

In Qin et al. (2013), a prediction-based reversible steganographic scheme based on image inpainting is proposed. Another reversible data-hiding scheme in encrypted image is proposed in Qin and Zhang (2015). This scheme has better decrypted image quality and higher image recovery accuracy. In Zhang and Wang (2006), a method of steganographic embedding in digital images is proposed, in which each secret digit in a $(2n+1)$-ary notational system is carried by n cover pixels and, at most, only one pixel is increased or decreased by 1. In other words, the $(2n+1)$ different ways of modification to the cover pixels correspond to $(2n+1)$ possible values of a secret digit. Because the directions of' modification are fully exploited, this method provides high embedding efficiency.

In Qin et al. (2014), a joint data-hiding and compression scheme is presented for digital images using side match vector quantization (SMVQ) and image inpainting. The two functions of data hiding and image compression can be integrated into one single module seamlessly. On the sender side, except for the blocks in the leftmost and topmost of the image, each of the other residual blocks in raster-scanning order can be embedded with secret data and compressed simultaneously by SMVQ or image inpainting adaptively according to the current embedding bit.

In Qin et al. (2015), a data-hiding scheme with reversibility based on exploiting modification direction (EMD) is proposed. One cover image is first chosen and prepared to generate two visually similar steganographic images. During the secret embedding, the pixels in the first steganographic image are modified by no more than one gray level to embed secret data using the traditional EMD method, while the pixels in the second steganographic image are adaptively modified through referring to the first steganographic image without any confusions in image recovery process. On the receiver side, secret data can be extracted easily and the original cover image can also be recovered from the two steganographic images correctly.

In Sajedi and Jamzad (2009b) an approach for selecting proper cover images in steganography is presented. This approach consists of two stages. The first stage is an evolutionary algorithm that extracts the signature of cover images against stego images in the form of fuzzy if–then rules. In the second stage, the fuzzy rules are used for selecting suitable cover images for steganography. This approach selects the appropriate cover images from an image database and using them produces more secure steganography.

Due to the various contents of images, the stego images produced by a steganography method may have different levels of undetectability against steganalyzers. In other words, a steganography method may cause less detectable statistical artifacts on some images compared to other images. In Sajedi and Jamzad (2010e), different features of images are analyzed to find the similarity between proper cover images for each steganography method. Similarity between images is modeled in the form of fuzzy rules. Subsequently for hiding secret data in a cover image, a reliable steganography method is suggested in Sajedi and Jamzad (2010e) that results to an undetectable stego image against steganalysis methods.

In the current paper, the idea of extracting fuzzy rules in a similar way of Sajedi and Jamzad (2009b, 2010e) is used to reveal the signature of steganography methods and enhance the performance of steganalysis algorithms.

Steganalysis algorithms try to distinguish stego images from clean images. Generally, a classifier is built based on stego and clean image. In condition of observing a new image, we do not have any information about the used steganography method or the payload. Therefore, a general steganalyzer is built using a set of clean images and a set of stego images generated by various steganography algorithms and different payloads.

On the other words in popular steganalysis methods, the important issue is to detect the existence of the hidden information. They do not consider different patterns of steganography algorithms. Therefore, the classifier should learn a complex function that can distinguish clean images from stego images with various steganography patterns.

Each steganography method employs a special mechanism to embed secret data in the images. Therefore, it puts a distinct pattern on the stego images. Discovering this pattern in the images leads us to hire a proper classifier to be constructed particularly to detect stego images which are the results of a special steganography algorithm.

In this paper, we present an approach that consists of two stages. In the first stage, we analyze an image database to discover the pattern or signature of stego images. By the pattern, we mean the effective features of stego images and their relative values. This pattern is constructed in the form of a set of fuzzy if–then rules that represent the similarity between stego images. In the second stage, the steganalyzer is trained to detect only one steganography method at once. After discovering patterns of the used steganography method from the stego image, the proper model is used to analyze it. This approach simplifies the problem of blind steganalysis to partially blind steganalysis.

The process of generating the signature of stego images is done by an Evolutionary Algorithm (EA). EAs have been used as rule generation and optimization tools in the design of fuzzy rule-based systems (Cordon et al., 2004; Hu et al., 2003).

To obtain accurate fuzzy rules, we employ an evolutionary rule generation algorithm based on Iterative Rule Learning (IRL) approach (Sajedi and Jamzad, 2009b, 2010e). The rules are generated incrementally so that the evolutionary algorithm optimizes one fuzzy rule at a time. The generated fuzzy rules are