Chapter 23
A Steganalytic Scheme Based on Classifier Selection Using Joint Image Characteristics

Jie Zhu
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China & School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Xianfeng Zhao
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China & School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Qingxiao Guan
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China & School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Yun Cao
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China & School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Gong Chen
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China & School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

ABSTRACT

Steganalysis relies on steganalytic features and classification techniques. Because of the complexity and different characteristics of cover images, to make steganalysis more applicable toward detecting stego images in real applications, we need to train different classifiers so as to match different images according to their characteristics. Selection of classifiers according to characteristics of images is the key point to improve accuracy of steganalysis. In our work, we study the methods of classifier selection based on characteristics of images including image size, quantization factor, or matrix. Besides, we also discuss other characteristics, such as texture, cover source, which makes an appreciable difference to steganalysis.

DOI: 10.4018/978-1-7998-2466-4.ch023

INTRODUCTION

Nowadays, steganography has been attracting much attention and widely studied by the researchers all over the world. It aims to not only embed the message in digital files for covert communication but also conceal the act of hiding the secret. Meanwhile, the goal of steganalysis is to find out the suspicious digital files, the related actor and even the embedded message. Steganalysis that aimed to recognize the existence of stego image files is based on feature space designed to represent the image and binary classifier trained for distinguishing the stego image file (Fridrich, 2009). It has been proved that the detector well trained by large training set and rich model can achieve outstanding accuracy with the assumption that the source of test set is the same as the training data (Kodovsky, Fridrich, & Holub, 2012, Fridrich & Kodovsky, 2012, Kodovský & Fridrich, 2012, Goljan, Fridrich, & Cogranne, 2014, Holub & Fridrich, 2015). Although, the complexity of the image properties itself has a huge influence on the accuracy of the staganalytic detector in the real-world application, such as the size, the quality factor, the camera producing the raw image file, the double compression. While the training data is different from testing set, the performance of the detector decreases significantly (Goljan, Fridrich, & Holotyak, 2006, Kodovský, Sedighi, & Fridrich, 2014, Ker & Pevný, 2014). This is what we call "cover source mismatch", which has been recognized as an open problem considering the steganalysis in practice (Ker, Bas, Böhme, Cogranne, Craver, Filler, Fridrich, & Pevný, 2013). The definition of Cover Source Mismatch is given by Kodovský et al., and the negative impact of CSM is widely recognized and can range from small decrease of performance to complete fail when one relatively accurate steganalysis algorithms on one image source detects the steganography on another source (Kodovský, Sedighi, & Fridrich, 2014).

In recent years, many researchers focus on cover source mismatch, and discuss different targeted approaches to address the problems of mismatch in different domain of digital image. Barni et al. (2010) considered the use of an image forensics tool for the steganalysis of images produced by different sources. The experiments are conducted to analyze two types of image: computer generated and camera images (Barni, Cancelli, & Esposito, 2010). Lubenko & Ker (2012) proved the simple classifiers have more robustness to train mismatch. Fridrich (2013) studied the effect of the cover quantization on steganalysis. However, the study is limited to the situation where the source of training set matches the testing set. Kodovský et al. (Kodovský, Sedighi, & Fridrich, 2014) studied the effectiveness of two simple approaches: training a single classifier on a mixture of sources and training a bank of classifiers with different sources first and then testing a given unseen source on the closest source used for training. Both can mitigate the negative effect of mismatch, however, selecting a closest source based on the camera is not realistic since it demands an unacceptable number of the classifiers because of various camera available and the custom quantization tables they used. Ker & Pevný (2014) presents an in-depth study of one particular instance of model mismatch, and demonstrates some effective methods to considerably reduce rather than completely remove the mismatch penalty. However, they only discussed a single type of steganography and detector.

In this paper, we proposed a novel steganalytic scheme based on classifier selection using image characteristic for JPEG domain. Motivated by Kodovský et al. (Kodovský, Sedighi, & Fridrich, 2014) and aimed at the real-world application, we firstly trained a bank of classifiers, each of which is called template, according to the essential characteristics of image: the size and the quantization table or JPEG quality factor. These characteristics can be directly read from image file to compute the closest trained classifier. Then, we select the most suitable classifier for every testing image after compute the similarity between the testing image and every template. The experiments settings are based on the real-world