Digital Investigation 12 (2015) 1–14

Contents lists available at ScienceDirect

Digital Investigation

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

A study on JPEG steganalytic features: Co-occurrence matrix vs. Markov transition probability matrix

Jicang Lu [a, c, *], Fenlin Liu [a, b], Xiangyang Luo [a, c]

[a] Zhengzhou Information Science and Technology Institute, Zhengzhou, Henan 450001, China
[b] State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, Henan 450001, China
[c] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:
Received 4 November 2014
Accepted 7 December 2014
Available online 27 December 2014

Keywords:
Steganalysis
Feature comparison
JPEG
Co-occurrence matrix
Markov transition probability matrix

ABSTRACT

Statistical feature selection is a key issue affecting the performance of steganalytic methods. In this paper, a performance comparison method for different types of image steganalytic features was proposed firstly based on the changing rates. Then, for two types of typical steganalytic features – co-occurrence matrix and Markov transition probability matrix, the performances of them were discussed and theoretically compared for detecting two types of well-known JPEG steganography that preserve DCT coefficients histogram and lead the histogram to shrink respectively. At last, a conclusion on the sensitivity comparison between components of these two types of features was derived: for the steganography that preserve the histogram, their sensitivities are comparable to each other; whereas for the other one (such as the steganography that subtract 1 from absolute value of the coefficient), different feature components have different sensitivities, on the basis of that, a new steganalytic feature could be obtained by fusing better components. Experimental results based on detection of three typical JPEG steganography (F5, Outguess and MB1) verified the theoretical comparison results, and showed that the detection accuracy of the fused new feature outperforms that of existing typical features.

Introduction

Image steganography and steganalysis is one of the important topics in multimedia security Ker et al. (2013). For various types of steganographic algorithms Cheddad et al. (2010), there currently have been many steganalytic algorithms Luo et al. (2008), which are mainly concentrated in two aspects: targeted steganalysis and universal blind steganalysis. The universal blind steganalysis is one of the important aspects of steganalysis, and the detection performance relies mainly on statistical features extracted from the detected objects. Currently, there are various types of statistical features for steganalysis, such as: image quality metrics İsmail Avcbaş et al. (2003), characteristic function (CF) moments Kodovský and Fridrich (2005), probability density function (PDF) moments Lyu and Farid (2006), co-occurrence matrix (CM) Liu et al. (2011), Markov transition probability matrix (MTPM) Fu et al. (2006), etc. However, current steganalyzers have usually selected features based on experiments, but lacking reliable theoretical basis, which makes the detection results somewhat fortuity and blindness. Therefore, the theoretical analysis and selection of more suitable statistical features for steganalysis is of great significance, both theoretical and practical, to decrease blindness of feature selection and improve reliability of detection results.

At present, the researches on steganalytic feature analysis and selection mainly include: feature dimensionality reduction based on components analysis, and feature selection based on performance comparison. The former

* Corresponding author. Zhengzhou Information Science and Technology Institute, Zhengzhou, Henan 450001, China.
E-mail address: lujicang@sina.com (J. Lu).

http://dx.doi.org/10.1016/j.diin.2014.12.001
1742-2876/

focuses on reducing dimensions of the feature by removing redundant components to improve the classification efficiency while maintain the detection accuracy, the frequently used methods include analysis of variance (ANOVA) İsmail Avcbaş et al. (2003), Bhattacharyya distance measure Xuan et al. (2006b), Mahalanobis distance measure Davidson and Jalan (2010), principle component analysis (PCA) Qin et al. (2009), Fisher criterion Lu et al. (2014), etc. In a certain sense, the dimensionality reduction is a feature components optimization problem, but does not care about the comparison between different types of features. In the aspect of steganalytic feature selection based on performance comparison, because it is difficult to establish mutual representation between different types of features, so, existing researches usually select features based on empirical analysis, even in the significantly successful steganalytic methods such as Fridrich and Kodovský (2012); Vojtěch Holub (2013); Cogranne et al. (2014). There are only a few methods based on theoretical analysis currently. Wang and Moulin (2007) made a landmark research on theoretical based feature comparison and selection. In Wang and Moulin (2007), a theoretical conclusion is derived for the first time by taking the wavelet high frequency subband as feature extraction source, which indicates that, the absolute CF moments of the coefficients histogram are more suitable for steganalysis compared with the absolute PDF moments. Then, we Luo et al. (2011) extended their conclusions to other multiple feature extraction sources such as wavelet prediction subband, wavelet prediction error subband and wavelet subband of noise, and further proved that, as long as the feature extraction source follow Gaussian distributions with mean of 0 before and after steganography, the absolute CF moments of wavelet coefficients histogram always outperform the absolute PDF moments. Especially, for the wavelet log prediction error subband, the theoretical basis that the first order PDF moments is superior to the first order CF moments under the condition of limit precision is presented, which verified the experimental conclusion in Wang and Moulin (2007).

This paper focuses on the problem of theoretical based performance comparison between different types of statistical features used for steganalysis, and the main contributions are as follows:

1) **A feature comparison method**. For the problem of performance comparison between different types of steganalytic features, a comparison method is proposed based on the changing rates of the features before and after embedding. Then, a feature fusion method is proposed based on the comparison results. By analyzing the changes of the features before and after embedding, the comparison conclusion is obtained: the feature vector fused by feature components with better sensitivities is always superior to the feature vector fused by relatively feature components with less well sensitivities.

2) **Comparison of co-occurrence matrix and Markov transition probability matrix**. Firstly, the detailed comparison method for steganalytic features CM and MTPM is presented based on the proposed feature comparison method. Then, for two types of well-known JPEG image steganography: the steganography that lead the DCT coefficients histogram to shrink (Referred to as histogram shrinking steganography, e.g., F5 Westfeld, 2001), and the steganography that preserve the DCT coefficients histogram (Referred to as histogram preserving steganography, e.g., Outguess Provos, 2001 and MB1 Sallee, 2003), the sensitive changing of these two types of features are compared based on theoretical analysis, and the performance comparison conclusions of them used for steganalysis are derived. At last, a new feature with better performance is obtained based on the fusion of selected feature components.

The rest of this paper is organized as follows. Co-occurrence Matrix (CM) and Markov Transition Probability Matrix (MTPM) will briefly introduce the features CM and MTPM that used for steganalysis. In A Comparison Method for Steganalytic Features, a performance comparison method will be proposed based on the analysis of features changing rate. The detailed theoretical based comparison between CM and MTPM will be presented in Feature Comparison: CM vs. MTPM, and then, a feature components fusion method is presented according to the obtained comparison results. Experimental Results and Analysis will report the experiments to verify and analysis the comparison results. The paper is concluded in Conclusions.

Co-occurrence matrix (CM) and Markov transition probability matrix (MTPM)

For natural images, there is always strong dependence between neighboring image data (such as pixels or coefficients). CM can reflect the distribution characteristics of the entire data space by measuring the joint distribution probability of neighboring data, and can illustrate the dependence well. In early times, it was usually used to describe the textual characteristics of digital images Haralick et al. (1973). MTPM is another type of statistical feature that is to characterize the strong dependence between neighboring image data. Sullivan et al. (2006) indicate that: although the MTPM is complex compared with the independent identical distribution that reflects single data distribution, but it is the least complex model incorporating dependencies between neighboring data.

In image steganography, the strong dependence between neighboring image data will usually be broken by the stochastic modifications of image data during embedding messages, which will lead to the change of features CM and MTPM. Based on the above phenomenon, these two types of features are extracted by Fridrich (2004) and Sullivan et al. (2005, 2006) for steganalysis, respectively, which are the first use of them in image steganalysis. Now, there are a large number of steganalytic algorithms based on these two types of features respectively. For example, in the CM feature-based steganalysis, the detection algorithms are proposed for frequency domain steganogrpahy in Xia et al. (2010); Kodovský and Fridrich (2012); Zong et al. (2012), while for spatial domain in Chen et al.