11

Improve Steganalysis by MWM Feature Selection

B. B. Xia, X. F. Zhao and D. G. Feng
Institute of Software Chinese Academy of Sciences
China

1. Introduction

Steganography is the art of invisible communication. It is derived from the ancient Greece thousands of years ago (Johnson & Jajodia, 1998; Kahn, 1996), and grow rapidly in the past few years along with the development of digital technology and the internet. Modern steganography has been widely used in various scenarios such as secret communication, digital rights management, data temper detection and recovery, etc (Provos & Honeyman, 2003). The main goal of modern steganography is to hide some secret messages into the so-called cover data, e.g. images, videos, audios, documents…, which produced the so-called stego data, and make the existence of the hidden messages unnoticeable to everyone expect the prospective receiver (Provos, 2001; Fridrich & Goljan, 2002; Fridrich, 2005). Usually an embedding key is also involved in the steganography scheme to provide security. The malicious can never tamper, remove, nor obtain the secret messages in the stego data, as long as the embedding key is kept unknown.

In contrast to steganography, steganalysis is developed to detect the presence of the secret messages, and furthermore estimate the length or even extract the content of the embedded messages. Although the secret message in stego data is always imperceptible to human's visual, the embedding process changes some statistics of the cover medium nevertheless. This can be utilized by steganalysis methods to distinguish stego mediums containing secret messages from the clean cover mediums (Lyu & Farid, 2002; Fridrich et al., 2002; Fridrich, 2005; Harmsen & Pearlman, 2003; Ker, 2005; Tzschoppe & Aauml, 2003; Xuan et al., 2005).

Steganalysis methods can be roughly divided into two categories: the targeted (or specific) steganalysis which detects a particular known steganography method, and the blind (or universal) steganalysis which can deal with a wide variety of steganography methods. Though the targeted methods often have slightly better accuracy and efficiency than the blind ones, it is quit reasonable to assume that the nature of the cover data and the embedding method used for steganography is unknown to the analysers beforehand. Therefore, blind steganalysis based on learning and classifying are more valuable from a practical point of view (Fridrich & Goljan, 2002; Provos & Honeyman, 2003; Tzschoppe & Aauml, 2003).

The typical framework of blind steganalysis is a procedure of two-class classification, which consists of training and classifying. First, a set of statistics called steganalysis features is

extracted from a pair of training set which contains cover and stego mediums respectively. A classifier is then trained by these extracted features. Then, given the medium under test, the steganalysis features is extracted similarly and input to the classifier to decide whether it contains hidden messages.

The choice of steganalysis features is crucial to the classification accuracy. As mentioned earlier, embedding messages in cover medium will change some of the statistics. It is obvious that choosing statistics which are sensitive to the steganography embedding process will provide better steganalysis accuracy. The statistic moments and transition probabilities have been proved to be more efficient than other choice for a wide variety of steganography methods, and so are used frequently in modern steganalysis (Davidson & Jalan, 2010; Fridrich et al., 2002; Lyu & Farid, 2002; Pevný et al., 2010a; Pevný & Fridrich, 2007).

To improve the steganography security, some embedding methods attempt to maintain the statistics of cover medium by means of minimizing the embedding distortion. Encoding methods such as matrix encoding or wet paper encoding are implemented to steganography process to reduce the embedding distortion (Fridrich et al., 2004; Westfeld, 2001). LSB(Least Significant Bit) matching method solves the imbalance problem introduced to the sample value histogram by the original LSB replacement method, and thus provide good security against steganalysis based on 1*st* order statistic features (Mielikainen, 2006). An adaptive steganography called HUGO (Highly Undetectable Steganography) is proposed recently. Before embedding secret messages into a cover image, HUGO determines a distortion measure for each pixel by calculating a weighted sum of difference between the features derived from cover and stego images. 1*st* and 2*nd* order transition probabilities of SPAM steganalysis features are chose as the features (Pevný et al., 2010a), which makes HUGO undetectable using steganalysis methods based on 1*st* and 2*nd* order statistic features. The details of this algorithm can be found in (Pevný et al., 2010b).

As steganography methods try to reduce embedding distortion and preserve representation of covers approximately for low-order statistic features, it is natural that steganalysis takes one more step in the same direction. That is to say, higher-order statistic features should be used as steganalysis features for better classification accuracy. However, this leads to a catastrophic growth of the amount of feature dimensions. Recently, Gul & Kurugollu propose a 1237 dimension feature set constructed by k-variate probability distribution function (PDF) estimates (Gul & Kurugollu, 2011). Fridrich et al. suggest a final HOLMES feature set that consists of 33,963 features to obtain better accuracy against HUGO (Fridrich et al., 2011).

The increasing dimensions of features bring new challenges to steganalysis. Training classifiers in high dimensions requires relatively large number of samples. With the significant growth of the feature dimensions, it becomes harder or even impossible to obtain sufficient samples. Furthermore, the computational complexity of training the classifier on a large-scale training set in high-dimensional spaces also becomes prohibitive.

Feature selection is a typical method to deal with the excessive feature dimensions. Feature selection not only reduces the number of dimensions, but also removes the inefficient or redundant features, leaves the efficient ones to the classifier for better training and classification. The theoretical ideal way of feature selection is exhaustive searching all the possible combination of feature dimensions, which can not be achieved in practical scenario.

Miche et al. (Miche et al., 2006) uses a fast classifier called K-Nearest-Neighbors combined with a forward selection method to achieve feature selection, which is still limited to deal with the relative low dimension feature sets. Dong et al. (Dong et al., 2008) make use of the Boosting Feature Selection (BFS) algorithm as the fusion tool to select a subset of original statistic features. Each dimension of the original features is treated as a weak classifier, and the output of BFS classification is calculated for each of them as an evaluation indicator. The final classifier is then constructed by every weak classifier with the evaluation indicator as their weight. Gul & Kurugollu (Gul & Kurugollu, 2011) also establish an evaluation indicator for each single dimension of the original features by means of calculating the co-variance between features and the embedding rates. After that, the original features are sorted corresponding to their co-variance in the decreasing order. The best K features are then determined to form the final classifier, by adding all the features one by one to the classifier and test the performance. Fridrich et al. (Fridrich et al., 2011) propose another method of ensemble classifier to reduce the dimension of steganalysis classifier. The ensemble classifier consist of weak base learners which constructed by randomly choose subsets of the original features. The dimensionality of each base learner is significantly smaller than the full dimensionality of the original feature set. The final decision is obtained by fuse the result of all base learners together under certain voting rule.

In this chapter, we present a novel methodology called MMD-weighted-MI (MMD, Maximum-Mean-Discrepancy; MI, Mutual-Information) feature selection to deal with high-dimensional steganalysis features. Before training the classifier, a MMD-weighted-MI (MWM) value is calculated and assigned to each dimension of the original features by evaluating the distribution of the extracted features using the MI and MMD indicators. The MI and MMD are both efficient measurements used exclusively in steganography benchmarking, but focus on different aspects of the feature distribution (Pevný & Fridrich, 2008). The MMD gives an overall view of a subset of the features, evaluates the difference of feature distribution between cover and stego training sets by means of generate a set of functions from the kernel ones and then calculate the maximum mean discrepancy. On the other hand, the MI, which is calculated only for single feature dimension due to its unacceptable complexity introduced by estimation of high-dimensional distribution, gives more details about how each dimension contributes to the classifier in steganalysis. When combined together as MWM values, these two indicators can give us a more comprehensive impression about difference between features extracted from the cover and stego training sets. After the MWM values are assigned, feature selection is simply implemented by choosing feature dimensions with high value.

The organization of this chapter is as follows: Section 2 introduces some basic concepts of MI and MMD, as well as elaborates their different effect in feature selection briefly. Section 3 describes the proposed approach of MWM feature selection. Experimental results are presented in Section 4. The chapter is finally concluded in Section 5.

2. Basic concepts of MI and MMD

In this section, we explain the basic concepts of MI and MMD. These two measurements have been used in steganography benchmarking due to their characteristic of evaluating the difference between two distributions, which makes them natural candidates for steganalysis feature selection.

Without loss of generality, images are chose as the cover medium in the following discussion. The same result holds for other form of mediums such as videos, audios and documents.

2.1 Mutual Information (MI)

Denote X the whole set of images corresponding to the steganalysis system. X can be divided into two non-overlap subsets, namely cover set C and stego set S respectively. Denote P and Q the distribution of the cover and stego set, with p and q as the probability distribution function (pdf) respectively. Then the difficulty of distinguishing stego images from cover ones can be measured using statistic called Kullback-Leibler divergence (Cachin, 1998)

$$KL(P||Q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx \tag{1}$$

where x denotes the sample medium drawn from the whole set of image X.

The KL divergence is a fundamental quantity for steganography benchmarking, which provides good estimate to the difference between cover and stego sets for certain features (Cover & Thomas, 2001). However, the asymmetry in calculating the KL divergence becomes a main drawback. From (1), it is obvious that

$$KL(P||Q) \neq KL(Q||P). \tag{2}$$

This computing asymmetry, without carefully treatment, could cause inconsistent in the quantitative evaluation for feature dimensions, and thus leads to inconveniency and ambiguity in feature selection. To overcome this difficulty, we use Mutual Information (MI) to substitute KL divergence

$$I(P,Q) = \sum_i \sum_j \phi(x_i, y_j) \log \frac{\phi(x_i, y_j)}{p(x_i) q(y_j)} \tag{3}$$

where x_i and y_i denote the steganalysis features extracted from images in cover and stego set respectively, $\phi(x_i, y_i)$ denote the joint probability distribution function, $p(x_i)$ and $q(y_i)$ denote the marginal probability distribution functions respectively. It is obvious that the definition of MI is symmetric

$$I(P,Q) = I(Q,P). \tag{4}$$

The MI can be represented as an expectation of KL divergence as below

$$I(X:Y) = E_Y\left(KL\left(p(x|y)||p(x)\right)\right) \tag{5}$$

where $p(x|y)$ denotes the conditional probability of image x drawn from the cover set, given the image y from the stego set, and $E_Y(\cdot)$ denotes the expectation for the random variable y. The relationship between the MI and the KL divergence in (5) suggests that MI maintains the characteristics of KL divergence in steganography benchmarking, provides a

fundamental quantity that estimates the difference between the distributions of the features obtained from the cover and stego set. In this way, the MI establishes a measurement of how much the features contribute to the final classifier. These properties and the computing symmetry shown in (4) make the MI an appropriate choice in evaluating the value of feature dimensions in steganalysis feature selection.

The calculation of the MI relies on the estimation of the distributions of P and Q, which is quite difficult or even impossible to achieve for high dimensional features in a practical point of view. Thus, we treat each dimension of the original features as a single feature and calculate MI separately. Histogram estimates are applied to each single feature to provide estimation of their distributions. The details can be found in Section 3.

2.2 Maximum Mean Discrepancy (MMD)

Given two distributions P and Q defined on the whole set of images X, the disparity of P and Q can be evaluated by a statistic called Maximum Mean Discrepancy (MMD) (Gretton et al., 2007). The main idea behind MMD is based on the statement that P and Q are the same distribution if and only if their probability distribution functions (pdf) p and q satisfy that

$$E_{x\sim p}\left(f(x)\right)=E_{x\sim q}\left(f(x)\right),\forall f\in C(X) \tag{6}$$

where $C(X)$ denotes the set of all continuous bounded functions on X, $E_{x\sim p}(\cdot)$ and $E_{x\sim q}(\cdot)$ denotes the expectation for the random variable x with p and q as the pdf respectively.

The number of functions in $C(X)$ is infinite, but only part of the functions in $C(X)$ can be utilized because of the finite number of samples in the training sets in practical steganalysis scenarios. Denote Γ a subset of $C(X)$, then the difference between distributions P and Q is evaluated by MMD values corresponding to Γ as

$$MMD\left[\Gamma,X_D,Y_D\right]=\sup_{f\in\Gamma}\left(\frac{1}{D}\sum_{i=1}^{D}f(x_i)-\frac{1}{D}\sum_{i=1}^{D}f(y_i)\right) \tag{7}$$

where $X_D=\{x_1,\ldots,x_D\}$ and $Y_D=\{y_1,\ldots,y_D\}$ are observations of the cover and stego distributions P and Q respectively.

The choice of Γ affects the performance of MMD significantly. It has to be rich enough to make p and q distinguishable, while still under the restriction of the finite number of images in cover and stego training sets. A typical construction of Γ is the Reproducing Kernel Hilbert Spaces (RKHS) built from the so-called *kernel* function. The kernel function is a symmetric, positive definite function used to generate the RKHS. Gaussian kernel has been proved to be a valuable choice (Pevný & Fridrich, 2008) as

$$k(x,y)=\exp\left(-\gamma\|x-y\|_2^2\right),\gamma>0\ . \tag{8}$$

In this case, the MMD values corresponding to Γ are obtained by an unbiased estimate based on U-statistics as

$$MMD_u[\Gamma, X_D, Y_D] = \left[\frac{1}{D(D-1)} \sum_{i \neq j} \left(k(x_i, x_j) + k(y_i, y_j) - k(x_i, y_j) - k(x_j, y_i) \right) \right]^{\frac{1}{2}}. \quad (9)$$

Fig. 1 shows the effectiveness of MMD values as steganography benchmarking. Training sets generated by various embedding methods (Fridrich et al., 2004; Westfeld, 2001; Mielikainen, 2006) and different embedding rates are applied to calculate MMD values. The false rates of the classifier corresponding to each training sets are also obtained, and normalized to be comparable to the MMD values. Note that the original MMD values are replaced by $-\log_{10}(MMD)$ for better visual. The result shows that MMD values are good estimations of the performance of classifiers in steganalysis.

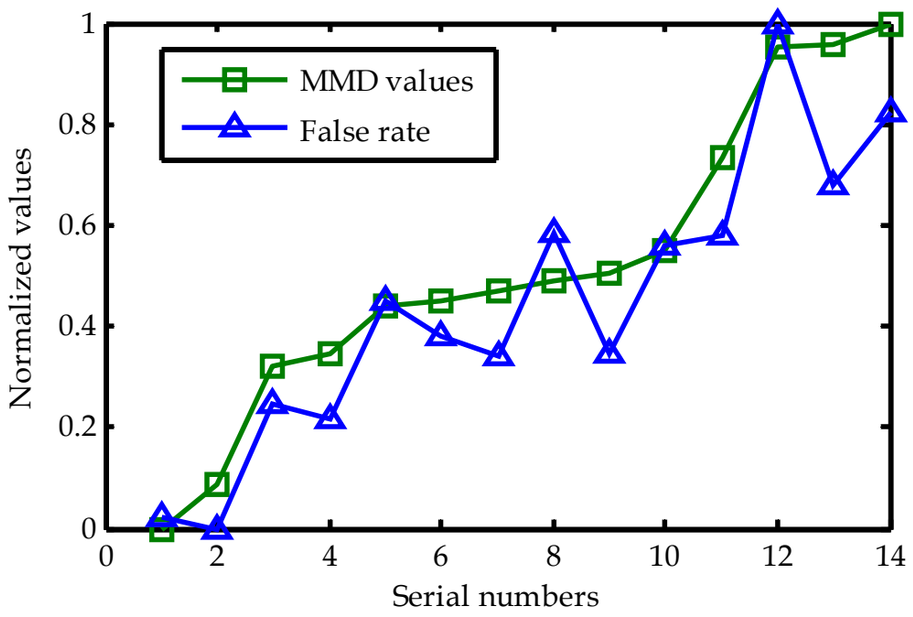


Fig. 1. Comparison between the MMD values (square marked) and the false rate of the classifiers (triangle marked)

The computational complexity of MMD with Gaussian kernel is $O(D^2)$, where D is the number of sample images. It is far more efficient in comparison to Support Vector Machines (SVM), which is a commonly used classifier in modern steganalysis. Further more, the MMD converges with error $1/\sqrt{D}$, yet almost independently on feature dimensions, which means that a sample set with roughly 10^3 images is sufficient for MMD to provide accurate estimations despite the feature dimensions. These advantages make MMD a natural choice to achieve feature selection for high-dimensional steganalysis features.

3. MMD-weighted-MI feature selection

The MI and MMD are both efficient measurements of evaluating the difficulty of distinguishing stego images from cover ones. Therefore, we apply them to steganalysis feature

selection methods. As they focus on different aspects of the feature distributions, we combine these two indicators into MMD-Weighted-MI for more comprehensive feature selection.

3.1 MI values for single feature dimension

As shown in Section 2.1, MI is a fundamental quantity for evaluating the value of feature dimensions in steganalysis feature selection. However, the calculation of MI relies on accurate estimation of high-dimensional distribution of the original features, which is difficult or even impossible to achieve from a practical point of view. To solve this problem, we calculate MI for each single dimension of the original features separately instead of treating them as high-dimensional features. Denote $x=\left(x^1,x^2,\cdots,x^d\right)$ the original feature extracted from the images in X, d is the total number of dimensions. Denote P^k and Q^k the marginal distribution of the k-th dimension in original features extracted from the cover and stego set respectively, the MI value for the k-th dimension of the original features is defined as

$$MI^k = I(P^k,Q^k) = \sum_i \sum_j \phi_k(x_i^k,y_j^k)\log\frac{\phi_k(x_i^k,y_j^k)}{p_k(x_i^k)q_k(y_j^k)}, k=1,2,\cdots d \tag{10}$$

where $\phi_k(x_i^k,y_j^k)$ denote the joint probability distribution function of the k-th dimension of the cover and stego set, $p_k(x_i^k)$ and $q_k(y_j^k)$ denote the marginal distributions respectively. Since $p_k(x_i^k)$ and $q_k(y_j^k)$ are both distributions of single random variables, it is simple to estimate their pdf by histogram estimation as

$$\tilde{p}_k(x)=\frac{n_j}{nh} \tag{11}$$

where n_j is the frequency fell into the j-th category, and h denotes the interval of categories. $\phi_k(x_i^k,y_j^k)$ is estimated by the joint histogram similarly. The choice of h affects the discrepancy and variance of the histogram estimation. Higher value of h leads to larger discrepancy and smaller variance, or vice versa. To achieve balance between discrepancy and variance, we set intervals dynamically corresponding to the dynamic range of each feature dimension in the proposed algorithm in Section 3.3.

Fig. 2 gives an example of MI calculated for each single dimension. The training set consists of cover images in jpg format and corresponding stego images generated by F5 steganography algorithm (Westfeld, 2001) with embedding rate at 0.05 bpac[1]. The Merging Markov and DCT features (Pevný & Fridrich, 2007) are chose as original steganalysis features. Fig. 2 shows that MI value for each single dimension varies significantly, and the feature dimensions with higher MI values contribute more than others in steganalysis classifier.

3.2 MMD-Weighted-MI (MWM)

The calculation of MI values of single feature dimensions treats each dimension as an independent feature. However, the correlation of different dimensions also plays an important

[1] bpac, bit per AC coefficients

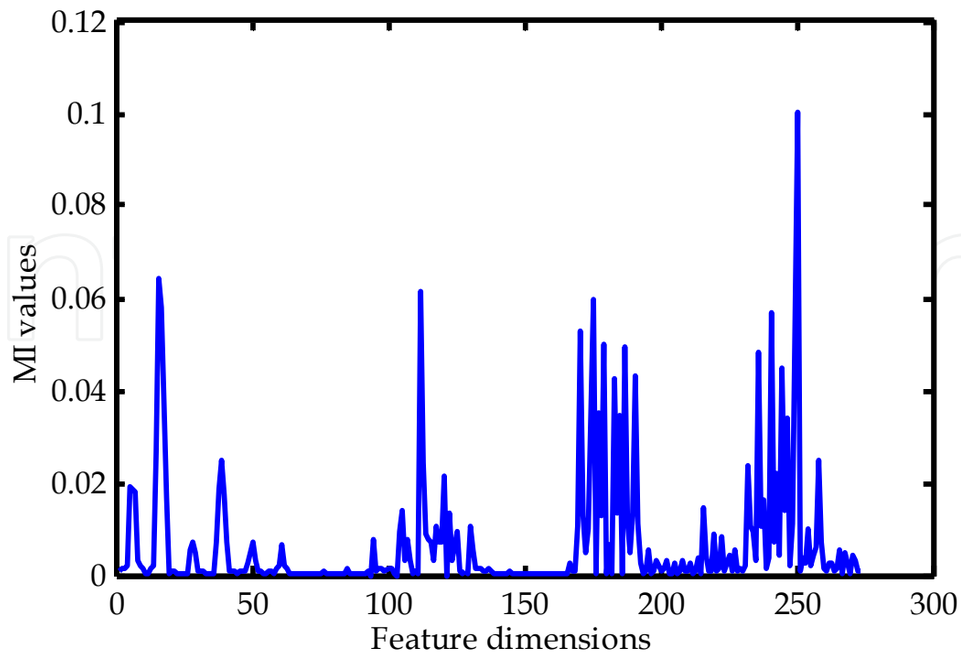


Fig. 2. MI values of Merging Markov and DCT features

role in training and classification. The absence of the information of feature correlation will bring troubles to the feature selection. Furthermore, the MI values of features drawn from different original feature sets have different dynamic range, which leads to an unfair comparison between feature sets if only MI values are used for feature selection. Fig. 3 shows the comparison of two feature sets as an example, where the MI values of most feature dimensions from the Partially Ordered Markov Model features sets (Davidson & Jalan, 2010) are relatively low than the Merging Markov and DCT features. The MI values of Y-Axis are limited to 0.04 for the purpose of clear observation though.

The raw MI values of single feature dimension are defective for feature selection due to the lack of correlation information. To overcome this difficulty, we introduce MMD as well for evaluating the feature dimensions. The MMD is numerically stable even in high-dimensional spaces, which makes it an excellent choice for providing information about correlation between feature dimensions. The computational complexity is also relatively low so that calculating MMD values for high-dimensional feature sets is feasible.

The combination of MI and MMD values provides a more comprehensive impression about the difference between the distribution of features extracted from the cover and stego training sets, which results in a new indicator called MMD-Weighted-MI (WMW) for better feature selection. Denote $MI(i,j)$ the MI value of the j-th dimension in the i-th feature set, and $MMD(i)$ the MMD value of the i-th feature set. Then a WMW value is assigned to each feature dimension as

$$WMW(i,j) = \frac{MI(i,j)}{MMD(i)} \tag{12}$$

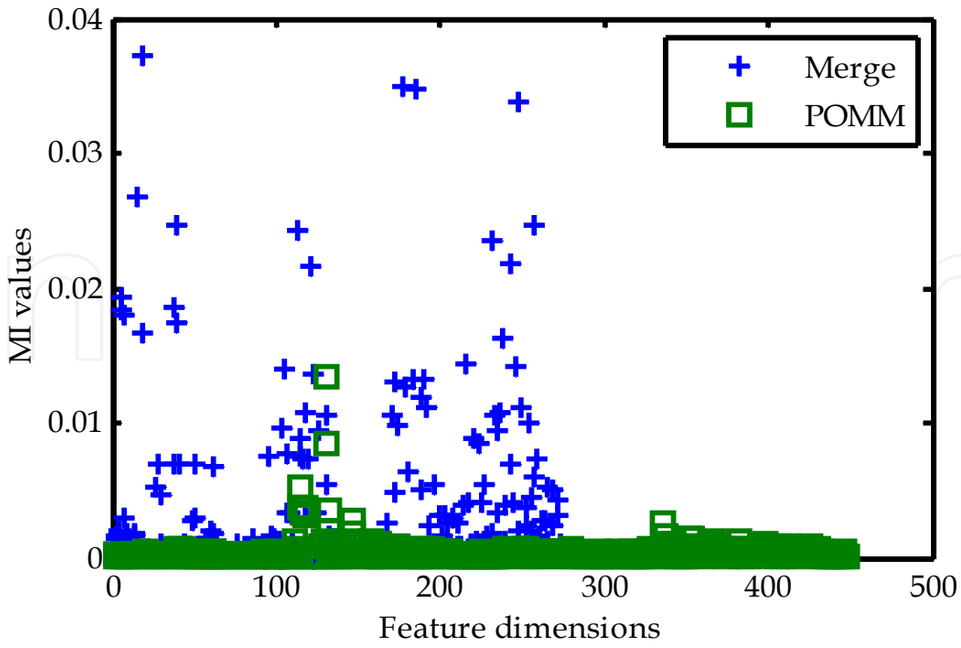


Fig. 3. Comparison of MI values between the Merging Markov and DCT features ('Merge', cross marked ones) and the Partially Ordered Markov Model features ('POMM', square marked ones)

Fig. 4 shows the MWM values derived from the two feature sets, namely the Merging Markov and DCT features and the Partially Ordered Markov Model features which is the

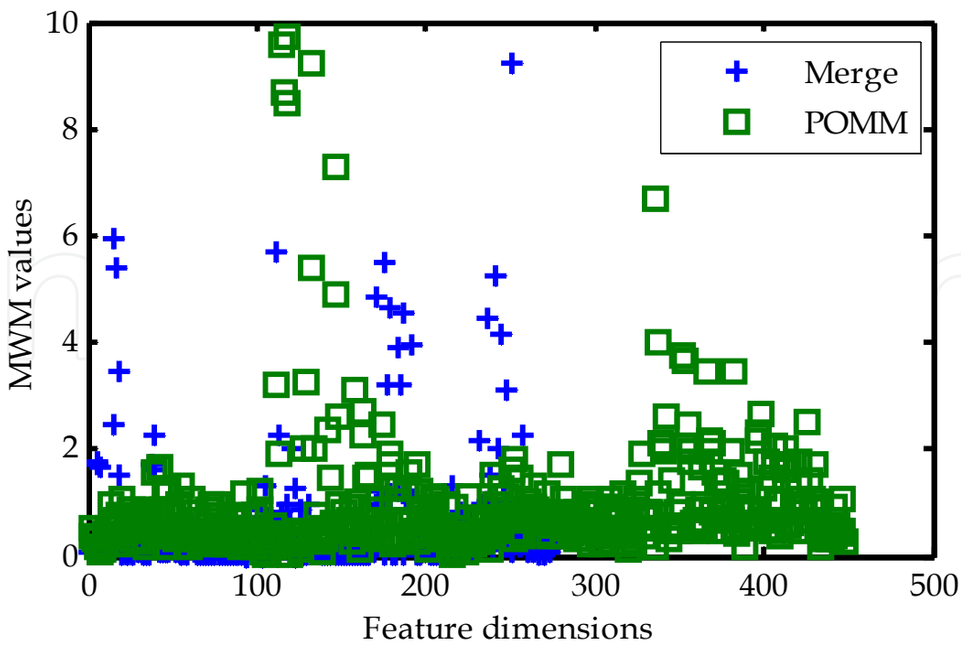


Fig. 4. Comparison of MI values between the Merging Markov and DCT features ('Merge', cross marked ones) and the Partially Ordered Markov Model ('POMM', square marked ones)

same as used in Fig. 3. Compared to the result of the raw MI values, MWM values achieve a fair comparison between different feature sets; make the dynamic range of the two feature sets comparable. This leads to a better feature selection for steganalysis and thus better accuracy of the classifier, which is supported by experiment results in Section 4.

3.3 Feature selection approach

The feature selection is implemented based on the MWM values and new steganalysis classifiers are constructed by the selected features. Fig. 5 shows the overview of our approach, and the details are presented as follows.

- Step 1. We choose several different steganalysis methods and extract the corresponding feature sets from the training set.
- Step 2. MI values are calculated and assigned to each feature dimension, and MMD values are calculated for each feature set as high-dimensional features. The MWM values are then generated based on MI and MMD values.
- Step 3. Feature dimensions with their MWM values larger than a given threshold are selected to assemble the new fused features. The serial numbers of the selected features are recorded as well. A classifier is trained with the fused features for steganalysis.

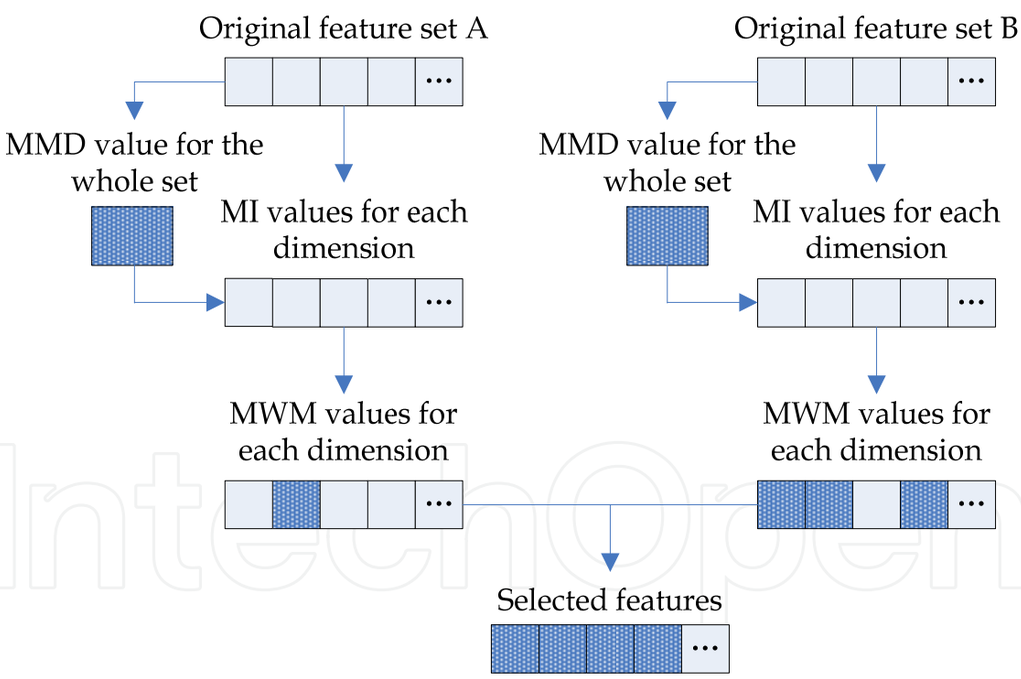


Fig. 5. Overview of the MWM feature selection

4. Experiment results

In this section, we experimentally investigate the performance of our WMW feature selection. We choose images of JPEG format as the cover images without loss of generality. The extensively used BOSSbase image database (Bas et al., 2010) is used as the source of cover images. Because the original images from BOSSbase are in the RAW format, we

converse them into JPEG format with a 98 JPEG quality factor. The stego images are generated using the following two typical steganography algorithms:

a. F5 steganography by Westfeld (Westfeld, 2001)
b. Perturbed Quantization (PQ) steganography by Fridrich et al. (Fridrich et al., 2004)

The embedding rates vary from 0.05 bpac to 0.15 bpac. We randomly choose 1000 pair of cover/stego images as the training sets and 300 other pairs as the testing sets.

Three typical steganalysis feature sets are chose to provide original features for WMW feature selection:

a. Markov Transition Probability features by Shi et al. (Shi et al., 2007), with 900 dimensions.
b. Merging Markov and DCT features by Pevný & Fridrich. (Pevný & Fridrich, 2007), with 274 dimensions.
c. Partially Ordered Markov Model features by Davidson & Jalan (Davidson & Jalan, 2010), with 448 dimensions.

The total number of dimensions involved in our experiments is 900+274+448=1622. We gradually increase the number of the chosen dimensions by WMW feature selection and test the performance of the corresponding classifiers by TR [2]. The intervals of the feature numbers are set differently because of the uneven density of the distribution of the MWM values. For the purpose of a clear view, we set interval to 5 dimensions within the first 150 features and 100 dimensions for the rest of them. The TR of the classifiers are tested and shown in Fig. 6, 7 and 8 for different embedding rate (0.05bpac, 0.1bpac and 0.15bpac) respectively.

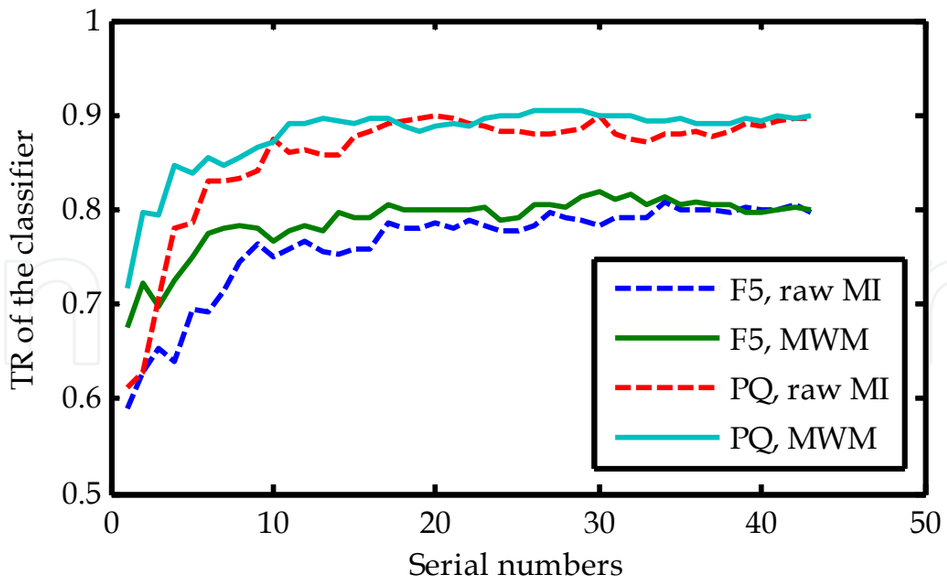


Fig. 6. Comparison of the performance between the classifier using WMW values (solid lines) and the raw MI values (dotted lines) for feature selection, with embedding rate 0.05bpac.

[2] TR, True Rate, the average of the True Positive rate (TP) and True Negative rate (TN)

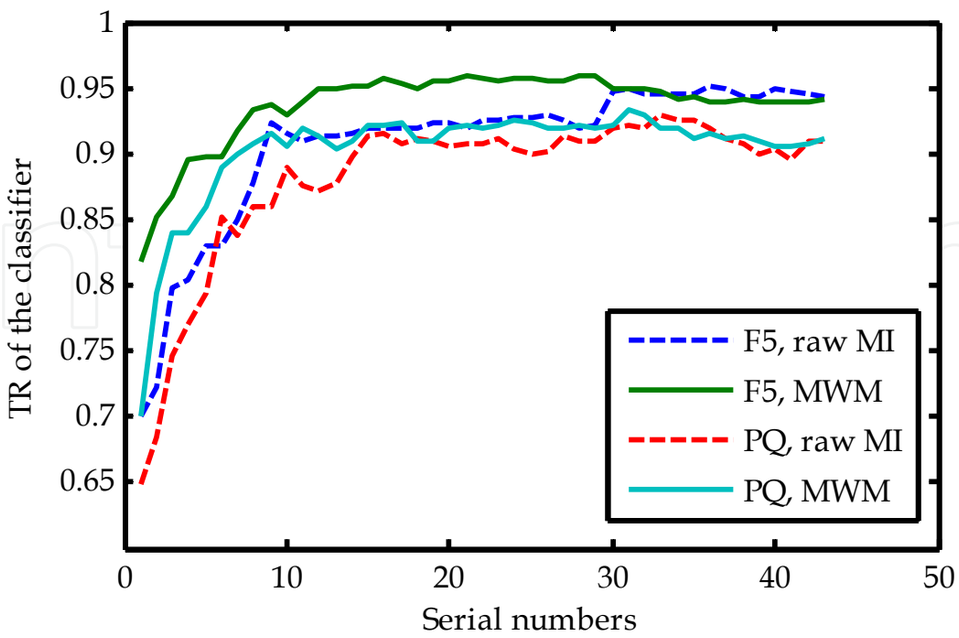


Fig. 7. Comparison of the performance between the classifier using WMW values (solid lines) and the raw MI values (dotted lines) for feature selection, with embedding rate 0.1bpac.

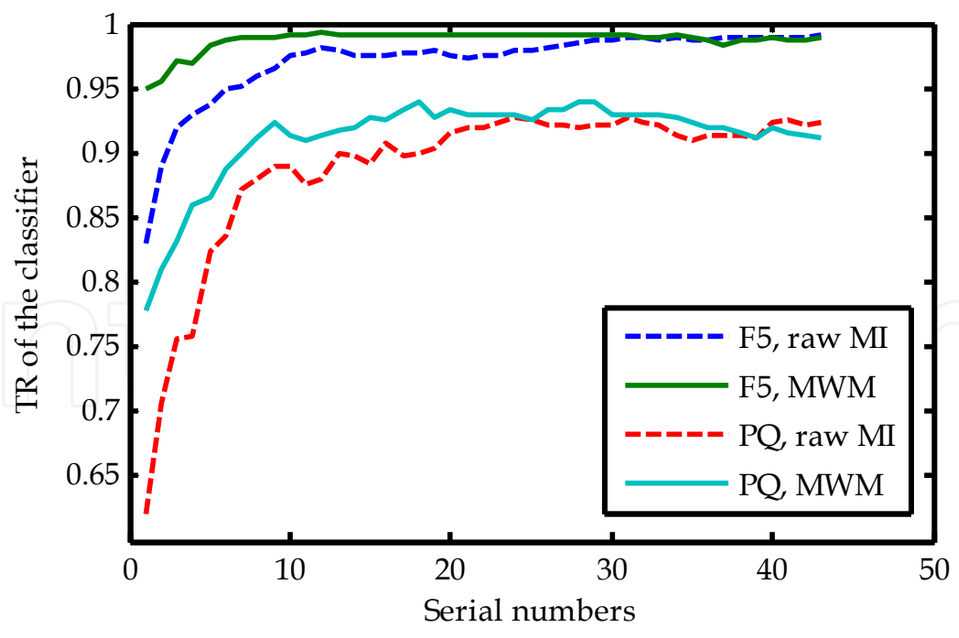


Fig. 8. Comparison of the performance between the classifier using WMW values (solid lines) and the raw MI values (dotted lines) for feature selection, with embedding rate 0.15bpac.

From Fig. 6, 7 and 8, we can observe that WMW feature selection provides higher TR of the classifier than feature selections using only the raw MI values. The reason of this has been discussed in Section 3.2. Note that the last TR value of each curve represents the performance of the classifier consist of all features without selection approach. It is then obvious that we can always achieve better accuracy of steganalysis using WMW feature selections than the original feature sets, whereas feature selection with raw MI values fail in some cases, e.g. F5 steganography with embedding rate 0.1 bpac in Fig. 7, and Perturbed Quantization steganography with embedding rate 0.15 bpac in Fig. 8. Table 1 shows the optimal accuracy of each case, and the TR of classifiers consist of the original feature set are also listed for comparison.

Embedding Cases	MPB	Merge	POMM	Total	Raw	MWM
F5, 0.05bpac	60.5%	80.7%	67.6%	79.7%	80.4%	**81.9**%
F5, 0.10bpac	73.1%	93.9%	85.5%	94.4%	95.2%	**96.0**%
F5, 0.15bpac	84.2%	98.3%	94.0%	98.9%	**99.2**%	**99.2**%
PQ, 0.05bpac	68.8%	88.4%	73.3%	89.9%	89.9%	**90.5**%
PQ, 0.10bpac	72.6%	89.2%	76.2%	91.2%	92.9%	**93.4**%
PQ, 0.15bpac	71.4%	90.9%	78.0%	92.4%	92.7%	**93.9**%

Table 1. Optimal accuracy of steganalysis for different embedding cases using different feature sets: Markov Transition Probability features ('MPB'), Merging Markov and DCT features ('Merge'), Partially Ordered Markov Model features ('POMM'), fused features contain all feature dimensions without feature selection ('Total'), feature selection using only raw MI values ('Raw'), and MWM feature selection ('MWM').

The best accuracy for each embedding case is marked in bold in Table 1, and from that we can assert that the MWM feature selection is always the better choice for steganalysis comparing to other methods.

5. Conclusion

In this chapter, we present a new approach of feature selection in steganalysis involving MI and MMD, which are both efficient indicators for evaluating the difference between cover and stego sets. Although the MI values are well understood theoretically, the computational difficulty of estimating the distribution of high-dimensional features makes it inconvenient in steganalysis feature selection. Thus, we treat each dimension as a single feature and calculate MI values separately.

This approach, however, abandons the correlation between feature dimensions, which makes raw MI values defective for feature selection. To solve this problem, MMD values are introduced in our approach as well. The MMD values are numerically stable even in high-dimensional spaces, and the computational complexity is relatively low. These advantages

make MMD a natural complementary to MI values, and thus leads to our proposed approach of feature selection based on MMD-weighted-MI values.

To test the performance of the MWM feature selection, we apply our method to three typical steganalysis feature sets to generate new classifiers, and estimate the accuracy of these fused classifiers against two widely used steganography algorithms. Experimental results shows that the MWM feature selection approach outperforms the feature selections with raw MI values, and guarantees better accuracy comparing to the original feature sets.

6. Acknowledgment

This work was supported by the NSF of China under 61170281, the NSF of Beijing under 4112063, the Strategic and Pilot Project of Chinese Academy of Sciences (CAS) under XDA06030600, and the Project of Institute of Information Engineering, CAS, under Y1Z0051101 and Y1Z0041101.

7. References

Bas, P.; Filler, T. & Pevný, T. (2010). BOSS. http://boss.gipsa-lab.grenobleinp.fr/BOSSRank/, July 2010

Cachin, C. (1998). An information-theoretic Model for Steganography. *Information Hiding 1998, Lecture Notes in Computer Science*, Vol. 1525, (1998), pp. 306–318

Cover, T. M. & Thomas, J. A. (2001). *Frontmatter and Index, in Elements of Information Theory*. John Wiley & Sons, Inc., ISBN 9780471062592, New York

Davidson, J. & Jalan, J. (2010). Steganalysis Using Partially Ordered Markov Models. Proc. *Information Hiding 2010, Lecture Notes in Computer Science*, Vol. 6387, (2010), pp. 118-132

Dong, J.; Chen, X.; Guo, L. & Tan, T. (2008). Fusion Based Blind Image Steganalysis by Boosting Feature Selection. *Digital Watermarking, Lecture Notes in Computer Science*, Vol. 5041, (2008), pp. 87-98

Fridrich, J. (2005). Feature-Based Steganalysis for JPEG Images and Its Implications for Future Design of Steganographic Schemes. *Information Hiding, Lecture Notes in Computer Science*, Vol. 3200, (2005), pp. 67-81

Fridrich, J. & Goljan, M. (2002). Practical Steganalysis of Digital Images: State of the Art. In: *Security and Watermarking of Multimedia Contents*, Vol. SPIE-4675, pp. 1–13, 2002

Fridrich, J.; Goljan, M. & Hogea, D. (2002). Steganalysis of JPEG Images: Breaking the F5 Algorithm. *5th International Workshop on Information Hiding*, Vol. 2578, pp. 310-323, Oct 07-09, 2002

Fridrich, J.; Goljan, M. & Soukal, D. (2004). Perturbed Quantization Steganography with Wet Paper Codes. *Proc. ACM Multimedia Workshop'04,* pp. 4-15, 2004

Fridrich, J.; Kodovský, J.; Holub, V. & Goljan, M. (2011) Steganalysis of Content-adaptive Steganography in Spatial Domain. *Information Hiding, Lecture Notes in Computer Science*, 2011

Gul, G. & Kurugollu, F. (2011). A New Methodology in Steganalysis: Breaking Highly Undetectable Steganography (HUGO). *Information Hiding, Lecture Notes in Computer Science*, Vol. 6958, pp. 71-84, 2011

Gretton, A.; Borgwardt, K.; Rasch, M.; Scholkopf, B. & Smola, A. (2007). A Kernel Method for the Two-sample-problem. *Advances in Neural Information Processing Systems*, Vol. 19, (2007), pp. 513–520, MIT Press, Cambridge

Harmsen, J. J. & Pearlman, W.A. (2003). Steganalysis of Additive Noise Modelable Information Hiding. In: *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents VI*, (2003), pp. 131–142

Johnson, N. F. & Jajodia, S. (1998). Exploring Steganography: Seeing the Unseen. In: *IEEE Computer Society*, Vol. 31, pp. 26–34, 1998

Kahn, D. (1996). The History of Steganography. *Information Hiding, Lecture Notes in Computer Science*, Vol. 1174, (1996), pp. 1-5

Ker, A. D. (2005). Steganalysis of LSB Matching in Grayscale Images. Signal Processing Letters, IEEE, Vol. 12, No. 6, (June 2005), pp. 441-444, ISSN: 1070-9908

Lyu S. & Farid, H. (2002). Detecting Hidden Messages Using Higher-order Statistics and Support Vector Machines. In: *5th International Workshop on Information Hiding,* pp. 340-354, 2002

Miche, Y.; Roue, B.; Lendasse, A. & Bas, B. (2006). A Feature Selection Methodology for Steganalysis. *Multimedia Content Representation, Classification and Security Lecture Notes in Computer Science*, Vol. 4105, (2006), pp. 49-56

Mielikainen, J. (2006). LSB matching revisited. *Signal Processing Letters, IEEE*, Vol. 13, No. 5, (May 2006), pp. 285-287, ISSN: 1070-9908.

Pevný, T.; Bas, P. & Fridrich, J. (2010). Steganalysis by Subtractive Pixel Adjacency Matrix. *IEEE Transaction on Information Forensics and Security*, Vol. 5, No. 2, (June 2010), pp. 215-224, ISSN: 1556-6013

Pevný, T.; Filler, T. & Bas, P. (2010). Using High-dimensional Image Models to Perform Highly Undetectable Steganography. *Information Hiding, 12th International Workshop, Lecture Notes in Computer Science*, Calgary, Canada, June 28–30, 2010

Pevný, T. & Fridrich, J. (2007). Merging Markov and DCT features for Multi-class JPEG steganalysis. Proc. *SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, Vol. 6505, pp. 3-1 – 3-14, 2007

Pevný, T. & Fridrich, J. (2008). Benchmarking for Steganography. *Information Hiding, Lecture Notes in Computer Science*, Vol. 5284, (2008), pp. 251-267

Provos, N. (2001) Defending Against Statistical Steganalysis. In: *Proceedings of the 10th USENIX Security Symposium*, pp. 323–336, 2001

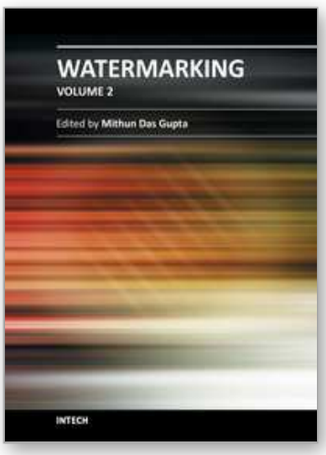
Provos, N. & Honeyman, P. (2003). Hide and seek: an introduction to steganography. *Security & Privacy, IEEE*, Vol. 1(3), (May-June 2003), pp. 32-44, ISSN 1540-7993

Shi, Y. Q.; Chen, C. H. & Chen, W. (2007). A Markov Process Based Approach to Effective Attacking JPEG Steganography. Proc. *Information Hiding '07*, vol.4437, (2007), pp. 249-264

Tzschoppe, R. & Aauml, R.B. (2003). Steganographic System based on Higher-order Statistics. In: *Proceedings of SPIE, Security and Watermarking of Multimedia Contents V*, (2003), Vol. 5020

Westfeld A. (2001). F5 - A Steganographic Algorithm High Capacity Despite Better Steganalysis. *4th International Workshop on Information Hiding*, Vol. 2137, (April 25-27, 2001), pp. 289-302

Xuan, G. R.; Shi, Y. Q. & Gao, J. J. (2005). Steganalysis based on Multiple Features formed by Statistical Moments of Wavelet Characteristic Functions. *Proc. Information Hiding 2005*, Vol. 3727, p.p. 262-277, 2005



Watermarking - Volume 2

Edited by Dr. Mithun Das Gupta

ISBN 978-953-51-0619-7
Hard cover, 276 pages
Publisher InTech
Published online 16, May, 2012
Published in print edition May, 2012

This collection of books brings some of the latest developments in the field of watermarking. Researchers from varied background and expertise propose a remarkable collection of chapters to render this work an important piece of scientific research. The chapters deal with a gamut of fields where watermarking can be used to encode copyright information. The work also presents a wide array of algorithms ranging from intelligent bit replacement to more traditional methods like ICA. The current work is split into two books. Book one is more traditional in its approach dealing mostly with image watermarking applications. Book two deals with audio watermarking and describes an array of chapters on performance analysis of algorithms.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

B. B. Xia, X. F. Zhao and D. G. Feng (2012). Improve Steganalysis by MWM Feature Selection, Watermarking - Volume 2, Dr. Mithun Das Gupta (Ed.), ISBN: 978-953-51-0619-7, InTech, Available from: http://www.intechopen.com/books/watermarking-volume-2/improve-steganalysis-by-mwm-feature-selection

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen