ON INFORMATION HIDING WITH INCOMPLETE INFORMATION ABOUT STEGANALYSIS

R. Chandramouli

Department of ECE
Stevens Institute of Technology

ABSTRACT

Suppose the information hider has a set of information hiding schemes to be employed in the presence of a set of steganalysis detectors. Let the information hider have only an incomplete information about the steganalysis detectors. That is, suppose the probability of the jth steganalysis detector will be used is p_j, $j = 1, 2 \ldots n$, then the information hider only knows the ordering of these probabilities, say, $p_1 \geq p_2 \geq \ldots \geq p_n$. Under this circumstance how can the information hider choose the *optimal* embedding strategy? This is the question addressed in this paper. Theoretical analysis and some numerical results are presented.

1. INTRODUCTION

While image steganography deals with techniques for hiding information in digital images the goal of steganalysis is to detect and/or estimate potentially hidden information given an image. We can broadly classify image steganalysis into two categories:

- **Passive steganalysis**: Detect the presence or absence of a secret message in an observed image or identify the type of embedding algorithm.
- **Active steganalysis**: Estimate/extract some properties of the message or the embedding algorithm. For example, extract a (possibly approximate) version of the secret message from a stego message. This falls under the category of image forensics.

Several image steganalysis approaches have been proposed for different embedding algorithms. Some of them exploit the image bit plane statistics [1], blind source separation based detectors [2], etc.

Most of the information hiding algorithms have been optimized for maximizing capacity without much consideration for resilience against steganalysis attacks. As a result many of these embedding algorithms can be reliably detected by steganalysis detectors designed for that particular algorithm. Therefore it becomes important to consider potential steganalysis detectors when designing embedding algorithms as discussed in [3] within the context of adaptive steganography. In many practical situations it is possible that one among a number of different steganalysis detectors (either different detection algorithms or one detection algorithm with different parameter settings) could be employed. Depending on the embedding algorithm used or its parameter settings the reliability of the detector output will vary.

From the information hider's perspective three possible scenarios arise:

- **Complete information case:** The information hider exactly knows the probabilities $\{p_j\}_{j=1}^{n}$ where p_j is the probability of the jth steganalysis detector being employed.
- **Incomplete information case:** The information hider does not know the $\{p_j\}$ values but only the ordering, say, $p_1 \geq p_2 \geq \ldots \geq p_n$ (without loss of generality).
- **No information case:** In this case, no knowledge of $\{p_j\}$ is available.

It is easy to see that the complete information and no information cases are at two extremes. It may be possible to estimate the probability set $\{p_j\}$ fairly accurately by using adaptive steganographic techniques [3] or choosing the probabilities subjectively based on some a prior information. We note that accurate estimation of the probabilities may need a lot of effort in several situations. An alternative to knowing these probabilities is to assume that no knowledge of $\{p_j\}$ is available. This will lead the information hider to strategies that optimize for the worst case risk and seek extreme payoffs. The incomplete information case is somewhat mid-way between these two extremes.

Let us consider the following situation in the prisoner's problem formulation of steganography [4]. Suppose Alice wants to send a secret message to Bob and Wendy the warden chooses a steganalysis detector from a set $\mathcal{S}$ unknown to Alice and Bob. Further, it is quite possible that some of

This work is partially supported by NSF DAS 0242417 and AFRL F30602-02-2-0193.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

the detectors in $\mathcal{S}$ are not as accurate as the others. Therefore, Wendy may choose to employ the most reliable detector with the highest probability (say, p_1) and so on. Now, Alice can try out the different embedding strategies available to her a small number of times and easily estimate only the ordering of the probabilities (e.g., $p_1 \geq p_2 \geq \ldots \geq p_n$ and $\sum_j p_j = 1$) but not the values themselves. Intuitively, the effort needed by Alice to estimate the ordering will be lower than that for accurately estimating the probability values.

The main problem addressed in this paper is the following. How must the information hider choose its strategy in the case of incomplete information about the steganalysis detector set? This problem is formulated mathematically and decision theoretic ideas are developed to solve it. A definition of *best strategy* for the information hider is also provided among other details. The ideas presented in this paper are organized as follows. Section 2 presents the mathematical formulation of the problem with detailed analysis. Numerical results are presented in Section 3. Concluding remarks are presented in Section 4.

2. MATHEMATICAL FORMULATION

Let $\mathcal{H}$ with cardinality m denote the set of data hiding algorithms available to the information hider. $\mathcal{S}$ is the set of steganalysis detectors and p_j is the probability with which the detector $j \in \mathcal{S}$ is employed. The information available to the information hider is that $p_1 \geq p_2 \geq \ldots \geq p_n$ where the cardinality of $\mathcal{S}$ is equal to n. The information hider gets a payoff equal to X_{hj} when embedding algorithm $h \in \mathcal{H}$ is used and detector $j \in \mathcal{S}$ is active. For instance the payoff matrix can be constructed as follows:

$$X_{hj} = \begin{cases} N_{hj}, & \text{if steganalysis detection fails} \\ 0, & \text{if steganalysis succeeds} \end{cases} \quad (1)$$

where N_{hj} is the embedded message size. We note that N_{hj} is a function of the steganalysis detection algorithm, detection error probabilities, embedded message length, image statistics, etc. Note that, in general, N_{hj} will be a small value if the jth steganalysis algorithm is accurate. Clearly, there are also other possible ways of defining X_{hj}.

The question faced by the information hider is the following: *how to choose the embedding algorithm such that the expected payoff is maximized?* The optimal strategy for the information hider h^* such that the expected payoff is maximized is given by,

$$h^* = \arg\max_h \sum_{j=1}^{n} p_j X_{hj}. \quad (2)$$

Note that if the payoff function described in Eq. (1) is used then the resultant maximum expected payoff is the expected value of the embedding capacity in the presence of steganalysis as discussed in [5].

2.1. Maximizing Expected Payoff—Loose Inequality Constraint

In this section we discuss a solution to the optimization problem in Eq. (2) when $p_1 \geq p_2 \geq \ldots \geq p_n$ (loose inequality constraint). We consider the case $m = 2$ for simplicity, i.e., let $\mathcal{H} = \{h_1, h_2\}$. Then the payoff matrix for the information hider is given by,

$$\begin{array}{l} \text{Strategy } h_1 \\ \text{Strategy } h_2 \end{array} \begin{bmatrix} X_{h_1 1} & X_{h_1 2} & \ldots & X_{h_1 n} \\ X_{h_2 1} & X_{h_2 2} & \ldots & X_{h_2 n} \end{bmatrix} \quad (3)$$

The results presented here can be extended to $m > 2$ case with little additional effort. We first present a useful definition.

Definition 1 *Embedding strategy $h_k \in \mathcal{H}$ is said to stochastically dominate $h_l \in \mathcal{H}$ if $E(h_k) \geq E(h_l)$ where $E(h_i) = \sum_{j=1}^{n} p_j X_{h_i j}$, $i = 1, 2$ is the expected value of the payoff to the information hider in using strategy h_i.*

From this definition we observe that the information hider looks to compute the stochastically dominating embedding strategy if it exists. We note that this type of optimal statistical decision problem was first introduced in [6] and the following theorem is based on the analysis presented there.

Theorem 1 *If $p_1 \geq p_2 \geq \ldots \geq p_n$ then h_1 stochastically dominates h_2 if $\sum_{j=1}^{k} X_{h_1 j} \geq \sum_{j=1}^{k} X_{h_2 j}$, $\forall\ k = 1, 2, \ldots n$.*

Proof. We see that,

$$E(h_1) - E(h_2) = \sum_{j=1}^{n} p_j (X_{h_1 j} - X_{h_2 j}). \quad (4)$$

Now we resort to the Abel's summation identity, namely,

$$\sum_{j=1}^{n} a_j b_j = \sum_{k=1}^{n-1} \left(\sum_{j=1}^{k} a_j \right) (b_k - b_{k+1}) + b_n \sum_{j=1}^{n} a_j.$$

Therefore applying the Abel's identity to Eq. (4) we get,

$$E(h_1) - E(h_2) = \sum_{j=1}^{n} p_j (X_{h_1 j} - X_{h_2 j}) \quad (5)$$

$$= \sum_{k=1}^{n-1} (\sum_{j=1}^{k} (X_{h_1 j} - X_{j_2 j}))(p_k - p_{k+1}) + p_n \sum_{k=1}^{n} (X_{h_1 j} - X_{h_2 j}) \quad (6)$$

Since $p_k - p_{k+1} \geq 0,\ \forall k = 1, 2 \ldots n$ and $p_{n+1} = 0$ the result follows from Eq.(6).

Therefore using Theorem 1 the information hider can choose the optimum embedding strategy (if it exists) given the payoff matrix. The next relevant question for the information hider is: what are the achievable maximum and minimum expected payoffs? Let $h \in \mathcal{H}$ be a generic embedding strategy. Then the following maximization/minimization problem is to be solved:

$$\begin{aligned} \text{optimize} \quad & \left[E(h) = \sum_{j=1}^{n} p_j X_{hj} \right], \\ \text{subject to} \quad & \sum_{j=1}^{n} p_j = 1, \\ & p_j \geq p_{j+1},\ j = 1, 2 \ldots n-1, \\ & p_j \geq 0,\ j = 1, 2 \ldots n. \end{aligned} \tag{7}$$

Theorem 2 *The maximum and minimum solutions to the optimization problem in Eq. (7) are given respectively by* $\max_{k=1,2,\ldots n}\{\frac{1}{k}\sum_{j=1}^{k} X_{hj}\}$ *and* $\min_{k=1,2,\ldots n}\{\frac{1}{k}\sum_{j=1}^{k} X_{hj}\}$.

Proof. First we define the following variables,

$$\begin{aligned} Z_{hk} &= \sum_{j=1}^{k} X_{hj} \\ q_j &= p_j - p_{j+1},\ j = 1, 2, \ldots (n-1) \\ q_n &= p_n. \end{aligned} \tag{8}$$

Then, after some algebraic manipulations it is easy to see that the optimization problem posed in Eq.(7) can be rewritten as follows:

$$\begin{aligned} \text{optimize} \quad & \left[E(h) = \sum_{j=1}^{n} q_j Z_{hj} \right], \\ \text{subject to} \quad & \sum_{j=1}^{n} j q_j = 1, \\ & q_j \geq 0,\ j = 1, 2, \ldots n. \end{aligned} \tag{9}$$

This is a linear programming problem in the standard form. Therefore, if a finite optimal solution exists to this problem then it will be a corner point. This means only one of the q_j's will be non-zero. Then, in order to satisfy the constraint $\sum_{j=1}^{n} j q_j = 1$ this non-zero q_j must be equal to $1/j$ which in turn implies that $E(h)$ is maximized (minimized) when Z_{hj}/j is maximized (minimized). Hence the result.

2.2. Maximizing Expected Payoff–Strict Inequality Constraint

Now consider the problem in Eq. (7) with the probability inequality is strict, i.e., $p_1 > p_2 > \ldots > p_n$. This case arises when the information hider has additional knowledge about the steganalysis strategy and therefore is able to tighten the inequality on the probabilities. The optimization (maximization or minimization) problem faced by the information hider is then given as follows:

$$\begin{aligned} \text{optimize} \quad & \left[E(h) = \sum_{j=1}^{n} p_j X_{hj} \right], \\ \text{subject to} \quad & \sum_{j=1}^{n} p_j = 1, \\ & (p_j - p_{j+1}) \geq c_j > 0,\ j = 1, 2 \ldots n-1, \\ & p_j \geq 0,\ j = 1, 2 \ldots n. \end{aligned} \tag{10}$$

Here c_j are some known constants.

Theorem 3 *The maximum and minimum solutions to the optimization problem in Eq. (10) are given respectively by* $\max_{k=1,2,\ldots n}\{\frac{1}{k}\sum_{j=1}^{k} X_{hj}(1-\sum_{j=1}^{n} j c_j)+\sum_{k=1}^{n} c_k Z_{hk}\}$ *and* $\min_{k=1,2,\ldots n}\{\frac{1}{k}\sum_{j=1}^{k} X_{hj}(1-\sum_{j=1}^{n} j c_j)+\sum_{k=1}^{n} c_k Z_{hk}\}$.

Proof. First we define the following variable,

$$\tilde{q}_j = p_j - p_{j+1} - c_j,\ j = 1, 2, \ldots n. \tag{11}$$

Then, after some algebraic manipulations similar to the case of loose inequality constraints we see that the maximization or minimization problem can be reformulated as:

$$\begin{aligned} \text{optimize} \quad & \left[E(h) = \sum_{j=1}^{n} \tilde{q}_j Z_{hj} + \sum_{j=1}^{n} c_j Z_{hj} \right], \\ \text{subject to} \quad & \sum_{j=1}^{n} j\tilde{q}_j + \sum_{j=1}^{n} j c_j = 1, \\ & \tilde{q}_j \geq 0,\ j = 1, 2, \ldots n. \end{aligned} \tag{12}$$

Following a similar argument as given previously, if a finite optimal solution exists to this problem then it will be a corner point. Hence the result.

3. NUMERICAL RESULTS

Consider two image information hiding algorithms h_1 and h_2 and four steganalysis detectors such that $p_1 \geq p_2 \geq p_2 \geq p_4$. Let the corresponding payoff matrix be given by,

$$\begin{bmatrix} 100 & 220 & 75 & 500 \\ 230 & 125 & 170 & 300 \end{bmatrix} \tag{13}$$

Let the payoff values correspond to the maximum message size that can be embedded without being detected by the corresponding steganalysis detector for a certain false alarm

and miss probability. These payoff numbers could be obtained for practical image steganalysis algorithms (e.g., refer to [7]). Then by applying Theorem 1 we observe that neither strategy stochastically dominates the other. The partial average payoffs for h_1 are given by,

$$\begin{aligned} X_{h_1 1} &= 100 \\ \frac{1}{2}(X_{h_1 1} + X_{h_1 2}) &= 160 \\ \frac{1}{3}(X_{h_1 1} + X_{h_1 2} + X_{h_1 3}) &= 131.6 \\ \frac{1}{4}(X_{h_1 1} + X_{h_1 2} + X_{h_1 3} + X_{h_1 4}) &= 223.75 \end{aligned} \quad (14)$$

and for h_2 they are,

$$\begin{aligned} X_{h_2 1} &= 230 \\ \frac{1}{2}(X_{h_2 1} + X_{h_2 2}) &= 177.5 \\ \frac{1}{3}(X_{h_2 1} + X_{h_2 2} + X_{h_2 3}) &= 175 \\ \frac{1}{4}(X_{h_2 1} + X_{h_2 2} + X_{h_2 3} + X_{h_2 4}) &= 206.25. \end{aligned} \quad (15)$$

Therefore by employing h_1 the information hider will achieve a maximum expected payoff equal to 223.75 and a minimum equal to 100. Similarly for h_2 the maximum is 230 and minimum is equal to 175. We note in this example that if the maximum expected payoff criterion is used to select the information hiding strategy then the different between h_1 and h_2 is not not very significant. However, if maximin criterion is used then h_2 seems to be significantly better than h_1.

4. CONCLUSIONS

A mathematical model that captures the relationship between information hiding and steganalysis detection is presented. The information hider operates under incomplete knowledge about the steganalysis detector. Mathematical conditions under which one hiding strategy is better has a simple structure. The maximum and minimum payoffs for the information hider are also derived. A numerical result is presented to illustrate the ideas.

5. REFERENCES

[1] J. Fridrich, M. Goljan, and R. Du, "Reliable detection of lsb steganography in grayscale and color images," *Proc. of the ACM Workshop on Multimedia and Security*, pp. 27–30, Oct. 2001.

[2] R. Chandramouli, "A mathematical framework for active steganalysis," *ACM/Springer Multimedia Systems Special Issue on Multimedia Watermarking*, vol. 9, pp. 303–311, Sept. 2003.

[3] R.Chandramouli and N.D. Memon, "Adaptive steganography," *Proc. SPIE Security and Watermarking of Multimedia Content*, vol. 4675, Jan. 2002.

[4] G.J. Simmons, "The prisoner's problem and the subliminal channel," *Proceedings of CRYPTO*, 1983.

[5] R. Chandramouli and N.D. Memon, "Steganography capacity: A steganalysis perspective," *Proc. SPIE Security and Watermarking of Multimedia Contents*, vol. 5020, pp. 173–177, Jan. 2003.

[6] P.C. Fishburn, *Decision and value theory*, John Wiley, 1963.

[7] S. Trivedi and R. Chandramouli, "Secret key estimation in sequential steganography," *Supplement on Secure Media, IEEE Trans. on Signal Processing*, Feb. 2005.