2010 International Conference on Multimedia Information Networking and Security

Cryptographic Secrecy Analysis of Matrix Embedding

Jiufen Liu, Jiayong Chen, Weiming Zhang, Tao Han
Department of Information Research,
Zhengzhou Information Science and Technology Institute, Zhengzhou, China
cjy1003@sina.com

***Abstract*—Matrix embedding has been used to improve the embedding efficiency of steganography, which is an efficient method to enhance the concealment security. The privacy security of matrix embedding has also been studied under the condition of known-stego-object attcack. However, with the development of steganalysis, the attacker could obtain the estimated cover by the cover restoration technique. Consequently, the privacy security under the stronger attack condition should be considered. In this paper we study the secrecy security of matrix embedding using information theory under the circumstance of known-cover attack from the point of the key equivocation. The relation among the wet ratio of covers, embedding rate, and key equivocation is presented. We also proposed a new differential attack to matrix embedding under the circumstance of chosen-stego attack.**

Keywords-component; Steganographic coding; wet paper codes; key equivocation; known-cover attack; unicity distance

I. INTRODUCTION

Information hiding is one of the hot spots in the domain of information security, of which steganography and watermarking are the main branches. The goal of steganography is embedding secret message into the multimedia data, such as digital images, audios, videos and texts, in order to realize the covert transmission.

The early studies about the security of steganography mainly focused on the statistical undetectability of the cover, i.e. the concealment security. However, the attacker may not only reveal the existence of the secret message but also extract the message by recovering the stego key. For instance, Fridrich et al. [1][2] analyzed problem of searching for stego key from the view of computational complexity. Based on the model of information theory, Zhang et al. [3][4] analyzed the relation among both kinds of securities and the relation between security and hidden capacity.

Matrix embedding [5][6][7][8][9] is an efficient coding method for steganography, which can reduce the number of embedding changes on the cover. Phillip [10] studied the secrecy security of matrix embedding under the condition of stego-object-only attack using the method of information theory. Based on message equivocation and key equivocation, He researched the secrecy security of matrix embedding under various key models.

So far, the relative research mainly focused on the security under the condition of stego-object-only attack. However, it is necessary to consider the security problem under the stronger condition of attacking, such as the known-stego-object attack. In general, it is difficult for the attacker to get the original cover, but if the attacker can process the stego object using the cover restoration technique and get a precise estimation of the original cover, the known-stego-object attack can be available [11]. Accordingly, both sides of communication have to consider the secrecy security of steganography under the stronger attack condition.

Based on the results in [10], we study the secrecy security of matrix embedding under the condition of known-cover attack by using the method of information theory. We prove that matrix embedding has weak security under the condition of chosen-ciphertext attack.

II. THE PRESENTATION OF PROBLEM

A. Symbols

Throughout the text, italic capital letters denote random variables, and boldface small letters denote the instances of random variables. Let Σ be the finite set of letters, Σ^n be a sequence of Σ, the length of which is n. Note that S is the cover sequence, C is the stego sequence, K is the shared key between sender and recepient, M is the secret message, T is the way used to choose the embedding positions. Security referred behind represents the secrecy security.

Considering that both sides of communication use matrix embedding to transfer secret message on the multimedia channel, such as digital images, audios, and videos, Figure 1 demonstrates the flow chart of communication.

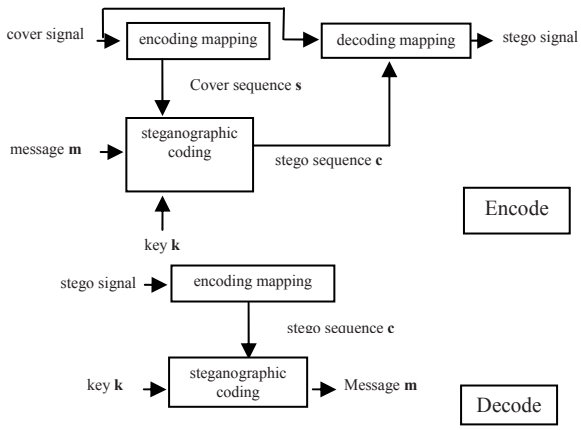

Figure. 1 Stegosystem using matrix embedding

Assume that both sides of communication want to transfer message $\mathbf{m}$ ($\mathbf{m} \in F_2^q$) with q bits, using a binary cover sequence $\mathbf{s}$ ($\mathbf{s} \in F_2^n$) whose length is n, where $0<q\leqslant n$. The sender maps the cover signal to a binary cover sequence

978-0-7695-4258-4/10 $26.00 © 2010 IEEE
DOI 10.1109/MINES.2010.150

by using the encoding mapping, embeds q bits message $\mathbf{m}$ into $\mathbf{s}$ and gets the stego sequence $\mathbf{c}$ ($\mathbf{c} \in F_2^n$), and then generates the stego signal. After getting the stego signal, the receiver gets the stego sequence $\mathbf{c}$ by using the decoding mapping, and extracts the secret message $\mathbf{m}$ by using the key $\mathbf{k}$, where $\mathbf{m}=\mathbf{kc}(\mathrm{mod}2)$. For example, the encoding mapping used widely is getting the LSB sequence of the cover signal, and the corresponding decoding mapping is using the stego sequence to take place of the LSB sequence of the cover signal.

Especially, as for a good encoding mapping function, the 2^n instances of the cover $\mathbf{s}$ are equiprobable. Assume that the message has been compressed, and then the 2^q instances of the message are equiprobable. The key $\mathbf{k}$ is a random matrix on $F_2^{q\times n}$.

If both sides of communication using wet paper codes, assume that there are l dry positions, and then the sender embeds message $\mathbf{m}$ with q bits by changing some dry positions of $\mathbf{s}$, where $q \leqslant l \leqslant n$. Note that the embedding rate is $r = q/n$.

The entropy is defined as $H(\cdot)$, where

$$H(S)= \sum_{\mathbf{s}\in F_2^n} \Pr(\mathbf{s}) \log_2 \Pr(\mathbf{s}) .$$

The mutual information is defined as $I(\cdot)$, where

$$I(M;C)= \sum_{\mathbf{m}\in F_2^n} \sum_{\mathbf{c}\in F_2^n} \Pr(\mathbf{m},\mathbf{c}) \log_2 \frac{\Pr(\mathbf{m},\mathbf{c})}{\Pr(\mathbf{m})\Pr(\mathbf{c})} .$$

As for the zero-distortion channel, the channel capacity is defined as $C=H(C|S)$, and the information transfer rate is defined as $R_m=H(M)/n$. Define $H(C|S)- H(M)$ as **embedding redundancy**. The **unicity distance** to a stegosystem is the number of objects, where the object is the original cover or the couple of original cover and stego cover, which makes the expectation of the number of pseudo-keys is 0.

A binary $[n,k]$ matrix embedding C is a linear subspace of F_2^n. Given the key $\mathbf{k}$, define the rank of $\mathbf{k}$ as q, and $\mathbf{k}$ is full rank. Then for any $\mathbf{b} \in F_2^n$, the vector $\mathbf{g} = \mathbf{kb} \in F_2^n$ is called the syndrome of $\mathbf{b}$. The set $C(\mathbf{m}) = \{\mathbf{b} \in F_2^n \mid \mathbf{m} \equiv \mathbf{kb}(\mathrm{mod}\,2)\}$ is called a coset. We have that cosets associated with different syndromes are disjoint. Let coset $C(\mathbf{m}) = \mathbf{b} + C$, where $\mathbf{b}$ is an arbitrary vector of $C(\mathbf{m})$. Therefore, there are 2^{n-k} disjoint cosets, with each consisting of 2^k vectors. Let $w(\mathbf{s})$ be the Hamming weight of the vector $\mathbf{s}$, and $d(\mathbf{s},\mathbf{c})$ be the Hamming distance between vectors $\mathbf{s}$ and $\mathbf{c}$.Let $e(\mathbf{m})$ be a coset leader, where

$$w(e(\mathbf{m})) = \min\{w(\mathbf{b}) \mid \mathbf{b} \in C(\mathbf{m})\}$$

coset $C(\mathbf{m})$ with the smallest Hamming weight and will be denoted as $e(\mathbf{m})$.

B. Matrix embedding

Matrix embedding was firstly proposed to improve the embedding efficiency of steganography, which is a typical application of linear covering codes. Using matrix embedding, we can embed message $\mathbf{m}$ with q bits into a binary cover sequence $\mathbf{s}$ by changing at most R bits. Then the embedding algorithm $Emb()$ is

$$Emb(\mathbf{s},\mathbf{m}) = \mathbf{s} + e(\mathbf{m} - \mathbf{ks}) = \mathbf{c} .$$

And the extracting algorithm $Ext()$ is

$$Ext(\mathbf{c}) = \mathbf{kc} ,$$

where

$$\mathbf{kc} = \mathbf{ks} + \mathbf{k}e(\mathbf{m} - \mathbf{ks}) = \mathbf{ks} + \mathbf{m} - \mathbf{ks} = \mathbf{m} .$$

Because $\mathbf{m}$ follows the uniform distribution, the average change number needed for the embedding process is equal to the average Hamming weight of all the coset leaders of a code C. The average change number is equal to the average distance to code. The distances of two arbitrary codewords of the same coset to code is equivalent, both equal to the Hamming weight of any coset leader of the coset, i.e. $d(\mathbf{s},C) = d(\mathbf{c},C) = w(e)$. Consequently, the average distance R_s of all the codewords is equal to the embedding average change number, i.e.

$$R_s = \frac{1}{2^n} \sum_{\mathbf{s}\in F_2^n} d(\mathbf{s},C) = \frac{1}{2^{n-k}} \sum_{i=1}^{2^{n-k}} w(e(\mathbf{s})) .$$

C. Wet paper codes

Using wet paper codes, the sender can embed message when some positions of the cover is restricted not to be changed, while the receiver can extract message without the information of the restricted positions.

Assume that the sender chooses l changeable bits s_j, $j \in L \subset \{1,2,\cdots n\}$, $|L| = l$, from a binary cover $\mathbf{s}$=(s_1, ..., s_n), while the remaining n-l bits can not be changed. The changeable positions are called dry positions and the unchangeable positions are called wet positions. The sender embeds the message into $\mathbf{s}$ by changing some s_j, $j \in L \subset \{1,2,\cdots n\}$ and gets $\mathbf{c}$, which satisfy

$$\mathbf{kc}=\mathbf{m} \qquad (1)$$

Let $\mathbf{v}=\mathbf{c}-\mathbf{s}$, then

$$\mathbf{kv}= \mathbf{m}-\mathbf{ks} \qquad (2)$$

Since n-l positions of $\mathbf{s}$ are not allowed to change, there are l unknowns v_j, $j \in L \subset \{1,2,\cdots n\}$, while the remaining n-l values v_i, $i \notin L$, are zeros. Thus, we can remove n-l unused columns from $\mathbf{k}$, and denote the obtained matrix as $\mathbf{h}$. We also remove n-l unused elements from $\mathbf{v}$, and denote the obtained vector as $\mathbf{u}$. We get the following equation from (2)

$$\mathbf{hu}= \mathbf{m}-\mathbf{ks}, \qquad (3)$$

where $\mathbf{h}$ is a binary $q\times l$ matrix and $\mathbf{u}$ is an unknown $l\times 1$ binary vector. The encoding of wet paper codes is completed by solving (3). The wet rate of the cover is denoted as $\alpha_{wet} = \frac{n-l}{n}$, where $0 \le \alpha_{wet} < 1$. The embedding rate of wet paper codes is denoted as $r_{wet} = \frac{q}{l}$. In fact, the random matrix embedding can be viewed as one of the special wet paper codes, of which the wet rate is 0.

III. SECRECY SECURITY ANALYSIS

The security of matrix embedding under the condition of stego-object-only attack has been discussed in [10]. This paper will discuss the security under the condition of known-cover attack and chose-stego-object attack separately.

A. Key equivocation

Theorem 1 [10]: For random steganographic codes and wet paper codes, under the condition of stego-object-only attack, the key equivocation function is bounded as

$$I(K;C) \le [q - H(M)] + [H(C) - H(S)].$$

Theorem 1 indicated that, under the condition of stego-object-only attack, if the embedded message has been compressed, i.e. $H(M) = q$, and the cover sequence obtained by the encoding mapping is random, i.e. $H(S) = n$, then $H(C) = n = H(S)$ and $I(K;C) = 0$. Here matrix embedding can achieve the perfect secrecy.

Lemma 1: For wet paper codes, under the condition of known-cover attack, we have

$$I(T;K,C,S,M) \ge \phi(n,q) \ge 0,$$

where $\phi(n,q) = q + \log C_n^q - \log \sum_{i=0}^{q} C_q^i C_{n-i}^{q-i}$.

Proof: Consider $I(T;K,C,S,M)$. Because the way of choosing embedding positions *T* and the embedded message *M* are not related to the key *K*. The receiver can extract the message without knowing *T*. Thus, for any given stego object *S*, we have

$$\begin{aligned} I(T;K,C,S,M) &= I(T;C,S) \\ &= H(T) - H(T \mid C,S) \end{aligned},$$

where $H(T) = \sum_{\mathbf{t} \in T} \Pr(\mathbf{t}) \log \frac{1}{\Pr(\mathbf{t})} = \log C_n^q$. The information about the way of choosing embedding positions *T* can only be obtained by comparing the difference between the stego object *C* and cover object *S*, consequently

$$\begin{aligned} H(T \mid C,S) &= \sum_{c,s} \Pr(\mathbf{c} \oplus \mathbf{s}) H(T \mid \mathbf{c} \oplus \mathbf{s}) \\ &= \sum_{i=0}^{q} \Pr(w(\mathbf{c} \oplus \mathbf{s}) = i) H(T \mid w(\mathbf{c} \oplus \mathbf{s}) = i) \end{aligned}$$

Because $\mathbf{c}$ is random on F_2^n , and the Hamming weight of $\mathbf{c} \oplus \mathbf{s}$ is $w(\mathbf{c} \oplus \mathbf{s})$. Then

$$\Pr(w(\mathbf{c} \oplus \mathbf{s}) = i) = \frac{C_n^i C_{n-i}^{q-i}}{\sum_{j=0}^{q} C_n^j C_{n-j}^{q-j}} = \frac{C_n^i C_{n-i}^{q-i}}{C_n^q \sum_{j=0}^{q} C_q^j} = \frac{C_n^i C_{n-i}^{q-i}}{C_n^q 2^q}$$

$$H(T \mid w(\mathbf{c} \oplus \mathbf{s}) = i) = \log C_{n-i}^{q-i}.$$

As $\sum_{i=0}^{q} \frac{C_n^i C_{n-i}^{q-i}}{C_n^q 2^q} = 1$, according to **Jensen Inequation**, we have

$$\begin{aligned} H(T \mid \mathbf{c} \oplus \mathbf{s}) &= \sum_{i=0}^{q} \frac{C_n^i C_{n-i}^{q-i}}{C_n^q 2^q} \cdot \log C_{n-i}^{q-i} \\ &\le \log \left(\frac{\sum_{i=0}^{q} C_n^i C_{n-i}^{q-i} C_{n-i}^{q-i}}{C_n^q 2^q} \right) = \log \left(\frac{C_n^q \sum_{i=0}^{q} C_q^i C_{n-i}^{q-i}}{C_n^q 2^q} \right) \\ &= \log \sum_{i=0}^{q} C_q^i C_{n-i}^{q-i} - q \end{aligned}$$

Then

$$\begin{aligned} I(T;K,C,S,M) &= H(T) - H(T \mid \mathbf{c} \oplus \mathbf{s}) \\ &\ge q + \log C_n^q - \log \sum_{i=0}^{q} C_q^i C_{n-i}^{q-i} \triangleq \phi(n,q). \end{aligned}$$

According to **Vandermonde Identical Equation**, $\sum_{i=0}^{n} C_n^i \cdot C_m^{k-i} = C_{n+m}^k$, for $0 \le i \le q$, we have

$$0 = q + \log C_n^q - \log \sum_{i=0}^{q} C_q^i C_{n-q}^{q-i} < \phi(n,q) < q + \log C_n^q - \log \sum_{i=0}^{q} C_q^i C_n^{q-i} = q.$$

As a result, $0 \le \phi(n,q) \le I(T;K,C,S,M)$.

Theorem 2: For random steganographic codes and wet paper codes, under the condition of known-cover attack, the key equivocation function is bounded as

$$I(K;S,C) \le [H(C \mid S) - H(M)] - \phi(n,q).$$

Specially, *I*(*K*;*S*,*C*) achieves the maximum value when $q = \frac{n}{2}$.

Proof: For the proof, first of all we have

$$\begin{aligned} &H(C,K,M,T,S) \\ &= H(K,M,T,S) + \underbrace{H(C \mid K,M,T,S)}_{=0} \\ &= H(K) + H(M) + H(S) + H(T) \end{aligned} \quad (4)$$

At the same time,

$$\begin{aligned} &H(C,K,M,T,S) \\ &= H(C) + H(S \mid C) + H(K \mid C,S) \\ &\quad + \underbrace{H(M \mid K,C,S)}_{=0} + H(T \mid K,C,S,M) \\ &= H(C) + H(S \mid C) + H(K \mid C,S) + H(T \mid K,C,S,M) \end{aligned} \quad (5)$$

As a result,

$$\begin{aligned} &H(K) + H(M) + H(S) + H(T) \\ &= H(C) + H(S \mid C) + H(K \mid C,S) + H(T \mid K,C,S,M) \end{aligned},$$

and

$$\begin{aligned} &[H(S) - H(S \mid C)] + [H(K) - H(K \mid C,S)] \\ &+ [H(T) - H(T \mid K,C,S,M)] = H(C) - H(M) \end{aligned}.$$

As

$$\begin{cases} I(C;S) = H(S) - H(S \mid C) \\ I(K;C,S) = H(K) - H(K \mid C,S) \\ I(T;K,C,S,M) = H(T) - H(T \mid K,C,S,M) \end{cases},$$

we have

$$I(C;S) + I(K;C,S) + I(T;K,C,S,M) = H(C) - H(M).$$

Then, from Lemma 1,

$$I(K;C,S) = H(C) - H(M) - I(C;S) - I(T;K,C,S,M)$$

$$= H(C \mid S) - H(M) - I(T; K, C, S, M)$$
$$\le [H(C \mid S) - H(M)] - \phi(n, q) \qquad (6)$$

$\phi(n,q)$ could not be further simplified to our known. The values of typical $\phi(n,q)$ is shown in Figure 2(n=30). When n is fixed, we have

(1) When $r \to 0$ or $r \to 1$, the key equivocation achieves the maximal upper bound, which is close to the hidden capacity;

(2) When $r \to 0.5$, the key equivocation achieves the minimal upper bound.

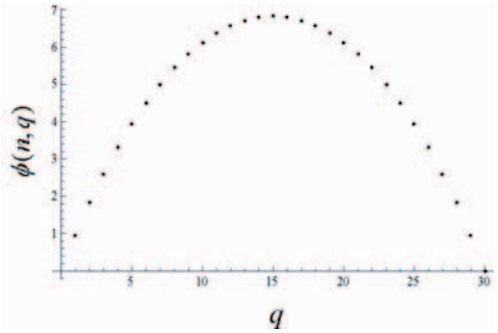

Figure. 2 $\phi(n,q)$ for n=30

Consider the relation between the embedding rate r_{wet} of wet paper codes and the key equivocation $I(K;C,S)$ for different wet rate α_{wet}. Generally speaking, we have

(1) $0.5 \le \alpha_{wet} < 1$. If $r_{wet} = 1$, the key equivocation achieves the minimum value; if $r_{wet} \to 0$, the key equivocation achieves the maximum value and is close to the embedding redundancy.

(2) $0 \le \alpha_{wet} < 0.5$. If $r_{wet} = 0.5(1 - \alpha_{wet})$, the key equivocation achieves the minimum value; If $r_{wet} \to 0$, the key equivocation achieves the maximum value and is close to the embedding redundancy.

(3) $\alpha_{wet} = 0$. If $r_{wet} = 0.5$, the key equivocation achieves the minimum value; If $r_{wet} \to 0$, the key equivocation achieves the maximum value and is close to the embedding redundancy. Under this circumstance, wet paper codes are steganographic codes.

Theorem 2 indicates that the channel redundancy should be decreased as much as possible to improve the secrecy of matrix embedding. In order to decrease the key equivocation, the sender should choose a suitable embedding rate according to the wet rate.

B. Unicity distance

Theorem 3: Under the circumstance of stego-object-only attack, unicity distance of the stegosystem is

$$N \ge \frac{H(K) - q}{H(C) - [H(M) + H(S)]}.$$

Proof: According to Theorem 1 in [10]

$$H(S \mid K, C, T)$$
$$= H(K) - H(K \mid C) + I(T; K, C) + H(M) + H(S) - H(C)$$

where $H(S \mid T, K, C) \le q$ and $I(T; K, C) = 0$, thus we have

$$H(K \mid C) = H(K) + H(M) + H(S) - H(C) - H(S \mid T, K, C)$$
$$\ge H(K) + H(M) + H(S) - H(C) - q$$

Denote N groups of stego objects as $\mathbf{c}^N = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N\}$, and then

$$H(\mathbf{k} \mid \mathbf{c}^N) \ge H(K) + N[H(M) + H(S) - H(C)] - q,$$

so we have

$$\log(\overline{K_p} + 1) \ge H(K) + N[H(M) + H(S) - H(C)] - q,$$

at the same time,

$$\overline{K_p} \ge 2^{H(K) - q + N[H(M) + H(S) - H(C)]} - 1.$$

So unicity distance is

$$N \ge \frac{H(K) - q}{H(C) - [H(M) + H(S)]}.$$

Theorem 4: Under the condition of known-cover attack, unicity distance of the key $\mathbf{k}$ is

$$N \ge \frac{H(K)}{H(C \mid S) - [H(M) + I(T; M, C, S, K)]}.$$

Proof: For the proof, denote N groups of covers and stego objects as $\mathbf{s}^N = \{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_N\}$ and $\mathbf{c}^N = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N\}$ separately. Given N pairs of covers and stego objects, the set of all the possible stego keys is

$$K(\mathbf{c}^N, \mathbf{s}^N) = \{\mathbf{k} \in K \mid \exists \mathbf{m}_i \in M, \mathbf{t}_i \in T, \Pr(\mathbf{m}_i) > 0 \text{ and } \mathbf{kc}_i = \mathbf{m}_i\}.$$

And the expectation of the number of pesudo keys is

$$\overline{K_p} = \sum_{(\mathbf{c}^N, \mathbf{s}^N) \in (C^N, S^N)} \Pr(\mathbf{c}^N, \mathbf{s}^N)[K(\mathbf{c}^N, \mathbf{s}^N) - 1]$$
$$= \sum_{(\mathbf{c}^N, \mathbf{s}^N) \in (C^N, S^N)} \Pr(\mathbf{c}^N, \mathbf{s}^N) K(\mathbf{c}^N, \mathbf{s}^N) - 1.$$

We have

$$H(K \mid \mathbf{c}^N, \mathbf{s}^N)$$
$$= \sum_{(\mathbf{c}^N, \mathbf{s}^N) \in (C^N, S^N)} \Pr(\mathbf{c}^N, \mathbf{s}^N) H(K \mid \mathbf{c}^N, \mathbf{s}^N)$$
$$\le \log \sum_{(\mathbf{c}^N, \mathbf{s}^N) \in (C^N, S^N)} \Pr(\mathbf{c}^N, \mathbf{s}^N) \mid K(\mathbf{c}^N, \mathbf{s}^N) \mid$$
$$= \log(\overline{K_p} + 1)$$

and $H(C \mid K, M, T, S) = 0$. As the same time

$$H(\mathbf{c}^N, \mathbf{s}^N, \mathbf{m}^N, \mathbf{t}^N, \mathbf{k})$$
$$= H(\mathbf{s}^N, \mathbf{m}^N, \mathbf{t}^N, \mathbf{k}) + \underbrace{H(\mathbf{c}^N \mid \mathbf{s}^N, \mathbf{m}^N, \mathbf{t}^N, \mathbf{k})}_{=0}$$
$$= N[H(S) + H(M) + H(T)] + H(K) \qquad (7)$$

Thus,

$$H(\mathbf{c}^N, \mathbf{s}^N, \mathbf{m}^N, \mathbf{t}^N, \mathbf{k})$$
$$= H(\mathbf{s}^N) + H(\mathbf{c}^N \mid \mathbf{s}^N) + H(\mathbf{k} \mid \mathbf{c}^N, \mathbf{s}^N)$$
$$+ H(\mathbf{m}^N \mid \mathbf{c}^N, \mathbf{s}^N, \mathbf{k}) + H(\mathbf{t}^N \mid \mathbf{m}^N, \mathbf{c}^N, \mathbf{s}^N, \mathbf{k})$$
$$\le NH(S) + NH(C \mid S) + H(\mathbf{k} \mid \mathbf{c}^N, \mathbf{s}^N)$$
$$+ H(\mathbf{t}^N \mid \mathbf{m}^N, \mathbf{c}^N, \mathbf{s}^N, \mathbf{k})$$

$$\le N[H(S)+H(C\mid S)+H(T)-I(T;M,C,S,K)] + H(\mathbf{k}\mid \mathbf{c}^N,\mathbf{s}^N) \qquad (8)$$

From equation (7) and equation (8),

$$H(\mathbf{k}\mid \mathbf{c}^N,\mathbf{s}^N)\ge H(K)-N[H(C\mid S)-H(M)-I(T;M,C,S,K)]$$

Then

$$\log(\overline{K_p}+1)\ge H(K)-N[H(C\mid S)-H(M)-I(T;M,C,S,K)].$$

Consequently, the expectation value of pesudo-keys is

$$\overline{K_p}\ge H(\mathbf{k}\mid \mathbf{c}^N,\mathbf{s}^N)\ge 2^{H(K)-N[H(C|S)-H(M)-I(T;M,C,S,K)]}-1.$$

So unicity distance is

$$N\ge \frac{H(K)}{H(C\mid S)-[H(M)+I(T;M,C,S,K)]}.$$

C. Differential attack

Theorem 5: Under the circumstance of chosen-stego-object attack, the attacker can restore the key, using *n* groups of differential equations.

Proof: For the proof, assume that by some way the attacker has already known both sides of communication use wet paper codes to transfer the secret message, and here wet paper codes can be viewed as an encryption algorithm. If the cover group is considered as plaintext and the stego group is considered as ciphertext, wet paper codes are a block cipher in fact, which uses the same key $\mathbf{k}$ to encrypt different groups of plaintexts. The length of plaintext group is *n*, the same as the length of ciphertext group. The goal of the attacker is to restore the key $\mathbf{k}$. Assume that the attacker has already had many plaintext-ciphertext pairs, and could choose the needful stego image block $\mathbf{c}$ and the corresponding message group $\mathbf{m}$, which amounts to making a chosen-ciphertext attack.

Then an attack is given under the above circumstance, of which the main idea is that some information of the key is obtained by using differential attack and a group of equivalent keys are found by solving a group of linear equations.

The following operations are discussed on F_2. Because the attacker can choose stego objects, he can get two groups of stego images $\mathbf{c}_1$ and $\mathbf{c}_1'$, where $\mathbf{c}_1-\mathbf{c}_1'=(1,0,0,\cdots,0)^{\mathbf{T}}$. Let the corresponding plain messages of them are $\mathbf{m}_1$ and $\mathbf{m}_1'$ separately, so differential equations can be obtained by

$$\mathbf{k}\mathbf{c}_1-\mathbf{k}\mathbf{c}_1'=\mathbf{m}_1-\mathbf{m}_1'.$$

Since $\mathbf{c}_1-\mathbf{c}_1'=(1,0,0,\cdots,0)^{\mathbf{T}}$, we have

$$\mathbf{k}(\mathbf{c}_1-\mathbf{c}_1')=\mathbf{k}_{j,1}=\mathbf{m}_1-\mathbf{m}_1'.$$

Similarly, for $\mathbf{c}_i-\mathbf{c}_i'=(0,0,\cdots\overset{i}{\overset{\sim}{1}},\cdots,0)^{\mathbf{T}}$, we can get differential equations

$$\mathbf{k}_{j,i}=\mathbf{m}_i-\mathbf{m}_i',$$

where $i=1,2,\cdots,n$. Because the attacker can get *q* bits of $\mathbf{k}$ by solving a group of equations, only *n* groups of differential equations need to be constructed, and then the key $\mathbf{k}$ is obtained.

Theorem 5 indicates that, the security of matrix embedding is very weak under the circumstance of chosen-stego-object attack.

IV. Conclusion

In this paper, the relation between the embedding rate of matrix embedding and the key equivocation is given under the circumstance of known-cover attack, as well as the relation among the wet rate, embedding rate and the key equivocation. But it should be pointed out that the cover-object cannot be exactly estimated, and the restoration of cover object is in fact a very difficult problem in steganography. We also conclude that for matrix embedding, only *n* groups of differential equations are needed to restore the key $\mathbf{k}$ under the condition of chosen-stego-object attack.

V. Acknowledgment

This work was supported by the Natural Science Foundation of China (60803155,60970141,60902102) and the National High Technology Research Program of China(863 Program) (2008AA01Z420).

References

[1] J. Fridrich, M. Goljan, D. Soukal. “Searching for the stego key”[C]. Security, Steganography and Watermaking of Multimedia Contents, 70-82, 2004.

[2] J. Fridrich, M. Goljan, D. Soukal. etc. Forensic Steganalysis: “Determining the Stego Key in Spatial Domain Steganography”[C]. Security, Steganography and Watermaking of Multimedia Contents, 2005, 631-642.

[3] Zhang Weiming, Li Shiqu. “Security Measurements of Steganographic Systems”. Proc. The Second International Conference of Applied Cryptogarphy and Network Security(ACNS’04). LNCS Vol.3089, Springer-Verlag, 194-204, 2004.

[4] Zhang Weiming, Li Shiqu, Cao Jia, Liu Jiufen. “Information-Theoretic Analysis for the Difficulty of Extracting Hidden Information”. Wuhan University Journal of Natural Sciences, Vol.10(1):315-318, 2004.

[5] J. Fridrich, P. Lisonˇek, and D. Soukal, “On steganographic embedding efficiency,” in Proc. 8th Int. Workshop on Information Hiding, 2006, vol. 4437 of Springer LNCS, pp. 282–296.

[6] J. Fridrich and D. Soukal, “Matrix embedding for large payloads,” IEEE Trans. Inf. Forensics Security, vol. 1, no. 3, pp. 390–394, Sep. 2006.

[7] J. Fridrich, M. Goljan, P. Lisonek and D. Soukal, “Writing on wet paper,” IEEE Trans. Signal Process., vol. 53, pp. 3923–3935, 2005.

[8] J. Fridrich, M. Goljan, and D. Soukal, “Wet paper codes with improved embedding efficiency,” IEEE Trans. Inform. Forensics Secur., vol. 1, no. 1, pp. 102–110, 2006.

[9] J. Fridrich, M. Goljan, and D. Soukal, “Efficient wet paper codes,” in Proc. 7th Int. Workshop on Information Hiding, LNCS 3727, pp. 204–218, 2005.

[10] Phillip A.Regalia. “Cryptographic Secrecy of Steganographic Matrix Embedding”. IEEE Transactions on Information Forensics Security, vol.3, no.4, December 2008.

[11] Zhang Weiming, Li Shiqu, Liu Jiufen. “Extracting Attack against the LSB Steganoraphy of the Space Domain Images”. Journal of Computers. Vol. 30(9): 1625-1631, 2007.