Security Measurements of Steganographic Systems

Weiming Zhang and Shiqu Li

Department of Applied Mathematics, University of Information Engineering,
P.O.Box 1001-747, Zhengzhou 450002, P.R. China
{nlxd_990, ShiquLi}@yahoo.com.cn

Abstract. Different security measurements for a steganographic system, i.e. security (detectability), robustness and secrecy (difficulty of extraction), are discussed in this paper. We propose a new measurement for the security of stegosystems using variational distance which can upper bound the advantage for passive attackers. It is proved that the hiding capacity, which is also the measurement for robustness, is limited by security. We think the extracting attack essentially is a kind of cryptanalysis and define the secrecy of stegosystems as an analogue of secrecy of cryptosystems. The relations of secrecy with capacity and security are analyzed in the terms of unicity distance. And it is shown that there is a tradeoff between secrecy and capacity while there is some kind of consistency between secrecy and security.

1 Introduction

This paper is about steganography which is the oldest branch of information hiding. The scientific study of steganography began with Simmons' "Prisoners' Problem" [1]. The survey about the history and current development of it can be found in [2] and [3]. A general model of a steganographic system (i.e. stegosystem) can be described as follows. The embedded data M is the message that Alice wants to send secretly to Bob. It is hidden in an innocuous message $\widetilde{X}$, usually named cover-object, in the control of a stego-key K, producing the stego-object X. And the receiver can extract M from X with the stego-key K.

The attacks to a stegosystem mainly include passive attack, active attack, and extracting attack. A passive attacker only wants to detect the existence of the embedded message, while an active attacker wants to destroy the embedded message. The purpose of an extracting attacker is to obtain the message hidden in the stego-object. So there are three kinds of security measurements for the different attackers respectively, i.e. detectability, robustness and difficulty of extraction. Usually the problem of steganography only concerns the detectability so in many literatures detectability is referred to as the security of a stegosystem. In this paper, we also call the detectability as security of a stegosystem and the difficulty of extraction as secrecy of it. But so far the definitions of the three security measurements are still tangly and relations of them are still unclear. The

M. Jakobsson, M. Yung, J. Zhou (Eds.): ACNS 2004, LNCS 3089, pp. 194–204, 2004.

main purpose of this paper is just to distinguish their definitions and analyze relations between them.

So far there have been several literatures that define the security (detectability) of stegosystems, such as [4,5,6,7], and the one of C.Cachin [5] is most influential. Cachin formulates the steganography problem as a hypothesis testing problem and defines the security using the statistic distance between the cover-object and stego-object which indeed catches the key of detectability. But, he uses the relative entropy as the security measurement which, to some extent, seems not appropriate. According to Cachin's definition the stegosystem is ε-secure when the relative entropy $D(\widetilde{X}||X) \leq \varepsilon$, and perfectly secure when $\varepsilon = 0$. Supposing the false alarm probability (the probability of a cover-object being mistaken as a stego-object) equals zero, Cachin uses the relative entropy to estimate the lower bound of missing probability (the probability of a stego-object being mistaken as a cover-object). However, it is evident that the adversary will not use a rule such that he makes the false alarm probability very small, because this means he will leak the illegal messages in a large probability. For instance, in Cachin's model, when the stegosystem is perfect security, the probability of the adversary finding the stego-object equals zero. But the fact is that even guessing randomly, he could success with probability $\frac{1}{2}$.

S.Katzenbeisser and F.A.Petitcolas [8] defines security in computational settings, and their definition still need a security measurement which is referred as to the advantage for a adversary, i.e. the probability of the adversary's successful detection minus $\frac{1}{2}$. This description for stegosystem's security is reasonable, but it is a description in words. And the definition of R.Chandramouli and N.D.Memon [9] can be though of as a mathematic version of description in [8], and their definition is related with the strategy of attackers. In fact we hope there is a metric that can reflect the adversaries' advantage, and in this paper we will propose such a metric with variational distance.

Information hiding with active attackers were analyzed by P.Moulin and J.A.O'Sullivan [10] and M.Ettinger [11]. They defines the robustness using "hiding capacity". Robustness is mainly concerned in watermarking problem, but as the measure of efficiency, capacity is also important for steganography. I.S.Moskoxitz et al. [7] proposed a two dimension security measure for steganography, i.e. *capability* $= (P, D)$ where P is the payload size and D is detectability threshold. In this paper, we prove that the capacity is limited by detectability, and for stegosystems with active attackers this shows a tradeoff between the security and robustness.

The security and robustness have been greatly concerned. However there is scarcely any literature about extracting attacks. We only know that R.Chandramouli ever studied how to extract the hidden message for some kind of scenario in [12], and J.Fridrich et al. recently presented a methodology for identifying the stego-key in [13]. In fact, for most of stegosystems the message is asked to be encrypted before it is embedded into the cover-object, so the secrecy is guaranteed by the cryptographic algorithm. So stegoanalysts only concern detection and think extraction is the task of cryptanalysts, while the latter only process

encrypted data. But how to extract the hidden message is a very difficult problem itself. We think the extracting attack essentially is a kind of cryptanalysis. When facing the model of "encrytion+hiding", a cryptanalyst has to analyze a "multiple cipher": he should extract the hidden messag (the ciphertexts) from stego-objects, and then extract the plaintexts from the hidden message. In this paper, we distinguish the secrecy of steganography from that of cryptography. If the message has been encrypted, the extraction attacker is successful as long as he can extract the cipertexts. So the secrecy of steganography is just the difficulty of extraction. Because extracting attack is a kind of cryptanalysis, we define the secrecy of steganography imitating Shannon's definition for unconditional security of cryptosystems [14], i.e. measuring the secrecy with mutual information $I(M;X)$ or $I(M;X,\widetilde{X})$. And we will analyze the relations between security, capacity and secrecy.

The rest of this paper is organized as follows: Section 2 defines the security of stegosystems with variational distance and estimates the upper bound of the advantage for passive adversaries. Section 3 proves the tradeoff between the security and capacity. Section 4 defines the perfect secrecy for only stego-object extracting attack and known cover-object extracting attack respectively, and analyzes the relations between capacity, security and secrecy in terms of unicity distance. The paper concludes with a discussion in Sect. 5.

2 Security of Stegosystems

2.1 Notations and Statement of Problem

We use the following notations. Random variables are denoted by capital letters (e.g. X), and their realizations by respective lower case letters (e.g. x). The domains over which random variables are defined are denoted by script letters (e.g. $\mathcal{X}$). Sequences of n random variables are denoted with a superscript n (e.g. $X^n = (X_1, X_2, \cdots, X_n)$ which takes its values on the product set $\mathcal{X}^n$). The probability mass function (p.m.f.) of random variable X is denoted by $P_X(x)$, and when no confusion is possible, we drop the subscript.

Definition 1. [15] *Let $\widetilde{X}$ and X are two random variables on a discrete universe $\mathcal{X}$, then the variational distance between $\widetilde{X}$ and X is defined to be*

$$VD(\widetilde{X}, X) = \max_{S \subseteq \mathcal{X}} |P_{\widetilde{X}}(S) - P_X(S)| \ .$$

Lemma 1. [16] *Let $\widetilde{X}$ and X are two random variables on a discrete universe $\mathcal{X}$, and $\mathcal{T}$ is another discrete universe, then for any function $f : \mathcal{X} \to \mathcal{T}$, $VD(f(\widetilde{X}), f(X)) \leq VD(\widetilde{X}, X)$.*

In this paper, $\widetilde{X}$ stands for cover-object, taking values in $\mathcal{X}$. M denotes the hidden message, K is the stego-key (embedding key). X, which is also defined in $\mathcal{X}$, denotes the stego-object. Here hidden message is what will ultimately

be embedded into the cover-object which usually is encrypted data. And the stego-key only refers to the embedding key excluding the encrytion key. E is the embedding algorithm, with which the sender Alice embeds m into $\widetilde{x}$ to get x using k, i.e. $x = E(\widetilde{x}, m, k)$. And D is the extracting algorithm used by receiver Bob, which satisfies $m = D(x, k) = D(E(\widetilde{x}, m, k), k)$. We denote a stegosystem by a set with 6 elements: $stegosystem(\widetilde{X}, X, M, K, E, D)$.

The present paper mainly follows the view of Cachin [5] who formulated the steganography problem with passive attackers as a hypothesis testing problem. Alice, who maybe uses a stegosystem, sends data to Bob. The passive adversary Wendy observes the data and makes a hypothesis testing. Here the original hypothesis H_0 is that the data is generated according to $\widetilde{X}$, i.e. Alice sent a cover-object. And the opposite hypothesis H_1 is that the data is generated according to X, i.e. Alice sent a stego-object. The probability that Wendy fails to detect a stego-object is called missing probability and denoted by β. And the probability that she thinks of a cover-texts as a stego-object is called false alarm probability and denoted by α.

2.2 Security of Stegosystem

Variational distance can reflect the statistic difference of two probability distributions as relative entropy does. What's more, Variational distance is a distance in the sense of mathematics and take values between zero and one. So with variational distance as the measurement, we can compare the security of different stegosystems. We define the security of a stegosystem as follows.

Definition 2. *A* $stegosystem(\widetilde{X}, X, M, K, E, D)$ *is called* ε*-secure, if*

$$VD(\widetilde{X}, X) \leq \varepsilon \ .$$

And when $\varepsilon = 0$*, the system is called perfectly secure.*

With relative entropy as the security measure, Cachin [5] yields a lower bound on the missing probability β, i.e. if $D(\widetilde{X}||X) \leq \varepsilon$ and the false alarm probability $\alpha = 0$, then $\beta \geq 2^{-\varepsilon}$. But, as the analysis in the Sect. 1, what we need is the estimation about the advantage for adversaries. To do this, we define the event of successful attack as

$$\begin{aligned} SUCC = &\{H_0 \text{ is true and Wendy accepts } H_0\} \\ &\cup \{H_1 \text{ is true and Wendy accepts } H_1\} \ . \end{aligned}$$

And its complementary event is defined to be

$$\begin{aligned} \overline{SUCC} = &\{H_0 \text{ is true and Wendy accepts } H_1\} \\ &\cup \{H_1 \text{ is true and Wendy accepts } H_0\} \ . \end{aligned}$$

It is reasonable for Wendy to suppose the prior probability of both H_0 and H_1 is that $P(H_0) = P(H_1) = \frac{1}{2}$, because the event that which kind of object

Alice will send is random for Wendy who wants to get some advantage through the observed data. So the advantage for the adversary (Adv) is defined by

$$Adv = |P(SUCC) - \frac{1}{2}| \ . \tag{1}$$

As for Adv, using the security measurement in definition 2 we can yield the following result.

Theorem 1. *If a stegosystem*$(\widetilde{X}, X, M, K, E, D)$ *is* ε*-secure, then the advantage for the adversary satisfies* $Adv \leq \frac{\varepsilon}{2}$. *And when the system is perfectly secure, i.e.* $\varepsilon = 0$, *then* $Adv = 0$.

Proof. Note that the probabilities of two type errors made by Wendy are just that $\alpha = P\{$*Wendy accepts* $H_1 | H_0$ *is true*$\}$, *and* $\beta = P\{$*Wendy accepts* $H_0 | H_1$ *is true*$\}$.
Combing these two equalities with the fact $P(H_0) = P(H_1) = \frac{1}{2}$, *we have* $P(\overline{SUCC}) = \frac{1}{2}(\alpha + \beta)$ *and then*

$$P(SUCC) = 1 - \frac{1}{2}(\alpha + \beta) \ . \tag{2}$$

The probabilities of the two type errors, α *and* β *can induce two* $0-1$ *random variables as follows:*

	0	1
$\widetilde{X}'$	α	$1-\alpha$
X'	$1-\beta$	β

$\widetilde{X}'$ *and* X' *can be get through a same function from* $\widetilde{X}$ *and* X, *so using Lemma 1 we can obtain that* $VD(\widetilde{X}', X') \leq VD(\widetilde{X}, X)$, *i.e.* $1-\varepsilon \leq \alpha+\beta \leq 1+\varepsilon$, *which with (2) implies that* $\frac{1}{2} - \frac{\varepsilon}{2} \leq P(SUCC) \leq \frac{1}{2} + \frac{\varepsilon}{2}$, *i.e.* $Adv \leq \frac{\varepsilon}{2}$. □

Theorem 1 shows that if a stegosystem is ε-security the advantage for a passive adversary using any decision rule over the adversary guessing randomly will not larger than $\frac{\varepsilon}{2}$. And if the stegosystem is perfectly secure, then any decision rule used by the adversary will not more effective than guessing randomly. That means that the knowledge the adversary get through observing data about whether Alice has sent stego-object or not is zero. So the metric given in Definition 2 accurately depicts the security of stegosystems.

3 Tradeoff between Security and Capacity

Moulin and O'Sullivan. [10] and Ettinger [11] view the information hiding problem as a capacity game between the users of a stegosystem and the active attacker. According to formulations in [10], a strategy of the sender is just a "covert channel", i.e. a conditional p.m.f $\widetilde{Q}(x, u|\widetilde{x}, k)$, subject to distortion D_1. Here U is an auxiliary random variable. $\widetilde{\mathcal{Q}}$ is the set of all such cover channels. The

attacker's output is denoted by Y, and a strategy of the attacker is described as a "attack channel", i.e. a conditional p.m.f $Q(y|x)$, subject to distortion D_2. And The set of all such attack channels is denoted by $\mathcal{Q}$. The hiding capacity is defined as the upper-bound of rates of reliable transmission of the hidden message. Moulin and O'Sullivan obtained a expression for the hiding capacity as follows:

$$C = \max_{\widetilde{Q} \in \widetilde{\mathcal{Q}}} \min_{Q \in \mathcal{Q}} [I(U;Y|K) - I(U;\widetilde{X}|K)] \ . \tag{3}$$

where $(U, \widetilde{X}, K) \to X \to Y$ is a Markov chain.

In this section, we discuss the relation between the detectability (security) and the capacity (robustness) of general information hiding problems. We think the detectability of a information hiding code should include two parts: one is the sensual detectability (transparency) which is needed by any information hiding problem such as watermarking, steganography and fingerprint, the other is statistic detectability which is just the security of steganography. The former means the stego-object is a good estimation of the cover-object, so it can be measured by the probability $p_e = P(X \neq \widetilde{X})$ which is relative with the conditional entropy $H(\widetilde{X}|X)$, and the latter can be measured by the advantage for adversaries which, as we have proved in Sect. 2, is relative with the varational distance $VD(\widetilde{X}, X)$. Theorem below shows that there is a tradeoff between the detectability and the capacity.

Lemma 2. [16] *Let X and $\widetilde{X}$ are random variables on a discrete universe $\mathcal{X}$, and $VD(\widetilde{X}, X) = \varepsilon$. Then $|H(X) - H(\widetilde{X})| \leq H(\varepsilon) + \varepsilon \log_2(|\mathcal{X}| - 1)$.*

Theorem 2. *For a stegosystem$(\widetilde{X}, X, M, K, E, D)$, if $P(X \neq \widetilde{X}) = p_e$, $VD(\widetilde{X}, X) = \varepsilon$ and the hiding capacity is C, then we have*

$$C \leq H(p_e) + H(\varepsilon) + (p_e + \varepsilon) \log_2(|\mathcal{X}| - 1) \ . \tag{4}$$

Proof.

$$\begin{aligned}
& I(U;Y|K) - I(U;\widetilde{X}|K) \\
\overset{(a)}{\leq}\ & I(U;X|K) - I(U;\widetilde{X}|K) \\
=\ & [I(U;\widetilde{X},X|K) - I(U;\widetilde{X}|X,K)] - [I(U;\widetilde{X},X|K) - I(U;X|\widetilde{X},K)] \\
=\ & I(U;X|\widetilde{X},K) - I(U;\widetilde{X}|X,K) \\
\leq\ & I(U;X|\widetilde{X},K) \\
\leq\ & H(X|\widetilde{X},K) \\
\leq\ & H(X|\widetilde{X}) \\
=\ & H(X) - I(X;\widetilde{X}) \\
=\ & [H(X) - H(\widetilde{X})] + H(\widetilde{X}|X) \\
\overset{(b)}{\leq}\ & H(\varepsilon) + \varepsilon \log_2(|\mathcal{X}| - 1) + H(p_e) + p_e \log_2(|\mathcal{X}| - 1) \\
=\ & H(p_e) + H(\varepsilon) + (p_e + \varepsilon) \log_2(|\mathcal{X}| - 1) \ .
\end{aligned}$$

Where (a) follows from the data processing inequality applied to the Markov chain $(U, \widetilde{X}, K) \to X \to Y$. *(b) is obtained from the Lemma 2 and Fano's inequality. And combining the inequality above with (3) just proves the theorem.* □

On account of the meaning of p_e and Theorem 1, it is reasonable for us to suppose that $p_e \leq \frac{1}{2}$ and $\varepsilon \leq \frac{1}{2}$. Under this condition, the right of (4) increases with p_e and ε. So Theorem 2 shows a tradeoff between the capacity and detectability. And the upper-bound of hiding capacity includes two symmetrical parts: the first part is a function of sensual detectability, i.e. $H(p_e) + p_e \log_2(|\mathcal{X}| - 1)$, and the second part is a function of statistic detectability (security), i.e. $H(\varepsilon) + \varepsilon \log_2(|\mathcal{X}| - 1)$. Given p_e, Theorem 2 means a tradeoff between the security and capacity, and for information hiding problems with active attackers this is just the tradeoff between the security and robustness.

4 The Relations between Capacity, Security, and Secrecy

Since the extracting attack to a stegosystem in principle is a kind of cryptanalysis, we define the secrecy of stegosystems simulating the one of Shannon's [14] for cryptosystems.

Definition 3. *a stegosystem*$(\widetilde{X}, X, M, K, E, D)$ *is perfectly secret for only stego-object extracting attack if* $I(M; X) = 0$*, and is perfectly secret for known cover-object extracting attack if* $I(M; X, \widetilde{X}) = 0$.

J.Zölner et al. [4] ever defined the security of stegosystem using $I(M; X, \widetilde{X})$, but what they wanted to describe was the detectability, which seemed not appropriate because of the difference between the security and secrecy.

In this section, we only discuss the steganographic problem without active attackers. And suppose that stego-key K is independent with M and X. In this scenario, the result of [10] combined with the discussion in [17] implies that the hiding capacity

$$C = \max_{P(X|\widetilde{X})} H(X|\widetilde{X}) \ . \tag{5}$$

We also suppose that both the source of cover-objects and the channel $P(X|\widetilde{X})$ are memoryless. This seems not realistic, but we can think that X and $\widetilde{X}$ are both stand for block data, and usually supposing blockwise memoryless is reasonable.

What the extracting attacker ultimately wants to obtain is just the stego-key. Therefore we analyze the relations between capacity, security and secrecy in the terms of unicity distance for the stego-key. And we begin with the known cover-object extracting attack.

Lemma 3. *For a stegosystem*$(\widetilde{X}, X, M, K, E, D)$*, if* K *is independent with* $\widetilde{X}$*, then* $H(K|\widetilde{X}, X) = H(K) + H(M|\widetilde{X}, K) - H(X|\widetilde{X})$.

Proof. Because X *can be determined by* $(\widetilde{X}, M, K)$*, and* M *can be determined by* (X, K)*, we have* $H(X|\widetilde{X}, M, K) = 0$*, and* $H(M|X, K) = 0$*. So*

$$\begin{aligned} H(\widetilde{X}, M, K) &= H(\widetilde{X}, M, K) + H(X|\widetilde{X}, M, K) \\ &= H(\widetilde{X}, X, M, K) \\ &= H(\widetilde{X}, X, K) + H(M|\widetilde{X}, X, K) \\ &= H(\widetilde{X}, X, K) \ . \end{aligned}$$

Since K *is independent with* $\widetilde{X}$*, using the chain rules we have*

$$\begin{aligned} H(\widetilde{X}, M, K) &= H(K) + H(\widetilde{X}|K) + H(M|\widetilde{X}, K) \\ &= H(K) + H(\widetilde{X}) + H(M|\widetilde{X}, K) \ , \end{aligned}$$

and

$$H(\widetilde{X}, X, K) = H(\widetilde{X}) + H(X|\widetilde{X}) + H(K|\widetilde{X}, X) \ .$$

Combining the three equalities above, we can get

$$H(K|\widetilde{X}, X) = H(K) + H(M|\widetilde{X}, K) - H(X|\widetilde{X}) \ .$$

□

Theorem 3. *For a stegosystem*$(\widetilde{X}, X, M, K, E, D)$*, if* K *is independent with* $\widetilde{X}$ *and* M*, and both source of cover-objects and cover channel are memoryless, then for given long enough sequence (the length is* n*) of pairs of cover-objects and stego-objects, the expectation of spurious stego-keys* $\overline{S}_n$ *for known cover-object extracting attack has the lower bound such that*

$$\overline{S}_n \geq \frac{2^{H(K)}}{2^{nC}} - 1 \ , \tag{6}$$

where $C = \max_{P(X|\widetilde{X})} H(X|\widetilde{X})$ *is the hiding capacity.*

Proof. For a given sequence of pairs of cover-objects and stegotexts $(\widetilde{x}^n, x^n) \in (\mathcal{X}^n \times \mathcal{X}^n)$*, defining the set of possible stego-keys as*

$$K(\widetilde{x}^n, x^n) = \{k \in \mathcal{K} | \text{there is } m^n \in \mathcal{M}^n \text{ such that } P(m^n) > 0,\ E(\widetilde{x}^n, m^n, k) = x^n\} \ .$$

So the number of spurious stego-keys for observed $(\widetilde{x}^n, x^n)$ *is* $|K(\widetilde{x}^n, x^n) - 1|$*, and the expectation of spurious stego-keys is given by*

$$\overline{S}_n = \sum_{(\widetilde{x}^n, x^n)} P(\widetilde{x}^n, x^n)(|K(\widetilde{x}^n, x^n) - 1|) = \sum_{(\widetilde{x}^n, x^n)} P(\widetilde{x}^n, x^n)|K(\widetilde{x}^n, x^n)| - 1 \ .$$

Using Jesen's inequality, we can get

$$\begin{aligned} H(K|\widetilde{X}^n, X^n) &= \sum_{(\widetilde{x^n}, x^n)} P(\widetilde{x}^n, x^n) H(K|\widetilde{x}^n, x^n) \\ &\leq \sum_{(\widetilde{x^n}, x^n)} P(\widetilde{x}^n, x^n) \log_2 |K(\widetilde{x}^n, x^n)| \\ &\leq \log_2 \sum_{(\widetilde{x^n}, x^n)} P(\widetilde{x}^n, x^n) |K(\widetilde{x}^n, x^n)| \\ &= \log_2(\overline{S}_n + 1) \ . \end{aligned}$$

On the other hand, Lemma 3 and the fact that source of cover-objects and cover channel are memoryless implies that

$$\begin{aligned} H(K|\widetilde{X}^n, X^n) &= H(K) + \quad H(M^n|\widetilde{X}^n, K) - H(X^n|\widetilde{X}^n) \\ &\geq H(K) - H(X^n|\widetilde{X}^n) \\ &= H(K) - nH(X|\widetilde{X}) \ . \end{aligned}$$

Combing the two inequalities above, we have $\log_2(\overline{S}_n + 1) \geq H(K) - nH(X|\widetilde{X})$, *i.e.*

$$\overline{S}_n \geq \frac{2^{H(K)}}{2^{nH(X|\widetilde{X})}} - 1 \ .$$

Since $C = \max_{P(X|\widetilde{X})} H(X|\widetilde{X})$, *we have*

$$\overline{S}_n \geq \frac{2^{H(K)}}{2^{nC}} - 1 \ .$$

□

Definition 4. *The unicity distance* n_0 *for a stegosystem with known cover-object extracting attackers is the length of pairs of cover-objects and stego-objects at which one expects that the expectation of spurious stego-keys equals zero. And the unicity distance* n_1 *for a stegosystem with only stego-object extracting attackers is the length of stego-objects at which one expects that the expectation of spurious stego-keys equals zero.*

It is easy to know that $n_1 \geq n_0$, because $H(K|X) \geq H(K|\widetilde{X}, X)$. What's more, in (6), let $\overline{S}_n = 0$ and we have

$$n_1 \geq n_0 \geq \frac{H(K)}{C} \ . \tag{7}$$

Inequality (7) with Theorem 2 implies that

$$n_1 \geq n_0 \geq \frac{H(K)}{H(p_e) + H(\varepsilon) + (p_e + \varepsilon)\log_2(|\mathcal{X}| - 1)} \ . \tag{8}$$

For a stegosystem, (7) shows a tradeoff between the secrecy and capacity, while (8) shows some king of consistency of secrecy with security.

5 Conclusion

In this paper, three kind of security measuremeasures of stegosystems are discussed together. The relations and differences between them are analyzed with information theoretic method. We substitute variational distance for relative entropy to measure the security (detectability) of a stegosystem. This new measurement can upper bound the advantage for passive attackers. And it is proved out that the capacity (i.e. the robustness for stegosystems with active attackers) is limited by security. So an interesting problem is what the expression of hiding capacity subject to some security level ε is. Recently, P.Moulin and Y.Wang derived the capacity expression for perfectly secure (i.e. $\varepsilon = 0$) steganographic systems [20].

Our definition for secrecy is an analogue of Shannon's for cryptosystems. And it is shown that there is a tradeoff between secrecy and capacity but some kind of consistency of secrecy with security. However, the lower bound for unicity distance in Sect. 4 is rough. And a more useful lower bound will be discussed with the redundancy of cover channel in our upcoming paper.

Extracting attack is a problem that cryptanalysts have to face. So far there have been many literatures about passive attacks (i.e. steganalysis) such as [18, 19], while there is few about extracting attack which should rely on the techniques of both steganalysis and cryptanalysis. Our further work will also include the study of different kinds of extracting attacks to stegosystems.

References

1. Simmons, G. J.: The prisoners' problem and the subliminal channel. in Advances in Cryptology: Proceedings of Crypto' 83. Plenum Press (1984) 51–67
2. Petitcolas, F. A., Anderson, R. J., Kuhn, M. G.: Information hiding-a survey. Proceedings of the IEEE, Special Issue on Idenntification and Protection of Multimedia Information, vol. 87. (1999) 1062–1078
3. Anderson, R. J., Petitcolas, F. A.,: On the limits of steganography. IEEE Journal of Selected Areas in communications, vol. 16. (1998) 474–481
4. Zölner, J., Federrath, H., Klimant, H., Pfitzmann, A., Piotraschke, R., Westfeld, A., Wicke, G., Wolf, G.: Modeling the security of steganographic systems. in Information Hiding, Second International Workshop, Lecture Notes in Computer Science, vol. 1525. Springer-Verlag, Berlin Heidelberg New York (1998) 344–354
5. Cachin, C.: An information-theoretic model for steganography. in Information Hiding: Second International Workshop, Lecture Notes in Computer Science, vol. 1525. Springer-Verlag, Berlin Heidelberg New York (1998) 306–318
6. Mittelholzer, T.: An information-theoretic approach to steganography and watermarking. In: Pfitzmann A. (eds.): 3rd International Workshop. Lecture Notes in Computer Science, vol. 1768. Springer-Verlag, Berlin Heidelberg New York (2000) 1–16
7. Moskowitz, I. S., Chang, L., Newman, R. E.: Capacity is the wrong paradigm. Available: http://chacs.nrl.navy.mil/publications/CHACS/2002/2002moskowitz-capacity.pdf (2002)
8. Katzenbeisser, S., Petitcolas, F. A.: Defining security in steganographic systems. Proc. Electronic Imaging, Photonics West, SPIE, San Jose, California (2002)

9. Chandramouli, R., Memon, N. D.: Steganography capacity: A steganalysis perspective. Proc. SPIE Security and Watermarking of Multimedia Contents. Available: http://www.ece.stevens-tech.edu/˜mouli/res.html (2003)
10. Moulin, P., O'Sullivan, J. A.: Information theoretic analysis of information hiding. IEEE Trans. on Information Theory, vol. 49. (2003) 563–593
11. Ettinger, M.: Steganalysis and game euilibria. in Information Hiding: Second International Workshop, Lecture Notes in Computer Science, vol. 1525. Springer-Verlag, Berlin Heidelberg New York (1998) 319–328
12. Chandramouli, R.: A mathematical framework for active steganalysis. In ACM Multimedia Systems Journal, Special Issue on Multimedia Watermarking. Available: http://www.ece.stevens-tech.edu/ mouli/res.html (2003)
13. Fridrich, J., Goljan, M., Soukal, D.: Searching for the stego key. Proc. EI SPIE San Jose, CA, Vol. 5306. (2004)
14. Shannon, C. E.: Communication theory of secrecy system. Bell Syst. Tech. J., vol. 28. (1949) 656–715
15. Cover, T. M., Thomas, J. A.: Elements of Information Theory. Wiley Series in Telecommunications, John Wiley & Sons Inc., 2nd edition. (1991)
16. Vadhan, S. P.: A study of statistical zero-knowledge proofs. Ph.D. dissertation. Department of Mathematics, Cambridge University (1999)
17. Somekh-Baruch, A., Merhav, N.: On the Capacity Game of Public Watermarking Systems. Available: http://tiger.technion.ac.il/˜merhav/papers/p71.ps (2002)
18. Westfeld, A., Pfitzmann, A.: Attacks on Steganographic Systems. In: Pfitzmann A. (eds.): 3rd International Workshop. Lecture Notes in Computer Science, vol. 1768. Springer-Verlag, Berlin Heidelberg New York (2000) 61–75
19. Fridrich, J. Goljan, M., Hogea, D.: Steganalysis of JPEG Images: Breaking the F5 Algorithm. 5th Information Hiding Workshop, Lecture Notes in Computer Science, vol. 2578. Springer-Verlag, Berlin Heidelberg New York (2003) 310–323
20. Moulin, P., Wang, Y.: New results on steganographic capacity. Proceeding of CISS 2004. University of Princeton, Princeton, New Jersey (2004)