INFORMATION TRANSMISSION AND STEGANOGRAPHY

Ingemar J. Cox[*]*, Ton Kalker* [†]*, Georg Pakura and Mathias Scheel*[‡]

ABSTRACT

Recently there has been strong interest in developing models of steganography based on information theory. Previous work has considered under what conditions the security of the stegosystem can be guaranteed and the number of bits that can then be embedded in a cover Work. This work implicitly assumes that the hidden message is uncorrelated with the cover Work, the latter simply being used to conceal the hidden message. Here, we consider the case in which the cover Work is chosen such that it is correlated with the covert message. In this situation, the number of bits needed to encode the hidden message can be considerably reduced. We discuss the information that can then be transmitted and show that it is substantially greater than simply the number of embedded bits. We also note that the security of the system as defined by Cachin need not be compromised. However, the Shannon security may be compromised, but it remains unclear to what extent. Experimental results are presented that demonstrate the fundamental concepts.

1. INTRODUCTION

The history of steganography can be traced back thousands of years, examples of which are described in [3]. Steganography seeks to provide a covert communication channel between two parties. In [1] the problem is framed as one in which two prisoners, Alice and Bob, are permitted to communicate between one another, while under the surveillance of a Warden. The Warden will prevent communication between Alice and Bob if any communications between them is determined to contain a hidden message.

In steganography, we have a hidden message that Alice wishes to transmit to Bob. This message is hidden in a cover Work, which might be an image, video, audio or text message, for example. The combination of cover Work and hidden message is refered to as the stegowork, or more specifically, the stegotext, stegoimage, etc depending on the particular instance of the cover Work. It is assumed that Alice and Bob share a secret key and a public function that takes as input the key and the stegowork and outputs the secret message. Alice sends Bob a transmitted Work which may either be a cover Work, i.e. there is no hidden message, or a stegoWork, i.e. there is a hidden message. The Warden, Eve, is free to examine all transmitted Works between Alice and Bob and must decide whether such transmissions include a hidden message.

Steganography differs from cryptography. Cryptography attempts to prevent a message between Alice and Bob being decoded by a third party who has intercepted the message. That is, in the latter case, it is known that Alice and Bob are conducting a private communication, but interception of the encrypted message hopefully does not allow the adversary to interpret the message. However, cryptography does not prevent the adversary from disrupting or destroying the communication channel between Alice and Bob, thereby preventing any further communication. Steganography attempts to hide the very fact that Alice and Bob are conducting a private communication. An adversary may know that the two parties are communicating, but this communication appears to the Warden to be a benign communication with no covert subtext.

Steganography differs from watermarking. In steganography, the cover Work is not considered to be of value to the two communicators, Alice and Bob. Thus, it is perfectly acceptable for example, for an image of a person in a grey suit to be altered to an image of a person in a blue suit, provided, of course, that such alteration does not raise the suspicion of the Warden. In contrast, in digital watermarking, the cover Work is considered to be valuable to at least one of the communicators and the fidelity of the cover Work must be preserved.

The adversary in a stegosystem can be assumed to be either active or passive. In the active case, the Warden is free to alter the Work transmitted by Alice, before delivering it to Bob. That is, the Warden is free to attempt to remove any possible hidden message from the stegowork before passing it on the Bob. In the passive case, the Warden is not permitted to alter the transmitted Work. Rather, the adversary must decide whether the transmitted Work contains a hidden message and if so, is then free to prevent receipt of the transmission to Bob. For the purposes of this paper, we assume a passive adversary.

Shannon [4] first considered secrecy systems from the viewpoint of information theory. Shannon identified three types of secret communication which he described as (i)

[*]University College London, Torrington Place, London
[†]HP Labs, Palo Alto, CA
[‡]University of Rostock, 18059 Rostock, Germany

"*concealment systems, including such methods as invisible ink, concealing a message in an innocent text, or in a fake covering cryptogram, or other methods in which the existence of the message is concealed from the enemy*", (ii) privacy sytems and (iii) cryptographic systems. On concealment systems, i.e. steganography, Shannon stated that such "*systems are primarily a psychological problem*" and did not consider them further.

Anderson and Petitcolas [5, 3] revisited the question of steganography from an information theoretic viewpoint, suggesting that indeed information theory could also be used to describe such systems. They considered the ideal scenario of perfect source encoding, say for music. In this case, a source decoder would decompress any random bit string into an acceptable musically piece. Thus, Alice could take an encrypted hidden message and then pass this message through the ideal source *decoder* to produce a stegotext that would appear to the Warden, Eve, as an acceptable cover Work. Eve would not be able to determine that such a stegotext contained a hidden message. On receipt, Bob would input the stegotext into the source *encoder* to again produce the encrypted hidden message which is then decrypted. This thought experiment reveals that under certain circumstances steganography may be impossible to detect.

In fact, we do not even need perfect compression in order to ensure that steganography is undetectable. Very low bit rate steganography is indeed impossible to detect (at least from a statistical perspective). For example, if Alice and Bob share a secret key, then a public hash function can be used to map a string of bits plus the key into an n-bit hash. Then for $n < 20$, say, it is perfectly feasible for Alice to search through a collection of approximately 1 million images and identify the image which hashes to the desired n-bit string. Since the image has not been altered in any way, it is not possible for the Warden to determine that the communication of the image contains a covert message.

Unfortunately, for $n > 20$, the size of the database quickly becomes prohibitive. For example, for $n = 40$, we must search a database of approximately one trillion items. The continuing increase in storage and computational power does not significantly help - to send an extra 20 bits, e.g. $n = 60$, requires a million-fold increase in capacity. Thus, to send messages of greater length will require sending multiple partial messages. For example, for $n = 20$ and a message size of 100k bits, Alice must send over 5,000 stegoWorks to Bob. This may well raise the suspicion of the Warden. Ideally, Alice would like to hide as much information as possible in a coverWork while maintaining the security of the system. Thus, a key question in steganography is how many bits can be safely embedded in a cover Work without raising the risk of detection by the Warden above some small probability.

Cachin [6] examined this problem from an information theoretic view point. Cachin assumes a passive adversary whose decision is based on a statistical hypothesis test as to whether the stegowork is drawn from the distribution of allowable benign communications, i.e. coverWorks, or otherwise. Under these conditions, Cachin defines conditions for both a *perfectly secure* and an ϵ-*secure* stegosystem. Section 2 summarizes this contribution. Cachin defines the conditions under which the probability is either zero or negligible that an adversary will detect the existence of a covert communication. However, this work does not indicate how many bits can be embedded. Sallee [7] provided an answer to this question, the results of which are discussed in Section 2.1.

One might now conclude that the question of how much information can be transmitted between Alice and Bob is answered. And if it is assumed that the hidden message and the cover Work are statistically independent, then this is indeed the case. However, what if the covert message and the cover Work are correlated? That is, the mutual information between the hidden message and the cover Work is non-zero? Clearly, if the cover Work has non-zero mutual information with the message, then there is leakage of information to the Warden. However, we argue that in many circumstances, this leakage may not be sufficient for the Warden to learn anything significant. For example, consider the case where Alice wishes to transmit a covert image to Bob. Given the covert image, Alice selects a cover Work (image) from a database such that the mutual information between the cover Work and covert image is non-zero. Thus, the number of bits needed to encode the covert image will be (much) less than would otherwise be needed. The Warden may learn that the class of covert messages is an image. However, it may not be possible to determine what that message is. And more importantly, if the number of bits needed to encode the covert image is less than the upper bound provided by Sallee, then the Warden will not be able to determine that a covert message is even present. However, the amount of information received by Bob is much greater that the number of bits needed to encode the hidden message! This is because the cover Work provides more than a cover. Rather, it defines a probability distribution which permits a very efficient source coding of the hidden message.

In this paper, we consider the situation in which the hidden message and the cover Work are correlated. In this case, the number of bit needed to communicate the hidden message may be much less than for the case where the cover Work is statistically independent. Source coding of correlated information sources has been studied [8] and these results are discussed in Section 3.1.

Section 3 describes our proposed steganography system and provides an estimate of the information that can be communicated between the two communicating parties. It is different from Cachin's and Sallee's results in that the

question we ask is not how many bits can safely be embedded in a cover Work, but rather, how much information can the receiving party learn. Section 4 then illustrates the steganographic principle with a demonstration of embedding a covert image within a cover image. Finally, Section 5 concludes with a summary and directions for future work.

2. AN INFORMATION-THEORETIC MODEL OF STEGANOGRAPHY

In order to discuss information theoretic models of steganography, we first provide a brief summary of some basic results in information theory.

Consider an ensemble $\mathcal{X} = (x, A_{\mathcal{X}}, P_{\mathcal{X}})$, where the outcome x is the value of a random variable. The values of x are drawn from an alphabet $A_{\mathcal{X}} = (a_1, a_2, \cdots a_l)$ with probabilities $P_{\mathcal{X}} = (P_1, P_2, \cdots P_l)$ such that

$$P(x = a_i) = P_i, P_i \geq 0 \text{ and } \sum_{a_i \in A_{\mathcal{X}}} P(x = a_i) = 1 \quad (1)$$

The Shannon information content of an outcome x is

$$h(x) = log_2 \frac{1}{P(x)} \quad (2)$$

and the entropy of the ensemble $\mathcal{X}$ is

$$H(\mathcal{X}) = \sum_{x \in A_{\mathcal{X}}} P(x) log_2 \frac{1}{P(x)} \quad (3)$$

The entropy provides a lower bound on the number of bits that are needed to encode x, for infintely long, independent, identically distributed (iid) sequences. The entropy $H(\mathcal{X}) \geq 0$ and is only zero if $P(x_i) = 1$, i.e. the signal is entirely deterministic.

The joint entropy between $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$H(\mathcal{X}, \mathcal{Y}) = \sum_{x,y \in A_{\mathcal{X}} A_{\mathcal{Y}}} P(x, y) log \frac{1}{P(x, y)} \quad (4)$$

where $P(x, y)$ is the joint probability of the outcomes x and y occuring.

$$H(\mathcal{X}, \mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) \text{ iff } P(x, y) = P(x)P(y) \quad (5)$$

i.e. x and y are independent of one another.

The conditional entropy of $\mathcal{X}$ given $y = b_k$, is the entropy of the probabilty distribution $P(x|y = b_k)$ and is given by

$$H(\mathcal{X}|y = b_k) = \sum_{x \in \mathcal{A}_{\mathcal{X}}} P(x|y = b_k) log \frac{1}{P(x|y = b_k)} \quad (6)$$

The conditional entropy of $\mathcal{X}$ given $\mathcal{Y}$ is the average over y of the conditional entropy of $\mathcal{X}$ given y, i.e.

$$H(\mathcal{X}|\mathcal{Y}) = \sum_{x,y \in \mathcal{A}_{\mathcal{X}}, \mathcal{A}_{\mathcal{Y}}} P(x, y) log \frac{1}{P(x|y)} \quad (7)$$

The conditional entropy measures the uncertainty in x given knowledge of y. Thus, to code x, given y, we only need $H(\mathcal{X}|\mathcal{Y})$ bits rather than $H(\mathcal{X})$.

A related measure is the mutual information between $\mathcal{X}$ and $\mathcal{Y}$ and is given by

$$I(\mathcal{X}; \mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X}|\mathcal{Y}) \quad (8)$$

The mutual information is always greater than or equal to zero and measures the average reduction in uncertainty of x given y.

The relative entropy or Kullback-Leibler divergence between two distributions $P(x)$ and $Q(x)$ that are defined over the same alphabet $A_{\mathcal{X}}$ is

$$D_{KL}(P \parallel Q) = \sum_x P(x) log \frac{P(x)}{Q(x)} \quad (9)$$

The relative entropy can be thought of as the difference between Huffman coding a source with pdf P using a table determined by P and an alternative Q (suboptimal choice of codeword lengths). Note that the relative entropy, $D_{KL}(P \parallel Q) \geq 0$ with equality iff $P = Q$.

2.1. Steganography, Steganalysis and Information Theory

In a stegosystem, Alice sends Bob an innocent looking message, the cover Work, inside of which may be hidden a secret message. Communication is over a public channel that allows the adversary, Eve, to inspect the message.

Alice may send either a cover Work with no hidden message or a cover Work with a hidden message. Eve must decide whether Alice and Bob are communicating covertly.

Let c denote the cover Work, which is drawn from a distribution, $P_{\mathcal{C}}$, that is known to Eve. Let s denote the stegotext, and $P_{\mathcal{S}}$ its distribution. If Eve's decision is based on comparing the known distribution of the cover Works, $P_{\mathcal{C}}$, with the suspected stegotext, then clearly if

$$D_{KL}(P_{\mathcal{C}} \parallel P_{\mathcal{S}}) = 0 \quad (10)$$

then $P_{\mathcal{C}} = P_{\mathcal{S}}$ and Cachin defines this as *perfectly secure*, i.e. it is impossible for Eve to distinguish between cover Works that contain or do not contain a hidden message.

If $D_{KL}(P_{\mathcal{C}} \parallel P_{\mathcal{S}}) \leq \epsilon$, then the system is said to be ϵ-secure.

Cachin analysed Eve's detection performance using the theory of hypothesis testing [9]. Eve must decide between the two hypotheses, H_0, representing the hypothesis that the transmission does not contain a hidden message and H_1, representing the hypothesis that the transmission does contain a hidden message. Given the observation space, C, there are two probability distributions, P_0 and P_1, such that if H_0 is true then the observed message, C, was generated

according to P_0. Conversely, if H_1 is true, then C was generated from the distribution P_1.

Eve can make two forms of error. First, accepting H_1 when H_0 was true, often referred to as a false positive, and second, accepting H_0 when H_1 is true, often referred to as a false negative. Let α and β denote the probabilities of type 1 and type 2 errors, respectively. The binary relative entropy, $d(\alpha, \beta)$, is given by

$$d(\alpha, \beta) = \alpha log \frac{\alpha}{(1-\beta)} + (1-\alpha) log \frac{(1-\alpha)}{\beta} \quad (11)$$

and

$$d(\alpha, \beta) \leq D(P_0 \parallel P_1) \quad (12)$$

This inequality can be used to determine a lower bound on the probability of type 2 errors, β, given a desired upper bound on the probability of a type 1 error, α. In particular, Cachin shows that if the probability of a type 1 error is $\alpha = 0$, i.e. Eve is not permitted to accuse Alice of transmitting a covert message when in fact she has not, then the probability of a type 2 error, β, i.e. of missing a covert communication is

$$\beta \geq 2^{-\epsilon} \quad (13)$$

That is, the probability of *not* detecting a covert communication is very high.

Sallee [7] extended the work of Cachin to ask what is the maximum message length that can be securely hidden.

Given an instance of cover text, c, drawn from a distribution, $P_{\mathcal{C}}$, Sallee seperates it into two distinct parts, c_a, which remains unchanged after embedding, and c_b, which is replaced by c_b', to encode the hidden message. For example, c_b, may be the least significant bit of each pixel and c_a the remaining higher order bits.

These two parts are assumed to be drawn from two distributions, $\mathcal{C}_a$ and $\mathcal{C}_b$. Given the distribution of $P_{\mathcal{C}}$, or a model thereof, we can estimate the conditional distribution, $P_{\mathcal{C}_b|\mathcal{C}_a}(\mathcal{C}_b|\mathcal{C}_a = c_a)$. Then, if $\mathcal{C}_b'$ is chosen to obey this conditional distribution, then the resulting stegotext, $\mathcal{C}' = (c_a, c_b')$ will have the same distribution, $P_{\mathcal{C}}$, as the cover work.

Sallee suggested using an arithmetic entropy encode/decoder [10]to accomplish this. Arithmetic coding is a method for very efficient compression of strings given a model of their distribution. However, if a random bit string (read hidden message) is fed into an arithmetic *decoder*, the output bit sequence will have the same distribution as the model distribution. This is a practical means for generating the distributions required by Cachin to ensure perfect security. Note the similarity between this and the ideas of Anderson and Petitcolas [3] regarding ideal compresison that were described earlier.

Sallee's method has a capacity equal to the entropy of the conditional probability distribution, $P_{\mathcal{C}_b|\mathcal{C}_a}$

$$H(\mathcal{C}_b|\mathcal{C}_a = c_a) = -\sum_{c_b} P_{\mathcal{C}_b|\mathcal{C}_a}(c_b|c_a) log P_{\mathcal{C}_b|\mathcal{C}_a}(c_b|c_a) \quad (14)$$

Essentially, $\mathcal{C}_b$ is an open communications channel without any restriction. Note that this capacity is independent of the distribution of the message to be hidden.

3. INFORMATION TRANSMISSION WITH CORRELATION BETWEEN COVER AND COVERT WORKS

Consider a message m drawn from a distribution, $P_{\mathcal{M}}$ and a cover Work c drawn from a distribution, $P_{\mathcal{C}}$. If m is independent of c, then the minimum number of bits needed to encode the message is the message's entropy, $H(m)$. However, if m and c are correlated, then the number of bits needed to encode m given c is the conditional entropy, $H(m|c)$. The conditional entropy, $H(m|c)$, may be much less than the entropy of the message, $H(m)$.

What if $m = c$, or more usefully, m is a deterministic function of the cover Work, c? Then, the conditional entropy is zero and there is no need to embed a secret message. At first sight, this would not appear to offer any form of covert channel. However, if Alice and Bob share a secret key, then even if the deterministic function is known publicly, this offers a perfectly secure channel, since the distribution of c is unchanged. This form of steganography was discussed in the introduction using a one-way hash function, though any receiver function will suffice.

The key question then is how much information can Alice transmit to Bob without being detected by Eve. Assuming that we split the covertext into two parts, then from [7], we know that given a covertext, c, the maximum size of the hidden message is given by Equation 14. Thus, if the hidden message is uncorrelated with the cover Work, then the maximium information transmitted is simply this number of bits. However, the information transmitted to Bob includes both the message *and* the cover Work. Traditionally, the cover Work has been ignored. It is simply a means by which to conceal the hidden message. However, this need not be the case.

Given a message m and cover Work, c with conditional entropy, $H(m|c)$, then we only need to encode $H(m|c)$ bits of information in c in order to encode m. For explanatory purposes, let's assume that the encoding procedure splits the covertext into two parts. Then, the information received by Bob is

$$H(c, m) = H(c_a) + H(c_b) = H(c_a) + H(m|c) = H(c) \quad (15)$$

which is potentially much greater than simply $H(m)$.

Thus, given a hidden message, we choose a cover Work from a set of cover Works, such that the correlation between the two permits a very efficient source coding of the hidden message.[1] We believe that the search for a correlated coverWork is significantly easier than finding a coverWork that hashes to a desired n-bit value. In the latter case, for large n, this is almost impossible. However, most images, for example, exhibit correlation with one another.

If the cover Work is highly correlated with the message, m, then the number of embedded bits needed will be very low. What does this imply regarding the secrecy of the covert channel?

First, Eve cannot distinguish between a cover text with no hidden message and a stegotext provided we ensure that the number of embedded bits is less than that given by Equation 14. Thus, provided this condition is met, then there is no reduction in security as defined by Chachin.

Shannon [4] defines *perfect security* as "*a system that after a cryptogram is intercepted by the enemy, the a posteriori probabilities of this cryptogram representing various messages be identically the same as the a priori probabilities of the same messages before the interception*". Thus, a system that exploits the mutual information between the hidden message and the cover Work would not appear to be perfectly secure, as defined by Shannon.

From the sender's perspective, the cover Work defines a probability distribution that permits a very efficient source coding of the hidden message. Without this, Alice would need at least $H(m)$ bits to encode the message, m. With the cover Work, Alice only needs $H(m|c)$, bits. A judicious choice of the cover Work, c, will then permit a very significant reduction in the number of bits that need to be embedded. These bits can then be encrypted using the secret key shared by Alice and Bob. The encryption does not increase the number of bits, but prevents the Warden from decoding the message, assuming it is detected. This pseudo-random encrypted bit sequence is then embedded into the cover Work. This can be accomplished using Sallee's method.

If Eve suspects that Alice is exploiting the conditional entropy between the cover Work and message, then what can Eve learn from examining the cover Work? Certainly, upon interception of the stegowork, the adversary, Eve, has learned something about the hidden message. For example, if the cover Work is an image, the adversary may confidently conclude that the hidden message is also an image. However, our earlier example demonstrated that even if the conditional entropy is zero, Eve may still not be able to learn anything about the message, since she does not have knowledge of the key shared between Alice and Bob. In fact, the cover Work informs the Warden of the probability distribution used by Alice to perform the source encoding. However, this is not sufficient to decode the message.

We do not claim that steganography based on coding that exploits the conditional entropy between the hidden message and the cover Work is perfectly secure in the Shannon sense. However, it can certainly be perfectly secure of ϵ-secure in the Chachin sense.

3.1. Encoding of correlated sources

The encoding of correlated sources has been well studied. Interestingly, Slepian and Wolf [8] showed that efficient noiseless coding of two correlated sources $\mathcal{X}$ and $\mathcal{Y}$ could be achieved even if the two source encoders do *not* have access to the other signal, provided both signals are available to the decoder.

More recently, Pradhan and Ramchandran [16, 17] extended these results to provide a constructive procedure for distributed source coding based on syndrome codes.

Together with Chou, they also recognized the duality between distributed source coding and data hiding [12]. However, while this paper demonstrated how to embed a hidden message in a cover Work using syndrome coding, it did not consider exploiting the mutual information between the cover Work and the hidden message. Rather, it can be considered an efficient implementation of results due to Costa [18] in which it was shown that the channel capacity of system with two noise sources, the first of which is entirely known to the transmitter, but neither of which is known to the receiver, is equivalent to the channel capacity of a system in which the known first noise source is absent. From a data hiding perspective, the first noise source represents the cover Work while the second noise source represents the distortion in the stegoWork prior to its receipt. This forms the foundation for considerable work on modeling watermarking as communication with side information [11, 12, 13].

Chou *et al* [12] also observed the similarity between distributive source coding, digital watermarking and that of writing on defective memories [19]. More recently, Fridrich *et al* [20] have applied these ideas to steganography. Their "wet paper" codes assume a cover Work consisting of n samples, k of which are "dry" and can be modified while the remaining $(n-k)$ bits are "wet" and must not be altered. They show that it is possible to embed k-bits of information into a cover Work without the decoder knowing which of the k samples have been modified. Once again, correlation between the hidden message and the cover Work is not considered and the capacity of the system is considered to be

[1] There is a similarity between this and digital watermarking, where, given a cover Work, it is common to choose a watermark from a set of watermarks such that the watermark is easy to embed. Such techniques are based on the modeling watermarking as communication with side information [11, 12, 13] and the watermarks are often referred to as dirty paper codes [14, 15]. However, digital watermarking does not use the correlation between the message and the cover work to reduce the number of bits needed to encode the message. Rather, the purpose is to reduce or eliminate the "noise" due to the cover Work and thereby improve the robustness and/or fidelity.

k-bits.

4. EXPERIMENTAL RESULTS

To demonstrate the concepts discussed in the previous section, we modified a steganographic method due to Chan *et al* [21]. They describe a procedure for embedding a covert image within a cover image. While their paper does not discuss relative entropy, relative entropy is exploited in order to reduce the number of bits needed to encode the covert image. In this example, we did not search for a cover image with high correlation with the hidden image, but rather, relied on the correlation that is present between 8×8 blocks across all images. It should be noted that this example is for illustrative purposes only and does not represent the most efficient means to implement our proposal.

The embedding procedure consists of:

1. Partition the cover image and covert image into 8×8 blocks, denoted c_i and m_j respectively
2. For all i and j, compute the error-matrix, $EM_{i,j}$ and the normalized error-matrix, $NEM_{i,j}$ defined as:

$$EM_{i,j} = m_i - c_j \tag{16}$$

and

$$NEM_{i,j} = EM_{i,j} - min(EM_{i,j}) \tag{17}$$

3. The range of errors, is refered to as the distance degree (DD) and is defined as

$$DD_{i,j} = max(EM_{i,j}) - min(EM_{i,j}) \tag{18}$$

where the min and max operations are over the 8×8 elements of the blocks.

4. For each hidden image block m_j, find the cover image block c_i such that $DD_{i,j}$ is a minimum. The location of the cover image block is referred to as the reference-block-index $RBI(j) = i$.
5. Given $DD_{RBI(j),j}$, the quantization error matrix is selected according to Table 1
6. Quantize the $NEM_{RBI(j),j}$
7. Embed the extra information of (i) the referenced block number, (ii) the quantization error matrix and (iii) the minimum element in the error matrix.
8. Embed this extra information in the LSB of the DCT coefficients, according to the method of [7].

The extraction procedure follows:

1. Extract the RBI

DD_{c_i,m_j}	QEM_{c_i,m_j}
(3 - 4)	2
(5 - 6)	3
(7 - 8)	4
(9 - 11)	5
(12)	6
(13),(26 - 27), (52 - 55), (104 - 111)	$\lfloor \frac{NEM}{7} \rfloor \times 7 + 3$
(14 - 15),(28 - 31), (56 - 63), (112 - 127)	$\lfloor \frac{NEM}{8} \rfloor \times 8 + 3$
(16 - 17),(32 - 35), (64 - 71), (128 - 143)	$\lfloor \frac{NEM}{9} \rfloor \times 9 + 4$
(18 - 19),(36 - 39), (72 - 79), (144 - 159)	$\lfloor \frac{NEM}{10} \rfloor \times 10 + 4$
(20 - 21),(40 - 43), (80 - 87), (160 - 175)	$\lfloor \frac{NEM}{11} \rfloor \times 11 + 5$
(22 - 23),(44 - 47), (88 - 95), (176 - 191)	$\lfloor \frac{NEM}{12} \rfloor \times 12 + 5$
(24 - 25),(48 - 51), (96 - 103), (192 - 207)	$\lfloor \frac{NEM}{13} \rfloor \times 13 + 6$
(208 - 223)	$\lfloor \frac{NEM}{14} \rfloor \times 14 + 6$
(224 - 239)	$\lfloor \frac{NEM}{15} \rfloor \times 15 + 7$
(240 - 255)	$\lfloor \frac{NEM}{16} \rfloor \times 16 + 7$

Table 1. Quantization error matrix.

Fig. 1. Cover image

Fig. 2. Image to be hidden

2. extract the QEM
3. extract the minimum element
4. reconstruct the secret image as

$$S_j = H_i' + QEM + min(EM_{i,j}) \qquad (19)$$

Figure 1 shows an image used as a cover image. Figure 2 shows the image that is to be hidden in Figure 1.

Using the method outlined above, the number of bits needed to encode the hidden image was 222144 or 0.8474 bits per pixel. This approach uses lossy compression and the relative entropy of image 1 given image 2, may be higher. Nevertheless, this example serves to illustrate the considerable reduction in the number of bits that must be embedded when the cover Work is correlated with the hidden message. Independent coding of the the hidden image would have required 8 bits per pixel.[2]

[2]For this example, a similar or smaller number of bits per pixel would be possible by simple lossy compression of the hidden image using say JPEG compression. However, this need not be the case and we emphasise that the experimental results described here are purely for illustrative purposes.

The number of bits that can safely be inserted in the cover Work according Equation 14 is 223110 or 0.8511 bits per pixel. Thus, the hidden message cane be embedded without risk of detection from a Warden.

Note that the resulting stegoImage has a 50.57dB signal-to-noise ratio compared with the original cover Work. Similarly, the recovered hidden image has a 38.93 dB SNR compared with the original hidden image, prior to embedding.

5. CONCLUSION

Previous work on modeling steganography using information theory has assumed that the hidden message is uncorrelated with the cover Work. In this scenario, the cover Work serves only to hide the covert message. However, it may often be the case that the sender of a stegotext, Alice, may be able to chose a cover Work that is correlated with the hidden message. In this case, the cover Work not only serves to hide the covert message, but also defines a probability distribution which permits a very efficient source coding of the message.

It is well known that if a message, m has entropy $H(m)$, this entropy defines a lower bound on the number of bits needed to reliably code the message. However, given a coverWork, c, that is correlated with the message, then it is also well-known that the message m requires only $H(m|c)$ additional bits, where $H(m|c)$ is the conditional entropy between the message and coverWork. This may be very much less than $H(m)$.

The reduction in the number of bits needed to encode the message is very beneficial, especially in ensuring security, as defined by Cachin. However, more importantly, we point out that the information received by Bob is much more than simply the number of encoded bits. Rather, Bob receives information that it equivalent to the entropy of the coverWork. This is much higher than previously thought possible.

We discussed the security issues related to steganography using mutual information between cover work and covert message. It is clear that from the perspective of detectability, we can still ensure that the system is perfectly secure or ϵ-secure as defined by Cachin. In fact, given that we need far fewer bits to encode the secret message, it should be easier to ensure such security. However, from the perpsective of perfect secrecy as defined by Shannon, the adversary learns a probability distribution defined by the cover Work which the hidden message is correlated with. Nevertheless, it is unclear how useful this knowledge is to the Warden.

We provided experimental results that are intended to illustrate that basic concepts of the method. Specifically,

we discussed hiding an image within another cover image. This example was purely illustrative and more sophisticated techniques based on the approach proposed here are the subject of future work. We also note that this appraoch is applicable to many different kinds of cover Works and covert messages, including text, audio, video, computer graphics, maps, and electronic schematics.

There is clearly a deep connection between coding of correlated sources, distributed source coding, digital watermarking and steganography. We (and others) have identified a number of these connections but a rigorous mathematical model needs to be developed further. The basic problem can be described as given a message m that we wish to hide, first find a covertext, c, that is correlated with the message. We than want to jointly encode m and c into a stegotext, s such that s has the same source model as c and m should be recoverable from s up to some distortion (given some secret shared between Alice and Bob). An optimum solution to this problem remains a goal of future work.

Acknowledgement

The author thanks Matt L. Miller of NEC and Nasir Memon of Polytechnic University for valuable discussions of this work. Also Professor L. M. Cheng of the City University of Hong Kong for assistance with the experiments. Ingemar Cox is currently BT Chair of Telecommunications and thanks British Telecom for their support. This research was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-03-1-3A46. The U.S. Government is authorized to reproduce and distribute reprints for Government purpose notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

6. REFERENCES

[1] G. J. Simmons, "The prisoner's problem and the subliminal channel," in *Proc. of CRYPTO'83*, 1984, pp. 51–67.

[2] Herodotus, *The Histories*, Penguin Books, London, 1996, translated by Aubrey de Sélincourt.

[3] R. J. Anderson and F. A. P. Petitcolas, "On the limits of steganography," *IEEE J. of Selected Areas in Communications*, vol. 16, no. 4, pp. 474–481, 1998.

[4] C. E. Shannon, "communications theory of secrecy systems," *Bell System technical Journal*, vol. 28, pp. 656–715, 1954.

[5] R. J. Anderson, "Stretching the limits of steganography," in *Proc. on the First Workshop on Information Hiding*, 1996, vol. 1174 of *Springer Lecture Notes in Computer Science*, pp. 39–48.

[6] C. Cachin, "An information-theoretic model for steganography," in *Proc. on the Second Workshop on Information Hiding*, 1998, vol. 1525 of *Springer Lecture Notes in Computer Science*, pp. 306–318.

[7] P. Sallee, "Model based steganography," in *Int. Workshop on Digital Watermarking*, T. Kalker, I. J. Cox, and Y. M. Ro, Eds., 2004, vol. 2939 of *Springer Lecture Notes in Computer Science*, pp. 154–167.

[8] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Information Theory*, vol. IT-19, no. 4, pp. 471–480, 1973.

[9] R. E. Blahut, *Principles and Practice of Information theory*, Addison-Wesley, 1987.

[10] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003.

[11] I. J. Cox, M. L. Miller, and A. L. McKellips, "Watermarking as communications with side information," *Proc. IEEE*, vol. 87, pp. 1127–1141, July 1999.

[12] J. Chou, S. S. Pradhan, and K. Ramchandran, "On the duality between distributed source coding and data hiding," in *Proc. Thirty-third Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 1999, vol. 2, pp. 1503–1507.

[13] B. Chen and G. W. Wornell, "An information-theoretic approach to the design of robust digital watermarking systems," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, Arizona, USA, March 1999, vol. 4, pp. 2061–2064.

[14] M. L. Miller, G. J. Doërr, and I. J. Cox, "Dirty-paper trellis codes for watermarking," in *Proc. IEEE Int. Conf. on Image Processing*, Rochester, New York, USA, Sept. 2002, vol. 2, pp. 129–132.

[15] M. L. Miller, G. J. Doërr, and I. J. Cox, "Applying informed coding and embedding to design a robust high-capacity watermark," *IEEE Trans. Image Processing*, vol. 13, pp. 792–807, June 2004.

[16] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," in *Proc. of the 1999 IEEE Data Compression Conference*, 1999.

[17] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," 2003, vol. 49, pp. 626–643.

[18] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[19] C. Heegard and A. El Gamal, "On the capacity of a computer memory with defects," *IEEE Trans. on Information Theory*, vol. 29, 1983.

[20] J. Fridrich, M. Goljan, and D. Soukal, "Perturbed quantization steganography with wet paper codes," in *ACM Multimedia Workshop*, 2004.

[21] C-K Chan, L. M. Cheng, K-C Leung, and S-L Li, "Image hiding based on block difference," in *8th Int. Conf. on Control, Automation, Robotics and Vision*, 2004.