Contemporary Engineering Sciences, Vol. 1, 2008, no. 1, 27 - 39

Steganographic Schemes for Noisy Communication Channels[1]

Guglielmo Morgari, Maria Spicciola, Silvia Deantonio

TELSY Elettronica, Torino, Italy
guglielmo.morgari@telsy.it

Michele Elia

Politecnico di Torino, Torino, Italy
michele.elia@polito.it

Abstract

Very noisy channels are particularly appealing supports for steganographic communications. The cost to pay is a low transmission rate of concealed messages; the advantage is an almost undetectable oblivious transfer. Since very noisy channels need error correcting codes, the code redundancy is utilized to insert message bits masked in the form of artificial channel errors. Two different insertion modes are proposed along with suitable criteria for distinguishing genuine from artificial errors at the receiver side. The resulting steganographic channels are still binary symmetric channels whose bit error probability depends on the error probability of the transmission channel, the decoding strategy of the error-correcting code, and the method for separating genuine and artificial errors. Two simple yet concrete examples illustrate both the proposed method and the complexity of computing the bit error probability affecting the steganographic channel.

Keywords: steganography, BCH codes, noisy channel

[1]A preliminary version of this paper was presented at the "International Conference on Advances in Interdisciplinary Statistics and Combinatorics" - October 12-14, 2007 - UNCG, Greensboro, USA.

1 - Introduction

The goal of steganography is to conceal the very existence of the messages. A typical steganographic artifice is to hide information in innocent cover messages by exploiting their high semantic redundancy, as occurs for example in voice messages or pictures. However, information-hiding techniques operating at a lower-level layer in a transmission chain have also been proposed. Following this second approach, we describe a steganographic scheme that exploits the redundancy of the error-correcting codes necessarily used over noisy channels, for example in the ubiquitous cell phone or wireless data access communications. The stratagem is to insert the steganographic bits as artificial channel errors on an honest communication channel, and to use a suitable criterion to discriminate between genuine and artificial errors, thus recovering the hidden information. The resulting steganographic or oblivious channel is a Binary Symmetric Channel (BSC), characterized by a bit error probability p_g, which can be used to send the hidden information using standard techniques and protocols. The paper is organized as follows. Section 2 presents the main concepts, namely an introduction of the general framework, a description of the steganographic channel, or stega-channel, and a characterization of the BSC model of the stega-channel. Section 3 discusses concealment issues. Section 4 analyses two simple yet concrete implementations of stega-channels, and explicitly evaluates their performances. Lastly, in Section 5 some observations are made on the feasibility of the scheme and its potential applications.

2 - Steganographic channel models

A digital communication chain connecting a source S with a user U is composed of a binary encoder E using an (n, k, d) linear code $\mathcal{C}$ with $d = 2t + 1$, a binary symmetric channel (BSC) with bit error probability p, and a decoder D. The chain E-BSC-D is referred to as the primary channel. A decoding rule $\mathcal{D}$ for the corrupted code words received is described by a full set $\mathcal{T}$ of coset leaders ℓ_i, $1 \leq i \leq 2^{n-k}$, which identify the cosets $\ell_i + \mathcal{C}$.
A steganographic channel is created by inserting artificial errors on the BSC channel. We will consider and compare two basic modes of artificial bit error insertion; both modes introduce a single artificial error per code word:

Mode 1: the stega-bit is inserted in a fixed position within a code word $\mathbf{c} \in \mathcal{C}$ according to the following rule

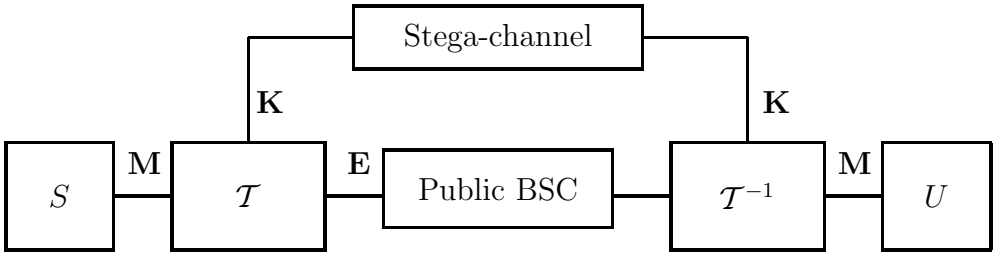

Figure 1: Communication Channel with Error-Correcting Codes and Stega-channel

a "**0**" stega-bit is inserted if there is no artificial error in the chosen position.

a "**1**" stega-bit is inserted if there is an artificial error in the chosen position.

Mode 2: the stega-bit is inserted as an artificial error in a random position within a code word $\mathbf{c} \in \mathcal{C}$ according to the following rule

a "**0**" stega-bit is inserted as an artificial error hitting a random position among those occupied by 0s in a code word $\mathbf{c} \in \mathcal{C}$.

a "**1**" stega-bit is inserted as an artificial error hitting a random position among those occupied by 1s in a code word $\mathbf{c} \in \mathcal{C}$.

In both Modes, the code $\mathcal{C}$ is used to recognize both error status and stega-bits; however, the separation of artificial from genuine errors follows different principles in different modes:

- In Mode 1 the stega-information is encoded in a known position within a code word, therefore it is easily recognized using the decoding rule $\mathcal{D}$.
- In Mode 2, the stega-information is carried by a code-word symbol artificially corrupted in a random position unknown to the stega-user, therefore the decoding rule $\mathcal{D}$ is not sufficient for identifying the artificial error and a detection criterion must be defined.

The stega-channel between sender and receiver is a binary communication channel characterized by a bit error probability p_g which depends on

1) the error correcting capabilities of the code $\mathcal{C}$,
2) the cover channel bit error probability p,
3) the decoding rule $\mathcal{D}$, and
4) the stega-bit detection criterion in Mode 2.

Let $\mathbf{c} \in \mathcal{C}$ denote a transmitted code word, and let $\mathbf{e}$ be an error pattern introduced by the BSC; the bit error probability p_g is thus defined as follows

$$p_g = \sum_{x=0,1} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{e}\in\mathbb{F}_2^n} p\{\hat{x} \neq x|\mathbf{c}, \mathbf{e}, x\} p_t(x) p(\mathbf{c}) p(\mathbf{e})$$

where

- $p_t(x)$ is the probability of sending a stega-bit x;
- $p(\mathbf{c})$ is the probability of sending a code word $\mathbf{c}$;
- $p(\mathbf{e}) = p^{w_H(\mathbf{e})}(1-p)^{n-w_H(\mathbf{e})}$ is the probability that an error pattern $\mathbf{e}$ of Hamming weight $w_H(\mathbf{e})$ occurs;
- $\hat{x}$ is the stega-bit detected.

Let $\mathcal{L}_i$ denote the detection rule of Mode i, $i = 1, 2$, that extracts the stega-bit $\hat{x}$ from a detected error pattern $\hat{\mathbf{e}}$; thus we may rewrite the expression for p_g as

$$p_g = \sum_{x=0,1} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{e}\in\mathbb{F}_2^n} p\{\mathcal{L}_i(\mathbf{c}+\mathbf{e}+\mathcal{D}(\mathbf{c}+\mathbf{e})) \neq x|\mathbf{c}, \mathbf{e}, x\} p_t(x) p(\mathbf{c}) p(\mathbf{e}) \quad .$$

Letting $\ell(\mathbf{e})$ denote the coset leader of the coset containing the error pattern $\mathbf{e}$, we have

$$p_g = \sum_{x=0,1} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{e}\in\mathbb{F}_2^n} p\{\mathcal{L}_i(\mathbf{e}+\ell(\mathbf{e})) \neq x|\mathbf{c}, \mathbf{e}, x\} p_t(x) p(\mathbf{c}) p(\mathbf{e}) \quad .$$

In particular, referring to Mode 1, the average can be computed over the transmitted code words and the equation simplifies to

$$p_g = \sum_{x=0,1} \sum_{\mathbf{e}\in\mathbb{F}_2^n} p\{\mathcal{L}_1(\mathbf{e}+\ell(\mathbf{e})) \neq x|\mathbf{e}, x\} p_t(x) p(\mathbf{e}) \quad .$$

However, exact computation of p_g is usually very difficult, as it requires the enumeration of a large number of error configurations. Thus estimations are practically unavoidable; in particular a good estimation is obtained as follows: assuming that a code word $\mathbf{c}$ is sent, two probabilities may be computed exactly:

1. $p_c(\mathbf{c})$, the probability that the stega-bit is undoubtedly received correctly, given that $\mathbf{c}$ was sent;

2. $p_e(\mathbf{c})$, the probability that the stega-bit is undoubtedly received incorrectly, given that $\mathbf{c}$ was sent.

The stega-bit error probability $p_g(\mathbf{c})$ given that $\mathbf{c}$ was sent is estimated as

$$p_g(\mathbf{c}) = p_e(\mathbf{c}) + \frac{1}{2}[1 - p_c(\mathbf{c}) - p_e(\mathbf{c})] = \frac{1}{2} - \frac{1}{2}p_c(\mathbf{c}) + \frac{1}{2}p_e(\mathbf{c})$$

since we may assume that in all cases left out of the definitions of $p_c(\mathbf{c})$ and $p_e(\mathbf{c})$ the stega-bit is incorrectly detected with probability $\frac{1}{2}$. Finally, the stega-bit error probability p_g is the average of $p_g(\mathbf{c})$ computed over all code words:

$$p_g = \frac{1}{2^k} \sum_{\mathbf{c} \in \mathcal{C}} p_g(\mathbf{c}) \quad . \tag{1}$$

It is worth noting that in Mode 1 $p_g(\mathbf{c}) = p_g$, thus no average is necessary and the computation of p_g is independent of the transmitted code word. On the contrary, in Mode 2 the computation of the averages in equation (1) can be obtained from the polynomial weight enumerator $W(x, y)$ of the code.

Remark 1. In Mode 2, a detection strategy may be devised such that the stega-channel is modelled as a binary erasure channel. In this case the performance may be greatly improved by using a code on this channel that is both capable of correcting and allowed to correct erasures and errors.
Remark 2. As a consequence of the definition of a stega-channel as a BSC, the steganographic information may be pre-processed, namely compressed, encrypted, and encoded using an error-correcting code; the resulting stream is the sequence of bits to be sent.
The required concealing capability is decisive to define the stega-channel transmission rate. It is also fundamental for defining the transmission protocol specifying the stega-channel.

3 - Concealment issues

If channel noise is negligible, no genuine errors occur and the stega-bits are easily recovered, but conversely the stega-channel is easily detected. Therefore, the secrecy of the stega-channel lies in the primary channel noise. In particular, the difficulty of detecting the stega-channel depends on the ratio $\rho = \frac{p}{a}$ between the rate p of genuine errors and the rate a of artificial errors. If ρ is large,

the existence of stega-bits is unlikely to be recognized, but conversely the transmission rate of the stega-channel is small for a given p. Thus the choice of ρ is a compromise between the achievable rate of the stega-channel and its detectability. The average number of genuine errors per code word is np, and its variance is $\sqrt{np(1-p)}$; the average number of total errors per code words is $np + a$, and the stega-bit is undetectable at level $\alpha > 0$ if

$$np + a \leq np + \alpha\sqrt{np(1-p)}$$

thus $a \leq \alpha\sqrt{np(1-p)}$.
Let R $bits/sec$ be the transmission rate of the cover channel, then the rate of the stega-channel is $\frac{Ra}{n}$ $bits/sec$.
For example, if $p = 10^{-2}$, $\alpha = 1/10$, and $n = 31$ then $a = 0.055$, that is one stega-bit may be inserted about every 20 code words. If the transmission rate is 10 $Mbits/sec$, the net rate of the stega-channel is $\frac{10000}{620}$ $kbits/sec = 16$ $kbits/sec$. However, if the level α is increased to 1, the rate a is increased to 0.5, which means that a stega-bit may be inserted every two code words. In this case the net rate of the stega-channel is 160 $kbits/sec$.

4 - Examples, Simulation, and Results

In this section we present some examples of stega-channels which are obtained considering both Modes 1 and 2 over noisy channels and using codes with different error correcting capabilities. The main scope is to illustrate the computation of the bit error rate p_g of the equivalent BSC stega-channel, and to asses the validity of the approximated expressions.

Example 1. Consider a cover channel using the repetition code $(3, 1, 3)$ which is a perfect single error-correcting code. Since the dimension of the code is small the computations may be exact. It is assumed that stega-bits 0 and 1 are equally probable, i.e. $p_t(0) = p_t(1) = \frac{1}{2}$.

Mode 1: For computational purposes it is not restrictive to assume that the stega-bit occupies the first position in every code word, and that the word $(0, 0, 0)$ is sent. We have

1. If 0 is sent, then it is incorrectly received only if an error is detected to hit the first entry of $(0, 0, 0)$, and this event occurs only when the error patterns are

$$(1, 0, 0) \quad , \quad (0, 1, 1)$$

then $p_e(0) = p(1-p)^2 + p^2(1-p) = p - p^2$, consequently the probability of correct detection is $p_c(0) = 1 - p_e(0) = 1 - p + p^2$ since no ambiguous configuration occurs.

2. If 1 is sent, then it is incorrectly received when no error is detected in the first code word position, and this event occurs only when the error patterns are

$$(1,0,0) \quad , \quad (0,1,1) \quad , \quad (0,0,1) \quad , \quad (0,1,0) \quad , \quad (1,1,0) \quad , \quad (1,0,1)$$

then $p_e(1) = 3p(1-p)^2 + 3p^2(1-p) = 3p - 3p^2$, consequently the probability of correct detection is $p_c(1) = 1 - p_e(1) = 1 - 3p + 3p^2$ since no ambiguous configuration occurs.

Summarizing, we have
$p_e = \frac{1}{2}(p_e(0) + p_e(1)) = 2p - 2p^2$ and $p_c = \frac{1}{2}(p_c(0) + p_c(1)) = 1 - 2p + 2p^2$, and the bit error probability of the BSC model of the stega-channel is $p_g = 2p - 2p^2$.

Mode 2: For computational purposes it is convenient to use a coset code

$$\{(1,0,0), \quad (0,1,1)\}$$

(that is a translate of the code) so that the all zeros and the all ones code words are excluded and transmission of a stega-bit 0 or 1 is possible without exception. Moreover we know that p_e and p_c do not depend on which stega-bit is sent, but both probabilities depend on the code word used, thus they are computed referring to a 0 stega-bit sent. Moreover, in this case some undecidable situations occur which may be seen as erasures, thus also an erasure probability p_u is computed. The standard array is

leaders	000	010	001	100
	100	110	101	000
	011	001	010	111

and the words in each row are decoded into the code word in the first position of the row.

1. If $(1,0,0)$ is the code word used and the 0 stega-bit is sent using the second entry, the eight possible situations corresponding to the genuine error patterns are reported in the following table, where column I contains the genuine error patterns, column II the received words, column

III the decoded code word, column IV the estimated decoded stega-bit, column V the estimated error position and the type of error with E indicating an erroneous stega-bit, U an erased stega-bit, and C a correct stega-bit, and column VI contains the probability of the error event

I	II	III	IV	V	VI
000	110	100	0	$2C$	$(1-p)^3$
111	001	011	1	$2E$	p^3
100	010	011	1	$3E$	$p(1-p)^2$
011	101	100	0	$3C$	$p^2(1-p)$
110	000	100	1	$1E$	$p^2(1-p)$
001	111	011	0	$1C$	$p(1-p)^2$
101	011	011	e	$0U$	$p^2(1-p)$
010	100	100	e	$0U$	$p(1-p)^2$

2. If $(0,1,1)$ is the code word used, the configurations are summarized in the following table as in the previous case

I	II	III	IV	V	VI
000	111	011	0	$1C$	$(1-p)^3$
111	000	100	1	$1E$	p^3
100	011	011	e	$0U$	$p(1-p)^2$
011	100	100	e	$0U$	$p^2(1-p)$
110	001	011	1	$2E$	$p^2(1-p)$
001	110	100	0	$2C$	$p(1-p)^2$
101	010	011	1	$3E$	$p^2(1-p)$
010	101	100	0	$3C$	$p(1-p)^2$

Summarizing, we have

$$
\begin{aligned}
p_e &= \tfrac{1}{2}(2p^3 + 3p^2(1-p) + p(1-p)^2) \\
p_c &= \tfrac{1}{2}(2(1-p)^3 + p^2(1-p) + 3p(1-p)^2) \\
p_u &= \tfrac{1}{2}(2p(1-p)^2 + 2p^2(1-p))
\end{aligned}
$$

The probability $p_g = p$ is obtained by equally splitting p_u between p_e and p_c.

Example 2. Consider a cover channel employing a BCH $(31, 16, 7)$ code correcting three errors. In this case, an exact computation is not feasible in practice, therefore we consider an approximate computation and compare the results with numerical simulations.

Mode 1: for computational purposes it is not restrictive to assume that the stega-bit is sent using the first position in a code word. In this mode the computation of p_e and p_c depends on whether the stega-bit is 0 or 1, which are supposed to be sent with the same probability 1/2. We have

1. If 0 is sent, then $p_e(0)$ is the probability that an error hits the position of the stega-bit and the total number of errors is at most three

$$p_e(0) = p(1-p)^{30} + 30p^2(1-p)^{29} + \frac{30 \cdot 29}{2}p^3(1-p)^{28} \quad .$$

   If 1 is sent, $p_e(1)$ is equal to the probability that one error hits the position of the stega-bit and at most three errors hit the other positions within the code word

$$p_e(1) = p(1-p)^{30} + 30p^2(1-p)^{29} + \frac{30 \cdot 29}{2}p^3(1-p)^{28} + \frac{30 \cdot 29 \cdot 28}{6}p^4(1-p)^{27}.$$

2. If 0 is sent, $p_c(0)$ is equal to the probability that at most three errors hit positions within the code word, excluding the stega-bit position

$$p_c(0) = (1-p)^{31} + 30p(1-p)^{30} + \frac{30 \cdot 29}{2}p^2(1-p)^{29} + \frac{30 \cdot 29 \cdot 28}{6}p^3(1-p)^{28} \quad .$$

   If 1 is sent, $p_c(1)$ is the probability that at most two errors hit positions within the code word, excluding the stega-bit position

$$p_c(1) = (1-p)^{31} + 30p(1-p)^{30} + \frac{30 \cdot 29}{2}p^2(1-p)^{29} \quad .$$

Summarizing, we have

$$\begin{aligned} p_e &= p(1-p)^{30} + 30p^2(1-p)^{29} + \tfrac{30 \cdot 29}{2}p^3(1-p)^{28} + \tfrac{30 \cdot 29 \cdot 28}{12}p^4(1-p)^{27} \\ p_c &= (1-p)^{31} + 30p(1-p)^{30} + \tfrac{30 \cdot 29}{2}p^2(1-p)^{29} + \tfrac{30 \cdot 29 \cdot 28}{12}p^3(1-p)^{28} \quad . \end{aligned}$$

Thus, the bit error probability p_g of the BSC model for Mode 1 is

$$\begin{aligned} p_g \;=\; & \frac{1}{2} - \frac{1}{2}(1-p)^{31} - \frac{29}{2}p(1-p)^{30} - \\ & \frac{405}{2}p^2(1-p)^{29} - \frac{1595}{2}p^3(1-p)^{28} + 1015p^4(1-p)^{27} \quad . \end{aligned}$$

Mode 2. In this mode we define a detection rule of the stega-bit as follows

- Use the code to recognize the positions of each error.

- Compute the number N_0 of errors affecting the zero positions and the number N_1 of errors affecting the one positions in the code word.

- The decision rule outputs the symbol identified by the subscript of the largest between N_0 and N_1.

- Ties may be resolved in two ways:

  - by a random choice

  - by inserting an erasure (for later use in the decision process); in this case p_u indicates the erasure probability.

The probabilities p_c, p_e, and p_u do not depend on the transmitted stega-bit, and are obtained by averaging the code word set, since they are conditioned on the transmitted code word $\mathbf{c}$. The computations of $p_c(\mathbf{c})$, $p_e(\mathbf{c})$, and $p_u(\mathbf{c})$ may be done referring to a finite (relatively small) number of genuine error configurations which are summarized in the following table (a_r indicate the artificial error position) under the assumption that a 0 stega-bit is transmitted.

It is direct to obtain $p_c(\mathbf{c})$, $p_e(\mathbf{c})$, and $p_u(\mathbf{c})$, with $n - k_0 = w_H(\mathbf{c})$, from Table 1. However, for the purpose of computing p_g it is necessary to average these probabilities over the whole code. These averages $\mathbf{E}_{\mathcal{C}}[p_j(\mathbf{c})]$ may be obtained from the weight enumerator polynomial $W(x, y) = \sum_{j=0}^{n} A_{n-j} x^{n-j} y^j$ of the code, since they are linear combinations of averages

$$E_{\mathcal{C}}[k_0(k_0-1)\dots(k_0-v)(n-k_0)(n-k_0-1)\dots(n-k_0-u)] = W_{y\cdots y,x\cdots x}(1,1)'$$

where $W_{y\cdots y,x\cdots x}(1,1)'$ is the multiple partial derivative of $W(x, y)$ with respect to the subset of subscript variables evaluated for $x = y = 1$.

x_{ar}	k_0-1 0-posis.	$n-k_0$ 1-posis.	event prob.
correct detection error configurations			
			$(1-p)^{31}$
	e		$(k_0-1)p(1-p)^{30}$
e	e		$(k_0-1)p^2(1-p)^{29}$
	ee		$\frac{(k_0-1)(k_0-2)}{2}p^2(1-p)^{29}$
	e	e	$(k_0-1)(31-k_0)p^2(1-p)^{29}$
e	ee		$\frac{(k_0-1)(k_0-2)}{2}p^3(1-p)^{28}$
e	eee		$\frac{(k_0-1)(k_0-2)(k_0-3)}{6}p^4(1-p)^{27}$
e	ee	e	$\frac{(k_0-1)(k_0-2)(n-k_0)}{2}p^4(1-p)^{27}$
wrong detection error configurations			
e		e	$(31-k_0)p^2(1-p)^{29}$
		ee	$\frac{(31-k_0)(31-k_0-1)}{2}p^2(1-p)^{29}$
e		ee	$\frac{(31-k_0)(31-k_0-1)}{2}p^3(1-p)^{28}$
e		eee	$\frac{(31-k_0)(31-k_0-1)(31-k_0-2)}{6}p^4(1-p)^{27}$
e	e	ee	$\frac{(k_0-1)(31-k_0)(31-k_0-1)}{2}p^4(1-p)^{27}$
erasure detection error configurations			
		e	$(31-k_0)p(1-p)^{30}$
e	e	e	$(k_0-1)(31-k_0)p^3(1-p)^{28}$

The approximated error probabilities of the BSC modelling a stega-channel with Mode 2 is

$$p_g = \frac{1}{2}-\frac{1}{2}(1-p)^{31}-\frac{29}{4}p(1-p)^{30}-\frac{403}{4}p^2(1-p)^{29}+\frac{29}{4}p^3(1-p)^{28}+\frac{405}{4}p^4(1-p)^{27}$$

The curves of p_g versus p for both Modes are shown in Figure 2, together with the curves obtained by simulation: the agreement is very good whenever $31p < 3$.

5 - Conclusions

A realization of a steganographic channel has been described which exploits the redundancy of error-correcting codes over communication channels. The steganographic information is sent as artificial errors on the channel, and it is undetectable provided that the artificial errors do not significantly affect the rate of genuine errors, in the sense that the artificial error rate is undistinguishable from the random variations of the genuine error rate.
An interesting observation is that the performance of the steganographic channel principally depends on the error correcting capabilities of the code, rather

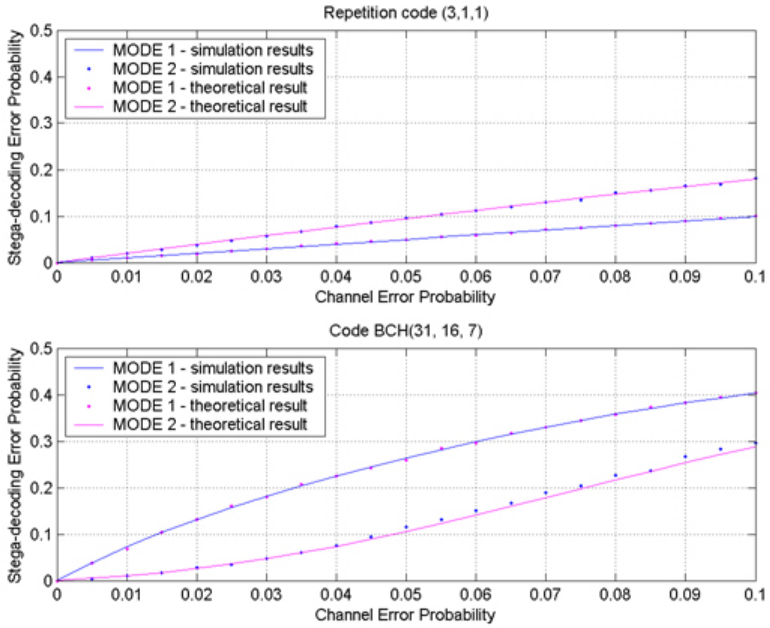

Figure 2: BSC model of Stega-channel: Curves of p_g versus p by simulation and formulas

than on the code itself, and it is also independent of the code rate to some extent.

Two modes of inserting artificial errors, called Mode 1 and Mode 2, have been considered, and the resulting steganographic channel, seen as BSC with a bit error probability p_g, has been completely characterized for each mode.

The feasibility of each Mode has been tested by simulating the overall transmission, and the estimated bit error probability p_g is in good agreement with that obtained by means of explicit formulas. The examples show that it is not possible to establish a rank order between Mode 1 and Mode 2 based on achievable p_g, since example 1 shows that Mode 2 is better than Mode 1, whereas example 2 shows the converse.

Lastly, the proposed steganographic scheme is applicable to any system that employs error correcting codes. In particular, it enables the watermarking of consumer products to be personalized whenever an error-correcting code is used, as for example in CD-Roms or in GSM cell phone systems.

References

[1] R.J. Anderson, F.A.P. Petitcolas, On the Limits of Steganography, *IEEE Journal of Selected Areas in Communications*, 1998.

[2] J. Fridrich, M. Goljan, D. Hogea, D. Soukal, Quantitative steganalysis of digital images: estimating the secret message length, *Multimedia Systems*, 2003.

[3] N.F. Johnson, S. Jajodia, Steganalysis, The Investigation of Hidden Information, *Proceedings of the 1998 IEEE Information Technology Conference*, Syracuse, New York, USA, September 1st - 3rd, 1998.

[4] S. Katzenbeisser, F.A.P. Petitcolas, *Information hiding techniques for steganography and digital watermarking*, Artech House Books, 1999.

[5] F.J. MacWilliams and N.A.J. Sloane, *The Theory of Error Correcting Codes,* North Holland, Amsterdam, 1977.

[6] A.J. Menezes, P.C. van Oorschot, S.A. Vanstone, *Handbook of Applied Cryptography*, New York: CRC, 1997.

[7] B. Schneier, *Applied Cryptography*, New York: Wiley, 1995.

[8] C.E. Shannon, Communication Theory and Secrecy Systems, *BSTJ*, vol. 28, 1949, pp.656-715.

[9] G.J. Simmons, *Contemporary Cryptology: The Science of Information Integrity*, New York: IEEE Press, 1992.

[10] P. Wayner, *Disappearing Cryptography - Information Hiding: Steganography and Watermarking*, second edition, Morgan Kaufman Publisher, 2002.

[11] A. Westfeld, A. Pfitzmann, Attacks on Steganographic Systems, *Proceedings of the Third International Workshop on Information Hiding*, p.61-76, Sept. 29- Oct. 01, 1999.

Received: February 9, 2008