Fifteenth International Workshop on Algebraic and Combinatorial Coding Theory
June 18-24, 2016, Albena, Bulgaria *pp. 250–254*

On Ryabko&Ryabko asymptotically optimal perfect steganographic scheme in a noisy channel

Valeria Potapova lera.potapova.93@mail.ru
Institute for Information Transmission Problems, Moscow, Russia

Abstract. We consider a scenario when a perfect steganographic system of [1] is used for data transmission over a noisy channel. We shown that errors in the resulting channel form the insertion-deletion channel.

1 Introduction

We consider a problem of transmission of hidden messages over a noisy channel. Usually steganography deals with the problem of hiding of secret messages into ordinary messages, which are called containers. We call the resulting message as stegoword and the main goal of a steganosystem is to make stegowords indistinguishable from empty containers, i.e. containers which do not contain a secret message. The corresponding model, introduced in [2], consists of two participants: Alice and Bob, who want to communicate being in prison, and third participant - a warden who should deliver a message from Bob to Alice (or vice versa) but the warden may not deliver it if the corresponding message looks suspicious to him. But the warden cannot just reject to deliver any messages (just to be on a safe side) as he "pays" some fine for any message (container) which he rejected to deliver but which appeared to be "empty". A stegosystem called *perfect* if it is not possible to distinguish between "empty" container and stegoword. In this case the warden have to deliver all messages.

There are two main criteria how to measure indistinguishability. Combinatorial one, when embedding (or hiding) a secret message into the container produces just a few changes in the container. I.e., Hamming distance between a container and the corresponding stegoword should be small enough. First practical algorithm of such type was proposed by Crandall under the name of matrix method [3]. This method can be explained as employing of linear covering codes, and even more complicated problem, when the warden becomes active and can alter "transmitted" messages, can be stated as a coding theory problem, see [4],[5].

Another model is a probabilistic one. Namely, we assume that containers are generated by some source with the probability distribution and embedding should not change this probability distribution. Such stegosystem called *perfect*, see [6]. A nice construction of perfect stegosystems for sources with independent letters was proposed and investigated in [1]. As usual in steganography it was considered under noiseless assumption, i.e. in the case of a passive warden. In this paper we consider more general case of a noisy channel, or, an active warden, like in [5], but with different criteria of indistinguishability. Our main result is to show that the resulting errors in secret messages produced by the active warden are equivalent to insertions and deletions – the well-known model introduced by Levenshtein [7].

2 Asymptotically optimal perfect steganographic systems

Consider the following probabilistic model [6]. There is some source μ of non-secret messages (containers). Containers are generated as strings of symbols which are i.i.d. random variables from some finite alphabet $\mathbb{A}$. The sender wants to use containers generated by μ for a hidden transmission of binary secret messages. These secret binary messages are independent and generated equiprobably by a source ω. In the channel the warden can intercept and then reads all messages. He tries to find out if a given the container is "empty" or not. If containers with secret information and without secret information are identically distributed, then the warden fails. Such steganographic system called *perfect*.

In the article "Asymptotically optimal perfect steganographic systems" of B. Ya. Ryabko and D. B. Ryabko [1] the construction of perfect steganographic scheme was proposed and some asymptotically tight bounds on their transmission rate were found. In addition, the authors designed simple encoding(embedding) and decoding algorithms. We shall give a sketch of the scheme below. Our goal is to consider the following scenario when Alice and Bob use the perfect scheme of [1] and the warden knowing that he cannot distinguish between "empty" container and stegoword tries to destroy communication between Alice and Bob by altering transmitted messages, i.e., by introducing some errors. It is natural to assume that the warden's power of producing errors is limited in the number of errors.

Consider the simplest case. The alphabet $\mathbb{A} = \{a, b\}$ is binary. We should transmit $y = y_1, y_2, ...$ by embedding it into the message $x = x_1x_2x_3....$ The

sequences x and y turn into a new sequence X. The receiver should be able to extract the secret message y from X, and the distribution of symbols in X should be equal to the distribution of symbols in x. Firstly, we divide symbols of x into pairs and rename them in the following way:

$aa = u$, $bb = u$, $ab = v_0$ and $ba = v_1$.

Then the pairs aa and bb are idle, we don't use them for embedding, but we change pairs v_k into pairs corresponding to $v_{y_1}, v_{y_2}, v_{y_3}, ...$ in the following way:
$(X_{2i-1}, X_{2i}) = (min\{x_{2i-1}, x_{2i}\}, max\{x_{2i-1}, x_{2i}\})$ if the corresponding $y_k = 0$ and
$(X_{2i-1}, X_{2i}) = (max\{x_{2i-1}, x_{2i}\}, min\{x_{2i-1}, x_{2i}\})$ if $y_k = 1$.

For example, $y = 0110...$ and $x = aababaaaabbaaaaabb....$ By renaming pairs we get
$x = uv_1v_1uv_0v_1uuu....$ We embed y and end up with the sequence
$X = uv_0v_1uv_1v_0uuu... = aaabbaaabaabaaaabb.$

Decoding is obvious. The receiver divides the sequence X into pairs and switches pairs ab and ba to 0 and 1 correspondingly. It was proved in [1] that the scheme is perfect, i.e., that the distribution of probabilities of sequences of symbols after embedding is the same as the source μ has.

Now consider the situation when X is transmitted over a noisy channel, or, the same, the warden is active. If a single error occurs in a pair aa or bb then we end up here with ab or ba. The decoder interprets this pair as a pair that contains some secret information (but actually it doesn't). So an insertion occurs in the secret message y. On the other hand, if a single error occurs in a pair ab or ba, the receiver obtains here aa or bb and after decoding it loses a secret symbol which was embedded in the original pair. In other words, there is a deletion in the message y. Hence single errors produced by the warden (or the channel) cause single insertion/deletion in the secret message y. The corresponding class of codes correcting single insertion/deletion is well-known [7].

This stego construction was generalized in [1] to nonbinary alphabet $\mathbb{A} = \{0, 1, \ldots, q-1\}$, which symbols are ordered as integers. Like in the case of binary alphabet, the pairs of equal symbols are idle, but the pairs of different symbols are changed in the following way:
$(X_{2i-1}, X_{2i}) = (min\{x_{2i-1}, x_{2i}\}, max\{x_{2i-1}, x_{2i}\})$ if the corresponding $y_k = 0$ and
$(X_{2i-1}, X_{2i}) = (max\{x_{2i-1}, x_{2i}\}, min\{x_{2i-1}, x_{2i}\})$ if $y_k = 1$.

In the nonbinary case curious things take place during the transmission via a noisy channel (or via a channel with an active warden). If a single errors

occur in pairs such as $\alpha\alpha$ where $\alpha \in \mathbb{A}$, then the insertion happens in the secret message. If a single error occurs in a pair $\alpha\beta$, then sometimes it ends up with deletion of a symbol in the secret message, sometimes with a regular reversal error, and sometimes nothing happens! More formally, let us assume that we transmit a pair $\alpha\beta$ and $\alpha < \beta$ in terms of the assigned order. It means that the secret symbol is 0. If the pair turns into $\alpha\alpha$ or $\beta\beta$ (that happens with the conditional probability $\frac{2}{q-1}$) then the symbol is deleted. If α turns into α' and $\alpha' > \beta$ or β turns into β' and $\alpha > \beta'$ then we have a regular reversal error. In all other cases the error doesn't bother the secret symbol. It is interesting to notice that the probability of the reversal depends on the pair. Say α is k-th letter in the alphabet and β is l-th letter ($k < l$). Then the probability of the reversal on conditions that an error takes place is $\frac{q-1-l}{q-1} + \frac{k}{q-1}$. We can see the same effect on pairs corresponding to 1.

More general stego scheme was proposed in [1] which uses blocks of length n (n is a parameter of the scheme). The idea is to use the lexicographical order on the blocks of length n and compositional classes. The model of errors that occur in secret message when the warden is much more complicated.

3 Conclusion and Acknowledgement

We have investigated the universal perfect steganographic system and its behavior during the transmission via a noisy channel or, the same, a channel with an active warden. If an error in transmitted stegoword happens during the transmission, an insertion/deletion takes place in the embedded secret message.

The research was carried out at the IITP RAS and supported by the Russian Science Foundation (project 14-50-00150).

References

[1] Ryabko B. Ya., Ryabko D.B "Asymptotically optimal perfect steganographic systems", Prob. Peredach. Inform., vol. 45, no. 2, pp.119-126, 2009.

[2] Simmons G. J., "The Prisoner's Problem and the Subliminal Channel", Advances in Cryptology, Proceedings of CRYPTO' 83 (Workshop on Communication Security), Plenum, New York, pp. 51-67, 1984.

[3] Crandall R.,"Some notes on steganography", 1998

[4] F. Galand, G. Kabatiansky, “Steganography via covering codes”, Proceedings. IEEE International Symposium on Information Theory, p.192, 2003.

[5] F. Galand, G. Kabatiansky, “Coverings, centered codes, and combinatorial steganography”, Prob. Peredach. Inform., vol. 45, no. 3, pp.289-294, 2009.

[6] Cachin C., ”An Information-Theoretic Model for Steganography”, Proc. 2nd Int. Workshop on Information Hiding, Lecture Notes Comput. Sci., 1525, Springer, Berlin, pp. 306-318, 1998.

[7] V.I. Levenshtein, “ Binary codes capable of correcting deletions, insertions, and reversals”, Soviet Physics Doklady, 10 (8), pp.707710, 1965. Prob. Peredach. Inform., vol. 1, no.1, pp. 12-25, 1965