Information-Theoretic Approaches to Steganography: Latest Achievements

Boris Ryabko*, Andrey Fionov, Katherine Eltysheva, Ivan Nechta,
Yulia Soldatova, Mikhail Zhilkin
Siberian State Univ. of Telecom. and Inform. Sci.
Kirov St. 86, Novosibirsk 630102 Russia

We consider the following problem of steganography: A must send a message to B in such a way that not only the content of the message but also the very fact of its transmission be concealed for anyone else. To accomplish this task, A chooses an innocuous file (whose transmission will not raise any suspicion), encrypts the message and embeds it in the file. The file is then sent to B who will be able to extract and decrypt the message. The file is said to be a container for hidden messages, the contents of the file be a covertext, and the covertext with embedded message be a stegotext. The security of the stegosystem rests upon inability to distinguish between covertext and stegotext.

The stated problem attracts much research due to its great practical importance for many applications. In this talk we consider the essence of information-theoretic approach to the problem of steganography. The covertext is assumed to be generated by a probabilistic source, the encrypted message be a sequence of equiprobable and independent bits. The first question is determining the embedding capacity of covertext, the second is the security of embedding. The security is addressed by the notion of perfect stegosystem [1]. A stegosystem is perfect if statistical structures of initial covertext and the covertext with embedded message are equivalent, or, in other words, the stegotext is one of the messages that might be generated by the same source as the covertext. There is not much literature on information-theoretic aspects of steganography. The most noticeable papers devoted to study of capacity, other bounds, and some codes are [2], [3], [4], [5]. Efficient constructions of perfect stegosystems for sources with known and unknown statistics were recently suggested by the authors [6], [7]. The ideas of these constructions and further work will be presented.

The constructions of perfect stegosystems and the methods of source coding theory have motivated the search for increasing security of practical stegosystems. The results on new methods of embedding data in raster images and executable files are detailed in the corresponding papers at this symposium.

The dual direction of research which also benefits from information-theoretic ideas, is steganalysis, i.e. the problem of detecting hidden data in practical (not perfect) stegosystems. As was stated in the beginning, the problem is reduced to finding deviations in statistical structures of covertexts and stegotexts. The use of universal source codes was recently suggested by the authors as a powerful tool for solving many statistical problems [8], [9]. These codes are implemented today as data compression systems, such as archivers. Data compression was successfully applied to steganalysis of digital images (BMP and JPEG files) [10]. The main results and further work in this direction will be presented.

* This work has been supported by Russian Foundation for Basic Research, under grant no. 09-07-00005-.

REFERENCES

[1] C. Cachin, "An information-theoretic model of steganography," in *Proc. 2nd Information Hiding Workshop, Lect. Notes in Compute. SCSI.*, Springer Verlag, vol. 1525, 1998, pp. 306–318.

[2] P. Moulin and J.A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 563–593, 2003.

[3] Tri Van Le and K. Kurosawa, "Efficient public key steganography secure against adaptive chosen stegotext attacks," *Cryptology ePrint Archive*, Report 2003/244, 2003, `http://eprint.iacr.org/2003/244`.

[4] Y. Wang and P. Moulin, "Perfectly secure steganography: capacity, error exponents, and code constructions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2706–2722, 2008.

[5] J. Shikata and T. Matsumoto, "Unconditionally secure steganography against active attacks," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2690–2705, 2008.

[6] B. Ryabko and D. Ryabko, "Information-theoretic approach to steganographic systems," in *IEEE Int. Symposium on Inform. Theory*, Nice, France, 2007. pp. 2461-2464.

[7] A. Fionov and B. Ryabko, "Simple ideal steganographic systems for containers with known statistics," in *XI Int. Symposium on Problems of Redundancy*, St.-Petersburg, July 2-6, 2007, pp. 184-188.

[8] B. Ryabko and J. Astola, "Universal codes as a basis for time series testing," *Statistical Methodology*, vol. 3, pp. 375-397, 2006.

[9] B. Ryabko and J. Astola, "Universal codes as a basis for nonparametric testing of serial independence for time series," *Journal of Statistical Planning and Inference*, vol. 136, no. 12, pp. 4119-4128, 2006.

[10] M. Zhilkin, N. Melentsova and B. Ryabko, "Data compression based method of revealing hidden information in steganographic systems," in *XI Int. Symposium on Problems of Redundancy*, St.-Petersburg, July 2-6, 2007, pp. 42-44.