A Method for Hiding Secret Messages using Minimum-Redundancy Codes

Srinivas.CH[1], D.Prabhakar[2], Jayaraman.K[3], Gopala Krishna.M[4]

[1,2]Associate Professor, Dept. of ECE, MIC College of Technology, Vijayawada, AP, India
[3]Assistant Professor, Dept. of ECE, S.M.K.FOMRA INSTITUTE OF TECHNOLOGY, Chennai, India
[4]Assistant Professor, Dept. of ECE, MIC College of Technology, Vijayawada, AP, India

Abstract:

This paper demonstrates an effective lossless data hiding scheme using minimum Redundancy codes. The binary secret message is concurrently embedded and encoded with a cover medium such as a video file, an audio file, or even a text file. The proposed scheme not only provides good data hiding capacity and data recovery capability, but also being efficient in space saving. Each symbol in a cover medium can carry one secret bit, and the cover medium can be reversed. And the experimental results show that the redundancy code can saves up to 38% of space compared with the cover medium. In this paper, the symbol or sequence of symbols associated with a given message will be called the "message code." The entire number of messages which might be transmitted will be called the "message ensemble." The mutual agreement between the transmitter and the receiver about the meaning of the code for each message of the ensemble will be called the "ensemble code."

Keywords: data hiding, redundancy codes, Secret messages, method.

I. Introduction

An optimum method of coding an ensemble of messages consisting of a finite number of members is developed. A minimum-redundancy code is one constructed in such a way that the average number of coding digits per message is minimized.

Server data formats are used to be the cover medium in data hiding, e.g. audio files, video files, image files, text files, and so on. Although the data structure of text files is similar to image files than the other data format mentioned above, most of image data hiding schemes are not suitable for text files. The main reason is that most image data hiding schemes embed secret information into cover image by slightly perturbing the pixel values. Since gray-scale or color images can tolerant a small amount modifications of pixel values, it will cause no perceptible distortions. On the contrary, any changes in the text file might lead to meaningless content.

Few studies have referred to hiding secret messages in text files. In [1], the data was embedded by modifying the inter-character space, but it resulted in some distortions in the shape of words. In [3], a technique was proposed for copyright protection that marks the text file by shifting lines up or down and words right or left; however, the technique might change the typesetting of the text file accordingly. In addition to the security problem, bandwidth consumption is also an important concern. The size of transmitted files can be reduced by either of two categories of data compression technology: lossless and lossy technologies. The lossy data compression technology is widely used in images, but it may be unsuitable for text files because any loss of data may lead the content meaningless.

II. Derived Coding Requirements

For an optimum code, the length of a given message code can never be less than the length of a more probable message code. If this requirement were not met, then a reduction in average message length could be obtained by interchanging the codes for the two messages in question in such a way that the shorter code becomes associated with the more probable message. Also, if there are several messages with the same probability, then it is possible that the codes for these messages may differ in length. However, the codes for these messages may be interchanged in any way without affecting the average code length for the message ensemble. Therefore, it may be assumed that the messages in the ensemble have been ordered in a fashion such that

$$P(1)\geq P(2)\geq \cdots P(N-1)\geq P(N)$$

and that, in addition, for an optimum code, the condition

$$L(1)\leq L(2)\leq \cdots L(N-1)\leq L(N)$$

holds. This requirement is assumed to be satisfied throughout the following discussion. It might be imagined that an ensemble code, could assign q more digits to the Nth message than to the (N-l)st message. However, the first L(N- 1) digits of the Nth message must not be used as the code for any other message. Thus the additional q digits would serve no useful purpose and would unnecessarily increase Lav. Therefore, for an optimum code it is necessary that L(N) be equal to L(N-1).

The kth prefix of a message code will be defined as the first k digits of that message code. Basic restriction (b) could then be restated as: No message shall be coded in such a way that its code is a prefix of any other message, or that any of its prefixes are used elsewhere as a message code.

Imagine an optimum code in which no two of the messages coded with length L(N) have identical prefixes of order L(N) -1. Since an optimum code has been assumed, then none of these messages of length L(N) can have codes or prefixes of any order which correspond to other codes. It would then be possible to drop the last digit of this entire group of messages and thereby reduce the value of Lav. Therefore, in an optimum code, it is necessary that at least two (and no more than D) of the codes with length L(N) have identical prefixes of order L(N)-1.

The procedure is applied again and again until the number of members in the most recently formed auxiliary message ensemble is reduced to two. One of each of the binary digits is assigned to each of these two composite messages. These messages are then combined to form a single composite message with probability unity, and the-coding is complete.

III. Hiding Terminology

As we have noted previously, there has been a growing interest, by different research communities, in the fields of steganography, digital watermarking, and fingerprinting. This led to some confusion in the terminology. We shall now briefly introduce the terminology which will be used in the rest of the paper and which was agreed at the first international workshop on the subject [4], [9] (Fig. 1).

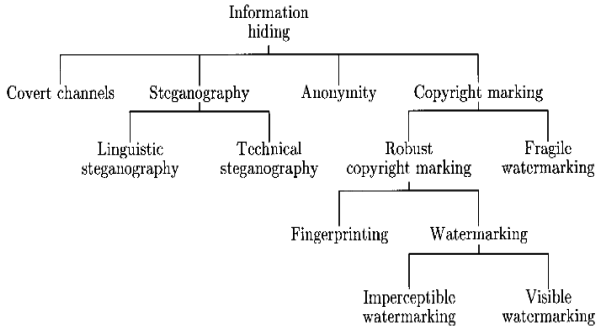

Fig.1: A classification of information-hiding techniques

The general model of hiding data in other data can be described as follows. The embedded data are the message that one wishes to send secretly. It is usually hidden in an innocuous message referred to as a cover text, cover image, or cover audio as appropriate, producing the stegotext or other stego-object. A stego-key is used to control the hiding process so as to restrict detection and/or recovery of the embedded data to parties who know it (or who know some derived key value).

As the purpose of steganography is having a covert communication between two parties whose existence is unknown to a possible attacker, a successful attack consists in detecting the existence of this communication. Copyright marking, as opposed to steganography, has the additional requirement of robustness against possible attacks. In this context, the term "robustness" is still not very clear; it mainly depends on the application. Copyright marks do not always need to be hidden, as some systems use visible digital watermarks, but most of the literature has focused on invisible (or transparent) digital watermarks which have wider applications. Visible digital watermarks are strongly linked to the original paper watermarks which appeared at the end of the thirteenth century to differentiate paper makers of that time. Modern visible watermarks may be visual patterns (e.g., a company logo or copyright sign) overlaid on digital images.

In the literature on digital marking, the stego-object is usually referred to as the marked object rather than stego-object. We may also qualify marks depending on the application. Fragile watermarks1 are destroyed as soon as the object is modified too much. This can be used to prove that an object has not been "doctored" and might be useful if digital images are used as evidence in court. Robust marks have the property that it is infeasible to remove them or make them useless without destroying the object at the same time. This usually means that the mark should be embedded in the most perceptually significant components of the object. Fingerprints (also called labels by some authors) are like hidden serial numbers which enable the intellectual property owner to identify which customer broke his license agreement by supplying the property to third parties. Watermarks tell us who is the owner of the object.

IV. Steganographic Technique

We will now look at some of the techniques used to hide information. Many of these go back to antiquity, but unfortunately many modern system designers fail to learn from the mistakes of their predecessors. By the sixteenth and seventeenth centuries, there had arisen a large literature on steganography and many of the methods depended on novel means of encoding information.

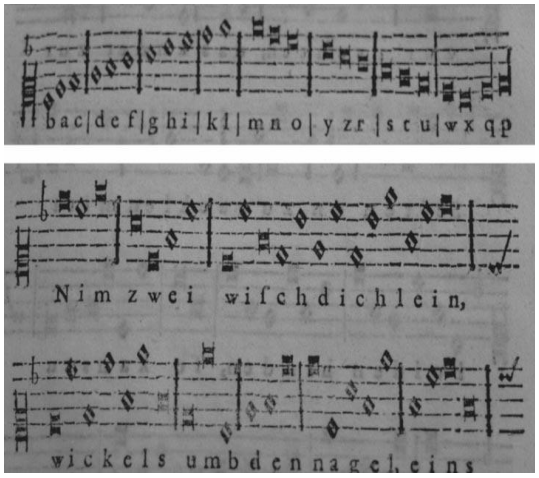

Fig. 2. Hiding information into music scores: Schott simply maps the letters of the alphabet to the notes

Schott (1608–1666) explains how to hide messages in music scores: each note corresponds to a letter (Fig. 4). Another method, based on the number of occurrences of notes and used by Bach, is mentioned in [10]. Schott also expands the "Ave Maria" code proposed by Trithemius (1462–1516) in *Steganographiæ*, one of the first known books in the field. The expanded code uses 40 tables, each of which contains 24 entries (one for each letter of the alphabet of that time) in four languages: Latin; German; Italian; and French. Each letter of the plain text is replaced by the word or phrase that appears in the corresponding table entry and the stego-text ends up looking like a prayer or a magic spell. It has been shown recently that these tables can be deciphered by reducing them modulo 25 and applying them to a reversed alphabet. In [2], Wilkins (1614–1672), Master of Trinity College, Cambridge, shows how "two

Musicians may discourse with one another by playing upon their instruments of musick as well as by talking with their instruments of speech" [2, ch. XVIII, pp. 143–150]. He also explains how one can hide secretly a message into a geometric drawing using points, lines or triangles. "The point, the ends of the lines and the angles of the figures do

each of them by their different situation express a several letter" [2, ch. XI, pp. 88–96].

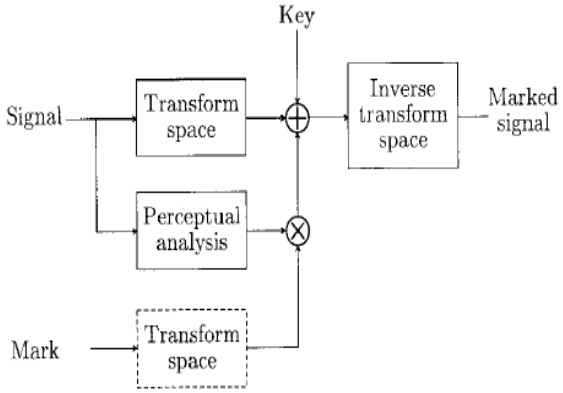

Fig. 3. A typical use of masking and transform space for digital watermarking and fingerprinting.

V. Conclusion

In this paper we gave an overview of information hiding in general and steganography in particular. We looked at a range of applications and tried to place the various techniques in historical context in order to elucidate the relationships between them, as many recently proposed systems have failed to learn from historical experience. We then described a number of attacks on information hiding systems, which between them demolish most of the current contenders in the copyright marking business. We have described a tool, StirMark, which breaks many of them by adding subperceptual distortion, and we have described a custom attack on echo hiding.

This led us to a discussion of marking in general. We described some of the problems in constructing a general theory and the practical requirements that marking schemes and steganographic systems may have to meet. We advanced the suggestion that it is impractical to demand that

VI. Acknowledgements

The authors would like to thank the anonymous reviewers for their comments which were very helpful in improving the quality and presentation of this paper.

References:

[1]. Tacticus, How to Survive Under Siege/Aineias the Tactician (Clarendon Ancient History Series). Oxford, U.K.: Clarendon, 1990, pp. 84–90, 183–193.
[2]. J. Wilkins, Mercury: Or the Secret and Swift Messenger: Shewing, How a Man May with Privacy and Speed Communicate His Thoughts to a Friend at Any Distance, 2nd ed. London, U.K.: Rich Baldwin, 1694.
[3]. D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," Commun. ACM, vol. 24, no. 2, pp. 84–88, Feb. 1981.
[4]. R. J. Anderson, Ed., Information Hiding: 1st Int.Workshop (Lecture Notes in Computer Science), vol. 1174. Berlin, Germany: Springer-Verlag, 1996.
[5]. S. Roche and J.-L. Dugelay, "Image watermarking based on the fractal transform," in Proc. Workshop Multimedia Signal Processing, Los Angeles, CA, 1998, pp. 358–363.
[6]. J.-P. M. G. Linnartz, "The "ticket" concept for copy control based on embedded signalling," in Computer Security—5th Europ. Symp. Research in Computer Security, (ESORICS'98) (Lecture Notes in Computer Science), vol. 1485, J.-J. Quisquater, Y. Deswarte, C. Meadows, and D. Gollmann, Eds. Berlin, Germany: Springer, 1998, pp. 257–274.
[7]. M. L. Miller, I. J. Cox, and J. A. Bloom, "Watermarking in the real world: An application to DVD," in Multimedia andSecurity—Workshop at ACM Multimedia'98 (GMD Report), vol.41, J. Dittmann, P. Wohlmacher, P. Horster, and R. Steinmetz, Eds. Bristol, U.K.: ACM, GMD—Forschungszentrum Informationstechnik GmbH, 1998, pp. 71–76.
[8]. J. C. Benaloh, Verifiable Secret-Ballot Elections, Ph.D. dissertation, Yale University, New Haven, CT, YALEU/DCS/TR-561,1987.
[9]. B. Pfitzmann, "Information hiding terminology," in Lecture Notes in Computer Science, vol. 1174. Berlin, Germany: Springer-Verlag, 1996.

Authors Profile:

Srinivas.CH, working as Associate professor in MIC college of Technology, has 10 years of industrial and teaching experience. He received his M.E degree in ECE from College of Engineering, Guindy, Anna University, Chennai in 2003.

D.Prabhakar, working as Associate professor in MIC College of Technology, has 8 years of Teaching Experience. He received his M.Tech degree in Radar & Microwave Engineering from Andhra University in 2003.

Jayaraman krishnamoorthy, working in Rajiv Gandhi Salay IT Highway (OMR), has worked as Assistant Professor in Shree Motilal Kanhaiyalal FOMRA INSTITUTE OF TECHNOLOGY. He received his M.Tech degree in VLSI from SATHYABAMA UNIVERSITY, Chennai.

Gopala Krishna.M, working as Assistant professor in MIC College of Technology, has 4 years of Teaching Experience. He received his M.Tech degree in VLSI.