Secure Hybrid Spread Spectrum System for Steganography in Auditive Media

Marcus Nutzinger*, Christian Fabian†, Marion Marschalek‡
Institute of IT Security Research
University of Applied Sciences St. Poelten
St. Poelten, Austria
{*mnutzinger,†lbfabian,‡is091512}@fhstp.ac.at

***Abstract*—Steganography is used to embed secret messages in cover media. This is especially important in areas, where the use of cryptography is prohibited. In this paper, we introduce a novel hybrid steganographic algorithm for digital audio data. We enhance a direct sequence spread spectrum (DSSS) system with aspects of frequency hopping to vary the carrier frequency of the binary phase shift keying (BPSK) signal, representing the secret message. Further, we adopt the number of chips per secret bit. These modifications give a more secure steganographic system, making guesses about the bit-rate or message length less feasible. Further, an adaptive amplitude control is used to guarantee inaudibility, even in silent passages. After an overview of the DSSS technique, we highlight our contributions to form a more secure steganographic system. The results from our working prototype are pointed out and further usage scenarios in relation to future research are given.**

***Index Terms*—steganography; digital audio; direct sequence spread spectrum; frequency hopping; bit-rate variation;**

I. Introduction

Steganography is the science which deals with hidden information exchange. In a communication, only the sender and the intended recipient are supposed to know about, and extract, the secret message [1]. In *technical steganography*, the use of different cover media, like digital audio and video files or digital images, is possible [2].

The purpose of steganographic techniques is to transmit data secretly and to identify or protect the owner of information. Steganographic algorithms are an important field of research, as numerous applications do exist. These range from copyright protection of digital media to tamper proving, authentication, integrity proving and secure data transmission [3]. The key requirements for steganography are imperceptibility, robustness against destruction and a high data rate [4]. The latter is quite contrary to the invisibility of hidden information, because the more it is embedded, the easier it is for an observer to detect parts of the secret message.

Various steganographic algorithms have been proposed, including echo-hiding, phase-coding, patchwork technique and spread spectrum [5]. Our research focuses on auditive data as cover media, such as voice calls or compressed and uncompressed digital sound files. To embed messages in digital audio data, the spread spectrum technique seems to be the most promising field of research [6].

The major contribution of this paper is the advancement of the DSSS technology in the field of steganography. To form a more secure algorithm, two additional variables are introduced in the communication process. By making the frequency of the carrier, which modulates the secret message, adjustable, guesses about the exact location of this frequency are hampered. Therefore, it becomes harder for an observer trying to demodulate an eavesdropped communication. The second variable introduced by our system deals with the representation of secret bits as chip sequences. Based on a configurable number of chips per bit, this number is varied for every embedded bit. As a consequence, the bit-rate does not stay continuous over the whole communication. This in turn makes it even harder for an observer to estimate the bit-rate or length of the secret message.

The rest of this paper is structured as follows: In Section two, related approaches, dealing with enhancements of the spread spectrum algorithm, as well as improvements gained through our hybrid system are described. Section three gives a brief overview about the DSSS technique as it is the underlying algorithm in our system. The fourth section consists of three main parts. The first describes one aspect of our system, namely the variation of the carrier frequency. In the second part the other aspect introduced by our system, which deals with the adjustment of the chip count per secret bit, is presented. Part three of Section four summarizes results which we got from testing our system using a prototype implemenation. At the end of this paper, ideas for improvements and future research are given.

II. Related Work

A lot of work has been done in the field of information hiding in the context of the spread spectrum technology. Due to space limitations, we only provide a short overview of related approaches.

Most of the work published is put under the focus of digital watermarking [7], [8], [9], [10], [11], [12], [13]. In contrast, our approach deals with another aspect of information hiding, namely steganography. The main difference between these two techniques is the fact, that watermarks are public and their usage is known by everybody. In the field of steganography, the existence of hidden information is only known by the communication partners [14]. Hence, our main focus lies on the prevention of wiretapping, whereas in watermarking

978-0-7695-4222-5/10 $26.00 © 2010 IEEE
DOI 10.1109/IIHMSP.2010.27

systems, the prevention of various types of attacks has also to be dealt with [2].

While our focus lies on auditive data as cover media, which could be Voice over IP (VoIP) streams, WAV-files or samples recorded from a soundcard, many publications in the field of watermarking only deal with digital images [15], [16], [17], [18], [19]. Our work distinguishes itself from other publications in the audio watermarking or steganography field in important aspects. Issues like the synchronisation of sender and receiver, as described in [20], are improved by our solution. Further, we do not significantly raise the noise level [21] while still giving good results. The carrier frequency as well as the bit-rate are varied, while still being suitable for a voice channel [11]. We also adopt an amplitude control for the BPSK signal, representing the chip sequence [12]. Instead of using two different algorithms in a row to form a hybrid system [8], [15], [16], [17], [18], [19], [22], we aim at improving the spread spectrum algorithm to achieve best results without losing the advantages of a homogenous system.

III. The Spread Spectrum Technique

Before entering the description of our proposed system, this section recalls some important aspects of the spreading and despreading processes.

In a DSSS system, a signal of low bandwidth is spread over a broad frequency range. Hence the power of the signal is decreased and thus the signal vanishes in the noise of the cover media [23]. To extract an embedded signal from the cover, the receiver needs knowledge about the spreading process. This knowledge can therefore be described as a kind of secret key, needed as input to the system.

At the sender, each bit is first converted to a sequence of chips, thus increasing the bandwidth while decreasing the signal power. Pseudo-noise (PN) chip sequences are generated using a linear feedback shift register (LFSR). The length of an output sequence depends on the number of stages. In general, m-sequences are preferred, depicting the maximum length of a generated sequence without repetitions [24].

After spreading, the resulting chips are modulated onto a carrier:

$$s(t) = A \cdot c(t) \cdot \cos(2\pi \cdot f_c \cdot t) \qquad (1)$$

$c(t)$ denotes the chip sequences and is built up of values 1 and -1. As a consequence, a BPSK modulation is done, introducing phase hops at each edge in $c(t)$.

At the receiver side, the BPSK signal has to be extracted from the noise of the received signal [23]. This is done by a coherent demodulation:

$$r(t) = s(t) \cdot \cos(2\pi \cdot f_c \cdot t + \varphi) \qquad (2)$$

f_c and φ have to be equal to those at the sender for the extraction to work. To get rid of resulting harmonics, the outcome is filtered before analyzation. For the extraction of the chip sequence, a correlator is used which matches the chips against the same PN-sequence used at the sender.

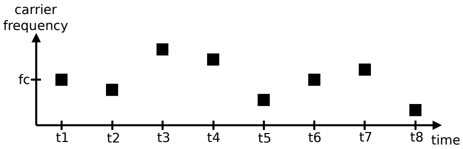

Figure 1. Varying the carrier frequency over time

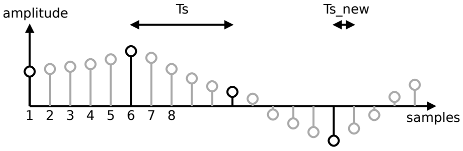

Figure 2. Oversampled audio signal

IV. The Hybrid System

Our system basically uses the spread spectrum technology as described in [25] and [26]. This section now goes into the novel contributions of our proposed algorithm. First, we highlight our enhancements towards a standard DSSS system. Afterwards, we present the results from our implemented prototype.

A. Frequency Hopping

The concept of frequency hopping is shown in Fig. 1. When varying the frequency of the BPSK carrier, not all available frequencies are useable. Taken into consideration the bit-rate, chip-rate and channel bandwidth, our system adapts the frequency to allow for good results at the receiver. Starting from a central frequency which fits best for the current channel properties, the frequency of the carrier is changed in a reasonable range.

At first, the audio signal has to be oversampled (see Fig. 2). If the audio signal has $f_s = 8kHz$, the oversampled signal has $f_s = 40kHz$. When adding the BPSK signal, not only samples 1 and 6 are available for the first two BPSK values, but further all the new sampling points in between. The generation of the carrier hence gets more flexible, because the distance between two sampling points does not have to be T_s anymore. It can be adjusted by $T_{s_{new}}$.

When decreasing the distance between two carrier sampling points, the carrier frequency is increased and the resulting signal length decreases. The reverse happens when increasing the distance. The carrier frequency is then calculated by

$$f_c = \frac{f_s}{k \cdot (1 \pm \frac{x}{n})}, \qquad (3)$$

where k denotes the number of sampling points in one sine period, x marks the distance between carrier sampling points and n refers to the oversampling factor.

Related approaches using a varying modulation carrier frequency are presented in [7] and [10]. Our method improves existing approaches in that the carrier frequency is not adopted according to signal features which an observer could analyze. Instead, a PN-sequence is generated which is used

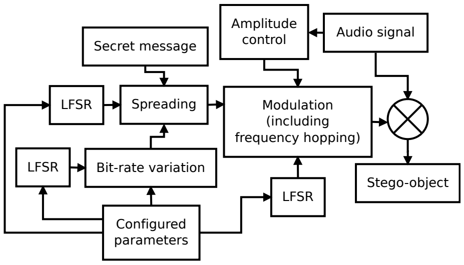

Figure 3. The embedding process

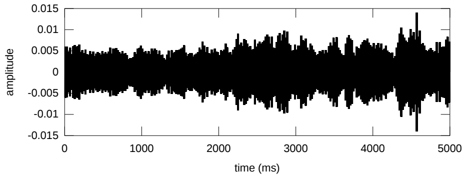

Figure 4. BPSK signal with amplitude control

to determine the distance of BPSK sampling points for the current signal portion. As a consequence, the LFSR parameters become part of the stego key.

B. Bit-rate Variation

In standard DSSS systems, the number of chips per bit stays constant which creates a target for steganalysis. This is especially true for systems using one PN-sequence and its inverted form to represent the bit states (*bipolar sequences*).

Our system allows for a variation of the bit-rate for each bit of the secret message. Starting from a configurable central number of chips per bit, the decimal interpretation of portions from a LFSR output is used to modify the bit-rate. In our prototype, portions of eight bits from a LFSR output are interpreted as signed decimal number. Hence variations from -128 below to 127 chips above the central number of chips are possible. Due to these variations, methods like the estimation of the bit-rate or the length of the secret message, as described in [27] and [28], are no longer feasible as the number of chips varies for every bit. As with frequency hopping, these LFSR parameters, together with the central number of chips per bit, become part of the stego key.

C. Embedding and Extraction

Fig. 3 shows an overview of the embedding process including frequency hopping and bit-rate variation. *Amplitude control* describes the regulation of the amplitude for the BPSK modulation carrier. This method prevents disruptions of the audio signal even in silent passages. In our prototype, the amplitude is changed 50 times per second, which keeps the signal-to-noise ratio (SNR) close to constant. Fig. 4 shows a sample BPSK signal created with amplitude control on the basis of a WAV-file.

A schematic of the extraction process is outlined in Fig. 5. Before starting to extract the secret message (*de-spreading*), the receiver has to synchronize itself with the sender. In our system, synchronization is achieved in two steps. First, the received signal is oversampled. The second step serves to approximate the start offset of the first secret bit. This is done by correlating the signal starting at the current offset with the PN-sequences used for bit representation. When a predefined treshold is reached, the current offset is assumed to be the start offset and the extraction of secret bits begins. Best results are produced when doing this synchronization not only for the start offset, but for every secret bit to approximate and readjust the phase angle.

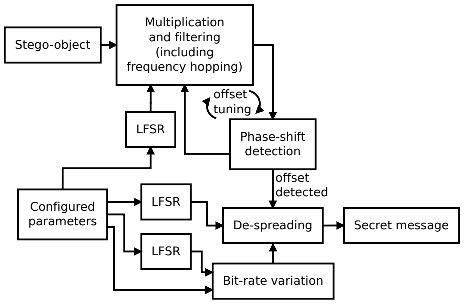

Figure 5. The extraction process

D. Testing our System

The system presented in this section has been implemented as a prototype using GNU Octave [29]. Tests with uncompressed WAV-files and MP3-compressed audio files, as well as encoding with the G.711 codec [30] and the GSM 06.60 codec [31] were successful. There have been none to negligible influences on the audio quality. The receiver was able to extract most of the transmitted bits forming the secret message.

V. Outlook and Future Work

This section gives ideas for improvements which came up during tests with our implemented prototype. Furthermore, future usage scenarios as well as possible research subjects are pointed out.

Currently, our prototype implementation works for most cases, with changing audio files, compression and encoding. Further work will be done in this area to optimize the different parameters used in the embedding and extraction processes to head for an error-free as well as undetectable transmission. In this context, an analyzation of the resulting audio quality as well as a comprehensive evaluation against other approaches will be accomplished.

Other publications, such as [25], describe the embedding of the BPSK signal in the frequency domain. Future research will focus on the meaningfulness of embedding parts of the signal in the frequency domain and other parts in the time domain. We want to figure out, if this makes it even more sophisticated for an observer to make assumptions about the bit-rate or the secret data length. Further, the influence of manipulations in the frequency domain on the quality of the resulting audio data has to be examined.

Another field of application for steganography, besides auditive media, are digital images. An idea for a further usage scenario is the implemenation of our system to work with digital images as cover media. As spread spectrum systems were used with digital images before [32], future research will show if our solution will also be feasible for operation in this area.

VI. Conclusion

In this paper, a novel, secure and hybrid steganographic algorithm using the DSSS technology in combination with frequency hopping and bit-rate variation is presented. As steganographic cover medium, digital audio data is used. Future work will investigate the usability of other cover media as well.

Our hybrid algorithm is built up on two enhancements of a standard DSSS system. Frequency hopping was introduced for the carrier signal, which expresses PN-sequences that represent bits from the secret message. Further, the concept of bit-rate variation was added, leading to a different number of chips for each secret bit. Both facets hamper an observer in making assumptions about the bit-rate or length of the embedded message.

An implemented prototype was used for tests of our system and the results yielded a good performance using standard audio codecs and compression techniques. Even after transmission over a GSM link, the receiver was able to extract most of the embedded bits using our prototype. These results justify for future research to optimize the embedding and extraction processes for further undetectability.

Acknowledgment

Our project is funded by the KIRAS program of the Austrian Research Promotion Agency (FFG). KIRAS funds projects in the field of security, with IT security being a subcategory in this context.

References

[1] S. Katzenbeisser and F. A. Petitcolas, Eds., *Information Hiding Techniques for Steganography and Digital Watermarking*. Norwood, MA, USA: Artech House, Inc., 2000.

[2] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding – a survey," *Proceedings of the IEEE, special issue on protection of multimedia content*, vol. 87, no. 7, pp. 1062–1078, Jul. 1999.

[3] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3-4, pp. 313–336, 1996.

[4] H. Wang and S. Wang, "Cyber warfare: steganography vs. steganalysis," *Commun. ACM*, vol. 47, no. 10, pp. 76–82, 2004.

[5] M. Arnold, M. Schmucker, and S. D. Wolthusen, *Techniques and Applications of Digital Watermarking and Content Protection*. Norwood, MA, USA: Artech House, Inc., 2003.

[6] X. He and M. Scordilis, *Spread Spectrum for Digital Audio Watermarking*, ser. Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks. Hershey, PA: Information Science Reference, 2008, ch. 2, pp. 11–49.

[7] T. Muntean, E. Grivel, and M. Najim, "Audio digital watermarking based on hybrid spread spectrum," in *CW '02: Proceedings of the First International Symposium on Cyber Worlds (CW'02)*. Washington, DC, USA: IEEE Computer Society, 2002, p. 150.

[8] J. Fridrich, "Combining low-frequency and spread spectrum watermarking," in *Proc. SPIE Int. Symposium on Optical Science, Engineering, and Instrumentation*, San Diego, July 19-24 1998, pp. 2–12.

[9] M. Jimenez Salinas and F. Garcia Ugalde, "Improved spread spectrum image watermarking in contourlet domain," 2008, pp. 1–6.

[10] N. Cvejic and T. Seppänen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," *Signal Process.*, vol. 84, no. 1, pp. 207–213, 2004.

[11] Q. Cheng and J. Sorensen, "Spread spectrum signaling for speech watermarking," in *ICASSP '01: Proceedings of the Acoustics, Speech, and Signal Processing, 2001. on IEEE International Conference*. Washington, DC, USA: IEEE Computer Society, 2001, pp. 1337–1340.

[12] H. Marvar and A. Florencio, "Improved spread spectrum: A new modulation technique for robust watermarking," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 898–905, April 2003.

[13] K. Yamamoto and M. Iwakiri, "Real-time audio watermarking based on characteristics of pcm in digital instrument," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 2, pp. 59–71, April 2010.

[14] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007.

[15] Z.-J. Lee, S.-W. Lin, S.-F. Su, and C.-Y. Lin, "A hybrid watermarking technique applied to digital images," *Appl. Soft Comput.*, vol. 8, no. 1, pp. 798–808, 2008.

[16] K. Bhandari, S. K. Mitra, and A. Jadhav, "A hybrid approach to digital image watermarking using singular value decomposition and spread spectrum," in *PReMI*, 2005, pp. 447–452.

[17] C.-T. Hsieh, Y.-K. Wu, and K.-M. Hung, "Hybrid watermarking scheme for halftone images," *International Journal of Advanced Science and Technology*, vol. 1, no. 1, pp. 9–20, December 2008.

[18] A. Sverdlov, S. Dexter, and A. Eskicioglu, "Robust dct-svd domain image watermarking for copyright protection: embedding data in all frequencies," in *Proceedings of the 13th European Signal Processing Conference (EUSIPCO '05)*, Turkey, September 2005.

[19] F. Deguillaume, S. Voloshynovskiy, and T. Pun, "Secure hybrid robust watermarking resistant against tampering and copy attack," *Signal Process.*, vol. 83, no. 10, pp. 2133–2170, 2003.

[20] J. Seok, J. Hong, and J. Kim, "A novel audio watermarking algorithm for copyright protection of digital audio," *ETRI Journal*, vol. 24, no. 3, pp. 181–189, June 2002.

[21] H. Matsuoka, "Spread spectrum audio steganography using sub-band phase shifting," in *IIH-MSP '06: Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 3–6.

[22] S. P. Mohanty, K. R. Ramakrishnan, and M. Kankanhalli, "A dual watermarking technique for images," in *MULTIMEDIA '99: Proceedings of the seventh ACM international conference on Multimedia (Part 2)*. New York, NY, USA: ACM, 1999, pp. 49–51.

[23] R. L. Pickholtz, D. L. Schilling, and L. B. Millstein, "Theory of spread spectrum communication - a tutorial," *IEEE Transactions on Communications*, pp. 855–884, 1982.

[24] S. W. Golomb, *Shift Register Sequences*. Laguna Hills, CA, USA: Aegean Park Press, 1981.

[25] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[26] D. Kirovski and H. Malvar, "Spread spectrum watermarking of audio signals," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 1020–1033, April 2003.

[27] Z. Dong and H. Hu, "The detection, symbol period and chip width estimation of dsss signals based on delay-multiply, correlation and spectrum analysis," *Engineering Letters*, vol. 15, no. 1, pp. 140–144, 2007.

[28] J. J. Fridrich, M. Goljan, D. Hogea, and D. Soukal, "Quantitative steganalysis of digital images: estimating the secret message length," *Multimedia Syst.*, vol. 9, no. 3, pp. 288–302, 2003.

[29] "GNU Octave," http://www.gnu.org/software/octave.

[30] "ITU-T Recommendation G.711. Pulse Code Modulation (PCM) of Voice Frequencies." ITU-T, Tech. Rep., November 1988.

[31] "ETSI GSM, ETS 300 726: Digital cellular telecommunications system; enhanced full rate (EFR) speech transcoding (GSM 06.60)," ETSI, Tech. Rep., 1999.

[32] L. M. Marvel, C. G. B. Jr., and C. T. Retter, "Spread spectrum image steganography," *IEEE Transactions on Image Processing*, vol. 8, no. 8, pp. 1075–1083, 1999.