Steganography in Digital Media

Steganography, the art of hiding of information in apparently innocuous objects or images, is a field with a rich heritage, and an area of rapid current development. This clear, self-contained guide shows you how to understand the building blocks of covert communication in digital media files and how to apply the techniques in practice, including those of steganalysis, the detection of steganography. Assuming only a basic knowledge in calculus and statistics, the book blends the various strands of steganography, including information theory, coding, signal estimation and detection, and statistical signal processing. Experiments on real media files demonstrate the performance of the techniques in real life, and most techniques are supplied with pseudo-code, making it easy to implement the algorithms. The book is ideal for students taking courses on steganography and information hiding, and is also a useful reference for engineers and practitioners working in media security and information assurance.

Jessica Fridrich is Professor of Electrical and Computer Engineering at Binghamton University, State University of New York (SUNY), where she has worked since receiving her Ph.D. from that institution in 1995. Since then, her research on data embedding and steganalysis has led to more than 85 papers and 7 US patents. She also received the SUNY Chancellor's Award for Excellence in Research in 2007 and the Award for Outstanding Inventor in 2002. Her main research interests are in steganography and steganalysis of digital media, digital watermarking, and digital image forensics.

Steganography in Digital Media

Principles, Algorithms, and Applications

JESSICA FRIDRICH
Binghamton University, State University of New York (SUNY)

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521190190

First published 2010

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-19019-0 Hardback

To Nicole and Kathy

Time will bring to light whatever is hidden; it will cover up and conceal what is now shining in splendor.

Quintus Horatius Flaccus (65–8 BC)

Contents

Preface

Steganography is another term for covert communication. It works by hiding messages in inconspicuous objects that are then sent to the intended recipient. The most important requirement of any steganographic system is that it should be impossible for an eavesdropper to distinguish between ordinary objects and objects that contain secret data.

Steganography in its modern form is relatively young. Until the early 1990s, this unusual mode of secret communication was used only by spies. At that time, it was hardly a research discipline because the methods were a mere collection of clever tricks with little or no theoretical basis that would allow steganography to evolve in the manner we see today. With the subsequent spontaneous transition of communication from analog to digital, this ancient field experienced an explosive rejuvenation. Hiding messages in electronic documents for the purpose of covert communication seemed easy enough to those with some background in computer programming. Soon, steganographic applications appeared on the Internet, giving the masses the ability to hide files in digital images, audio, or text. At the same time, steganography caught the attention of researchers and quickly developed into a rigorous discipline. With it, steganography came to the forefront of discussions at professional meetings, such as the Electronic Imaging meetings annually organized by the SPIE in San Jose, the IEEE International Conference on Image Processing (ICIP), and the ACM Multimedia and Security Workshop. In 1996, the first Information Hiding Workshop took place in Cambridge and this series of workshops has since become the premium annual meeting place to present the latest advancements in theory and applications of data hiding.

Steganography shares many common features with the related but fundamentally quite different field of digital watermarking. In late 1990s, digital watermarking dominated the research in data hiding due to its numerous lucrative applications, such as digital rights management, secure media distribution, and authentication. As watermarking matured, the interest in steganography and steganalysis gradually intensified, especially after concerns had been raised that steganography might be used by criminals.

Even though this is not the first book dealing with the subject of steganography [22, 47, 51, 123, 142, 211, 239, 250], as far as the author is aware this is the first self-contained text with in-depth exposition of both steganography and

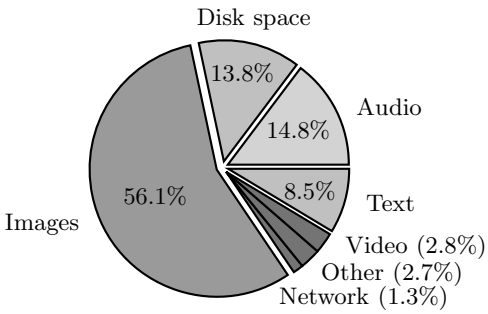
steganalysis for digital media files. Even though this field is still developing at a fast pace and many fundamental questions remain unresolved, the foundations have been laid and basic principles established. This book was written to provide the reader with the basic philosophy and building blocks from which many practical steganographic and steganalytic schemes are constructed. The selection of the material presented in this book represents the author's view of the field and is by no means an exhaustive survey of steganography in general. The selected examples from the literature were included to illustrate the basic concepts and provide the reader with specific technical solutions. Thus, any omissions in the references should not be interpreted as indications regarding the quality of the omitted work.

This book was written as a primary text for a graduate or senior undergraduate course on steganography. It can also serve as a supporting text for virtually any course dealing with aspects of media security, privacy, and secure communication. The research problems presented here may be used as motivational examples or projects to illustrate concepts taught in signal detection and estimation, image processing, and communication. The author hopes that the book will also be useful to researchers and engineers actively working in multimedia security and assist those who wish to enter this beautiful and rapidly evolving multidisciplinary field in their search for open and relevant research topics.

The text naturally evolved from lecture notes for a graduate course on steganography that the author has taught at Binghamton University, New York for several years. This pedigree influenced the presentation style of this book as well as its layout and content. The author tried to make the material as self-contained as possible within reasonable limits. Steganography is built upon the pillars of information theory, estimation and detection theory, coding theory, and machine learning. The book contains five appendices that cover all topics in these areas that the reader needs to become familiar with to obtain a firm grasp of the material. The prerequisites for this book are truly minimalistic and consist of college-level calculus and probability and statistics.

Each chapter starts with simple reasoning aimed to provoke the reader to think on his/her own and thus better see the need for the content that follows. The introduction of every chapter and section is written in a narrative style aimed to provide the big picture before presenting detailed technical arguments. The overall structure of the book and numerous cross-references help those who wish to read just selected chapters. To aid the reader in implementing the techniques, most algorithms described in this book are accompanied with a pseudo-code. Furthermore, practitioners will likely appreciate experiments on real media files that demonstrate the performance of the techniques in real life. The lessons learned serve as motivation for subsequent sections and chapters. In order to make the book accessible to a wide spectrum of readers, most technical arguments are presented in their simplest core form rather than the most general fashion, while referring the interested reader to literature for more details. Each chapter is closed with a brief summary that highlights the most important facts. Readers

Cover type	Count
Audio	445
Disk space	416
Images	1689
Network	39
Other files	81
Text	255
Video	86



Number of steganographic software applications that can hide data in electronic media as of June 2008. Adapted from [122] and reprinted with permission of John Wiley & Sons, Inc.

can test their newly acquired knowledge on carefully chosen exercises placed at the end of the chapters. More involved exercises are supplied with hints or even a brief sketch of the solution. Instructors are encouraged to choose selected exercises as homework assignments.

All concepts and methods presented in this book are illustrated on the example of digital images. There are several valid reasons for this choice. First and foremost, digital images are by far the most common type of media for which steganographic applications are currently available. Furthermore, many basic principles and methodologies can be readily extended from images to other digital media, such as video and audio. It is also considerably easier to explain the perceptual impact of modifying an image rather than an audio clip simply because images can be printed on paper. Lastly, when compared with other digital objects, the field of image steganography and steganalysis is by far the most advanced today, with numerous techniques available for most typical image formats.

The first chapter contains a brief historical narrative that starts with the rather amusing ancient methods, continues with more advanced ideas for data hiding in written documents as well as techniques used by spies during times of war, and ends with modern steganography in digital files. By introducing three fictitious characters, prisoners Alice and Bob and warden Eve, we informally describe secure steganographic communication as the famous prisoners' problem in which Alice and Bob try to secretly communicate without arousing the suspicion of Eve, who is eagerly eavesdropping. These three characters will be used in the book to make the language more accessible and a little less formal when explaining technical aspects of data-hiding methods. The chapter is closed with a section that highlights the differences between digital watermarking and steganography.

Knowing how visual data is represented in a computer is a necessary prerequisite to understand the technical material in this book. Chapter 2 first explains basic color models used for representing color in a computer. Then, we describe the structure of the most common raster, palette, and transform image formats,

including the JPEG. The description of each format is supplied with instructions on how to work with such images in Matlab to give the reader the ability to conveniently implement most of the methods described in this book.

Since the majority of digital images are obtained using a digital camera, camcorder, or scanner, Chapter 3 deals with the process of digital image acquisition through an imaging sensor. Throughout the chapter, emphasis is given to those aspects of this process that are relevant to steganography. This includes the processing pipeline inside typical digital cameras and sources of noise and imperfections. Noise is especially relevant to steganography because the seemingly useless stochastic components of digital images could conceivably convey secret messages.

In Chapter 4, we delve deeper into the subject of steganography. Three basic principles for constructing steganographic methods are introduced: steganography by cover selection, cover synthesis, and cover modification. Even though the focus of this book is on data-hiding methods that embed secret messages by slightly modifying the original (cover) image, all three principles can be used to build steganographic methods in practice. This chapter also introduces basic terminology and key building blocks that form the steganographic channel – the source of cover objects, source of secret messages and secret keys, the data-hiding and data-extraction algorithms, and the physical channel itself. The physical properties of the channel are determined by the actions of the warden Eve, who can position herself to be a passive observant or someone who is actively involved with the flow of data through the channel. Discussions throughout the chapter pave the way towards the information-theoretic definition of steganographic security given in Chapter 6.

The content of Chapter 5 was chosen to motivate the reader to ask basic questions about what it means to undetectably embed secret data in an image and to illustrate various (and sometimes unexpected) difficulties one might run into when attempting to realize some intuitive hiding methods. The chapter contains examples of some early naive steganographic methods for the raster, palette, and JPEG formats, most of which use some version of the least-significant-bit (LSB) embedding method. The presentation of each method continues with critical analysis of how the steganographic method can be broken and why. The author hopes that this early exposure of specific embedding methods will make the reader better understand the need for a rather precise technical approach in the remaining chapters.

Chapter 6 introduces the central concept, which is a formal information-theoretic definition of security in steganography based on the Kullback–Leibler divergence between the distributions of cover and stego objects. This definition puts steganography on a firm mathematical ground that allows methodological development by studying security with respect to a cover model. The concept of security is further explained by showing connections between security and detection theory and by providing examples of undetectable steganographic schemes built using the principles outlined in Chapter 4. We also introduce the concept

of a distortion-limited embedder (when Alice is limited in how much she can modify the cover image) and show that some well-known watermarking methods, such as spread-spectrum watermarking and quantization index modulation, can be used to construct secure steganographic schemes. Finally, the reader is presented with an alternative complexity-theoretic definition of steganographic security even though this direction is not further pursued in this book.

Using the definition of security as a guiding philosophy, Chapter 7 introduces several design principles and intuitive strategies for building practical steganographic schemes for digital media files: (1) model-preserving steganography using statistical restoration and model-based steganography, (2) steganography by mimicking natural phenomena or processing, (3) steganalysis-aware steganography, and (4) minimal-impact steganography. The first three approaches are illustrated by describing in detail specific examples of steganographic algorithms from the literature (OutGuess, Model-Based Steganography for JPEG images, stochastic modulation, and the F5 algorithm). Minimal embedding impact steganography is discussed in Chapters 8 and 9.

Chapter 8 is devoted to matrix embedding, which is a general method for increasing security of steganographic schemes by minimizing the number of embedding changes needed to embed the secret message. The reader is first motivated by what appears a simple clever trick, which is later generalized and then reinterpreted within the language of coding theory. The introductory sections naturally lead to the highlight of this chapter – the matrix embedding theorem, which is essentially a recipe for how to turn a linear code into a steganographic embedding method using the principle of syndrome coding. Ample space is devoted to various bounds that impose fundamental limits on the performance one can achieve using matrix embedding.

The second chapter that relates to minimal-impact steganography is Chapter 9. It introduces the important topic of communication with a non-shared selection channel as well as several practical methods for communication using such channels (wet paper codes). A non-shared selection channel refers to the situation when Alice embeds her message into a selected subset of the image but does not (or cannot) share her selection with Bob. This chapter also discusses several diverse problems in steganography that lead to non-shared selection channels and can be elegantly solved using wet paper codes: adaptive steganography, perturbed quantization steganography, a new class of improved matrix embedding methods, public-key steganography, the no-shrinkage F5 algorithm, and the MMx algorithm.

While the first part of this book deals solely with design and development of steganographic methods, the next three chapters are devoted to steganalysis, which is understood as an inherent part of steganography. After all, steganography is advanced through analysis.

In Chapter 10, steganalysis is introduced as the task of discovering the presence of secret data. The discussion in this chapter is directed towards explaining

general principles common to many steganalysis techniques. The focus is on statistical attacks in which the warden reaches her decision by inspecting statistical properties of pixels. This approach to steganalysis provides connections with the abstract problem of signal detection and hypothesis testing, which in turn allows importing standard signal-detection tools and terminology, such as the receiver operating characteristic. The chapter continues with separate sections on targeted and blind steganalysis. The author lists several general strategies that one can follow to construct targeted attacks and highlights the important class of quantitative attacks, which can estimate the number of embedding changes. The section on blind steganalysis contains a list of general principles for constructing steganalysis features as well as description of several diverse applications of blind steganalyzers, including construction of targeted attacks, steganography design, multi-class steganalysis, and benchmarking. The chapter is closed with discussion of forensic steganalysis and system attacks on steganography in which the attacker relies on protocol weaknesses of a specific implementation rather than on statistical artifacts computed from the pixel values.

Chapter 11 contains examples of targeted steganalysis attacks and their experimental verifications. Experiments on real images are used to explain various issues when constructing a practical steganography detector and to give the reader a sense of how sensitive the attacks are. The chapter starts with the Sample Pairs Analysis, which is a targeted quantitative attack on LSB embedding in the spatial domain. The derivation of the method is presented in a way that makes the algorithm appear as a rather natural approach that logically follows from the strategies outlined in Chapter 10. Next, the approach is generalized by formulating it within the structural steganalysis framework. This enables several important generalizations that further improve the method's accuracy. The third attack, the Pairs Analysis, is a quantitative attack on steganographic methods that embed messages into LSBs of palette images, such as EzStego. The concept of calibration is used to construct a quantitative attack on the F5 embedding algorithm. The chapter is closed with description of targeted attacks on ± 1 embedding in the spatial domain based on the histogram characteristic function.

Chapter 12 is devoted to the topic of blind attacks, which is an approach to steganalysis based on modeling images using features and classifying cover and stego features using machine-learning tools. Starting with the JPEG domain, the features are introduced in a natural manner as statistical descriptors of DCT coefficients by modeling them using several different statistical models. The JPEG domain is also used as an example to demonstrate two options for constructing blind steganalyzers: (1) the cover-versus-all-stego approach in which a binary classifier is trained to recognize cover images and a mixture of stego images produced by a multitude of steganographic algorithms, and (2) a one-class steganalyzer trained only on cover images that classifies all images incompatible with covers as stego. The advantages and disadvantages of both approaches are discussed with reference to practical experiments. Blind steganalysis in the spatial domain is illustrated on the example of a steganalyzer whose features are

computed from image noise residuals. This steganalyzer is also used to demonstrate how much statistical detectability in practice depends on the source of cover images.

Chapter 13 discusses the most fundamental problem of steganography, which is the issue of computing the largest payload that can be securely embedded in an image. Two very different concepts are introduced – the steganographic capacity and secure payload. Steganographic capacity is the largest rate at which perfectly secure communication is possible. It is not a property of one specific steganographic scheme but rather a maximum taken over all perfectly secure schemes. In contrast, secure payload is defined as the number of bits that can be communicated at a given security level using a specific imperfect steganographic scheme. The secure payload grows only with the square root of the number of pixels in the image. This so-called square-root law is experimentally demonstrated on a specific steganographic scheme that embeds bits in the JPEG domain. The secure payload is more relevant to practitioners because all practical steganographic schemes that hide messages in real digital media are not likely to be perfectly secure and thus fall under the squre-root law.

To make this text self-contained, five appendices accompany the book. Their style and content are fully compatible with the rest of the book in the sense that the student does not need any more prerequisites than a basic knowledge of calculus and statistics. The author anticipates that students not familiar with certain topics will find it convenient to browse through the appendices and either refresh their knowledge or learn about certain topics in an elementary fashion accessible to a wide audience.

Appendix A contains the basics of descriptive statistics, including statistical moments, the moment-generating function, robust measures of central tendency and spread, asymptotic laws, and description of some key statistical distributions, such as the Bernoulli, binomial, Gaussian, multivariate Gaussian, generalized Gaussian, and generalized Cauchy distributions, Student's t-distribution, and the chi-square distribution.

As some of the chapters rely on basic knowledge of information theory, Appendix B covers selected key concepts of entropy, conditional entropy, joint entropy, mutual information, lossless compression, and KL divergence and some of its key properties, such as its relationship to hypothesis testing and Fisher information.

The theory of linear codes over finite fields is the subject of Appendix C. The reader is introduced to the basic concepts of a generator and parity-check matrix, covering radius, average distance to code, sphere-covering bound, orthogonality, dual code, systematic form of a code, cosets, and coset leaders.

Appendix D contains elements of signal detection and estimation. The author explains the Neyman–Pearson and Bayesian approach to hypothesis testing, the concepts of a receiver-operating-characteristic (ROC) curve, the deflection coefficient, and the connection between hypothesis testing and Fisher information. The appendix continues with composite hypothesis testing, the chi-square test,

and the locally most powerful detector. The topics of estimation theory covered in the appendix include the Cramer–Rao lower bound, least-square estimation, maximum-likelihood and maximum a posteriori estimation, and the Wiener filter. The appendix is closed with the Cauchy–Schwartz inequality in Hilbert spaces with inner product, which is needed for proofs of some of the propositions in this book.

Readers not familiar with support vector machines (SVMs) will find Appendix E especially useful. It starts with the formulation of a binary classification problem and introduces linear support vector machines as a classification tool. Linear SVMs are then progressively generalized to non-separable problems and then put into kernelized form as typically used in practice. The weighted form of SVMs is described as well because it is useful to achieve a trade-off between false alarms and missed detections and for drawing an ROC curve. The appendix also explains practical issues with data preprocessing and training SVMs that one needs to be aware of when using SVMs in applications, such as in blind steganalysis.

Because the focus of this book is strictly on steganography in digital signals, methods for covert communication in other objects are not covered. Instead, the author refers the reader to other publications. In particular, linguistic steganography and data-hiding aspects of some cryptographic applications are covered in [238, 239]. The topic of covert channels in natural language is also covered in [18, 25, 41, 161, 182, 227]. A comprehensive bibliography of all articles published on covert communication in linguistic structures, including watermarking applications, is maintained by Bergmair at `http://semantilog.ucam.org/biblingsteg/`. Topics dealing with steganography in Internet protocols are studied in [106, 162, 163, 165, 177, 216]. Covert timing channels and their security are covered in [26, 34, 100, 101]. The intriguing topic of steganography in Voice over IP applications, such as Skype, appears in [6, 7, 58, 147, 150, 169, 251]. Steganographic file systems [4, 170] are useful tools to thwart "rubber-hose attacks" on cryptosystems when a person is coerced to reveal encryption keys after encrypted files have been found on a computer system. A steganographic file system allows the user to plausibly deny that encrypted files reside on the disk. In-depth analysis of current steganographic software and the topics of data hiding in elements of operating systems are provided in [142]. Finally, the topics of audio steganography and steganalysis appeared in [9, 24, 118, 149, 187, 202].

Acknowledgments

I would like to acknowledge the role of several individuals who helped me commit to writing this book. First of all and foremost, I am indebted to Richard Simard for encouraging me to enter the field of steganography and for supporting research on steganography. This book would not have materialized without the constant encouragement of George Klir and Monika Fridrich. Finally, the privilege of co-authoring a book with Ingemar Cox [51] provided me with energy and motivation I would not have been able to find otherwise.

Furthermore, I am happy to acknowledge the help of my PhD students for their kind assistance that made the process of preparing the manuscript in TEX a rather pleasant experience instead of the nightmare that would for sure have followed if I had been left alone with a TEX compiler. In particular, I am immensely thankful to TEX guru Tomáš Filler for his truly significant help with formatting the manuscript, preparing the figures, and proof-reading the text, to Tomáš Pevný for contributing material for the appendix on support vector machines, and to Jan Kodovský for help with combing the citations and proofreading. I would also like to thank Ellen Tilden and my students from the ECE 562 course on Fundamentals of Steganography, Tony Nocito, Dae Kim, Zhao Liu, Zhengqing Chen, and Ran Ren, for help with sanitizing this text to make it as free of typos as possible.

Discussions with my colleagues, Andrew D. Ker, Miroslav Goljan, Andreas Westfeld, Rainer Böhme, Pierre Moulin, Neil F. Johnson, Scott Craver, Patrick Bas, Teddy Furon, and Xiaolong Li were very useful and helped me clarify some key technical issues. The encouragement I received from Mauro Barni, Deepa Kundur, Slava Voloshynovskiy, Jana Dittmann, Gaurav Sharma, and Chet Hosmer also helped with shaping the final content of the manuscript. Special thanks are due to George Normandin and Jim Moronski for their feedback and many useful discussions about imaging sensors and to Josef Sofka for providing a picture of a CCD sensor. A special acknowledgement goes to Binghamton University Art Director David Skyrca for the beautiful cover design.

Finally, I would like to thank Nicole and Kathy Fridrich for their patience and for helping me to get into the mood of sharing.