Information Hiding—A Survey

FABIEN A. P. PETITCOLAS, ROSS J. ANDERSON, AND MARKUS G. KUHN

Information-hiding techniques have recently become important in a number of application areas. Digital audio, video, and pictures are increasingly furnished with distinguishing but imperceptible marks, which may contain a hidden copyright notice or serial number or even help to prevent unauthorized copying directly. Military communications systems make increasing use of traffic security techniques which, rather than merely concealing the content of a message using encryption, seek to conceal its sender, its receiver, or its very existence. Similar techniques are used in some mobile phone systems and schemes proposed for digital elections. Criminals try to use whatever traffic security properties are provided intentionally or otherwise in the available communications systems, and police forces try to restrict their use. However, many of the techniques proposed in this young and rapidly evolving field can trace their history back to antiquity, and many of them are surprisingly easy to circumvent. In this article, we try to give an overview of the field, of what we know, what works, what does not, and what are the interesting topics for research.

Keywords— *Copyright marking, information hiding, steganography.*

I. INTRODUCTION

It is often thought that communications may be secured by encrypting the traffic, but this has rarely been adequate in practice. Æneas the Tactician, and other classical writers, concentrated on methods for hiding messages rather than for enciphering them [1]; although modern cryptographic techniques started to develop during the Renaissance, we find in 1641 that Wilkins still preferred hiding over ciphering [2, ch. IX, p. 67] because it arouses less suspicion. This preference persists in many operational contexts to this day. For example, an encrypted e-mail message between a known drug dealer and somebody not yet under suspicion, or between an employee of a defence contractor and the embassy of a hostile power, has obvious implications.

So the study of communications security includes not just encryption but also traffic security, whose essence lies in hiding information. This discipline includes such technologies as: spread spectrum radio, which is widely used in tactical military systems to prevent transmitters being located; temporary mobile subscriber identifiers, used in digital phones to provide users with some measure of location privacy; and anonymous remailers, which conceal the identity of the sender of an e-mail message [3].

Manuscript received February 2, 1998; revised December 1, 1998. The work of F. A. P. Petitcolas was supported by Intel Corporation under the grant "Robustness of Information Hiding Systems." The work of M. G. Kuhn was supported by the European Commission under a Marie Curie Training Grant.

The authors are with the University of Cambridge Computer Laboratory, Security Group, Cambridge CB2 3QG U.K.

Publisher Item Identifier S 0018-9219(99)04946-4.

Table 1
Number of Publications on Digital Watermarking During the Past Few Years According to INSPEC, January 1999 (Courtesy of J.-L. Dugelay [5])

Year	1992	1993	1994	1995	1996	1997	1998
Publications	2	2	4	13	29	64	103

An important subdiscipline of information hiding is steganography. While cryptography is about protecting the content of messages, steganography is about concealing their very existence. It comes from Greek roots ($\sigma\tau\varepsilon\gamma\alpha\nu\acute{o}$-$\varsigma, \gamma\rho\acute{\alpha}\varphi$-$\varepsilon\iota\nu$), literally means "covered writing" [151], and it is usually interpreted to mean hiding information in other information. Examples include sending a message to a spy by marking certain letters in a newspaper using invisible ink, and adding subperceptible echo at certain places in an audio recording.

Until recently, information-hiding techniques received much less attention from the research community and from industry than cryptography, but this is changing rapidly (Table 1), and the first academic conference on the subject was organized in 1996 [4]. The main driving force is concern over copyright; as audio, video, and other works become available in digital form, the ease with which perfect copies can be made may lead to large-scale unauthorized copying, and this is of great concern to the music, film, book, and software publishing industries. There has been significant recent research into digital "watermarks" (hidden copyright messages) and "fingerprints" (hidden serial numbers); the idea is that the latter can help to identify copyright violators, and the former to prosecute them.

In another development, the DVD consortium has called for proposals for a copyright marking scheme to enforce serial copy management. The idea is that DVD players available to consumers would allow unlimited copying of home videos and time-shifted viewing of TV programs but could not easily be abused for commercial piracy. The proposal is that home videos would be unmarked, TV broadcasts marked "copy once only," and commercial videos marked "never copy"; compliant consumer equipment would act on these marks in the obvious way [6], [7].

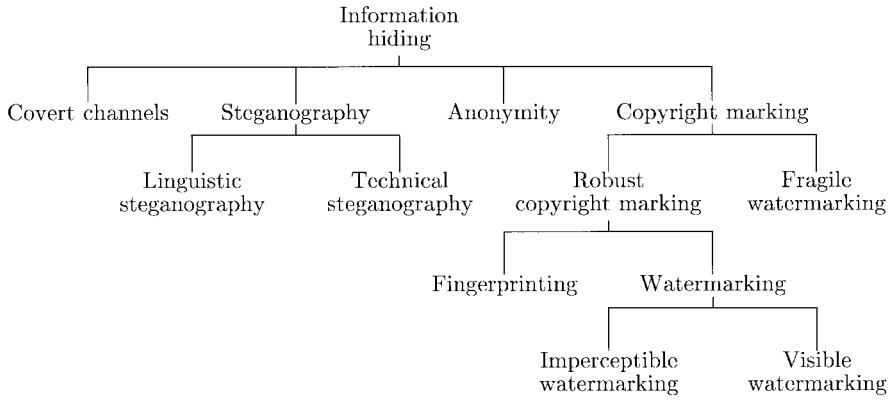


Fig. 1. A classification of information-hiding techniques based on [10]. Many of the ancient systems presented in Sections III-A and III-B are a form of "technical steganography" (in the sense that messages are hidden physically) and most of the recent examples given in this paper address "linguistic steganography" and "copyright marking."

There are a number of other applications driving interest in the subject of information hiding (Fig. 1).

- Military and intelligence agencies require unobtrusive communications. Even if the content is encrypted, the detection of a signal on a modern battlefield may lead rapidly to an attack on the signaller. For this reason, military communications use techniques such as spread spectrum modulation or meteor scatter transmission to make signals hard for the enemy to detect or jam.
- Criminals also place great value on unobtrusive communications. Their preferred technologies include prepaid mobile phones, mobile phones which have been modified to change their identity frequently, and hacked corporate switchboards through which calls can be rerouted.
- Law enforcement and counter intelligence agencies are interested in understanding these technologies and their weaknesses, so as to detect and trace hidden messages.
- Recent attempts by some governments to limit online free speech and the civilian use of cryptography have spurred people concerned about liberties to develop techniques for anonymous communications on the Internet, including anonymous remailers and Web proxies.
- Schemes for digital elections and digital cash make use of anonymous communication techniques.
- Marketeers use e-mail forgery techniques to send out huge numbers of unsolicited messages while avoiding responses from angry users.

We will mention some more applications later. For the time being, we should note that while the ethical positions of the players in the cryptographic game are often thought to be clear cut (the "good" guys wish to keep their communications private while the "bad" eavesdropper wants to listen in), the situation is much less clear when it comes to hiding information. Legitimate users of the Internet may need anonymous communications to contact abuse helplines or vote privately in online elections [8]; but one may not want to provide general anonymous communication mechanisms that facilitate attacks by people who maliciously overload the communication facilities. Industry may need tools to hide copyright marks invisibly in media objects, yet these tools can be abused by spies to pass on secrets hidden in inconspicuous data over public networks. Finally, there are a number of noncompetitive uses of the technology, such as marking audio tracks with purchasing information so that someone listening to a piece of music on his car radio could simply press a button to order the CD.

The rest of this paper is organized as follows. First, we will clarify the terminology used for information hiding, including steganography, digital watermarking, and fingerprinting. Secondly, we will describe a wide range of techniques that have been used in a number of applications, both ancient and modern, which we will try to juxtapose in such a way that the common features become evident. Then, we will describe a number of attacks against these techniques. Finally, we will try to formulate general definitions and principles. Moving through the subject from practice to theory may be the reverse of the usual order of presentation, but it appears appropriate to a discipline in which rapid strides are being made constantly, and where general theories are still very tentative.

II. 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).

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 stego-text 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).

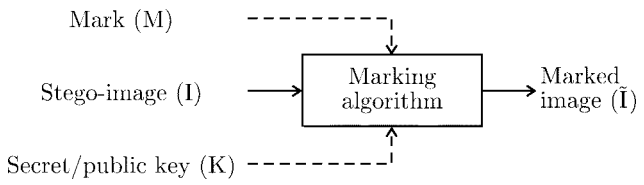


Fig. 2. Generic digital watermark embedding scheme. The mark M can be either a fingerprint or a watermark.

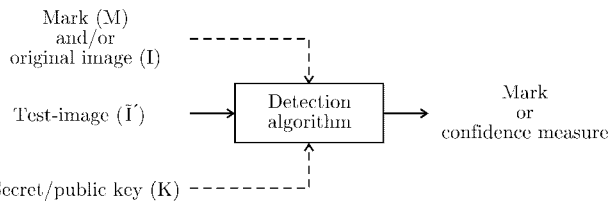


Fig. 3. Generic digital watermark recovery scheme.

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 [11], 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 [12]. 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 watermarks[1] 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 [13].

Authors also make the distinction between various types of robust marks. 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.

Fig. 2 illustrates the generic embedding process. Given an image I, a mark M, and a key K (usually the seed of a random number generator), the embedding process can be defined as a mapping of the form: $I \times K \times M \rightarrow \tilde{I}$ and is common to all watermarking methods.

The generic detection process is depicted in Fig. 3. Its output is either the recovered mark M or some kind of confidence measure indicating how likely it is for a given mark at the input to be present in the image $\tilde{I}'$ under inspection.

[1]Fragile watermarks have also wrongly been referred to as "signature," leading to confusion with digital signatures used in cryptography.

There are several types of robust copyright marking systems. They are defined by their inputs and outputs.

- *Private marking* systems require at least the original image. *Type I* systems, extract the mark M from the possibly distorted image $\tilde{I}'$ and use the original image as a hint to find where the mark could be in $\tilde{I}'$. *Type II* systems (e.g., [14]–[16]) also require a copy of the embedded mark for extraction and just yield a "yes" or "no" answer to the question: does $\tilde{I}'$ contain the mark M? $(\tilde{I}' \times I \times K \times M \rightarrow \{0,1\})$. One might expect that this kind of scheme will be more robust than the others since it conveys very little information and requires access to secret material [13]. *Semiprivate marking* does not use the original image for detection $(\tilde{I}' \times K \times M \rightarrow \{0,1\})$ but answers the same question.

  The main uses of private and semiprivate marking seem to be evidence in court to prove ownership and copy control in applications such as DVD where the reader needs to know whether it is allowed to play the content or not. Many of the currently proposed schemes fall in this category [17]–[23].
- *Public marking* (also referred to as blind marking) remains the most challenging problem since it requires neither the secret original I nor the embedded mark M. Indeed, such systems really extract n bits of information (the mark) from the marked image: $\tilde{I}' \times K \rightarrow M$ [24]–[28]. Public marks have much more applications than the others and we will focus our benchmark on these systems. Indeed, the embedding algorithms used in public systems can usually be used in private ones, improving robustness at the same time.
- There is also *asymmetric marking* (or public key marking) which should have the property that any user can read the mark, without being able to remove it.

In the rest of the paper, "watermark" will refer to "digital watermark" unless said otherwise.

III. Steganographic Techniques

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.

A. Security Through Obscurity

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. In his 400-page book *Schola Steganographica* [29],

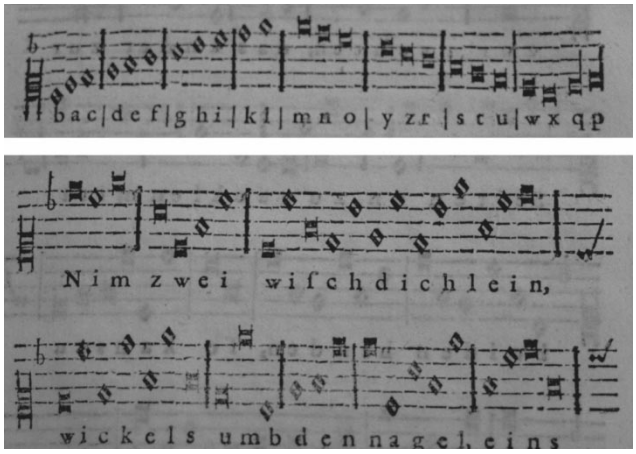


Fig. 4. Hiding information into music scores: Schott simply maps the letters of the alphabet to the notes. Clearly, one should not try to play the music [29, p. 322].

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 [30]. 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].

A very widely used method is the acrostic. In his book *The Codebreakers* [31], Kahn explains how a monk wrote a book and put his lover's name in the first letters of successive chapters. He also tells of prisoners of war who hid messages in letters home using the dots and dashes on $i, j, t,$ and f to spell out a hidden text in Morse code. These "semagrams" concealed messages but have an inherent problem, that the cover text tends to be laborious to construct and often sounds odd enough to alert the censor. During both World Wars, censors intercepted many such messages. A famous one, from World War I, was a cablegram saying "Father is dead," which the censor modified into "Father is deceased." The reply was a giveaway: "Is Father dead or deceased?" [31, pp. 515–516].

Although steganography is different from cryptography, we can borrow many of the techniques and much practical wisdom from the latter, more thoroughly researched discipline. In 1883, Kerckhoffs enunciated the first principles of cryptographic engineering, in which he advises that we assume the method used to encipher data is known to the opponent, so security must lie only in the choice of key[2] [32]. The history of cryptology since then has repeatedly shown the folly of "security-by-obscurity"—the assumption that the enemy will remain ignorant of the system in use.

Applying this wisdom, we obtain a tentative definition of a secure stego-system: one where an opponent who understands the system, but does not know the key, can obtain no evidence (or even grounds for suspicion) that a communication has taken place. In other words, no information about the embedded text can be obtained from knowledge of the stego (and perhaps also cover) texts. We will revisit this definition later, to take account of robustness and other issues; but it will remain a central principle that steganographic processes intended for wide use should be published, just like commercial cryptographic algorithms and protocols. This teaching of Kerckhoffs holds with particular force for marking techniques intended for use in evidence, which implies their disclosure in court [33].

That any of the above "security-by-obscurity" systems ever worked was a matter of luck. Yet many steganographic systems available today just embed the "hidden" data in the least significant bits of an audio or video file—which is trivial for a capable opponent to detect and remove.

B. Camouflage

The situation may be improved by intelligent use of camouflage. Even if the method is known in principle, making the hidden data expensive to look for can be beneficial, especially where there is a large amount of cover traffic.

Since the early days of architecture, artists have understood that works of sculpture or painting appear different from certain angles and established rules for perspective and anamorphosis [34]. Through the sixteenth and seventeenth centuries, anamorphic images supplied an ideal means of camouflaging dangerous political statements and heretical ideas [35]. A masterpiece of hidden anamorphic imagery—the *Vexierbild*—was created in the 1530's by Shö, a Nürnberg engraver, pupil of Dürer (1471–1528); when one looks at it normally one sees a strange landscape, but looking from the side reveals portraits of famous kings.

In his *Histories* [36], Herodotus (c. 486–425 B.C.) tells how around 440 B.C. Histiæus shaved the head of his most trusted slave and tattooed it with a message which disappeared after the hair had regrown. The purpose was to instigate a revolt against the Persians. Astonishingly, the method was still used by some German spies at the beginning of the twentieth century [37]. Herodotus also tells how Demeratus, a Greek at the Persian court, warned Sparta of an imminent invasion by Xerxes: he removed the wax from a writing tablet, wrote his message on the wood underneath, and then covered the message with wax. The

[2] Il faut qu'il n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi [32, p. 12].

tablet looked exactly like a blank one (it almost fooled the recipient as well as the customs men).

A large number of techniques were invented or reported by Æneas the Tactician [1], including letters hidden in messengers' soles or women's earrings, text written on wood tablets and then whitewashed, and notes carried by pigeons. The centerpiece is a scheme for winding thread through 24 holes bored in an astragal: each hole represents a letter and a word is represented by passing the thread through the corresponding letters. He also proposed hiding text by making very small holes above or below letters or by changing the heights of letter-strokes in a cover text. These dots were masked by the contrast between the black letters and the white paper. This technique was still in use during the seventeenth century, but it was improved by Wilkins who used invisible ink to print very small dots instead of making holes [2] and was reused by German spies during both World Wars [31, p. 83]. A modern adaptation of this technique is still in use for document security [38].

Invisible inks were used extensively. They were originally made of available organic substances (such as milk or urine) or "salt armoniack dissolved in water" [2, ch. V, pp. 37–47] and developed with heat; progress in chemistry helped to create more sophisticated combinations of ink and developer by World War I, but the technology fell into disuse with the invention of "universal developers" which could determine which parts of a piece of paper had been wetted from the effects on the surfaces of the fibers [31, pp. 523–525]. Nowadays, in the field of currency security, special inks or materials with particular structure (such as fluorescent dyes or DNA) are used to write a hidden message on bank notes or other secure documents. These materials provide a unique response to some particular excitation such as a reagent or laser light at a particular frequency [39].

By 1860 the basic problems of making tiny images had been solved [40]. In 1857, Brewster suggested hiding secret messages "in spaces not larger than a full stop or small dot of ink" [41]. During the Franco–Prussian War of 1870–1871, while Paris was besieged, messages on microfilm were sent out by pigeon post [42], [43]. During the Russo–Japanese War of 1905, microscopic images were hidden in ears, nostrils, and under finger nails [40]. By World War I, messages to and from spies were reduced to microdots by several stages of photographic reduction and then stuck on top of printed periods or commas in innocuous cover material such as magazines [37], [44].

The digital equivalent of these camouflage techniques is the use of masking algorithms [16], [26], [45]–[47]. Like most source-coding techniques (e.g., [48]), these rely on the properties of the human perceptual system. Audio masking, for instance, is a phenomenon in which one sound interferes with our perception of another sound [49]. Frequency masking occurs when two tones which are close in frequency are played at the same time: the louder tone will mask the quieter one. Temporal masking occurs when a low-level signal is played immediately before or after a stronger one; after a loud sound stops, it takes a little while before we can hear a weak tone at a nearby frequency.

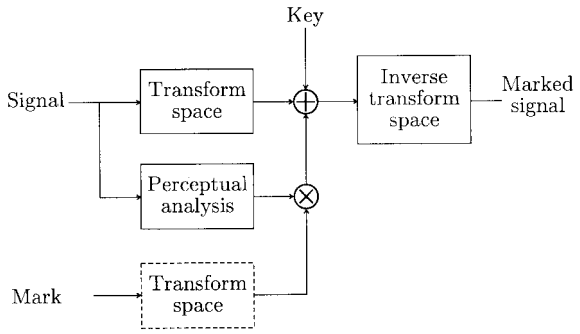


Fig. 5. A typical use of masking and transform space for digital watermarking and fingerprinting. The signal can be an image or an audio signal. The perceptual analysis is based on the properties of the human visual or auditory systems, respectively. $\oplus$ corresponds to the embedding algorithm and $\otimes$ to the weighting of the mark by the information provided by the perceptual model.

Because these effects are used in compression standards such as MPEG [50], many systems shape the embedded data to emphasize it in the perceptually most significant components of the data so it will survive compression [26], [46] (Fig. 5). This idea is also applied in buried data channels where the regular channels of an audio CD contain other embedded sound channels [51]; here, an optimized noise shaper is used to reduce to minimize the effect of the embedded signal on the quality of the cover music.

For more details about the use of perceptual models in digital watermarking, the reader is referred to [52] and [53].

C. Hiding the Location of the Embedded Information

In a security protocol developed in ancient China, the sender and the receiver had copies of a paper mask with a number of holes cut at random locations. The sender would place his mask over a sheet of paper, write the secret message into the holes, remove the mask and then compose a cover message incorporating the code ideograms. The receiver could read the secret message at once by placing his mask over the resulting letter. In the early sixteenth century, Cardan (1501–1576), an Italian mathematician, reinvented this method which is now known as the Cardan grille. It appears to have been reinvented again in 1992 by a British bank, which recommended that its customers conceal the personal information number used with their cash machine card using a similar system. In this case, a poor implementation made the system weak [54].

A variant on this theme is to mark an object by the presence of errors or stylistic features at predetermined points in the cover material. An early example was a technique used by Bacon (1561–1626) in his *biliterarie* alphabet [55, p. 266], which seems to be linked to the controversy as to whether he wrote the works attributed to Shakespeare [56]. In this method, each letter is encoded in a 5-bit binary code and embedded in the cover text by printing the letters in either normal or italic fonts. The variability of sixteenth-century typography acted as camouflage.

Further examples come from the world of mathematical tables. Publishers of logarithm tables and astronomical ephemerides in the seventeenth and eighteenth century used

to introduce errors deliberately in the least significant digits (e.g., [57]). To this day, database and mailing list vendors insert bogus entries in order to identify customers who try to resell their products.

In an electronic publishing pilot project, copyright messages and serial numbers have been hidden in the line spacing and other format features of documents (e.g., [58]). It was found that shifting text lines up or down by one-three-hundredth of an inch to encode zeros and ones was robust against multigeneration photocopying and could not be noticed by most people.

However, the main application area of current copyright marking proposals lies in digital representations of analog objects such as audio, still pictures, video, and multimedia generally. Here there is considerable scope for embedding data by introducing various kinds of error. As we noted above, many writers have proposed embedding the data in the least significant bits [22], [59]. An obviously better technique, which has occurred independently to many writers, is to embed the data into the least significant bits of pseudorandomly chosen pixels or sound samples [60], [61]. In this way, the key for the pseudorandom sequence generator becomes the stego-key for the system and Kerckhoffs' principle is observed.

Many implementation details need some care. For example, one might not wish to disturb a pixel in a large expanse of flat color, or lying on a sharp edge; for this reason, a prototype digital camera designed to enable spies to hide encrypted reports in snapshots used a pseudorandom sequence generator to select candidate pixels for embedding bits of cipher text and then rejected those candidates where the local variance of luminosity was either too high or too low.

One scheme that uses bit tweaking in a novel way is Chameleon. Ideally, all distributed copies of a copyright work should be fingerprinted, but in applications such as pay TV or CD, the broadcast or mass production nature of the medium appears to preclude this. Chameleon allows a single cipher text to be broadcast while subscribers are given slightly different deciphering keys, which produce slightly different plain texts. The system can be tuned so that the deciphered signal is only marked in a sparse subset of its least significant bits, and this may produce an acceptably low level of distortion for digital audio. The precise mechanism involves modifying a stream cipher to reduce the diffusion of part of its key material [62].

Systems that involve bit twiddling have a common vulnerability, that even very simple digital filtering operations will disturb the value of many of the least significant bits of a digital object. This leads us to consider ways in which bit tweaking can be made robust against filtering.

D. Spreading the Hidden Information

The obvious solution is to consider filtering operations as the introduction of noise in the embedded data channel [63] and to use suitable coding techniques to exploit the residual bandwidth. The simplest is the repetition code—one simply embeds a bit enough times in the cover object so that evidence of it will survive the filter. This is inefficient in coding theoretic terms but can be simple and robust in some applications.

Another way to spread the information is to embed it into the statistics of the luminance of the pixels, such as [64] and [65]. Patchwork [64], for instance, uses a pseudorandom generator to select n pairs of pixels and slightly increases or decreases their luminosity contrast. Thus the contrast of this set is increased without any change in the average luminosity of the image. With suitable parameters, Patchwork even survives compression using JPEG. However, it embeds only 1 bit of information. To embed more, one can first split the image into pieces and then apply the embedding to each of them [27], [66].

These statistical methods give a kind of primitive spread spectrum modulation. General spread spectrum systems encode data in the choice of a binary sequence that appears like noise to an outsider but which a legitimate receiver, furnished with an appropriate key, can recognize. Spread spectrum radio techniques have been developed for military applications since the mid 1940's because of their antijamming and low-probability-of-intercept properties [67]–[69]; they allow the reception of radio signals that are over 100 times weaker than the atmospheric background noise.

Tirkel *et al.* were the first to note that spread spectrum techniques could be applied to digital watermarking [70], and later a number of researchers have developed steganographic techniques based on spread spectrum ideas which take advantage of the large bandwidth of the cover medium by matching the narrow bandwidth of the embedded data to it (e.g., [63], [71]–[73]).

In [15], Cox *et al.* present an image watermarking method in which the mark is embedded in the n most perceptually significant frequency components $V = \{v_i\}_{i=1}^{n}$ of an image's discrete cosine transform to provide greater robustness to JPEG compression. The watermark is a sequence of real numbers $W = \{w_i\}_{i=1}^{n}$ drawn from a Gaussian distribution and is inserted using the formula $\tilde{v}_i = v_i(1 + \alpha w_i)$. If I is the original image and $\tilde{I}$ the watermarked image, that is the image whose main components have been modified, the presence of the watermark is verified by extracting the main components of I and those with same index from $\tilde{I}$ and inverting the embedding formula to give a possibly modified watermark W'. The watermark is said to be present in $\tilde{I}$ if the ratio $W \cdot W'/\sqrt{W' \cdot W'}$ is greater than a given threshold.

The authors claim that $O(\sqrt{n/\ln n})$ similar watermarks must be added before they destroy the original mark. This method is very robust against rescaling, JPEG compression, dithering, clipping, printing/scanning, and collusion attacks. However, it has some drawbacks. Most seriously, the original image is needed to check for the presence of a watermark.

The second problem is the low information rate. Like Patchwork, this scheme hides a single bit and is thus suitable for watermarking rather than fingerprinting or steganographic communication. The information rate of such schemes can again be improved by placing separate marks in the image, but at a cost of reduced robustness.

Information-hiding schemes that operate in a transform space are increasingly common, as this can aid robustness against compression, other common filtering operations, and noise. Actually, one can observe that the use of a particular transform gives good results against compression algorithms based on the same transform.

Some schemes operate directly on compressed objects (e.g., [72]). Some, steganographic tools, for example, hide information in `gif` [74] files by swapping the colors of selected pixels for colors that are adjacent in the current palette [75]. Another example is MP3Stego [76], which hides information in MPEG Audio Layer III bit streams [50] during the compression process. However, most schemes operate directly on the components of some transform of the cover object like discrete cosine transform [15]–[17], [77]–[79], wavelet transforms [16], [80], and the discrete Fourier transform [46], [81].

A novel transform coding technique is echo hiding [82], which relies on the fact that we cannot perceive short echoes (of the order of a millisecond). It embeds data into a cover audio signal by introducing two types of short echo with different delays to encode zeros and ones. These bits are encoded at locations separated by spaces of pseudorandom length. The cepstral transform [83] is used to manipulate the echo signals.

E. Techniques Specific to the Environment

Echo hiding leads naturally to the broader topic of information-hiding techniques that exploit features of a particular application environment. One technology that is emerging from the military world is meteor burst communication, which uses the transient radio paths provided by ionized trails of meteors entering the atmosphere to send data packets between a mobile station and a base [84]. The transient nature of these paths makes it hard for an enemy to locate mobiles using radio direction finding, and so meteor burst is used in some military networks.

More familiar application-specific information-hiding and marking technologies are found in the world of security printing. Watermarks in paper are a very old anticounterfeiting technique (Fig. 6); more recent innovations include special ultraviolet (UV) fluorescent inks used in printing travellers' cheques. As the lamps used in photocopiers have a high UV content, it can be arranged that photocopies come out overprinted with "void" in large letters. Inks may also be reactive; one of the authors has experience of travellers' cheques coming out "void" after exposure to perspiration in a money belt. Recent developments address the problem of counterfeiting with scanners and printers whose capabilities have improved dramatically over the last few years [85].

Many other techniques are used. For a survey of optically variable devices, such as diffraction products and thin film interference coatings, see [86] and [87]; the design of the U.S. currency is described in [88] and [89], and the security features of the Dutch passport in [90]. Such products tend to combine overt marks that are expensive to reproduce (holograms, kinegrams, intaglios, and optically variable inks) with tamper-evidence features (such as laminates and reactive inks) and secondary features whose presence may not be obvious (such as microprinting, diffraction effects visible only under special illumination, and alias band structures—dithering effects that normal scanners cannot capture) [91], [92]. In a more recent application called subgraving, variable information (such as a serial number) is printed on top of a uniform offset background. The printed area is then exposed to an excimer laser: this removes the offset ink everywhere but underneath the toner. Fraudulent removal of the toner by a solvent reveals the hidden ink [93].

Fig. 6. Monograms figuring TGE RG (Thomas Goodrich Eliensis—Bishop of Ely, England—and Remy/Remigius Guedon, the paper maker). One of the oldest watermarks found in the Cambridge area (c.1550). At that time, watermarks were mainly used to identify the mill producing the paper—a means of guaranteeing quality. (Courtesy of Dr. E. Leedham-Green, Cambridge University Archives. Reproduction technique: beta radiography.)

Increasingly, features are incorporated that are designed to be verified by machines rather than humans. Marks can be embedded in the magnetic strips of bank cards, giving each card a unique serial number that is hard to reproduce [94]; they are used in phone cards too in some countries. Magnetic fibers can be embedded randomly in paper or cardboard, giving each copy of a document a unique fingerprint.

The importance of these technologies is not limited to protecting currency and securities. Forgery of drugs, vehicle spares, computer software, and other branded products is said to have cost over $24 billion in 1995, and to have directly caused over 100 deaths worldwide [95]. Security printing techniques are a significant control measure, although many fielded sealing products could be much better designed given basic attention to simple issues such as choice of pressure-sensitive adhesives and nonstandard

materials [96]. Fashion designers are also concerned that their product might be copied and wish to find techniques to enable easy detection of counterfeit clothes or bags. As a greater percentage of the gross world product comes in the form of digital objects, the digital marking techniques described here may acquire more economic significance.

Also important are covert channels: communication paths that were neither designed nor intended to transfer information at all. Common examples include timing variations and error messages in communication protocols and operating system call interfaces [97], [98]. Covert channels are of particular concern in the design and evaluation of mandatory access control security concepts, where the operating system attempts to restrict the flow of information between processes in order to protect the user from computer viruses and Trojan horse software that transmits secrets to third parties without authorization.

The electromagnetic radiation produced by computers forms another covert channel. It not only interferes with reception on nearby radio receivers but can also convey information. For instance, the video signal emitted by CRT or liquid-crystal displays can be reconstructed using a simple modified TV set at several hundred meters distance [99]. Many military organizations use especially shielded "Tempest" certified equipment to process classified information, in order to eliminate the risk of losing secrets via compromising emanations [100].

We have shown in [101] how software can hide information in video screen content in a form that is invisible to the user but that can easily be reconstructed with modified TV receivers. More sophisticated ways of broadcasting information covertly from PC software use spread spectrum techniques to embed information in the video signal or CPU bus activity.

It is possible to write a virus that searches a computer's hard disk for crypto-key material or other secrets and proceeds to radiate them covertly. The same techniques could also be used in software copyright protection: software could transmit its license serial number while in use, and software trade associations could send detector vans around business districts and other neighborhoods where piracy is suspected—just like the "TV detector vans" used in countries with a mandatory TV license fee. If multiple signals are then received simultaneously with the same serial number but with spreading sequences at different phases, this proves that software purchased under a single license is being used concurrently on different computers and can provide the evidence to obtain a search warrant.

IV. Limitations of Some Information-Hiding Systems

A number of broad claims have been made about the "robustness" of various digital watermarking or fingerprinting methods. Unfortunately, the robustness criteria and the sample pictures used to demonstrate it vary from one system to the other, and recent attacks [102]–[106] show that the robustness criteria used so far are often inadequate. JPEG compression, additive Gaussian noise, low-pass filtering, rescaling, and cropping have been addressed in most of the literature but specific distortions such as rotation have often been ignored [81], [107]. In some cases the watermark is simply said to be "robust against common signal processing algorithms and geometric distortions when used on some standard images." This motivated the introduction of a fair benchmark for digital image watermarking in [108].

Similarly, various steganographic systems have shown serious limitations [109].

Craver *et al.* [110] identify at least three kinds of attacks: robustness attacks, which aim to diminish or remove the presence of a digital watermark; presentation attacks, which modify the content such that the detector cannot find the watermark anymore [e.g., the Mosaic attack (see Section IV-C)]; and the interpretation attacks, whereby an attacker can devise a situation which prevents assertion of ownership. The separation between these groups is not always very clear though; for instance, StirMark (see Section IV-B1) both diminishes the watermark and distorts the content to fool the detector.

As examples of these, we present in this section several attacks which reveal significant limitations of various marking systems. We will develop a general attack based on simple signal processing, plus specialized techniques for some particular schemes, and show that even if a copyright marking system were robust against signal processing, bad engineering can provide other avenues of attacks.

A. Basic Attack

Most simple spread spectrum-based techniques and some simple image stego software are subject to some kind of jitter attack [104]. Indeed, although spread spectrum signals are very robust to amplitude distortion and to noise addition, they do not survive timing errors; synchronization of the chip signal is very important and simple systems fail to recover this synchronization properly. There are more subtle distortions that can be applied. For instance, in [111], Hamdy *et al.* present a way to increase or decrease the length of a musical performance without changing its pitch; this was developed to enable radio broadcasters to slightly adjust the playing time of a musical track. As such tools become widely available, attacks involving sound manipulation will become easy.

B. Robustness Attacks

1) StirMark: After evaluating some watermarking software, it became clear to us that although most schemes could survive basic manipulations—that is, manipulations that can be done easily with standard tools, such as rotation, shearing, resampling, resizing, and lossy compression—they would not cope with combinations of them or with random geometric distortions. This motivated the design of StirMark [104].

StirMark is a generic tool for basic robustness testing of image watermarking algorithms and has been freely

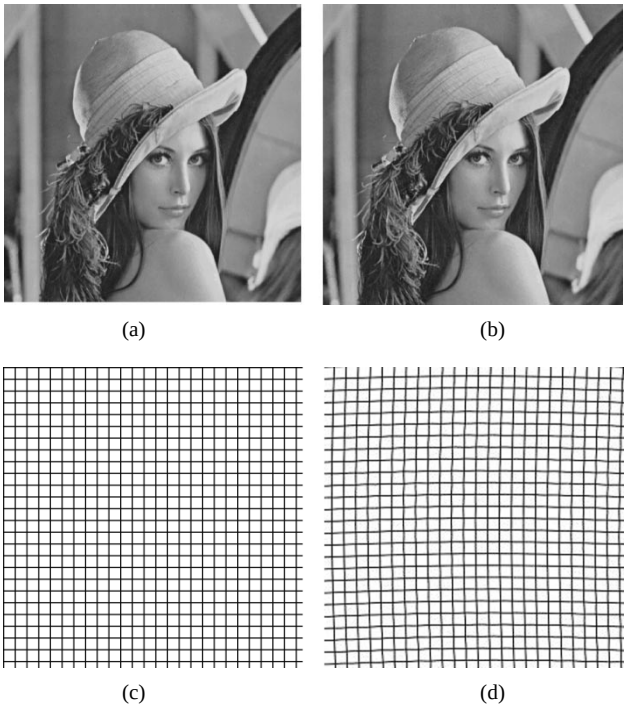


Fig. 7. When applied to images, the distortions introduced by StirMark are almost unnoticeable. "Lena" (a) before and (b) after StirMark with default parameters. (c), (d) For comparison, the same distortions have been applied to a grid.

available since November 1997.[3] It applies a minor unnoticeable geometric distortion: the image is slightly stretched, sheared, shifted, bent and rotated by an unnoticeable random amount. A slight random low-frequency deviation, which is greatest at the center of the picture, is applied to each pixel. A higher frequency displacement of the form $\lambda \sin(\omega_x x) \sin(\omega_y y) + n(x, y)$—where $n(x, y)$ is a random number—is also added. Finally, a transfer function that introduces a small and smoothly distributed error into all sample values is applied. This emulates the small nonlinear analogue/digital converter imperfections typically found in scanners and display devices. Resampling uses the approximating quadratic B-spline algorithm [112]. An example of these distortions is given in Fig. 7.

StirMark can also perform a default series of tests which serve as a benchmark for image watermarking [108]. Digital watermarking remains a largely untested field and very few authors have published extensive tests on their systems (e.g., [113]). A benchmark is needed to highlight promising areas of research by showing which techniques work better than others.

One might try to increase the robustness of a watermarking system by trying to foresee the possible transforms used by pirates; one might then use techniques such as embedding multiple versions of the mark under suitable inverse transforms; for instance, O'Ruanaidh *et al.* [81] suggest using the Fourier–Mellin transform.

However, the general lesson from this attack is that given a target marking scheme, one can invent a distortion (or a combination of distortions) that will prevent detection of the watermark while leaving the perceptual value of the previously watermarked object undiminished. We are not limited in this process to the distortions produced by common analog equipment or usually applied by end users with common image processing software. Moreover, the quality requirements of pirates are often lower than those of content owners who have to decide how much quality degradation to tolerate in return for extra protection offered by embedding a stronger signal. It is an open question whether there is any digital watermarking scheme for which a chosen distortion attack cannot be found.

2) Attack on Echo Hiding: As mentioned above, echo hiding encodes zeros and ones by adding echo signals distinguished by two different values for their delay τ and their relative amplitude α to cover an audio signal. The delays are chosen between 0.5 and 2 ms, and the relative amplitude is around 0.8 [82]. According to its creators, decoding involves detecting the initial delay and the autocorrelation of the cepstrum of the encoded signal is used for this purpose. However, the same technique can be used for an attack.

The "obvious" attack on this scheme is to detect the echo and then remove it by simply inverting the convolution formula; the problem is to detect the echo without knowledge of either the original object or the echo parameters. This is known as "blind echo cancellation" in the signal processing literature and is known to be a hard problem in general.

We tried several methods to remove the echo. Frequency invariant filtering [114], [115] was not very successful. Instead we used a combination of cepstrum analysis and "brute force" search.

The underlying idea of cepstrum analysis is presented in [83]. Suppose that we are given a signal $y(t)$ which contains a simple single echo, i.e., $y(t) = x(t) + \alpha x(t - \tau)$. If Φ_{xx} denotes the power spectrum of x then $\Phi_{yy}(f) = \Phi_{xx}(f)(1+2\alpha\cos(2\pi f\tau)+\alpha^2)$ whose logarithm $\log \Phi_{yy}(f) \approx \log \Phi_{xx}(f)+2\alpha\cos(2\pi f\tau)$. Taking its power spectrum raises its "quefrency" τ, that is, the frequency of $\cos(2\pi\tau f)$ as a function of f. The autocovariance of this later function emphasizes the peak that appears at "quefrency" τ.

We need a method to detect the echo delay τ in a signal. For this, we used a slightly modified version of the cepstrum: $C \circ \Phi \circ \ln \circ \Phi$, where C is the autocovariance function $(C(x) = E((x - \overline{x})(x - \overline{x})^*))$, Φ the power spectrum density function, and $\circ$ the composition operator. Experiments on random signals as well as on music show that this method returns quite accurate estimators of the delay when an artificial echo has been added to the signal. In the detection function we only consider echo delays between 0.5 and 3 ms (below 0.5 ms the function does not work properly and above 3 ms the echo becomes too audible).

Our first attack was to remove an echo with random relative amplitude, expecting that this would introduce enough modification in the signal to prevent watermark recovery. Since echo hiding gives best results for α greater

[3] For more information see http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/.

than 0.7, we could use $\hat{\alpha}$—an estimator of α—drawn from, say, a normal distribution centered on 0.8. It was not really successful, so our next attack was to iterate: we reapplied the detection function and varied $\hat{\alpha}$ to minimize the residual echo. We could obtain successively better estimates of the echo parameters and then remove this echo. When the detection function cannot detect any more echo, we have found the correct value of $\hat{\alpha}$ (as this gives the lowest output value of the detection function).

3) Other Generic Attacks: Some generic attacks attempt to estimate the watermark and then remove it. Langelaar *et al.* [105], for instance, present an attack on white spread spectrum watermarks. They try different methods to model the original image and apply this model to the watermarked image $\tilde{I} = I + W$ to separate it into two components: an estimated image $\hat{I}$ and an estimated watermark $\hat{W}$, such that the watermark W does not appear anymore in $\hat{I}$, giving $\rho(\hat{I}, W) \approx 0$. The authors show that a 3×3 median filter gives the best results. However, an amplified version of the estimated watermark needs to be subtracted because the low-frequency components of the watermark cannot be estimated accurately, leading to a positive contribution of the low frequencies and a negative contribution of the high frequencies to the correlation. Only a choice of good amplification parameters can zero the correlation.

In some cases, the image to be marked has certain features that help a malicious attacker gain information about the mark itself. An example of such features is where a picture, such as a cartoon, has only a small number of distinct colors, giving sharp peaks in the color histogram. These are split by some marking algorithms. The twin peaks attack, suggested by Maes [103], takes advantage of this to recover and remove marks. In the case of grayscale images, a simple example of digital watermarking based on spread spectrum ideas is to add or subtract randomly a fixed value d to each pixel value. So each pixel's value has a 50% chance of being increased or decreased. Let n_k be the number of pixels with gray value k and suppose that for a particular gray value k_0 the dth neighboring colors do not occur, so $n_{k_0-d} = n_{k_0+d} = 0$. Consequently, the expected numbers of occurrences after watermarking are: $\tilde{n}_{k_0-d} = \tilde{n}_{k_0+d} = n_{k_0}/2$ and $\tilde{n}_{k_0} = 0$. Hence, using a set of similar equations, it is possible in certain cases to recover the original distribution of the histogram and the value of the embedded watermark.

C. The Mosaic Attack

There is a presentation attack which is quite general and which possesses the initially remarkable property that we can remove the marks from an image and still have it rendered exactly the same, pixel for pixel, as the marked image by a standard browser.

It was motivated by a fielded system for copyright piracy detection, consisting of a watermarking scheme plus a web crawler that downloads pictures from the net and checks whether they contain a client's watermark.

Our mosaic attack consists of chopping an image up into a number of smaller subimages, which are embedded one after another in a web page. Common web browsers render juxtaposed subimages stuck together as a single image, so the result is identical to the original image. This attack appears to be quite general; all marking schemes require the marked image to have some minimal size (one cannot hide a meaningful mark in just one pixel). Thus, by splitting an image into sufficiently small pieces, the mark detector will be confused [104]. One defense would be to ensure that the minimal size would be quite small and the mosaic attack might therefore not be very practical.

But there are other problems with such "crawlers." Mobile code such as Java applets can be used to display a picture inside the browser; the applet could descramble the picture in real time. Defeating such techniques would entail rendering the whole page, detecting pictures and checking whether they contain a mark. Another problem is that pirated pictures are typically sold via many small web services, from which the crawler would have to purchase them using a credit card before it could examine them.

D. Interpretation Attacks

StirMark and our attack on echo hiding are examples of the kind of threat that dominates the information-hiding literature—namely, a pirate who removes the mark directly using technical means. Indeed, the definition commonly used for robustness includes only resistance to signal manipulation (cropping, scaling, resampling, etc.). However, Craver *et al.* show that this is not enough by exhibiting a "protocol" level attack in [116].

The basic idea is that as many schemes provide no intrinsic way of detecting which of two watermarks was added first. If the owner of the document d encodes a watermark w, publishes the marked version $d + w$, and has no other proof of ownership, then a pirate who has registered his watermark as w' can claim that the document is his and that the original unmarked version of it was $d + w - w'$. Their paper [117] extends this idea to defeat a scheme which is noninvertible (an inverse needs only be approximated).

Craver *et al.* argue for the use of information-losing marking schemes whose inverses cannot be approximated closely enough. Our alternative interpretation of their attack is that watermarking and fingerprinting methods must be used in the context of a larger system that may use mechanisms such as timestamping and notarization to prevent attacks of this kind.

Environmental constraints may also limit the amount of protection which technical mechanisms can provide. For example, there is little point in using an anonymous digital cash system to purchase goods over the Internet, if the purchaser's identity is given away in the headers of his e-mail message or if the goods are shipped to his home address.

E. Implementation Considerations

The robustness of embedding and retrieving algorithms and their supporting protocols is not the only issue. Most

real attacks on fielded cryptographic systems have come from the opportunistic exploitation of loopholes that were found by accident; cryptanalysis was rarely used, even against systems that were vulnerable to it [54].

We cannot expect copyright marking systems to be any different, and the pattern was followed in the first attack to be made available on the Internet against one of the most widely used picture marking schemes. This attack exploited weaknesses in the implementation rather than in the underlying marking algorithms, even although these are weak (the marks can be removed with StirMark).

Each user has an ID and a two-digit password, which is issued when he registers with the marking service and pays a subscription. The correspondence between ID's and passwords is checked using obscure software and, although the passwords are short enough to be found by trial and error, the published attack first uses a debugger to break into the software and disable the password-checking mechanism. As ID's are public, either password search or disassembly enables any user to be impersonated.

A deeper examination of the program allows a villain to change the ID, and thus the copyright mark, of an already marked image as well as the type of use (such as adult versus general public content). Before embedding a mark, the program checks whether there is already a mark in the picture, but this check can be bypassed fairly easily using the debugger with the result that it is possible to overwrite any existing mark and replace it with another one.

Exhaustive search for the personal code can be prevented without difficulty, but there is no obvious solution to the disassembly attack. If tamper-resistant software [118] cannot give enough protection, then one can always have an online system in which each user shares a secret stego-key with a trusted party and uses this key to embed some kind of digital signature. Observe that there are two separate keyed operations here: the authentication (such as a digital signature) and the embedding or hiding operation.

Although we can do public key steganography—hiding information using a public key so that only someone with the corresponding private key can detect its existence [119]—we still do not know how to do the hiding equivalent of a digital signature; that is, to enable someone with a private key to embed marks in such a way that anyone with the corresponding public key can read them but not remove them. Some attempts to create such watermarks can be found in [120]. But unless we have some new ideas, we appear compelled to use either a central "mark reading" service or a tamper-resistant implementation, just as cryptography required either central notarization or tamper-evident devices to provide a nonrepudiation service in the days before the invention of digital signatures.

However, there is one general attack on tamper-resistant mark readers due to Cox *et al.* [121]. The idea is to explore, pixel by pixel, an image at the boundary where the detector changes from "mark absent" to "mark present" and iteratively construct an acceptable image in which the mark is not detected. Of course, with a programmable tamper-proof processor, one can limit the number of variants of a given picture for which an answer will be given, and the same holds for a central mark reading service. But in the absence of physically protected state, it is unclear how this attack can be blocked.

V. A Basic Theory of Steganography

This leads naturally to the question of whether we can develop a comprehensive theory of information hiding, in the sense that Shannon provided us with a theory of secrecy systems [122] and Simmons of authentication systems [123]. Quite apart from intellectual curiosity, there is a strong practical reason to seek constructions whose security is mathematically provable. This is because copyright protection mechanisms may be subjected to attack over an extraordinarily long period of time. Copyright subsists for typically 50–70 years after the death of the artist, depending on the country and the medium; this means that mechanisms fielded today might be attacked using the resources available in 100 years' time. Where cryptographic systems need to provide such guarantees, as in espionage, it is common to use a one-time pad because we can prove that the secrecy of this system is independent of the computational power available to the attacker. Is it possible to get such a guarantee for an information-hiding system?

A. Early Results

An important step in developing a theory of a subject is to clarify the definitions. Intuitively, the purpose of steganography is to set up a secret communication path between two parties such that any person in the middle cannot detect its existence; the attacker should not gain any information about the embedded data by simply looking at cover text or stego-text. This was first formalized by Simmons in 1983 as the "prisoners' problem" [124]. Alice and Bob are in jail and wish to prepare an escape plan. The problem is that all their communications are arbitrated by the warden Willie. If Willie sees any cipher text in their messages, he will frustrate them by putting them into solitary confinement. So Alice and Bob must find a way to exchange hidden messages.

Simmons showed that such a channel exists in certain digital signature schemes: the random message key used in these schemes can be manipulated to contain short messages. This exploitation of existing randomness means that the message cannot even in principle be detected, and so Simmons called the technique the "subliminal channel." The history of the subliminal channel is described in [125], while further results may be found in [123], [126]–[128].

In the general case of steganography, where Willie is allowed to modify the information flow between Alice and Bob, he is called an active warden; but if he can only observe it he is called a passive warden. Further studies showed that public key steganography is possible (in this model, Alice and Bob did not exchange secrets before going to jail, but have public keys known to each

other)—although the presence of an active warden makes public key steganography more difficult [129].

This difficulty led to the introduction, in [130], of the supraliminal channel, which is a very low bandwidth channel that Willie cannot afford to modify as it uses the most perceptually significant components of the cover object as a means of transmission. For example, a prisoner might write a short story in which the message is encoded in the succession of towns or other locations at which the action takes place. Details of these locations can be very thoroughly woven into the plot, so it becomes in practice impossible for Willie to alter the message—he must either allow the message through or censor it. The effect of this technique is to turn an active warden into a passive one. The same effect may be obtained if the communicating parties are allowed to use a digital signature scheme.

B. The General Role of Randomness

Raw media data rates do not necessarily represent information rates. Analog values are quantized to n bits giving, for instance, a data rate of 16 bit/sample for audio or 8 bit/pixel for monochrome images. The average information rate is given by their entropy; indeed, the entropy of monochrome images is generally around 4–6 bits per pixel. This immediately suggests the use of this difference to hide information. So if C is the cover text and E the embedded text, transmitted on a perfect n bit channel, one would have: $H(E) \leq n - H(C)$ bit/pixel, so all the gain provided by compression is used for hiding. One could also take into account the the stego-text S and impose the constraint that no information is given about E, even knowing S and C_k (a part of C typically the natural noise of the cover text): the transinformation should be zero $T(E;(C_k,S))=0$. In this case, it can be shown that $H(E) \leq H(C_k|S)$ [131]. So the rate at which one can embed cipher text in a cover object is bounded by the opponent's uncertainty about the cover text given knowledge of stego-text. But this gives an upper bound on the stego-capacity of a channel, when for a provably secure system we need a lower bound. In fact, all the theoretical bounds known to us are of this kind. In addition, the opponent's uncertainty and thus the capacity might asymptotically be zero, as was noted in the context of covert channels [132].

This also highlights the fact that steganography is much more dependent on our understanding of the information sources involved than cryptography is, which helps explain why we do not have any lower bounds on capacity for embedding data in general sources. It is also worth noting that if we had a source which we understood completely and so could compress perfectly, then we could simply subject the embedded data to our decompression algorithm and send it as the stego-text directly. Thus, steganography would either be trivial or impossible depending on the system [119].

Another way of getting around this problem is to take advantage of the natural noise of the cover text. Where this can be identified, it can be replaced by the embedded data (which we can assume have been encrypted and are thus indistinguishable from random noise). This is the philosophy behind some steganographic systems [60], [133], [134] and early image marking systems [22] (it may not work if the image is computer generated and thus has very smooth color gradations). It can also be applied to audio [51], [135]; here, randomizing is very important because simple replacement of the least significant bit causes an audible modification of the signal [51]. So a subset of modifiable bits is chosen and the embedding density depends on the observed statistics of the cover signal [135] or on its psychoacoustic properties [51].

It is also possible to exploit noise elsewhere in the system. For example, one might add small errors by tweaking some bits at the physical or data link layer and hope that error-correction mechanisms would prevent anyone reading the message from noticing anything. This approach would usually fall foul of Kerckhoffs' principle that the mechanism is known to the opponent, but in some applications it can be effective [136].

A more interesting way of embedding information is to change the parameters of the source encoding. An example is given by a marking technique proposed for DVD. The encoder of the MPEG stream has many choices of how the image can be encoded, based on the tradeoff between good compression and good quality—each choice conveys one or more bits. Such schemes trade expensive marking techniques for inexpensive mark detection; they may be an alternative to signature marks in digital TV where the cost of the consumer equipment is all important [137].

Finally, in case the reader should think that there is anything new under the sun, consider two interpretations of a Beethoven symphony, one by Karajan the other one by Bernstein. These are very similar, but also dramatically different. They might even be considered to be different encodings, and musicologists hope to eventually discriminate between them automatically.

C. Robust Marking Systems

In the absence of a useful theory of information hiding, we can ask the practical question of what makes a marking scheme robust. This is in some ways a simpler problem (everyone might know that a video is watermarked, but so long as the mark is unobtrusive this may not matter) and in other ways a harder one (the warden is guaranteed to be active, as the pirate will try to erase marks).

As a working definition, we mean by a robust marking system one with the following properties.

- Marks should not degrade the perceived quality of the work. This immediately implies the need for a good quality metric. In the context of images, pixel based metrics are not satisfactory, and better measures based on perceptual models can be used [108], [138].
- Detecting the presence and/or value of a mark should require knowledge of a secret.
- If multiple marks are inserted in a single object, then they should not interfere with each other; moreover if different copies of an object are distributed with

different marks, then different users should not be able to process their copies in order to generate a new copy that identifies none of them.
- The mark should survive all attacks that do not degrade the work's perceived quality, including resampling, re-quantization, dithering, compression, and especially combinations of these.

Requirements similar to these are found, for example, in a recent call for proposals from the music industry [139]. However, as we have shown with our attacks, there are at present few marking schemes, whether in the research literature or on commercial sale, that are robust against attacks involving carefully chosen distortions. Vendors, when pressed, claim that their systems will withstand most attacks but cannot reasonably be engineered to survive sophisticated ones. However, in the experience of a number of industries, it is:

> ... a wrong idea that high technology serves as a barrier to piracy or copyright theft; one should never underestimate the technical capability of copyright thieves [140].

Our current opinion is that most applications have a fairly sharp tradeoff between robustness and data rate which may prevent any single marking scheme meeting the needs of all applications. However, we do not see this as a counsel of despair. The marking problem has so far been over abstracted; there is not one "marking problem" but a whole constellation of them. Most real applications do not require all of the properties in the above list. For example, when monitoring radio transmissions to ensure that adverts have been played as contracted, we only require enough resistance to distortion to deal with naturally occurring effects and prevent transfer of marks from one advert to another [141]; where our concern is to make proprietary images available to scholars, as in the "Vatican Library Accessible Worldwide" project, IBM came up with a simple solution using visible watermarks—which leave the documents still perfectly suitable for research purposes but discourage illegal publication for profit [11].

VI. 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 any one marking scheme satisfy all of these requirements simultaneously, that is, that "the marking problem," as sometimes described in the literature, is overspecified.

That does not, of course, mean that particular marking problems are insoluble. Both historical precedent and recent innovation provide us with a wide range of tools, which if applied intelligently should be sufficient to solve most of the problems that we meet in practice.

Acknowledgment

Some of the ideas presented here were clarified by discussion with R. Needham, J. Daugman, P. Rayner, M. Kutter, and S. Craver. Special thanks to the Whipple Science Museum Library, the Rare Book Section of the Cambridge University Library, and the Cambridge University Archives for their help.

References

[1] A. 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 and Security—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.
[10] F. L. Bauer, *Decrypted Secrets—Methods and Maxims of Cryptology*. Berlin, Heidelberg, Germany: Springer-Verlag, 1997.
[11] G. W. Braudaway, K. A. Magerlein, and F. Mintzer, "Protecting publicly-available images with a visible image watermark," in *Optical Security and Counterfeit Deterrence Techniques*, vol. 2659, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1996, pp. 126–133.
[12] B. Rudin, *Making paper—A Look Into the History of an Ancient Craft*. Vällingby, Sweden: Rudins, 1990.
[13] I. J. Cox and M. L. Miller, "A review of watermarking and the importance of perceptual modeling," in *Human Vision and Electronic Imaging II*, vol. 3016B, E. Rogowitz and T. N. Pappas, Eds. San Jose, CA: IS&T and SPIE, 1997.
[14] G. Caronni, "Assuring ownership rights for digital images," in *Reliable IT Systems (VIS'95)*, H. Brüggermann and W. Gerhardt-Häckl, Eds. Germany: Vieweg, 1995, pp. 251–263.
[15] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "A secure, robust watermark for multimedia," in R. J. Anderson, Ed., "Information hiding: First international workshop," in *Lec-*

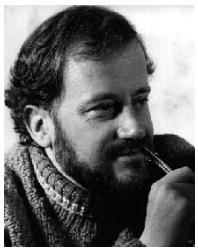
ture Notes in Computer Science, vol. 1174. Berlin, Germany: Springer-Verlag, 1996, pp. 183–206.
[16] C. I. Podilchuk and W. Zeng, "Digital image watermarking using visual models," in *Human Vision and Electronic Imaging II*, vol. 3016, B. E. Rogowitz and T. N. Pappas, Eds. San Jose, CA: IS&T and SPIE, 1997, pp. 100–111.
[17] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 357–372, May 1998.
[18] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Int. Conf. Acoustic, Speech and Signal Processing (ICASP)*, vol. 5. Seattle, WA: IEEE, 1998, pp. 2969–2972.
[19] G. Nicchiotti and E. Ottaviano, "Non-invertible statistical wavelet watermarking," in *Proc. 9th Europ. Signal Processing Conf. (EUSIPCO'98)*, Rhodes, Greece, Sept. 8–11, 1998, pp. 2289–2292.
[20] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, no. 3, pp. 385–403, May 1998.
[21] D. Tzovaras, N. Karagiannis, and M. G. Strintzis, "Robust image watermarking in the subband or discrete cosine transform domain," in *Proc. 9th European Signal Processing Conf. (EUSIPCO'98)*, Rhodes, Greece, Sept. 8–11, 1998, pp. 2285–2288.
[22] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, Austin, TX, 1994, pp. 86–90.
[23] R. B. Wolfgang and E. J. Delp, "A watermarking technique for digital imagery: Further studies," in *Proc. IEEE Int. Conf. Imaging, Systems, and Technology*, Las Vegas, NV, June 30–July 3, 1997, pp. 279–287.
[24] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Signal Processing*, vol. 66, no. 3, pp. 283–301, May 1998.
[25] A. Herrigel, J. J. K. O'Ruanaidh, H. Petersen, S. Pereira, and T. Pun, "Secure copyright protection techniques for digital images," in *Information Hiding: Second Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 169–190.
[26] M. D. Swanson, B. Zu, and A. H. Tewfik, "Robust data hiding for images," in *Proc. IEEE 7th Digital Signal Processing Workshop (DSP 96)*, Loen, Norway, Sept. 1996, pp. 37–40.
[27] G. C. Langelaar, J. C. A. van der Lubbe, and R. L. Lagendijk, "Robust labeling methods for copy protection of images," in *Storage and Retrieval for Image and Video Database V*, vol. 3022, I. K. Sethin and R. C. Jain, Eds. San Jose, CA: IS&T and SPIE, pp. 298–309.
[28] J. Zhao and E. Koch, "Embedding robust labels into images for copyright protection," in *Int. Congr. Intellectual Property Rights for Specialised Information, Knowledge and New Technologies*, Vienna, Austria, Aug. 1995.
[29] G. Schott, *Schola Steganographica: In Classes Octo Distributa* (Whipple Collection). Cambridge, U.K.: Cambridge Univ., 1680.
[30] J. Reeds, "Solved: The ciphers in book III of Trithemius' steganographia," *Cryptologia*, vol. XXII, no. 4, pp. 291–317, Oct. 1998.
[31] D. Kahn, *The Codebreakers—The Story of Secret Writing*. New York: Scribner, 1996.
[32] A. Kerckhoffs, "La cryptographie militaire," *J. Sciences Militaires*, vol. 9, pp. 5–38, Jan. 1883.
[33] R. J. Anderson, "Liability and computer security: Nine principles," in *Computer Security—3rd Europ. Symp. Research in Computer Security (ESORICS'94)* (*Lecture Notes in Computer Science*), vol. 875, D. Gollmann, Ed. Berlin, Germany: Springer-Verlag, pp. 231–245.
[34] J. Baltrušaitis, "Anamorphoses ou thaumaturgus opticus," in *Les Perspectives Dépravées*. Paris, France: Flammarion, 1984, pp. 5 and 15–19.
[35] A. Seckel, "Your mind's eye: Illusions & paradoxes of the visual system," presented at National Science Week, Univ. Cambridge, Cambridge, U.K., Mar. 1998. [Online]. Available WWW: http://www.illusionworks.com/.
[36] Herodotus, *The Histories*. London, U.K.: J. M. Dent & Sons, 1992, ch. 5 and 7.
[37] B. Newman, *Secrets of German Espionage*. London, U.K.: Robert Hale, 1940.
[38] WitnesSoft and ScarLet security software. (1997, Apr.). Aliroo home page. [Online]. Available WWW: http://www.aliroo.com/.
[39] J. C. Murphy, D. Dubbel, and R. Benson, "Technology approaches to currency security," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314, R. L. van Renesse, Ed.. San Jose, CA: IS&T and SPIE, 1998, pp. 21–28.
[40] G. W. W. Stevens, *Microphotography—Photography and Photofabrication at Extreme Resolutions*. London, U.K.: Chapman & Hall, 1968.
[41] D. Brewster, "Microscope," in *Encyclopædia Britannica or the Dictionary of Arts, Sciences, and General Literature*, vol. XIV, 8th ed. Edinburgh, U.K.: Britannica, 1857, pp. 801–802.
[42] G. Tissandier, *Les Merveilles de la Photographie*. Paris, France: Librairie Hachette & Cie, 1874, ch. 6, pp. 233–248.
[43] J. Hayhurst. (1970). The pigeon post into Paris 1870–1871. [Online]. Available WWW: http://www.windowlink.com/jdhayhurst/pigeon/pigeon.html.
[44] J. E. Hoover, "The enemy's masterpiece of espionage," *The Reader's Dig.*, vol. 48, pp. 49–53, May 1946.
[45] J. F. Delaigle, C. De Vleeschouwer, and B. Macq, "Watermarking algorithm based on a human visual model," *Signal Processing*, vol. 66, no. 3, pp. 319–335, May 1998.
[46] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," in *Proc. 1996 IEEE Int. Conf. Multimedia Computing and Systems*, Hiroshima, Japan, June 17–23, 1996, pp. 473–480.
[47] F. Goffin, J.-F. Delaigle, C. D. Vleeschouwer, B. Macq, and J.-J. Quisquater, "A low cost perceptive digital picture watermarking method," in *Storage and Retrieval for Image and Video Database V*, vol. 3022, I. K. Sethin and R. C. Jain, Eds. San Jose, CA: IS&T and SPIE, 1997, pp. 264–277.
[48] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, pp. 1385–1422, Oct. 1993.
[49] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 3rd ed. London, U.K.: Academic, 1989.
[50] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information—Part 3: Audio.*, British Standard, BSI, implementation of ISO/IEC 13818-3:1995, Oct. 1995.
[51] A. Werner, J. Oomen, M. E. Groenewegen, R. G. van der Waal, and R. N. Veldhuis, "A variable-bit-rate buried-data channel for compact disc," *J. Audio Eng. Soc.*, vol. 43, no. 1/2, pp. 23–28, Jan./Feb. 1995.
[52] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and videos," this issue, pp. 1108–1126.
[53] M. Swanson and A. Tewfik, "Perceptual watermarking of audio signals," submitted for publication.
[54] R. J. Anderson, "Why cryptosystems fail," *Commun. ACM*, vol. 37, pp. 32–40, Nov. 1994.
[55] F. Bacon, *Of the Advancement and Proficiencie of Learning or the Partitions of Sciences*, vol. VI. Oxford, U.K.: Leon Lichfield, 1640, pp. 257–271.
[56] P. Leary, *The Second Cryptographic Shakespeare: A Monograph wherein the Poems and Plays Attributed to William Shakespeare are Proven to Contain the Enciphered Name of the Concealed Author, Francis Bacon*, 2nd ed. Omaha, NE: Westchester House, 1990.
[57] N. R. Wagner, "Fingerprinting," in *IEEE Symp. Security and Privacy*, Oakland, CA, Apr. 25–27, 1983, pp. 18–22.
[58] J. Brassil, S. Low, N. Maxemchuk, and L. O'Garman, "Electronic marking and identification techniques to discourage document copying," in *Proc. Infocom*, Toronto, Canada, June 1994, pp. 1278–1287.
[59] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proc. IEEE Int. Conf. Images Processing*, Lausanne, Switzerland, Sept. 1996, pp. 219–222.
[60] S. Walton, "Image authentication for a slippery new age," *Dr. Dobb's J. Software Tools*, vol. 20, no. 4, pp. 18–26, Apr. 1995.
[61] K. Matsui and K. Tanaka, "Video-steganography: How to secretly embed a signature in a picture," *J. Interactive Multimedia Association Intellectual Property Project*, vol. 1, no. 1, pp. 187–205, Jan. 1994.

[62] R. J. Anderson and C. Manifavas, "Chameleon—A new kind of stream cipher," in *Fast Software Encryption—4th Int. Workshop (FSE'97)* (*Lecture Notes in Computer Science*), vol. 1267, E. Biham, Ed. Berlin, Germany: Springer-Verlag, pp. 107–113, 1997.
[63] J. R. Smith and B. O. Comiskey, "Modulation and information hiding in images," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 207–226.
[64] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, nos. 3, 4, pp. 313–336, 1996.
[65] I. Pitas, "A method for signature casting on digital images," in *Proc. Int. Conf. Image Processing*, vol. 3, Sept. 1996, pp. 215–218.
[66] G. C. Langelaar, J. C. van der Lubbe, and J. Biemond, "Copy protection for multimedia data based on labeling techniques," in *Proc. 17th Symp. Information Theory in the Benelux*, Enschede, The Netherlands, May 1996.
[67] R. C. Dixon, *Spread Spectrum Systems with Commercial Applications*, 3rd ed. New York: Wiley, 1994.
[68] R. A. Scholtz, "The origins of spread spectrum communications," *IEEE Trans. Commun.*, vol. 30, pp. 822–853, May 1982.
[69] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread spectrum communications—A tutorial," *IEEE Trans. Commun.*, vol. 30, pp. 855–884, May 1982.
[70] A. Z. Tirkel, G. A. Rankin, R. M. van Schyndel, W. J. Ho, N. R. A. Mee, and C. F. Osborne, "Electronic watermark," in *Digital Image Computing, Technology and Applications (DICTA'93)*, Macquarie University, Sydney, Australia, 1993, pp. 666–673.
[71] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video," in *Proc. IEEE Int. Conf. Image Processing (ICIP'96)*, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 243–246.
[72] F. Hartung and B. Girod, "Watermarking of MPEG-2 encoded video without decoding and re-encoding," in *Multimedia Computing and Networking 1997*, vol. 3020, M. Freeman, P. Jardetzky, and H. M. Vin, Eds. San Jose, CA: IS&T and SPIE, 1997, pp. 264–273.
[73] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," in *Proc. IEEE Int. Conf. Multimedia Computing and Systems.*, Hiroshima, Japan, June 17–23, 1996, pp. 473–480.
[74] CompuServe, Inc., OH. (1987, June). Graphics interchange format (GIF) specification. [Online]. Available WWW: http://icib.igd.fhg.de/icib/it/defacto/company/compuserve/gif87a/.
[75] G. Jagpal, "Steganoraphy in digital images," Ph.D. dissertation, Selwyn College, Cambridge Univ., Cambridge, U.K., May 1995.
[76] F. A. P. Petitcolas. (1998, Aug.). MP3Stego. [Online]. Available WWW: http://www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/.
[77] E. Koch and J. Zhao, "Toward robust and hidden image copyright labeling," in *Proc. IEEE Workshop Nonlinear Signal and Image Processing*, Neos Marmaras, Greece, June 1995, pp. 452–455.
[78] J. J. K. O'Ruanaidh, W. J. Dowling, and F. M. Boland, "Watermarking digital images for copyright protection," *Proc. Inst. Elect. Eng. Vision, Signal and Image Processing*, vol. 143, no. 4, pp. 250–256, Aug. 1996.
[79] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Transparent robust image watermarking," in *Proc. IEEE Int. Conf. Image Processing*, vol. III, 1996, pp. 211–214.
[80] D. Kundur and D. Hatzinakos, "A robust digital image watermarking method using wavelet-based fusion,' in *Proc. IEEE Int. Conf. Image Processing*, Santa Barbara, CA, Oct. 1997, pp. 544–547.
[81] J. J. K. O'Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 303–317, May 1998.
[82] D. Gruhl, W. Bender, and A. Lu, "Echo hiding," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 295–315.
[83] B. P. Bogert, M. J. R. Healy, and J. W. Tukey, "The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking," in *Symp. Time Series Analysis*, M. Rosenblatt, Ed. New York: Wiley, 1963, pp. 209–243.
[84] D. L. Schilling, Ed., *Meteor Burst Communications: Theory and Practice* (Wiley Series in Telecommunications). New York: Wiley, 1993.
[85] D. Gruhl and W. Bender, "Information hiding to foil the casual counterfeiter," in *Information Hiding: Second Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 1–15.
[86] R. L. van Renesse, "Security design of valuable documents and products," in *Optical Security and Counterfeit Deterrence Techniques*, vol. 2659. San Jose, CA: IS&T and SPIE, 1996, pp. 10–20.
[87] T. Matsumoto, "Protection of documents," submitted for publication.
[88] R. L. van Renesse, Ed., *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314. San Jose, CA: IS&T and SPIE, 1998.
[89] ——, *Optical Security and Counterfeit Deterrence Techniques*, vol. 2659. San Jose, CA: IS&T and SPIE, 1996.
[90] D. van Lingen, "The new Dutch passport," in *Optical Security and Counterfeit Deterrence Techniques*, vol. 2659, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1996, pp. 67–73.
[91] S. Spannenburg, "Optically- and machine-detectable security elements," in *Optical Security and Counterfeit Deterrence Techniques*, vol. 2659, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1996, pp. 76–96.
[92] R. L. van Renesse, "Verifying versus falsifying banknotes," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314. San Jose, CA: IS&T and SPIE, 1998.
[93] J. D. Brongers, "Search for effective document security by 'inventioneering'," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1998, pp. 29–38.
[94] "Anti-counterfeit trials begin with watermark technology," *Financial Technol. Int. Bulletin*, vol. 9, no. 2, pp. 6–7, Oct. 1993.
[95] I. M. Lancaster and L. T. Konntnik, "Progress in counterfeit deterrence: The contribution of information exchange," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1998, pp. 2–8.
[96] R. Johnson and A. Garcia, "Vulnerability assessment of security seals," *J. Security Administration*, vol. 20, no. 1, pp. 15–27, June 1997.
[97] V. Gligor, "A guide to understanding covert channel analysis of trusted systems," National Computer Security Center, Ft. George G. Meade, MD, Tech. Rep. NCSC-TG-030, Nov. 1993.
[98] B. Lampson, "A note on the confinement problem," *Commun. ACM*, vol. 16, no. 10, pp. 613–615, Oct. 1973.
[99] W. van Eck, "Electromagnetic radiation from video display units: an eavesdropping risk?," *Comput. Security*, vol. 4, no. 4, pp. 269–286, Dec. 1985.
[100] D. Russel and G. Gangemi, *Computer Security Basics*. Sebastopol, CA: O'Reilly & Associates, 1991, ch. 10.
[101] M. G. Kuhn and R. J. Anderson, "Soft tempest: Hidden data transmission using electromagnetic emanations," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 124–142.
[102] J.-P. M. G. Linnartz and M. van Dijk, "Analysis of the sensitivity attack against electronic watermarks in images," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 258–272.
[103] M. Maes, "Twin peaks: The histogram attack on fixed depth image watermarks," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 290–305.
[104] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 218–238.
[105] G. C. Langelaar, R. L. Lagendijk, and J. Biemond, "Removing spatial spread spectrum watermarks by nonlinear filtering," in *9th European Signal Processing Conference (EUSIPCO'98)*, Rhodes, Greece, Sept. 8–11 1998, pp. 2281–2284.

[106] R. Barnett and D. E. Pearson, "Frequency mode LR attack operator for digitally watermarked images," *Electron. Lett.*, vol. 34, no. 19, pp. 1837–1839, Sept. 1998.
[107] M. Kutter, "Watermarking resisting to translation, rotation, and scaling," in *Proc. SPIE Multimedia Systems and Applications*, vol. 3528, Boston, MA, Nov. 1998, pp. 423–431.
[108] M. Kutter and F. A. P. Petitcolas, "A fair benchmark for image watermarking systems," in *11th Int. Symp. Electronic Imaging*, vol. 3657. San Jose, CA: IS&T and SPIE, Jan. 25–27, 1999.
[109] N. F. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 273–289.
[110] S. Craver, B.-L. Yeo, and M. Yeung, "Technical trials and legal tribulations," *Commun. ACM*, vol. 41, no. 7, pp. 44–54, July 1998.
[111] K. N. Hamdy, A. H. Tewfik, T. Chen, and S. Takagi, "Time-scale modification of audio signals with combined harmonic and wavelet representations," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 1, Munich, Germany, pp. 439–442.
[112] N. A. Dodgson, "Quadratic interpolation for image resampling," *IEEE Trans. Image Processing*, vol. 6, pp. 1322–1326, Sept. 1997.
[113] G. W. Braudaway, "Results of attacks on a claimed robust digital image watermark," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314, R. L. van Renesse, Ed. San Jose, CA: IS&T and SPIE, 1998.
[114] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Int. ed. Englewood Cliffs, NJ: Prentice-Hall, 1989, ch. 12, pp. 768–834.
[115] R. W. Schafer, "Echo removal by discrete generalized linear filtering," Massachusetts Inst. Technol., Cambridge, MA, Tech. Rep. 466, Feb. 1969.
[116] S. Craver, N. Memon, B.-L. Yeo, and M. M. Yeung, "Can invisible watermark resolve rightful ownerships?," in *Storage and Retrieval for Image and Video Database V*, vol. 3022, I. K. Sethin and R. C. Jain, Eds. San Jose, CA: IS&T and SPIE, 1997, pp. 310–321.
[117] ——, "Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 573–586, May 1998.
[118] D. Aucsmith, "Tamper resistant software: An implementation," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 317–333.
[119] R. J. Anderson, "Stretching the limits of steganography," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 39–48.
[120] F. Hartung and B. Girod, "Fast public-key watermarking of compressed video," in *Proc. IEEE Int. Conf. Image Processing (ICIP'97)*, vol. I, Santa Barbara, CA, Oct. 1997, pp. 528–531.
[121] I. J. Cox and J.-P. M. G. Linnartz, "Public watermarks and resistance to tampering," in *Proc. IEEE Int. Conf. Image Processing (ICIP'97)*, Santa Barbara, CA, Oct. 26–29, 1997.
[122] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1949.
[123] G. J. Simmons, Ed., *Contemporary Cryptology—The Science of Information Integrity*. New York: IEEE Press, 1992.
[124] ——, "The prisoners' problem and the subliminal channel," in *Proc. IEEE Workshop Communications Security CRYPTO'83*, Santa Barbara, CA, 1983, pp. 51–67.
[125] ——, "The history of subliminal channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 452–462, May 1998.
[126] R. Anderson, S. Vaudenay, B. Preneel, and K. Nyberg, "The Newton channel," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 151–156.
[127] ——, "Subliminal channels: Past and present," *Europ. Trans. Telecommun.*, vol. 5, no. 4, pp. 459–473, July/Aug. 1994.
[128] ——, "Results concerning the bandwidth of subliminal channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 463–473, May 1998.
[129] R. J. Anderson and F. A. P. Petitcolas, "On the limits of steganography," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 474–481, May 1998.
[130] S. Craver, "On public-key steganography in the presence of an active warden," IBM Res. Division, T. J. Watson Res. Center, Yorktown Heights, NY, Tech. Rep. RC20931, July 1997.
[131] J. Zöllner, H. Federrath, H. Klimant, A. Pfitzmann, R. Piotraschke, A. Westfeld, G. Wicke, and G. Wolf, "Modeling the security of steganographic systems," in *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525, D. Aucsmith, Ed. Berlin, Germany: Springer-Verlag, 1998, pp. 344–354.
[132] I. Moskowitz and M. Kang, "Covert channels—Here to stay?," in *Compass'94*, 1994, pp. 235–243.
[133] C. Maroney. (1997, Mar.). Hide and seek. [Online]. Available WWW: http://www.cypher.net/products/hideseek.html.
[134] (1997, Mar.). Stegodos. [Online]. Available WWW: ftp://ftp.funet.fi/pub/crypt/steganography/ stegodos.zip.
[135] E. Franz, A. Jerichow, S. Möller, A. Pfitzmann, and I. Stierand, "Computer based steganography: How it works and why therefore any restriction on cryptography are nonsense, at best," in *Information Hiding: 1st Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1174, R. J. Anderson, Ed. Berlin, Germany: Springer-Verlag, 1996, pp. 7–21.
[136] P. Wayner, *Disappearing Cryptography—Being and Nothing on the Net*. Chestnut Hill, MA: AP Professional, 1996.
[137] B. M. Macq and J.-J. Quisquater, "Cryptology for digital TV broadcasting," *Proc. IEEE*, vol. 83, pp. 944–956, June 1995.
[138] J. Fridrich and M. Goljan, "Comparing robustness of watermarking techniques," in *11th Int. Symp. Electronic Imaging*, vol. 3657. San Jose, CA: IS&T and SPIE, 1999.
[139] Int. Federation of the Phonographic Industry, "Request for proposals—Embedded signalling systems," Int. Federation of the Phonographic Industry, London, U.K., June 1997.
[140] J. Gurnsey, *Copyright Theft*. Aldershot, U.K.: Aslib Gower, 1995.
[141] R. Willard, "ICE (Identification Coding, Embedded)," in *74th Conv. Audio Engineering Society Preprints*, Berlin, Germany, Mar. 16–19, 1993, Preprint 3516 (D2-3).
[142] W. Niblack and R. C. Jain, Eds., *Storage and Retrieval for Image and Video Database III*, vol. 2420. San Jose, CA: IS&T and SPIE, 1995.
[143] B. E. Rogowitz and T. N. Pappas, Eds., *Human Vision and Electronic Imaging II*, vol. 3016. San Jose, CA: IS&T and SPIE, 1997.
[144] D. Chaum, Ed., *Proc. IEEE Workshop Communications Security CRYPTO'83*, Santa Barbara, CA.
[145] M. Lomas, B. Crispo, B. Christianson, and M. Roe, Eds., *Security Protocols: Proc. 5th Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1361. Berlin, Germany: Springer-Verlag.
[146] E. Biham, Ed., *Fast Software Encryption—4th Int. Workshop (FSE'97)* (*Lecture Notes in Computer Science*), vol. 1267. Berlin, Germany: Springer-Verlag.
[147] J.-J. Quisquater, Y. Deswarte, C. Meadows, and D. Gollmann, Eds., *Computer Security—5th Europ. Symp. Research in Computer Security, (ESORICS'98)* (*Lecture Notes in Computer Science*), vol. 1485. Berlin, Germany: Springer-Verlag, 1998.
[148] J. Dittmann, P. Wohlmacher, P. Horster, and R. Steinmetz, Eds., *Multimedia and Security—Workshop at ACM Multimedia'98* (*GMD Report*), vol. 41. Bristol, U.K.: ACM, GMD—Forschungszentrum Informationstechnik GmbH, 1998.
[149] D. Aucsmith, Ed., *Information Hiding: 2nd Int. Workshop* (*Lecture Notes in Computer Science*), vol. 1525. Berlin, Germany: Springer-Verlag, 1998.
[150] I. K. Sethin and R. C. Jain, Eds., *Storage and Retrieval for Image and Video Database V*, vol. 3022. San Jose, CA: IS&T and SPIE, 1997.
[151] *The Oxford English Dictionary* (corrected reissue). Oxford, U.K.: Clarendon, 1933.

Fabien A. P. Petitcolas graduated from the École Centrale, Lyon, France and received the Diploma in computer science from the University of Cambridge, U.K.

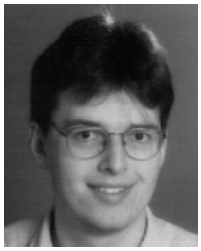
He is currently a research student at the Computer Laboratory, University of Cambridge, U.K. His research topic is the robustness of information-hiding systems.



Ross J. Anderson received the B.A., M.A., and Ph.D. degrees from the University of Cambridge, U.K.

He currently teaches and directs research in computer security and software engineering at the University of Cambridge.

Dr. Anderson was the Program Chair of the First International Workshop on Information Hiding, held at Cambridge in May–June 1996. He is a Fellow of the RSA and the IMA and is a Chartered Engineer. He is also the Editor-in-Chief of *Computer and Communications Security Reviews*.



Markus G. Kuhn received the Diploma from the University of Erlangen–Nürnberg, Germany, and the M.Sc. degree from Purdue University, West Lafayette, IN, both in computer science.

He is currently with the Computer Laboratory at the University of Cambridge, U.K.. His research interests include the security of tamper-resistant hardware, intellectual property protection mechanisms, and global-scale distributed databases.