Weaknesses of copyright marking systems

Fabien A.P. Petitcolas and Ross J. Anderson

University of Cambridge, Computer Laboratory
Pembroke Street, Cambridge CB2 3QG, UK

{fapp2, rja14}@cl.cam.ac.uk

1. ABSTRACT

Hidden copyright marks have been proposed as a solution for solving the illegal copying and proof of ownership problems in the context of multimedia objects. We show that the first generation of systems does not fulfil the expectation of users through a number of attacks that enable the information hidden by them to be removed or otherwise rendered unusable. We also propose a possible *benchmark* to compare these systems on a fair basis.

1.1 Keywords

Digital watermarking, fingerprinting, attacks, benchmark.

2. INTRODUCTION

The ease with which digital media could be copied led people to propose techniques for embedding hidden copyright marks and serial numbers in still images, video and audio. We formed the view that useful progress might come from trying to attack all these first generation schemes. In the related field of cryptology, progress was iterative: cryptographic algorithms were proposed, attacks on them were found, better algorithms were proposed, and so on. Eventually, theory emerged: fast correlation attacks on stream ciphers and differential and linear attacks on block ciphers, now help us understand the strength of cryptographic algorithms in much more detail than before.

Electronic copyright management schemes have been proposed as a solution to the copying problem. These schemes might be imposed in applications such as Digital Versatile Disk (DVD) and video-on-demand where the idea is that DVD players would refuse to copy files containing suitable copyright marks. But such schemes suffer from a number of drawbacks. They rely on the tamper-resistance of consumer electronics – a notoriously unsolved problem [1]. The tamper-resistance mechanisms being built into DVD players are fairly rudimentary and the history of satellite TV piracy leads us to expect the appearance of ‘rogue’ players which will copy everything[1]. Electronic copyright management schemes also conflict with applications such as digital libraries, where ‘fair use’ provisions are strongly entrenched. Another problem, according to Samuelson, is that ‘*Tolerating some leakage may be in the long run of interest to publishers*’ [2]. A European legal expert put it even more strongly: that copyright laws are only tolerated because they are not enforced against the large numbers of petty offenders [3].

Similar issues are debated within the software industry; some people argue, for example, that a modest level of amateur software piracy actually enhances revenue because people may ‘try out’ software they have ‘borrowed’ from a friend and then go on to buy it. Bill Gates’ view is significant: ‘*Although about three million computers get sold every year in China, people don't pay for the software. Someday they will, though. And as long as they're going to steal it, we want them to steal ours.* […] *Then we'll somehow figure out how to collect sometime in the next decade.*’ [4]

For all these reasons, we may expect leaks in the primary copyright protection mechanisms and wish to provide independent secondary mechanisms that can be used to trace and prove ownership of digital objects. Here too marking techniques are expected to be important.

3. COPYRIGHT MARKS

There are two basic kinds of mark: *fingerprints* and *watermarks*. One may think of a fingerprint as an embedded serial number while a watermark is an embedded copyright message. The first enables us to trace offenders, while the second can provide some of the evidence needed to prosecute them. It may ever, as in the DVD proposal, form part of the primary copy management system; but it will more often provide an independent back-up to a copy management system that uses overt mechanisms such as digital signatures.

[1] As a matter of fact techniques to bypass the territorial lock of certain DVD implementations are already available on the Internet.

In [5], we discussed various applications of fingerprinting and watermarking, their interaction, and some related technologies. Here, we are concerned with the robustness of the underlying mechanisms. What sort of attacks are possible on marking schemes? What sort of resources are required to remove marks completely, or to alter them so that they are read incorrectly? What sort of effect do various possible removal techniques have on the perceptual quality of the resulting audio or video?

The basic problem is to embed a mark in the digital representation of an analogue object (such as a film or sound recording) in such a way that it will not reduce the perceived value of the object while being difficult for an unauthorised person to remove. A first pass at defining robustness in this context may be found in a recent request for proposals for audio marking technology from the International Federation for the Phonographic Industry, (IFPI) [6]. The goal of this exercise was to find a marking scheme that would generate evidence for anti-piracy operations, track the use of recordings by broadcasters and others and control copying. The IFPI robustness requirements are as follows:

- the marking mechanism should not affect the sonic quality of the sound recording;
- the marking information should be recoverable after a wide range of filtering and processing operations, including two successive D/A and A/D conversions, steady-state compression or expansion of 10%, compression techniques such as MPEG and multi-band non-linear amplitude compression, adding additive or multiplicative noise, adding a second embedded signal using the same system, frequency response distortion of up to 15 dB as applied by bass, mid and treble controls, group delay distortions and notch filters;
- there should be no other way to remove or alter the embedded information without sufficient degradation of the sound quality as to render it unusable;
- given a signal-to-noise level of 20 dB or more, the embedded data channel should have a bandwidth of at least 20 bits per second, independent of the signal level and type (classical, pop, speech).

Similar requirements could be drawn up for marking still pictures, videos and multimedia objects in general. However, before rushing to do this, we will consider some systems recently proposed and show attacks on them that will significantly extend the range of distortions against which designers will have to provide defences, or greatly reduce the available bandwidth, or both.

4. ATTACKS

This leads us to the topic of attacks and here we present some quite general kinds of attack that destroy, or at least reveal significant limitations of, several marking schemes: PictureMarc 1.51 [7], SysCoP [8], SureSign [9], JK_PGS (É.P.F.L. algorithm, part of the European TALISMAN project), EIKONAmark [10], [11], Echo Hiding [20], Giovanni [18] and the N.E.C. method [13]. We suspect that systems that use similar techniques are also vulnerable to our attacks.

4.1 The jitter attack

Our starting point in developing a systematic attack on marking technology was to consider audio marking schemes. A simple and devastating attack on these schemes is to add jitter to the signal by removing samples or duplicating other. In fact most simple spread-spectrum based techniques are subject to this kind of attacks. Indeed, although spread-spectrum signals are very robust to distortion of their amplitude and to noise addition, they do not survive timing errors: synchronisation of the chip signal is very important and simple systems fail to recover this synchronisation properly. So, in general time scaling based attacks are very efficient against audio marking systems.

4.2 StirMark

Following this attack and after evaluating some watermarking software, it became clear that although many 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. This motivated the design of StirMark, initially implemented by Markus G. Kuhn and enhanced and maintained by the first author [14].

StirMark is a generic tool developed for simple robustness testing of image marking algorithms and other steganographic techniques. StirMark simulates a resampling process, i.e. it introduces the same kind of errors into an image as printing it on a high quality printer and then scanning it again with a high quality scanner. It applies a minor geometric distortion: the image is slightly stretched, sheared, shifted and/or rotated by an unnoticeable random amount and then resampled using Nyquist interpolation.

With those simple geometrical distortions we could confuse most marking systems available on the market.

More distortions – still unnoticeable – can be applied to a picture. We applied a global ‘bending’ and ‘random displacement’ to the image: in addition to the general bi-linear property explained previously, a slight deviation is applied to each pixel, which is greatest at the centre of the picture and almost null at the corners and to which is added a higher frequency displacement of the form $\lambda \sin(\omega_x x)\sin(\omega_y y) + n(x,y)$ – where n is a random number – is added (Fig. 1).

Finally a transfer function that introduces a small and smoothly distributed error into all sample values is applied. This emulates the small non-linear analogue/digital converter imperfection typically found in scanners and display devices.

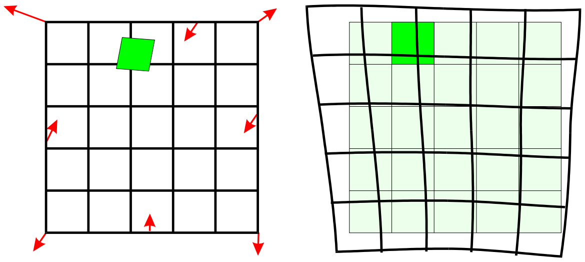


Fig. 1. We exaggerate here the distortion applied by StirMark 2 to still pictures.

In order for these distortions – which are practically unnoticeable as one can see from Fig. 2 – to be most effective, a medium JPEG compression is applied after the distortions.

We suggest that image-watermarking tools, which do not survive StirMark – with default parameters – should be considered unacceptably easy to break. This immediately rules out the majority of commercial marking schemes.

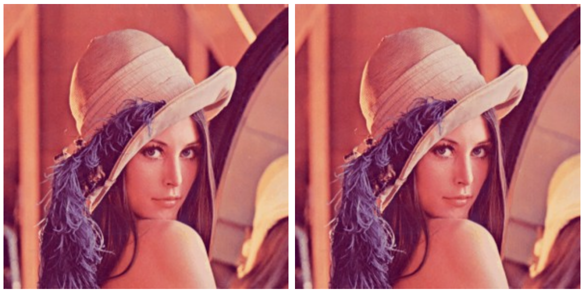


Fig. 2. ‘Lenna’ before and after StirMark used with default parameters.

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 Ó Ruanaidh and Pereira suggest using the Fourier-Mellin transform. However, the general theme of the attacks described above is that given a target marking scheme, we invent a distortion (or a combination of distortions) that will remove it or at least make it unreadable, while leaving the perceptual value of the previously marked object undiminished. We are not limited in this process to the distortions produced by common analogue equipment, or considered in the IFPI request for proposals cited above.

Table 1. Robustness tests for various digital watermarking products. Marks range from 0 to 20. The average is given in the bottom row. For each product 5 test images have been used and for each image 42 transformations have been applied using StirMark 2. Details for these transformations as well as the set of images used are given in the StirMark 2 package. Each image has been watermarked using the best parameters that do not give obvious and annoying distortions; the goal was fairness. Note that these results only reflect the robustness of the watermarking algorithm itself and not the whole copyright marking system, which might include registration process and other procedures.

	Digimarc 1.51	**SureSign 3.0 Demo**	**EikonaMark 3.01**	**JK_PGS 1.0 (Sun)**	**Giovanni 1.1.0.2**	**SysCoP 1.0R1**
GIF Conversion	20.00	20.00	20.00	20.00	12.00	16.00
Scaling	14.00	20.00	0.00	0.00	12.67	0.00
Cropping	20.00	20.00	0.00	8.00	3.00	0.00
Rotation & cropping	16.00	11.33	0.00	0.00	2.00	0.00
Rotation & scaling	16.67	12.00	0.00	0.67	2.00	0.00
JPEG	11.20	14.40	18.00	9.20	2.40	11.60
Filtering	20.00	20.00	20.00	20.00	12.00	16.00
Horizontal flip	20.00	20.00	0.00	0.00	0.00	0.00
StirMark 1.0	16.00	16.00	0.00	0.00	0.00	0.00
StirMark 2.2	0.00	0.00	0.00	0.00	0.00	0.00

It is an open question whether there is any marking scheme for which a chosen distortion attack cannot be found.

A number of methods claim to be 'robust'; but the criteria as well as the pictures used to 'prove' their robustness vary from one system to the other. This is not practical at all for comparison. So StirMark has been extended to provide a set of default transformations to a pre-defined set of pictures, such that any marking system can be tested and rated. This full benchmark test (piece of software freely available [14]) allows fair comparison of various digital watermarking techniques. Some results are presented in Table 1.

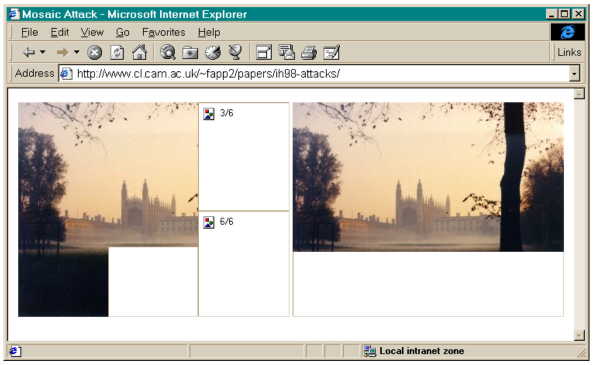
4.3 The mosaic attack

This point is emphasised by a 'presentation' attack, which is of quite general applicability and which possesses the initially remarkable property that a marked image can be unmarked and yet still rendered pixel for pixel in exactly the same way as the marked image by a standard browser.

The attack was motivated by a fielded automatic 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 watermark.

It consists of chopping an image up into a number of smaller subimages, which are embedded in a suitable sequence in a web page. Common web browsers render juxtaposed subimages stuck together, so they appear identical to the original image (Fig. 3). 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. The best that one can hope for is that the minimal size could be quite small and the method might therefore not be very practical.

There are other problems with such 'crawlers'. Java applets, ActiveX controls, etc. can be embedded to display a picture inside the browser; the applet could even de-scramble the picture in real time. Defeating such techniques would entail rendering the web page, detecting pictures and checking whether they contain a mark. An even more serious problem is that much current piracy is of pictures sold via many small services, from which the crawler would have to purchase them using a credit card before it could examine them. A crawler that provided such 'guaranteed sales' would obviously become a target.


Mosaic Attack - Microsoft Internet Explorer
File Edit View Go Favorites Help
Links
3/6
6/6
Local intranet zone

Address http://www.cl.cam.ac.uk/~fapp2/papers/ih98-attacks/

Fig. 3. Screen-shot of a web browser while downloading an image after the *mosaic attack*. This attack chops a watermarked image into smaller images, which are stuck back together when the browser renders the page. We implemented software 2Mosaic that reads a JPEG picture and produces a corresponding mosaic of small JPEG images as well as the necessary HTML code automatically [15]. In some cases downloading the mosaic is even faster than downloading the full image! In this example we used a 350×280-pixel image watermarked using PictureMarc 1.51. Photography: Kings' College Chapel, courtesy of John Thompson, JetPhotographic, Cambridge.

4.4 A general attack on audio marking

Audio restoration techniques have been studied for several years and have proved to be very useful to remove localised degradations (clicks, crackles, scratches, etc.) from old recordings [16], [17]. After finding the local degradations, these methods basically ignore the bad samples and interpolate the signal using the neighbouring ones.

Our attack is based on this idea: the signal is reconstructed block by block using the original data. The method we used assumes that the recorded data x is the realisation of a stationary autoregressive (AR) process of order p, i.e.

$$x_n = \sum_{k=1}^{p} a_k x_{n-k} + e_n \qquad n = p+1,\ldots,N \qquad (1)$$

where $\mathbf{e} = \left[e_{p+1},\ldots,e_N\right]^T$ is the 'excitation' noise vector. We suppose that we want to reconstruct a block of l consecutive samples starting at sample $m+1$ and assume to be unknown. Estimators for both a and x are chosen such that they minimise the quadratic error $E = \mathbf{e}^T\mathbf{e}$ which is a function of the unknown samples

$\mathbf{x}_u = [x_{m+1}, \ldots, x_{m+l}]^T$ and the unknown AR parameters $\mathbf{a} = [a_1, \ldots, a_p]^T$.

Minimisation of E is non-trivial since it involves non-linear fourth order unknown terms but a sub-optimal solution to the above problem can be used.

First E is minimised with respect to $\mathbf{a}$ by taking an arbitrary initial estimate for $\mathbf{x}_u$ (typically zero) in order to obtain an estimate $\hat{\mathbf{a}}$ of $\mathbf{a}$. If we note $\mathbf{x}_1 = [x_{p+1}, \ldots, x_N]^T$, then equation (1) can be written $\mathbf{e} = \mathbf{x}_1 - \mathbf{B}(\mathbf{x})\mathbf{a}$ and $\hat{\mathbf{a}}$ is given by:

$$\mathbf{B}^T\mathbf{B}\hat{\mathbf{a}} = \mathbf{B}^T\mathbf{x}_1 \qquad (2)$$

Then E is minimised with respect to $\mathbf{x}_u$ and using $\hat{\mathbf{a}}$. Equation (1) is written as $\mathbf{e} = \mathbf{D}_k(\mathbf{x})\mathbf{x}_k + \mathbf{D}_u(\mathbf{x})\mathbf{x}_u$ where $\mathbf{x}_k$ is the vector of known samples. After minimisation, the reconstructed block $\hat{\mathbf{x}}_u$ is given by:

$$\mathbf{D}_u{}^T\mathbf{D}_u\hat{\mathbf{x}}_u + \mathbf{D}_u\mathbf{D}_k\mathbf{x}_k = 0 \qquad (3)$$

These two steps can be iterated to get better results but it seems that one iteration is usually enough. For the attacks we just increase *m* in steps of the block length *l* and compute for each step an estimated block which is appended to the others. We end up with a fully reconstructed signal.

Other and better interpolation algorithms are available, but the least square AR interpolation technique, we briefly summarised, gives satisfactory results if the blocks are relatively small, up to 80 samples [16], [17].

Although we used it only against BlueSpike's method [18], this attack is quite general and could also be used against image marking too. Similar algorithms for image reconstruction are given in [19].

4.5 Attack on echo hiding

Echo hiding hides information in sound by introducing echoes with very short delays [20]. It relies on the fact that we cannot perceive short echoes (say 1 ms) and embeds data into a cover audio signal by introducing an echo characterised by its delay τ and its relative amplitude α. By using two types of echo it is possible to encode ones and zeros. For this purpose the original signal is divided into chunks separated by spaces of pseudo-random length; each of these chunks will contain one bit of information.

The echo delays are chosen between 0.5 and 2 milliseconds and the best relative amplitude of the echo is around 0.8. According to its creators, decoding involves detecting the initial delay and the auto-correlation of the cepstrum of the encoded signal is used for this purpose.

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 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 [21]. 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 we note Φ_{xx} the power spectrum of x then $\Phi_{yy}(f) = \Phi_{xx}(f)(1 + 2\alpha\cos(2\pi f\tau) + \alpha^2)$ whose logarithm is approximately $\ln\Phi_{yy}(f) \approx \ln\Phi_{xx}(f) + 2\alpha\cos(2\pi f\tau)$. This is a function of the frequency *f* and taking its power spectrum raises its 'quefrency' τ, that is the frequency of $\cos(2\pi f\tau)$ as a function of *f*. The auto-covariance[2] of this later function emphasises the peak that appears at 'quefrency' τ.

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 milliseconds. 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 than 0.7 we could use $\tilde{\alpha}$ – an estimation of α – drawn from, say a normal distribution centred on 0.8. It was not really successful, so our next attack was to iterate: we re-apply the detection function and vary $\tilde{\alpha}$ to minimise the residual echo. We could obtain successively better estimators of the echo parameters and then remove this echo. When the detection function cannot detect any more echo, we have got the correct

[2] $\mathrm{C}(x) = \mathrm{E}((x - \bar{x})(x - \bar{x})^*)$

value of $\widetilde{\alpha}$ (as this gives the lowest output value of the detection function).

4.6 Protocol considerations

The main threat addressed in the literature is an attack by a pirate who tries to remove the watermark directly. As a consequence, the definition commonly used for robustness includes only resistance to signal manipulation (cropping, scaling, resampling, etc.). Craver *et al.* show that this is not enough by exhibiting a 'protocol' level attack [22].

The basic idea is that many schemes provide no intrinsic way of detecting which of two watermarks was added first: the process of marking is often additive, or at least commutative. So if the owner of the document d encodes a watermark w and publishes the marked version $d + w$ and has no other proof of ownership, 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 $d + w - w'$.

Craver *et al.* argue for the use of information-losing marking schemes whose inverses cannot be approximated closely enough. However, 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 time-stamping and notarisation to prevent attacks of this kind.

Registration mechanisms have not received very much attention in the copyright marking literature to date. The existing references such as [23], [24], [26] and [27] mainly focus on protecting the copyright holder and do not fully address the rights of the consumers who might be fooled by a crooked reseller. Moreover a good registration and trading mechanism cannot be based on a weak marking technique.

4.7 Implementation considerations

The robustness of embedding and retrieving techniques is not the only issue. Most 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 [28].

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 the most widely used picture marking scheme, PictureMarc, which is bundled with Adobe Photoshop and Corel Draw. This attack [29] exploited weaknesses in the implementation rather than the underlying marking algorithms, even although these are weak (the marks can be removed using StirMark).

Each user has an ID and a two-digit password, which are issued when she registers with Digimarc and pays for a subscription. The correspondence between IDs and passwords is checked using obscure software in the implementation and although the passwords are short enough to be found by trial and error, the attack first uses a debugger to break into the software and disable the password checking mechanism. We note in passing that IDs are public, so either password search or disassembly can enable any user to be impersonated.

A deeper examination of the program also allows a villain to change the ID and thus the copyright 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 by making it longer, but there is no obvious solution to the disassembly attack. If tamper resistant software [30] cannot give enough protection, then one can always have an online system in which each user shares a secret embedding 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 (which can be done with a signature) and the embedding or hiding operation.

4.8 Robustness against insiders

Although we can do public-key steganography – hiding information so that only someone with a certain private key can detect its existence [31] – 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. But if the stego key is widely released (e.g. as part of a global law enforcement or in equipment) it is very likely to leak over time.

Another problem is that a public decoder can be used by the attacker; he can remove a mark by applying small changes to the image until the decoder cannot find it anymore. This was first suggested by Perrig in [27]. In [32] a more theoretical analysis of this attack is presented as well as a possible countermeasure: randomising the detection process. One could also make the decoding process computationally expensive. However neither approach is really satisfactory in the absence of tamper-resistant hardware.

Unless a breakthrough is made, applications that require the public verifiability of a mark (such as DVD) appear doomed to operate within the constraints of the available tamper resistance technology (one could use a number of marks with keys revealed in succession[3]), or to use a central 'mark reading' service. This is evocative of cryptographic key management prior to the invention of public key techniques.

5. CONCLUSION

We have demonstrated that the majority of copyright marking schemes in the literature are vulnerable to attacks involving the introduction of sub-perceptual levels of distortion. In particular, many of the marking schemes in the marketplace provide only a limited measure of protection against attacks. Most of the image marking systems are defeated by StirMark, a simple piece of software that we have placed in the public domain [14]. We have also shown specific attacks some audio marking systems.

This experience confirms our hypothesis that steganography would go through the same process of evolutionary development as cryptography, with an iterative process in which attacks lead to more robust systems.

Our experience in attacking the existing marking schemes has convinced us that any system which attempted to meet all the accepted requirements for marking (such as those set out by IFPI) would fail: if it met the robustness requirements then its bandwidth would be quite insufficient. This is hardly surprising when one considers that the information content of many music recording is only a few bits per second, so to expect to embed 20 bits per second against an opponent who can introduce arbitrary distortions is very ambitious.

Our more general conclusion from this work is that the 'marking problem' has been over-abstracted; there is not one 'marking problem' but a whole constellation of them. We do not believe that any general solution will be found. The trade-offs and in particular the critical one between bandwidth and robustness, will be critical to designing a specific system.

We already remarked in [5] on the importance of whether the warden was active or passive – that is, whether the mark needed to be robust against distortion. In general, we observe that most real applications do not require all of the properties in the IFPI list. For example, when auditing radio transmissions, we only require enough resistance to distortion to deal with naturally occurring effects such as multipath. Many applications will also require supporting protocol features, such as the timestamping service that we mentioned in the context of reversible marks.

So we do not believe that the intractability of the 'marking problem' is a reason to abandon this field of research. On the contrary; practical schemes for most realistic application requirements are probably feasible and the continuing process of inventing schemes and breaking them will enable us to advance the state of the art rapidly.

Finally, we suggest that the real problem is not so much inserting the marks as recognising them afterwards. Thus progress may come not just from devising new marking schemes, but in developing ways to recognise marks that have been embedded using the obvious combinations of statistical and transform techniques and thereafter subjected to distortion. The considerable literature on signal recognition may provide useful starting points.

6. ACKNOWLEDMENTS

The first author is grateful to Intel Corporation for financial support under the grant 'Robustness of Information Hiding Systems'.

7. REFERENCES

1 Ross J. Anderson and Markus G. Kuhn. Tamper Resistance – A Cautionary Note. In *Second USENIX Workshop on Electronic Commerce*, pages 1–11, Oakland, CA, USA, November 1996. ISBN 1-880446-83-9.

2 Pamela Samuelson. Copyright and Digital Libraries. *Communications of the ACM*, pages 15–21, 110, 38(4), April 1995.

3 Alastair Kelman. Electronic Copyright Management – The Way Ahead. Security Seminars, University of Cambridge, 11 February 1997.

4 *The Bill & Warren Show*. Fortune, page 44, 20th July 1998. Public dialogue between Bill Gates, founder and

[3] This is what happens for bank note printing in some countries: notes have a number of 'anti-copy' features, which are publicised in succession. Forgers are less likely to reproduce them since they do not know their existence.

CEO of Microsoft Corporation, and Warren Buffett, chairman of Berkshire Hathaway Inc.

5 Ross J. Anderson and Fabien A.P. Petitcolas. On The Limits of Steganography. *IEEE Journal of Selected Areas in Communications (J-SAC) – Special Issue on Copyright & Privacy Protection*, pages 474–481, 16(4), May 1998. ISSN 0733-8716.

6 International Federation of the Phonographic Industry. Request for Proposals – Embedded Signalling Systems Issue 1.0. 54 Regent Street, London W1R 5PJ, June 1997.

7 Geoffrey B. Rhoads. Steganography methods employing embedded calibration data. Digimarc Corporation. US Patent 5,636,292, 3 June 1997.

8 E. Koch and J. Zhao. Towards Robust and Hidden Image Copyright Labeling. In *Workshop on Nonlinear Signal and Image Processing*, pages 452–455, Neos Marmaras, Greece, 20–22 June 1995. IEEE.

9 Signum Technologies – SureSign digital fingerprinting. http://www.signumtech.com/, October 1997.

10 Alpha Tec Ltd. EIKONAmark. http://www.generation.net/~pitas/sign.html, October 1997.

11 I. Pitas. A method for signature casting on digital images. In *International Conference on Image Processing*, volume 3, pages 215–218, September 1996.

12 Ross J. Anderson, editor. *Information hiding: first international workshop*, volume 1174 of *Lecture notes in Computer Science*. Springer Verlag, Berlin, Germany, May 1996. ISBN 3-540-61996-8.

13 Ingemar J. Cox, Joe Kilian, Tom Leighton and Talal Shamoon. A Secure, Robust Watermark for Multimedia. In Anderson [12], pages 183–206.

14 Fabien A.P. Petitcolas and Markus G Kuhn. StirMark 2. http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/, November 1997.

15 Fabien A.P. Petitcolas. 2Mosaic. http://www.cl.cam.ac.uk/~fapp2/watermarking/2mosaic/, October 1997.

16 Saeed Vahed Vaseghi. *Algorithms for restoration of archived gramophone recordings*. PhD thesis, Emmanuel College, University of Cambridge, UK, February 1988.

17 Simon J. Godsill, Peter J.W. Rayner and Olivier Cappé. Digital audio restoration. In Mark Kahrs and Karlheinz Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Electroacoustics*. Kluwer Academic Publishers, 1998.

18 Giovanni audio marking software. Blue Spike company. http://www.bluespike.com/, May 1998.

19 Raymond Veldhuis. *Restoration of lost samples in digital signals.* International Series in Acoustics, Speech and Signal Processing. Prentice Hall, Hertfordshire, UK, 1990.

20 Daniel Gruhl, Walter Bender and Anthony Lu. Echo hiding. In Anderson [12], pages 295–315.

21 Bruce P. Bogert, M.J.R. Healy and John W. Tukey. The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Ceptstrum and Saphe Cracking. In M. Rosenblatt, editor, *Symposium on Time Series Analysis*, pages 209–243, New-York, USA, 1963. John Wiley & Sons, Inc.

22 Scott Craver, Nasir Memon, Boon-Lock Yeo and Minerva M. Yeung. Resolving Rightful Ownerships with Invisible Watermarking Techniques: Limitations, Attacks, and Implications. *IEEE Journal of Selected Areas in Communications (J-SAC) – Special Issue on Copyright & Privacy Protection*, pages 573–586, 16(4), May 1998. ISSN 0733-8716.

23 Marc Cooperman and Scott A. Moskowitz. Steganographic method and device. The DICE Company. US Patent 5,613,004, 18 March 1995.

24 Alexander Herrigel, Adrian Perrig and Joseph J.K. Ó Ruanaidh. A Copyright Protection Environment for Digital Images. In *Verläßliche IT-Systeme '*97, Albert-Ludwigs Universität, Freiburg, Germany, October 1997.

25 David Aucsmith, editor. *Information hiding: second international workshop*, volume 1525 of *Lecture Notes in Computer Science*, Portland, Oregon, USA, 1998. Springer Verlag, Berlin, Germany. (to appear)

26 Alexander Herrigel, Joseph J.K. Ó Ruanaidh, Holger Petersen, Shelby Pereira, and Thierry Pun. Secure copyright protection techniques for digital images. In Aucsmith [25], pages 170–191.

27 Adrian Perrig. A copyright protection environment for digital images. Diploma dissertation, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, February 1997.

28 Ross J. Anderson. Why cryptosystems fail. *Communications of the ACM*, 37(11):32–40, November 1994.

29 Anonymous (zguan.bbs@bbs.ntu.edu.tw). Learn cracking IV – another weakness of PictureMarc. news://tw.bbs.comp.hacker mirrored on http://www.cl.cam.ac.uk/~fapp2/watermarking/image_watermarking/digimarc_crack.html, August 1997. Includes instructions to override any Digimarc watermark using PictureMarc.

30 David Aucsmith. Tamper resistant software: An implementation. In Anderson [12], pages 317–333.

31 Ross J. Anderson. Stretching the limits of steganography. In Anderson [12], pages 39–48.

32 Jean-Paul M.G. Linnartz and Marten van Dijk. Analysis of the sensitivity attack against electronic watermarks in images. In Aucsmith [25], pages 259–273.