Evaluating the Feasibility of Digital Watermarking To Enforce Music Copyright

CS 588: Final Project

Rob Farraher
Ken Pickering
Lim Vu

Date Due: 12/04/01

Introduction

With the tremendous increase in the usage and creation of digital media today, there is an ever-growing need for new methods of protecting intellectual property rights. This need results from the ease with which any person can make identical copies of a digital file to share and distribute. In particular, people abuse copyright protections everyday by participating in music and file sharing networks such as Gnutella and Morpheus. As most are aware, MP3 technology began this file-sharing phenomenon by providing a means for copyrighted music to be freely available in an easily downloadable, compressed media. Major American companies realize the potential in downloadable music but run into the problem of copy-protecting digital music after a user has downloaded it. If one user licenses and downloads a song, companies must somehow prohibit others from copying that music. As a result, researchers have developed several watermarking schemes to protect the integrity of digital music and thwart anyone from illegally copying and distributing that music.

Digital watermarking technology is scalable to the value of the contents it protects. Artists and corporations can choose from merely authenticating their work to digitally tracking it using varying watermarking technologies [1]. Watermarking also provides the possibility of fully utilizing digital media while continuing to uphold copyright regulations. However, all current digital watermarking schemes today have inherent weaknesses that make them vulnerable to attacks. Also, another innate obstacle must be overcome before digital watermarking is a viable form of copyright protection: the universal acceptance of a particular working watermark protection scheme. Software companies must agree to provide digital music extractors that use an agreed upon

watermarking scheme. Hardware companies also must agree to provide digital music players that implement the agreed upon watermark scheme. Lastly, there must also be a general consumer acceptance. Thus, due to lack of a successful watermarking scheme as well as difficulty in gaining universal acceptance of a single watermark, we believe digital watermarking is currently infeasible.

For this project, we will provide a general overview of digital watermarking technology. One particular digital watermarking initiative we plan to study will be SDMI (Secure Digital Music Initiative). In addition, we will examine the conditions that must occur in the general market for a watermarking scheme to be successfully implemented. Overall, we believe this project will provide a complete analysis of the state of digital watermarking today as well as explain the difficulties that exist in bringing it to fruition.

Overview

Digital watermarking is an active field of research that covers not just protection of audio but all digitally created content. It involves embedding a digital signal into any content that can later be detected and extracted. Digital watermarks can be placed with two general categories; robust and fragile. Robust watermarks can withstand manipulation of the content it is embedded within. Thus, compression, copying, distortion, or any kind of general manipulation of the digital content leaves the watermark intact. Fragile watermarks are the opposite in that any modification of the data will leave the watermark undetectable [2]. Watermarks can also be classified into another two major categories: Perceptible and Imperceptible. Perceptible watermarks are visible on the media. Imperceptible are invisible to the viewer of the document, and instead are

hidden in some invisible way. The main focus of this project is to look at imperceptible watermarks, since they are the type used to protect digital music in most cases. Since you want to user to be able to listen to music with no obvious distortion, or even knowledge the watermark exists. There are several applications for watermarks that include:

- Ownership assertion – Authors can assert that their digital work belongs to them by embedding a watermark created with their cryptographic private key. Thus, the author is the one and only person that can produce an image with the same watermark as in the copies they have distributed.
- Fingerprinting – Authors can prevent unauthorized copying and distribution of their work by embedding a distinct watermark into every copy of their data. Hence, if any illegal copies are found, the originating copy will always be known and offenders traceable.
- Authentication and integrity verification – Authors can embed what is similar to a cryptographic hash into their digital work. This hash is invisible and inseparable from the data in the form of a watermark.
- Content labeling – Watermarks are capable of carrying information. Thus, an artist can embed information describing the digital work they have created that someone later can extract.
- Usage control – Artists can insert a watermark that indicates the number of copies permitted by a user. Thus, with specialized hardware, the author can enforce the user to a limited number of copies of the digital work.
- Content protection – A visible watermark can be inserted into a digital work so that a preview or free showing is commercially worthless [2].

Ownership Assertion

One use of watermarks is to assert ownership or fingerprint documents for tracking/record keeping purposes. These are usually perceptible watermarks and a visual part of the media, often embedded directly into the document's content.

Ownership assertion watermarks are much like digital signatures used in emails. It makes sure the recipient knows that the producer of the document is, in fact, who he/she says he/she is. The basic system involves the creator using a private key to produce the unique watermark within the document. The recipient can verify the authenticity of the document by checking the sender's public key. This differs from digitally signing emails in that the actual contents of the message are not "hidden" or changed at all. The only thing different about the original and the sent piece of information is the addition of a watermark.

Fingerprinting

Fingerprinting is used in much the same way. Different keys are used to produce watermarks in the media being sent out to various sources. These keys are noted. If the document is copied or leaked out, the source can be determined because the watermark is still inside it. Usually, these can be either perceptible or imperceptible. If you are aware someone is leaking information, and wanted to find the source, invisible watermarks could be used. The people copying the information would have no clue as to the existence of the watermark. A visible watermark could be used to prohibit copying in advance, since any distributed copies would be clearly marked with the whatever unique identifier the creator wanted to include.

Authentication/Verification

The authentication and verification of digital media is an important application of digital watermarking. With the insecurity of networks today, the authentication of data transferred over either a secure or insecure channel is important. Thus, the content source as well as the content itself must be verified. For example, a commercial content provider decides to sell rare, unattainable music over the Internet. The provider's guarantee to its customers is that they are providing legitimate copies of this rare music require a strong authentication mechanism. Another example that demonstrates the need for authentication is the distribution of anti-virus updates. Malicious attackers can potentially intercept data transmissions or pretend to be trusted file servers and replace the anti-virus update with a damaging virus. However, with an authentication mechanism in place this malady could be prevented. Consequently, a strong argument can be made for authentication through watermarking. Qualities that can define a watermark, such as robustness, persistence, and unobtrusiveness, make it a superior candidate.

Authentication of data can also be done with conventional cryptography; however, the advantage gained with watermarks is that authentication is inseparably bound to the content. Thus, eliminating the logistical problems involved with conventional cryptographic authentication. With a watermark authentication, the originator of the content is verified when a watermark is extracted using the originator's unique key. The content itself is verified by checking the extracted watermark's integrity. The types of watermark used in authentication must have the ability to detect and localize any changes made to the content it is protecting.

Content Labeling

Content labeling involves embedding a descriptive annotation into content using a watermark. This binds content description with the content itself. The annotation adds value to the content in terms of beneficial information conveyed. For instance, a digital picture could contain information such as author, resolution, subject, and title. This information is inseparable from the content, thus, eliminating the need to handle extra description data. Also, the content annotation can be used to sort and organize any content with a label. A digital watermark being used as a content label must not degrade the quality of the content. Thus, this limits the amount of data that can be inserted into the content.

Usage Control

Artists can insert a watermark that indicates the number of copies permitted by a user. Thus, with specialized hardware or software, the author can enforce the user to a limited number of copies of the digital work. The existence of a watermark may certify the legitimacy of a particular document or digital media. However, if the production or replication of such a document is not controlled, then anyone with access to the material may make a copy that seems as legitimate as the original, hence without copy protection, efforts toward legitimacy are moot.

Consider a watermark that is able to adjust the number of copies of the document it is contained within can be made. Obviously, this number would have to be adjusted each time a copy was made. A master watermarked document may have a certain

number of allowed copies, none of which (the copies) are capable of being copied. When the master document is copied, the count of allowed copies is decremented and the watermarked saved. When the number drops below zero, the watermark becomes perceptible and the document is no longer viewable/usable. Copies of the master watermark could be spawned with a copy count of zero making them viewable but not able to make copies without rendering their watermark's perceptibility and making themselves unusable.

If some sort of control mechanism as suggested above were implemented, there would have to be an associated viewer that would have the unique ability of checking the watermark and using/viewing the document. The security of this type of system would rely on the secrecy of both the viewer software/hardware's algorithm for viewing the protected document and the means by which a digital copy is made. If a direct digital copy can be made, then anyone could copy a document with a count of zero (which should not able to be copied) and it would seem just as legitimate as the original. This type of system relies on the exclusive control of copying and interpretation of documents by the implementer of the watermarking scheme.

While it may be possible to ensure that only certain software or hardware can interpret a particular format, copying in a digital world is difficult to restrict. This is not satisfactory for purely digital media as they can be copied very easily and efficiently, but if a certain type of format, not purely digital, were implemented, this may be possible. The difficulty in controlling usage of media through digital watermarking is hindered by the inability to restrict absolute copying of purely digital media, if in some way the digital

media were incorporated into an form that was not able to make perfect copies, the watermark could succeed.

If perhaps a scheme were implemented where not only did a certain file need to have a valid watermark but that watermark needed to be specific to a certain person/player. For instance if Alice was licensed to make X copies to hardware that recognizes her own watermark and Y copies to hardware that recognizes anyone else's watermark, the Y copies would have to be more than just a direct digital copy of Alice's file. Since the copies need to have the watermark of the person for which they are being copied, Alice's copy of the watermarked file wouldn't play for any device but Alice's. In this way, copying could be restricted to the number of copies allowed could be controlled by the copying/watermark adjusting mechanism. However, this scheme would also require that an un-watermarked copy of the file could never be obtained, either by removing the watermark or by obtaining the file before the watermark were applied. If the un-watermarked file were ever encountered it could be copied and played by players not requiring a watermark. Also, any user could make copies before adding their own watermark, assuming they have that capability of adding their own watermark.

In summary, in a user-specific watermarking scheme were implemented, if users can add their own watermark or use files that do not have watermarks then they must never be able to access the un-watermarked file or remove the watermark. This scheme also has the pitfalls of consumer rejection, as people may not desire the restriction or the necessity that their music has to be watermarked specific to their devices before they can access it.

Content Protection

Content protection is when a perceptible watermark is inserted into a digital work so that a preview or free showing is commercially worthless [2]. In this way, a digitally watermarked document can be ruined by the perceptibility of the watermark. Thus, restricting unauthorized commercial reproduction or distribution. This is useful for certain documents that require very high quality such as video or audio. If a certain video has a watermark, it can be viewed for certain uses such as the demonstration of a particular hardware setup without the risk of the audio or video document being used for anything other than the display of the hardware. This is true as the document has sufficient degradation. This use of the watermark also requires that it be irremovable and robust. Thus, the degradation of the content is permanent and its value worthless.

A Basic Watermark: Echo-Hiding

Most of the techniques used in digital watermarking for audio files utilize echo-hiding or some kind of noise addition in the time domain. This exploits the temporal and/or spatial masking models of the human auditory system (HAS). Anthony Lu, Daniel Grohl, and Walter Bender explain, "The HAS perceives over a range of power greater than one billion to one and a range of frequencies greater than one thousand to one. Sensitivity to additive random noise is also acute." However, even with the high acuity of the HAS, it has "holes" in its perceptive range where data may be hidden. Lu explains further, "While the HAS has a large dynamic range, it often has a fairly small differential range. As a result, loud sounds tend to mask out quiet sounds. Additionally, while the HAS is sensitive to amplitude and relative phase, it is unable to perceive

absolute phase. Finally, there are some environmental distortions so common as to be ignored by the listener in most cases [7]."

Lu goes on to describe the encoding process used with echo hiding. First, the properties of an echo are defined; initial amplitude, decay rate, one "offset", and zero "offset" (offset + delta). The offset (delay) of the echo is of significant importance because the human ear at a certain point will not hear the original signal and echo but a single distorted signal. Thus, the encoding process would degrade the quality of the file significantly. Below figure one, illustrates these parameters.

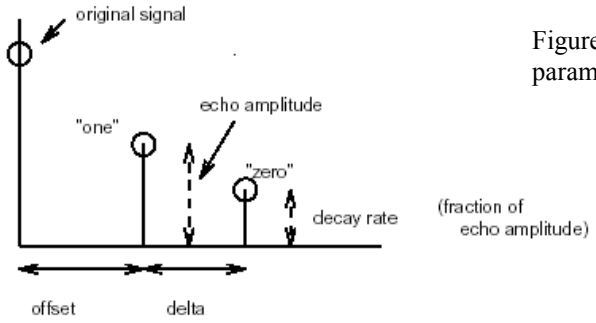

Figure 1: adjustable parameters [7]

The encoding process takes place by using discrete time exponentials (with delay between impulses) using only two impulses. One pulse copies the original signal; the other pulse represents the echo. To process a binary one or zero the two system functions in figure two are used. Thus, the binary data to be inserted as a watermark is encoded with either system function (kernel). Hence, with the delay between the original signal and echo are dependent upon the data and system function used.

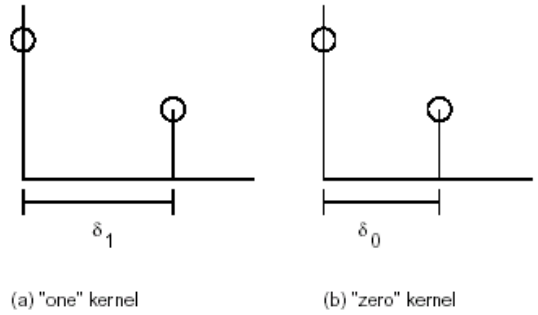

Figure 2: Echo system functions (kernels) [7]

Figure three below demonstrates the encoding process overall.

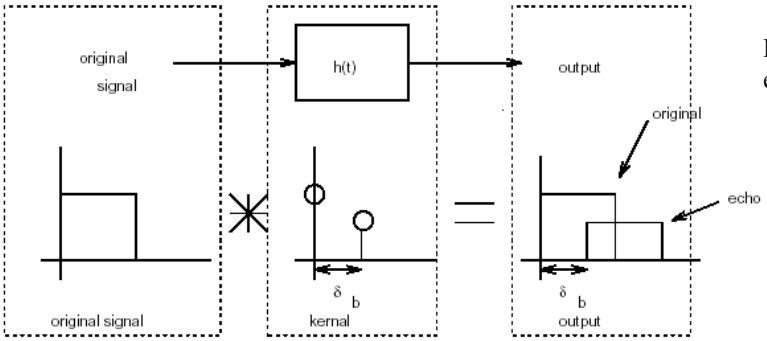

Figure 3: The echo-hiding encoding process [7]

Extraction or detection of the embedded watermark requires detecting the spacing between the echoes. Thus, the magnitude of the autocorrelation of the encoded signal's cepstrum must be examined [7].

New digital watermarking schemes and techniques are being developed everyday. Even though this are of research is relatively new, significant progress has already been made. Hopefully, in the future, a truly robust, imperceptible, and persistent watermarking scheme will be developed.

Cracking SDMI

SDMI provides an excellent example of the infeasibility of implementing a working watermarking scheme for copyright protection. A large conglomerate of over

150 international companies that boasts members from the worldwide recording, consumer electronics, and information technologies companies known as the Secure Digital Music Initiative (SDMI) has banded together to discover an open specification for protecting digital music distribution. Their plan involves production of software that will recognize their watermarking schema, and only play if a valid license is owned, as well as production of hardware devices that can only play purchased music [3]. It has been the largest attempt yet to implement a watermarking scheme in order to protect digital music.

Recently, SDMI held an open challenge to teams of researchers from Princeton, Rice, and Xerox research labs to evaluate their encryption techniques. The Princeton team discusses their luck at breaking the five types sent to them (known as A, B, C, and F). They define the types of attacks they used into two types. Type-1 exploits weaknesses found in a blind attack, while Type-2 uses the embedded mechanism to look at known inputs and the corresponding outputs. Type-2 is similar to a plain-text/cipher-text attack in cryptography, while Type-1 is more akin to a brute-force method. By blocking certain frequencies from a few of the watermarked files, they were able to bypass the embedded protection [4]. After SDMI completed its test implementations of several prototype-watermarking schemes, they allowed the public three weeks to attack and break the watermarks proposed. They provided the public with three files for each watermark; an unwatermarked file (sample-1), file one watermarked (sample-2), and another watermarked file (sample-3). Below in figure four, a pictorial example of the SDMI provided files is given. The testers were legally bound to never disclose their attacks on the challenges and were prevented from doing research beyond the three-week limit.

This weakened the reality of the SDMI challenges because would-be attackers would not be limited to such constraints [5].

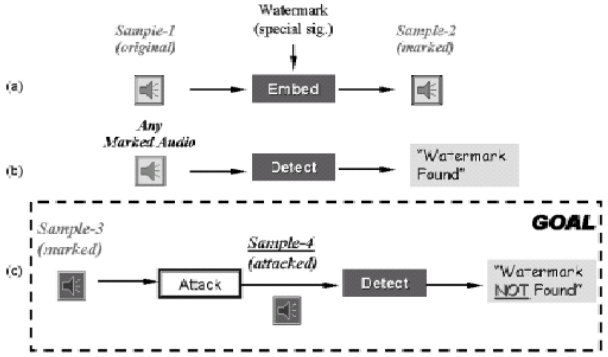

Figure 4: Four each of the four watermark challenges, Sample-1, sample-2, and sample-3 are provided by SDMI. Sample-4 is generated by participants in the challenge and submitted to SDMI for testing. (Excerpt and figure provided by [5])

In September 2000, Professor Edward Felten and others from Princeton University responded to SDMI's challenge to analyze provided watermarked files. However, unlike the other challengers in the SDMI competition, Felten decided to not abide by SDMI's legal constraints of non-disclosure and published his findings. After his attempt to disclose his findings in a USENIX conference, SDMI followed through and sued him. Felten instead preemptively litigated citing that SDMI interfered with his academic research [6].

In challenge A, the challenge was to take the provided un-watermarked file, the watermarked file, and another watermarked file and learn enough about the watermarking technology to remove from the difference between the original (1) and the watermarked file (2) to remove the watermark from the third file (3). Professor Felten found the watermark implanted via a Fast Fourier Transform (FFT) and used the following to remove the watermark from the third file [5].

```
while(framesLeftInSong()) {
        Y = FFT(nextFrame(markedMusicFile));
        X = FFT(nextFrame(unmarkedMusicFile));
        H = elementDivider(Y,X);
```

```
        Z = FFT(nextFrame(otherMarkedFile);
        R = elementwiseDivider(Z,H);
        outputFrame(IFFT(R));
}
```

What the above code does is compare the FFT of the music files 1 and 2 (original and watermarked file) in a divide, computes the FFT of the different watermarked file (3), and divides the difference found between 1 and 2 from 3. Then the inverse FFT is performed on the result. This output was confirmed by SDMI as being a successful removal of the watermark [5].

Felten successfully broke challenge B using a similar attack as in A. He first did a short-time FFT of both the watermarked (2) and unwatermarked (1) files and discovered notches within the 2 at around 2800Hz and 3500Hz. Felten's attack involves filling these notches with random but bounded coefficient values in the other provided watermarked sample (3). This output was the confirmed by SDMI as a successful attack [5]. Below in figure five is an excellent illustration of Felton's attack on challenge B.

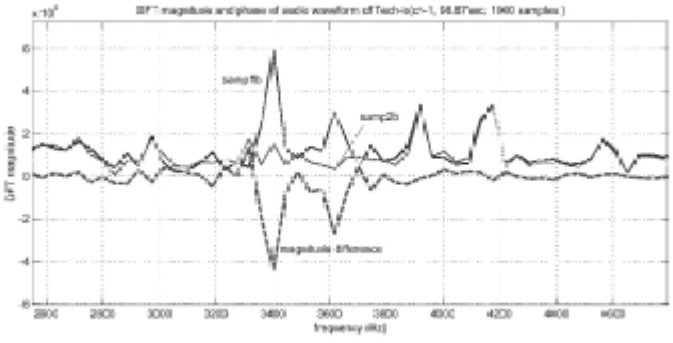

Figure 5: Challenge B: FFT magnitudes of the matching watermarked (samp1b) and unwatermarked (samp2b) files and their difference for 1000 samples at 98.67 sec. (Excerpt and figure provided by [5])

Challenge C was successfully attacked by shifting the pitch of a watermarked sample about a quartertone and then by passing this signal through a band-stop filter. This attack was performed around the frequencies were Felten noticed narrowband bursts (1350Hz). In addition to removing the watermark, the attack also passed the "golden-

ear" test performed by SDMI after the challenge. The "golden-ear" test is done to determine if there is no degradation of the audio file [5].

Challenge F, the last in the series proposed by SDMI, used methods learned when breaking watermark A. Felten was able to mount a successful attack with warping the time axis by inserting a periodically varying delay [5].

With Felten successfully attacking and rendering SDMI useless, the relative infancy of digital watermarking as a copyright protection is revealed. The large-scale of SDMI also proves that even with strong incentive and financial backing, digital watermarking is very difficult to implement properly.

Watermarking Solutions

The main problem with watermarking MP3's lies in their wide acceptance and the fact that MP3 is an "open" format. As long as people are able to produce and play MP3's without watermark verification, several problems arise. The "plaintext" or unwatermarked version will be available for comparison to the watermarked version. Also, if people can rip unwatermarked songs from CD's, then they can easily be distributed. If MP3 players will play these unwatermarked versions, then piracy is still very much a problem. In order for a system such as this one to work, there are a few solutions, although they may not be extremely feasible.

One such solution would be to switch to a proprietary format that has watermarks built into the system. When someone rips a song in this new format, it would automatically insert a watermark for only that person's use. The only players available for the format would have watermark usage control built in. Microsoft Windows has a

built in media format (.wmv) that produces compressed music similar to that of MP3. This is one implementation of a possible solution, although they make watermarking "optional" when you rip CD's. This is just as bad as the old system, because no user would inject a watermark if they did not have to, since they may want to exchange music with their friends.

The only other real solution to this problem would be to produce all new CD's with a very difficult watermarking algorithm built in. An individual algorithm could be used for each different type of CD made. Although it would not be "hack proof", making a series of difficult algorithm would reduce the amount of piracy. The overhead for this system would be pretty large, as the players would have to know which CD's use which algorithm and which CD's are licensed for the user. However, record companies would have to realize that copy protecting any CD produced before the switch to the new scheme would be a lost cause.

Conclusion

The current system as it stands cannot work. As long as players play unprotected songs, and people can produce unwatermarked MP3's, there can be no protection. Also, as long as the unwatermarked copies are available, any music purchased online with a watermark can be compromised as well. Overall, these problems state the basic need for universal acceptance of one watermarking scheme by the recording, electronic, and computer industries as well as consumers.

Works Cited

[1] Acken, John M. "How Watermarking Adds Value to Digital Content." Communications of the ACM, July 1998, vol 41, no 7.

[2] Memon, Nasir and Ping Wah Wong. "Protecting Digital Media Content." Communications of the ACM, July 1998, vol 41, no 7.

[3] Secure Digital Music Initiative (SDMI). "Guide to the SDMI Portable Device Specification, Version 1.0." Available at http://www.sdmi.org

[4] Wu, Min et al. "Analysis of Attacks on SDMI Audio Watermarks." Proceedings from 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. 7-11 May, 2001, Salt Palace Convention Center, Salt Lake City, Utah, USA: 1369-1372.

[5] Felten, Edward et al. "Reading Between the Lines: Lessons from the SDMI Challenge." Proceedings from the 10th USENIX Security Symposium. 13 – 17 August, 2001, Washington, DC, USA.

[6] McCullagh, Declan. "Copyright Law Foes Lose Big." Wired News. 29 November 2001. Terra Lycos Network. 29 November 2001 <http://www.wired.com/news/politics/0,1283,48726,00.html>

[7] D. Gruhl, W. Bender, and A. Lu. "Echo-hiding", Information hiding: 1st International Workshop, R.J. Anderson, Ed., vol. 1174 of Lecture Notes in Computer Science, 1996, Isaac Newton Institute, England, pp. 295-315.