What is Acceptable Quality in the Application of Digital Watermarking: Trade-offs of Security, Robustness and Quality

Scott Moskowitz,
Blue Spike, Inc. 16711 Collins Avenue No. 2505, Miami Beach, Florida 33160
scott@bluespike.com

Abstract

Quality is subjective. Quality can be objectified by the industry standards process represented by such consumer items as compact disc ("CD") and digital versatile disc ("DVD"). What is lacking is a means for not only associating the creation of valued intangible assets and extensions of recognition but establishing responsibility for copies that may be digitized or pass through a digital domain. Digital watermarking exists at a convergence point between piracy and privacy. Watermarks serve as a receipt for information commerce. There is not likely to be a single digital watermark encoding scheme that best handles the trade-offs between security, robustness, and quality but several architectures to handle various concerns. The most commercially useful watermarking schemes are key-based, combining cryptographic features with models of perception. Most importantly, in audio watermarking there currently exists mature technologies which have been proven to be statistically inaudible.

1. Introduction

The efficacy of copyright management systems will depend largely on keeping "security" out of view from consumers while enabling clear responsibility to be attributed to the media content being traded. Consumers have repeatedly rejected access restriction and registration protocols as currently deployed in favor of open peer-to-peer systems. Meanwhile, the digital watermark research literature is littered with assertions concerning "quality" which have been made without comprehensive "golden ears" listening tests, such as those conducted for the Secure Digital Music Initiative's ("SDMI") Phase 2 standards process or similar tests that have been conducted in the visual applications field. Security and quality are complex and subjective.

Complicating matters is the inherent difficulty with implementations of digital rights management ("DRM") systems on consumer PCs that typically lack realistic provisions for authenticating digital objects. Ignoring historical precedent and legal province of "fair use" and the "first sale doctrine" serves to obscure the economic value attributed to content. In an ideal digital watermarking system, maintenance of the intended perceptible quality must be weighed against the technical reality of trade-offs with security and robustness against attack. Determining tampering or attributing responsibility for copies are inherent features of economic activity. Successful commercialization requires a focus on the perception of value; the file format must be relegated to convenience [1].

Without an audit trail, or the creation of receipts for content, a means of settling responsibility for particular digital objects will prevent successful commerce in an information economy. The general need for commercial deployment of workable digital watermarking schemes is best represented by the widespread acceptance of Napster™, and its progeny, including, Music City™, KaZaA™, et al.

The presence of a content identification watermark is the hook to facilitate commercial markets surrounding the *use* of music, and other media, by consumers. Some of these uses include: monitoring of broadcast playback by performing rights organizations ("PROs"), premium services for peer-to-peer music distribution networks (a commercial Napster), and consumer content identification services (like Gracenote™/CDDB for individual tracks). The cost on a computational resource basis is lower than competing identification systems using so-called "signal fingerprinting" and onerous application of DRMs that obscure any *a priori* willingness of general consumers to pay for content [2]. Furthermore, the cost is borne by each client in a distributed manner, avoiding processing and bandwidth bottlenecks, similar to the way that Napster distributed storage.

In this paper, a description of several of the decoding system applications, and why watermarks are a necessary feature of any workable market for the commercial

exchange of content will be highlighted. Included is a comparable statistical measure of the actual maturity of audio digital watermarking having been proven to meet the most stringent, if not subjective, standards of sonic quality.

2. Broadcast Monitoring

At present, a variety of technologies are used to monitor the playback of sound recordings on broadcast outlets. Digital watermarking is a better alternative to all of the deployed technologies because it couples automated detection with extremely high reliability. A single PC-based monitoring station can continuously monitor up to 16 channels of audio broadcasts 24 hours a day with no human interaction. The results of the monitoring are assembled at a central server and made available to interested licensees, such as the PROs, for a fee equivalent to the price they currently pay for monitoring data. Unlike currently deployed systems, there is an extremely low statistical chance of misdetection. Additionally, the system can distinguish between otherwise identical versions of a song, which are watermarked for different distribution channels, further improving the quality of the reported data.

Deployment of such a system requires two things: a monitoring infrastructure and the watermarks to be present in the content. Leading monitoring companies have developed and deployed extensive infrastructures that have been designed to identify certain encoded audio and video signals as they are distributed. Watermarking music or video is planned by all major entertainment companies, those who possess closed networks, as well as, those involved in advertising.

3. Peer-to-Peer File Sharing

The immense popularity of peer-to-peer file sharing ("P2P"), in combination with recent legal rulings, presents a challenge: how to commercialize a file-sharing network. Watermark-based content identification is the solution. Each track is to be identified by the client's computer using a watermark detector. Ideally, the detector may be upgraded or replaced by a plurality of watermarking algorithms, if said algorithms are generated in combination with an upgradeable cryptographic key for such use. A so-called "steganographic cipher key" performs identification and authentication functions without revealing the unwatermarked original media content. The identity or authenticity of the track is then used to filter the server search engine, so that each subscription level only provides access to "allowed" content. Signal fingerprints or web trawlers cannot independently establish responsibility for any given digital object at comparable measures of computational overhead as embedded watermarks but can be used to reduce forensic searches for particular files.

As there are many embedding techniques and compression algorithms; so there should be support for many types of watermarking embedders and detectors. That a key-based watermark process essentially maps or concatenates a cryptographic signature in such a manner as to mimic the perceptibility of any given media object, emphasis on authenticity of digital objects is likely to assist in accurately determining what consumers are willing to pay for. These keys may also be used to watermark portions of specific areas of a signal or even save signal characteristics to the key to assist in detection or decoding watermark message data. Collectively, the ability to tamperproof or restore a suspect digital object with a watermark key is invaluable to maintain authorized information-based markets. Here is how it works in action:

3.1. Encoding

Encoding happens at the mastering level of each sound recording, as currently contemplated by the major label music companies as well as the major studios for video. Downstream, "transactional" watermarks are also considered. Each song is assigned a unique ID from the identifier database, and that ID is encoded in the sound recording after all other mastering processes are completed, but prior to the song being prepared for a specific distribution channel. To enhance imperceptible encoding of those few audio or video recordings that require special processing, human-assisted watermark key generation is readily available.

3.2. Decoding

Decoding happens each time a new song is made available on a P2P user's computer. A highly efficient background process decodes each sound recording, and queries P2P's main server as to the status of the selected track. The server would respond that the sound recording falls into one of the following categories:

Uncontrolled: The sound recording either does not contain a watermark, or the copyright owner has chosen to make the song freely available to all users. In this example, the sound recording will be freely available to pass through the P2P server.

Premium: The sound recording is part of a subscription package and is made available only to the premium subscriber of that subscription package.

Restricted: The sound recording is not authorized to be shared on the main server and will not be available for file sharing purposes.

4. A Real World Example

Alice is a Napster user. She has a hard drive directory of audio files which her Napster application monitors. She rips a new CD into that folder and starts the Napster application. The application reads the watermark on each track to identify those tracks. The new tracks, like all on her computer, are available for her own, unlimited, use.

When Alice connects to the server, her computer broadcasts the identity of all of the sound recordings in her shared folder. These, are a mix of uncontrolled, premium, and restricted content, as determined by the server at that time. For the new tracks that were recently added to her folder, the server identifies that one song is premium, and the others are uncontrolled.

Bob is a Napster user, and is looking for music. He is a premium subscriber. The Napster server makes the uncontrolled and premium music on Alice's computer available to Bob.

Carl is another Napster user, but not yet a subscriber. He sees only the uncontrolled music when he logs on to the Napster server.

This system provides minimum impact on consumers, while maintaining the safeguards necessary for the sharing of copyrighted material. Each user is not prevented from using restricted songs on their own computer, since in most cases they will have purchased them legally, for instance on CD or by subscription. Those songs are simply not available to others against the wishes of the copyright owner. No other approach to the rampant problem of unfettered file sharing is technically reasonable. When combined with technologies such as a content-specific cipher, which encrypts data in such a manner as to retain perceptibility but distort the media content in a tiered fashion (a predetermined key or key pair combined with a transfer function), copyright owners can estimate the highest optimized mix of quality thresholds demanded by consumers over a network in real time.

Users, in this scenario, purchase individualized keys (essentially tied to their public key or some equivalent digital credential for purchase options) based on observable music, video, or images, with reasonably open access that improve the quality of the music, or other media, as consumers "click through" to higher quality thresholds. A reduction in server overhead and cost, as well as maintenance of recognizable but secure media files, combined with digital watermarking, represent the state of the art in addressing file sharing. This also allows for multiple subscription levels based on content types and quality settings. The need to store multiple versions, both compressed and uncompressed, as per requirements for typical DRM systems, in an encrypted state is likewise reduced. Commercially, owners or aggregators of content will be able to estimate payment and bandwidth resources in real time. A natural extension is to provision paths of packets, that comprise media content, demanded between users, to efficiently provision bandwidth at the highest market price.

In the event that the sound recordings are not available with watermarking, application of signal recognition (fingerprinting) offers additional coverage. A unique abstract of the selected sound recording is taken and its signal characteristics are compared to an associated database. This comparison will identify the name of the performance if the sound recording is included in the database. Simple hashes or checksums of the audio file are ineffective given the range of reasonable alterations conceivable. Predetermination of the types or amount of signal manipulations expected on the audio file can be used to create a better, more robust "signal abstract" (which may be stored publicly, privately, or at a certification authority to point out authorized versions of the recording) than currently available signal fingerprinting applications. Application to other forms of media is obvious.

The signal recognition application is primarily useful for legacy, unwatermarked, material. This specifically limits the scope of the signal fingerprint database, which is crucial to maintaining the feasibility of fingerprinting. At present, no entity has demonstrated fingerprint technology that can economically scale to cover the daily increase in available media content. Nor can it be expected that "versioning" of the content in question will decrease in the future. With versioning of media content, more personalized exchange of any particular digital object is likely to require a means to independently authenticate objects without requiring predetermination of all possible manipulations of the media object in question.

5. Consumer Song Identification

Gracenote (formerly CDDB) offers a hugely successful system to identify physical CD's based on their Table of

Contents. The hole in the system is that it is useless for content that arrives as an individual digital track. An MP3 found on a peer-to-peer system can arrive without any linkage to the distributor or artist. Watermarking can fix this, allowing an anonymous track to be reassociated with its creator, and facilitating sales by all of the members of the value chain.

An inexpensive watermark detector would be added as a feature or plug-in to all popular music players, just as the present Gracenote software is included. Any incoming track could be detected and then decoded, and a resulting query could be made to a server which not only identifies the track, but places it in a sales context for the up-sell of all manner of associated items, from other tracks by the same artist, to concert tickets and merchandise.

Best of all, the consumer's identification act also provides critical data on the use and popularity of each track. Here the watermark is crucial, because it can distinguish between identical tracks obtained from different sources, thus informing the viability and market potential of different modes or even channels of distribution. Finally, if the distribution channel is correctly identified, the consumer can be up-sold the appropriate items. For example, if they recorded the song from an Internet broadcast, sell them the CD.

6. "Audio Quality" by Statistics: SDMI

Much has been ignored or misunderstood in the research literature concerning acceptable quality parameters for digital watermarking systems. Given the generally higher sensitivity to distortion in the human auditory system, and its relevance to any psychoacoustic modeling, this paper offers opinions based on the most extensive audibility testing endured over the past six (6) years. This testing has been conducted on a number of different encoding schemes: least significant bit (LSB), adaptive quantization, amplitude masking, and several variations of mature psychoacoustic masking has yielded statistical proof that at least one audio watermarking technical is "inaudible" and technically mature. Most of this audibility testing has been conducted under confidentiality agreements with little if any provision for publicly benchmarked results. Moreover, automated watermarking systems, not the far more flexible application of key-based systems, have been exclusively emphasized for unknown reasons. The exception was the lengthy, heavily publicized, and comprehensive SDMI Phase 2 listening tests. The results presented herein were prepared by an independent doctoral statistician hired by the SDMI organization.

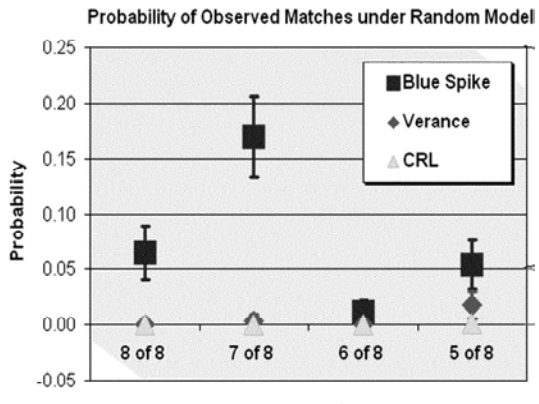

Figure 1. Values above 0.05 indicate agreement with a random model. A digital watermark was less likely to have been detected. Values under 0.05 indicate disagreement with a random model. A digital watermark was more likely to have been detected [3].

While it might be obvious that most commercially valuable music is loud and compressed enough to make any watermarking system acceptable from a sonic quality perspective, all of the significant commercial tests which have been conducted have been focused almost exclusively on classical pieces with very little data hiding space. Unfortunately, most testing has also focused on robustness without provisioning for key-based systems that can authenticate audio files and carry enough data in the key to assist in determining the original recording's scale or other signal features, without requiring the original unwatermarked file. Watermarking is a mature, flexible analog to its real world counterpart: that significant feature of commerce-- the receipt. Without provably secure watermarks, or receipts, it is not likely any technology will satisfy the expectations of rights owners, consumer electronics manufacturers, information technology vendors and the public at large.

7. Conclusion

Consumers have created and embraced particular usage models for music, which includes CD copying, file-swapping, and format indifference. They expect to be able to play music on any of a number of device platforms, from stereos to computers to cell phones. Any system of music distribution that ignores or significantly impedes these models will meet with limited success.

More pointedly, the economics of DRM are questionable at best [4, 5]. The cost of recognition, promoting or otherwise creating demand for information content is separate from responsibility once that information content has been transacted. Access restriction threatens the viability of the historic reality that a few copyrights account for a lion's share of revenues. In 1999, for instance, only 0.03% of compact discs accounted for over a quarter of all revenues [6]. In 2000, 0.35% of all albums released accounted for over half of all revenues: 88

releases represented slightly over 25% of revenue [7]. Similar market realities apply to all forms of entertainment, including video, limiting any supposition that we can predetermine the success of any given media content release [8].

Arguments that "superdistribution" will replace market realities lack any real world examples; in fact, financial success generally boasts models seeking monopolistic or oligopolistic control of profitable intellectual property. As with physical media distribution emphasis is better placed on enabling differentiations between authorized and pirated versions of a given media content file copy or stream. Concatenating a digital signature to a media file, a key-based digital watermark, is the most appropriate means to enable markets for the open, accessible exchange of media content. Ultimately, key-based digital watermarks enable a balance to be struck between privacy and piracy. Moreover, they assist in providing transparency to replace statistical models currently relied upon by market participants. Essentially enabling receipts for information commerce. It is the conduit through which the business of music, and media in general, will be conducted, now and in the future.

References

[1] "Deciphering Music's Digital Devolution", Billboard, Timothy White, April 28,2001, p. 10.

[2] "It's the Pricing, Stupid!", Stereophile, October 25,2001 {online edition}.

[3] **"**An Evaluation of the SDMI Listening Test", Eugene Ericksen, PhD, SDMI Foundation, December 1, 2000.

[4] "Music industry still in first gear online", Reuters, January 7, 2002, {http://news.cnet.com/news/0-1005-200-8395107.html}.

[5] 'Will Fixation on Security Silence the Trumpets of Fame", Digital Mogul, Scott Moskowitz and Peter Cassidy, Volume 3 Report 7 {online edition},

[6] "The Heavenly Jukebox", Atlantic Monthly, Charles C. Mann, September 2000 {online edition}.

[7] "SoundScan Numbers Show .35% Of Albums Account For More Than Half Of All Units Sold", Billboard, April 28,2001, p. 66 .

[8] "Will 'Harry' have Legs?" The Wall Street Journal, November 30,2001, p. W4.