DIGITAL WATERMARKING: FOSTERING AND ENHANCING LEGITIMATE PEER-TO-PEER (P2P) ECOSYSTEMS

DIGIMARC CORPORATION
9405 SW Gemini Drive
Beaverton, Oregon 97008
www.digimarc.com

DIGIMARC

DIGIMARC

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

The music, movie and distribution industries are at a crossroad.

On one hand, digital distribution of music and movie files offers new channels to market, and new revenue opportunities for the industry, while satisfying the desire of millions of consumers to digitally consume entertainment content. On the other, this same copyrighted content can easily be converted to compressed files, known as "ripping," and these compressed "ripped" files are easy to distribute in peer-to-peer (P2P) environments without systems in place to identify copyrighted content or to ensure the content owners receive proper payment.

Digital watermarks provide a solution to this problem, enabling content identification that gives the media & entertainment industry a means to enforce their rights while offering consumers access to legitimate content, when and where they want. Using digital watermarks, content owners can embed secure copyright data into the music or movie content. Since the digital watermark data inherently survives the ripping process and format conversions, copyrighted songs can be identified on a P2P system even after that content has been ripped.

In such a system, the content owners embed the copyright data, and P2P software detect the copyright data on the user's computer during the indexing process that P2P software currently needs to perform. The copyright data held in the digital watermark can be used to enable multiple usage models, including (1) identifying copyrighted songs such that only non-copyrighted songs can be shared, (2) enabling copyrighted songs to be secured and licensed on P2P systems if agreed to by the content owner, and (3) enhancing content that can be shared. Using digital watermarks is more efficient than using audio fingerprints and both more efficient and more secure than using file names and header data. Digital watermarking can also easily be applied to enhance other distribution methods in both audio and video.

P2P Ecosystem Issues

Music CDs provide an unprotected digital master that users can turn into quality compressed audio files (a.k.a. ripping), such as MP3 files, and share those files on P2P

systems without proper compensation to the content owners (e.g. record labels, musicians, song writers, etc.). Even with protection, CDs and digitally distributed songs can be recorded and digitized from the analog outputs (a.k.a. analog hole), or captured within the digital buffers of sound cards, and turned into unprotected and compressed audio files that are easy to share. Thus, P2P systems have difficulty determining which audio files are copyrighted and not allowed to be shared or require license rights and which audio files are non-copyrighted or allowed to be shared.

Successful commercial deployments of digital watermarking by the music, movie, broadcasting and advertising industries are already having a significant impact on reducing piracy in pre-release music and movies, and improving the ability to monitor, track and manage digital media. Fingerprinting, or pattern recognition, also can be used with digital watermarking to filter legacy distributed music and re-associate basic meta data (artist, song, copyright) that has been lost during the ripping process. Digital watermarking opens the door to new and legitimate business models, new protection schemes and enhanced consumer experiences by providing additional related content in a "connected media" fashion that truly enhances the entertainment experience.

2 OVERVIEW OF DIGITAL WATERMARKS

Digital watermarks can enable P2P systems to determine copyrighted from non-copyrighted audio files within the existing distributed P2P system architecture. Digital watermarks can even enhance P2P systems, enabling the P2P providers to work with record labels and other audio retailers or content owners to legitimately sell copyrighted songs and related items.

As background, digital watermarks are digital data elements that are embedded into actual content—not carried in the header—so the elements survive analog conversion and standard processing, such as ripping to MP3. Digital watermarks may be embedded into, and read from, video, audio and still images for the applications described in this paper. For video content, watermarks in the audio, video or both the audio and video tracks may be used. The digital watermark data is not perceptible to the human ear (or eye), but can be read by computers. The digital data is secured through secret keys similar to encryption keys.

The digital data can include copyright control information, content classification flags for filtering, content identification (content ID) and forensic identification, for example. This data can enable numerous applications, including copyright communication, content filtering, copy protection, broadcast monitoring, Internet monitoring, forensic tracking, authentication, digital asset management, digital rights management (DRM), and enhanced e-commerce. Copyright communication, content filtering, DRM and enhanced e-commerce are most applicable to P2P systems.

Digital watermarks are in extensive use around the world, with billions of digitally watermarked objects and hundreds of millions of detectors in use for copy protection, copyright notification, authentication and forensic tracking applications. Digital watermark technology providers include Digimarc, Activated Content, Dolby/Cinea, Thomson, Philips, Signum, Verance, and Verimatrix. Major record labels currently use digital watermarks to forensically track most pre-release CDs. The system has led to a significant reduction in illegitimate use of pre-release music, and a similar system used by the movie industry has led to arrests by the FBI for pre-release Academy Award screener copies of movies.

3 ARCHITECTURE TO LEGITIMIZE AND ENHANCE P2P SYSTEMS

Digital watermark systems have an embedder that adds the digital data into the content, and a detector that reads the digital data. In a P2P system, the architecture enables content owners (e.g., record labels) to embed a secure copyright digital watermark (DWM) into digital content files, and allows the software client provided by P2P providers to detect the digital watermark in content files when indexing those files that the user has allowed to be shared. The architecture details are described below and shown in Figure 1.

The content owners use a digital watermark embedder to embed the copyright DWM in the audio master, such that the DWM is included in all retail versions of the song. The embedder is only available to legitimate producers of music, using standard distribution methods for security protection systems. The presence of a secure copyright DWM identifies the song as copyrighted work, and contains a content ID to identify the song. For example, the copyright DWM acts as a flag to identify the presence of a copyrighted work, and the content ID is used to provide content specific information and services as detailed below. A copyright flag and content ID may be integrated into a single digital watermark or carried by separate digital watermark signals.

The digital watermark detector is securely integrated with P2P software that is downloaded by the user. The detector is used to look for a DWM on each audio file that the user has enabled the P2P software to share when the P2P software indexes these songs to identify them for sharing. This indexing processing already occurs in P2P systems, and is usually run in the background and overnight. Adding DWM detection to the system will have negligible effects to the user (e.g., requiring a fraction of a second of additional processing per song on modern PCs or mobile devices), especially given that it fits with the existing distributed and run-in-the-background architecture for P2P software.

The copyright DWM with the content ID can be used by the P2P provider to determine rights information, as well as other information about that song and related music, via a remote and potentially distributed database (i.e., to secure and enhance the content). This database can be stored where the P2P provider stores the index database used to determine where to find songs for search results. In fact, the database may comprise several databases where some parts of the database, such as song lyrics, may be stored by the content owners, and can be used as additional revenue opportunities for all participants. These details can be determined by the P2P provider and content owners.

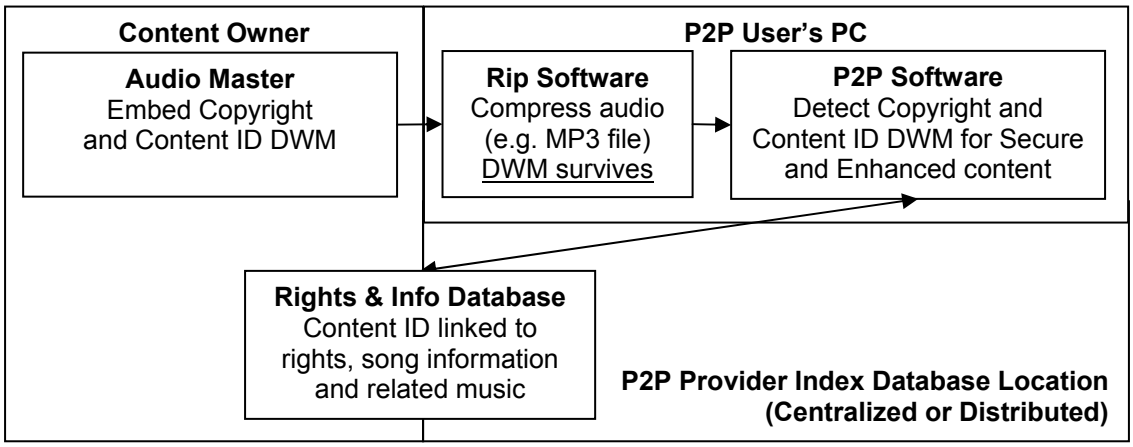

Figure 1: Overview of copyright digital watermark architecture

4 POTENTIAL USAGE MODELS AND BENEFITS

This architecture enables three usage models, including (1) copyright communication, (2) licensed content and (3) enhanced content, as describe below and shown in Figure 2.

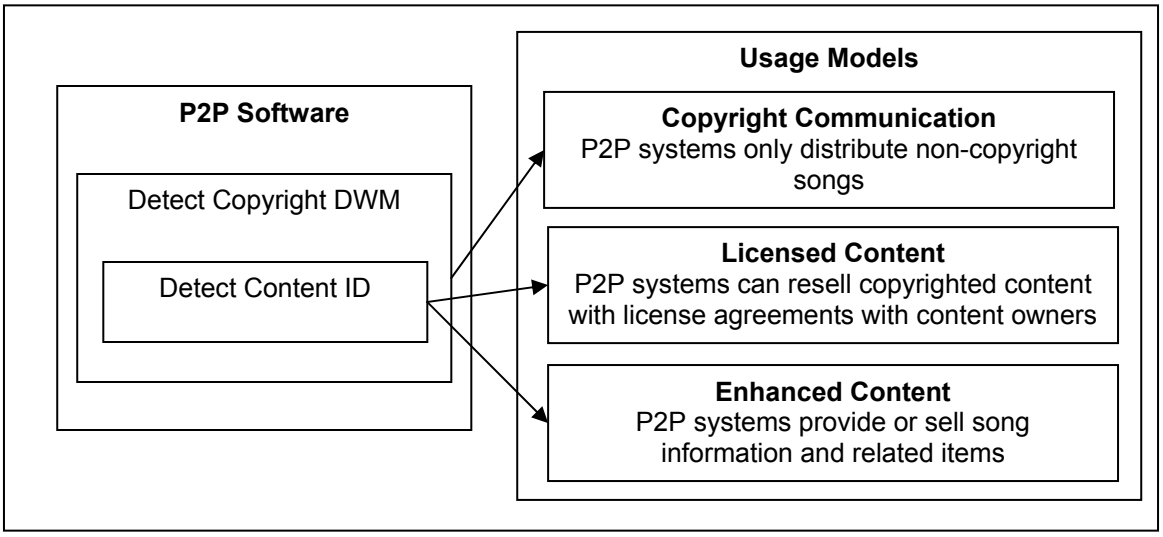

Figure 2: Overview of usage models

4.1 Usage Model 1: Copyright Communication

The copyright communication usage model would enable P2P providers to prevent the listing of content where the copyright DWM has been detected during the indexing process. The copyright communication usage model does not require use of a content ID, thus being completely distributed and easiest to implement.

The benefits of this usage model are:

- Consumers benefit by minimizing the risk of copyright infringement and potential legal liability by using P2P for legitimately licensed and non-copyrighted material
- Content owners benefit by having less piracy and increased revenues from legitimately licensed content
- P2P providers benefit by minimizing risk of a lawsuit while enabling non-copyright and legitimately licensed distribution
- Other manufacturers and service providers benefit because consumers have less fear of using computers on the Internet, which leads to more usage

4.2 Usage Model 2: Licensed Content

The licensed content usage model would enable P2P providers to detect the copyright DWM during the indexing process, look for a content ID, and use the content ID to look up the rights information about that content, and secure a copy of the content (or possibly a higher quality version) in a digital rights management (DRM) package, if allowed. The DRM package enables P2P providers to sell the content to consumers after they have licensing agreements with the labels.

This usage model requires a slightly more complex architecture than for copyright communication, but also includes more benefits, including:

- Consumers benefit by having more access to music, leading to increased sales, as well as minimizing the risk of a lawsuit
- Content owners benefit by selling more content and having even less piracy with improved content availability, as well as enabling non-copyrighted content distribution
- P2P providers benefit by selling more music and having customers enjoy more product capabilities, as well as minimize the risk of a lawsuit while enabling non-copyrighted and legitimate copyrighted content distribution
- Other manufacturers and service providers benefit by further increasing usage of computers, mobile devices and the Internet, and increased digital music sales leads to more content being played on more portable players

Some P2P systems may exclusively use copyright communication, whereas some P2P systems may use copyright communication for songs with a copyright DWM but no content ID or no agreement with the content owner, and licensed songs with a content ID and agreement for distribution with the content owner.

4.3 Usage Model 3: Enhanced Content

The enhanced content usage model would enable P2P providers to detect the copyright DWM during the indexing process and use the content ID to link that user and others searching for that content to related content and information.

The enhanced content usage model can work synergistically with the licensed content usage model, thus only linking the users to more information for songs that the P2P provider has the legal right to sell. When both securing and enhancing the content, only the extra step of linking the content ID to related content and information is required beyond the process to secure the content. This combination usage model enables even more benefits, including:

- Consumers benefit by having more access to music and related information, leading to increased sales, as well as minimizing risk of a lawsuit and using P2P for non-copyrighted material
- Content owners benefit by selling more content and having even less piracy with the improved content and related information availability, as well as enabling non-copyrighted distribution
- P2P providers benefit by selling even more music, related items and having consumers enjoy even more product capabilities, as well as minimizing risk of a lawsuit and enabling non-copyrighted distribution
- Other manufacturers and service providers benefit by increase usage of computers, mobile devices and the Internet and even more content to be played on more portable players

Once content is identifiable in any format through a digital watermark, e-commerce can be enhanced. The digital watermark is read and linked to more information about the artist or content and "buy now" opportunities. For example, a song played on a cell phone can be identified with a digital watermark, the content ID can be used to look up information about the artist, and other songs from that artist which can be purchased with the click of a "buy now" button. The possibilities are numerous; however, the process must start with efficiently identifying the content with a digital watermark.

5 DIGITAL WATERMARKING AND FINGERPRINTING ARE COMPLEMENTARY TECHNOLOGY SOLUTIONS

Given the distributed architecture of the P2P system, digital watermarks are synergistic with audio fingerprinting (a.k.a., robust audio hashes) and more secure and efficient than using metadata, such as file name, or song title or copyright flag in the file header, to identify copyrighted content. Further, since digital watermarks need not be identical for every copy of a particular song, they enable more flexible business models and allow the digital watermark to link to information and services that are tailored to a particular distribution or usage context.

Audio fingerprinting requires that, for each song to be identified, a relatively computationally intense search of a remote and larger database must be performed to find the closest matching ID to identify the song. Audio fingerprinting can be retroactively fit to legacy content. Digital watermarking does not require any database for copyright communication, and only a simple database lookup for securing and enhancing the content via the content ID once embedded in new content. Using both fingerprinting for legacy content and digital watermarking for newly released content in combination strengthens the total ecosystem by providing the most user friendly and comprehensive solution.

Metadata is easy to change and a copyright flag in the song header is simple to remove. In contrast, digital watermarks have no or minimal database issues (as described in the previous paragraph), and they are secured with a secret key, as used in encryption. In other words, digital watermarks require expertise to attack that only a few researchers in the world possess. This greatly raises the cost of piracy, thus reducing its likelihood, especially if legitimate content is easy to obtain.

6 CONCLUSIONS

Digital watermarks provide an easy-to-use and -implement approach to legitimize P2P systems, and even enhance them. A copyright DWM is embedded by the content owner, and the copyright DWM is detected on the user's computer or mobile device to properly identify audio files that the user has requested to be shared. This identification can lead to copyright communication, usage rights and licensing opportunities that legitimize P2P systems, and even enhance them with sales of additional content and related items. This DWM architecture is synergistic with audio fingerprints and more efficient and secure than techniques using file names, song titles or copyright flags, but can be used as a complementary layered approach with such systems.

Furthermore, the copyright digital watermark embedded by the content owners can be utilized for similar benefits in other distribution chains. A similar system can be implemented by movie studios, even potentially using the same audio copyright DWM such that the P2P system searches for the same copyright DWM in all audio and video.

Digital watermarks can enable P2P systems to determine copyrighted from non-copyrighted audio files within the existing distributed P2P ecosystem architecture. Digital watermarking can even enhance P2P systems, enabling the P2P providers to interoperate freely with existing value chain entities such as record labels and other audio retailers or to a larger extent working directly with content owners to market legitimate copyrighted songs and other market related materials.

DIGIMARC

For more information, please contact:

Reed Stager
Executive Vice President
DIGIMARC CORPORATION
503.469.4684 or rstager@digimarc.com
www.digimarc.com