OULU 2008
C 311

ACTA

UNIVERSITATIS OULUENSIS


*Mikko Löytynoja*

# DIGITAL RIGHTS MANAGEMENT OF AUDIO DISTRIBUTION IN MOBILE NETWORKS

*FACULTY OF TECHNOLOGY,*
*DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING,*
*UNIVERSITY OF OULU;*

ACTA UNIVERSITATIS OULUENSIS
C Technica 311

*MIKKO LÖYTYNOJA*

# DIGITAL RIGHTS MANAGEMENT OF AUDIO DISTRIBUTION IN MOBILE NETWORKS

Academic dissertation to be presented, with the assent of the Faculty of Technology of the University of Oulu, for public defence in Raahensali (Auditorium L10), Linnanmaa, on December 15th, 2008, at 12 noon

OULUN YLIOPISTO, OULU 2008

Acta Univ. Oul. C 311, 2008

Supervised by
Professor Tapio Seppänen

Reviewed by
Professor Tommi Mikkonen
Professor Martti Mäntylä

ISBN 978-951-42-8936-1 (Paperback)
ISBN 978-951-42-8937-8 (PDF)
http://herkules.oulu.fi/isbn978951428937 8/
ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)
http://herkules.oulu.fi/issn03553213/

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2008

**Löytynoja, Mikko, Digital rights management of audio distribution in mobile networks**
Faculty of Technology, Department of Electrical and Information Engineering, University of Oulu, P.O.Box 4500, FI-90014 University of Oulu, Finland
*Acta Univ. Oul. C 311, 2008*
Oulu, Finland

***Abstract***

Nowadays, content is increasingly in digital form and distributed in the Internet. The ease of making perfect copies of the digital content has created a need to develop a means to protect it. Digital rights management (DRM) relates to systems designed to protect the intellectual property rights of the digital content. The DRM systems try to enable a secure distribution of digital content to the users and to prevent the unauthorized copying, usage, and distribution of the content. This is usually done in practice using encryption and digital watermarking techniques.

This thesis concentrates on the problem of protecting and distributing multimedia content securely in mobile environment. The research objectives are: (1) to design an overall DRM architecture which allows an easy content distribution to the user in mobile environment; (2) to develop protection methods that can be used in mobile devices with limited computational capabilities to prevent unauthorized usage of the audio content; (3) to create methods for managing and enforcing the user's rights and restrictions to the content usage; (4) to study a method for providing the users with an easy access to new digital content and services. The research is carried out by first developing an overall DRM platform to mobile environment. The experimental prototype of the platform is implemented on server side to PC environment and the client runs on a mobile phone. The platform is used to test the functionality and complexity of the content protection methods developed which are based on digital watermarking and encryption techniques.

The main results of the thesis are: (1) a DRM platform for mobile devices that supports peer-to-peer networking and license negotiation; (2) audio protection methods utilizing digital watermarking and encryption techniques which support content superdistribution and content preview; (3) methods for counting offline how many times content has been played on the user's terminal using watermarking and hash chains; (4) a method for adding metadata, such as a web link, into audio content, so that it survives digital to analog to digital transformation and recording with a mobile phone.

*Keywords:* audio watermarking, content protection, digital rights management, DRM

# Acknowledgements

The research related to this thesis has been carried out at the MediaTeam Oulu research group of the Department of Electrical and Information Engineering at the University of Oulu, Finland.

I would like first to acknowledge Professor Tapio Seppänen for supervising my thesis and his excellent guidance during my postgraduate studies. I also wish to express my gratitude for all the past and present members of the information hiding team, particularly Dr. Nedeljko Cvejic, Anja Keskinarkaus, Marko Brockman, Eeva Lähetkangas, and Timo Koskela, who have contributed to the original publications. Especially I would like to thank Dr. Mika Rautiainen, Eero Väyrynen, and Jari Forstadius for all the interesting and stimulating discussions and arguments we have had. I would also like to thank Professor Timo Ojala and all my colleagues in MediaTeam for creating an encouraging and pleasant working environment. Special thanks go to Dr. Pertti Väyrynen for proofreading the manuscript.

I would like to thank the official reviewers, Professor Tommi Mikkonen and Professor Martti Mäntylä, for their suggestions and constructive criticism.

For the financial support for this thesis, I would like gratefully thank Graduate School in Electronics, Telecommunications, and Automation (GETA), the Nokia Foundation, the Finnish Foundation for Technology Promotion (TES), and the Finnish Funding Agency for Technology and Innovation (TEKES) and industrial partners in Stego, Stardust, and Zirion projects.

I am grateful to my parents Lea and Martti for everything they have done for me. Finally, I would like to deeply thank my wife Laura for patience, encouragement, and for just being there in everyday life.

Oulu, October 2008 Mikko Löytynoja

# List of symbols and abbreviations

| | |
|---|---|
| A | Cover signal |
| A' | Modified signal |
| $A_w$ | Watermarked signal |
| *b* | Binary sequence |
| *i* | Discrete index |
| *h*() | Hash function |
| K | Key |
| *K* | Watermark strength |
| *n* | Number of iterations |
| $t_i$ | Hash value |
| AACS | Advanced Access Content System |
| AES | Advance encryption standard |
| BER | Bit error rate |
| CA | Certification authority |
| CD | Compact disk |
| CSS | Content Scrambling System |
| DA/AD | Digital to analog/analog to digital |
| DES | Data encryption standard |
| DMCA | Digital Millennium Copyright Act |
| DRM | Digital rights management |
| DSSS | Direct sequence spread spectrum |
| DVD | Digital versatile disc |
| FFT | Fast Fourier transformation |
| FH | Frequency hopping |
| HAS | Human auditory system |
| HD DVD | High definition DVD |
| ID | Identification |
| IP | Intellectual property |
| IPMP | Intellectual property management and protection |
| ITU | International telecommunication union |
| LNS | License negotiation system |
| LSB | Least significant bit |
| MD | Message digest |
| MP3 | MPEG-1 audio layer 3 |
| MPEG | Moving Picture Experts Group |

| | |
|---|---|
| MOS | Mean opinion score |
| NEMO | Networked environment for media orchestration |
| ODG | Objective difference grade |
| ODRL | Open Digital Rights Language |
| OMA | Open Mobile Alliance |
| P2P | Peer-to-peer |
| PC | Personal computer |
| PCM | Pulse-code modulation |
| PEAQ | Perceptual evaluation of audio quality |
| PKI | Public key infrastructure |
| RDD | Rights data dictionary |
| REL | Rights expression language |
| RSA | Rivest, Shamir, & Adleman |
| SDG | Subjective difference grade |
| SHA | Secure hash algorithm |
| SPIE | Society of Photo-Optical Instrumentation Engineers |
| TPM | Trusted platform module |
| U.S. | Unites States |
| WIPO | World intellectual property organization |
| WLAN | Wireless local area network |
| XML | Extensible Markup Language |
| XrML | eXtensible Rights Markup Language |

# List of original publications

I Löytynoja M, Seppänen T & Cvejic N (2003) Experimental DRM architecture using watermarking and PKI. Proc. 1st International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet, Helsinki, Finland: 47–52.

II Löytynoja M, Koskela T, Brockman M & Seppänen T (2006) Mobile DRM-enabled multimedia platform for peer-to-peer applications. Proc. IEEE International Symposium on Multimedia, San Diego, CA, USA: 139–144.

III Löytynoja M, Cvejic N, Lähetkangas E & Seppänen T (2005) Audio encryption using fragile watermarking. Proc. Fifth International Conference on Information, Communications and Signal Processing, Bangkok, Thailand: 881–885.

IV Löytynoja M, Cvejic N & Seppänen T (2007) Audio protection with removable watermarking. Proc. Sixth International Conference on Information, Communications and Signal Processing, Singapore: 1–4.

V Löytynoja M & Seppänen T (2005) Hash-based counter scheme for digital rights management. Proc. IEEE International Conference on Multimedia & Expo, Amsterdam, Netherlands: 121–124.

VI Löytynoja M, Cvejic N & Seppänen T (2006) Watermark-based counter for restricting digital audio consumption. International Journal of Signal Processing, 3(1): 17–23.

VII Löytynoja M, Cvejic N, Keskinarkaus A, Lähetkangas E & Seppänen T (2006) Mobile commerce from watermarked broadcast audio. Proc. IEEE International Conference on Consumer Electronics, Las Vegas, NV, USA: 5–6.

VIII Löytynoja M, Keskinarkaus A, Cvejic N & Seppänen T (2008) Watermark-enabled value added services to broadcast audio. Proc. Second IEEE Conference on Digital Ecosystems and Technologies, Phitsanulok, Thailand: 388–396.

# Contents

# 1 Introduction

*"Any person can invent a security system so clever that she or he can't think of how to break it."*

Schneier's Law

## 1.1 Background

The digital watermarking is an art of hiding information in digital content by utilizing imperfections of the human perceptual system. The embedded information is tightly coupled with the host content so that it is hard to remove the embedded data. Digital watermarking is closely related to information hiding.

First references to information hiding date back to ancient Greek, where concealed messages where used in communication rather than encryption to hide the very existence of communication. Conventional paper watermarking appeared around 1282 in Italy, where watermarks were made with thin wires inserted in papers molds to make paper slightly thinner. The purpose of the early watermarks is not known for sure, but in the eighteenth century watermarks were started to be used as trademarks and means for anti-counterfeiting. (Cox *et al.* 2002.)

The first example of technology similar to the digital watermark was in 1954 when Muzak Corporation filed a patent for watermarking musical works (Hembrooke 1961). This system was used by Muzak Corporation until 1984. (Cox *et al.* 2002.) According to Cox *et al*. (2002), the first reference to digital watermarking is probably by Szepanski (1979), who presented a machine-readable pattern that could be embedded in documents. Holt *et al*. (1988) acquired a patent for a method for embedding an identification code in audio content. The first use of the term digital watermarking was done by Komatsu & Tominaga (1989), according to Cox *et al.* (2002). The term became in wider use in early 1990s.

In the mid 1990s, the interest in digital watermarking research started to increase gradually. The first Information Hiding Workshop was held in 1996 with digital watermarking as one of its primary topics. The Society of Photo-Optical Instrumentation Engineers (SPIE) started a conference specializing in Security and Watermarking of Multimedia Content in 1999.

Around the same time, digital rights management (DRM) started to emerge as a field of its own as the Internet began to commercialize. The DRM relates to

systems that are designed to protect intellectual property rights of the digital content. The DRM systems try to enable a secure distribution of digital content to the users and to prevent the unauthorized copying, usage, and distribution of the content. This is usually done using encryption and digital watermarking techniques.

The first generation of the DRM focused entirely on security and it used encryption as a means to prevent copying and access to the content from users that had not paid for it. The second generation of the DRM covers broader capabilities of DRM such as content description, identification, trading, protection, monitoring, and tracking. (Iannella 2001.) The DRM should not be seen only as a technology against piracy, however. The DRM can, for instance, make content purchasing easier and thus increase commerce. By increasing the security felt by the content providers the quality of the digital content will increase, which will create more revenues due to more satisfied customers. (Becker *et al.* 2003.)

The DRM systems can be visualized as a three-legged stool where the seat is the DRM system and the legs of the stool compose of technical methods, business models, and law. The technical elements cannot work in isolation, however; it would be unwise to use complex intrusive security methods to protect low valued content or vice versa. In order to operate appropriately, the three elements must be in balance. (Becker *et al.* 2003.)

Most DRM solutions developed thus far are variations on a common theme, now called as DRM reference architecture, illustrated in Fig. 1. Such architecture has three major components: content server, license server, and client. The content server's function is to store the actual content and to prepare it for DRM-based distribution. The license server's responsibility is to generate the licenses for content consumption to the user. The licenses are used to define what the user is allowed to do with the content, and they additionally contain keys needed to remove protection. The client comprises of the DRM controller, currently more often called a DRM Agent, which is the main actor in the DRM system. It is responsible for interacting with the user, as well as with the rendering application and the license server. (Rosenblatt *et al.* 2002.)

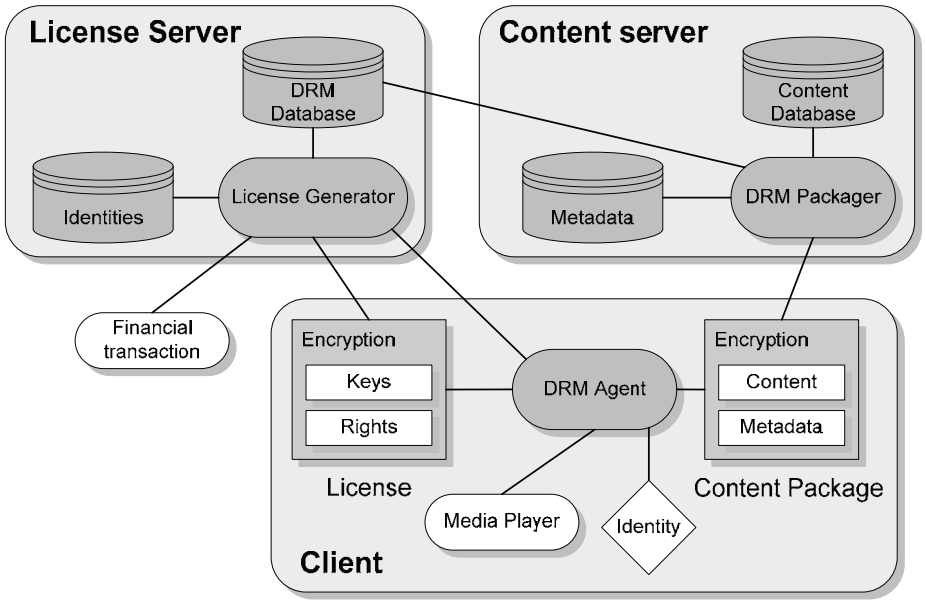


**Fig. 1. The DRM reference architecture. (Revised from Rosenblatt *et al.* 2002.)**

According to Institute for Policy Innovation analysis, global music piracy causes $12.5 billion of economic losses every year (Siwek 2007). This has increased the music industry's interest in the DRM solutions. One of the big problems with DRM systems is that the solutions have not been interoperable, because the field of DRM lacks standardization (Becker *et al.* 2003). This has lead to a situation where most of the major music distributors use different DRM techniques in their stores and the users have been restricted to buy their content from stores that support the DRM solution used by their portable music players. One exception is Open Mobile Alliance's (OMA) mobile DRM standard which is currently in its second version and is used in most mobile phones.

Another problem with the DRM systems is that once anyone manages to break their protection, the same method of violation is immediately available to all users. This forces the DRM system manufacturers to constantly upgrade and patch their protection methods. For example, the Advanced Access Content System (AACS), the content protection method of the new high definition video HD DVD and Blu-ray discs was broken in less than a year and subsequently the updated versions were also broken very fast (Wikipedia n.d.-b).

The very restricting nature of the current DRM solutions has lead to wide opposition of the DRM, e.g. see (Wikipedia n.d.-a). Hence, some adversaries of the DRM interpret the DRM to mean digital restriction management. In response to this, there is a current trend in the music distribution in the Internet to relax the copyright enforcement. Many major music stores have been begun to sell DRM-free music. In most cases, this means that the music sold is not encrypted and can be played with any music player. Some stores, however, embed customer's fingerprint into the content to find out who is responsible for leaking the pirated media. This fingerprinting can be done with digital watermarking techniques.

## 1.2 Research problem and objectives

The research problem of this thesis is: How to protect and distribute multimedia content securely in mobile environment? Specifically, the work concentrates on the following sub-problems:

1. What kind of overall architecture is needed to allow an easy content distribution to the user in mobile environment?
2. What kind of protection methods can be used in mobile devices with limited computational capabilities to prevent unauthorized usage of the audio content?
3. How the user's rights and restrictions to the content usage are managed and enforced?
4. How to provide the users an easy access to new digital content and services?

The main objective of the research is to create an experimental DRM platform to mobile environment which comprises both client and server side components. The platform is then expanded by developing several content protection and distribution methods that are used in the platform to answer the research problems.

## 1.3 Research scope and approach

This thesis focuses on developing protection techniques for audio content distribution. The protection methods developed can be applied also to other media types, although the used digital watermarking algorithms are media type specific. This means that the principles of the protection methods could be implemented using different watermarking algorithms which fulfill the robustness, capacity,

and other requirements set by the protection method. With respect to the three-legged stool model of the DRM systems, this thesis concentrates on the technology leg of the stool.

The research carried out here uses a constructive approach. The protection methods developed utilize existing technologies and research results where possible. The research is carried out by first developing the overall DRM platform, which is based on the DRM reference architecture (Fig. 1). The experimental prototype of the platform is implemented on server side to PC environment and the client runs on a mobile phone. The platform is used to implement and test the functionality and complexity of the content protection methods developed which are based on digital watermarking and encryption techniques.

Fig. 2 illustrates a flowchart of a typical DRM system, regarding how audio content is protected, distributed, and consumed. The DRM component residing in the user's device manages and enforces the rights and restrictions set by the license. The developed methods focus on the content distribution, license management and enforcement, and content protection parts of Fig. 2.

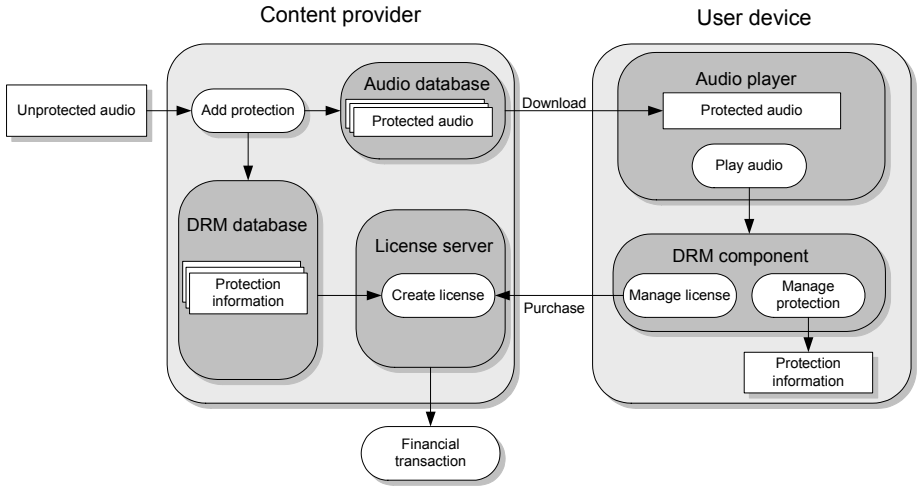


**Fig. 2. Overall flowchart for protection, distribution and consumption of audio.**

The focus of digital watermarking research is mainly to improve either the capacity or robustness of the watermarking algorithms. The watermarking algorithms are applied to develop realistic applications in real environment.

Developing real applications point out new requirements for the algorithm design which would otherwise be overlooked.

The watermarking algorithms developed are tested against selected attacks that are most probable in the selected applications. The robustness requirements depend heavily on the application use case, for instance, whether malicious attacks against the watermarks are expected or not. In cases where the malicious attacks are not expected, the robustness is only tested against the unintentional attacks that are inherent in the application. The audio quality tests are done with both objective tests and subjective listening tests.

## 1.4 Contributions and summary of original publications

The following is a summary of the contributions of this thesis:

1. A DRM platform for mobile devices that supports peer-to-peer (P2P) networking and license negotiation.
2. Novel audio protection methods utilizing digital watermarking and encryption techniques.
3. Methods for counting how many times content has been played on a user's terminal.
4. A method for adding metadata, such as a web link, into audio content so that it survives digital to analog to digital (DA/AD) transformation.

Paper I describes a DRM platform which uses digital watermarking and encryption to protect audio files. The platform uses public key infrastructure (PKI) to authenticate the users and protect the license files. The system allows the protected content to be superdistributed between users and offers a freely playable teaser of the content. The author of this thesis was responsible for the development of the architecture and implementation of the system. Dr. Cvejic was responsible for the watermarking algorithm and Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper II describes a mobile DRM platform that supports P2P networking. The platform provides strong protections methods utilizing encryption and watermarking. The users are allowed to superdistribute both the protected content and the actual DRM player software with Bluetooth or P2P networking. The license files that define what the user is allowed to do with the protected content are acquired with the help of a license negotiation subsystem. The author was responsible for the design of the platform and protection methods. Mr. Koskela

was responsible for the license negotiation. Mr. Brockman implemented the platform. Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper III introduces a method for increasing the robustness of a watermark by embedding another watermark into the content. The second watermark is fragile, i.e. it is destroyed if user tries to modify the content, and it contains information that the user does not want to lose. The method was demonstrated with an application in which the fragile watermark contains decryption keys for the audio content. The author was responsible for the design of the method and implementation of the experiments. Dr. Cvejic was responsible for the watermarking algorithm and Ms. Lähetkangas participated in the design of the application. Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper IV introduces an audio protection method that uses annoying audible watermarks. The digital watermark is embedded into audio content with high enough power that it is audible to the user. This watermarked audio file can then be freely downloaded by the users and it acts as a teaser to the actual content. If the user decides to buy the content, the watermark can be removed to restore the audio quality to the original level. The author was responsible for the development of the method and testing. Dr. Cvejic helped with the watermarking algorithm and Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper V introduces a method for counting how many times the user has played protected content with a DRM-enabled player. The method utilizes hash function to count the number of plays in the case where the user cannot be assumed to have online connection and the counter must be stored on the user's terminal. The basic method utilizes a modified hash chain method that prevents the malicious user from increasing the counter value. The paper, additionally, describes methods for improving the basic method. The author was responsible for the development of the method and Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper VI describes watermarking methods for counting how many times the user has played an audio file. The paper presents three different methods for using watermarks to store a counter value. The first one embeds the watermark with specific intervals, and as the counter value increases, the watermarked parts in the content increase as well. The second method uses rewritable watermarks to store the counter value. Finally, the third method embeds multiple watermarks in the

same location of the audio file, and the number of watermarks gives the counter value. The author was responsible for the development of the protection method and implementation of the experiments. Dr. Cvejic was responsible for the watermarking algorithm and Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper VII introduces a method for linking analog audio such as radio broadcast music to web pages with digital watermarks. The method uses two watermarks; one for synchronization and another to embed the actual data. Audio which is played through loudspeakers can be recorded with a mobile phone and the web link is extracted from the recording. The author was responsible for the design of the method and implementation of the application. Dr. Cvejic and Ms. Lähetkangas were responsible for the design of watermarking algorithm with the help of Ms. Keskinarkaus who also carried out the testing. Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

Paper VIII introduces an improved version of the method for embedding metadata in analog audio with digital watermarks. The watermarked audio is recorded with a mobile phone while it is played with loudspeakers and the embedded data is extracted with application. The improved version of the method uses better error correction and synchronization methods, and the paper presents much more thorough testing. The author was responsible for the design of the method and implementation of the application. Ms. Keskinarkaus were responsible for the design of watermarking algorithm with the help of Dr. Cvejic and author. Ms. Keskinarkaus and author carried out the testing. Prof. Seppänen was the supervisor and also participated in the finalization of the manuscript.

# 2 Literature review

This chapter provides an overview of the literature related to DRM and audio watermarking. The chapter is divided into two parts. Section 2.1 deals with DRM related research and some DRM systems and standardization work. Section 2.2 investigates concepts related to digital watermarking, concentrating on audio watermarking, in particular.

## 2.1 Digital rights management

As the digital content started to become more widespread, DRM systems were developed to fight ever increasing content piracy. The possibility of making perfect copies of the digital content makes the piracy very tempting. The DRM systems try to prevent this using encryption and watermarking techniques, in particular.

Hietanen *et al.* (2008) conducted a survey among P2P network users to find out their attitudes towards copyrights. The results indicate that the users of the P2P networks are aware of the fact that they are breaking the copyright laws, and about half of them even think that what they are doing is morally wrong. The survey asserts that the biggest profit of using the P2P networks, naturally, was the access to a large base of digital content free of charge. The risk of getting caught was seen remote for the respondents.

Fetscherin & Schmid (2003a, 2003b) made an empirical survey of the use of DRM systems in music, film, and print industry. The results of the survey point out that the content providers were using a variety of content protection methods, passwords and encryption being the most common ones, each with its own goal. The respondents were satisfied with their level of protection but planned to enforce it in the future. Half of the respondents of the music industry were skeptical about the benefit of DRM in reducing piracy, whereas the respondents in film and print industry believed in the DRM to do it.

### *2.1.1 DRM systems*

The DRM systems usually comprise of encryption and key management, access control, copy control, identification, tracing, and billing mechanisms. The access control is done using a flexible set of usage rules that define what the user can do with the content. These are defined with rights expression languages (REL),

discussed in more detail in the next subsection. The access rules are defined in the license, which typically also contain the keys needed to access the content. Copy control usually utilizes encryption and watermarking methods, and it is used to prevent making unauthorized copies of the content, which is generally very hard to achieve in practice. Device compliancy helps the copy control. If the multimedia device does not allow free access to the content, copying is much harder than in an open PC environment. Thus the devices can have tamperproof hardware or trusted platform module (TPM), where the keys and licenses are stored, to prevent the user access to them. Finally, identification and tracking can be used as a last resort to track the source of pirated copies and enable legal action. (Hartung & Ramme 2000, Jonker & Linnartz 2004.)

One of the main principles of DRM systems is the separation of content from the rights. This allows the content to be distributed freely since the users are not able to consume the content without a valid license. The distribution can be done using offline delivery packaged in CD, DVD, etc., online distribution with downloading or streaming, or superdistributed from other users. (Subramanya & Yi 2006.)

The major architectural components present in most DRM solutions are shown in Fig. 1 (p.15). Jamkhedkar & Heileman (2004) created an alternative layered model of the DRM dividing the functionality of the DRM system in three blocks. The purpose of the layered framework is to reduce the complexity of systems and to enable vendors to develop systems focusing on different aspect of the DRM that could be made to interoperate. To start with, the upper layers deal with the end-to-end functions of the application. Its responsibility is to negotiate the used content formats, financial transactions, identification of the content, and protection methods used between the client and the server. The middle layers, in their turn, handle the rights expression and interpretation and it is the key element for a universal interoperable DRM. Finally, the lower layers ensure rights enforcement and the different content protection methods are implemented here. They later analyzed DRM interoperability from the viewpoint of layered architecture (Heileman & Jamkhedkar 2005). In (Jamkhedkar *et al.* 2007), middleware based on the layered model is proposed.

There are two major DRM standardization works being done, MPEG IPMP by the Moving Picture Experts Group (MPEG) and OMA DRM by the Open Mobile Alliance (OMA). The intellectual property management and protection (IPMP) is the MPEG's term for DRM. The MPEG IPMP standardization started with the MPEG-4 IPMP mechanism. It provided hooks, i.e. control points, which

could be used to plug in proprietary protection methods. Additionally, the protected content has associated IPMP system ID, which is used to identify what DRM tool is needed to consume the content. The content can also contain private data that is used by the DRM system. (Hartung & Ramme 2000, Rump 2004.) The MPEG-4 IPMP was later extended to allow interoperability and secure communication between different DRM tools. The interoperability is achieved so that the player identifies the DRM tools needed, and if some are unavailable, it can try to download them. (Rump 2004.) The MPEG-4 IPMP extension also provides ways for DRM tools to exchange messages, which was analyzed by Pang & Wu (2005). Furthermore, several content protection methods have been proposed to MPEG-4 IPMP (Lacy *et al.* 1999, Kim & Hong 2004, Senoh *et al.* 2004).

The MPEG-21 IPMP is based on extending the work done in MPEG-4 IPMP. In the MPEG-21, DRM functionality is in the core of the standardization; 4 out of 18 technical parts of the standard are dealing specifically with DRM issues: part 4 defines the IPMP components, part 5 provides an REL, part 6 specifies a rights data dictionary (RDD), and part 11 defines the best practices for evaluating persistent association technologies. The MPEG-21 IPMP (part 4) includes ways to download DRM tools from remote locations, exchanging messages between the tools, and between these tools and the terminal. It also addresses authentication of IPMP tools. It has provisions for integrating rights expressions according to the RDD and the REL. The REL (part 5) provides the syntax to create rights expressions governing the content, while the RDD (part 6) provides the semantics. Finally, part 11 deals with technologies that link information to identify and describe content with the content itself, such as watermarking and fingerprints, i.e. identifying the content based on its properties. (Bormans *et al.* 2003, Rump 2004.) Consequently, Sheppard (2007) describes an experimental implementation of the MPEG-21 IPMP components, including an interface for DRM tools and MPEG-21 terminals to communicate based on the MPEG REL, dynamic construction of licenses that permit a user to carry out an action, and a cryptographic architecture bound to the existence of licenses.

The other major DRM standard, the OMA DRM, is currently in its second version. The OMA DRM 1.0 is a simple protection system for low cost media objects (such as a ringtone or a background image) and supports three basic functionalities:

1. forward lock to prevent an unprotected media object from leaving the device,
2. combined delivery for distributing a protected media object jointly with the rights to consume it, and
3. separate delivery for a separate distribution of a protected media object and the rights to consume it.

The newer OMA DRM 2.0 focuses on mobile content protection. It provides more robust protection for the content using public key infrastructure (PKI). Each OMA DRM 2.0-compliant device has a unique set of keys that are used to protect the licenses. Additionally, with the OMA DRM 2.0, the user is able to create domains, which enables sharing of protected content between OMA DRM 2.0-compliant devices within a domain, such as home. Currently, OMA DRM 2.1 is under development and it will address detailed metering of content usage, binding content to secure removable media and secure content exchange to non-OMA DRM systems. (Buhse & van der Meer 2007.)

With respect to the OMA DRM, Soriano *et al.* (2005) present enhancements to the OMA DRM 2.0 by introducing a method for paying licenses via mobile operators, separate the DRM agent into device and user agents, and apply watermarking and fingerprinting as an additional means of copy detection. The DRM user agent, which manages the licenses, can move from device to another, thus the protected content can be used in different devices. Yu *et al.* (2005), for their part, propose a DRM scheme based on the TPM which enhances the security of the OMA DRM 2.0 and provides interoperability and compatibility between the TPM and the OMA DRM.

The mobile DRM platform developed in this work is much simpler than OMA DRM and MPEG-21 IPMP, especially with respect to the REL. The DRM standards and the mobile DRM platform are all based in the DRM reference model presented earlier. The standards have to be developed by taking into account many sorts of different applications that could be built using them. This makes them complex, whereas the platform developed only contains features which are necessary for testing the content protection methods created. This, on the other hand, allows easier development of protection methods since the platform can be updated if new features are required. Furthermore, one unique feature of the mobile DRM platform is that it includes a system which negotiates with the license servers a suitable license for the user. The protection methods created can be adapted to be used with the DRM standards.

An overview of some other commercial DRM solutions available can be found from the literature (Liu *et al.* 2003, Jonker & Linnartz 2004, Michiels *et al.* 2005, Serrão *et al.* 2007). With respect to other DRM related research, Merabti & Llewellyn-Jones (2006) address the DRM from the point of view of ubiquitous computing. They propose a distributed trust paradigm for the DRM systems in the ubiquitous environment. Adelsbach *et al.* (2005), in their turn, describe a multilateral DRM infrastructure where the users will be able to verify that the seller is actually the owner of the content. Additionally, they propose a method for describing the ownership of perceptibly similar content which are modifications of the original content. A DRM enabled P2P architecture has been proposed in (Sung *et al.* 2006, Lou *et al.* 2007).

### 2.1.2 Rights expression languages and DRM interoperability

Several rights expression languages (REL) have been developed over the years with the eXtensible rights Markup Language (XrML) and the Open Digital Rights Language (ODRL) becoming the most popular. Some time ago, the XrML was accepted as the standard REL for the MPEG-21 IPMP, and ODRL was selected as a basis for the OMA DRM. However, these RELs have not been widely used in DRM applications. Instead, most commercial DRM systems do not use REL, despite the fact that the use of the REL would make DRM interoperability much easier. The problem with the current RELs is that they are too complex to be easily implemented in DRM systems. (Jamkhedkar *et al.* 2006.)

Both the XrML and ODRL are extensible XML-based markup languages to specify rights and conditions to control the use and distribution of digital content and to access services. There are four elements at the core of the XrML (resource, principal, right, and condition) which determines the right to perform action on certain content or service and also who can do it and under what conditions. (Wang *et al.* 2002.) The ODRL, on the other hand, consists of three similar entities (principal, permission, and asset) which differ from the XrML ones in that the permission includes the conditions. (Polo *et al.* 2004.)

A shortcoming of the XrML is that it does not support negation, which means that, if some action is not specifically permitted, it is forbidden. This limits what kind of licenses can be defined. What is more, the XrML does not have formal semantics. Instead, the XrML specification presents the semantics in two ways, first is a description of the language, and second, a description of an algorithm to determine if permission follows from a set of licenses. Unfortunately, the

specification does not specify in which order the licenses should be tested, because of which the output depends on the implementation of the algorithm. As a result, the MPEG-21 REL has been changed to address these problems. (Halpern & Weissman 2008.)

The XrML based MPEG-21 REL is defined with extensions using the XML Schema and XML namespaces. The core contains definitions of the REL's authorization model. Next, the standard extension contains definitions of concepts generally and broadly useful and applicable to DRM usage scenarios. Finally, the multimedia extension contains definitions of DRM concepts related to multimedia content. (Wang 2004.) The MPEG-21 RDD provides the semantic to the REL. The MPEG-32 RDD is based on a logical structure, referred to as the context model, which is used to construct natural language ontology for terms used in rights management. Its baseline dictionary contains approximately 2,000 terms. Additionally, the MPEG-21 RDD includes 14 RDD actions that can be used in REL grants. Fourteen actions have been selected to cover the most common actions a user might wish to undertake with respect to digital content. (Wang *et al.* 2005.)

Concerning the DRM systems interoperability, the systems must be able to translate the content providers' rules from one system to another and to ensure that devices and content services that use different DRM systems enforce the rules in the same way (Geer 2004). The first option is to directly translate the license from one REL into another. This is sometimes impossible since there is no conceptually equivalent element in the target language. Nevertheless, this can be avoided by using profiles which are a limited subset of the source REL. The second way of translating RELs is to use an intermediate schema which can describe the elements in both RELs. (Cooper & Montague 2005.)

There are a number of papers which propose methods for making different DRM systems interoperable. Guth *et al.* (2003), for example, describe a generalized contract schema CoSa which provides a means for the generic representation of digital contracts formulated in arbitrary RELs. They also implemented an interpreter which can translate ODRL licenses into CoSa. Polo *et al.* (2004), in their turn, developed tools to translate expressions from ODRL to MPEG-21 REL and vice versa. Moreover, Safavi-Naini *et al.* (2004) examined the translation of licenses between MPEG-21 REL and OMA REL.

In contrast, networked environment for media orchestration (NEMO) is an experimental framework for the discovery, access, composition, and orchestration of media-related online services. It does not enable DRM system interoperability

*per se*; however, it provides a way for different DRM systems used by content providers and consumers to communicate with one another (Koenen *et al.* 2004). Other interoperability methods are proposed in (Kravitz & Messerges 2005, Chang *et al.* 2007, Chen & Huang 2007, Jeong *et al.* 2007, Nam *et al.* 2007a, Nam *et al.* 2007b).

### 2.1.3 Implementing DRM system

The most commonly used protection method employed in the DRM systems is the cryptography and digital watermarking, the latter of which is discussed in more detail in Section 2.2. Encrypting can be divided into two major types: symmetric and public-key algorithms. In symmetric algorithm, the same key is used in both encryption and decryption and the key is shared with the sender and the receiver. Some widely used symmetric algorithms are Data Encryption Standard (DES) and its variation 3DES (Schneier 1996) and Advanced Encryption Standard (AES) (NIST 2001) which has replaced the 3DES as the encryption standard by the U.S. government. The concept of public-key algorithms was first suggested by Diffie & Hellman (1976) to solve the problem of the shared secret. The public-key algorithms use public/private key pairs in which one key is used to encrypt and another to decrypt. The public-key algorithm known most commonly is RSA, named after its creators Ron Rivest, Adi Shamir, and Len Adleman. Since public-key algorithms are much slower than symmetric algorithms, they are usually combined so that the message is first encrypted with symmetric algorithm and the key is encrypted with public-key algorithm. (Schneier 1996.) As a rule, it is important to note that you should never try to create your own encryption algorithm unless you are expert in cryptography; otherwise, you are almost guaranteed to result in unsecure algorithm.

Another important cryptographic method used in DRM is digital signatures. These are used to sign the licenses so that the DRM system can be sure that the license has not been forged by a malicious user. The digital signatures use one-way hash functions which take variable length input and turn it into fixed length output called the hash value. One-way function is a function which is easy to compute but very hard to reverse. This means that it must fulfill the following requirements: First, given a hash value, it should be hard to find a message that produces that hash. Second, given a message, it should be hard to find a message that produces same hash. Finally, it should be hard to find two messages that produce the same hash.

Some well-known hash algorithms are Message Digest algorithm 5 (MD5) and Secure Hash Algorithm 1 (SHA-1) and their variations. The actual digital signature is created by using the hash function on the object being signed and encrypting the hash value with the signer's private key. The signature can be verified by calculating the hash value from the object, decrypting the signature with the signer's public key, and comparing these two values. (Schneier 1996.)

Since it is impossible to know the owner of the key pair, certificates are used to link the identity of the key pair owner and the public/private key pair. The certificates are issued by a certification authority, which is a trusted party by the sender and the receiver, by using a digital signature to bind the identity and the key pair. This concept is known as public-key infrastructure (PKI). (Adams & Lloyd 1999.) With respect to the PKI, Serrão *et al.* (2007) have analyzed the key management in some open DRM systems.

Regarding the copy control, the trusted computing module (TPM) (Pearson 2003) consists of secure hardware designed to minimize the opportunity for software-based attacks. It provides a function called remote attestation, which allows a content provider to refuse services to a platform in an unapproved state. A TPM based DRM system is introduced in (Cooper & Martin 2006) which does not limit the user to select their operating-system and applications. A key component in the architecture is a security manager that enforces mandatory access controls on shared devices, restricted information flows between virtual machines and DRM policy on protected objects. Additionally, the system provides content usage count restrictions by storing a counter in the TPM. With respect to that, Shapiro & Vingralek (2002) review several methods that could be used to store persistent data on hostile environment such as a user's terminal.

Kwok (2002) studied the license management models of some commercial DRM music systems and proposes an enhanced license management model which allow both in online and offline purchases. Ziolkowski & Stoklosa (2007), for their part, introduce an agent based DRM system in which the DRM agent is capable of an autonomous searching of content from different providers and determining which offer is the best in terms of price and usage conditions. On the other hand, Abbadi & Mitchell (2007) propose a DRM system in which the identification of a device ownership is done on the basis of a mobile phone. Morin & Pawlak (2008), in their turn, have studied a exception aware DRM system. The exception is defined to be any legitimate access attempt to a protected content, which has not been anticipated by the resource owner when

protecting it. They propose a credential based system in which some rights should be waived in case of exception.

Superdistribution is a multi-level networking approach where digital content is distributed by the users, for example, via e-mail, P2P platforms, Bluetooth, memory sticks or other distribution channels. It adopts the idea of viral marketing for the content industry and utilizes existing social networks. (Küpper *et al.* 2007.) With respect to that, Chong & Deng (2006) present a DRM system which superdistributes encrypted layered content and allows the user to select the desired quality level of the content. The mobile DRM platform developed allows both the content and the DRM player software to be superdistributed directly from device to another or with P2P networks.

Concerning the restrictions in the licenses, Mundt (2005) proposes a location dependent DRM system in which the constraint for using protected content can be location based. The user's location is tracked with a satellite positioning system. Similar location based DRM system is discussed in (Surminen *et al.* 2007). It utilizes wireless local area networks (WLAN) to verify the location of the user. The system enables restricting the use of the content for example in a certain building where the WLAN access points are located. Xiaosong *et al* (2007), in their turn, propose a P2P framework with DRM components. Their method utilizes selective content poisoning in which all distribution agents and customers will send poisoned content to unpaid peers while sharing clean content among each other. This creates a huge difference in the download performance between paid customer and unpaid peers. Finally, Venkatachalam *et al.* (2004) discuss a music identification system based on acoustic DNA called MusicDNA. Acoustic DNA is a fingerprint calculated from audio which identifies the sound recording. MusicDNA fingerprint is based on energy values computed from frequency bands which form two vectors that are normalized and concatenated to form a 30-component fingerprint.

### *2.1.4 Copyright legislation and critical view of DRM*

The history of the intellectual property (IP) starts officially with 1709 Statute of Anne, known as the first copyright law. The first major international treaty concerning industrial property was made in 1883. Next, in 1886 the international copyright act was adopted and later revised in 1912 and 1957. Later, the copyright laws changed to digital era in 1998 with the U.S. digital millennium copyright act (DMCA) and in 2001 with the European Copyright Directive. Both laws make

illegal circumventing the technical security measures, and manufacturing and trading equipment which help to circumvent them. (Borda 2005.)

Copyright does not protect the ideas themselves, instead it protects the way in which they are expressed and it arises automatically with the creation of a work. The copyright expires after a certain number of years after the death of the author and after that the work is in the public domain. Copyright has exceptions of which the most prominent is for fair use which allows for a limited reproduction of the creative work without permission. However, the major problem with the fair use is that the difference between it and copyright infringement is vague. (Owens & Akalu 2004.)

The enforcement and recognition of intellectual property rights is an international concern and the World Intellectual Property Organization (WIPO), a specialized agency of the United Nations, was created to promote and protect IP around the world. Much of the recent legislation around the world, including the DMCA and its European counterpart, which relate to the DRM have been a result of two WIPO treaties. (Owens & Akalu 2004.) The anti-circumvention laws, enacted based on the WIPO Copyright Treaty, are used to give a legal protection to the DRM systems. Regarding the anti-circumvention laws, Akalu & Kundar (2004) review the trials concerning the Content Scrambling System (CSS) used in DVDs. It portrays the difference in interpretation of the copyright laws in the U.S. and Europe. De Rosnay (2002) discusses the requirements the copyright laws impose for a good DRM system.

Erickson & Mulligan (2004) study the concepts and architecture of policy specification and enforcement, and how usage control policies are evaluated in DRM systems. They also discuss the challenge of coding copyright law in DRM systems. In (Felten & Halderman 2006), Sony-BMG's content protection methods XCP and MediaMax are studied which use active protection methods to prevent reading the CD. The XCP installs in the user's computer as a rootkit, which hides it from the user and introduced a vulnerability to malicious programs which was abused shortly after its release. In the light of the XCP incident, Bishop & Frincke (2006) raise questions about situations where the interests of the content providers and the users differ, whose policy and expectation dominate, and what sorts of defenses are appropriate, and in which situations.

Feigenbaum *et al.* (2002) discuss privacy problems in the DRM systems. They propose a list of privacy engineering principles for the DRM systems which would improve the privacy protection. On the other hand, Camp (2003) argues that the DRM suppresses creativity and innovation by restricting the use of the

content too strictly. Moreover, Stini *et al.* (2006) claim that the DRM artificially reduces the value of the digital content by limiting the actions that the user can perform on it. Thus, a user who legally purchases DRM protected content not only has to pay for the content, but also gets an inferior product compared to someone who acquired an illegal unprotected copy for free. They propose that the DRM systems should concentrate on managing and protecting content ownership rather than managing and protecting the content itself. Finally, Lesk (2003) envisions a world with a perfect DRM system, and how would such world differ from the current one.

## 2.2 Digital audio watermarking

Digital watermarking is a method for hiding additional information into digital content by utilizing imperfections of the human perceptual system. Watermarking has been used in applications like broadcast monitoring, owner identification, proof of ownership, authentication, fingerprinting, copy and access control, and information carrier (Cvejic & Seppänen 2008).

### *2.2.1 Digital watermarking concepts*

A watermarking system is illustrated in Fig. 3. The watermark, which is typically a binary sequence *b*, is embedded in the cover signal A in a watermark embedder. The watermark and the cover signal are coupled tightly together so that the watermark cannot be removed easily from the cover signal. The embedding is usually done using a watermarking key K so that the security does not depend on algorithm being secret, called Kerckhoffs' principle. The embedder outputs watermarked signal $A_w$ which cannot be perceptually distinguished from the cover signal. The watermarked signal passes through a watermark channel in which the signal can be processed and/or attacked. This transforms the watermarked signal to A'. Finally, the watermark is extracted in a watermark detector. The detector uses the watermarking key to either extract the embedded message or to detect if a known watermark *b* was embedded in it. (Barni & Bartolini 2004.)

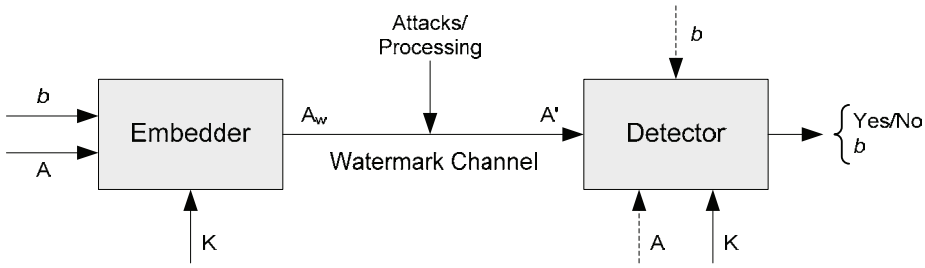


**Fig. 3. Block diagram of watermarking system.**

The watermarking schemes can be divided into three categories based on detector requirements. If the detector requires the original cover signal, the scheme is called non-blind. The blind scheme, in which the cover signal is not needed, can be further divided into two, with and without synchronization information. In the blind schemes without synchronization, the watermark embedding is done using a feature points from the cover signal. (Cvejic & Seppänen 2008.)

The main requirements of the watermarking system can be visualized as a "magic triangle" (Fridrich 1999) shown in Fig. 4. The requirements, that is, imperceptibility, capacity, and robustness, are contradictory. The operating point of the watermarking system has to be selected based on requirements of the application. Sacrificing the capacity of the watermark by using error correction, the robustness can be increased. Using a lower embedding strength the watermark will become less perceptible, but the robustness will lower, too. If malicious attacks are not expected in the application, the robustness requirements are somewhat relaxed. To sum up, all the requirements cannot be maximized at the same time. (Cvejic & Seppänen 2008.)

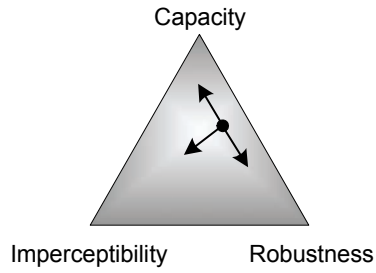


**Fig. 4. Magic triangle—three contradictory requirements of watermarking.**

Mintzer & Braudaway (1999) were among the first to suggest embedding multiple watermarks in the same cover media. The reasoning for using more than one watermark is that watermarks may be needed for different purposes in the application, one for owner identification and the other for authentication, for example. The different watermarks can have varying robustness and capacity requirements. Sheppard *et al* (2001) were the first to classify multiple watermarking schemes. The three classes are called re-watermarking, segmented watermarking and composite watermarking. In the re-watermarking, the multiple watermarks are embedded in on top of each other, while in the segmented watermarking, in its turn, the space available for watermarking is divided and allocated to different watermarks. Finally, the composite watermarking builds a single composite watermark from a collection of watermarks. Consequently, Liu & Smith (2004) examined a mathematical model to embed two watermarks with different robustness or capacity requirements, either using re-watermarking or segmented watermarking. They demonstrate that the re-watermarking achieves higher data hiding rates than the segmented watermarking. Additionally, their results show that optimal data hiding requires embedding of the more robust watermark first. On the one hand, Kundur & Hatzinakos (1998) suggested using a known reference watermark to characterize attacks before blind extraction of data watermark. They were also the first to suggest using attack characterization in watermarking.

Petitcolas *et al.* (1999) proposed the following set of requirements for a robust watermarks, which are a good guideline for building a watermarking system:

- The watermark should not degrade the perceived quality of the cover signal.
- The watermark detection should require knowledge of a watermarking key.
- Multiple watermarks embedded in a single object should not interfere with each other.
- The users should not be able to modify the embedded information.
- The watermark should survive all attacks that do not degrade the cover signal's perceived quality.

Furthermore, perceptible watermarks can be used to protect the content in a more active manner. They are usually used in applications in which the content is freely available to the users and the unauthorized usage is prohibited. By using perceptible watermarking to embed a copyright mark, the owner of the content can be identified perceptually while preserving the perceived quality of the

content. (Huang & Wu 2004.) The perceptible watermark usually should be permanent to increase robustness against unauthorized attempts to remove it. However, there are applications in which the watermark needs to be removable or reversible. For instance, in (Mintzer *et al.* 1997), visually watermarked images are posted on the Internet. The images serve as a teaser and are freely downloadable by the users. The user is able to remove the visual watermark with a program available for a fee. Although the reversible watermarking has been widely researched, the research has mainly focused on reversible watermarking of images. A review of the reversible watermarking research status is found in (Feng *et al.* 2006). Nevertheless, a perceptible removable watermarking has been rarely studied in the literature. Hu *et al.* (2006), in their turn, propose a removable visible watermarking method in which the embedded watermark is user keys dependent. The watermark removal is not complete, however, and some residual distortions are left after removal. In (Hu & Jeon 2006), a reversible visible watermarking method is presented which is able to recover the original image losslessly.

### *2.2.2 Human auditory system and watermark quality*

Audio watermarking is harder than image watermarking since the human auditory system (HAS) has a wider dynamic range than the human visual system. The HAS perceives sounds over a range of power greater than $10^9$:1 and range of frequencies greater than 1000:1. In addition, the HAS is very sensitive to additive random noise, it can be detected as low as 80 dB below ambient level. However, the differential range of the HAS is fairly small, i.e. loud sounds generally tend to mask out quiet sounds. Moreover, the HAS cannot perceive the absolute phase, only the relative phase. There are also some environmental distortions which are so common that the listener usually ignores them. (Bender *et al.* 1996.)

The HAS can be modeled using a set of overlapping bandpass filters for frequency bands called critical bands, in which the humans cannot perceive a difference. In fact, at low frequencies, the bandwidth of a critical band is about 100 Hz, while at frequencies above 500 Hz the bandwidth is about 20 % of center frequency. As a consequence, two properties of the HAS mainly used in audio watermarking are frequency or simultaneous masking and temporal masking. The frequency masking is a phenomenon in which a low level signal (e.g. a pure signal) can be made inaudible by a stronger signal. A masking threshold is defined to be a sound pressure level below which a signal is inaudible. Hence, the level of

the masking signal on a critical band, in which the masked signal is, determines the masking threshold. On the other hand, without a masking signal, a signal is audible if its level is above the threshold in quiet. Furthermore, the masking effect is not limited to simultaneous signals. This phenomenon, called the temporal masking, can be divided to premasking and postmasking in time. The premasking effect appears just before the masking signal is switched on and its duration is only about 20 ms. Likewise, the postmasking effect is present after the masking signal has been switched off and its duration is much longer, from 50 to 200 ms, depending on the duration of the masking signal. (Zwicker & Fastl 1999.)

The MPEG-1 audio standard (ISO 1993) provides two implementations of psychoacoustic model, i.e. model of human perception of sounds. Psychoacoustic Model 1 is less complex than Psychoacoustic Model 2, and thus makes more compromises to simplify the calculations. To begin with, MPEG-1 audio layer 3 (MP3) compression divides the frequency into 32 subbands. Further, the Psychoacoustic Model 1 uses 1024 point Fourier transformation to calculate frequency components. The frequency components are grouped based on critical bands, on which tonal and noise-like components of the signal are separated. The model identifies tonal components based on the local peaks of the audio power spectrum. Next, the remaining spectral values are summed into a single non-tonal component per critical band. The model determines the masking thresholds by first applying an empirically determined masking function for each critical band. Second, the threshold in quiet is calculated as a lower bound for audibility. Finally, a minimum masking threshold is selected within each 32 subbands as a subband threshold. (Pan 1995.)

The quality of watermarked audio can be evaluated either using subjective or objective measures. In fact, evaluating watermarked audio quality is similar to the problem of evaluating audio codecs, hence the same principles and test methods can be applied. The subjective testing can be done in two ways, testing the transparency of the watermarked samples or rating the quality of the watermarked samples with respect to the reference samples. The testing for transparency can be performed by a so-called two-alternative-forced-choice test. In the test, a number of sample pairs are randomly chosen from the set of possible combinations [(O,O), (O,W), (W,O), (W,W)], where ”O” denotes the original and ”W” the watermarked sample. The test subject is asked whether both items were equal or not. As a result, a test for non-transparency is performed by trying to reject the transparency hypothesis. (Arnold 2002.) Moreover, the subjective rating of the sample quality can be made according to International Telecommunication Union

recommendations (ITU 1997, 2003). The test subjects are asked to rate the audio quality of a sample sequence using a 5-point impairment scale: (0: imperceptible, −1: perceptible but not annoying, −2: slightly annoying, −3: annoying, −4: very annoying). Finally, subjective difference grade (SDG) is calculated using the test results as:

$$SDG = Score_{signal\ under\ test} - Score_{reference\ signal} \quad (1)$$

The objective quality evaluation has also been standardized by ITU (2001) and is called perceptual evaluation of audio quality (PEAQ). It should replace the ITU-R BS.1116 (ITU 1997) standard, and it is very sensitive and enables the detection of even small distortions, and therefore should be only used with high quality audio material. In the PEAQ, both the reference signal and the signal under test are processed by an ear model to calculate objective difference grade (ODG). (Arnold 2002.) Alternatively, Creusere *et al.* (2008) have developed an objective quality metric which is based on model output variables of ITU-R BS.1387-1 (ITU 2001) and that can accurately quantify subjective quality over audio fidelities, ranging from highly impaired to perceptually lossless. In contrast, Lin & Abdulla (2008) propose using objective quality measures adopted in speech processing for objective quality evaluation in audio watermarking. The quality measures provide a faster and more efficient method of evaluating compared to the perceptual models.

### 2.2.3 Attacks against watermarks and synchronization

An attack against watermark can be defined as any processing which prevents the intended purpose of the watermarking technique for a given application. Hence, attacks can be divided into two groups, malicious or intentional and unintentional attacks. An attack is considered successful if it maintains the perceived quality of the content while circumventing the watermarking technique. Arnold (2003) categorized attacks into four different groups of removal, desynchronization, embedding, and detection attacks, respectively. The first two attacks prevent the watermark from being detected. The embedding attack embeds a new watermark to the content and results in a false watermark being extracted. For one thing, the aim of a copy attack is to copy a watermark from one cover signal to another. For another, in overmarking a second watermark is embedded in an already marked carrier signal. Finally, in the detection attack, the watermark is extracted

unauthorized. The unauthorized detection of the watermark is usually used as a preceding step before removal attack.

Consequently, Steinebach *et al.* (2001) classify removal/desynchronization attacks in the following groups:

- Dynamics – These change the loudness profile of an audio file.
- Filter – Filters cut off or increase a selected part of the spectrum.
- Ambience – This group consists of audio effects simulating the presence of a room.
- Conversion – Audio material is often subject to format changes.
- Lossy compression – Audio compression algorithms based on psychoacoustic effects.
- Noise – Noise can be the result of most of the attacks described above.
- Modulation – Modulation effects like vibrato, chorus, amplitude modulation or flanging are usually not used in postproduction.
- Time stretch and pitch shift – These either change the length of an audio event without changing its pitch or change the pitch without changing the length.
- Sample permutations – This group consists of algorithms not used for audio manipulation in usual environments. These are specialized ways to attack audio watermarks.

Lang *et al.* (2005) have developed audio benchmarking suite called StirMark benchmark for audio. StirMark contains a wide range of processing attacks against audio watermarks divided into add/remove attacks, filter attacks, and modification attacks based on processing type. Besides performing single attacks, StirMark includes another attack mode, called profile attack, to run more than one attack in a serial order. Unfortunately, at the time of writing, StirMark benchmark for audio is no longer available for free download; however, researchers can contact the author of the StirMark[1] to evaluate their algorithms. Steinebach *et al.* (2002) present a test environment for noisy, analog channels. They analyze if it is possible to simulate an acoustic, noisy DA/AD environment with filters, quantization, and noise generators. Based on the results, they identify the parameters relevant for watermarks to successfully survive noisy acoustic channels. Additionally, they describe a design concept for a DA/AD simulation. Xiang & Huang (2006), in their turn, analyze the performance of quantization-

[1] http://wwwiti.cs.uni-magdeburg.de/~alang/smba.php

based audio watermarking against the DA/AD attack and conclude that it is very sensitive to it. They also investigate the effect of DA/AD transformation to the audio signal.

The fingerprinting application of the digital watermarking, i.e. an application where the ID of the content user is embedded in the content, is suspect to a collusion attack. In such an attack, the attacker has access to a set of watermarked content with unique a message embedded in each of them. The attacker averages the content files in order to remove the watermark from them. Liu *et al.* (2005) concentrates on the collusion attack and fingerprinting methods which are collusion resistant and can detect a number of attackers. The drawback of the anti-collusion codes is that they usually require a very large number of bits to be embedded, and thus high capacity for the algorithm, if the number of different fingerprints is large. Robert & Picard (2005), in their turn, discuss the vulnerability of the watermarking system to the so-called mask attack. The mask attack utilizes the knowledge on the perceptual properties deduced from the masking models to estimate and remove the embedded watermark. As a result, they suggest that for audio watermarking using high frequencies and dominant tonal components should be avoided.

Semifragile watermarking has the property that it is robust against certain attacks but fragile against others. Thus, it can be used in authentication applications, in which the watermark should be robust against attacks that do not alter the perceived view of the content such as lossy compression. Regarding that, in (Tu & Zhao 2003), a semifragile audio watermarking method is described.

Synchronization is one of the most challenging elements of a watermarking system. The synchronization is an essential element of every digital communication system; however, in watermarking it poses unusual and challenging new problems since the primary goal is not the communication of the watermark but that of the multimedia information. Sharma & Coumou (2006) review methods suggested for watermark synchronization and propose a feature-based synchronization paradigm. Zaidi *et al.* (2006), for their part, study the impact of additive noise, from which the part that is signal-like has been removed, to the watermarking. This desynchronization noise is shown to more accurately characterize the attack impact on the watermark. They, additionally, investigated optimal attacker and defender strategies in a game theory context.

Finally, Saied-Bouajina *et al.* (2004) view the analyzed audio watermarking scheme as a communication system. They investigate an error correction strategy based on the modification of the decoder metric. This modification is related to

the *a priori* knowledge of the probability density function of the channel noise. As a result, they propose a generalized Gaussian distribution model, which is exploited to design an optimal channel Viterbi decoding structure.

### 2.2.4 Audio watermarking algorithms

This subsection overviews the audio watermarking algorithms proposed in the literature over the past few years. A broad range of different embedding techniques have been suggested from the least significant bit (LSB) scheme through using transform domains to spread spectrum methods.

The LSB method is one of the first techniques proposed for audio watermarking. The main advantage of the LSB method is a very high capacity and a low computational complexity of the algorithm, while its main disadvantage is considerably low robustness. In fact, the LSB is one of the simplest algorithms: the watermark embedder selects a subset of all available cover audio samples chosen by a key and substitutes the original bit values with the watermark values. The detector simply extracts the watermark by reading the value of the selected bits from the watermarked audio. The robustness can be increased by using higher LSB layers which introduces more distortions to the cover audio, however. This can be reduced by flipping the lower LSB layer bits to minimize the embedding error. (Cvejic & Seppänen 2005.)

Echo hiding was first described in (Bender *et al.* 1996). In the echo hiding, the watermark is embedded into a cover audio signal by adding echoes, which are shifted and downscaled copies of the original signal. Thus, the echo is perceived by the HAS as an added resonance. The embedder sets the time delay between the original signal and the echo by selecting a kernel used to add the echo corresponding to the watermark bit. Regarding that, Kim & Choi (2003) propose a method that uses both backward and forward kernels. The forward kernel adds echoes before its original signal exists. It performs best when the kernels are symmetric, i.e. they have the same delay from the original signal. Ko *et al.* (2005), in their turn, recommend spreading the echo in time domain using pseudorandom sequence. Embedding watermark with various delays from the original signal creates complicated echoes which, as a consequence, can provide a more natural sound quality than using only a single echo or several multiple echoes.

A patchwork scheme was also presented in (Bender *et al.* 1996) for images. It is a statistical method based on hypothesis testing and relies on a large data set.

The embedding process artificially modifies a certain statistic in a cover audio to be many deviations away from the expected value. The two major steps in the algorithm are: first, choose two patches pseudorandomly and second, add a small constant value to the sample values of one patch and subtract the same value from the sample values of another. The extraction is done with the subtraction of the sample values between two patches, and if the results differ enough from the expected value, the watermark has been detected. Moreover, the watermark embedding is usually done in transform domain in order to spread the watermark in time and to increase robustness. Yeo & Kim (2003) presented a modified patchwork algorithm which uses adaptive modification value and additive embedding, i.e., watermark is added to the cover signal.

Spread spectrum watermarking is the most widely used audio watermarking scheme. There are two types of spread spectrum technologies: frequency hopping (FH) and direct sequence spread spectrum (DSSS). In frequency hopping, the narrow band signal is not spread into a wideband signal; instead, the signal's carrier frequency hops from channel to channel at different times. Hence, the FH system has to use more power than DSSS system to achieve the same signal-to-noise ratio. However, the FH methods have usually larger capacity than DSSS methods. (Cvejic & Seppänen 2004) is an example of FH watermarking algorithm. The DSSS, on the other hand, uses a pseudorandom sequence to spread a narrowband signal into a signal with a much wider bandwidth. The watermark detector usually calculates correlation between the received signal and the pseudorandom sequence to extract the watermark. (Cvejic & Seppänen 2008.) In particular, many different direct sequence spread spectrum based watermarking algorithms have been proposed in the literature (Kirovski & Malvar 2001, Seok & Hong 2001, Esmaili *et al.* 2003, Liu & Inoue 2003). Most of the watermarking methods employed in this thesis use the spread-spectrum watermarking in either way.

The watermarks can be embedded in the cover audio in time domain or some transform domain such as Fourier domain. To begin with, Bassia *et al.* (2001) describe a method which divides the cover audio into segments and modifies the amplitudes of the audio samples, whereas in (Lemma *et al.* 2003), a watermarking method based on envelope modulation is presented. The short-time envelope of the audio signal is modified in such a way that the change is imperceptible to the human listener. The watermark is extracted from the energy of the envelope using a correlation with reference signal. Lie & Chang (2006), for their part, propose a low frequency amplitude modification algorithm. The amplitude modification is

employed to scale amplitudes in a group manner instead of sample-by-sample manner. No basic form of the watermark signal is set in advance; instead, segments of host audio signals are modified based on the principle of differential amplitudes. Further, in (Xiang & Huang 2007), a multibit algorithm based on the two statistical features in time domain, histogram shape and modified mean value has been studied. Wu *et al.* (2005), in their turn, describe a self-synchronization algorithm for audio watermarking, in which the synchronization codes and the hidden data are embedded in the low frequency coefficients of wavelet domain. Finally, Wang & Zhao (2006) propose synchronization invariant audio watermarking scheme based on wavelet and cosine transform. The algorithm is based on the quantization of coefficients of the both wavelet and cosine transform.

Algorithms that embed watermark in the phase of the cover audio utilize the insensitivity of the HAS to a constant relative phase shift. Thus, the phase coding method works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments. (Bender *et al.* 1996.) Takahashi *et al.* (2005) propose a multiple watermarking method based on phase modulation that can embed three watermarks in audio content by applying frequency division multiplexing.

With respect to other watermarking techniques, Li *et al.* (2005, 2006) propose audio watermarking algorithm based on music edge detection. The key point lies in the fact that music edges like drum sounds will not be changed much in order to maintain high auditory quality. Next, Erküçük *et al.* (2006) suggest using linear chirps, i.e. signals whose frequency changes linearly with time, as a watermark signal. Different chirp rates are used as symbols to different messages. Yet in (Wang *et al.* 2007), use of support vector machines for audio watermarking has been studied. The algorithm embeds template information and watermark signal into the original audio. In the extraction, the corresponding features of template and watermark are extracted and the template is used as a training sample to train the support vector machine, which is used to extract the watermark. Lastly, Larbi & Jaidane-Saidane (2005) discuss using watermarking as a preprocessing step for further audio processing systems. The watermark conveys no information; rather, it is used to modify the statistical characteristics of an audio signal in order to stationnarize the host signal.

The watermarking algorithms used in the original publications are based on the work done in (Cvejic 2004). The algorithms have been modified, however,

according to the application specific requirements. In addition, (Cvejic & Seppänen 2008) contains an excellent compilation of references to the audio watermarking publications.

# 3 Research contributions

In this chapter, a summary of contributions described in detail in the original publications is given. Additionally, it discusses how the original publications answer the set research problems. Papers I and II describe the DRM platform that was developed. In Papers III and IV, audio protection methods utilizing digital watermarking techniques are described. Papers V and VI, in their turn, focus on the issue of counting how many times the user has played protected content to enforce restrictions set by the license. Finally, in Papers VII and VIII, a method for linking analog audio to network based services is presented.

## 3.1 Mobile DRM platform

This section describes the DRM platform which was developed to test individual protection methods, as well as to research content superdistribution and rights management methods. The first version of the platform presented in Paper I uses PKI and watermarking to protect the audio files. Paper II describes the second version of the platform which adds support for the mobile devices and content superdistribution. Additionally, license management and content protection methods are improved. With respect to the first research problem in this thesis, the platform developed has flexible and extendable support for content protection methods with varying complexity and security; in addition, it provides mechanism for content and DRM player superdistribution. It should be noted, that the key management and distribution has not been thoroughly studied in this thesis. It has been implemented in simplified way and should be researched more in the future.

The architecture of the DRM platform developed is similar to that in the DRM reference architecture illustrated in Fig. 1 (p.15). It has three main components: content server, license server, and client. The license server and the content server may be deployed into same physical server or they can reside on different locations. The platform also uses services provided by certification authority (CA) which is part of the PKI to provide the users an encryption key pair and a certificate to link the identity and the key pair. Both servers are implemented with Java and they are run on a PC. On the first version of the platform, the client run also on a PC, whereas on the second version a support for Symbian smart phones was added.

In the first version of the platform, audio is the only media type supported. Audio content is protected by embedding a spread spectrum watermark in it. The watermark has three functions. First, it is used to identify the audio content so that correct license can be acquired for it. Second, it informs the multimedia player that the content is protected and it should not be played without a valid license. Third, the watermark can be used to store the ID of the content owner, distributor, or end-user that downloaded the content. The audio is additionally encrypted using AES algorithm after the watermark has been embedded. The beginning of the audio is not encrypted and it is used as a teaser, i.e. a preview to the content.

When the user wants to play the protected audio, the DRM player extracts the watermark from the teaser part of the content. The player then checks whether the user has a valid license to play the audio. If the user does not have a license and she wants to get one, the player connects to the license server, whose address was stored in the watermark, to buy one. The licenses are XML files that are signed with the license server's private key. The license defines what the user is allowed to do with the content and what restrictions apply, such as validity period or limited number of plays. Additionally, the license file contains the decryption key to the content, which has been encrypted with the user's public key. The player checks that the license is valid and the user has not tampered it by comparing the signature on the license to the license server's certificate created by the CA.

The second version of the platform adds supports to images and video content. Default method for content protection is encryption with AES algorithm without a watermark. The media file has a header that contains information about the protection method used and unique identifier associated with the content, which is employed to identify the file on the server. The platform can be extended to use other protection methods that have been developed, which are described in the following sections.

Since the content is protected from unauthorized usage without a valid license, it can be distributed to the users using multiple methods. The content server keeps a list of content available for direct download from the server, and this is the primary method of importing new content to the system. After the user have received content, she can superdistribute it to other users, for instance, using mobile P2P networking. Additionally, both the content and the DRM player application can be shared over Bluetooth connection. This eases the superdistribution since the user does not need to download the DRM player from the server and the data transfer is free of charge. To enhance the content

superdistribution, rewards can be offered to the users who help to distribute the content to other users.

The user acquires licenses for the content with the help of a license negotiation system (LNS). The LNS operates in client/server model. The LNS client is used by the user to specify the parameters of the license she wants to acquire. The parameters define the rights and restrictions of using the content. The client sends the license request to the LNS server which contacts the license servers and asks for license offers. The user then selects the offer that is most suitable for her and after paying the offer, the license server generates the license for the user.

The key elements of the DRM architecture are content protection and content distribution. It is important that the DRM architecture is flexible and it can be extended to support a protection method with varying complexity and security. This helps to improve the system security; when the protection methods are broken, new ones can be introduced. This is the important lesson that was learned from DVD copy protection method CSS. The protection method used in DVD was not designed to be updated or exchanged; therefore, when it was broken in 1999 all the current and upcoming DVDs were compromised. The support for different protection methods is realized in the second version of the DRM platform presented in Paper II by using an extra header in the protected media to identify what protection method is used. The DRM client running in the user's device identifies the method used and launches a corresponding method to unprotect the content for consumption if the user has a valid license.

Another important part of the DRM system is the easy content distribution. This is especially important with mobile devices where the user potentially have to pay much for the network connection. Superdistributing the content is a cost effective way to distribute content. It reduces the need for content servers since the users spread the content between each other. Hence, it is useful to separate the content and the protection method from each other to allow the users to send the content to their buddies. Besides superdistribution of the content, the DRM platform developed supports the sending the DRM client to another user which, again, makes usage more easy.

## 3.2 Audio protection with digital watermarking

The second research problem in this thesis concerned developing audio protection methods that are suitable for mobile devices with limited computational capabilities. In this section, two different methods for using watermarking techniques to protect audio are described. Usually watermarks are used to protect the data by embedding copyright mark or user's fingerprint in the content, but the actual watermark does not prevent the user from playing the content. In these methods, the watermark prevents the user from playing the music without a permit. In Paper III, a fragile watermark is used to protect the data in a robust watermark that is also embedded in the same content. Paper IV presents an audible watermarking method that allows the user to preview the protected content before buying it.

Paper III presents a method for increasing the robustness of a watermark by using a second fragile watermark. The idea of the method is to try to persuade the user not to try to remove the robust watermark by embedding a fragile watermark in the same content, which contains valuable data to the user that she does not want to lose. As an example, a lightweight scrambling scheme for audio is described.

In the protection scheme, two watermarks are embedded in the audio file. First, robust watermark is embedded in the content. The watermark could contain, for example, the fingerprint of the user, the ID of the content owner, or a copyright mark for the DRM player. Second, the audio file is divided into segments that are 0.02–0.04 second long. Finally, the segments are randomly permutated and in each segment the position and the length of the next segment in embedded with the fragile watermarking method. Fig. 5 illustrates how the first four segments have been permuted and the link to the next segment which is embedded in the fragile watermark.

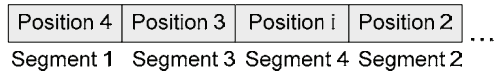


**Fig. 5. Scrambled audio file segments and the embedded data. (III.) Copyright© [2005] IEEE. Reprinted from Proc. Fifth International Conference on Information, Communications and Signal Processing, Audio encryption using fragile watermarking, Löytynoja M, Cvejic N, Lähetkangas E & Seppänen T.**

The fragile watermarking algorithm that is used in the protection scheme embeds the data in the fourth to the sixth LSB layers using a two-step approach. First, the watermark bit is embedded by changing the bit value of the selected LSB layer. Second, the impulse noise caused by the watermark is shaped by changing the bits on lower LSB layers so that the audio sample value is as close to the original as possible. The used LSB layers and audio samples are selected using a pseudorandom sequence given by the watermarking key.

The robust watermark is embedded using a frequency hopping method in a Fourier domain. The embedding is done on a set of fast Fourier transformation (FFT) coefficients that are selected pseudorandomly. Two FFT coefficients are selected from a frequency subband. A mean value of FFT magnitudes in the subband is calculated and the magnitude of the one coefficient selected is set to be *K* decibels above the mean value and the other coefficient is set to be *K* decibels below the mean value. Depending whether bit 1 or 0 is wanted to be embedded, the lower FFT coefficient is set to be above or below the mean value. The same is done for the higher FFT coefficient. Fig. 6 illustrates embedding of bit 1. The value *K* is selected so that it is below the frequency masking threshold of the HAS. The frequency masking threshold is calculated using a model derived from the Psychoacoustic Model 1. When extracting the watermark the higher coefficient is subtracted from the lower one and if the difference is positive bit 1 was extracted, otherwise, the extracted bit was 0.

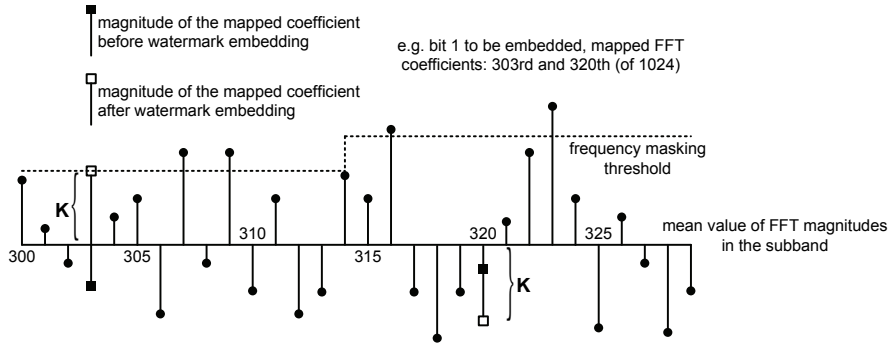


**Fig. 6. Watermark embedding with FH method. (Cvejic & Seppänen 2004.) Reprinted from Signal Processing, 84(1), Cvejic N & Seppänen T, Spread spectrum audio watermarking using frequency hopping and attack characterization, 207-213, Copyright (2004), with permission from Elsevier.**

Paper IV describes a method for protecting audio files using removable watermarks. In the protection scheme, an audible removable watermark is embedded in the host audio. The watermarked data can be any arbitrary data, since the actual data is not used to protect the audio. The audio file can be freely distributed to users who can listen to it as a preview of the actual content. If the user decides to buy the original un-watermarked version of the audio, the watermark can be removed by downloading only the watermarking key.

The watermark is embedded in the host audio in a way that it is clearly audible, but allows the user to have an idea what the un-watermarked content would sound like. For instance, using the algorithm presented in Paper IV produces narrowband noise that changes frequency rapidly. The watermarking method used is format independent, which allows the scheme to be used with lossy compression and supports transcoding of audio into another format.

The watermark is embedded using a modified version of the FH algorithm described above in this section. In this method, the watermark strength $K$ is selected to be *above* the frequency masking threshold of the HAS. This ensures that the watermark is audible. The $K$ must not be too large, however, so that the audio quality is not too poor for the users to be able to preview the content.

The watermark extraction is done in the same way as described above. After the extraction, the watermark is removed before the audio is played to the user. The watermark removal is done by calculating the mean magnitude of the FFT coefficient in a subband, as done in the embedding phase. Next, the magnitude of the two selected FFT coefficients are set to the mean value. This removes the watermark from the content, although with some residual distortions. The experiments show that the remaining distortions are imperceptible to the users and that using lossy compression does not affect the quality either. Additionally, the watermark does not worsen the compression results.

When developing a protection method for mobile devices, it is especially vital to consider the computational complexity of the method. For instance, many watermarking algorithms are computationally heavy, which restricts their use in mobile clients because they are not possible to use in real time. Paper III presented a lightweight content scrambling method that additionally helps to protect the copyright watermark embedded in the same content.

The fragile watermarking method used in Paper III does not survive almost any content processing, thus it is not suitable method, for example, in scenarios where lossy compression is used. For more versatile use, a semifragile watermarking method should be developed which is robust against common

processing operations like lossy compression. The robust watermarking method used in Paper III is robust against tested signal processing attacks with the worst bit error rate (BER) being $2.29 \cdot 10^{-3}$ with 32 kbps MP3 compression attack. It should be noted that this can be significantly reduced by using error correction codes. The audio quality was evaluated with subjective listening tests; the lowest impairment scale value recorded during experiments was −2 and the average mean opinion score (MOS) for the tested audio excerpts was −0.3 using ITU's 5-point impairment scale (see p.36).

To make content superdistribution more useful, the user should be able to preview the content. The protection method described in Paper IV is designed to allow the user to listen the content with degraded quality before making decision about buying it. Similarly, the developed DRM platform supports an option where the user is allowed to get a preview license to the content, although it can be limited to a certain part and restricting the number of plays. Attacks against the watermark were not tested because they do not make the watermark imperceptible. On the contrary, removing attempts of a watermark will further distort the host signal. The audio quality after removing the watermark was evaluated to be excellent. The worst evaluated SDG value was −0.5 when using MP3 compression before the watermark removal. Additionally, the watermark removal is not affected by the use of lossy MP3 compression and the watermark does not worsen the compression results, either.

### 3.3 Counting number of plays on user's terminal

The third research question in this thesis deals with managing and enforcing the rights and restrictions of the content usage. One of the possible restrictions for content consumption that can be set in the license is the number of times the user is allowed to play the content. The easiest way of counting the number of plays is to use counter located on an external server; however, it has the downside that the user must have online access in order to play the content. Placing the counter to the user's terminal allows the malicious user to try attack against it to increase the number of allowed plays. This section focuses on methods that store the counter on the user's terminal and aim to make the attacks against the counter impractical. Paper V utilizes hash chains to count the usages, and Paper VI focuses on watermarking methods to implement the counter.

The counter scheme described in Paper V utilizes hash chains that were originally proposed to be used as a micropayment mechanism. The scheme uses a

cryptographically strong hash function which is iterated on the counter value. In other words, when the user buys a license for *n* playbacks, an initial random value is generated. The hash function is applied *n* times recursively

$$t_{i+1} = h(t_i), \tag{2}$$

for $i = 0,1,\ldots,n-1$, where $t_0$ is the initial value. The license is divided into two parts, one part is digitally signed by the license server and the other is not. The initial value and the number of playbacks *n* are stored in the unsigned part of the license. Additionally, the result of the recursive hash calculation and *n* are stored in the signed part of the license. Fig. 7 shows an example of the XML license file which is used to store the counter data.

```
<License>
<Signature>loDet09DTwestFS</Signature>
<SignedPart>
        <Count count="5">2Q9ksZ4jd8</Count>
        …
</SignedPart>
<UnsignedPart>
        <Counter count="3">uCkZm6vuk=</Counter>
</UnsignedPart>
</License>
```

**Fig. 7. XML license showing the counter fields. (V.) Copyright© [2005] IEEE. Reprinted from Proc. IEEE International Conference on Multimedia & Expo, Hash-based counter scheme for digital rights management, Löytynoja M & Seppänen T.**

When the user plays the content, the value stored in the unsigned part of the license is hashed and the value in the license is replaced with the new one. The counter count is decreased additionally. The number of plays left can be verified by hashing the counter value in the unsigned part of the license recursively until the hash value matches the value stored in the signed part of the license. If the number of hash operations is less than the value stored in the signed part and it optionally matches the value stored in the unsigned part, the counter has not been tampered.

Since the counter value must be updated each time the user plays the content, the counter is vulnerable to replay attack. In the replay attack, the user makes a backup copy of the license file and later, when she has run out of allowed playbacks, replaces the license from the backup. This resets the counter value to

the one in backup. The replay attack can be made harder by storing the counter to an external server, when the user is online, and by checking if the local copy has been tampered. Another possible method is to use file date attributes to check if the license has been modified by the user.

Paper VI presents three different methods for storing the counter in watermarks. The first method embeds the counter watermark in the audio content at specific intervals. The interval is divided into the number of segments, which give the maximum counter value. The number of un-watermarked segments between watermarks gives the counters current value. This is demonstrated in Fig. 8 where the counter's maximum value is five. Each time the user plays the audio file, one segment is watermarked to increase the counter value.

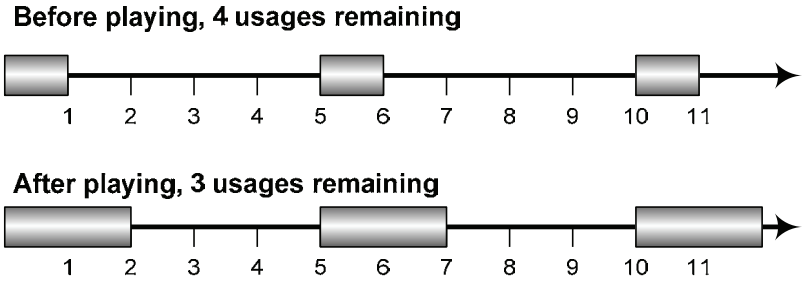


**Fig. 8. Counter based on watermarks embedded at intervals. (VI.) Copyright© 2006 World Academy of Science, Engineering and Technology. Reprinted with permission.**

The interval length can be either fixed or it can depend on the maximum value of the counter. The former alternative has the disadvantage that the counter's maximum value must be known in order to avoid ambiguous counter values. In the latter case, the length of the interval can be much higher if the maximum counter value is high, this can limit how large the counter can be.

The second method uses a single watermark to store the counter value. The counter can be embedded in a one location of the file or it can be repeated multiple times to increase security and robustness. The DRM player must be able to either remove and replace the watermark or the watermark data must be updateable. This might help the malicious user to attack the counter easier. As another disadvantage, the number of updates to the counter could be limited if the watermark cannot be fully removed.

The third counter method embeds multiple watermarks in the same content location. The number of the watermarks gives the counter value, thus the

maximum number of usages must be stored in either watermark or license file. Multiple watermarks can be embedded in the same location if they have a low cross-correlation. The disadvantage of embedding multiple watermarks is that as the number of watermarks increases the quality of the content lowers. This limits how many watermarks can be embedded in the same location.

The watermark based counter methods were demonstrated using a spread spectrum algorithm in which the watermark is summed to the host audio. The watermark is shaped using a filter that is an approximation of the threshold in the quiet curve of the HAS. The algorithm utilizes also the temporal masking properties of the HAS. The experiments revealed that the watermark can be updated when using the second method around 20–25 times before the residual distortions become too large.

The audio quality evaluated with subjective listening tests gave −0.4 as the average MOS for the tested audio samples, and the worst value was −2 using the 5-point impairment scale. The worst signal processing attacks tested against the watermark were time scale modification and MP3 compression with the average BER around $2 \cdot 10^{-2}$, which can be significantly reduced with error correction codes.

The replay attack can be used also against the watermark based counters presented. Storing the counter value in the actual content does make the replay attack more impractical, however. Making backup copies of the content need much more storage space than backing up much smaller license files.

Ideally, the content protection methods should be transparent to the user. Currently, many DRM systems are intrusive and too restrictive, which causes users to prefer pirated content without a cumbersome DRM mechanism. This leads into a situation where only the lawful user is harassed with the protection and not vice versa, which was the original intent. The DRM platform developed tries to distract the user only when necessary. The users are allowed to copy the protected content to other devices unless the license defines otherwise. The user's identity is defined with a public encryption key, which is used to encrypt the licenses. The license negotiation system provides the user easy way to compete different content sellers.

Papers V and VI describe offline methods for enforcing a limited play count set by the license. These methods, of course, limit content to be used in one device at the time, but at least the content can be moved to another device along with the counter data. The methods proposed differ from the ones presented in (Shapiro & Vingralek 2002, Cooper & Martin 2006) in that they do not require

TPM. An online counter would be necessary to allow the user to have the same content on multiple devices at the same time if a limited number of playbacks is required. The watermarking algorithm used to implement the counters described in Paper VI was chosen so that the watermark embedding would be a relatively light operation in which the shaped watermark is summed to the cover media. The watermark extraction can be a somewhat more complex operation as the watermark is only needed to be extracted from one place of the content, as compared with the embedding, which needs to be done over the whole content.

### 3.4 Linking analog music to network services with watermarks

The fourth research problem was concerned with providing the users an easy access to new digital content and services. A digital watermarking method can be made to survive DA/AD transformation. This property is utilized in the method for linking music played through loudspeakers to network based services by recording a sample of the music with a mobile phone and extracting a watermarked link to the service from the recording. In Paper VII, the first version of the method is described. The second version of the method presented in Paper VIII uses better error correction and synchronization and it is much more comprehensively tested than the first version.

The digital watermarking methods can be used to create value added services in addition to the more traditional content protection applications. The audio linking application allows the user to easily access network based services from, for example, music played in the radio. The user needs only record a short sample of the audio with a mobile phone which contains the watermark extraction software. The application extracts a Web link from the recording and directs the user to the network service.

The application embeds two different watermarks in the audio, the first is used to synchronize the extraction of the second watermark which contains the actual data. The synchronization watermark is a spread spectrum watermark that is embedded in frequencies 250–700 Hz. The data watermark is embedded using a modified version of the FH algorithm presented in Section 3.2. The watermarking algorithm uses fixed coefficients on frequencies 2.3–2.7 kHz, i.e. it does not use frequency hopping, additionally it uses three coefficient pairs to increase the watermark power.

Synchronization is done by calculating the cross-correlation between the synchronization watermark and the recording. To increase the correlation peak,

the synchronization watermark is repeated in the content and the correlation is calculated over multiple repetitions.

The mobile phones that were used to test the application could only record audio in mono 8 kHz PCM format. To reduce the effect of downsampling to the watermark extraction, the original audio used was in mono 48 kHz format. Besides downsampling, there are other severe attacks present in the use scenario, DA/AD conversion, background noise, and synchronization noise.

In the first version of the application, 78 bits are embedded in approximately ten seconds of music. The embedded bits consist of 15 bits, that are used to mark the start of the message, and 36 bit payload, which is Hamming (7,4) coded and interleaved to increase the robustness. Watermark extraction is done by first synchronizing the FH algorithm with the synchronization watermark. After that the data watermark is extracted as described in Section 3.2. Next, the message start bit sequence is searched and errors are corrected before the actual data is read.

In the second version of the application, the error correction method was changed to turbo coding with coding rate ½. The data watermark in the second version consists of 35 bits that are interleaved and turbo coded which makes the total length 74 bits. The message start is indicated using a fourth FFT coefficient pair in the FH algorithm by changing its value from 1 to 0 or vice versa.

The experiments showed that the success ratio of the message extraction was 83 % when the subjective quality of the audio SDG was graded on average to be perceptible but not annoying. Malicious attacks against the watermark were not tested because the watermark contains value added data for the user. The key element in successful message extraction is the correct synchronization. The processing power of the mobile phone limits the number of calculations that can be done in message extraction, so a trade-off must be done between reliable extraction and computing time.

Papers VII and VIII present a way to access network based services from music played in the radio, for example. The application does not require the system that is used to play the music to be aware of the watermark. Naturally, the robustness of the watermark and the quality of the audio can be increased if the audio system is optimized not to harm the watermark. One alternative way to implement a similar application would be to identify the music from their acoustic fingerprint (Venkatachalam *et al.* 2004) and then fetch the network link from a database. The benefit of using watermarks to carry the information is that there is no need for content database. Related to other literature about DA/AD

transformation and watermarking, the work done by Steinebach *et al.* (2002) is focused on analyzing and modeling the DA/AD attack. Xiang & Huang (2006) illustrate the main challenge of the DA/AD attack, combination of variety of attacks. Additionally, they point out the weakness of the quantization-based audio watermarking to the DA/AD attack.

# 4 Conclusions

Digital rights management is most probably here to stay; however, as a concept it is somewhat flawed. It is impossible to assume that the protection would not be broken, if the attacker has the decrypting algorithm, protected media, and usually even the keys needed to unprotect the content. This is especially so, since the whole purpose of the content is that it can be consumed by the user. This does not mean that DRM has no merits, however. Sometimes it is enough to prevent copying certain amount of time, for example with computer games most of the sales happen during the first two weeks. If the DRM is able to prevent piracy this time, it has reached its goal. Yet it is important to realize that given enough time all protection methods will break eventually. The business and legal methods, however, can be utilized to protect the content after the technological measures has been broken.

In this thesis, mobile DRM platform and protection methods were studied to create methods for protecting and distributing multimedia content securely in mobile environment. The thesis addressed the problem by focusing on the following sub-problems: A mobile DRM platform was implemented to research what kind of overall architecture is needed to allow easy content distribution to the user in mobile environment. On top of the DRM platform, several content protection and distribution methods were developed to study audio protection methods for mobile devices with limited computing capabilities. Additionally, license management and enforcement methods were investigated. Digital watermarking was applied to provide the users an easy way to access network based services and digital content by embedding a Web link into analog audio content.

Many of the developed methods utilize digital watermarking, and the main purpose of the research was not to create and enhance the watermarking algorithms, but to apply them in realistic scenarios and to show that digital watermarking can be used to implement those scenarios. Applying the algorithms in realistic applications provided new requirements for algorithm design, the most important ones being the computational complexity and watermark synchronization. This resulted in improving the algorithms to fulfill those requirements. Admittedly, the developed methods are far from perfect, each of them have their own weaknesses that can be used to break them. However, they do illustrate that the digital watermarking is a useful tool to create DRM solutions. An especially useful way to use digital watermarks is fingerprinting,

which should be researched in the future to be used with the methods developed. It is an excellent complimentary technology for encryption, since it can be used to track the person responsible for uploading the content into P2P networks after the cryptographic protection has failed. It has a further advantage that the watermark extraction algorithm is not needed in the client, which makes breaking the watermark more difficult. Embedding different watermark into each version of the same content has the drawback that attacker can use collusion attack which is hard to resist.

Currently, the mobile devices begin to be powerful enough to perform more complex calculations fast enough to utilize digital watermarking in different applications. The encryption algorithms are already fast enough to decrypt content in real-time. It is vital to carefully analyze the use scenario where the watermarking is to be used to optimize the limited computing power of the mobile devices into the most efficient use. Additionally, the attacks present in the scenario need to be researched to strengthen the algorithm's robustness against them. This has also the benefit of lowering the complexity of the algorithm if certain attacks are not expected in the use scenario.

In most watermarking applications studied in this thesis, the synchronization to the watermark has been critical for successful data extraction. Often the synchronization is a computationally heavy operation and the precision of its result affects the watermark robustness directly. Oftentimes, the easiest way to attack a watermark is to desynchronize the watermark detector rather than try to remove the watermark from the content. If the detector cannot find the watermark from the content, it does not matter whether or not the watermark survived the attack and is still in the content. Different synchronization methods should be researched more in the future.

The audio protection method presented in Paper IV has already been developed further and an invention disclosure form has been filed for the method developed. Possibly it will further lead to a patent application.

# References


Abbadi IM & Mitchell CJ (2007) Digital rights management using a mobile phone. Proceedings of the Ninth International Conference on Electronic Commerce, Minneapolis, MN: 185–194.

Adams C & Lloyd S (1999) Understanding the public-key infrastructure: Concepts, standards, and deployment considerations. Indianapolis, IN: New Riders Publishing.

Adelsbach A, Rohe M & Sadeghi A-R (2005) Towards multilateral secure digital rights distribution infrastructures. Proceedings of the 5th ACM Workshop on Digital Rights Management, Alexandria, VA: 45–54.

Akalu R & Kundar D (2004) Technological protection measures in the courts. IEEE Signal Processing Magazine 21(2): 109–117.

Arnold M (2002) Subjective and objective quality evaluation of watermarked audio tracks. Proceedings of the Second International Conference on Web Delivering of Music, WEDELMUSIC 2002: 161–167.

Arnold M (2003) Attacks on digital audio watermarks and countermeasures. Proceedings of the Third International Conference on Web Delivering of Music, 2003 WEDELMUSIC: 55–62.

Barni M & Bartolini F (2004) Data hiding for fighting piracy. IEEE Signal Processing Magazine 21(2): 28–39.

Bassia P, Pitas I & Nikolaidis N (2001) Robust audio watermarking in the time domain. IEEE Transactions on Multimedia 3(2): 232–241.

Becker E, Buhse W, Günnewig D & Rump N (eds) (2003). Digital rights management: Technological, economic, legal and political aspects. Berlin, Germany, Springer-Verlag.

Bender W, Gruhl D, Morimoto N & Lu A (1996) Techniques for data hiding. IBM Systems Journal 35(3–4): 313–336.

Bishop M & Frincke DA (2006) Who owns your computer? [digital rights management]. IEEE Security & Privacy 4(2): 61–63.

Borda ME (2005) Digital rights protection – a great challenge of the new millennium. Proceedings of the 7th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services 2005 1: 207–214.

Bormans J, Gelissen J & Perkis A (2003) MPEG-21: The 21st century multimedia framework. IEEE Signal Processing Magazine 20(2): 53–62.

Buhse W & van der Meer J (2007) The open mobile alliance digital rights management [standards in a nutshell]. IEEE Signal Processing Magazine 24(1): 140–143.

Camp L (2003) Access denied [digital rights management]. IEEE Security & Privacy 1(5): 82–85.

Chang F-C, Wu C-L & Hang H-M (2007) A switchable DRM structure for embedded device. Proceedings of the Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2007 2: 37–40.

Chen X & Huang T (2007) Interoperability issues in DRM and DMP solutions. Proceedings of the IEEE International Conference on Multimedia and Expo: 907–910.

Chong DJT & Deng RH (2006) Privacy-enhanced superdistribution of layered content with trusted access control. Proceedings of the ACM Workshop on Digital Rights Management, Alexandria, VA: 37–44.

Cooper A & Martin A (2006) Towards an open, trusted digital rights management platform. Proceedings of the ACM Workshop on Digital Rights Management, Alexandria, VA: 79–88.

Cooper B & Montague P (2005) Translation of rights expressions. Proceedings of the 2005 Australasian Workshop on Grid Computing and E-research, Newcastle, New South Wales, Australia 44: 137–144.

Cox IJ, Miller ML & Bloom JA (2002) Digital watermarking. San Francisco, CA: Morgan Kauffman Publishers.

Creusere CD, Kallakuri KD & Vanam R (2008) An objective metric of human subjective audio quality optimized for a wide range of audio fidelities. IEEE Transactions on Audio, Speech, and Language Processing 16(1): 129–136.

Cvejic N (2004) Algorithms for audio watermarking and steganography. PhD thesis. University of Oulu, Department of Electrical and Information Engineering.

Cvejic N & Seppänen T (2004) Spread spectrum audio watermarking using frequency hopping and attack characterization. Signal Processing 84(1): 207–213.

Cvejic N & Seppänen T (2005) Increasing robustness of LSB audio steganography by reduced distortion LSB coding. Journal of Universal Computer Science 11(1): 56–65.

Cvejic N & Seppänen T (eds) (2008). Digital audio watermarking techniques and technologies: Applications and benchmarks. Hershey, PA, Information Science Reference.

Diffie W & Hellman M (1976) New directions in cryptography. IEEE Transactions on Information Theory 22(6): 644–654.

Erickson JS & Mulligan DK (2004) The technical and legal dangers of code-based fair use enforcement. Proceedings of the IEEE 92(6): 985–996.

Erküçük S, Krishnan S & Zeytinoğlu M (2006) A robust audio watermark representation based on linear chirps. IEEE Transactions on Multimedia 8(5): 925–936.

Esmaili S, Krishnan S & Raahemifar K (2003) Audio watermarking time-frequency characteristics. Canadian Journal of Electrical and Computer Engineering 28(2): 57–61.

Feigenbaum J, Freedman MJ, Sander T & Shostack A (2002) Privacy engineering for digital rights management systems. Revised Papers from the ACM CCS-8 Workshop on Security and Privacy in Digital Rights Management: 76–105.

Felten EW & Halderman JA (2006) Digital rights management, spyware, and security. IEEE Security & Privacy 4(1): 18–23.

Feng J-B, Lin I-C, Tsai C-S & Chu Y-P (2006) Reversible watermarking: Current status and key issues. International Journal of Network Security 2(3): 161–171.

Fetscherin M & Schmid M (2003a) The application of digital rights management systems in the music industry – an empirical investigation. Proceedings of the Third International Conference on Web Delivering of Music, 2003 WEDELMUSIC: 115–121.

Fetscherin M & Schmid M (2003b) Comparing the usage of digital rights management systems in the music, film, and print industry. Proceedings of the 5th International Conference on Electronic Commerce, Pittsburgh, PA: 316–325.

Fridrich J (1999) Applications of data hiding in digital images (tutorial slides). Proceedings of the Fifth International Symposium on Signal Processing and Its Applications, ISSPA '99 1: 9.

Geer D (2004) Digital rights technology sparks interoperability concerns. Computer 37(12): 20–22.

Guth S, Neumann G & Strembeck M (2003) Experiences with the enforcement of access rights extracted from ODRL-based digital contracts. Proceedings of the 3rd ACM Workshop on Digital Rights Management, Washington, DC: 90–102.

Halpern JY & Weissman V (2008) A formal foundation for XrML. Journal of the ACM 55(1): 1–42.

Hartung F & Ramme F (2000) Digital rights management and watermarking of multimedia content for m-commerce applications. IEEE Communications Magazine 38(11): 78–84.

Heileman GL & Jamkhedkar PA (2005) DRM interoperability analysis from the perspective of a layered framework. Proceedings of the 5th ACM Workshop on Digital Rights Management, Alexandria, VA: 17–26.

Hembrooke EF (1961) Muzak Corporation, assignee. Identification of sound and like signals. US patent 3,004,104.

Hietanen H, Huttunen A & Kokkinen H (2008) Criminal friends of entertainment: Analysing results from recent peer-to-peer surveys. SCRIPT-ed 5(1): 31–49.

Holt L, Maufe BG & Wiener A (1988) Emi Plc Thorn, assignee. Encoded marking of a recording signal. UK patent GB2196167.

Hu Y & Jeon B (2006) Reversible visible watermarking and lossless recovery of original images. IEEE Transactions on Circuits and Systems for Video Technology 16(11): 1423–1429.

Hu Y, Kwong S & Huang J (2006) An algorithm for removable visible watermarking. IEEE Transactions on Circuits and Systems for Video Technology 16(1): 129–133.

Huang C-H & Wu J-L (2004) Attacking visible watermarking schemes. IEEE Transactions on Multimedia 6(1): 16–30.

Iannella R (2001) Digital rights management (DRM) architectures. D-Lib Magazine 7(6), DOI: 10.1045/june2001-iannella

ISO International Organization for Standardization (1993) Information technology – coding of moving pictures and associated audio for digital storage up to about 1.5 mbits/s – part 3: Audio. ISO/IEC IS 11172-3.

ITU International Telecommunication Union (1997) Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. ITU-R BS.1116-1.

ITU International Telecommunication Union (2001) Methods for objective measurement of perceived audio quality. ITU-R BS.1387-1.

ITU International Telecommunication Union (2003) Subjective assessment of sound quality. ITU-R BS.1284-1.

Jamkhedkar PA & Heileman GL (2004) DRM as a layered system. Proceedings of the 4th ACM Workshop on Digital Rights Management, Washington, DC: 11–21.

Jamkhedkar PA, Heileman GL & Martínez-Ortiz I (2006) The problem with rights expression languages. Proceedings of the ACM Workshop on Digital Rights Management, Alexandria, VA: 59–68.

Jamkhedkar PA, Heileman GL & Martínez-Ortiz I (2007) Middleware services for DRM. Proceedings of the 2nd International Conference on Communication Systems Software and Middleware, COMSWARE 2007: 1–8.

Jeong Y, Park J, Kim J & Yoon K (2007) DRM content adaptation scheme between different DRM systems for seamless content service. Proceedings of the IEEE International Conference on Multimedia and Expo: 867–870.

Jonker W & Linnartz JP (2004) Digital rights management in consumer electronics products. IEEE Signal Processing Magazine 21(2): 82–91.

Kim HJ & Choi YH (2003) A novel echo-hiding scheme with backward and forward kernels. IEEE Transactions on Circuits and Systems for Video Technology 13(8): 885–889.

Kim K & Hong J (2004) MPEG4 IPMP authoring system for protection of object based contents. Proceedings of the 6th International Conference on Advanced Communication Technology 1: 499–503.

Kirovski D & Malvar H (2001) Robust spread-spectrum audio watermarking. Proceedings on the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01) 3: 1345–1348.

Ko B-S, Nishimura R & Suzuki Y (2005) Time-spread echo method for digital audio watermarking. IEEE Transactions on Multimedia 7(2): 212–221.

Koenen RH, Lacy J, Mackay M & Mitchell S (2004) The long march to interoperable digital rights management. Proceedings of the IEEE 92(6): 883–897.

Komatsu N & Tominaga H (1989) Authentication system using concealed images in telematics. Memoirs of the school of science & engineering, Waseda University 52: 45–60.

Kravitz DW & Messerges TS (2005) Achieving media portability through local content translation and end-to-end rights management. Proceedings of the 5th ACM Workshop on Digital Rights Management, Alexandria, VA: 27–36.

Kundur D & Hatzinakos D (1998) Improved robust watermarking through attack characterization. Optics Express 3(12): 485–490.

Kwok SH (2002) Digital rights management for the online music business. ACM SIGecom Exchanges 3(3): 17–24.

Küpper A, Ahrens S, Hess T & Freese B (2007) Superdistribution of digital content – overview, opportunities and challenges. Proceedings of the ITI 5th International Conference on Information and Communications Technology, ICICT 2007: 173–179.

Lacy J, Rump N, Shamoon T & Kudumakis P (1999) MPEG-4 intellectual property management & protection. Proceedings of the 17th Conference of Audio Engineering Society.

Lang A, Dittmann J, Spring R & Vielhauer C (2005) Audio watermark attacks: From single to profile attacks. Proceedings of the 7th Workshop on Multimedia and Security, New York, NY: 39–50.

Larbi SD & Jaidane-Saidane M (2005) Audio watermarking: A way to stationnarize audio signals. IEEE Transactions on Signal Processing 53(2): 816–823.

Lemma AN, Aprea J, Oomen W & van de Kerkhof L (2003) A temporal domain audio watermarking technique. IEEE Transactions on Signal Processing 51(4): 1088–1097.

Lesk M (2003) The good, the bad, and the ugly: What might change if we had good DRM. IEEE Security & Privacy 1(3): 63–66.

Li W, Xue X & Lu P (2005) Robust audio watermarking based on rhythm region detection. Electronics Letters 41(4): 218–219.

Li W, Xue X & Lu P (2006) Localized audio watermarking technique robust against time-scale modification. IEEE Transactions on Multimedia 8(1): 60–69.

Lie W-N & Chang L-C (2006) Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification. IEEE Transactions on Multimedia 8(1): 46–59.

Lin Y & Abdulla WH (2008) Perceptual evaluation of audio watermarking using objective quality measures. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2008: 1745–1748.

Liu KJR, Trappe W, Wang ZJ, Wu M & Zhao H (2005) Multimedia fingerprinting forensics for traitor tracing. New York, NY: Hindawi Publishing Corporation.

Liu Q, Safavi-Naini R & Sheppard NP (2003) Digital rights management for content distribution. Proceedings of the Australasian information security workshop conference on ACSW frontiers, Adelaide, Australia 21: 49–58.

Liu Y-W & Smith JO (2004) Multiple watermarking: Is power sharing better than time sharing? Proceedings of the IEEE International Conference on Multimedia and Expo, ICME '04 3: 1939–1942 Vol.3.

Liu Z & Inoue A (2003) Audio watermarking techniques using sinusoidal patterns based on pseudorandom sequences. IEEE Transactions on Circuits and Systems for Video Technology 13(8): 801–812.

Lou X, Hwang K & Zhou R (2007) Integrated copyright protection in peer-to-peer networks. Proceedings of the 27th International Conference on Distributed Computing Systems Workshops, ICDCSW '07: 28–28.

Merabti M & Llewellyn-Jones D (2006) Digital rights management in ubiquitous computing. IEEE Multimedia 13(2): 32–42.

Michiels S, Verslype K, Joosen W & De Decker B (2005) Towards a software architecture for DRM. Proceedings of the 5th ACM Workshop on Digital Rights Management, Alexandria, VA, USA: 65–74.

Mintzer F & Braudaway GW (1999) If one watermark is good, are more better? Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '99 4: 2067–2069.

Mintzer F, Lotspiech J & Morimoto N (1997) Safeguarding digital library contents and users: Digital watermarking. D-Lib Magazine 3(12).

Morin J-H & Pawlak M (2008) Exception-aware digital rights management architecture experimentation. Proceedings of the International Conference on Information Security and Assurance, ISA 2008: 518–526.

Mundt T (2005) Location dependent digital rights management. Proceedings of the 10th IEEE Symposium on Computers and Communications, ISCC 2005: 617–622.

Nam D-W, Jeong Y, Park J & Yoon K-S (2007a) DRM content adaptation between different DRM systems for seamless content service. Proceedings of the IEEE International Symposium on Consumer Electronics, ISCE 2007: 1–4.

Nam D-W, Lee J-S, Kim J-H & Yoon K-S (2007b) Interlock system for DRM interoperability of streaming contents. Proceedings of the IEEE International Symposium on Consumer Electronics, ISCE 2007: 1–4.

NIST National Institute of Standards and Technology (2001) Advanced encryption standard (AES). Federal Information Processing Standards Publication 197 (FIPS 197).

Owens R & Akalu R (2004) Legal policy and digital rights management. Proceedings of the IEEE 92(6): 997–1003.

Pan D (1995) A tutorial on MPEG/audio compression. IEEE Multimedia 2(2): 60–74.

Pang H & Wu Y (2005) Evaluation of MPEG-4 IPMP extension. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '05) 2: 1161–1164.

Pearson S (ed) (2003). Trusted computing platforms. Upper Saddle River, NJ, Prentice Hall.

Petitcolas FAP, Anderson RJ & Kuhn MG (1999) Information hiding-a survey. Proceedings of the IEEE 87(7): 1062–1078.

Polo J, Prados J & Delgado J (2004) Interoperability between ODRL and MPEG-21 REL. Proceedings of the First International ODRL Workshop, Vienna, Austria: 65–76.

Robert A & Picard J (2005) On the use of masking models for image and audio watermarking. IEEE Transactions on Multimedia 7(4): 727–739.

Rosenblatt B, Trippe B & Mooney S (2002) Digital rights management: Business and technology. New York, NY: M&T Books.

de Rosnay MD (2002) Digital rights management systems and European law: Between copyright protection and access control. Proceedings of the Second International Conference on Web Delivering of Music, WEDELMUSIC 2002: 117–124.

Rump N (2004) Can digital rights management be standardized? IEEE Signal Processing Magazine 21(2): 63–70.

Safavi-Naini R, Sheppard NP & Uehara T (2004) Import/export in digital rights management. Proceedings of the 4th ACM Workshop on Digital Rights Management, Washington DC, USA: 99–110.

Saied-Bouajina S, Larbi S, Hamza R, Slama LB & Jidane-Saidane M (2004) An error correction strategy for digital audio watermarking scheme. Proceedings of the First International Symposium on Control, Communications and Signal Processing: 739–742.

Schneier B (1996) Applied cryptography. New York, NY: John Wiley & Sons.

Senoh T, Ueno T, Kogure T, Shen S, Ji N, Liu J, Huang Z & Schultz CA (2004) DRM renewability & interoperability. Proceedings of the First IEEE Consumer Communications and Networking Conference, CCNC 2004: 424–429.

Seok JW & Hong JW (2001) Audio watermarking for copyright protection of digital audio data. Electronics Letters 37(1): 60–61.

Serrão C, Dias M & Delgado J (2007) Key management in open DRM platforms. Proceedings of the Third International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution, AXMEDIS '07: 47–54.

Shapiro W & Vingralek R (2002) How to manage persistent state in DRM systems. Revised Papers from the ACM CCS-8 Workshop on Security and Privacy in Digital Rights Management: 176–191.

Sharma G & Coumou DJ (2006) Watermark synchronization: Perspectives and a new paradigm. Proceedings of the 40th Annual Conference on Information Sciences and Systems: 1182–1187.

Sheppard NP (2007) On implementing MPEG-21 intellectual property management and protection. Proceedings of the 2007 ACM Workshop on Digital Rights Management, Alexandria, Virginia, USA: 10–22.

Sheppard NP, Safavi-Naini R & Ogunbona P (2001) On multiple watermarking. Proceedings of the 2001 Workshop on Multimedia and Security: New Challenges, Ottawa, Ontario, Canada: 3–6.

Siwek SE (2007) The true cost of sound recording piracy to the U.S. economy. Institute for Policy Innovation.

Soriano M, Flake S, Tacken J, Bormann F & Tomas J (2005) Mobile digital rights management: Security requirements and copy detection mechanisms. Proceedings of the Sixteenth International Workshop on Database and Expert Systems Applications: 251–256.

Steinebach M, Lang A, Dittmann J & Neubauer C (2002) Audio watermarking quality evaluation: Robustness to DA/AD processes. Proceedings of the International Conference on Information Technology: Coding and Computing: 100–103.

Steinebach M, Petitcolas FAP, Raynal F, Dittmann J, Fontaine C, Seibel S, Fates N & Ferri LC (2001) Stirmark benchmark: Audio watermarking attacks. Proceedings of the International Conference on Information Technology: Coding and Computing: 49–54.

Stini M, Mauve M & Fitzek FHP (2006) Digital ownership: From content consumers to owners and traders. IEEE Multimedia 13(4): 1–6.

Subramanya SR & Yi BK (2006) Digital rights management. IEEE Potentials 25(2): 31–34.

Sung J-Y, Jeong J-Y & Yoon K-S (2006) DRM enabled P2P architecture. Proceedings of the 8th International Conference of Advanced Communication Technology, ICACT 2006 1: 487–490.

Surminen MJ, Sheppard NP & Safavi-Naini R (2007) Location-based DRM using WiFi access points. Proceedings of the International Symposium on Communications and Information Technologies, ISCIT '07: 882–886.

Szepanski W (1979) A signal theoretic method for creating forgery-proof documents for automatic verification. Proceedings of the Carnahan Conference on Crime Countermeasures, Lexington, KY, USA: 101–109.

Takahashi A, Nishimura R & Suzuki Y (2005) Multiple watermarks for stereo audio signals using phase-modulation techniques. IEEE Transactions on Signal Processing 53(2): 806–815.

Tu R & Zhao J (2003) A novel semi-fragile audio watermarking scheme. Proceedings of the 2nd IEEE Internatioal Workshop on Haptic, Audio and Visual Environments and Their Applications, HAVE 2003: 89–94.

Wang X (2004) MPEG-21 rights expression language: Enabling interoperable digital rights management. IEEE Multimedia 11(4): 84–87.

Wang X, DeMartini T, Wragg B, Paramasivam M & Barlas C (2005) The MPEG-21 rights expression language and rights data dictionary. IEEE Transactions on Multimedia 7(3): 408–417.

Wang X, Lao G, DeMartini T, Reddy H, Nguyen M & Valenzuela E (2002) XrML – extensible rights markup language. Proceedings of the 2002 ACM Workshop on XML Security, Fairfax, VA: 71–79.

Wang X, Qi W & Niu P (2007) A new adaptive digital audio watermarking based on support vector regression. IEEE Transactions on Audio, Speech, and Language Processing 15(8): 2270–2277.

Wang XY & Zhao H (2006) A novel synchronization invariant audio watermarking scheme based on dwt and dct. IEEE Transactions on Signal Processing 54(12): 4835–4840.

Venkatachalam V, Cazzanti L, Dhillon N & Wells M (2004) Automatic identification of sound recordings. IEEE Signal Processing Magazine 21(2): 92–99.

Wikipedia (n.d.-a) Digital rights management. [cited 2008 September 17th]; Available from: http://en.wikipedia.org/wiki/Digital_rights_management#Controversy.

Wikipedia (n.d.-b) Security of advanced access content system. [cited 2008 September 17th]; Available from: http://en.wikipedia.org/wiki/Security_of_AACS.

Wu S, Huang J, Huang D & Shi YQ (2005) Efficiently self-synchronized audio watermarking for assured audio data transmission. IEEE Transactions on Broadcasting 51(1): 69–76.

Xiang S & Huang J (2006) Analysis of D/A and A/D conversions in quantization-based audio watermarking. International Journal of Network Security 3(3): 230–238.

Xiang S & Huang J (2007) Histogram-based audio watermarking against time-scale modification and cropping attacks. IEEE Transactions on Multimedia 9(7): 1357–1372.

Yeo I-K & Kim HJ (2003) Modified patchwork algorithm: A novel audio watermarking scheme. IEEE Transactions on Speech and Audio Processing 11(4): 381–386.

Zaidi A, Boyer R & Duhamel P (2006) Audio watermarking under desynchronization and additive noise attacks. IEEE Transactions on Signal Processing 54(2): 570–584.

Zheng Y, He D, Wang H & Tang X (2005) Secure DRM scheme for future mobile networks based on trusted mobile platform. Proceedings of the 2005 International Conference on Wireless Communications, Networking and Mobile Computing 2: 1164–1167.

Ziolkowski B & Stoklosa J (2007) Mobile agent-based digital rights management scheme. Proceedings of the 6th International Conference on Computer Information Systems and Industrial Management Applications, CISIM '07: 213–218.

Zwicker E & Fastl H (1999) Psychoacoustics: Facts and models. Berlin, Germany: Springer-Verlag.

# Original publications

I Löytynoja M, Seppänen T & Cvejic N (2003) Experimental DRM architecture using watermarking and PKI. Proc. 1st International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet, Helsinki, Finland: 47–52.

II Löytynoja M, Koskela T, Brockman M & Seppänen T (2006) Mobile DRM-enabled multimedia platform for peer-to-peer applications. Proc. IEEE International Symposium on Multimedia, San Diego, CA, USA: 139–144.

III Löytynoja M, Cvejic N, Lähetkangas E & Seppänen T (2005) Audio encryption using fragile watermarking. Proc. Fifth International Conference on Information, Communications and Signal Processing, Bangkok, Thailand: 881–885.

IV Löytynoja M, Cvejic N & Seppänen T (2007) Audio protection with removable watermarking. Proc. Sixth International Conference on Information, Communications and Signal Processing, Singapore: 1–4.

V Löytynoja M & Seppänen T (2005) Hash-based counter scheme for digital rights management. Proc. IEEE International Conference on Multimedia & Expo, Amsterdam, Netherlands: 121–124.

VI Löytynoja M, Cvejic N & Seppänen T (2006) Watermark-based counter for restricting digital audio consumption. International Journal of Signal Processing, 3(1): 17–23.

VII Löytynoja M, Cvejic N, Keskinarkaus A, Lähetkangas E & Seppänen T (2006) Mobile commerce from watermarked broadcast audio. Proc. IEEE International Conference on Consumer Electronics, Las Vegas, NV, USA: 5–6.

VIII Löytynoja M, Keskinarkaus A, Cvejic N & Seppänen T (2008) Watermark-enabled value added services to broadcast audio. Proc. Second IEEE Conference on Digital Ecosystems and Technologies, Phitsanulok, Thailand: 388–396.



Original publications are not included in the electronic version of the dissertation.

293. Koski, Anna (2008) Applicability of crude tall oil for wood protection

294. Gore, Amol (2008) Exploring the competitive advantage through ERP systems. From implementation to applications in agile networks

295. Kirillin, Mikhail (2008) Optical coherence tomography of strongly scattering media

296. Tölli, Antti (2008) Resource management in cooperative MIMO-OFDM cellular systems

297. Karkkila, Harri (2008) Consumer pre-purchase decision taxonomy

298. Rabbachin, Alberto (2008) Low complexity UWB receivers with ranging capabilities

299. Kunnari, Esa (2008) Multirate MC-CDMA. Performance analysis in stochastically modeled correlated fading channels, with an application to OFDM-UWB

300. Särkkä, Jussi (2008) A novel method for hazard rate estimates of the second level interconnections in infrastructure electronics

301. Mäkelä, Juha-Pekka (2008) Effects of handoff algorithms on the performance of multimedia wireless networks

302. Teräs, Jukka (2008) Regional science-based clusters. A case study of three European concentrations

303. Lahti, Markku (2008) Gravure offset printing for fabrication of electronic devices and integrated components in LTCC modules

304. Popov, Alexey (2008) TiO2 nanoparticles as UV protectors in skin

305. Illikainen, Mirja (2008) Mechanisms of thermomechanical pulp refining

306. Borkowski, Maciej (2008) Digital ?-? Modulation. Variable modulus and tonal behaviour in a fixed-point digital environment

307. Kuismanen, Kimmo (2008) Climate-conscious architecture—design and wind testing method for climates in change

308. Kangasvieri, Tero (2008) Surface-mountable LTCC-SiP module approach for reliable RF and millimetre-wave packaging

309. Metsärinta, Maija-Leena (2008) Sinkkivälkkeen leijukerrospasutuksen stabiilisuus

310. Prokkola, Jarmo (2008) Enhancing the performance of ad hoc networking by lower layer design

Book orders:
OULU UNIVERSITY PRESS
P.O. Box 8200, FI-90014
University of Oulu, Finland

Distributed by
OULU UNIVERSITY LIBRARY
P.O. Box 7500, FI-90014
University of Oulu, Finland

UNIVERSITY OF OULU P.O.B. 7500 FI-90014 UNIVERSITY OF OULU FINLAND

# ACTA UNIVERSITATIS OULUENSIS



ISBN 978-951-42-8936-1 (Paperback)
ISBN 978-951-42-8937-8 (PDF)
ISSN 0355-3213 (Print)
ISSN 1796-2226 (Online)

UNIVERSITY of OULU
OULUN YLIOPISTO