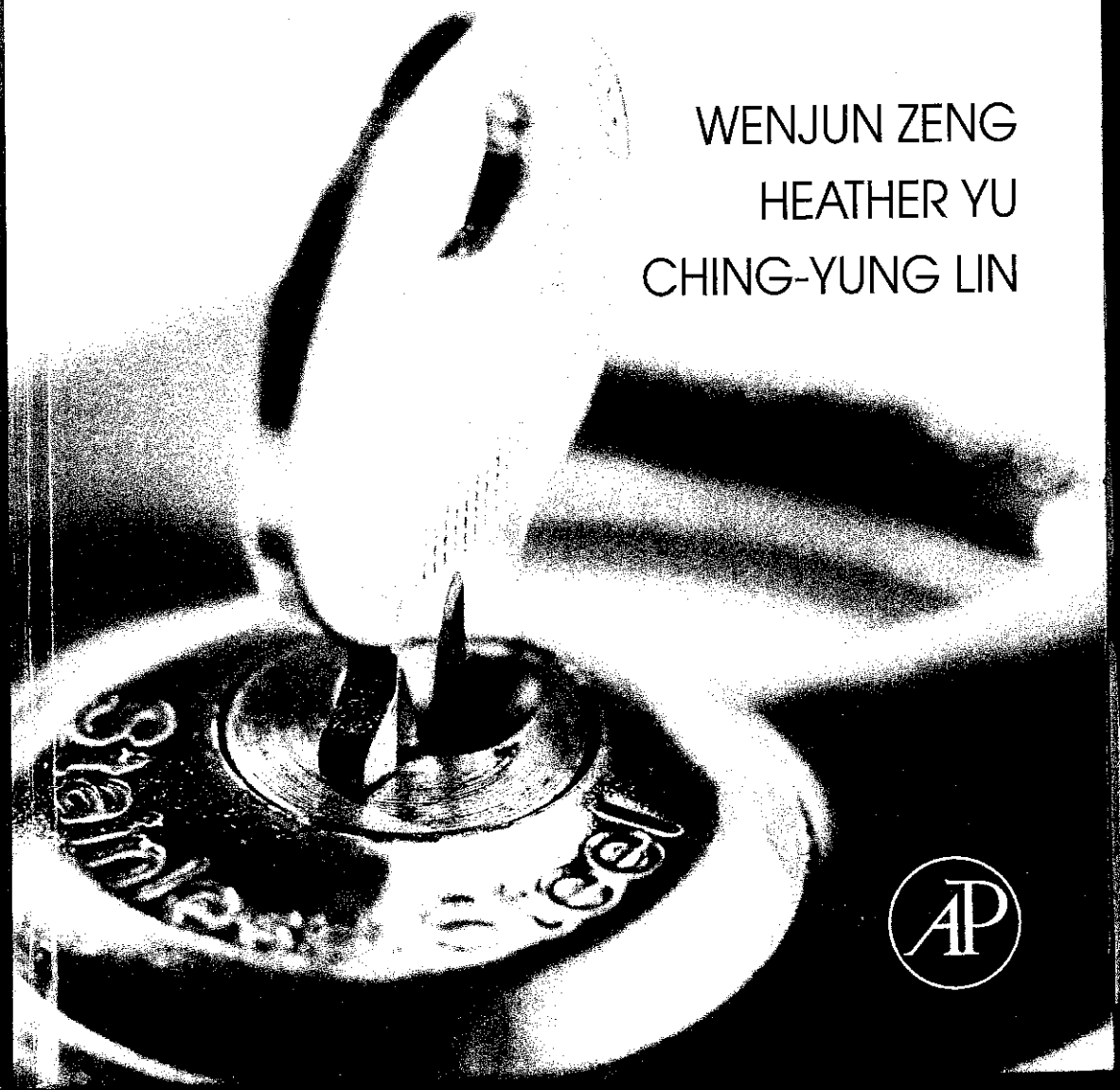
Multimedia Security Technologies for Digital Rights Management

WENJUN ZENG
HEATHER YU
CHING-YUNG LIN



Academic Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data
Multimedia security technologies for digital rights management/edited by Wenjun Zeng, Heather Yu, and Ching-Yung Lin.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-12-369476-8 (casebound : alk. paper)
ISBN-10: 0-12-369476-0 (casebound : alk. paper) 1. Computer security. 2. Multimedia systems–Security measures. 3. Intellectual property. I. Zeng, Wenjun, 1967- II. Yu, Hong Heather, 1967- III. Lin, Ching-Yung.

QA76.9.A25M875 2006
005.8–dc22
2006003179

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN 13: 978-0-12-369476-8
ISBN 10: 0-12-369476-0

For information on all Academic Press publications
visit our Web site at www.books.elsevier.com

Printed in the United States of America
06 07 08 09 10 9 8 7 6 5 4 3 2 1

Table of C

Introduction—Digital Rights Management

Scott Moskowitz

1.1 PROPERTY AND VALUE

Real property is familiar to most people. We live in houses, work in offices, shop at retailers, and enjoy ball games at stadiums. In contrast with "personality," which includes personal effects and intellectual property, real estate derives from *realty*—historically, land and all things permanently attached. Rights, whether for real property or intellectual property, have communal roots. Security, however, is a term with very subjective meaning. Simply "feeling secure" is not necessarily equivalent with the expectations or actual protections provided. Securing real property can mean locking a door or, for the significantly more paranoid, deploying tanks on one's lawn. Although it can be argued that intellectual property is related to real property, there are inherent and significant differences—the obvious one being that intellectual property is not physical property. The most controversial aspect of intellectual property is the ease at which it can be and is shared. Divergent viewpoints on this issue exist. At the extremes, "information is free," while others assert theft. We will leave the ability to define "piracy" to economists, lobbyists, policymakers, and even jurists with such interests. Clearly, we need to consider the law and the cost of copy protection when making technical decisions about designing the appropriate system. A particular set of problems will need definitions in order for agreement on any "secure" solutions. For this reason, any resource on "Digital Rights Management" (DRM) should include appropriate context. While other chapters of this book focus on technology topics and the development of the burgeoning market for DRM products and services,

this chapter covers a number of topics identifying the importance of rights management technologies.

1.2 "ORIGINAL WORK"

It is prudent to provide a cursory outline of copyrights, not in the interests of providing any form of legal advice, but to delineate the impact of how copyright protection has evolved with respect to U.S. copyright law.[1] Copyright is established in the U.S. Constitution. The single occurrence of the word "right" in the Constitution appears in Article 1, Section 8, Clause 8: "[t]o promote the Progress of Science and useful Arts, by securing for limited times to authors and inventors the exclusive *right* to their respective writings and discoveries." As with all U.S. laws, the U.S. Congress first enacts legislation, while the courts provide judicial oversight and interpretation of law. Over time, legislation has been adopted making copyright more consistent with advances in the technology landscape. Lobbying efforts by a variety of stakeholders have provided additional impetus for change for economic reasons. Litigating "copyright infringements" represent additional efforts at defining copyright and its associated protections. However, when one has a copyright, what exactly does that mean? Essentially, a copyright is a form of contract between the creator of the original work and the public. While based on the recognition of property rights, in general, the creator agrees to make his work publicly available in consideration of legal recognition under the law. The Constitution promulgated copyright in the interests of promoting science and the arts for the benefit of society. Subsequent changes, challenges, and context have become arguably more public with the huge success of the Internet and networking technologies in general.

To be a bit more specific, a "work," the copyrighted value to be protected, is "created" when it is fixed in a copy or phonorecord for the first time: where a work has been prepared over a period of time, the portion of it that has been fixed at any particular time constitutes the work as of that time, and where the work has been prepared in different versions, each version constitutes a separate work. A "derivative work" is a work based upon one or more pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of editorial revisions, annotations, elaborations, or other modifications which, as a whole, represent an original work of authorship is a derivative work. As electronics and digital editing software become the inexpensive tools of the

[1] For international copyright issues, one helpful resource is http://caselaw.lp.findlaw.com/data/constitution/article01/39.html.

Information Age, copyright is thought to need additional protections. We do not argue the merits of such a belief, but provide the following milestones as to how we got here from there.

1.3 LOOKING BACK AT THE COPYRIGHT ACT OF 1976

Including a list of burgeoning "copyright protection" software companies, the National Information Infrastructure Copyright Act of 1995 made recommendations to the Copyright Act of 1976 and addressed the potential problems with open networks such as the "Internet." It is a fairly interesting point to start a historical timeline from which rights management technologies have evolved as several of the companies listed in that report made subsequent impacts in the field. For our purposes, it is not necessary to interpret the large body of legal arguments, but it is helpful to provide what limits have been argued and how far the perception of technology impacts DRM. After all, the copyright holder is not the only party with legal rights. While copyright previously concerned "sweat of the brow," what is referred to as "Feist," a modicum of creativity has become the more stringent standard for establishing copyright. An early case, *Lotus Corporation v. Borland* is somewhat emblematic of the early fights over copyright protection of intellectual property.

> In Feist [Feist Publications, Inc. v. Rural Telephone Serv. Co., 499 U.S. 340 (1991)], the court explained:
> The primary objective of copyright is not to reward the labor of authors, but to promote the Progress of Science and useful Arts. To this end, copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.
>
> Feist, 499 U.S. at 349-50. We do not think that the court's statement that "copyright assures authors the right to their original expression" indicates that all expression is necessarily copyrightable. While original expression is necessary for copyright protection, we do not think that it is alone sufficient. Courts must still inquire whether original expression falls within one of the categories foreclosed from copyright protection by 102(b) [1].

Section 107 of the Copyright Act of 1976 provides additional guidance for the wide range of stakeholders who may need to access or manipulate copyrighted works. Perhaps inevitably, reverse engineering and related attempts at circumventing "security" increase the perception that copies of the original work may require layered security and additional legal protections. The least understood aspect of copyright and its place "to promote the Progress of Science and useful Arts" regards "fair use." Bounded by several factors, the relative weights are not provided by the Copyright Act of 1976, and fair use may indeed be the one legal issue that presents the most difficult challenges in engineering solutions to piracy.

> Four factors must be considered: (1) the purpose and character of the use, including whether such use is of a commercial nature or is for non-profit educational purposes; (2) the nature of the work; (3) the amount and the substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use on the market value of the copied work [2].

The one case at the heart of the most extreme debates in copyright circles may be *Sony Corporation v. Universal City Studios* (1984), concerning the sale of videocassette recorders (VCRs). The U.S. Supreme Court ruled that "[b]ecause recorders were 'widely used for legitimate, unobjectionable purposes,' the recording did not constitute direct infringement of the studio's copyrights .... Absent such direct infringement, there could be no contributory infringement by *Sony* [3]." The key factor being that there was value in personal recording. While citing the concept of fair use, which protects consumers from *some forms* of copyright infringement, the debate did not end with this ruling. Indeed, the concept of fair use has been extended to areas not previously anticipated, including reverse engineering of copyrighted software.

Additionally, the Copyright Act of 1976 laid several other "foundations," though they are still unsettled in the minds of the stakeholders involved. Besides extending the length of copyright protection, library photocopying was changed to make possible preservation and inter-library loans without permission. Section 107 is at the heart of the types of issues for evaluation of DRM system design, even if less than all stakeholders' rights are considered. Fair use is a doctrine that permits courts to avoid rigid application of the copyright statute when to do otherwise would stifle the very creativity that copyright law is designed to foster. One author addresses this notion of relativity in the early days of the Internet Age.

> The doctrine of fair use recognizes that the exclusive rights inherent in a copyright are not absolute, and that non-holders of the copyright are entitled to make use of a copyrighted work that technically would otherwise infringe upon one or more of the exclusive rights. Although fair use originated 'for purposes such as criticism, comment, news reporting, teaching, ... scholarship, or research,' it also applies in other areas, as some of the examples below illustrate. However, courts seem more willing to accept an assertion of fair use when the use falls into one of the above categories. Perhaps more than any other area of copyright, fair use is a highly fact-specific determination. Copyright Office document FL102 puts it this way: 'The distinction between "fair use" and infringement may be unclear and not easily defined. There is no specific number of words, lines, or notes that may safely be taken without permission. Acknowledging the source of the copyrighted material does not substitute for obtaining permission.' The document then quotes from the 1961 Report of the Register of Copyrights on the General Revision of the U.S. Copyright Law, providing the following examples of activities that courts have held to be fair use:—Quotation of excerpts in a review or criticism for purposes of illustration or

comment;—Quotati
tration or clarificati
content of the work
in a news report;—
a damaged copy;—
illustrate a lesson;—
reports;—Incidental
located in the scene

Several other mor
provide a broader co

Digital Millennium
is the provision,
tion of access rest
the copyright ow
in place to prote
However, it is sti
sures can be circu
computer, as in th
action that is inhe
ian of Congress c
Congress.

Digital Theft Deter
Congress increase
from that of $500 t
from $100,000 to

Librarian of Cong
of Congress issue
Circumvention Pr
two exemptions in
by filtering softw
programs and dat
permit access bec
ommendation can

Dmitri Skylyarov A
grammer for Elco
Reader DRM. Alth
tinued with the pr
the DMCA. As on
observers viewed
could be pushed i
"not guilty" in late

comment;—Quotation of short passages in a scholarly or technical work for illustration or clarification of the author's observations;—Use in a parody of some of the content of the work parodied;—Summary of an address or article with brief quotations, in a news report;—Reproduction by a library of a portion of a work to replace part of a damaged copy;—Reproduction by a teacher or student of a small part of a work to illustrate a lesson;—Reproduction of a work in legislative or judicial proceedings or reports;—Incidental and fortuitous reproduction in a newsreel or broadcast, of a work located in the scene of an event being reported [4].

Several other more recent legal and legislative actions should be mentioned to provide a broader consideration of what the fuss is really all about.

Digital Millennium Copyright Act, the "DMCA" (1998). Key among its impact is the provision, known as Section 1201, of a prohibition on circumvention of access restriction controls or technological protections put in place by the copyright owner. If a copyright owner puts an access restriction scheme in place to protect a copyright, unauthorized access is essentially illegal. However, it is still unclear how to define "access restriction" if such measures can be circumvented by holding the shift key at start-up of a personal computer, as in the case of one access restriction workaround or any consumer action that is inherent to the use of general computing devices. The Librarian of Congress conducted a proceeding in late 2000 to provide guidance to Congress.

Digital Theft Deterrence and Copyright Damages Improvement Act (1999). Congress increased damages that can be assessed on copyright infringements from that of $500 to $750 to $20,000 to $30,000. Willful infringement increased from $100,000 to $150,000.

Librarian of Congress Issues Exemptions to the DMCA (2000). Librarian of Congress issues exemptions to the DMCA, Section 1201(a)(1), the Anti-Circumvention Provision, for "classes of works" that adhere to fair use. These two exemptions include: "Compilations consisting of lists of websites blocked by filtering software applications; and Literary works, including computer programs and databases, protected by access control mechanisms that fail to permit access because of malfunction, damage, or obsoleteness." The full recommendation can be found at http://www.loc.gov/copyright/1201/anticirc.html.

Dmitri Skylyarov Arrested under DMCA Provisions (2001). The Russian programmer for ElcomSoft was accused of circumventing Adobe Systems' eBook Reader DRM. Although Adobe later reversed course, government attorneys continued with the prosecution of the case, presumably to test the interpretation of the DMCA. As one of the first criminal cases brought under the DMCA, many observers viewed this as a test case for how far allegations under the DMCA could be pushed into actual indictments. A federal jury returned a verdict of "not guilty" in late 2002.

U.S. Supreme Court Hears Challenge to Sonny Bono Copyright Term Extension Act, the "CTEA" (2002). In copyright debates Lawrence Lessig, a well-known constitutional scholar, has been active in promulgating such mechanisms as the "Creative Commons." His representation of the plaintiffs in *Eric Eldred v. John Ashcroft* extended his experience in the copyright debate. Ultimately, the Supreme Court ruled against the plaintiffs, affirming the constitutionality of the CTEA and affirming Congress's role in intellectual property. Retrospectively, the CTEA extended existing copyrights by 20 years—to 70 years from the life of an author, from 50 years. As well, adding 20 years of protection to future works. Protection was extended from 75 to 95 years for "works made for hire," a common contractual framework used by many corporations.

***MGM v. Grokster* (2005).** It is unclear how many rounds of dispute resolution between technology innovators and content owners will go before the courts or Congress. For this reason, it may take some time to understand fully the impact of the *MGM v. Grokster* decision. The most widely quoted aspect of the ruling, thus far, concerns who should determine when a device is "promoted" to infringe copyright. The Supreme Court essentially decided:

> For the same reasons that *Sony* took the staple-article doctrine of patent law as a model for its copyright safe-harbor rule, the inducement rule, too, is a sensible one for copyright. We adopt it here, holding that one who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties. We are, of course, mindful of the need to keep from trenching on regular commerce or discouraging the development of technologies with lawful and unlawful potential. Accordingly, just as *Sony* did not find intentional inducement despite the knowledge of the VCR manufacturer that its device could be used to infringe, 464 U.S., at 439, n. 19, mere knowledge of infringing potential or of actual infringing uses would not be enough here to subject a distributor to liability. Nor would ordinary acts incident to product distribution, such as offering customers technical support or product updates, support liability in themselves. The inducement rule, instead, premises liability on purposeful, culpable expression and conduct, and thus does nothing to compromise legitimate commerce or discourage innovation having a lawful promise [5].

In the world of physical media distribution, there are many channels available, both for broadcast and for physical carriers. Specialized retailers compete for consumer sales by differentiating their efforts from other more generalized retailers. Written content and imagery attracts consumers to publications such as magazines; and spoken content and music selection attracts consumers to radio. The number of possible combinations of content and editorial material provides for rich broadcast opportunities, which have the effect of attracting advertising

dollars to the broadcasters. The parallels with online streaming or pay-per-click-type schemes are not a coincidence. Total spending on advertising has continued to grow over time, although the ability to reach a profitable, aggregated group of consumers has grown more difficult. The ability to reach paying audiences is the obvious aim of advertising.

The argument that there is too much entertainment vying for consumers' dollars is beginning to meet the more complicated issue of how to measure actual time for said consumption, while deploying efforts at protecting copyrighted material. Supply meets demand whether measured in units of time (e.g., minutes on a cellular phone), bandwidth (e.g., amount of data per unit of time), or copyrighted CDs, books, and DVDs. Some agreement on the unit of measurement obviously needs consideration. When supply is controlled, as with generalized DRM, the ability to measure demand may become distorted. Though the conclusions are contentious, the arguments can be made from a variety of viewpoints. Simply, can technical controls for accessing copyrighted material cost less than the cost of implementation and maintenance of these same controls? How are new devices and services handled given legacy control systems or even open systems? Is there value in securing copyrights with DRM? What rights of revocation exist, and who should determine the scope and form of revocation? How much open access should be provided to consumers? Is there value in providing copyrighted works for free? What constitutes a consumer's property in contrast with a content provider's property?

1.4 COMMUNICATION THEORY—WHO SCREAMS LOUDEST?

When considering the security of multimedia data, several issues pose challenges. First, multimedia data is compressible and easily transferable. Second, advances in digital signal processing have made the ability to digitize analog waveforms both economic and more commercially viable. Third, ownership and responsibility for any copies made of digitized content are typically a double-edged sword. Manufacturing has been made inexpensive to the owners and licensors, increasing profit margins, but content has increasingly been copied without regard to the interests of those rights holders. More on these issues will be discussed below.

1.4.1 Shannon's Gift

Before delving into technical aspects of DRM, attention must be paid to communications and cryptography. Cryptography has impacted history at several points. World War II was emblematic of the tight relationship between codes, militaries, governments, and politics—before the first microprocessors, but at a time of great technical innovation. The work in cracking the codes of that war was

supplemented later by a growing interest in the underlying nature of communications. Largely unknown to the public, the seminal work of Claude E. Shannon in *The Mathematical Theory of Communication* and *Communication Theory of Secrecy Systems* provides helpful analysis in what can be expected theoretically. Developments based on communication theory, including cryptographic systems, are pervasive in modern society. The impact on our daily lives is incalculable. Telephones, financial markets, and even privacy itself have changed in dramatic, often unpredictable, ways. The demand for codes to assist with the secure transport of sensitive data was matched by the increasing importance of computerized networks for dispersal and distribution of such data.

At some point, confidentiality, one of several primitives designed into data security systems, was met by increasing calls for restrictions on the deployment of cryptographic protocols. Separately, but just as important, authentication, data integrity, and non-repudiation—additional primitives of cryptography—assisted in the growth of business over electronic networks. Public key cryptography provides all four of these primitives, in a manner making distribution of codes and ciphers economically feasible for all persons wishing to secure their communications. The landmark failure of the U.S. government's Clipper chip [6] in 1993 was only the beginning of an increased public interest in cryptography. With the proliferation of more bandwidth and anonymity, in many cases based on so-called strong encryption, commercial concerns were also heightened. Here, we deal specifically with copyrighted works such as images, audio, video, and multimedia in general. A basic notion that should be considered in understanding DRM may well be how to balance privacy with notions of piracy. Ironically, the emphasis on protecting privacy has been trumped in many ways by the goal of securing against piracy. Should personal secrets be shared to satisfy the demands of copyright holders? Put another way, is a social security number used to secure a purchase for a song download a fair exchange of value asserted by the copyright holder?

Shannon's conceptualization of communication theory provides a fitting background to copy protection techniques to be explored in this book. Actual performance of real-world systems should be matched against theory to encourage appropriate expectations. Communication theory at its most basic level is about the transmission of information between a sender and a receiver. The information typically has meaning or context. Obviously, there are limitations to communication systems as explored by Shannon and others. The channel and destination of the information being transmitted provide additional parameters to a communication system. Here, we eliminate the simplified arrangements for a noiseless communication channel where the inputs and outputs are equivalent. By noiseless we mean no "chance variables" occur, and thus no redundancy or other error correction is needed to communicate messages.

The ratio of the actual rate of information transmission to capacity in a given channel is called the efficiency of the coding scheme. Efficiency to both the sender

and the receiver can hav
istic scheme is analyzed,
it is proven that there ar
sets of errors (which can
Because binary data is ei
each bit of data in the a
with 1 or 0 being the limi
elements of the coin flip c
munication, the entropy
source, the entropy of th
tion, correction, and conc
with the context of the ir
and the entropy of the s
may be successfully repr
part, Shannon's legacy.
cern ourselves with how
approximate the original
processing and in a philos
replica, but is an exact f
mated waveform. The nat
value of the coding schen
economics of deployment

In a discrete channel,
a "chance variable," wh
precise digitization syste
communicate informatio
hardware systems. As w
senders and receivers can
of functions that facilitate
ted. Similarly, the ensem
the presence of noise in
observers. So long as the
associated message can
by a cryptographic algo
tationally easy to discov
cipher.

The key is thus a stat
receiver can be assured th
sage. The data transmis
communication channel.
tion error (e.g., what is r
be small relative to the

and the receiver can have subjective measurements as well. When a more realistic scheme is analyzed, namely efficient transmission in the presence of noise, it is proven that there are still a finite number of errors (perceptibly "noise") or sets of errors (which can be mathematically generalized to create noise filters). Because binary data is either a "1" or a "0" in a given channel, we can say that each bit of data in the abstract may be completely random by flipping a coin, with 1 or 0 being the limited choices. That is not to say that entropy of any of the elements of the coin flip can be ignored. However, in order to ensure effective communication, the entropy of any chance variables, the entropy of the information source, the entropy of the channel, etc. must be taken into account. Error detection, correction, and concealment form a large body of work in dealing specifically with the context of the information, the channel and nature of the transmission, and the entropy of the source impacts the channel capacity. That information may be successfully reproduced and can be expressed mathematically is, in large part, Shannon's legacy. This applies to cell phones and DVDs. Here, we concern ourselves with how a perceptible signal can be digitized, or "sampled," to approximate the original analog waveform. However, as is well known in signal processing and in a philosophical sense, the digitized signal can never be a perfect replica, but is an exact facsimile of an otherwise analog and infinitely approximated waveform. The natural limit is quantization itself; however, the limit of the value of the coding scheme in terms of practical use is human perception and the economics of deployment.

In a discrete channel, entropy measures in an exact way the randomness of a "chance variable," which itself may be random. The development of very precise digitization systems representing an "ensemble of functions" used to communicate information has been reduced into a multitude of software or hardware systems. As we delve into cryptography, here, we quickly note that senders and receivers can exchange secrets, or "keys," associated with an ensemble of functions that facilitate agreement over the integrity of the data to be transmitted. Similarly, the ensemble of functions assures transmission of the message in the presence of noise in the channel. Keys may be mistaken as noise by other observers. So long as the sender and receiver can agree to the key, the "secret," the associated message can be authenticated. The key is ciphered (i.e., processed by a cryptographic algorithm) in a manner to mimic randomness not computationally easy to discover even if the other observers are in possession of the cipher.

The key is thus a state or index of an ensemble of functions from which the receiver can be assured that the sender of the message did indeed transmit the message. The data transmission's discrete rate may not exceed the capacity of the communication channel. Finally, relating back to sampled signals, the quantization error (e.g., what is related to data conversion between analog to digital) must be small relative to the information transmitted in order to establish sufficiently

small probabilities that the received signal is the communication intended by the sender. Statistically isolating "perturbing noise" from other errors and bounding upper and lower limits of capacity in a communication channel are presently computationally easy.

The introduction of digital CDs resulted from agreements over trade-offs of the general technologies so far described. As a medium for music, it is fitting to observe this medium for rich discussions on DRM. The CD is itself a discrete communication channel. The reflective material sandwiched between transparent plastic, which can be read by a CD player, is converted into a series of binary data (1s and 0s) as physical pits on the reflective material substrate. This data stream has pre-determined sampling rates and quantization values (16 bits, 44.1 kHz per second, for a Red Book Specification Audio Compact Disc). Again, data bits which have pre-determined locations or modality on the physical CD, are fed through an ensemble of functions which filter the digitized sample information stream into analog audio signal data. This data, of course, may be compressed for more economic use of bandwidth. We hear a song, the binary information sent out to an amplifier to be transduced, but, there is no "perceptually obvious" relationship with the music rendered. The data are presented according to the Red Book standard. We hear the music with our psychoacoustic abilities, our ears, and ultimately, our brains process the music and may associate the music information with some other independent or unrelated information.

Any such "associated information" may be different for every listening experience, every time for every individual listener. We would call this associated information "value added" or "rich" because it can be associated, with other independent information that may have no relationship with the primary communicated information which is the same for all listeners. The "hits" are hits for each individual in different ways that are aggregated in such a manner that they can be called hits—the memorable song for a high school prom, the one played when waking up, or any number of events associated with the copyrighted work in unintended ways, impacting the value attributed to such a work. Money is one obvious measure of success. Acting out a song may reflect the meaning intended by its creator or it may not. What matters with regards to DRM are the decisions made by creators and consumers of copyrighted works to create, seek, and consume with a fixed and limited amount of time and money determined by the harsh realities of the marketplace. Recognizable and potentially valuable multimedia can be rendered by general computing devices. Multimedia having many different interpretations depending on what stake the party has in the work. After all, creators, too, may give their work away for free.

We have generalized that it is computationally feasible to reproduce information, allowing senders and receivers to share the gestalt of information that may be transmitted. We ignore the specifics of digital filters and error correction to stress the point that, conceptually, data can be communicated and

communicated securely. If the communication channel is too expensive, based on bandwidth or overall available transmission capacity or, as is central to this book, the cost of protection, it ceases to play a role in enabling security of data. Additionally, if the bandwidth requirements for reproduction are sufficiently high, certain other types of data are not computationally feasible to economically transmit over communication channels. As more information is digitized and, by extension, digitally copied, even if there are imperceptible differences with the original analog waveform, the limit to data transmission becomes closely linked to bandwidth [7].

Interestingly enough, Shannon does address "intelligibility criterion" of information transmissions in providing "fidelity evaluation functions." Because systems must be economically practical, and information is ultimately deemed authentic or genuine by the creator or source of the information (assuming the source is trusted or the information can be verified), human perception does play a role in establishing a close enough proximity of replicated data information, when "exact recovery" is infeasible, given the presence of noise in communications channels. The five examples Shannon provides for measuring acceptable fidelity of a proposed information channel include root mean square (i.e., "RMS," to assist in determining coordinate information of the data), frequency weighted root mean square (essentially weighting different frequency components prior to RMS, which is similar to passing the distance between data through a shaping filter and calculating the average power of data output), absolute error criterion (over the period of zero to a discrete time), human perception (which cannot be defined explicitly, though we can observe how noise is received by our senses and our brain, sufficiently subjective parameters), and the discrete case (differencing input from output and dividing by the total amount of input data).

1.4.2 Kerckhoffs' Limits

In cryptography, the content or bits comprising the message must not be changed in order to provide acceptable levels of confidence in a secure system. However, systems themselves cannot guarantee security. A human can compromise a system by providing passwords or systems may generate weak pseudo-random numbers, making the most seemingly strong "cryptographic algorithm" ("cipher") unsecure. A "keyed" algorithm defines an ensemble of functions with the specific member of the ensemble identified by a unique key. With respect to encryption, the set of all keys defines a plurality of encryption functions. Each element is instantiated by a specific key. Though there may be randomness ("entropy") within the input, the use of the randomness only relates to the manner in which the function operates as a Turing machine (e.g., a general computing device). The random

choice of a key to specify the element in the plurality of encryption functions is essential.

As Shannon stressed, communications is concerned with "operations on ensembles of functions," not with "operations on particular functions." Cryptography, too, is about ensembles of functions. The basic difference with coding (i.e., communications) is the exchange of the key. The ensemble of functions occupies a finite set, so that the input and output can be secured by associating the data to be transmitted with a randomly generated key that is pre-determined by both parties by some mutually agreed to means—the cryptographic algorithm or cipher. Kerckhoffs' law is the foundation by which such determinations are made; it is assumed that the adversary possesses the cipher, and thus the security must rest in the key. Auguste Kerckhoffs provided five additional principles, including (1) system indecipherability, (2) the key must be changeable, (3) the system should be compatible with the means of communication, (4) portability and compactness of the system is essential, and (5) ease of use. Of these principles, ease of use and whether security rests with the key have historically made for difficult engineering challenges within DRM. In cases where DRM systems must come in contact with other DRM systems, these challenges are heightened. Some have argued that it is not possible to tamperproof cryptographic systems to sufficiently prevent hacks [8]. This has obvious impacts on DRM.

1.5 CRYPTOGRAPHY—MUCH TO DO

With a basic understanding of communications theory and its relationship with cryptography, we can describe two conventional techniques for providing key-based confidentiality and authentication currently in use: symmetric and asymmetric encryption. Both systems use non-secret algorithms to provide encryption and decryption and keys that are used by the algorithm. This is the basis for Kerckhoffs' law: all security should reside in the key, as it is assumed the adversary will have access to the cryptographic algorithm. In symmetric systems, such as AES, the decryption key is derivable from the encryption key without compromising the security of the message. To assure confidentiality and authenticity, the key should be known only to the sending and receiving entities and is traditionally provided to the systems by secure physical communication, such as human courier. Other systems where a common key may be developed by the sender and receiver using non-secure communications are widely deployed. In such systems, each party to a communication generates a numerical sequence, operates on the sequence, and transfers the result to the other party. By further operation using the transferred result and the locally generated sequence, each party can develop the identical encryption key, which cannot be obtained from the transferred results alone. As implemented for use over the Internet, common encryption systems are

those denoted by the Secure Socket Layer (SSL) and IP Security Protocol (IPSEC) protocols.

In asymmetric encryption systems, a first party to a communication generates a numerical sequence and uses that sequence to generate non-reciprocal and different encrypting and decrypting keys. The encrypting key is then transferred to a second party in a non-secure communication. The second party uses the encrypting key (called a public key because it is no longer secure) to encrypt a message that can only be decrypted by the decrypting key retained by the first party. The key generation algorithm is arranged such that the decrypting key cannot be derived from the public encrypting key. Similar methods are known for using non-reciprocal keys for authentication of a transmission. There are also digital signature algorithms. In some cases, as with RSA, encryption and digital signature functionality are properties incorporated by the same algorithm. In a manner parallel with the real-world handwritten signatures, the non-secure public key can be used to tamperproof a message (i.e., providing nonrepudiation) that has been digitally signed using a secure "private" or secret key known only to the originating party—the signer. Thus, the receiving party has assurance that the origination of the message is the party who has supplied the "public" decrypting key. So, how does this relate to DRM? We have devised several areas of interest to establish commonality of the elements typically considered in designing a DRM system, namely authentication, data integrity, non-repudiation, and confidentiality. However, DRM is inherently constrained from legal, economic, and political constraints, as well as consumer expectations—not strictly cryptography or more generally communication theory. Mentioned previously, some argue it is not possible to tamperproof software programs given the inherent foundations of communications. Within the DRM product and service space, terminology and practicality can vary widely. Here, we generalize DRM by discussing "wrapping" and "embedding," so-called "digital watermark," technology.

1.6 DIGITAL RIGHTS MANAGEMENT—WRAPPING AND EMBEDDING

It is not prudent to limit our discussion solely on word choice. Essentially, the terms may not always reflect the utility or functionality of the protections being described. Rights are typically matched by responsibilities. DRM offers up examples of how stakeholders may not share common interests [9]. Copy protection and content extensions generally apply to digitized content, while "scrambling," a scheme related to encryption, may be applied to an analog signal. Such analog scrambling is evident in analog cable and analog cell phone systems. Encryption, as discussed previously, scrambles content, but the number of 1s and 0s may be different after the encryption process. In some scenarios, prior to enabling access to content it must be decrypted, with the point being that once the content has been encrypted,

it cannot be used until it is decrypted. Encrypted audio content itself might sound like incomprehensible screeching, while an encrypted image or video might appear as random noise when viewed. The encryption acts as a transmission security measure—access control. One approach has commonly been called "conditional access" when someone or something has the right to access the media. In many scenarios, identifying information or authentication of that party must first be completed prior to decryption of the content or description of the intended scope of use. There may be layered access restrictions within the same scheme. In either case, the transmission protection ends when the content is to be observed.

Encryption is poorly applied in at least two specific areas with respect to copy protection of content. First, so-called "pirates" have historically found ways to crack the protection as it is applied to content. The effect is essentially equivalent to obtaining the decryption key without paying for it. One such technique is "differencing," where an unencrypted version of the content is compared with an encrypted version of the same to discover the encryption key or other protections. Differencing is also a weakness in many digital watermark systems. In some watermark systems, the requirement to maintain original unwatermarked material for comparing and recovering embedded code from a suspect copy of content introduces other problematic issues such as additional data storage requirements at the detection side. Why store watermarked content for protection purposes when unwatermarked content may exist at the same site for decoding said watermarks? Second, and perhaps more complicated to address, is that once a single legitimate copy of content has been decrypted, a pirate is now free to make unlimited copies of the decrypted copy. In effect, in order to make, sell, or distribute an unlimited quantity of content, the pirates could simply buy one copy, which they are authorized to decrypt, and make as many copies as desired. These issues were historically referred to as the "digital copy problem"; others prefer "digital piracy."

Copy protection also includes various methods by which an engineer can write software in a clever manner to determine if it has been copied and, if so, to deactivate the software. The same engineer may be a "rogue engineer" who essentially has the backdoor key to deactivate the copy protection. This is typically the result of a poorly chosen encryption algorithm or means for obtaining a key. Also included are undocumented changes to the storage format of the content. Copy protection was generally abandoned by the software industry, since pirates were generally just as clever as the software engineers and figured out ways to modify their software and deactivate the protection. The cost of developing such protection was also not justified considering the level of piracy that occurred despite the copy protection. That being said, the expansion of software product activation keys, online registration schemes, and registered version upgrades indicates increased interest and benefit in securing even software programs. Software watermarking schemes, including those using "steganographic ciphers," have correspondingly increased over the past few years [10].

Content extension refers to any system attaching some extra information indicating whether a copy of the original content can be made or some other logic with regards to the use and accessibility of the content. A software or hardware system must be specifically built around this scheme to recognize the additional information and interpret it in an appropriate manner. An early example of such a system is the Serial Copyright Management System (SCMS) included in Digital Audio Tape (DAT) hardware. Under this system, additional information is stored on the track immediately preceding each sound recording indicating whether or not it can be copied. The hardware reads this information and uses it accordingly. By wrapping content, we are generally referring to "content extensions." We further formalize concepts below.

When we discuss watermarks, we are addressing steganography, or hiding information in plain view, in combination with cryptographic techniques. They need not be mutually exclusive and in many cases complement each other. Watermarks [11] are a unique technology that embed and protect a "code" by placing "transactional information" intrinsically within the electronic work. The transaction information can specify time, date, recipient, and supplementary information known by the transmitter at the time of the transfer to the recipient. Review of the electronic copy of the media at a later instance reveals the historical record of the electronic copy. Safeguarding from manipulation or deletion, unauthorized modification of the transactional information results in degradation of the perceptual quality of the work. Tampering with watermarked media is, thus, quickly identifiable. More advanced schemes include watermark code which itself interacts with the system. This code, with or without interaction with a key, can upgrade content security systems and can be characterized by a variety of interactions between the protection scheme, associated keys, watermark information, and content to be protected. Before delving into finer detail, we note that it is unclear that any wrapped, embedded, or generally "DRM'd" content has remained wrapped or inaccessible. In parallel, we have not observed clear examples where copyright holders have yet to eschew traditional distribution channels to achieve economic success solely through DRM distribution schemes.

1.6.1 Who Is in Control—Active and Reactive Controls

Protection of copyrighted works may be a proactive control that reduces the potential of loss at the time of an event, while a reactive control provides an audit trail after the fact to conclude what happened and by whom. The two types of controls are complementary and, in many cases, can and should be used concurrently. Such consideration as the time value of the content, that period in which the content is worth most for protection, is subjective and varies among media types, distribution channels, and market forces. Yesterday's newspaper arguably suffers far greater reduction of economic value than long-running hits on Broadway during

the time it takes a new edition of the newspaper to appear (changes in critiques of the Broadway work, notwithstanding). Uniqueness over data or data copies assists in establishing responsibility for the data. Similar to the physical world use of receipts for transactions over the "same" material, watermarks act as a control for receipts of digitized data. However, time also plays a significant role in value.

Active controls provide a first line of defense in times of a breach in security. With regard to data security risks, there are several types of commonly established information security controls, generally categorized as physical, procedural, and logical controls. Physical controls are generally building access and alarm systems. Procedural controls include policies, operating procedures, training, and audits. Logical controls are placed at the computer system level and include application and operating system-level access controls, lists, and perimeter protection with firewalls, router security, and intrusion detection systems. With respect to the copying of copyrighted media, the most common type of active controls is "security wrappers," often called ("active") DRM [12]. A wrapper wraps the digital media around a digital structure to prevent extraction of the media from the stored data object. Generally, the wrapper includes encryption of the media, "meta-data" about the media, and may include other logic, encrypted or not. A simplistic explanation follows here.

First, content is encoded with associated meta-data, followed by encryption of the meta-data and media, and any additional non-encrypted data may be placed. Finally, additional information, oftentimes a software wrapper, that must be run to extract the media is added. The data object is stored directly within the software wrapper. That is, the media is blanketed with multiple layers of controls. To obtain the media in a perceptually similar form, the wrappers must be removed. Hence, this is an active control. However, to be useful, the wrapper must be removed, making the media extremely vulnerable at the time of use (viewing, playback, etc., when the media is "in the clear" and susceptible to unauthorized use). The software wrapper may also require active coordination by a third party during the unwrapping process. For instance, the software wrapper may require interaction with the content provider to obtain keys to decrypt the content. This communication requirement adds additional complexity to the process and, if required, places additional constraints when the active DRM-protected media is part of a larger workflow. Watermarks need not be incorporated in the previous example. Instead, meta-data are placed external to the content for operational requirements, and both the meta-data and the media are encrypted. The meta-data, for instance, may provide cryptographic authentication of the media or may provide keys for an external cryptographic operation that must be performed again including upgrades to the system in parts or in its entirety. The placement of watermarks as an additional reactive control provides complementary benefits.

Reactive controls do not actively prevent misappropriation or data transfer from happening. However, the benefits of reactive controls are multifold. To support

recovery of losses, so-called "tracing traitors" or "identifying pirates," reactive controls provide an audit trail for actuarial or forensic analysis. As an ancillary benefit to their forensic capability, reactive controls act as a deterrent. Knowing that a copy can be traced back to a pirate is common in traditional commerce. Reactive controls may also assist with authentication or indicate tampering. Furthermore, valuable actuarial information may be obtained through reactive controls, providing marketing information intrinsic to the data objects or distribution channels being utilized. Reactive controls are complementary to active controls and may be used concurrently.

Watermarks are a "reactive" DRM control technique. Unlike wrappers, watermarks are maintained throughout the data workflow. As watermarks are intrinsically embedded into the content, they cannot be removed during processing of the copy of the digital media. Ideally, attempts at removal result in degradation of the content and a corresponding devaluation of the content's economic worth. As the copy of the media is moved through its expected and unexpected workflow, there are no stages requiring removal of the watermark as the media retains its same perceptual qualities. As watermarks do not modify the copy to a new format via wrapping or encryption, processing and workflow used prior to the incorporation of watermarks in the media do not require modification. Previous processes continue to stay the same without the incorporation of new steps or technology. Moreover, the watermark is retained in each step of the workflow rather than being stripped off as is required in many security controls employing encryption or wrappers. Once the wrappers are stripped off, they are ineffective, and the only protection mechanism remaining is the reactive controls. Watermarks can be designed to survive format and data transformations between digital and analog domains for varying degrees of persistence. This persistence assists with analysis of data that exists in different formats or channels.

1.6.2 Traceability and Active Controls

Watermarks, being a part of the media rather than external to it, are not susceptible to problems with wrappers. Moreover, when used in conjunction with an active control, watermarks are not removed during the unwrapping process. By indelibly mapping transaction information to the characteristics of the media, watermarks are presently deployed in several active control environments to manage authentication of the media and enable such features as copy management and even system upgradeability. In a manner parallel to physical money, active controls are comparable to copy protection features, including ink type, paper stock, fiber, angles of artwork that distort in photocopier machines, inserted magnetic strips, and composite art. When all of these security features are reduced to digital data, active and reactive controls can be similarly compared. These controls are intended to make money difficult to reproduce, while the serial number is intended to enable

audits of specific transactions. Responsibility over individual media copies via watermarks can be used to enable policies regarding use limitations, first and third party transfers, and any number of active controls.

Though active controls provide a first line of defense, they have many inherent deficiencies. By the very nature of a wrapper, it must be unwrapped to use. Similar to a crab moving out of its shell, at the point of unwrapping the media has no effective protection mechanism. In practice, several technologies have been used to actively protect the media, including physical protection. However, these additional controls have limited effectiveness given the sophistication of hackers, complexity of the wrapper, and inconveniences presented to users. Once hacks have been successfully made, it is relatively easy for less sophisticated users to deploy the same hack with little effort. Wrappers increase overall processing requirements depending on operating systems or file formats limiting persistent protection. Inconvenience is the most significant problem for the users of the media. Unless each step of the workflow is able to unwrap "securely," the process leaves exposed media vulnerable. Active controls limit the movement of information, as each process requires the unwrapping technology associated with it.

1.6.3 Binding Transactions, Not Just a Handshake

The placement of transactional information directly into media works has many benefits. First and foremost, it creates an audit trail embedded directly into the work. This information can include time, place, and the identities of the transferring party and the transferee of the electronic media. Whereas system logs on computers can state prior actions that have taken place on a server, these logs cannot be used to analyze two copies of the same media and state the past history of the works. Yet today, it is not uncommon that multiple copies of the same media are transferred to multiple parties, including internal and external parties. System logs are insufficient to determine cause during a forensic analysis of media discovered at an unauthorized location unless each copy is serialized. System logs also make analysis of first and third party responsibility an unsupported process, if applied alone. In practice, a unique serial or transaction number, rather than the actual, copyable information, is placed as a search index to map back to additional transaction information (e.g., name, date, time, distribution channel, transaction id, etc.) stored in a database. Such hierarchy, or layering of "unique digitized data," is beneficial for workflow separation [13] and assigning responsibility over data as it moves within and beyond an organization's electronic systems.

As a single work (or other electronic media) may be digitally copied into multiple digital works at little or no marginal cost, digital watermarks ensure that each digital work is uniquely serialized. Similar to physical money with serial numbers, each

unit is unequivocally different and perceptually equivalent from other copies of the same source. Properly deployed, digital watermarks enable inherent audit trails of digital data in any number of electronic transactions or workflows. For instance, Person A has a copyrighted work with their identity embedded as the watermark "A". In transferring a copy of the digital work to Person B, Person A imprints a watermark with identity "B" into a new copy of the work. This process can be repeated from Person B to Person C and so forth. Similarly, additional transactional information or a unique serial or transaction number may be placed into the work via a watermark. In the process, each electronic copy is digitally unique yet perceptually the same. Hence, each copy incorporates an internally embedded audit trail of its transactional history. The same work may also have been transferred by the same person to two different entities. In this scenario, a work sent to "B" is uniquely different, but perceptually equivalent to a work sent to "C." As the data is digital rather than physical, a recipient may create exact copies. Because of the watermark, each new copy must contain the previous embedded audit trail relating to its past history. Each work, independent of what the watermark contains and the number of watermarks incorporated into the copy, is perceptually the same. From an auditing and forensic point of view, these are unique. A copy with watermark "A, B" relates to a work that was last authorized for transfer from Person A to Person B and was not obtained directly from "C" or from "A."

1.7 NOW, THE FUTURE

Looking backward at the progress of technology, as with any hindsight, is much simpler than projecting forward. The concepts discussed here do not represent the definitive "last word," but an introduction to an important aspect of the technology landscape. DRM is a subject with so many competing stakeholders that new paradigms or business models do not necessarily appear obvious [14], and the viewpoints are not mutually exclusive. However, business is primarily an exercise in seeking profits. Measuring profitability or even accountability are invaluable starting points, but by no means is money the only perspective nor should it be, especially with regards to copyright. It is not just copyrighted multimedia that is impacted by advances and debates over DRM. Arguably, all intellectual property will be subjected to similar pressures. A valuable and fungible asset in the economy, besides time, is trust. Trust itself shapes many of the compromises that are needed in further commercializing networks [15]. An important aside: if we knew what the "blockbusters" would be, we would forgo the agents, promotion, distribution channels, specialty retailers, and all other middlemen and offer the valuable works from the back of our cars. *Caveat emptor.*

ACKNOWLEDGMENTS

Thanks for all of the rich insight and valuable comments received over the past decade. Special thanks goes to Yair Frankel and my family.

REFERENCES

[1] Lotus Development Corporation v. Borland International, Inc., 49 F. 3d 807, 818 (1st Cir. 1995).

[2] P. Durdik. Reverse Engineering As A Fair Use Defense To Software Copyright Infringement, Jurimetrics J., 34:451–470, Summer 1994.

[3] C. Miller. New Technology and Old Protection: The Case for Resale Royalties on the Retail Sale of Used CDs. Hastings Law Journal, 46:217–241, November 1994.

[4] Originally, T. Carroll. A Frequently Asked Questions Document Discussing Copyright Law ftp://ftp.aimnet.com/pub/users/carroll/law/copyright/faq/part2. Updated on September 11, 2002, http://www.tjc.com/copyright/FAQ/.

[5] Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd. (04-480) 380 F. 3d 1154 (Sup Ct. 2005).

[6] Basically, a chip for encryption with a backdoor for the government.

[7] S. Moskowitz. *Bandwidth as Currency*, IEEE MultiMedia, pp. 14–21, January–March 2003.

[8] Barak, Goldreich, et al. *On the (Im)possibility of Obfuscating Programs*. An extended abstract appeared in *CRYPTO 2001*, pp. 1–43, August 16, 2001.

[9] R.J. Anderson. Cryptography and Competition Policy—Issues with 'Trusted Computing,' 2003 Wenk Lecture related to R.J. Anderson. TCPA/Palladium FAQ, at http://www.cl.cam.ac.uk/users/rja14/tcpa-faq.html/.

[10] Method for Stega-Cipher Protection of Computer Code, U.S. Patent No. 5,745,569, Patent and Trademark Office, 1998.

[11] There are many forms of digital watermarks. We generically use watermark to mean forensic or traceable watermark.

[12] "Active DRM" is our preferred terminology to distinguish between DRM technologies providing active controls and DRM technologies providing reactive controls.

[13] Workflow separation refers to the steps, or identifiable points, data moves as it is being prepared or processed.

[14] W. Fisher. Promises to Keep: Technology, Law, and the Future of Entertainment, PTKIntroduction.doc Draft, pp. 1–21, March 22, 2003. Also see http://www.harvard.edu/Academic_Affairs/coursepages/tfisher/Music.html/

[15] A. Odlyzko. Privacy, Economics, and Price Discrimination on the Internet [Extended Abstract], http://www.dtc.umn.edu/~odlyzko, pp. 1–16, July 27, 2003.

Dig
Ma

Marina

2.1 INTRODUCT
MANAGEME

In recent years there
content from analog
tion from vinyl recor
formats; from books
MPEG-2, HD-TV, ar
Video format and di
enhanced end-user e
distributions, as well

Content owners c
source of revenue a
files containing copy
an "implicit" form of
specter of unlimited
compensation to the

Digital Rights M
distribution of digit
technology doesn't t
but instead provide
words, DRM system
but the existence of
protected.