Digitale Wasserzeichen für den Urheberrechtsschutz

Philipp Tomsich/Stefan Katzenbeisser

Institut für Softwaretechnik, Technische Universität Wien
Favoritenstrasse 9-11/E188
phil@ifs.tuwien.ac.at, skatzenbeisser@acm.org

Schlagworte: Digital Watermarking, dezentrale kryptographische Protokolle, manipulationsgeschütze Hardware, Copyright Protection, digitale Medien

Abstract: Mit der Vordringen von digitalen Technologien und dem Internet in den Unterhaltungsbereich und der zunehmenden Ausrichtung auf elektronische Distributionskanäle für den Vertrieb von multimedialen Inhalten, entstehen neue Herausforderungen für den Schutz der Urheberrechte. Elektronisch übermittelte Inhalte erzeugen zunehmende Konflikte und Probleme auf diesem Gebiet angesichts der leichten und originalidenten Reproduktion digitaler Medien. Eine weitverbreitete und vielversprechende Möglichkeit um Bild-, Video- und Tondaten zu verfolgen stellt die Verwendung von digitalen Wasserzeichen dar. Im Folgenden werden die Voraussetzungen für ein dezentrales Watermarking Protokoll vorgestellt, das auf Basis von manipulationsgeschützter Hardware und einer Public-Key Infrastruktur die Identifikation und Verfolgung von urheberrechtlich geschützten Inhalten gestattet.

1. Problemstellung

Mit der zunehmenden Verfügbarkeit und dem Vertrieb von urheberrechtlich geschützten Inhalten in digitaler Form entstehen neue Herausforderungen für den Schutz des geistigen Eigentums der Urheberrechtsinhaber. Die Möglichkeit einfach und billig Kopien herzustellen gefährdet die Film-, Musik- und Unterhaltungsindustrie. Folglich kann die Entwicklung und der Einsatz von effektiven technischen Hilfsmitteln zum Schutz der Urheberrechte als eine zentrale Voraussetzung für die Akzeptanz elektronischer Vertriebskanäle von Seiten der Anbieter von multimedialen Inhalten angesehen werden.

Zwei grundsätzlich unterschiedliche Ansätze zur Reduktion des erhöhten Risikos von Urheberrechtsverletzungen bei digitalen Medien können in der Literatur identifiziert werden. Ein **Kopierschutz** versucht den Zugang zu urheberrechtlich geschützten Materialien dermaßen einzuschränken, dass die Erstellung von Kopien gänzlich verhindert wird. Beispiele für die Verwendung eines Kopierschutzes sind Lizenzserver und

technische Mechanismen an den verwendeten Datenträgern. Das Design von technischen Verfahren zum Schutz vor unauthorisierten Kopieren in einem offenen System (wie dem Internet) scheint sehr schwierig. Etwa wurde der Kopierschutz auf DVDs ein paar Wochen nach Markteinführung bereits gebrochen[1].

Der **Urheberrechtsschutz** hingegen, bringt Informationen über die Nutzungsrechte, den Käufer und den Urheber direkt in einem digitalen Objekt derart an, dass für einen Betrachter kein Qualitätsverlust bemerkbar ist. Weiters ist sehr schwierig diese Information wieder aus dem Objekt zu entfernen, ohne es gänzlich zu verstören. Sollte das Urheberrecht eines solcherart modifizierten Werks jemals umstritten sein, kann diese Information extrahiert werden, um den rechtmäßigen Urheber festzustellen. Wenn auch die Identität des Käufers codiert wird, kann sogar der Verbreitungsweg von etwaigen aufgefundenen unauthorisierten Kopien nachvollzogen werden. Der bekannteste und verbreiteste Weg um Information in Bild-, Ton- und Videodaten anzubringen stellt die Verwendung von digitalen Wasserzeichen[2] dar.

Ein Watermarkingverfahren besteht stets aus zwei Algorithmen, wobei der eine zur Anbringung des Wasserzeichens und der andere zur Extraktion eines bereits bestehenden Wasserzeichens verwendet wird. Je nach der Art des Extraktionsalgorithmus unterscheidet man *private watermarking*, bei dem durch den Vergleich einer mit einem Wasserzeichen versehnen Kopie und einer Mutterkopie festgestellt werden kann, ob ein Wasserzeichen vorhanden ist; *blind watermarking*, bei dem lediglich an Hand eines kryptographischen Schlüssels und des Mediums das Wasserzeichen extrahiert werden kann; und *semiblind watermarking*, das zusätzlich das unmarkierte Original für die Extraktion des Wasserzeichens benötigt.

Die Extraktion eines Wasserzeichens bleibt auch dann möglich, wenn das markierte Medien verändert wird. Diese Eigenschaft wird als Robustheit des Wasserzeichens bezeichnet und ist eine zentrale Voraussetzung für eine erfolgreiche Anwendung von Watermarking in realen Applikationen. Modifikationen, die dabei toleriert werden müssen reichen von der Verwendung von verlustbehafteten Kompressionsalgorithmen über Über-

---

[1] *B. Schneier,* DVD Encryption Broken, Cryptogram Newsletter 11/1999. *J. A. Bloom, I. J. Cox*, et. al., Copy Pretection for DVD Video, Proceedings of the IEEE, vol. 87, no. 7, 1999.

[2] *S. Katzenbeisser, F. A. P. Petitcolas*, Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, 2000.

tragungsfehler bis hin zu gezielten Verfälschungen, die einzig dazu dienen sollen, die Extraktion des Wasserzeichens zu erschweren oder gar zu verhindern.

2. Digitale Watermarkingprotokolle

Viele Forscher haben in der Vergangenheit auch simple Watermarkingprotokolle als ein Allheilmittel für die Wahrung der Urheberrechte gesehen. Um dies zu illustrieren, betrachte man ein einfaches Protokoll zum Schutz der Urheberrechte, das drei Personen umfasst: Alice, Bob und Carol; hierbei sei Alice der rechtmäßige Inhaber der Urheberrechte eines digitalen Objekts, Bob verletze die Urheberrechte von Alice und Carol agiere als Schiedsrichter. In diesem Fall könnte Watermarking wie folgt verwendetet werden: Alice bringt ihr Wasserzeichen an, in dem ihre Identität eindeutig festgehalten wird, weiters bringt sie ein Wasserzeichen mit der Identität von Bob an, bevor sie das derart markierte digitale Objekt an Bob verkauft (und die unmarkierte Mutterkopie bei sich sicher aufbewahrt). Sollte nun Bob das markierte Objekt illegal weitergeben und Alice auf eine so entstandene Kopie stoßen, so kann sie in Anwesenheit von Carol ihre Wasserzeichen extrahieren und eine derart durch Carol überprüfbare Anschuldigung gegen Bob erheben. Unglücklicherweise ist dieses einfache Protokoll in mehrerlei Hinsicht fehlerhaft und anfällig für eine Reihe verschiedener Angriffe:

Umkehrbarkeit. Bob hat die Möglichkeit sein eigenes Wasserzeichen dem markierten Objekt hinzuzufügen und zu behaupten, dass er der eigentliche Urheber sei. Nun wäre es leicht zu argumentieren, dass ein solcher Angriff schon deshalb scheitern müsse, weil Alices Original Bobs Wasserzeichen nicht enthalte und damit eine eindeutige Reihenfolge für die Anbringung der Wasserzeichen durch einen Schiedsrichter feststellbar sei. Jedoch haben *Craver et.al.*[3] gezeigt, dass unter bestimmten Voraussetzungen ein zweites Wasserzeichen derart eingefügt werden kann, sodass es anscheinend auch in der sicher verwahrten Originalkopie von Alice vorhanden erscheint. In einem solchen Angriff, zieht Bob sein Wasserzeichen von dem markierten Objekt ab (d.h. er markiert das Objekt mit dem Inversen seines Wasserzeichens) und behauptet es handle sich bei dem Resultat um das eigentliche Original. Bei einer Überprüfung wird nun Bobs Wasserzeichen im Original von Alice festgestellt. Es gibt keine

---

[3] *S. Craver, N. Memon, B. L. Yeo, M. M. Yeung,* Can invisible watermarks resolve rightful ownership?, 1997.

Möglichkeit festzustellen welche Kopie nun das tatsächliche Original darstellt, wenn der Watermarkingalgorithmus für solche Angriffe anfällig ist und das Inverse eines Wasserzeichens berechenbar ist. Aus diesem Grund sollten nicht-umkehrbare, auf Hash-Funktionen basierte Watermarkingverfahren verwendet werden, oder ein zentral generierter Zeitstempel im Wasserzeichen zusätzlich angebracht werden.

Öffentliche und geheime Informationen. Wenn Alice ihren Besitzanspruch öffentlich vor einem Gericht zu beweisen versucht, muss sie jenen geheimen Schlüssel bekanntgeben den sie zur Erstellung des Wasserzeichens verwendet hat. Da dieser Schlüssel Anteilung der Informationen des Wasserzeichens in einem Datenobjekt steuert, ist es oftmals mit Kenntnis des Schlüssels möglich, das Wasserzeichen zu entfernen. Deshalb werden theoretisch alle Wasserzeichen von Alice angreifbar, die den selben Schlüssel verwenden. Daher bleiben nur zwei alternative Varianten: einerseits kann Alice jede Kopie ihres Werks mit einem neuen Schlüssel markieren, bevor sie diese weitergibt; andererseits könnte ein asymmetrisches Watermarkingverfahren angewandt werden (analog zu asymmetrischen Cryptoalgorithmen), in dem das Watermark mit einem geheimen Schlüssel angebracht wird, jedoch mit einem öffentlichen Schlüssel feststellbar ist. Unglücklicherweise ist die Konstruktion solcher assymetrischen Watermarkingalgorithmen schwierig[4].

Verkäufer-Käufer Konflikte. Angenommen ein digitales Werk wird nicht direkt durch Alice verkauft, sondern über einen Zwischenhändler namens Carol. In einem solchen Fall kann Bob immer behaupten, dass nicht er, sondern Carol eine illegale Kopie in Umlauf gebracht hat und Bob fälschlicherweise als Kunden in einem Wasserzeichen identifiziert hat um vom eigenen Fehlverhalten abzulenken. In einem solchen Fall ist es für Carol unmöglich das Gegenteil zu behaupten.

Kopierangriffe. Kürzlich haben *Kutter* et al.[5] gezeigt, dass in manchen Watermarkingsystemen eine dritte Partei ein Wasserzeichen von einem markierten Objekt in ein unmarkiertes Objekt kopieren kann, ohne den geheimen Schlüssel für die Anbringung des Wasserzeichens zu kennen.

Aus diesen Gründen kann mit einem einfaches Watermarkingprotokoll alleine kein vollständig sicherer technischer Urheberrechtsschutz realisiert werden. Jedoch können eine Reihe von Anforderungen aus den oben genannten Angriffen an ein sicheres Watermarkingprotokoll abge-

---

[4] *S. Craver*, Zero Knowledge Watermark Detection, 2000.

[5] *M. Kutter*, *S. Voloshynovsky*, *A. Herrigel*, The Watermark Copy Attack, 2000.

leitet werden. Darüber hinaus muss die Skalierbarkeit des Systems und Akzeptanz sowohl beim Kunden als auch beim Händler gewährleistet werden, woraus sich eine starke Bevorzugung von dezentralen Verfahren ableitet, die nicht von einer einzigen Stelle kontrolliert werden können und nicht alle zu markierenden Dokumente einen einzelnen Flaschenhals durchlaufen müssen.

3. Dezentrales Watermarking unter Verwendung manipulationsgeschützter Hardware

Software, die im Hauptspeicher eines Computers ausgeführt wird ist anfällig für Manipulation und Beobachtung. So kann beispielsweise ein Angreifer den bei der Erzeugung eines Watermarks verwendeten geheimen Schlüssel aus dem Hauptspeicher kopieren, während das Programm ausgeführt wird. Spezielle Hardware bietet hingegen einen weitaus besseren Schutz, indem die möglichen Zugangs- und Anknüpfungspunkte des Systems begrenzt werden. In Verbindung mit sicherheitsgeprüften Kommunikationsprotokollen und zertifizierter Software können Angreifer effektiv aus dem System ausgesperrt werden. Im Verlauf der letzten Jahre wurden auf diese Art eine Reihe von manipulationsgeschützten Hardwaregeräten entwickelt und eingesetzt, wobei Smartcards für Identifikationszwecke und Finanzanwendungen sicher die weitest verbreiteten und populärsten sind.

Die Sicherheit eines manipulationsgeschützten Geräts leitet sich aus der speziellen integrierten Bauart ab, die einen Prozessor und Speicher vereint. Wird das äußere Gehäuse des Chips entsprechend gefertigt, kann der physische Zugriff auf die Datenpfade ausgeschlossen werden, da weder ein direkter Zugang zu den Speicherzellen des Speichermoduls, noch zu den elektrischen Signalen zwischen Speicher und Prozessor besteht. Jeder Zugriff auf das dermaßen geschützte Gerät erfolgt über dafür bestimmte Protokolle, die von einem kleinen Betriebssystem zur Verfügung gestellt werden und das die notwendigen Sicherheits- und Authentifikationsmechanismen zur Verfügung stellt.

Spezielle Hardware, die kryptographische Protokolle und Funktionen implementiert, kann beim heutigen Stand der Technik einfach und billig produziert werden. Um ein digitales Wasserzeichen anzubringen ist nur wenig Speicher und Rechenleistung notwendig, vorausgesetzt die Multimediaobjekte werden blockweise bearbeitet. Eine Public-Key Infrastruktur kann ebenfalls mit geringem Aufwand realisiert werden. Angesichts der fortlaufenden Forschungen auf dem Gebiet der Watermarkingalgo-

rithmen ist jedoch zu erwarten, dass ein gewisses Maß an freier Programmierbarkeit vorhanden bleiben muss, um bei Bedarf neue sicherere Verfahren zu implementieren. Betrachtet man jedoch die heutigen Kosten für 8bit Mikrocontroller, würde die notwendige Infrastruktur für digitales Watermarking auf Basis von manipulationsgeschützter Hardware einen vernachlässigbaren Kostenanteil an Informationsgeräten wie Set-Top Boxen darstellen. Weiters kann durch das Hinzufügen von Funktionen für die hardwareseitige Auslagerung von Public-Key Verschlüsselungsalgorithmen ein solcher „Watermarking Chip" zu einem idealen Co-Prozessor für e-Commerce Anwendungen werden.

3.1. Watermark Key Generation

Um sicherzustellen, dass der geheime Schlüssel für das Anbringen der Wasserzeichen nicht ausgelesen werden kann, wird eine manipulationsgeschützte Hardware derart verwendet, dass der Schlüssel die Spezialhardware niemals unverschlüsselt verlässt. Mittels asymmetrischer Verschlüsselung kann sichergestellt werden, dass jeweils nur ein Watermarking-Chip auf den Schlüssel zugreifen kann. Als sicheres Übertragungs- und Speicherformat wird ein *watermark key envelope* eingesetzt, welches den codierten Schlüssel und Information über die Identität des Inhabers des Watermarks enthält. Da diese Information nie die Hardware unverschlüsselt verlässt, kann nicht einmal der Inhaber des Watermarks diese Information manuell verändern; dieser ist auf die Validierung seiner Zugriffe durch einen zertifizierten Watermarking-Chip angewiesen[6].

3.2. Watermark Insertion

Das Einfügen eines Wasserzeichens in ein digitales Objekt erfolgt ebenfalls innerhalb der manipulationsgeschützten Hardware unter Verwendung eines nicht-invertierbaren Watermarkingverfahrens. Dabei wird das unmarkierte Objekt an den Watermarking-Chip übergeben; diese Übertragung erfolgt unter Verwendung asymmetrischer Kryptographie, um sicherzustellen, dass die übertragenen Daten auf dem Weg zur Hardware nicht ausgelesen und verändert werden können. Weiters wird ein *watermark key envelope* übertragen, welches den geheimen Schlüssel für die Anbringung des Wasserzeichens enthält. Die Hardware führt nun eine Berechtigungsprüfung durch, überprüft die Integrität der erhaltenen Daten

---

[6] *P. Tomsich, S. Katzenbeisser*, Towards a secure and de-centralized watermarking infrastructure, 2000.

und erzeugt anschließend ein mit einem Wasserzeichen versehenes digitales Objekt. Zusätzlich kann als Fingerprint die Identität des Käufers eingebracht werden.

3.3. Watermark Verification

Die Überprüfung des Wasserzeichens in einem digitalen Objekt erfolgt wiederum in der manipulationsgeschützten Hardware. Als Eingabedaten werden ein *watermark key envelope* und ein digitales Objekt benötigt. Um sicherzustellen, dass mittels des Schlüssels im *watermark key envelope* keine neuen Wasserzeichen angebracht werden können, enthält dieses auch eine Anmerkung, dass es nur für die Verifikation verwendet werden darf. Da diese Anmerkung mitverschlüsselt ist, kann sie nicht nachträglich entfernt werden, sondern wird vom Watermarking-Chip intakt empfangen. Dieser wiederum überprüft, ob der Schlüssel auch tatsächlich nur für die Verifikation eines Wasserzeichens verwendet wird. Die Ausgabe des Geräts besteht bei der Verifikation aus WAHR oder FALSCH, womit die Existenz eines bestimmten Wasserzeichens entweder bestätigt oder widerlegt wird.

3.4. Streitschlichtungsprotokoll

Streitfälle können aus zwei Gründen entstehen. Einerseits kann Bob behaupten, dass ein digitales Objekt, welches tatsächlich von Alice gestohlen wurde, sein Eigentum wäre. Andererseits kann sich ein Streitfall daraus ergeben, dass Bob abstreitet unauthorisierte Kopien eines digitalen Objekts hergestellt zu haben. Die Abarbeitung des Streitschlichtungsprotokolls umfasst mindestens drei Parteien: Alice, Bob und einen Schiedsrichter Carol. Die Anwesenheit Carols ist nur erforderlich, um Fairness zu garantieren, indem sie die Wasserzeichen in ihrem Gerät überprüft und damit etwaige Anschuldigungen, dass im Überprüfungsprozess geschummelt würde, ausschließt. Damit Carol die Verifikation der Wasserzeichen durchführen kann, fordert sie von den Streitparteien Verifikationsdaten an, die für ihr Gerät verschlüsselt wurden. Nun kann Carol überprüfen, welche Wasserzeichen in den vorgelegten Objekten vorhanden sind.

Tatsächlich versucht das Streitschlichtungsprotokoll nicht den tatsächlichen Urheber bzw. Eigentümer eines digitalen Objekts zu identifizieren, sondern dient einzig dazu eine Reihung der Ansprüche mehrerer Parteien vorzunehmen. Dies ist analog zu der Reihung von Ansprüchen auf eine Erfindung, wie sie im Patentrecht verwendet wird. Der tatsächli-

che Urheber kann nur dann ermittelt werden, wenn er auch an dem Streitschlichtungsprotokoll teilnimmt.

Um diese Reihung vorzunehmen, muss Carol auch die von den Streitparteien vorgelegten „Originale" überprüfen. Dabei können die folgenden vier Fälle auftreten:

- *Bobs Original enthält Alices Wasserzeichen, aber Alices Original enthält Bobs Wasserzeichen nicht*: In diesem Fall hat Bob offensichtlich sein Wasserzeichen eingefügt, nachdem Alice das ihre angebracht hat. Die Ansprüche von Alice sind daher älter.
- Alices Original enthält Bobs Wasserzeichen, aber Bobs Original enthält Alices Wasserzeichen nicht: Dieser Fall ist analog zum erstgenannten. In diesem Fall sind Bobs Ansprüche älter.
- *Sowohl Bobs als auch Alices Wasserzeichen sind vorhanden*: Dies ist eine klassische „Deadlock" Situation, wie sie von Umkehrungsangriffen produziert würde. Es kann keine Entscheidung getroffen werden. Dies sollte jedoch bei der Verwendung eines nicht-invertierbaren Watermarkingalgorithmus nicht auftreten.
- *Weder Bobs noch Alices Wasserzeichen sind vorhanden*: In diesem Fall kann keine Entscheidung getroffen werden. Entweder haben Bob oder Alice eine unmarkierte Originalkopie von jeweils anderen entwendet oder beide haben unabhänging voneinander ihre Wasserzeichen in ein Objekt eingefügt, das eigentlich einer dritten Person gehört.

Während es in den ersten beiden Fällen eine Streitschlichtung möglich ist, kann in den letzteren beiden keine Entscheidung gefällt werden. Jedoch sind selbst die einfachen beiden Fälle komplexer als sie erscheinen: es besteht die Möglichkeit, dass sowohl Alice als auch Bob ein Objekt vom ursprünglichen Autor kopiert haben, der sich gar nicht an diesem Protokoll beteiligt. In diesem Fall würde das „Original" noch ein weiteres Wasserzeichen enthalten, das jedoch nicht ohne Beteiligung des unbekannten Urhebers erkannt wird. Carol kann diese Möglichkeit nicht ausschließen ohne Wasserzeichen aller möglichen Parteien zu prüfen.

4. Schlussbetrachtung

Die Anbringung von digitalen Wasserzeichen stellt eine gute Möglichkeit dar um digitale Inhalte zu markieren. Aufgrund einer Vielzahl von Angriffsmöglichkeiten gegen einfache Watermarkingprotokolle wird jedoch eine komplexe Infrastruktur benötigt um sichere Aussagen über den Urheber und Rechteinhaber eines Objekts zu treffen. Eine Möglich-

keit dies dezentral zu tun besteht in der Verwendung von asymmetrischer Kryptographie und manipulationsgeschützter Hardware.