

Partielle Verschlüsselung von MPEG Audio

Martin Steinebach, Sascha Zmudzinski

Fraunhofer Institut für Integrierte Publikations- und Informationssysteme (IPSI)

{[@ipsi.fraunhofer.de">steinebach|zmudzinski](mailto:steinebach|zmudzinski)}@ipsi.fraunhofer.de

Zusammenfassung

Um Multimedia-Daten effizient zu schützen, können Verschlüsselungsstrategien verfolgt werden, die auf bekannte Eigenschaften dieser Daten Rücksicht nehmen. Oft genügt es hier, einen geringen Teil von Multimedia-Daten zu verschlüsseln, um sie vollständig zu schützen. Dabei ist der Schutz zwar geringer als bei einer vollständigen Verschlüsselung, für verschiedene Anwendungsszenarien allerdings oft ausreichend. Wir stellen ein Verfahren zum Verschlüsseln von MPEG Audio Layer 2 Dateien vor, welches sich auf die Skalenfaktoren beschränkt. Dazu liefern wir neben der technischen Umsetzung und einer Diskussion des Sicherheitsniveaus auch beispielhafte Anwendungsszenarien. In diesen wird durch unser Verfahren die Authentizität, Vertraulichkeit und Integrität geschützt.

1 Motivation

Audiomaterial liegt heute fast nur noch in digitaler Form vor, sei es im Internet, in Produktionsstudios oder in der Unterhaltungselektronik. Dies führt zu effizienten und umfangreichen Verarbeitungsmöglichkeiten, aber gerade deshalb auch zu einem Verlust an Vertrauenswürdigkeit: Jeder, der über einen Computer und die entsprechende Software verfügt, kann tiefgreifende Änderungen an dem Material vornehmen, ohne dass dabei nachvollziehbare Spuren zurückbleiben.

Multimedia-Daten können daher unterschiedlichen Sicherheitsrisiken ausgesetzt sein. Im Zusammenhang mit Audiodaten werden in erster Linie Integrität, Vertraulichkeit und Authentizität betrachtet.

- Die **Authentizität** beschreibt die Echtheit oder Glaubwürdigkeit eines Objektes, wobei hier allgemein Personen, Gegenstände oder Informationen gemeint sein können. Um die Authentizität festzustellen, wird eine Authentifikation durchgeführt. Dabei wird überprüft, ob die Identität des Objektes belegt werden kann.
- Mit **Integrität** ist die Unversehrtheit von Informationen und Daten gemeint. Solange beispielsweise eine Audiodatei nicht verändert oder manipuliert worden ist, ist die Integrität vorhanden.
- Die **Vertraulichkeit** beschreibt den Schutz von Informationen und Daten und deren Übertragung vor unberechtigtem Einblick Dritter. In einer extremen Ausprägung der Vertraulichkeit bei Daten-Übertragungen soll sogar die bloße Existenz der Übertragung für Dritte verborgen bleiben.

Allgemein werden bezüglich der Sicherheit von Informationen oder Daten über Integrität, Vertraulichkeit und Authentizität hinaus noch weitere Schutzziele betrachtet. In [E01] werden beispielsweise noch Verfügbarkeit und Verbindlichkeit aufgeführt, in [S00] Zugriffsschutz,

Nachweisbarkeit und Unleugbarkeit. Diese Aspekte werden von uns hier nicht weiter betrachtet und können durch den Mechanismus der partiellen Verschlüsselung nicht gewährleistet werden. Sie werden im allgemeinen eher von übergeordneten Systemen und Protokollen gewährleistet.

Viele Konzepte zur Durchsetzung von Schutzzielen basieren auf konventionellen Verschlüsselungsverfahren, wie beispielsweise AES, RSA oder DES in seinen verschiedenen Ausprägungen. Sie werden unterstützt von weiteren kryptografischen Verfahren, wie z.B. Hash-Funktionen, Message-Authentication etc. Wir stellen im Folgenden ein Verfahren zur Verschlüsselung für MPEG Audio vor, das im Unterschied zu konventionellen Verschlüsselungsverfahren nur Teile des zu schützenden Mediums verschlüsselt: die *partielle Verschlüsselung*. Wir zeigen, wie sich damit die Schutzziele Authentizität, Vertraulichkeit und Integrität erreichen lassen und für welche Anwendungsszenarien sie geeignet ist.

2 Grundlagen

In diesem Abschnitt beschreiben wir das allgemeine Konzept der partiellen Verschlüsselung, den Aufbau von MPEG Audio und verwandte Arbeiten zu dem Thema.

2.1 Partielle Verschlüsselung

Konventionelle Verschlüsselungsverfahren basieren auf Algorithmen, die eine Original-Datei in eine verschlüsselte Version der Datei konvertieren. Das Ergebnis des Verschlüsselungsprozesses ist dabei von einem geheimen Schlüssel abhängig. Der Inhalt der verschlüsselten Datei sieht dabei aus wie zufälliges Rauschen, ein Rückschluss auf die Original-Nachricht oder den verwendeten Schlüssel ist bei *sicheren* Verfahren in vertretbarer Zeit nicht möglich. Das nachträgliche Ändern auch nur eines Bits der verschlüsselten Datei führt dazu, dass der anschließende Entschlüsselungsprozess fehlschlägt: der Inhalt der entschlüsselten Datei ist zerstört und sieht seinerseits aus wie Rauschen. Diese Eigenschaft sicherer Verfahren ist unabhängig davon, ob es sich um ein symmetrisches oder asymmetrisches Verschlüsselungsverfahren handelt. Solche konventionellen Verfahren verschlüsseln dabei immer eine komplette Datei als Ganzes.

Die Idee der *partiellen* Verschlüsselung besteht darin, nur die relevanten Teile einer Datei einer Verschlüsselung zu unterziehen. Dadurch müssen deutlich weniger Informationen verschlüsselt werden, die Effizienz der Verschlüsselung steigt also an. Man spricht hierbei auch von *selektiver* Verschlüsselung [SeMa02] [GDS+03] [LoVS00] [SkUh02]. In einer Audiodatei würden beispielsweise die Teile verschlüsselt, die einen Einfluss auf die Klangqualität oder die Verständlichkeit der wiedergegebenen Datei besitzen. Es werden also Teile derart verschlüsselt, dass die Audiodatei nicht oder nur eingeschränkt zu verstehen oder wiederzuerkennen ist. Dabei soll jedoch die Abspielbarkeit der Audiodatei nicht beeinträchtigt werden

Ein Spezialfall der partiellen Verschlüsselung ist die *transparente* Verschlüsselung. Wichtig ist hier, dass die Qualität zwar reduziert wird, das unmodifizierte Original aber noch gut erkannt werden kann. Derart verschlüsselte Inhalte können als Preview in reduzierter Qualität öffentlich zur Verfügung gestellt werden. Einem Kunden wird anschließend der passende Schlüssel gegen Entgelt angeboten, um die Datei in voller Klangqualität zu erhalten.

Motivation für eine partielle Verschlüsselung im Vergleich mit einer vollständigen Verschlüsselung sind im Wesentlichen drei Aspekte:

- Multimedia-Daten können auch verschlüsselt noch abspielbar bleiben, da die Datei auch nach der Verschlüsselung noch das ursprüngliche Dateiformat hat. Somit werden Übertragungswege nicht gestört, die eine automatische Fehlerkontrolle beinhalten und bei herkömmlicher Verschlüsselung von einem Dateifehler ausgehen oder die Übertragung verweigern würden.
- Die Stärke der Verschlüsselung kann parametrisiert werden, um die Klangqualität gezielt in dem gewünschten Maße zu beeinträchtigen. Für Anwendungen mit Geheimhaltungsaspekten ist ein hohes Maß an Unkenntlichkeit erforderlich, während z.B. für verschlüsselte Musikstücke eine leichte bis mittelstarke Veränderung der Klangqualität ausreicht. Im ersten Fall darf die Sprachinformation nicht mehr heraushörbar sein, im Zweiten Fall geht es oft nur um eine Beeinträchtigung der Qualität. Ein idealer Algorithmus ermöglicht es, beide Anforderungen wahlweise zu erfüllen.
- Die partielle Verschlüsselung benötigt weniger Rechenleistung beim Ver- und Entschlüsseln als die vollständige Verschlüsselung, da nur ein kleiner Teil der Datei verschlüsselt wird. Dies ist z.B. im Hinblick auf Echtzeitfähigkeit von Bedeutung. Auch bei mobilen Endgeräten kann dies relevant sein, denn gerade bei tragbaren Geräten spielt der Energieverbrauch eine entscheidende Rolle.

Um die oben genannten Ziele zu erreichen, muss man identifizieren, welches die relevanten Teile einer Mediendatei im Sinne der partiellen Verschlüsselung sind. Für MPEG Audiodateien soll dies im Folgenden gezeigt werden.

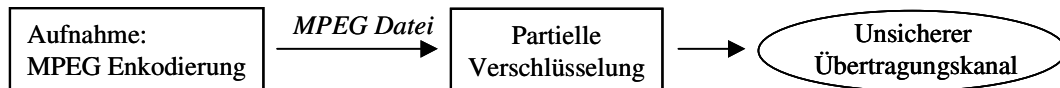


Abb. 1: partielle Verschlüsselung

Der Einsatz von partieller Verschlüsselung ist dabei jedoch nur sinnvoll, wenn sie nicht parallel zu verlustbehafteten Kompressionsverfahren eingesetzt wird. Die Komplexität solcher Kompressionsalgorithmen übersteigt die von konventionellen Verschlüsselungsverfahren bei weitem.

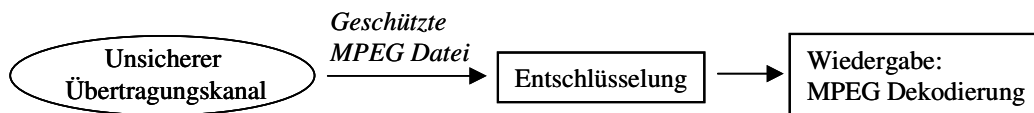


Abb. 2: Entschlüsselung

Daher fällt die Reduzierung des Aufwandes durch den Einsatz von partieller Verschlüsselung nicht ins Gewicht [SkUh02], sie ist vom Standpunkt der Ressourceneinsparung also überflüssig. Sie eignet sich vielmehr besonders für Szenarien, bei denen das Audiomaterial bereits in komprimierter Form vorliegt, beispielsweise für den Fall von MPEG Audio (siehe hierzu Abbildungen 1 und 2).

2.2 MPEG Audio

MPEG ist die Bezeichnung von Kompressionsstandards für Multimediadateien und auch Abkürzung des Organisationsnamens der Entwickler dieser Standards: der *Motion Pictures Experts Group*. MPEG Audio ist der Audioteil des MPEG-Datenstroms.

Nach dem internationalen Standard [ISO93] gibt es drei verschiedene „Layer“ des Kodierungssystems mit ansteigender Encoderkomplexität. Die Decoder sind abwärtskompatibel, somit kann ein „Layer-*n*“ Decoder auch alle Layer mit niedrigerer Ordnung decodieren. MPEG Audio Layer 3 ist unter dem Namen „MP3“ sehr bekannt. Im Rahmen dieser Arbeit stellen wir kurz den gegenüber dem Layer 3 deutlich einfacheren Layer 2 vor (MP2). Dabei sind die Prinzipien der Formate sehr ähnlich, durch flexiblere Handhabung des Datenstroms und zusätzliche Kompression wird MP3 allerdings effizienter. Das MP2-Format kommt beispielsweise bei Digital Audio Broadcast (DAB) oder Digital Video Broadcast (DVB) zum Einsatz.

Ein MP2-Datenstrom besteht aus vielen kurzen Teilabschnitten, sogenannten Frames, welche eigenständige Einheiten bilden, und in denen das Audiosignal eines bestimmten Zeitabschnitts gespeichert ist (siehe hierzu Abb. 3). Die Dauer der pro Frame gespeicherten Tonfolgen kann von Datei zu Datei unterschiedlich sein, abhängig von der Abtastfrequenz und dem verwendeten Layer. Zum Beispiel beträgt die Dauer bei Layer-2 und bei CD-Qualität 26 ms.

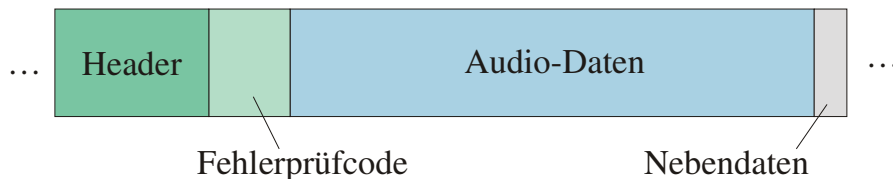


Abb. 3: Aufbau eines MP2-Frames

Im Wesentlichen besteht jeder MP2-Frame aus einem Header und den eigentlichen Audiodaten. Im Header eines jeden Frames befinden sich alle nötige Informationen um die Daten des Frames interpretieren zu können. Ein Header eines MPEG Audio Datenstroms besteht immer aus 32 Bit, unabhängig vom verwendeten Layer. Jeder Header beginnt mit einer Sequenz aus 12 gesetzten Bits zur Synchronisation, d.h. zum Erkennen eines neuen Frames.

An den Header schließen sich die sog. Audio-Daten des MP2-Frames an. Diese enthalten mehrere Datenblöcke, die jeweils eine unterschiedliche Bedeutung besitzen. Im Folgenden stellen wir kurz die Bedeutung dieser Daten-Blöcke vor:

- **Abtastwerte:** Jeder MP2-Frame beschreibt die Audio-Information im Frequenzbereich. Die Abtastwerte (*samples*) enthalten – bis auf einen Skalierungsfaktor - die Information, wie sich die Gesamtenergie auf die einzelnen Frequenzbänder verteilt. Diese stellen den größten Teil der Datenmenge eines MP2-Frames dar.
- **Skalenfaktoren:** Einzelne Gruppen von Abtastwerten müssen bei der Dekodierung mit einem gemeinsamen Skalierungsfaktor multipliziert werden, um den korrekten Wert der Energien der Frequenzbänder zu erhalten. Aus den Skalenfaktoren (*scale factors*) kann dieser Faktor errechnet werden. Ein Skalenfaktor besteht immer aus 6 aufeinanderfolgenden Bit, es ergeben sich somit 64 mögliche Skalenfaktoren. Ein Skalenfaktor darf dabei nicht den Wert 63 besitzen (binär: *111111*), da für diesen Wert keine Interpretation gemäß [ISO93] definiert ist. Der Grund hierfür ist, dass zwei aufeinanderfol-

gende 6-bit Skalenfaktoren, bei denen jeweils alle Bits gesetzt sind, beim Dekodieren sonst mit dem Beginn eines neuen Headers verwechselt werden könnten (12-bit Sync-Sequenz, siehe oben).

- Allokation: Die hier gespeicherten Daten enthalten Informationen über die Abtastwerte eines Frequenz-Unterbandes eines Tonkanals: ob Abtastwerte zu dem Frequenz-Unterband gespeichert wurden und wenn ja ob es sich um mehrere Abtastwerte oder einen Gruppen-Abtastwert handelt.
- SFAI: Jedes Frame ist in 3 Teile von je 12 Frequenz-Unterbändern unterteilt. Wenn Abtastwerte vorliegen, also die entsprechende Allokationsangabe ungleich „0“ ist, ist hier zu erkennen, wie viele Skalenfaktoren gespeichert sind und auf welchen Teil sie sich beziehen:

Jeder MP2-Frame wird beendet von den sog. Nebendaten. Diese bestehen genau aus der Anzahl Bit, welche zwischen dem Ende der Abtastwerte und dem Beginn des nächsten Headers liegen, sie werden gemäß [ISO93] nicht verwendet, jedoch wird durch diese Bit erreicht, dass ein Frame immer eine ganzzahlige Anzahl Byte hat.

2.3 Bezug zu anderen Arbeiten

Partielle Verschlüsselung für Audiodaten wird bereits in verschiedenen anderen Veröffentlichungen diskutiert. Beispielsweise in [SeTM03] wird ein Verschlüsselungs-Verfahren für MP3-Audiodateien dargestellt: Es werden hierbei Teile der sog. Side-Information der MP3-Frames verschlüsselt. Das derart geschützte Medium klingt anschließend wie eine Tiefpassgefilterte Audiodatei.

In [GARX01] wird ein weiteres Verfahren für MP3-Audiodateien vorgestellt. Es basiert auf der Vertauschung der Codewörter, Regionen und Granulen der MP3-Frames. Zusätzlich wird hier ein Verfahren zum Erzeugen von geschützten MP3-Dateien vorgestellt, das in der DCT-Domäne ausgewählte Frequenzbänder permutiert. Die Verschlüsselung ist dabei komplexer als im oben vorgestellten Verfahren, da die zu schützenden Audiodaten zuerst in die DCT-Domäne transformiert werden müssen: Eine zu schützende MP3-Datei muss also zuerst teilweise dekodiert und anschließend wieder enkodiert werden. Weitere Verfahren für Audiodaten finden sich u.a. in [THZW00] und [HeAl99].

Speziell für Sprachdaten werden Verfahren beschrieben zur partiellen Verschlüsselung von Dateien nach ITU G.729 Standard [SeMa02] oder für MPEG-4 CELP [GDS+03]. Verfahren zur partiellen Verschlüsselung sind somit für vielerlei unterschiedliche Audioformate bekannt.

Darüber hinaus findet sich eine weitaus größere Anzahl von Verfahren für MPEG Video. Beispielsweise setzt das von Kunkelmann in [Ku98] beschriebene Verfahren für Videodaten im MPEG Format auf den Koeffizienten der einzelnen Bildblöcke an. Diese werden permutiert, um eine Verschlüsselung des resultierende Bildes zu erreichen. Dabei kann in Abhängigkeit des Schutzziels die Menge und Position der Koeffizienten variiert werden. Ist eine starke partielle Verschlüsselung gewünscht, so werden Koeffizienten, welche zu niedrigen Frequenzen gehören verschlüsselt. Für eine transparente Verschlüsselung wählt man die Koeffizienten, welche hohen Frequenzen zugehörig sind. Ein weiteres Beispiel findet sich in [DiSt97].

Vielen Publikationen ist gemeinsam, dass die Autoren keine Analyse der Sicherheit ihrer Ver-

fahren gegen systematische Angriffe durchführen. Ohne derartige Untersuchungen lassen sich keine Aussagen über das gebotene Sicherheitsniveau treffen und potentielle kommerzielle Anwendungsszenarien ableiten. Über die verschiedenen kryptoanalytischen Angriffe auf partielle Verschlüsselung findet man bei [ToMo02] eine systematische Beschreibung: Unterschieden werden hierbei *statistical attacks*, *perceptual attacks* und *system-use attacks*.

Für das hier vorgestellte MP2-Verfahren werden wir weiter unten näher auf die Sicherheit eingehen und daraus denkbare Anwendungsszenarien ableiten.

3 Partielle Verschlüsselung für MP2 Audio

In diesem Kapitel stellen wir ein Verfahren zur partiellen Verschlüsselung für MP2 Audio. Dabei diskutieren wir, welche Elemente eines MP2-Frames geeignet sind für eine partielle Verschlüsselung. Daraus leiten wir ein Verfahren speziell für MP2-Skalenfaktoren ab und diskutieren anschließend dessen Sicherheit.

3.1 Auswahl relevanter Parameter

Im folgenden sollen die einzelnen Parameter eines MP2-Frames bezüglich ihrer Eignung für eine Verschlüsselung betrachtet werden. Kriterien für die Beurteilung sind die Abspielbarkeit der Datei (Interpretationsfähigkeit des Datenstroms) *nach* der partiellen Verschlüsselung, die Auswirkungen auf die akustischen Eigenschaften sowie der prozentuale Anteil des zu verschlüsselnden Teils am MP2-Frame und somit der Energie und Zeitaufwand der partiellen Verschlüsselung.

Für die ersten beiden Kriterien liefert Anhang E2 der ISO-Spezifikation [ISO93] viele Anhaltspunkte (siehe Tabelle 2).

Tab.2: Auswirkung einzelner Bitmanipulationen bei MP2 Frames

Parameter	Bit-Nummer	Auswirkungen
Bit Allokation	alle	katastrophal (nicht-abspielbar)
SFAI	alle	katastrophal (nicht-abspielbar)
Skalenfaktoren	5 (=MSB)	sehr störend
	4	sehr störend
	3	sehr störend
	2	störend
	1	wahrnehmbar, nicht störend
	0 (=LSB)	nicht wahrnehmbar
Abtastwerte	8-16 (=MSB)	sehr störend
	5-7	störend
	3,4	wahrnehmbar, nicht störend
	0-2 (=LSB)	nicht wahrnehmbar

Weitere Betrachtungen zu diesem Thema finden sich beispielsweise in [StDi03]. Wir kommen daher zu folgendem Ergebnis bzgl. der Eignung:

- Header: Werden die Header verschlüsselt, wird der Datenstrom falsch interpretiert, somit drohen Programmabsturz des Abspielprogramms sowie „Nichtabspielbarkeit“ der Datei. Eine Verschlüsselung der Header wäre nicht besonders sicher, denn mit ausreichender Kenntnis der Struktur des MP2-Datenstroms kann der Header auch relativ schnell und unkompliziert durch einfaches Probieren ermittelt werden, da die Header aller Frames einer Datei nahezu identisch sind.
- Allokation: Bei einer Verschlüsselung der Allokation würde der Datenstrom falsch interpretiert werden, alle Bitfehler in der Allokation wirken katastrophal.
- SFAI: Auch hier gilt nach einer Verschlüsselung der SFAI würde der Datenstrom nicht mehr interpretiert werden können bzw. falsch interpretiert werden und wie z.B. in [ISO93] beschrieben, wirken alle Bitfehler katastrophal.
- Skalenfaktoren: Je nach verändertem Bit wird der Fehler in [ISO93] als leicht störend bis sehr störend, aber in jedem Fall hörbar beschrieben. Da nur die Faktoren mit welchen die Abtastwerte multipliziert werden, verändert werden, bleibt die Datei unter einigen später betrachteten Bedingungen abspielbar. Da die Skalenfaktoren nur ca. 3-10% des gesamten Datenstroms bilden, jedoch einen sehr großen Einfluss auf die gesamte Datei haben, sind sie sehr gut für eine Verschlüsselung geeignet.
- Abtastwerte: Da die Abtastwerte den größten Teil der Datei bilden, ist es unter Beachtung der Zielstellung ein ressourcensparendes Verfahren zu implementieren nicht sehr sinnvoll, diese zu verschlüsseln. Möglich wäre allerdings eine Selektion innerhalb der Abtastwerte.

Aufgrund dieser Voraussetzungen haben wir ein Verfahren gewählt, dass die Skalenfaktoren eines MP2-Frames verschlüsselt. Ein Vorteil hierbei ist, dass sich die Auswirkungen auf Klangqualität und Verständlichkeit der Datei gut auf das Anwendungsszenario anpassen lassen: Je nach Zahl der verschlüsselten Bits des jeweiligen Skalenfaktors sind die Auswirkungen auf die Klangqualität „nicht störend“ oder „sehr störend“ wahrnehmbar.

3.2 Partielle Verschlüsselung der Skalenfaktoren

Um ein geeignetes Verschlüsselungsverfahren für MP2-Skalenfaktoren zu entwickeln, müssen vorher die Bedingungen für die Abspielbarkeit der verschlüsselten Datei näher untersucht werden. Wenn die Audiodatei auch nach der Verschlüsselung abspielbar sein soll, muss für die Verschlüsselung der Skalenfaktoren folgendes gelten:

- Die Länge der Skalenfaktoren (sechs Bit) darf sich durch die Verschlüsselung nicht verändern, sonst würde die Datei nicht mehr interpretiert werden können bzw. falsch interpretiert werden. Die Folgen einer falschen Interpretation können bis hin zum Absturz des Abspielprogramms oder sogar bis zum Systemabsturz reichen, je nach Abspielprogramm.
- Ein verschlüsselter Skalenfaktor darf nicht den Wert 63 annehmen (es dürfen nicht alle sechs Bit den Wert 1 haben), da dieser Wert gemäß der ISO-Spezifikation [ISO93] nicht vorkommen darf. Wie die Abspielprogramme mit solchen Werten umgehen, hängt von der jeweiligen Implementierung der Fehlerbehandlung ab. Es wäre eine interne "Korrektur" auf einen beliebigen anderen (erlaubten) Wert denkbar oder eine In-

terpolation über die vorherigen und/oder nachfolgenden Werte, bis hin zum evtl. Programm- und/oder Systemabsturz. Leider enthalten die entsprechenden MPEG-Spezifikationen [ISO93] und [ISO95] diesbezüglich keine Angaben über eine Fehlerbehandlung.

Um beide Forderungen zu erfüllen, schlagen wir ein Verfahren vor, dass die sechs Bits jedes Skalenfaktors um null bis fünf Positionen pseudo-zufällig verschiebt, bzw. rotiert (Bit-Shifting). So wird beispielsweise aus einem Skalenfaktor mit dem Wert 35 (binär: *100011*) durch Bit-Shifting um zwei Positionen nach rechts der Wert 56 (binär: *111000*) generiert etc. Der Schlüssel generiert hierbei eine Pseudo-Zufallssequenz für die Anzahl der Positionen, um die die Skalenfaktoren jeweils geshiftet werden. Jeder Skalenfaktor wird dabei um einen anderen Wert geshiftet.

Durch das gewählte Verfahren des Bitshiftings ist gewährleistet, dass ein verschlüsselter Skalenfaktor niemals den unzulässigen Wert 63 (binär: *111111*) annehmen kann: dazu müsste er bereits vor der Verschlüsselung diesen Wert besessen haben, was nach Voraussetzung ausgeschlossen ist.

Darüber hinaus kann man die Stärke der Verschlüsselung sehr elegant dem Anwendungsszenario anpassen: Wenn man nur einige der niedrigwertigen Bits des Skalenfaktors verschlüsselt, kann man weniger starke Auswirkungen auf die Klangqualität und die Verständlichkeit der Datei erreichen. So lässt sich beispielsweise eine transparente partielle Verschlüsselung erreichen (siehe oben).

Beim Entschlüsselungsprozess wird der oben beschriebene Prozess umgekehrt. Ebenso wie bei konventionellen Verschlüsselungsverfahren erhält man nach der Entschlüsselung die 1:1-Originaldatei zurück.

3.3 Sicherheit des Verfahrens

Im Folgenden gehen wir auf die Frage ein, welches Sicherheitsniveau das von uns vorgestellte Verschlüsselungsverfahren bietet. Kriterien für die Sicherheit speziell bei partiellen Verschlüsselungsverfahren kann man nach [LoVS00] wie folgt formulieren:

- Sicherheit gegenüber statistischen Angriffen
- Sicherheit gegenüber System-Use Angriffen
- Sicherheit gegenüber Perceptual Attacks

Wir erläutern diese Kriterien und untersuchen, in wieweit unser vorgestelltes Verfahren ihnen gerecht wird. Diese Bewertung wird vor dem Hintergrund der zu erreichenden Schutzziele Authentizität, Vertraulichkeit und Integrität eingeordnet.

3.3.1 Sicherheit gegenüber statistischen Angriffen

Gemeint ist hiermit, dass ein Angreifer statistische Abhängigkeiten zwischen Elementen des verschlüsselten Mediums oder Kenntnisse über typische unverschlüsselte Medien ausnutzt. Ziel des Angriffs ist es, das unverschlüsselte Medium zu rekonstruieren oder zumindest den Suchraum für eine Rekonstruktion zu reduzieren. Dieses Kriterium gilt allgemein für Angriffe auf Verschlüsselungsverfahren. Unser Verfahren basiert auf pseudozufälligem Bit-Shifting der MP2-Skalenfaktoren. Jeder dieser Skalenfaktoren besitzt eine Länge von sechs Bit. Somit sind

pro Skalenfaktor sechs verschiedene Werte für das Bit-Shifting möglich. Die Zahl möglicher Werte erhöht sich zusätzlich dadurch, dass evtl. nicht alle sechs Bits eines Skalenfaktors verschlüsselt werden, sondern nur einige der niedrigen Bits.

Die Anzahl der Skalenfaktoren innerhalb *eines* Frames ist abhängig vom MPEG-Encoder und vom jeweiligen Audiomaterial und liegt in der Größenordnung von einigen Dutzend. Die Gesamtanzahl *aller* Frames in einer MP2-Datei wiederum hängt im Wesentlichen von der Stärke der Kompression, bzw. der gewählten Bitrate ab. Jede Sekunde Audio besteht aus ca. 42 Frames.

Durch die Wahl sicherer Verfahren bei der Generierung der Pseudo-Zufallssequenz ist gewährleistet, dass jeder Skalenfaktor gleichwahrscheinlich um null bis fünf Stellen geshiftet wird. Eine Rekonstruktion des wahren Wertes eines Skalenfaktors ist somit nur durch Ausprobieren aller Kombinationen möglich. Für MP2-Audiodateien sind dies nach obigen Angaben einige hunderttausend Kombinationen pro Minute.

Allerdings kann bei realen, unverschlüsselten MP2-Dateien nicht vorausgesetzt werden, dass deren Skalenfaktoren in aufeinanderfolgenden Frames unabhängig voneinander sind. Im zu schützenden Audiomaterial bestehen in der Regel Abhängigkeiten zwischen zeitlich aufeinanderfolgenden Frames: Töne und Klänge dauern oftmals länger an als die zeitliche Dauer eines einzelnen Frames und werden u.U. auf die gleiche Art enkodiert. Insbesondere also, wenn man bei transparenter Verschlüsselung (siehe Abschnitt 2.1) nicht alle aufeinanderfolgenden Frames verschlüsselt, lassen sich u.U. die wahren Werte durch Vergleich mit zeitlich benachbarten Frames zurückschließen.

3.3.2 Sicherheit gegenüber System-Use-Angriffe

Eine Analyse der Sicherheit von Multimediadaten in verlustbehafteten Kompressionsformaten muss den Durchlauf einer Datei durch das gesamte Übertragungssystem betrachten (siehe Abbildungen 1 und 2). Die Sicherheit gegen *System-Use*-Angriffe betrachtet, wie das Verschlüsselungsverfahren mit dem zugehörigen Kompressionsverfahren kooperiert. In [LoVS00] wird die Zusammenarbeit von Kompressionsverfahren und Verschlüsselung in *kooperativ*, *neutral* oder *antagonistisch* eingeteilt, je nachdem, ob die Sicherheit durch die Enkodierung/ Dekodierung erhöht oder erniedrigt wird. Dieses Kriterium ist demnach ein Spezialfall der Forderung nach Sicherheit gegen statistische Angriffe, berücksichtigt jedoch speziellen Gegebenheiten der Komponenten des Übertragungssystems (siehe Abschnitt 3.3.1).

Für Audiodaten muss beispielsweise untersucht werden, welchen Einfluss die Wahl des jeweiligen MPEG-Enkodierers (Fraunhofer Codec, Xing etc.) auf die statistische Verteilung der Skalenfaktoren *vor* der partiellen Verschlüsselung besitzt. Ebenso könnten theoretisch *nach* der partiellen Verschlüsselung intelligente Fehlerkorrekturmechanismen des MPEG-Dekodierers (Player-Software) die Klangqualität des unverschlüsselten Originals annähernd rekonstruieren und somit die Verschlüsselung umgehen. Bei realen MP2-Dateien sind die 63 möglichen Werte der Skalenfaktoren nicht gleichwahrscheinlich sind (je nach verwendetem MP2-Enkodierer). Durch pseudozufälliges Bit-Shifting lässt sich daher nicht erreichen, dass die verschlüsselten Skalenfaktoren gleichverteilt alle 63 erlaubten Werte annehmen. Alternativ wäre hierbei (unter Vermeidung des verbotenen Werts 63) eine pseudozufällige Substitutionstabelle für jeden Frame oder eine XOR-Verknüpfung mit einer Pseudo-Zufallssequenz denkbar.

3.3.3 Sicherheit gegenüber Perceptual Attacks

Diese Klasse von Angriffen nutzt aus, dass in der Audiodatei trotz MPEG-Enkodierung noch Informationen enthalten sind, die redundant oder für die menschliche Wahrnehmung (engl. *perception*) irrelevant sind. Ein Angreifer kann eventuell aus diesen irrelevanten oder redundanten Anteilen eine Audiodatei rekonstruieren, die sich ähnlich wie das unverschlüsselte Original anhört, ohne dabei die Datei exakt zu entschlüsseln. Ein solcher Angriff müsste alle möglichen Kombinationen der Skalenfaktoren erzeugen und die somit entstandene Audiodatei müsste auf ihre (mit dem Ohr wahrnehmbare) Klangqualität überprüft werden.

Dabei ist eine Rekonstruktion des *exakten* Wertes eines Skalenfaktors u.U. nicht notwendig: Möchte man beispielsweise erreichen, dass der Inhalt einer vertraulichen, verschlüsselten Audiodatei besser zu verstehen ist, müssen „sehr störende“ Verzerrungen (im Sinne von Tabelle 2) vom Angreifer entfernt werden. Dazu genügt es, die drei höchstwertigen Bits der Skalenfaktoren zu erraten. Dies reduziert die Anzahl der Kombinationen um das Achtfache.

3.3.4 Schutz von Vertraulichkeit, Authentizität und Integrität

Wie bereits ausgeführt, ist es bei Angriffen auf die Vertraulichkeit nach den vorigen Abschnitten möglich, den Suchraum zu reduzieren. Die Verständlichkeit muss nur soweit hergestellt werden, dass der Inhalt einer Sprachdatei zu verstehen ist. Für Kommunikation mit der Forderung nach Vertraulichkeit im Hochsicherheitsbereich ist das Verfahren daher nicht geeignet.

Bei *transparenter* Verschlüsselung ist der Bruch der Vertraulichkeit, d.h. die Rekonstruktion des Originalinhaltes aufwändiger. Beim Umgehen der Verschlüsselung werden höhere Ansprüche gestellt, da es auf die Rekonstruktion der *vollen* Klangqualität ankommt. Nach Tabelle 2 kann nur die Rekonstruktion des niedrigwertigsten Bits vernachlässigt werden, da i.d.R. der Angriff erst dann erfolgreich ist, wenn die Datei mit „nicht störend wahrnehmbaren“ Artefakten rekonstruiert werden kann. Der Aufwand für *perceptual Attacks* besonders bei transparenter Verschlüsselung ist also weiterhin hoch: Für jede Sekunde Audiomaterial sind mit den Annahmen aus Abschnitt 3.3.1 einige tausend Einstellungen der Skalenfaktoren durchzuprobieren und auf ihre Klangqualität zu überprüfen.

Für den Schutz der Authentizität und Integrität mittels bei partieller Verschlüsselung gilt, dass ein Angreifer das Original nur insoweit rekonstruieren kann, als die vermeintlich entschlüsselte Datei keine wahrnehmbaren Artefakte enthält. Dies ist jedoch i.d.R. durch unterschiedliche Kombinationen von vermeintlichen Skalenfaktoren möglich. Demnach lässt sich somit nicht vollständig auf den geheimen Schlüssel zurückschließen. Einem Angreifer ist es somit anschließend nicht möglich, eine eigene Datei so zu verschlüsseln, dass sie als Original authentifiziert werden kann: Er kann sich nicht sicher sein, dass durch die Entschlüsselung mit dem wahren Schlüssel nicht Klangartefakte erzeugt werden.

Auch sind Angriffe auf die Integrität der Dateien für einen Angreifer schwierig. Dazu müsste er Teile der Datei durch eigene Inhalte ersetzen. Diese müsste er vorher mit dem geheimen Schlüssel verschlüsseln und im geschützten Medium austauschen. Da sich der Schlüssel nicht exakt bestimmen lässt, kann der Angreifer nicht sicher sein, dass auf der Empfängerseite durch die Entschlüsselung nicht evtl. zusätzliche Klangartefakte provoziert werden, die die Manipulation aufdecken.

4 Anwendungsszenarien

Wie bereits in Abschnitt 1 erwähnt, können durch die partielle Verschlüsselung verschiedene Schutzziele befriedigt werden. Wir stellen nun verschiedene Anwendungsszenarien vor, in denen der Einsatz partieller Verschlüsselung denkbar ist und Vorteile gegenüber anderen Mechanismen mit sich bringt.

Vorher wollen wir noch einmal explizit herausstellen, wie die Schutzziele durch partielle Verschlüsselung erreicht werden können:

- Die **Vertraulichkeit** ist intuitiv zu verstehen: Durch die Änderung in den Audiodateien werden die Audioinformationen beim Abspielen unverständlich. Es genügt, die Skalenfaktoren zu modifizieren, um beispielsweise gesprochenen Text unverständlich zum machen. Nur Sender und Empfänger sind in der Lage, durch einen Schlüssel die Audiodaten verständlich abzuspielen.
- Die **Authentizität** basiert darauf, dass nur durch den Besitz eines bestimmten Schlüssels die Datei so modifiziert werden kann, dass durch die Entschlüsselung mit dem passenden Gegenstück eine Datei von hoher Qualität entsteht. Erhält der Empfänger nach dem Entschlüsseln eine gestörte Datei, so kann die Datei entweder beim Transport beschädigt worden sein oder der Schlüssel passt nicht zu dem des erwarteten Senders.
- Die **Integrität** basiert auf einem der Authentizität ähnlichen Umstand. Wird eine Datei partiell verschlüsselt und ein Angreifer modifiziert diese verschlüsselte Datei, kann er ohne Kenntnis des passenden Schlüssels seine Änderungen nicht so durchführen, dass beim Entschlüsseln eine Datei von guter Qualität entsteht. Dies ist analog zum Schutz von vollständig verschlüsselten Dateien zu sehen: Auch hier kann die Chiffre zwar problemlos modifiziert werden, aber nicht auf einer Art, die nach dem Entschlüsseln eine scheinbar unversehrte Datei mit leicht verändertem Inhalt erzeugt. Die partiell verschlüsselte Datei kann also nachträglich nicht verändert werden, ohne dass sich dies nach der Entschlüsselung durch Klangartefakte bemerkbar macht.

Durch den möglichen Schutz der Vertraulichkeit und der Authentizität kann ein Mechanismus erzeugt werden, welcher das Urheberrecht schützt. So kann der Autor von Audioaufnahmen beispielsweise seine Inhalte partiell verschlüsselt übermitteln, so dass diese vom Empfänger mit dem entsprechenden Schlüssel wieder entschlüsselt werden können. Durch eine möglichst starke Verzerrung soll einem potenziellen Angreifer das Hören der Inhalte unmöglich gemacht werden. Darüber hinaus erlaubt die partielle Verschlüsselung dem Urheber, die Klangqualität und die Verständlichkeit der verschlüsselten Datei zu steuern. Dies ist sinnvoll für Anwendungsszenarien, in denen eine transparente partielle Verschlüsselung zum Einsatz kommt.

4.1 Digitaler Rundfunk

Digitaler Rundfunk im Digital Audio Broadcast Format (DAB) verwendet eine Audiokomprimierung, die mp2 entspricht [Lau96]. Dadurch können effiziente, auch beispielsweise in portablen Endgeräten einsetzbare partielle Verschlüsselungen dazu verwendet werden, gewisse Programme nur einem ausgewählten Hörerkreis zur Verfügung zu stellen. An den Übertragungswegen muss nichts geändert werden, das Programm wird vor der Übertragung ver- und nach dem Empfang entschlüsselt. Einzige Zusatzvoraussetzung ist das Synchronisieren des Schlüssels im Endgerät mit dem laufenden Programm.

Der Einsatz der partiellen Verschlüsselung ist hier wie folgt denkbar:

1. Das Audiomaterial wird entweder durch spezialisierte Hardware nach mp2 gewandelt oder liegt bereits als mp2 vor.
2. Die mp2 Dateien werden senderseitig partiell verschlüsselt
3. Die verschlüsselten Informationen werden über DAB übertragen
4. Der Empfänger entschlüsselt die mp2 Daten
5. Die entschlüsselten mp2 Daten werden wiedergegeben

Hier kann es zu Fällen kommen, bei denen ein einzelner Rechner erst nach mp2 wandelt und danach eine partielle Verschlüsselung durchführt. Somit fallen entstehende Gewinne bei der Effizienz kaum ins Gewicht. Allerdings geht es in diesem Szenario hauptsächlich um eine Verschlüsselung, die eingesetzt werden kann, ohne die DAB Systeme zu stören. Vollständig verschlüsselte mp2 Ströme würden zu Fehlern bei der Handhabung der Datenströme führen.

4.2 Media on Demand

Im Bereich Media on Demand, bekannteste Vertreter sind Video on Demand und Audio on Demand, spielen Schutzmechanismen eine große Rolle. Hier kann durch partielle Verschlüsselung eine Lösung gefunden werden, deren Komplexität auch für Endgeräte wie Handys niedrig genug ist, um empfangene Daten in Echtzeit zu entschlüsseln. Somit ist z.B. ein Video on Demand für Handys umsetzbar, der gleichzeitig ausreichend sicher und schonend für die Betriebsdauer der Akkus ist.

Im Gegensatz zum vorherigen Szenario sind hier für jeden Kunden individuell verschlüsselte Dateien sinnvoll. Jeder Kunde bekommt eine Datei zugesandt, welche sich nur mit seinem Schlüssel entschlüsseln lässt. Der Ablauf sieht dementsprechend wie folgt aus:

1. Die Datei wird bereits komprimiert auf dem System des Anbieters gelagert
2. Der Kunde fordert eine Datei an
3. Die Datei wird mit dem individuellen Schlüssel des Kunden partiell verschlüsselt und an diesen weitergeleitet
4. Der Kunde entschlüsselt die Datei und kann nun die Mediendaten konsumieren

Hier spielt die hohe Effizienz der partiellen Verschlüsselung eine große Rolle: Der Anbieter stellt seine Medien einer großen Menge von Kunden zur Verfügung, und kann durch niedrigeren Rechenaufwand beim Verschlüsseln der individuellen Kopien an notwendiger Hardware sparen: Die gleiche zur Verfügung stehende Rechenleistung reicht für deutlich mehr parallel ausgeführte Verschlüsselungsvorgänge.

4.3 Weitere Anwendungsfelder

Die oben genannten Anwendungsfelder können nur einen geringen Teil der Möglichkeiten der partiellen und transparenten Verschlüsselung aufzeigen. Wir beschreiben hier daher noch zwei weitere Anwendungen in einer geringeren Detailliertheit. Sowohl der Integritätsschutz als auch die transparente Verschlüsselung werden hier eingesetzt.

4.3.1 Previews bei der Distribution von Mediendaten

Bei der Verteilung von Medien kann transparente Verschlüsselung zum Einsatz kommen. Vorstellbar ist es, Audiodateien frei zur Verfügung zu stellen, durch transparente Verschlüsselung allerdings mit einer absichtlich reduzierten Qualität. Erst wenn der Kunde ein Entgelt bezahlt, bekommt er den Schlüssel, der die Datei in voller Qualität herstellt. Dadurch ist kein mehrmaliges Übertragen der Mediendaten notwendig. Der Kunde könnte sich sogar nach und nach von einer CD Musikstücke freischalten, die ihm gefallen.

Ein mögliches Geschäftsmodell könnte beinhalten, individuell CDs zu erstellen, auf denen sich eine Vielzahl von Musikstücken befinden, welche durch unterschiedliche Schlüssel transparent verschlüsselt würden. Ein Kunde kann dann online den zu der spezifischen CD und dem Musikstück passenden Schlüssel erwerben, um die Musik in voller Qualität hören zu können.

4.3.2 Integritätsschutz von Medienarchiven

Die partielle Verschlüsselung ermöglicht das Erkennen von Manipulationen an digitalen Medien. Werden Teile der verschlüsselten Datei nachträglich verändert, beispielsweise durch Austauschen einer Gruppe von MP3-Frames einer Sprachdatei, so gelingt die Entschlüsselung der manipulierten Frames nicht mehr fehlerfrei.

Die entschlüsselte Datei ist zwar (per Definition) weiterhin abspielbar, die manipulierten Frames sind aber durch wahrnehmbare Artefakte deutlich wahrnehmbar, da der Entschlüsselungsprozess das ausgetauschte Audiomaterial fehlerhaft rekonstruiert. Die partielle Verschlüsselung erlaubt hierbei eine (frame-) genaue Lokalisierung der Manipulationen - im Unterschied zu Schutzmechanismen, die auf der konventionellen Verschlüsselung der gesamten Datei beruhen.

Auf diese Weise lassen sich archivierte Mediendaten mit einem Schutz vor unberechtigter Manipulation versehen. Wählt man eine mittlere Stärke für die Verschlüsselung, so können die Audiodaten – bei reduzierter Klangqualität – trotzdem mit Standardsoftware abgespielt. Die geschützten Dateien können also z.B. im Arbeitsalltag eines Archivs problemlos probegehört werden.

5 Zusammenfassung und Ausblick

Wir stellen ein Verfahren zum partiellen und transparenten Verschlüsseln von mp2 Dateien vor, welches auf dem Verschieben von Bitpositionen von Skalenfaktoren beruht. Dadurch erhalten wir eine mp2-Datei, welche ohne den passenden Schlüssel in Abhängigkeit von Parametern keine verständlichen oder auch nur leicht gestörte Audioinformationen enthält.

Wir beschreiben den Einsatz dieser Verfahren für die Szenarien Digitaler Rundfunk, Media on Demand, Previews von Mediendaten und dem Integritätsschutz in Medienarchiven. Dabei

zeigt sich, dass sich der Algorithmus durch seine flexible Parametrisierung an alle Anwendungsszenarien anpassen lässt.

Der bisher prototypisch umgesetzte Algorithmus muss in einer Reihe von Tests auf Schwachstellen hin untersucht werden. Dazu sind beispielsweise weitere umfangreiche Hörtests bezüglich der Verständlichkeit verschlüsselter Informationen notwendig.

Sind die Versuche abgeschlossen, muss der Algorithmus in neue Geschäftsmodelle integriert werden. Potentielle Anwendungsszenarien müssen umgesetzt werden, um Praxiserfahrungen mit dem Verfahren zu erhalten.

Zur Erhöhung der Sicherheit des Verfahrens werden wir alternative Verschlüsselungsverfahren vorstellen. Darüber hinaus existieren bereits erste Ergebnisse dazu, das vorgestellte MP2-Verfahren auf MP3-Daten zu erweitern, um damit weitere Anwendungsfelder zu erschließen.

Danksagung

Wir danken an dieser Stelle Herrn Thorsten Boelke von der Hochschule Anhalt für die am Fraunhofer IPSI im Rahmen eines Praktikums erfolgte Implementierung des hier vorgeschlagenen Verfahrens.

Literatur

- [Buch99] Buchmann; *Einführung in die Kryptographie*, Springer, Berlin, ISBN 3-540-66059-3, 1999
- [DiSt97] Dittmann, Jana; Steinmetz, Arnd: *A Technical Approach to the Transparent Encryption of MPEG-2 Video*, in Katsikas, Sokratis (Ed.), *Communications and Multimedia Security*, Vol.3 (pp. 215-226), London, Weinheim, New York: Chapman & Hall, 1997
- [Ecke01] Claudia Eckert; *IT-Sicherheit: Konzepte - Verfahren - Protokolle*, Oldenbourg Wiss., München, ISBN 3-486-25298-4, 2001
- [GARX01] L. Gang, A.N. Akansu, M. Ramkumar und X. Xie, *Online music protection and MP3 compression*, in Proc. of Int. Symposium on Intelligent Multimedia, Video and Speech Processing, May 2001, pp. 13–16.
- [GDS+03] Gibson, J.D.; Dong H.; Servetti, A.; Gersho A.; Lookabaugh, T; de Martin, J.C.; *Selective Encryption and Scalable Speech Coding for Secure Video over Mobile ad hoc Networks*, technical report, Engineering Center, University Colorado, USA
- [HeAl99] Jürgen Herre und E. Allamanche, *Compatible scrambling of compressed audio*, Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 1999, pp. 27–30
- [ISO93] ISO/IEC 11172-1, First Edition 1993: *Information technology - Coding of moving pictures and associated audio for digital storage and media at up to about 1,5 Mbit/s, Part 3 – Audio*, 1993

- [ISO95] ISO/IEC 13818-3: 1995 (E): *Information technology – Generic coding of moving pictures and associated audio information – Part 3: Audio*
- [Kunk98] Thomas Kunkelmann, *Sicherheit für Videodaten*, Vieweg, ISBN: 3528056800, 1998
- [Laut96] Thomas Lauterbach, *Digital Audio Broadcasting*, Franzis-Verlag, 1996
- [LoVS00] Tom Lookabaugh, Indrani Vedula, Douglas C. Sicker; *Selective Encryption and MPEG-2*, 2000
- [MvOV01] Menezes, Van Oorschot, Vanstone; *Handbook of Applied Cryptography*, CRC Press, ISBN 0849385237, 2001
- [Ste00] Steinmetz; *Multimedia-Technologie. Grundlagen, Komponenten und Systeme*. ISBN: 3540673326, Springer, Heidelberg, 2000
- [Schn96] Schneier; *Angewandte Kryptographie: Protokolle, Algorithmen und Source-Code in C*, Addison-Wesley, Bonn, ISBN 3-89319-854-7, 1996
- [SeMa02] Antonio Servetti, Juan Carlos de Martin, *Perception-based selective encryption of G.729 Speech*, 2002
- [SeTM03] Antonio Servetti, Cristiano Testa, Juan Carlos de Martin, *Frequency-selective partial encryption of compressed audio*, Proceedings of IEEE ICASSP, Hong Kong, April 2003, vol. 5, pp. 668-671.
- [SkUh02] Champskud J. Skrepth, Andreas Uhl, *selective encryption of visual data*, Proc. 6th Joint Working Conference on Communications and Multimedia Security, September 26-27, 2002, Portoroz, Slovenien, pp. 213-226
- [StDi03] M. Steinebach, J. Dittmann, *Capacity-optimized mp2 audio watermarking*, Security and Watermarking of Multimedia Contents V, Santa Clara, CA, USA, Proceedings of SPIE, Edward J. Delp III, Ping Wah Wong (Eds.), Vol. 5020, pp. 44 - 54, ISBN 0-8194-4820-6, 2003
- [THZW00] N.J. Thorwirth, P. Horvatic und J. Zhao R. Weis, *Security methods for MP3 music delivery*, Proc. Asilomar Conf. on Signals, Systems and Computers, October 2000, vol. 2, pp. 1831–1835.
- [ToMo02] Torrubia, A. und Mora, F., *Perceptual Cryptography on MPEG-1 Layer III Bit-Streams*, in: *ICCE 2002 Digest of Technical Papers*, (2002), 324-325.