

Psycho-akustische Modelle für StirMark Benchmark – Modelle zur Transparenzevaluierung

Andreas Lang, Jana Dittmann, Martin Steinebach
Universität Magdeburg, Arbeitsgruppe Multimedia and Security
Fraunhofer IPSI, MERIT

Abstract: Der immer größere Einsatz robuster digitaler Wasserzeichen erfordert eine immer bessere Evaluierung bezüglich der Qualität. Die meisten Untersuchungen, wie z. B. in StirMark Benchmarking, unterziehen die Wasserzeichen verschiedenen Angriffen und überprüfen damit die Robustheit. Ein weiteres entscheidendes Qualitätskriterium ist dabei auch die Transparenz. Ein Angriff kann nur dann als erfolgreich gelten, wenn die Qualität der Daten erhalten bleibt. Eine Möglichkeit der automatischen Qualitätsevaluierung ist die Verwendung von psycho-akustischen Modellen. In diesem Beitrag werden drei Varianten der Anbindung von psycho-akustischen Modellen an StirMark for Audio vorgestellt und diskutiert. Dabei findet einmal der Vergleich des Originals mit der angegriffenen Datei statt und als Ergebnis wird eine Einschätzung der Transparenz gegeben. Bei der zweiten Variante findet eine Iteration statt, die ebenfalls einen Vergleich zwischen Original und angegriffener Datei vornimmt, dann jedoch die Angriffsparameter neu definiert, um somit eine Optimierung zwischen maximaler Angriffsstärke und minimaler Transparenz zu erhalten. Weiterhin stellen wir eine Variante vor, die bei der Durchführung des Angriffes die Angriffsparameter so setzt, dass die Qualität gewährleistet wird.

1 Motivation

Der Einsatz von robusten digitalen Wasserzeichen zum Nachweis der Urheberschaft oder zum Einbetten von Kundeninformationen in digitale Medien findet derzeit immer größere Bedeutung und Verwendung. Dabei kommen Verfahren bzw. Algorithmen zum Einsatz, die auf ganz unterschiedliche Weise Robustheit gegen übliche Medien-transformation oder gezielte Angriffe und Transparenz oder unterschiedliche Kapazitäten bieten [Dit00]. Eine Evaluierung der Verfahren findet dabei meist in kleineren Szenarien und mit unterschiedlichen Testparametern statt. StirMark Benchmark, ein Werkzeug zur Evaluierung digitaler Wasserzeichen, hat sich als Ziel gesetzt, dieses Problem zu lösen, damit eine einheitliche und nachvollziehbare Evaluierung gewährleistet wird [Pet00, SPR_01]. Hierbei verwendet StirMark eine Sammlung von Angriffen und Angriffskombinationen, derzeit für Bild [Stib] und Audio [Stia] implementiert, um das digitale Wasserzeichen zu verändern, nicht detektierbar zu machen oder gar zu löschen [KP99, SLD02].

Bereits bei der Evaluierung der von der Secure Digital Music Initiative (SDMI) vorgeschlagenen Wasserzeichenverfahren kam es zu Diskussion zwischen erfolgreichen Angreifern und den Entwicklern der Wasserzeichen. Erstere hatten erfolgreiche Angriffe entdeckt, letztere sahen die Verluste in der Qualität der Audiodaten als so hoch an, dass sie

diese nicht als erfolgreich gelten lassen wollten [SDM].

Wie in [DL03] diskutiert, ist die Transparenz ein entscheidendes Erfolgskriterium für Angriffe auf die Robustheit. Das bedeutet, dass ein Angriff auf das Wasserzeichen nur dann als erfolgreich gelten kann, wenn nicht nur die Funktion des Wasserzeichens nicht mehr gegeben ist, sondern der eigentliche Angriff die Qualität des Mediums nicht signifikant verschlechtert hat. Bisher werden deshalb nach simulierten Angriffen subjektive Tests durchgeführt [NSSP03] oder es werden Modelle herangezogen, die die Transparenz automatisch versuchen zu evaluieren [RHT97]. In unserem Beitrag konzentrieren wir uns, ausgehend von unseren Erkenntnissen aus [DL03], auf Audiomaterial. In diesem Fall dürfte ein erfolgreicher Angriff keine hörbaren Veränderungen am Datenmaterial vornehmen. Die Verwendung von psycho–akustischen Modellen [Pan95, RHT97] ermöglicht es, das Datenmaterial so zu verändern, dass es für den größten Teil der Bevölkerung nicht möglich ist zu hören. Hierbei werden Eigenschaften des menschlichen Gehörs ausgenutzt. Anwendungen der psycho–akustischen Modelle sind bei der Datenkompression [GPS] der Fall. Die Transparenzevaluierung mit automatischen Modellen hat deutliche Fortschritte gemacht. Es gibt Softwarepakete, die eine Qualitätsuntersuchung am Datenmaterial vornehmen und dann eine Entscheidung über die Qualitätsunterschiede treffen [Ope]. Die Problematik der subjektiven Hörtests ist dadurch nicht umgangen. Es sollten trotzdem weiterhin Hörtests durchgeführt werden, da u. a. das Hörempfinden bei jeder Person anders ist und eine differenzierte Betrachtung im Einzelfall noch immer nicht perfekt möglich ist.

2 Überblick: Stirmark for Audio – existierende Angriffe

In der derzeitigen Version 0.3 von StirMark for Audio (SMFA) [Stia] sind Angriffe enthalten, die den Signalverlauf eines Musikstückes direkt verändern. Ob der eigentliche Angriff dabei hörbar ist oder nicht hängt von der Stärke des Angriffs ab.

Es sind Angriffe enthalten, die sowohl im Zeitbereich als auch im Frequenzraum Manipulationen vornehmen. Eine detailliertere Beschreibung aller verfügbaren Angriffe ist auf der Internetseite von StirMark for Audio [Stia] zu finden. An dieser Stelle sollen nur auszugswise Angriffe diskutiert werden, da dies den Rahmen dieses Beitrages übersteigen würde. In der folgenden Tabelle 1 sind einige exemplarisch ausgewählte Angriffe mit Angriffsparameteranzahl und möglichen Parameterwerten zu sehen, die sich an der Hörbarkeitsschwelle befinden. Dabei sei angemerkt, dass die Parameterwerte extrem von dem Musikstück, den Lautsprechern und dem D/A-Wandler im Computer abhängen. Ebenfalls sind die Parameterwerte durch eine sehr einfache Testumgebung evaluiert worden. Die angegebenen Zahlenwerte sind derzeit noch angriffsspezifisch. Das bedeutet, dass nicht verallgemeinert gesagt werden kann, welcher der Angriffsstärke wieviel dB (dezi Bel) entsprechen.

In der Tabelle 1 bedeutet ein *evtl.* in der Spalte *Hörbar*, dass es vom Inhalt des Datenmaterials abhängt, ob der jeweilige Angriff hörbar ist oder nicht.

Wie der Tabelle 1 zu entnehmen ist, gibt es Angriffe, die das Datenmaterial verändern, ohne hörbar zu werden. Als Beispiel dafür ist der Invertierungsangriff zu nennen. Hier wird die Phase des Signals um 180° gedreht. Wenn dies in allen Kanälen geschieht, ist diese Manipulation nicht wahrnehmbar. Tests mit diesem Angriff haben auch gezeigt, dass diese

Angriff	Parameter	Hörbar	Kurzbeschreibung
Invert	keinen	nein	invertiert das Audiosignal
LSBZero		eher nicht	setzt alle LSB auf Null
Normalize		evtl.	normalisiert das Signal
Smooth		evtl.	„glättet“ das Signal
AddBrumm	Stärke von	evtl. >90	fügt Brummen hinzu
AddNoise	0 – 32768	evtl. >140	verrauscht das Signal
ExtraStereo	Stärke von	evtl. >25	erhöht Stereoanteil
ZeroLength	0 – 1000	evtl. >20	verlängert die Nulldurchgänge
CopySample	mehrere	evtl.	kopiert Samples um
FFT-HL-Pass		evtl.	Hoch-, Tiefpass im FFT Raum
FlippSample		evtl.	vertauscht Sample-Bereiche
RC-High-Pass	Frequenz	evtl.	Hochpassfilter
Resampling	Samplerate	evtl.	ändert die Samplerate

Tabelle 1: Ausgewählte Angriffe aus StirMark for Audio

triviale Form der Manipulation ein einfaches digitales Wasserzeichen verändern kann. Andere vorgestellte Angriffe wie z. B. *AddBrumm*, *CopySample*, *ExtraStereo* oder *FlippSample* können das Musikstück hörbar verändern und die Tonqualität somit verschlechtern. Entscheidend hierbei ist die Wahl der Stärke des Angriffs und das Musikstück selbst, da diese Angriffe kontextunabhängig arbeiten. Mit Kontextunabhängigkeit ist hier gemeint, dass eine Vielzahl der Angriffe unabhängig vom Inhalt der Audiodaten arbeiten. Welche Angriffe dies betrifft und wie die jeweilige Arbeitsweise ist, würde den Rahmen dieses Beitrages überschreiten. Deshalb sei an dieser Stelle wieder auf die Internetseite [Stia] verwiesen.

Hat das Musikstück eine relativ hohe Grundlautstärke, so ist die Manipulation schwerer wahrnehmbar, als wenn es ein von Grund auf leiseres Musikstück ist. Da das Musikstück vorgegeben ist, können nur noch an den Angriffsparametern Einstellungen vorgenommen werden, damit die Transparenz gewährleistet wird. Wie zuvor schon beschrieben ist dafür hauptsächlich der Parameter der Angriffstärke verantwortlich. Für die vorgestellten Beispiele bedeutet dies, dass, dass der *AddBrumm* Angriff ab einer Stärke von ≈ 90 und der Angriff *ExtraStereo* ab ≈ 25 wahrnehmbar wird. Wird die Stärke zu hoch gewählt, so ist der Angriff hörbar, die Transparenz nicht mehr gewährleistet und der Angriff als solcher nicht erfolgreich – unabhängig davon, was mit dem digitalen Wasserzeichen geschehen ist. Ist die Stärke wiederum zu schwach und ist der Angriff transparent durchgeführt, bedeutet dies, dass die Qualität des Datenmaterials nur sehr wenig verändert (verschlechtert) wurde. Das Wasserzeichen würde dann nicht ausreichend stark angegriffen sein, was zur Folge hat, dass es noch immer auslesbar oder Teile auslesbar wären. Allerdings stellt sich dann die Frage, ob und wenn ja wie stark das digitale Wasserzeichen angegriffen wurde. Ein Ziel ist es, den Stärkeparameter so einzustellen, dass der Angriff so stark wie möglich, aber trotzdem noch immer transparent ausgeführt wird. Bei den Angriffen, die mehrere Parameter zur Einstellung haben, ist nicht alleine die Stärke entscheidend, sondern auch die zusätzlichen Parameter. Als Beispiel hierfür sind der *Addbrumm*, *FlippSample* oder

CopySample zu nennen. Bei *AddBrumm* bspw. wird zusätzlich zur Angriffsstärke noch die Frequenz angegeben.

3 StirMark for Audio – Transparenz der Angriffe

Im SMFA sind eine Vielzahl von Angriffen möglich, die das Datenmaterial verändern. Dabei kann anhand von Parametern festgelegt werden, wie stark der jeweilige Angriff ausgeführt werden soll. Je stärker die Änderung am Datenmaterial, desto stärker wird das digitale Wasserzeichen angegriffen und die Qualität wird schlechter. Mit Transparenz bei den Angriffen ist gemeint, dass ein Höchstmaß an erlaubter Verschlechterung zugelassen wird. Mit anderen Worten: Durch einen Angriff auf das Datenmaterial wird ein gewisser Qualitätsverlust als zulässig erachtet (SDMI challenge). Unter Angriffen, die „nicht hörbar“ sind, ist zu verstehen, dass ein Maß an zulässiger Qualitätsverschlechterung erlaubt ist. In SMFA ist derzeit noch kein psycho-akustisches Modell enthalten, welches anhand der Veränderung des Datenmaterials entscheidet, ob der Angriff hörbar wäre oder nicht. Da nur Angriffe sinnvoll sind, die transparent sind, muss SMFA entscheiden können, ob ein Angriff zu stark sein könnte oder nicht. Das soll nicht bedeuten, dass keine subjektiven Hörtests mehr nötig sein sollen – vielmehr sollen dadurch die Angriffe mit zu stark gewählter Parametereinstellung herausgefiltert werden können. Die entscheidende Rolle dabei spielt die Wahl des psycho-akustischen Modells oder des Algorithmus, der diese Entscheidung trifft. Ebenfalls ist es essentiell, dass bei der Verwendung eines psycho-akustischen Modells eine Fourier-Transformation unumgänglich ist. Dabei wird die Audiodatei nicht als Ganzes betrachtet, sondern in viele sogenannte „Fenster“ geteilt, welche dann einzeln betrachtet und bearbeitet werden. Die Fenstergröße kann dabei variieren im Vielfachen einer Zweierpotenz. Typische Werte sind dabei 512, 1024 oder 2048.

3.1 Überblick über psycho-akustische Modelle

Es gibt verschiedene psycho-akustische Modelle [EZ01, Esk97] sowie Qualitätsbewertungen für Audio. Mit Qualitätsbewertungen sind hier Verfahren bzw. mathematische Berechnungen gemeint, die bspw. Signal-Rausch-Abstand (SNR) oder Just Noticeable Difference (JND) [Ste00] ermitteln. Hier soll angemerkt werden, dass die JND eigentlich schon zu der Psycho-akustik gezählt wird, da Hörcharakteristiken bewertet werden.

Mit diesen Verfahren ist es möglich die Qualität als solches zu ermitteln oder Hörbarkeitsschwellen zu definieren. Bei SNR erhält man einen Wert zurück, der Auskunft über das Verhältnis von Rauschen zum Nutzsignal gibt. Dabei ist natürlich die Frage zu klären, was Rauschen ist und wie es sich vom Nutzsignal unterscheidet. Das JND lässt sich sehr gut zur Ermittlung von Wahrnehmungsschwellen einsetzen. So kann man bspw. mit dem JND-Verfahren feststellen, ab wann ein Unterschied am Musikstück hörbar wäre. Allerdings ist das nicht kontextabhängig, was dazu führt, dass die ermittelten Werte nur Richtlinien sind. Beide Verfahren sind aber für den Einsatz im SMFA nur begrenzt sinnvoll, da die eigentliche Hörcharakteristik nur bedingt berücksichtigt wird. Es sollte aber als zusätzliches Entscheidungskriterium integriert werden.

Mit Hilfe psycho-akustischer Modelle ist es möglich, den Anteil eines Signals herauszufin-

den, der für den größten Teil der Bevölkerung nicht hörbar ist. Dabei ist mit „nicht hörbar“ der Anteil eines Musikstückes gemeint, der auf Grund des Inhaltes und der Maskierung durch das psycho-akustische Modell nicht wahrnehmbar ist. Die folgende Abbildung 1 wird die Maskierung des nicht hörbaren Signale verdeutlicht und schematisch dargestellt.

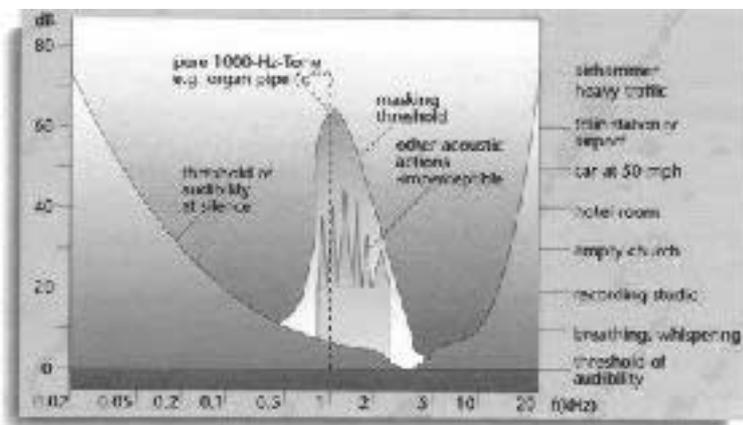


Abbildung 1: Maskierung bei psycho-akustischen Modellen [Pic]

In der Abbildung 1, welche ein Beispiel für simultane Maskierung darstellt, ist der Frequenzgang der Hörschwelle als Hauptlinie abgebildet. Der 1 kHz Ton stellt in diesem Falle den so genannten Maskierer dar. Die Linie in der Abbildung 1 „masking threshold“, welche Maskierungsschwelle genannt wird, maskiert die Signale, welche auf Grund des 1 kHz Tones nicht wahrnehmbar sind. In psycho-akustischen Modellen spielt die Zeit auch eine wichtige Rolle. Nicht nur die in der Abbildung verdeutlichten Signale werden maskiert, sondern auch über den zeitlichen Verlauf hinweg, denn durch das zeitlich versetzte Auftreten von Maskierern (in diesem Fall der 1 kHz Ton) kommt es zu Vor- und Nachverdeckungen. Details zu diesem Thema sind in der Literatur zu finden [ISO, ITUa, ITU, EZ01].

Es gibt von der ISO [ISO] ein entwickeltes psycho-akustisches Modell, welches die Wahrnehmbarkeit von Signalen beim Menschen widerspiegelt. Dabei werden die zuvor diskutierten Aspekte berücksichtigt. Praktisch umgesetzt und angewendet werden sie oft und unterschiedlich. Die folgende Liste stellt eine kurze Auswahl von Modellen dar, welche für den Einsatz in SMFA relevant sein können. Im Folgenden wollen wir die ausgewählten psycho-akustischen Modelle kurz diskutieren.

- GSPYCHO [GPS]: Ist ein Open Source Modell, welches von dem MP3 Encoder LAME (Ain't an Mp3 Encoder) [Lam] verwendet wird.
- Fraunhofer Psycho-akustisches Modell für MP3 [Psy].

- Xiph Vorbis OGG Codec [Vora]

Die Umsetzung eines psycho-akustischen Modells von Fraunhofer gehört zu den Ersten, die als Anwendungsfeld die Datenkompression nach MPEG ISO Layer 1/2/3 erachteten. Leider ist der Quelltext nicht öffentlich und somit für die Integration in SMFA nicht in Frage kommend.

Die Umsetzung, welche das GPSYCHO Modell verwendet, unterliegt der GNU LIBRARY GENERAL PUBLIC LICENSE (GLPL2) und ist somit vom Quelltext her komplett offen. Weiterhin sind die Programmierschnittstelle und eine ganze Reihe von Beispielen für diese Umsetzung im Internet zu finden, was eine leichtere und schnellere Implementierung gewährleistet.

Die Implementierung von Vorbis basiert auf einer anderen mathematischen Idee und Umsetzung [Vorb]. Da der Quelltext ebenfalls komplett offen liegt, sollte auch hier weiter untersucht werden, wie komplex eine Anbindung an SMFA erfolgen kann.

3.2 Konzepte zur Evaluierung der Transparenz

Zielsetzung ist in SMFA ein psycho-akustisches Modell zu integrieren, damit eine automatische Qualitätsüberprüfung stattfinden kann. Neben den einzelnen Angriffen auf die Robustheit soll auch die Transparenz, d. h. die Audioqualität, evaluiert werden. Damit eine modulare Integration erfolgen kann, werden drei verschiedene Ansätze gezeigt und deren Vor- und Nachteile diskutiert.

- A) Die Angriffsparameter werden anhand eines psycho-akustischen Modells vor dem Angriff maximiert und anschließend angewandt
- B) Es wird ein Vergleich der Qualität des Datenmaterials vor bzw. nach dem Angriff durchgeführt und eine Qualitätsentscheidung getroffen.
- C) Während des Angriffs werden die Angriffsparameter dynamisch eingestellt, damit eine gewünschte Qualität erhalten bleibt.

Alle Varianten A, B und C haben ihre Vor- und Nachteile. Wir wollen im Folgenden die Varianten kurz vorstellen und bezüglich des Zeitaufwandes, den Erfolg auf Transparenz sowie die automatische oder manuelle Parameterwahl eingehen.

Bei der Variante A wird ein psycho-akustisches Modell dem SMFA vorangestellt. Die Angriffsparameter und die markierte Audiodatei werden dem Modell übergeben. Anhand der Eigenschaften des Datenmaterials können der oder die Angriffsparameter maximiert werden ohne, dass die Änderung hörbar wird. In der Abbildung 2 ist schematisch die Anbindung zu sehen.

Die markierte Audiodatei wird dabei nicht als Ganzes dem psycho-akustischen Modell übergeben, sondern in mehreren kleinen Frames/Fenstern. Oftmals werden dabei 512, 1024 oder 2048 Samplewerte verwendet. Bei der Anwendung des psycho-akustischen Modells kann verglichen werden, wie laut oder wie stark ein Störsignal sein kann, so dass es

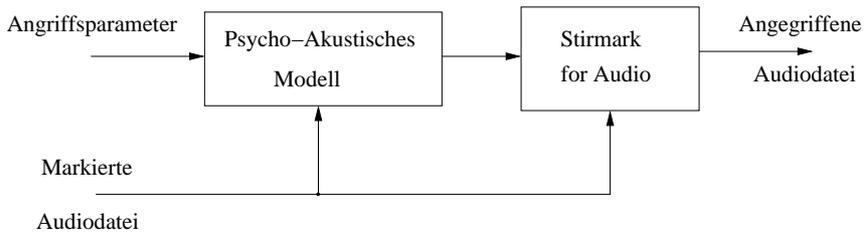


Abbildung 2: Möglichkeit A der Anbindung eines psycho-akustischen Modells an StirMark

für den Hörer transparent ist.

Bei der Variante B würde die Implementierung wie folgt aussehen: Die Eingangsdatei wird wie bisher auch dem SMFA übergeben und angegriffen. Die dabei entstehende Ausgangsdatei wird nun mit der Eingangsdatei unter Verwendung eines psycho-akustischen Modells verglichen. Das Ergebnis soll Auskunft darüber geben, ob die Unterschiede zwischen Eingangs- und Ausgangsdatei hörbar sind oder nicht. Die Entscheidung dafür trifft ein psycho-akustisches Modell. Es findet ein Vergleich der originalen- und der angegriffenen Datei statt. In dem psycho-akustischen Modell werden die Unterschiede untersucht und die Stärke der hörbaren Veränderung ausgegeben. In der Abbildung 3 ist schematisch die Anbindung des psycho-akustischen Modells an SMFA dargestellt.

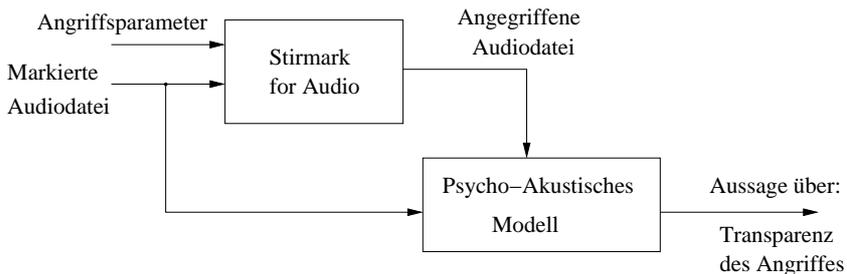


Abbildung 3: Möglichkeit B der Anbindung eines psycho-akustischen Modells an StirMark

Wenn eine Datei als eventuell hörbar eingestuft wird, sollte unbedingt ein subjektiver Hörtest folgen, da sich die Parameter der Angriffe genau an der Empfindungsschwelle des Gehörs befinden.

Bei der Variante C wird während der Durchführung des Angriffs ein psycho-akustisches Modell zur Qualitätssicherung verwendet. Dabei werden die Angriffsparameter dynamisch verändert, was zur Folge hat, dass die Stärke des Angriffs vom Datenmaterial und dem psycho-akustischen Modell abhängt. Das bedeutet, wenn das Modell nach einem Angriff angibt, dass das Ergebnis hörbar wäre, wird der Stärkeparameter des Angriffs verringert und der gesamte Angriff wiederholt. Es findet eine Iteration statt, die das digitale Wasser-

zeichen mit dem optimalen Stärkeparameter angreift. In der Abbildung 4 ist schematisch die Anbindung des psycho-akustischen Modells an SMFA dargestellt.

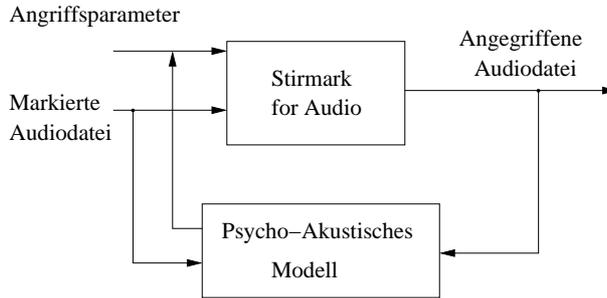


Abbildung 4: Möglichkeit C der Anbindung eines psycho-akustischen Modells an StirMark

Werden die Varianten A, B und C bzgl. Zeitaufwandes, den Erfolg auf Transparenz sowie die automatische oder manuelle Parameterwahl miteinander verglichen, so ist erkennbar, dass die Variante C wesentlich mehr Zeit in Anspruch nimmt, da der Angriff bis zum gefundenen Optimum immer wiederholt werden muss. Bei der Variante B wird der Angriff wie gewohnt durchgeführt, der Benutzer erhält danach einen Hinweis, ob der Angriff transparent durchgeführt wurde oder nicht. Zusammengefasst lässt sich sagen, dass bei Variante B und Variante C ein Qualitätsvergleich der markierten und manipulierten Datei erfolgt. Wird die Variante A verwendet, ist es möglich, innerhalb der Maskierungen des psycho-akustischen Modells den Angriff maximal durchzuführen. In der Tabelle 2 sind die Vor- und Nachteile der drei vorgestellten Varianten zusammengefasst.

	Variante A	Variante B	Variante C
Rechenaufwand	normal	normal	wesentlich höher
>Angriffsstärke, <Transparenz	möglich	nicht gegeben	möglich
optimale Angriffsparameter	automatisch	manuell	automatisch

Tabelle 2: Zusammenfassung der Variante A, B und C

Mit der Bezeichnung *normal* in der Zeile *Rechenaufwand* ist gemeint, dass die Verwendung eines psycho-akustischen Modells kaum einen zeitlichen Einfluss auf den eigentlichen Angriff hat. Lediglich die Ermittlung der Maskierungen verlängert die Angriffsdauer. Der Rechenaufwand ist ein nicht zu unterschätzender Aspekt, denn das Herausfinden des optimalen Angriffsparameters erfordert mehrere Durchläufe pro Angriff. Das heißt, der benötigte Zeitaufwand nimmt proportional zu. Dabei legt das Annäherungsverfahren an das gewünschte Ziel fest, wie schnell es gefunden wird (dabei sei an numerische Iterationsalgorithmen gedacht, die hier nicht weiter betrachtet werden sollen). Bei der Variante

A wird durch das psycho-akustische Modell die maximale Angriffsstärke ermittelt, mit welcher der Angriff durchgeführt werden soll. Der Rechenaufwand ist dabei ähnlich der von Variante B. Bei der Variante B kann der optimale Angriffsparameter nur durch manuelle Änderung gefunden werden. Allerdings sollte das Ergebnis durch subjektive Tests überprüft und dem tatsächlichen Hörempfinden angepasst werden. Bei der Variante C wird durch die Iteration der optimale Angriffsparameter ermittelt, der dem psycho-akustischen Modell entspricht.

3.3 Angriffe und psycho-akustische Modelle

Voranstellen möchten wir, dass die Anbindung eines psycho-akustischen Modells die subjektive Evaluierung nicht ersetzen kann. Aber die Angriffsparameter, die bisher willkürlich und blind gesetzt wurden, können datenmaterialabhängig optimiert werden. Dadurch kann der Angriff maximiert und die Transparenz minimiert werden. Im Folgenden wollen wir die Ansätze A, B und C hinsichtlich der Brauchbarkeit und Anwendbarkeit diskutieren.

Bei der vorgestellten Variante B können alle durchgeführten Angriffe jeweils auf Transparenz untersucht werden. Wird hingegen die Variante C angewendet, kann bei verletzter Transparenz nicht jeder Angriff erneut durchgeführt werden. Grund dafür ist, dass manche Angriffe keine Parameter besitzen. Nachfolgend sind einige ausgewählte Angriffe aufgelistet, die keinen Parameter haben.

- Invert
- LSBZero
- Normalize

Bei den Angriffen, die nur einen Parameter besitzen, hat die Anwendung des Verfahrens C den Vorteil, dass dieser eine Parameter durch die Iteration optimiert werden kann. Angriffsbeispiele hierfür sind:

- Extrastereo
- RC-High-Pass
- Resampling

Als problematisch werden die Angriffe erachtet, die bei Verwendung der Variante C, mehr als einen Parameter benötigen. Das Problem hierbei ist, dass durch die Änderung eines Parameters die Transparenz verletzt werden kann. Jedoch kann durch den zweiten oder, wenn vorhanden, dritten Parameter diese Hörbarkeit wieder korrigiert werden. Als Beispiel ist hierfür der CopySample-Angriff zu nennen. Nicht nur die Anzahl der Samples die kopiert werden, sondern auch die Länge und die Anzahl der Wiederholungen wirken sich ungünstig auf die Transparenz aus. Die folgende Aufzählung zeigt weitere Angriffe, die diese Problematik mit sich bringen.

Angriff	Variante A	Variante B	Variante C
AddBrumm	X	X	X
AddNoise	X	X	X
CopySample	–	x	x
Invert	–	X	–
ExtraStereo	x	X	X
FFT-HL-Pass	x	X	X
FlippSample	–	x	x
LSBZero	–	X	–
Normalize	–	X	–
Smooth	–	X	–
RC-High-Pass	x	X	X
Resampling	–	x	x
ZeroLength	–	x	x
...

Tabelle 3: Verwendung der Varianten A, B und C für Angriffe

- Addbrumm
- CopySample
- FlippSample

Eine weitere Problematik stellen dabei all die Angriffe dar, welche die Länge des Datenmaterials verändern. Wenn sich die Länge des Musikstückes ändert, wird der Vergleich zwischen der originalen und der angegriffenen Datei schwieriger. Bei der Anwendung des psycho-akustischen Modells muss diese Verschiebung berücksichtigt werden. Angriffe, die die Länge des Datenmaterials verändern, sind auszugsweise hier aufgelistet.

- CopySample
- ZeroLength

Ein weiterer Aspekt ist die Frage nach der Brauchbarkeit der drei vorgestellten Varianten A, B und C. In der folgenden Tabelle 3 ist zu entnehmen, welche Variante für welchen Angriff angewandt werden kann. Ein „X“ meint, dass diese Variante sich sehr gut für den entsprechenden Angriff eignet. Mit dem „x“ ist gemeint, dass diese Variante nur bedingt oder nur zum Teil angewandt werden kann. Ein „–“ drückt aus, dass diese Variante bei dem entsprechenden Angriff nicht angewandt werden kann.

Darüber hinaus wird derzeit die Datenmaterialabhängigkeit in Bezug auf die Angriffe untersucht. Dabei soll evaluiert werden, ob die Varianten A, B und C für alle Audio Charakteristika anwendbar sind.

Zur Bewertung der Auswirkungen der Angriffe auf die wahrgenommene Qualität der Audiodaten ist eine normierte Metrik notwendig. Dazu kann beispielsweise PEAQ-Norm

[ITUb] herangezogen werden, die eine einheitliche und objektive Beurteilung von Audioqualitäten verspricht. Alternativ existieren einfachere Verfahren, wie beispielsweise der Vergleich des SNR, die aber nur grobe Vergleichsmöglichkeiten bieten.

4 Zusammenfassung

Bei der Entwicklung und beim Einsatz digitaler Wasserzeichen ist eine Qualitätsevaluierung unumgänglich. Dabei werden die digitalen Wasserzeichen den verschiedensten Angriffen und Kombinationen von Angriffen unterzogen. Jeder einzelne Prozess (Einbetten, Angreifen) hat eine Qualitätsverschlechterung zur Folge. Wird die Qualität des Datenmaterials zu schlecht, zählt ein Angriff als nicht erfolgreich. Eines der wichtigsten Kriterien ist dabei die Transparenz. Damit die Frage geklärt werden kann, ob ein Angriff transparent ist oder nicht, haben wir drei Konzepte entwickelt, die psycho-akustische Modelle in StirMark for Audio (SMFA) integrieren. Die drei verschiedenen Varianten der Integration A, B und C werden derzeit in SMFA implementiert und bilden eine Grundlage für die Qualitätsevaluierung. Bei der Variante A wurde vor dem eigentlichen Angriff die Stärke unter Zuhilfenahme eines psycho-akustischen Modells ermittelt und anschließend angewandt. Bei der Variante B wurde die Ausgangsdatei mit der Eingangsdatei unter Zuhilfenahme eines psycho-akustischen Modells verglichen, ob der Angriff transparent durchgeführt wurde. In der vorgestellten Variante C findet ebenfalls ein derartiger Vergleich statt, jedoch werden bei verletzter Transparenz die Angriffsparameter neu gesetzt und der Angriff wiederholt. Alle drei Varianten können subjektive Hörtests nicht ersetzen, aber sie können dahingehend vereinfacht werden, dass nur noch die Angriffe subjektiv evaluiert werden müssen, die sehr kritisch sind bzw. die sich an der Grenze der Hörbarkeit befinden.

Literatur

- [Dit00] Jana Dittmann. *Digitale Wasserzeichen. Grundlagen, Verfahren, Anwendungsgebiete*. Springer, Berlin, April 2000. ISBN 3-540-66661-3.
- [DL03] Jana Dittmann and Andreas Lang. Psycho Akustische Modelle für Stirmark Benchmark – Transparenz von Angriffen auf Audio Wasserzeichen. 2. *Thüringer Medienseminar der FK TG, Rechte Digitaler Medien, 20.–21.05., Erfurt, Germany*, 2003.
- [Esk97] Georg Eska. *Schall und Klang. Wie und was wir hören*. Birkhäuser, 1997. ISBN 3-764-35728-2.
- [EZ01] Hugo Fastl Eberhard Zwicker. *Psychoacoustics. Facts and Models*. Springer, Berlin, 2. edition, November 2001. ISBN 3-540-65063-6.
- [GPS] GPSYCHO: psycho-acoustic and noise shaping model <http://lame.sourceforge.net/gpsycho/gpsycho.html>.
- [ISO] ISO/IEC Joint Technical Committee 1 Subcommittee 29 Working Group 11. Information technology Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s Part 3: Audio. ISO/IEC 111723, 1993. ISO, CD 11172-3, Annex D-Psychoacoustic Models.

- [ITUa] ITU-R-International Telecommunication Union – Telecommunication P. 862 – Perceptual evaluation of speech quality (PESR).
- [ITUb] ITU-R-International Telecommunication Union – Radiocommunication BS. 1378 - Methods for objective Measurements of Percieved Audio Quality (PEAQ).
- [KP99] Martin Kutter and Fabien A. P. Petitcolas. A fair benchmark for image watermarking systems. In Ping Wah Wong and Edward J. Delp, editors, *Security and Watermarking of Multimedia Contents*, volume 3657, pages 226–239, San Jose, California, U.S.A., 25–27 January 1999. The Society for Imaging Science and Technology (IS&T) and the International Society for Optical Engineering, SPIE.
- [Lam] Lame: <http://lame.sourceforge.net/>.
- [NSSP03] C. Neubauer, M. Steinebach, F. Siebenhaar, and J. Pickel. Robustness evaluation of transactional audio watermarking systems. *SPIE Security and Watermarking of Multimedia Contents V*, 5020–02, 2003.
- [Ope] OPERATM Voice/Audio Quality Analyzer: http://www.opticom.de/3_products/3opera-set.html.
- [Pan95] Davis Pan. A Tutorial on MPEG/Audio Compression. *IEEE MultiMedia*, 2(2):60–74, 1995.
- [Pet00] Fabien A. P. Petitcolas. Watermarking schemes evaluation. *IEEE signal processing*, 17 no. 5:58–64, September 2000.
- [Pic] Source of masking: <http://www.dab.lu/images/mpeg.jpg>.
- [Psy] Psycho-akustisches Modell vom Fraunhofer IIS-A: <http://www.iis.fraunhofer.de/amm/techinf/basics.html>.
- [RHT97] M. G. Ramos, S. S. Hemami, and M. A. Tamburro. Psychovisually-based multiresolution image segmentation. *IEEE image processing*, 3, October 1997.
- [SDM] SDMI challenge FAQ: <http://www.cs.princeton.edu/sip/sdmi/faq.html>.
- [SLD02] Martin Steinebach, Andreas Lang, and Jana Dittmann. StirMark Benchmark: Audio watermarking attacks based on lossy compression. *Photonics West 2002*, 4675, 19–25 January 2002.
- [SPR⁺01] M. Steinebach, F. A. P. Peticolas, F. Raynal, J. Dittman, C. Fontaine, C. Seibel, and N. Fatès. StirMark Benchmark: audio watermarking attacks. In *Multimedia Security*, Las Vegas, Nevada, U.S.A., April 2001. IEEE International Conference on Information Technology: Coding and Computing (ITCC'2001).
- [Ste00] Jonathan Y. Stein. *Digital Signal Processing*. John Wiley & Sons, September 2000. ISBN 0-471-29546-9.
- [Stia] StirMark for Audio: <http://amsl-smb.cs.uni-magdeburg.de>.
- [Stib] StirMark for Image: <http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/>.
- [Vora] Xiph OGG Homepage: <http://www.xiph.org/ogg/vorbis/>.
- [Vorb] Xiph OGG description: <http://www.xiph.org/ogg/vorbis/doc/vorbis-spec-intro.html>.