http://virtualgoods.tu-ilmenau.de/2004/

Reviewed Papers

Session 1: Watermarking for Virtual Goods

1. StirMark and profiles: from high end up to preview scenarios
   Andreas Lang, Jana Dittmann,
   1-12

2. Synchronization of Video Watermarks for Oblivious Detection after Geometrical Distortions
   Uwe Wessely
   13-23

   Martin Steinebach, Sascha Zmudzinski
   24-35

Session 2: Culture and Business for Virtual Goods

4. On-line music distribution: a case study
   Francis Rousseaux, Alain Bonardi, Romain Poncelet
   36-46

5. Secure Music Content Standard - Content Protection with CodeMeter
   Marcellus Buchheit, Rüdiger Kügler
   47-58

   Lutz Niehüser, Johannes Bräutigam
   59-71

Session 3: The Value of Virtual Goods

7. Modelling the eVerlage Payment Protocols
   Uwe Petermann
   72-83

8. How to Pay in LicenseScript
   Cheun Ngen Chong, Sandro Etalle, Pieter Hartel
   84-90

9. Personalized Previews: An Alternative Concept of Virtual Goods Marketing
   Patrick Aichroth, Stefan Puchta, Jens Hasselbach
   91-100

Session 4: Digital Protection and Digital Rights for Virtual Goods

10. Enabling Digital Content Protection on Super-Distribution Models
    Carlos Serrão, Joaquim Marques
    101-112

11. Licensing Structured Data with Ease
    Yee Wei Law, Cheun Ngen Chong, Sandro Etalle, Pieter Hartel, Ricardo Corin
    113-124

12. Interoperability Challenges for DRM Systems
    Andreas U. Schmidt, Omid Tafreschi, Ruben Wolf
    125-136
StirMark and profiles: from high end up to preview scenarios


Andreas Lang, Jana Dittmann
Otto-von-Guericke-University of Magdeburg
{andreas.lang, jana.dittmann}@iti.cs.uni-magdeburg.de

Abstract

Recently digital watermarking has become an accepted technology to ensure security requirements like authenticity and integrity of the data as well as to embed annotations into virtual goods to enhance the perceptual content. Based on the variety of existing watermarking algorithms it is unfortunately on the one hand very difficult to determine the algorithms parameter to ensure the appropriate quality of the watermarking signal on the other hand. The application framework for a specific algorithm itself is in most cases not well defined. In this paper, we introduce an approach to define audio watermarking evaluation profiles for two problems. The basic idea is to determine and define adequate audio profiles by describing typical watermarking parameters and their typical characteristics and requirements for predefined application scenarios. We discuss parameters for the embedding/detection process as well as for the evaluation of the digital audio watermarks. The selected profiles reflect most of the real world applications. To implement the profiles we use the StirMark for Audio environment, which helps to simulate the expected modifications or attacks in the different application scenarios. Besides attack profiles we also consider non-malicious modifications.

1 Motivation

Digital watermarking technology and the evaluation of digital watermarks have developed into an important research area. It is often used to annotate virtual goods with additional information or for copyright protection. The embedding process makes a small but measurable modification to the audio signal. Furthermore the evaluation process of the digital watermarking introduces small signal modification. When a customer uses the audio file to hear the music or the spoken text by using a player, amplifier and loudspeaker, then the audio signal also gets modified to a small degree. To define and to specify profiles, which reflect typical characteristics of such a process can help to classify embed/retrieve processes and to use evaluation processes. The profiles consider the watermarking parameters (especially robustness, transparency, capacity, fragility, security) and define which attacks can be used for each profile and which cannot or should not be used.

StirMark Benchmark [1] is an evaluation environment for digital audio watermarks by using specified attacks and attack parameters. In this paper we want to introduce different profiles as well as the StirMark for Audio [2] (SMBA) attacks which can be used to create application scenarios or rather profiles. As known from IT security general vulnerabilities can be classified into specification or design problems, implementation and configuration faults [14]. Based on this observation benchmarking should analyze the watermarking algorithms. But in respect to these three vulnerability classes it might become difficult for the benchmarking itself to identify which class causes the problem during the benchmarking, see for example in section 2 we discuss that issue further. Therefore one the one hand during the design of attacks for the benchmarking all three vulnerabilities classes should be a design aspect. On the other hand due to the incomplete knowledge of the algorithm itself (design, implementation and configuration) it becomes difficult
for the final benchmarking report to classify the problem into the caused vulnerability class automatically.

The least element of SMBA benchmark is a single attack. Using only a single attack is in most cases not the best way to evaluate digital robust audio watermarks. For example if a virtual good has an embedded digital audio watermark, then we have two ways that the digital watermark can be attacked:

A) The normal use of the content, e.g. playing the sound or compress it by using a lossy compression algorithm.

B) A potential attacker with the goal to attack the digital watermark by using special audio tools.

In case (B) the attacker tries each possible attack or attack combinations to destroy, remove or weaken the watermark and a lower quality is accepted in a specified range. In case (A) the “normal use” can also be interpreted as an attack against the digital watermark, e.g. the MP3 or OGG lossy compression [3] or to transmit the audio content over the Internet [4]. When the virtual goods are played using an audio player, each audio device will attack the embedded watermark in the system. E.g. the amplifier changes the energy and adds noise and distortion and the loudspeaker have a specified frequency characteristic, which affects the watermark.

To evaluate the given digital watermark against more than the basic attacks, we need profiles, which reflect the “real world attacks” or application scenarios of them. The identified profiles are useable for both, the embedding/detection process of the digital watermark as well as for the evaluation process of the watermark. In these two cases the original signal of our audio file will be manipulated or changed. The properties of this alteration can be summarized in the profiles. To specify the robustness of digital audio watermarks we offer different robustness parameters of an audio watermark. Often the following scenarios are used to evaluate the robustness of a digital robust audio watermark:

- lossy compression (mp3) with a specified bit rate [4]
- time or pitch scaling up to a defined value
- digital-analogue conversion (DA/AD) [5]

Most researchers evaluate their watermarking algorithms using one or more of these scenarios. The paper is structured as follows: based on the motivation on the research area and their importance for virtual goods by summarizing the challenges of the benchmarking suite StirMark for audio, we introduce in section two our predefined evaluation profiles: Low quality robust, Low quality fragile, High quality robust, High quality fragile, Annotation, Key space, Coalition resistance, Long time, DA/AD, Hidden communication, Calculation time, Lossy compression rates and Degree of fragility. Furthermore, these selected profiles are discussed with respect to the virtual goods music and speech. Section three contains the definition of the detailed evaluation parameter and the detailed parameter setting of the StirMark attacks. Section four describes the implementation and the paper concludes with a summary of further research problems to ensure watermarking quality in virtual goods scenarios.

2 Evaluation profiles
In this section we introduce our predefined evaluation profiles and discuss them in terms of virtual goods (music, speech). Motivated from the variety of benchmarking activities, see for example image based approaches from [1], [12] or [13] we can identify different profiles for audio, which we want to describe and to classify. Our main general classification is in respect to robustness or
fragility, and transparency by identifying high or low quality constraints by the applications.
Transparency has to be determined after the watermark embedding as well as after the benchmarking attacks.
To address further application specific properties sub-profiles become necessary for example to
determine the robustness or other parameter like security, capacity or performance in a more
precise manner. The following enumeration gives a first overview of our work combined with a
short description about the profiles, but of course it is by fare not exhaustive.

Main profiles, based on the robustness/fragility and transparency classification:

- **Low quality robust**: Is useful for scenarios, where the robustness of the embedded
  information is more important than the transparency. Examples are: preview of
  virtual goods, advertisement, telephone, Internet radio or logging function. For the
  evaluation process this means, that the evaluation or attack can be executed stronger
  to attack the digital robust watermark as hard as possible.

- **Low quality fragile**: Is useful for scenarios, where the quality is not an essential
  parameter, but the fragility to manipulations is important. Examples are the same like
  for **low quality robust**: preview of virtual goods, advertisement, telephone or logging
  function. The degree of fragility can vary depending on the application scenario;
  therefore sub-profiles become necessary.

- **High quality robust**: Reflects the high quality scenarios with an embedded robust
  audio watermark. Examples are: CD or DVD audio data, cinema application and
  concert or theatre scenarios. Transparency and robustness are most important
  parameters. In general modifications caused by watermark embedding or evaluation
  processes have impacts on the audio signal quality and the transparency has to be
  determined. If the transparency or the robustness is affected then the benchmarking
  detects a vulnerability of the watermarking algorithm.

- **High quality fragile**: Reflects the high quality scenarios for fragile digital audio
  watermarks. Examples are: CD or DVD audio data, cinema application and concert
  or theatre scenarios. The embedded information can be used to identify manipulation
  on the audio content but the transparency of the watermark is very important and has
  to be ensured. As same as for the low quality fragile profile the degree of fragility
  can vary depending on the application scenario and sub-profiles become necessary.

Sub-profiles to determine watermarking parameter in more detail:

- **Annotation**: Is useful for annotation like combining information with the content. It
  can be useful for low or high quality robust or fragile watermarks. Examples are:
  karaoke applications, affiliates programs or just additional information. The most
  important watermark parameter for this sub-profile is to evaluate the capacity to
  embed enough information into the audio content.

- **Key space**: This profile is important for applications where security is important in
  respect to attacks to the watermarking key. Similar to crypto analysis the key space
  has to be large and free from weak keys. In this profile, the used key space is
  evaluated or defined depending on the key length as well as by considering weak
  keys [6].

- **Coalition resistance**: This profile is important for customer identifications also
  called fingerprint watermarks. Due to the nature of using different watermarks on the
  same or similar content specific attacks called coalition or collusion attacks are know
  and have to be evaluated during benchmarking [9].

- **Long time**: Can be important by streaming application or very large content files.
  Examples are Internet radio, radio streams or long audio files. Important is that the
  length of the audio file can be a special case for the embed algorithm, because of
some internal variables or loops during the embedding/retrieval process. One first goal of the profile is to evaluate vulnerabilities caused by implementation like coding mistakes cause variables to overflow. A second goal is to determine the security of the watermark for example the watermark period caused by the pseudo random noise generator as general design vulnerability.

- **DA/AD**: Digital analogue conversion is useable for broadcast application, cinema or concert. The DA/AD conversion - here more specifically as over the air transmission - is on one hand a kind of attack to try to disable or destroy the watermarks and on the other hand a typical usage scenario for each virtual good. The user must perform a conversion to hear the file and benefit from the virtual goods. Many researchers use this kind of scenario to evaluate their robustness of digital robust audio watermarks [5].

- **Hidden communication**: This profile evaluates the security by searching directly for the embedded or hidden message. Statistical analysis e.g. chi-square-test [7] or RS-stegoanalyse [8] are used here to perform a decision about possible embedded information.

- **Calculation time**: An essential parameter of the profile is the performance of the embedding and detection rate by indicating ranges for embedding and detection time frames. The profile evaluates the embedding, retrieval or evaluation time (speed). This profile is useable for real time application or when an algorithm will be developed in hardware as a watermarking or evaluation device.

- **Lossy compression rates**: This profile can be used to determine the resistance or fragility to specific encoder models and to specific compression rates. Here the profile supports actual known audio encoder models like MP3, OGG, WMA or VQF [4]. The following audio compression ratios can be selected for evaluation: 8, 16, 24, 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160, 192, 224, 256 and 320 kbps as fixed bit rates or variable bit rates with quality steps or minimum and maximum bit rates. Depending on the used encoder model different fixed and variable bit rates can be used.

- **Degree of fragility**: As already indicated in the low and high quality fragile profile definition the degree of fragility can vary in different applications. Therefore this sub-profiles allows to determine this parameter in more detail. The idea is to allow the user to specify three categories. The first one is highly fragile and no bit change is allowed. The second choice is a semi-fragile category where the user can select from a set of single or combined attacks from SMBA. The third is the content-fragile category where all content-preserving transformation known from SMBA will be selected.

Based on our first classification of possible main profiles and sub-profiles it I now possible to built application oriented combinations of main and sub-profiles. Here a simple top down combination might be useful like high quality robust and DA/AD, or for hybrid watermark benchmarking a combination of high quality robust and high quality fragile profiles.

From our tests and experiences we recognized that the type of audio content is very important for the profiles. Therefore we distinguish between two main types of virtual audio goods: music and spoken text. The parameters for these types are different and have to be treated separately. To understand the differences, consider the following example.

When using the StirMark for Audio tool to evaluate digital robust audio watermarks, the attack parameter specifies how strong an attack affects the audio file. For example, the attack VoiceRemove removes the voice from an audio file by analyzing the two audio channels, identifying any monophonic audio signals and removing them. Voices are mostly on both channels
with a spoken text, while music usually has different voices on the two channels. Therefore, this attack produces different results with spoken text and music sources. Another example is the ZeroRemove attack. This attack removes samples whenever the value is 0 (zero). The result with a spoken text is that all pauses are removed and a listener perceives it as bad content. If music is used for this attack, then a hearer does not recognize it.

In terms of the profiles, we have to differ between the profile parameters for different audio content. The following figure 1 gives a short overview how the profiles and the parameters of these profiles can be divided. First, we can differ between an embedding/retrieval and evaluation process. For the embed process it is possible to use e.g. the Key space profile or the long time profile. For both embedding and evaluation, there are similar profiles available. E.g. the high and low quality profiles for robust and fragile digital watermarks as well as calculation time.

![Figure 1: Profiles and their subjection](image)

The following Table 1 summarizes how to use the profiles for embedding/retrieval or evaluation processes.

<table>
<thead>
<tr>
<th></th>
<th>Embed or Retrieval</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low quality robust</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Low quality fragile</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>High quality robust</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>High quality fragile</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Annotation</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Key space</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Coalition resistance</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Long time</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DA/AD</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Hidden communication</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Calculation time</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Lossy compression rates</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Degree of fragility</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1: Profiles for embed/retrieval or evaluation process**

The parameters of the profiles are the properties of the attacks and reflect the scenario itself and do not only depend on the profile itself. As we have identified that the type of audio content is also an important consideration in order to use the right parameter setting. This means, that the parameters must change, if the audio content varies, e.g. a spoken text or a song like rock, pop or jazz.
2.1 Profile example
In this subsection, we want to describe and discuss only one profile to improve the understanding. We select the low quality robust profile.

The low quality robust profile is a profile where the quality or transparency is not the important parameter. We make a differentiation between embed/retrieval and evaluation process:

**Embed/retrieval process:**
This profile can be used for online internet shops to offer a preview of the content or for advertisement. An Internet online shop sells virtual goods like music or video. The buyer does not buy any random media object e.g. audio files: He wants to know the song or wants to preview it. The online shop can not offer the original audio file to the buyer for preview because the buyer can copy it and wouldn’t be interested in buying it anymore. There are two simple solutions to offer the buyer a kind of preview.

- offer a short version of the original file (e.g. only a few seconds)
- offer the complete file but in lower quality

If the online shop offers a short version of the audio file, then the buyer can not hear the complete audio file and an important part could be missing. If the shop sells different versions of the same song, then a short preview audio file can not help the buyer to decide which one is the right for him.

To offer different song versions in lower quality can help the buyer find the right song. This requires a preview version in lower quality for all virtual goods in the catalogue. An addition information tag is embedded into the virtual goods by using digital watermarks with the information seller shop, artist, title etc. [9]. The embedding process reduces the quality of the audio file because of the changes made during the embedding. If the digital content is sold, then the shop can search for the files in the Internet and try do detect and retrieve the watermark to protect their rights. This watermark has to survive many processes e.g. lossy compression (e.g.mp3) to transmit the audio data [4] or attacks which try to remove or destroy the watermark. So the embedded watermark must be very robust against manipulation.
For this profile the transparency of the embedding process is not really important, because the shop offers a low quality version of there songs.

**Evaluation process:**
Using the low quality robust profile for evaluation digital robust audio watermarks, the attack parameter can set to attack very strong. This evaluation reduces the quality up to a specified level and the transparency is not very important. E.g. the attack AddNoise adds a noise signal to the marked audio file controlled by the attack parameters. If this attack is set to add “too much” noise to the marked audio file, the transparency is bad, because the noise is hear able. By using a higher attack parameter, the digital watermark is attacked more and the evaluation process works harder.

3 Evaluation parameters
In this section, we want to introduce the detailed evaluation parameter and the parameter setting for StirMark. The followed tables show the watermarking parameters (robustness, transparency, fragility, security etc.) as well as the implemented attacks in StirMark for Audio and the appraisal of the parameters and profiles. More detailed information about the attacks is available under [2]. There is also a description of the attacks and the usage of StirMark for Audio.
The selected main profiles *high quality robustness* (HQR), *high quality fragile* (HQF), *low quality robustness* (LQR) and *low quality fragile* (LQF) and their relevance for StirMark for audio attacks are shown in the table 2. Of course, the attack parameter (if needed) can specify the attack strength and transparency. We will only show which attack can be relevant and not what parameter values are possible. The parameter values depend on e.g. the audio content type. There are all single attacks listed and in the cells mean an “X” that this attack is or can be relevant for the profile.

<table>
<thead>
<tr>
<th>Relevant attacks</th>
<th>HQR</th>
<th>HQF</th>
<th>LQR</th>
<th>LQF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AddBrumm</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AddDynNoise</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>AddFFTNoise</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>AddNoise</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>AddSinus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Amplify</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>BassBoost</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Compressor</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CopySample</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CutSample</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Echo</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Exchange</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ExtraStereo</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FFT_HLPassQuick</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FFT_Invert</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FFT_RealReverse</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FFT_Stats1</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FlippSample</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Invert</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>LSBZero</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Normalize</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Nothing</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>PitchScale</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>RC-HighPass</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>RC-LowPass</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Resampling</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Smooth</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Smooth2</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Stats1</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Stats2</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TimeScale</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>VoiceRemove</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>ZeroCross</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>ZeroLength</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>ZeroRemove</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

**Table 2: Attacks and their usage for the main profiles**

Each attack is usable for the *low quality robustness* profile when the right attack parameters are used. The transparency is not very important here so the evaluation process can attack an embedded watermark very strong or the embedding process can embed the watermark very robust or with very high capacity.
For both high quality profiles it is not possible to use each attack to evaluate a digital watermark because same attacks are not transparent enough. The attacks marked with an “X” can be used in both high quality profiles, when the attack parameters are not hurt the transparency or quality. These parameters depend on the audio content and we are still in the process of characteristic dependent parameter determination.

The following itemization discusses the other watermarking parameters [10], which have to be considered in the main profiles as well as detailed by using the specific and relevant sub-profiles.

- **Fragility**: This watermark parameter is only relevant for the two profiles high quality fragile and low quality fragile. Here is the fragility itself the important parameter and the transparency differs by using the corresponding profile. If the watermark in such a profile should be evaluated, then the fragility should be determined with the sub-profile degree of fragility. A high fragile watermark breaks by changing one bit and no modification or evaluation is possible.

- **Transparency**: Is the important parameter for the high quality robust as well as high quality fragile profile. In these both profiles the transparency of the embedding/retrieval and evaluation process is very important and should not be injured.

- **Capacity**: Is relevant for the high quality robust, high quality fragile, low quality robust, low quality fragile, annotation, long time, hidden communication and calculation profiles. These profiles measure and compare the capacity of the watermarking algorithms. This profile is only useable for embedding/retrieval profiles and the important value is: how many bits per second are possible in the specific profile. In a low quality profile, more bits per second are possible as in a high quality profile, if the same embedding algorithm is used.

- **Security**: There are relevant security attacks available [11]. These attacks are relevant for the profiles coalition resistance, key space or hidden communication. Attacks to the watermarking protocol are currently not included in our observations.

- **Invertible**: An invertible watermark is a watermark, which can completely remove from the marked audio file without the original audio file [10]. This is important for using the high quality robust or high quality fragile profile. Through the embedding process the quality is lowered only a little bit – in small steps. If the watermark can be removed completely, then the original audio file is available and the highest possible quality is disposable.

- **Complexity**: This parameter is important for the calculation time profile. If an embedding/retrieval or evaluation process is complex, then it takes more time for embedding, retrieval or evaluation. The complexity itself is important for real time application and hardware producers.

From the sub-profiles the long time, hidden communication, calculation time and key space or coalition resistance profiles are not supported by our current implementation. The lossy compression profile and the degree of fragility could use existing implementations of compression models for lossy compression or existing modifications from single attacks of SMBA for the later profile.

The profile hidden communication can be integrated into StirMark for Audio but not as attack like the existing attacks. This profile does not create an output file, so it must be calculating the probability of existence of steganographic information. For this profile we want to use existing mechanisms, which are used for steganography for images like chi-square test or other statistical analysis (see for example [7]).

The profile long time is important for streaming application or very large audio files and currently a simulation is not implemented within StirMark for Audio. This profile cannot be simulated by the implemented attacks or evaluation processes. But the idea is to confront the embedding or retrieval algorithm with audio files, which are very long. For example, a typical music file in CD quality (2
channels, 44100 sample rate and 16 bit) plays about 3-5 minutes. These are about 30-53 mega bytes file size. The long time profile uses audio files, which has much more than 700 mega bytes file size or the data stream does not end (streaming). The interesting point is how the embedding or retrieval process works with such a file size. Expected errors can be segmentation fault if a variable gets an overflow or some internal variables loops or runs over their value space.

The profile calculation time is useable for real time application or when an algorithm will be developed in hardware as a watermarking, watermark detection/retrieval or evaluation device. Different watermarking algorithms can be compared by using the same audio files and same watermarking parameters (capacity, robustness, security etc.). The time during embed, retrieval or evaluation process of digital watermarks is measured for audio files with different file size and quality levels.

If the watermarking algorithm is known, then the key space profile [6] tries to find the right key to read the embedded message. By decreasing the key space this attack is able to find the right key. If the embedded message is not encrypted before or during embedding, the time is shorter than the message is embedded encrypted. The algorithm is discussed in [6]. This profile is important, because if the key can be found by a brute force search in a short time interval, the watermark algorithm is broken and not useable any more.

The coalition resistance profile could use in a first implementation a very simple attack by comparing differences from differently marked originals as described in [9].

4 Implementation

To create the profiles, we want to introduce the StirMark process and discuss them how it can be used to create the profiles. The followed Figure 2 shows the general usage of SMBA. A digital robust audio watermark will be embedded into the original audio file. Then the marked audio file can be evaluated by using SMBA. This process works with a lot of different attacks, which can be configured by using attack parameters. After this process the audio signal is modified a little bit and the watermark detector tries to detect and retrieve the watermark and gives results about the prosperous attack.

![Figure 2: Usage of StirMark for Audio](image)

One question is what the SMBA process is doing? The simple’s way is a brute force attack by using every possible attack against the watermark. StirMark has also default attack parameters, which can be used to evaluate very quickly or changed to optimize the attack strength or transparency. Each attack can be a single evaluation process with the goal to know where the watermarking algorithm is weak or with which attack the watermark can be broken. These single attacks are atomic signal modification processes. The followed Figure 3 shows the called “single attack process”. The watermarked audio file undergoes many attacks and produced for each attack a separate audio file. Each of these audio files is only modified by a single attack (e.g. add noise, change the pitch or change the amplitude). This is useful to find a plainly weakness of a watermark algorithm.
To create an evaluation profile, which reflects the real world or a scenario, which comprised more than one attacks, it is necessary to combine single attacks in the right order and with the right attack parameters. The recursive execute of attacks is our key to create the scenarios or profiles. The Figure 4 illustrates the principles. The watermarked audio file will be attacked by a single attack, which was discussed before. The result of the single attack is the input for the next attack and so on. The effect is only one audio file, which is modified in a special order of the given attacks.

The complicacy is to know which attack with which parameters are needed to get the properties needed for the profile. The general idea is that everybody can create his/her own profile or his/her specified real or virtual scenario if all parameter information is known.

5 Conclusion

This paper had envisaged 4 main profiles and 9 sub-profiles profiles for embedding, retrieving and/or evaluation of digital audio watermarks. The main profiles and sub-profiles are useable to compare the algorithms themselves by using the StirMark for Audio tool and simulate real world usage scenarios. We had showed which attacks are important to simulate the profiles with SMBA and we discussed them in general. To implement the profiles we showed firstly how StirMark works and presented how the attacks can be combined to create the profiles. In particular we have summarized that the attack and attack parameter are highly content characteristic dependent and therefore we have introduced speech and music profiles.

Beside the variety of research in image benchmarking the research for audio benchmarking needs still improvements and therefore our steps in the direction of audio profiles are an essential part. Our approach shows, that new ways to represent efficiently the characteristics of the audio watermarking algorithms are under development. Our future work is focused on further scenarios to define additional useful sub-profiles and to determine or estimate scenario dependent attack parameter.
A comparison of image profiles and audio profiles to have consistent profile definition are further essential steps in an overall benchmarking approach. Here we will concentrate on combining our approach with the image based watermarking evaluation framework WET [12]. Furthermore the classification of the vulnerability class might be of high interest. From the general behavior of the algorithms it might be possible to perform an estimation of the problem to identify design, implementation of configuration vulnerabilities. In our further research we will therefore try to recognize patterns for those estimators.

ACKNOWLEDGEMENT
Effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number FA8655-04-1-3010. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

Furthermore basic ideas presented here come from the previous works done in StirMark and we would like to thank all involved persons for their stimulating discussions and their work. Furthermore we would like to thank our student Ryan Spring for his help in structuring and correcting the paper.

References

