Transparency Benchmarking on Audio Watermarks and Steganography

Christian Kraetzer[a], Jana Dittmann[a] and Andreas Lang[a]

[a]Research Group Multimedia and Security, Department of Computer Science,
Otto-von-Guericke-University of Magdeburg, Universitaetsplatz 2, 39106 Magdeburg, Germany

ABSTRACT

Abstract: The evaluation of transparency plays an important role in the context of watermarking and steganography algorithms. This paper introduces a general definition of the term transparency in the context of steganography, digital watermarking and attack based evaluation of digital watermarking algorithms. For this purpose the term transparency is first considered individually for each of the three application fields (steganography, digital watermarking and watermarking algorithm evaluation). From the three results a general definition for the overall context is derived in a second step. The relevance and applicability of the definition given is evaluated in practise using existing audio watermarking and steganography algorithms (which work in time, frequency and wavelet domain) as well as an attack based evaluation suite for audio watermarking benchmarking - StirMark for Audio (SMBA). For this purpose selected attacks from the SMBA suite are modified by adding transparency enhancing measures using a psychoacoustic model. The transparency and robustness of the evaluated audio watermarking algorithms by using the original and modified attacks are compared. The results of this paper show that transparency benchmarking will lead to new information regarding the algorithms under observation and their usage. This information can result in concrete recommendations for modification, like the ones resulting from the tests performed here.

1. MOTIVATION AND INTRODUCTION

According to Fridrich[1] data hiding is a highly multidisciplinary field that combines image and signal processing with cryptography, communication theory, coding theory, signal compression, and the theory of perception. This holds true for the domain of data hiding in images and can be transferred into the audio domain, which is considered in this paper. Steganographic methods and digital watermarking algorithms, which are distinguishable practical applications of data hiding, have different characteristics by which they can be described. Among those characteristics transparency is one of the most important ones since it has a strong impact on the usability and acceptance of an algorithm and therefore on a possible usage in a commercial application. This has especially to be considered for watermarking algorithms where different thresholds of perceptibility can be desired in different fields of application. The importance of transparency in this context implies that it has to be one of the main goals in algorithm benchmarking. Since the transparency is depending on other characteristics of an steganographic or watermarking algorithm (like the robustness desired or the capacity used) and external factors (like the content used for embedding) a fair benchmarking approach has to consider these too. In this paper a notation addressing this problem is introduced and discussed. The applicability of the notation is tested for transparency benchmarking of implemented steganographic and digital watermarking algorithms as well as the attack based benchmarking approach of SMBA.

This paper is organised as follows: Section 2 introduces function based transparency definitions for the transparency in the contexts of steganography, digital watermarking and attack based evaluation. From those definitions an universal formal working definition is derived which is used in this paper. Section 3 describes the evaluation suite SMBA and its approach of psychoacoustic modelling which will be used in the evaluations. In section 4 the complete test scenario consisting of test objectives, test set-up and the test procedure is introduced. Section 5 gives the results of the tests performed. The paper closes in section 6 with a summary and an outlook on future work.

2. TRANSPARENCY DEFINITION

In literature various approaches to the aspect of transparency in data hiding can be found. Cox et al.[2] proposes a rating of the performance of a watermarking system considering the terms of fidelity and quality. Fidelity is in his works defined as the perceptual similarity between the unmarked and watermarked works at the point at which they are presented to a consumer (i.e. after all possible degenerations through transmission processes). Quality is an absolute measure of appeal. An example given is that of an high-quality image or audio clip, which simply looks or sounds good and therefore has no obvious processing artifacts.[2]
In a different approach Fridrich ([1] and[3]) and Dittmann[4] grade the performance of a watermarking algorithm by considering (among others) the two aspects of undetectability (embedded information is undetectable if the image with the embedded message is consistent with a model of the source from images are drawn, i.e. the concept of undetectability is inherently tied to the statistical model of the image source) and invisibility (perceptual transparency; information is embedded in a way such that the average human subject is unable to distinguish between carriers that do contain hidden information and those that do not).
Considering only the three characteristics capacity, transparency, and robustness of a data hiding method, it is obvious that there is a trade-off between these three characteristics. No algorithm can provide maximum capacity and maximum transparency at the same time. This principle is shown in figure 1.

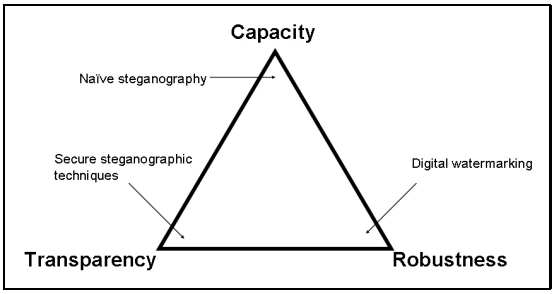


Figure 1: Trade-off between capacity, transparency, and robustness.[3]

In the corners of the triangle shown in figure 1 can be found the ideal positions for secure steganographic techniques, naïve steganography and digital watermarking. Based on this approach (and focusing only on the three characteristics capacity, transparency, and robustness) the meaning of transparency in three different contexts (steganography, digital watermarking and the evaluation of digital watermarking techniques using attacks) is discussed in the following sections and a formal definition of the transparency term is given which can be applied for all three approaches. Furthermore two other problems to be considered (which are indicated by Kutter et al.[5]) are addressed: first, a reliable assessment of the introduced distortion has to be provided (either by subjective tests or by a quality metric), and second, in application, where operations like the scaling of an image might increase the perceptibility of the watermark, the perceptibility has to be evaluated after each operation.

2.1. Transparency in the context of steganography

As indicated by the approach from Fridrich introduced above the transparency and capacity are the two characteristics most important to steganographic methods. In the focus of this paper the secure steganographic techniques (i.e. high transparency) are more relevant than the naïve steganography (high capacity). The later is neglected in the following considerations. In general the transparency T of steganographic algorithms can be described as a function of the cover O (context dependency), the steganographic algorithm and its set of parameters A_S, the message embedded M, the used capacity C (in this case in bit/s), the transparency determination technique used T_M (from which the unit is derived) and the alteration of the statistical behaviour of the material associated with the embedding D. The function is given as equation: $T = f(O, A_S, M, C, T_M, D)$. This paper determines for given implementations of steganographic algorithms how different parameterisations and the content on which the algorithm is run influence the transparency.

2.2. Transparency in the context of digital watermarking

Digital watermarking algorithms have primary to be evaluated with respect to their robustness but transparency evaluations are also crucial in this field since different degrees of transparency can be desired for a watermark, depending on the intended application, to ensure a certain quality level. This is the reason why sophisticated digital watermarking algorithms exploit in one way or another the effects of perceptual modelling.[2]
The trade-off between watermark robustness and watermark perceptibility already mentioned above is revisited in Kutter et al.[5] where it is stated that for a fair benchmarking and performance evaluation one has to ensure that the methods under investigation are tested under comparable conditions (i.e. under consideration of robustness and transparency). Considering the trade-off between the characteristics a high embedding strength does not only increase the robustness of the watermarking algorithm but has a negative impact on its transparency. In general the transparency T of a watermarking algorithm can be described as a function of the cover O (context dependency), the watermarking algorithm and its set of parameters A_W, the watermark embedded M, the used capacity C, the robustness of the algorithm R, T_M and D which are the same as in steganography. The function is given as equation: $T = f(O, A_W, M, C, R, T_M, D)$. This paper determines for given implementations of watermarking algorithms how different parameterisations (i.e. the increasing of the embedding strength) influence the transparency.

2.3. Transparency in the context of attack based watermark evaluation

Attacks on digital watermarks known from watermarking benchmarking can only be considered successful if the embedded watermark is destroyed and the modification made in the attack process is transparent. This means that after the evaluation of an algorithm statements concerning the detection of the watermark and the transparency of the modification have to be provided. This is investigated for given algorithms in this paper. In general the transparency T of an attack based watermarking evaluation can be described as a function of the cover (or covers) used O (context dependency), the set of used evaluation methods (attacks) of the evaluation suite and their parameters E, the transparency determination technique used T_M, as well as the alteration of the statistical behaviour of the material associated with the embedding D. The function is given as equation: $T = f(O, E, T_M, D)$.

2.4. Formal working definition of the transparency term

For our observations the transparency of an algorithm is considered to be a function with seven arguments. This function is derived from the function for digital watermarking introduced above since the transparency evaluation on digital watermarks is the most complex of the three contexts considered in this work. The transparency T of a watermarking algorithm is here described as a function of the cover O (context dependency), the used algorithms and methods as well as their sets of parameters A, the information/message embedded M, the used capacity C (in this case in bit/s), the robustness of the algorithm R, the transparency determination technique used T_M (subjective and objective) and the alteration of the statistical behaviour of the material associated with the embedding D. Even if perceptual transparency and statistical (steganalytical) transparency are both parts of the overall complexity term they have to be considered independently (the focus of this paper is set on perceptual transparency; steganalytical transparency is neglected here and might propose a good topic for future research). Additionally in this function a parameter T_T for the desired transparency threshold is introduced. This parameter allows addressing fields of application where the perception of the modification is intended (e.g. in applications like visual/perceptual watermarks). The function is given as equation: $T = f(O, A, M, C, R, T_M, D, T_T)$

3. EVALUATION SUITE SMBA AND PSYCHOACOUSTIC MODELLING

The SMBA evaluation suite and its evaluation approach based on attacks are very well described in a large number of publications (e.g. Lang et al.[6]) and are therefore not introduced in detail here. For our investigations on transparency in the context of steganography we use the basic profile[7] P_B-Transparency. For the evaluation of transparency in the case of digital watermarking we use P_B-Transparency and P_B-Robustness. For the research on attack based watermark evaluation we use P_B-Transparency, P_B-Robustness and the approaches of perceptual attack tuning described below.

3.1. Basic Profile Transparency (P_B-Transparency)

In general two common approaches for the evaluation of modifications on audio material (like the attacks of SMBA) exist. The first one is the evaluation with subjective tests (listening tests). This method is very time consuming and requires many human test subjects. The second approach is the use of so called objective perceptual measurement techniques. As the measure of choice the Objective Difference Grade (ODG; as described by Kabal[8]) was chosen, because it is considered to be the only measure directly verifiable against listening test data determining the SDG (Subjective Difference Grade[9] as described by Thiede[10]). The objective perceptual measurement techniques do not have the restrictions of the subjective tests with an audience, but on the other hand still lack acceptance. The reason for this fact is simple: until a model is found which is capable of simulating all the phenomena of the human hearing satisfactorily, objective measures are to be considered error prone. Nevertheless they are a good indicator, which has to be supported by tests with a human auditory, if necessary.
For the evaluations required for this paper the usage a objective measurement technique (ODG in this case) is chosen as the transparency determination technique T_M due to the large set of test files under evaluation. For the computation of the ODG values the open source software tool EAQUAL[11] is used. The ODG values computed by EAQUAL are in the range [0,−4] (imperceptible to very annoying).

3.2. Basic Profile Robustness (P_B-Robustness)

The basic profile robustness (P_B-Robustness) uses 40 attacks provided by SMBA.[12] After the attacking process, the watermarking algorithm tries to detect the watermark information. For our current application we classify the robustness of watermarking algorithms into three distinguishable classes: robust, fragile and non-classifiable. In our paper we define that a watermarking algorithm A_W is considered robust against an attack e_i ($e_i \in E$) if in less then 10% of all marked files the watermark embedded is not detectable. A_W is considered fragile against the attack e_i if in more then 90% of all marked files the message or watermark embedded is not detectable. The robustness of A_W against an attack e_i is considered non-classifiable if the percentage of successful attacks lies between 10% and 90% of all audio files. The results for each of the algorithms A_W considered are given in the form: (number of e_i against which A_W is robust / number of e_i where the robustness of A_W is considered non-classifiable / number of e_i against which A_W is fragile). An example for this form would be (3/27/10) which would indicate that the corresponding algorithm A_W is considered robust against three attacks, in 10 cases A_W is fragile and no definitive answer can be given for 27 attacks. Considering the test set O of 389 files the thresholds of 10% and 90% are equivalent to 39 and 350 files respectively.
As discussed by Kraetzer[13] out of the attack set of SMBA a subset of three attacks (AddBrumm, BassBoost and AddSinus) is chosen from the list of perceptual tuneable attacks. They are examined in two different fields of application (adjusting single attack parameters and multi-parameterisation of attacks).

3.3. Approaches of Perceptual Attack Tuning

Basics concerning the usage of psychoacoustics in audio watermarking and of a model of the HVS (human visual system) in image watermarking, like masking, pooling and the description of basic perceptual models, can be found in Cox et al.[2] where a perceptual model is generally addressed as a function, that gives a measure of the maximum acceptable distance between the original work, and the watermarked work. The idea of perceptual attack tuning is based on using a perceptual model as a distance threshold. The following three ways for integration of a psychoacoustic model into SMBA are introduced by Dittmann et al.[14]:
Pre attack alignment: Uses a psychoacoustic model to pre-compute the maximal strength of the attack, before the attack itself is launched.
Post attack alignment: After an attack the data is compared by a psychoacoustic model to a copy of the original. The psychoacoustic model is not used directly to calibrate or influence the attack, it only makes quality assessments.
Simultaneous/Iterative alignment: While the attack is running the attack parameters are adjusted (context aware) by the psychoacoustic model to guarantee the quality of the data. The psychoacoustic module used in this case, evaluates the quality after an attack (like in item 2 above). If the attack is considered not successful (i.e. audible) it re-launches the attack with the parameters set to a lower level. This process runs in an iteration

loop until the psychoacoustic model considers the attack to be successful. This method is, due to the iterations, the most time and computation power consuming of the three.
Currently SMBA supports the first approach with an implementation based on the works of Zwicker et al.[15] to be used in our tests.

4. TEST SCENARIO

In this section the complete test scenario is introduced. It consists of the test objectives, the test set-up (including the test files used O, the steganography algorithms A_S, watermarking algorithms A_W, the set of attacks (including the perceptually modified attacks) E, the parameters and mechanisms used in the evaluation (T_M)) and the test procedure.

4.1. Test Objectives

In this paper a basic transparency evaluation for implemented steganographic and watermarking algorithms is performed and the impact of transparency enhancing methods on an attack based robustness evaluation approach is discussed. For this steganographic algorithms are evaluated in terms of embedding transparency (using the P_B-Transparency introduced in section 3.1), detector output and content dependency. Watermarking algorithms are evaluated here by considering P_B-Transparency and P_B-Robustness, allowing for the strong connectivity between transparency and robustness in digital watermarking. The embedding transparency of a set of watermarking algorithms is measured and by using SMBA their robustness is evaluated. The performance of the perceptually tuned SMBA attacks is evaluated also using P_B-Transparency and P_B-Robustness to show how an increasing of the transparency might decrease the impact on the robustness of a watermark investigated. The results are compared to the results of the unmodified (blind) version of the attacks.

4.2. Test Set-up

Test files: For the test set-up an audio test set of 389 files is used (as covers O), which is divided into four main categories (music, sounds, speech and SQAM). All audio files are PCM coded WAVE files with 44100 Hz sampling rate, 16 bit quantisation and 2 channels (stereo) (audio CD format). They have a duration of about 30 seconds. In the category music are 267 files which are distributed into ten sub-categories (metal (20 files), pop (20), reggae (20), blues (20), jazz (20), techno (20), hiphop (20), country (20), synthetic (20) and classical). The sub-category classical music (with 87 audio files) is again sub-divided into choir (8 files), string quartet (18), orchestra (21), single instruments (19) and opera (19). The main category sounds is divided into four sub-categories (computer generated (12 files), natural (8), silence (2) and noise (11)). The main category speech has four sub-categories (male (24 files), female (20), computer generated (20) and sports (11)). The main category SQAM,[16] which is well known[17] for testing has 16 audio files (9 voice and 7 for instrumental).

Algorithms chosen for the tests: Among the algorithms used for transparency evaluation in this paper are Open Source steganography algorithms as well as LSB, Spread Spectrum and Wavelet watermarking algorithms implemented at Universities (Purdue University, USA and Otto-von-Guericke University, Germany). These algorithms were taken from the Audio WET[18] system.

A_S **chosen:** For evaluating the transparency of steganographic algorithms A_S, we used the following four algorithms:

- Publimark[19] (version 0.1.2): This algorithm is described in detail by Lang and Dittmann.[7]
- Steghide[20] (versions 0.4.3 and 0.5.1): When embedding in audio files the secret data is first compressed and encrypted (Rijndael with a key size of 128 bits in cipher block chaining mode). Then a sequence of positions of audio samples in the cover file is created for embedding, based on a pseudo-random number generator initialised with a passphrase. A graph-theoretic matching algorithm is used to find pairs of positions such that exchanging their values has the effect of embedding the corresponding part of the secret data. Because most of the embedding is done by exchanging sample values it is implied that the first-order statistics (i.e. the number of times a value occurs in the file) is not changed.

- LSB (version Heutling051208): A steganographic algorithm developed at the University of Magdeburg. The message is embedded into the LSBs of all audio samples which are not identified as silence. This algorithm is described in detail by Vogel et al..[21]

A_W **chosen:** For evaluating digital audio watermarking algorithms A_W, we used four different algorithms (Least Significant Bit, Spread Spectrum and two wavelet based algorithms (VAWW (Viper Audio Water Wavelet) and 2A2W (AMSL Audio Water Wavelet)). Those four algorithms are described in detail by Lang and Dittmann.[7]

Profile parameters and metrics for P_B-Transparency: For the evaluation of the embedding transparency of the selected A_S and A_W the ODG value between o_i (the original file) and o_{iE} (the file with the embedded steganographic message or watermark) is computed using the Open Source software tool EAQUAL (Evaluating of Audio QUALity[11]). A global transparency value for the A considered is given as the means (sum of all absolute values of the ODGs divided by the number of files) of the ODG values between all o_i (all 389 files of the test set) and the corresponding o_{iE}. Since the ODG value is defined within the range of [0,-4] and EAQUAL in some cases returns values slightly larger than 0 those values are considered to be rounding errors. In the presentation of the results in section 5 the absolute value of the ODG ($|ODG|$) will be used for a better understanding. For all used algorithms A where a message M to be embedded can be supplied then $M = UniversityOfMagdeburg$. For the transparency evaluation on steganographic algorithms the following algorithms and parameter sets A_S are considered: Publimark (version 0.1.2) with standard keys and without additional parameters, Steghide (version 0.5.1) with 20 different parameter sets (with and without encryption for 10 different compression levels (none, $c = 1$ to $c = 9$)), Steghide (version 0.4.3) with four different parameter sets (encryption on/off and ECC on/off) and the LSB steganographic algorithm (version Heutling051208; where no parameters can be set). For both Steghide versions and the LSB steganographic algorithm the passphrase used for embedding was also set to "UniversityOfMagdeburg" and therefore equal to M. For the transparency evaluation on watermarking algorithms the following algorithms and parameter sets A_W are considered: LSB watermark (version 0.3) with four differnt parameter sets ($key = 22$ ECC ON/OFF and $key = 22/\varnothing$), VAWW (build 051013) with two different parameter sets ($key = 22$, $threshold = 40$ and $scalar = 0.1$ or $scalar = 0.2$) in contradiction to all other A no M can be given to this algorithm, 2A2W (build 051013) with $key = 22$, $encoding\,method = binary$ and $watermarking\,method = ZeroTree$, Spread Spectrum (version 0.3) in four different parameter sets ($key = 22$, $embed\,strength = 5000$, ECC ON/OFF and two frequency bands for embedding (9-11 kHz and 17-19 kHz)).

Profile parameters for P_B-Robustness: The 40 attacks from SMBA used here (see Lang et al.[7]) are run with their default parameters[22] on the results from watermarking all 389 files of the test set. After the attacking process, the watermarking algorithm tries to detect the embedded watermark. For the watermarking algorithm VAWW the watermark is considered detected if the absolute value of the correlation value returned by the algorithm is between 50% and 100% of the correlation value returned when running the detector on the unmarked file.

Parameters for the evaluation of the perceptually tuned attacks: The three modified attacks are run with their default parameters and the psychoacoustics module enabled.

4.3. Test Procedure

Under the assumptions that O is the set of test files introduced and A are the algorithms identified in section 4, $M = UniversityOfMagdeburg$ (where possible), the capacity C is constant, the robustness R for the watermarking algorithms is considered separately, as T_M the ODG computed by EAQUAL is used, and D as well as T_T are not considered in this work, the following test goals for this paper are identified for the evaluation of transparency for steganography and watermarking: first for all algorithms a_i ($a_i \in A$) the embedding transparency is determined (for the A_S ($A_S \subset A$) also the detection rate is computed). Second, all results are evaluated based on the type of O (context dependency). R for the watermarking algorithms is measured using the P_B-Robustness to fulfil this requirement for a fair benchmarking of watermarking algorithms. Selected SMBA attacks (normal and perceptually tuned attacks) are run on material (watermarked by using selected parameterisations for all A_W which are considered fit for this test - due to assumptions for the spread spectrum algorithm made after the transparency and robustness evaluations it is dismissed form this evaluation) and the impact of the perceptual modelling on transparency and robustness is determined.

5. TEST RESULTS

The results of the three different evaluations (transparency for steganography and digital watermarking, robustness and the impact of psychoacoustic modelling on attack based watermark evaluation) are presented in this section.

5.1. Evaluation of Transparency for Steganography and Digital Watermarking

Steganography: Table 1 presents the average embedding transparency of all A_S tested. For the A_S where more than one parameterisation was tested the results are given in the form [x .. y] which indicates the range of the values with the upper and lower bounds (x and y).

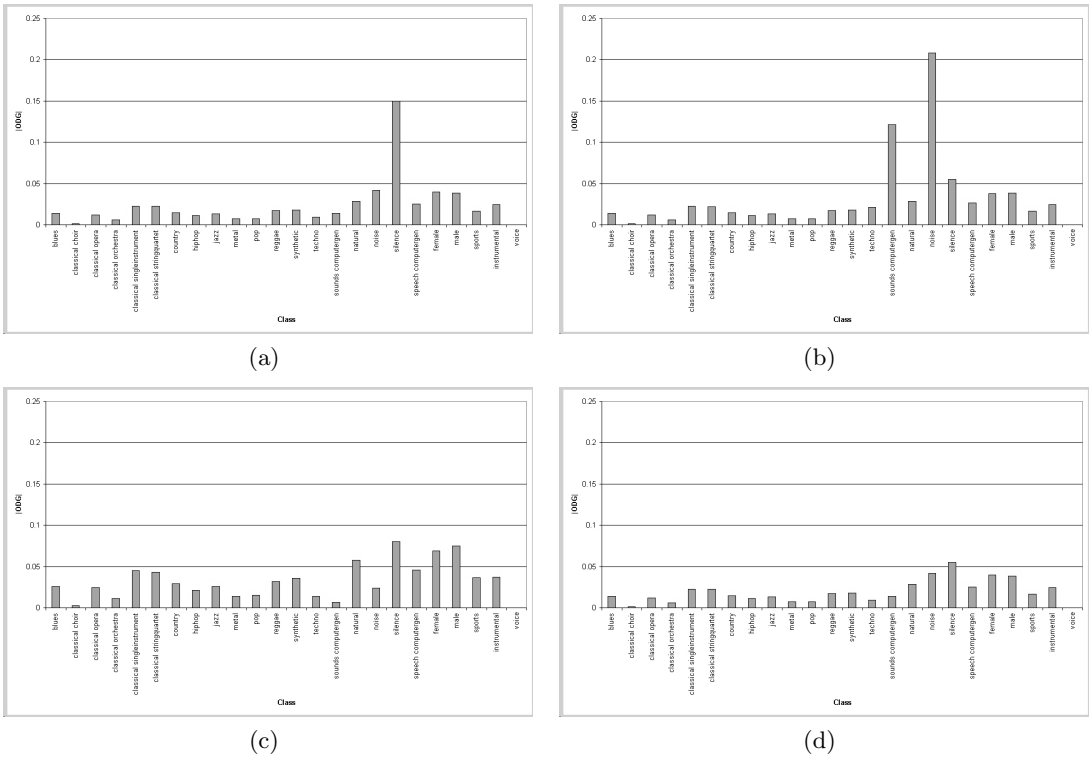
A_S	**Parameters**	**avg. embed. transp.** [$\|ODG\|$]
Publimark (version 0.1.2)		0.0180
Steghide (version 0.4.3)	Enc. ON/OFF, ECC ON/OFF	[0.0255 .. 0.0275]
Steghide (version 0.5.1)	Enc. std./OFF, compr. off/level 1-9	[0.0232 .. 0.0265]
LSB (version Heutling051208)		0.01797

Table 1: Computed average ODG values for all A_S and their parameters

From these results can be seen that all four A_S used with all parameters tested have a very similar embedding transparency (which is in all cases about 0.02 and therefore to be considered very transparent). Differences have to be found in details, when considering the detection process and the context dependency of the algorithm. Publimark shows with its average $|ODG|$ the second best result of all algorithms. It can embed into all 389 and does retrieve the correct message from all 389 marked files. The largest $|ODG|$ value was reached for the file `sounds___silence___silence.wav` ($|ODG| = 0.26$ which is about four times the value of the second largest result) which has a significant impact on the average result for the category sounds/silence (see figure 2). Steghide 0.4.3, which was tested with four different parameter sets, could embed in all 389 files but was not able to detect in seven cases (out of the $4*389 = 1556$ detection attempts performed). These seven cases were all speech signals and occurred with different settings for encryption and ECC set (`speech___female___21Abschnitt.wav` for all four sets of parameters, `speech___male___thomas-D2.wav` for both cases when encryption was disabled and `speech___male___christian2-D2.wav` when encryption and ECC are disabled). Steghide 0.4.3 gives with all parameterisations used high $|ODG|$ values for the two files `sounds___computergen___unnamed-irinter.wav` (average $|ODG| = 1.39$) and `sounds___noise___phonenumber.wav` (1.87). The third highest value is equal to 0.10. Therefore those two files have a very strong influence on the average for their category. Steghide 0.5.1 was tested with a large number of parameters to show their influence on the performance of the algorithm. All 389 files were embedded with all 20 parameter sets, which resulted in an abnormal programme termination during the embeding in the file `sounds___silence___silence.wav` for all 20 parameters. The retrieval of the message is successful in all other ($388 * 20 = 7760$) cases. Steghide 0.5.1 also gives with all parameterisations used a high $|ODG|$ value for the file `sounds___noise___phonenumber.wav` (1.87). The file `sounds___computergen___unnamed-irinter.wav` is also still the file which does produce the second largest value ($|ODG| = 0.18$) but here it is only slightly larger than the next following result (0.10). Concerning the overall performance of Steghide 0.5.1 it can be stated from the results represented in figure 2, that it performs on speech signals noticeably worse than any other algorithm tested. In general it can be stated for both versions of Steghide that the parameters used have no noticeable impact on the transparency of the algorithm. The LSB steganography algorithm, which shows the best average $|ODG|$, was expectedly not able to embed in the file `sounds___silence___silence.wav` (which contains only digital silence, i.e. the PCM value for all samples equals 0) for it uses a silence detection to identify samples in the file where no embedding should take place. The steganographic message was retrieved successfully from all other files.

When considering the classes of audio files the algorithms considered become more distinguishable. As can be seen in figure 2 all four A_S (Steghide versions 0.4.3 and 0.5.1 are represented by the results from one parameter set used, since all parameter sets return very similar results for the class representation) are performing nearly

identical on the music (columns 1 to 14) and SQAM (columns 23 & 24) categories but show strong differences in the noise (col. 15 to 18) and speech (col. 19 to 22) categories.



(a) (b)

(c) (d)

Figure 2: Visualisation of the $|ODG|$ values for the A_S (embedding) sorted by the classes ((a) Publimark, (b) Steghide 0.4.3 (SEnc, ECC on), (c) Steghide 0.5.1 (SEnc, Comp. off) and (d) LSB stego).

When considering only the transparency results for the music and speech categories (the SQAM category is considered to small with only 16 files and in sounds the fluctuation between the algorithms is to strong) for all algorithms it is implied that the impact of the embedding in speech is stronger than in music (resulting in higher $|ODG|$ values for the complete classes). For fixed transparency thresholds this could have influence on the capacity of the algorithms and should be researched in future research.

Digital Watermarking: Table 1 presents the average embedding transparency of all A_W tested. Like in the case of the A_S above value ranges are given in the form [x .. y] if the algorithm is used with more than one parameter and there is no significant difference between the values returned.

A_W	**Parameters**	**avg. embed. transp.** [$\|ODG\|$]
LSB (0.3)	key=22/empty ECC ON/OFF	[0.01770 .. 0.02039]
Spread Spectrum (0.3)	ECC ON High	0.68411
Spread Spectrum (0.3)	ECC ON Middle	2.26359
Spread Spectrum (0.3)	ECC OFF High	0.81742
Spread Spectrum (0.3)	ECC OFF Middle	2.39787
VAWW	s=0.1	1.89296
VAWW	s=0.2	2.74462
2A2W		1.08502

Table 2: Computed average $|ODG|$ values for all A_W and their parameters

Noticeable from the results in table 2 is that the average $|ODG|$ values are in three of the four cases considerably higher than the average values for the steganography algorithms. Only the LSB watermarking tool used here does produce results which are in the same range like the ones of the steganographic algorithms. This is consistent with the statement made in sections 2.1 and 2.2 and shows the practical relevance of the assumptions made. It is also obvious that the parameterisation has a strong influence un the perceptual quality of the results for the Spread Spectrum algorithm and VAWW.

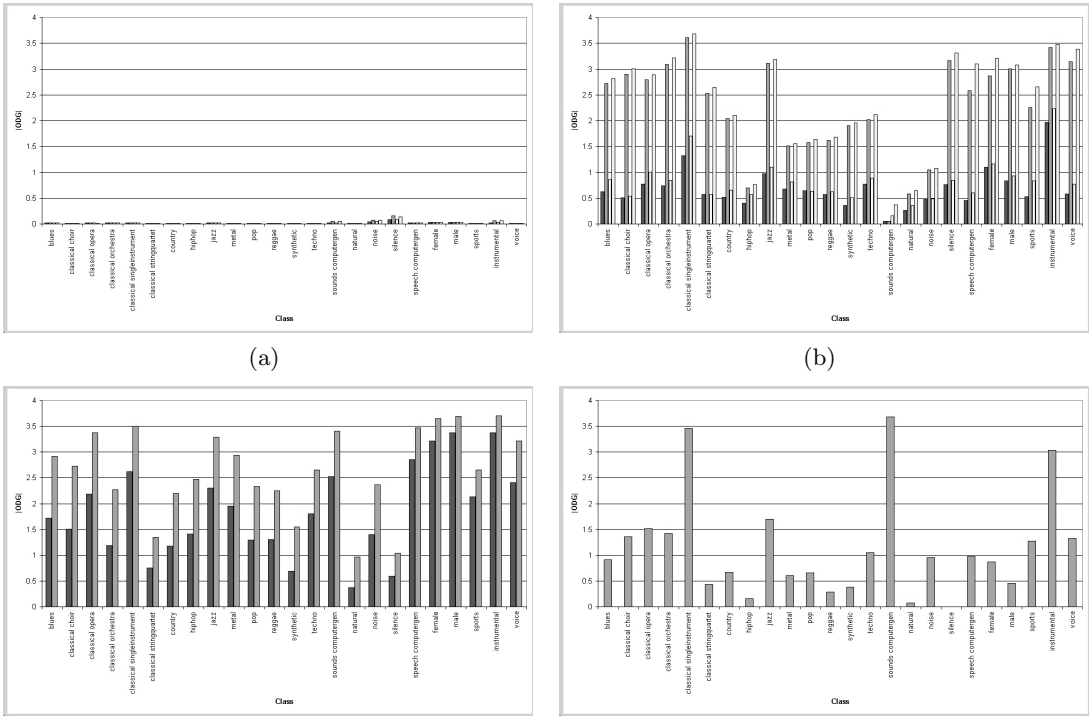


Figure 3: Visualisation of the $|ODG|$ values for the A_W (embedding) sorted by the classes ((a) LSB watermarking (four parameter sets), (b) Spread Spectrum (from left to right: ECC High, ECC Middle, no ECC High, no ECC Middle), (c) VAWW ($s = 0.1$ and $s = 0.2$) and (d) 2A2W.

The results presented in figure 3 also show that the A_W (except LSB (0.3)) perform with a much worse transparency than the A_S (please note that the scaling on the y-axis had to be changed from $y_{max} = 0.25$ to 4 for the larger codomain of the watermarking algorithms). Only the LSB watermarking algorithm (version 0.3) returns results in the same range like the A_S (like expected) even with the same distribution among the classes. The usage of an ECC does not influence the transparency but the usage of a key leads to significantly better (half the average $|ODG|$ value) results in four sub-categories (sounds/computer generated, sounds/noise, sounds/silence and SQAM/instrumental). The LSB algorithm was capable of detecting an embedded (and not modified) watermark successfully in 1551 out of the 1556 ($4 * 389$) cases. In the case of the spread spectrum algorithm (version 0.3) both parameters modified (ECC and embedding band) show am influence on the transparency. While the usage of the ECC leads to slightly worse results the selection of the Middle embedding band ($9 - 11\,kHz$) has a very strong influence and leads to a strongly distorted signal. No clear impact of one category on the embedding transparency can be determined here. The spread spectrum algorithm (version 0.3) was capable of detecting an unmodified watermark only in 685 out of the 1556 cases. This fact and the bad results in the embedding transparency evaluation (which are definitely not characteristically for spread spectrum algorithms like stated by Cox et al.[2]) lead to the assumption that here is an error in the implementation. It should be revised and tested

again. When observing the results for the VAWW algorithm (build 051013) it is obvious that the parameter s has an influence on the embedding transparency (although it is no linear dependency; the doubling of s did not lead in all cases to a dould average $|ODG|$). No clear category dependency on the transparency can be determined for this algorithm. VAWW was capable of detecting an embedded (and not modified) watermark successfully in 746 out of the 778 ($2 * 389$) cases. The results for the 2A2W algorithm (build 051013) do also show no clear category dependency of the embedding transparency. 2A2W was capable of detecting an embedded (and not modified) watermark successfully in 386 out of the 389 cases. When comparing the results of VAWW and 2A2M they show strong similarities when considering the sub-classes. This might result from the fact that they do both embed in the wavelet domain and should be considered in future research. When considering all categories of audiofiles presented here it is noticable that certain categories generally lead either to very good or very bad results (see the results for music/classical/single instrument, music/jazz, sounds/computergen sounds/natural, sounds/silence and SQAM/instrumental. Further research on the basics of context dependency in audio should try to determine the reasons for this fact.

Evaluation of Robustness: A detailed evaluation of R using P_B-Robustness of the algorithms used here can be found in Lang et al..[7] From the results presented there the conclusion can be drawn that the robustness of some A_W (spread spectrum, VAWW and 2A2W) is context dependent (shown for the categories and sub-categories identified here) and for some it is not (LSB). It has to be mentioned that the VAWW algorithm shows a very high robustness (20/8/12 or 19/8/13 using the notation introduced in section 3.2) compared to the other A_W (next one (2A2W) has 7/7/26). The very bad robustness results for the spread spectrum algorithm support the idea that the implementation tested here contains a programming error.

5.2. Transparency evaluation for attacks on digital watermarks

Here three selected attacks form the SMBA suite are analysed with respect to attack transparency and impact to the watermark. Table 3 shows the selected SMBA attacks and their modified counterparts for selected parameterisations, listing their average $|ODG|$ values on the set of test files as well as the number of successfully detected watermarks (det.). Since it is indicated by the preceding tests that the spread spectrum watermarking algorithm (version 0.3) evaluated here has with a high probability problems in its implementation it is neglected here. All other A_W are represented in this section with only one parameterisation since it is assumed that the parameters of the algorithm has no impact on the difference between the normal and modified attack.

Attack	A_W	**average** $\|ODG\|$	**average** $\|ODG\|$	**det.**	**det.**
		normal	psy.ac.	normal	psy.ac.
AddBrumm	LSB k22 ECC off	0.18470437	0.67874036	0	0
AddSinus	LSB k22 ECC off	1.335989717	0.989511568	0	1
BassBoost	LSB k22 ECC off	0.493521851	2.520539846	113	1
AddBrumm	VAWW ($s = 0.1$)	1.867634961	2.014498715	384	388
AddSinus	VAWW ($s = 0.1$)	2.345732648	2.301748072	385	388
AddBrumm	2A2W	0.989203085	1.449562982	384	165
AddSinus	2A2W	1.833367609	1.654781491	385	164

Table 3: Comparison between the selected SMBA attacks and their modified counterparts

From table 3 the following conclusions for P_B-Robustness can be drawn: In the cases of the LSB and the VAWW A_W the psychoacoustics module seems to have no impact on the number of successfully detected watermarks (which gives in the cases of the AddSinus attacks the ideal modification - more transparent and no impact on the robustness), except for the tests of the BassBoost attack where the transparency values for the modified attack can be considered very bad. In this case a problem in the multi-parameterisation approach used for this attack is assumed to be responsible for the high values. An approach for tuning this attack with a single parameter optimisation (like in the cases of AddBrumm and AddSinus) is already introduced in.[13] For the other two algorithms the BassBoost is not considered due to this fact.

In the case of the 2A2W algorithm the modification of the other two attacks has a clear impact on the number of detected watermarks.
When considering the results for the perceptual transparency it can generally stated that the modification has an impact on the transparency. In all cases the transparency of the AddSinus attack was improved but the transparency of the AddBrumm attack was decreased. This unexpected behaviour of AddBrumm may be caused by different evaluation strategies used by the psychoacoustic model for the attacks and the T_M used (EAQUAL).

6. SUMMARY

Concluding can be stated for the steganography algorithms A_S that all four algorithms perform very transparent and nearly identical when considering the average $|ODG|$. If the average $|ODG|$ values for the different categories of covers and the performance on selected files (especially `sounds___silence___silence.wav` and `sounds___noise___phonenumber.wav`) are considered the results of the algorithms become clearly distinguishable. Since the steganalytical transparency is neglected in this paper but might be also a possibility to identify the output of a certain algorithm (even without the presence of the original file which is always required in the ODG computations) it does propose a good topic for future research. Differences noticed in the impact of the embedding on music and speech signals could result in different embedding capacities for these two categories of contexts and a defined transparency threshold. Since the capacity C was kept fixed in this paper this is not evaluated here but should be researched in future research.
The A_W (except LSB watermarking) return bad results (in terms of transparency) if compared to A_S, but this was expected since watermarking algorithms are more focused on robustness then on transparency. When considering the context sub-categories for VAWW and 2A2W these two algorithms show strong similarities. In future research it might be interesting to determine whether it is possible to use this way to identify the domain where an embedding took place. Based on the bad results of the spread spectrum algorithm a recommendation for the revising of the implementation was given. If the problems encountered in the test of this algorithm are fixed it should be evaluated again. As a general benchmarking result it can be stated that the LSB watermarking algorithm is with considerable distance the most transparent and the VAWW algorithm is definetly the most robust algorithm/parameter combination tested here. Considering the performance on all context categories it can be concluded that all A_W show good results on the category sounds/natural and bad results on music/classical/singleinstrument and SQAM/instrumental. The reasons for this context dependency and its possible benefits should be determined in future research.
When observing single attacks like in section 5.2 it can be stated that the trade-off between the two characteristics impact on robustness and transparency seems to be the same like in data hiding. The example of the AddSinus attack on files marked with 2A2W shows that an increasing of the transparency decreases the robustness in this case. A further improvement of the psychoacoustic module for SMBA might lead to better results for a even larger number of attacks.

Acknowledgements

The work about single SMBA attacks described in this paper has been supported in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT. The information in this document is provided as is, and no guarantee or warranty is given or implied that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

Effort for transparency evaluation of the audio attacks is 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.

REFERENCES

1. J. Fridrich, "Applications of data hiding in digital images," *Tutorial for the ISPACS 1998 conference in Melburne, Australia* , 1998.
2. I. Cox, M. L. Miller, and J. A. Bloom, *Digital watermarking*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.
3. J. Fridrich, "Applications of data hiding in digital images," *Tutorial for the ISSPA 1999 conference in Brisbane, Australia* , 1999.
4. J. Dittmann, *Digitale Wasserzeichen*, Springer Verlag, 2000.
5. M. Kutter and F. Hartung, "Introduction to watermarking techniques.," in *Information Hiding: Techniques for Steganography and Digital Watermarking*, S. Katzenbeisser and F. A. P. Petitcolas, eds., Artech House, 2000. ISBN 1-58053-035-4.
6. A. Lang, J. Dittmann, R. Spring, and C. Vielhauer, "Audio watermark attacks: From single to profile attacks.," in *Proc. of ACM Multimedia and Security Workshop 2005*, (New York), Aug. 2005.
7. A. Lang and J. Dittmann, "Profiles for evaluation and their usage in audio wet," in *IS&T/SPIE's 18th Annual Symposium, Electronic Imaging 2006: Security and Watermarking of Multimedia Content VIII, Vol. 6072*, P. W. Wong and E. J. Delp, eds., *SPIE Proceedings*, (San Jose, California USA), Jan. 2006.
8. P. Kabal, "An examination and interpretation on ITU-R BS.1387: Perceptual evaluation of audio quality," tech. rep., McGill University, Telecommunications and Signal Processing Labatory, Department of Electrical and Computer Engineering, Mc Gill University, Canada, Dec. 2003. Version 2: 2003-12-08.
9. ITU - International Telecommunication Union, *ITU-R Rec. BS.1116 - Methods for the Subjective Assessment of small Impairments in Audio Systems including Multichannel Sound Systems*, issued by the ITU-R in 1994, updated in 1997. http://www.itu.int/ITU-R/.
10. T. Thiede, *Perceptual Audio Quality Assessment using a Non-Linear Filter Bank*. PhD dissertation, Technische Universität Berlin, Fachbereich Elektrotechnik, 1999.
11. z. Alexander Lerch.
12. A. Lang, *StirMark Benchmark for Audio (SMBA) Website.* Otto-von-Guericke Universität Magdeburg, Institute of Technical and Business Information Systems, Research Group Advanced Multimedia and Security, 2005. http://amsl-smb.cs.uni-magdeburg.de/smfa/main.php March 1st, 2005 6:45 PM.
13. C. Krätzer, *Improving Attack Transparency of Audio Watermarks by Using Psychoacoustic Methods.* Diploma thesis, Department of Computer Science, Otto-von-Guericke-Universität Magdeburg, Germany, Research Group Multimedia and Security, Department of Computer Science, Otto-von-Guericke-Universität Magdeburg, P.O. Box 4120, 39016 Magdeburg, Germany, 2005.
14. J. Dittmann, C. Krätzer, and A. Lang, "Attack tuning - attack transparency models and their impact to geometric attacks.," in *ECRYPT Research Report on Watermarking Fundamentals (D.WVL.2), Proceedings of the WAVILA Workshop on Watermarking Fundamentals*, M. Barni, J. Dittmann, J. Herrera-Joamcomarti, S. Katzenbeisser, and F. P.-G. (eds.), eds., (Barcelona (Spain)), June 2005. ISBN: 3-929757-89-3.
15. H. Fastl and E. Zwicker, *Psychoacoustics. Facts and Models.*, Springer, Berlin, second ed., 1999. ISBN 3-540-65063-6.
16. http://sound.media.mit.edu/mpeg4/audio/sqam/.
17. G. T. Waters, "Sound quality assessment material recordings for subjective tests," users' handbook for the ebu - sqam compact disc, European Broadcasting Union, Avenue Albert Lancaster 32, 1180 Bruxelles (Belgique), 1988.
18. http://141.44.30.156/wet/.
19. G. L. Guelvouit. http://perso.wanadoo.fr/gleguelv/soft/publimark/.
20. S. Hetzl. http://steghide.sourceforge.net/.
21. T. Vogel, J. Dittmann, R. Hillert, and C. Kraetzer, "Design und Evaluierung von Steganographie für Voice-over-IP," in *To appear in Sicherheit 2006 GI FB Sicherheit, GI Proceedings*, (Magdeburg, Germany), Feb. 2006.
22. A. Lang and R. Spring, "Stirmark for audio - a suite of attacks against audio watermarks." SMBA Internal documentation - unpublished, Dec. 2004.