Journal of Computers Vol. 28, No. 5, 2017, pp. 82-93
doi:10.3966/199115992017102805008

A Fair Watermarking Benchmark for Word Document

Qing Chen[1*], Xiao-xi Xing[1]

[1] School of Optical-Electrical & Computer Engineering, University of Shanghai for Science & Technology, Shanghai 200093, China
qchen@usst.edu.cn, wealthxiaoxi@163.com

Received 23 June 2016; Revised 7 February 2017; Accepted 8 February 2017

Abstract. Aiming at solving the problem of lack of common criteria for performance estimate and fair comparison of different text watermarking algorithms, this paper presents a fair watermarking benchmark for evaluating the performances of Word text watermarking. The benchmark focuses on evaluation and comparison of the most important watermarking attributes -- robustness and imperceptibility. First, the general framework of digital watermarking system is described. Then, the necessary parameters for proper benchmarking and the rating criteria for text visual quality measurement are analyzed and identified. Moreover, the paper surveys some popular image watermarking benchmarks and uses them as references to propose a new attack classification and attack pattern specialized for of Word text watermarking systems. Finally, the proposed benchmark is used to evaluate the performance of different text watermarking algorithms. The experimental results illustrate that the fair evaluation and comparison of watermark performances between different methods are efficient and reliable.

Keywords: attack classification, benchmark, performance, text watermarking

1 Introduction

The rapid growth of multimedia processing and transmission in digital form has brought the requirements of multimedia security and content protection. A lot of watermarking methods have been proposed as solutions to traditional copyright protection technologies [1-5, 14-19]. Earlier research on watermarking mainly focused on the study of various watermarking algorithms, and often neglected the issue of proper evaluation and benchmark. Most algorithm estimates had their own test sample and environment as well as performance analysis and estimate criteria of robustness, which made impossible to compare different algorithms. If an efficient and consistent benchmark can be used to evaluate performances and compare the pros and cons of various watermarking systems, new strategies for improving algorithm and promoting application of watermarking will be explored based on the estimation [6, 20, 21]. Benchmarks are tools that standardize the process of evaluating and comparing different watermarking systems and measure their performance against various types of attacks [7]. Internally, the benchmark system has fixed data sets of attacks, carriers, keys, messages, etc. to perform a fair comparison.

Some benchmarks have been constructed for evaluating audio, image, video, and 3D mesh watermarking techniques [22-26]. There are three main types of representative watermarking benchmarks: StirMark, CheckMark, and Optimark. StirMark [7, 27-28] is software package for basic robustness estimate of image watermarking algorithms based on the same test environment and the same test materials. It contains different attacks on image watermarking systems with the goal of deleting, removing or destroying the digital watermark or perturbing the synchronization. Those attacks are divided into three categories: signal processing, geometric transformations, and special transforms. However, StirMark doesn't include the execution times of watermarking embedding and detection which are important watermarking measure parameters. CheckMark [8-29], the second generation of watermarking benchmark, regroups the test classes and adds some new ones to Stirmark attempting to better evaluate watermarking technologies. It proposes Weighted PSNR and Watson's metric as fair

* Corresponding Author

criteria for comparing the visibility of different watermarking schemes. Besides, the results are outputted in a flexible XML format. Another benchmarking tool Optimark [30] has a graphical user interface for still image watermarking algorithm estimate. It includes multiple trials using different keys and messages to evaluate performance of watermarking detection or decoding. It provides evaluation of the detection and decoding performance metrics, the average embedding and detection times, the algorithm payload, and the algorithm breakdown limit for a certain attack and a certain performance criterion. It also provides the user with the option of both customized and preset benchmarking sessions. The evaluation results are summarized in multiple levels using a set of user defined weights on the selected attacks and images.

Since the very different characteristics of the text watermarking carriers, the algorithm designs and attack types of text watermarking are quite different from that of image and video watermarking algorithms. Few similar benchmarks have been proposed for performance analysis and evaluation of text watermarking. With the increasing applications of text watermarking, it is desired to build benchmark to evaluate and fairly compare text watermarking schemes. For watermarking evaluation, robustness and imperceptibility are the most important performance features. In general, there is a tradeoff between the two features. Hence, for fair benchmarking and performance evaluation, the investigated algorithms should be tested under comparable conditions not only by the robustness evaluation, but also by subjective or quantitative distortion evaluation introduced through the watermarking process. The above-mentioned benchmarks are not appropriate for text documents. However, those principles and thoughts can be used as references for establishing reasonable evaluation criteria for text watermarking.

The paper presents a fair watermarking benchmark for Word document to evaluate the performances of text watermarking. Our contributions are: (1) We described a general attack model of digital text watermarking system, and analyzed and identified the necessary parameters for proper benchmarking and the rating criteria for text visual quality measuring. (2) We surveyed some popular image watermarking benchmarks and used them as references to propose new attack classification and attack patterns specialized for Word document watermarking systems. (3) We addressed the important and often neglected issue of evaluating text watermarking techniques, established a text watermarking benchmark focused on evaluate and comparison of the two most important attributes -- robustness and imperceptibility of text digital watermarking algorithm. (4) The proposed benchmark was used to evaluate and compare the performance of two different text watermarking algorithms. The procedure of the comparison demonstrated fair evaluation and comparison of watermark performances between different methods was efficient and usability.

2 Framework and Parameters

2.1 Digital Watermarking Framework

The generic attack model of text watermarking system comprises of three main parts, i.e. message embedding, attack channel and message extraction [9, 31]. The message embedding is a two-step process. First, the message M is transcoded and/or encrypted as a watermark W. Then, the watermark is embedded imperceptibly into the host text H to obtain the watermarked text H_w. The extracted watermark W^* will be gotten from the attacked watermarked data H^*_w. Denote watermark encoder by *Enc*, watermark embedder by *Emb*, attacks by *Att*, watermark extractor by *Ext* (which may produce two kinds of outputs: decode *Dec* or detection *Det*.) A description of generic attack model M_{att} of text watermarking is given as follow:

$$M_{att}=<\{H,M,K,H_w\},\{Alg,Det,Par, H^*_w\}, \{H^*_w,K,M^*\}>. \quad (1)$$

$$Enc: W \leftarrow Enc(H,M,K_{enc}) \quad (2)$$

$$Emb: H_w \leftarrow Emb(H,W,K_{emb}). \quad (3)$$

$$Att: A \leftarrow Att(Alg,Det,Par). \quad (4)$$

$$H^*_w \leftarrow H_w+A. \quad (5)$$

$$Ext: ([W^*]) \leftarrow Ext\,(H^*_w, K_{ext}). \quad (6)$$

$$Dec: ([M^*]) \leftarrow Dec(W^*). \quad (7)$$

$$Det: ([\text{yes/no}]) \leftarrow Dec(W^*_t). \quad (8)$$

K_{enc}, K_{emb}, and K_{ext} represent the set of secret keys used in watermark encoding, embedding and extraction, respectively. The Alg, Det, and Par denote the text watermarking algorithm, detector and parameters imitated by attacker, respectively.

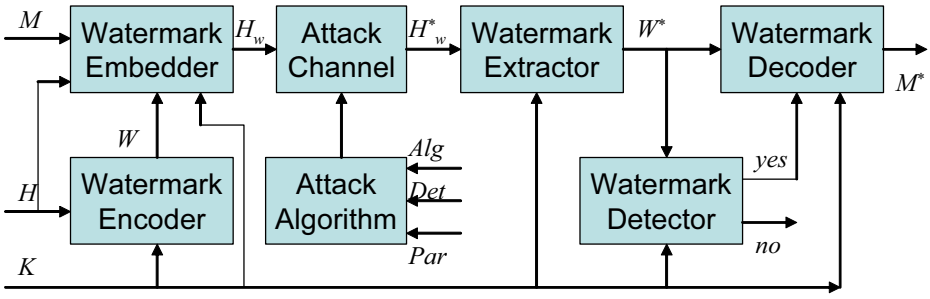


Fig. 1. Generic attack model of text watermarking system

2.2 Watermarking Parameters

It is now well accepted that an effective watermarking scheme must successfully deal with the most important requirements of imperceptibility and robustness. Generally, the relationships between imperceptibility and robustness are mutual influent and restrictive [32]. We list the essential parameters and variables which impact imperceptibility and robustness as follow:

Amount of embedded information. The amount of embedded information may influence the watermark robustness. In general, for a certain algorithm, the more information needs to embed, the lower is the watermark robustness. Unlike image watermarking, the amount of embedded information may not influence visual quality of embedded documents depending on the chosen embedding attributes and algorithms of the watermarking.

Watermark embedding strength. The strength of watermark embedding represents the difference between the attribute value of the original document and that of embedded document. It influences both the robustness and imperceptibility of watermark system. Increasing robustness requires a larger embedding strength and in turn decreasing the imperceptibility for a certain algorithm.

The number of characters of the document. It directly influences the robustness and capacity. For some text watermarking algorithms that use the cyclic embedding method, the larger the number of characters of the document is, the more the cyclic time and the stronger the robustness are, or higher capacity can be achieved at the cost of robustness.

Redundancy parameter of channel coding. For improving robustness, many algorithms exploited channel coding to increase reliability of watermark extraction. The algorithms with larger redundancy parameter of channel coding may obtain stronger resistance to various attacks. However, they may sacrifice embedding data rates and some functions of watermarking systems.

Secret key. In some scenarios, the embedded information should be bound to confidentiality. Secret key is often used to generate pseudo-random sequence as a watermark or to scramble an image watermark in order to protect the watermark from easy identification by illegal users. The watermark structure is an essential factor affecting robustness. In comparison with character watermark, image watermark can withstand certain distortion because it can be identified by human eyes. The key space, that is the range of all possible values of the secret information, must be large enough to make exhaustive search attacks impossible [33].

3 Visual Quality Metrics

Imperceptibility, the basic characteristic of watermarking system, requires that the embedded data and the original one should be perceptually undistinguishable to avoid causing the opponent's interest so as to increase the likelihood of the marked data to be attacked. Digital watermarking technology is closely related to the features of the human visual system (HVS). To ensure the visual quality of watermarked word documents, many researchers have exploited HVS to design embedding algorithm and restrict the watermark embedding strength. The HVS model is not only essential for reasonably selecting the location and strength of watermark embedding, but also significant for evaluating visual degradation due to watermark embedding and establishing a fair benchmark of text watermarking.

For the position shift, Maxemchuk et al. proposed that human eyes could not distinguish horizontal word displacements less than 1/150 inch (corresponding to about 0.5 pounds of character-shift in WORD) and vertical line displacements less than 1/300 inch (corresponding to about 0.25 pounds of line-shift in WORD) [10]. For the color change [11], the RGB components varying from (0,0,0) to (60,60,60) are very close to the default value of Word black, which are undistinguishable by human eyes.

Therefore, we take those values as the Just Noticeable Differences (JND). When the changes caused by watermark embedding are under the JND thresholds, it is generally believed that this watermarking has imperceptibility. According to the JNDs, the visual quality ratings are divided into five different levels (Table 1), which are similar to the MOS levels, the standard of classical image quality evaluation method. That is: Excellent – Imperceptible; Good – Perceptible, not annoying; Fair – Slight annoying; Poor –Annoying; Bad – Very Annoying.

Table 1. Quality ratings on a scale from 1 to 5

Rating	Word spacing	Line spacing	RGB value	Quality
5	0.3	0.1	(30,30,30)	Excellent
4	0.5	0.25	(60,60,60)	Good
3	0.7	0.35	(90,90,90)	Fair
2	1.0	0.5	(120,120,120)	Poor
1	>1.0	>0.5	(120,120,120)+	Bad

4 Possible Attacks

Watermarking attacks refer to the processes that may attenuate watermark detection or damage the representation of watermark information. Evaluating the robustness is very important for improving watermarking algorithms.

4.1 Attacks Classifications

Inspired by the principles and thoughts of image watermarking benchmark, the study of attack classification suitable for text watermarking can be analyzed and summarized under the help of the research findings of image watermarking. The first generation of watermarking benchmark – Stirmark [32] listed a number of attacks that mainly divided into geometric attacks and image processing. The second generation of benchmark Checkmark classified the attacks of watermarking as removal attacks, geometrical attacks, cryptographic attacks and protocol attacks. The goal of removal attacks is to totally remove the watermark from the stego data. Different from removal attacks, the purpose of geometrical attacks is not to remove the embedded watermark itself, but to distort the watermark through spatial or temporal alterations of the embedded carrier. Cryptographic attacks analyze the key and determine the content of an encrypted watermark and then subtract them from the watermarked works, which makes the detector unable to detect the watermark any more. The concept of protocol attacks is that an attacker can embed another watermark into the watermarked digital work to lead to ambiguity in the ownership of the digital work.

Borrowing the methods of existing attack classification of image watermarking, the attacks on text watermarking are categorized into four groups: removal attacks, geometrical attacks, cryptographic attacks and protocol attacks.

Removal Attacks.

Remark 1. Attacks based on algorithm: Text watermarking is mainly embedded watermark by fine-tuning document formats or changing the document contents.
(A) Attacks based on fine-tuning formats
Algorithms based on fine-tuning document formats usually embed watermark by changing the features of the document, such as line spacing, word spacing and RGB value, under the condition of imperceptibility constrained by human visual system [34].

Line spacing attacks: these attacks try to change the vertical line spacing of the document so that the watermark information, which is embedded by changing line spacing, may be removed.

Word spacing attacks: these attacks try to modify the word spacing of the document so that the watermark information, which is embedded by modifying word spacing, can be destroyed.

Character feature attacks: these attacks commonly try to process the character features, such as changing the RGB components, inserting imperceptible underline, etc, to embed watermark information. Therefore, attacking character features may achieve the goal of removing the embedded watermark.
(B) Attacks based on changing contents

The algorithms which embed watermark based on changing Word document contents have stronger robustness, and can resist format attacks. Usually watermark can be embedded by replacing the words in source text with their synonyms according to designed synonym replacement table of embedding algorithm or changing syntax of source text into equivalent syntax.

Synonym replacement attacks: try to use synonyms to replace the suspicious text words in the embedded document to obtain the goal of removing the watermark information.

Natural language processing attacks: these attacks try to analyze the equivalent syntax [35], and replace the suspicious embedded contents with equivalent syntax to fulfill the purpose of deleting embedded watermark information.
Remark 2. Statistical averaging and collusion: Furthermore, the attacker may obtain multiple copies of the same document that is embedded with different watermarks and combine them (such as average the documents, or extract small parts of all documents and reassembling them) in order to prevent watermark extraction, which is often called collusion attack.
Geometrical attacks. Geometrical attacks mainly are text edits (such as cut, delete and paste etc.) applied to text contents, which may disorder spatial or time series arrangement of watermark and make watermark undetectable. They are also called asynchronous attacks.
Cryptographic attacks. The aim of cryptographic attacks is to breaking the security methods in watermarking schemes and thus discovering a solution to eliminate the embedded watermark information or to embed misleading watermarks. However, implementation of these attacks is restrained due to their high computational complexity.
Remark 1. Oracle attack: From the view of security engineering, an oracle attack is an attack that utilizes the available weakness in text watermarking system as an "oracle" which can give a simple indication to show whether the attacker has reached, or is nearing, their goal. When a public decoder is available, an attacker can eliminate a watermark by small modification test to the embedded document until the decoder cannot find watermark anymore.
Remark 2. Brute force attacks: The brute force attack, also called exhaustive key search, is a cryptanalytic attack that can be used against almost any encrypted data. Such an attack might be used in the scenario lacked exploitable weaknesses in text watermarking system that would make the task easier. It contains systematically inspecting all possible passwords or keys until the right one is figured out. In the most exhaustive case, this search might traverse the whole search space. Since many watermarking algorithms utilize a secret key, it is very important to choose keys with a secure length.
Protocol attacks: The aim of protocol attack is not to destroy watermark or impair its detection, but to attack the concept of its application. Craver et al [36] proposed the first protocol attack and pointed out that embedding watermark into the embedded documents again might create ambiguity with respect to the real ownership of the document. Ownership dispute occurs when an attacker inserts a watermark in publicly available content and claims the legitimate copyright owner of the copyright protected content.

4.2 Attacks Patterns

Attack patterns are composed of attack strength, attack range, and attack mode.

The attack strength refers to the difference between the value of original attribute and the attacked one. If we change the value of line spacing and word spacing large enough (for example larger than the embedding modification) to attack a document, the watermark will be destroyed.

The attack range refers to the ratio of the attacked content to the total document, which is generally represented as a percentage.

The attack mode refers to the centralized attack mode (the attack is applied to successive text content) or the distributed attack mode (the attack is applied to separate text contents)

5 Performance Evaluation

In order to fairly evaluate the performance of different watermarking schemes, the testing setup conditions, testing performance criteria, testing objects need to be unified. There are many factors impacting the robustness. For obtaining reasonable evaluation the robustness of watermarking algorithm, certain parameters must be fixed to control the testing environment. The different testing setup conditions are listed in Table 2.

Table 2. Graphs and their testing setup conditions

Graph Type	Parameter			
	Quality	Robustness	Attack	Bits
Attack vs. Robust	Fixed	Variable	Variable	Fixed
Strength vs. Quality	Variable	Fixed	Variable	Fixed
Bits vs. Robust	Fixed	Variable	Fixed	Variable
ROC	Fixed	Variable	Variable/ fixed	Fixed

In order to illustrate feasibility and effectiveness of the proposed performance evaluation method, the paper implements a comparative scenario for two different text watermarking algorithms. One (represented as algorithm ①) is based on word spacing to embed the watermark [34]. The other (algorithm ②) is based on the font color RGB value to embed watermark [12].

The estimates of text watermarking systems were performed on various word documents with different structures and features using different keys. Some 5-page documents used in the experiments have following characteristics: (a) a document with one column format containing Arabic numerals, English and Chinese characters; (b) a document formed by inserting characteristic (a) with images and formulas; (c) a document formed by inserting characteristic (a) with tables; (d) a document formed by inserting characteristic (b) with tables; (e) a document formed by changing characteristic (d) to two columns. The robustness of text watermarking is measured by bit error rate (BER). The visual quality of the watermarked hosts will be evaluated according to the quality ratings (Table 1).

5.1 Robustness vs. Attack

Format attacks (Table 3): Attack patterns represent to change the line spacing value, word spacing value, and the RGB value according to the percentage of them in the JND thresholds. The other features represent the font, font size and so on. The experimental results show that the robustness of the two text watermarking algorithms is very strong under various attacks except that the attack attribute used is the same as embedding attribute in the algorithm.

Table 3. Robustness of two algorithms under format attacks

BER / Attack pattern	Attack method							
	Line spacing		Word spacing		RGB value		Other features	
	①	②	①	②	①	②	①	②
10%	0	0	1	0	0	1	0	0
30%	0	0	1	0	0	1	0	0
50%	0	0	1	0	0	1	0	0
80%	0	0	1	0	0	1	0	0
100%	0	0	1	0	0	1	0	0

Note. The 0 represents the attack is failed and the watermark can be correctly detected and extracted; the 1 indicates the attack is successful; other values between 0 and 1 are the values of BER. The same is as following tables.

Content attacks (Table is omitted): The synonym replacement attacks and natural language processing attacks were performed on the two text watermarking algorithms with attack range from 10% to 100%. Due to exploiting the cyclical embedding in both algorithms, the detection and extraction of watermark could be successful until the attack proportion up to 80% but failed under the attack range of 100%. Moreover, the experimental data show that the BER of algorithm ① is slightly lower than the that of algorithm ②.

Protocol attacks (Table is omitted): Another algorithm ③, based on underline attribute to embed watermark [13], is added to test the protocol attacks. In the test, one of the three algorithms is used to embed the first watermark (which is associated with the original copyright); and then the other two algorithms are used to embed another watermark (which are second watermark to fuzzify the copyright of the original watermark) into the embedded documents under different attack patterns; finally, check whether the first watermark can be detected from the attacked documents. The experimental results show that all of the first watermarks can be extracted and the three algorithms do not influence each other.

5.2 Visual Quality vs. Embedding Strength

This experiment measures visual qualities of the watermarked documents under different embedding strengths. The algorithms which are based on the changing content to embed watermark have poor imperceptibility, so measurements are only performed on the algorithms based on changing document formats to embed the watermark information. Table 4 shows the relationship between the embed strength and the visual quality. The results show that the visual qualities of three algorithms degraded with the increase of the attack strength. With the same embedding strength, the visual qualities of algorithm ② are better than the others.

Table 4. Visual quality vs. embedding strength

Quality / Strength	Embedding method		
	Word spacing ①	RGB value ②	Underline ③
10%	4.67	4.7	4.67
20%	4.33	4.4	4.33
30%	4	4.1	4
40%	3.67	3.8	3.67
50%	3.33	3.5	3.33
60%	3	3.2	3
70%	2.75	2.9	2.67
80%	2.5	2.6	2.33
90%	2.25	2.3	2
100%	2	2	1

5.3 Embedding Data Rate vs. Robustness

The relationship between embedding data rate and robustness to delete attacks on various documents, such as documents inserted images and formulas (Table 5), formatted in double columns (table is omitted), and inserted tables (Table 6), were tested. The results show that higher embedding data rate will lead to the lower watermarking robustness, and the robustness of the algorithm ① is stronger than that of algorithm ②. Additionally, the experimental results show that robustness to delete attack on various documents has no remarkable differences. However, inserting table into document will have certain influence on the test results.

Table 5. "Bits vs. Robustness" test — document with images and formulas

Attack strength \ BER	Embedding data rate									
	9byte		12byte		16byte		24byte		36byte	
	①	②	①	②	①	②	①	②	①	②
10%	0	0	0.375	0.367	0.355	0.363	0.357	0.372	0.393	0.396
20%	0	0	0.136	0.136	0.338	0.381	0.258	0.283	0.224	0.224
30%	0	0	0	0	0	0	0.491	0.491	0.222	0.224
40%	0.304	0.324	0.307	0.320	0.319	0.291	0.301	0.301	0.311	0.331
50%	0.458	0.458	0.312	0.332	0.232	0.222	0.322	0.322	0.438	0.438
60%	0	0	0	0	0.102	0	0.482	0.491	0.385	0.373
70%	0	0	0	0	0.25	0.319	0.283	0.298	0.383	0.494
80%	0.403	0.420	0.411	0.422	0.403	0.455	0.418	0.493	0.463	0.544
90%	0	0	0	0	0	0	0	0	0.334	0.464

Table 6. "Bits vs. Robustness" test—document with tables

Attack strength \ BER	Embedding data rate									
	9byte		12byte		16byte		24byte		36byte	
	①	②	①	②	①	②	①	②	①	②
10%	0	0	0.277	0.297	0.355	0.312	0.367	0.272	0.393	0.293
20%	0	0	0.239	0.239	0.338	0.324	0.305	0.333	0.324	0.364
30%	0	0	0	0	0	0	0.338	0.487	0.347	0.357
40%	0.426	0.486	0	0.320	0	0.319	0.324	0.346	0.371	0.391
50%	0.531	0.571	0	0.332	0	0.393	0.400	0.415	0.483	0.440
60%	0	0	0	0	0	0	0	0	0.473	0.493
70%	0	0	0	0	0	0	0	0	0	0
80%	0	0	0	0.422	0	0	0	0	0	0
90%	0	0	0	0	0	0	0	0	0	0

5.4 ROC, Receive Operating Characteristic

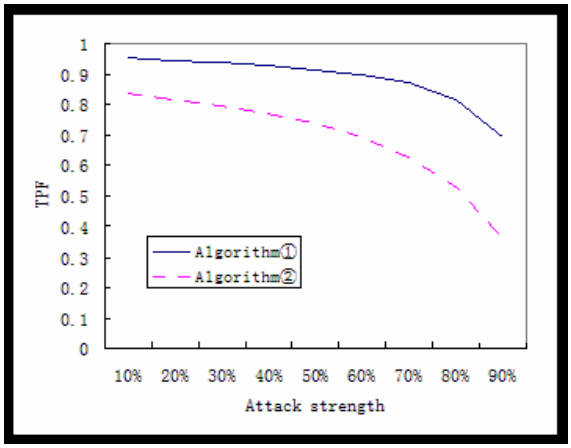
The purpose of hypothesis testing is to determine whether enough statistical evidence exists to enable us to conclude that a hypothesis about a parameter is supported by the data. For the possible outcomes of a hypothesis test, there are two types of errors: (a) false positive; (b) false negative. The definition of true positive-fraction and false-positive fraction are as follow:

$$TPF = \frac{TP}{TP + FN}. \tag{9}$$

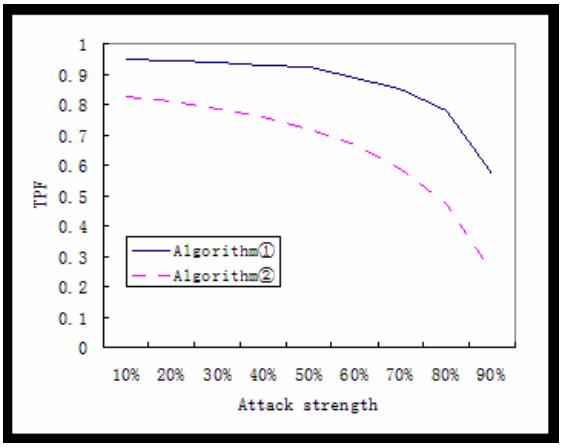
$$FPF = \frac{FP}{TN + FP}. \tag{10}$$

Where TP and FP are the numbers of true-positive and false-positive test results, respectively; and FN and TN are the numbers of false negative and true negative test results, respectively.

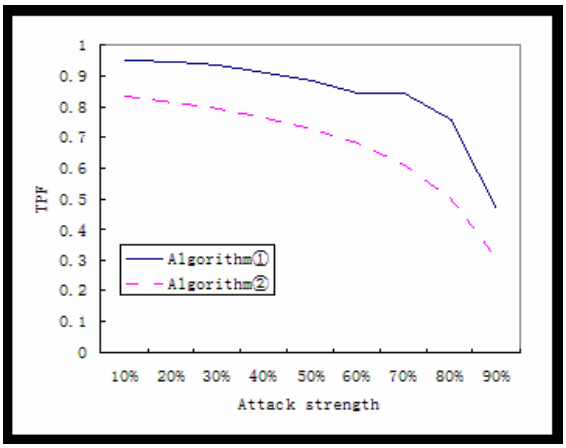
The watermarked and non-watermarked word documents were tested. The "delete" attack was used and the attack strengths were from 10% to 90% in steps of 10% on the documents with images and formulas (Fig. 2(a)), with tables (Fig. 2(b)), and in double columns (Fig. 2(c)). The graphs show that the algorithm ① possessed higher detection reliability under a variety of word documents with different contents and compositions.



(a) document with images and formulas



(b) documents with tables



(c) document in double column

Fig. 2. The TFP graph

In the experiments, the value of FPF was 0 because no watermark information could be extracted from the non-watermarked word documents.

From the above various experiments of performance evaluation, both the attack strength and the embedding data rate will influence the robustness of algorithms. The relationship between embedding strength and visual quality will interrelate and restrict each other. With the increase of attack strength or embedding data rate, the robustness of watermarking system will decrease. Visual quality will degrade as the embedding strength increase. Therefore, when designing, the text watermarking algorithms should satisfy the trade-off between imperceptibility and robustness with careful consideration of the restriction of human visual system.

6 Benchmark

Evaluating the performance of watermarking can help us analyze the strength and weakness of different algorithms and will play a very important role on improvement of watermarking techniques. In order to obtain a fair assessment test, benchmark needs a series of standards, including unified test environment, parameters and attack methods. The performances of different algorithms can be compared according to the test results.

Watermarking performance evaluation benchmark for Word formatted text is designed and implemented under the software environment of VC++6.0. The benchmark consists of watermark

embedding, attack, extraction and reports. Reports give out the operation steps and the results of each step.

6.1 Contents

Basic Strategy of Document Database for a fair estimate of text watermarking, it is required to establish a benchmark document library for testing different watermarking algorithms with the same set of standard documents. This library should cover a broad range of text documents with various contents and types, such as poem, fiction, report etc. And the contents include: Chinese, English, number, formula, tables, graphs, images, highlighting, etc.

6.2 Procedures

- Select the host document for watermark embedding;
- Input secret key. In the case of ensuring the visual quality, watermark is embedded according to the maximum capacity;
- Estimate the visual quality of the watermarked word document;
- Perform a series of attacks on the watermarked word documents under different attack patterns;
- Extract and detect watermark to determine the robustness of the watermarking algorithm at different attack patterns;
- Output the report of robustness and execution time.

7 Conclusion

Aiming at solving the problem of lack of common criteria for performance estimate and fair comparison of different text watermarking algorithms, this paper proposes a benchmark for text watermarking performance evaluation. Theoretical analysis and experimental verification are carried out. According to the characteristics of the word and human visual system, a visual quality rating standard is suggested. The classification and patterns of the attacks on text watermarking are discussed with reference to the image benchmarks. Under the same test environment, “attack strength vs. robustness”, “embed strength vs. visual quality” and “embedding data rate vs. robustness” are tested to evaluate individual performance and fair comparison among different text watermarking algorithms. The experimental results show the effectiveness and practicability of the benchmark for evaluating and comparing text watermarking algorithm. It will be beneficial to the improvement of text watermarking techniques and promotion of watermarking applications. The parameter payload, fidelity, robustness and security are often used to evaluate robust watermarking schemes. In the proposed benchmark, we paid little attention on security, a rather high-level requirement and need further study in depth.

Acknowledgements

This work was supported by the State 863 projects (No.2012AA050206) and Shanghai Natural Science Foundation (No.12ZR1420800), and National Project Cultivating Fund of University of Shanghai for Science and Technology (16HJPY-MS06).

References

[1] A.K Singh, M. Dave, A. Mohan, Multilevel encrypted text watermarking on medical images using spread-spectrum in DWT domain, Wireless Pers. Commun. 83(3)(2015) 2133-2150

[2] D. Renza, D. M. Ballesteros, H. D. Ortiz, Text hiding in images based on QIM and OVSF, IEEE Latin America Transactions 14(3)(2016) 1206-1212.

[3] X.X. Xing, Q. Chen, L.X. Fu, Research of multiple text watermarks technique in electric power system texts, Sensors and

Transducers 157(10)(2013) 331-337.

[4] G. Hua, J.W. Huang, Y.Q. Shi, J.T. Goh, L.L. Vrizlynn, Twenty years of digital audio watermarking- a comprehensive review, Signal Processing 128(11)(2016) 222-242.

[5] X.M. Zhou, L. Tan, S. Xu, An erasable watermarking scheme based on text exact authentication, Rev. Téc. Ing. Univ. Zulia. 39(6)(2016) 237-248.

[6] B.B. Zaidan, A.A. Zaidan, H.A. Karim, N.N. Ahmad, A new digital watermarking evaluation and benchmarking methodology using an external group of evaluators and multi-criteria analysis based on “large-scale data”, Software - Practice and Experience 47(10)(2017) 1365-1392.

[7] M. Kutter, F.A.P. Petitcolas, A fair evaluation methods for image watermarking systems, Elec. Imag. 9(4)(2000) 445-455.

[8] S. Voloshynovskiy, S. Pereira, V. Iquise, T. Pun, Attack modelling: towards a second generation watermarking benchmark, Signal Processing 81 (2001) 1177-1214.

[9] X. Zhou, S. Wang, S. Xiong, J. Yu, Attack model and performance evaluation of text digital watermarking, Journal of Computers 5(12)(2010) 1933-1941.

[10] J.T. Brassil, S. Low, N.F. Maxemchuk, Copyright protection for the electronic distribution of text documents, Proceedings of the IEEE 87(7)(1999) 1181-1196.

[11] M. Du, Q.Y. Zhao, Text watermarking algorithm based on human visual redundancy Advances in Information Sciences and Service Sciences 3(5)(2011) 229-235.

[12] F.F. Cai, Y. Liu, X.L. Yin, Text watermarking scheme for word documents, Computer Science 39(11)(2012) 39-40. (in Chinese)

[13] W. Fang, M. Shu, Fragile text watermarking based on changing the characters’ underlining Computer Applications and Software 25(11)(2008) 171-173.

[14] O. Tayan, Y.M. A review of recent advances on multimedia watermarking security and design implications for digital Quran computing, in: Proc. International Symposium on Biometrics and Security Technologies, 2014.

[15] D. Trick, W. Berchtold, M. Schäfer, M. Steinebach, 3D watermarking in the context of video games, IEEE International Workshop on Multimedia Signal Processing, MMSP 2013, 2013.

[16] N.N. Patil, J.B. Patil, Performance analysis of a novel text watermarking technique for devanagari text, in: Proc. International Conference on Intelligent Communication, Control and Devices, Advances in Intelligent Systems and Computing, 2017.

[17] S.R. Zhang, Z. Yao, X.C. Meng, C.C. Liu, New digital text watermarking algorithm based on New-defined characters, in: Proc. 2014 International Symposium on Computer, Consumer and Control, 2014.

[18] J.P. Chen, F.X. Yang, H.Y. Ma, Q.R. Lu, Text watermarking algorithm based on semantic role labeling, in: Proc. Third International Conference on Digital Information Processing, 2016.

[19] L. Chen, F.C. You, The study on digital watermarking based on word document, in: Proc. 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), 2013.

[20] T.H.N. Le, K.H. Nguyen, H.B. Le, Literature survey on image watermarking tools, watermark attacks, and benchmarking tools, in: Proc. 2nd International Conference on Advances in Multimedia, 2010.

[21] K. Kamiya, T. Mori, K. Iwamura, Development of benchmark tool for digital watermarking, in: Proc. 4th International Conference on Intelligent Information Hiding and Multimedia Signal, 2008.

[22] D. Megías, J. Herrera-Joancomartí, J. Serra, J. Minguillón, A benchmark assessment of the WAUC watermarking audio

algorithm, in: Proc. SPIE - v 6072, 2006

[23] R. Kaur, G.S. Kalra, M. Kansal, A user friendly GUI based benchmark for image watermarking, in: Proc. Turing 100-International Conference on Computing Sciences, 2012.

[24] P.A. Hernandez-Avalos, C. Feregrino-Uribe, R. Cumplido, J.J. Garcia-Hernandez, Towards the construction of a benchmark for video watermarking systems: temporal desynchronization attacks, in: Proc. IEEE International 53rd Midwest Symposium on Circuits and Systems, 2010.

[25] P. Schaber, S. Kopf, C. Wesch, W. Effelsberg, CamMark: A camcorder copy simulation as watermarking benchmark for digital video, in: Proc. the 5th ACM Multimedia Systems Conference, 2014.

[26] K. Wang, G. Lavoué, F. Denis, A. Baskurt, X. He, A benchmark for 3D mesh watermarking, in: Proc. International Conference on Shape Modeling and Applications, 2010.

[27] Y. Zhao, X. Zheng, N. Li, A robust audio digital watermarking in complex cepstrum domain based on chaos theory against STIRMARK, in: Proc. IEEE 8th World Congress on Intelligent Control and Automation, 2010.

[28] P.A. Hernandez-Avalos, C. Feregrino-Uribe, R. Cumplido, J.J. Garcia-Hernandez, Towards the construction of a benchmark for video watermarking systems: Temporal desynchronization attacks, in: Proc. IEEE Int. Midwest Symposium on Circuits and Systems (MWSCAS), 2010.

[29] J.C. Vorbruggen, F. Cayre, The Certimark benchmark: architecture and future perspectives, in: Proc. IEEE Int. Conf. on Multimedia and Expo, 2002.

[30] T.H.N. Le, K.H. Nguyen, H.B. Le, Literature survey on image watermarking tools, watermark attacks and benchmarking tools, in: Proc. IEEE Second Int. Conf. on Advances in Multimedia, 2010.

[31] M. Tanha, S.D.S. Torshizi, M.T. Abdullah, F. Hashim, An overview of attacks against digital watermarking and their respective countermeasures, in: Proc. International Conference on Cyber Security, Cyber Warfare and Digital Forensic, Cyber, 2012.

[32] S. Anfinogenov, Design of digital watermarking system robust to the number of removal attacks, in: Proc. Federated Conference on Computer Science and Information Systems, 2013.

[33] S. Shoaib, R.C. Mahajan, Authenticating using secret key in digital video watermarking using 3-level DWT, in: Proc. International Conference on Communication, Information and Computing Technology, 2015.

[34] Q. Chen, Y.F. Zhang, L.M. Zhou, X.D. Ding, Z. Fu, Word text watermarking for IP protection and tamper localization, in: Proc. 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce, 2011.

[35] M.L. Mali, N.N. Patil, J.B. Patil, Implementation of text watermarking technique using natural language watermarks, in: Proc. International Conference on Communication Systems and Network Technologies, 2013.

[36] S. Craver, N.D. Memon, B.L. Yeo, May invisible watermarks resolve rightful ownerships, in: Proc. SPIE Storage & Retrieval for Image and Video Databases, 1997.