Journal of Computer Sciences

Review Articles

State-of-the-Art in Techniques of Text Digital Watermarking: Challenges and Limitations

[1]Nasr Addin Ahmed Salem Al-Maweri, [1]Roslizah Ali, [1,2]Wan Azizun Wan Adnan, [1]Abd Rahman Ramli and [1]Sharifah Mumtazah Syed Ahmad Abdul Rahman

[1]Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia
[2]Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Article history
Received: 08-02-2015
Revised: 30-07-2015
Accepted: 19-03-2016

Corresponding Author:
Nasr Addin Ahmed Salem Al-Maweri
Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia

Email: senassr_maweri@yahoo.com

Abstract: Data protection from malicious attacks and misuse has become a crucial issue. Various types of data, including images, videos, audio and text documents, have given cause for the development of different methods for their protection. Cryptography, digital signatures and steganography are the most well known technologies used to protect data. During the last decade, digital watermarking technology has also been utilized as an alternative to prevent media forgery and tampering or falsification to ensure both copyright and authentication. Much work has been done to protect images, videos and audio but only a few algorithms have been considered for text document protection with digital watermarking. However, our survey observed that available text watermarking algorithms are neither robust nor imperceptible and as such remain unsecured methods of protection. Hence, research to improve the performance of text watermarking algorithms is required. This paper reviews current watermarking algorithms for text documents and categorizes text watermarking methods based on the way the text was treated during the watermarking process. It also discusses merits and demerits of available methods as well as recent proposed methods for evaluating text watermarking systems and the need for further research on digital text watermarking.

Keywords: Digital Watermarking, Text Watermarking, Text Protection, Security, Performance Evaluation

Introduction

Data representation on computer systems takes various types of file formats that are continually exchanged over the internet. Text documents, images, audio and video files represent digital media variously presented and easily obtained on global computerized networks. Adding confidentiality, security, authenticity and integrity levels to data transmission has become a vital issue due to the increased use of illegal software and malicious attacks during data transmission. Digital watermarking plays an important role in overcoming such problems. The majority of data sent via the internet is in the form of text. Hence, text protection using various methods has become a necessity and textual watermarking has been given greater attention by researchers as a solution. Cryptography is also considered an effective method of text protection. However, this method only protects data during transmission with the limitation that once data is decrypted, it becomes available and consequently protection ends in this point. Digital signatures are another means that utilize very small inserts into original data. Those inserted information are easy to remove. Fingerprinting is yet another method that only concerns text protection by providing information regarding copyright violators (Zhou *et al*., 2009).

One may, however, become confused over the concepts of steganography and watermarking. Both act as data hiding methods. But, the main difference arises in the purpose and the way of hiding the information. In watermarking, the purpose of hiding information in the media files such as images, text or videos is to protect the media content and to verify the identity, authentication and ownership. This is achieved by embedding an encoded watermark to the original content

in a visible or invisible way to prove the originality, authentication, copyright and ownership of the media file. However, in steganography, the main purpose is to transfer a secret message between two or more communicated parties through digital medium. The secret message is concealed by changing the used media file in a way that makes it difficult for anyone to figure out a present of secret message in the transferred media. Steganography is used by intelligent services, governments and armies to deliver secrets safely.

Text documents are exchanged every single moment between individuals, organizations, departments, managers and staff either via internet or intranet. Furthermore, government departments use text documents to exchange sensitive information over the internet. Misuse, fabrication, forgery or leaking such data can be of grave concern for governments and corporations as it might place them in critical circumstances. Subsequently, text documents should possess robust and secure methods of protection against malicious attacks and other abuses. However, digital watermarking algorithms that are currently used for multimedia contents such as images, audio and videos might not be applicable to text documents, (Zhou *et al*., 2009). Hence, digital watermarking research is perhaps better directed towards the design and development of enhanced and applicable watermarking algorithms for text documents.

Nowadays, digital text watermarking is taking place to protect the documents in the digital world due to the massive amount of communication via the internet using text form. Proving the ownership, copyright, access authentication, medical reports identification, privacy annotation, forensics, Smartphone texting protection, web content filtering and online content searching are some real-world applications currently involved with text watermarking. Text watermarking can be applied to many new directions in addition to the conventional know applications like copyright, ownership, broadcast monitoring and redistribution control.

Medical field now can takes the advantage of the watermarking technology by inserting the watermarks to a patient small document which will be associated to the original data reported for the patient. This will enhance the privacy of the medical information. Watermarks can be also used for the purpose of authentication of doctors and staff to specific patient reports. Not far from medical field, text watermarking can be used in forensics to gather evidences for criminal proceedings.

Smart phones, in this era, are widely used to communicate by texting, long texting messages and chatting applications are the most used applications. As new research direction, one should ask, why not text must be watermarked before sending or receiving. This may introduce a high level of security, privacy and reliability of communication. It also can be used as evidences in the case of internet crimes.

Text watermarking can be used in the market as well. Internet market promoting is one attractive application that could be thought of when choosing new directions for text watermarking. Companies and advertisers can use watermarks to make their products easier to find and locate in the internet.

Other new direction for text watermarking is to prepare the internet websites for parents caring which can be used to filter their interested contents.

Emails can be also watermarked, for the purpose of authentication, copy right or leakage control. Emails watermarking can be applied either by using image-based watermarking methods or treating the text as text. However, the later is preferred since the form of transmitted information between individuals is not welcome to be reformed, (Singh and Chadha, 2013). The text watermarking technology covers and not limited to the above explained applications and the fast evolution of the digital world might open more directions for digital text watermarking.

The scope of this paper is therefore limited to the digital watermarking of text documents, the study is organized as follows: The first section defines the importance of digital text watermarking whereas the general concept and architecture of text watermarking systems are described in the second section. The third section presents recent text watermarking techniques and offers a categorization of these techniques discussing differences, advantages and drawbacks of different approaches. The fourth section covers the importance of evaluating text watermarking algorithms and available benchmarking tools. Finally, the fifth section concludes the paper with research trends and motivators in the field of text watermarking.

Digital Text Watermarking

Currently, digital watermarking has been adopted as an effective technique that supports the protection of digitalized data from forgery, fabrication, leaking and tampering. Text watermarking systems hold the same concept for digital watermarking as any other media. But watermarking text documents differs in the methods and features used to embed the watermarks. Digital watermarking works by hiding a secured signal called a 'watermark' as a separate signal that represents the actual data meant to be exchanged via network or internet. The watermark is used to ensure the data's authenticity and integrity.

Text is known to be extensively used in the digital world. Digital books, websites content, online magazines, journals, newspapers, emails, articles and

much more contains large amount of valuable and important text to its producers. The text used in such environments is susceptible to users' misbehaviors including copyright violations, tampering, leaking etc. Hence, digital text watermarking importance has been increased consequently to protect text documents from such behaviors.

The following sections review the concept of digital watermarking systems in detail.

Digital Watermarking Architecture

Any digital watermarking system comprises two main phases, namely, *watermark embedding* and *watermark detection or extraction*. Some watermarking schemes include an initial process that generates the watermark signal before use by a component called *watermark generator*. The watermark is the actual data that is meant to be inserted in the text which could be text, logo, or image. After inserting the watermark, the original text will be considered as watermarked text. During the transfer, any attack can be applied to the text. Attacking the watermarked text is executed intentionally by malicious attacker. This kind of attacks include any misbehavior from unauthorized user to destroy the hidden watermark such as deleting text, removing features, noising, reordering sentences, paraphrasing, etc. It can be executed as well unintentionally by non-malicious attackers. This kind of attacks include normal behaviors from users such as reformatting, copying and pasting, file extensions conversions, etc. The extracted watermark is the retrieved watermark from the watermarked text. Figure 1 describes the general framework of the typical digital watermarking system.

The following sections review the major components of the digital watermarking system.

Watermark Generator

Most digital watermarking systems have a pre-specified and constant watermark signal which is used directly in the embedding process. Some systems do, however, use a different watermarking algorithm requiring that the watermark is generated first. Some algorithms use pseudo random generators to prepare the watermark signal while other algorithms generate the watermark from the main data according to a specified strategy with extracted features from targeted content (Lew and Woo, 2013b; Qadir and Ahmad, 2006; Yang *et al*., 2005; Jalil *et al*., 2010a). The use of a secret key is sometimes utilized during the watermark signal generation to introduce a level of security (Malkin and Kalker, 2007). This key is used to scramble the watermarks in some algorithms while some use it to secure the embedding location in the text or to encrypt the watermark message before really embedding it to the text. The responsible component for generating the watermark is called the *'watermark generator'* from which its output is used as input during the embedding phase.

Watermark Embedding

The second and most important phase in a digital watermarking system is the process of inserting the watermark signal into the original content, text, image, audio or video. This process is called *watermark embedding* or *insertion*. Embedding algorithms varies from one environment to another according to the type of the watermarked content.

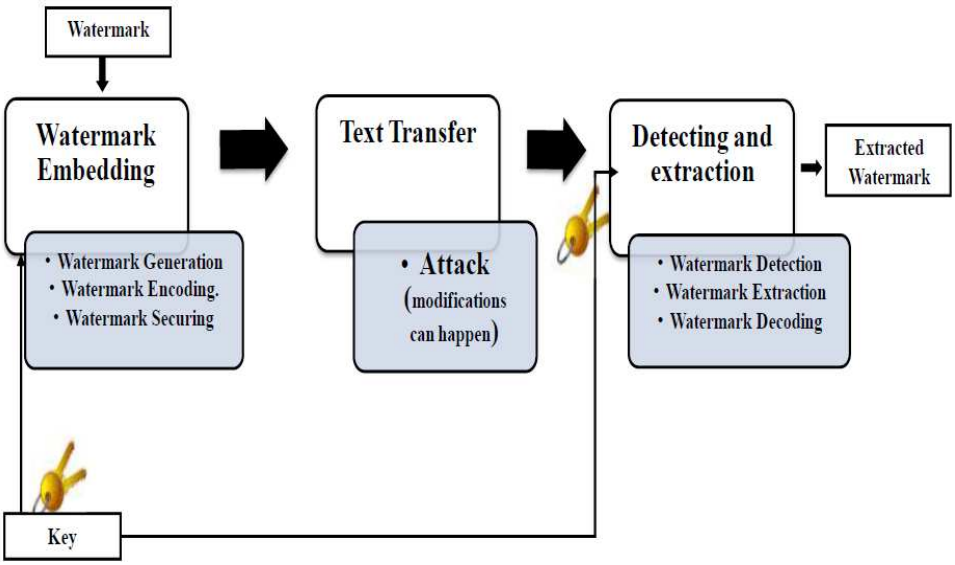

Fig. 1. Digital watermarking system

Algorithms used to embed a watermark in a text document differ from those used in images and videos. Typically, the embedding process utilizes a secret key to secure the watermark from intentional detection or removal by malicious attack. Hence, any insertion of a watermark into original data must be associated with this particular key. The same key must be used to detect the watermark later during extraction. The general watermarking scheme is represented by the following formula:

$$X = S + W \quad (1)$$

Where:
X = The watermarked data signal
S = The original data signal designated for the watermark
W = The watermark signal

This equation is common in all digital watermarking systems. It can also generalizes the idea of watermarking text by adding watermark signal to the original text in some techniques such as adding spaces, adding character extensions and in image-based techniques.

Watermark Detection and Extraction

The data watermarked during the embedding phase is delivered to the user via network or internet. At this point, it is necessary to have a function that validates the watermark signal. Two different terms are used for this phase, detection and extraction and some researchers even consider their functionality synonymous, (Muharemagic and Furth, 2006). However, they can be different terms in others perspectives, such that, watermark detection, is the component of the watermarking system that detects and verifying the availability or absence of the watermark signal in watermarked data. The extraction phase component retrieves the watermark signal's bit stream from the watermarked data and validates its integrity. The general watermarking scheme for watermark extraction can be written as:

$$\acute{W} = X - S \quad (2)$$

where, $\acute{W}$ is the extracted watermark; X is the data signal containing the watermark' and S is the original data. The secret key used to embed the watermark is used once again to detect and extract the watermark at the receiver.

The above extraction formula explains the general extraction concept which denotes non-blind text watermarking. However, the text watermarking can be blind as well where the original document should not be necessary to extract the watermark. This can be seen in some techniques such as zero-watermarking and content watermarking techniques.

Types of Watermarks

Digital watermarks can be represented as different types according to categories described in Fig. 2. They are categorized as either visible or invisible as perceived by the user depending on application requirements. Further classification depends on the extraction method utilized to retrieve the watermark. Here, the watermarking is said to be either blind or non-blind. The difference arises from any need for the original multimedia content during the extraction stage (Potdar *et al*., 2005). Another category concerns levels of strength required for the watermark. This category includes robust as well as fragile or semi-fragile watermarks. Finally, watermarks can be segregated according to their capacity to be either zero-bit or non-zero-bit, depending on the watermarking system's design and application requirements. In zero-bit watermarking the algorithm is mainly designed to detect only the presence or absence of the watermark by returning 1 or 0. But, in non-zero-bit watermarking it indicates embedding and detecting a watermark stream with multiple bits.

Digital Watermarking Applications

Traditionally, watermarking has been used to certify a document's authenticity for passports, money and certificates. A new direction in research evolved after first emerging in the early 90’s and concerns how watermarking is applied to digital media, specifically for digitally copyrighted content and proof of ownership, (Cox *et al*., 2002).

Copyright protection is the first and best-known application for digital watermarking and plays an important role in protecting text documents from ownership violations such as illegal copying and redistribution. Furthermore, it assures a text document's originality as well as confident distribution by embedding owner information such as an identification number, time-stamp, or unique serial number. Later, this information is extracted to verify the copyright or ownership (Levy and Rodriguez, 2006).

Watermarking is also used for document authentication. Embedding a unique watermark ensures that the document is being used by an authorized user or guarantees the originality of the document and also prevents forgery in the event of any falsification of the document (Levy and Rodriguez, 2006). The use of fragile watermarks identifies any modification of the document, because any change destroys the watermark. The extracted watermark can then be used as proof that the document was tampered. This type of application is called integrity or tamper detection.

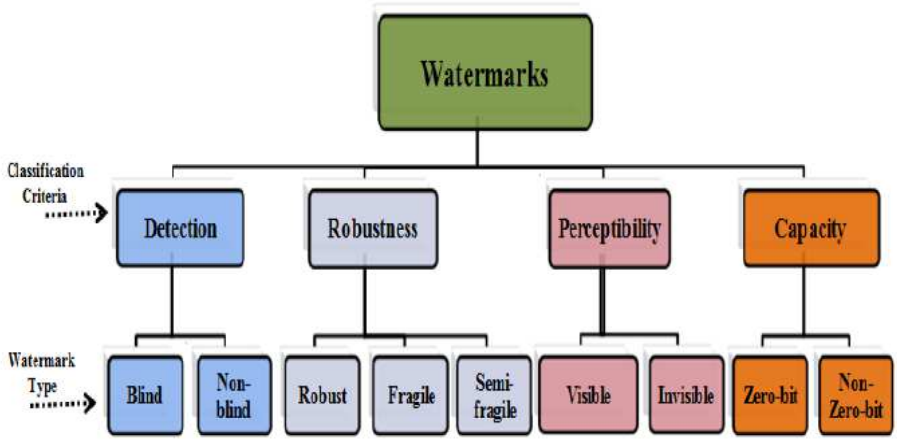

Fig. 2. Watermark types

Although broadcast monitoring is commonly used for videos; recent watermarking techniques have also been applied to monitor broadcast news stories, advertisements and internet promotions in the form of text. Access time, misuse and duration of use can be reported by linking the watermark to the broadcasting system (Levy and Rodriguez, 2006). Metadata insertion is yet another application for digital watermarking where text documents are used to hide a large amount of data. Normally, a link or reference to the original data or metadata, such as the author's name, the date of creation and document related information or description is embedded within the text content by using a watermark code. The extracted watermark is later used to refer to original information in the database or data management system (Levy and Rodriguez, 2006).

Watermark applications vary and there are endless possibilities due to the rapid growth of global user requirements and the flexible applicability of digital watermarking.

Digital Watermarking Design Constraints

In general, when designing a watermarking system, many considerations must be taken into account. The common properties of most watermarking systems are imperceptibility, robustness, capacity and security (Kaur *et al*., 2009). Achieving high degrees of capability for one or more property depends on the application. Nevertheless, an ideal watermarking system provides optimal trade-offs between these properties according to application requirements.

In text watermarking, imperceptibility, as in any other media, refers to how much perceptual quality is maintained in the original text content. Thus, increased imperceptibility implies that watermarking does not introduce any noticeable distortion due to artifacts after watermarking. Generally speaking, the attainment of high imperceptibility usually affects the robustness of the watermarking algorithm which is where its strength might otherwise be compromised.

Much like images and videos, a text watermarking algorithm must be robust enough to withstand malicious attacks. In this regard, robustness refers to the algorithm's resistance of textual violations that can potentially remove the watermark partially or completely. Copying to another environment such as the web or to any text editors, including file format conversion as well as text reordering, retyping, deletions, insertions and so forth are examples of such violations (Liu, 2006). In other words, the potential ability to detect the presence of an embedded watermark in the text by a violator during any such attack indicates the robustness of the watermarking system.

Watermark capacity is yet another issue with an important role when designing a text watermarking system, just as it is for images, videos and audio. Capacity, sometimes known as payload size, refers to the size of the watermark which is usually measured in bits. It also refers to the number of bits needed to encode the watermark. As watermark's bits number increases its robustness may decreases. This can happen once the watermark length is large and the misdetection can lead to inaccurate watermark retrieval. This is due to the increased probability of defect on the watermark if a small area in the text is attacked. Hence, watermarks with a smaller payload size considering repeated embedding are preferred. A related issue, referred to as granularity, is the amount of text needed to insert watermarking bits. In text, unlike images, fine granularity is preferred for repeated watermarks throughout the text. Consequently, robustness increases but there is a trade-off with imperceptibility.

Another constraint is watermark security. Preventing users from detecting hidden information visually or from

removing the watermark's signal from the original text provides a certain level of security. But once the watermark is detected by malicious attackers the signal should remain confusing or at least difficult to decode. Currently, association with a secret key achieves a level of security for the watermark signal. Protecting the watermark message from the ability of attackers to reveal the secret message is referred to as watermark security level. The watermark is defined secure when its level of preventing unauthorized users from accessing the watermark channel as well as decoding it is high. Such prevention was recently achieved by using either scrambling or cryptography algorithms to encode the watermark data. In scrambling, the watermark is modified in somehow to have a new reversible form. This scrambled watermark cannot be revealed without the key previously used by the authorized user. In cryptography, the watermark message is encoded using any crypto algorithm such as DES and RSA using private and public keys which will be used later to decode the encrypted message. Unfortunately however, research in text digital watermarking security is not extensively explored and needs more attention (Singh and Chadha, 2013; Liu, 2006; Furon and Bas, 2013).

Other issues for consideration when designing a text watermarking system concern the embedding and detection performance in terms of computational complexity, reliability and accuracy.

Digital Text Watermarking Techniques

Text watermarking algorithms have different schemes and strategies. Developing robust algorithms for text watermarking is not an easy task since few algorithms are currently developed for text documents and most do not address many of the known limitations. Any text watermarking algorithm can be classified under one of the following categories, namely: Format and structure, content, zero watermarking and text-as-image. Figure 3 describes these categories.

Format and Structure Watermarking

The design of this type of watermarking algorithm involves embedding the watermark in the general formatting or 'properties' of the text document. In format and structure watermarking, the layout of the text is modified to hide the watermark bits. Spaces in between the words and lines, letter extensions, curved letters and letters diacritics in some language such as Arabic and Persian and any other property that change the shape or the look of the text in an unnoticeable way. In the last decade, a variety of techniques that use the format of the text to carry the embedded watermark signal or 'payload' have emerged. A recent work was proposed by Ruia and Jinqiaob (2013) that belong to format and structure category. This algorithm executes multiple watermarking for mixed Chinese and English text. In their method Microsoft word was utilized for implementing the watermarking system.

Simply, the *LanguageID* property was used to embed two bits of the watermark, while the *Noproofing* property was used to embed 1 bit. By modifying *Noproofing* property value in the chosen character from False to True, a bit with the value 1 was considered to be the hidden bit. Changing the value of *LanguageID* between *wdBasque, wdVoda and wdEastern* indicates embedding 01, 10 and 11 respectively. In this method it was claimed that the capacity was increased. However, it was 0.5 bit/char. To further enhance the security, MD5 hashing method was used to encode the watermark before embedding. The demerit of this method is that it can be easily cracked and the watermark can be destroyed by just changing the *LanguageID* property values or by resting the *Noproofing* property to the default value. Although the robustness showed some weakness, the algorithm provides a high level of imperceptibility.

Jaiswal and Patil (2013) proposed a new method that used the text document properties. This algorithm was concerned with watermarking HTML web pages. In this approach the watermark was generated by converting the watermark information into Unicode form, then into HEX form. After that, the HEX digits were used to generate the HTML tag. This tag later will be inserted to the source code of the web page. The main advantage of this algorithm is that it resulted in a good imperceptibility, since the hidden data did not affect the content of the web page. But, when it comes to robustness, the algorithm can be attacked by only having access to the source code of the page and removing the tags that are suspected in the source code. Moreover, the watermark information security issue was not taken into account.

Mir (2014) designed a mixed watermarking method that used natural language syntactic and semantic analysis to generate the watermark. The watermark then was encoded using hashing algorithm. The hashed watermark then converted into white spaces. Those white spaces were utilized to achieve the embedding process as in other algorithms that used spaces. This method was applied in web pages content only which added some security and robustness level. Using this approach in text document such as word, pdf will have the same disadvantages of using white spaces in watermarking. These disadvantages include unpleasant large gaps between words or lines and weak resistance to normal amendment in text (deletion, reformatting, insertion, reordering, retyping or file conversion).

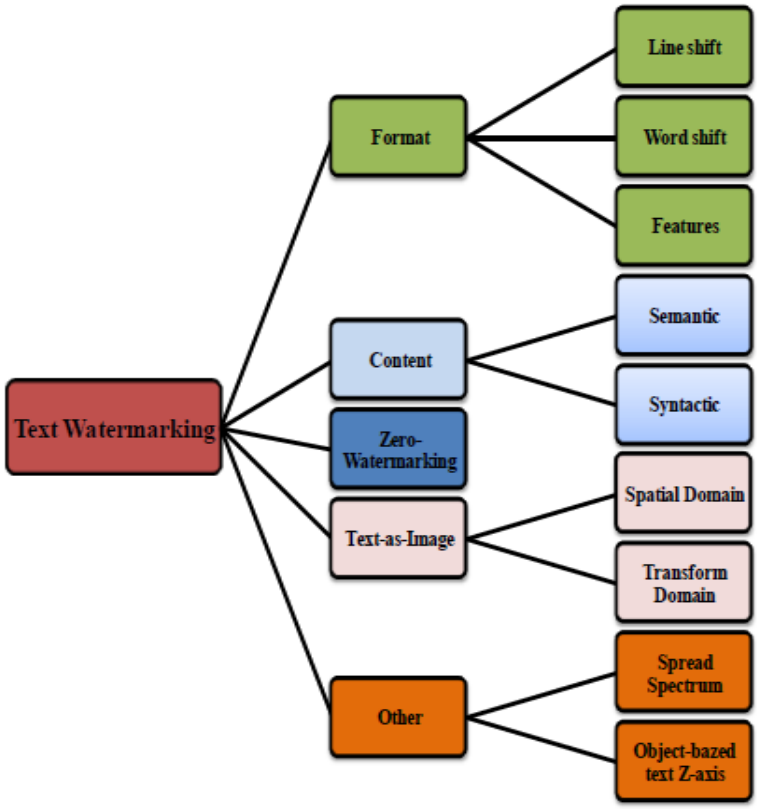

Fig. 3. Text watermarking categories

UniSpaCh was before proposed by (Por *et al*., 2012) as a text-based data hiding method introduced with Unicode space characters. In UniSpaCh, the Microsoft word document was employed as the medium for the developed algorithm. It used inter-word spaces, inter-sentence spaces, end of line and end of paragraph spaces to hide the secret message. A combination of normal spaced characters and Unicode spaced characters were implemented to encode secret data bits that survive show or hide features in Microsoft office. In this method, the ordinary space character was combined with selected Unicode space characters such as thin, six-per-em, hair punctuation, to be utilized as the carriers to hide the bits. This method increases the spaces between characters and paragraphs.

UniSpaCh utilized white spaces to embed a secret bit stream that could either be a watermark or secret message. The merit of this algorithm derived from the mixing of Unicode and ordinary space characters to introduce more efficient embedding. However, the suggested method remained vulnerable to attacks that used statistic analysis on Unicode characters. To protect the hidden information from such attacks, they suggested using a secret key and changes in periodic mapping of spaces as well as encryption of the hidden data for greater protection. Furthermore, Expert readers can notice the abnormality in some places where spaces between words is too high.

Another algorithm that belongs in the format and structure category was proposed by (Gutub *et al*., 2010). Here, watermarking is based on the text-format and intended only for Arabic language e-text. Extendable characters in Arabic were utilized to embed watermark bits by using the ‘kashida’ extension. In this algorithm, watermark bits were embedded by extending capable characters in a word. The method was introduced because the ‘kashida’ extension does not affect text content. The modification to the letters kashida will increase the length of the letter in the word according to the number of bits embedded. During embedding, a secret key generator was used to add a key seed to original watermark bits for the purpose of confusion, making the hidden data hard to guess at. Researchers attempted to achieve high watermarking capacity with an improved level of security for data hidden within the Arabic e-text. This algorithm was intended for use to prove ownership and e-text authentication. However, the proposed method did not survive re-typing and failed to show a high degree of imperceptibility because the text appeared unusual due to the presence of numerous extended characters which made readers skeptical. (Kabir *et al*., 2013) recently evaluated watermarking

algorithms for Arabic text including the 'kashida' method. In addition, they compared 'sukun' diacritics and space methods with 'kashida'. They found higher imperceptibility with 'sukun' diacritics while the other techniques proved more robust and had greater capacity. (Alginahi *et al*., 2013) tried again to enhance using 'kashidah' extension in Arabic text by adding the 'kashidah' in pre specified characters. The presence of the extended characters indicates bit 1 is hidden while the absence of the extend 'kashidah' in the chosen characters indicates zero. This method, in somehow, enhanced the perceptual quality compared to the previous 'kashidah' methods. Using a secret key and repeating the watermark N times added some enhancement to the robustness as well.

Jalil and Mrirza (2010a) proposed a method using an image watermark embedded in the text without physically effecting textual content. This method implements some features of the text in the embedding process to prove copyright for the text's author. The method converts an image format to alphabetical forms before integrating the derived alphabet with original text features (prepositions and double letters) to generate a key. The generated key is then registered with a Certifying Authority (CA). Subsequently, the key is used to extract the image watermark from the text for the purpose of verifying copyright or ownership of the text. Jalil and Mrirza (2010b) extended this work by using a mixed watermark (i.e., image-plus-text) that improved its robustness. Initially, the watermark image was converted to text before combining it with watermarked text. Both watermarks are then utilized with text features (double letters and prepositions) to generate a key as in the previous work Jalil and Mrirza (2010a). However, these methods were found unsuitable for tracking or redistribution control because both algorithms generated keys registered with a third party that can lead to error. This is because different keys used in distribution control to extract watermarks present an obstacle since the determination of which key to use for a specific document is almost impossible. Using a wrong key leads to the extraction of an incorrect watermark.

Jaseena and John (2011) utilized Zunera Jalil algorithm with improvements by adding encryption after embedding the watermark and then encrypted the text document for further security. With respect to this proposal, the work of encrypting the text is completely isolated from the watermarking component since the key generated from the text is not physically within the document but registered with the third party. Hence, encrypting the text document is an independent matter related to the organization's or owner's use to secure text documents. However, it can be said that encryption of text guarantees greater security. The above two methods can be also categorized under zero watermarking type.

Most watermarking techniques under this category such as inter-word spaces, document layout, text features or format, still do not survive normal usage. Watermarks are easily removed by reformatting, justifying, conversion and even sometimes from simple copying of the text to another environment. However, it can be said it promotes high imperceptible watermarking. Text features must be investigated by researches to find more suitable watermarking objects.

Because of the shortage found in the published works regarding the performance evaluation and reporting the results, we tried to evaluate each of the discussed techniques from our point of view. Embedding capacity, computation time or complexity and distortion rate are evaluated considering ranking levels (low, modest and high) for capacity, (low, medium and large) for computation time, (zero, low, modest and high) for distortion rate. The evaluation levels for each technique are estimated according to the methods of embedding used.

Table 1 summarizes the performance evaluation for the watermarking techniques under format and structure category. Robustness evaluation is reported in Table 6.

Content Watermarking

The content of a text document is used by the watermarking algorithm to hide the watermark in this technique. Few algorithms that use natural language to watermark a text have been published. Here, the syntactic or semantic analysis of a text's content is used to insert the watermark. Sentence structure, verbs, nouns, adjectives, pronouns, prepositions, grammar rules, synonyms and much more are the usable objects in the text content to hide the watermark message. In natural language watermarking, algorithms try to preserve the meaning but change the syntax or the semantic, so all content watermarking share the concept of changing the original content of the text such as sentence or words reordering, substituting the words or playing with grammar rules and so on. (Atallah *et al*., 2001) proposed the use of sentence syntactic structure to embed watermark bits. In this method a secret key was utilized to secure the watermark bits. The document was analyzed first, then, the syntax tree was generated. To embed the watermark bits, a criterion was assigned to select the sentences which will be used for hiding the bits. The embedding process was achieved by modifying the transformation in the syntax tree. A similar approach was proposed by (Liu *et al*., 2005) in which syntactic structure was used for Chinese language text. Here, a neural network was combined with syntactic structure to build syntax trees. The text terms were segmented first then tagged using AutoTag, a Chinese language tool. Then a secret key was generated to protect the watermark bits. According to some pre-specified factors, candidate terms are selected for semantic substitution.

Table 1. Format and Structure Techniques Evaluation

Algorithm	Embedding capacity	Computation time	Distortion rate
Ruia and Jinqiaob (2013)	low	large	zero
Jaiswal and Patil (2013)	high	medium	zero
Mir (2014)	modest	medium	modest
Por *et al*. (2012)	modest	medium	high
Gutub *et al*. (2010)	low	large	high
Alginahi *et al*. (2013)	low	large	modest
Jalil and Mrirza (2010b)	modest	large	zero
Jaseena and John (2011)	modest	large	zero

After that, the watermark embedding starts by substituting the synonyms. Such a syntactic tree was also suggested by (Kim, 2008) to build text watermarking for Korean language texts. In this method, the text was analyzed to construct the syntactic tree. By using this tree, the target constituents were selected. The embedding of the watermark bits was performed by those targets movement. The limitation of this method arises where it could be applied only in agglutinative languages such as Korean and Turkish where text reordering may not defect the meaning.

Meral *et al*. (2009) proposed another method that embedded watermark bits via binary change on Wordnet and the dictionary of the syntax tree. This method claimed to avoid semantic drops. In this method, morpho-syntactic tools were used to implement their algorithm. The embedding of the watermark bits was performed according to the available sentences in the text that could be altered. These modifications, such as adverb displacement, conjunct order change and verbs replacement. The way of altering the sentence forward or backward indicated the embedded bit either 1 or 0 respectively. Like other natural language watermarking methods, this work utilized syntactic parser to transform the text into its hierarchy as syntactic tree before watermarking. This method, however, attained a low capacity where it was only able to hide 0.5 to 1 bit/sentence. Chiang *et al*. (2004) developed an approach by utilizing the semantic replacement of textual content while preserving the same meaning in order to embed watermark bits. In this method, the text was initially segmented and then tagged after which a secret key was generated and subsequently, candidate words were selected. Finally, semantic replacement was executed. This method was also implemented in Chinese language text. Topkara *et al*. (2006) proposed a similar semantic approach by enhancing synonym replacements. This method used heuristics according to a generated priority of alternatives rather than conformity to a language model that attempted to reasonably maximize ambiguity. Their algorithm tried to enhance the capacity as well as decreasing the distortions in the watermarked text. Another work belongs to content watermarking was implemented by (Vybornova and Macq, 2007) who used presuppositions to hide watermark bits to protect text integrity and prove ownership. In this method, the text was analyzed in prior to building a sentences hierarchy. Those sentences were then transformed based on presupposition triggers. To embed the watermark bits sentences forced to have even presuppositions indicated '1', while forcing the sentences to have odd presuppositions indicated '0'. A secret key was utilized to protect the encoded bits. Mali *et al*. (2013) implemented another method that constructs the watermark from the content of the text. This method was designed to work in web pages. In their method, grammatical rules were employed to generate the watermark message. A combination of the author ID and the count of conjunctions, modal verbs and pronouns found in the text were used to produce the key. This key is saved later with Certifying Authority (CA) after passing it in AES encryption algorithm to be encrypted. By comparing this method to other natural language watermarking algorithms, it has more imperceptibility but lower robustness. Another natural language watermarking work was proposed by (Halvani *et al*., 2013). In this study, four approaches were designed to achieve the embedding process either by lexical or syntactic transformation. This work was implemented specifically to work in German language. The first syntactic transformation was utilizing enumeration modulation which embeds the watermark bits using constituent movements. The second method used conjunctions modulation constituent movement as well but this time with two nouns separated by arbitrary conjunction. Then a method of prefix expansion that is based on modifications on the negations of the words was used. The fourth method belongs to the lexical transformation, here; the watermark was embedded by modifying the words of repeated letters, connected anglicisms and inflected adjectives. What distinguishes these proposed methods over the other methods was the adaptability to some other languages such as English, Spanish or French. But, this method carries the same problems that natural language watermarking algorithms suffer from. Another algorithm that was introduced using encryption with watermarking was proposed by (Lew and Woo, 2013a). As in other Natural language based watermarking, this method used semantic technique to implement the embedding

component. The only enhancement was by adding Pseudo Random Number Generators (PRNG) for the purpose of seeding a key. The seed was combined with watermark message and then encrypted using RSA algorithm to secure the message. After that the semantic algorithm was used to embed the encrypted watermark. This method claimed in the results that it achieved better results in the robustness than Zunera Jalil Image-plus-text algorithm, since experiments were run to compare both algorithms.

A reversible text watermarking algorithm was proposed by (Fei and Tang, 2013). This algorithm used both context collation and prediction error techniques. Context collation was used to decrease the ambiguity of the content semantics when it caused during synonyms replacement. Chaos mapping was then utilized to convert overflow and underflow information resulted in the error prediction. The concept of this method was executed in Chinese text by calculating context collation degree of the sentences and words. The synonyms replacement was used to implement the embedding process under the condition that those synonyms meet the threshold condition to calculate the prediction error. The main steps to embed the watermark include word segmentation, words choosing, finding collation, finding synonyms, embedding watermark according to the prediction error. This approach enhanced the previous natural language methods in avoiding semantic drops and ambiguity. The novelty of this algorithm arises in reversing the watermarked text into its original status throughout watermark extraction function.

Dealing with Natural Languages is a complex matter since not all languages are rich enough to support semantic or syntactic manipulation. Furthermore, most such algorithms are not preferred by organizations where the original content is the utmost importance such that these approaches manipulate the meaning and the originality of the text content.

Table 2 summarizes the performance evaluation for the watermarking techniques under content Watermarking category. Robustness evaluation is reported in Table 6.

As observed from the table, content watermarking causes low embedding capacity and high distortion rate to the watermarked text compared to the other categories.

Zero-Watermarking

Due to distortions introduced to original text documents by most available algorithms, research in text watermarking has taken a new direction that attempts to achieve higher levels of imperceptibility since abnormal alterations to a text's structure or content are easily discerned. Zero-watermarking was therefore proposed to avoid such problems. In this technique, practically no watermark bits are embedded within the text's content. Instead, the algorithm extracts a stream of bits from a text's features to constitute the watermark. Since there are no actual modifications to the text, the generated watermarks from the text are registered with a Certifying Authority (CA).

Kuang and Xu (2011) developed a zero watermarking scheme that utilized TF-IDF to extract textual features (frequency of words). The document was segmented into words and then features were extracted. The watermark was generated as a combination of a time-stamp and extracted features. Instead of registering the watermark itself, the resulting hash value, SHA-1 or MD5, was registered. The hash value was then utilized for the purpose of detection. This algorithm was developed for copyright protection but can also be used to detect tampering.

Yingjie *et al*. (2010) investigated applying zero watermarking to the Chinese language. They proposed the use of sentence entropy and word frequency to constitute the watermark after trimming useless data from the text, such as punctuations. The watermark was later constructed by ordering the most significant components of sentence entropy. Before registering the watermark with a third party it was encrypted. Their method showed an ability to tolerate both deletions and insertions. However, it was limited to the Chinese language.

Table 2. Content watermarking techniques evaluation

Algorithm	Embedding capacity	Computation time	Distortion rate
Atallah *et al*. (2001)	low	medium	high
Liu *et al*. (2005)	low	medium	high
Kim (2008)	low	medium	high
Meral *et al*. (2009)	low	medium	high
Chiang *et al*. (2004)	low	medium	high
Topkara *et al*. (2006)	low	medium	high
Vybornova and Macq (2007)	low	medium	high
Mali *et al*. (2013)	low	medium	low
Halvani *et al*. (2013)	low	medium	high
Lew and Woo (2013b)	low	medium	high
Fei and Tang (2013)	low	medium	high

A new approach was released by (Jalil *et al*., 2010a) that generated the watermark payload by utilizing text content. They proposed using a keyword that determined where to extract textual features. Depending on the keyword as specified by the text's author, a count of word letters before and after the keyword was stored as different variables that were then concatenated to form the watermark's data. The keyword was added as a prefix to the accumulated data within the watermark vector. This watermark, as in other zero watermarking algorithms, was then saved with a certifying authority for the purpose of detecting any text tampering. Hence, this algorithm's watermarks were intended to be fragile. Another tamper detection algorithm was proposed by (Jalil *et al*., 2010a) who used word length as the basis to implement watermark generation. Each word of more than four letters was selected and the first character of each word was chosen and then concatenated to construct the watermark. Another zero watermarking algorithm was proposed by (Kaur and Babbar, 2013). This algorithm focused on tamper detection. To fulfill that, multiple occurrences of letters in words was used to select watermark patterns. The pattern was constructed by choosing the first letter from each word that has multiple occurrences of letters. Those patterns were combined later to generate the watermark message which will be registered with third party. This method is limited to applications that require detecting malicious modifications. Hence, the watermark was designed to be fragile. This makes the algorithm unsuitable in other applications where robustness is required.

Qi and Liu (2013) proposed a zero based watermarking approach for Chinese text. In this approach, the entropy for the frequency of different part of speech as well as expectation of text value were both calculated to generate the watermark message and then register it with the CA. This method's results claimed to have high robustness to space deletion, format, conversion, font changing, paragraph justification and synonym substitution. But, as many other zero watermarking algorithms, text insertion, deletion and rewriting may lead to incorrect watermark detection.

A recent algorithm based on zero watermarking was developed by (Ba-Alwi *et al*., 2014). This algorithm used Markov model of order three to analyze the text document and generate the watermark key which will be later registered with third party. In Markov model the three unique consecutive letters were employed to construct a probabilistic pattern. This pattern will be formed from all the occurrences of Markov model output. The resulted sequence of numbers then was encoded using MD5 hashing algorithm to secure the generated watermark. The experiments in this method showed 94.02% accuracy. However, insertion attack of 50% to the watermarked text decreased the accuracy to 63.54 of accuracy. While the algorithm achieved 96.41% by applying 5% of deletion attack, the accuracy was decreased to 67.30% by deleting 50%.

The technique of double letters was recently used again by (Prasannakumar and Balachandrudu, 2013). In This work, another algorithm that used image-plus-text watermark combined with the list of double letters was developed. The generated watermark is registered later with CA. In this algorithm the embedding process was executed as follow: Watermark was split into text and image, preprocessing will convert the image into pure text. The text produced from the image and the watermark were merged with the occurrence of the double letters from the original text to generate the final watermark key. According to the published results, the overall detection accuracy achieved was 92%. However, it was 86% in the presence of dispersed tampering attack. The disadvantage of this algorithm is that it produces long length key which make it difficult for the third party to maintain the key.

A new interesting direction in zero watermarking fields was proposed by (He and Gui, 2013). They proposed a new approach to develop an automated attack on chaotic based text zero watermarking. In this method, lexical analysis was used to implement the attack model. Adding words, deleting words, modifying words and adjusting word order were investigated in the proposed algorithm. This algorithm was intended to designing attack model for Chinese text that uses zero watermarking. Two attacks components were designed to implement the algorithm, named, Syn-attack and Birthday-attack based on synonym substitution and for the purpose of destroying the watermark information. This method results showed high imperceptibility while achieving the goal of watermark information destruction. This method opens new direction for researchers to design such automated attacks.

Although, zero watermarking has shown marked improvement in imperceptibility since the amendments do not affect textual content. However, the technique remains limited to only a few applications where robustness is not required. This is because intentional modification of the text document affects robustness and consequently leads to the extraction of different features and incorrect watermark detection.

Table 3 summarizes the performance evaluation for the watermarking techniques under zero Watermarking category. Robustness evaluation is reported in Table 6.

As noticed from the table zero watermarking introduces no zero impact on the distortion rate. However, its computation time might be higher compared to the other categories.

Table 3. Zero Watermarking Techniques Evaluation

Algorithm	Embedding capacity	Computation time	Distortion rate
Kuang and Xu (2011)	low	medium	zero
Yingjie *et al*. (2010)	low	medium	zero
Jalil *et al*. (2010b)	modest	large	zero
Kaur and Babbar (2013)	low	large	zero
Qi and Liu (2013)	low	large	zero
Ba-Alwi *et al*. (2014)	low	large	zero
Prasannakumar and Balachandrudu (2013)	low	large	zero

Image-Based Watermarking

In addition to modifying the format of the text document, various alternative techniques have been used. Some researchers have implemented their text watermarking algorithms after for scanned text documents. In such case the image of the text is used during the watermarking process. Hence, this text watermarking algorithm falls under 'image-based' watermarking where traditional techniques used for image watermarking sometimes can be applied but considering playing with the texture of the image or the image pixels. In this type of watermarking modifications explained before in format and structure watermarking such as line shift word shift or playing with letter extensions can be also used. However, this time, the text document is treated as image and the watermarking using those methods is executed in the spatial domain. Using other domains might introduce additive noise and distortions to the text image.

Borrowing on the concept that the human visual system has less sensitivity for slight amendments in character colors (Du and Zhao, 2011) developed an algorithm that used the lower four bits of RGB color components of each character for the purpose of embedding watermark bits. First, watermark data was encrypted to form the watermark vector by using Huffman encoding. The watermark vector was then segmented into 12 bit segments after which each segment was embedded by changing the lower 4 bits of each character. Numerous repetitions of the embedding process seemed to strengthen the algorithm, thus, making the process more robust. This algorithm also attained greater capacity (average capacity is 8.6 bit rate). Although the algorithm appeared to be robust, researchers have since proved that using the spatial domain in images does not provide a high level of robustness (Manoharan *et al*., 2010). Playing with text colors or converting the image from colored to a binary image, or changing the extension of the type definitely leads to watermark bit modification and loss, which showed the weakness of utilizing RGB for watermarking text images.

Another study by (Afrakhteh *et al*., 2010) also treated the text as image and reported the development of an algorithm for watermarking printed documents. The watermark image was embedded in the spatial domain where the text image was divided into partitions. The watermark's image was embedded by using multiplication with a regulation factor ranging from 0 to 1. Normalization was used during extraction for more accuracy. However, the watermark is insecure and could be detected by sniffers when they attempt to modify the regulation factor to estimate a watermark image. Rotation, Scaling and Cropping (RSC) is still a challenge for this algorithm, since detecting the accurate watermark image remains difficult.

Puhan *et al*. (2007) proposed a method to watermark binary document images for the purpose of authentication. This method used insignificant pixels, background and isolated pixels. The image was divided into blocks and the embedding of a hashed message was done within each block after lossless compression. Embedding the message bits was executed pixel-wise and the method introduced a level of robustness against Holliman-Memon and parity attacks. However, according to their paper, imperceptibility was not taken into consideration because there was so much noise introduced to the watermarked document's background when compared to the original document, which made the document image doubtful. Moreover, applying any kind of noise filters to the watermarked document may disturb the watermark's integrity. The content of the document was, however, retained since the proposed algorithm did not modify the texture of the image representing the text, although the quality of the original document was distorted after watermarking.

A different approach was implemented by (Huang and Yan, 2006). They treated the text as a binary image as mentioned in previous studies. However, they proposed a different method for hiding and constructing watermark data. Sine waves were utilized with inter-word spaces in the text lines to embed the watermark. Inter-word spaces were used as sampling points to construct the sine wave and the watermark was embedded in the form of sine wave, specifically, in the phase of sampled waves. The sine wave of the watermark was then used to manipulate inter-word spaces to embed the watermark signal. However, this method introduced slight changes to spaces in the text and thus held similar limitations as previously mentioned

algorithms that used spaces between words to embed watermark signals, regardless of signal form.

Another study by (Kim *et al*., 2003) treated text as image. They proposed using inter-word space statistics to embed payload bits by using a preprocessing system for the embedding of pixels represented by white spaces between words. Left or right word shifting was conducted in the algorithm to hide a 0 or 1 bit. To make the algorithm tolerable and robust, word classification based on word width was used to segment the text document. Bits were later embedded in each segment. This protocol made the algorithm markedly different from conventional word-shift algorithms.

Wan *et al*. (2006) tried to solve the problem of watermarking capacity for text documents. They proposed a technique that maximized embedding capacity which was calculated before actual embedding. Inter-word spaces were also utilized in this algorithm to hide watermark bits. Vertical and horizontal profiling was executed on the text image to calculate the sum of pixels that would eventually indicate characters and inter-word spaces. After ascertaining maximum embedding size, the watermark was generated by a random sequence bit generator according to calculated capacity. A hash value of character components was then XORed with the bit sequence to form the final payload. This algorithm showed better performance when detecting text tampering.

Recently few methods were emerged which treated the text as an image. Xia *et al*. (2013) proposed two approaches for watermarking Chinese text documents. The first method used the text spacing to embed bit 1 or zero. Spaces were calculated using the number of pixels between the letters and grouped into two groups B1 and B2 after characters segmentation was executed. For inserting bit 1 or 0 in each group, if the group length of spaces was in the former is greater than latter, bit one is inserted, otherwise bit zero is inserted. The embedding concept was implemented by shifting the letters left or right in the spatial domain. In the second method, the characters were segmented, each character boundary was found to compute the character height. By using the average of the character height, a reference line was assigned horizontally. Embedding 1 or 0 was achieved by shifting the letter upper or lower the reference line. Before extracting the watermark from the scanned document, noising was minimized by *binarizing* and *deskewing* of the scanned image. Comparing this algorithm to the previous line and word shift algorithms showed more capacity but less robustness to document copy process.

Taking the advantage of Persian and Arabic language richness of curved letters (Yazdani *et al*., 2013) developed a new method to watermark Persian text documents. As in the previous discussed methods, this method treated the text document as a binary image. In this method the curved letters such as 'ﺝ, ﭺ, ﭽ' were selected from the text. Embedding the watermark bits was executed by moving out specific pixel in the curved letter to hide bit 1 while leaving the letter unchanged meant 0 was hidden. The main advantage of using such technique in binary text document for Arabic and Persian text was that it provides more capacity because both languages contain many letters that meet such characteristic. This method showed high robustness compared to other similar methods. However, its disadvantage arises in the difficulty of using it in text that are printed and scanned. Furthermore, the output is limited to a fixed font.

As this review demonstrates, many conventional algorithms for text watermarking have been designed to treat text documents as binary images. Their major drawback are: (a) The degradation of original text quality; and (b) normal use and transformation of watermarked documents such as file conversion, compression, or image manipulation can destroy watermark data and lead to false detection.

Table 4 summarizes the performance evaluation for the watermarking techniques under Image-based Watermarking category. Robustness evaluation is reported in Table 6.

As noticed from the table Image-based watermarking introduces high impact on the distortion rate. However, it promotes more embedding capacity compared to the other categories.

Table 4. Image-based Watermarking Techniques Evaluation

Algorithm	Embedding capacity	Computation time	Distortion rate
Du and Zhao (2011)	high	low	low
Afrakhteh *et al*. (2010)	high	low	modest
Puhan *et al*. (2007)	high	low	high
Huang and Yan (2006)	modest	large	high
Kim *et al*. (2003)	modest	large	high
Wan *et al*. (2006)	modest	large	high
Xia *et al*. (2013)	high	large	modest
Yazdani *et al*. (2013)	modest	medium	low

Reversible Vs. Irreversible Watermarking

From another perspective, digital watermarking techniques for text can be categorized as reversible or irreversible. Once the original document data are extremely important and must be recovered to the previous state as it was before watermarking, the algorithm of watermarking should be implemented in such way that allows almost zero impact on the original text, thus, the hidden data can be extracted and at the same time the original data can be recovered. This type of algorithms can be sometimes referred to as invertible or lossless. The opposite of such concept is called irreversible: In which the original documents data are not needed to be accurately recovered and the algorithm of watermarking can be designed according to that. Research papers in this direction have not shown large specific works for watermarking text since it is most known to concern about image watermarking (Caldelli *et al*., 2010). However, few researchers specified this term in text watermarking. One example is the watermarking algorithm developed by (Fei and Tang, 2013) discussed previously in content watermarking.

Another algorithm focused on this concept was previously proposed by (Pamboukian and Kim, 2006). In this approach a reversible data hiding method was proposed for binary images and applied in text. To recover the original image a pre-process that generates compressed data is applied to the image. This data is embedded along with the watermark and used upon extraction. Using the compressed data allows the algorithm to recover the original image accurately. The compressed data are predicted using neighborhood pixels. This algorithm had to use 453 pixels for embedding 128 bit which indicates good capacity. However, it was limited to large images and images that have no random noise.

According to the method used for embedding does not restrict the algorithm to a specific category. Hence, the previous techniques mentioned under other categories can fall under more than one category. Consequently, any of the above algorithms that convey with the purpose of saving the original document as before watermarking can be described as reversible algorithms such as those techniques under zero watermarking category.

Furthermore, those techniques that implemented based on natural language and tend to modify the content can be described as irreversible and so on.

Other Techniques

Researchers have attended all the above mentioned categories and common textual properties such as spaces, double letters, word frequencies, word-shift, line-shift and other features. Nonetheless, other works in text watermarking techniques have led to the exploration of novel concepts that are difficult to define or even categorize.

An interesting research which highlights a new direction in text watermarking was proposed by (Cheng *et al*., 2013). This research proposed using multiple algorithms to watermark text documents at once. This new technique was called polymorphism watermarking. In polymorphism watermarking one or two existing algorithms of text watermarking can be applied to the text to collaborate in increasing robustness and covering each other weaknesses. For instance, a word shift method can be used besides synonym substitution method to watermark the same text document with same watermark bits. This research was launched first by QingCheng Li and needs further investigation and attention in the future of digital text watermarking.

Abdullah and Wahab (2008), for example, developed a text watermarking algorithm that employed the z-axis of text objects. This method's main targeted theme was the object-based text environment. In this method, z-ordering was utilized to embed watermark bits. Every text object's z-order was checked with a sum generated by the watermark key. The key was then processed with hash function and added to a random sequence from the watermark. The combination with the key sum was checked with the z-order in each text object in terms of parity (i.e., even or odd). According to some rules, 0 or 1 bits were inserted in the text by modifying the z-order of the text's objects. Later, the same method was used to extract the watermark from the text document.

Qadir and Ahmad (2006) suggested a different technique that used text watermarking for document transfer from sender to receiver. The watermark was inserted in the sender's signal and the watermark's added signal was extracted on the receiver's side. A spread spectrum form was used to represent the text signal. Pseudo Random Noise (PRN) was used to spread watermark signal bits and then adds to the original text signal in an invisible manner. Before this, the watermark generation phase was executed. During this phase, a Document Analyzer (DAS) was used to extract some features from the text. Document analysis output was combined with the author's signature to generate key1. A user chosen watermark represented Key2. Key1 and Key2 were encrypted and combined to form the complete watermark signal which was used as input to the spread function over PRN. Subtracting the PRN signal from the received signal and reversing the embedding process were then used to detect the watermark.

Previously, (Alattar and Alattar, 2004) also used the spread spectrum to generate watermark sequence bits with BCH error coding to protect watermark data from noise. This technique, as in other algorithms, treated the text as an image. Word-shift and line-shift were used to embed watermark bits. Their concern was to deal with text images where the text had irregular spacing. In addition, detecting the watermark from printed documents was also proposed.

Table 5. Other watermarking techniques evaluation

Algorithm	Embedding capacity	Computation time	Distortion rate
Abdullah and Wahab (2008)	low	low	low
Qadir and Ahmad (2006)	high	low	high
Alattar and Alattar (2004)	low	medium	high

Table 6. Available text watermarking techniques performance comparison (Imperceptivity and Robustness)

		Robustness				
Technique	Imperceptibility	Limitations	Text insertion	Text deletion	Reformat	Retyping
Ruia and Jinqiaob (2013)	High	Survive	Low survive	Survive No	Survive	Chinese language
Jaiswal and Patil (2013)	High	__	__	__	__	Limited to html webpage
Mir (2014)	Low	Low survive	Low survive	Low survive	No Survive	Limited to webpage
Por *et al*. (2012)	High	Low survive	Low survive	Low survive	No Survive	
Gutub *et al*. (2010)	Low	Survive	Low survive	Survive	No Survive	Arabic letters
Jalil and Mrirza (2010a) -INVISIBLE…	High	Low survive	Low survive	Survive	Survive	Does not work in distribution control
Jalil and Mrirza (2010b)-ImgePlusText	High	Low survive	Low survive	Survive	Survive	Does not work in distribution control
Jaseena and John (2011)	High	Low survive	Low survive	Survive	Survive	Does not work in distribution control
Atallah *et al*. (2001)	Low	Low survive	Low survive	Survive	Low survive	Only applicable when original text content is not important.
Liu *et al*. (2005)	Low	Low survive	Low survive	Survive	Low survive	Limited to Chinese Language
Kim (2008)	Low	Low survive	Low survive	Survive	Low survive	Only applicable in agglutinative languages.
Meral *et al*. (2009)	Low	Low survive	Low survive	Survive	Low survive	Applicable in rich morpho-syntactic languages like Turkish.
Chiang *et al*. (2004)	Low	Low survive	Low survive	Survive	Low survive	Limited to Chinese Language.
Topkara *et al*. (2006)	Low	Low survive	Low survive	Survive	Low survive	Only applicable when original text content is not important.
Vybornova and Macq (2007)	Low	Low survive	Low survive	Survive	Low survive	Only applicable when original text content is not important.
Mali *et al*. (2013)	High	Low survive	Low survive	Survive	Survive	Limited to webpage
Halvani *et al*. (2013)	High	Low survive	Low survive	Survive	Survive	
Lew and Woo (2013b)	High	Low survive	Low survive	Survive	Survive	
Fei and Tang (2013)	High	Low survive	Low survive	Survive	Survive	
Kuang and Xu (2011)	High	Low survive	Low survive	Survive	Survive	
Yingjie *et al*. (2010)	High	Low survive	Low survive	Survive	Survive	Limited to Chinese Language.
Jalil *et al*. (2010a) -Zero watermarking	High	Low survive	Low survive	Survive	Survive	
Kaur and Babbar (2013)	High	Low survive	Low survive	Survive	Survive	Limited to tamper detection
Qi and Liu (2013)	High	Low survive	Low survive	Survive	Survive	Limited to Chinese text
Ba-Alwi *et al*. (2014)	High	Low survive	Low survive	Survive	Survive	
Prasannakumar and Balachandrudu (2013)	High	Low survive	Low survive	Survive	Survive	
Du and Zhao (2011)	Low	Low survive	Low survive	Low survive	Low survive	
Afrakhteh *et al*. (2010)	Low	Low survive	Low survive	Low survive	Low survive	
Puhan *et al*. (2007)	Low	Low survive	Low survive	Low survive	Low survive	
Huang and Yan (2006)	Low	Low survive	Low survive	Low survive	Low survive	
Kim *et al*. (2003)	Low	Low survive	Low survive	Low survive	Low survive	
Wan *et al*. (2006)	Low	Low survive	Low survive	Low survive	Low survive	
Xia *et al*. (2013)	Low	Low survive	Low survive	Low survive	Low survive	
Yazdani *et al*. (2013)	High	Low survive	Low survive	Low survive	Low survive	
Abdullah and Wahab (2008)	High	Low survive	Low survive	Survive	Low survive	

Table 5 summarizes the performance evaluation for the rest of discussed watermarking techniques. Robustness evaluation is reported in Table 6.

Text Watermarking Systems Evaluation

Digital watermarking performance analysis has not enjoyed any specified standard or benchmarking tool. However, a few benchmark tools were proposed for performance analysis but are limited to image watermarking (Stirmark, Optimark and Certimark), (Muharemagic and Furth, 2006). These recommendations at least provided a step towards performance analysis and evaluation and most researchers have introduced important metrics to evaluate watermarking systems. These comprise

imperceptibility, robustness, capacity and security (Zhou *et al*., 2010; Lu *et al*., 2008).

As watermarking do not have any recognized standard method of evaluation, most researches have tested watermarked text by using several text samples to apply possible attack methods to measure the accuracy of detected watermarks (Kaur *et al*., 2009; Jaiswal and Patil, 2013). Since it remains difficult for computers to analyze similarity in text documents, imperceptibility was practically evaluated by using subjective rather than objective methods. In subjective evaluation, differences between watermarked and original documents are analyzed by experts and normal users using the Human Visual System (HVS), (Levy and Rodriguez, 2006).

Robustness of watermark effectiveness is usually evaluated by applying normal expected text operations to the watermarked document such as (i) copying to another editing environment; (ii) text insertion, deletion or reordering; or (iii) by reformatting the layout of the text and file conversion between word, pdf, notepad and WordPad.

One needs to differentiate between the evaluation of watermark robustness and security. In robustness, the watermark must survive expected normal operations in text, while a secured watermark must be evaluated for its tolerance to intentional attacks purposely designed to detect, reveal and remove it. In addition, capacity evaluation in text watermarking is not a complicated matter as most works in text watermarking evaluate this metric by measuring the number of bits needed to carry the watermark message.

Alternatively, if the watermark text is treated as an image, the evaluation methods for images are used to measure the above metrics instead. Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE) or Bit Error Rate (BER) and Just Noticeable Difference (JND) are some examples. Furthermore, benchmarking tools such as Stirmark, Optimark and Certimark can also be used (Levy and Rodriguez, 2006; Jalil *et al*., 2011). Detection Error Rate is also considered by some researchers as a measure that evaluates the reliability of watermark embedding and detection and shows the accuracy of the watermarking system (Muharemagic and Furth, 2006).

Speaking about the evaluation of text watermarking systems has put into account evaluating some of the techniques which were explained previously. Table 1 compare between those algorithms in terms of imperceptibility and robustness. Three levels of performance were considered for robustness No, Low survive or survive. For imperceptibility, two levels (High or Low) were considered.

Conclusion

Regardless of the technology used, either watermarking or digital signature, document protection in a visible way for users remains vulnerable. Watermarking techniques that treat text documents as plain text such as inter-word spaces, document layout, text features or format, do not survive normal usage. Watermarks are easily removed by reformatting, justifying, conversion and even sometimes from simple copying of the text to another environment. Algorithms that deal with text content (Natural Language) tend to change original document content and sometimes even the actual meaning of sentences, which is most unwelcomed and hence, difficult to apply. Legal documents, poems and official letters are not suitable for this method.

Although zero watermarking has shown a greater degree of imperceptibility, it has limited application and remains difficult to implement. Most conventional works in text watermarking deal with text documents as binary images. Watermarking of text content using this method has drawbacks that include the degradation of original text quality in addition to loss of the watermark through normal usage such as file conversion, compression or image manipulation.

Low robustness, poor imperceptibility and reduced security are significant drawbacks and marked weaknesses found in the above methods. Hence, further research is essential for the development of more efficient and effective algorithms to watermark text documents.

Available benchmarking tools for image watermarking are not applicable for text watermarking systems' evaluation as the latter cannot be treated as an image. Furthermore, we noted that no standard methodology exists for the benchmarking of text watermarking algorithms. Hence, there remains considerable need to develop text watermarking benchmarking tools and establish standard methods of evaluation.

Digital text watermarking is still under research and currently evolving. Some interesting research directions include: Designing algorithms that recover reordering and retyping attacks, investigating the performance of polymorphism watermarking and designing algorithms that survive screen shots attacks and printing. Another interesting direction is implementing a benchmarking system that automates attacks and measures the performance metrics.

Acknowledgement

This research was supported by the minisry of higher education malaysia and universiti putra malaysia under exploratory.

Author's Contributions

Nasr addin Ahmed Salem Al-maweri: Designed the research plan and orgnized the study, paericipated in all experiments, coordinated the data analysis and contrubuted to the writing of the manuscript.

Roslizah Ali, Wan Azizun Wan Adnan, Abd Rahman Ramli and Sharifah Mumtazah Syed Ahmad Abdul Rahman: Participaed in all experiments, coordinated the data analysis and contruvuted to he writing of the manuscript.

Ethics

The authors confirm that they have read and approved the manuscript and there is no ethical issue involved. This work is original and contains unpublished materials.

References

Abdullah, M. and F. Wahab, 2008. Key based text watermarking of e-text documents in an object based environment using z-axis for watermark embedding. World Acad. Sci. Eng. Technol.

Afrakhteh, M., S. Ibrahim and M. Salleh, 2010. Printed document authentication using watermarking technique. Proceedings of the 2nd International Conference on Computational Intelligence, Modelling and Simulation, Sept. 28-30, IEEE Xplore Press, Bali, pp: 367-370. DOI: 10.1109/CIMSiM.2010.70

Alattar, A.M. and O.M. Alattar, 2004. Watermarking electronic text documents containing justified paragraphs and irregular line spacing. Proceedings of the Security, Steganography and Watermarking of Multimedia Contents, (WMC' 04), SPIE, pp: 685-685. DOI: 10.1117/12.527147

Alginahi, Y., M. Kabir and O. Tayan, 2013. An enhanced kashida-based watermarking approach for Arabic text-documents. Proceedings of the International Conference on Electronics, Computer and Computation, Nov. 7-9, IEEE Xplore Press, Ankara, pp: 301-304. DOI: 10.1109/ICECCO.2013.6718288

Atallah, M.J., V. Raskin, M. Crogan, C. Hempelmann and F. Kerschbaum *et al*., 2001. Natural language watermarking: Design, analysis and a proof-of-concept implementation. Proceedings of the 4th International Workshop on Information Hiding, Apr. 25-27, Springer, Pittsburgh, PA, USA, pp: 185-199. DOI: 10.1007/3-540-45496-9_14

Ba-Alwi, F.M., M.M. Ghilan and F.N. Al-Wesabi, 2014. Content authentication of English text via internet using zero watermarking technique and Markov model. Int. J. Applied Inform. Syst., 7: 25-36. 10.5120/ijais14-451128

Caldelli, R., F. Filippini and R. Becarelli, 2010. Reversible watermarking techniques: An overview and a classification. EURASIP J. Inform. Security, 2010: 134546-134546. DOI: 10.1155/2010/134546

Cheng, Y., J. Zhang, X. Liu, Q. Li and Z. Chen, 2013. Research on polymorphism in digital text watermarking. Proceedings of 5th International Conference on Intelligent Networking and Collaborative Systems, Sept. 9-11, IEEE Xplore Press, Xi'an, pp: 166-172. DOI: 10.1109/INCoS.2013.33

Chiang, Y.L., L.P. Chang, W.T. Hsieh and W.C. Chen, 2004. Natural Language Watermarking using Semantic Substitution for Chinese Text. In: Digital Watermarking, Kalker, T., I. Cox and Y. Ro (Eds.), Springer Berlin Heidelberg, pp: 129-140.

Cox, I.J., M. Miller, J. Bloom and C. Honsinger, 2002. Digital watermarking. J. Electron. Imag., 11: 414-414. DOI: 10.1117/1.1494075

Du, M. and Q. Zhao, 2011. Text watermarking algorithm based on human visual redundancy. Adv. Inform. Sci. Service Sci., 3: 229-235.

Fei, W. and X. Tang, 2013. Reversible text watermarking algorithm using prediction-error expansion method. Proceedings of the International Conference on Computer, Networks and Communication Engineering (NCE' 13), Atlantis Press.

Furon, T. and P. Bas, 2013. A New Measure of Watermarking Security Applied on QIM. In: Information Hiding, Kirchner, M. and D. Ghosal (Eds.), Springer Berlin Heidelberg, pp: 207-223.

Gutub, A.A.A., F. Al-Haidari, K.M. Al-Kahsah and J. Hamodi, 2010. E-text watermarking: Utilizing 'kashida' extensions in Arabic language electronic writing. J. Emerg. Technol. Web Intell., 2: 48-55. DOI: 10.4304/jetwi.2.1.48-55

Halvani, O., M. Steinebach, P. Wolf and R. Zimmermann, 2013. Natural language watermarking for german texts. Proceedings of the 1st ACM Workshop on Information Hiding and Multimedia Security, Jun. 17-19, ACM, Montpellier, France, pp: 193-202. DOI: 10.1145/2482513.2482522

He, L. and X.L. Gui, 2013. An active attack on chaotic based text zero-watermarking. Proceedings of the IEEE Conference Anthology, Jun. 1-8, IEEE Xploe Press, China, pp: 1-4. DOI: 10.1109/ANTHOLOGY.2013.6784740

Huang, D. and H. Yan, 2006. Interword distance changes represented by sine waves for watermarking text images. IEEE.

Jaiswal, R. and N. Patil, 2013. Implementation of a new technique for web document protection using Unicode. Proceedings of the International Conference on Information Communication and Embedded Systems, Feb. 21-22, IEEE Xplore Press, Chennai, pp: 69-72. DOI: 10.1109/ICICES.2013.6508287

Jalil, Z., M.A. Jaffar and A.M. Mirza, 2011. A novel text watermarking algorithm using image watermark. Int. J. Innovative Comput. Inform. Control, 7: 1255-1271.

Jalil, Z. and A.M. Mirza, 2010a. Text watermarking using combined image-plus-text watermark. Proceedings of the 2nd International Workshop on Education Technology and Computer Science, Mar. 6-7, IEEE Xplore Press, Wuhan, pp: 11-14. DOI: 10.1109/ETCS.2010.494

Jalil, Z. and A.M. Mrirza, 2010b. An Invisible Text Watermarking Algorithm Using Image Watermark. In: Innovations in Computing Sciences and Software Engineering, Sobh, T. and K. Elleithy (Eds.), Springer, New York, ISBN-10: 9048191122, pp: 147-152.

Jalil, Z., A.M. Mrirza and H. Jabeen, 2010a. Word length based zero-watermarking algorithm for tamper detection in text documents. Proceedings of the 2nd International Conference on Computer Engineering and Technology, Apr. 16-18, IEEE Xplore Press, Chengdu, pp: 378- 382. DOI: 10.1109/ICCET.2010.5486185

Jalil, Z., A.M. Mrirza and M. Sabir, 2010b. Content based zero-watermarking algorithm for authentication of text documents. Int. J. Comput. Sci. Inform. Security, 7: 212-217.

Jaseena, K.U. and A. John, 2011. Text watermarking using combined image and text for authentication and protection. Int. J. Comput. Applic., 20: 8-13.

Kabir, M.N., O. Tayan and Y.M. Alginahi, 2013. Evaluation of watermarking approaches for Arabic text documents. Int. J. Comput. Sci. Informat. Security.

Kaur, D.P., J. Kaur and K. Deep, 2009. Digital image watermarking: Challenges and approach for a robust algorithm. Int. J. Electron. Eng., 1: 95-97.

Kaur, S. and G. Babbar, 2013. A zero-watermarking algorithm on multiple occurrences of letters for text tampering detection. Int. J. Comput. Sci. Eng., 5: 294-301.

Kim, M.Y., 2008. Text watermarking by syntactic analysis. Proceedings of the 12th WSEAS International Conference on Computers, (ICC' 08), World Scientific and Engineering Academy and Society, Heraklion, Greece, pp: 904-909.

Kim, Y.W., K.A. Moon and I.S. Oh, 2003. A text watermarking algorithm based on word classification and inter-word space statistics. Proceedings of the 7th International Conference on Document Analysis and Recognition, Aug. 3-6, IEEE Xplore Press, Korea, pp: 775-779. DOI: 10.1109/ICDAR.2003.1227767

Kuang, Q. and X. Xu, 2011. A new zero-watermarking scheme based on features extraction for authentication of text. J. Convergence Inform. Technol., 6: 155-165. DOI: 10.4156/jcit.vol6.issue11.18

Levy, K. and T. Rodriguez, 2006. Digital watermarking framework: Applications, parameters and requirements. Auerbach Publications, Taylor and Francis Group.

Lew, C.H. and C.S. Woo, 2013a. Design and implementation of text based watermarking combined with Pseudo- Random Number Generator (PRNG) for cryptography application. Proceedings of the 12th International Conference on Applied Computer and Applied Computational Science, (ACS' 13), Wseas LLC.

Lew, C.H. and C.S. Woo, 2013b. Using combined pseudo-random number generator with digital text-based watermarking for cryptography application. Proceedings of the International Multi Conference of Engineers and Computer Scientists, (ECS' 13).

Liu, Z., 2006. New Trends and Challenges in Digital watermarking Teccnology: Applications for Printed Materials. In: Multimedia Watermarking Techniques and Applications, Kirovski, D. (Ed.), Auerbach Publications, Boca Raton, ISBN-10: 0849372135, pp: 257-306.

Liu, Y., X. Sun and Y. Wu, 2005. A natural language watermarking based on Chinese syntax. Proceedings of the 2st International Conference on Advances in Natural Computation, Aug. 27-29, Springer, Changsha, China, pp: 958-961. DOI: 10.1007/11539902_119

Lu, P., Z. Lu, Z. Zhou and J. Gu, 2008. An optimized natural language watermarking algorithm based on TMR. Proceedings of the 9th International Conference for Young Computer Scientists, Nov. 18-21, IEEE Xplore Press, Hunan, pp: 1459-1463. DOI: 10.1109/ICYCS.2008.398

Mali, M.L., N.N. Patil and J.B. Patil, 2013. Implementation of text watermarking technique using natural language watermarks. Proceedings of the International Conference on Communication Systems and Network Technologies, Apr. 6-8, IEEE Xplore Press, Gwalior, 482-486. DOI: 10.1109/CSNT.2013.106

Malkin, M. and T. Kalker, 2007. A cryptographic method for secure watermark detection. Proceedings of the 8th International Conference on Information Hiding, Jul. 10-12, Springer, Alexandria, VA, USA, pp: 26-41. DOI: 10.1007/978-3-540-74124-4_3

Manoharan, J., D.C. Vijila and A. Sathesh, 2010. Performance analysis of spatial and frequency domain multiple data embedding techniques towards geometric attacks. Int. J. Security, 4: 28-37.

Meral, H.M., B. Sankur, A.S. Zsoy, T. Günger and E. Sevinç, 2009. Natural language watermarking via morphosyntactic alterations. Comput. Speech Lang., 23: 107-125. DOI: 10.1016/j.csl.2008.04.001

Mir, N., 2014. Copyright for web content using invisible text watermarking. Comput. Human Behav., 30: 648-653. DOI: 10.1016/j.chb.2013.07.040

Muharemagic, E. and B. Furth, 2006. Survey of Watermarking Techniques and Applications. In: Multimedia Watermarking Techniques and Applications, Kirovski, D. (Ed.), Auerbach Publications, Boca Raton, ISBN-10: 0849372135, pp: 91-130.

Pamboukian, S.V.D. and H.Y. Kim, 2006. Reversible data hiding and reversible authentication watermarking for binary images. Proceedings of the 6th Brazilian Symposium on Information and Computer System Security, (CSS' 06).

Por, L.Y., K. Wong and K.O. Chee, 2012. Unispach: A text-based data hiding method using unicode space characters. J. Syst. Software, 85: 1075-1082. DOI: 10.1016/j.jss.2011.12.023

Potdar, V.M., S. Han and E. Chang, 2005. A survey of digital image watermarking techniques. Proceedings of the 3rd IEEE International Conference on Industrial Informatics, Aug. 10-12, IEEE Xplore Press, pp: 709-716. DOI: 10.1109/INDIN.2005.1560462

Prasannakumar, K.R.P. and K.E. Balachandrudu, 2013. Text watermarking using combined image and text. Int. J. Eng. Res. Technol.

Puhan, N.B., A.T.S. Ho and F. Sattar, 2007. Erasable authentication watermarking in binary document images. Proceedings of the 2nd International Conference on Innovative Computing, Information and Control, Sept. 5-7, IEEE Xplore Press, Kumamoto, pp: 288-288. DOI: 10.1109/ICICIC.2007.289

Qadir, M.A. and I. Ahmad, 2006. Digital text watermarking: Secure content delivery and data hiding in digital documents. IEEE Aerospace Electr. Syst. Magaz., 21: 18-21. DOI: 10.1109/MAES.2006.284353

Qi, X. and Y. Liu, 2013. Cloud model based zero-watermarking algorithm for authentication of text document. Proceedings of the 9th International Conference on Computational Intelligence and Security, Dec. 14-15, IEEE Xplore Press, Leshan, pp: 712-715. DOI: 10.1109/CIS.2013.155

Ruia, X. and C.X.S. Jinqiaob, 2013. A multiple watermarking algorithm for texts mixed Chinese and English. Proc. Comput. Sci., 17: 844-851.

Singh, P. and R.S. Chadha, 2013. A survey of digital watermarking techniques, applications and attacks. Int. J. Eng. Innovative Technol., 2: 165-175.

Topkara, U., M. Topkara and M.J. Atallah, 2006. The hiding virtues of ambiguity: Quantifiably resilient watermarking of natural language text through synonym substitutions. Proceedings of the 8th Workshop on Multimedia and Security, Sept. 26-27, ACM, Geneva, Switzerland, pp: 164-174. DOI: 10.1145/1161366.1161397

Vybornova, O. and B. Macq, 2007. Natural language watermarking and robust hashing based on presuppositional analysis. Proceedings of the IEEE International Conference on Information Reuse and Integration, Aug. 13-15, IEEE Xplore Press, Las Vegas, IL., pp: 177-182. DOI: 10.1109/IRI.2007.4296617

Wan, I.A., S.A.M. Gilani and S.A. Shah, 2006. Utilization of maximum data hiding capacity in object-based text document authentication. Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Dec. 18-20, IEEE Xplore Press, Pasadena, CA, USA., pp: 597-600. DOI: 10.1109/IIH-MSP.2006.265073

Xia, Z., S. Wang, X. Sun and J. Wang, 2013. Print-scan resilient watermarking for the Chinese text image. Int. J. Grid Distributed Comput., 6: 51-62. DOI: 10.14257/ijgdc.2013.6.6.05

Yang, H., A.C. Kot and J. Liu, 2005. Semi-fragile watermarking for text document images authentication. Proceedings of the IEEE International Symposium on Circuits and Systems, May 23-26, IEEE Xplore Press, pp: 4002-4005. DOI: 10.1109/ISCAS.2005.1465508

Yazdani, V., M.A. Doostari and H. Yazdani, 2013. A new method to Persian text watermarking using curvaceous letters. J. Basic Applied Scientific Res., 3: 125-131.

Yingjie, M., G. Tao, G. Zhihua and G. Liming, 2010. Chinese text zero-watermark based on sentence's entropy. Proceedings of the International Conference on Multimedia Technology, Oct. 29-31, IEEE Xplore Press, Ningbo, pp: 1-4. DOI: 10.1109/ICMULT.2010.5631421

Zhou, X., S. Wang, S. Xiong and J. Yu, 2010. Attack model and performance evaluation of text digital watermarking. J. Comput., 5: 1933-1941. DOI: 10.4304/jcp.5.12.1933-1941

Zhou, X., W. Zhao, Z. Wang and L. Pan, 2009. Security theory and attack analysis for text watermarking. Proceedings of the International Conference on E-Business and Information System Security, May 23-24, Wuhan, IEEE Xplore Press, pp: 1-6. DOI: 10.1109/EBISS.2009.5138072