Computers, Materials & Continua
DOI:10.32604/cmc.2022.019686
Article

Tech Science Press

An Optimal Text Watermarking Method for Sensitive Detecting of Illegal Tampering Attacks

Anwer Mustafa Hilal[1,*], Fahd N. Al-Wesabi[2,3], Mohammed Alamgeer[4], Manar Ahmed Hamza[1], Mohammad Mahzari[5] and Murad A. Almekhlafi[6]

[1]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, AlKharj, Saudi Arabia
[2]Department of Computer Science, King Khalid University, Muhayel Aseer, Kingdom of Saudi Arabia
[3]Faculty of Computer and IT, Sana'a University, Sana'a, Yemen
[4]Department of Information Systems, King Khalid University, Muhayel Aseer, Kingdom of Saudi Arabia
[5]Department of English, College of Science & Humanities, Prince Sattam Bin Abdulaziz University, AlKharj, Saudi Arabia
[6]Faculty of Engineering, Department of Electrical Engineering, Communication Engineering, Sana'a University, Yemen
[*]Corresponding Author: Anwer Mustafa Hilal. Email: a.hilal@psau.edu.sa
Received: 22 April 2021; Accepted: 26 May 2021

Abstract: Due to the rapid increase in the exchange of text information via internet networks, the security and authenticity of digital content have become a major research issue. The main challenges faced by researchers are how to hide the information within the text to use it later for authentication and attacks tampering detection without effects on the meaning and size of the given digital text. In this paper, an efficient text-based watermarking method has been proposed for detecting the illegal tampering attacks on the Arabic text transmitted online via an Internet network. Towards this purpose, the accuracy of tampering detection and watermark robustness has been improved of the proposed method as compared with the existing approaches. In the proposed method, both embedding and extracting of the watermark are logically implemented, which causes no change in the digital text. This is achieved by using the third level and alphanumeric strategy of the Markov model as a text analysis technique for analyzing the Arabic contents to obtain its features which are considered as the digital watermark. This digital watermark will be used later to detecting any tampering of illegal attack on the received Arabic text. An extensive set of experiments using four data sets of varying lengths proves the effectiveness of our approach in terms of detection accuracy, robustness, and effectiveness under multiple random locations of the common tampering attacks.

Keywords: Text analysis; text-watermarking; tampering detection; text authentication

1 Introduction

Text authentication and tampering detection of the digital text in various languages and applications have gained great importance because the transfer of digital contents via the internet

through various applications such as eCommerce, eBanking, eLearning and other internet applications, and communication technologies are growing rapidly. Most of these digital texts are very sensitive to changes in terms of contents, structure, syntax, and semantic. Malicious attackers may temper these digital contents during the transfer process, which will result to wrong decisions [1]. Extensive research work to develop algorithms and techniques for accomplishing information security such as content authentication, integrity verification, tampering detection, owner identification, access control, and copyright protection are in progress. Steganography and digital watermarking are the main techniques used to solve these problems in information security. In digital watermarking, information such as text, binary image, video, and audio is embedded as a watermark key in those digital contents [2].

For information security, many algorithms and techniques are available such as the authentication of content, verification of integrity, detection of tampering, identification of owners, access control, and copyright protection. To overcome these issues, steganography and automated methods of watermarking are commonly used [3]. A technique of digital watermarking (DWM) can be inserted into digital material through various details such as text, binary pictures, audio, and video. A fine-grained text watermarking procedure is proposed based on replacing the white spaces and Latin symbols with homoglyph characters [4].

Several conventional methods and solutions for text watermarking were proposed [5,6] and categorized into different classifications such as linguistic, structure and image-based, and format-based binary images [7]. To insert the watermark information into the document, most of these solutions require certain upgrades or improvements to the original text in digital format material. Zero-watermarking without any alteration to the original digital material to embed the watermark information is a new technique with smart algorithms that can be used. Also, this technique can be used to generate data for a watermark in the contents of a given digital context [1,7–9].

Restricted research has centered on the appropriate solutions to verify the credibility of critical digital media online [10–12]. The verification of digital text and the identification of fraud in research earned great attention. In addition, text watermarking studies have concentrated on copyright protection in the last decade, but less interest and attention has been paid to integrity verification, identification of tampering and authentication of content due to the existence of text content based on the natural language [13].

Proposing the most appropriate approaches and strategies for dissimilar formats and materials, especially in Arabic and English languages, is the most common challenge in this area [14,15]. Therefore, authentication of content, verification of honesty, and detection of tampering of the sensitive text constitute a big problem in various applications and require necessary solutions.

Some instances of such sensitive digital text content are Arabic interactive Holy Qur'an, eChecks, tests, and marks. Different Arabic alphabet characteristics such as diacritics lengthened letters and extra symbols of Arabic make it simple to modify the key meaning of the text material by making basic changes such as modifying diacritic arrangements [16]. The most popular soft computation and natural language processing (NLP) technique that supported the analysis of the text is HMM.

We suggest an intelligent hybrid approach named "an efficient text-watermarking method based on the third level and alphanumeric strategy of Markov model (ETWMM)" for Arabic text authentication and tampering detection. Towards this purpose, text-watermarking and Markov model techniques have been integrated. In this method, the third level of the alphanumeric strategy of the Markov model was used for text analysis and extract the features of the given

Arabic text to use it as a watermark key. Without any alterations or effects on the original text size, the watermark created is logically integrated into the original Arabic history. The embedded watermark would later be used to identify all manipulation on Arabic text obtained after transmission of text through the Internet and whether it is authentic or not.

The primary objective of the ETWMM method is to provide the high accuracy of tampering detection of illegal attack on Arabic text which is transmitted via the Internet network.

The remainder of the article is structured as follows: In Section 2, we explain the existing works done so far. In Section 3, we discussed the suggested method (ETWMM). The simulation and implementation are provided in Section 4, results discussion is provided in Section 5, and finally, we conclude the article in Section 6.

2 Related Work

According to the processing domain of NLP and text watermarking, these existing methods and solutions of text watermarking reviewed in this paper are classified into linguistical, structural, and zero-watermark methods [1,7,13].

Natural language is the foundation of approaches to linguistic text watermarking. The mechanism of those methods embedding the watermark is based on changes applied to the semantic and syntactic essence of plain text [1].

To enhance the capability and imperceptibility of Arabic text, a method of text watermarking has been suggested based on the location of the accessible words [17]. In this method, any word-space is used to mask the Boolean bit 0 or 1 that physically modifies the original text.

A text steganography technique was proposed to hide information in the Arabic language [18]. The step of this approach considers Harakat's existence in Arabic diacritics such as Kasra, Fatha, and Damma as well as reverses Fatha to cover the message.

A Kashida-marks invisible method of watermarking [19], based on the features of frequent recurrence of document security and authentication characters, was proposed. The method is based on a predetermined watermark key with a Kashida placed for a bit 1 and a bit omitted.

The method of steganography of the text has been proposed based on Kashida extensions on the characters 'moon' and 'sun' to write digital contents of the Arabic language [20]. In addition, the Kashida method characters are seen alongside characters from Arabic to decide which hidden secret bits are kept by specific characters. In this form, four instances are included in the kashida characters: moon characters representing '00'; sun characters representing '01'; sun characters representing '10'; and moon characters representing '11'.

A text steganographic approach [21] based on multilingual Unicode characters has been suggested to cover details in English scripts for the use of the English Unicode alphabet in other languages. Thirteen letters of the English alphabet have been chosen for this approach. It is important to embed dual bits in a timeframe that used ASCII code for embedding 00. However, multilingual ones were used by Unicode to embed between 01, and 10, as well as 11. The algorithm of text watermarking is used to secure textual contents from malicious attacks according to Unicode extended characters [22]. The algorithm requires three main steps, the development, incorporation, and extraction of watermarks. The addition of watermarks is focused on the development of predefined coding tables while scrambling strategies are often used in generating and removing the watermarking key is safe.

The substitution attack method focused on preserving the position of words in the text document has been proposed [23]. This method depends on manipulating word transitions in the text document. Authentication of Chinese text documents based on the combination of the properties of sentences and text-based watermarking approaches have been suggested [24,25]. The proposed method is presented as follows: a text of the Chinese language is split into a group of sentences, and for each word, the code of a semantic has been obtained. The distribution of semantic codes influences sentence entropy.

A zero-watermarking method has been proposed to preserve the privacy of a person who relies on the Hurst exponent and the nullity of the frames [26]. For watermark embedding, the two steps are determined to evaluate the unvoiced frames. The process of the proposed approach bases on integrating an individual's identity without notifying any distortion in the signals of medical expression.

A zero-watermarking method was proposed to resolve the security issues of text documents of the English language, such as verification of content and copyright protection [27]. A zero-watermarking approach has been suggested based on the authentication Markov model of the content of English text [28,29]. In this approach, to extract the safe watermark information, the probability characteristics of the English text are involved and stored to confirm the validity of the attacked text document. The approach provides security against popular text attacks with a watermark distortion rate if, for all known attacks, it is greater than one. For the defense of English text by copyright, based on the present rate of ASCII non-vowel letters and terms, the conventional watermark approach [30] has been suggested.

According to the suggested methods, content authentication and tampering detection of digital English contents that have been ignored by researchers in the literature for many reasons. English text is natural language-dependent. On the other hand, hiding the watermark information is complicated since there is no location to hide it within the text as pixels in the case of an image, waves in audio, and frames in a video.

3 The Proposed Approach

In this paper, the authors propose an efficient method called ETWMM. Hence, the third level of alphanumeric strategy consisting of a model performing as a soft computing tool and NLP in cooperation between the zero-watermarking technique and the Markov model. The third level of alphanumeric strategy has been selected in our proposed ETWMM method to improve the accuracy of random detection of tampering attacks and reduced the gap of interrelationships between the given numbers, special symbols and Arabic alphabets as compared with other levels. Markov model is used in our proposed ETWMM method for text analysis to extract the features of the given Arabic text with getting rid of external watermark information and without any modifications of the original text to embed the watermark key. Unlike the previous work, the proposed approach ETWMM can effectively detect any tampering whenever tampering volume was very low or very high. In addition, ETWMM can be improved to determine the place of tempering occurrence. This feature can be considered an advantage over Hash function method.

The following subsections explain in detail two main processes that should be performed in ETWMM. Which are watermark generation and embedding process, and watermark extraction and detection process.

3.1 Watermark Generation and Embedding Process

The three sub-algorithms included in this process are pre-processing algorithm, text feature extraction and watermark generation algorithm, and watermark embedding algorithm as illustrated in Fig. 1.

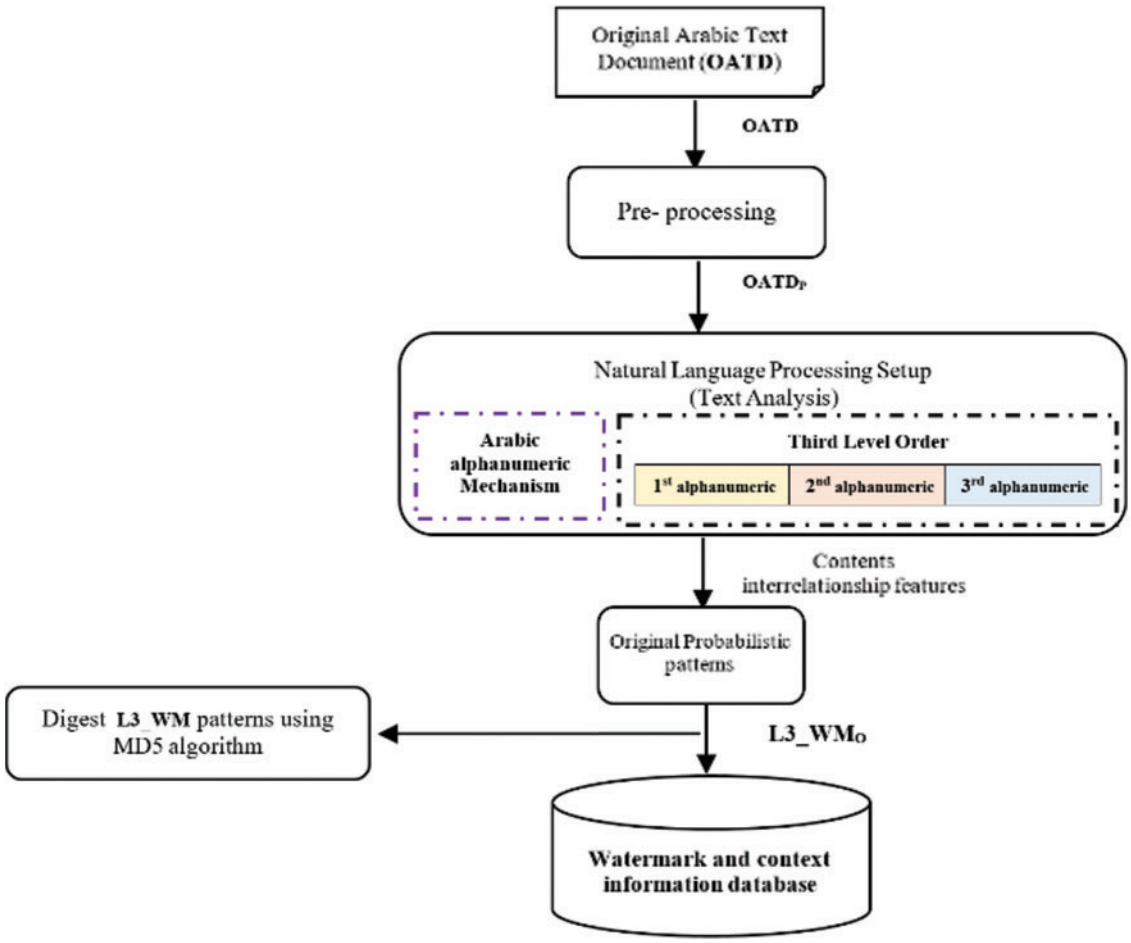

Figure 1: Text-watermark generation and embedding processes of ETWMM

3.1.1 Pre-Processing Algorithm

Preprocessing of the original Arabic text is one of the key steps in both the watermark generation and extraction processes to remove extra spaces and new-lines, and it will directly influence the tampering detection accuracy and watermark robustness. The original Arabic text (OATD) is required as input for Pre-processing process.

3.1.2 Text Feature Extraction and Watermark Generation Algorithm

This algorithm includes two sub-processes are building the Markov matrix, and text features extraction and watermark generation processes.

— Building the Markov matrix is the starting point of Arabic text analysis and watermark generation process using the Markov model. A Markov matrix that represents the possible

states and transitions available in the given text is constructed without reputations. In this approach, each triple alphanumeric within a given Arabic text represents a present state, and each unique alphanumeric a transition in the Markov matrix. During the building process of the Markov matrix, the proposed algorithm initializes all transition values by zero to use these cells later to keep track of the number of times that the i^{th} triple alphanumeric is followed by the j^{th} alphanumeric within the given Arabic text document.

Pre-processing and building Markov matrix algorithm executes as presented below in Algorithm 1.

Algorithm 1: Algorithm of preprocessing and building Markov matrix of ETWMM

```
PROCEDURE PB_MM (OATD)

Input: original Arabic text (OATD)

Output: Markov matrix with zeros initial value (OATDP)

1.  BEGIN
2.  // perform pre-processing process
3.  for each alphanumeric in OATD
4.          // remove new lines and spaces letters
5.       OATDP ← trim ("space" or "newLine")
6.  // Build list of non values text alphanumerics
7.  L3_mm = { }
8.  for each alphanumeric in OATDP
9.      if alphanumeric not in a1_list
10.         L3_mm ← L3_mm U {alphanumeric}
11.     for ps = 1 to L3_mm.length – 3
12.         for ns = 1 to L3_mm.length
13.             L3_mm[ps][ns] = 0
14. return L3_mm
```

where OATD is an original Arabic text, $OATD_P$ is a pre-processed Arabic text, L3_mm refers to states and transitions matrix with zeros values for all cells, ps: refers to the current state, ns: refers to next state.

According to the above, a method is presented to construct a two-dimensional matrix of Markov states and transitions named L3_mm[i][j], which represents the core part of the Markov model for Arabic text analysis. The length of L3_mm[i][j] a1_mm[i][j] matrix of ETWMM is fixed in which the number of states and transitions equal to the total of alphabetic Arabic characters, numbers from 0 to 9, and special symbols and Harakat (Fatha, Kasra, Dommah, … etc).

— Text features extraction and watermark generation process: after the Markov matrix was constructed, the Arabic text analysis process should be performed to extract features of the given Arabic text and generate watermark patterns. In this algorithm, the number of appearances of possible next states transitions for each current state of a single alphanumeric will be calculated and constructed as transition probabilities by Eq. (1) below.

$$L3\_mm[ps][ns] = \sum_{i,j=1}^{n-3} L3transitions(i,j) \tag{1}$$

where n is number of states, i: is i^{th} current state of a single alphanumeric, j: is j^{th} next state transition.

The following example of the Arabic text sample describes the mechanism of the transition process of the present state to other next states.

"يقفز الثعلب البني السريع فوق الثعلب البني البطيء للوصول إلى الثعلب البني الميت."

When using the third level order of alphanumeric mechanism of Hidden Markov model, every unique triple of Arabic alphanumeric is a present state. Text analysis is processed as the text is read to obtain the interrelationship between the present state and the next states. Fig. 2. below illustrates the available transitions and analysis results of the above sample of Arabic text.

Figure 2: The states representation of the given Arabic text sample using ETWMM

As a result of analysing the given Arabic sentence based on third level order and alphanumeric mechanism of the Markov model, we represent 47 unique present states and their 61 possible transitions as illustrated in Fig. 3.

State ID	Present States	Next Transition(s)
1	("يقف")	[ز]
2	("قفز")	[ ]
3	("فز ")	[ا]
4	("ز ا")	[ل]
...	.....	.....
...	.....	.....
45	("الم")	[ي]
46	("لمي")	[ت]
47	("ميت")	[.]

Figure 3: Sample of an Arabic text states and their transitions in ETWMM

As illustrated in Fig. 3 above, we assume "يقف" is a present state of three continuous alphanumerics, and the available next transition is "ز". We observe that one transition only available in the given Arabic text sample.

Text analysis and watermark generation algorithm are presented formally and executed as illustrated in Algorithm 2.

Algorithm 2: Watermark generation algorithm of ETWMM

PROCEDURE WMG(OATD$_P$)

```
1.  Input: OATDP
2.  Output: fm
3.  BEGIN
4.  PB_MM (OATDP)
5.  pa = first  trip  alpha(OATDP)
6.  pd2 = OATDP – [pa] // begin with 2nd triple of alphanumeric
7.  fm = L3  mm
8.  for each a in pd2
9.        fm[pa][ca] = fm[pa][a] + 1
10.       pa = ca
11. return fm
```

where pa: previous triple of alphanumeric, ca: current triple of alphanumeric.

The algorithm of text feature extraction and watermark generation based on the third level order of the alphanumeric mechanism of the Markov model proceeds as illustrated in Fig. 4.

St	Avalible Transitions in Given Arabic Text																																				
	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط	ظ	ع	غ	ف	ق	ك	ل	م	ن	ه	و	ي	0	1	2	3	4	5	6
يقف	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
قفز	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
فز	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ز ا	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
الم	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
لمي	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ميث	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4: Text feature extraction and watermark generation of the given Arabic text using ETWMM

3.1.3 Watermark Embedding Algorithm

In our proposed ETWMM method, the watermark key will be generated as a result of text analysis and text feature extraction process by finding all non-zero values in the Markov matrix. All of these non-zero values will be concatenated sequentially to generate the original watermark pattern L3_WM$_O$, as given in Eq. (2) and illustrated in Fig. 5.

$$L3\_WM_O\& = L3\_mm[ps][ns] \tag{2}$$

1 – 1 – 1 – 1 – 1 – 1 – 1

Figure 5: Sample of the generated watermark patterns L3_WM$_O$ in a decimal form using ETWMM

The algorithm of watermark embedding based on the third level of alphanumeric strategy of the Markov model is presented formally and executed as illustrated below in Algorithm 3.

Algorithm 3: Watermark embedding algorithm of ETWMM

```
PROCEDURE WME (PET)

- Input: pre-processed text (OATDP)
- Output: original watermark patterns L3_WMO

1. BEGIN
2. WMG (OATDP)
3. for ps = 1 to L3_arrList.Length - 3,
4.     for ns = 1 to L3_arrList.Length,
5.         if L3_mm[ps][ns] != 0
6.             L3_WMO &= L3_MM [ps] [ns]
7. return L3_WMO
```

where $L3\_WM_O$ is the original watermark pattern.

3.2 Watermark Extraction and Detection Process

Before the detection of pre-proceed attacked Arabic text ($AATD_P$), attacked watermark patterns ($L3\_EWM_A$) should be generated, and matching rate of patterns and watermark distortion should be calculated by ETWMM for detecting any tampering with the authentication of the given Arabic text.

Two core algorithms are involved in this process, which are watermark extraction and watermark detection. However, $L3\_EWM_A$ will be extracted from the received ($AATD_P$) and matched with $L3\_WM_O$ by detection algorithm.

$AATD_P$ should be provided as the input for the proposed watermark extraction algorithm. The same process of watermark generation algorithm should have been performed to obtain the watermark pattern for ($AATD_P$) as illustrated in Fig. 6.

3.2.1 Watermark Extraction Algorithm

$AATD_P$ is the main input required to run this algorithm. However, the output of this algorithm is $L3\_EWM_A$. The watermark extraction algorithm is presented formally as illustrated in Algorithm 4.

Algorithm 4: Watermark extraction algorithm of ETWMM

```
PROCEDURE WMEX (AATDP)

- Input: pre-processed text (AATDP)
- Output: attacked watermark patterns (L3_EWMA).

1. BEGIN
2. WMG (AATDP)
3. for ps = 1 to L3_arrList'.Length - 3,
4.     for ns = 1 to L3_arrList'.Length,
5.         if L3_MM'[ps][ns] != 0,
6.             L3_EWMA &= L3_MM'[ps] [ns],
7. return L3_EWMA
```

where $L3\_EWM_A$ is the attacked watermark.

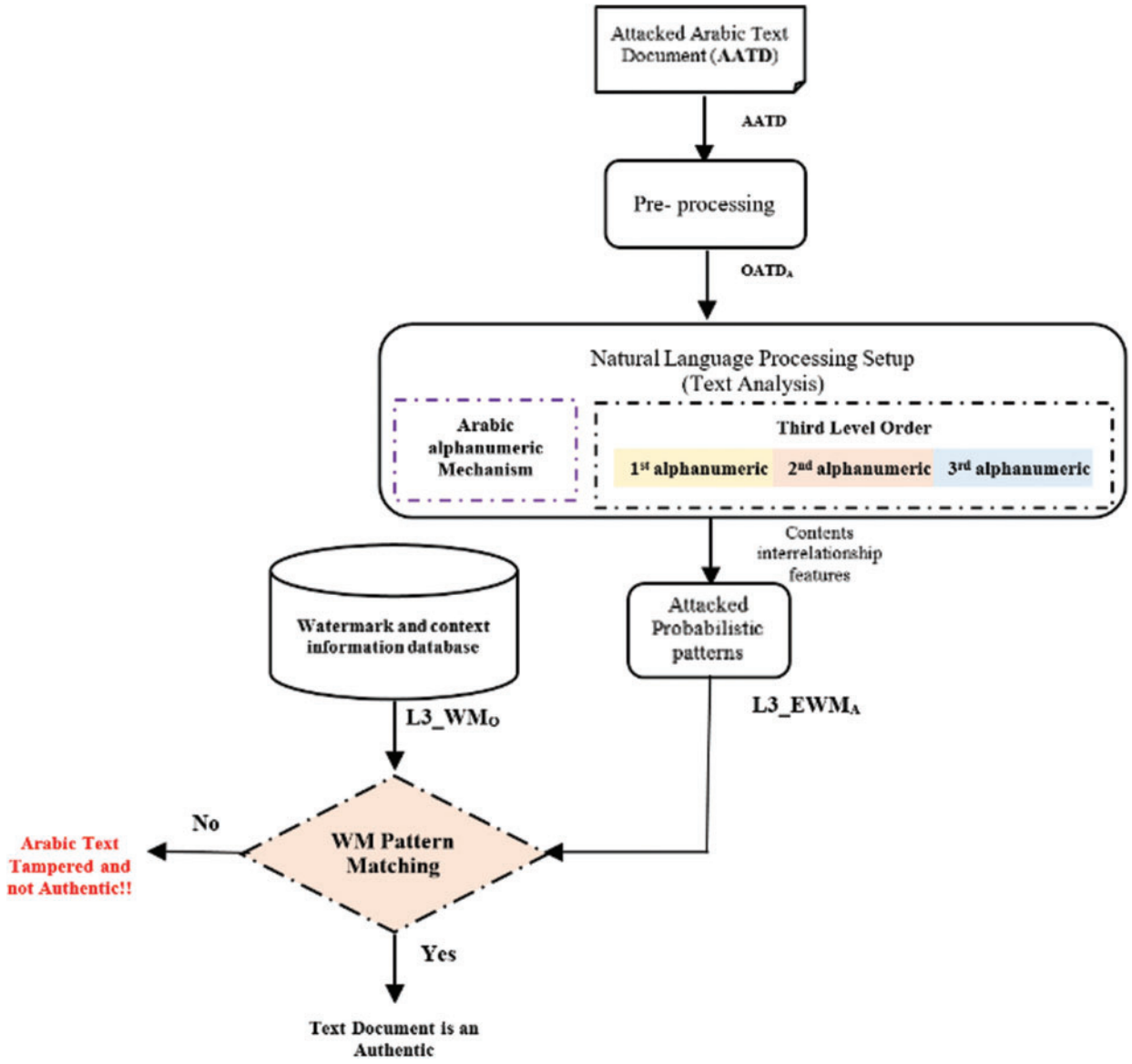

Figure 6: Text-watermark detection process of ETWMM

3.2.2 Watermark Detection Algorithm

$L3\_EWM_A$ and $L3\_WM_O$ are the main inputs required to run this algorithm, while the output of this algorithm is the notification Arabic text document, which can be authentic or tampered. The detection process of the extracted watermark is achieved in two main steps:

- Primary matching is achieved for $L3\_WM_O$ and $L3\_EWM_A$. If these two patterns appear identical, then an alert will appear as "Arabic text is an authentic and no tampering occurred." Otherwise, the notification will be "Arabic text is tampered," and then it continues to the next step.
- Secondary matching is achieved by matching the transition of each state in a generated pattern. This means that $L3\_EWM_A$ of each state is compared with the equivalent transition of $L3\_WM_O$ as given by Eqs. (3) and (4) below

$$L3\_PMR_T(i,j) = \left| \frac{L3\_WM_O[i][j] - (L3\_WM_O[i][j] - L3\_EWM_A[i][j])}{L3\_WM_O[i][j]} \right| \tag{3}$$

where $L3\_PMR_T$ represents tampering detection accuracy rate value in transition level.

$$L3\_PMR_S(i) = \left| \frac{\sum_{j=1}^{n-3} (L3\_PMR_T(i,j))}{Total\ \ State\ \ Pattern\ \ Count(i)} \right| \tag{4}$$

where $L3\_PMR_S$ is value of tampering detection accuracy rate in state level.

After the pattern matching rate of every state has been produced, we have to find the weight of every state stored in the Markov matrix as presented in Eq. (5) below.

$$L3\_sw = \left| \frac{L3\_PMR_S(i) * Transitions\ frequency\ (i)}{total\ number\ of\ transitions} \right| \tag{5}$$

where L3_sw refers to total tampering detection accuracy and watermark robustness.

The final L3_PMR of $AATD_P$ and $OATD_P$ are calculated by Eq. (6).

$$L3\_PMR = \left| \frac{\sum_{i=1}^{n-3} L3\_PMRS(i)}{N} \right| * 100 \tag{6}$$

The rate of watermark distortion represents the rate of tampering attacks occurring on the contents of the attacked Arabic context, which is denoted by L3_WDR and calculated by Eq. (7).

$$L3\_WDR = 1 - L3\_PMR * 100 \tag{7}$$

The steps involved in the watermark detection algorithm are illustrated in algorithm Algorithm 5.

Algorithm 5: Watermark detection algorithm of ETWMM

PROCEDURE WMD ($L3\_WM_O$, $L3\_EWM_A$)

- Input: pre-processed text ($L3\_WM_O$, $L3\_EWM_A$)
- Output: L3_PMR, L3_WDR

1. BEGIN
2. WMG ($L3\_WM_O$)
3. WMEX ($L3\_EWM_A$)
4. **IF** $L3\_EWM_A = L3\_WM_O$
5. Print "Arabic document is authentic and no tampering occurred"
6. L3_PMR = 100
7. **Else**
8. Print "Arabic document is not authentic and tampering occurred"
9. **for** i = 1 **to** L3_arrList'.Length - 3,
10. **for** j = 1 **to** L3_arrList'.Length
11. **IF** $L3\_WM_O[i][j] != 0$
12. pattern Count + = 1
13. $L3\_PMR_T(i,j) = \left|\frac{L3\_WM_O[i][j] - (L3\_WM_O[i][j] - L3\_EWM_A[i][j])}{L3\_WM_O[i][j]}\right|$
14. transPMRTotal + = $L3\_PMR_T$
15. **Else IF** $L3\_EWM_A[i][j] != 0$
16. patternCount + = $L3\_EWM_A[i][j]$
17. $L3\_PMR_S(i) = \left|\frac{\sum_{j=1}^{n-3}(L3\_PMR_T(i,j))}{Total\ StatePatternCount(i)}\right|$
18. $sWeight = \frac{L3\_PMR_S(i) * Transitions\ frequency(i)}{total\ no\ of\ transitions}$
19. L3_SW + = stateWeight
20. $L3\_PMR = \frac{\sum_{i=1}^{n-3}(L3\_SW) * Total\ number\ of\ transitions}{Total\ number\ of\ transitions} * 100$
21. L3_WDR = 1 – L3_PMR * 100
22. **return** L3_PMR, L3_WDR

where L3_SW refers to the weight value of states correctly matched, L3_WDR: refers to the value of watermark distortion rate ($0 < L3\_WDR_S <= 100$).

States	Original WM Patterns	Extracted WM Patterns	Destroyed WM Patterns	Primary Matching Rate	L3_PMRt(I,j) of Transition Level								L3_PMRs(I,j) of State Level
					TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	
"يقف"	1	2	2	-	0.5	-	-	-	-	-	-	-	0.5
"قفز"	1	1	1	1	1	-	-	-	-	-	-	-	1
"فز "	1	1.1	1.1	-	0.5	-	-	-	-	-	-	-	0.5
"ز ا"	1	3	3	-	0.33	-	-	-	-	-	-	-	0.33
.....	-	-	-	-	-	-	-	-	-	-	-	-	0
.....	-	-	-	-	-	-	-	-	-	-	-	-	0
"الم"	1	1	1	1	1	-	-	-	-	-	-	-	1
"لمي"	1	2	2	-	0.5	-	-	-	-	-	-	-	0.5
"ميت"	1	1	1	1	1	-	-	-	-	-	-	-	1
"يت."	1	1	1	1	1	-	-	-	-	-	-	-	1
Total L3_PMR =													**0.7288**

Figure 7: Sample of watermark extraction and detection process using ETWMM

The results of the watermark extraction and detection process are illustrated in Fig. 7.

4 Implementation and Simulation

To evaluate the tampering detection accuracy and robustness of our ETWMM method, several scenarios of simulation and experiments are performed. This section depicts an implementation, simulation, and experimental environment, experiment parameters, experimental scenarios of standard Arabic datasets, and results discussion.

4.1 Simulation and Implementation Environment

The self-developed program has been developed to test and evaluate the tampering detection accuracy and robustness of ETWMM. The implementation environment of ETWMM is: CPU: Intel Core i7-4650U/2.3 GHz, RAM: 8.0 GB, Windows 10–64 bit, PHP programming language with VS Code IDE.

4.2 ETWMM Simulation and Experiment

This subsection presents the tampering detection accuracy and robustness evaluation of ETWMM. Many simulations and experiment scenarios are performed as shown in Tab. 1, for all forms of attacks and their volumes.

Table 1: Detection accuracy and robustness evaluation of ETWMM under all attacks with various volumes

Attack volume (%)	Insertion	Deletion	Reorder
5	87.57	91.13	93.35
10	83.02	93.10	82.74
20	68.80	82.77	75.21
s50	49.12	59.92	61.74

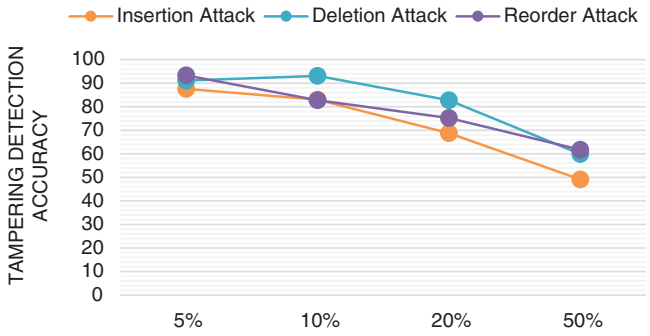

Figure 8: Detection accuracy and robustness evaluation under all attacks with various volumes

In case of attack volume effects against all dataset sizes, results shown in Tab. 1 above and Fig. 8 below, high effect detected under reorder and deletion attacks in case of small, mid, and high attack volumes. Results shows also reorder attacks more sensitive than deletion attacks in case of medium attack volumes, however, deletion attacks more sensitive than reorder attack in

cases of low and high attack volumes. This mean ETWMM gives the best detection accuracy at all under insertion attack in all scenarios of attack volumes.

5 Comparison and Result Discussion

The detection accuracy and robustness results were critically analyzed. This subsection displays an effect study and a comparison between ETWMM and baseline approaches named Hybrid of Natural Language Processing and Zero-Watermarking Approach (HNLPZWA) [5] and Zero-Watermarking Approach based on Fourth level order of Arabic Word Mechanism of Markov Model (ZWAFWMMM) [6]. It also contains a discussion of their effect under the major factors, namely dataset size, attack types and volumes.

5.1 Attacks Type-Based Comparison

Tab. 2 shows a comparison of the different attack's type of effect on detection accuracy and robustness of ETWMM, ZWAFWMMM, and HNLPZWA approaches against all dataset sizes and all scenarios of attack volumes.

Table 2: Comparison based on attacks type effect

Method	HNLPZWA	ZWAFWMMM	ETWMM
Insertion	80.50	80.02	83.81
Deletion	70.45	66.25	82.60
Reorder	72.13	81.73	78.26

Tab. 2 and Fig. 9 show how the detection accuracy and robustness of ETWMM, HNLPZWA, and ZWAFWMMM approaches is influenced by analyzing the attack types. In all cases of deletion and reorder attack, ETWMM approach outperforms FAWMW and HNLPZWA in terms of general detection accuracy and robustness in all scenarios of all attacks. HNLPZWA and ZWAFWMMM give the best detection accuracy and robustness rate in case of insertion attack. This means that ETWMM approach is strongly recommended and applicable for content authentication and tampering detection of Arabic text transmitted via the internet especially under deletion and reorder attacks.

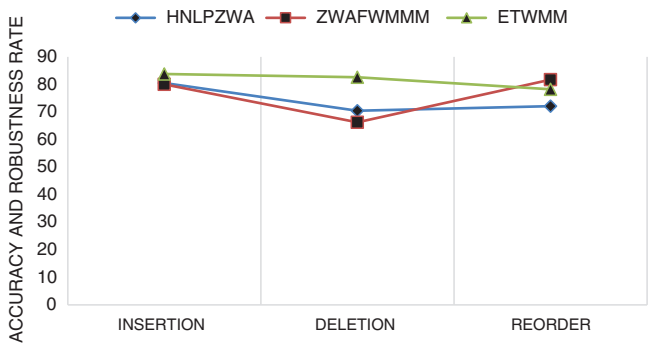

Figure 9: A compression based on attacks type effect of our ETWMM method and baseline approaches

5.2 Attacks Volume-Based Comparison

Tab. 3 provides a comparison of the different attack volumes' effect on detection accuracy and robustness against all dataset sizes and all scenarios of attack volumes. The comparison is performed using ETWMM, ZWAFWMMM, and HNLPZWA approaches.

Table 3: Comparison based on attacks volume effect

Attack volume (%)	HNLPZWA	ZWAFWMMM	ETWMM
5	84.98	82.09	90.68
10	76.21	72.74	86.29
20	61.46	57.71	75.59
50	39.57	13.66	56.93

Tab. 3 and Fig. 10 show how the detection accuracy and robustness of ETWMM, HNLPZWA, and ZWAFWMMM approaches is influenced by analyzing the attack volumes. In Fig. 10, it can be seen that if the attack volume increases, the detection accuracy and robustness decrease. It is seen also, ETWMM approach outperforms both HNLPZWA and ZWAFWMMM approaches in terms of general detection accuracy and robustness in all scenarios of low, mid, and high volumes of all attacks. This means that ETWMM approach is strongly recommended and applicable for tampering detection of Arabic text under all volumes of all attacks.

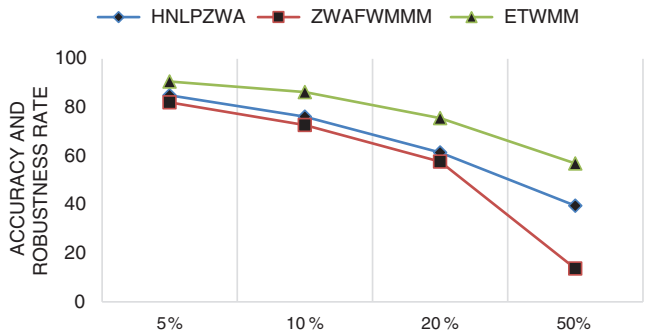

Figure 10: A compression based on attacks volume effect of our ETWMM method and baseline approaches

5.3 Dataset Size-Based Comparison

This section presents a comparison of the various dataset size effect on detection accuracy and robustness against all forms of attacks within their multiple volumes as shown in Tab. 4.

Fig. 11 shows how the detection accuracy and robustness of ETWMM, HNLPZWA, and ZWAFWMMM approaches is influenced by analyzing the dataset size. In Fig. 11, it can be seen that in all cases of baseline HNLPZWA and ZWAFWMMM approaches, the detection accuracy and robustness increased with decreasing dataset size and decreased with increasing dataset size. However, in the proposed approach, we can say the detection accuracy and robustness are increased with increasing document size. On the other hand, results show that ETWMM approach outperforms both HNLPZWA and ZWAFWMMM approaches in terms of general

detection accuracy and robustness under all scenarios of dataset sizes. This means that the proposed ETWMM approach is strongly recommended for tampering detection with low, mid, and large document size.

Table 4: Comparison based on dataset size effect

Dataset size	HNLPZWA	ZWAFWMMM	ETWMM
[ASST]	71.33	69.53	75.19
[AMST]	69.93	68.13	76.98
[AHMST]	66.90	65.11	81.83
[ALST]	63.94	62.07	75.49

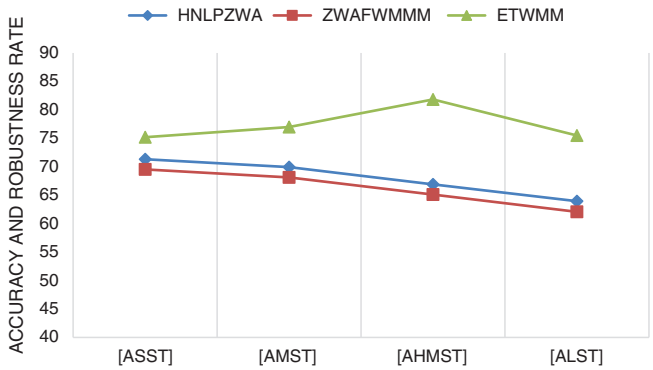

Figure 11: A compression based on dataset effect of our ETWMM method and baseline approaches

6 Conclusion

In this paper, ETWMM method has been proposed for detecting illegal tampering attacks on the Arabic text by combining text-watermarking and natural language processing techniques. A text analysis process should be performed to extract the features of the given Arabic text and generate a watermark key. The generated watermark will be embedded logically in the original Arabic context without modifications and effect on the size of the original text. The embedded watermark will be used later after transmission of text via the Internet to detect any tampering that occurred on received Arabic text and it is authentic or not. ETWMM approach has implemented using PHP zero program developed in VS code IDE and experiments using various standard datasets under different volumes of insertion, deletion, and reorder attacks. We have compared ETWMM with HNLPZWA and ZWAFWMMM. Comparison results show that ETWMM outperforms HNLPZWA and ZWAFWMMM in terms of robustness and accuracy of tampering detection because using the third-level order and alphanumeric mechanism of HMM leads to better robustness and accuracy of tampering detection in which the third level of interrelationships between alphanumeric is stronger than other levels of alphanumeric or words mechanisms of HMM. Also, results show that ETWMM is applicable to all Arabic alphabetic letters, special characters, numbers, and spaces. Although ETWMM is an efficient approach, and

it is designed only for all scenarios of deletion and reorder attacks. For future work, we will consider detection accuracy under all scenarios of insertion attacks. Moreover, we also intend to evaluate the performance using other information security and artificial intelligence techniques.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under Grant Number (RGP.1/53/42), Received by Mohammed Alamgeer. www.kku.edu.sa

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] F. N. Al-Wesabi, "A smart English text zero-watermarking approach based on third-level order and word mechanism of Markov model," *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1137–1156, 2020.

[2] M. Abd-Eldayem, "A proposed security technique based on watermarking and encryption for digital imaging and communications in medicine," *Egyptian Informatics Journal*, vol. 14, no. 1, pp. 1–13, 2013.

[3] F. N. Al-Wesabi, "A hybrid intelligent approach for content authentication and tampering detection of Arabic text transmitted via internet," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 195–2011, 2021.

[4] S. G. Rizzo, F. Bertini and D. Montesi, "Fine-grain watermarking for intellectual property protection," *EURASIP Journal on Information Security*, vol. 10, no. 1, pp. 804, 2019.

[5] F. N. Al-Wesabi, "Proposing high-smart approach for content authentication and tampering detection of Arabic text transmitted via internet," *IEICE Transactions in Information Systems*, vol. E103, no. 10, pp. 2104–2112, 2020.

[6] F. N. Al-Wesabi, K. Mahmood and N. Nemri, "A zero watermarking approach for content authentication and tampering detection of Arabic text based on fourth level order and word mechanism of Markov model," *Journal of Information Security and Applications*, vol. 52, no. 1, pp. 1–15, 2020.

[7] P. Selvama, S. Balachandran, S. Pitchai and R. Jayabal, "Hybrid transform based reversible watermarking technique for medical images in telemedicine applications," *ELSEVIER Optik*, vol. 145, no. 5, pp. 655–671, 2017.

[8] N. Hurrah, A. Parah, N. Loan, A. Sheikh, M. Elhoseny *et al.,* "Dual watermarking framework for privacy protection and content authentication of multimedia," *ELSEVIER Future Generation Computer Systems*, vol. 94, pp. 654–673, 2019.

[9] A. Panah, R. Van, T. Sellis and E. Bertino, "On the properties of non-media digital watermarking: A review of state-of-the-art techniques," *IEEE Access*, vol. 4, pp. 2670–2704, 2016.

[10] C. Qin, C. Chang and T. Hsu, "Fragile watermarking for image authentication with high-quality recovery capability," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 11, pp. 2941–2956, 2013.

[11] S. Parah, J. Sheikh and G. Bhat, *StegNmark: A Joint Stego-Watermark Approach for Early Tamper Detection*, vol. 660. Switzerland: Springer International Publishing, pp. 427–452, 2017.

[12] S. Hakak, A. Kamsin, O. Tayan, M. Yamani and G. Gilkar, "Approaches for preserving content integrity of sensitive online Arabic content," *Information Processing and Management*, vol. 56, no. 2, pp. 367–380, 2019.

[13] M. Taleby, Q. Li, X. Zhu, M. Alazab and J. Zhang, "A Novel intelligent text watermarking technique for forensic identification of information on social media," *Computers and Security*, vol. 90, pp. 1–14, 2020.

[14] S. Parah, J. Sheikh, J. Akhoon and N. Loan, "Electronic health record hiding in images for smart city applications: A computationally efficient and reversible information hiding technique for secure communication," *ELSEVIER Future Generation Computer Systems*, vol. 108, no. 6, pp. 935–949, 2020.

[15] R. Ahmed and L. Elrefaei, "Arabic text watermarking: A review," *International Journal of Artificial Intelligence & Applications*, vol. 6, no. 4, pp. 1–16, 2015.

[16] K. Hameed, A. Khan, M. Ahmed and A. G. Reddy, "Towards a formally verified zero watermarking scheme for data integrity in the internet of things based-wireless sensor networks," *ELSEVIER Future Generation Computer Systems*, vol. 167, pp. 1–16, 2018.

[17] R. Alotaibi and L. Elrefaei, "Improved capacity text watermarking methods based on open word space," *Journal of King Saud University–Computer and Information Sciences*, vol. 30, no. 2, pp. 236–248, 2018.

[18] M. Memon and A. Shah, "A novel text steganography technique to Arabic language using reverse fat5th5ta," *Pakistan Journal of Engineering, Technology and Sciences*, vol. 1, no. 2, pp. 106–113, 2015.

[19] Y. Alginahi, M. Kabir and O. Tayan, "An enhanced Kashida-based watermarking approach for increased protection in arabic text-documents based on frequency recurrence of characters," *International Journal of Computer and Electrical Engineering*, vol. 6, no. 5, pp. 381–392, 2014.

[20] A. Shaker, F. Ridzuan and S. Pitchay, "Text steganography using extensions Kashida based on moon and sun letters," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 8, pp. 286–290, 2017.

[21] A. Rahma, W. Bhaya and D. Al-Nasrawi, "Text steganography based on Unicode of characters in multilingual," *Journal of Engineering Research and Applications*, vol. 3, no. 4, pp. 1153–1165, 2013.

[22] N. Al-maweri, W. Adnan, A. Rahman, S. Khair and S. Syed, "Robust digital text watermarking algorithm based on Unicode characters," *Indian Journal of Science and Technology*, vol. 9, no. 48, pp. 1–14, 2016.

[23] M. Bashardoost, M. Rahim, T. Saba and A. Rehman, "Replacement attack: A new zero text watermarking attack," *3D Research*, vol. 8, no. 1, pp. 1–9, 2017.

[24] Y. Liu, Y. Zhu and G. Xin, "'A zero-watermarking algorithm based on merging features of sentences for chinese text," *Journal of the Chinese Institute of Engineers*, vol. 38, no. 3, pp. 391–398, 2015.

[25] P. Zhu, W. Song, A. Li, Y. Zhang and R. Tao, "A text zero watermarking algorithm based on chinese phonetic alphabets," *Wuhan University Journal of Natural Sciences*, vol. 21, no. 4, pp. 277–282, 2016.

[26] Z. Ali, M. Shamim, G. Muhammad and M. Aslam, "New zero-watermarking algorithm using hurst exponent for protection of privacy in telemedicine," *IEEE Access*, vol. 6, pp. 3457–3470, 2018.

[27] O. Tayan, Y. Alginahi and M. Kabir, "An adaptive zero-watermarking approach for text documents protection," *International Journal of Image Processing Techniques*, vol. 1, no. 1, pp. 33–36, 2014.

[28] M. Ghilan, F. Ba-Alwi and F. N. Al-Wesabi, "Combined Markov model and zero watermarking to enhance authentication of Arabic text," *Journal of Computational Linguistics Research*, vol. 5, no. 1, pp. 26–42, 2014.

[29] F. N. Al-Wesabi, A. Alsakaf and K. U. Vasantrao, "A zero text watermarking algorithm based on the probabilistic patterns for content authentication of text documents," *International Journal of Computer Engineering & Technology*, vol. 4, no. 1, pp. 284–300, 2013.

[30] H. Ahmed and M. Khodher, "Comparison of eight proposed security methods using linguistic steganography text," *Journal of Computing & Information Sciences*, vol. 12, no. 2, pp. 243–251, 2016.