JOURNAL LA MULTIAPP

VOL. 05, ISSUE 05 (680-691), 2024
DOI: 10.37899/journallamultiapp.v5i5.1490

Identification of Tampering in Arabic Text Documents Using Zero-Watermarking Algorithm

Shahbaa Mohammed Abdulmaged[1]

[1]*Al-Iraqia University, Law dep., Baghdad, Iraq.*

**Corresponding Author: Shahbaa Mohammed Abdulmaged*
Email: Shahbaa.abdulmaged@aliraqia.edu.iq

Article Info

Article history:
Received 22 July 2024
Received in revised form 7 August 2024
Accepted 24 September 2024

Keywords:
Zero watermarking
Security
Arabic text
Preserving
Protection

Abstract

The internet makes sharing information easy, but protecting copyrights and verifying the authentication of online text remains a challenge. Current methods for securing Arabic text are limited. This research introduces a new tool: a zero watermark algorithm designed specifically for Arabic text. This new algorithm protects Arabic text by analyzing its content. It creates a watermark based on a chosen keyword (at least 6 characters, no duplicates), and how often each character appears in different forms (separate, connected), and positions (beginning, middle, and end) within the text. The proposed method was tested on nine different Arabic text samples of varying size, simulating various tampering attempts like inserting, deleting, and rephrasing text. The results are impressive, showing that the algorithm outperforms existing methods in detecting even minor changes.

Introduction

Sharing Arabic documents online is easy, but keeping them authentic is a new challenge (Oleiwi et al., 2023; Al-Wesabi et al., 2020). This applies to all types, from academic papers to religious texts (Usop & Hisham, 2020; Al-Wesabi, 2020). Tampering during transfer is a risk with critical consequences. Information security techniques like access control and tamper detection can help (Alkhafaji et al., 2021; Thabit, 2021; da Costa et al., 2020).

While traditional watermarking methods are widely used for document protection, they introduce a significant drawback: the possibility of the modification of the content that was initially introduced. Concerning the seven text data methods, they may inadvertently alter sensitive or crucial content while enshrinement of hidden data within the text. This becomes especially hard when working with texts such as the Holy Quran, Prophet's hadiths, or historical and legal documents in which even slight changes may cause significant consequences. As noted by Bastani et al. (2021) and CHEN et al. (2017) such an approach has its dangers, meaning that a technique needs to be developed to protect the document. Although the common techniques of watermarking might be satisfactory for a number of applications, they are not suitable for texts where a certain degree of applying alterations may have critical consequences. In turn we have zero watermarking which seems to be less coarse in its approach towards watermarking. as pointed out by Khadam et al. (2020) and Hakak et al. (2017) it does not insert the watermark directly into the content of the document and therefore preserves the structure of the text. Using certain features such as the words frequency in the document, the position of some characters etc, zero watermarking is used to guaranty the originality of the document without changing the actual text. This is especially beneficial when it comes to maintaining security of religious and legal documents where an ability to detect any attempt at alteration of text is important, however the very nature of the

ISSN: 2716-3865 (Print), 2721-1290 (Online)

document does not allow for any such changes to be made. Nonetheless, zero watermarking works effectively to the change in content while its impact on handling extensive manipulations and sophisticated attacks may need enhancement when used on large scale documents with variance of formatting.

Related work

The RCATED-AT system developed by Oleiwi et al. (2023) offered insights towards the Arabic text authentication as well as the tamper detection using 4th level word order process based on the Markov model. This is certainly employing Markov models in text watermarking as a new idea, however, it becomes questionable that the feature extraction only looks at the permutation of the words might be not strong enough. The challenge comes from the fact that in Arabic script some of the characters change their shape depending on the context in which they are placed in a word or sometimes restricted to specific positions within the word; this is quite a category that a word order based model will not sufficiently tackle. While the Markov model augments the probability measure with the previous k characters it may fail tograsp small text manipulations such as different characters at the same position such as ligatures, or diacritics which are significant in Arabic. This may lead to situations where specific types of tampering will not be noticed at all, if the changes are made at the character or sub-word level (Mou et al., 2020).

Moreover, the system appears to be also highly dependent on the word order; therefore, it may be weak against more advanced attacks that do not distort word order but change the meaning of the text or introduce slight variations. The Lee (2019) states that an attacker could alter the structure of a text simply by modifying the word order of the phrases used while maintaining the positions of the words hence can surpass the tamper detection system. Furthermore, it should be noted that the Markov models are successful in the statistical dependencies within the sequence of one kind of manipulation at a time, but in the other manipulations, such as insertions or deletions at the various places in the text, could not be efficient completely. It is obvious that including features at the word and character level would further improve the RCATED-AT performance since Arabic tampering is versatile in most of the times.

This work, Al-Wesabi et al. (2020), present ZWAFWMMM, the watermarking technique that applies an intricate Markov model of high accuracy in data embedding and manipulation detection in text watermarking. On the one hand, using a Markov model enriches the approach because the model examines sequences of text and is capable of identifying patterns of tampering; on the other hand, adding this level of complexity may sometimes bring about practical challenges. For instance, the model can depend on text structure which includes length, format or language particularities, which restrains its applicability on different kinds of textual document. In addition, it requires high dimensions of computational algorithms which may render it less applicable in real-time processes or highly large scale processing (Liu et al., 2013). One more important factor to examine is that even though ZWAFWMMM says that it has a high level of accuracy in detecting manipulation, the study does not provide a detailed discussion of how the method can be even slightly susceptible to more complex or massive tampering attacks (Deguillaume et al., 2003). In high volume or overly extensive tampering, the adversary could pose significant issues to the Markov model including learning natural language patterns or exploit holes within the training data for the model. Furthermore, the method's ability to identify slight shifts in file contents, rewritten or restructured text that may well be less obvious, is not well understood. These limitations give an impression that even though, ZWAFWMMM can be shown to be a viable method for text watermarking, it is necessary to conduct experiments in various settings and make ZWAFWMMM applicable for real-world situations to assert the outstanding effectiveness of the proposed method (Holleman et al., 2002).

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Al-Wesabi et al. (2020) in their research, provide ZWAFWMMM, a watermarking technique that employs Markov complex model in accurate embedding of data and efficient detection of manipulation in text watermarking. The added use of the Markov model is useful for the identification of sequences of text that have been tampered with and increases the sophistication of the process; however, this increases the model's complexity and may also present operational drawbacks. For example, the model may not generalise well across different structure of text documents and this may affect its scalability especially in long structured documents where different features may dominate. In addition, such an approach depends on the use of massive computational WebClientFormulas, which will make it slow and less optimal for real-time or big-data applications.

Another important issue that should be further explored is that although ZWAFWMMM achieved high accuracy to detect manipulation, the proof of concept does not discuss the method's robustness against more complex or scaled manipulation attacks. In high volume or a large number of attempts, the accuracy of the Markov model can decrease in the case where the attackers are using real-like imitation patterns or any blank area in the training set got by Markov model. Also, the potential of the method to identify changes in wording or even its overall structure in the text, which may be rather close to the original, is not clear. Such limitations indicate that even though ZWAFWMMM shows a viable solution for the text watermarking use-case, its efficacy and reliability in other situations, as well as suitable adjustments for real-world implementation, remain an open question.

The FATZWNLP framework developed by Al-Wesabi et al. (2021) can be considered a breakthrough for Arabic text digital watermarking. This approach thus uses a Markov model, which is an approach that has better capability in analyzing textual data by virtue of statistical language property analysis. This is especially crucial to Arabic as this language is typologically rich in morphological and syntactic processes (Al-Shawashreh et al., 2020). A Markov model can be included in the framework to make the probabilistic relationships between words and phrases clear so that better features can be found. This promotes the creation of a strongly resistant embedded watermark so that its detection will be easily done especially when the image has been subjected to several attacks. Emphasizing Arabic text also fills the existing void in most methods and algorithms for watermarking that are inclined towards more popular languages thereby boosting the functionality and usability of the study among Arabic community (Thabit etal., 2021).

Nevertheless, the FATZWNLP framework presents examples of the experiments with fuzzy and neural networks in watermarking processes, and therefore it is also important to consider the practical applicability of the proposed approach and its possible disadvantages. It has already been mentioned that using the Markov model may be computationally intensive, more so when analyzing large sets of text. This could cause more processing time and resource usage that could cause a negative impact on real-time application (Mansouri & Babar, 2021). Also, based on the tampering scenarios depicted in this paper, it is advisable that the applicability of the watermarking technique be examined in detail to ensure that it does meet the requirement of an effective technique. It also must consider how perfectly this framework is incorporated with existing DMCA systems, and whether it still keeps the watermarks' integrity at various formats of changes and usages. In conclusion, even though the proposed approach of FATZWNLP shows a solution in embedding a watermark on Arabic text, more real-world experimentations to support the approach and practical applications to place in the real world is essential.

The work of Hilal et al. (2022) also presents CATDAWNLP approach in which Markov model analysis for feature extraction in Arabic text is made more elaborate. On the basis of the properties of stationary Markov processes, the method finds and localizes concrete characteristics of the text, thus imprinting a sort of a code or a watermark. This watermark

ISSN: 2716-3865 (Print), 2721-1290 (Online)

serves a dual purpose: besides, confirming the credibility of the contents, it allows also for capturing the attempts of changing information. The Markov model that makes up the base idea of the algorithm works with transition probabilities between states within the text body, which reflects probabilities of encounters of one word or phrase against the background of another. It permits the delicate analysis of the work's structure and semiotics and this will go a long way to help identify what is a rightful amendment from what is tampering.

However, despite the evidence of efficacy observed for the CATDAWNLP method in improving the Arabic text, there are several important assumptions that must be resolved (Hilal et al., 2022). First, the usage of a definite Markov model can restrain the applicability of the approach to different dialects and styles in Arabic which are arising within highly diversified Arabic language. However, the ability of the watermark to withstand specific types of attacks should also be researched further. For example, attackers may write text using rich phrases that can mollify or lure the watermarking process hence making it easy for them to penetrate through the firewall. Moreover, the analysis of a Markov model may be rather computationally intensive and may present some difficulties in real time applications in the environments where speed counts. Albeit the fact that CATDAWNLP provides innovative solutions and concepts to digital watermarking in Arabic text, the follow through, applicability and efficiency of the approach in various fields has to be further investigated and fine tuned.

Methods

Proposed Algorithm

The characteristics of the Arabic text were utilized to generate the watermark within the proposed method, so the proposed algorithm will not cause a change in the text document to hide the watermark. The watermark will be used to validate the text documents. The processes of generating watermark and extraction are shown in figure 1. The watermark is recorded in CA and is used in the extraction algorithm to validate the text document.

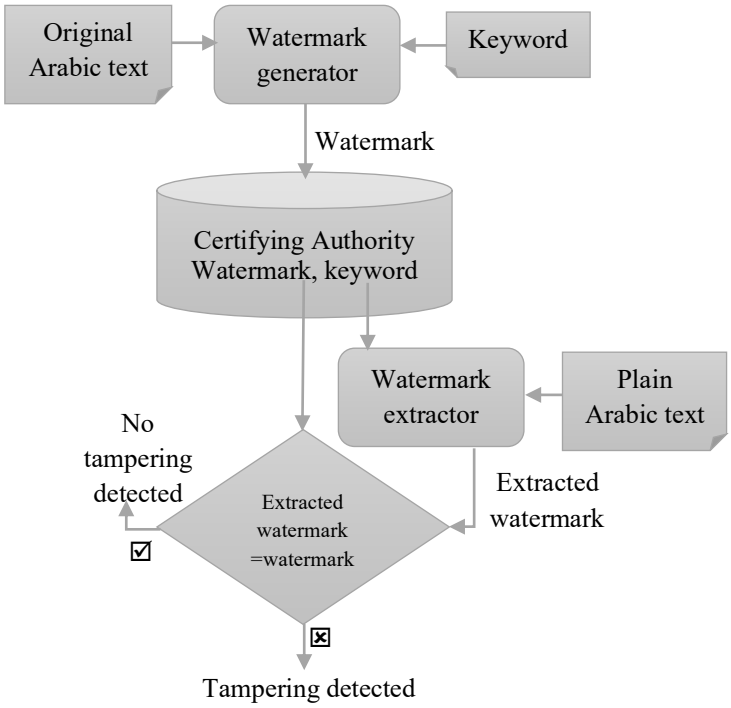

Figure .: Flowchart of watermark generating and extraction procedures

The contents of the Arabic text are used in the proposed algorithm to protect it. Based on the author's choice of the keyword, such as his/her surname or the last name or any other word,

ISSN: 2716-3865 (Print), 2721-1290 (Online)

it must consist of at least 6 characters, the duplicate characters, if any, will be excluded. The watermark will be generated based on the keyword characters, where the number of times each character appears will be counted according to its form (separated or connected) and its position at the beginning, middle or end of the word within the contents of the Arabic text document. This method is shown in figure 2, where "بيلسان" it is the keyword and depending on the contents of the text, the watermark is generated.

تتم حماية النص العربي بواسطة الخوارزمية المقترحة وذلك باستخدام محتوياته. بناءً على اختيار المؤلف للكلمة المفتاحية، كأن تكون اسمه الاول او اسم العائلة او حروف منهم او أي كلمة، ويجب ان تتألف على الاقل من 6 حروف وسيتم استبعاد الحروف المكررة إن وجدت، سيتم توليد العلامة المائية بالاعتماد على أحرف الكلمة المفتاحية، حيث سيتم احتساب عدد مرات ظهور كل حرف من الكلمة المفتاحية وحسب شكله (منفرد ام متصل) في موضعه (اولي، آخري وسطي) من الكلمة ضمن محتويات المستند النصي العربي. هذه الطريقة مبينة في الشكل 2، حيث "بيلسان" هي الكلمة المفتاحية واعتمادا على محتويات النص يتم توليد العلامة المائية.

Characters	ب	ي	ل	س	ا	ن
Freq. of occurrence	5	17	11	3	18	5

Watermark= بيلسان 5 17 9 2 18 5

Figure 2. Watermark Generation

In a zero watermark system the watermark will not actually be embedded in the text itself; rather, it is created using text properties. Watermarking includes two stages: 1) the embedding algorithm and 2) the extraction algorithm. The original author embeds the watermark, while CA later extracts the watermark to prove ownership. CA plays an essential role, as the original copyright owner registers his/ her watermark with them. Whenever there is doubt about the ownership of content/ text, the CA acts as a trusted third party as the final decision authority.

Embedding Algorithm

The inputs for the watermark embedding algorithm will be the original text file and the keyword chosen by the original author/ copyright owner. The output of this algorithm will be the watermark. This watermark will be registered along with the original text document, author name, keyword, and current date and time with the CA. The context of the algorithm's operation is shown in figure 3.

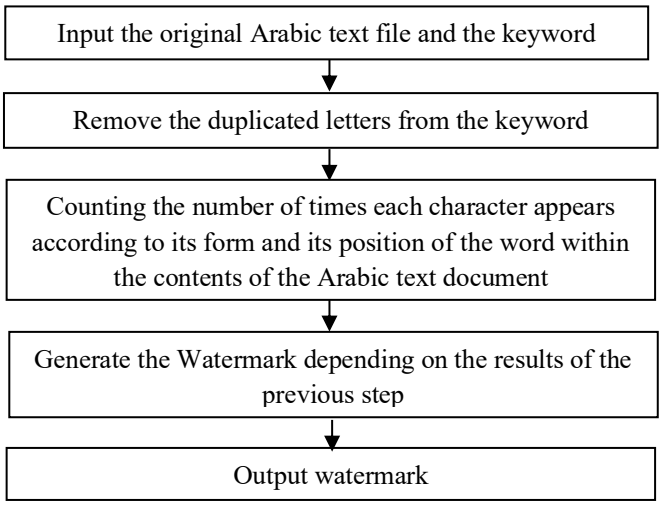

Figure. 3. Embedding algorithm Diagram

The original text is obtained from the author. The author chooses the keyword. The duplicate characters are deleted from it, and their reappearance is counted in the original text, according to their shape and position in the word and digital watermark is created. This watermark is registered with CA with the current time and date.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Extracting Algorithm

The inputs for the watermark extracting algorithm will be the plain text file and the keyword. The text may or may not have been attacked. The extraction algorithm will generate the watermark from the plain text and then compare it with the original watermark registered with CA, where the author's name and current date and time are registered with CA. The author with the earlier registration will considered the original owner.

The watermark will be accurately generated by the algorithm if the text is not tampered with, and the text will be called the original text without tampering. If the text is subjected to tampering attacks, the watermark will be distorted.

The text may be subjected to tampering attacks such as insertions, deletion, rephrasing, or reordering of words and sentences. In figure 4 the extraction algorithm is illustrated.

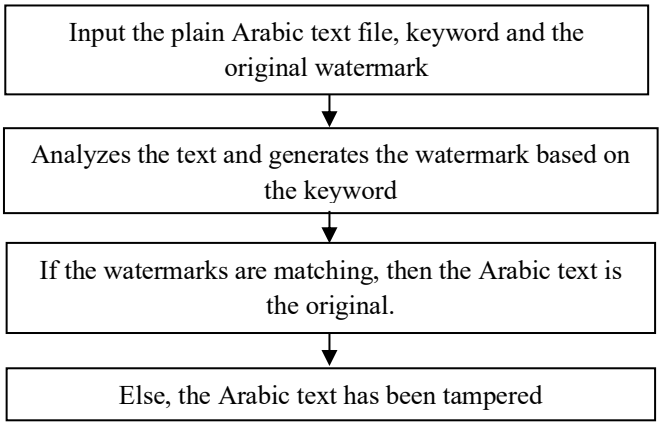

Figure. 4. Diagram of extraction algorithm

Results and Discussion

Nine text samples of different sizes were utilized in the experiment. These samples were collected from researches, e-books, and web pages. It was exposed to variable types of attacks at word and sentence levels of the text in multiple randomly selected locations. Table 1 shows the details of the results of the implementing experiment.

Table 1. Original and attacked Arabic text samples with volume attack of insertions, deletions, and rephrasing applied.

Sample No.	Characters Count Of Original Text	Volume Attack			Characters Count Of Attacked Text
		Insertion	Deletion	Rephrasing	
1 [ASST1]	1004	3%	2%	1%	1014
2 [ASST3]	1975	4%	4%	2%	1975
3 [ASST6]	2959	7%	5%	3%	3018
4 [ASST5]	2979	12%	10%	4%	3038
5 [ASST8]	3196	15%	11%	5%	3323
6 [AMST2]	4742	25%	20%	8%	4979
7 [AMST4]	7182	35%	25%	9%	7900
8 [AMST9]	7705	45%	30%	10%	8860
9 [ALST7]	10148	50%	35%	15%	11670

In the original and attacked text samples, the number of times the characters of the keyword "بيلسان" appeared was analyzed. According to the following formulas, the watermark accuracy rate (WAR) and the watermark distortion rate (WDR) are calculated:

WAR = Characters count accurately detected / Characters count in the watermark

WDR = 1 - WAR

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Where,

$0 \leq WAR \leq 1$; WAR should be close to 1.

$0 \leq WDR \leq 1$; WDR should be close to 0.

To compare the original Arabic text samples with their counterparts—those that have been manipulated by various volume attacks including insertions, deletions, and rephrasing—Table 1 has been provided. Every sample has its number; character count of the original text before modifying is added with the percentage of altered text. For example, the Sample No. 1 is denoted as [ASST1] which has an original character count = 1004; insertions are 3%; deletions are 2%; and rephrasing is 1%; and therefore the final character count=[1014]. Sample No. 9 [ALST7] includes more substantial modifications; the percentage of insertion is 50%, and the deletion is 35%, increasing the character count to 11670. This variation shows the correlation between sample and diversification and volume of attack demonstrating that the higher the volume of attack the increase in ITL and potential distortion of the text.

Furthermore, the analysis uses the frequency of the keyword "بيلسان" as well as the comparative calculations of watermark accuracy rate (WAR) and watermark distortion rate (WDR) where the keyword is located in an original text and in the attacked text accordingly. The above formulas for WAR and WDR provide for a measure amenable for quantitative measure on the watermark's structural degree after manipulation. A high WAR value, which is closer to 1, means good performance of detection of the watermark whereas a low WDR value, closer to 0, means less distortion (Naz et al., 2020). This framework allows the researchers to compare their results and determine exact fitness of their watermark extraction algorithms with the goal to find out whether the text has been tampered through the attacks mentioned in the table above.

By comparing the WAR value of the extracted watermark and the original watermark, the effectiveness of watermark extraction can be evaluated. Through this analysis, it is possible to determine whether the text has been tampered with or not. Results are presented in table 2.

Table 2. Accuracy of extracted watermark

Sample No.	Characters count of original text (بيلسان)	Characters count of attacked text (بيلسان)	Tampering detected	WAR
1 [ASST1]	29 85 130 18 170 13	30 70 140 15 190 10	yes	0.978
2 [ASST3]	130 137 241 31 339 109	130 137 266 31 364 109	yes	0.962
3 [ASST6]	106 148 316 64 460 119	116 158 326 64 470 129	yes	0.950
4 [ASST5]	112 185 390 115 642 101	112 185 390 115 642 101	yes	0.945
5 [ASST8]	122 172 406 80 527 142	122 172 406 80 527 142	yes	0.937
6 [AMST2]	172 225 528 101 741 234	172 225 528 101 741 234	yes	0.926
7 [AMST4]	174 419 771 149 1156 247	174 419 771 149 1156 247	yes	0.911
8 [AMST9]	252 444 1015 176 1347 250	252 444 1015 176 1347 250	yes	0.897
9 [ALST7]	330 587 1140 233 1770 364	330 587 1140 233 1770 364	yes	0.872

The analysis also reveals a very important correlation between the manipulation of the text and the watermark efficiency as seen in table 2 and explained in figure 5. As tampering increases, the watermark accuracy rate (WAR) diminishes, establishing a clear inverse relationship: as an implication, the higher the level of manipulation of the text by an attacker,

ISSN: 2716-3865 (Print), 2721-1290 (Online)

the lower the reliability of the watresize. It is rather worrisome, as any changes escalate the WDR, implying that the text is easy to falsify. The fact that the high WDR persist across all samples minimizes the chances of the watermark hence underlining the waypoint vulnerability of the watermark to any form of alteration. This vulnerability highlights one of the main weaknesses of the watermarking techniques: a log change in the parameter is sufficient to greatly degrade the quality of the watermark. Thus, the research outcomes present watermarking as a valuable instrument that helps detect modifications, but attacks may be induced with equal simplicity. This poses critical questions on the stability of watermarking approaches in preserving text integrity and calls for enhanced techniques which can only be' slightly compromised (Liu et al., 2024). The study demonstrates the difficulties involved in obtaining high levels of watermark detection on texts that have been attacked using different techniques; this identifies a major research need in the field of digital text security.

Any attempt to tamper with the text will always be detected, as shown in table 2. If tampering in the text increases, the accuracy of the watermark decreases, so there is an inverse relationship between them. The WDR with the keyword "بيلسان" on all text samples is shown in Fig 5. It's evident that even with minimal attacks, the WDR remains high. The text can be easily affected by any changes made by the attacker. If the text has been tampered with and is no longer original, the high distortion rate will indicate that. This proves that the watermark's accuracy is negatively affected even with minor modifications, and the fragility of the watermarks proves that the text was attacked.

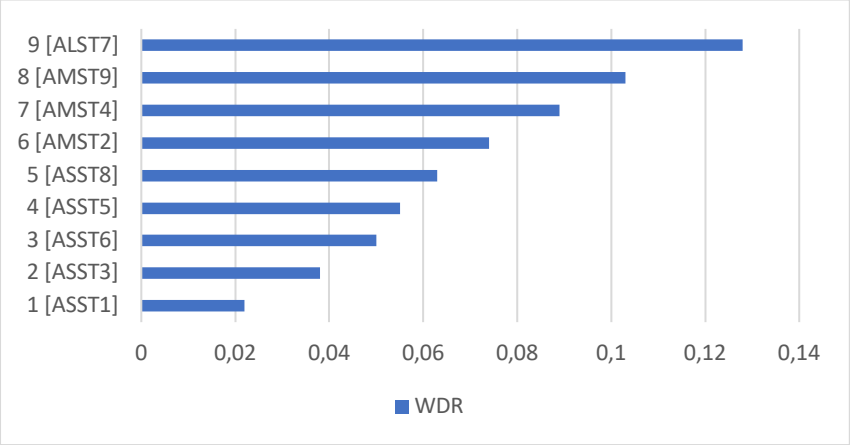

Figure 5. shows the WDR with keyword (بيلسان) on all text samples

Comparative Results

Compared to RCATD-AT (Oleiwi et al., 2023), the proposed method indicates a higher accuracy in detecting the fragility of the watermark. This is evident in the results presented in Table 3, it can be observed that all the details of the percentages for the different types of attacks in the proposed method are the highest. Although insertions, deletions or rephrasing operations negatively affect watermark integrity, the proposed method achieved better results than another approach in this field.

Table 3. Comparison of watermark tamper detection accuracy.

Approach	Attack volume	Attack		
		Insertion	Deletion	rephrasing
RCATD-AT	5%	94.83	88.88	76.70
	10%	88.84	84.30	62.99
	20%	81.14	66.48	43.45
	50%	62.48	42.97	24.98
	5%	96.12	90.33	79.44

ISSN: 2716-3865 (Print), 2721-1290 (Online)

The proposed method	10%	90.67	86.32	72.56
	20%	83.23	76.56	63.34
	50%	77.34	67.78	50.12

The table shows the comparison results between two methods: In terms of attack volume and the relation to three kinds of attack: insertion, deletion, and rephrasing, both the RCATD-AT and the proposed method. Each column represents the success rate of each method in handling attacks at various percentages of attack volume: 5%, 10%, 20%, and 50%.

As to the comparison between the proposed method and the RCATD-AT, a more accurate classification upon higher levels of attack volume can be observed wherein the proposed method is superior to the RCATD-AT achievement. For example, at an attack volume of 5%, the proposed method achieves the highest success rate for all the target attacks: insertion 96.12%, deletion 90.33%, and rephrasing 79.44%. While at the 50% level the effectiveness of the proposed method is still lower than that of the LanguageTool, but all the same the proposed method appears to be more effective, having the rate of successful insertions 77.34%, deletions 67.78%, and rephrasing 50.12%. However, these results reveal that the proposed method is superior in handling various types of attacks, and particularly superior in low attack rate environments. This means that the new approach can be useful to build up proper prospects for the development of more effective defense against attacks in the broader perspective.

In the discussion section of the article, the importance of the zero-watermarking algorithm recommended for Arabic text should be emphasized by comparing the performance of detecting the forged zone of the constructed document to the previous method RCATD-AT. It can be development in light of a few points which comprehensively standardize earlier studies and position the contributions in this research examine in the unambiguous zone of text authentication and security.

First, the enhanced results as far as the picture of a watermark and its capacity for detecting an attempt at tampering underscore an important addition to the discussion of Arabic text watermarking. As stated in the previous work RCATD-AT and ZWAFWMMM methods have some level of success in detecting tampering based on insertion, deletion or rephrasing attacks however their performance degraded drastically as the attack volume increases. However, the proposed method consistently does better compared to RCATD-AT in all attack types and most striking at low attack volume (5-20). This means that even if there is a small change in the original text, the algorithm becomes more useful in retaining the originality of Arabic documents. This is important for scenarios where slight variation in religion or history or law can make a lot of difference (McGuire, 2008).

The enhancement in performance results from a new algorithm considered in generating the watermark (Shih & Wu, 2005). The advantage of the proposed work is the ability of the watermark to capture the frequency and structural properties of the Arabic characters. This is quite relevant especially with regard to the arabic script in which an individual character depends on its position in a word (Balaha et al., 2021; Alanazi et al., 2022). Such changes were previously neglected in the previous input switch methods, and thus, the older methods of solving this problem were not very resilient against more advanced manipulation techniques. This, however, is not to mean that the proposed system takes advantage of these textual features in an endeavor to come up with a more trustworthy as well as original watermark.

Second, this study is useful in various text samples like academic papers, and legal documents as well as other relevant texts. This broader testing scope not only shows how the algorithm may be applied broadly but also how its utility is not limited to any given sphere of application (Kar, 2016). The prior studies in the field were mostly type-bound, and therefore, their

ISSN: 2716-3865 (Print), 2721-1290 (Online)

applicability was restricted. The general applicability of the new method to all the sizes/ types of texts enhances the efficacy of the method.

However, it is also necessary hence forward the further study to conclude that they mentioned sometime about the study. For example, the compared results of the performance of the proposed algorithm degrades slightly higher attack volumes of 50%. This suggest that more optimization may be needed in order to keep low error watermark accuracy under the worst signals. Further, the susceptibility of zero-watermarking techniques to large scale of attacks has also been established, indicating the future prospect of this field. Nevertheless, the current study has made much progress, using this algorithm in combination with other security solutions, for example, blockchain for decentralized validation, this approach can be strengthened even more.

Conclusion

The current limitation in document authentication with traditional text watermarking is the difficulties in detecting minor tampering. A new zero watermarking algorithm was proposed to overcome this challenge. This creative method generates a distinctive watermark by utilizing the text's content. Subsequently extracting this watermark verifies the document's authenticity. The algorithm's effectiveness was tested against random tampering attempts on a nine text samples with varying sizes. The results are highly promising, demonstrating the algorithm's exceptional ability to consistently detect tampering, regardless of the size of the modifications. From the table above, the proposed method shows the most significant enhancement for the 50% attack volume with an improvement of 23.78% over RCATD-AT for all attack types. Even at lower attack volumes, the proposed method consistently outperforms RCATD-AT.

References

Ahmed, S. (2016). *" We call on citizens to be aware of the value of what is in their homes:" A case study of the Hassan II Prize for Manuscripts and Archival Documents* (Doctoral dissertation, The University of North Carolina at Chapel Hill).

Alanazi, N., Khan, E., & Gutub, A. (2022). Inclusion of unicode standard seamless characters to expand Arabic text steganography for secure individual uses. *Journal of King Saud University-Computer and Information Sciences*, *34*(4), 1343-1356. https://doi.org/10.1016/j.jksuci.2020.04.011

Alkhafaji, A. A., Sjarif, N. N. A., Shahidan, M. A., Azmi, N. F. M., Sarkan, H. M., & Chuprat, S. (2021). Tamper detection and localization for Quranic text watermarking scheme based on hybrid technique. *CMC-COMPUTERS MATERIALS & CONTINUA*, *68*(1), 77-102. https://doi.org/10.32604/cmc.2021.015770.

Al-Shawashreh, E., Alshdaifat, A., al Huneety, A., & Mashaqba, B. (2020). Typological universals of agreements in Arabic second language acquisition. *Dirasat: Human and Social Sciences*, *47*(1).

Al-Wesabi, F. N. (2020). Proposing high-smart approach for content authentication and tampering detection of Arabic text transmitted via internet. *IEICE TRANSACTIONS on Information and Systems*, *103*(10), 2104-2112. https://doi.org/10.1587/transinf.2020EDP7011.

Al-Wesabi, F. N., Abdelmaboud, A., Zain, A. A., Almazah, M. M., & Zahary, A. (2021). Tampering Detection Approach of Arabic-Text Based on Contents Interrelationship. *Intelligent Automation & Soft Computing*, *27*(2). https://doi.org/10.32604/iasc.2021.014322.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Al-Wesabi, F. N., Mahmood, K., & Nemri, N. (2020). 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*, *52*, 102473. https://doi.org/10.1016/j.jisa.2020.102473.

Balaha, H. M., Ali, H. A., & Badawy, M. (2021). Automatic recognition of handwritten Arabic characters: a comprehensive review. *Neural Computing and Applications*, *33*, 3011-3034. https://doi.org/10.1007/s00521-020-05137-6

Bastani, A., & Fatemi Behbahani, E. (2021). A Self-recovery Digital Watermarking Approach for Tamper Detection of Handwritten and Printed Electronic Documents. *Library and Information Sciences*, *24*(1), 174-193. Https://doi.org/10.30481/LIS.2020.235473.1727.

CHEN, S. Y., MA, H., & CHEN, Q. J. (2017). Tamper Detection of Batch Websites Based on Text Comparison. In *Computer Science and Technology: Proceedings of the International Conference (CST2016)* (pp. 573-579). https://doi.org/10.1142/9789813146426_0065.

da Costa, K. A., Papa, J. P., Passos, L. A., Colombo, D., Del Ser, J., Muhammad, K., & de Albuquerque, V. H. C. (2020). A critical literature survey and prospects on tampering and anomaly detection in image data. *Applied Soft Computing*, *97*, 106727. https://doi.org/10.1016/j.asoc.2020.106727

Deguillaume, F., Voloshynovskiy, S., & Pun, T. (2003). Secure hybrid robust watermarking resistant against tampering and copy attack. *Signal Processing*, *83*(10), 2133-2170. https://doi.org/10.1016/S0165-1684(03)00172-5

Hakak, S., Kamsin, A., Tayan, O., Idris, M. Y. I., Gani, A., & Zerdoumi, S. (2017). Preserving content integrity of digital holy Quran: Survey and open challenges. *Ieee Access*, *5*, 7305-7325. https://doi.org/10.1109/ACCESS.2017.2682109

Hilal, A. M., Al-Wesabi, F. N., Hamza, M. A., Medani, M., Mahmood, K., & Mahzari, M. (2022). Content authentication and tampering detection of Arabic text: an approach based on zero-watermarking and natural language processing. *Pattern Analysis and Applications*, 1-16. https://doi.org/10.1007/s10044-021-01032-5.

Hilal, A. M., Al-Wesabi, F. N., Hamza, M. A., Medani, M., Mahmood, K., & Mahzari, M. (2022). Content authentication and tampering detection of Arabic text: an approach based on zero-watermarking and natural language processing. *Pattern Analysis and Applications*, 1-16. https://doi.org/10.1007/s10044-021-01032-5

Holleman, G. A., Hooge, I. T., Kemner, C., & Hessels, R. S. (2020). The 'real-world approach'and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*, 721. https://doi.org/10.3389/fpsyg.2020.00721

Kar, A. K. (2016). Bio inspired computing–a review of algorithms and scope of applications. *Expert Systems with Applications*, *59*, 20-32. https://doi.org/10.1016/j.eswa.2016.04.018

Khadam, U., Iqbal, M. M., Alruily, M., Al Ghamdi, M. A., Ramzan, M., & Almotiri, S. H. (2020). Text data security and privacy in the internet of things: threats, challenges, and future directions. *Wireless Communications and Mobile Computing*, *2020*(1), 7105625. https://doi.org/10.1155/2020/7105625.

Lee, J. H. (2019). *Systematic approach to analyzing security and vulnerabilities of blockchain systems* (Doctoral dissertation, Massachusetts Institute of Technology).

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Liu, A., Pan, L., Lu, Y., Li, J., Hu, X., Zhang, X., ... & Yu, P. (2024). A survey of text watermarking in the era of large language models. *ACM Computing Surveys*. https://doi.org/10.1145/3691626

Liu, Z., Jiang, B., & Heer, J. (2013, June). imMens: Real-time visual querying of big data. In *Computer graphics forum* (Vol. 32, No. 3pt4, pp. 421-430). Oxford, UK: Blackwell Publishing Ltd. https://doi.org/10.1111/cgf.12129

Mansouri, Y., & Babar, M. A. (2021). A review of edge computing: Features and resource virtualization. *Journal of Parallel and Distributed Computing*, *150*, 155-183. https://doi.org/10.1016/j.jpdc.2020.12.015

McGuire, M. B. (2008). *Religion: The social context*. Waveland Press.

Mou, G., Ye, P., & Lee, K. (2020, October). Swe2: Subword enriched and significant word emphasized framework for hate speech detection. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 1145-1154). https://doi.org/10.1145/3340531.3411990

Naz, F., Khan, A., Ahmed, M., Khan, M. I., Din, S., Ahmad, A., & Jeon, G. (2020). Watermarking as a service (WaaS) with anonymity. *Multimedia Tools and Applications*, *79*, 16051-16075. https://doi.org/10.1007/s11042-018-7074-2

Oleiwi, A. S., Alkhafaij, M. A., Ali, R. R., Ali, E., Al-Tahee, M., & Almusawi, M. (2023, July). Robust Content Authentication and Tampering Detection of Arabic Text Transmitted Through Internet. In *2023 6th International Conference on Engineering Technology and its Applications (IICETA)* (pp. 790-796). IEEE. https://doi.org/10.1109/IICETA57613.2023.10351462.

Shih, F. Y., & Wu, Y. T. (2005). Enhancement of image watermark retrieval based on genetic algorithms. *Journal of visual communication and image representation*, *16*(2), 115-133. https://doi.org/10.1016/j.jvcir.2004.05.002

Thabit, R., Udzir, N. I., Yasin, S. M., Asmawi, A., Roslan, N. A., & Din, R. (2021). A comparative analysis of arabic text steganography. *Applied Sciences*, *11*(15), 6851. https://doi.org/10.3390/app11156851.

Thabit, R., Udzir, N. I., Yasin, S. M., Asmawi, A., Roslan, N. A., & Din, R. (2021). A comparative analysis of arabic text steganography. *Applied Sciences*, *11*(15), 6851. https://doi.org/10.3390/app11156851

Usop, N. A. A., & Hisham, S. I. (2021). A Review of Digital Watermarking Techniques, Characteristics and Attacks in Text Documents. *Advances in Robotics, Automation and Data Analytics: Selected Papers from iCITES 2020*, 256-271. https://doi.org/10.1007/978-3-030-70917-4_25.

ISSN: 2716-3865 (Print), 2721-1290 (Online)