Arabic text watermarking tuned for medical e-record semi-authentication

Adnan Gutub * and Esraa Almehmadi

Computer Engineering Department, Umm Al-Qura University, Makkah, Saudi Arabia

** Corresponding Author: aagutub@uqu.edu.sa*

Submitted : 27-07-2022

Revised : 05-09-2022

Accepted : 17-09-2022

ABSTRACT

This study focuses on Arabic text watermarking semi-verification utilizing recent counting-based secret-sharing to partially validate the ownership and correctness for all sensitive medical e-records. The benefit of this approach is its semi-trust of e-records, as needed services provided, even if complete text medical report is a bit delayed for full verification. The proposed work presents two approaches of text-watermarking which are Seed Shares approach and WM Shares approach. The capacity results of our formulations are achieved without effect of secret bits hidden. The security is achieved via using XOR operation and CBSS technique. The comparisons result shows that the accuracy percentage of the seed shares approach is much better than the WM in terms of tampering attack. The exploration tested its proposed strategy on standard Arabic benchmark of 42 Nawawi Hadiths which provided promising remarks, which can also be useful for Urdu and Farsi texting.

Keywords: Watermarking; Steganography; Count-Based Secret Sharing; Medical Security.

INTRODUCTION

Among the primary responsibilities of the healthcare sector collection of patient data is keeping the collected data confidential, used lawfully and sensitively secure. After data collection, the next step is the simultaneous sharing of sensitive data with the patients and the intended staff of the hospital caring (Gutub, 2022b). As the information is confidential and private needing information sharing, the security of private materials is considered basic requirement in the healthcare systems, which implies that medical systems must use up-to-date confidentiality technologies (Gutub, 2022c). The approach of healthcare process has gradually been shifting from hospital-centered to patient-centered, as now innovatively following distributed approach. This current urgency is facilitated through health record distributed digitalization, as novel approach accepted worldwide being inevitable and mostly top vital. As understood, patient data is confidential and sensitive, however, it needs to be shared with all the intended persons and healthcare media including patients and their relatives, medical staff, and other linked systems institutions (Gutub, 2022b). Confidentiality is desirable to be safeguarding the information (Gutub, 2022c), as people/systems other than the planned ones should be fully strained or stopped from access, i.e. as exactly needed by the healthcare procedure. So, safekeeping all medical data is an obligation of today's technology to be derived from legislative guidelines as giving urgency rights to patient and their duties, linked to the intended medical specialists. The todays' advanced healthcare e-organizations are grounded by data allocations classified to be posted on open unsecured networks like the internet, i.e. to be available for doctors or specialists as well as healthcare system. E-medical reports are of these secret materials measured as key central items within telemedicine arenas, to be utilized whenever needed for specialized investigation and therapeutic diagnostics. Any alteration, or unavailability, of health e-reports will drastically disturb the professional judgment affecting the overall civil-health society. For these motives, it is compulsory to offer the safekeeping secrecy settings to exchange medical e-data in order to provide highest honesty and genuineness, i.e. of the medical reports during the e-communication. Watermarking is a possible scheme to validate the e-medical materials, particularly with its progression procedures. The chief goal of this investigation is to highlight tuning imperceptible watermarking, as supplementary in demonstrating trustful rights, by adopting counting-based secret-sharing for e-record semi-authentication (Gutub et al., 2019). Counting-based secret-sharing (CBSS) has been experienced in related explorations to make use of incomplete reliable watermarking, as to verify identity for data-bits in (Almehmadi & Gutub, 2021). This Arabic text research (Almehmadi & Gutub, 2021) adopted the owner's secret password from CBSS, as seed bits, to yield shares to be XORed as watermarking pattern providing seamless embedding stream remarking successfully on Arabic text files different than CSNTSteg (Thabit et al., 2022) functionality. The intended CBSS method tracks the data hiding models, as bits from steganography to demonstrate secrecy, while having no consequences on medical e-report visibility. Ultimately, when validating, the implanted bits are gathered again reforming the CBSS shares to form the password seed combination as justifying the watermarking originality, despite the fact that twosome tampering is discovered taking place on Arabic text medical reports (Gutub & Fattani, 2007). Note that watermarking shares are produced by the system proposed in (Almehmadi & Gutub, 2021) to process embedding, similar to Arabic-text steganography, and hence giving novel system direction of watermarking that satisfies following properties. First, the ability to use partial verification attribute if some parts of the watermarked medical e-reports are available, i.e. to show the ownership. Second, not only is the performance of the extracting process enhanced but the complexity delay is reduced as well. Because ownership is now provable by merely processing parts of the watermarked text, as at least three parts are reconstructed for extracting the watermarked text successfully. Whereas, other techniques depend on the processing of the entire text. Testing this study has been run as partial alteration percentage on Arabic benchmark Nawawi 40-Hadiths, where integrity verification has been ranked. Note that this proposed work currently focuses on Arabic, which is spoken by over 1.6 billion Muslim peoples as well as conserved by Holy Quran revelation, as primary book of Islam. The presented work is adequate for similar languages such as Urdu and Farsi (Gutub & Fattani, 2007). The utilization of data hiding to authenticate information or for security purposes involves many other areas and linking various techniques for serving the current usage of e-media files (Gutub et al., 2010). Our emphasis of research (Almazrooie et al., 2020) is securing watermarking from discovery as hidden against illegal or unwanted interference. Whereas study (Thyagarajan et al., 2020), aims to help in the certification of copyrights, this watermarking is targeted for private data injection, as for the identification of owner's medical incomplete authenticity benefiting from the recently presented CBSS strategy, as illustrated in the forthcoming sections.

The paper is sequenced by, Section 2 contains the related work. Section 3 depicts an overview of the proposed model in the context of watermarking Arabic text files. Section 4 describes results and discussion. Section 5 and Section 6 show different related comparisons, analysis and remarks concluding the work.

RELATED WORK

This section explores related works of cybersecurity concerning the linked healthcare system.

Cybersecurity for healthcare systems: The primary goal of cybersecurity is the protection of content and the integrity of the computing processes from those to whom the content does not belong while interacting with a network. The aim is to protect the content of its owner or authorized person from all other actors in the network who are not authorized for accessing that content and can cause a cyber-attack (Gutub, 2022a). The job of Cybersecurity in the healthcare domain is quite similar as it has to protect the content of patient's data which is confidential and prone to attacks from hackers whom can leak that particular content. The other misuses of the content include the theft for clinical fraud, such as fake medical documents for their harmful benefits and other illegal monetary benefits. While preventing all these infringements and cybercrimes, Cybersecurity allows retaining the confidential content of the patient. As the approach of eHealth systems has increased exponentially in recent times, which are developed on cloud-based technology that demands the provision of efficacious security and confidentiality for the patient data, which is a challenge for the field of Cybersecurity (Anand et al., 2020). In order to achieve the goal of security and confidentiality, healthcare organizations are required to provide an environment that ensures the confidentiality and safety of preserving sensitive medical reports. To protect the eHealth system of an organization, the Cybersecurity system should be able to control and disclose all the protection elements of that particular eHealth organization (Thyagarajan et al., 2020).

Watermarking in healthcare: Hospitals around the globe have started using Electronic Patient Records (EPR) which is a healthcare system designed for sharing information of the patients which is utilized for the improvements in medical service of the patients as well as for conducting the research. The bigger challenge halting the efficiency of such a system is its vulnerable security of the EPR shared. In order to tackle this flaw, researches develop a dual watermarking technique that is based on the compression-then-encryption cryptographic method utilized to secure the data yielding more features. The method produced expected successful results going through a series of well-designed experiments. Analysis conducted comparing the related systems revealed significant improvements in security and robustness due to the proposed technique. A combination of multiple techniques such as

RDWT-RSVD and SPIHT-STE are utilized to enhance the robustness and security of the system. For encoding, the turbo code concept is applied to the data before inserting it in the watermark of the image. The generated mark image is then inserted in the medical cover image revealing better outcomes regarding robustness, impalpability, and security. The proposed method shows better results in terms of BER and NC when matched with the currently existing systems. Also, the experiment-based analysis of all the systems including the proposed system marks the most efficient for the EPR security for smart healthcare. The proposed method can be further enhanced by adding machine and deep learning techniques to it. This method can also be tested on various other applications (Zear et al., 2018). Algorithms used for multiple digital watermarking combining Discrete Wavelet and Discrete Cosine Transforms or DWT and DCT, and SVD or Singular Value Decomposition, have been put forward for various healthcare applications such as teleophthalmology, teleconsultancy, telediagnosis, telemedicine, and services. In order to decrease the outcomes of thefts related to medical identity, various watermarks are utilized in the algorithms. The phase of embedding the medical cover image contains disintegration of the image into the third level DWT. DCT transforms the low-frequency bands (LH2 and LL3), then the coefficients of the DCT are combined with SVD. Furthermore, two image watermarks were transformed by DCT and SVD. The informational values of the watermark images are embedded in the unit value of the medical cover image. An extraction algorithm was used to extract the watermarks. To improve the robustness of the extracted watermarks, a neural network called back-propagation is applied which decreases the effects of noise produced during production. After application, results are obtained by differentiating the improvement factor and the various cover image modalities. The statistics of the series of experiments that are conducted show that this method can sustain various signal processing attacks, and can show excellent robustness and impalpability. Comparative analysis of this technique with others was also performed. A subjective method has been utilized to measure the visual quality of the proposed techniques (Jalil, 2010). In (Sahu et al., 2022) It has been proposed to use a unique local binary pattern-based reversible data hiding (LBP-RDH) approach to keep the perceived transparency and hiding power fairly symmetric. In this research paper (Suresh et al., 2022), SVD-based systems and their susceptibility to FPP are thoroughly researched, dissected, and explained. The healthcare watermarking (WM) model is illustrate in Figure 1.

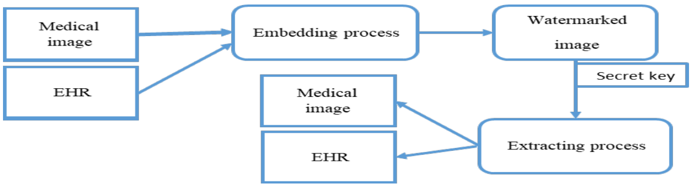

Figure 1. The WM model for healthcare system

PROPOSED APPROACH

The Seed Shares approach: One of the earlier research studies (Almehmadi & Gutub, 2021) uses CBSS to enhance the watermarking system for the Arabic text. Embedding process is divided into three phases. In the first phase, secret bits/password is further categorized into two sections: watermark bits and seeds bits. Based on secret seeds, secret shares are developed. Moreover, one-bit method is used to replace the one bit in the given seed and it generates a new share. Consequently, the generated stream is converted into single stream from left to right, generating the share stream as an output. After ending the first phase, the generated share stream with the repeated watermarking bits with the help of XOR operation. This XOR operation produces the secret combined bits ready to be prepared in the Arabic text in the form of 0/1 series. After ending the second phase, secret combined bits are embedded in the Arabic text and producing the resultant text that is watermarked. The CBSS lies on the same fact where secret shares are used for the reconstruction of the shares. This requirement is fulfilled for the certain number of parts k. In other words, those outcomes can constitute the k parts which yields watermarks efficiently. The extension of the CBSS in our proposed work is to introduce the security via calculating the accuracy percentage of the generated watermarks. Next, the watermark generated by intruder is compared with the proposed watermark using Arabic text. CBSS system ensures the procedure to calculate the precision percentage of extracted watermarks. Figure 2 shows the Seed shares approach.

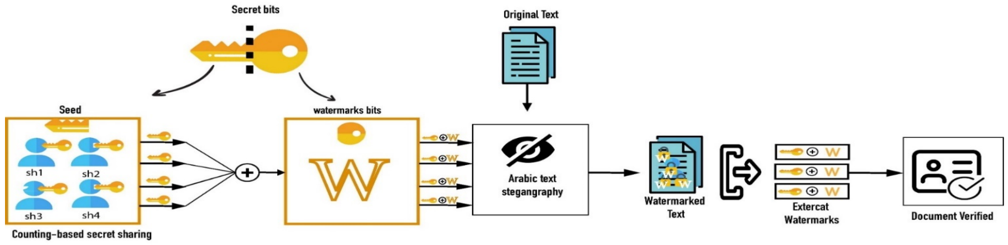

Figure 2. Our proposed Seed shares approach

The WM Shares approach: The advantage of using watermark share approach is its embedding process simplicity avoiding complexities as outlined in Figure 3. The reason is it produce shares of the watermark and merging it with XOR operation by using the repeated seeds. The process gets repeated until sufficient spots in the text are covered. To extract data from any section of the text is challenging, because the hint for starting and ending the shares are not found. To resolve this issue, all possibilities of the share positions should be tested. For this purpose, the XOR output from the consecutive two parts must be combined to see whether the share position is correct or not.

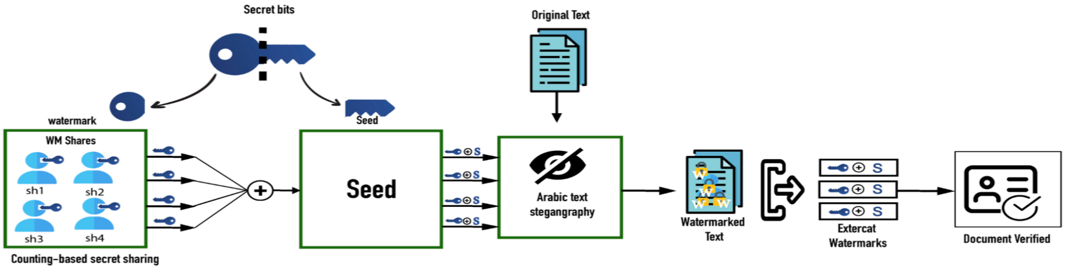

Figure 3. Our proposed WM shares approach

Accordingly, in the first phase, k is set where it decides the use of number of bits, then, removal of extra bits takes place from the original input. The least value of k is 4 because least combination that is required is two (2) and combination is tend to be produce two (2) sets. In the second phase, complexity is involved than first phase as positions of the bits are checked for each possibility and result splits in the k parts. After that, both parts are merged and k-1 combinations are achieved. To calculate the percentage, the repeated results are saved in the temporary list. With temporary list, most repeated combinations are acquired for possible positions. Using these acquired positions, a value is extracted for each item in the list. The value is achieved by dividing it to the repeated count. Lastly, the list is filtered using maximum value and by adding percentage, outcome is obtained, as algorithm fully outlined below.

```
1: Create shares:
Input: watermark
Output: shares
for every zero in the watermark
        change this zero only in the watermark to one
        add the changed watermark to shares
end for
```

```
2: Create embedding bits:
Input: key, watermark
Output: embedding stream
shares ← shares of watermark
key ← repeat key for shares count

return xor( key, shares )
```

```
3: Determine k value
k is the count multiplied key count can be contains the bits
Input: key, extracted bits
Output: k
x ← length of extracted bits
y ← length of key
if x < y * 3 then
        return invalid k
else if x > y then
        return k is x / y
end if
return k is ⌊ x / y ⌋
```

```
4: Combine two shares:
Input: share1, share2
Output: combined bits

for every bit in the share1
        if the bit = 1 and same position bit in share2 = 1 then
                add 1 to output
        else
                add 0 to output
        end if
end for
```

6: Extract watermark from bits (with percentage)

```
Input: extracted bits, key
Output: watermarks
k ← k value
bits ← (k * length of key) bits from extracted bits
key ← repeat key for k times
x ← length of key
list ← temporary list
i ← 0 (iterative position)
While i < bits length do
        bits ← rotate bits to left for one step
        xnor (key, bits)
        rotate xnor out to right for i steps
        add maximum repeated key length combined parts to list with repeated count
        i ← i + 1
end while
for each item in list
--------------------------------------------------------------------------------------------------------------------------
if item repeated count / item repetition count in list > max then
                max ← item repeated count / item repetition count in list
        end if
end for
filter list for max only
% ← 100 / (k * (k – 1)) * max
return the list and %
```

RESULTS AND DISCUSSION

The proposed watermark technique, based on two suggested approaches that is discussed in the previous section, is implemented. Specifically, a data analysis is performed remarking attractive results on standard cover, Nawawi Hadith-42, with no formatting or special characters such as comma and extra space. These results are summarized using Excel sheets with the set statistical formulas, as originally listed in Table 1. The aim of this phase is to calculate the differences of two watermarking systems in precision to extract the watermark. The password for embedding procedure is (11010001110100010011101011).

Table 1. The testbench Hadith samples statistics number of letters and number of words.

Text number	Hadith cover text	Number of letters	Number of words
1	الأعمال بالنيات	280	55
2	حديث جبريل	955	187
3	بني الاسلام على خمس	224	46
4	عمل اهل الجنة وعمل اهل النار	536	108
5	من عمل عملا ليس عليه أمرنا فهو رد	123	27
6	الحلال بين والحرام بين	408	84
7	الدين النصيحة	157	30
8	أمرت أن اقاتل الناس حتى يشهدوا ان لا اله الا الله	260	51
9	مانهيتكم عنه فاجتنبوه	216	41
10	إن الله تعالى طيب لا يقبل إلا طيبا	379	77
11	دع مايريبك	172	38
12	من حسن اسلام المرء	109	24
13	لا يؤمن أحدكم حتى يحب لأخيه	159	35
14	لا يحل دم امرئ	165	32
15	من كان يؤمن بالله واليوم الآخر فليقل خيرا او ليصمت	211	42
16	لاتغضب	114	25
17	ان الله كتب الإحسان على كل شي	196	39
18	اتق الله حيثما كنت	179	37
19	احفظ الله يحفظك	375	77
20	اذا لم تستح	161	34
21	آمنت بالله ثم استقم	149	34
22	أرأيت أذا صليت المكتوبات	211	42
23	الطهور شطر الإيمان	288	54
24	إني حرمت الظلم على نفسي	914	180
25	أهل الدثور بالأجور	503	102
26	كل سلامي من الناس عليه صدقة	285	57
27	البر حسن الخلق	140	30
28	وإياكم ومحدثات الأمور فإن كل بدعة ضلالة	380	69
29	ألا أدلك على أبواب الخير	719	144
30	ان الله فرض فرائض فلا تضيعوها	219	44
31	ازهد في الدنيا يحبك الله	224	46
32	لاضرر ولا ضرار	103	24
33	لو يعطى الناس بدعواهم	160	32
34	من رأى منكم منكرا	169	34
35	كل المسلم على المسلم حرام	337	66
36	نفس عن مؤمن كربة	503	104
37	إن الله تعالى كتب الحسنات والسيئات	356	72
38	من عادى لي وليا فقد آذنته بالحرب	341	69
39	إن الله تجاوز لي عن أمتي الخطأ	125	26
40	كن في الدنيا كأنك غريب أو عابر سبيل	243	50
41	لا يؤمن أحدكم حتى يكون هواه تبعا لما جئت به	138	31
42	يا ابن آدم لو بلغت ذنوبك عنان السماء ثم استغفرتني غفرت لك	291	61

The resultant secret combined bit is produced by secret shares and XOR. Recall the standard cover of Nawawi Hadith 42 benchmark has been prepared first by omitting Kashida letter with no formatting. In this context, same testbench is used for all suggested approaches using Arabic text. As a result, different test samples are established for ordinary Arabic text as shown in Figure 4. It is evident from results that the accuracy of seed share approach is better than the watermark share approach in 100% Hadiths. In Figure 5, the comparison is carried out between seed share and watermark approaches in terms of accuracy percentage of extracted watermark. The orange line represents the accuracy of seed shares in Arabic text and blue line shows the accuracy of watermarks based on watermark share approach. Number of watermark extraction is directly proportional to the Kashida number as shown in Hadith#32 (لا ضرر ولاضرار), as number of Kashida is low therefore, watermark extraction percentage is also zero when it is compared to password length. On the other hand, extraction percentage is less than 100% and one correct watermark is obtained. Accuracy of watermark extraction increases as number of zeroes in passwords increases. The reason is number of zeroes produces larger number of shares and it tends to generates accurate watermarks. If password length is taken as 16, then number of zeros will become 4 $\log_2$ (16) = 4.

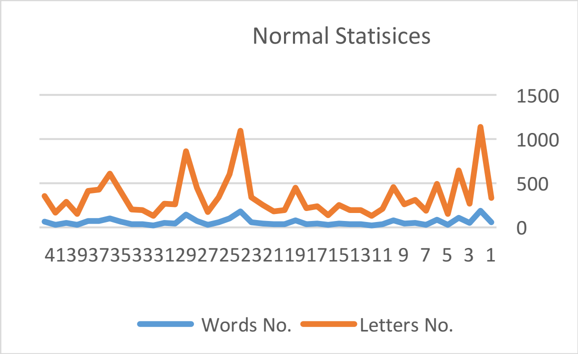

Figure 4. Normal statistics of Hadiths

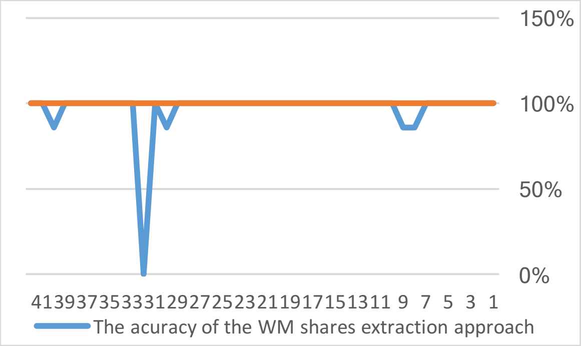

Figure 5. Accuracy comparisons between two Arabic- text-watermarking approaches

COMPARISONS AND ANALYSIS

In section 2, various watermarking techniques are presented. Some of the techniques are closely linked to this research paper, as (Gutub et al., 2010) and (Almehmadi & Gutub, 2021), while some are apart. Therefore, a complete comparison will be carried out in this section based on seed shares watermarking approach (Almehmadi & Gutub, 2021), and seed watermarking system (Gutub et al., 2010). The last approach is important in terms of embedding characters of Kashida. The whole approach is divided into two phases. In the first phase, random values are generated from e-text for security purpose. Security here means the obtained watermark is distinguished from the intruder generated watermark. This is called e-text due to randomly adding the Kashida characters in the Arabic text. In the second phase, another section of Kashida characters are inserted to watermarking text. This comparison is based on the fact that difference needs to be identified when Arabic text file gets corrupted either intentionally or unintentionally. Here, intentional corruption of Arabic text file means that copyright ownership is hindered. Partial extraction and enhancement performance are the driving force of this research. This is the reason CBSS is selected technique for watermark system. The comparison is performed in details in the subsequent sections.

Capacity comparison: The science and practice of secure transmission is known as steganography. On the other hand, steganalysis is the research of revealing the steganographic method (Sahu & Sahu, 2020). In order to obtain large capacity with little distortion, this research in (Sahu & Swain, 2019) suggests a unique method of data concealment employing several stego images. In our proposed work, a greater number of secret bits embedding capacity is supported. The reason behind it is all Kashida location possibilities are used without compromising secret hidden data. This is opposite to the existing work (Gutub et al., 2010) where Kashida location is not fully utilized. Counting-based secret sharing, a seed and XOR are used to enhance security. Precision of proposed work is higher than the previous ones where it must extract three shares at least (Gutub et al., 2010). In Figure 6, grey line shows the proposed work that is higher than existing work i.e., blue line in all 42 Hadith stego-covers. 161 letters in cover 20 are found in 42 Hadith stego-cover whereas, the maximum Kashida in the first approach is obtained as 46 and in the second approach/proposed method, it becomes 58. While in the steganography cover 27, total letters are 140 and the maximum Kashida in the first approach is obtained as 46 and in the second approach and proposed method, it is 2 and 5 respectively. It is evident that number of Kashida in cover are close to each other than the total number of letters, which are different in the cover.

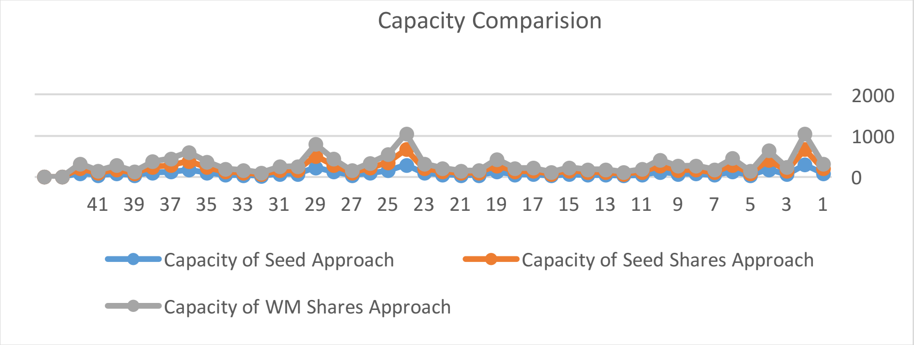

Figure 6. Capacity comparisons of WM approaches

Security comparison: To measure the security, data must be protected from the unauthorized access if it is transmitted or stored (Vaishakh et al., 2019). An eavesdropper is not able for the detection of hidden text and can be called a security breach (Gutub & Ghamdi, 2019). For the security of the Arabic text using steganography, signal to noise ratio should be used as a security metric (Rahman, 2019). This approach is used in testing method as well (Rahman, 2019). The formula to calculate signal to noise ratio is by using MSE where I is considered as Image and noiseless monochrome is represented using mn with noise estimation K, where MAXI is 2B -1 and MATLAB signal to noise ratio is used to implement the above metric (Miller et al., 2022). For this purpose, text is converted into image using Photoshop with resolution 300 pixels/cm. These images are saved without compression resulting in high quality. After the conversion, three types of images are obtained, read and comparison is made among these images with original image. In the same process, cover objects are taken with the stego objects that consists of secret bits. Table 2 lists the security results indicating that if signal to noise ratio is higher, distortion is less. One of the interesting facts is security of the original text turns out to be same where signal to noise ratio is rounded off to the 18 using various techniques. Therefore, it is remarked that security achieved using proposed work is higher, as it does not decrease the capacity. The difference of security can be observed in the watermark approach that contain small value of signal to noise ratio. The 12 bits of password is used by Hadith no. 15. 211 and 42 letters are used by cover media and its secret password becomes [11010001 11010001 00111010 11].

Table 2. Security comparisons of three watermarking approaches

PSNR of		
seed approach	**seed shares approach**	**WM shares approach**
18.1194	17.7919	17.6418

Robustness comparison: Robustness is the ability to endure attacks for hidden data that includes rotation, cropping, noise, compression etc. The two most important components of any authentication-based watermarking approach are tamper detection and localization (Sahu, 2022). Just consider an example where large number of text and content is being transferred via internet and robustness is the relevant term to deal with such scenario. Many research articles, tempering attack is the kind of text attack that has many possible forms such as insert, delete, copy paste, font format, print and retype (Rahman, 2019) and (Aman et al., 2017). Seed approach, watermark approach and proposed approach, Arabic text is tested to evaluate the robustness property against text attacks. The overall text can be used to insert new words or sentences in the original text. There are two ways to insert and delete the processes called localized and dispersed (Gutub & Ghamdi, 2019). To perform localized insertion, one place i.e., beginning, middle or end is considered along with watermarked text. However, in dispersed insertion, the words/sentence can be added randomly using watermark text. Moreover, if we consider localized deletion, words and sentences can be deleted randomly in watermark text. For dispersed deletion, different random places are targeted for deletion in the watermark text and it is the most common attack of the watermark text. In Table 3, accuracy percentage of the watermark is tested for the robustness using the four types of attacks mentioned above. To measure the robustness, it is seen if we can extract the right watermark from the watermark that is attacked. Using Hadith no. 10, robustness of proposed work is evaluated and in Table 3, tested attacks and corresponding results are mentioned.

Table 3. Accuracy percentage comparisons

WM Approaches	Tempering Attacks			
	Localized Insertion	Dispersed Insertion	Localized deletion	Dispersed deletion

Seed	0%	0%	0%	0%
Seed Shares	75 %	42%	100%	78%
WM shares	67%	25%	100%	53%

REMARKS AND CONCLUSION

This presented Arabic text WM approach is related to others briefly linked in terms of capacity, speed and security, as shown in Table 4. By making use of the seed approach presented in (Gutub et al., 2010), the work will suffer low capacity and speed due to process-checking all the watermarked text that exists. Whereas, with this stated approach, Kashida locations are utilized for disguising watermarking locations burdening capacity and security to be undermined. In seed shares technique proposed by Almehmadi & Gutub (2021), high capacity is acquired with medium speed, due to checking at least two shares extracted from the available watermarked text, but considered also complicated process though enhancing security to medium, using CBSS seed bits increasing, as extra layer of difficulty, i.e. for the intruder to extract the watermark added by XOR operation.

On the other hand, repeated WM is tested but decreased level of security. Therefore, an approach used by WM shares showed high capacity and speed, although checking three extracted shares, due to increasing CBSS watermarked bits. It also involved XOR operation with the seed utilizing WM shares to retrieve required secrecy. In Bi-location technique put forth in (Gutub & Alaseri, 2020), it selects one location and leaves the next for hiding Kashida, suffering low capacity and medium speed because of scanning most parts of the watermarking that are not spread sequentially in the entire text. This remarked clear trade-off between capacity and security paying to achieve security by compromising capacity. Likewise, the 2/3 location technique in (Gutub & Alaseri, 2020), selected two Kashida locations, then leaved one, for hiding the watermark commented medium capacity and medium speed in similar trade-off manner. A different technique called Ps-Kashida that is proposed in (Alanazi et al., 2021), hides the secret at different positions, i.e. in spaces between words, depending on letters accepting or rejecting Kashidas. The work enjoyed high capacity on the price of speed, needing to check all the watermarked text. These watermarking analyses need to be tested further in future studies, as their performance tuned to utilizing both strategies of image watermarking semi-authentication (Gutub, 2022a) and dynamic smart preference for medical image confidentiality (Gutub, 2022b) which can innovate enhancement attractive opportunities.

Table 4. Comparisons with others

Approach	**Capacity**	**Speed**	**Security**
Seed	Low	Low speed	Low security
Seed shares	High	Medium speed	Medium security
WM shares	High	High speed.	Better security
Bi-location	Low	Medium speed	medium security
2/3 location	Medium	Medium speed	Medium
Ps-Kashida	High	Low speed	Medium

ACKNOWLEDGMENT

Special thanks to my genius graduate students of Computer Engineering course: Advanced Topics in Information Security (1400519), as their collaborative work remarked attractive contribution and completion used within this paper. Thanks to our great school, Umm Al-Qura University – Makkah, for supporting me by the sabbatical leave year (2022-2023) to focus working out this Cybersecurity research hoping to continue offering the MS graduate program as well as allowing for PhD, approved by Ministry of Education within Saudi Arabia, opening doors for progressive scientific activities.

Agreement: We declare that this work is original and not considered to be published in any other publication media.

Funding: There is no funding source.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Conflict of Interest: The authors declare that they have no conflict of interest.

Data Availability: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

REFERENCES

Alanazi, N., Khan, E. & Gutub, A. 2021. Involving spaces of Unicode standard within irreversible Arabic text steganography for practical implementations. Arabian Journal for Science and Engineering, 46(9): 8869-8885.

Almehmadi, E. & Gutub, A. 2021. Novel Arabic e-text watermarking supporting partial dishonesty based on counting-based secret sharing. Arabian Journal for Science and Engineering, 47(2): 2585-2609.

Almazrooie, M., Samsudin, A., Gutub, A., Salleh, M., Omar, M. & Hassan, S. 2020. Integrity verification for digital Holy Quran verses using cryptographic hash function and compression. Journal of King Saud University-Computer and Information Sciences, 32(1): 24-34.

Aman, M., Khan, A., Ahmad, B. & Kouser, S. 2017. A hybrid text steganography approach utilizing Unicode space characters and zero-width character. International Journal on Information Technologies and Security, 9(1): 85-100.

Anand, A., Singh, A., Lv, Z. & Bhatnagar, G. 2020. Compression-then-encryption-based secure watermarking technique for smart healthcare system. IEEE Multi Media, 27(4): 133-143.

El Rahman, S. 2019. Text steganography approaches using similarity of English font styles. International Journal of Software Innovation (IJSI), 7(3): 29-50.

Gutub, A., Al-Juaid, N. & Khan, E. 2019. Counting-based secret sharing technique for multimedia applications. Multimedia Tools and Applications, 78(5): 5591-5619.

Gutub, A. & Alaseri, K. 2020. Hiding shares of counting-based secret sharing via Arabic text steganography for personal usage. Arabian Journal for Science and Engineering, 45(4): 2433-2458.

Gutub, A., Al-Haidari, F., Al-Kahsah, K. & Hamodi, J. 2010. e-Text watermarking: utilizing 'Kashida' extensions in Arabic language electronic writing. Journal of Emerging Technologies in Web Intelligence, 2(1): 48-55.

Gutub, A. & Al-Ghamdi, M. 2019. Image based steganography to facilitate improving counting-based secret sharing. 3D Research, 10(1): 1-36.

Gutub, A. 2022a. Watermarking images via counting-based secret sharing for lightweight semi-complete authentication. International Journal of Information Security and Privacy (IJISP), 16(1): 1-18.

Gutub, A. 2022b. Dynamic smart random preference for higher medical image confidentiality. Journal of Engineering Research (JER), in press. http://doi.org/10.36909/jer.17853

Gutub, A. 2022c. Enhancing Cryptography of Grayscale Images via Resilience Randomization Flexibility. International Journal of Information Security and Privacy (IJISP), in press. http://doi.org/10.4018/IJISP.307071

Gutub, A. & Fattani, M. 2007. A Novel Arabic Text Steganography Method Using Letter Points and Extensions. International Journal of Computer, Electrical, Automation, Control and Information Engineering, 1(3):502-505.

Jalil, Z. 2010. Copyright protection of plain text using digital watermarking (Doctoral dissertation, FAST National University of Computer & Emerging Sciences, Islamabad, Pakistan.).

Miller, M., Bloom, J., Fridrich, J., & Kalker, T. 2022. Digital Watermarking and Steganography. https://www.mathworks.com

Sahu, M., Padhy, N., Gantayat, S. & Sahu, A. 2022. Local binary pattern- based reversible data hiding. CAAI Transactions on Intelligence Technology.

Suresh, G., Narla, V., Gangwar, D. & Sahu, A. 2022. False-Positive-Free SVD Based Audio Watermarking with Integer Wavelet Transform. Circuits, Systems, and Signal Processing. 1-26.

Sahu, A. 2022. A logistic map based blind and fragile watermarking for tamper detection and localization in images. Journal of Ambient Intelligence and Humanized Computing, 13(8): 3869-3881.

Sahu, A. & Sahu, M. 2020. Digital image steganography and steganalysis: A journey of the past three decades. Open Computer Science, 10(1): 296-342.

Sahu, A. & Swain, G. 2019. A novel multi stego-image based data hiding method for gray scale image. Pertanika Journal of Science & Technology, 27(2): 753-768.

Thabit, R., Udzir, N., Yasin, S., Asmawi, A. & Gutub, A. 2022. CSNTSteg: Color Spacing Normalization Text Steganography Model to Improve Capacity and Invisibility of Hidden Data. IEEE Access 10:65439-65458.

Thyagarajan, C., Suresh, S., Sathish, N. & Suthir, S. 2020. A typical analysis and survey on healthcare cyber security. Int. Journal of Scientific & Technology Research, 9(3): 3267-3270.

Vaishakh, K., Pravalika, A., Abhishek, D., Meghana, N. & Prasad, G. 2019. A semantic approach to text steganography in sanskrit using numerical encoding. In Recent Findings in Intelligent Computing Techniques. 181-192.

Zear, A., Singh, A. & Kumar, P. 2018. Multiple watermarking for healthcare applications. Journal of Intelligent Systems, 27(1): 5-18.