Computer Systems
Science and Engineering

Tech Science Press

DOI: 10.32604/csse.2023.037305

ARTICLE

Digital Text Document Watermarking Based Tampering Attack Detection via Internet

Manal Abdullah Alohali[1], Muna Elsadig[1], Fahd N. Al-Wesabi[2], Mesfer Al Duhayyim[3], Anwer Mustafa Hilal[4,*] and Abdelwahed Motwakel[4]

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

[2]Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Abha, 62217, Saudi Arabia

[3]Department of Computer Science, College of Sciences and Humanities-Aflaj, Prince Sattam bin Abdulaziz University, Al-Aflaj, 16733, Saudi Arabia

[4]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia

*Corresponding Author: Anwer Mustafa Hilal. Email: a.hilal@psau.edu.sa

Received: 29 October 2022 Accepted: 02 February 2023 Published: 20 May 2024

ABSTRACT

Owing to the rapid increase in the interchange of text information through internet networks, the reliability and security of digital content are becoming a major research problem. Tampering detection, Content authentication, and integrity verification of digital content interchanged through the Internet were utilized to solve a major concern in information and communication technologies. The authors' difficulties were tampering detection, authentication, and integrity verification of the digital contents. This study develops an Automated Data Mining based Digital Text Document Watermarking for Tampering Attack Detection (ADMDTW-TAD) via the Internet. The DM concept is exploited in the presented ADMDTW-TAD technique to identify the document's appropriate characteristics to embed larger watermark information. The presented secure watermarking scheme intends to transmit digital text documents over the Internet securely. Once the watermark is embedded with no damage to the original document, it is then shared with the destination. The watermark extraction process is performed to get the original document securely. The experimental validation of the ADMDTW-TAD technique is carried out under varying levels of attack volumes, and the outcomes were inspected in terms of different measures. The simulation values indicated that the ADMDTW-TAD technique improved performance over other models.

KEYWORDS

Content authentication; tampering attacks; detection model; security; digital watermarking

1 Introduction

Security problems of digital text in different formats and languages have anticipated greater emphasis on communication technologies, particularly concerning copyright protection, content authentication, and integrity verification [1,2]. Different applications, like e-Banking and e-commerce, impose numerous problems with content transmitted through the Internet. Most digital media transported through the Internet is in text form and can be extremely sensitive regarding structure, semantics, syntax, and content [3]. Malicious attackers tempers this digital content during transmission, and thereby the amended content might lead to a wrong decision [4,5]. Various solutions for data security were introduced for different purposes that involve data hiding, encryption, integrity verification, unauthorized access control, and copyright protection [6]. Digital watermarking (DW)and Steganography were the two popular technologies for hiding data for numerous purposes like copyright protection and content authentication. Shkilev et al. [7] presented an Evolutionary Optimization powered Watermark for Tampering Attack Detection in the Digital Document (EO-WTAD3) method. A digital watermarking technique was developed for authentication and article copyright security through data mining notion. Additionally, the EO-WTAD3 method employs the idea of data mining. Likewise, the fractional gorilla troops optimization (FGTO) method could be implemented for evaluating the optimum condition. Jana et al. [8] introduced a self-embedding fragile watermarking method for tamper identification and recovery dependent upon local image features employing the Absolute Moment Block Truncation Coding (AMBTC) method and fuzzy logic (FL). By employing the similarity matrix, the blocks have been considered reliant on similarity among neighbored pixels. This solution requires a few transformations or modifications on digital text content to embed watermark data within text. Fig. 1 represents the overall process of the tampering detection process.

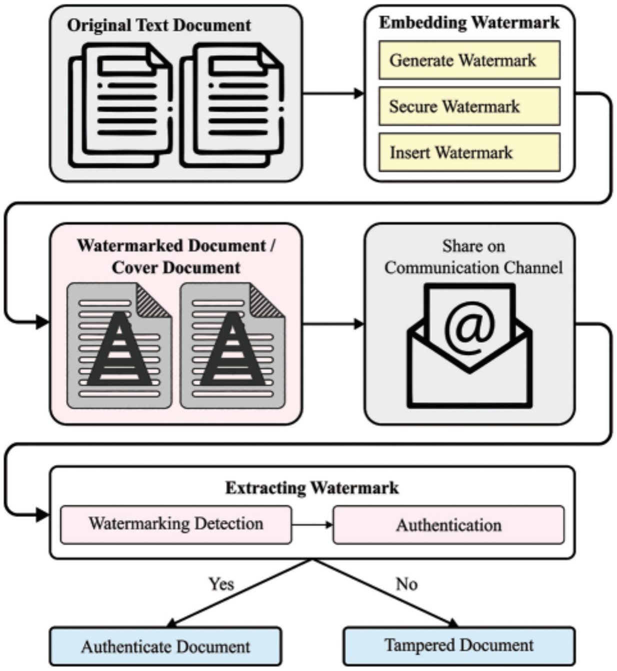

Figure 1: Overall tampering detection process

Researchers have greater attention and developed many solutions for copyright protection, content authentication, integrity verification, and tampering recognition of audio, video, and digital images [9,10]. On the other hand, limited studies were conducted in these fields to identify the appropriate solution for integrity verification of delicate online digital media [11]. Currently, tampering detection and authentication of digital texts have great attention [12,13]. Furthermore, a study in text watermarking focuses on copyright protection; however, it provided less attention and interest in content authentication, integrity verification, and tamper detection owing to the nature of text content that is regarded as a natural language dependent [14]. The general problems in these areas are to present the appropriate technique and solution for distinct languages, formats, and contents, namely Arabic or English [15]. Therefore, tamper detection, content authentication, and integrity verification of sensitive text are crucial challenges in different applications and require suitable solutions.

This study develops an Automated Data Mining based Digital Text Document Watermarking for Tampering Attack Detection (ADMDTW-TAD) via the Internet. In the presented ADMDTW-TAD technique, the DM concept is exploited to identify the appropriate characteristics of the document to embed a larger size of watermark information. Once the watermark is embedded with no damage to the original document, it is then shared with the destination. The watermark extraction process is performed to get the original document securely. The experimental validation of the ADMDTW-TAD technique is carried out under varying levels of attack volumes, and the outcomes are examined in terms of different measures.

In short, the paper contributions are listed as follows:

- Develop an ADMDTW-TAD technique for tampering text document detection in English, which is transmitted via the Internet.
- Employ the DM concept for identifying proper features of the document to embed a larger size of watermark information
- Perform watermark embedding and watermark extraction processes for effectively transmitting the text document over the Internet.
- Examine the performance of the proposed model on varying attack volumes, attacks, and datasets.

The rest of the study is planned as follows. Section 2 offers a brief review of tampering attack detection models. Next, Section 3 introduces the proposed model. Later, Section 4 provides the performance validation of the proposed model. Finally, Section 5 concludes with major key research findings and possible enhancements.

2 Related Works

In [16], the authors modelled a method related to word mechanism and first-level order of Markov method called Robust English Text Watermarking and Natural Language Processing Approach (RETWNLPA) to scale up the accurateness of tampering identification of delicate English text. This method detects and embeds the watermark logically without modifying simple text files. The 1st level order of word mechanism relies upon the hidden Markov model (HMM) employed for examining the relationship among English text. The derived features were employed as watermark data and compiled with text-zero watermarking approaches. In [17], the authors devised a Hybrid of NLP and Zero-Watermarking Approach (HNLPZWA) for content tampering detection and authentication of Arabic textual data. The method above embeds and identifies watermarks without modifying the text file for embedding a watermark key. The fifth level order of word system related to HMM was compiled with digital zero-watermarking approaches for enhancing tampering recognition accuracy problems of the earlier literature devised by the authors. The fifth level order of HMM was employed as an NLP method for examining the Arabic text.

In [18], the authors developed a combined method termed CAZWNLP (a combined approach of zero-watermarking and NLP) for tampering identification of English text interchanged via the Internet. The text-watermarking technologies are used in the 3rd gram of alphanumeric of HMM for boosting the accuracy and performance of tampering recognition problems which can be restricted by prevailing works studied in the literature. To examine an English text and extract the textual features of given contexts, the 3rd-grade level of HMM was employed in this technique as NLP. In [19], the authors presented a Smart-Fragile Approach related to Soft Computing and Digital Watermarking (SFASCDW) text zero-watermarking method for content tampering detection and authentication of English text. To scale up the watermark sturdiness of the presented method, a 1st level-order alphanumeric system related to HMM was compiled with digital zero-watermarking approaches. The author used the 1st level order and the alphanumeric system to examine English text as a soft computing method.

In [20], the authors modelled the Hybrid Structural component, and word length (HSW) related zero watermarking methodologies where the context of text files is not changed for watermark embedding. It includes 2 steps they are extraction and watermark embedding. The watermark and the text were registered to certifying authority, and they can be further utilized for pattern matching and identifying tampering in text files.

3 The Proposed Model

In this study, we have developed a new ADMDTW-TAD approach for the identification of tampering attacks on the Internet. In the presented ADMDTW-TAD technique, the DM concept is exploited to identify the appropriate characteristics of the document to embed a larger size of watermark information. The presented secure watermarking scheme intends to transmit digital text documents over the Internet securely. Once the watermark is embedded with no damage to the original document, it is then shared with the destination. The watermark extraction process is performed to get the original document securely.

3.1 Watermark Embedding Process

The study discusses how to watermark data can be embedded in the document. The confidential messages are encoded AES technique with 256 bits; Encryption can be employed to secure the message [21]. The encrypted messages are transported to the next stage, wherein the watermark is produced. The encrypted messages are transformed into binary numbers. Algorithm 1 divides the numbers into 4 equivalent parts and saves them into different parameters (a, b, c, d). The length (L) of numbers can be measured and added zero later in the first place. Logarithm Base 10 decreases the variable value (a, b, c, d). The inverse function to exponential can be utilized in mathematics for finding the Anti-log.

The logarithm of “x” was the exponent to other numbers that are predetermined; base “b” should be raised for “x” number is given as follows:

$$ln\left(x\right) = log_e\left(x\right) \tag{1}$$

$$e = \lim_{n}\left(1 + n\right)^n \tag{2}$$

The inverse of the logarithm is named an anti-logarithm evaluated by rising base “b” to logarithm “y”.

$$x = log^{(-1)}\left(y\right) = b^y \tag{3}$$

The aim of using $\log_{10}(x)$ the output is often nearer 0 and 1. Then, the original document takes as input. Data mining was used to find appropriate properties from the MS Word document, which has a set of objects with different methods and attributes. The word document includes document and application classes. The special property of MS Word documents is applicable for 2 purposes. Firstly, big data is saved without adversely affecting the entire document. Next, MS Word mutual command won't affect watermark data.

The watermark data can be split into equivalent groups entrenched in this property, and later the second-level embedding begins. During second-level embedding, the MS Word document margin from the layouts is targeted. The values of margin-bottom, margin-left, margin-top, and margin-right can be replaced and modified by four parameters correspondingly.

The watermarked document can be produced in PDF, and the verification procedure, while the document format was transformed, the document margin could not be changed. Once the MS Word document is converted into PDF or PDF to Word, the Layout and Margins of the document remain the same. Afterwards embedding the watermark, we convert MS Word documents into PDFs and save them or share them through the cloud. Algorithm 1 shows the entire process of watermark embedding.

Algorithm 1: Watermark Embedding

```
Inputs: $Key\ (K)$, $Document\ (T)$, $Secret\ message\ (SM)$,
Output: Watermarked Document $\left(\acute{T}\right)$
Initiate:
Data: $E_m$, $N_n$, $WSP$, $B_s$, $SP$, $DM_{argin,}$ $W_{log}$, $W_{log1}$, $W_{log2}$, $W_{log3}$
$W_{log4}$,
Variable Declaration:
     $E_m = Encoded\ message$
    $N_n = Natural\ numbers$
    $WSP = MS\text{-}word\ special\ properties$
    $B_s = Binary\ string$
    $DM_{argin} = Document\ margins$
    $SP = Suitable\ properties$
    $W_{log} = log\ variables$
Initialization:
                              $SP = 0$
                          $E_m \leftarrow (SM,\ K)$
                            $B_s \leftarrow E_m$
                            $N_n \leftarrow B_s$
    Data Mining (WSP)
    for ($i\ to\ WSP$) do
        WSP = [$SP$]
        SP $W_{\log}$
    end for
        Verify document margins $\left(DM_{argin}\right)$
        Set $DM_{argin}$
            Top $W_{log1}$
            Left $W_{log2}$
```

(Continued)

Algorithm 1 (continued)

```
            Right $W_{log3}$
            Bottom $W_{log4}$
            $Convert\ (T) \Rightarrow PDF$
                    $PDF \Rightarrow PDF\acute{T}$
        end
```

3.2 Watermark Extraction Process

Verification or Watermark extraction extracts the watermark (confidential data) from the watermarked document. This is the reverse procedure of watermark embedding. A PDF in the cloud can be provided as input and converted into MS Word. Initially, the value is gained from the special property, and anti-log was used for retrieving the actual value. The 4 temporary parameters are utilized for storing value and later concatenate to individual parameters "*M*".

The system identifies the right, top, bottom, and left margins of the document layout automatically and stores them into individual parameters such as ("T", "B", "L", "R"). Anti-log was used for retrieving the original value and concatenate to ''D". The "*M*" and "*D*" variables concatenate. The result is generated as a number string.

Algorithm 2: Watermark Extraction

```
Inputs: Key (K), Watermarked Document $\left(\acute{T}\right)$,
Output: Secret message (M)
Initiate:
Data: $WSP,\ E_m,\ N_n,\ M,\ D,\ B_s,\ T,\ B,\ L,\ R,\ W_{log},\ DM_{argi},$
Variable Declaration:
            $WSP = MS$-Word special properties
            $E_m$ = Encrypted message
            $N_n$ = Natural numbers
            $B_s$ = Binary string
            $DM_{argin}$ = Document margins
            $W_{log} = log$ variables
        Retrieve $(\acute{T})$ from Cloud
      $PDF \Rightarrow MS$-Word
        Retrieve from WSP
        Verify $(WSP)$
for $(i \Rightarrow WSP)$ do
                    $W_i = [WSP]$
                    $M \leftarrow W_{log}$
end for
          Verify document margins $\left(DM_{argin}\right)$
        $T \Rightarrow DM_{argin}$ [Top]
        $B \Rightarrow DM_{argin}$ [Bottom]
        $L \Rightarrow DM_{argin}$ [Left]
        $R \Rightarrow DM_{argin}$ [Right]
```

(Continued)

Algorithm 2 (continued)

```
            $D \Rightarrow T,\ B,\ L,\ R$
            $Anto - \log \ \Rightarrow D$
            $E_m = M + D$
            $N_n = E_m$
            $B_s = N_n$
            $C_{har} = B_s$
            $D_m = AES\,(C_{har})$
            $SM = D_m$
    $SM$ ⇒Secret Message
end
```

Further, convert number string into binary and later revert into character. AES with 256-bit is utilized for encryption that can be employed to decrypt the cover message. The same Key will be applied to decode text that provides secret messages concealed in the document. Algorithm 2 defines the whole process of verification or watermark extraction.

4 Results and Discussion

This section examines the effective tampering attack detection results of the ADMDTW-TAD model. Table 1 and Fig. 2 provide the performance validation of the ADMDTW-TAD model under different attacks and volumes in terms of TDA. The experimental outcomes demonstrated that the ADMDTW-TAD model had enhanced TDA values under all attacks and attack volumes. For instance, on attack volume of 5%, the ADMDTW-TAD model has attained a TDA of 94.67% under insertion attack, 93.10% under deletion attack, and 89.07% under reorder attack.

Table 1: Overall TDA outcome of ADMDTW-TAD approach with varying attacks and volume

Attack volume (%)	Attacks		
	Insertion	Deletion	Reorder
5	94.67	93.10	89.07
10	91.34	85.24	80.05
20	82.19	71.54	67.47
50	65.65	36.56	45.57

Meanwhile, on attack volume of 10%, the ADMDTW-TAD methodology has attained a TDA of 91.34% under insertion attack, 85.24% under deletion attack, and 80.05% under reorder attack. Moreover, on attack volume of 20%, the ADMDTW-TAD approach has attained a TDA of 82.19% under insertion attack, 71.54% under deletion attack, and 67.47% under reorder attack. Finally, on attack volume of 50%, the ADMDTW-TAD technique has achieved a TDA of 65.65% under insertion attack, 36.56% under deletion attack, and 45.57% under reorder attack.

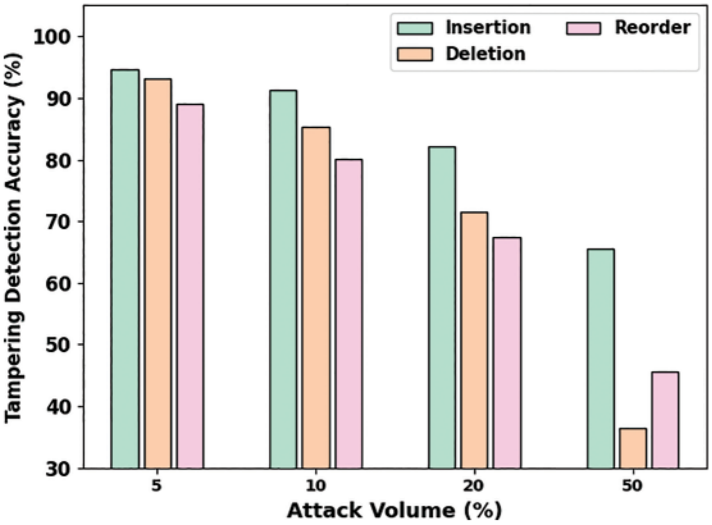

Figure 2: Overall TDA outcome of ADMDTW-TAD approach with varying attack volumes

Table 2 and Fig. 3 provide the performance validation of the ADMDTW-TAD model under different datasets and in terms of $accu_y$. The experimental outcomes demonstrated that the ADMDTW-TAD model had gained enhanced $accu_y$ values under different datasets. For instance, on the ESST dataset, the ADMDTW-TAD model has attained an $accu_y$ of 75.92%. Meanwhile, under the EMST dataset, the ADMDTW-TAD approach has attained an $accu_y$ of 74.26%. Moreover, on the EHMST dataset, the ADMDTW-TAD method has reached an $accu_y$ of 73.07%. Finally, on the ELST dataset, the ADMDTW-TAD technique has achieved an $accu_y$ of 72.91%.

Table 2: Overall TDA outcome of ADMDTW-TAD approach with varying datasets

Datasets	Tampering detection accuracy (%)
ESST	75.92
EMST	74.26
EHMST	73.07
ELST	72.91

Table 3 and Fig. 4 represent the overall TDA outcomes of the ADMDTW-TAD model with existing models under four datasets [16]. The results indicated the improved efficacy of the ADMDTW-TAD model over other models. For instance, on the ESST dataset, the ADMDTW-TAD model has reached an increased TDA of 75.92% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have obtained a reduced TDA of 68.28%, 70.54%, and 74.94% respectively. Meanwhile, on the EMST dataset, the ADMDTW-TAD model has reached an increased TDA of 74.26% while the HNLPZWA, ZWAFWMMM, and RETWNLPA methods have obtained reduced TDA of 64.42%, 68.21%, and 73.06% correspondingly. Eventually, on the EHMST dataset, the ADMDTW-TAD approach has reached an increased TDA of 74.07% while the HNLPZWA, ZWAFWMMM, and RETWNLPA methods have found reduced TDA of 58.95%, 65.45%, and 71.86% correspondingly. Furthermore, on the ELST dataset, the ADMDTW-TAD model has reached an increased TDA of

72.91% while the HNLPZWA, ZWAFWMMM, and RETWNLPA methods have gained reduced TDA of 54.26%, 61.45%, and 71.41% correspondingly.

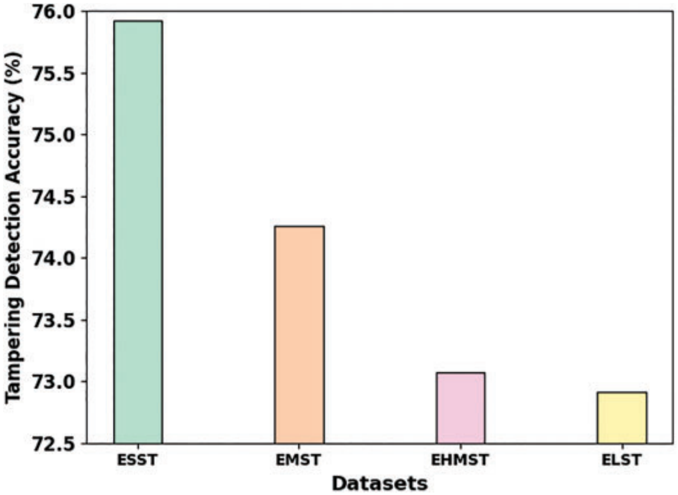

Figure 3: Overall TDA outcome of ADMDTW-TAD approach with varying datasets

Table 3: Comparative TDA outcomes of ADMDTW-TAD model with existing systems under varying datasets [16]

Datasets	HNLPZWA	ZWAFWMMM	RETWNLPA	ADMDTW-TAD
ESST	68.28	70.54	74.94	75.92
EMST	64.42	68.21	73.06	74.26
EHMST	58.95	65.45	71.86	74.07
ELST	54.26	61.45	71.41	72.91

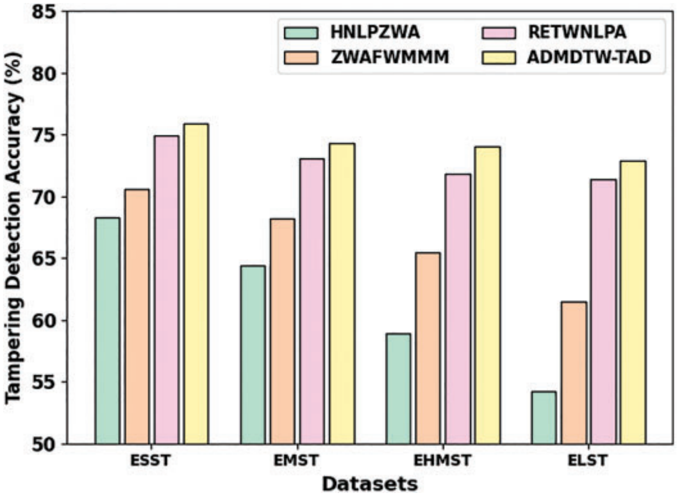

Figure 4: Comparative TDA outcomes of the ADMDTW-TAD model under varying datasets

Table 4 and Fig. 5 signify overall TDA outcomes of the ADMDTW-TAD approach with prevailing techniques under varying attacks. The outcomes indicated the improved efficacy of the ADMDTW-TAD method over other techniques. For example, on insertion attack, the ADMDTW-TAD approach has reached an increased TDA of 84.80% while the HNLPZWA, ZWAFWMMM, and RETWNLPA methods have attained reduced TDA of 72.61%, 79.79%, and 83.51% correspondingly. In the meantime, on deletion attack, the ADMDTW-TAD model has reached an increased TDA of 70.36% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have obtained reduced TDA of 61.13%, 65.52%, and 69.46%, respectively. Eventually, on reordering attack, the ADMDTW-TAD method has reached an increased TDA of 60.86% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have obtained reduced TDA of 37.61%, 44.37%, and 59.17% correspondingly.

Table 4: Comparative TDA outcome of ADMDTW-TAD approach with other methods under varying attacks

Attack type	HNLPZWA	ZWAFWMMM	RETWNLPA	ADMDTW-TAD
Insertion	72.61	79.79	83.51	84.80
Deletion	61.13	65.52	69.46	70.36
Reorder	37.61	44.37	59.17	60.86

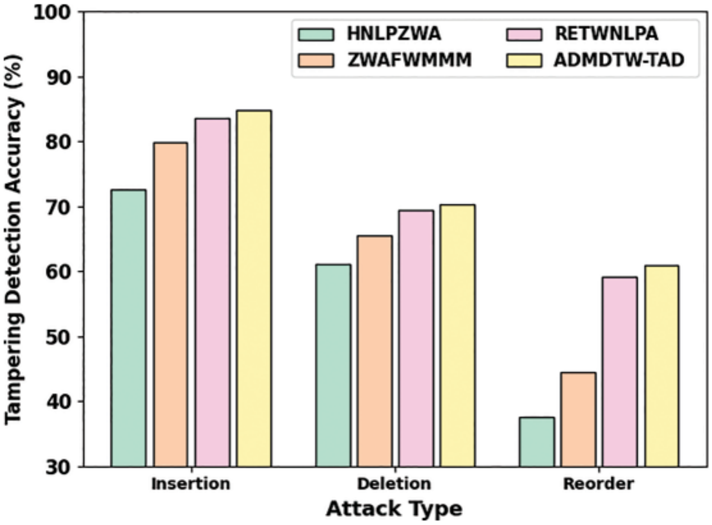

Figure 5: Comparative TDA outcome of ADMDTW-TAD approach under varying attacks

Table 5 and Fig. 6 denote the overall TDA outcomes of the ADMDTW-TAD method with existing methods under varying attack volumes. The outcomes indicated the improved efficacy of the ADMDTW-TAD model over other models. For example, on attack volume of 5%, the ADMDTW-TAD technique has reached an increased TDA of 90.58% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have gained reduced TDA of 81.93%, 82.92%, and 89.32%, respectively. In the meantime, on attack volume of 10%, the ADMDTW-TAD approaches have reached an increased TDA of 81.55% while the HNLPZWA, ZWAFWMMM, and RETWNLPA methods have gained a reduced TDA of 71.98%, 74.84%, and 81.68% correspondingly. Eventually, on attack volume of 20%, the ADMDTW-TAD technique has reached an increased TDA of 68.81% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have gained reduced TDA of 58.64%, 58.61%, and 66.54% correspondingly. Also, on attack volume of 50%, the ADMDTW-TAD model has reached an increased TDA of 47.25% while the HNLPZWA, ZWAFWMMM, and RETWNLPA models have obtained reduced TDA of 13.31%, 38.79%, and 46.55%, respectively.

Table 5: Comparative TDA outcomes of ADMDTW-TAD system with recent methods under varying attack volumes

Attack volume (%)	HNLPZWA	ZWAFWMMM	RETWNLPA	ADMDTW-TAD
5	81.93	82.92	89.32	90.58
10	71.98	74.84	81.68	81.55
20	58.64	58.61	66.54	68.81
50	13.31	38.79	46.55	47.25

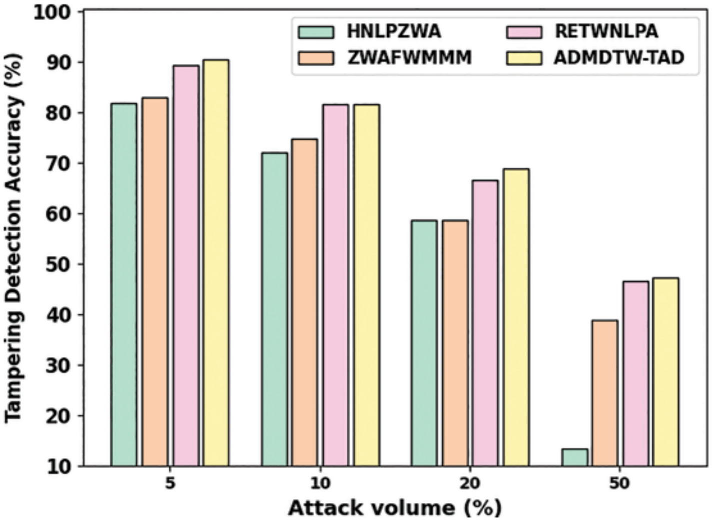

Figure 6: Comparative TDA outcomes of ADMDTW-TAD system under varying attack volumes

These results show the improved outcomes of the ADMDTW-TAD model over other existing models.

5 Conclusion

This study has developed a new ADMDTW-TAD technique for tampering Attack Detection in English text documents via the Internet. In the presented ADMDTW-TAD technique, the DM concept is exploited to identify the appropriate characteristics of the document to embed a larger size of watermark information. The presented secure watermarking scheme intends to transmit digital text documents over the Internet securely. Once the watermark is embedded with no damage to the original document, it is then shared with the destination, where the watermark extraction process is performed to get the original document securely. The experimental validation of the ADMDTW-TAD technique is carried out under varying levels of attack volumes, and the outcomes are examined in terms of different measures. The simulation values indicated that the ADMDTW-TAD technique had reached improved performance over other existing models. Deep learning models can be employed to accomplish enhanced content authentication.

Acknowledgement: This work was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University.

Funding Statement: This work was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Research Groups Program Grant No. (RGP-1443-0051).

Author Contributions: Manal Abdullah Alohali conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft; **Muna Elsadig** conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft; **Fahd N. Al-Wesabi** conceived and designed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft; **Mesfer Al Duhayyim** performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft; **Anwer Mustafa Hilal** performed the experiments, performed the computation work, authored or reviewed drafts of the article, and approved the final draft; **Abdelwahed Motwakel** performed the experiments, analyzed the data, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.

Availability of Data and Materials: Data sharing not applicable to this article as no datasets were generated during the current study.

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

References

[1] X. Tang, S. Liu, W. Che and W. Tang, "Tampering attack detection in analog to feature converter for wearable biosensor," in *2022 IEEE Int. Symp. on Circuits and Systems (ISCAS)*, Austin, TX, USA, pp. 1150–1154, 2022.

[2] M. Bashardoost, M. S. 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.

[3] N. Vryzas, A. Katsaounidou, L. Vrysis, R. Kotsakis and C. Dimoulas, "A prototype web application to support human-centered audiovisual content authentication and crowdsourcing," *Future Internet*, vol. 14, no. 3, pp. 1–17, 2022.

[4] F. Saeed and A. Dixit, "Combined markov model and zero watermarking for integrity verification of English text documents," *Intelligent Communication, Control and Devices*, pp. 857–864, 2020. https://doi.org/10.1007/978-981-13-8618-3_88.

[5] L. Laouamer and O. Tayan, "Performance evaluation of a document image watermarking approach with enhanced tamper localization and recovery," *IEEE Access*, vol. 6, pp. 26144–26166, 2018.

[6] A. E. Afify, A. Emran and A. Yahya, "A tamper proofing text watermarking shift algorithm for copyright protection," *Arab Journal of Nuclear Sciences and Applications*, vol. 52, no. 3, pp. 126–133, 2019.

[7] R. Shkilev, A. Kormiltseva, M. Achaeva, A. Tarasova and M. Matquliyeva, "Fortifying textual integrity: Evolutionary optimization-powered watermarking for tampering attack detection in digital documents," *Fusion: Practice and Applications*, vol. 14, no. 2, pp. 97, 2024.

[8] M. Jana, B. Jana and S. Joardar, "Local feature-based self-embedding fragile watermarking scheme for tampered detection and recovery utilizing AMBTC with fuzzy logic," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 19, pp. 9822–9835, 2022.

[9] T. Saba, M. Bashardoost, H. Kolivand, M. S. M. Rahim, A. Rehman *et al.,* “Enhancing fragility of zero-based text watermarking utilizing effective characters list,” *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 341–354, 2020.

[10] S. A. Parah, J. A. Sheikh and G. M. Bhat, “StegNmark: A joint Stego-watermark approach for early tamper detection,” *Intelligent Techniques in Signal Processing for Multimedia Security*, pp. 427–452, 2017. https://doi.org/10.1007/978-3-319-44790-2_19.

[11] S. Trivedy and A. K. Pal, “A logistic map-based fragile watermarking scheme of digital images with tamper detection,” *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 41, no. 2, pp. 103–113, 2017.

[12] A. Eid, A. Emran and A. Yahya, “Tamper proofing text watermarking shift algorithm for copyright protection,” *International Journal of Hybrid Information Technology*, vol. 11, no. 3, pp. 13–22, 2018.

[13] S. A. Shah, I. A. Khan, S. Z. H. Kazmi and F. H. B. M. Nasaruddin, “Semi-fragile watermarking scheme for relational database tamper detection,” *Malaysian Journal of Computer Science*, vol. 34, no. 1, pp. 1–12, 2021.

[14] S. Hakak, A. Kamsin, O. Tayan, M. Y. I. Idris and G. A. Gilkar, “Approaches for preserving content integrity of sensitive online Arabic content: A survey and research challenges,” *Information Processing & Management*, vol. 56, no. 2, pp. 367–380, 2019.

[15] M. T. Ahvanooey, Q. Li, H. J. Shim and Y. Huang, “A comparative analysis of information hiding techniques for copyright protection of text documents,” *Security and Communication Networks*, vol. 2018, pp. 1–22, 2018.

[16] A. M. Hilal, F. N. Al-Wesabi, A. Abdelmaboud, M. A. Hamza, M. Mahzari *et al.,* “A hybrid intelligent text watermarking and natural language processing approach for transferring and receiving an authentic English text via internet,” *The Computer Journal*, vol. 65, no. 2, pp. 423–435, 2022.

[17] F. N. Al-Wesabi, “Proposing high-smart approach for content authentication and tampering detection of Arabic text transmitted via internet,” *IEICE Transactions on Information and Systems*, vol. 103, no. 10, pp. 2104–2112, 2020.

[18] F. N. Al-Wesabi, S. Alzahrani, F. A. M. Al-Yarimi, M. Abdul, N. Nemri *et al.,* “A reliable nlp scheme for English text watermarking based on contents interrelationship,” *Computer Systems Science & Engineering*, vol. 37, no. 3, pp. 297–311, 2021.

[19] M. Alamgeer, F. N. Al-Wesabi, H. G. Iskandar, I. Khan, N. Nemri *et al.,* “Smart-fragile authentication scheme for robust detecting of tampering attacks on English text,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 2497–2513, 2022.

[20] F. Saeed and A. Dixit, “Hybrid hsw based zero watermarking for tampering detection of text contents,” in *Int. Conf. on Computer Networks, Big Data and IoT*, Madurai, India, pp. 820–826, 2018. https://doi.org/10.1007/978-3-030-24643-3_96.

[21] U. Khadam, M. M. Iqbal, M. A. Azam, S. Khalid, S. Rho *et al.,* “Digital watermarking technique for text document protection using data mining analysis,” *IEEE Access*, vol. 7, pp. 64955–64965, 2019.