The ISC Int'l Journal of Information Security
ISeCure

October 2025, Volume 17, Number 3 (pp. 1–6)
http://www.isecure-journal.org

Presented at the ISCISC'2025 in TEHRAN, Iran.

Time-Based Steganography in Text **

Zahra Ghoraeian [1], Mohammad Reza Sadeghi [1,*], and Samaneh Mashhadi [2]
[1] *Department of Mathematics and Computer Science, Amirkabir University of Technology – Tehran Polytechnic, Tehran, Iran.*
[2] *School of Mathematics and Computer Science, Iran University of Science and Technology, Narmak, Tehran, Iran.*

ARTICLE INFO.

Keywords:
Data Confidentiality, Data Hiding, Information Security, Text Steganography, Text Watermarking

Type:

doi:

ABSTRACT

Preserving data confidentiality is crucial in today's digital world where data exchange is increasingly becoming digital. This paper presents a novel text steganography algorithm. Initially, the secret message is converted into a bit stream. This bit stream is then shuffled using a random sequence to enhance security. Finally, the data is converted into a specific "time" (including date and hour), and this generated time is embedded within a suitable cover text. The results demonstrate that the proposed algorithm is robust against a variety of attacks, including retyping, OCR, printing and photocopying, compression, document feature modification, non-Unicode environment conversion, and semantic paraphrasing. The algorithm is language-independent and applicable to all languages. The scheme exhibits high transparency against visual and machine attacks and has a capacity of 18 bits per time. The embedding of information bits using a random sequence enhances the scheme's resistance against detection attacks.

1 Introduction

In today's world where data exchange is increasingly becoming digital, information security is of paramount importance. Cybercrimes and data breaches have become increasingly prevalent, causing significant financial and reputational damage to individuals, organizations, and governments. To combat these challenges, there is a need for secure and efficient methods to protect information from unauthorized access, disclosure, and manipulation. Cryptography and information hiding are two primary approaches to addressing these concerns. Information hiding embeds data within another object, such as an image, audio, or text, to conceal it from unauthorized viewers [1]. Information hiding is further categorized into two main classes: steganography and watermarking. The primary goal of steganography is to preserve the confidentiality of information.

This paper presents a novel approach to text steganography. In this approach, the message text is converted into a bitstream, which is then shuffled using a random sequence. The resulting bitstream is used to generate a time interval comprising the day, hour, minute, and second. The "generated time" is subsequently embedded into a suitable text, and the steganographic text is transmitted. To the best of our knowledge, a similar encoding approach has not

* Corresponding author.
**The ISCISC'2025 program committee effort is highly acknowledged for reviewing this paper.
Email addresses: zahra.ghoraeian@aut.ac.ir, msadeghi@aut.ac.ir, smashhadi@iust.ac.ir
ISSN: 2008-2045

been reported anywhere in the literature.

The remaining sections of the paper are organized as follows: Section 2 provides a review of the related literature to the proposed approach. Section 3 presents the proposed scheme along with an example. Section 4 presents the analysis of the experimental results and discussions related to the proposed approach. Finally, Section 5 concludes this paper with some future work.

2 Related Work

This section provides essential definitions and a comprehensive review of existing text steganography techniques, laying the groundwork for understanding our proposed method. A steganography system consists of four main components: the secret message, the cover media, the steganographic algorithm, and the key. The presence of a key is not mandatory; if present, it turns the system into a private steganography system. The cover media can be of various types, such as images, audio, video, or text. In this paper, the cover medium is text.

Various classifications have been proposed for steganography techniques. A classification based on the distinction between the covertext and the stego-text offers greater coherence and generalizability. From this perspective, text steganography techniques can be technically divided into three categories [2]: a) Algorithms where the covertext and stego-text are identical in terms of textual content, and no difference can be observed between the two texts. These techniques are also called structural techniques or font format and feature code-based techniques. b) Algorithms where the covertext and stego-text have visual differences. c) Algorithms where there is no covertext and the stego-text is generated directly from the secret message.

Considering this, steganography algorithms are primarily divided into three main categories: structural, linguistic, and statistical (random) [2]. Over the past decade, research in text steganography algorithms has primarily focused on structure-based approaches, while less attention has been paid to language-based and random/statistical methods. This imbalance is due to the inherent limitations of these methods, such as low embedding capacity, changes in the meaning of the covertext, and the need for additional prerequisites such as dictionaries and datasets [2].

2.1 Structural Techniques

Structural algorithms are the most common methods in text steganography [2]. These methods embed secret information by making changes to the visual structure of the text without altering its content. Elements such as font, font size, spacing between words and lines, and text color are some of the elements that are manipulated in this method. Methods based on structural techniques are vulnerable to retyping, reformatting, OCR attacks, and sometimes copy-paste, and the embedded secret information is lost [3]. The advantages of these methods include ease of implementation and high capacity [4]. Structural techniques are categorized into six subcategories: open space [5], line shift [6], word shift [7], zero-width characters [8], font feature changes [9], emoticons, and combined and innovative methods [10].

2.2 Linguistic Techniques

Linguistic or natural language processing-based algorithms modify the semantic [11] and syntactic [12] features of text content. These methods utilize linguistic rules to embed information. Linguistic methods can be divided into two categories: semantic and syntactic techniques [13]. Generally, techniques that manipulate the meaning of the text (or alter elements to preserve semantic information) or affect certain character features are highly influenced by context (e.g., prose versus technical writing) as well as the reader's skill. Consequently, designing countermeasures and attacks against them is challenging [14]. These methods are resistant to visual attacks, retyping, OCR, and copy-paste [2].

2.3 Random and Statistical Algorithms

Random and statistical algorithms utilize the statistical properties of the secret message to automatically generate a new covertext and hide the secret information within the text. This process is computationally intensive. These algorithms can be divided into two categories: cover-based techniques and compression-based techniques. These schemes have a high capacity and are resistant to OCR, retyping, and statistical attacks, but they are vulnerable to natural language processing attacks [6].

2.3.1 Compression-based methods

Compression-based methods utilize lossless compression algorithms such as Huffman, LZW, and arithmetic coding to hide secret messages within covertext. These methods have high computational complexity and are inefficient for hiding secret messages in short texts, but they offer high invisibility, optimal capacity, and low robustness against structural attacks. [15] provides an example of compression-based methods.

2.3.2 Random Cover Generation Technique

These techniques create a covertext based on the letters of the secret message. Initially, a function called Emb() generates a cover message (CM) based on the letters of the secret message, and then embeds the bits of the secret message into it. For example, the AH4S technique uses the omega network structure to hide the bits of the secret message in a generated cover message. This method converts the letters of the secret message into related letters and uses a dictionary to find the appropriate cover word. Finally, a long and unfamiliar text is produced for the short secret message, making it difficult to identify. Random cover-based techniques have low transparency, low capacity for hiding information, and the cover message generation is complex [16]. Examples of text generation are presented in [17]. The primary objective of this research is to introduce a novel steganography method based on the generation of "time" in common text files. The detailed steps of the embedding and extraction algorithms will be explained in the next section.

3 The proposed scheme

This method involves converting the hidden text (message) into a specific time. For example, the hidden text can be represented as "Monday, 23:11:30". This time is then embedded within a text containing various dates and times. For instance, it can be inserted into an online shopping list as the purchase time or in a list of student assignment submission times. In this method, days of the week are encoded using two bits (Table 1).

Additionally, a table of alphabets, numbers 0 to 9, and punctuation marks (a total of 48 characters) is prepared. Thus, each character is encoded using six bits (Table 2). It is obvious that for steganography in other languages, it is only necessary to replace the letters of the desired language in the alphabet table (Table 2). This table can be adapted for languages with up to 64 characters.

Table 1. Two-bit Encoding of Days of the Week

Day	Binary Code
Saturday and Sunday	00
Monday and Tuesday	01
Wednesday and Thursday	10
Friday	11

3.1 Embedding Algorithm

The embedding algorithm consists of 5 steps.

(1) Grouping characters: Divide the characters of the hidden message into groups of three.

Table 2. Character Encoding Table

Character index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Character	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
Character index	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Character	q	r	s	t	u	v	w	x	y	z	.	)	(	!	?	,
Character index	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Character	"	'	:	;	0	1	2	3	4	5	6	7	8	9	space	/

(2) Converting groups to binary code: Convert each group, based on its constituent characters, into a sequence of 0s and 1s using Table 2. For example, "you" is encoded as 011001001111010101.

(3) Extracting the day of the week and time: From the first two bits of each resulting 18-bit string, extract the day of the week using Table 1. The next 16 bits are used to represent the time (in seconds) as a binary number. That is, the 16 bits are converted to a decimal number. This number represents the desired time in seconds.

(4) Converting time in seconds to real time: The extracted binary number (seconds) from the previous step is converted to real time (hours, minutes, and seconds).

(5) Inserting time into the text: Finally, the resulting time is inserted into a suitable text, such as an online shopping list or a list of student assignment submission times.

3.2 Extracting Algorithm

The extraction algorithm, consisting of 5 steps, is the inverse of the insertion process (Figure 1).

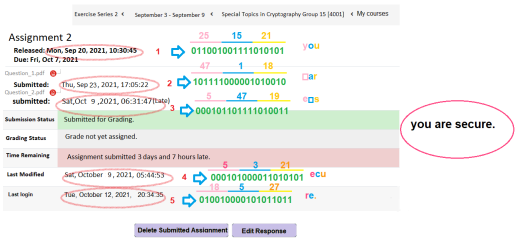


Figure 1. Extracting Algorithm Steps

(1) Identifying times: In the received text, identify and record all times, including day, hour, minute, and second, in the order they appear.

(2) Converting time to a binary sequence:

(a) Converting the day: For each time, convert the day to a 2-bit sequence according to Table 1.

(b) Converting hours, minutes, and seconds to seconds: Convert the hours, minutes, and seconds of each time to seconds using the following formula:

$$\text{Total seconds} = (H \times 3600) + (M \times 60) + S$$

In this formula, H denotes hours, M denotes minutes, and S denotes seconds.

(c) Convert the total seconds to a 16-bit binary number.

(d) Combining bits: Combine the 2-bit sequence resulting from the day conversion with the 16-bit sequence resulting from the second conversion to create an 18-bit sequence.

(3) Converting to 6-bit groups and then to characters:

(a) Dividing into 6-bit groups: Divide the resulting bit sequence into 6-bit groups.

(b) Converting to decimal: Convert each 6-bit group to its equivalent decimal number.

(c) Converting to a character: Using Table 2, convert the decimal number of each group to its corresponding character.

(4) Forming the hidden message: Place the characters obtained from the previous step side by side in order to extract the hidden message.

4 Algorithm Analysis

In the following, the features of the proposed algorithm are examined. Focusing on three main aspects, including capacity, resistance, and transparency, the strengths and weaknesses of the method are identified and its performance is evaluated compared to other existing methods.

4.1 Capacity

This method can hide 18 bits of information (equivalent to 3 characters) for each date (including day of the week and time). Therefore, it is important that the cover text, without losing transparency and without creating doubt, can accommodate several dates. For example, suggested texts can include items such as an online shopping list, the time and date of submitting assignments by students, and similar items, where it is normal to include a large number of dates (Figure 2).

4.2 Resistance

Compared to structure-based steganography algorithms, this scheme is resistant to a wide range of attacks, including retyping, OCR, printing and copying, compression, changing document features, and transferring to a non-Unicode environment. Also, unlike linguistic steganography methods, this scheme is resistant to semantic paraphrasing and changes in word order. However, it is still vulnerable to visual attacks such as deletion.

Transaction Details

The third week of January
From January 10, 2021 to January 16, 2021

No.	Purchaser	Amount (USD)	Transaction reference number	Transaction time
1.	John Doe	$3.14	612345	Monday 03:17:29
2.	Jane Smith	$4.89	789012	Tuesday 18:42:56
3.	Michael Johnson	$7.23	834567	Wednesday 09:23:11
4.	Emily Brown	$9.56	912345	Thursday 21:55:43
5.	David Wilson	$11.89	678901	Friday 14:38:07
6.	Olivia Davis	$13.21	543210	Saturday 06:12:34
7.	Daniel White	$15.54	876543	Sunday 23:09:52
8.	Sophia Rodriguez	$17.87	901234	Tuesday 11:47:18
9.	Matthew Lee	$20.20	567890	Wednesday 02:36:59
10.	Ava Moore	$22.53	432109	Thursday 19:21:44
11.	Benjamin Martin	$24.86	789012	Friday 05:54:27
12.	Charlotte Taylor	$27.19	654321	Saturday 22:30:15
13.	James Anderson	$29.52	901234	Sunday 13:16:41
14.	Abigail Hall	$31.85	567890	Monday 04:51:03
15.	William Harris	$34.18	432109	Tuesday 20:25:38
16.	Victoria Nelson	$36.51	876543	Wednesday 11:09:22
17.	Ethan Carter	$38.84	901234	Thursday 03:43:57
18.	Ella Adams	$41.17	567890	Friday 20:18:32
19.	Alexander Wright	$43.50	432109	Saturday

Figure 2. An example of a shopping list (Transaction Details)

4.3 Transparency (Imperceptibility)

If the appropriate text context is selected, the scheme has visual transparency and there is no distinction between the original cover text and the hidden text.

5 Conclusion

In this paper, an innovative algorithm for text steganography was presented. This algorithm is a type of text generation algorithm and is not dependent on a specific language and can be used in all languages. The proposed algorithm is resistant to most attacks (except for visual attacks involving deletion and replacement) and has good transparency. The algorithm boasts a capacity of 18 bits per date. A key consideration for the algorithm's application is the selection of an appropriate initial textual context for embedding the dates. While finding or creating such a context requires careful consideration, it presents an opportunity to integrate the hidden information naturally and seamlessly within the document.

References

[1] Nor Ashila Roslan, Nur Izura Udzir, Ruzliza Mahmod, and Abdallah Gutub. Systematic literature review and analysis for arabic text steganography method practically. *Egyptian Informatics Journal*, 23(4):177–191, 2022.

[2] Mojtaba Taleby Ahvanooey, Qing Li, Jiarui Hou, Ameer Raza Rajput, and Yang Chen. Modern text hiding, text steganalysis, and applications: a comparative analysis. *Entropy*, 21(4):355, 2019.

[3] Igor Stojanov, Aneta Mileva, and Ivana Stojanovic. A new property coding in text steganography of microsoft word documents. 2014.

[4] R Din, Rafsanjani Ab Thabit, Nur Izura Udzir, and Sonny Utama. Traid-bit embedding process on arabic text steganography method. *Bulletin of Electrical Engineering and Informatics*, 10(1): 493–500, 2021.

[5] Marwah Alaa Majeed, Rosli Sulaiman, and Zaitun Shukur. New text steganography technique based on part-of-speech tagging and format-preserving encryption. *KSII Transactions on Internet & Information Systems*, 18(1), 2024.

[6] Santosh Panwar, Manoj Kumar, and Sheetal Sharma. Text steganography based on parallel encryption using cover text (pect). In *4th International Conference on Internet of Things and Connected Technologies (ICIoTCT), 2019: Internet of Things and Connected Technologies*, pages 303–313. Springer International Publishing, 2020.

[7] R Gurunath and D Samanta. A new 3-bit hiding covert channel algorithm for public data and medical data security using format-based text steganography. *Journal of Database Management (JDM)*, 34(2):1–22, 2023.

[8] Lokesh Kumar Tyagi, Aayush Gupta, and Adel Mohamed. Unveiling the invisible an in-depth analysis of text steganography techniques, challenges, and advancement. In *2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, pages 177–183. IEEE, 2023.

[9] Muhammad Askari, Arfan Mahmood, and Zunaira Iqbal. A novel font color and compression text steganography technique. In *2023 International Conference on Communication, Computing and Digital Systems (C-CODE)*, pages 1–6. IEEE, 2023.

[10] Umer Khadam, Muhammad Mubashir Iqbal, Leonardo Mostarda, and Fida Ullah. An efficient framework for text document security and privacy. In *Security and Privacy in Social Networks and Big Data: 6th International Symposium, SocialSec 2020, Tianjin, China, September 26–27, 2020, Proceedings 6*, pages 132–140. Springer Singapore, 2020.

[11] M Shazzad-Ur-Rahman, Anik Singha, Nipa Ibne Akhtar, M Fahim Ashhab, Kaif Ali, et al. An efficient bengali text steganography method using bengali letters and whitespace characters. In *Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*, pages 477–487. Springer Singapore, 2021.

[12] M Khairullah. A novel text steganography system in financial statements. *IJDTA International Journal of Database Theory and Application*, 7(5):123–132, 2014.

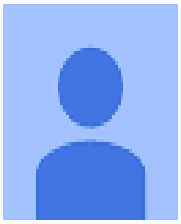
[13] Chin-Yung Chang. *Transformations for linguistic steganography*. PhD thesis, University of Southern California, 2023.

[14] Sven Wendzel, Luca Caviglione, Wojciech Mazurczyk, Aneta Mileva, Jana Dittmann, Christian Krätzer, Jörg Schwenk, and Sarah Zillien. A generic taxonomy for steganography methods. *Authorea Preprints*, 2023.

[15] Mustafa Vefa Arısoy. Lzw-cie: a high-capacity linguistic steganography based on lzw char index encoding. *Neural Computing & Applications*, 34 (21):19117–19145, 2022.

[16] Mohammad Shirali-Shahreza. Pseudo-space persian/arabic text steganography. In *2008 IEEE Symposium on Computers and Communications*, pages 864–868. IEEE, 2008.

[17] Wei Peng, Tao Wang, Zheng Qian, Shu Li, and Xinxing Zhang. Cross-modal text steganography against synonym substitution-based text attack. *IEEE Signal Processing Letters*, 30:299–303, 2023.



Zahra Ghoraeian is currently a Ph.D. candidate in Applied Mathematics at the Department of Mathematics and Computer Science, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. Her research interests include data hiding, information security, and cryptography.

Mohammad-Reza Sadeghi is a Professor at the Department of Mathematics and Computer Science, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. His main research areas cover cryptography, coding theory, and mathematical aspects of information security.

Samaneh Mashhadi is an Associate Professor at the School of Mathematics and Computer Science, Iran University of Science and Technology, Narmak, Tehran, Iran. Her research interests include information hiding, steganography, and related topics in computer security.