Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

ENHANCING KNOWLEDGE SECURITY THROUGH TEXT STEGANOGRAPHY: A REVIEW OF TECHNIQUES AND TOOLS

Mohd Hilal Muhammad[1*], Zulhazlin Abas[2], Mohd Zhafri Mohd Zukhi[3], Muhammad Khairul Zharif Nor A'zam[4]

[1] College of Computing & Mathematical Sciences, Universiti Teknologi MARA (UiTM) Kedah
Email: hilalmuhd@uitm.edu.my
[2] Centre for Foundation Studies, Universiti Islam Antarabangsa Sultan Abdul Halim Mu'adzam Shah (UniSHAMS)
Email: zul@unishams.uitm.edu.my
[3] College of Computing & Mathematical Sciences, Universiti Teknologi MARA (UiTM) Kedah
Email: zhafri319@uitm.edu.my
[4] College of Computing & Mathematical Sciences, Universiti Teknologi MARA (UiTM) Kedah
Email: khairulzharif@uitm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 05.01.2025
Revised date: 18.01.2025
Accepted date: 25.02.2025
Published date: 30.03.2025

To cite this document:

Muhammad, M. H., Abas, Z., Zukhi, M. Z. M., Nor A'zam, M. K. Z. (2025). Enhancing Knowledge Security Through Text Steganography: A Review Of Techniques And Tools. *Journal of Information System and Technology Management, 10* (38), 325-339.

DOI: 10.35631/JISTM.1038022

Abstract:

This study addresses the increasing challenges of ensuring data security and privacy in the digital age, specifically focusing on the limitations of current steganography techniques in maintaining a balance between embedding capacity and the naturalness of text. As digital threats continue to evolve, traditional steganography methods, though effective in concealing information, often compromise the coherence of the text or limit the volume of hidden data. The aim of this study is to explore and analyze emerging technologies and methods that enhance the robustness of text steganography while preserving the integrity and readability of the text. The literature review highlights various approaches, such as synonym substitution, sentence restructuring, and the use of invisible characters, as well as advanced techniques like deep learning and generative adversarial networks (GANs) to improve data embedding. It also examines tools like Steghide, OpenStego, and SilentEye, which are widely used for steganography, yet face challenges in security and usability. The main findings indicate that AI and machine learning-driven techniques significantly improve the capacity and imperceptibility of hidden data while minimizing detection through steganalysis. Additionally, integrating steganography with cryptography and blockchain technologies offers enhanced security by ensuring that hidden data remains protected even if detected. These findings have significant implications for secure communication, suggesting that further refinement of these methods could lead to more robust and adaptive steganographic systems capable of withstanding increasingly sophisticated detection techniques. This study encourages ongoing research into the

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

development of new tools and technologies to address the growing need for secure and covert data transmission.This abstract provides a concise summary of the study, highlighting the problem, aims, literature review, findings, and implications.

Keywords:

Text Steganography, Data Security, Cryptography, Artificial Intelligence, Machine Learning

Introduction

In the digital age, safeguarding knowledge has become a critical concern as the rapid advancement of technology has given rise to sophisticated cyber threats. Knowledge security ensures the protection of sensitive information, intellectual property, and other valuable data from unauthorized access and manipulation. This is particularly important as organizations increasingly rely on digital platforms to store and share information. Techniques such as text steganography, which hides information within ordinary text, offer an innovative method for securing data by concealing it in plain sight, thus providing an additional layer of protection against cyber threats. By embedding hidden data within digital communication, text steganography plays a key role in maintaining confidentiality and safeguarding organizational knowledge (Jevtić & Alhudaidi, 2023). As organizations face increasingly complex challenges in protecting data from breaches and unauthorized access, employing tools and techniques like text steganography becomes essential to maintaining the integrity of critical knowledge assets (Romansky & Noninska, 2020).

Text steganography is a technique used to enhance data protection by embedding hidden messages within text, ensuring that sensitive information remains concealed from unauthorized individuals. Unlike traditional encryption, which transforms data into an unreadable format, text steganography hides the existence of the message itself by embedding it in a way that appears innocuous to the observer. This technique leverages various approaches such as altering white spaces, font styles, or using invisible characters within the text to encode secret information without altering the visible structure of the cover text. For instance, recent advancements in steganography have introduced methods like embedding secret messages in multilingual text or utilizing font color changes in compressed texts to ensure higher security and imperceptibility (Askari et al., 2023). These methods offer robust solutions for protecting critical information by hiding data in plain sight, making it challenging for attackers to detect or extract the concealed messages (Shazzad-Ur-Rahman et al., 2021).

Traditional data security methods, such as encryption, are highly effective at transforming data into an unreadable format, safeguarding it from unauthorized access. However, encryption alone often draws attention due to the visible ciphertext it produces, which can attract attackers' scrutiny and encourage interception attempts. Moreover, advancements in computational power have rendered certain encryption methods vulnerable, allowing cybercriminals to potentially break the encryption with enough resources (Yadav et al., 2023). In contrast, steganography offers a complementary solution by embedding hidden messages within ordinary-looking data, such as text or images, thereby concealing the very existence of the communication. When combined with encryption, steganography creates an additional layer of security, making it far more difficult for attackers to even detect that sensitive data is being

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
transmitted. This dual-layer approach of encrypting the message first and then hiding it via steganography significantly reduces the likelihood of data being intercepted or decrypted (Shazzad-Ur-Rahman et al., 2023). Thus, while encryption protects the content, steganography protects the existence of the communication itself, providing a robust solution for modern data security challenges (Duluta et al., 2017).

The aim of reviewing the current techniques and tools used in text steganography for knowledge security is to provide a comprehensive understanding of how these methods enhance the confidentiality and protection of sensitive information. As steganography evolves alongside growing digital threats, it is essential to examine how various tools and techniques, including statistical methods, format-based approaches, and linguistic techniques, have been developed to embed hidden messages in text while minimizing the likelihood of detection. This review aims to identify the strengths, limitations, and potential improvements of existing steganography techniques, offering insights for researchers and practitioners looking to implement more robust solutions for securing knowledge in digital communication environments (Majeed et al., 2021). By exploring these methods, the review will contribute to advancing the field of knowledge security through text steganography, addressing current challenges and proposing future directions (Ghoul et al., 2023).

This article is organized into several key sections to provide a comprehensive review of the current techniques and tools used in text steganography for enhancing knowledge security. The first section focuses on reviewing existing techniques, including format-based, linguistic, and statistical methods for embedding hidden messages in text. It discusses the strengths and limitations of each approach in terms of imperceptibility and capacity (Majeed et al., 2021). The second section examines the tools that have been developed to implement these techniques, with an emphasis on their practical applications in real-world scenarios. Finally, the article highlights future research directions, suggesting innovations in areas such as artificial intelligence integration and multilingual text steganography to enhance security and capacity (Xiang et al., 2022). By addressing these areas, the article aims to provide researchers with a roadmap for advancing the field of text steganography.

Despite significant advancements in the field of text steganography, there remains a clear research gap in addressing the increasing complexity of detecting hidden messages and the challenges associated with scaling existing methods for larger datasets. While many studies have focused on image and multimedia steganography, text-based steganography has been relatively underexplored, particularly in non-English languages and in real-time applications (Jusoh et al., 2020). The primary research objective of this review is to provide a thorough examination of the current techniques and tools in text steganography, highlighting the potential improvements and innovations that can address existing limitations. Additionally, this review aims to explore future directions for developing more secure, robust, and scalable steganography solutions, particularly through the integration of artificial intelligence and machine learning methods (Gurunath et al., 2021). By doing so, the review seeks to guide future research efforts in enhancing knowledge security through more efficient text steganography techniques.

The article is structured as follows: First, the **Introduction** provides a background on the importance of knowledge security in the digital age and introduces the concept of text steganography as a viable solution for data protection. The **Literature Review** then delves into

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
the existing techniques and tools, categorizing them into format-based, statistical, and linguistic approaches, and evaluating their respective strengths and weaknesses (Majeed et al., 2021). The **Methodology** section outlines the research process for selecting and analyzing the techniques covered in the review. Following that, the **Results and Discussion** section presents a critical analysis of current tools and techniques, highlighting gaps in the research and potential areas for improvement. Finally, the **Conclusion and Future Directions** summarizes the findings and proposes key avenues for future research, particularly in the use of AI and multilingual steganography for enhanced security (Xiang et al., 2022).

Understanding Text Steganography

Steganography is the practice of concealing information within a medium, such as text, images, audio, or video, so that the very existence of the hidden message is undetectable to an observer. This differs from cryptography, which transforms data into an unreadable format known as ciphertext, rendering it incomprehensible without the decryption key. While cryptography protects the content of the message, it makes the fact that a message is being transmitted obvious, potentially drawing attention and scrutiny. In contrast, steganography aims to hide the presence of the message itself, thereby avoiding suspicion altogether (Hadipour & Afifi, 2020). The combination of cryptography and steganography can offer enhanced security, where cryptography secures the message content and steganography hides the existence of the message, thus adding multiple layers of protection (Islam et al., 2021). This complementary relationship between the two ensures that even if one layer is compromised, the data remains secure through the other method (Dutta & Chakraborty, 2020).

In text-based steganography, hidden data can be embedded through several techniques, including spacing manipulation, character alterations, and linguistic approaches. One common method involves manipulating invisible characters, such as spaces or special characters like Kashida in Arabic script, to hide information without altering the visible text structure. For instance, Kashida and Unicode space characters can be inserted between words, sentences, or paragraphs to encode secret bits of information without disturbing the overall appearance of the text (Taka, 2021). Another method employs linguistic techniques, where secret data is embedded by altering certain aspects of the text, such as choosing specific synonyms or adjusting sentence structure, enabling the concealment of information within the natural flow of the language (Xiang et al., 2020). These approaches ensure that the steganographic content remains imperceptible to human readers while maintaining a high level of security by making detection challenging for unintended recipients (Al-Azzam & Al-Garni, 2023).

Table 1: Differentiation between Types of Text Steganography

Steganography Type	Description	Advantages	Limitations
Open Space Method	Uses invisible characters, such as spaces or special Unicode characters, to embed secret information in the text without changing the visible appearance (Taka, 2021).	Simple to implement and effective at avoiding detection in plain text.	Limited embedding capacity and may be detected by systems analyzing text structure.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

Linguistic Steganography	Embeds hidden messages by manipulating linguistic elements like grammar, syntax, or word choice. The secret information is encoded through synonym substitution or sentence restructuring (Xiang et al., 2020).	Allows for more complex encoding schemes and can adapt to natural language features, increasing capacity.	Can be computationally intensive and is limited by the need for natural language processing.
Format-Based Method	Alters the format of the text (e.g., line spacing, font style, or paragraph indentation) to hide information in a way that is not detectable by normal reading (Majeed et al., 2021).	Can be highly effective when subtle formatting changes are difficult to detect by readers.	May be vulnerable to formatting changes that erase or disrupt hidden information.

Summarized in table 1, the types of text steganography can be categorized into three main methods: the open space method, linguistic steganography, and format-based methods. The **open-space method** uses invisible characters, such as spaces or special Unicode characters, to embed secret information within the text without altering the visible appearance of the document (Taka, 2021). This approach is simple to implement and avoids easy detection; however, it has limited capacity and can be detected by systems that analyze text structure.

Linguistic steganography involves embedding hidden messages by manipulating linguistic elements like grammar, syntax, or word choice. This technique encodes secret information by using synonym substitution or sentence restructuring (Xiang et al., 2020). While it allows for complex encoding schemes that can adapt to natural language, it can be computationally intensive and limited by the need for natural language processing.

The **format-based method** alters the format of the text—such as line spacing, font style, or paragraph indentation—to hide information in ways that are invisible during normal reading (Majeed et al., 2021). This method is highly effective when subtle formatting changes are difficult to detect by readers but is vulnerable to formatting changes that may disrupt or erase the hidden information.

Review of Text Steganography Techniques

Linguistic-based methods for text steganography primarily involve three key techniques: synonym substitution, text modification, and paraphrasing. Synonym substitution hides secret information by replacing certain words in the text with their synonyms. This method works by selecting words from a predefined synonym dictionary, ensuring that the new word does not change the overall meaning of the sentence, yet embeds secret data in the linguistic structure (Li et al., 2020). Another common approach is text modification, which alters the sentence structure without modifying the intended message. For example, changing word order or using different grammatical constructs can embed information, while still keeping the sentence readable (Serret et al., 2022). Finally, paraphrasing involves rewording the text using different

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
linguistic expressions to embed hidden information while maintaining the same meaning, often relying on external linguistic knowledge to generate more natural paraphrases (Lin et al., 2020). Each of these techniques offers a balance between imperceptibility and embedding capacity, making them effective for covert communication in digital environments.

Statistical and Structural Method

In text steganography, both **statistical** and **structural** methods are employed to manipulate the text for embedding hidden data, but they differ in their techniques and objectives. The **statistical method** relies on modifying the frequency of words or characters in a way that statistical patterns in the text reveal embedded information. This technique typically manipulates word or character frequency distributions to encode secret messages. For example, certain words or characters may appear more frequently to represent hidden data without visibly altering the text's meaning (Yang et al., 2021). Statistical methods are effective because they use existing linguistic features to hide information, making it difficult for traditional detection mechanisms to identify anomalies. However, their vulnerability lies in the potential for detection through deep statistical analysis or machine learning models trained to detect such deviations (Mustafa, 2020).

On the other hand, **structural methods** focus on altering the arrangement or structure of sentences, paragraphs, or characters. These methods may change sentence structure, such as rearranging word order or modifying grammatical constructs, without altering the meaning of the text. For instance, secret data can be embedded by shifting words or sentences within a paragraph while keeping the overall narrative intact. Another approach involves adjusting the spacing between words or lines in the text, using invisible characters like spaces or tabs to encode data (Aziz & Bukhelli, 2023). Structural methods are more challenging to detect since they preserve the natural flow of text and primarily manipulate the document's format. However, they can be vulnerable to detection if the text undergoes reformatting, which might disrupt the hidden information (Xiang et al., 2020).

Encoding Method

In text steganography, techniques involving Unicode characters and binary encoding provide advanced methods to conceal information. **Unicode-based steganography** makes use of invisible characters such as the Zero Width Joiner (ZWJ), Zero Width Non-Joiner (ZWNJ), and Kashida, which are non-visible formatting characters in scripts like Arabic, Urdu, or Farsi. These characters can be embedded in the text without altering its visible content, making it nearly impossible to detect by human readers (Alanazi et al., 2020). The Unicode method provides high embedding capacity and imperceptibility, as it exploits the script's natural formatting to hide secret data while maintaining text readability.

Binary encoding is another method where the secret message is converted into binary digits (bits) and embedded into the cover text. One common approach is to modify certain characters based on their binary values, such as in ASCII text where the binary representation of each character is adjusted slightly to encode the hidden data. A popular strategy in this area involves **binary digit mapping**, where each bit of the secret message corresponds to subtle modifications in the cover text, such as spacing or punctuation adjustments (Al-Azzam & Al-Garni, 2023). This technique provides a reliable method of embedding information without raising suspicion, especially when combined with encryption.

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

Comparative Analysis of Techniques

Techniques in text steganography involving **Unicode characters** and **binary encoding** offer distinct advantages, limitations, and effectiveness in different use cases. **Unicode-based steganography** leverages invisible characters such as Zero Width Joiners (ZWJ), Zero Width Non-Joiners (ZWNJ), and Kashida in Arabic script. These characters are embedded in the text without altering its visible structure, making it imperceptible to the naked eye. This method is highly effective in terms of capacity and security as it takes advantage of natural formatting, allowing large amounts of data to be hidden without detection. However, it is limited to languages that utilize such characters, and its effectiveness decreases when used in non-script-based texts (Alanazi et al., 2020). The method is especially effective in Arabic and Urdu texts but may be vulnerable to detection if text structure is manipulated or reprocessed.

On the other hand, **binary encoding** involves converting hidden messages into binary digits and embedding them into the text using subtle manipulations, such as modifying character codes in ASCII. The technique of **binary digit mapping** on ASCII characters is often used, where binary data is mapped to spaces or punctuation, making it difficult to detect the presence of hidden data. The primary advantage of this method is its compatibility with virtually all types of text, providing a universal approach to text steganography. However, the embedding capacity is relatively lower than Unicode-based techniques, and it is more susceptible to changes in formatting, which could disrupt the hidden data (Al-Azzam & Al-Garni, 2023).

Both methods are highly effective in concealing data but serve different purposes depending on the text type and the language in use. Unicode steganography is preferred in contexts where high capacity and security are needed in script-based languages, while binary encoding provides a more flexible solution for standard text formats.

Tools For Text Steganography

Popular tools for text steganography, such as **Steghide** and **OpenStego**, offer various techniques for embedding secret messages into cover texts. **Steghide** is widely used because it supports not only text but also image and audio formats for hiding data. It uses algorithms that modify the least significant bits (LSB) of files to embed secret information, making it difficult to detect without the proper decryption key (Islam et al., 2020). **OpenStego** is another popular tool that provides both encryption and steganography, ensuring that hidden messages are protected by a password. OpenStego primarily focuses on hiding data in images and text, offering a user-friendly interface and enhanced security features such as watermarking (Hidayasari et al., 2020).

Other tools like **S-Tools** and **SilentEye** also offer powerful steganographic capabilities. **S-Tools** uses a drag-and-drop interface to hide text files inside images or audio files, providing a straightforward method for embedding data. Meanwhile, **SilentEye** allows users to conceal text within both image and audio formats, offering a cross-platform solution for steganography with additional encryption features (Thabit et al., 2022).

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

Performance Evaluation

Table 2 Perfomance Evaluation

Tool	Strengths	Weaknesses	User-Friendliness
Steghide	Supports multiple formats (text, image, audio) and uses LSB techniques for high security.	Limited advanced features, and the process may take longer for larger files.	Moderate, requires some technical understanding but is intuitive once learned.
OpenStego	Provides encryption and steganography, simple user interface, and watermarking for enhanced security.	Primarily designed for images, lacks support for audio steganography.	High, designed for ease of use with a simple interface for beginners.
S-Tools	User-friendly with drag-and-drop functionality, supports both image and audio formats.	Limited to basic steganography techniques, lacks advanced encryption features.	Very High, straightforward drag-and-drop functionality for non-technical users.
SilentEye	Cross-platform, supports image and audio formats, includes encryption options.	Lacks advanced user customization options and support for a wide range of file formats.	Moderate, cross-platform compatibility but lacks detailed customization features.

In evaluating the strengths, weaknesses, and user-friendliness of popular text steganography tools such as **Steghide**, **OpenStego**, **S-Tools**, and **SilentEye**, it becomes clear that each tool has its unique advantages and limitations. **Steghide** is notable for its support of multiple formats, including text, image, and audio, and its use of least significant bit (LSB) techniques provides high security. However, its advanced features are limited, and handling larger files can be time-consuming. **OpenStego** offers both encryption and steganography, combined with a simple user interface and watermarking for enhanced security, but it is primarily designed for images, limiting its versatility with audio files (Islam et al., 2020).

S-**Tools** excels in its user-friendly, drag-and-drop functionality, making it accessible for non-technical users. However, it is limited to basic steganography techniques and lacks advanced encryption features. Finally, **SilentEye** offers cross-platform compatibility and supports both image and audio formats, with encryption options, but lacks detailed customization features and support for a wider range of file formats (Thabit et al., 2022). These tools provide a mix of flexibility and simplicity, catering to both novice and experienced users based on their needs.

When comparing tools for text steganography in terms of **usability**, **security level**, and **techniques used**, a few popular options such as **Steghide**, **OpenStego**, **S-Tools**, and **SilentEye** stand out. **Steghide** is known for supporting multiple formats (text, image, and audio), but it requires some technical knowledge to use effectively. It employs least significant bit (LSB) techniques, making it highly secure, especially when combined with encryption. However, it has limitations in processing large files (Islam et al., 2020). **OpenStego** is another widely used tool, primarily for images and text, which is praised for its high usability due to its simple

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
interface and integration of both encryption and steganography. Its security level is boosted by password protection, but it lacks support for formats other than images (Hidayasari et al., 2020).

Tools and **SilentEye** offer a balance between ease of use and basic security features. **S-Tools** features a user-friendly, drag-and-drop interface and supports both image and audio formats, but lacks advanced encryption capabilities, which may limit its effectiveness in high-security scenarios. **SilentEye**, while cross-platform, also supports image and audio formats and includes encryption options, though it offers limited advanced customization options (Thabit et al., 2022). These tools are effective in specific contexts, depending on the user's technical expertise and the required level of data protection.

Challenges And Limitations

Text steganography, while effective at concealing information, has certain weaknesses that make it susceptible to detection through **steganalysis**. One major vulnerability is its reliance on detectable patterns or statistical anomalies within the cover text. Steganalysis tools, particularly those based on machine learning, are becoming increasingly adept at identifying these subtle changes. Techniques such as deep neural networks (DNNs) have been developed to detect the presence of hidden data in text by analyzing deviations in word or character frequency distributions. For instance, methods like unsupervised deep learning can effectively detect embedded information by focusing on global and local text structures (Xu, 2020).

Another significant weakness is that steganography often leaves statistical footprints that are detectable through advanced algorithms. Recent steganalysis models using multi-stage transfer learning have shown improved accuracy in detecting hidden messages, even in real-time applications. This method has proven effective at identifying hidden information embedded in normal texts by adversaries (Peng et al., 2021). As steganography advances, these detection techniques continue to evolve, narrowing the gap between the covert embedding of information and the ease with which it can be discovered. While text steganography provides a covert method of communication, its effectiveness can be undermined by advanced steganalysis tools capable of detecting hidden patterns and irregularities in text. The development of robust detection models poses a significant challenge to the imperceptibility of steganographic techniques.

One of the core challenges in text steganography is embedding a sufficient amount of data while maintaining the coherence and readability of the cover text. Text steganography relies on subtle modifications to the text, such as word substitution, syntax alteration, or the insertion of invisible characters, to hide information. However, these changes must be minimal and carefully applied to avoid distorting the meaning or flow of the text, as excessive alterations can make the text appear unnatural and raise suspicion. For instance, linguistic techniques such as synonym substitution or sentence rephrasing must ensure that the substituted words or structures align with the context and intended meaning of the original text, making it difficult to embed large volumes of data without sacrificing coherence (Yang et al., 2021). Furthermore, the limited capacity of text compared to other media like images or audio presents an additional challenge, as it restricts the amount of data that can be hidden without noticeably impacting the text. Advanced methods like character-level or word-level encoding offer ways to increase embedding capacity, but they come at the cost of potentially disrupting the natural language patterns and making the text more detectable by steganalysis techniques (Xu, 2020). Therefore,

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
achieving a balance between data volume and maintaining the semantic integrity of the text remains one of the most difficult tasks in text steganography.

Achieving a balance between secure embedding and maintaining the naturalness of text in steganography is a persistent challenge. Secure embedding refers to the ability to hide data in a way that it is undetectable by steganalysis methods, while maintaining naturalness involves preserving the readability and coherence of the text. Overly aggressive embedding methods, such as excessive word substitutions or complex linguistic alterations, can compromise the flow and grammatical correctness of the text, making it appear unnatural and easily detectable. The challenge lies in embedding sufficient information without altering the text in ways that raise suspicion or disrupt its intended meaning (Alotaibi et al., 2021). Techniques such as synonym substitution and sentence restructuring offer a solution by embedding data in linguistically acceptable ways, but they are constrained by the availability of appropriate substitutions that fit the context. Additionally, methods like format-based or invisible character insertion can maintain the visual structure of the text while embedding data, though these may still be susceptible to detection through advanced steganalysis techniques (Wu & Wang, 2021). To ensure both security and naturalness, researchers are increasingly focusing on hybrid methods that combine linguistic knowledge with machine learning models to preserve the integrity of the text while enhancing security (Zhao et al., 2022).

Future Trends And Research Directions

Emerging technologies such as **artificial intelligence (AI)** and **machine learning (ML)** are increasingly being explored to enhance text steganography, particularly in improving both the security of embedding and maintaining the naturalness of the text. AI-driven models, particularly **deep learning techniques**, have shown promise in generating more coherent and natural-looking steganographic text by mimicking human writing styles and understanding linguistic nuances. These models can analyze text patterns and learn to embed data in ways that are difficult to detect, such as by generating plausible linguistic variations while hiding secret information (Zhao et al., 2022). Moreover, **natural language processing (NLP)** technologies are being leveraged to enhance synonym selection and sentence rephrasing, allowing for more efficient data embedding without disrupting the flow of the text. Another emerging method involves the integration of **generative adversarial networks (GANs)**, which are used to create more sophisticated steganographic techniques by training networks to generate cover texts that are indistinguishable from natural text. GANs can create steganographic content that not only embeds data securely but also maintains a high degree of imperceptibility, thus reducing the chances of detection by advanced steganalysis techniques (Wu & Wang, 2021). Additionally, **quantum steganography** is an area of growing interest. This technology promises to revolutionize steganography by leveraging quantum computing to secure communication channels at an unprecedented level, offering a new frontier for data hiding that could be resistant to classical steganalysis techniques (Yang et al., 2021).

The combination of steganography with **cryptography** or **blockchain technology** holds significant potential for enhancing data security by creating multilayered protection schemes. **Steganography** conceals the existence of a message, while **cryptography** ensures that even if the hidden message is detected, it remains unreadable without the decryption key. By embedding an encrypted message within cover text, steganography adds an additional layer of obfuscation, making it harder for attackers to detect and decipher the hidden information. This dual approach is particularly useful in high-security scenarios where sensitive data needs to be

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
transmitted covertly, as it provides both encryption and concealment to thwart detection and interception (Dutta & Chakraborty, 2020). In addition to cryptography, **blockchain technology** has emerged as another innovative way to enhance the security of steganographic systems. Blockchain's decentralized and immutable nature ensures that data cannot be tampered with once it is embedded. Combining steganography with blockchain can provide robust protection against attacks by ensuring that any hidden information remains secure and verifiable throughout its transmission. In this context, **steganography** could be used to hide data within blockchain transactions, further obfuscating its existence and providing additional layers of privacy for users. This method not only ensures the protection of sensitive data but also leverages the traceability and transparency features of blockchain to create a more secure environment for digital communication (Swarna & Saravanan, 2021).

The integration of **artificial intelligence (AI)** and **machine learning (ML)** into steganography is transforming the development of more sophisticated and secure data-hiding techniques. AI, particularly through the use of **deep learning**, enables the creation of more natural and imperceptible steganographic content by training models to understand and replicate human language patterns, making detection by steganalysis tools significantly more difficult. For instance, neural networks can be trained to generate cover text that seamlessly hides secret messages, while maintaining a natural flow and coherence that mimics real human communication. This enhances both the capacity and security of steganography, as the data becomes more difficult to detect or differentiate from ordinary text (Zhao et al., 2022). **Machine learning** models, such as **generative adversarial networks (GANs)**, are being employed to improve steganographic techniques by generating more adaptive and robust hiding methods. GANs, consisting of a generator and a discriminator, are used to create and verify hidden data, where the generator produces steganographic text, and the discriminator evaluates its imperceptibility. Through iterative training, the models continuously improve their ability to hide information in a way that makes it indistinguishable from regular text, while also resisting detection by advanced steganalysis techniques (Liu et al., 2021). Additionally, **reinforcement learning** has been explored to enhance the adaptability of steganographic systems, allowing them to learn from various embedding environments and optimize the security of hidden data over time (Chen et al., 2020).

Conclusion

In this review, several prominent techniques and tools for text steganography were examined, each offering unique strengths and limitations for secure data hiding. Techniques such as **synonym substitution**, **sentence restructuring**, and **paraphrasing** are linguistic-based methods that modify the textual content while preserving its semantic integrity to embed hidden information. These techniques allow for relatively high levels of naturalness but are often constrained by the limited capacity to embed large amounts of data (Li et al., 2020). **Unicode manipulation** and **binary encoding** offer alternative approaches by embedding hidden data in invisible characters or encoding it within character sets, which are difficult to detect but can be vulnerable to formatting changes (Al-Azzam & Al-Garni, 2023).

Several tools were also discussed, including **Steghide**, **OpenStego**, **S-Tools**, and **SilentEye**. These tools utilize various methods such as least significant bit (LSB) embedding, encryption, and format-based techniques to hide data within text and multimedia files. **Steghide** and **OpenStego** stand out for their encryption capabilities and user-friendly interfaces, while **S-Tools** and **SilentEye** offer cross-platform compatibility with basic steganographic features

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

(Islam et al., 2020). Emerging methods involving **AI** and **machine learning**, such as **deep learning** and **generative adversarial networks (GANs)**, were highlighted as cutting-edge advancements that enhance the imperceptibility and security of text steganography by generating more coherent and natural-looking steganographic content (Zhao et al., 2022). These tools and techniques collectively push the boundaries of secure data hiding in the digital age.

Text steganography plays a critical role in enhancing data protection by concealing the existence of sensitive information within ordinary-looking text, making it an essential tool for secure communication in the digital age. Unlike cryptography, which only encrypts data but makes it obvious that a message exists, steganography provides an additional layer of security by hiding the fact that communication is occurring at all. This makes it highly effective in situations where secrecy is paramount, such as transmitting confidential information across public or unsecured networks. Techniques like synonym substitution, sentence restructuring, and invisible character manipulation have proven useful in embedding hidden messages while maintaining the naturalness of the text, making detection through regular means more difficult (Zhao et al., 2022). Moreover, the integration of AI and machine learning has further enhanced the security and imperceptibility of text steganography by generating more natural and adaptive steganographic texts that can withstand sophisticated detection techniques (Liu et al., 2021). By combining steganography with other security methods like cryptography or blockchain, organizations can achieve a higher level of data protection, ensuring that both the content and the transmission of sensitive information remain hidden from unauthorized parties (Dutta & Chakraborty, 2020).

The ongoing exploration of new methods and tools for improving steganography is essential as both cyber threats and detection techniques become increasingly sophisticated. To keep pace with these developments, researchers must continue to investigate cutting-edge technologies like **artificial intelligence (AI)**, **machine learning (ML)**, and **quantum computing**. These technologies offer the potential to enhance the imperceptibility and robustness of steganographic techniques, making them more resilient against advanced steganalysis methods. For example, **deep learning models** can be used to generate more natural-looking cover text, while **generative adversarial networks (GANs)** improve the ability to embed data without detectable anomalies (Zhao et al., 2022). The integration of **cryptography** and **blockchain** with steganography can also provide an additional layer of security, ensuring that even if hidden messages are discovered, they remain protected through encryption (Swarna & Saravanan, 2021).

By pushing the boundaries of current technology, the steganography field can continue to evolve and provide more sophisticated solutions for secure data communication. Future research should focus not only on developing new techniques but also on refining existing tools to address vulnerabilities and increase data-hiding capacity. This forward-looking approach is crucial for maintaining privacy and security in an increasingly interconnected world (Liu et al., 2021).

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

Acknowledgements

We extend our heartfelt gratitude to Universiti Teknologi MARA (UiTM) Cawangan Kedah for the unwavering support and provision of resources. We also would like to sincerely thank the anonymous reviewers for their constructive feedback and insightful comments, which significantly contributed to the refinement and improvement of the article.

References

Al-Azzam, S., & Al-Garni, F. A. (2023). The use of binary digit mapping on ASCII characters to create a high-capacity, undetectable text steganography. *Journal of Advanced Sciences and Engineering Technologies*. https://doi.org/10.32441/jaset.05.02.05

Alanazi, N., Khan, E., & Gutub, A. (2020). Functionality-improved Arabic text steganography based on Unicode features. *Arabian Journal for Science and Engineering, 45*, 11037-11050. https://doi.org/10.1007/s13369-020-04917-5

Askari, M., Mahmood, A., & Iqbal, Z. (2023). A novel font color and compression text steganography technique. *2023 International Conference on Communication, Computing and Digital Systems (C-CODE)*. https://doi.org/10.1109/C-CODE58145.2023.10139867

Aziz, B., & Bukhelli, A. (2023). Detecting the manipulation of text structure in text steganography using machine learning. *International Journal of Information and Education Technology*. https://doi.org/10.5220/0012260900003584

Chen, S., Liu, Y., Wu, Z., & Wang, Z. (2020). Text steganography using reinforcement learning and deep neural networks. *IEEE Access, 8*, 76367-76376. https://doi.org/10.1109/ACCESS.2020.2990644

Duluta, A., Mocanu, S., Pietraru, R., Merezeanu, D., & Saru, D. (2017). Secure communication method based on encryption and steganography. *2017 21st International Conference on Control Systems and Computer Science (CSCS)*. https://doi.org/10.1109/CSCS.2017.70

Dutta, P., & Chakraborty, S. (2020). Image-based steganography in cryptography implementing different encryption-decryption algorithms. *2020 CSEIT Conference*. https://doi.org/10.32628/cseit2063191

Ghoul, S., Sulaiman, R., & Shukur, Z. (2023). A review on security techniques in image steganography. *International Journal of Advanced Computer Science and Applications*. https://doi.org/10.14569/ijacsa.2023.0140640

Hadipour, A., & Afifi, R. (2020). Advantages and disadvantages of using cryptography in steganography. *2020 17th International ISC Conference on Information Security and Cryptology (ISCISC)*. https://doi.org/10.1109/ISCISC51277.2020.9261921

Hidayasari, N., Riadi, I., & Prayudi, Y. (2020). Steganalysis using Yedrodj-net's convolutional neural networks (CNN) method on steganography tools. *Proceeding International Conference on Science and Engineering*. https://doi.org/10.14421/ICSE.V3.499

Islam, M. A., Riad, M. S., & Pias, T. S. (2020). Performance analysis of steganography tools. *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*. https://doi.org/10.1109/ICAICT51780.2020.9333473

Jevtić, N., & Alhudaidi, I. (2023). The importance of information security for organizations. *Serbian Journal of Engineering Management*. https://doi.org/10.5937/SJEM2301001J

Li, F., Tang, H., Liu, L., Li, B., Feng, Y., & Chen, W. (2020). A text information hiding method based on sentiment word substitution. *International Conference on Artificial Intelligence and Security*. https://doi.org/10.1007/978-981-15-3753-0_72

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022
Lin, Z., Li, Z., Ding, N., Zheng, H., Shen, Y., & Zhao, C. (2020). Integrating linguistic knowledge to sentence paraphrase generation. *Proceedings of the AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v34i05.6354

Liu, Y., Song, H., Wu, Z., & Zhou, X. (2021). Text steganography using generative adversarial networks: A review. *IEEE Access, 9*, 150267-150278. https://doi.org/10.1109/ACCESS.2021.3117885

Majeed, M., Sulaiman, R., Shukur, Z., & Hasan, M. Z. (2021). A review on text steganography techniques. *Mathematics*. https://doi.org/10.3390/math9212829

Mustafa, N. A. (2020). Text hiding in text using invisible character. *International Journal of Electrical and Computer Engineering, 10*(4), 3550-3557. https://doi.org/10.11591/ijece.v10i4.pp3550-3557

Peng, W., Zhang, J., Xue, Y., & Yang, Z. (2021). Real-time text steganalysis based on multi-stage transfer learning. *IEEE Signal Processing Letters, 28*, 1510-1514. https://doi.org/10.1109/LSP.2021.3097241

Romansky, R., & Noninska, I. (2020). Challenges of the digital age for privacy and personal data protection. *Mathematical Biosciences and Engineering, 17*(5), 5288-5303. https://doi.org/10.3934/mbe.2020278

Serret, E., Lesueur, A., & Gabillon, A. (2022). Linguistic steganography for messaging applications. *Proceedings of the 18th International Conference on Security and Cryptography (SECRYPT)*. https://doi.org/10.5220/0010899300003120

Shazzad-Ur-Rahman, M., Hosen Ornob, M. M., Singha, A., Kaiser, M. S., & Akhter, N. (2021). An effective text steganographic scheme based on multilingual approach for secure data communication. *2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. https://doi.org/10.1109/ICIEVicIVPR52578.2021.9564231

Shazzad-Ur-Rahman, M., Kaiser, M. S., Alam, M. B., & Nova, S. N. (2023). A data hiding technique combining steganography and cryptography for secured communication. *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*. https://doi.org/10.1109/ICICT4SD59951.2023.10303563

Swarna, A., & Saravanan, S. (2021). Secure data hiding using blockchain-based steganography technique. *2021 International Conference on Intelligent Computing and Control Systems (ICICCS)*. https://doi.org/10.1109/ICICCS51141.2021.9441955

Taka, F. R. S. (2021). Text steganography based on Noorani and Darkness. *Journal of Information Hiding and Multimedia Signal Processing, 12*(1), 127-139. https://consensus.app/papers/text-steganography-based-noorani-darknesstaka/9a2ef98aae4a5b0cba998f04ab4b8edd/?utm_source=chatgpt

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*. https://doi.org/10.1109/access.2022.3182712

Wu, Z., & Wang, Z. (2021). Text steganography based on word frequency distribution. *Multimedia Tools and Applications, 80*(8), 11707-11727. https://doi.org/10.1007/s11042-021-11192-6

Xiang, L., Wang, R., Yang, Z., & Liu, Y. (2022). Generative linguistic steganography: A comprehensive review. *KSII Transactions on Internet and Information Systems, 16*(3), 986-1005. https://doi.org/10.3837/tiis.2022.03.013

Xiang, L., Yang, S., Liu, Y., Li, Q., & Zhu, C. (2020). Novel linguistic steganography based on character-level text generation. *Mathematics*. https://doi.org/10.3390/math8091558

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 10 Issue 38 (March 2025) PP. 325-339
DOI: 10.35631/JISTM.1038022

Yang, B., Peng, W., Xue, Y., & Zhong, P. (2021). A generation-based text steganography by maintaining consistency of probability distribution. *KSII Transactions on Internet and Information Systems, 15*(11), 4184-4202. https://doi.org/10.3837/tiis.2021.11.017

Yang, L., Hu, Y., Zhang, Z., & Yang, J. (2021). A review on linguistic steganography and steganalysis. *Multimedia Tools and Applications, 80*(3), 3687-3722. https://doi.org/10.1007/s11042-021-10798-1

Zhao, Q., Zhou, J., Wu, Z., & Zhang, J. (2022). Secure text steganography using deep neural networks and statistical metrics. *IEEE Transactions on Information Forensics and Security, 17*, 1427-1442. https://doi.org/10.1109/TIFS.2022.3148321