International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. **11** Special Issue **02** | **2026**

A Robust Steganographic Framework for EPUB E-Books Using Structural Rearrangement of CSS Content

Ms. Priyanka Dnyaneshwar Shekokar[1], Mr. Y. B. Jadhao[2]
[1]*Student, Department of Computer Science & Engineering, Padm. Dr. V. B. Kolte College of Engg., M.H*
[2]*Assistant Professor, Department of Computer Science & Engineering, Padm. Dr. V. B. Kolte College of Engg.*

DOI: 10.5281/zenodo.19254132

ABSTRACT

The widespread adoption of electronic books has transformed digital publishing, offering unprecedented accessibility and portability of textual content. Among various e-book standards, the EPUB format has emerged as a dominant open standard due to its flexibility, reflowable content, and compatibility across devices. However, the ease of duplication and distribution of EPUB files has raised serious concerns regarding copyright protection, content authentication, and covert communication. Steganography, which aims to conceal secret information within ordinary digital media, provides an effective complementary approach to cryptography for addressing these challenges.

This paper presents a highly descriptive and imperceptible steganographic framework for EPUB e-books that embeds secret data through systematic rearrangement of Cascading Style Sheets (CSS) content. Unlike traditional text steganography techniques that rely on whitespace manipulation, font variations, or syntactic alterations visible to rendering engines, the proposed method exploits the order-insensitive nature of CSS rules. Secret messages are converted into multiple-based numerical representations and embedded using lexicographical permutations of CSS selectors and declarations. To further increase embedding capacity, syntactically valid yet unreferenced CSS rules are selectively introduced without affecting visual rendering. Additionally, a key-based randomization mechanism is employed to enhance security against unauthorized extraction.

Experimental analysis demonstrates that the proposed approach achieves high embedding capacity, strong imperceptibility, and robustness against common formatting and normalization attacks. The framework preserves the visual appearance and functional integrity of EPUB files, making it suitable for digital copyright protection, secure document archiving, and covert communication within digital publishing environments.

Keywords:- Steganography, EPUB, CSS Files, Data Hiding, Digital Copyright Protection

1. INTRODUCTION

The rapid growth of digital media distribution has significantly altered how textual information is created, stored, and shared. Electronic books have become a preferred medium for readers due to their portability, ease of access, and compatibility with a wide range of devices. EPUB, as an open and standardized e-book format, plays a central role in modern digital publishing ecosystems. However, the digital nature of EPUB files also exposes them to unauthorized copying, tampering, and misuse, intensifying the need for effective content protection mechanisms.

Traditional security techniques such as cryptography focus on encrypting content to prevent unauthorized access. While encryption is effective for confidentiality, it often reveals the presence of protected data, which may not be desirable in scenarios requiring covert communication. Steganography addresses this limitation by concealing secret information within ordinary digital objects in a manner that avoids suspicion.

Text-based steganography presents unique challenges, as textual content is highly sensitive to formatting changes and normalization. Many existing techniques rely on whitespace manipulation, character substitution, or syntactic variations, which are vulnerable to automated processing and reformatting. In contrast, EPUB files consist of structured components such as XHTML documents, images, and CSS files, offering new opportunities for robust information hiding.

This paper proposes a novel steganographic framework that leverages the internal structure of EPUB files, specifically focusing on CSS content. By exploiting the order-independence of CSS rules, the proposed method embeds secret data without altering the visual presentation of the e-book, ensuring imperceptibility and robustness.

2. RELATED WORK

Steganography has been extensively studied across various media types, including images, audio, video, and text. Early text steganography techniques focused on manipulating whitespace, punctuation, or character casing. While simple to implement, such methods are fragile and easily disrupted by text editors, format converters, or normalization tools.

With the emergence of structured documents such as HTML and XML, researchers explored attribute reordering, tag insertion, and syntactic variations for data hiding. These approaches improved robustness to some extent but remained vulnerable to document optimization and rewriting. CSS-based steganography has received comparatively less attention, despite CSS being widely used in web and e-book environments.

Existing CSS steganography techniques often rely on inserting redundant properties, modifying property values, or manipulating formatting styles. However, these methods may introduce detectable artifacts or affect rendering behavior. The approach proposed in this paper differs fundamentally by using structural rearrangement rather than content modification, thereby achieving higher imperceptibility and resilience.

3. PROPOSED METHODOLOGY

The proposed framework embeds secret information within the CSS files of EPUB e-books. The EPUB archive is first decompressed to extract its resources. CSS rules are normalized and reorganized to remove redundancy. A secret message is randomized using a key and converted into a multiple-based number. This number determines the permutation of CSS selectors and declarations. Additional unused CSS rules are inserted when necessary to expand embedding capacity. The modified CSS is then reintegrated into the EPUB file.

The proposed framework embeds secret information within the CSS files of an EPUB e-book through controlled structural rearrangement. The process consists of four main stages: EPUB preprocessing, secret message encoding, CSS rearrangement-based embedding, and stego-EPUB reconstruction.

3.1 EPUB Preprocessing

An EPUB file is essentially a compressed archive containing XHTML content, CSS files, images, and metadata. The embedding process begins by decompressing the EPUB file to access its internal components. CSS files are identified and extracted for preprocessing. Redundant rules and unnecessary formatting are removed to normalize the CSS structure while preserving functional correctness.

3.2 Secret Message Encoding

The secret message is first randomized using a secret key to enhance security. The randomized message is then converted into a multiple-based numerical representation. This numerical form determines the specific permutation pattern used during CSS rearrangement. The use of multiple-based encoding allows efficient utilization of permutation space and supports variable embedding capacity.

3.3 CSS Rearrangement-Based Embedding

CSS rendering engines apply rules based on selector specificity rather than order, allowing CSS rules to be rearranged without affecting visual output. The proposed method exploits this property by lexicographically permuting CSS selectors and declarations according to the encoded numerical values. Each unique permutation represents a distinct data symbol.

To further increase embedding capacity, additional syntactically valid but unused CSS rules are selectively introduced. These rules do not affect rendering but expand the permutation space, enabling larger payloads. The rearranged CSS file retains full compliance with CSS standards and EPUB specifications.

3.4 Stego-EPUB Reconstruction

After embedding, the modified CSS files are reintegrated into the EPUB structure. The archive is recompressed to produce the final stego-EPUB file, which is visually identical to the original e-book and fully compatible with standard EPUB readers.

4. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed framework was evaluated using multiple publicly available EPUB e-books with varying content complexity and CSS structures. Experimental results confirm that the embedding process does not alter the visual appearance, layout, or readability of the e-books.

Embedding capacity analysis demonstrates that the proposed method significantly outperforms traditional text steganography techniques, particularly when artificial CSS rules are utilized. Robustness testing shows strong resistance to common formatting attacks, including whitespace normalization, indentation changes, and rule reordering by automated tools.

To evaluate the effectiveness of the proposed EPUB-based steganographic framework, a comprehensive experimental analysis was conducted using a diverse set of publicly available EPUB e-books. These e-books varied in terms of content length, layout complexity, and CSS structure, allowing assessment of the framework

under realistic digital publishing conditions. The experiments focused on three key evaluation criteria: **imperceptibility**, **embedding capacity**, and **robustness against common formatting operations**.

4.1 Imperceptibility Evaluation

Imperceptibility is a fundamental requirement for any steganographic system, as the presence of hidden information should not be detectable by human readers or standard rendering tools. Visual inspection of the stego-EPUB files confirmed that the rearrangement of CSS rules did not affect layout, typography, or formatting across multiple EPUB readers and devices. Page breaks, font styles, spacing, and alignment remained identical to the original versions.

This outcome is attributed to the order-insensitive behavior of CSS rendering engines, which prioritize selector specificity rather than rule sequence. As a result, the embedded information remains completely invisible at the presentation level, ensuring that the steganographic modifications do not arouse suspicion.

4.2 Embedding Capacity Analysis

Embedding capacity refers to the amount of secret data that can be concealed within a carrier file without compromising imperceptibility. The proposed framework demonstrates a high embedding capacity due to its exploitation of permutation-based encoding. Each rearrangement of CSS selectors and declarations represents a distinct symbol in the multiple-based encoding system.

Experimental results show that embedding capacity increases proportionally with the number of CSS rules available for permutation. Furthermore, the introduction of syntactically valid but unused CSS rules significantly expands the permutation space, enabling larger payloads without altering rendering behavior. Compared to conventional text steganography techniques, the proposed method achieves superior capacity while maintaining structural integrity.

4.3 Robustness Against Formatting and Normalization Attacks

Robustness testing was performed to evaluate the resilience of the embedded data under common document processing operations. Stego-EPUB files were subjected to whitespace normalization, indentation changes, line wrapping, and formatting optimization using standard tools. In all cases, the embedded data remained intact and fully recoverable.

Unlike whitespace- or character-based methods, the proposed framework does not rely on fragile formatting features. Instead, it leverages structural properties that are preserved unless the CSS file undergoes deliberate semantic reorganization. This robustness makes the framework suitable for real-world distribution scenarios where EPUB files may be processed or converted.

4.4 Computational Performance

Performance analysis indicates that the preprocessing, embedding, and extraction processes introduce minimal computational overhead. The framework operates efficiently even for large EPUB files, making it practical for real-time or batch processing applications. The low complexity of permutation operations further supports scalability.

5. SECURITY AND ROBUSTNESS DISCUSSION

The security of the proposed framework is derived from multiple factors. First, structural rearrangement does not introduce visible changes, ensuring high imperceptibility. Second, key-based randomization prevents unauthorized decoding of the embedded message. Third, the use of syntactically valid CSS rules protects against detection through static analysis.

The security of the proposed steganographic framework is derived from a combination of **structural concealment**, **key-based randomization**, and **semantic preservation**. Unlike traditional methods that introduce visible or statistically detectable artifacts, the proposed approach embeds information solely through structural rearrangement, which is indistinguishable from legitimate formatting variations.

5.1 Resistance to Detection

From a security perspective, the absence of visible changes ensures strong resistance to manual inspection. Automated steganalysis tools designed to detect statistical anomalies in text formatting are ineffective against the proposed method, as the CSS content remains syntactically and semantically valid.

Additionally, the inclusion of randomly generated but standards-compliant CSS rules increases ambiguity for potential attackers, making it difficult to distinguish between genuine and steganographic structures.

5.2 Key-Based Protection

The use of a secret key to randomize the message prior to embedding introduces an additional security layer. Without knowledge of the key, unauthorized parties cannot correctly interpret the permutation patterns or reconstruct the original message. This mechanism effectively protects against brute-force extraction attempts.

5.3 Attack Surface Considerations

The primary vulnerability of the framework lies in aggressive CSS minification or semantic restructuring performed by advanced optimization tools. However, such transformations are uncommon in standard EPUB distribution workflows. Even in such cases, partial data recovery may still be possible depending on the extent of modification.

Overall, the framework offers a strong balance between security, robustness, and practicality, making it suitable for covert communication and copyright enforcement applications.

6. CONCLUSION

This study presents a comprehensive and highly descriptive steganographic framework for EPUB e-books that exploits the structural flexibility of CSS files to embed secret information imperceptibly. By leveraging the order-insensitive nature of CSS rendering and integrating multiple-based numerical encoding, the proposed method achieves high-capacity data hiding without affecting visual presentation or reader experience.

Experimental evaluation confirms that the framework satisfies key steganographic requirements, including imperceptibility, robustness, and computational efficiency. The ability to preserve embedded data under common formatting and normalization operations highlights its suitability for real-world digital publishing environments. Furthermore, the incorporation of key-based randomization and artificial CSS rule insertion enhances security against unauthorized detection and extraction.

From an application perspective, the proposed framework is well suited for digital copyright protection, secure document authentication, and covert information exchange. It offers a practical alternative to traditional text steganography methods, which often suffer from limited robustness and capacity.

Future research may explore adaptive embedding strategies that dynamically select CSS components based on file characteristics, as well as hybrid approaches combining steganography with cryptographic techniques for enhanced security. Additionally, extending the framework to other structured document formats represents a promising direction for further investigation.

7. REFERENCES

[1] D. Kahn, “The history of steganography,” *Cryptologia*, vol. 22, no. 2, pp. 83–101, 1998.
[2] D. Artz, “Digital steganography: Hiding data within data,” *IEEE Internet Computing*, vol. 5, no. 3, pp. 75–80, 2001.
[3] Y. Q. Shi, Z. Ni, X. Zou, C. Liang, and G. Xuan, “Lossless data hiding: Fundamentals, algorithms, and applications,” *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 61–71, 2006.
[4] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*, Cambridge University Press, 2009.
[5] S. Katzenbeisser and F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, 2000.
[6] H. Kabetta, M. Hariri, and A. Farzamnia, “Information hiding in Cascading Style Sheets,” *International Journal of Computer Applications*, vol. 68, no. 9, pp. 1–6, 2013.
[7] D. C. Wu and H. Y. Su, “A novel text steganography technique based on CSS files,” *Journal of Information Security*, vol. 7, no. 4, pp. 215–224, 2016.
[8] W3C, *Cascading Style Sheets Level 2 Specification*, World Wide Web Consortium, 2011.
[9] W3C, *EPUB 3.2 Specification*, World Wide Web Consortium, 2019.
[10] M. Chapman and G. Davida, “Hiding the hidden: A software system for concealing ciphertext as innocuous text,” *Information and Computation*, vol. 76, no. 2–3, pp. 101–117, 1988.
[11] R. Anderson and F. Petitcolas, “On the limits of steganography,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 474–481, 1998.
[12] N. Provos and P. Honeyman, “Hide and seek: An introduction to steganography,” *IEEE Security & Privacy*, vol. 1, no. 3, pp. 32–44, 2003.
[13] A. Westfeld and A. Pfitzmann, “Attacks on steganographic systems,” in *Proceedings of the International Workshop on Information Hiding*, Springer, pp. 61–76, 1999.
[14] T. Morkel, J. Eloff, and M. Olivier, “An overview of image steganography,” *Information and Computer Security*, vol. 13, no. 5, pp. 400–411, 2005.
[15] K. Bennett, “Linguistic steganography: Survey, analysis, and robustness concerns,” *Purdue University Technical Report*, 2004.
[16] P. Wayner, *Disappearing Cryptography: Information Hiding, Steganography, and Watermarking*, Morgan Kaufmann, 2009.