

THE PERFORMANCE OF TEXT STEGANOGRAPHY BASED ON SYMBOL

Muhammad Syahmi Naquib Saumi and Muhamad Arif Hashim
College of Computing, Informatics and Mathematics,
Universiti Teknologi MARA, Perlis Branch
2020828334@student.uitm.edu.my and muhamadarif487@uitm.edu.my

ABSTRACT - Data security has always been a priority in our lives. Humans have devised several methods for transferring data to one another while assuring that other individuals are unable to obtain this data which is data hiding. Steganography is the technique of concealing a private message in a non-secret message, where communication is essential, by concealing data in other data such as text, picture, audio, and video. Eventually, the data will be concealed from view or discovery, making it impossible for humans to notice that the file contains a hidden word. This study focuses on steganography, which is a technology that encrypts data inside text files, also known as cover files, to conceal the existence of sensitive information. The goal of this research is to elaborate text steganography using symbols as a hidden word. Then, explores the concept of text steganography, its significance in the field of information security, and its potential applications. Various methods of text steganography are discussed, including encoding and decoding techniques using Jupyter Notebook and Python language. Finally, the importance of text steganography in safeguarding sensitive information and the need for further research to enhance its effectiveness and robustness in an ever-evolving digital landscape.

Keywords: Data security, text steganography, data hiding, symbol text

1. INTRODUCTION

Steganography has piqued the interest of researchers as an effective auxiliary method to cryptography (Wang & Gao, 2019). Text steganography is a method of disguising a secret text message as a covering message within another text or constructing a cover message related with the initial secret message. There are several types of text steganography. Subsequently, the data will be hidden from being seen or found. It is also applicable to other forms of media such as text, speech, communication channels, and binary reports (Maheswari et al., 2022). Steganography and cryptography are two different techniques because cryptography only focuses on encrypting data not hiding. Since then, steganography has got a lot of attention nowadays. Data from nations or organizations that require confidentiality must be conveyed appropriately and securely to the right persons (Memis et al., 2022). People currently often utilize the Internet, which leads to cyberattacks in any technology. As a result, new methods for data privacy have emerged.

2. METHODOLOGY

There are six phases in methodology phases which are initiation, planning, development, evaluation, analysis and documentation. The method that was used to carry out this research is a system that can encode and decode text. This is a process of developing a text steganography system that can hide the text or we call it the data in a text cover file. All the software required, which are the Jupyter Notebook application and a Python language, used to develop the performance of text steganography based on symbols. The secret message to be concealed is chosen by the sender. This message can be any form of sensitive or confidential information. An encoding strategy is used to ensure that the concealed message stays intact and recoverable. This method specifies how the secret message is expressed within the cover text through the use of certain encoding rules or algorithms.

3. RESULTS AND DISCUSSION

When the system was built by encoding the concealed message inside the cover text file, it was tested with Hex Editor to compare two files. The cover text file (.txt), which contains the secret message and is the first step in the testing, is filled with varying numbers of symbols. The capacity technique, that calculates the file size with concealed symbols after encoding, is then used to calculate the performance. Figure 1 below shows Hex Editor is used to compare files.

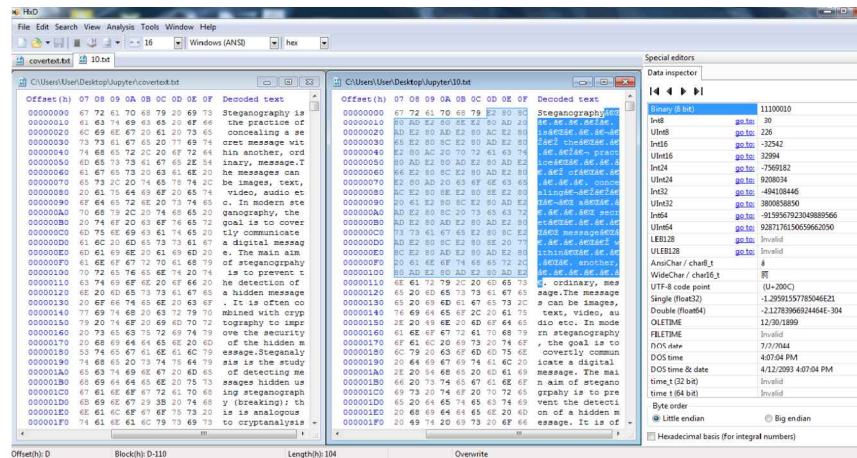


Figure 1 File comparison between cover text file and encoded text file.

The covertxt.txt file (on the left) in Figure 1 above did not contain any hidden symbols. The 10.txt file, on the other hand, is encoded with ten secret symbols, as shown in the highlighted above.

4. NOVELTY OF RESEARCH/PRODUCT

Steganography using symbols in text can help a secret message stand out and be unique. It can increase security in which the secret communications are less likely to be discovered and decrypted by attackers who are skilled in symbol-based steganography techniques when special symbol sets and encoding techniques are used. A further degree of protection is added through the use of unique symbols and encoding techniques.

5. CONCLUSION

In conclusion, this study succeeds in creating a user-friendly graphic user interface and a symbol text steganography system. The capacity criterion measure, which compares the size of the original text file with the encoded text file, was successfully met as a consequence of this research.

REFERENCES

- Maheswari, K. G., Siva, C., Nalinipriya, G., Hemaroshini, M., Jayathraa, V., & Reethika, R. (2022). An Innovative Model for Secure Environment Using Steganography. *8th International Conference on Smart Structures and Systems, ICSSS 2022*. <https://doi.org/10.1109/ICSSS54381.2022.9782259>
- Memis, O. K., Yildiz, S., & Kirci, P. (2022). Designing a Steganography System. *Proceedings - 16th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2022*, 812–815. <https://doi.org/10.1109/TCSET55632.2022.9766938>
- Wang, K., & Gao, Q. (2020). A Coverless Plain Text Steganography Based on Character Features. *IEEE Access*, 7, 95665–95676. <https://doi.org/10.1109/ACCESS.2019.2929123>