Volume 9, Issue 9, September – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

# From Cryptography to Steganography: Detecting Hidden Data in the Digital World

Eman Daraghmi
Palestine Technical University – Kadoorie
Tulkarem, Palestine

Ahmed Hamoudi
Palestine Technical University - Kadoorie
Tulkarem, Palestine

**Abstract:- Steganography is a method used to conceal information, while steganalysis focuses on detecting hidden data. In today's digital landscape, steganography is often used across open communication channels, embedding files, videos, messages, and images within other files to obscure their content from unintended viewers. However, cybercriminals exploit these techniques to covertly transmit data to various devices. Traditional endpoint antimalware tools are not typically designed to search for hidden data, making the detection of steganographic content challenging. The ease with which cybercriminals can transmit data using this method highlights its potential threat. This paper reviews various steganalysis tools and explores the integration of antivirus programs for real-time detection to enhance data confidentiality. A proof-of-concept for one of the steganalysis tools is also provided.**

***Keywords:-*** *Steganographyt; Cryptography; Cryptanalysis; Forensics.*

## I. INTRODUCTION

The confidentiality and integrity of data are paramount for both organizations and individuals. Secure data transmission methods are critical to preventing tampering, unauthorized exchange, or deletion of sensitive information. The transfer of confidential data via unsecured internet channels presents significant challenges in protecting this information. Hackers often seek to break cryptographic protections to gain access to private data. Cryptography, by encoding information into ciphertext, makes it unreadable to unauthorized users, thus safeguarding sensitive data such as credit card numbers, bank account details, and business communications. This encrypted data can only be decrypted with the appropriate key used during encryption [1], [2], [3], [4], [5], [6], [7].

Cryptographic methods can be classified into two main types: symmetric key and asymmetric key approaches, both of which transform plaintext into ciphertext. Encryption is akin to using a secret language that only the sender and recipient can understand, ensuring privacy even if the message is intercepted (see Fig.1). In symmetric key cryptography, the same key is used for both encryption and decryption, while in asymmetric cryptography, different keys are used—one for encryption (public key) and another for decryption (private key). Although both methods provide strong protection, asymmetric encryption is often preferred in scenarios where secure key exchange is difficult.

However, cryptography alone may not always be sufficient to protect sensitive data from more sophisticated attacks. While cryptography obscures the content of messages, it does not conceal the fact that communication is taking place. This is where steganography comes into play as an additional layer of security. Steganography, unlike cryptography, hides the existence of the communication itself by embedding hidden messages within seemingly innocuous files, such as images, audio, or video files. This makes it far less likely that an observer will even realize that sensitive information is being exchanged.

The increasing use of steganography by cybercriminals has raised significant concerns in the field of cybersecurity. Cybercriminals exploit steganography to transmit malicious payloads, command-and-control instructions, or sensitive stolen data without triggering traditional detection mechanisms. The covert nature of steganography makes it especially difficult for traditional antimalware tools to detect these hidden communications, further complicating the task of securing digital environments. For this reason, steganalysis—the study and practice of detecting steganography—has become an essential component of modern cybersecurity strategies.

Steganalysis seeks to uncover hidden information by analyzing the characteristics of digital files and identifying patterns or anomalies that indicate the presence of concealed data. Various tools and techniques have been developed to assist with this process, ranging from simple signature-based detection methods to advanced statistical analysis. Despite these advancements, detecting steganography remains a challenging task, particularly as cybercriminals continue to develop more sophisticated methods to embed data in ways that evade detection.

Given the rising threat posed by steganography in the realm of cybersecurity, this paper aims to review existing steganalysis tools and explore their effectiveness when combined with real-time antivirus solutions. The integration of steganalysis and antivirus tools could significantly improve an organization's ability to detect and respond to hidden threats, providing a more comprehensive approach to data confidentiality. Additionally, the paper presents a proof-of-concept demonstration of one of the steganalysis tools in

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

action, showcasing its capabilities in uncovering hidden data within digital media files.

**The goal of this paper** is to review existing steganography detection techniques, examine various steganalysis tools, and explore how these tools can be integrated with real-time antivirus programs to enhance data confidentiality and security. Additionally, this paper presents a proof-of-concept demonstration of one such steganalysis tool, showcasing its effectiveness in identifying hidden data within digital media. Through this study, we aim to contribute to the growing field of cybersecurity by highlighting the importance of detecting steganography and offering solutions for strengthening organizational data protection.

*A. Cryptography*

*Cryptography is the practice of securing data by converting it into an unreadable format known as ciphertext. This process ensures the confidentiality, integrity, and authenticity of sensitive information during transmission, storage, or access. Cryptographic techniques are essential for protecting communication over untrusted networks, like the internet, preventing unauthorized users from intercepting or understanding the message. The two primary categories of cryptography are symmetric-key encryption, where a single key is used for both encryption and decryption, and asymmetric-key encryption, which uses a pair of keys: a public key for encryption and a private key for decryption [8], [9] [10].*

Symmetric encryption methods, such as AES (Advanced Encryption Standard) and DES (Data Encryption Standard), are efficient and widely used for encrypting large volumes of data. Asymmetric encryption algorithms, like RSA (Rivest-Shamir-Adleman) and ECC (Elliptic Curve Cryptography), are often employed for securing key exchanges and digital signatures, ensuring both confidentiality and integrity [10].

Cryptography serves various purposes beyond just securing communication. It plays a vital role in ensuring data authenticity through digital signatures and hashes, non-repudiation through public key infrastructure (PKI), and secure storage through encrypted files and databases. However, as encryption algorithms evolve, so too do the methods that aim to break them, such as cryptanalysis [11].

*B. Cryptanalysis*

*Cryptanalysis is the science of analyzing cryptographic algorithms and systems to discover weaknesses that can be exploited. The ultimate goal of cryptanalysis is to decrypt the ciphertext without having access to the secret key, thus gaining unauthorized access to the original message. Cryptanalysis often involves both theoretical and practical approaches to finding vulnerabilities in encryption methods, such as weak key generation, algorithmic flaws, or improper implementation.*

One of the most common cryptanalysis techniques includes brute force attacks, where every possible key combination is tested until the correct key is found. Other methods, such as differential cryptanalysis or linear cryptanalysis, focus on analyzing patterns in the ciphertext that can reveal information about the encryption process [11].

Mathematical and statistical tools are frequently used to carry out cryptanalysis. By understanding the structure and characteristics of the cipher being used, cryptanalysts can uncover subtle flaws or overlooked design weaknesses that make an algorithm vulnerable to attack. Advances in quantum computing have posed a potential threat to traditional encryption algorithms, as quantum cryptanalysis could break current standards like RSA and ECC much faster than classical computers [12].

*C.* Steganography

Steganography is a data-hiding technique that focuses on concealing the existence of information by embedding it within other, typically innocuous, digital media files such as images, videos, audio files, and even text documents. Unlike cryptography, which makes data unreadable but does not conceal its existence, steganography hides the presence of the data itself, making it more difficult for potential attackers to detect that communication is taking place [13]. Steganography derives from the Greek words *steganos* ("covered") and *graphein* ("writing"), which together mean "concealed writing." The technique has a long history dating back to ancient times. Early examples include hiding messages within wax tablets or shaving the head of a slave, tattooing a message on the scalp, and allowing the hair to regrow, thereby concealing the message. The concept was formally documented by Johannes Trithemius in his 1499 book *Steganographia*, which explored methods for hiding secret messages [14][14], [15].

The effectiveness of steganography lies in its ability to blend sensitive information with normal digital files without altering their outward appearance. For example, in image steganography, data can be embedded in the least significant bits (LSB) of pixel values, which allows minor modifications to the image without creating noticeable visual changes to the human eye. This makes steganography ideal for use in covert communications, especially over public and unsecured networks [16]. Modern steganographic systems often embed data in multimedia objects like images, audio, or video files, which are frequently shared across the internet [17]. Steganography can be classified into several categories based on the type of cover medium used:

- *Image Steganography*

This method hides secret data within digital images. It is one of the most popular forms of steganography due to the abundance of image files shared online. The least significant bit (LSB) technique is commonly used, embedding hidden information in the lower bits of each pixel, which minimally affects the overall appearance of the image.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

➢ *Audio Steganography*

In this technique, secret data is embedded within audio files. By slightly altering the frequency spectrum or the phase of an audio signal, hidden information can be encoded. Techniques like LSB, phase coding, spread spectrum, and echo hiding are used in audio steganography.

➢ *Video Steganography*

Similar to image steganography, video steganography involves hiding data in the individual frames of video files. The added advantage of video files is that they consist of multiple images (frames) and sound, providing larger data-hiding capacity. Popular formats like MPEG, AVI, and MP4 are commonly used for embedding secret data in video streams.

➢ *Text Steganography*

This method conceals data within text files. It is considered the oldest form of steganography. Modern techniques involve hiding data by manipulating text formatting, altering letter positions, or using linguistic methods to embed information in the structure of the text.

*D. Practical Applications of Steganography*

Steganography has a variety of legitimate applications in securing sensitive information, such as storing encrypted passwords within files or securely transmitting confidential data over insecure communication channels. It can also be used for watermarking digital media, protecting intellectual property rights by embedding ownership information that is invisible to users but detectable with specific tools [19].

However, steganography also has its darker applications. Cybercriminals often use steganographic techniques to hide malicious code or covertly exfiltrate sensitive data from compromised systems without detection. Combined with encryption, steganography presents significant challenges for cybersecurity systems, making detection efforts more complex. Despite this, advancements in steganalysis have made it possible to detect hidden data and mitigate threats through specialized tools and methodologies [20].

This paper aims to explore various steganography detection techniques and tools to uncover hidden messages within common digital files (text, multimedia, and documents). By leveraging advanced steganalysis methods, it is possible to enhance the detection of hidden content, ensuring that steganographic attacks can be neutralized before they compromise data security.

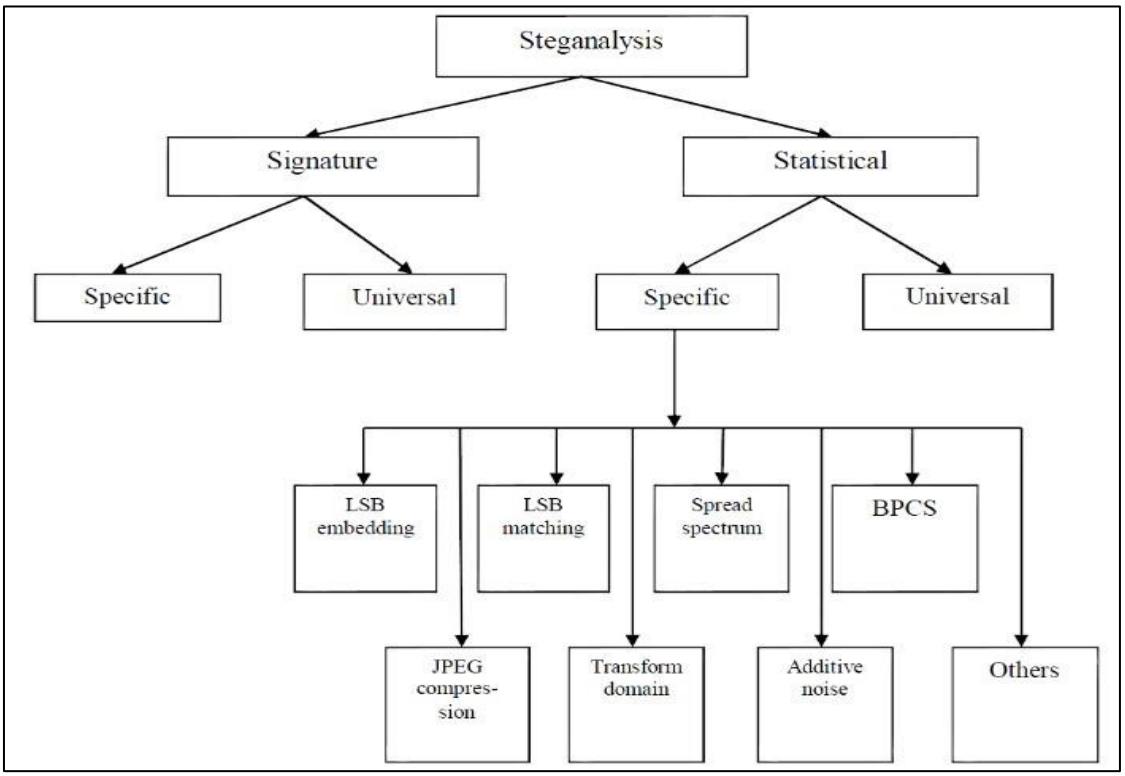


**Fig. 1. Techniques of Steganography**

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

## II. STEGANALYSIS TOOLS AND TECHNIQUES

Steganalysis refers to the process of detecting, extracting, and manipulating hidden data within steganographic files. The field of steganalysis focuses on revealing the existence of covert messages embedded in digital media. Unlike cryptography, where the encrypted message is visible but unreadable, steganography hides the message entirely, making detection more challenging. Steganalysis is therefore crucial for uncovering these hidden messages, particularly in cases where steganography is used for malicious purposes [19].

Two primary types of steganalysis exist: **signature-based** and **statistical-based** techniques. The distinction lies in the method used for detection: either by identifying specific signatures left by steganography tools or by analyzing statistical anomalies in the media.

### *A. Signature Steganalysis*

Signature-based steganalysis detects hidden messages by identifying specific signatures or patterns introduced into the digital media by steganographic algorithms. When data is embedded within an image, audio, or video file, certain steganography tools leave identifiable markers, such as degradation of the file quality or repeated patterns at the end of the file. These markers can be detected using signature-based steganalysis tools, making it possible to determine whether steganography was employed.

For instance, some tools may append a specific string of characters to the end of a file when injecting hidden data. These signature strings act as flags, enabling forensic investigators to uncover hidden messages. By comparing known signatures from different steganographic methods with the target media, the detection process can be automated. Signature-based detection methods are efficient, but they depend heavily on the availability of known signatures and may be ineffective against steganographic methods that leave no visible trace [19], [20] (see Fig.2).

### *B. Statistical Steganalysis*

Statistical steganalysis is more sophisticated and involves analyzing the statistical properties of a digital file to detect hidden information. Rather than relying on specific signatures, this method examines anomalies or irregularities in the file's structure. For example, the **Least Significant Bit (LSB) substitution** method alters the least significant bits of pixel values in an image to store hidden data. This change may introduce subtle distortions in the statistical distribution of pixel values, which statistical steganalysis can detect [21].

Other statistical methods include analyzing the frequency of color changes, detecting the noise patterns in audio files, or examining the statistical distributions of video frames. These techniques are highly effective at uncovering steganographic content even when no obvious signatures are present. However, they may require substantial computational resources and expertise in mathematical modeling to detect small statistical deviations introduced by steganography [21], [22], [23], [24], [25], [26].

### *C. Universal Statistical Steganalysis*

Universal statistical steganalysis refers to detection methods that are adaptable and do not require prior knowledge of the specific steganographic technique being used. This makes it a powerful tool for detecting hidden messages regardless of the embedding method employed. Unlike signature-based steganalysis, which depends on known patterns, universal statistical techniques analyze the underlying statistical properties of the file in question, allowing them to detect steganography even when the method is unknown [8], [9], [11], [16], [17], [19], [20], [22], [24], [25].

For instance, by analyzing the natural statistics of an image and comparing them with the altered statistics after data embedding, universal statistical methods can expose hidden information. Such approaches are versatile and can be applied to various media types (e.g., images, audio, video), making them a critical tool for steganalysis in general-purpose cybersecurity applications.[8], [9], [11], [16], [17], [19], [20], [22], [24], [25].

## III. STEGANOGRAPHY TOOLS AND TECHNIQUES

Steganography involves embedding secret information within seemingly harmless files. The goal is to prevent third parties from realizing that any hidden data exists at all. A variety of techniques and tools have been developed to achieve this, allowing users to hide data in images, audio files, videos, and other digital formats.

### *A. Steganography Techniques*

Digital steganography conceals information by embedding it within a carrier file, which could be an image, audio clip, video file, or even a simple text document. The cover file remains functional and retains its original appearance, making it difficult to detect without specialized analysis. Fig 3 provides an illustration of how digital steganography works [12], [13], [14], [21], [23], [26], [27].

Popular techniques used in digital steganography include:

- *Least Significant Bit (LSB) Substitution*

This is one of the simplest and most widely used methods. It works by replacing the least significant bits of each byte in an image or audio file with bits from the hidden message. Since the changes are so minor, they do not affect the overall quality of the media.

- *Transform Domain Techniques*

These methods hide data in the transformed coefficients of a media file, such as the frequency domain of an image or the discrete cosine transform (DCT) of a video. The hidden information is embedded in the less perceptible regions of the file, making detection more challenging.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

➢ *Spread Spectrum*

In audio steganography, this method spreads the hidden data over a wide frequency range in the audio signal, making it resistant to noise and compression techniques.

➢ *Statistical Methods*

These methods hide data by manipulating the statistical properties of the media file, such as changing the distribution of pixel values in an image or altering the bitrates in an audio file.

*B. Steganography Tools*

Several software tools have been developed to facilitate the embedding of hidden data in various types of files. Table I lists some of the most common steganography tools and the file formats they support:

Table 1. Steganography Previous Tools

| Program | Supported files | Notes |
|---|---|---|
| Anubis | BMP, JPEG, formats. | It appends the message to the end of the file. |
| BMPSecrets | images files | - |
| DeepSound | audio files | It combines the files and encrypt them with AES 256 |
| MP3Stego | MP3 audio files | - |
| Open Stego | BMP, PNG | Open source |
| S-Tools | BMP, GIF, WAV | - |
| Steg | BMP, PNG, JPEG, GIF | Symmetric and asymmetric key cryptography, runs on Win/Linux/Mac |
| StegaMail | BMP, PNG | 56bit encryption, zLib compression |
| Steghide | JPEG, BMP | Open source (GNU General Public License) |

These tools offer users the ability to hide sensitive data within digital media files, whether for legitimate purposes such as secure data transmission or malicious purposes like hiding malware. It is important to note that while these tools make it easy to embed hidden messages, they also present challenges for cybersecurity experts tasked with detecting steganographic content in a digital forensic investigation.

## IV. EXPERIMENT AND RESULTS ANALYSIS

We developed a Python-based tool designed to monitor directory changes in real time. This tool leverages the **Python Watchdog** library, which tracks file creation or modification events within a specified directory. Whenever a file is created or altered, the tool immediately notifies the user.

For the experiment, we combined this Python tool with a **Docker image** containing several steganography detection tools. The focus was on **JPG files**, given the widespread use of image steganography. Specifically, we employed the **Stegovirtas** tool to detect hidden data within these image files. The results showed successful identification of steganographic content, proving the effectiveness of our combined approach.

**Table 2. List of Steganography tools**

| Steganalytic Tools | Tools Analyzed | Detection Ap-proach | Extraction Approach |
|---|---|---|---|
| Steganography Analyzer real time Scanner | Performs Analysis agaonst Network Packets | Signature | |
| StegBreak | Can detect steganography created by the following tools: Jsteg-shell, JPhide, and | | Dictionary |
| Stego-Suite | Detects Steganographic content inside Image and Audio file | | Dictionary |

Fig. 4 illustrates the steganography detection workflow, while Fig. 5 and Fig. 6 demonstrates how the tool analyzed an image containing steganography created with the JPHIDE tool.

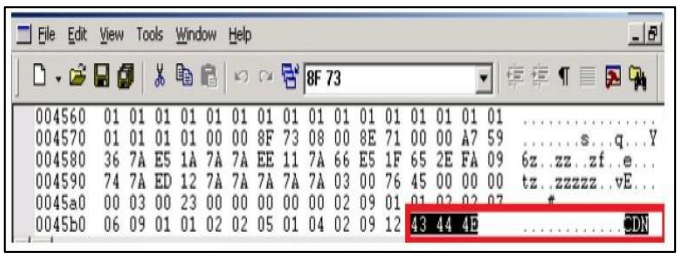


**Fig. 2. Signature Embedded to the End of the File**

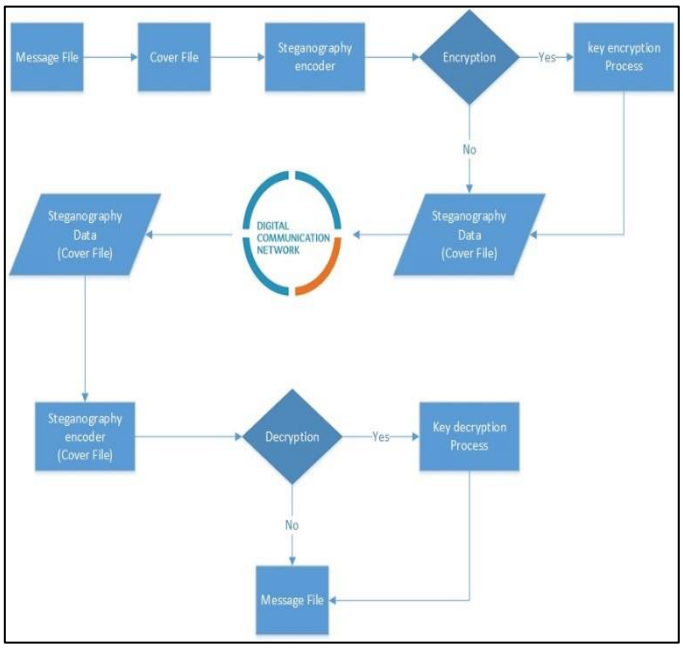


**Fig. 3. Technique of Digital Steganography Process**

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

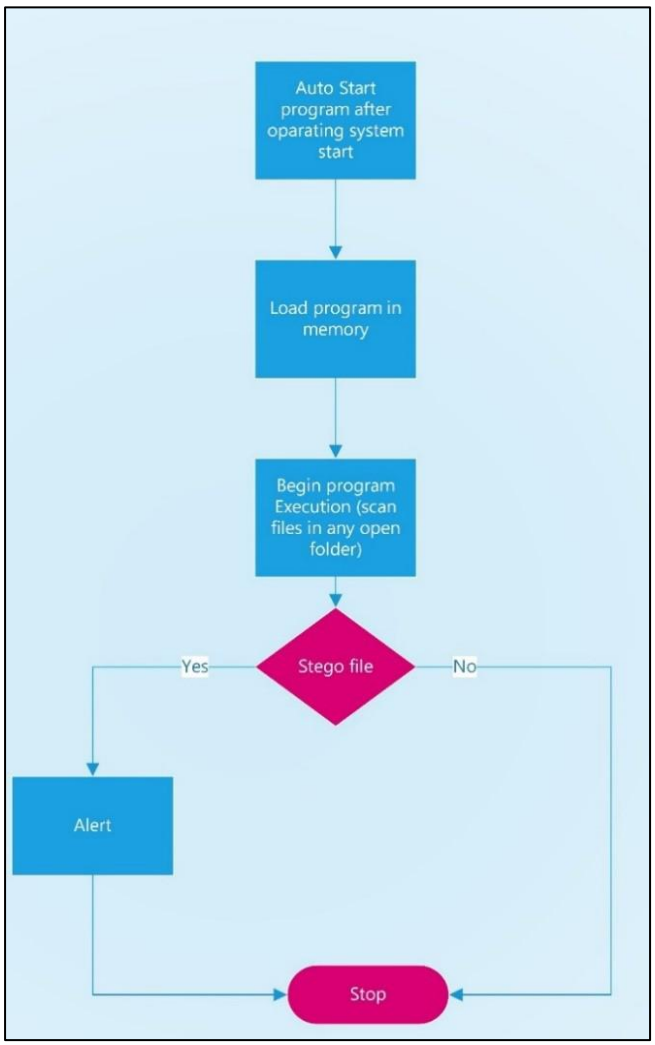


**Fig. 4. Schematic Diagram**

## V. CONCLUSION AND FUTURE WORK

Steganography, particularly when combined with cryptography, is an effective means of secure communication. Although the technique is not widely used, its potential for cybercriminal activity is significant. This paper reviews current steganographic methods and detection tools, along with a practical experiment. However, further research is needed to enhance detection across various file types. Future work could focus on integrating steganalysis tools with antivirus software for real-time detection, providing stronger data confidentiality.

## ACKNOWLEDGMENT

The authors would like to thank Palestine Technical University – Kadoorie for their support.

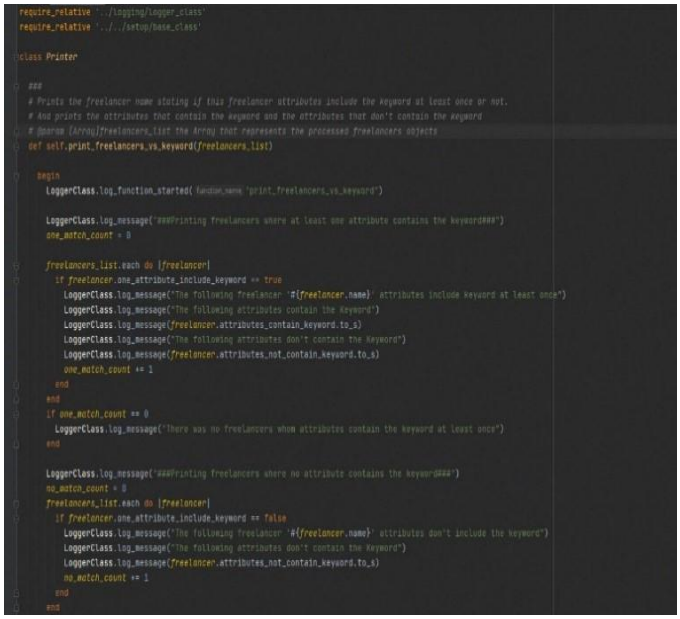

**Fig. 5. Example of Steganography for Images**

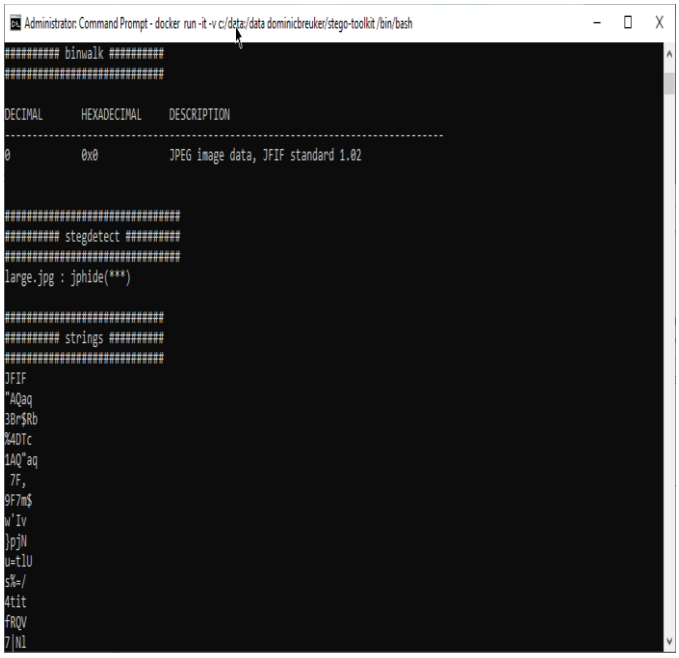


**Fig. 6. Steg Detection**

## REFERENCES

[1]. Z. Alsaed et al., “Role of Blockchain Technology in Combating COVID-19 Crisis,” Appl. Sci., vol. 11, no. 24, p. 12063, Dec. 2021, doi: 10.3390/app112412063.

[2]. E. Daraghmi, Z. Qaroush, M. Hamdi, and O. Cheikhrouhou, “Forensic Operations for Recognizing SQLite Content (FORC): An Automated Forensic Tool for Efficient SQLite Evidence Extraction on Android Devices,” Appl. Sci., vol. 13, no. 19, p. 10736, Sep. 2023, doi: 10.3390/app131910736.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24SEP937

[3]. E. Y. Daraghmi, C. H. Hsiao, and S. M. Yuan, "A New Cloud Storage Support and Facebook Enabled Moodle Module," in 2014 7th International Conference on Ubi-Media Computing and Workshops, Ulaanbaatar, Mongolia: IEEE, Jul. 2014, pp. 78–83. doi: 10.1109/u-media.2014.12.

[4]. E. Y. Daraghmi and Y. S. Ming, "Using graph theory to re-verify the small world theory in an online social network word," in Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services, Bali Indonesia: ACM, Dec. 2012, pp. 407–410. doi: 10.1145/2428736.2428811.

[5]. E. Daraghmi, C.-P. Zhang, and S.-M. Yuan, "Enhancing Saga Pattern for Distributed Transactions within a Microservices Architecture," Appl. Sci., vol. 12, no. 12, p. 6242, Jun. 2022, doi: 10.3390/app12126242.

[6]. E.-Y. Daraghmi, M.-C. Wu, and S.-M. Yuan, "A Multilayer Data Processing and Aggregating Fog-Based Framework for Latency-Sensitive IoT Services," Appl. Sci., vol. 11, no. 4, p. 1374, Feb. 2021, doi: 10.3390/app11041374.

[7]. Y. Salem and E. Daraghmi, "GDPR-BLOCKCHAIN COMPLIANCE FOR PERSONAL DATA," J. Theor. Appl. Inf. Technol., vol. 99, no. 24, pp. 5867–5877, 2021.

[8]. Y. Kumar, R. Munjal, and H. Sharma, "Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures," IJCSMS Int. J. Comput. Sci. Manag. Stud., vol. 11, no. 03, pp. 2231–5268, 2011.

[9]. R. Gupta, "Information Hiding and Attacks: Review," Int. J. Comput. Trends Technol., vol. 10, no. 1, pp. 21–24, 2014, doi: 10.14445/22312803/ijcttv10p105.

[10]. E. Y. Daraghmi, C.-F. Lin, and S. M. Yuan, "Mobile Phone Enabled Barcode Recognition for Preferences Monitoring," in Advances in Computer Science and Education Applications, vol. 202, M. Zhou and H. Tan, Eds., in Communications in Computer and Information Science, vol. 202. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 297–302. doi: 10.1007/978-3-642-22456-0_43.

[11]. T. W. Edgar and D. O. Manz, "Science and Cyber Security," in Research Methods for Cyber Security, Elsevier, 2017, pp. 33–62. doi: 10.1016/b978-0-12-805349-2.00002-9.

[12]. A. Salem, M. Sabbih, H. Al-tamimi, and A. Ahmed, "Secure Image Steganography Through Multilevel Security," Int. J. Comput. Sci. Inf. Secur., vol. 11, no. 1, pp. 80–103, 2020.

[13]. Y. Zheng, F. Liu, X. Luo, and C. Yang, "A Method Based on Feature Matching to Identify Steganography Software," in 2012 4th International Conference on Multimedia and Security (MINES 2012), 2012, pp. 989–994. doi: 10.1109/MINES.2012.26.

[14]. Merriam-Webster, "Definition of steganography." 2020. [Online]. Available: https://www.merriam-webster.com/dictionary/steganography

[15]. E.-Y. Daraghmi and A. Hamoudi, "THE DEVELOPMENT OF A BLOCKCHAIN-BASED SYSTEM FOR ELECTRONIC VOTING," . Vol., no. 17.

[16]. Y. Castelan and B. Khodja, "MP3 Steganography Techniques," in Proceedings of the 4th Annual ACM Conference on Research in Information Technology (RIIT 2015), 2015, pp. 51–54. doi: 10.1145/2808062.2808074.

[17]. A. El-Sayed, G. Attiya, and A. Fkirin, "Steganography Literature Survey, Classification and Comparative Study," Commun. Appl. Electron., vol. 5, no. 10, pp. 13–22, 2016, doi: 10.5120/cae2016652384.

[18]. E. Daraghmi, "Augmented Reality Based Mobile App for a University Campus," 2017, doi: 10.13140/RG.2.2.36356.24962.

[19]. K. Karampidis, E. Kavallieratou, and G. Papadourakis, "A Review of Image Steganalysis Techniques for Digital Forensics," J. Inf. Secur. Appl., vol. 40, pp. 217–235, 2018, doi: 10.1016/j.jisa.2018.04.005.

[20]. M. Kaur and G. Kaur, "Review of Various Steganalysis Techniques," Int. J. Comput. Appl., vol. 5, no. 2, pp. 1744–1747, 2014.

[21]. J. Makwana and S. G. Chudasama, "Dual Steganography: A New Hiding," Int. J. Adv. Res. Electr. Electron. Instrum. Eng., vol. 5, no. 4, pp. 3184–3188, 2016, doi: 10.15662/IJAREEIE.2016.0504109.

[22]. P. Hayati, V. Potdar, and E. Chang, "A Survey of Steganographic and Steganalytic Tools for the Digital Forensic Investigator," in Proceedings of the Workshop of Information Hiding and Digital Watermarking in Conjunction with IFIPTM, New Brunswick, Canada, 2007.

[23]. Shawahniibrahim, "Directory-watcher." 2020. [Online]. Available: https://github.com/shawahniibrahim/Directory-watcher

[24]. DominicBreuker, "Steganography Toolkit." 2020. [Online]. Available: https://github.com/DominicBreuker/stego-toolkit

[25]. Bannsec, "StegoVeritas." 2020. [Online]. Available: https://github.com/bannsec/stegoVeritas

[26]. A. Latham, "Steganography." 1999. [Online]. Available: http://linux01.gwdg.de/ alatham/stego.html

[27]. S. A. Laskar and K. Hemachandran, "A Review on Image Steganalysis Techniques for Attacking Steganography," Int. J. Eng. Res. Technol., vol. 3, no. 1, pp. 3400–3410, 2014.