

TOKYO SUMMIT-7

INTERNATIONAL INNOVATIVE
STUDIES & CONTEMPORARY
SCIENTIFIC RESEARCH
CONGRESS

April 21-23, 2023
Tokyo / Japan

EDITORS

DR. IVAN PAVLOVIĆ

ASSOC. PROF. MONIKA LOPUSZANSKA-DAWID

ISBN - 978-625-367-067-2

PROCEEDINGS BOOK

TOKYO SUMMIT – VII
TOKYO 7TH INTERNATIONAL INNOVATIVE
STUDIES & CONTEMPORARY SCIENTIFIC
RESEARCH CONGRESS

April 21-23, 2023 / Tokyo - Japan

EDITORS

DR. IVAN PAVLOVIC

ASSOC. PROF. MONIKA LOPUSZANSKA-DAWID

All rights of this book belong to

IKSAD Publishing House Authors are responsible both ethically and juridically

IKSAD Publications - 2023©

Issued: 20.04.2023

ISBN - 978-625-367-064-1

CONGRESS ID

CONGRESS TITLE

TOKYO SUMMIT – VII
TOKYO 7TH INTERNATIONAL INNOVATIVE STUDIES &
CONTEMPORARY SCIENTIFIC RESEARCH CONGRESS

DATE AND PLACE

April 21-23, 2023 / Tokyo - Japan

ORGANIZATION

IKSAD INSTITUTE

EDITORS

DR. IVAN PAVLOVIC
ASSOC. PROF. MONIKA LOPUSZANSKA-DAWID

PARTICIPANTS COUNTRIES (**36** countries)

Total Accepted Article: 181

Total Rejected Papers: 37

Accepted Article (Türkiye): 84

Accepted Article (Other Countries): 97

ISBN- 978-625-367-064-1

SCIENTIFIC COMMITTEE

Dr. Keisuke Wakizaka
İstanbul Gelisim University

Prof. Gou FENG
Xiamen University

Prof. Dr. Sevcan YILDIZ
Akdeniz University

Prof. Samson OYEYINKA
University of Ilorin

Assoc. Prof. Dr. Sehrana KASIMİ
Azerbaijan MEA

Prof. Dr. Umran TURKYILMAZ
Ankara Hacı Bayram Veli University

Assoc. Prof. Dr. Froilan D. Mobo
Philippine Merchant Marine Academy

Assoc. Prof. Dr. Aparna Srivastava
Noida International University

Dr. Ly Dai HUNG
VASS Vietnamese Institute of Economics

Dr. Abdussalam Ali Ahmed
Bani Waleed University

SCIENTIFIC COMMITTEE

Dr. Vikas Prajapati
University of Baroda

Dr. Xiangyi KONG
Chinese Academy of Medical Sciences

Dr. Luna Moni DAS
Vasanta College for Women

Dr. Pham Ngoc NHAN
University of Economic Ho Chi Min City

Dr. Kahkashan Khan
Malaviya University

Dr. Minji YANG
Busan University

Dr. Yicheng WU
Minzu University

Dr. Machunwangliu KAMEI
University of the People, California

Dr. Zhi Huan MENCHUANG
Renmin University

Cengiz TOPDEMIR
Awarded Mathematician

SCIENTIFIC COMMITTEE

**Kanokwan Somwong
Chiang Mai University**

**Ankit Gupta
University of Lucknow**

**Janaka Wijesinghe
Uva Wellassa University**

**Dr. Aygun MEHERREMOVA
Baku State University**

**Dr. Gulshen MEHERREMOVA
Azerbaijan University of Languages**

ORGANIZING COMMITTEE

Dr. WU Yicheng

Minzu University

Dr. Ly Dai HUNG

Vietnamese Institute of Economics

Dr. Mariam RASULAN

IKSAD Institute

Dr. Xiangyi KONG

Chinese Academy of Medical Sciences

Dr. Pham Ngoc NHAN

University of Economic Ho Chi Min City

Prof Dr Morakeng Edward Kenneth Lebaka

University of South Africa

Dr. Maria HOOKS

Methodis Hospital

Dr. Minji YANG

Busan University

**TOKYO 7TH INTERNATIONAL INNOVATIVE STUDIES & CONTEMPORARY
SCIENTIFIC RESEARCH CONGRESS
HAMMING CODING TO IMPROVE THE INTEGRITY OF TEXT-TO-IMAGE
STEGANOGRAPHY
METİNDEN GÖRÜNTÜYE STEGANOĞRAFİNİN BÜTÜNLÜĞÜNÜ ARTIRMAK İÇİN
HAMMING KODLAMASI**

Assist.Prof.Dr. Hüseyin Bilal MACİT

Burdur Mehmet Akif Ersoy University, Bucak ZTYO, Department of Information Systems and
Technologies

ORCID ID: <https://orcid.org/0000-0002-5325-5416>

ABSTRACT

The steganographic technique hides a significant data into an insignificant data and transmits it without arousing any suspicion in the transmission medium. Even an ordinary multimedia message can secretly carry dozens of pages of documents. Image files are frequently used in steganography due to their high payload. However, images are always open to manipulation and they can easily lose confidential data. Attackers often try to destroy confidential information by manipulating one or more LSB bits. In this study, Hamming coding is applied to increase the integrity of the confidential text data in the text-to-image steganographic method. The proposed method has been tested with various input parameters on an image dataset. Some artificial attacks were carried out on stego images, and the amount of lost and recovered data was evaluated.

Keywords: Data security, Steganography, Hamming Code, LSB.

ÖZET

Steganografik teknik, önemli bir veriyi önemsiz bir başka veriye gizleyerek iletim ortamında şüphe çekmeden iletir. Sıradan bir multimedya mesajı bile onlarca sayfalık dokümanı gizlice taşıyor olabilir. Görüntü dosyaları, yüksek taşıma kapasiteleri nedeniyle steganografide sıkça kullanılırlar. Ancak görüntüler manipülasyona her zaman açıktır ve gizli veriyi kolayca kaybedebilirler. Saldırganlar genellikle bir veya birkaç LSB bitini manipüle edilerek gizli bilgiyi yok etmeye çalışırlar. Bu çalışmada, görüntü içine metin gizlenen steganografik yöntemde gizli verinin sağlamlığını arttırmak için Hamming kodlaması uygulanmıştır. Yöntem, bir görüntü veri seti üzerinde çeşitli parametreler ile test edilmiştir. Stego görüntüler üzerinde bazı yapay saldırılar gerçekleştirilmiş, kaybedilen ve kurtarılan veri miktarları değerlendirilmiştir.

Anahtar kelimeler: Veri güvenliği, Steganografi, Hamming Kodu, LSB.

INTRODUCTION

The amount of data produced and transmitted in the world every day is more than the previous day. Almost all of the transmitted data is transmitted over the internet. Today, the Internet connects individuals, institutions and governments. Although there are advantageous areas of use such as communication, business, shopping and entertainment, the biggest disadvantage of the internet is still security. The fact that data has to be transported over insecure communication lines threatens their privacy, integrity and confidentiality (Subramanian et al., 2021). Attackers eavesdropping on the communication line can access or even modify data. The methods used to ensure data security are classified as in Figure 1.

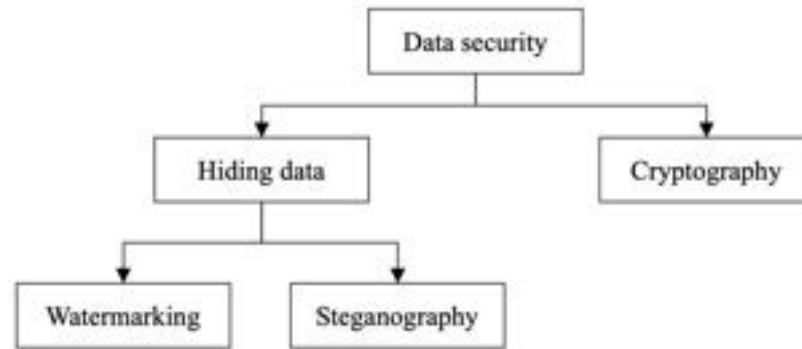


Figure 1. General data security methods (Mandal et al., 2022).

The prominent solution today to ensure data security is cryptography. Data secured by cryptography technique is transmitted encrypted along the transmission line. Thanks to today's high-performance processors, data can be encrypted with more complex algorithms. But there is no unbreakable encryption method. In some cases, data needs to be freely distributed on the transmission line without compromising its integrity. In this case, digital watermarking or steganography techniques are used instead of cryptography. In a digital watermarking technique, the watermark is often used for authentication and copyright protection. However, steganography mostly aims to hide information (Mandal et al., 2022). Table 1 shows the characteristic differences between steganography, watermarking, and cryptography.

Table 1. Comparison among data security systems (Majeed et al., 2021)

Characteristics	Steganography	Watermarking	Cryptography
Goal	Protects secret data from discover	Protects legitimacy of the media	Disorganizes the content of data
Cover choosing	Free	Limited	No cover use
Challenges	Imperceptibility, security and payload	Robustness	Robustness
Key use	Sometimes	Rarely	Always
Output	Stego-media	Watermarked media	Ciphertext
Visibility	Never acceptable	Sometimes necessary	Not a problem
Fail	If notices	If substituted	If decrypted

Steganography is the process of hiding a secret data into another cover data. Confidential data and the cover data can be a text, image, video or audio. The aim is to transmit confidential data without any suspicion over an insecure transmission medium. The name Steganography was put by Johannes Trithemus who lived between years 1462 and 1516. The word “Steganographia” is composed of Greek words “στεγανό-ς” which means secret and “γραφ-ειν” which means writing (Macit ve Koyun, 2020). Looking at history, it is seen that steganography was used even 3000 years ago. For example, around 440 BC, Demaratus sent a warning message on a wax tablet about an impending attack. The message was written on a wooden support and then protected with wax (Mandal et al., 2022). Today, steganography usage areas are military, intelligence agencies, medical, multimedia, smart IDs, corporate, advanced data structure, multimodal biometric data, radar systems, remote sensing, documents tracking tools, digital elections, electronic money, etc (Dalal and Juneja, 2021).

There are three factors that affect the performance of a steganographic system. These are imperceptibility, robustness and payload (Macit and Koyun, 2020). Imperceptibility refers to how well confidential data is camouflaged. Robustness is the resistance of confidential data against unauthorized steganalysis attacks. The payload specifies the amount of confidential data that the method can hide in the cover data. Today, the most popular steganography techniques use images as cover data due to high data load (Sachin and Rashmi, 2020). For this reason, steganalysis attacks are also usually performed on images. The first purpose of the steganalist is to determine whether the image is an ordinary image or a stego-image. If the steganist believes that the image is a stego-image, he or she launches various attacks to expose or corrupt confidential data. A general steganography process is shown in figure 2.

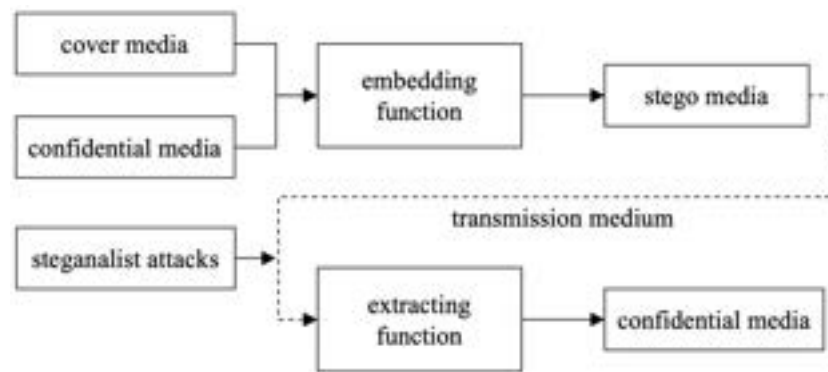


Figure 2. Common image steganography process

There are two general approaches to image steganography; spatial domain and transform domain. In spatial domain techniques, the pixel values of the cover data are manipulated directly to embed the hidden data. These techniques can be applied quickly and easily (Dalal and Juneja, 2021). Commonly used spatial domain image steganography methods are Least Significant Bit (LSB), palette based, Pixel Value Difference (PVD), multi-bitplane based techniques, histogram shifting, Random Pixel Embedding (RPE), Edges based Information Embedding (EBE), Multi-Pixel Differencing (MPD), etc. In the transform domain, image steganography transforms the cover image into the frequency domain. Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Complex Wavelet Transform (CWT), Integer Wavelet Transform (IWT), Dual-tree Complex Wavelet Transform and DFT (Discrete Frequency Transform) methods are used for this transformation (Pilania, 2021).

In this study, a method is proposed to improve the integrity of LSB steganography. The proposed method applies Hamming coding in the data hiding phase. Hamming coding was developed by Richard Hamming in 1950 to improve the error correction methods of computers. Hamming codes are a class of linear block code (Hamming, 1950). Hamming coding detects whether there is a change in the bits of the data reaching the receiver during the transmission process and can correct the 1-bit error (Singh, 2016). An example Hamming encoding for a 4-bit data block is shown in figure 3.

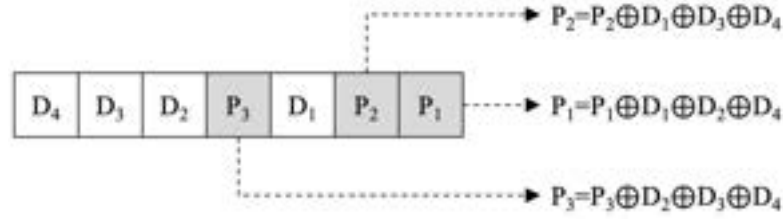


Figure 3. Hamming coding for 4 bit data

METHOD

Let $x_{i,j}$ be a pixel of the cover image I . The process of hiding the d data for this pixel with LSB steganography is done as in equation 1.

$$I' = I - (I \bmod 2^n) + d \quad (1)$$

The I' is the stego image, and n represents the number of bits modified to embed the hidden data. When the Stego image reaches the receiver, the receiver can read the confidential data as in equation 2.

$$d = I' \bmod 2^n \quad (2)$$

LSB-based image steganography techniques are easy to implement but more sensible to steganalysis attacks. The steganist usually applies algorithms such as adding noise or smoothing to the image to corrupt the confidential data. Thus, confidential data is tried to be destroyed by corrupting one or more LSBs in each pixel. In studies conducted with this technique, it is generally seen that confidential data is encrypted (Dalal and Juneja, 2021). However, when even 1 bit of the encrypted data is changed, it may be impossible to reach plain text again. This study proposes a method that increases robustness against steganalysis attacks by using Hamming coding for image steganography using text as confidential data.

Embedding phase

Each character of the confidential text data consisting of c characters is converted to a 7-bit binary array according to ASCII standards. Thus, c arrays of $t_i = b_7b_6b_5b_4b_3b_2b_1$ are obtained. $t = \{t_i | 0 \leq i < c\}$, $t_i \in \{0000000, \dots, 1111111\}$

(3)

The number of Hamming parity bits r required for a bit string consisting of m bits is calculated as in equation 4.

$$m + r + 1 \leq 2^r \quad (4)$$

Accordingly, for each t_i

$$7 + r + 1 \leq 2^r \rightarrow r = 4 \quad (5)$$

The Hamming parity bits are placed in the indexes corresponding to the exponents of 2 from right to left. Adding Hamming code to t_i produces th_i ;

$$th_i = b_7b_6b_5p_4b_4b_3b_2p_3b_1p_2p_1 \quad (6)$$

The parity bits for th_i are calculated using the XOR operator, respectively, as in equations 7-10:

$$p_1 = p_1 \oplus b_1 \oplus b_2 \oplus b_4 \oplus b_5 \oplus b_7 \quad (7)$$

$$p_2 = p_2 \oplus b_1 \oplus b_3 \oplus b_4 \oplus b_6 \oplus b_7 \quad (8)$$

$$p_3 = p_3 \oplus b_2 \oplus b_3 \oplus b_4 \quad (9)$$

$$p_4 = p_3 \oplus b_5 \oplus b_6 \oplus b_7 \quad (10)$$

For example, for the "a" ASCII character in the i .line of the text, $t_i = 1100001$. In this case, $th_i = 11000000110$. Now the resulting 11-bit th_i can be embedded in the LSBs of the image. For monochrome cover image C , which consists of C_1 rows and C_2 columns, each pixel is expressed with 8 bits.

$$C = \{x_{ij} | 0 \leq i < C_1, 0 \leq j < C_2, x_{ij} \in \{0,1,2, \dots, 255\}\} \quad (11)$$

x_{ij} refers to a pixel of the C image. The C image is transferred to a one-dimensional V vector by horizontal hatching.

$$V = \{v_i | 0 \leq i < C_1C_2, v_i \in \{0,1,2, \dots, 255\}\} \quad (12)$$

th_i 's are embedded in each element of V respectively.

$$LSB(v_i) = th_i(k) \quad k \in \{0, \dots, 11\} \quad (13)$$

In equation 13, k is the pointer to the next bit of th_i . The value of k is incremented by 1 at each step. After reaching the last bit of th_i ($k = 11$), k again displays the first bit for the next Hamming encoded character ($k = 1$). After the last element of the th vector is processed, the data hiding phase of the algorithm is completed. The resulting V' vector can be converted back to $C_1 \times C_2$ monochrome image C' and sent from the transmission medium.

Extracting phase

When the receiver receives the stego image, it extracts the confidential data. This phase is called the extracting phase. In the proposed method, when the receiver receives the C' image with $C_1 \times C_2$ resolution, it converts the image to the V' vector.

$$V' = \{v'_i | 0 \leq i < C_1C_2, v'_i \in \{0,1,2, \dots, 255\}\} \quad (14)$$

Each v'_i contains a hidden bit whose accuracy is questionable. th'_i vectors are created by reading the LSB of each v'_i in turn. Each th'_i is a vector of 11 binary digits representing one character of the hidden text.

$$th'_i(k) = LSB(v_i), k \in \{0, \dots, 11\} \quad (15)$$

To check the integrity of the secret data, the XOR operation in the equations 16-19 is performed on the parity bits of each th'_i respectively.

$$K_1 = p_1 \oplus b_1 \oplus b_2 \oplus b_4 \oplus b_5 \oplus b_7 \quad (16)$$

$$K_2 = p_2 \oplus b_1 \oplus b_3 \oplus b_4 \oplus b_6 \oplus b_7 \quad (17)$$

$$K_3 = p_3 \oplus b_2 \oplus b_3 \oplus b_4 \quad (18)$$

$$K_4 = p_3 \oplus b_5 \oplus b_6 \oplus b_7 \quad (19)$$

K series is the Hamming control series and indicates whether there is an error in th'_i . When the K series is converted to decimal and if $(K_4K_3K_2K_1)_{10} = 0$, th'_i is intact. However, if $(K_4K_3K_2K_1)_{10} > 0$, the value read from this series indicates which bit of th'_i is corrupt. In this case, the bit found to be corrupt is inverted. At the end of this process, the parity bits of th'_i are removed and a 7-bit series is obtained. This series is the ASCII code of the i .character.

RESULTS AND CONCLUSION

To test the performance of the proposed method, 512x512 pixel monochrome Lena, Cameraman and Baboon test images were used.



Figure 4. Test images

The payload for test images is calculated in equation 20.

$$payload = \frac{row.col}{bitarraysize} = \frac{512 \times 512}{11} = 23831 \text{ characters} \quad (20)$$

Even if only 1 character long data is hidden on the cover image, this causes an invisible difference between the cover image and the stego-image. A good steganography method makes minimal changes to the cover image even when using maximum payload. For example, if the LSBs of all pixels is assumed to change in the classical LSB method, the image will change by $1/256$. In this study, Peak to Signal Noise Ratio (PSNR) measurement was performed to calculate how close the C' taken by the receiver from the transmission medium is to the cover image C in vector. PSNR is calculated as the logarithm of the Mean Square Error (MSE). MSE is the sum of the squares of the pixel differences between the two images (Setiadi, 2021).

$$PSNR = 10 \log_{10} \left(\frac{max^2}{MSE} \right) \quad (21)$$

In equation 21, the max value is the maximum intensity value a pixel can get. The test images are 8 bit monochrome images. So, $max = 2^8 - 1 = 255$. $PSNR = \infty$ is calculated when C and C' images are identical. $PSNR > 40$ is expected for an imperceptible steganography method. To test the proposed method, hidden texts containing random characters between 1000 and 23000 characters were created. The PSNR results measured in the 3 test images are nearly identical (Figure 5).

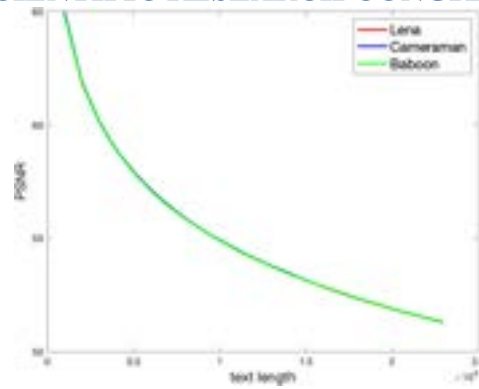


Figure 5. PSNR results of C versus C'

The absolute histogram difference between C and C' is given in the graphics in figure 6 when steganography is applied to the test images with the generated character strings.

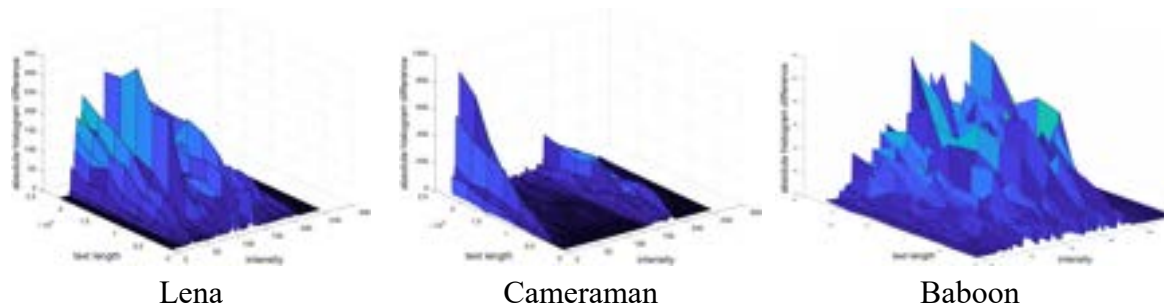
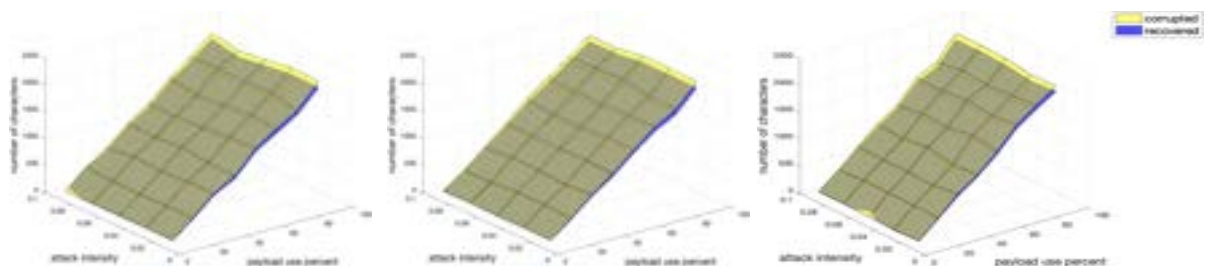


Figure 6. Surface graphs of absolute histogram difference of C versus C' for each test image

The histogram differences shown in Figure 6 clearly show that the proposed method causes different characteristic changes in different images. This situation complicates the steganalysis and increases the robustness of the method. The proposed method is designed on the assumption that the stego image will suffer partial degradation in the transmission line. To test the performance of the method, an artificial steganalysis attack was performed on the stego image. For this, salt&pepper noise was added to the image with different intensity parameters ranging from 1% to 10%. The results obtained in stego-images as a result of each attack applied with different parameters are shown with the surface graphics in figure 7. The yellow surfaces show the number of corrupted characters detected by the Hamming coding as a result of the attack. Blue surfaces indicate the number of characters recovered with Hamming parity bits. As seen in the figures, despite the increase in the intensity of the attack, most of the corrupted characters were successfully recovered.



Lena

Camerman

Baboon

Figure 7. Surface graphs of corrupted and recovered charater counts for test images

The ratio of the characters that could not be recovered to the total number of corrupted characters is shown in the graphics in Figure 8. As can be seen, only 5% of the corrupted characters could not be recovered. The proposed method offers approximately 95% recover guarantee.

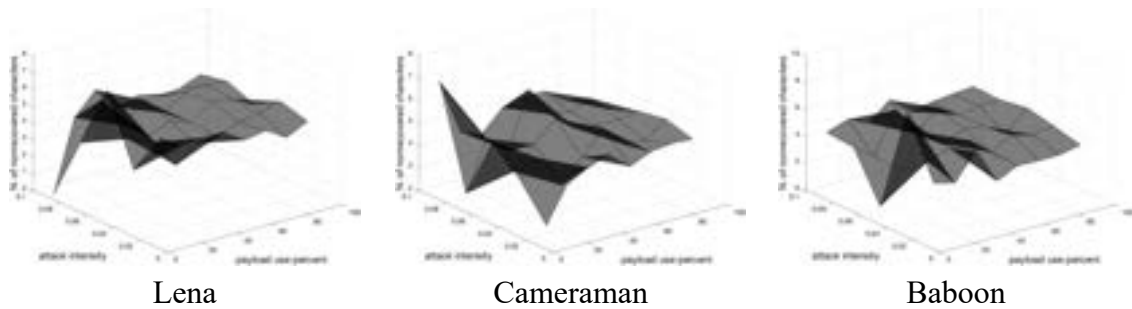


Figure 8. Surface graphs of percentage of non-recovered character counts to corrupted characters

In this study, Hamming coding method is applied to improve the integrity of confidential data in text-to-image steganography. An artificial salt&pepper attack was applied to evaluate the performance of the method. Even in the case of the change in the total 10% of the stego image, which was the worst case scenario in the tests, only ~0.5% of the hidden text was lost. The method provides both imperceptibility when evaluated in terms of PSNR results and high integrity when evaluated in terms of recover performance.

REFERENCES

- Dalal, M., Juneja, M. (2021). Steganography and Steganalysis (in digital forensics): a Cybersecurity guide, Multimedia Tools and Applications, (80), 5723–5771, <https://doi.org/10.1007/s11042-020-09929-9>
- Hamming, R.W. (1950). Error Detecting and Error Correcting Codes, Bell System Technical journal, (29)2, 147–160.
- Macit, H.B., Koyun, A., (2020). A New Imperceptible Steganography Method for Grayscale Images, Journal of Engineering Sciences and Design, 8(2), 357-365.
- Majeed,M.A., Sulaiman, R., Shukur, Z., Hasan, M.K. (2021). A Review on Text Steganography Techniques, Mathematics ,(9)2829, <https://doi.org/10.3390/math9212829>
- Mandal, P.C., Mukherjee, I., Paul, G., Chatterji, B.N. (2022). Digital image steganography: A literature survey, Information Sciences, (609), 1451–1488, <https://doi.org/10.1016/j.ins.2022.07.120>
- Pilania, U., Tanwar, R., Gupta, P., Choudhury, T. (2021). A roadmap of steganography tools: conventional to modern, Spatial Information Research, 29(5), 761–774, <https://doi.org/10.1007/s41324-021-00393-7>.

**TOKYO 7TH INTERNATIONAL INNOVATIVE STUDIES & CONTEMPORARY
SCIENTIFIC RESEARCH CONGRESS**

Sachin, D., Rashmi, G. (2020). Analysis of various data security techniques of steganography: A survey, Information Security Journal: A Global Perspective, DOI: 10.1080/19393555.2020.1801911

Setiadi, D.R.I.M. (2021). PSNR vs SSIM: imperceptibility quality assessment for image steganography, Multimedia Tools and Applications, (80), 8423–8444, <https://doi.org/10.1007/s11042-020-10035-z>

Singh, A.K. (2016). Error Detection and Correction by Hamming Code, 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication, 35-37.

Subramanian, N., Elharrouss, O., Al-Maadeed, S., Bouridane, A. (2021). Image Steganography: A Review of the Recent Advances, IEEE Access, (9), <https://doi.org/10.1109/access.2021.3053998>

TOKYO SUMMIT-VII

7TH INTERNATIONAL CONFERENCE ON INNOVATIVE STUDIES OF CONTEMPORARY SCIENCES


REF: INVITATION FOR A CONFERENCE

INVITATION FOR A CONFERENCE

İlgili makama;

TOKYO SUMMIT – VII, TOKYO 7TH INTERNATIONAL INNOVATIVE STUDIES & CONTEMPORARY SCIENTIFIC RESEARCH CONGRESS, April 21-23,2023 tarihleri arasında Manhattan, Tokyo, Japan’da 36 farklı ülkenin akademisyen / araştırmacılarının katılımıyla gerçekleşmiştir. Kongre kapsamında sunumu yapılan 181 bildirinin 84 adeti Türkiye’den katılımcılar tarafından; 97 bildiri ise 35 ülkeden katılımcılar tarafından sunulmuştur. Kongre 16 Ocak 2020 Akademik Teşvik Ödenęi Yönetmeliğine getirilen “Tebliğlerin sunulduğu yurt içinde veya yurt dışındaki etkinlięin uluslararası olarak nitelendirilebilmesi için Türkiye dışında en az beş farklı ülkeden sözlü teblię sunan konuşmacının katılım sağlaması ve teblięlerin yarıdan fazlasının Türkiye dışından katılımcılar tarafından sunulması esastır.” deęişikliğine uygun düzenlenmiştir. Bilgilerinize arz edilir,

Saygılarımla



Dr. Mustafa Latif Emek
On behalf of the Organizing Committee