www.dergipark.gov.tr
ISSN:2148-3736
El-Cezerî Fen ve Mühendislik Dergisi
Cilt: 6, No: 3, 2019 (668-680)

El-Cezerî Journal of Science and Engineering
Vol: 6, No: 3, 2019 (668-680)
DOI: 10.31202/ecjse.571030

ECJSE

Research Paper / Araştırma Makalesi

The Creation of Maze in Order to Hide Data, and the Proposal of Method Based on AES Data Encryption Algorithm

Erdal GÜVENOĞLU[1]*, Can RAZBONYALI[2]

[1] Maltepe University, Faculty of Engineering, Department of Computer Engineering, Istanbul, Turkey.
[2] Maltepe University, Vocational School, Computer Programming Department, Istanbul, Turkey.

Received/Geliş: 28.05.2019 **Accepted/Kabul:** 20.07.2019

Abstract: The field of science, which enables the transmission of private data in multimedia environments such as display, audio and video, is called steganography. The widespread use of the internet in our daily lives has brought about the problem of transmitting confidential information. Hence, the need for steganography science has also increased. In this study, a labyrinth is produced in the size of the environment in which the data can hide with the help of a private key. Color images with 24-bit depth were used as data hiding environment in different measures and different formats. For the production of maze, the Depth First Search algorithm, which ensures both the creation of the maze and the finding of solution path together with this, was utilized. It is made the information hide inside the pixels on the image, corresponding to the coordinates of the solution path obtained from the maze. The Lowest Bit Insertion method was used to hide information. In order to make the proposed method more robust, the data, which will be hidden before the data hiding process was encrypted with the AES encryption algorithm. This makes it difficult for third parties to obtain confidential data. The proposed method was tested with detailed security analysis and it was observed that data hiding / data extraction processes were successful.

Keywords: Steganography, Data Hiding, Data Security, Maze Generation, Stego Image.

Veri Gizlemek Amacıyla Labirent Oluşturma ve AES Veri Şifreleme Algoritması Tabanlı Bir Yöntem Önerisi

Özet: Görüntü, ses ve video gibi multimedya ortamlarında gizli verilerin iletilmesini sağlayan bilim dalına steganografi adı verilmektedir. Günlük hayatımızda internetin yaygın olarak kullanılması, kişiye özel bilgilerin iletilmesi problemini de beraberinde getirmiştir. Dolayısı ile steganografi bilimine olan ihtiyaç da artmıştır. Bu çalışmada kişiye özel bir anahtar yardımı ile veri gizlenecek ortam büyüklüğünde bir labirent üretilmektedir. Veri gizleme ortamı olarak farklı ölçülerde ve farklı formatlarda 24 bit derinliğe sahip renkli resimler kullanılmıştır. Labiret üretimi için hem labirentin oluşturulmasını hem de çözüm yolunun birlikte bulunmasını sağlayan Derin Öncelikli Arama algoritmasından faydalanılmıştır. Labirentten elde edilen çözüm yolu koordinatlarına denk gelecek şekilde görüntü üzerindeki piksellerin içerisine bilgi gizlenmektedir. Bilgi gizlemek için En Düşük Bite Ekleme yöntemi kullanılmıştır. Çalışmada önerilen yöntemin daha güçlü kılınması için, veri gizleme işleminden önce gizlenecek veriler Gelişmiş Şifreleme Standardı şifreleme algoritması ile şifrelenmiştir. Böylelikle gizli verilerin üçüncü şahıslar tarafından elde edilmesi güç hale gelmektedir. Öne sürülen yöntem detaylı güvenlik analizleriyle test edilmiş ve veri gizleme/veri çıkarma işlemlerinin başarılı olduğu gözlenmiştir.

Anahtar Kelimeler: Steganografi, Veri Gizleme, Veri Güvenliği, Labirent Üretimi, Şifreleme.

How to cite this article
Güvenoğlu, E., Razbonyalı, C.,*"The Creation of Maze in Order to Hide Data, and the Proposal of Method Based on AES Data Encryption Algorithm"*,El-Cezerî Journal of Science and Engineering, 2019, 6(3); 668-680.

Bu makaleye atıf yapmak için
Güvenoğlu, E., Razbonyalı, C., *"Veri Gizlemek Amacıyla Labirent Oluşturma ve AES Veri Şifreleme Algoritması Tabanlı Bir Yöntem Önerisi"*, El-Cezerî Fen ve Mühendislik Dergisi 2019, 6(3); 668-680.

ORCID NO (Güvenoğlu, E.: 0000-0003-1333-5953; Razbonyalı, C.: 0000-0001-6235-5323)

1. Introduction

Steganography is a means of preserving another medium of data into a carrier medium of data in order to ensure the security of the data. Traditionally, video, audio, and picture files that are used in the internet environment are used as the carrier medium [1]. The steganography word was derived from Greek and means "covered writing" [2-3]. The purpose of it is to hide a confidential article or its existence. In this approach, the medium in which the information is hidden is called cover-data or cover-object, and the medium that is formed is called stego-text or stego-object [4-5].

Steganography, despite being close to Encryption, is different from Encryption. While encryption deals with the preservation of the content of the message steganography deals with the concealment of the presence of the message. Thus, steganography is not a method of encryption, but a complementary element to encryption [6-7]. The biggest advantage of steganography in comparison with encryption is that one who looks at the cover data can not realize that it is important information in the medium that it is looking at. Hence, the attacks on these data can be prevented with hiding of the data in an appropriate medium. On the contrary, even if it is difficult to decipher a encrypted information, it is of great interest because of the mystery it contains. It can also be understood that the data is encrypted; and its password is tried to be broken by attack techniques over time. As a result, while encryption deals with the preservation of the content of the message steganography deals with the concealment of the presence of the message [8]. The block diagram of the general steganography system for image files is shown in Figure 1.

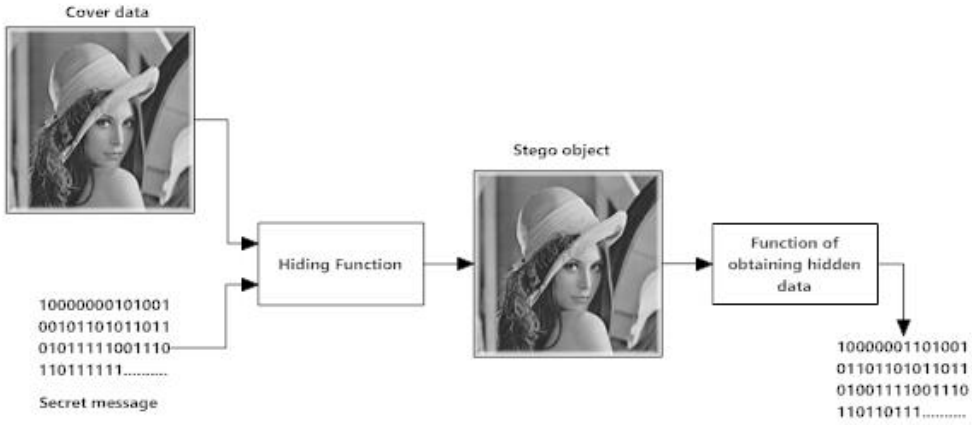


Figure 1. Block diagram of the steganography system for image files.

The most common technique used to hide information in steganography is LSB (Least Significant Bit). This method is based on the insertion of bit values of the hidden data into the least significant bit of the pixels in the cover data. An example of adding a character belonging to the LSB method is shown in Figure 2.

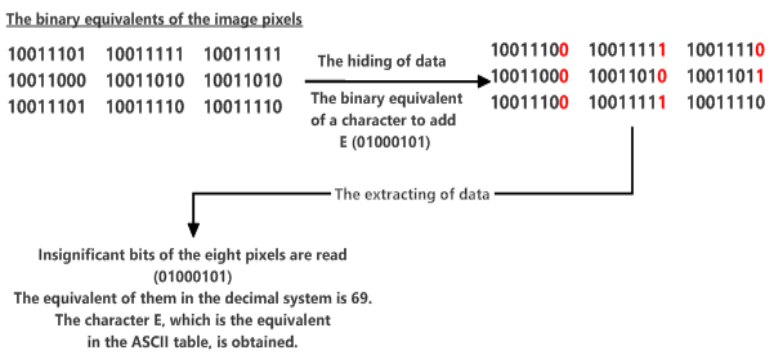


Figure 2. A sample LSB application.

In this work, a maze was produced by using DFS (Depth First Search) algorithm. The Data were hidden in the least significant bit of the pixel values on the solution path of the maze by using the LSB method. The maze and solution path are produced based on a secret user key. Therefore, there are an unlimited number of maze and solutions. In order for the method to be stronger, the data which will be hide had been encrypted by the AES (Advanced Encryption Standard) encryption algorithm before having been hidden.

2. Maze Generators

The structures which one can get stuck in it because of the multiplicity and complexity of the passages in it, is called the maze. According to Greek mythology, the palace made by the architect Daedalus for the King of Crete Minos was also called the maze. In archaeological excavations in countries such as Crete, Egypt, it was encountered to the remains of buildings made in maze form [9]. Nowadays the Mazes have become a source of inspiration for many computer games and academic studies.

There are many methods in the literature in order to create mazes. Each of these methods has different characteristics from each other. A maze basically consists of cells, walls, starting cell and end cell. Logically, a maze should have a complex path [10]. Many of the great mazes have a rectangular structure. A rectangular maze is composed of m cell width and n cell height and expressed as $m \times n$. If there is only one way between the two cells of the maze, it is called the perfect maze [11]. When security is considered, there must be only one exit of a maze [10]. The basic components of a maze are shown in Figure 3 [11].

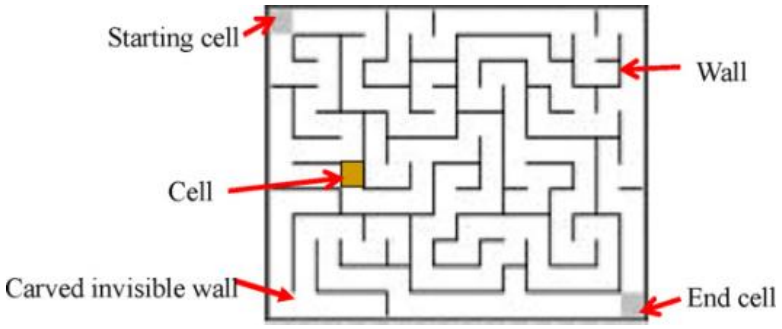


Figure 3. Maze structure.

In this study, the mazes have to be created very quickly due to the size of the picture files. For this reason, the DFS algorithm which makes both the creating of a maze and also the resolution of the path quickly was used. When it comes to image security, it is important that the maze has to be produced extremely complex and it has to be difficult to solve. The width and height of the maze to be produced in study are the same as the size of the image to be encrypted.

3.Depth First Search Maze Algorithm

The DFS is an algorithm which can be used to create maze. The stack data structure is used for implementation of the method. Basically, any point is chosen as the starting point [12]. In this study, the left upper corner is defined as the starting point and the lower right corner as the end point [12-13]. A grid which consists of n rows and m columns is defined in order to implement the method.

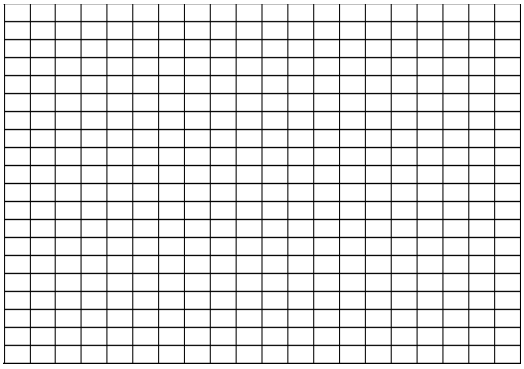


Figure 4. Segmentation in N × M dimension.

Since no cell has been visited yet, a random neighbor cell is selected. The cell information is added to the stack by removing the wall between these two adjacent cells. If there is a dead-end path, it is returned by using the cell information in the stack. All visited cells are marked. If it is not a dead-end, a path joining two adjacent cells is drawn. These processes are continued until all cells are visited [12]. An example of a maze obtained by using the DFS algorithm is shown in Figure 5.

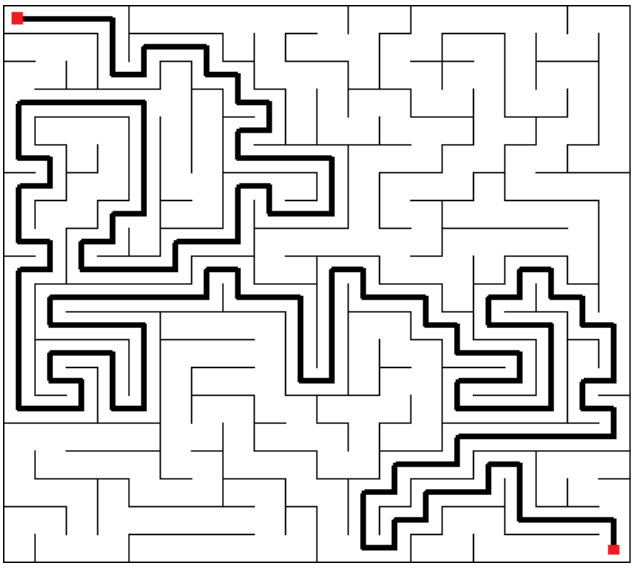


Figure 5. Example of a maze produced with DFS.

The neighbouring cells are randomly selected in the DFS algorithm. However, if the selection of neighbouring cells is chosen based on a random number generator (RNG) and a seed value (seed), the same labyrinth will always be obtained. The secret password which was determined in this study was used as the initial value for the RNG.

4. Advanced Encryption Standard

The robustness of the method used for attacks against encryption algorithms indicates the power of the relevant algorithm. The most fundamental reason for the emergence of the AES encryption algorithm is that the DES (Data Encryption Standard) encryption algorithm is vulnerable to attack. The DES is 56-bit key length, and was broken in 1997. Cryptographers Joan Daemen and Vincent Rijmen developed the Rijndael algorithm with 128, 192, 256-bit key length options [14-15]. This algorithm has been accepted as the data encryption standard in order to provide data security in electronic environment with the name of Advanced Encryption Standard (AES). AES still maintains its reliability today and is used in different fields for security in the world of informatics.

The Rijndael algorithm is realized by the cyclical processes. These cyclical processes are based on the key value that the algorithm has. The key is applied to the data by renewing after each cycle. The data are expressed as order. The order indices begin with a zero index. The algorithm uses the same key, when both encrypting 128-bit of data, and decrypting encrypted text. The 128-bit length data are divided into (4 × 4) matrices and included in the algorithm. Each element of this matrix consists of 8 bits and each row or column sum of bits has 32 bits. Each line of this matrix is called "word", and the matrix itself is called the "status matrix". The number of cycle increases with key length in AES encryption. Increasing the number of cycle allows the data to be more reliable. However, both the number of process and the memory space used increase. The key that generates the encryption is converted to the matrix called "the status". Since the AES encryption algorithm performs its operations on the status matrices, Data has to be converted to the most suitable form. The beginning of the encryption is done by making the summation of the status matrix belonging to the plain text and the status matrix belonging to the key [15-17].

5. Proposed Method

In the proposed method, a maze which is in size of image in which data can be hidden is created. The purpose of the method is to hide the text data to the lowest bits of the image pixels corresponding to the coordinates on the solution path of the generated maze by using a secret password. The DFS algorithm was used to obtain the maze and solution path. The data is encrypted with the AES algorithm before being hidden, in order that the method can be stronger. The secret key used for AES is also the starting point for the RNG used in the creation of maze. In this case, providing that the user key is kept secret the solution of the generated maze will not be known. The data hiding block diagram of the proposed method is shown in Figure 6.

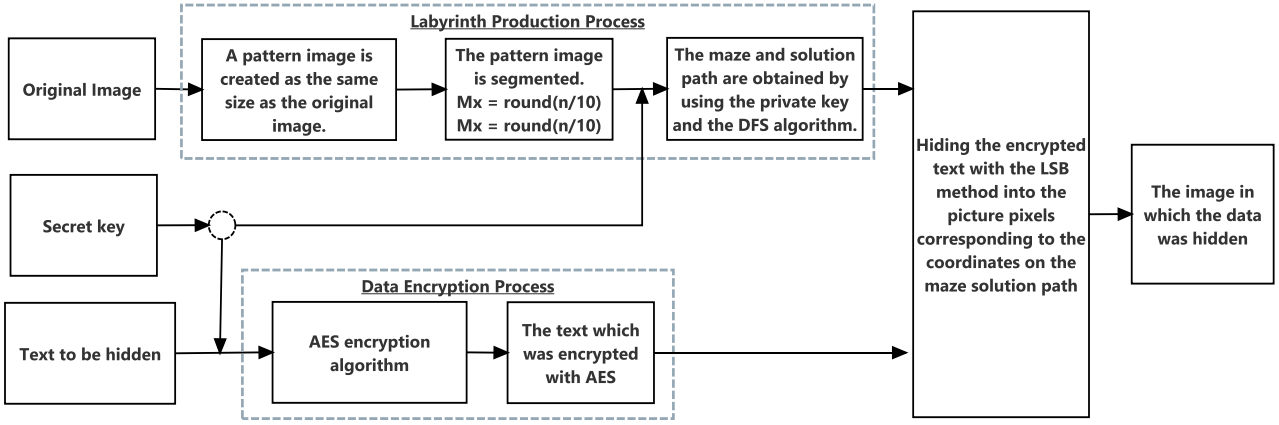


Figure 6. The data hiding block diagram of the proposed method.

The *L* pattern image who has a white background whose dimensions are $n \times m$ in a way that the original image is "*I*" and width *n* and height *m* of the image, was created. The *L* pattern image is the surface on which the maze is to be created. A grid is obtained on the *L* pattern image as shown in figure 4 before the maze is formed. The width (*Mx*) and height (*My*) of each cell in the grid are calculated by equation (1).

$$Mx = \text{round}(n/10)$$
$$My = \text{round}(m/10) \quad (1)$$

The maze is obtained by using the DFS algorithm with the aid of the personal key of the user. In this method, the center of gravity of the cell in the upper left corner as the starting point of the maze and the center of gravity of the cell in the lower right corner as the end point are selected. Different starting and end points can also be selected. The drawing of the maze solution path is made between

the coordinates corresponding to the center of gravity of each cell [18].

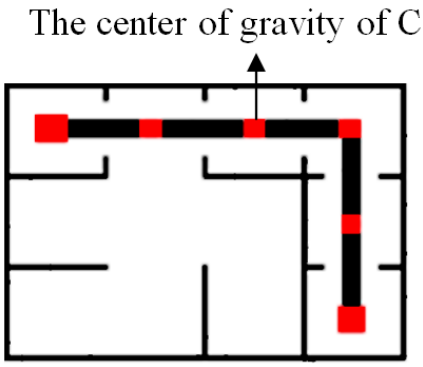


Figure 7. Presentation of center of gravity belonging to cell

The maze image and the original image are the same size. The data encrypted with AES are added to the lowest bits of the original image pixels corresponding to all coordinates between the gravity centers of cells on the pattern image. Binary equivalency of each encrypted character is taken before the addition of the data. Due to the structure of the LSB method, one character can be placed in eight pixels. Therefore, the number of characters (NC) that can be hidden in the original image is obtained by equation (2), where NPSP denotes the number of pixels on the solution path.

$$NC = NPSP/8 \quad (2)$$

In order to obtain the hidden data in the image, the maze and solution path are first obtained with the help of the secret key. The lowest bits are divided into eight groups by taking the binary values of the image pixels corresponding to the solution path coordinates. Every eight bits corresponds to a character. The equivalent character in the ASCII table were obtained by being taken the equivalents of those eight-each group in decimal system. The block diagram of obtaining secret data from the data hidden image is shown in Figure 8. The obtaining of the hidden original data is only possible if the private key is known.

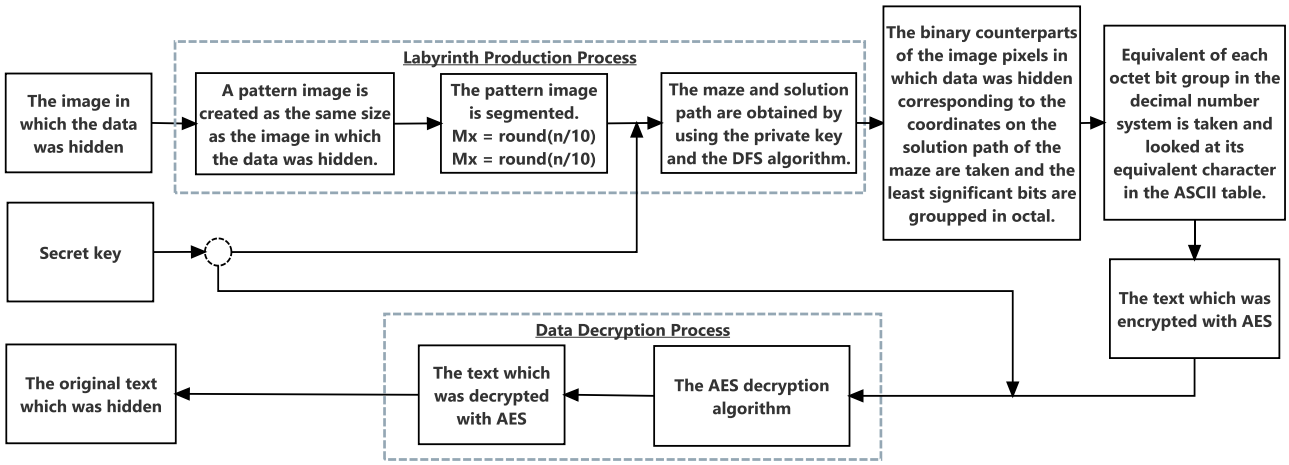


Figure 8. Block diagram showing the obtained hidden data.

6. Simulation and Security Analysis

In order to test the proposed method in a sample image, a colourful "baboon.bmp" image of 512 × 512 size and which is commonly used in image processing field is used. Assuming that private key

used with the application is “2a?4%”, the obtained maze and the result image in which data was hidden are shown in Figure 9.

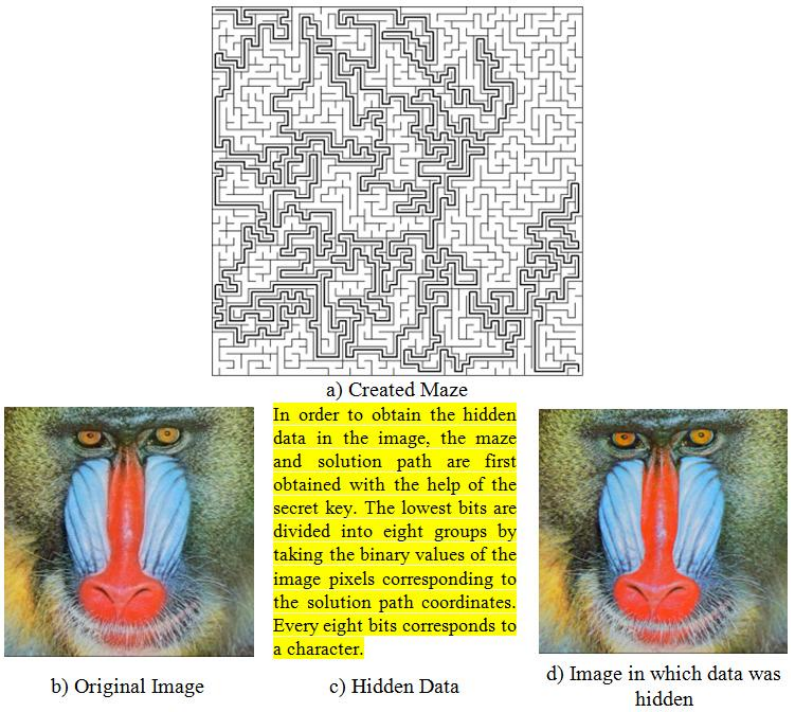


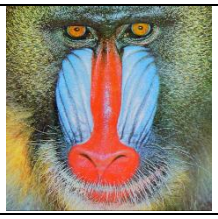
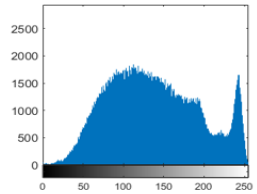
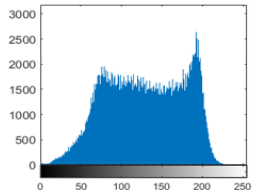
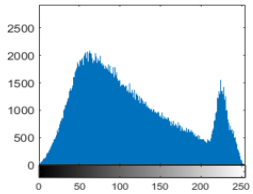
Figure 9. Example application images.

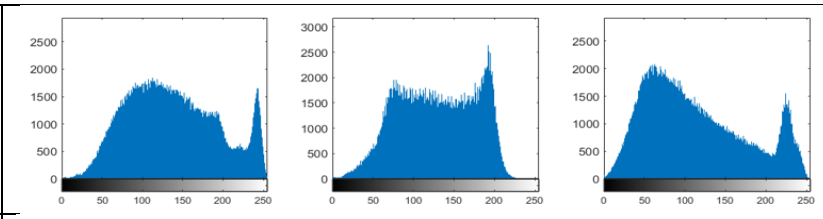
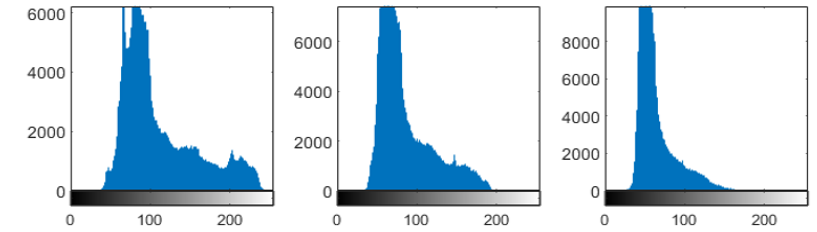
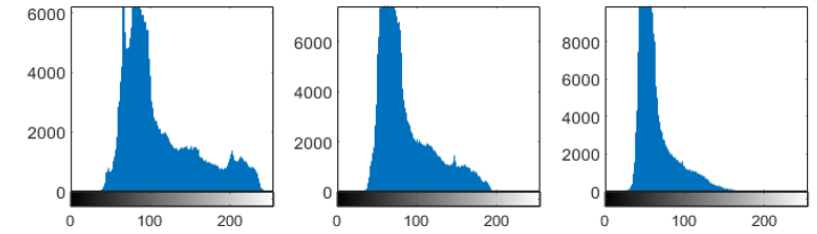
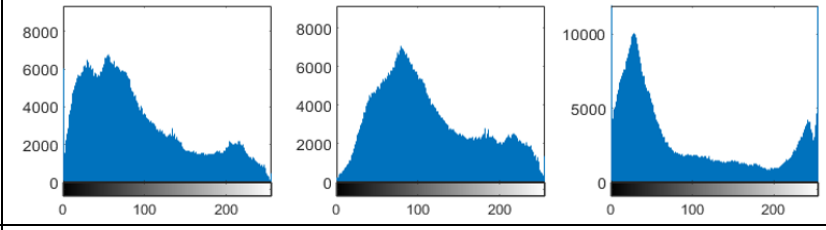
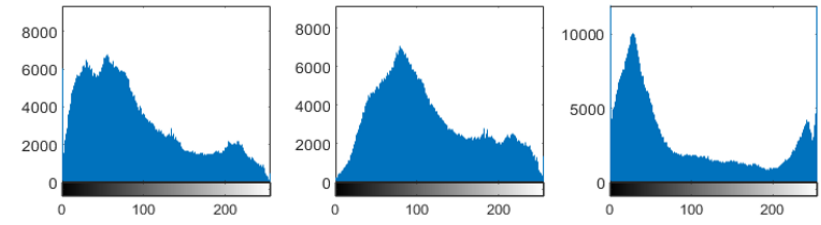
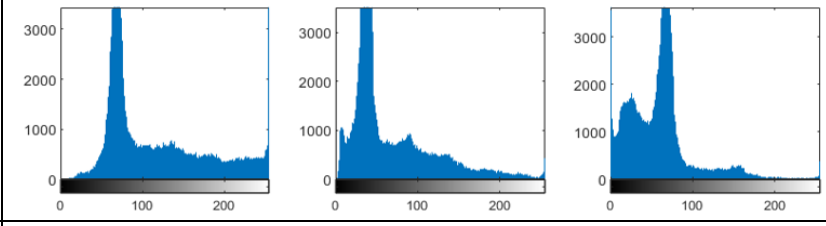
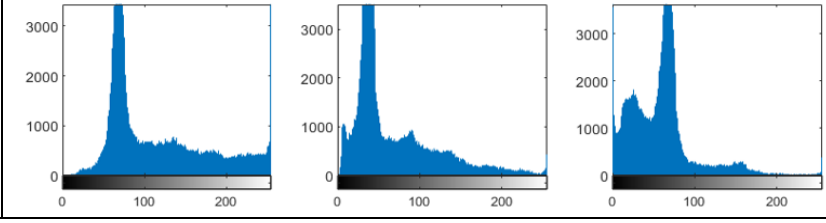
When the application was performed according to the user's secret key “2a?4%” it was seen that the number of pixels that can be hidden with the help of the obtained maze is 11241 and the number of characters that can be hidden is 1405. The security analysis of the proposed method is discussed below. For this purpose, the results were given in detail by being performed the histogram, correlation, entropy, structural similarity and time analysis.

6.1. Histogram Analysis

Histograms are used to show the distribution of gray level pixel values in the image [19]. Pixel distribution of the image in which the data is hidden is crucial. For this reason, the change in the pixels should be at a level that cannot be perceived visually. The original image histogram and the image histogram, in which the data was hidden, should be very close to each other in order to prevent the images in which the data was hidden from being exposed to attack. The histogram distributions on the colored images of the proposed method are shown in Table 1.

Table 1. Histograms of the original, and the images in which the data was hidden.

			Histograms		
			Red Histogram	Green Histogram	Blue Histogram
baboon.bmp	Input Image				

	Output Image		
el-cezeeri.jpg	Input Image		
	Output Image		
scene.bmp	Input Image		
	Output Image		
peppers.png	Input Image		
	Output Image		

When the image histograms in Table 1 in which the data was hidden are examined, it is seen that there are no differences as the level of human eye can perceive in the original image histograms. Therefore, it can be said that the proposed method is successful and resistant against to attacks.

6.2. Correlation Analysis

Statistical correlation is a measure which is used to measure the power of a linear relationship between two random variables [20]. The correlation coefficient, x and y which are a random variables consisting of n elements are calculated by equation (3).

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)D(y)}} \tag{3}$$

Here, the parameters given in Equation (3) are obtained with the help of equations of (4), (5) and (6).

$$cov(x,y) = \frac{1}{n}\sum_{i=0}^{n}[x_i - E(x)][y_i - E(y)] \tag{4}$$

$$D(x) = \frac{1}{n}\sum_{i=0}^{n}[x_i - E(x)]^2 \tag{5}$$

$$E(x) = \frac{1}{n}\sum_{i=0}^{n} x_i \tag{6}$$

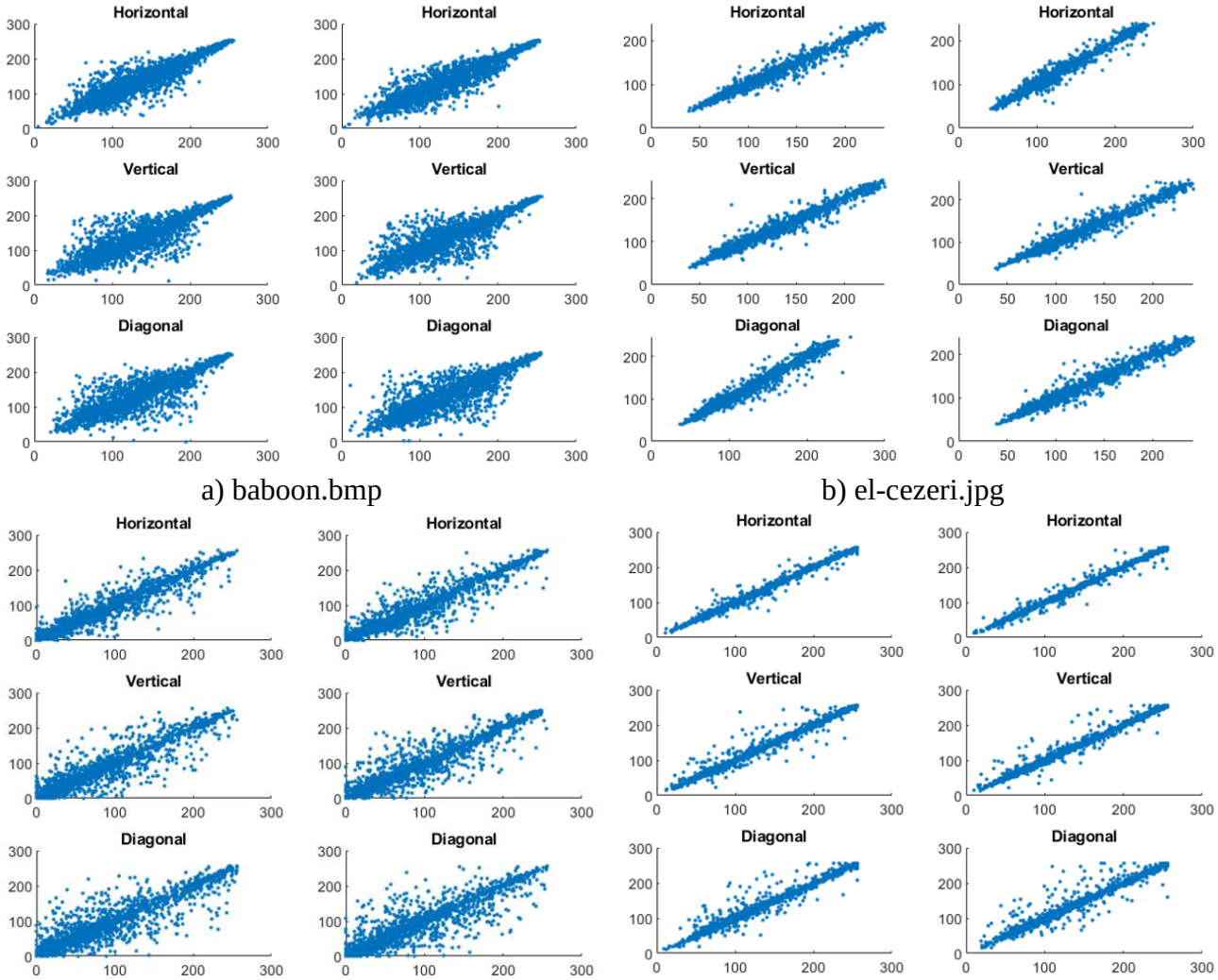
In this study, 4000 random pixels adjacent (horizontal, vertical, and diagonal) were selected for each of the original images and images in which the data was hidden in the calculation of the correlation of pixels in the image in which data was hidden. The results of the original image and of the image in which the data was hidden are given in Table 2.

Table 2. The results of the original images and its images in which the data was hidden.

Images	Correlation Aspects	Original Image	The image in which data was hidden
baboon.bmp	Horizontal	0.9231	0.9231
	Vertical	0.8660	0.8660
	Diagonal	0.8543	0.8543
el-cezeri.jpg	Horizontal	0.9859	0.9859
	Vertical	0.9842	0.9842
	Diagonal	0.9800	0.9800
scene.bmp	Horizontal	0.9524	0.9524
	Vertical	0.9323	0.9323
	Diagonal	0.9018	0.9018
peppers.png	Horizontal	0.9943	0.9943
	Vertical	0.9903	0.9903
	Diagonal	0.9854	0.9854

There is a strong and linear relationship between neighbouring pixels in any image. A high correlation coefficient is characterized as close to +1 and -1. In other words, that the correlation coefficient is very close to -1 and +1, means that the relationship between the pixels is strong [21]. On the other hand, that the correlation coefficient is close to 0, means that the relationship between the pixels is very weak. In this context, when Table 2 is considered, it is seen that the original image correlation coefficients and the image correlation coefficients in which data was hidden are the same at randomly selected 4000 pixels. Thus, along with no indication is detected that the data are hidden in the image but it is also prevented the image to become open to attacks. For better understanding of the results, the correlation distributions between neighbouring pixels are shown in Figure 10. The correlation distributions of the original image in the left column, while in the right

column the image in which the data was hidden, are given. When paid attention to these distributions, it was observed that they are small enough not to be perceived by eye.



a) baboon.bmp b) el-cezeri.jpg

c) scene.bmp d) peppers.png

Figure 10. Correlation distributions of original and data hidden images between the neighboring pixels.

6.3. Entropy Analysis

Entropy is defined as random occurrence and disorder in a system. The entropy of a image is calculated by equation (7).

$$H(m) = \sum_{i=0}^{MxN-1} P(m_i) log_2 \frac{1}{P(m_i)} \quad (7)$$

Where $P(m_i)$ represents the probability states of each m_i pixel in an image and M × N is the total number of pixels. In a gray level image $(m_0 = 0, m_1 = 1, \ldots m_{255} = 255)$, the probabilities of each gray value are obtained from the histogram of the image. The ideal entropy value is 8 for a random image and the entropy value is much lower than 8 for images with a lower randomness ratio. If the entropy value is much lower than 8, for example close to 0, it is predicted that there may be a threat against to security [21]. When thinking that there are messages in the images in which the data was hidden, the ideal entropy value is expected to be 8. Table 3 shows the entropy test results of the original image, and the image in which data was hidden.

Table 3. The results of entropy test.

		Entropy	
Image	**Size**	**Original image**	**The image in which data was hidden**
baboon.bmp	512×512	7.7624	7.7624
el-cezeri.jpg	603×696	6.9610	6.9610
scene.bmp	740×1110	7.8144	7.8145
peppers.png	384×512	7.3785	7.3786

When Table 3 is considered, it is seen that the entropy value of the image in which data was hidden is very close to 8 and is identical to the original image one to one. Therefore, it can be said that the proposed method is very resistant against to attacks.

6.4. Structural Similarity Test

Structural Similarity (SSIM) is a method used to measure the similarity between two images. It is also used to measure the quality difference between the original and processed images. When X is the original image and y is the processed image, μ_x and μ_y are the mean of x and y, σ_x^2 and σ_y^2 are the variances of x and y, the σ_{xy} is covariances of x and y and the C_1 and C_2 represent the constant variables using for balancing the images, the SSIM is obtained by equation 8 [22].

$$S(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{8}$$

That the SSIM value is close to or equal to 1, means that the original and processed image are structurally very similar to each other; That it is equal to or very close to 0, means that there is little or no structural similarity between the images. The structural similarity test results between the original image and the image that data was hidden are shown in Table 4.

Table 4. Structural similarity test results.

Image	**The similarity ratio between the original image and the image in which the data was hidden**
baboon.bmp	100%
el-cezeri.jpg	100%
scene.bmp	100%
peppers.png	100%

When Table 4 is taken into account, it is seen that the similarity ratio between the original image and the image in which data was hidden is 100%, and it is resistant against to attacks.

7. Conclusion

In this study, it was presented that a steganography method based on AES encryption algorithm, and creating of maze in order to hide data into the image and obtain hidden data. The proposed method was tested with the test methods available in the literature and it was seen that the performance ratio is high. In the method, when the starting point (upper left) and end point (right lower) of the maze are considered different maze and solution paths are obtained depending on the user's personal key. It is also clear that the maze key space that will be obtained will be much wider when the start and end points are determined by the user. In such a case, that the hidden text can be obtained from the image in which data was hidden is possible only if the user's personal key is known. It was supported by the AES encryption algorithm, which is accepted in the literature, in

order to make it much more difficult for third parties to obtain hidden texts. In the light of the information getting from the experimental results, it is possible that the proposed method can be used to hide text into image.

References

[1] Hui-Lung L., Chia-Feng L., Ling-Hwei C., "A perfect maze based steganographic method", Journal of Systems and Software, 2010,83(12), 2528-2535.

[2] Bradley H., Craigie W.A., Onions Ct., Ahm J., "The Oxford English Dictionary: Being a Corrected Re-issue", Clarendon Press, Oxford, (1970).

[3] Johnson N.F., Jajodia S., "Exploring Steganography: Seeing the Unseen", IEEE Computers, 1998,31(2), 26-34.

[4] Mohd B.J., Abed S., Al-Hayajneh T., Alouneh S. "FPGA hardware of the LSB steganography method", 2012 International Conference on Computer, Information and Telecommunication Systems (CITS), Amman, 1-4, (2012).

[5] Sharma S., Gupta A., Trivedi M.C., Yadav V.K. "Analysis of Different Text Steganography Techniques: A Survey", 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, 130-133, (2016).

[6] Anderson R.J., "Information Hiding: First International Workshop", Springer-Verlag, Cambridge, U.K., (1996).

[7] Petitcolas F.A.P., Anderson R.J., M.G, K., "Information hiding-a survey", Proceedings of the IEEE, 1999,87(7), 1062 - 1078.

[8] Abboud G., Marean J., Yampolskiy R.V. "Steganography and Visual Cryptography in Computer Forensics", Fifth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE), Oakland, CA, 25-32, (2010).

[9] Polap D. "Designing mazes for 2D games by artificial ant colony algorithm", Proceedings of the International Symposium for Young Scientists in Technology, Engineering and Mathematics. SYSTEM 2015, Catania, Italy, 63-70, (2015).

[10] Lee H.L., Lee C.F., Chen L.H., "A perfect maze based steganographic method", Journal of Systems and Software, 2010,83(12), 2528-2535.

[11] Sukumar T., Santha K.R., "Maze Based Data Hiding Using Back Tracker Algorithm", International Journal of Engineering Research and Applications (IJERA), 2012,2(4), 499-504.

[12] Kozlova A., Brown J.A., Reading E. "Examination of representational expression in maze generation algorithms", IEEE Conference on Computational Intelligence and Games (CIG), Tainan, Taiwan, 532-533, (2015).

[13] Sun, Y., Wang, J., Duan, X. "Research on path planning algorithm of indoor mobile robot", International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), Shengyang, China, 1108-1111, (2013).

[14] Dumane A.R., Narole N.G., Wanjari P. "Design of advanced encryption standard on soft-core processor", World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave), Coimbatore, India, 1-5, (2016).

[15] Pendli V., Pathuri M., Yandrathi S., Razaque A. "Improvising performance of Advanced Encryption Standard algorithm", Second International Conference on Mobile and Secure Services (MobiSecServ), Gainesville, FL, USA, 1-5, (2016).

[16] Garcia D.F. "Performance Evaluation of Advanced Encryption Standard Algorithm", Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), Sliema, Malta, 247-252, (2015).

[17] Güvenoğlu, E., "Resim Şifreleme Amacıyla Dinamik S Kutusu Tasarımı İçin Bir Yöntem", El-Cezerî Fen ve Mühendislik Dergisi, 2016,3(2), 179-191.

[18] Guvenoglu, E., "Maze Based Image Encryption Algorithm", International Research Journal of Engineering and Technology (IRJET) 2015,2(8), 1578-1585.

[19] Naveenkumar S.K., Panduranga H.T., Kiran. "Triple image encryption based on integer transform and chaotic map", International Conference on Optical Imaging Sensor and Security (ICOSS), Coimbatore, India, 1-6, (2013).
[20] Shannon C.E., "A Mathematical Theory of Communication", The Bell System Technical Journal, 1948,27, 379–423,623–656.
[21] Munir R. "Security analysis of selective image encryption algorithm based on chaos and CBC-like mode", 7th International Conference on Telecommunication Systems, Services, and Applications (TSSA), Bali, Indonesia, 142-146, (2012).
[22] Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P., "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, 2004,13(4), 600-612.