*THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE*

# A Steganography Model Based on Least Significant Bit Algorithm for Hiding Text Data in Digital Files

**Kultany Hillary**
Student, School of Computing and Information Technology,
Jomo Kenyatta University of Agriculture & Technology, Nairobi, Kenya
**Dr. Calvins Otieno**
Lecturer, School of Computing and Information Technology,
Jomo Kenyatta University of Agriculture & Technology, Nairobi, Kenya
**Wilson Cheruiyot**
Professor, School of Computing and Information Technology,
Jomo Kenyatta University of Agriculture & Technology, Nairobi, Kenya

***Abstract:***
*Steganography in the modern-day sense of the word usually refers to information or a file that has been concealed inside a digital file. With the rise of the Internet, one of the most fundamental factors is the security of information. It involves communication taking place without being known, by providing a method of writing hidden information in such a way that no one, apart from the sender and intended recipient, suspects the existence of any information/obscurity. This is the reason why in the encoding process, steganography tends to replace this redundant data with the secret message. Thus, this paper was enhanced and a new LSB algorithm developed in which it came out with a better result for the quantity of data hidden (Stego image). The paper was evaluated by using Least Significant Bit (LSB) algorithm, providing promising results in terms of quantity and image quality.*

***Keywords:*** *Least Significant Bit (LSB), Steganography, encoding, obscurity, Stego image, digital file*

## 1. Introduction

Computer network is a determining factor for the performance of a modern company. Managing networks is very important as managing any other aspect of the company's performance. Network security is an organization's strategy and provisions for ensuring the security of its assets and all network traffic is supplied to better meet the organization's set of security mechanisms and provide safe and reliable network that is called a secure computer network (Modiri and Arbasi, 2013). Security is a series of security dimensions designed to express and manage specific aspects of network security (Mashayekhi, Ashoorian, and Riahi, 2008). Security thinking in network is to achieve three important factors that together constitute the security triangle. These include confidentiality, integrity and being constantly available. The three basic principles form the information security in the network or outside it so that all necessary measures taken for the security of the network or the equipment made, are all due to the need to apply these three parameters in the maintenance and exchange (Swati and Mahajan, 2012).

### *1.1. Related Work*

Steganography concept was earlier used by Greece to alert the invading o Persia by Herodotus in 440 BC. A message was covered by wax on a tablet. Technology changed thus majority of text data are being hidden and relayed through digital files of images, audio and video. Several techniques are in place capable for encoding data in an image.
These techniques have been categorized based on their algorithms into two: (1) transform domain based; (2) spatial domain based (Jing and Thanh, 2010). Spatial domain-based steganography method is used by the LSB algorithm (Nithyanandam et al., 2011). LSB is most widely used technique for hiding data (Wien and Tung, 2012). The techniques available are based mainly on LSB where the message bits are changed directly by the cover file. Several LSB Steganography numbers of techniques have been proposed (Kousik et al., 2012; Swati and Mahajan, 2012).

In the study (Ki and Kee, 2012) tried working on increasing the payload capacity and image quality in steganography using techniques such as image interpolation and edge detection. Kanso et al., worked using steganalytic approach such as PSNR test, Histogram test, RS attack to test reliability of his algorithm used for steganography. In the study (Taruna, 2014) suggested a technique using keyless randomization to encode sensitive data in variable and multiple LSB's. Text steganography is a process in a steganography that hides the message behind other cover text file. Text steganography is not very popular cause text files have very less redundant data. Images provide an excellent medium for data hiding. The lesser constraints there are on how much information it can encode before it becomes a suspect the more comprehensive an image is. The JPHide/JPSeek package uses the coefficients in a JPEG to conceal secret information.

In the paper (Channali and Jadhav, 2009) present that the secret information can't be detectable in the steganography, because, the cover file hides the secret data which has no trigger for eavesdropper's suspicion to discover. Capacity, undetectability, robustness, security, resistance and invisibility against attacks are the main core features in steganography. The steganography is a form of security through ambiguity, by hiding the text data file (stego object) in another object (host file) (Zaidoon, Zaidan, Zaidan and Hamdan, 2010). The physical properties of the file should have minimal difference before and after text data file encoding. The text, image, video and music could be the cover object for steganography.

In the study (Marghny, AL-Aidroos and Bamatraf, 2012) suggested a technique to encode secret message into the original image by a dynamic LSB substitution scheme. This method is achieved by utilizing the similarity in the smooth area and not the edge area as in the simple techniques, with the use of the LSB substitution methods as a fundamental stage. This method preserves the quality of stego image with the increase of the data capacity. Furthermore, (Marghny et al., 2012) proposed an efficient steganographic method to encode message over gray scale images. The human eye is generally perceptive to the change in the smooth area than the edge area using pixel value difference, as well as using the LSB substitution method as a fundamental stage making this method more based. This method achieved the visual quality and more security with increased embedding capacity. Marghny et al. (2012) proposed a dynamic LSB substitution technique by dividing the cover image into parts and smooth areas. This technique can encode large amount of data as well as imperceptibility of stego image based on the pixel value differencing for secret communication. Experimentation results show that the quality of the stego image is satisfactory with the proposed method. Moreover, statistical analysis being carried can be resisted by steganalysis systems.

In the study, (Chen and En 2009) Suggested that the image quality is improved by hiding more information in the edge portions while maintaining the same encoding capacity. This advantage results because the human eyes rarely perceive trivial difference in the edge regions. With various individual demands, the encoding capacity is adjustable accordingly. The proposed method is more secure with the addition of the improvement on image quality. Steganography is a widely used technique to encode tremendous capacity of secret data while maintaining its visual quality. Least significant bit substitution makes changes to the cover image by simply substituting secret bits for the LSBs of the cover pixel being the most known techniques applied in steganography. This paper proposes a novel data hiding method based on LSB in which binary secret information can directly be encoded into the cover image. The encoding capacity can be enhanced as much as possible through encoding 2 secret bits into each cover pixel by modifying 3 LSBs with the guidance of a reference table. Experimental results confirm the reliability of the proposed method outperforming the related schemes in terms of embedding capacity with visual quality (Liu, Chin and Tzu, 2017).

## 2. Methodology

### *2.1. The Research Designs*

The paper adopted an experimental research method to measure the effect of using selected varied and random pixels during the encoding process on imperceptibility and hiding capacity. A notable advantage of experimental research is the fact that it enables other researchers to easily replicate the experiment and be able to validate the results. It is therefore considered an accurate method of research (Shuttle worth, 2011), as the researcher can effectively establish a causal relationship between variables by manipulating independent variable(s) to assess the effect upon dependent variable(s).

### *3.2. Steganography Framework*

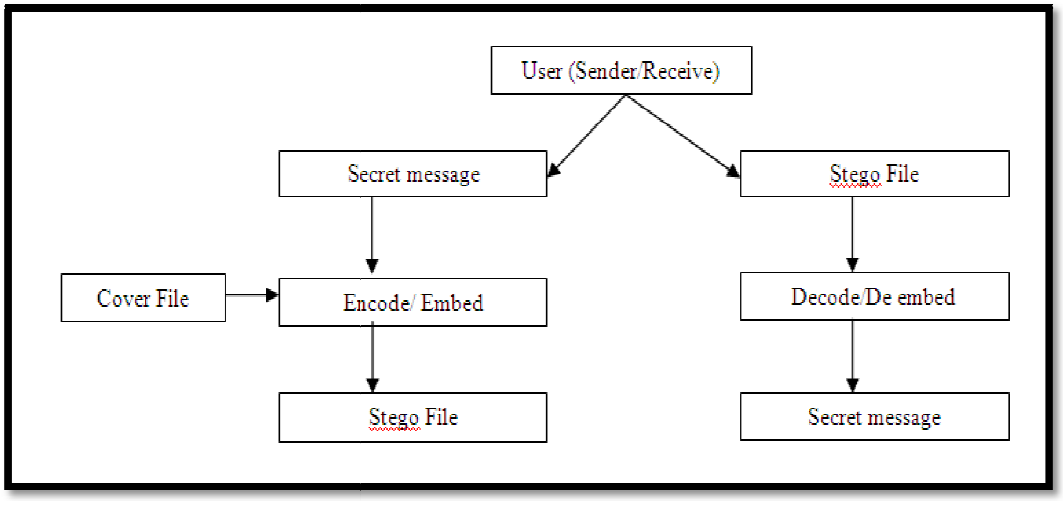


*Figure 1: Proposed Steganography Model*

*3.3. Description*

- Step 1: The sender first uses LSB algorithm for encoding the secret file message.
- Step 2: The sender sends data text file document to the encoding module for data encoding.
- Step 3: In encoding module, the data is encoded into carrier file that's the stego file.
- Step 4: The stego file is send to the receiver using a data transmission medium. E.g. Web or e-mail.
- Step 5: The recipient receives the stego file and places the file in the decoding module.
- Step 6: The decoding module decodes the original text document using the LSB algorithm decoder and retrieves the original message.
- Step 7: Acknowledgement is relayed to the sender

**4. Least Significant Bit Algorithm**

The most and reliable common method used for data hiding in images is the Least Significant Bit algorithm in which the LSB of the pixel values is substituted with the secret data to be encoded in a binary way. LSB coding is the best, effective and reliable way to encode information in a digital text file. This technique works by substituting some of the important information in a given pixel with information data file from the image. First, the sender uses LSB algorithm for encoding the secret file message which it becomes delivered to the encoding module for data encoding then through a carrier file it becomes delivered to the receiver and places it on decoding module which decodes using the LSB algorithm decoder to the original text document and image.

The data hiding and the data extracting will be done in modules as mentioned.

*4.1. Data Encoding Module*

It's the initial stage where the sender sends the information as well as the image file which act as the carrier image to transfer data to the destination. In this module the text message file will be encoded based on the principle of least significant bit algorithm. After encoding it becomes transmitted to the receiver.

The encoder algorithm is as given below:

- Start
- for i = 1, ..., len (msg) do
- p = LSB(pixel of the image)
- if p != message bit then
- pixel of the image = message bit
- end if
- end for
- Stop

*4.2. Data Transmission Module*

The stego image that the text file is encoded becomes transmitted easily acting as a carrier file for the secret information.

*4.3. Data Decoding Module*

The receiver receives the carrier image which is the stego file and sends it to the decoding module where the same least significant bit algorithm is implemented for decoding the least significant bits from the image and merge in an order to come up with the original message bits. After decoding it becomes accessed as the original text file and initial image.

The decoder algorithm is:

- Start
- for i = 1, ..., len(image string) do
- message string = LSB (pixel string of the image)
- end for
- Stop

**5. Validation and Results**

A series of experiments have been conducted for comparing stego image, cover results with different grayscale images values.

*5.1. Validation Window*

C#.Net language was used during the implementation phase. The screen shots of the validation window are as shown below.

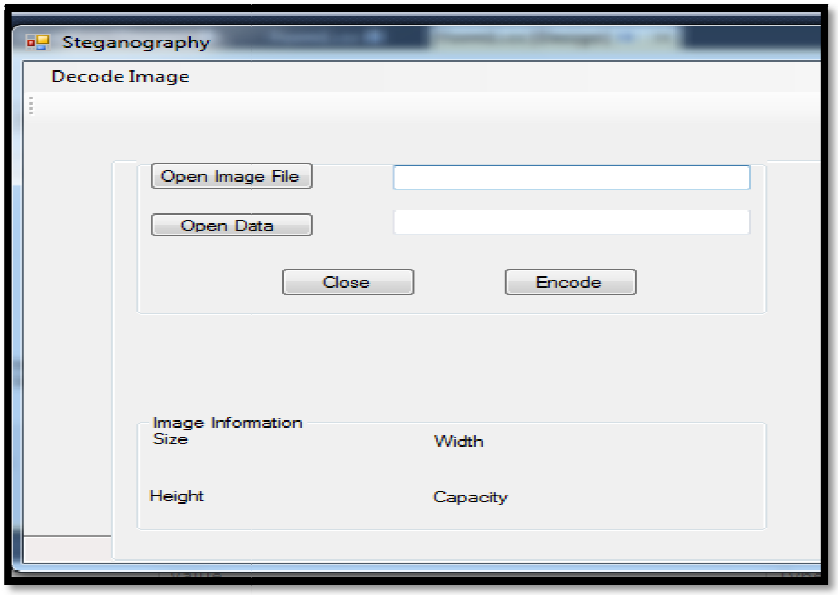


*Figure 2: Validation Window: Encoding form*

In fig. 2 above, encoding window is as shown. In this window, in order to hide the information, you select the file which is to be secured from browse button. When the image is selected, it automatically displays image information in the bottom left. When you press encode button text file is encoded and hidden into image and saved into your preferred location.

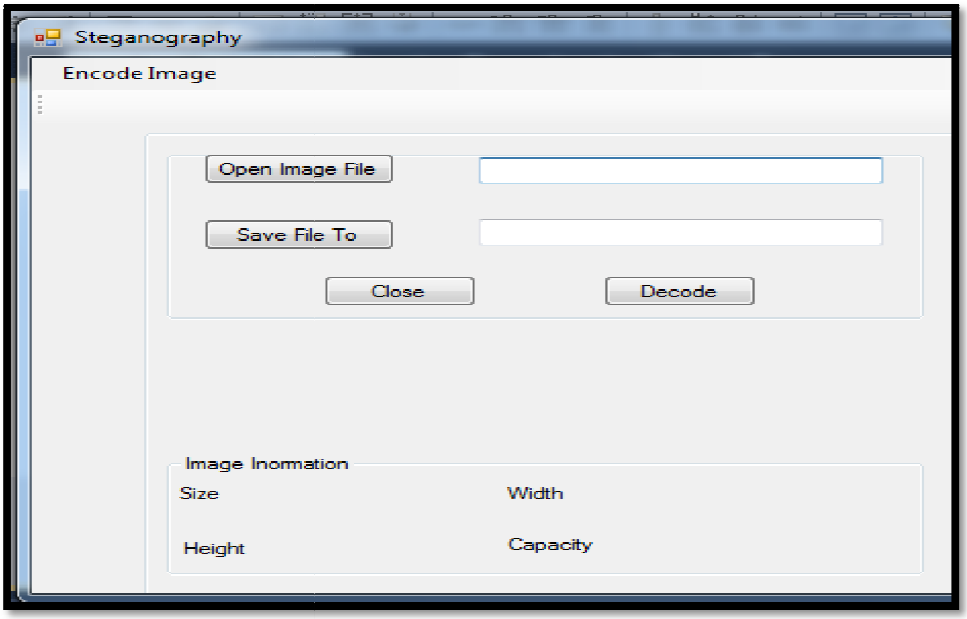


*Figure 3 Validation Window: Decoding form*

In fig. 3 above, decoding window is as shown in which first of all you select the image which is encoded and has hidden information file by using browse button. Then preferred location is selected where the original information is to be saved. In this window there is a button decode. When decode button is pressed it saves original file and image into preferred location.
Fig. 4 and Fig. 5 below show the resulted images.

*5.2. Cover Image and Stego Image*

The cover image and the stego image for the different standards grayscale images are as shown in (Fig. 4).

| Input Size | Output for LSB | Exp. Maximum Output for LSB |
|---|---|---|
| 221*166 | 81.5kb | 94.66kb |

*Table 1: Result for Determining Maximum Size That Can Be Hidden*

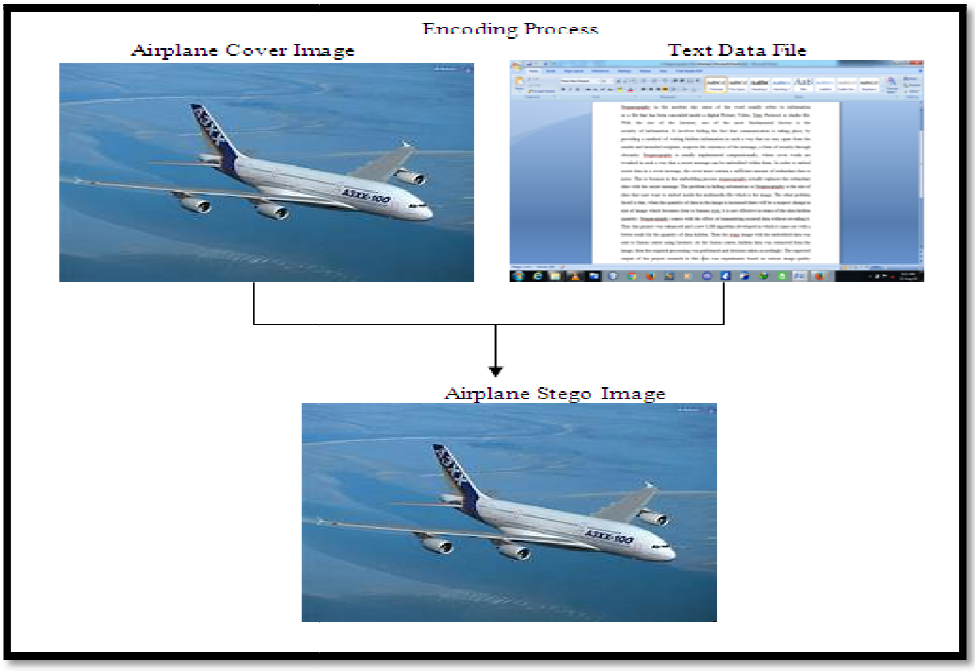


*Figure 4: Encoding Airplane Image with Text File Data to Stego Image*

The Airplane image is encoded with the text file data, it outputs an Airplane stego image as illustrated in Fig. 4. The output size for LSB algorithm is 81.5kb, while the maximum size is 94.66kb, from the value size difference it's clear that the Algorithm can hide more text data. Thus, when exposed both Airplane Stego image and original image looks alike which does not show any distortion. The stego image will not cause attention to itself. Hence being transmitted to the recipient without being detected of any information.

*5.3. Stego Image, Original File and Cover Image*

The stego image and the original cover image for the different standard grayscale image is as shown in (Fig. 4).

| Input Size | Maximum Output for LSB | Output for LSB |
|---|---|---|
| 221*166 | 94.66kb | 81.5kb |

*Table 2: Result for Determining Original Output*

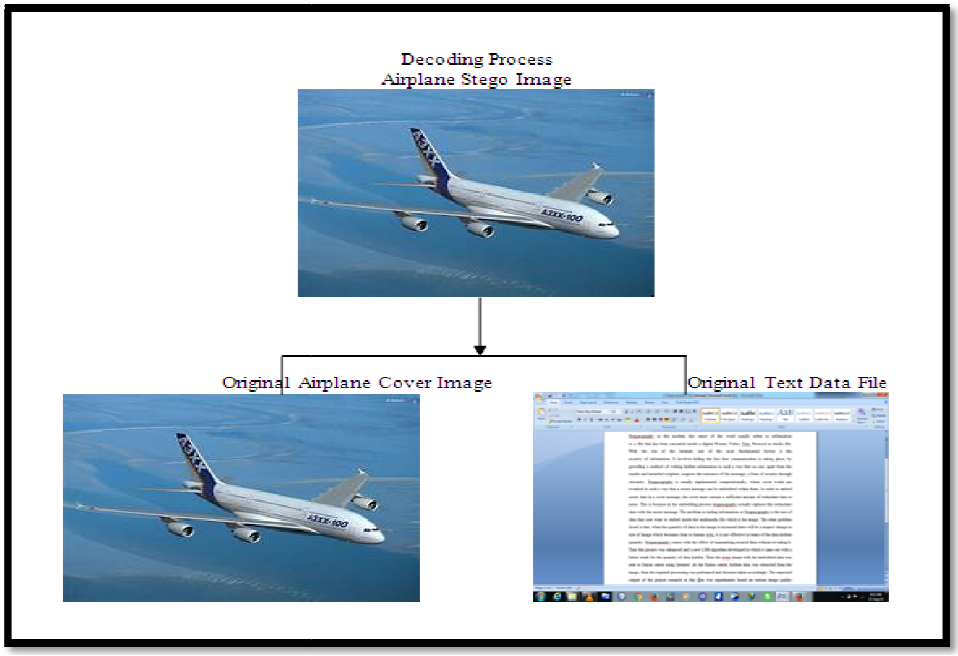


*Figure 5: Decoding Airplane Stego Image to Original Text File Data and Cover Image*

When the Airplane Stego image is decoded, it outputs an Original Airplane Cover Image and Text Data File as illustrated in Fig. 5 as shown above. Thus, when exposed both Airplane Stego-image and original image looks alike which does not show any distortion. The stego image will not cause attention to itself. Hence being transmitted to the recipient without being detected of any information.

As illustrated in the Table 3, the paper observed that the higher the size of the image, the higher the quantity of data that it can hide i.e. 1024pixel by 768pixel payload capacity is very high. But however, this may not be true in a case where the image has a short width and height i.e. 550pixel by 550pixel which payload capacity is low. This is because the edges based data embedding algorithm require an image with adequate height and width because the data quantity embedded determines by bitmap pixels. In this situation, when selecting the image to be used, then two considerations need to be considered i.e. the height and width of the image.

| Image Size(KB) | Bitmap Size(KB) | Width(Pixels) | Height(Pixels) | Height*Width | Payload (KB) |
|---|---|---|---|---|---|
| 4.88 | 4.5 | 221 | 166 | 36686 | 94.66 |
| 13.2 | 5.2 | 235 | 177 | 41595 | 107.85 |
| 33.5 | 32.4 | 320 | 320 | 102400 | 264.99 |
| 125 | 124.2 | 550 | 550 | 300000 | 786.32 |
| 238 | 233.2 | 800 | 600 | 480000 | 1246.86 |
| 239 | 236.7 | 1024 | 768 | 786432 | 2045.99 |

*Table 3: Payload Capacity Distribution*

**6. Discussion**

However, with the rise of the Internet, one of the most fundamental factors is the security of information. It involves pretending as if the communication is not taking place, by providing a method of hiding data text file information in such a way that no one, apart from the intended recipient and the one sending, suspects the existence of any information/obscurity. The paper was able to point out several aspects pertaining to hiding text data in digital files. The results were closely tied to the objectives as validity testing was in the affirmative since the two modules for encoding and decoding was able to secure a result of "pass", meaning it complies with the desired results. All the tests carried out above were considered to be above average. The output size for LSB algorithm was 81.5kb, while the maximum size was 94.66kb, from the value size difference it's clear that the Algorithm can hide more text data. The results from the payload capacity table 3 further observed that the higher the size of the image, the higher the quantity of data that it can hide i.e. 1024pixel by 768pixel payload capacity is very high. But however, this may not be true in a case where the image has a short width and height i.e. 550pixel by 550pixel which payload capacity is low. This is because the edges-based data encoding algorithm require an image with adequate height and width because the data quantity encoded determines by bitmap pixels. In this situation, when selecting the image to be used, then two considerations need to be considered i.e. the height and width of the image. This is because in the encoding process least significant bit algorithm actually replaces some of the information data in a given pixel with the secret information data file from the image. Thus, this paper was enhanced and a new LSB algorithm developed in which it came out with a better result for the quantity and quality of data hidden (Stego image).

**7. Conclusion**

The paper sought to enhance and develop a new LSB algorithm in which it came out with a better result for the quantity of data hidden (Stego image). The conclusions were closely tied to the objectives as validity testing was in the affirmative since the two modules for encoding and decoding was able to secure a result of "pass", meaning it complies with the desired results. The paper was evaluated by using Least Significant Bit (LSB) algorithm, providing promising results in terms of quantity and image quality.

**8. References**


i. Channali, S. and Jadhav A. (2009). Steganography an art of hiding data, International Journal on computer science and engineering, 3, 2009, 137-141.
ii. Chen, P. and En W. W. (2009). "A Modified Side Match Scheme for Image steganography", International Journal of Applied Science and Engineering no 7, 1: 53-60.
iii. Jing, M.G. and Thanh, N. L. (2010) "Secret Communication Using JPEG Double Compression", Signal Processing Letters, IEEE, Vol. 17, Issue No. 10, Pages No. 879 – 882.
iv. Kanso, A. and Hala S. Own. Steganographic algorithm based on a chaotic map. Communication Nonlinear Science Numerical Simulation, 17, pp 3287– 3302.
v. Ki, H. J. and Kee, Y. Y. (2012). Data hiding using edge detector for scalable images, Springer.
vi. Kousik, D., Mandal, J. K. and Paramartha D. (2012). "Hash Based Least Significant Bit Technique for Video Steganography (HLSB)", International Journal of Security, Privacy and Trust Management (IJSPTM), Vol. 1, Issue No. 2.
vii. Liu, Y. C., Chin, C. C. and Tzu, Y. (2017). A Revisit to LSB Substitution Based Data Hiding for Embedding More Information. 11-19. 10.1007/978-3-319-50209.

viii. Marghny, M. H., AL-Aidroos N. M. and Bamatraf, M. A. (2012). Innovative Multi-Level Secure Steganographic Scheme based on Pixel Value Difference, International Journal in Foundations of Computer Science & Technology, vol. 2, no. 6, pp. 1-13.
ix. Mashayekhi, N., Ashoorian, M., Riahi M.N.(2008). Providing the security matrix as a layer in NGN networks, the Third National Conference on Information and Communication Technology, Tehran.
x. Modiri, N. and Arbasi, H.(2013). Providing multiple layers to increase the layered e network security. The first national conference on new approaches in computer engineering and information retrieval. Gilan.
xi. Nithyanandam, P., Ravichandran, T., Santron N. M. and Priyadarshini, E. (2011). "A Spatial Domain Image Steganography Technique Based on Matrix Embedding and Huffman Encoding", International Journal of Computer Science and Security (IJCSS), Vol. 5, Issue No. 5.
xii. Shuttleworth, M. (2011). Experiment Resources. Accessed: September 25, 2011. [Online] Available: http://www.experiment-resources.com.
xiii. Swati, T. and Mahajan, R. P. (2012). "A Secure Image Based Steganographic Model Using RSA Algorithm and LSB Insertion", International Journal of Electronics Communication and Computer Engineering (IJECCE), Vol. 3, Issue No. 1.
xiv. Taruna, S. D. (2014). Message Guided Random Audio Stegano graphy Using Modified LSB Technique. International Journal of Computers & Technology. 12(5): 3464–8.
xv. Wien, H., and Tung, S. C. (2012). "A Novel Data Embedding Method Using Adaptive Pixel Pair Matching", IEEE Transactions on Information Forensics and Security, Vol. 7, Issue No. 1, Pages No. 176 - 184.
xvi. Zaidoon K. A., Zaidan, A. A., Zaidan, B. B. and Hamdan. O. A. (2010). Overview: Main fundamentals for Steganography, Journal of computing, Vol. 2, Issue.3, 2010,158-165.