Khulna University Studies **Volume 14 (1 & 2) : 95-106 : January-December 2017**

SCIENCE AND ENGINEERING

SECURING BANGLA TEXT COMMUNICATION USING IMAGE STEGANOGRAPHY WITH DYNAMIC SUBSTITUTION IN IOT ENVIRONMENT

Anupam Kumar Bairagi, Saikat Mondal* and Debashish Chakroborti

Computer Science and Engineering Discipline, Khulna University, Khulna9208, Bangladesh

KUS: 16/30:010816

Manuscript submitted: August 01, 2016 Accepted: November 07, 2017

Abstract: Privacy and security of information are the prime concerns of today's internet users. They want to get better and reliable services without sacrificing privacy from the current paradigm like the Internet of Things (IoT). In order to ensure IoT as a truthful service platform for the vast users, we need to protect the information that evolves throughout the internet and in the storage. In this paper, we propose a steganographic method for Bangla text communication over the unsecured internet based on RGB color image with the help of secret key for better protection of information. The proposed system consists of four main components namely preprocessing, embedding, extraction, and post-processing. We use dynamic positioning in case of substitution of (Bangla) text into the image. The secret key is transmitting throughput the image so that there is no extra hassle for communicating the secret key. We justify our proposed approach by using simulation with respect to imperceptibility, capacity, and robustness. We compare the result of the proposed method with other existing methods and get a better result over several existing efficient methods.

Keywords: Security, IoT, steganography, Bangla, dynamic substitution

Introduction

With the advancement of current paradigm like the Internet of Things (IoT), a huge number of devices like smart phones, laptops, tablets, and others are connecting to the internet for accessing services. These devices are producing a mammoth amount of data and communicating these data among devices and with Clouds. Gartner Inc. forecasts that there will be 4.9 billion connected things in the world in 2015 and the number will rise to 25 billion in 2020 (Gartner, 2014) and the data generated by these devices will be around 403ZB by 2018 (Cisco Inc., 2013). Most of these data needs to be transmitted over on the internet for storing and further processing. So securing these data is desperately necessary when it is traveling from one place to another through the unsecured channel.

Cryptography and steganography are the two key mechanisms that can be used to protect data in communication. Cryptographic algorithms protect the information by changing the original one to a meaningless thing. Many cryptographic techniques are using currently in different real systems. However, the encrypted message causes itself suspicion to the intruders (Shirali-Shahreza, 2006). On the other hand, steganography protects the information from unauthorized parties by hiding it in the digital media like text, image,

*Corresponding author: < saikatmondal@cse.ku.ac.bd>

DOI: https://doi.org/10.53808/KUS.2017.14.1and2.1630-E

video, audio etc. Therefore, steganography hides the fact that any secret communication taking place (Jamil, 1999).

Secret communication is the main goal of steganography and images are the most common cover media used for carrying secret information. The reason is that image resolution in most of the case is more than human perception and data can be kept secret in the noisy area of the image. Recent times, the researchers have proposed many digital image steganographic techniques. Nevertheless, in case of an ideal method, the cover image and stego-image (image with secret information) should be look alike and even can be found an insignificant difference in case of steganalysis.

Image steganography has engrossed researchers' concern due to its' immense data carrying capability without noticeable distortion to the carrier. The simplest and most popular image steganography approach is least significant bit (LSB) substitution. It hides information in the LSB(s) into the carrier image. The hiding capacity can be significantly improved by means of up to 4 LSBs in each pixel by substitution. An LSB substitution based image steganography technique of data concealing is proposed in (Younes and Jantan, 2008), where binary representation of secret data is used to substitute the LSB of each byte within the encrypted image randomly. The security is improved in two levels as data hides in encrypted image. Another image steganography method based on LSB is proposed in (Karim, Rahman and Hossain, 2011), where the authors hide secrets in different positions of LSB in the green or blue channel of the cover image based on the secret key to enhance the security of the traditional LSB substitution technique. The experimental result shows better image quality from the proposed method but LSB is more suspicious than other position in the image to the intruders.

A pixel in color image consists of red (R), green (G) and blue (B) three components. These components are also known as the channel and comprise of 8 bits each. We can conceal three bits if we simply substitute LSB of three channels. Different modified approaches of LSB substitution steganography are proposed with RGB images for providing better security of information in the different research papers. A robust RGB channel based steganography scheme is proposed in the paper (Bairagi, Mondal and Debnath, 2014) with the help of secret key imparting better information security. The technique hides data into the dynamic position of the channels and is determined by the value of channels and secret key. The technique is less vulnerable due to its' dynamic positioning in case of hiding. The use of secret key adds another level of security in the system. Here LSB of the three channels is used as an indicator for finding out whether there is some hidden data or not in the pixel. The research paper (Juneja and Sandhu, 2014) proposed a new improved method for information security by using hybrid feature detection technique in RGB color images. The procedure uses two-component based LSB and adaptive LSB substitution technique for hiding encrypted data with the help of AES in random pixel location of the image. The approach provides better imperceptibility and capacity along with better resistance to various attacks like histogram analysis, Chi-Square attack, and RS analysis.

A new steganography technique depending on improved bit-plane complexity segmentation (BPCS) is proposed in the paper (Sun, 2015). The paper introduces run-length irregularity and border noisiness to work out the problem of black-and-white border

complexity of the regular method. This technique is superior to general BPSC steganography.

Text steganography was proposed for Telegu and Arabic languages in the literature (Alamiti, 2010) and (Shirali-Shahrili, 2010) respectively. We found one literature (Jarin, 2015) for securing Bangla text with the help of image steganographic technique.They used 4-LSB positions in every component of 24-bit color image. The capacity of the proposed method was increased due to the utilization of four substitution positions while compromising the quality of the image. This makes the process vulnerable and they had not presented any security analysis in supporting of their method. Bangla speaking peoples are involving with the digital world rapidly and interacting with each other using Bangla text message. Due to the lack of effective mechanism for securing Bangla text communication over the internet, we are interested to propose such a simplified securing technique that can be utilized in IoT environments.

We propose a steganographic technique here based on RGB image with the help of secret key for communicating Bangla text securely over the internet that can be utilized in IoT environment. The main contribution of this work consists of four algorithms (Algorithm 1, Algorithm 2, Algorithm 3, and Algorithm 4) for securing Bangla text. With the help of first two algorithms, the sender can send his/her text by hiding inside RGB color image. On the contrary, the receiver can get back original text from the image by utilizing last two algorithms. The proposed technique has been analyzed using Bangla texts. The effectiveness of the proposed method has been verified in the result and discussion section, which clearly showed the good quality of the stego-image with higher hiding capacity than the existing works. It is also shown mathematically that this technique can resist visual attack and histogram attack.

Materials and Methods

Shielding information in critical infrastructure is urgent for our existence in the cloud. Here we use an image steganography technique for securing our information. As substitution technique modifies the LSBs, it is easy to reveal the embedded message if low transparency causes suspicious. Thus hiding message into dynamic positions of the deeper layer of the pixel in the image makes it more cumbersome for the adversaries. Moreover, data in deeper layers are less vulnerable to unintentional attacks. The overview of the proposed method is shown in the Fig. 1. It consists of four main blocks: preprocessing, hiding, extraction, and post-processing.

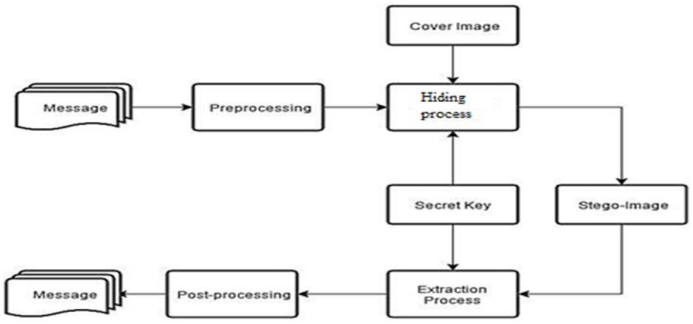

Fig.1: Proposed Steganographic Model

Preprocessing module is responsible to process Bangla text (message) to binary bits so that it can be used to hide into the cover image. Bangla text will be the inputs in preprocessing module. There is 89 Bangla character and the Unicode value of this is range from 0981 to 09FA. There are 122 numbers in this range and we can represent it by 7 bits. So we reduced this Unicode value (which is actually the hexadecimal value) to decimal value and then further reduce to 0 - 121 by subtracting 2433. The Algorithm 1 performs the preprocessing task.

Algorithm 1 pre-processing of the Bangla text

01: Input: Bangla text
02: Initialize CharNumber = 1
03: Convert the CharNumberth character Unicode value to hexadecimal value
04: Convert the hexadecimal value to decimal value
05: Subtract 2433 from that decimal value to represent it into a confined range of 0-121 to represent by using 7 bits
06: Convert this value to a 7-bit number and store into a string named message
07: set CharNumber = CharNumber + 1
08: Goto step 3 until all the character is processed
09: Output will be the bits and stored into the message

The output of the Algorithm 1 will be used as the input in thehiding process. The message bits are hidden into the RGB image with the help of secret key (SK). We use R, G and B channel for embedding the secret message and only LSB of these three channels as an indicator. The position of embedding is dynamic and it depends on the channel quality and message bit. The Algorithm 2 is applied for hiding Bangla text into the carrier image.

Algorithm 2 data hiding process

01: Input: message, Secret key, SK and a cover image, Img
02 Calculate the multiplicative inverse of the SK (miSK=SK-1) with the help of extended Euclid's algorithm
03: Hide the miSK in the image
04: Initialize the position in the image to pixel from the end of hiding miSK, position in secret message bit to mPos from the start
05: Calculate three positions (rP, gP, bP) in three channels of the pixel with the help of corresponding channel value and SK
06: if bit value in the position bP of B channel is equal to the secret data bit in position mPos in the message then
07: Change LSB of B channel to indicate the presence of data
08: Increase mP os by 1
09: else {Hint: Indicate for no data hidden}
10: Change differently of step 7
11: end if

12: if bit value in the position gP of G channel is equal to the secret data bit in position mPos in the message then
13: Change LSB of G channel to indicate the presence of data
14: Increase mP os by 1
15: else {Hint: Indicate for no data hidden}
16: Change differently of step 13
17: end if
18: if bit value in the position rP of R channel is equal to the secret data bit in position mPos in the message then
19: Change LSB of R channel to indicate the presence of data
20: Increase mP os by 1
21: else {Hint: Indicate for no data hidden}
22: Change differently of step 19
23: end if
24: Increase the value of the pixel by 1
25: if the pixel number does not exceed the size of Img or mPos does not exceed the length of the message then
26: Go to step 5
27: else {Hint: Stop the process of hiding}
28: Go to step 30
29: end if
30: Output: stego-image (sImg) that is cover image with message

The receiver will useblue channel of the RGB image for extracting the hiddenBangla text after getting the stego-image. Stego-image will be used as the input for this extraction process. We need the same secret key (SK) to extract the hidden message from stego-image, sImg. For the reason, we first read the miSK from sImg and calculate the secret key SK from there with the help of extended Euclid's algorithm. We check LSB of R, G, and B channels to find whether there is any hidden information or not. If there is any information then we calculate the position of that with the help of corresponding channel value and SK and extract. The Algorithm 3 is utilized for extracting hidden informationafter receiving stego-image in the receiver end.

Algorithm 3 data extraction process

01: Input: Stego-image (sImg)
02: Initialize the position in the sImg to pixel from the end of extraction miSK, position mPos to 1 for storing the extracted bits
03: Read the miSK from sImg
04: Calculate the multiplicative inverse of the miSK (i.e. SK) with the help of extended Euclid's algorithm
05: if LSB of the B channel in the pixel is 1 then
06: Calculate position bP in B channel of the pixel with the help of corresponding channel value and SK

07: Retrieve the stored data from bP position of B channel
08: Store it into a string message and increase mPos by 1
09: else {Hint: There is no data hidden in B channel}
10: Go to step 12
11: end if
12: if LSB of the G channel in the pixel is 1 then
13: Calculate position gP in G channel of pixel with the help of corresponding channel value and SK
14: Retrieve the stored data from gP position of G channel
15: Store it into a string message and increase mPos by 1
16: else {Hint: There is no data hidden in G channel}
17: Go to step 19
18: end if
19: if LSB of the R channel in the pixel is 1 then
20: Calculate position rP in R channel of the pixel with the help of corresponding channel value and SK
21: Retrieve the stored data from rP position of G channel
22: Store it into a string message and increase mPos by 1
23: else {Hint: There is no data hidden in R channel}
24: Go to step 26
25: end if
26: Increase the value of the pixel by 1
27: if the pixel does not exceed the size of sImg or mPos does not exceed the number of hidden bits then
28: Go to step 5
29: else {Hint: Stop the process of hiding}
30: Go to step 32
31: end if
32: Output: message

The unit is responsible for constructing the original Bangla message from the extracted bits that we get from the third module. The output of the extraction process will be the input of the post-processing component. The component is responsible for contracting the Bangla text from the extracted message bits. The process is shown in the Algorithm 4.

Algorithm 4 post-processing for getting Bangla text

01: Input will be the message
02: Initialize CharNumber = 1 and bPos = 1
03: Separate every seven bits from bP os of the message into f BChar
04: Convert fBChar to the decimal value (dBChar)
05: Set dBChar = dBChar + 2433
06: Convert dBChar to hexadecimal value (hBChar)

07: Convert hBChar to Bangla character (BChar)
08: Store BChar in a string named BMessage in position CharNumber
09: Set CharNumber = CharNumber + 1
10: Set bPos = bPos + 7
11: if bPos does not exceed the length of the message then
12: Go to step 3
13: else {Hint: Stop the post-process}
14: Go to step 16
15: end if
16: Output: BMessage

Results and discussion

To assess the algorithms, MATLAB R2010a modules are defined independently for pre-processing, embedding, extracting and post-processing of Bangla text. The experiments are performed on RGB images collected from well-known image dataset (Weber, 1993) and details of the images and experiments are shown in Table 1. For the experiment, we have used different amounts of Bangla text (A = 25010, B = 32330, C = 39040, D = 48800 character) with all the different images. We have used Steganography Studio 1.0.2 tool for analysing the image.

The performance of steganographic or data hiding technique can be measured by using three factors namely imperceptibly, capacity and robustness. The highest the stego-image quality thenit will less suspicious to the attacker. The distortion of the image can be determined by the parameter like mean square error (MSE), peak-signal-to-noise (PSNR) etc. The lesser the distortion, lesser the MSE and PSNR will be higher. If C and S are the cover image and stego-image respectively of size M X N, MSE and PSNR can be calculated by using the equation (1) and (2) respectively.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(S_{ij} - C_{ij})^2 \qquad (1)$$

Where, C_{ij} and S_{ij} are the cover image pixel and stego-image pixel values respectively at i^{th} row and j^{th} column. Here C_{max} represents the actual maximum pixel value in the image.

$$PSNR = 10\log_{10}\frac{C_{\max}^2}{MSE} \qquad (2)$$

The experimental results for the first factor are shown in the Fig. 2 and Fig. 3 by using the proposed method with the different amount of data in diverse images. From these two figures, we see that with the increased amount of data, the PSNR value is decreasing and MSE value is increasing for all the images. The result for the second factor is shown in Table 1. This is measured as the bit per pixel (BPP) and the experimental result shows that on average the proposed method can hide 1.60 bits in a single pixel. The robustness of the proposed method is tested through visual analysis and histogram analysis. A sample 24-bit RGB stego-image produced from the proposed method is shown in the Fig. 4 along with the corresponding cover image. The method successfully resists visual attacks as there

could not find any visual difference between these two images. The result of histogram analysis of the images (Fig. 4) is shown in the Fig. 5 and we find no significant difference among them.

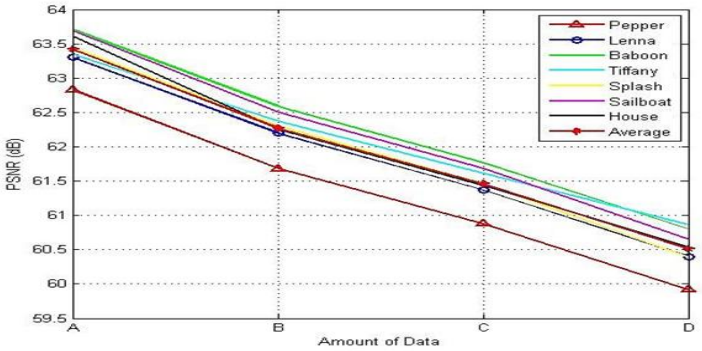

Fig.2: PSNR with different amount of data in different RGB images

Table 1: Detail of images and experiments

Cover Image	Type	Size	BitPer Pixel	Maximum Char
Pepper	PNG	512X512	1.55	50790
Lenna	PNG	512X512	1.64	53740
Baboon	PNG	512X512	1.69	55378
Tiffany	PNG	512X512	1.40	45875
Splash	PNG	512X512	1.58	51773
Sailboat	PNG	512X512	1.76	57672
House	PNG	512X512	1.56	51118
Average			1.60	52428

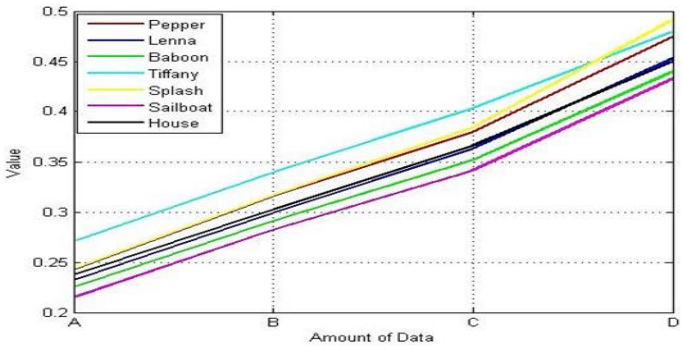

Fig.3: MSE with different amount of data indifferent RGB images

The comparison of the proposed method with existing techniques ((Bairagi et al., 2014; Juneja and Sandhu, 2014; Sun, 2015; Jarin, 2015)) in case of imperceptibility with 32330 Bangla character is shown in the Fig. 6. By comparing the results from the figure, we

can conclude that the proposed approach gives better quality images (imperceptibility) for all the standard images than the existing techniques.

Fig. 4: Comparison between cover image and stego-image

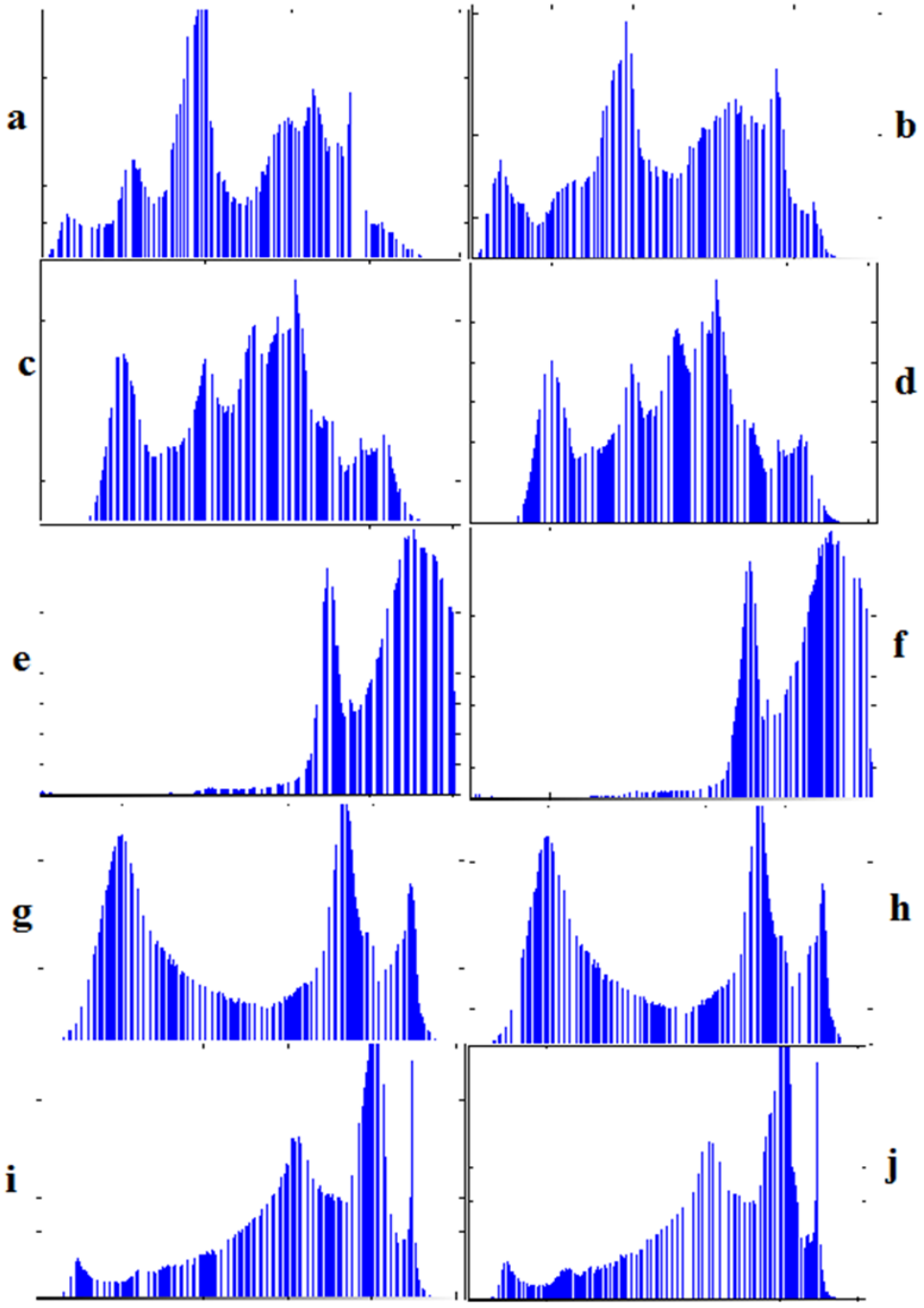

Fig. 5: Comparison of histogram between cover image and stego-image

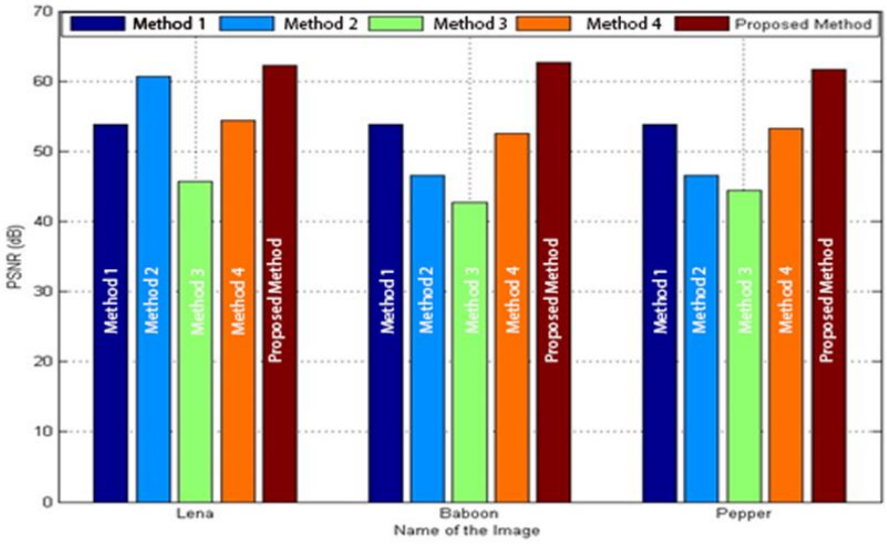

Fig. 6 Comparison of PSNR value of the proposed method with different existing method

Method 1: Bairagi, A. Mondal, S and Debnath, R. 2014. *A Robust RGB Channel Based Image Steganography Technique using a Secret Key*. 16th International Conference on Computer and Information Technology (ICCIT 2013), Khulna, Bangladesh

Method 2: Juneja, M. and Sandhu, P. 2014. Improved LSB based Steganography Techniques for Color Images in Spatial Domain. International Journal of Network Security, 16(6):452-462

Method 3: Sun, S. 2015. A New Information Hiding Method Based on Improved BPCS Steganography. Advances in Multimedia

Method 4: Jarin, Hossain, S.M. and Islam, R. 2015. Introducing Image Steganography in Bangla LanguaguageCommunication. International Journal of Computer Applications, 110(8)

Conclusion

Protecting trusted information is exigent for the users in any environment and it is now more necessary for the pervasive adaptation of new IoT products and services. Here, a text protection technique has been proposed, especially for securing Bangla language communication over the internet using image steganographic approach with the help of RGB color image as media and a secret key. The ambiguity of selecting the position in the channels increases the complexity of steganalysis. Both mathematical and experimental analyses have been performed to find the strength of the proposed method. The experimental results show that the proposed method produces good quality stego-image with high hiding capacity. We compared our proposed method with several efficient existing methods to verify the robustness, storage capacity, PSNR, MSE for different existing images and Bangla texts. It is also shown here mathematically that this technique can resist visual attack and histogram attack. The experimental results are evaluated on images from well-

known datasets and the comparison shows better performance of the proposed method than the other existing works. In the future, our target is to work on a technique that will protect information from intentional and unintentional attacks while transmitting secret data over the internet.

References

Inc, G. 2014.*Gartner says 4.9 billion connected things will be in use in 2015*. Press Release, Available on URL: http://www. gartner.com/newsroom/id /2905717

Inc, C. 2013.*Cisco Visual Networking Index: Forecast and Methodology, 2013-2018*, CISCO White Paper, Available on URL: http://www.cisco.com/c/en/us/ solutions/collateral/ service-provider/global-cloud-index gci/CloudIndex.WhitePaper.html

Shirali-Shahreza, M. 2006.*Stealth Steganography in SMS, in Proc. of IFIP International Conference On Wireless and Optical Communication Net- work*. Bangalore, India

Jamil, T. 1999.*Steganography: The art of hiding information in plain sight*. IEEE Potentials, 18(1): 10-12

Younes, A. and Jantan, A. 2008.*New Steganography Approach for Image Encryption Exchange by using Least Signicant Bit Insertion*. International Journal of Computer Science and Network Security, 8(6): 247-254

Karim, M. Rahman, S. and Hossain, I. 2011.*A New Approach for LSB based Image Steganography using Secret Key*. Proceedings of 14th Inter- national Conference on Computer and Information Technology (ICCIT 2011), Dhaka, Bangladesh

Bairagi, A. Mondal, S and Debnath, R. 2014.*A Robust RGB Channel Based Image Steganography Technique using a Secret Key*. 16th International Conference on Computer and Information Technology (ICCIT 2013), Khulna, Bangladesh

Juneja, M. and Sandhu, P. 2014.*Improved LSB based Steganography Tech- niques for Color Images in Spatial Domain*. International Journal of Network Security, 16(6):452-462

Sun, S. 2015.*A New Information Hiding Method Based on Improved BPCS Steganography*. Advances in Multimedia

Alameti, S. Pothalalah, S. and Babu, A.2010. *A New Approach to Telegu Text Steganography by Shifting Inherent Vowel Signs*. International Journal of Engineering Science and Technology, 2(12): 7203-7214

Shirali-Shahreza, H. and Shirali-Shahreza, M. 2010.*Arabic/Persian Text Steganography Utilizing Similar Letters with Different Codes*. The Ara- bian Journal for Science and Engineering, 35(1B)

Jarin, S.Hossain, S.M. and Islam,R. 2015.*Introducing Image Steganography in Bangla Languaguage Communication*. International Journal of Computer Applications, 110(8)

Weber, G. 1993.*The USC-SIPI image database: version 4,1993*. Available on: URL http://sipi.usc.edu/database/database.php?volume=misc