Data Hiding in Color Images: A High Capacity Data Hiding Technique for Covert Communication

Shabir A. Parah[1], Javaid A. Sheikh[2], G. M. Bhat[3]

[1,2]P.G. Department of Electronics University of Kashmir, Hazratbal, Srinagar,
J and K, India-19006.
[1]shabireltr@gmail.com [2]sjavaid_29ku@yahoo.co.in
[3]University Science Instrumentation Centre (USIC) University of Kashmir, Hazratbal, Srinagar, J and K, India-190006.
[3]gmbhat_ku@yahoo.co.in

Abstract

A high capacity data hiding technique using color images as cover medium and referred to as *4R-4G-4B* technique has been investigated and presented in this paper. The color image is firstly divided into its constituent bit planes followed by data embedding. To thwart the adversary different embedding algorithms have been used for embedding data in Red, Green and Blue planes. Additional layer of security to the embedded data is added by embedding secret data at the pseudorandom locations determined by Main Address Vector (MAV) and Complementary Address Vector (CAV). The comparison of our method with an existing technique shows that proposed technique is capable of providing better quality stego-images even if the embedded data is slightly more. A 2.7dB increase in PSNR in case of proposed technique substantiates the argument.

1. Introduction

Ever since the proliferation of digitized media (audio, image and video) and exponential rise in the usage of internet worldwide, one of the most important factors of communication and information technology has been the security of information being transmitted [1]. Cryptography has been used as a potential tool for information security, but the disguised look of the encrypted data attains the attention of adversaries and as such new techniques need to be developed for enhancing the information security. In recent years steganography [2, 3, 4] has got attention from research community across the world as an alternate option for communicating securely and without letting the adversary know that communication is taking place. The motivating force behind the whole process is that hiding information in the cover medium (which can be, a picnic photograph) is less suspicious than communicating an encrypted picture.

Steganography is the art and science of invisible communication. The history of non-electric steganography dates back to 440 BC, when *Histaeus* intended to communicate to his son-in-law in Greece [5]. The tremendous growth of internet and availability of low cost digital devices has prompted the contemporary research community to revisit steganography and use it as a potential tool for security of information being carried by insecure channels provide by networks like LAN (Local area networks), WAN (Wide area networks) and internet.

2. Color Image based covert communication

Covert communication can be carried out using almost all digital file formats as cover media. Among many available file formats audio and image files have high degree of redundancy but use of audio files as cover medium is less popular compared to images. This is because of multi-plane structure of digital images.

The color and size of a cover image is of paramount importance so for as covert communication is concerned. The three fundamental colors known as primary colors are red (R), green (G) and blue (B). Color images of late have been explored by research community for data hiding process. Various information hiding systems using color images have been presented in literature [6, 7, 8, 9]. A comprehensive review on recent trends in color image processing and data hiding is presented in [10]. A secure and high capacity stegnographic technique using color images as cover medium is presented in [11]. Cryptography and stegnographic concepts have been used to achieve fair degree of perceptual transparency and security. Owing to the multi plane (RGB) structure, color images prove to be very good cover medium from payload point of view.

A secure and high capacity data hiding system exploiting the multi-plane feature of color images is proposed and investigated in this paper. The data is embedded in three constituent planes of the cover image. The Main Address Vector (MAV) and Complementary Address Vector (CAV) determine the locations where secret data is embedded. The data in the three different planes is embedded using different embedding strategies to thwart the adversary. Besides, data in one of the planes has been embedded using crypto domain embedding, which further enhances the security of the system.

3. Proposed Covert Communication System Using Color Images

The proposed high capacity data hiding system uses color images as cover medium. A block diagram of the proposed data hiding system for covert communication is shown in Fig. 1. The RGB color space cover image in which data is to be hidden is broken down into its three constituent color planes: Red plane, Blue plane and Green plane. The message vector containing the data to be hidden into the cover image is divided into three equal message vectors, viz. M1, M2 and M3. The data embedder E1embeds the message vector M1 in the red plane of the cover image under the control of an embedding key K1; derived from master key K (besides containing embedding key K1 master also contains encryption key K2 used for encrypting blue plane). Similarly embedder E2 embeds message vector M2 in the blue plane and embedder E3 embeds message vector M3 in the green plane of the cover image. Although same Key is used to embed data in all the cover images but different embedding strategies have been used to hide data in the various constituent planes of the cover image.

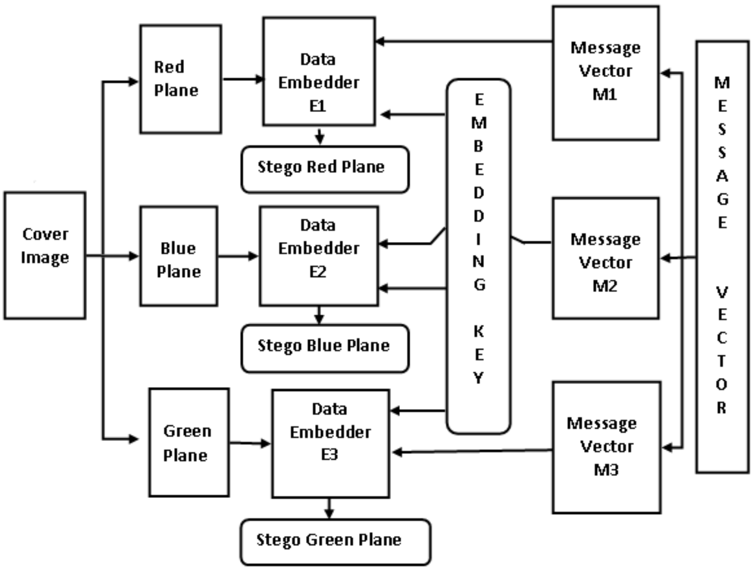


Fig.1 Data embedding in three constituent planes of the cover image

The proposed system hides 37.5% (262144 × 3) bits of data in every constituent color plane of the RGB cover image. The implemented technique uses four bit planes (LSB and first three Intermediate Significant Bit planes) of every constituent color plane for hiding data and as such can be referred to as *4R-4G-4B* technique. The embedding techniques pertaining to various color planes are described below.

3.1Embedding in red plane and address vector generation

The data embedding strategy in red plane of the cover image is depicted in Fig. 2. The data to be embedded in the red plane has been divided into four equal data vectors L1, L2, L3 and L4. The cover image (Red plane) has been broken down into constituent bit planes, as the bit depth of every pixel of red plane is eight. Since color test images of size 512 × 512 have been used as cover medium, the number of pixels contained by each plane of test image is 262144.

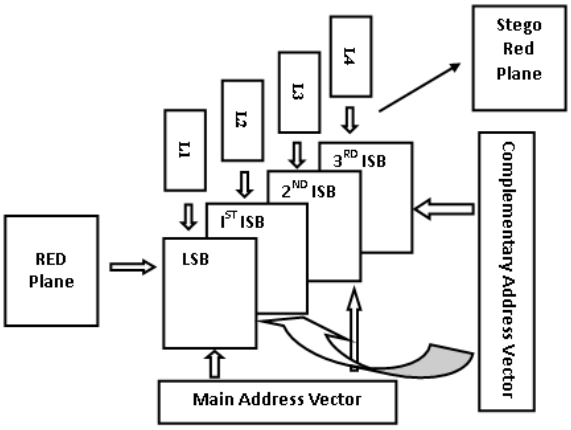


Fig. 2 Data embedding in red plane

To address all the pixel locations, an 18 bit Key has been used to generate a Pseudorandom Address Vector (PAV), called as Main Address Vector (MAV), containing as many as 262144 addresses [12]. PAV has been used to generate Complementary Address Vector (CAV). For example, for a 3-bit PAV generator capable of generating seven distinct addresses, let us assume the initial state of the generator to be 111(7). The successive clock cycles will make the generator to run through 011(3), 101(5), 010(2), 001(1), 100(4), 110(6) and back to 111(7). CAV for such a PAV has been obtained by subtracting every address of the address vector from 8. A typical pseudo random address vector and its complementary address vectors are show in Table. 1.

Table 1 A typical Main Address Vector and its Complementary Vector.

Main Address vector (MAV)	7	3	5	2	1	4	6
Complementary Address Vector (CAV)	1	5	3	6	7	4	2

The basic principle utilized for the generation of CAV from PAV is that, if all the addresses of PAV are represented by a n bit vector, then CAV is generated by subtracting PAV addresses from 2^n. The implemented technique uses the above mentioned concept to generate address vector containing 262143 addresses using an 18 bit seed word. The complementary address vector has been obtained by subtracting all the entries of the address vector from 262144. Since 37.5% or $262144 \times 3 = 786432$ bits of data are to be embedded in four bit planes of the red plane, the length of each data vector L1 through L4 equals $786432 \div 4 = 196608$ bits. Therefore in all four bit planes of the red plane, out of a total of 262144 locations the secure data bits are embedded in 196608 locations. The locations where data is embedded are determined by the contents of main address vector in case of 1st and 3rd bit planes of red plane, whereas for 2nd and 4th bit planes, the locations are determined by the contents of complementary address vector. It is worthwhile to mention that only first 196608 locations pointed to by corresponding address vectors in all the four concerned bit planes, are used for data embedding.

3. 2 Embedding in green plane

The data embedding strategy in green plane of the cover image is depicted in Fig. 3. Like red plane this plane also hides 37.5% of secure data; however the embedding strategy is quite different. The data to be embedded in this plane is divided into four data vectors L1, L2, L3 and L4. Keeping in view the fact that image quality deteriorates more when secure data is embedded in higher order bit planes, the lengths of data vectors L1 and L2 are choosen to be equal to 262144. Further, the lengths of data vectors L3 and L4 is equal to $262144 \div 2 = 131072$. In other words data contained by both vectors L3 and L4 is equal to that contained in either L1 or L2 alone.

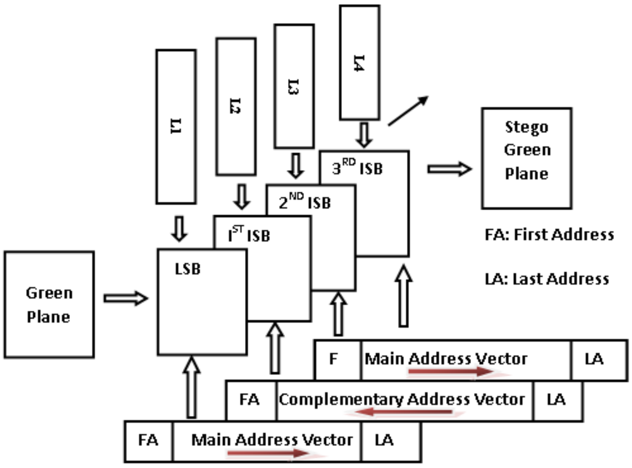


Fig.3. Data embedding in green plane

In green plane, alternate top-down and down top embedding has been used to embed data: data is embedded in the first bit plane of green plane at locations determined by addresses contained by the Main Address Vector (MAV) by traversing MAV from first to last location (top down embedding). Data embedding in the second bit plane of green plane has been carried out using Complementary Address Vector in reverse order i.e. from last address location to first one (bottom top embedding). Embedding in third and fourth bit planes is determined respectively by first and last half address locations of the MAV as depicted in Fig. 3.

3. 3 Embedding in blue plane

The data hiding strategy for blue plane is different from red and green planes. The cover medium i.e. blue plane is firstly encrypted by scrambling it as per address locations pointed to by the Main Address Vector. The scrambling process is carried out using the encryption key K2. The data embedding is carried out in encrypted blue plane to thwart the adversary. The data embedding has been carried out using same strategy as in case of green plane, i.e. LSB and first Intermediate Significant Bit plane (ISB) are fully embedded with data in accordance with main and CAV, respectively. Embedding in 3rd and 4th ISBs of encrypted blue plane is carried out partially; with embedding locations determined by first and second half of the MAV, respectively. Once the embedding is complete in all the specified planes of blue plane, the encrypted cover medium (blue plane) containing hidden data is converted back to its original form by decrypting it using same key as that used for encryption. The whole embedding process as such has been depicted in the Fig. 4. Fig. 5 shows the concatenation process of three constituent stego RGB planes to obtain corresponding color stego-image.

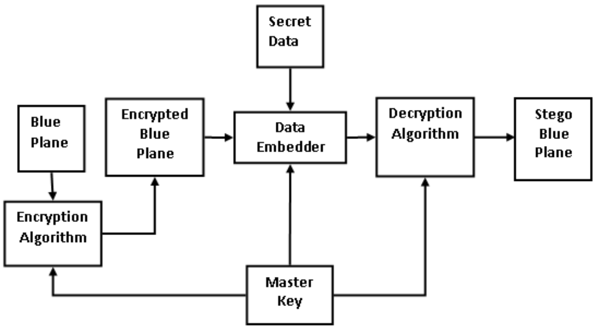


Fig. 4 Data Embedding in Blue plane

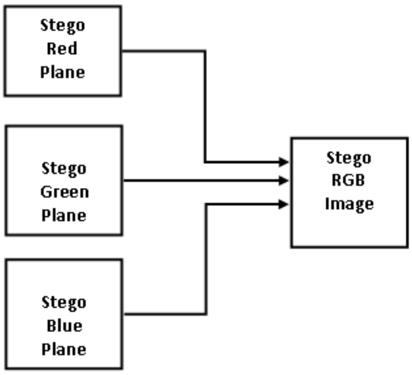


Fig. 5 Concatenation of three color planes

Figure 6 shows various color test images and their respective stego images obtained using *4R-4G-4B* technique. Table 2 shows various image indices pertaining to the various color test images and their respective stego versions.

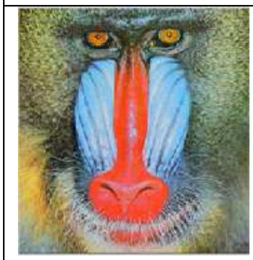
Cover Image	Stego-Image	Cover Image	Stego-Image
			

Fig. 6 Various color images and their stego-versions

Table 2 Observed image indices

S. No.	Cover Image	PSNR (dB)	NAE	NCC
1.	Lena	43.86	0.0108	0.9957
2.	Baboon	43.86	0.0087	0.9964
3.	Peppers	43.91	0.0129	0.9958
4.	Jet	43.89	0.0086	0.9967
5.	Girl	44.31	0.0253	0.9933
6.	Scene	43.84	0.0106	0.9966
7.	Milkdrop	43.94	0.0172	0.9950

4. Results and Discussions

Data hiding using images has found used for various applications including those in covert communications, watermarking and fingerprinting etc. with every application demanding its own set of requirements. Data hiding for covert communication is supposed to be accompanied by the attributes like high data hiding capacity (payload), better perceptual transparency and high security. The proposed system has been developed with an aim to achieve good perceptual quality of stego-images for a high payload, besides an adequate level of security for embedded data. The proposed *4R-4G-4B* based high capacity data hiding system is capable of hiding 235996 bits or 37.5% of data in a given cover image. To enhance security of the embedded data and hence thwart an adversary, a part of it has been embedded in the encrypted blue plane. This means that in order to extract data, the adversary should have the knowledge of encryption key besides the knowhow to generate Main Address Vector as well as Complementary Address Vector. Further the direction of utilization of the Address vectors is also fiddled with, to make the detection much more difficult. The cover images chosen for hiding the data are 512 × 512 RGB bitmap images. The subjective quality analysis of the proposed system shows good results in terms of the quality of stego image. The efficacy of the proposed system has been checked and verified in terms of various image indices like Peak Signal to Noise Ratio (PSNR), Normalized Absolute Error (NAE) and Normalized Cross Correlation (NCC) calculated as follows:

$$PSNR = 10 \times \log_{10} \frac{255^2}{MSE} \qquad (1)$$

$$MSE = \frac{1}{MN} \sum_{j=1}^{M} \sum_{k=1}^{N} (x_{j,k} - x'_{j,k})^2 \qquad (2)$$

Where *MSE* represents Mean Square Error and is calculated between original cover image $x_{j,k}$ and its corresponding stego image $x'_{j,k}$, M and N represent the number rows and columns of an image respectively.

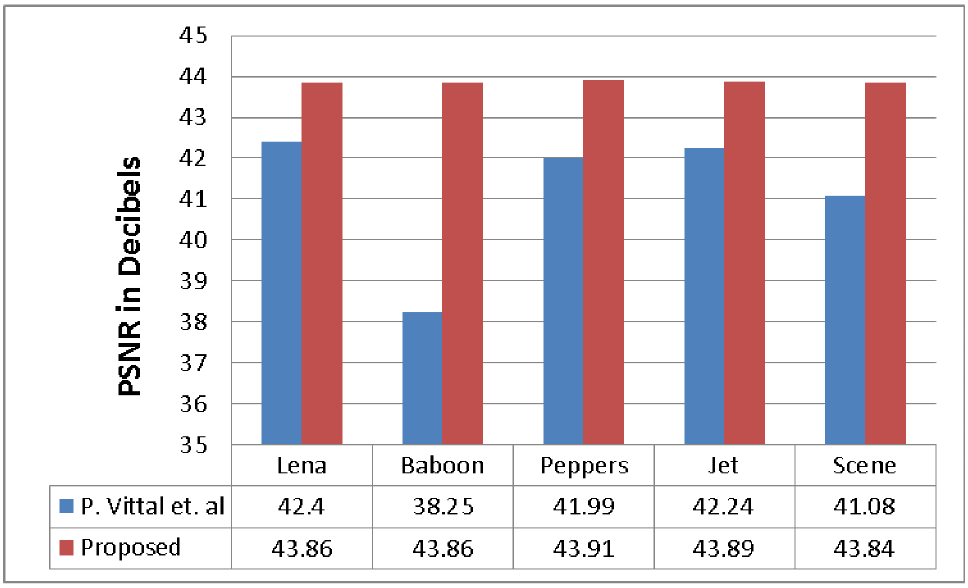
$$NAE = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N} |x_{j,k} - x'_{j,k}|}{\sum_{j=1}^{M} \sum_{k=1}^{N} |x_{j,k}|} \qquad (3)$$

$$NCC = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N} x_{j,k} * x'_{j,k}}{\sum_{j=1}^{M} \sum_{k=1}^{N} x^2_{j,k}} \qquad (4)$$

The various image indices are presented in Table 2. As has been shown, the PSNR of the proposed system lies in the range of 43 to 44dB, well above the 36 dB threshold [1], NAE values are low and NCC close to unity indicating that the adversary cannot get any indication of data hidden in the cover medium. The image indices presented in the Table 2 have been found by calculating mean values of their corresponding red, green and blue plane indices. The results obtained in the proposed scheme have been compared with those reported in [11]. The comparison clearly shows that PSNR values obtained in the proposed system are higher than those in [11], even though data hidden in the cover images is more. Table 3 shows the comparison results between the proposed system and [11]. Figures 7 and 8 present a graphical comparison of various parameters between proposed technique and those presented in [11].

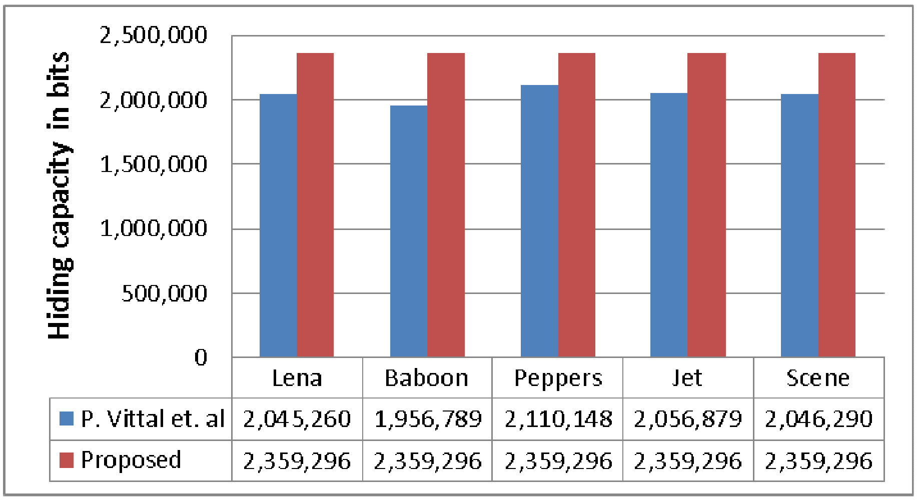
Table 3 Comparison of proposed scheme with [11]

Host Image	**Capacity (in bits)**		**PSNR(dB)**	
	P. Vittal *et. al*	Proposed	P. Vittal *et. al*	Proposed
Lena	20,45,260	23,59,296	42.40	43.86
Baboon	19,56,789	23,59,296	38.25	43.86
Peppers	21,10,148	23,59,296	41.99	43.91
Jet	20,56,879	23,59,296	42.24	43.89
Scene	20,46,290	23,59,296	41.08	43.84
Average	20, 43,703	23,59,296	41.19	43.87



	Lena	Baboon	Peppers	Jet	Scene
🟦 P. Vittal et. al	42.4	38.25	41.99	42.24	41.08
🟥 Proposed	43.86	43.86	43.91	43.89	43.84

Fig. 6 PSNR comparison between [11] and proposed system



	Lena	Baboon	Peppers	Jet	Scene
🟦 P. Vittal et. al	2,045,260	1,956,789	2,110,148	2,056,879	2,046,290
🟥 Proposed	2,359,296	2,359,296	2,359,296	2,359,296	2,359,296

Fig. 8 Payload (Hiding Capacity) comparison between [11] and proposed system

5. Conclusion

A high capacity and secure data hiding system suitable for covert communication using color images as cover medium has been presented. The proposed technique uses four bit planes of every constituent color plane for data hiding and as such is named as *4R-4G-4B* technique. The data to be embedded in the cover image is divided into three equal parts and the cover image is broken into three constituent color planes. The constituent (RGB) planes serve as cover media for the three data vectors. Different embedding strategies have been used to embed data in the three color planes to thwart an adversary. The proposed technique uses the concept of Pseudorandom Address Vector and a Complementary Address Vector to embed data in various bit planes of the cover media. To test the efficacy of the scheme a number of bitmap color test images (512 × 512) have been used as cover media. In every test image, 2359296 bits of data equivalent to 37.5% of the size (in bits) of cover image has been embedded. The efficiency of the proposed system has been checked in terms of various image indices like PSNR, NAE and NCC. As has been shown, the proposed technique provides PSNR in the range of 43 to 44dB, indicating good imperceptibility. The low values of NAE and values of NCC close to unity are an indication of the fact that adversary can't suspect the presence of any hidden data. The implemented technique has been compared with that reported in [11]. It is clear that proposed technique provides an average increase of about 2.7

dB in PSNR even when the average data hidden in the cover images is increased by about 316222 bits, a clear indication of an improved performance of the proposed technique.

References

1. N. I Wu and M.S. Hwang, "Data Hiding: Current Status and Key Issues" International Journal of Network Security, Vol.4, No.1, PP.1–9, Jan. 2007
2. A. Cheddad, J. Condell, K. Curran and P. M. Kevitt, "Digital image steganography: Survey and analysis of current methods" Signal Processing vol. 90. pp. 727–752, 2010
3. K.H. Jung and K.Y. Yoo, "Data hiding method using image interpolation", Computer Standards and Interfaces 31 (2) (2009) 465–470
4. Z. Li, X. Chen, X. Pan and X. Zeng, "Lossless data hiding scheme based on adjacent pixel difference", in: Proceedings of the International Conference on Computer Engineering and Technology, pp. 588–592, 2009
5. J. Silman, , "Steganography and Steganalysis: An Overview", SANS Institute, 2001.
6. C. C. Chang, C.-Y. Lin, and Y.-H. Fan, "Lossless data hiding for color images based on block truncation coding," Pattern Recognition, vol. 41, no. 7, pp. 2347–2357, 2008.
7. M. Xenos, K. Hantzara, E. Mitsou, and I. Kostopoulos, "A model for the assessment of watermark quality with regard to fidelity", Journal of Visual Communication and Image Representation, vol. 16, no. 6, pp. 621–642, 2005.
8. S. C. Pei and J. J. Ding, "Reversible integer color transform", IEEE Transactions on Image processing,vol.16,no.6,pp. 1686–1691, 2007.
9. R. Lukac and K. N. Plataniotis, "Secure color imaging," in Color Image Processing: Methods and Applications, R.Lukac and K. N. Plataniotis, Eds., chapter 8, pp. 185–202, CRC Press, Boca Raton, Fla, USA, 2007
10. A. Tremeau, S. Tominaga, and K. N. Plataniotis, "Color in Image and Video Processing:Most Recent Trendsand Future Research Directions" EURASIP Journal on Image and Video Processing, Hindawi Publishing Corporation, Article ID 581371, pp. 1-26 doi:10.1155/2008/58137, 2008
11. S. P. Vitthal., S. B. Rajkumar., A. R Panhalkar, "A Novel Security Scheme for Secret Data using Cryptography and Steganography" I. J. Computer Network and Information Security, vol. 2, pp. 36-42, 2012
12. S. A. Parah, J. A Sheikh and G. M. Bhat, "On the realization of a secure, high capacity data embedding technique using joint top-down and down- top embedding approach" Elixir Comp. Sci. & Engg. (49), pp. 10141-10146 , 2012.

Shabir A. Parah has completed his M. Sc and M. Phil in Electronics from University of Kashmir, Srinagar in the year 2004 and 2010 respectively in the field of Signal processing and embedded systems. He is presently perusing Ph. D in the field of image based data hiding. He is working as Assistant Professor in the department of Electronics and I. T University of Kashmir, Srinagar. His fields of interest are Signal Processing, embedded Systems, Secure Communication and Digital design. Mr. Shabir A. Parah has guided about fifteen projects. He has published about thirty research papers in International and National journals and conference proceedings.

Dr. Javaid A. Sheikh has completed his M.Sc., M. Phil and Ph. D in Electronics from University of Kashmir, Srinagar in the year 2004, 2008 and 2012 respectively in the field of communications and Signal Processing. He is working as Assistant Professor in the department of Electronics and I. T University of Kashmir, Srinagar. His fields of interest are Wireless Communications, design and development of efficient MIMO OFDM based wireless communication techniques, Spread Spectrum modulation, Digital Signal Processing, Electromagnetics. Besides teaching and research, Dr. Javiad A. Sheikh has guided about thirty five projects. He has published about thity research papers in International and National journals and conference proceedings.

Prof. G. M. Bhat obtained his M.Sc. (Electronics) from the University of Kashmir, Srinagar (India) in 1987, M.Tech. (Electronics) from Aligarh Muslim University (AMU), Aligarh (India) in 1993 and Ph.D. Electronics Engg. from AMU, Aligarh, (India) in 1997. The major field of research of Dr. Bhat is Signal Processing Techniques and Secure Message Communication. He has served as Assistant Professor, Associate professor and now as Professor & Director, University Science Instrumentation Centre (USIC), University of Kashmir. He has published more than fifty research papers on his area of interest. He has worked in the area of Mobile Radio

Communication, Spread Spectrum Communication and Neural Networks and has guided many research degrees leading to the award of M.Phil and Ph.D. His present research interests include Secure Message communication, Neural networks and Signal Processing techniques for communication.