Journal of Information Engineering and Applications
ISSN 2224-5782 (print) ISSN 2225-0506 (online)
Vol. 2, No.11, 2012

High Capacity Data Embedding using joint Intermediate Significant Bit (ISB) and Least Significant Bit (LSB) Technique

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

[1,2]Department of Electronics and Instrumentation Technology, University of Kashmir Srinagar-190006, India

[3]University Science Instrumentation Centre (USIC), University of Kashmir, Srinagar-190006, India

Corresponding Author: Shabir A. Parah Email: Shabireltr@gmail.com

Abstract: The success of the Internet, coupled with availability of relatively inexpensive digital devices has created an environment in which it has become very easy to obtain, replicate and distribute digital content without any loss in quality. In such a scenario, data hiding has received significant attention from the research community round the globe, as it has been found useful in various areas like copyright protection, copy control, fingerprinting, content authentication and information security. Least Significant Bit based data hiding techniques have been used as effective means to hide the data to be secured, but they are less robust in nature. This paper presents a high capacity data hiding technique in which the data to be secured is embedded in Intermediate Significant Bits in addition to Least Significant Bits of cover image. The data to be embedded is broken down in data blocks of variable length and each block is embedded in the cover media in such a way that highest length data vector is embedded in lower order bit plane and vice-versa. This work shows attractive results with respect to imperceptibility and capacity when compared with a few reported techniques.

Key Words: Intermediate Significant Bit, Embedding, Imperceptibility, Least Significant Bit.

I. Introduction

The rapid development of secure data transmission technology has increased the horizon of communication via internet. This has resulted in serious challenges pertaining to integrity and security of data being communicated. Owing to this covert communication, nowadays, is being used as a potent way to avert the data security and integrity challenges. One of the most ancient applications of data hiding is covert communication and it traces back to ancient Greek period. Data hiding is generally considered to be the art of keeping message secret and is also referred to as stenography [1]. The chief aim of steganography is to hide information inside cover medium in such a way that it is not possible to detect the existence of secret message [2]. One of the most sought after issues in steganography is that the very presence of a hidden message must be concealed. Stegnography and cryptography belong to spy craft family. Although steganography has been studied as part of cryptography for many decades, the focus of steganography is secret communication. In fact, the modern formulation of the problem goes by the name of the *prisoner's problem*. Here Alice and Bob are trying to hatch an escape plan while in prison. The problem is that all communication between them is examined by a warden, Wendy, who will place both of them in solitary confinement at the first hint of any suspicious communication. Hence, Alice and Bob must trade seemingly inconspicuous messages that actually contain hidden messages involving the escape plan. Further the duo ensures that the medium carrying information about their plan should pass through Wendy a less number of times so as to avert any suspicion. For this they try to put as much information in the medium (cover) as possible.

There are two versions of the problem that are usually discussed — one where the warden is passive, and only observes messages, and the other where the warden is active and modifies messages in a limited manner to guard against hidden messages.

This paper tries to address first problem where adversary is passive. As such emphasis has been given to high data hiding capacity coupled with imperceptibility. Rest of the paper has been organized as follows. Section II provides information about some application areas and requirements of data hiding system. In section III literature survey regarding high capacity data hiding techniques has been presented. Section IV provides complete description of the proposed work. The results of computer simulation tests carried on the proposed technique and comparison of the results has been presented in section V. The paper concludes in section VI.

II. Data Hiding: Applications and Requirements

Data hiding that encompasses both digital watermarking and steganography has been found useful in following areas:

i. Copyright Protection
ii. Copy control
iii. Content Authentication
iv. Broadcast monitoring
v. Fingerprinting
vi. Metadata binding
vii. Covert communication

A data hiding system is characterized by three important characteristics that contend with each other, they are capacity, security, and robustness. This situation is shown in Fig.1 by conflict triangle. Capacity of a data hiding system is also referred to as payload and refers to the amount of information that can be hidden in the cover medium; security refers to an eavesdropper's inability to detect hidden information. Robustness accounts for the amount of modification the stego-medium can withstand before an adversary can destroy hidden information [3].

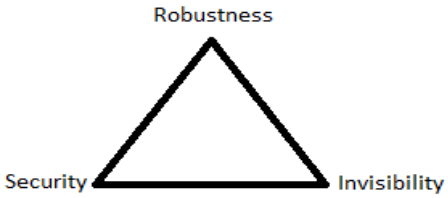


Fig. 1: Conflict Triangle

Broadly information or data hiding encompasses both watermarking and steganography. The chief goal of a watermarking system is to achieve a high level of robustness that is, it should be impossible to remove a watermark without degrading the data object's quality. However, steganography strives for high security and capacity, which often entails that the hidden information is fragile. This leads to destruction of any secret data in the cover medium even with trivial modifications made to stego-medium. The security issue in a stegnographic system is taken care of using security key which can be related to Kerckhoff's assumption. Kerckhoff's assumptions state that one should assume that the method used to encrypt the data is known to an unauthorized party and that the security lies in the choice of a key [4]. The data hiding system is therefore considered truly secure if mere knowledge of exact algorithms for embedding and extracting the data does not help an unauthorized party to detect the presence of the hidden data or remove it. The embedding key therefore forms a pivotal part in the strength of a data hiding system. In fact security strength of a data hiding system depends on the strength of Key (which is a function of key length) [5]. In addition to the security parameter, the other two parameters robustness and Invisibility, or imperceptibility (a function of payload) of a data hiding system conflict with each other.

III. Literature Review

The growth of multimedia technologies has resulted in an enormous research effort being laid on the authentication and copyright protection of digital data being transferred via internet and other communication channels. In the last few years lot of research attention has been paid in this direction, [6] gives an idea about acceleration of research in this area.

Least significant bit (LSB) data hiding is easiest and one of the earliest data hiding techniques. Two LSB techniques are described in [7]. The first replaces the LSB of the image with a PN sequence, while the second adds a PN sequence to the LSB of the data. In [8] a few direct sequence spread spectrum techniques are proposed to embed a watermark in host signals. One of these, LSB-based, is a statistical technique that randomly chooses n pairs of points (a_i, b_i) in an image and increases the brightness of a_i by one unit while simultaneously decreasing the brightness of b_i . Another PN sequence spread spectrum approach is proposed in [9] where the authors hide data by adding a fixed amplitude PN sequence to the image.

In [10] development of stegnographic techniques for gray scale images has been reported. The schemes are reported to have high hiding capacity and good imperceptibility properties. [11] Reports a high capacity data embedding scheme based on average covariance. The MSB of the payload are embedded into cover image based on average covariance of cover image. The authors have reported PSNR of 46.31% for hiding capacity of 12.50%.

A watermarking technique based on Intermediate Significant Bit (ISB) replacement has been presented in [12]. The authors have reported that embedding information in the intermediate significant bits improves robustness compared to robustness when data is hidden in least significant bits. [13] Reports a high capacity embedding technique based on spatial domain. The host image is partitioned into non-over lapping blocks, with each block containing three 3x3 pixels. In every block these pixels receive special treatment, with an aim to decrease the noise and deviations from the original picture values. The authors have reported an embedding capacity ranging between 20-26% when PSNR is limited between 27db to 30db range

IV. Proposed Data Hiding Technique

The block diagram of proposed high capacity data embedding scheme is shown in Fig. 2. Prior to data embedding the cover image is broken into its constituent bit planes as shown in Fig. 3. Since the perceptual quality of the covermedia directly depends on the amount of data embedded in the covermedia besides the significant bit plane in which the data is embedded, the proposed algorithm divides the data to be embedded in number of blocks equal to the bit planes in which the data is to be embedded [14,15]. The embedding strategy is also depicted in Fig. 3. The data vector of length L to be embedded in the cover medium is divided into four varying length data vectors, viz: *L1, L2, L3* and *L4* as shown. The highest length data vector *L1* is embedded in first bit plane under the control of a private key that is capable of embedding data randomly in various locations of the said bit plane. Similarly the other data vectors *L2, L3* and *L4* of relatively decreasing lengths are also embedded in the second third and fourth bit planes of the cover medium. The lesser data embedding in the higher order bit planes ensures better perceptual quality of the stego images. It is important to mention here that data embedding in second, third and fourth bit planes is also carried out using pseudorandom key to thwart the adversary. The embedding process is carried out in data embedder, that outputs an image containing secret data and is generally termed as stego-image. The security of data embedded is a function of Key Length. The used pseudo random number generator (PRNG) used has been chosen such that it is capable of addressing all the locations in any given bit plane.

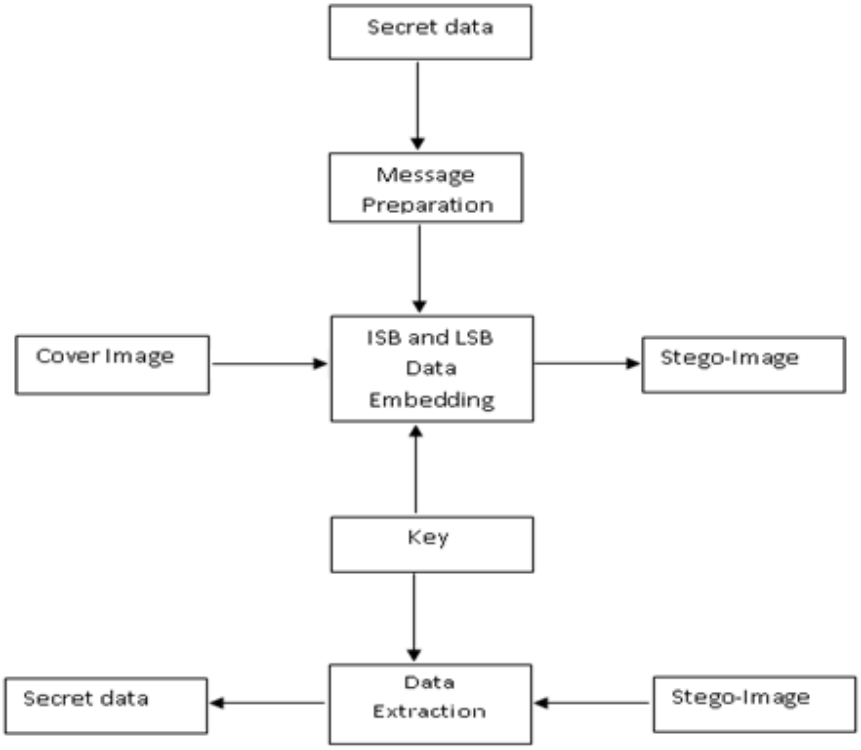


Fig. 2: Proposed high capacity data hiding and corresponding extraction system

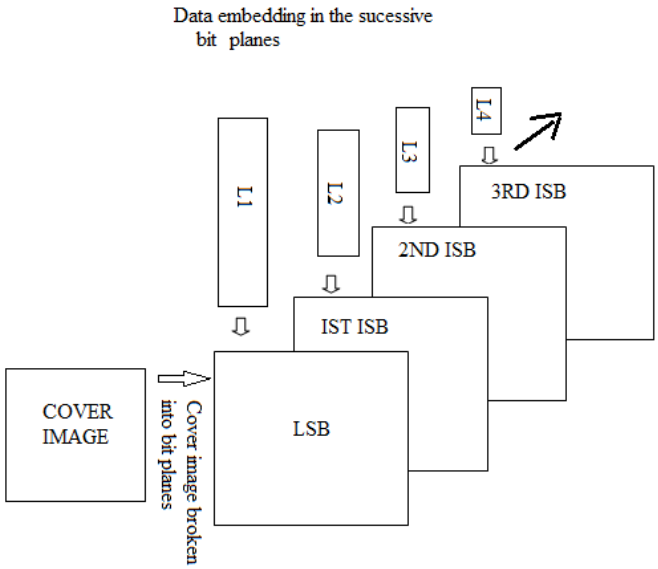


Fig. 3: Data embedding strategy

At the receiving end same key has been used to extract data from the stego image. Since cover image is not needed for the retrieval of secret data the proposed system falls in the category of blind detection. The complete algorithm has been summarized in Table 1.

V. Results and Analysis

A high capacity data hiding system has been presented. The implemented system uses a number of gray scale test images as cover medium to embed the information to be secured. The area of focus on the implemented system is to embed maximum information in the cover images keeping the image degradation minimum so that it could not be perceived that something has been embedded in the host image. The test images chosen are standard grayscale images (512 x512 size) as shown in Fig. 4. Table 2, shows every test image with its corresponding stego image, besides showing payload and PSNR in each case. In case of all the test images the embedding capacity has been fixed at 25% of the cover image except test image 'lake' where the payload of 31.25% has been choosen. Table 3 presents a subset of host images with amount of percentage data embedded and corresponding peak signal to noise ratio PSNR. Further a comparison of the proposed data hiding scheme with that of Zeki et. al is presented. Table 4, shows a graphical comparison of the proposed scheme with [13]. The hiding Capacity (HC) and PSNR have been calculated as follows.

Hiding Capacity (HC): Hiding capacity also referred to as or payload, is the size of the data that can be embedded in the cover image, without deteriorating integrity of the cover image. Capacity is represented by bits per pixel (bpp). It is given by (total number message bits/total number of image bits) multiplied by 100. If n and N respectively denote total message bits and image bits the hiding capacity is given by

Hiding Capacity (HC) = (n/N)*100

Peak Signal to Noise Ratio (PSNR): It is measure of objective equality of an image. It gives an idea about how much deterioration has embedding caused to the image. It is represented as

$$PSNR = 10\log_{10}\frac{255^2}{mse}db$$

Where mse is mean square error and is given by

$$mse = \left[\frac{1}{N*N}\right]^2\sum_{i=1}^{N}\sum_{j=1}^{N}(X_{ij}-\bar{X}_{ij})$$

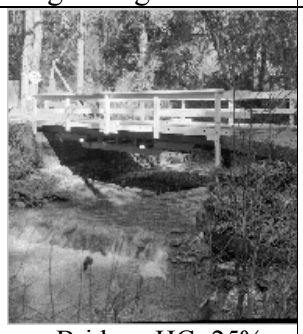
Where N and M are image dimensions, X_{ij} and $\bar{X}_{ij}$ represent original and stego images respectively.

Table1. Data Embedding and extraction Algorithm

Data embedding and extraction Algorithm
Embedding Algorithm
• *Take cover image* • *Prepare cover image by clearing all the locations where data is to be embedded.* • *Take alpha numeric or image based secret message to be communicated to the receiver.* • *Prepare secret message by converting it in a binary vector* • *Divide the message vector in as many parts as is the number of bit planes in which data is to be embedded.* • *Generate a pseudo random vector using a given seed that is capable of addressing all the bit locations in which data is to be embedded* • *Embedded the data in the selected bit planes and at the addresses determined by PRNG.* • *The output of embedder is stego image*
Extraction Algorithm
• *At the receiving end, apply the received stego image to data extractor* • *Use same key used at transmitter to the data extractor.* • *At the output of extractor we get secret message, embedded at the transmitter.* • *Convert the received bit stream in to the corresponding alphanumeric text or image using an inverse transformation at the receiver*

Fig. 4: Various Test Images

Table 2: Host images and their stego versions

Details of host image, its stego version PSNR and percentage of data embedded			
Host image	Stego Image	Host Image	Stego Image
 Bridge	 Bridge , HC=25% PSNR=36.52db	 Boats	 Boats, HC=25% PSNR=35.84db

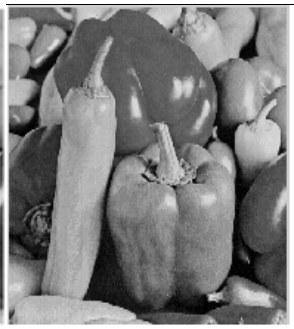
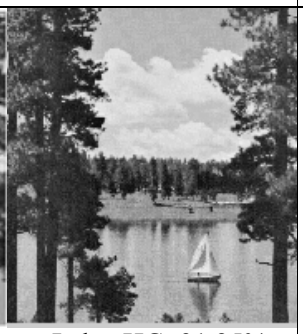
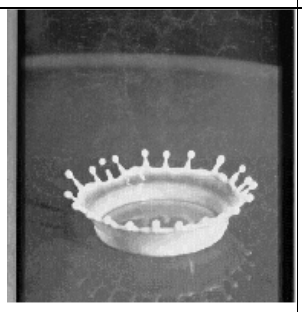
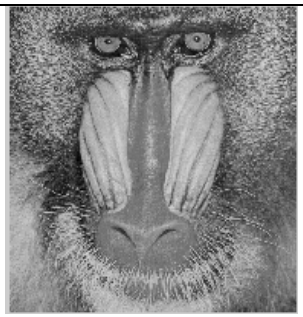
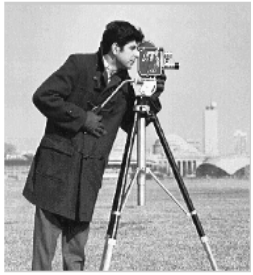
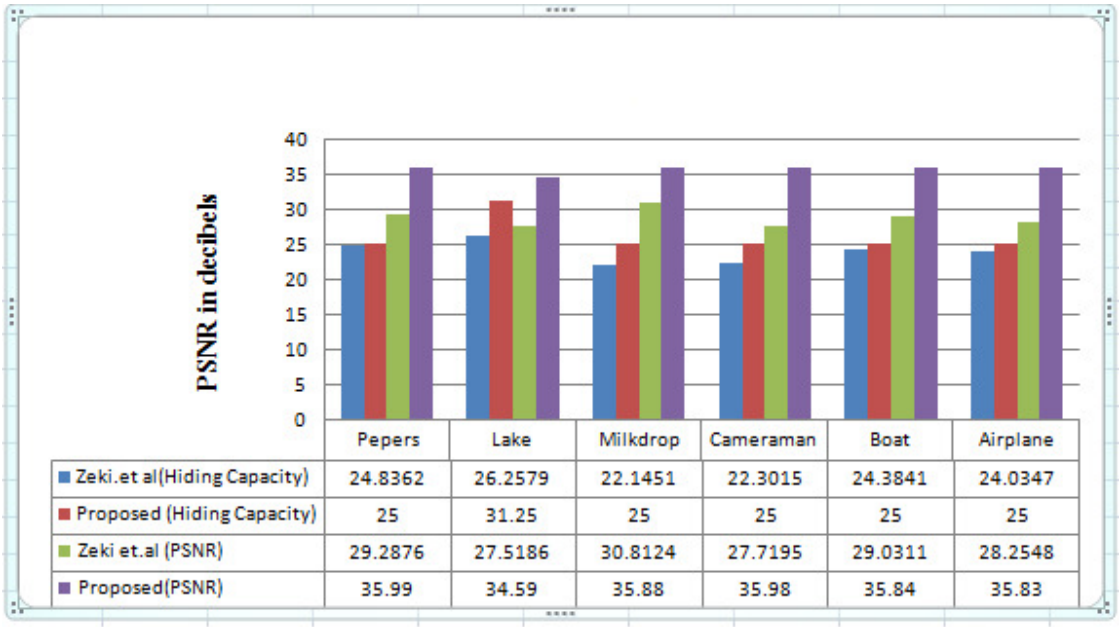
 Peppers	 Peppers, HC=25% PSNR=35.99db	 Airplane	 Airplane ,HC=25% PSNR=35.83db
 Lake	 Lake, HC=31.25% PSNR=34.59db	 Barbara	 Barbara, HC=25% PSNR=36.00db
 Milk drop	 Milk drop, HC=25% PSNR=35.88db	 Baboon	 Baboon, HC=25% PSNR=36.04db
 Cameraman	 Cameraman. HC=25% PSNR=35.98	 Lena	 Lena, HC=25% PSNR=36.08

Table 3: Capacity versus PSNR of various test images of proposed technique.

Host Image	Capacity (%)		PSNR(db)	
	Zeki[13]	Proposed	Zeki[13]	Proposed
Peppers	24.8362	25	29.2876	35.99
Lake	26.2579	31.25	27.5186	34.59
Milk drop	22.1451	25	30.8124	35.88
Cameraman	22.3015	25	27.7195	35.98
Boats	24.3841	25	29.0311	35.84
Airplane	24.0347	25	28.2548	35.83

Table 4: Comparison between proposed technique and that of Zeki. et.al [13].



VI. Conclusion

A high capacity data hiding technique is presented in this paper. The image in which the data is embedded has been broken into its constituent bit planes. The data to be embedded is divided into as many varying length data vectors as the number of bit planes in which the data is to be embedded. The implemented technique embeds data in four bit planes viz. LSB and first three ISBs. The data is embedded under control of a key that embeds data pseudo randomly in various bit planes, thus providing an adequate security to the data carried by the cover image. The technique has been implemented using MATLAB 7. The results obtained in the proposed method have been compared with some existing algorithms viz-a-viz hiding capacity and PSNR. The proposed technique provides an improvement of about 5db to 8db in PSNR when compared with that of Zeki et.al. The results clearly show that the proposed technique has a better performance.

References:

1. Cachin, C. (**2004**). "An information-theoretic model for steganography," *Information and Computation,.* 192, 41-56.

2. Petitcolas F. P., Anderson R. J., and Kuhn N. G.(**1999).** "Information Hiding—A Survey" *Proceedings of The IEEE*, , 87(7), 1062-1078

3. Miller M. L., Doerr G. J. and Cox I. J. **(2004).** "Applying Informed Coding and Embedding to Design a Robust, High capacity Watermark", *IEEE Transactions on Image Processing*, ,13(6), 792-807.

4. Bhat G. M, Parah S. A. et.al (**2009).** "VHDL Modeling and Simulation of Data Scrambler and Descrambler for Secure Data Communication", *Indian Journal of Science and Technology,* ,Vol 2, No. 10, pp. 41-43.

5. Bhat G. M, Parah S. A. et.al (**2010**). "FPGA Implementation of Novel Complex PN Code Generator based data Scrambler and Descrambler", *Maejo Int. J. Sci. Technolgy*. , 4(01), 125-135

6. Wong P. W. and Delp E. J. (**2000).**"Security and Watermarking of Multimedia Contents II,". *Society of Photo-optical Instrumentation Engineers*, volume 3971

7. Schyndel R.G., Tirkel A.Z., & Osborne C.F. **(1994).** "A digital watermark" .*Proceeding of IEEE International Conference on Image*", 2, 86-90.

8. Bender W., Gruhl D., et.al, (**1996).** "Techniques for data hiding". *IBM Systems Journal,* , 35, 313-316.

9. P. Wolfgang, & Delp, (**1996).** "A watermark for digital images." *International Conference on Image Processing Proceedings, ICIP*,96,219-222.

10. Wu N. I. and Hwang M. S. (**2010**). "Data Hiding: Current Status and Key Issues," *International Journal of Network Security,* 4(1),1-9.

11. Sathisha N. et. al (**2011**). "Embedding Information In DCT Coefficients Based On Average Covariance" *International Journal of Engineering Science and Technology (IJEST),* ,3 (4), 3184-3194.

12. Zeki A M. and Manaf A. A. (**2009**). "A Novel Digital Watermarking Technique Based on ISB (Intermediate Significant Bit)" , *World academy of science Engineering and Technology*,50,989-996

13. Zeki M. A. et. al,(**2011**). "High watermarking capacity based on spatial domain technique" *Information technology journal*: 10(7),1367-1373

14. Parah S. A. et. al **(2012)**, "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), 10141-10146*

15. Parah S. A. et. al **(2012)**," Data Hiding in Intermediate Significant Bit Planes, A High Capacity Blind Steganographic Technique," *In proc. Of IEEE intl. conf. INCOSET 2012*, 192-197. ISBN : 978-1-4673-5144-7/12.

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 Signal processing and 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 twenty six 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 twenty five research papers in International and National journals and conference proceedings.

Prof. G. Mohiuddin 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 many 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.