Genetic Algorithm Based Image Cryptography to Enhance Security

Rakhi Choudhary
Department of Computer Science& IT
University of Jammu
Jammu, India

Prof. Pawanesh Abrol
Department of Computer Science& IT
University of Jammu
Jammu, India

Abstract: **In today's digital world security, integrity, confidentiality of the organization's data is the main concern. In this study, a new approach for enhancing the security of images using the concept of a new genetic algorithms (GA) is used to output new method of encryption by means of crossover and mutation operation of genetic algorithm. The proposed encryption method in this study is tested on various images which differ in size and type, and effort has been made to improve the efficiency which has shown positive result. Image encryption is widely used to ensure security of the data. Many image related problems can be solved using genetic algorithms through modelling a simplified version of genetic processes. This paper proposed a method based on Genetic Algorithm which is used to generate key by the help of random number generator. Key generation will go through a number of steps to increase the complexity of key. In our approach we have implemented genetic algorithms with bits flipping to encrypt and decrypt data stream. The encryption process is applied over a binary file so that the algorithm can be applied over any type of image. This paper deals with the implementation of GA in MATLAB. A MATLAB code is developed for encryption and decryption of image using cipher. In this study,GA is implemented at keys as well as image level for enhancing the security of image.**

*Index terms:***Encryption, Decryption, Image, Genetic algorithm, Crossover, mutation**

I. INTRODUCTION

Genetic Algorithm has come into existence from the studies of cellular automata, conducted by John Holland and his colleagues at the University of Michigan. A Genetic Algorithm is a technique used in computer science to find approximate solutions to optimization problems. GAs are a particular class of evolutionary algorithms that use approaches inspired by evolutionary biology such as inheritance, mutation, natural selection, and recombination (or crossover). Once we have the genetic representation and the fitness function defined, GA proceeds to initialize a population of solutions randomly, and then make it better through repetitive application of mutation, crossover, and selection operators. Researchers have selected GA as a solution to optimization in various fields in recent years. GA as a solution to optimization problem started gaining acceptance towards the end of the last century as used to solve optimization problems in construction. Its built in parallelism facilitates the uses of distributed processing machines such as Distribution Network Planning. Problems which appear to be particularly applicable for solution by GA include Scheduling and State Assignment Problem. GA approach to Solve Map Color Problem has also been experimented. Researchers have shown interest in GA approach for solving scheduling types of problems, like job shop scheduling problem. It can be quite effective to put together GA with other optimization methods. Hybrid GA approach is also being adopted to derive higher quality solutions in relatively shorter time for complex combinatorial real world optimization problems such as traveling salesman problem (TSP) [1]. Of late, researchers are also trying to analyse the power of GA in various field of research like molecular research and genetic research to identify unknown genes of similar function from expression data [2]. This paper implements the GA in MATLAB. It uses a high level language which is useful to develop many applications such as image processing, control systems, signal processing and communications. Therefore, development of GA in MATLAB will be useful for many applications which required data security.

II. LITERATURE SURVEY

Lot of work is already defined by many researchers on the area of image encryption and decryption with the help of genetic algorithm.Some of the works describes a method which depends on two phases. Substitution phase- here location of pixels and their values are altered from the adjacent pixels to reduce the correlation between the pixels. Modification phase- here pixels values are altered and input image gets encrypted. Both phase works on binary patterns[3]. The method in which rows and columns of image are anyhow dislocated. Then image is partitioned into four equally sub images. Now two pixels are randomly chosen from the set of population. After that crossover and mutation is performed and image get

reconstructed again. Now entropy is checked for the image, if it increases then image is used for the next step and lastly randomness is measured by entropy, histogram analysis and coefficient correlation[4]. The concept of molecular genetics and image pattering deals with molecular genetic that is meiosis, fertilization, translation and mutation. At second step data is encrypted in the form of the image. To get highest level of data security; both the concepts work together. A key is generated, and then with the help of that key so many sub keys are generated that will be used for encryption. Then several patterns are developed for respective keys. Rather than image, it works on knowledge of image pattern, which directly gives the information on the basis of key pattern of image. Demerits- it takes larger percentage of overall time for coding[5]. The genetic algorithm, image encryption and video encryption with the help of physical model can work on cryptography. Proposed method worked upon signal and image processing[6]. Integration of two methods are done first is- symmetrical system using the GA and another is asymmetrical system using RSA cryptography. With the help of this method a strong key can be generated and can be made non-repeating too, because of this it is not so easy to break it at all. This method provides high level of security of data and information. Author takes GA as a base for developing key and generates a new block cipher system. This algorithm is much better than DES, AES, ELLIPTIC CURVE, RSA, NTRU etc. Merit- no one can break the code of key yet[7]. Images on internet or any other transmission medium can be secured by e- security with the help of GA and pseudorandom sequence, just to encrypt and decrypt the data and generates a pseudorandom sequence with the help of nonlinear feedback shift register{NLFSR}, then apply crossover operator to encrypt data over pseudorandom sequence, then mutation is applied on the binary patterns. Merit- speed of algorithm is good at the time of encryption, safe and reliable because of lack of knowledge about pseudorandom sequence and mutation string[8]. A symmetric key cryptosystem with the help of GA is defined firstly plain text is converted in the form of matrix which is key matrix and text matrix, secondly additive matrix is produced by adding both the text and key matrix as well. Now substitution function is performed to produce intermediate cipher on additive matrix and then crossover and mutation is applied. Merit- easy and simple to operate this technique, high security because of key generation and additive cipher technique[9].GA is used to produce a new method of encryption which also covers strong properties of crossover and mutation, here implementation is done on an image with some width and height. Author uses visual C++ 6.0 programming for the implementation and recorded that noise is 0%. Merit- no data is lost in encryption and decryption process[10]. When a high level of security is required, symmetric and asymmetric methods doesn't work. To obtain a perfect result within least time, author proposed an algorithm in which GA is combined with cryptography and output of that combination is an optimal solution. Here author gives some of the limitations of hash functions and digital signature. MATLAB 7.8.0 platform is used. By the help of random number generator, pseudo random numbers are generated. Data transmission in the form of image over the internet, analysis of encryption and decryption process in respect with time and an algorithm is proposed on graphs as well. Values of colors are declared on basis of bit level gray scale. 0 is for black and 255 is for white. Merit- this algorithm gives better throughput and solution with in required time[11]. A new method is proposed in which a key is generated by pseudorandom number generator and these random numbers will be developed with the help of current time of the computer system. Now GA is applied on it and image encryption is done with the help of [AES] symmetric key algorithm. Merit-this algorithm increases the efficiency and decrease the computational time, irregularity of key increases the complexity of key[12]. The concept of genetic algorithms with pseudorandom function is used for encryption and decryption of data stream .The encryption process is applied over a binary file so that the algorithm can be applied over any type of txt and multimedia data as well[13].The method based on Genetic Algorithm is used for generating key by the help of random number generator to make the key complex. Key generation goes through a number of steps and main criteria for key selection is the fitness value of the population[14]. The quality of the random numbers produced by the current algorithm is superior, then the key that is produced will always be much excellent. Author uses a threshold value for this selection. Basically to check the randomness of the sample coefficient of correlation has been used[15].

III. EXPERIMENTAL SETUP

The encryption and decryption algorithm are implemented in MATLAB R2013a.MATLAB is a high level language and a comfortable platform for numerical calculations, visualization and programming. This platform can be used to create models, develop algorithms, and applications. Using built in math functions of this platform makes it possible for us to achieve the results much faster than spreadsheets or conventional programming languages, like C/C++ or Java. With the help of MATLAB compiler one can generate executable code for any computer system for different environments. . This platform has wide range of applications in the areas including signal processing and communications, image and video processing, control systems, test and measurement, computational finance, and computational biology. A new script is to be opened and the program is to be written into new script and the file has to be saved in destination folder which is required for image encryption and decryption.

A. Proposed Work

In order to measure the strength of this encryption process against illegal decryption attempts, we consider the selecting two strings from one block that have very good correlation properties. To decrypt one block, one has to find the 16 bit with possible number 2^{16} and to decrypt the whole image $2^{16*no.of\ blocks}$ guesses have to be made. The strength is the most essential feature that a good qualityencryption algorithm should possess. If the encryptionalgorithm is unable to prevent all types of attack includingstatistical and brute force attacks, it will not be sufficientfor protecting the image data. The experiments were carried outfor defining the competency of the proposed technique. Inthis research, the proposed technique is applied on the images which have different formats and sizes.

B. The Proposed Algorithm

1. Generate the sequence of pseudorandom binary numbers using the random number generator for 64 bit key generation.
2. Convert the Hexadecimal key into its binary equivalent.
3. Convert the binary pseudorandom sequence into decimal and then binary sequence ranging from 0 to 63.
4. Divide key into 16X4 keys and follow the following steps.
5. Read 16 consecutive bytes from the data file.

6. Modify the consecutive bytes using transposition and transformation.
7. Take two consecutive bytes of the data stream as P1 and P2.
8. Perform crossover and then mutation on two consecutive bytes of the data stream as Q1 and Q2 by using number Yi.
9. Encrypt data as C1 and C2, where,
   Crossover (By circular shift)
   Xi = Yi ⊕ (Circular shift on Yi)

   Xi+1 = Yi+1 ⊕ (Circular shift on Yi+1)
   Flip bit mutation {By NOT (¬) operation}
   C1 = ¬(Q1 & Xi)
   C2 = ¬(Q2 & Xi+1)
10. Repeat steps 6 to 9 until complete 16 bits are encoded.
11. The decryption process are just reverse of the encryption process as mentioned above.

IV. RESULTS ANALYSIS

The analysis of cryptographic image was carried out using genetic algorithm in MATLAB. The images used in the experiment are mentioned in the Table (1) and following observations were made.

Table I. Showing the Details of Input Images Used in the Experiment

Sr. No.	*Input image Name*	*Image Type*	*Dimensions*	*No. Of Pixels*
1	cameraman	tif	256x256	65536
2	Rice	tif	256x256	65536
3	bridge	tif	259x194	49664
4	nuts	tif	300x168	50400
5	sunflower	tif	250x250	62500
6	baboon	bmp	256x256	65536
7	boy	bmp	203x249	50547
8	Lena	bmp	512x512	262144
9	bird	bmp	485x303	146955
10	bunch	bmp	443x332	147076
11	butterfly	png	409x256	104704
12	whitedog	png	332x300	99600
13	rose	png	442x333	147186
14	cat	png	470x313	147110
15	dog	png	480x319	153120
16	football	jpg	320x256	81920
17	building	jpg	438x336	147168
18	flower	jpg	300x225	67500
19	baby	jpg	580x387	224460
20	veg	jpg	757x572	433004

Table II. Showing the Formulas Against Each Security Assesment Parameters Used in the Experiment

Sr. No.	*Name of Parameter*	*Full Form*
1	$\text{SSIM} = \frac{(2 \times \bar{x} \times \bar{y} + C1)(2 \times \sigma_{xy} + C2)}{(\sigma_x^2 + \sigma_y^2 + C2) \times ((\bar{x})^2 + (\bar{y})^2 + C1)}$	Structural Similarity Index
2	$\text{NCC} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j) \times y(i,j))}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j))^2}$	Normalised Cross Co-relation
3	$\text{SC} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(y(i,j))^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j))^2}$	Structural Content
4	$\text{AD} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j) - y(i,j))$	Average Difference
5	$\text{NAE} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\lvert y(i,j) - x(i,j)\rvert}{\sum_{i=1}^{M}\sum_{j=1}^{N}\lvert y(i,j)\rvert}$	Normalised Absolute Error

Table III. Showing the Observed Values Against Parameters as Mentioned In Table (2) During the Experiment

Sr. No.	*Image Name*	*Elapsed Time (Sec)*	*SSIM*	*NCC*	*AD*	*SC*	*NAE*
1	cameraman	295.24	0.0071	0.87	-17.07	0.74	0.71
2	Rice	297.40	0.0099	0.97	-12.86	0.67	0.64
3	bridge	226.71	0.16	0.80	1.19	0.95	0.59
4	nuts	225.82	0.0155	0.96	-25.62	0.60	0.79
5	sunflower	296.97	0.087	0.86	-19.20	0.72	0.78
6	baboon	294.28	0.0152	1.55	-54.01	0.25	1.14
7	boy	257.49	0.0117	0.99	-35.53	0.52	0.90
8	Lena	1315.89	0.0171	0.92	-5.73	0.78	0.57
9	bird	707.29	0.0133	1.07	-20.63	0.59	0.64
10	bunch	698.88	0.0165	1.18	-27.65	0.47	0.74
11	butterfly	495.16	0.0234	0.72	29.45	1.47	0.45
12	whitedog	467.81	0.0135	0.78	20.25	1.15	0.47
13	rose	701.39	0.0093	1.42	-56.93	0.25	1.34
14	cat	703.52	0.0087	0.85	-58.64	0.50	1.28
15	dog	728.09	0.0123	1.00	-25.29	0.58	0.77
16	football	461.66	0.013	1.27	-40.80	0.36	0.98
17	building	848.82	0.0133	0.92	-8.16	0.76	0.61
18	flower	378.15	0.0133	1.07	-49.68	0.41	1.10
19	baby	1332.79	0.0073	0.86	-58.22	0.47	1.30
20	veg	2757.22	0.0103	1.10	-36.99	0.45	0.95

A. *Analysis of Security Paremeters*

1) SSIM (Structural Similarity Index): SSIM is used for measuring the similarity between two images. The SSIM is indication of image perception similarity. Higher the value of SSIM more perceptionally similar the images are concerned. In this research the value of SSIM for original image against encrypted images was observed to be very small. It clearly indicates the efficiency of the encryption process. The maximum observed was 0.16 as shown in fig 1.(SSIM) index gives the visual impact of shifts in image luminance, changes in photograph contrast, as well as any other remaining errors, collectively identified as structural changes. The observed SSIM value was much less than 1.0, which is indication of strong encryption. The normal value should have been 1.0 for similar images.

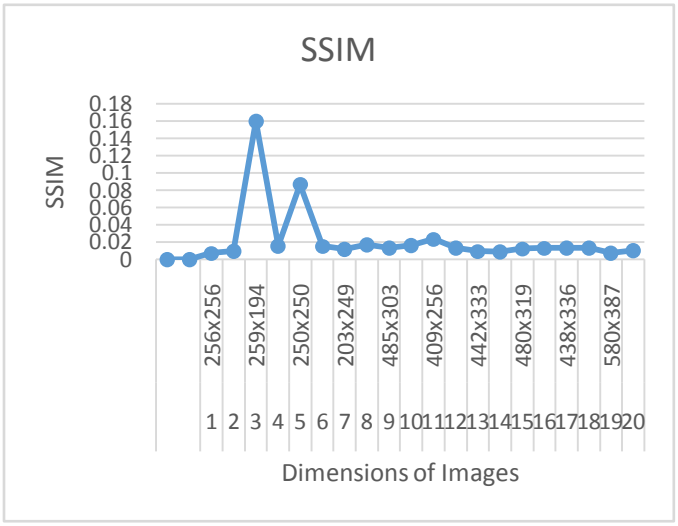

Figure 1. SSIM Against Each Image

2) Structural Content (SC): SC is also correlation based measure and measures the similarity between two images. The structural content was observed to be maximum against images with large size. While the minimum value of structural content was recorded for images with small size with value < 0.4 > as shown in fig 2.The observed value was greater than 0.5 in most cases, clearly indicating large difference between original and encrypted image. The normal value should have been equal to 1 for similar images.

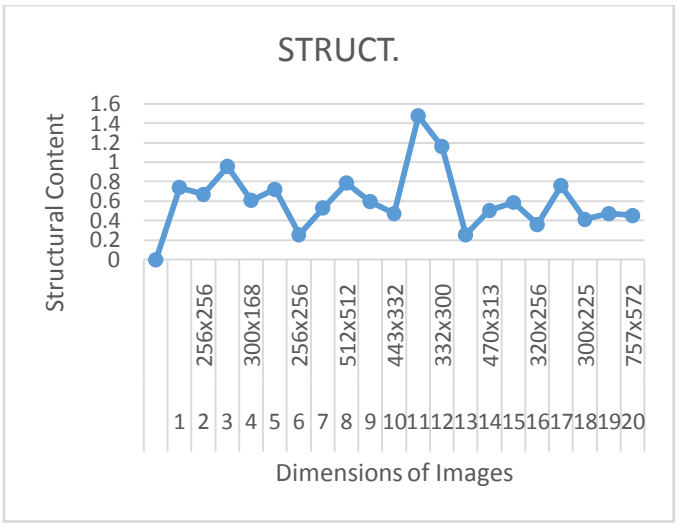

Figure 2. Structural Content Against Each Image

3) Normalised Cross Correlation(NC): The closeness between two digital images can also be quantified in terms of correlation function. The large value of NCC means that image is good in quality. The observed NCC values were less than 1.0 in most of the cases, clearly indicating dissimilarity between original and encrypted image. The normal value between two similar images should equal to 1.

4) Average Differenc(AD): AD is simply the average of difference between the reference signal and test image. Average Difference is measurement of differences between two images. The observed AD was less than 0; again clear indicator of the fact that there is large difference between the original and encrypted image. The normal AD should be 0 for similar images.

5) Normalised Absolute Error(NAE): The observed NAE was greater than 0.5 as shown in fig 3, which is clear indicator of encrypted image being different compared to the original one, which in normal case would have value of 0.

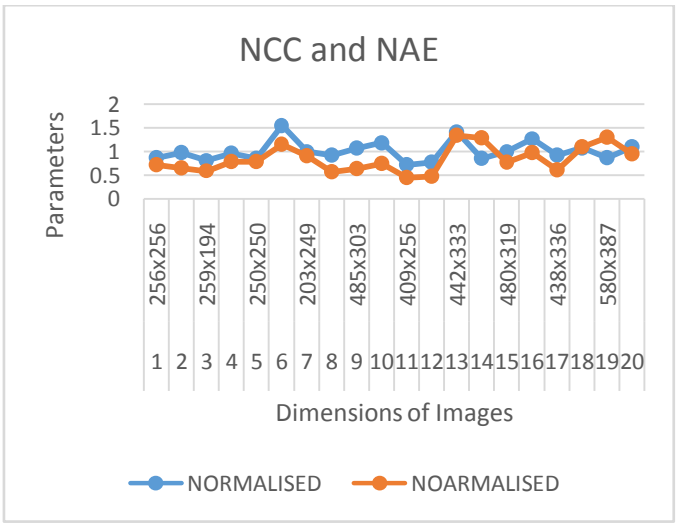

Figure 3. NCC and NAE Against Each Image

6) Elapsed Time: The speed is one of the important factor against a good encryption algorithm. In this research algorithm speed was measured for images of various types and sizes.It was observed that total elapsed time including encryption as well as decryption for larger images was more compared to small size images as shown in fig 4.

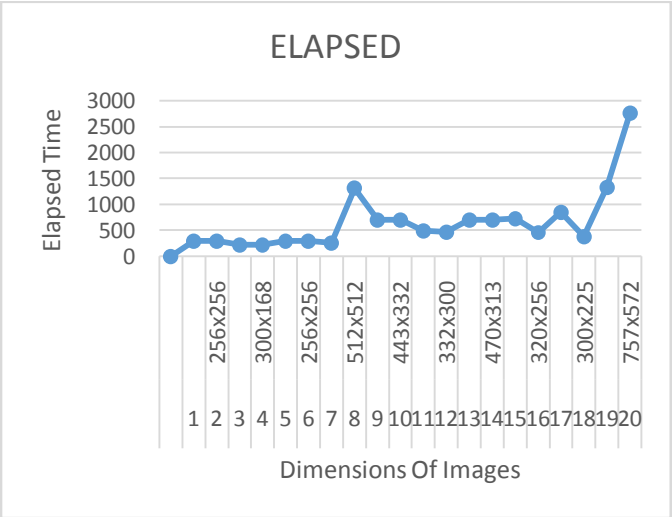

Figure 4. Elapsed Time Against Each Image

B. *Histogram Analysis*

An image histogram is a commonly used method of analysis in image processing applications. One of the benefits of the histogram is that it is able to present the distribution for a large set of data .The image histogram provides a clear illustration of how the pixels in an image are distributed by graphing the no. of pixels at each color intensity level. From security prospective it is essential to make sure that the encrypted and original images possess different statistics. The histogram analysis represents the ways that pixels in an image are distributed at each intensity level. The Figures show the results of the experiment on the original image, its corresponding encrypted image and their histograms. The findings show that the histogram of the encrypted image is significantly different from the respective histograms of the original images. Fig 5, fig 6, fig 7 and fig 8 makes it clear that there is a lot of difference between histograms of original and encrypted image.

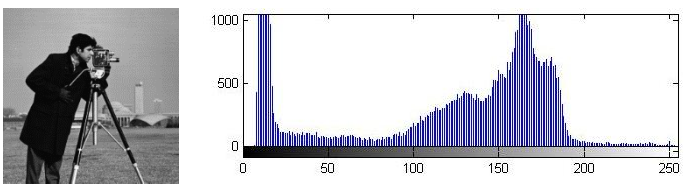

Figure 5. Cameraman Original Image and its Hisogram

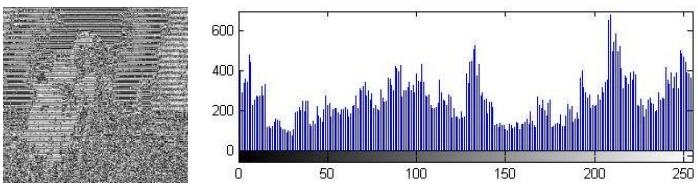

Figure 6. Cameraman Encrypted Image and its Hisogram

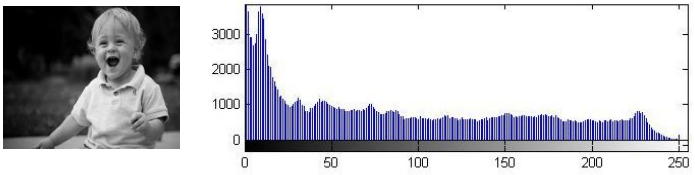

Figure 7. Baby Original Image and its Hisogram

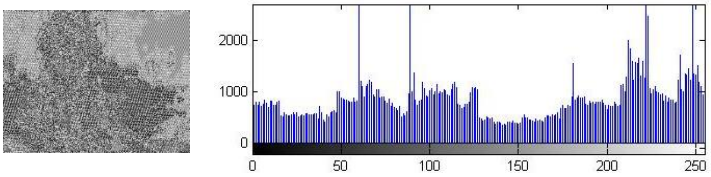

Figure 8. Baby Encrypted Image and its Hisogram

V. CONCLUSION

All the above observations clearly justifies the fact that the encrypted image being much more different than the original image. In other words, strong encryption of image is indicated indirectly indicating strong security of image which justifies the argument that greater the dissimilarity values for encrypted image, greater is the security.

REFRENCES

[1] Gondro C, Kinghorn BP, "A simple Genetic Algorithm multiple sequence alignment", Genetics and Molecular Research, Volume6(4), pp. 964–982, 2007

[2] Cuong C To, Vohradsky J., "A parallel Genetic Algorithm for single class pattern classification and its application for gene expression profiling in Streptomyces coelicolor", BMC Genomics, Volume8,2007.

[3] R. Afarin., S.Mozaffari2013.”Image encryption using genetic algorithm and binary patterns” at MVIP and IEEE.

[4] R. Afarin., S.Mozaffari2013.,”Image encryption using genetic algorithm” at MVIP and IEEE .

[5] M Prashant, R Siddhartha and Rajeev Kumar2011.”Formulation of an encryption algorithm on the basis of molecular genetics and image patterns” at IEEE.

[6] A. Tragha, F. Omary, A. Mouloudi2006.”ICIGA: improved cryptography inspired by genetic algorithms” at IEEE.

[7] Abdel-karim S.O. Hassan, Ahmed F. Shalash and Naglaa F. Saudy MAY 2014.,”Modifications on RSA cryptosystem using genetic optimization” at IJRRAS 19 (2).

[8] P.Singh, G. Gosawi, S. Dubey 4 (2014). ”GA: A technique for cryptography real time data transmission” at binary journal of data miming and networking 37-40.

[9] Sindhuja K, P. Devi2014,414-416.,"A symmetric key encryption technique using GA" at IJCSIT, volume 5(1).

[10] Mohammad A.F.Al-Husainy2006, 516-519..,"Image encryption using GA" at ITJ 5 (3).

[11] Dr. D. Singh,P.Rani, Dr. R. Kumar2013.," To design a GA for cryptography to enhance the security" at issue 2April 2013.

[12] A.Soni, S.Aggarawal2012.,"Using GA for symmetric key generation in image encryption" at IJARCET 2012.

[13]Suvajit Dutta, Tanumay Das, SharadJash, DebasishPatra, Dr. Pranam Paul, "A Cryptography Algorithm Using the Operations of Genetic Algorithm & Pseudo Random Sequence Generating Functions**"** - International Journal of Advances in Computer Science and Technology Volume 3, No.5, May 2014.

[14]AartiSoni, Suyash Agrawal,"Key Generation Using Genetic Algorithm for Image Encryption**"** International Journal of Computer Science and Mobile Computing IJCSMC, Vol. 2, Issue. 6, June 2013.

[15] S. Goyat 2012, "Genetic Key Generation for Public Key Encryption Cryptography", (IJSCE) ISSN: 2231-2307, Volume-2nd, Issue-3rd, July 2012.

RakhiChoudhary,M.Tech Department of Computer Science and IT, Jammu University

Prof. PawaneshAbrol, Ph.D. (Computer Science), MBA (HR), 30 publications in journals and conferences.