Volume 4, Issue 10, October 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Cryptography and the Optimization Heuristics Techniques

P. Saveetha
Associate Professor/Dept of IT
Nandha College of Technology
Erode, India

Dr. S. Arumugam
Principal
Nandha Engineering College
Erode, India

K. Kiruthikadevi
Assistant Professor/Dept of CSE
Nandha College of Technology
Erode, India

Abstract—***Network security is an ever green field and cryptography is one part of the network security. Cryptography is converting of original message into cipher text which is an unreadable form. The cryptanalysis is recovering of the original message from the cipher text. Many algorithms are available in cryptography architecture. In this paper the computational complexity of the Cryptanalysis is studied and that can be reduced by using Genetic algorithm, and tabu search is compared. The applicability of Genetic algorithms for searching the key space of encryption scheme is presented. In Vigenere cipher, guessing the key size is done by applying Genetic Algorithm. The frequency analysis is used as an essential factor in objective function.***

Keywords— Poly-alphabetic cipher; Vigenere cipher Particle swarm optimization; Simulated Annealing; genetic algorithm; Tabu search; firefly algorithm.

I. INTRODUCTION

Cryptanalysis is the method to get the plaintext and or key from a cipher text. Cryptography is the art and science of rendering data unreadable for all but the intended recipients [1, 6]. The purpose of a cryptosystem is to keep no others from reading that original data. Plaintext refers to the original message before any encryption and cipher text refers to the encrypted text. If the unknown person finds the key, he can be able to recover the original message. If the key size is larger, a search method is preferable rather than trying every key [9].

There are two types of cryptographic schemes namely symmetric cryptography and asymmetric cryptography. The symmetric scheme uses the same key for encoding and decoding process. Two keys is used in asymmetrical cryptography , one for encoding process, known as the public key, and the other for decoding process, known as the private key. Asymmetric cryptography is often used for key distribution and digital signature but its processing speed slow. The symmetric cryptography is normally used to encode the private data for its high performance.

The growth of the natural language programming is very suitable in the field of cryptography. Many types of natural language programming are available.

The genetic algorithm, simulated annealing, tabu search, particle swarm optimization, firefly algorithm are analyzed and the applications of the natural language programming are studied here. Genetic Algorithm (GA) is the best optimization - heuristics techniques to reduce the time for key search and computational complexity while breaking any type of the cipher. In artificial intelligence also genetic algorithm, tabu search and PSO are used for various applications.

Use of the GA and frequency analysis is the best way for finding length of the keyword. Other papers [12, 13, 14, 15, 16, 17, 18] done more research on genetic algorithms for Cryptanalysis. Substitution ciphers are vulnerable to frequency analysis. The Genetic algorithm is used to search of a key space and to guess the key.

Mr. Vinod Saroha., et.Al, [30] replace the letter with any other letter and transposition method only change position of characters.

Ekta Agrawal et.Al.,[31] analyzes the security information with enhanced speed of encryption and decryption process. Some variation makes in converting the plain text into cipher text for making data more secure so that the unauthorized user cannot access the data and cannot understand the cipher text easily but same cipher text for identical alphabets.

Sonia Dhull.,et.al., 2013 present a perspective on Poly-alphabetic techniques which are currently used only for encryption purpose.

The rest of the paper is organized as follows. Section 2 describes about the genetic algorithm and its steps and different surveys with their limitations. Section 3 details the tabu search with the results and discussion. Section 4 presents the simulated annealing technique with the results and discussion. Section 5 evaluated the results and discussion on firefly algorithm.

Section 6 presents the particle swarm optimization. Section 7 and 8 describes the algorithms for tabu search and genetic algorithms. The final two sections that are 9, 10 will summarize results of GA with one example.

II. GENETIC ALGORITHM

Genetic Algorithm is a general method of solving problems to which no obvious solution exists. It is based on the idea of the evolution of a species in nature and so the various components of the algorithm are roughly analogous to aspects of natural evolution. Common mathematical tasks amenable to genetic solutions include computing a curve to fit a set of data or approximating Non-deterministic Polynomial (NP) problems. Often these operators consist of flipping a single random bit of one individual or swapping two randomly selected substrings from a pair of [2] parents to generate a new child. To simulate Darwinian survival of the fittest some representation of the fitness of the individuals must be generated. Joseph Alexander Brown [19] has shown the result in his paper as how to find the weak keys using genetic algorithm technology. The work has also shown the comparison between known plain text and chosen plain text. The work has shown the result as the genetic algorithm can be used in other system attacks.

In genetic algorithm we have two main operations: Crossover and mutation. Crossover is a technique in which the bits are interchanged [22]. In crossover we have many types such as:

a. Single-point crossover
b. Two-point crossover
c. Uniform crossover
d. Multi-point crossover
e. Three-point crossover

Mutation is a process used to maintain the diversity from one generation to the next generation. The main purpose is to preserve and introduce diversity [23].

There are two types in mutation [6]:

a) *Flip Mutation:* Flip Mutation causes one bit to be randomly selected within the chromosome and then flipped, a 1 is changed to a 0 and a 0 is changed to a 1.

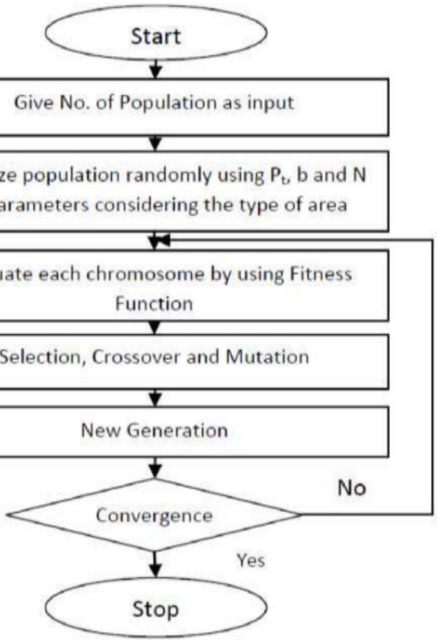

Start
Give No. of Population as input
Initialize population randomly using P_t, b and N as parameters considering the type of area
Evaluate each chromosome by using Fitness Function
Selection, Crossover and Mutation
New Generation
Convergence
No
Yes
Stop

Fig. 1. Flow chart for Genetic Algorithm

Before Mutation: 1010110000110101
After Mutation: 1011110000110101

b) Swap Mutation: In Swap Mutation two random bits are selected and exchanged within the chromosome.

Before Mutation: 1010110000110101
After Mutation: 1011110000110100

The advantages of GA are: parallelism, liability, wider solution space, complexity in the fitness landscape, and easy discovery of global optimum. It encounters some limitation too, they are: the problem to identifying the fitness function, computational time, definition of representation of the problem and occurrence of premature convergence. This is also a secured algorithm which can be used in data transfer in network and cloud computing.

III. TABU SEARCH

Fred Glover proposed tabu search in 1983 to allow local search methods to overcome local optima. Tabu search is to perform local search in local optimum by allowing non-improving moves. We maintain a list called tabu list to record the recent history of the search.

It is also one of the optimization heuristics techniques which prevent the search from returning to a previously explored region of the solution space immediately. Hence it maintains a list of previous solutions, which are called ‘tabu’; hence the name of the technique. Performance of the algorithm [11] depends on the size of the tabu list. In the tabu list two randomly chosen key elements are swapped to generate candidate solutions. In each iteration one worst solution in the tabu list will be eliminated with newer one. Verma et al [20] compared tabu search and genetic algorithm and concluded that tabu search is more powerful than genetic algorithm in cryptanalysis of mono-alphabetic substitution cipher. Ho yean[11] compared tabu search and genetic algorithm in the cryptanalysis of transposition ciphers, Poly-graphic substitution cipher(Hill cipher) and product cipher(the AES cipher).The work has shown that transposition cipher is more vulnerable in tabu search and genetic algorithm attack and the poly-alphabetic substitution cipher is also susceptible in tabu search and genetic algorithm. Advanced Encryption Standard (AES) cipher is vulnerable to weak password attacks.

IV. SIMULATED ANNEALING

In 1983, Krik- patrick et al proposed the simulated annealing technique which is derived from thermo-dynamic considerations with annealing interpreted as optimization procedure. Simulated annealing (SA) is used to solve the combinatorial optimization problem. It is the process of slowly cooling a heated metal in order to attain minimum energy state.

Nalini.N et al [21] have shown the experimental result as that the simulated annealing is good at performance in cryptanalysis of Simplified Data Encryption Standard.

V. FIREFLY ALGORITHM

Firefly algorithm was developed by Xin- She Yang at 2007 which is a population based meta-heuristic algorithm. In this algorithm we create a virtual swarm of fireflies where each firefly is allotted a random position in space. Each firefly represents the solution and its brightness depends on the fittest function value. The firefly with best fittest function value will be brighter. Other fireflies will be attracted towards the brighter firefly.

Saibal K.Pal et al [24] have found that the fitness value will be increased by increasing the total number of generations and also observed that the algorithm works best for large input cipher text lengths.

C.S.Rai et al [25] have shown that firefly algorithm is more powerful in solving the noisy non-linear optimization and it also improves its own space from the previous stage.

VI. PARTICLE SWARM OPTIMIZATION

Particle swarm optimization is developed by Kennedy and Elbert in 1995; it is used to solve many optimization problems. The algorithm consists of swarm of particles. The fitness of each particle represents the quality of the position. The particles will fly over the search space in a velocity and that velocity is the best position and best solution found by its neighbors.

C.S.Rai et al [25] worked on have observed that Particle Swarm Optimization is better in terms of speed of convergence.

VII. ALGORITHM FOR GA

GA is applied in four [4] steps:-

1. Initialize algorithm variables: G the maximum number of generations to consider, M the solution pool size and any other problem dependent variables.
2. Generate an initial solution pool containing M candidate solutions.
3. For G iteration, using the current pool:
   a. Select a breeding pool from the current solution pool and make pairings of parents.
   b. For each parental pairing, generate a pair of children using a suitable mating function.
   c. Apply a mutation operation to each of the newly created children. Evaluate the fitness function for each of the children.
   e. Based on the fitness of each of the children and the fitness of each of the solutions in the current pool, decide which solutions will be placed in the new solution pool. Copy the chosen solutions into the new solution pool.
   f. Replace the current solution pool with the new one. So, the new solution pool becomes the current one.

4. Choose the fittest solution from the final generation.

VIII. TABU SEARCH ALGORITHM

Tabu search algorithm [10] is presented here:-

1. Input: Intercepted cipher text, the key size P, and the language statistics.
2. Initialize parameters: The size of the tabu list STABU, the size of the list of possibilities considered in each iteration SPOSS, and the maximum number of iterations MAX.
3. Initialize the tabu list with random and distinct keys and calculate the cost for each key in the tabu list.
4. For I =1,..., MAX do:
   a. Find the best key with the lowest cost in the current tabu list, KBEST.
   b. For j=1,..., SPOSS do:
   i. apply the perturbation mechanism to produce a new key KNEW.
   ii. Check if KNEW is already in the list of possibilities generated for this iteration or the tabu list. If so, return to step 4(b) i.
   iii. Add KNEW to the list of possibilities for this iteration.
   c. From the list of possibilities for this iteration, find the key with the lowest cost, PBEST.
   d. From the tabu list, find the key with the highest cost, TWORST.

e. While the cost of PBEST is less than the cost of TWORST: i. Replace TWORST with PBEST.
ii. Find the new PBEST.
iii. Find the new TWORST.

5. Output the best solution from the tabu list, KBEST (the one with the least cost).

IX. CRYPTANALYSIS AND GENETIC ALGORITHM

Cryptanalysis is the method to get the plaintext and /or key from a cipher text. Genetic Algorithm (GA) is the best optimization heuristics techniques to reduce the computational complexity while breaking a [2], [8] Vigenere cipher. Method for breaking a Vigenere cipher includes two steps:

1. Finding the length of the keyword
2. Dividing the message into many simple substitution cryptograms.

Use of the GA and frequency analysis is the best way for finding length of the keyword. Substitution ciphers are vulnerable to frequency analysis. Use of a GA to conduct a directed random search of a key space and to [7] guess the key size can suppress the normal frequency data by using more than one alphabet to encrypt the message.

Jun Song et al [5] has shown the result that genetic algorithm was used to crypt analyze the four round DES. The fitness function of genetic algorithm is flexible and also can be applied to complex block cipher. It is hard to define the fitness function in crypt analyze of four round DES. Hence they produced adequate optimum keys based on fitness function. It is difficult to produce high fitness function value for each search and it can be solved by using multi plaintext/cipher text pairs in experiment. The auxiliary fitness function can also be used to improve the evaluation measure.

Guang Gong et al [26] proposed the selective discrete- fourier spectra attack on stream ciphers. They have divided work into two parts. First part is about DFT analysis and second part is new fast selective DFT attack which is related to fast algebraic attacks. They have used linear feedback shift register for more efficient analysis. Hence the attack imposes strong Boolean functions defined in the spectral immunity of the sequence. The use of standard operators in genetic algorithm based pseudorandom series will consume 5% to 10% lesser time.

Genetic algorithm attacks lead to the creation of security against substitution permutation network. This work is proved by Bethany Delman et al [7]. We can say that genetic algorithm can be used to break non-classical cipher which is an emerging field of research.

While using Genetic algorithm search technique in DES, we can find many weak keys. The genetic algorithm does not see a cipher in the same way a cryptographer using mathematical methods does [6].

The Matthews [27] et al attacks the columnar transposition cipher and which was successful only for key length 7. The Clark et al[28] attack was successful on a block length of 8, slightly successful on a block length of 16, and unsuccessful on a block length of 32. The cipher text attacked seemed to play a larger role in the attack's success.

D.Sahoo [29]et al shown the result on mobile host network that integration of genetic algorithm and secure key exchange makes the security system robust in order to verify authenticity of any client nodes entering into the network.

Joseph Alexander Brown et al [6] showed the result on use of genetic algorithm in substitution permutation network cipher. Location of weak keys was found from the process of selecting key classes using genetic algorithm. Then weak keys are combined with substitution permutation network properties to handle the round keys existent. It is illuminated using key class selection method.

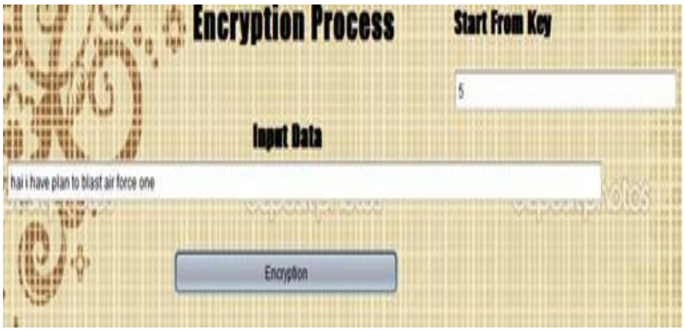
X. POLYALPHABETIC SUBSTITUTION CIPHER

An extension in mono alphabetic cipher is called Poly-alphabetic Substitution Cipher. The message is broken into blocks of equal length; the block is encrypted using a different simple substitution cipher key. The message is "Growth of work is success" and keyword is permutation, then the encryption process is shown.

*Keyword: PERMUTATIONPERMUTATIO
*Plaintext: GROWTHOFWORKISSUCCESS
*Cipher text: VVFINAOYECEZMJEOVCXAG
The decryption will be,
*Keyword: PERMUTATIONPERMUTATIO
*Cipher text: VVFINAOYECEZMJEOVCXAG
*Plaintext: PERMUTATIONPERMUTATIO

For both encryption and decryption vigenere table is used. The individual letter of the keyword and the intersection letter given by the individual letter of original message present at the column are chosen and the cipher text is recovered. The genetic algorithm is applied in stream cipher and results of encryption and decryption were studied.

A transformation plays the role of individual in the GA, because it provides solution to the spectral immunity. The key's length finds only the key, i.e. the transformation with the fixed length 'n'. The first simple solution which seems the most evident is to associate separate genes with elements of the transformation 'T', i.e. associate the i th gene with the number pi. Obviously in that case genes will depend on every other. If some gene (i.e) bits is equal to j in GA, then no other gene of this individual can be equal to j, because all numbers between 1 and 'n' are presented only once in the transformation 'T'. The dependence of bits results in important limitations to operators of mutation and crossover.

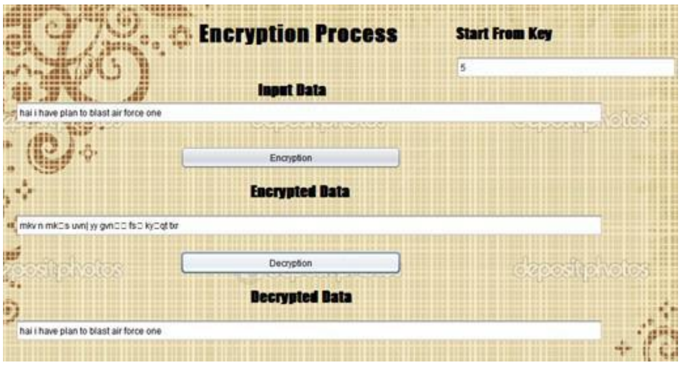


Encryption Process

Start From Key: 5

Input Data: hai i have plan to blast air force one

Fig. 3. Input Message to be Encoded



Encryption Process

Start From Key: 5

Input Data: hai i have plan to blast air force one

Encrypted Data: mkv n mk□s uvnj yy gvn□□ fs□ ky□qt bx

Decrypted Data: hai i have plan to blast air force one

Fig. 4. Secret data obtained in Decoding process with GA

series because all of them work with a string of independent bits for effective privacy. This leads to prejudice in the efficiency of using genetic algorithms. But this association is subconsciously comprehensible and doesn't need any additional computations to form bits in GA based pseudo-random series.

To compare GA with pseudo-random series process with fast discrete Fourier spectra approach on stream cipher is implemented in the java platform. The experiment of encoding and decoding process is conducted on the input secret data.

Fig 3 and 4 performs the encoding and decoding process on the stream cipher to enhance the security. The encrypted data in fig 4 uses the different spectral immunity values, so that the same characters in the secret data contains the different encrypted symbols. Encoding process on the stream cipher using the GA with pseudo-random series method consumed lesser memory and time when compared with the existing Fast Discrete Fourier Spectra approach on stream cipher.

XI. CONCLUSION

The natural language programming is very suitable in the field of cryptography. In terms of Performance and measure, the Genetic algorithm produces good results in transposition cipher bad result in substitution cipher and modern ciphers. Genetic algorithm produces better results when compared with tabu search algorithm. Based on the survey it has shown that Genetic algorithm and tabu search gives better performance when compared to the other optimization techniques. A result shows the probable of GA with pseudorandom series provide automated assisted un-breaking of ciphers. The GA splits the key set effectively from the small input texts on stream ciphers during decoding process when compared with the Fast Discrete Fourier Spectra approach on stream cipher.

REFERENCES

[1] Albassal A. and Wahdan A., "Genetic Algorithm Cryptanalysis of the Basic Substitution Permutation Network," in Proceedings of the 46th IEEE International Midwest Symposium on (MWSCAS'03), pp. 471-475, 2003.

[2] Ragheb Toemeh and Subbanagounder Arumugam ," Applying Genetic Algorithms for Searching Key- Space of Polyalphabetic Substitution Ciphers" , The International Arab Journal of Information Technology, Vol. 5, No. 1, January 2008.

[3] Dimovski A. and Gligoroski D., "Attack on the Poly-alphabetic Substitution Cipher Using a Parallel Genetic Algorithm," Technical Report, Swiss- Macedonian Scientific Cooperation Trough SCOPES Project, Ohrid, Macedonia, March 2003.

[4] Clark A. and Dawson E., "Optimisation Heuristics for the Automated Cryptanalysis of Classical Ciphers," Journal of Combinatorial Mathematics and Combinatorial Computing , vol. 28, pp. 63-86, 1998. Joseph Alexander Brown, Sheridan Houghten, Member, IEEE, and Beatrice Ombuki-Berman, "Genetic Algorithm Cryptanalysis of a Substitution Permutation Network", 2009 IEEE.

[5] Bethany Delman, "Genetic Algorithms in Cryptography", Master's thesis, Rochester Institute of Technology,

July 2004.

[6] Spillman R., Janssen M., Nelson B., and Kepner M., Use of a Genetic Algorithm in the Cryptanalysts of Simple Substitution Ciphers , Cryptologia, vol. 16, no. 1, pp. 31- 4, January 1993.

[7] Stallings W., Cryptography and Network Security: Principles and Practices, 4th edition, Pearson Education, 2007.

[8] Glover Fred, Taillard Eric and Werra Dominique de, "A User's Guide to Tabu Search" Annals of Operations Research, Vol. 41, pp. 3-28, 1993.

[9] Ho yean Li,Azan samsudin, and Bahari Balaton, “Heuristic Cryptanalysis of Classical and Modern Ciphers” pp 710-715, IEEE 2005.

[10] Julio César Hernández,”Easing collision finding in cryptographic primitives with genetic algorithms”, pp.535-539, IEEE 2002.

[11] Matthew Russell, John A.Clark, Susan Stepney, “Using Ants to Attack a Classical Cipher”, GECCO2003, LNCS 2723, 2003, pp.146-147.

[12] Ralph Morelli, Ralph Walde, William Servos, “A Study of Heuristic Approaches for Breaking short Cryptograms” , International Journal on Artificial Intelligence Tools, World Scientific Publishing Company, 2004 Vol.13.No.1(2004)45-64.

[13] Abbas Ghaemi Bafghi, Babak Sadeghiyan, “Finding Suitable Differential Characteristics for Block Ciphers with Ant Colony Technique” , IEEE 2004, pp.418-423.

[14] John A.Clark, Jeremy L.Jacob, Susan Stepney, “The Design of S-Boxes by Simulated Annealing” ,IEEE 2004, pp.1533-1537.

[15] Eng.Ayman M.B.Albassal1, Abdel-Moneim A.Wahdan, “Genetic Algorithm Cryptanalysis Of The Basic Substitution Permutation Network”,IEEE 2004, PP.471-475.

[16] Mohamed Amine Garici1,Habiba Drias2, “Cryptanalysis of Substitution Ciphers Using Scatter Search”, IWINAC 2005, LNCS 3562, 2005, pp.31- 40.

[17] Joseph Alexander Brown, Sheridan Houghten, Member, IEEE, and Beatrice Ombuki-Berman, ”Genetic Algorithm Cryptanalysis of a Substitution Permutation Network” 978-1-4244-2769-7/09 ©2009

[18] Bajeh.A.O, ”Optimization: Comparative study of tabu search and genetic algorithm”, International Journal for Computer Applications.

[19] Nalini.N and Ragevendra Rao. G,”Cryptanalysis of Simplified Data Encryption Standard (SDES) using Genetic Algorithm”, submitted to International Journal of Information and Computer Security (IJICS), Inter-science publishers.

[20] Harsh Bhasin, Ramesh Kumar, Neha Kathuria, ”Cryptography Using Cellular Automata”, International Journal of Computer Science and Information Technologies.

[21] Rashi Vohra, Brajesh Patel, ”An Efficient Choas-based Optimization Algorithm Approach for Cryptography.”, International Journal of Communication Network Security.

[24] Jitin Luthra,Saibal K Pal, ”A Hybrid Firefly Algorithm Using Genetic Operators for the Cryptanalysis of a Mono-Alphabetic Substitution Cipher”, World congress on information and communication Technologies, 2011.

[25] Saibal k Pal,C S Rai,Amrit Pal Singh, ”Comparative Study of Firefly Algorithm and Particle Swarm Optimization for Noisy Non-Linear Optimization Problem.”, I.J.Intelligent System and Applications, 2012.

[26] Guang Gong,” Fast Discrete Fourier Spectra Attacks on Stream Ciphers”, IEEE transaction on Information theory, august 2011.

[27] Matthews, R.A.J, “ The use of genetic algorithms in cryptanalysis”, Cryptologia, 1993.

[28] Clark, A, “Modern optimisation algorithms for cryptanalysis”, In Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, November 29 - December 2, (pp. 258-262).

[29] D.Sahoo, S.Champati Rai, S.Pradhan, ”Threshold Cryptography & Cryptanalysis of Four-Round DES Based on Genetic Algorithm” Genetic Algorithm Based on Secure Key Exchange for mobile hosts” .

[30] Mr. Vinod Saroha., Suman Mor., Anurag Dagar., “Enhancing Security of Caesar Cipher by Double Columnar Transposition Method,” International Journal of Advanced Research in Computer Science and Software Engineering., 2012.

[31] Ekta Agrawal, Dr. ParashuRam.Pal “Refined Polygram Substitution Cipher Method: A Enhanced Tool for Security,” International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 1, July 2012.