The Applicability of Genetic Algorithm in Cryptanalysis: A Survey

Asif Hameed Khan
Faculty of Engineering and Technology,
Jamia Hamdard University
New Delhi, India

Auqib Hamid Lone
Faculty of Engineering and Technology,
Jamia Hamdard University
New Delhi, India

Firdoos Ahmad Badroo
Faculty of Engineering and Technology,
Jamia Hamdard University
New Delhi, India

ABSTRACT
The Cryptanalysis has been the most fascinating area for science fraternity. The application of Genetic Algorithm (GAs) to the field of cryptanalysis is rather unique as no robust model for cryptanalysis using Genetic Algorithm exists. Genetic Algorithm (GAs) are the class of heuristic algorithm which are known for optimization and search problem. The paper presents the systematic review of Genetic Algorithm applied to the Cryptanalysis.

Keywords
Cryptanalysis, Genetic Algorithm (GAs), Cryptography.

1. INTRODUCTION
Cryptanalysis is the science of recovering the plaintext of a message without access to the key [1]. It is a method of transforming cipher text into a plaintext without knowing the key or algorithm [2]. Cryptanalysis of classical ciphers includes Permutation ciphers, Mono-alphabetic cipher, Poly-alphabetic cipher, Transposition cipher, Merkle-Hellman Knapsack cipher, Chor-rivest knapsack cipher and Vernam cipher. However the cryptanalysis of these traditional cipher through GAs is still an emerging issue.

GAs are based on the evolutionary ideas of Natural selection and genetics [3]. The algorithm are known for the useful solutions to optimization and search problems[48][4-10]. The algorithm have been successfully applied to Vertex-Cover problem [4][5], Maximum-Clique problem [6][7], Regression testing [8], N-puzzle problem [9], Traveling Salesman Problem [10].

The performance of GAs depends upon the proper setting of various algorithm components and parameters [14-17]. The operators used in GAs are crossover, mutation, selection amongst many others. The present work explores the related work done and applicability of GAs in a field of cryptanalysis.

The rest of the paper has been organized as follows: Section 2 presents a brief overview of GAs. Section 3 gives a Background. Section 4 discusses the Literature Review. The last section explains the future scope and the conclusion.

2. GENETIC ALGORITHM
GAs are the search heuristic that mimics the process of natural evolution [3, 12]. It is based on the Darwin's principle of Natural selection. According to this theory the chromosomes with the best fitness function should survive and create new offspring (survival of the fittest). GAs gives useful solution to optimization and search problem. It is a rapidly growing area of Artificial Intelligence.

The GAs starts with the population which is nothing but a chromosomes which can be decimal or binary or even hexadecimal. The GAs operator are applied to population in order to optimize the results [48]. The new population is formed from the old population with better fitness value. The population can be crafted using a following operators :

Population size : The population size is generally taken 5 to 100 [13-16]. Numerical experiments show that too large and too Small number of chromosomes in the population can lead to poor solutions [17].

Fittness Function: The fitness function decides whether the given solution is achieving the aims [18]. The proper crafting of the fitness function is a crux of the solutions.

Selection: The chromosomes are selected from the population for reproduction. The chromosomes with higher fitness value is more likey to be selected. [18]. There are many ways to do selection process some of them are Tournament selection, Roulette wheel selection, Stochastic universal selection, Truncation selection etcetera.

Crossover: The operator created a new offspring from the population by exchanging subsequences of two chromosomes to create two offspring.

The formula for the number of crossovers is as follows:

Number of crossover=(No. of cells in chromosomes*No. of chromosomes*crossover rate)/200 . There are many types of crossovers. Some of them are as follows:

1) **Single-point crossover**: One crossover point is selected, binary string from beginning of chromosome to the crossover point is copied from one parent, the rest is copied from the second parent [12]

2) **Two-point crossover** : Two crossover point are selected, binary string from beginning of chromosome to the first crossover point is copied from one parent, the part from the first to the second crossover point is copied from the second parent and the rest is copied from the first parent [12].

3) **Multi-point crossover :** Here more than two random points govern the formation of new chromosomes [12].

4) **Uniform crossover** : Bits are randomly copied from the first or from the second parent [12]

Mutation: Mutation operator flips the bit in chromosomes. The purpose of mutation is to maintain the diversity within the population. The mutation can be calculated by the following formula:

Number of Mutation=(No. of cells in a chromosomes*No. of chromosomes*mutation rate)/200.

3. BACKGROUND

The cryptanalysis has been the most fascinating areas of research. Many research exists on this topic. In the cryptanalysis using GAs, the focus were on classical ciphers like mono-alphabetic and poly-alphabetic, Permutation cipher, , Merkle-Hellman Knapsack cipher, Chor-Rivest Knapsack cipher, Transposition cipher, Vernam cipher and Substitution cipher.

The first paper was published by Spillman, Janssen, Nelson and Kepner in 1993 [19] .In this paper Cryptanalysis of simple substitution cipher is done using GAs. Soon afterwards, In 1993 another paper was published by R.A.J Matthews published a paper in which order-based GAs to attack a simple transposition cipher[20]. The work by Spillman applied GAs to a Merkle-Hellman Knapsack Cryptosystem [21]. In 1994, Clark includes GAs as one of the three optimization algorithms applied to cryptanalysis [22]. Feng-Tse Lin and Chen-Yan kao in 1995 proposed a cipher text attack on a vernam cipher [23]. The work by Clark, Dawson and Bergen [24] was fitness function used in [21], as well as a modified version of the same fitness function. The work by Clark, Dawson and Nieuwland [25] uses a parallel GAs for cryptanalysis. the paper published in 1997 by Clark and Dawson [26] is, overall, a slightly more detailed, longer version of [26]. In 1997,The paper was published by Kolodziejczyk [27], Which is an extension of [21], it focuses on the Merkle-Hellman knapsack system with enhancement in intitial parameters [21]. The work by Clark and Dawson in 1998 [28] compares three optimization algorithms applied to the cryptanalysis of a simple substitution cipher. Yaseen and Sahasrabuddhe in 1999 published an important work which proposed a GAs based on the Chor-Rivest public key cryptosystem [29]. In 2003, Grundlingh and Van Vuuren[30], attacks two classical ciphers with a GAs approach[30].

4. LITERATURE REVIEW

In order to have clear cut, unbiased, complete and broader prospective many sources have been explored. The Literature Review has been carried out according to the guidelines proposed by Kitchenham [11]. The extensive Literature review has been carried out in the following databases:

1. ACM digital Library
2. IEEE xplore
3. Science Direct
4. Wiley online Library
5. Springer

The reason behind exploring these databases is their rich library of journals with high impact factors. The review also takes into account conference proceedings.

The search term was 'Genetic algorithm and Cryptanalysis'. The search was filtered to include the papers and conferences of previous 10 years. This was done to limit the scope of research to the present trends instead of exploring unverified and undeveloped techniques.

The Results are summarized as follows:

Search the Keyword <Genetic Algorithm and Cryptanalysis> displayed 213 results in ACM digital library (19 oct 2015).These papers were ordered with respect to the year of publication from 2005 to 2015.On filtering by ACM publication the results were narrowed to 30.On further examination of these 30 papers with respect to their relevance of abstract 4 papers were finally filtered out and analyzed rigorously.

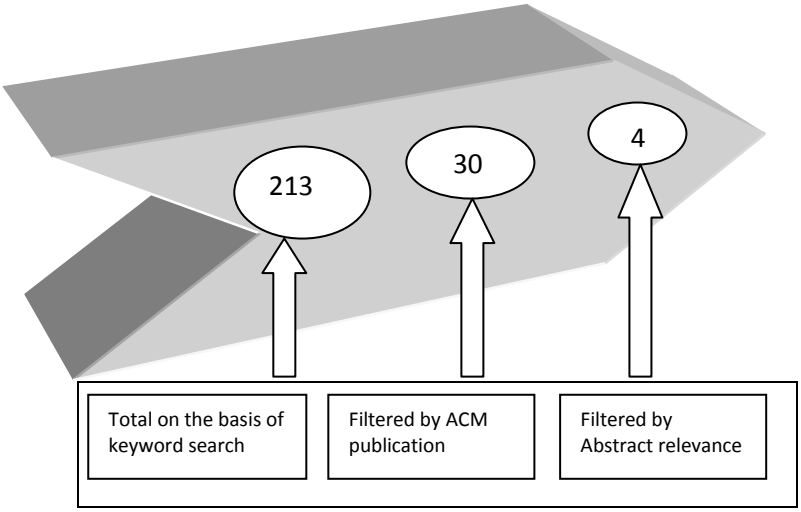

Fig 1. Selection of papers from ACM digital Library

Search the Keyword <Genetic Algorithm and Cryptanalysis> displayed 213 results in IEEE XPLORE(19 oct. 2015).These papers were ordered with respect to the year of publication from 2005 to 2015..On further examination of these papers with respect to their relevance of abstract 8 papers were finally filtered out and analyzed rigorously.

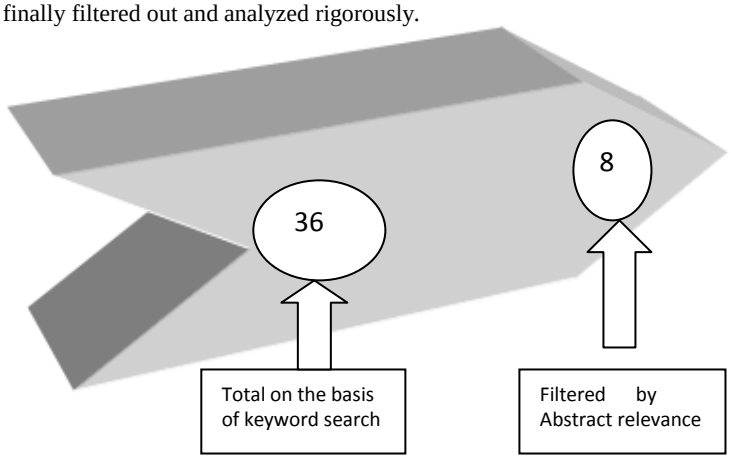

Fig.2 Selection of papers from IEEE XPLORE database

Search the Keyword <Genetic Algorithm and Cryptanalysis> displayed 102 results in SCIENCEDIRECT(19 oct. 2015).These papers were ordered with respect to the year of publication from 2005 to 2015..On further examination of these papers with respect to their relevance of abstract 2 papers were finally filtered out and analyzed rigorously.

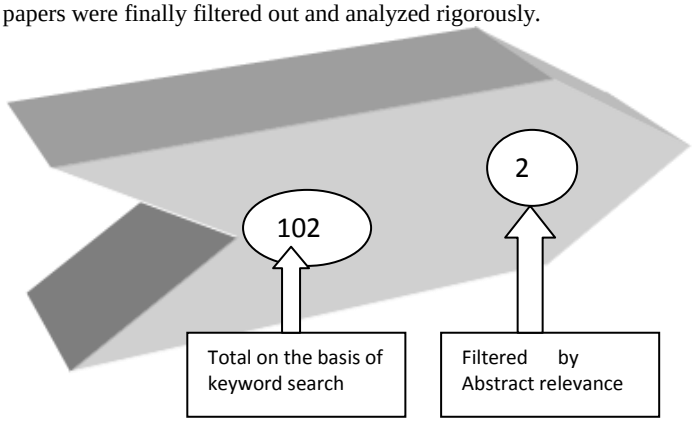

Fig. 3 Selection of papers from SCIENCE DIRECT database

Search the Keyword <Genetic Algorithm and Cryptanalysis> displayed 144 results in SPRINGER(19 oct. 2015).These papers were ordered with respect to the year of publication from 2005 to 2015..On further examination of these papers with respect to their relevance of abstract 3 papers were finally filtered out and analyzed rigorously.

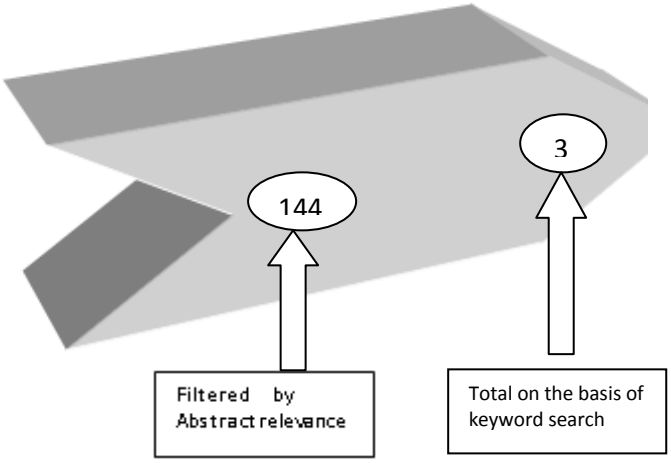

Fig.4 Selection of papers from SPRINGER database

Search the Keyword <Genetic Algorithm and Cryptanalysis> no results were found in WILEY Online library (19 oct. 2015).

4.1 Data extraction

As per the guidelines by Kitchenham [11], data were collected in the most efficient way. The quality of the papers was the foremost criteria in the paper selection .Several different techniques were identified from 18 papers for detailed analysis.

4.2 Research Questions

The systematic review intends to classify the work related to GAs in a field of cryptanalysis by proposing the set of questions .The aim of the work is to identify the gaps in existing methodologies and propose an answer to fill the gaps. The questions we wish to answer are as follows:

RQ1: What is the present state of the art?

RQ2: What are the advantages, if any of using Gas over traditional methods for cryptanalysis?

4.3 Answers to Research Questions

RQ1: During the review, it was found the GAs have not been explored in detail. In general SPRINGER had the highest number of pertaining to GAs.

RQ2: GAs have been used in diverse fields and applications . However, the focus of the review was to find the applicability of GAs in cryptanalysis. Research reveals GAs provide a robust way for cryptanalyzing traditional symmetric ciphers .But the applicability of GAs in cryptanalyzing Asymmetric ciphers is still in infancy and is not explored in detail.

4.4 Summary of Review

Reference No.	Author	Technique Used
[40]	A.Bhateja and S.kumar	GA with novel fitness function. A Roulette wheel method is used with two point crossover and cross mutation.
[41]	J.Luthra and S.K.Pal	Integration of operators of GAs with firefly algorithm for cryptanalysis.
[42]	J.Song et al	An efficient fitness measure is used to find some optimum keys with high fitness value
[43]	J.A.Brown et al	Division of keys into groups, each of which is analyzed to determine which groups are weaker.
[44]	S.S.Omran et al	The frequency analysis is used as an essential factor in objective function.
[45]	S.S.Omran et al	Attack on Mon alphabetic cipher via frequency analysis of cipher text.
[46]	N.Nalini and G.R.Rao	Simulated Annealing and Tabu search for Simple DES in comparison of GAs.
[47]	Y.H.Li et al	Tabu search and GAs for Intelligent Key-Search attack.
[31]	E.Y.Ma and C.Obimbo	GAs and harmony Search.
[32]	N.Nalini and G.R.Rao	Optimization Heuristic based on GAs, Tabu Search and Simulated Annealing.
[33]	H.C.J.Castro and I.P.Vinuela	GAs applied to Block cipher XTEA

[34]	J.Song et al	GAs applied to Two-round DES.
[35]	P.K.Bergmann	Simplified GAs.
[36]	D.Oranchak	Dictionary based attack using GAs.
[37]	J.Song et al	Attack against Transposition Cipher using improved Simulated Annealing GAs.
[38]	T.Mekhaznia	Particle Swarm Particle, Differential Evolution and GAs.
[39]	R.Vimalathithan and L.M.Valarmathi	Genetic Swarm Optimization is applied to cryptanalysis

5. FUTURE SCOPE AND CONCLUSION

It is observed that GAs have been applied to the cryptanalysis ,However most of the attack were on the classical ciphers like Mono alphabetic, Poly alphabetic, Substitution ciphers and many more. Also techniques like Simulated Annealing, Tabu Search, Particle Swarm Particle optimization have been applied with GAs for cryptanalysis.

The applicability of GAs exist in a cryptanalysis but very few research exists on this topic. No robust model is still accomplished.

GAs have many application [4-10][2].The work provide a systematic review in a order to understand the scope of GAs in a field of Cryptanalysis. Application of GAs, for cryptanalysis of asymmetric cipher is a field that needs to be explored.

6. ACKNOWLEDEMENT

We thank our mentor (Harsh Bhasin) and friends for motivation and support in presenting this paper.

7. REFERENCES

[1] Schneier,B.1996. Applied Cryptography,Second Edition:Protocols,Algorithms and Source Code in C.

[2] Delman,B.2004. Genetic Algorithm in Cryptography.Doctorol thesis,Rochester Institute of Technology.

[3] Holland,J.H.1992. Adaptation in Natural and Artificial Systems.

[4] M.Milanovic,"Solving the generalised vertex cover problem by Genetic Algorithm", Computing and Informatics,2010.

[5] H.Bhasin,M.Amini,"The applicability of Genetic Algorithm to Vertex cover", International Journal of Computer Application,2015.

[6] Bazgan,C.,Luchian,H.1995. A genetic Algorithm for maximal Clique Problem. In proceeding of the International Conference in Ales,France.

[7] H.Bhasin et al,"Hybrid Genetic algorithm for Maximum Clique Problem", International Journal of Application of Innovation in Engineering & Management,2013.

[8] H.Bhasin,Manoj,"Regression testing using Coupling and Genetic Algorithms", International Journal of Computer Science and Information Technologies,2012.

[9] H.Bhasin,N.Singla,"Genetic based algorithm for N-Puzzle problem", International Journal of Computer Application,2012.

[10] Y.Liao et al,"Evolutionary algorithm to Traveling Salesman Problems", Computer & Mathematics with Applications,2012.

[11] Kitchenham,B. et al, "systematic literature review in software engineering", Information and Software technology, Elsevier,2008.

[12] Bhasin,H.2015.Algorithms: Design and Analaysis.

[13] Alander.1992. On optimal population size of genetic algorithm. In Proceedings of the IEEE computer systems and software engineering.

[14] Diaz-Gomaz, Hougen.2007. Initial population for genetic algorithms: A metrics approach. In proceedings of 2007 International conference of Genetic and Evolutionary Methods.

[15] Piszcz, Soule.2006. Genetic Programming: optimal population sizes for varying complexity problems. In Proceedings of the Genetic and Evolutionary Computation Conference.

[16] Koumousis and Katsaras, " Asaw tooth Genetic Algorithm combining the effects of variable population size and re-initialization to enhance performance", IEEE Transaction on evolutionary computation,2006.

[17] Goldberg,D.E et al.2000. Bayesuan Optimization Algorithm, population sizing and time to convergence, University of Illinois,USA.

[18] Melanie,M.1996. An introduction to a Genetic Algorithm:MIT press paperback edition.

[19] R.Spillman etal,"Use of Genetic algorithm in the cryptanalysis of simple substitution ciphers", Cryptologia,1993.

[20] A.J.R.Matthews,"The Use of Genetic Algorithim in Crytanalysis," Cryptologia,1993.

[21] R.Spillman,"Cryptanalysis of Knapsack Ciphers using Genetic Algorithim", Cryptologia,1993.

[22] Clark,A.1994. Modern optimisation Algorithims for Cryptanalysis. In proceedings of the second Australian and New Zealand Conference on Intelligent Information Systems.

[23] Lin,F.1995. A Genetic Algorithim for Ciphertext-only attack in cryptanalysis. In proceeding of the IEEE International Conference on systems,Man and Cybernetics.

[24] A.Clark etal,"Combinational Optimization and Knapsack Cipher",Cryptologia,1996.

[25] Clark,A.etal.1996. Cryptanalysis of Polyalphabetic Substitution Ciphers using a Parallel Genetic Algorithim.In proceedings of IEEE International Symposium on Information and its Application.

[26] A.Clark,E.Dawson, "Parallel Genetic Algorithm for cryptanalysis of the polyalphabetic substitution cipher", Cryptologia,1997.

[27] Kolodziejczyk.1997. The Application of Genetic algorithm in cryptnalaysis opf Knapsack cipher. In proceeeding of the Fourth International Conference on Pattern Recognition and Information Processing.

[28] A.Clark,E.Dawson, "Optimization Heuristics for the automated cryptanalysis of classical ciphers", Journal of Combinatorial Mathematics and Combinatorial Computing,1998.

[29] Yaseen,I.F.T., Sahasrabuddhe,H.V.1998. A Genetic Algorithm for the cryptanalysis of Chor-rivest knapsack public key cryptosystem(PKC). In proceeding of the third International Conference on Computational Intelligence and Multimedia Applications

[30] Grundlingh, R.W, Vuuren, V.H.J, "Using Genetic Algorithm to break a simple cryptographic cipher", Retrieved from http://dip.sun.ac.za/nvuuren/abstracts/abstr-genetic,htm ,2003

[31] E.Ma, C.Obimbo,"An Evolutionary Computation Attack on One Round TEA",Elsevier,2011.

[32] N.Nalini, G.R.Raghavendra,"Attacks of simple block ciphers via efficient heuristics",Elsevier,2007.

[33] H.C.J.Castro, I.P.Vinuela,"New result on the genetic cryptanalysis of TEA and reduced-round versions od XTEA",New Generation Computing,2005.

[34] Song, J.et al.2007.Cryptanalysis of the two-round DES using Genetic algorithm. In proceedings of the second international conference on Advances in computation and intelligence.

[35] Bergmann, P.K.2008. Cryptanalysis using Genetic algorithm. In proceeding of tenth annual Conference on Evolutionary Computation.

[36] Oranchak,D.2008. Evolutionary algorithm for decryption of Monoalphabetic homophonic substitution ciphers encoded as constraint satisfaction problems. In proceeding of the tenth annual Conference on Genetic and Evolutionary Computation.

[37] Song,J.et al.2008. Cryptanalysis of Transposition cipher using Simulated Annealing Genetic Algorithm. In proceeding of third International Symposium on Advances in Computation and Intelligence.

[38] Mekhaznia,T.2013. Nature inspired heuristics for attack of simplified DES algorithm. In proceeding of the sixth International Conference on Security of Information and Networks.

[39] R.Vimalathithan,L.M.Valarmathi, "Cryptanalysis of simplified-DES using computational intelligence",WSEAS Transactions on Computers,2011.

[40] Bhateja,A.,Kumar,S.2014. Genetic Algorithm with elitism for cryptanalysis of Vignere cipher. In proceeding of International Conference on issues and challenges in Intelligent Computing Techniques(ICICT).

[41] Luthra,J.,Pal,S.K.2011. A Hybrid Firefly algorithm using Genetic operators for the cryptanalysis of monoalphabetic substitution cipher. In proceeding of World Congress on information and Communication Technologies(WICT).

[42] Song,J.et al.2007. Cryptanalysis of Four-round DES based on Genetic Algorithm. In proceeding of International Conference on Wireless Communications, Networking and Mobile Computing.

[43] Brown,J.A.et al.2009. Genetic algorithm cryptanalysis of a substitution permutation network. In proceeding of IEEE Symposium on Computational Intelligence in Cyber Security(CICS).

[44] Omran,S,S.et al.2011. A Cryptanalytic attack on Vignere cipher using Genetic Algorithm. In proceeding of conference on Open Systems(ICOS).

[45] Omran,S.S.et al.2010. Using Genetic Algorithm to break a mono-alphabetic substitution cipher. In proceeding of IEEE Conference on Open Systems(ICOS).

[46] Nalini,N.,Rao,G.R.2006. Cryptanalysis of simplified Data Encryption Standard via optimization Heuristics. In proceeding of Third International Conference on Intelligent Sensing and information Processing(ICISIP).

[47] Li,Y.H.et al.2005. Heuristics cryptanalysis pof classical and modern ciphers. In proceeding of jointly held with IEEE Seventh Malaysia International Conference on Communication and IEEE Thirteenth International Conference on Networks.

[48] Goldberg,D.E.1989. Genetic Algorithm in search, optimization and machine learning.