IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Design And Implementation Of Metaheuristic Inspired Cryptanalysis

[1]Sai Naveen Thota, [2]Hrushikesh Reddy Neravetla, [3]Akash Reddy Bhoomidi, [4]Niketh Yadav Arke,

[5]Dr. B. Srinivas Rao

[1]Student, [2]Student, [3]Student, [4]Student, [5]Professor

[1]Department of Computer Science and Engineering

[1]Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

Abstract: Cryptanalysis is the process of key recovery, or plaintext recovery without the knowledge of the key in Cryptology. It helps us to improve the Cryptographic system by finding any weak point and thus work on the algorithm to create a more secure secret code. In the process of Automated Cryptanalysis we decrypt the cipher text with many possible keys to obtain candidate plaintexts. The basic type of algorithm suitable for this process is a Brute Force attack. This attack is only feasible when key-space is searchable on computational resources available to an attacker. On the other hand, only the most complex algorithms achieve really high accuracy of the plaintext recognition. Nature-inspired algorithms are a set of optimized problem-solving Metaheuristic derived from natural processes that can be used for Automated Cryptanalysis. Genetic algorithm use a stochastic approach to find the best solution in the large search space of the problem. In our research work, we use Genetic algorithm to Cryptanalyze Vigenere and Columnar Transposition Ciphers. A comparision study between those two ciphers is made by testing the algorithm with various file sizes of encrypted plaintexts with two encryption algorithms and analyze which is more efficacy to Genetic algorithm.

Index Terms **– Metaheuristics, Cryptanalysis, Vigenere Cipher, Columnar Transposition Cipher, Genetic Algorithm.**

I. INTRODUCTION:

Cryptanalysis is the procedure of optimal key restoration or partial plaintext restoration without the information of the key. It allows us to highly recognize the cryptosystems and additionally allows us to enhance the system by locating any susceptible factor and for that reason work on the algorithm to create a greater secure secret code. In the procedure of automatic cryptanalysis, we decrypt the cipher textual content with many feasible keys to achieve candidate plaintexts. The primary type of algorithm appropriate for automatic cryptanalysis is a brute-force attack. Plaintext recognition is the most critical part of the algorithm from the overall performance point of view. This attack is best viable while key space is searchable on computational resources to be had by an attacker.

On the opposite hand, only the maximum complicated algorithms attain high accuracy of the plaintext recognition. Optimization algorithms are the relatively efficient algorithms which awareness on finding answers to notably complicated optimization problems. Nature-inspired algorithms are a set of metaheuristic problem-solving methodologies and strategies derived from natural processes. Some of the famous examples of nature-inspired optimization algorithms include: Genetic algorithm, Particle swarm optimization, Ant colony optimization and so on. By the usage of these nature inspired algorithms, the problem can be solved with much less computational efforts and time complexity. These algorithms use a stochastic technique to discover the pleasant answer in the huge search space of the problem.

1.1 Vigenere Cipher:

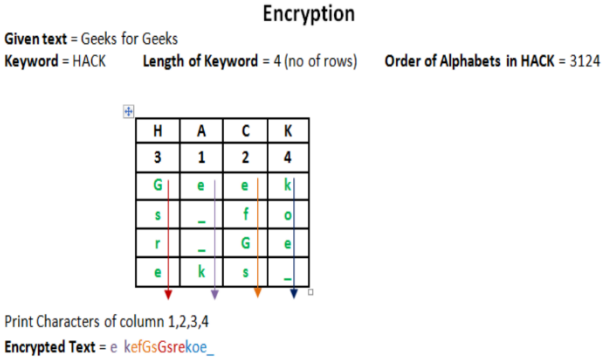
Vigenere Cipher is a way of encrypting alphabetic textual content. It makes use of an easy form of polyalphabetic substitution. A polyalphabetic cipher is a cipher based on substitution, the use of more than one substitution alphabets. The encryption of the original textual content is done using the Vigenere square or Vigenere table.

Encryption: $E(i) = (P(i) + K(i)) \% 26$

Decryption: $P(i) = (E(i) – K(i) + 26) \% 26$

1.2 Columnar Transposition Cipher:

The Columnar Transposition Cipher is a form of transposition cipher. It involves writing the plaintext out in rows, and then reading the ciphertext off in columns one by one.



1.3 Genetic Algorithm:

Genetic Algorithms(GAs) are adaptive metaheuristic search algorithms that belong to the larger part of evolutionary algorithms. Genetic algorithms are primarily based totally on the ideas of natural selection and genetics. These are smart exploitation of random search provided with historic information to direct the search into the area of higher overall performance in solution space. They are typically used to generate high-quality answers for optimization problems and search problems. They simulate "survival of the fittest" amongst individual of consecutive generation for fixing a problem.

Fitness Measure: N-gram Score (using Quadgrams) Quadrams determine how similar text is to English. For Example: For text GRIET, quadrams are GRIE and RIET.

$p(GRIET)=p(GRIE)*p(RIET)$

$p(GRIE)=count(GRIE)/N$

In the above equation, count() is the number of times quadram has occurred and N is the total number of quadrams in the training sample.

Applying log we get fitness score.

Fitness Score = $\log(p(GRIET)) = \log(p(GRIE))+\log(p(RIET))$

Higher the log probability, higher the fitness score.

II. RELATED WORK:

[1] Title: Metaheuristic Techniques for Automated Cryptanalysis of Classical Transposition Cipher. Author: Ashish Jain, Prakash C. Sharma, Santosh K. Vishwakarma & Nirmal K. Gupta.
Overviw: This paper compares the overall performance of those new and distinct metaheuristic strategies. Different metaheuristic optimization strategies had been supplied with inside the literature for automatic cryptanalysis of classical transposition cipher. It is noteworthy that most of the supplied metaheuristics the overall performance of genetic algorithm method is satisfactory with appreciate to all of the measures. Publisher: IEEE paper published in 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW) on 24 August 2017.

[2] Title: Metaheuristic Techniques in Attack and Defense Strategies for Cybersecurity.
Authors: Agustín Salas-Fernández, Broderick Crawford & Ricardo Soto.
Overview: This study objectives to inspect the metaheuristics carried out to optimize artificial intelligence strategies with inside the detection of threats or optimization of attacks via way of means of using unique measures: detection or attack technique, reason and the type of metahauristics involved. The evaluation changed into performed in applicable literature databases including Web of Science, SCOPUS, SciELO, ACM and Google Scholar. Publisher: Chapter published on 01 June 2021 as part of Studies in Computational Intelligence book series (SCI Volume 972): Artificial Intelligence for Cyber Security.

[3] Title: Design and implementation of algorithm for DES cryptanalysis.
Authors: Harshali D. Zodpe, Prakash W. Wani and Rakesh R. Mehta.
Overview: Cryptanalysis of block ciphers includes massive computations that are independent of each other and may be instantiated concurrently in order that the answer area is explored at a quicker rate. This paper provides the layout for Hardware implementation of Data Encryption Standard (DES) cryptanalysis on FPGA the use of exhaustive key search. Two architectures viz. Iterative and Loop unrolled DES structure are implemented. The goal of this work is to make cryptanalysis quicker and better. Publisher: IEEE paper published in 2012 12th International Conference on Hybrid Intelligent Systems (HIS) on 28 January 2013.

[4] Title: Design of Metaheuristic Based on Machine Learning: A Unified Approach.
Authors: Amir Nakib, Mohamed Hilia, Frederic Heliodore and El-Ghazali Talbi.
Overview: In this work, a framework primarily based totally on maximum likelihood estimation and mutual information is proposed to layout a metaheuristic. A multilevel decomposition of metaheuristics is proposed that permit to have a unified vision on this optimization approach. Then, a new layer based on machine learning is introduced to take benefit from the evolution of the algorithm to evolve it to the considered problem to relieve customer's efforts interested by designing and implementing metaheuristics. Publisher: IEEE paper published in 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW) on 24 August 2017.

III. RESEARCH METHODOLOGY

Using the Meta-heuristics to attack the substitution and transposition ciphers requires a way to decide the viability of a key, known as fitness function. The fitness functions are applied to consider the validity of a certain key relying on the form of the frequency evaluation. The goal of frequency evaluation is to evaluate the frequencies with inside the decrypted textual content with the frequencies determined in English literature. In different words, the frequency evaluation authorizes us to assess exactly genuine suits among a certain textual content and the language wherein the unique textual content was written.

Despite the fact that Meta-heuristics may be not commonly assured to discover the most appropriate global solution, they are able to often discover a sufficiently correct solution in a decent quantity of time. So, they may be an alternative to exhaustive search, which could take exponential time. Meta-heuristics regularly comprise a few forms of randomness to escape from local minima.

We have chosen Genetic Algorithm to perform Cryptanalysis of Vigenere and Columnar Transposition Ciphers as it has excellent parallel capabilities and provides answers that improve over time. For each cipher we provide various sizes of text files that contains plain texts and generate cipher texts.Now these cipher text files are provided as input to the Genetic algorithm to perform Cryptanalysis and the output is the plaintext that is generated by genetic algorithm.At the end the original plain text and the generated plain text are compared to find to how much extent Genetic algorithm is able to crack the Cipher text without the knowledge of key. Similarity percentage is calculated based on the number of characters matched.

IV. RESULTS AND DISCUSSION

Similarity percentage is calculated using formula, (Number of matched characters / Total number of characters in plain text)*100. Similarity for text file with size 1 KB(Kilobyte) that contained Cipher text encrypted by Vigenere Cipher is 100%. It is same with Cipher text encrypted by Columnar Transposition Cipher too. Similarity percentage for text file with size of 10 KB that contained Cipher text encrypted by Columnar Transposition Cipher is 79.33%. Similarity percentage for text file with size of 10 KB that contained Cipher text encrypted by Columnar Transposition Cipher is 66.97%. A table is drawn between the text file sizes and the similarity of text that is cryptanalyzed accurately. Along with Similarity percentage, time taken to perform the process is considered and analyzed for the two Ciphers. Time taken is considered in seconds. From the below table, graphs are drawn for easy comparision of Vigenere and Columnar Transposition Ciphers with with respect to similarity and time taken for Cryptanalyzing.

4.1 Figures and Tables:

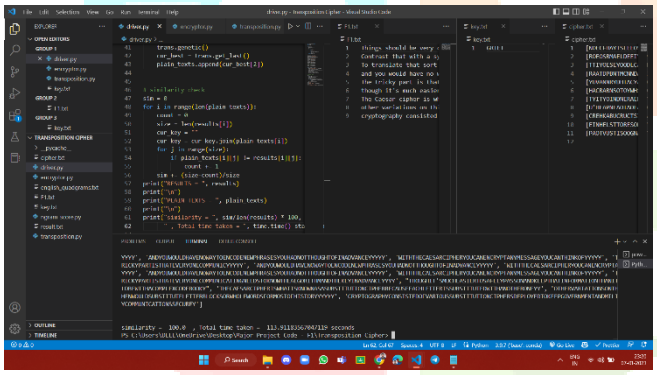


Fig.1: Output for 1 KB cipher text encrypted by Columnar Transposition Cipher

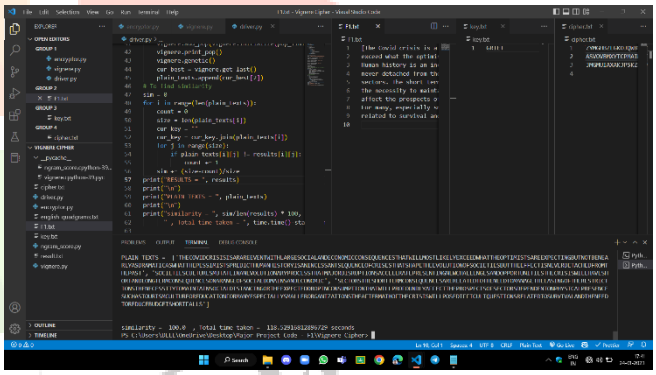


Fig.2: Output for 1 KB cipher text encrypted by Vigenere Cipher

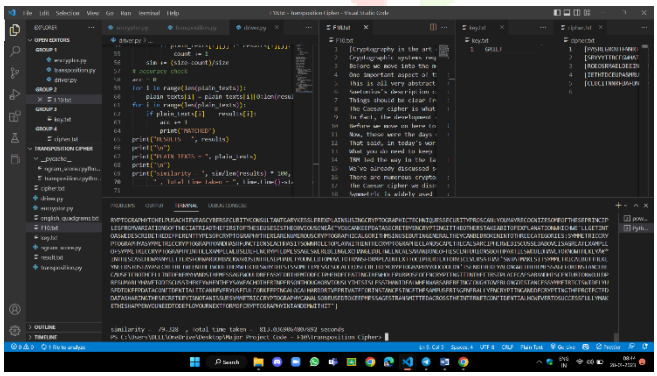


Fig.3: Output for 10 KB cipher text encrypted by Columnar Transposition Cipher

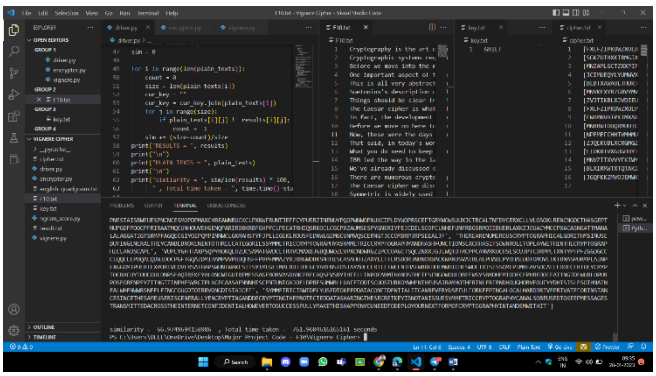


Fig.4: Output for 10 KB cipher text encrypted by Vigenere Cipher

Table-1

	Vigenere Cipher		Transposition Cipher	
File Size (in KB)	**Similarity (in Percentage)**	**Time Taken (in Sec.)**	**Similarity (in Percentage)**	**Time Taken (in Sec.)**
1	100.00	118.53	100.00	113.91
2	98.00	111.04	100.00	163.96
3	94.01	172.16	100.00	198.50
4	91.01	255.92	100.00	200.28
5	83.72	398.49	95.08	243.34
6	80.39	396.91	91.57	313.02
7	78.98	474.11	90.00	319.60
8	72.44	544.78	88.34	658.51
9	67.81	722.99	87.91	692.11
10	66.97	761.96	79.33	813.04

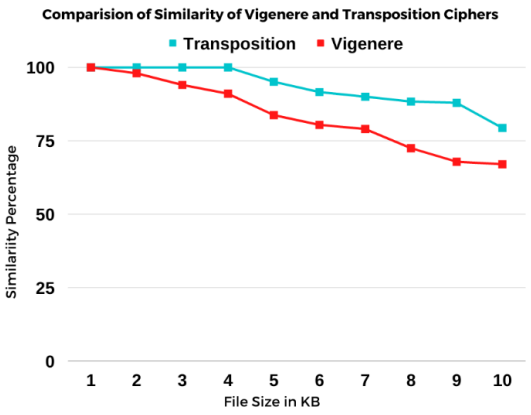


Fig.5: Similarity percentage graph of Vigenere and Transposition Ciphers with respect to different file sizes.

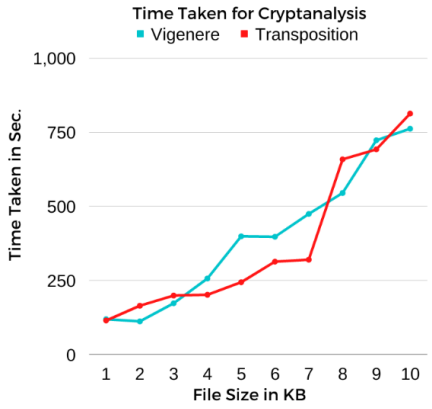


Fig.6: Time graph of Vigenere and Transposition Ciphers With respect to different file sizes.

V. Acknowledgment

From the above graphs, it can be observed that Columnar Transposition Cipher is weaker that Vigenere Cipher. Similarity percentage of cracking Vigenere cipher text is lesser that Columnar Transposition Cipher text. Hence, it can be concluded that Vigenere (Polyalphabetic substitution) Cipher is more stronger that Columnar Transposition Cipher.

VI. References

[1] Harshali D. Zodpe; Prakash W. Wani; Rakesh R. Mehta, (2013), "Design and implementation of algorithm for DES cryptanalysis Publisher: IEEE", IEEE paper published in 2012 12th International Conference on Hybrid Intelligent Systems (HIS) on 28 January 2013.

[2] Amir Nakib; Mohamed Hilia; Frederic Heliodore; El-Ghazali Talbi, (2017), "Design of Metaheuristic Based on Machine Learning: A Unified Approach", IEEE paper published in 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW) on 24 August 2017.

[3] Ashish Jain, Prakash C. Sharma, Santosh K. Vishwakarma & Nirmal K. Gupta, (2021), "Metaheuristic Techniques for Automated Cryptanalysis of Classical Transposition Cipher", 1 st International Online Conference on Smart Systems: Innovations in Computing.

[4] Agustín Salas-Fernández, Broderick Crawford & Ricardo Soto, (2021), "Metaheuristic Techniques in Attack and Defense Strategies for Cybersecurity", Chapter published on 01 June 2021 as part of Studies in Computational Intelligence book series (SCI Volume 972): Artificial Intelligence for Cyber Security.