ARTICLE
JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol.30; No.1.| 2022

Key Generation for Vigenere Ciphering Based on Genetic Algorithm

Asraa Abdullah Hussein [1*], Noor Kadhim Ayoob [2]

[1]College of Science for Women, University of Babylon, esraa_zd@yahoo.com, Babylon, Iraq.
[2]College of Science for Women, University of Babylon, noor.kadhum@gmail.com, Babylon, Iraq
***Corresponding author email:** esraa_zd@yahoo.com

Received:	1/9/2021	Accepted:	8/1/2022	Published:	31/3/2022

ABSTRACT

Cryptography is a science securing of information. Encryption requires impregnable keys to encrypt or decrypt data these keys should be unpredictable and not easily to break. In this research we use genetic algorithm to generate keys for vigenere cipher. The best key is used to perform encryption. The keys created by genetic algorithm are tested for randomness by using the entropy test. The entropy calculation shows that randomness of key generated based on genetic processing is better than chosen key in the classical vigenere cipher.

Keyword: vigenere, genetic algorithm, entropy, key generation.

1. INTRODUCTION

One of the uses of computer technologies in human lives is to protect private information from random access by unauthorized people and cryptographic techniques is one of the methods used to do this task [1,2].

Cryptography depends on converting data into a non-understandable code and then returns it to original using keys that exist at specific people who can access and identify data[3]. In general, cryptographic methods are classified into two main types: the first type depends on one-key that exists at the transmitter and receiver, this type of encryption is called the symmetric method. The second type is called the public key ciphering that use two keys, a declared key for all and a special key for the recipient of the message [4].

Vigenere cipher was created in 1553 by Italian cryptologist Giovanni Battista Belazzo. It is one of the classic symmetric key cryptographic algorithms in which encryption and decryption process use the same key [5].

Genetic algorithm is search and optimization algorithm appeared in 1975 by John Holland and become one of the most important methods for supporting and improving the performance of

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com
مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

other methods, for example, GA can increase the performance of any classifier by choosing the important features for these classifiers [6,7]. In the field of cryptography, the genetic algorithm can make the encryption method stronger by finding the best keys for those methods. The genetic algorithm can also be a double-edged weapon, as it can be used by attackers to break codes[8].

This paper discusses the possibility of using a genetic algorithm as a way to find the optimal key for vigenere coding. In the rest of this work, the researches related to the research topic are presented in the section 2. Sections 3 and 4 discuss the details of the method adopted in this work and present the results. A set of conclusions are presented in section 5.

2. RELATED WORK

Researchers work has been going on for decades and to day providing scientific research aimed at serving human life and facilitating the work of institutions.This section deals with a part of researchers work in the field of generate random key for cryptography as shown in the table1 :

Table 1: Related researches

NO.	Methods	year	Measure used in evolution
[9]	G.A. is used to obtain a best secret key in polyalphabetic substitution cipher.	2008	measure the differences between frequencies for each letter in the plain text file and in the encrypted files
[10]	Vigenere cipher with guessing the key by applying Genetic Algorithm.	2008	n-gram statistics
[11]	Vigenère cipher with varying key	2012	Index of Coincidence
[12]	Strong Key Machanism Generated by LFSR based Vigenère Cipher	2012	Frequency Analysis
[13]	Chaos theory to measure stream RC4 and Vigenere Cipher algorithms	2015	Entropy
[14]	Vigenere Cipher: Trends, Review and Possible Modifications	2016	Index of Coincidence, Entropy
[15]	Randomize a Vegenere Cipher Key Based on the Key Procedure of RC4	2019	Estimation time

ISSN: 2312-8135 | Print ISSN: 1992-0652 | info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

3. PROPOSED METHOD

Vigenere is one of the traditional methods for encrypting data. According to this method, data is encrypted using the following equation:

$$\text{Cipher} = (\text{text} + \text{Key}) \bmod 26 \qquad (1)$$

Due to the importance and sensitivity of the selection of the key in the process of coding and decoding, the genetic algorithm is responsible for generating key randomly without any user intervention, thus the key created by genetic algorithm makes vigenere stronger and safer than usual. The genetic algorithm works on discovering the perfect key through the following phases:

1) creating first generation:

The genetic algorithm has the advantage of proposing a set of solutions organized into generations. The first generation is usually generated randomly. The chromosome is the main part of the solution suggested by the genetic algorithm and it is usually a vector. In this paper, any generation consists of 20 chromosomes, each chromosome produced by the genetic algorithm represents a proposed key for vigenere. The chromosome consists of 26 genes which means that there is a gene for each character of alpha. The value of genes is binary so we can imagine the chromosome as a vector of 0 and 1 as shown in figure 1:

Fig. 1: chromosome interpretation

According to the previous chromosome, the characters no. :1, 2, 4, 5, 6, 13, 14, 15, 19, 20, 22, 25 and 26 are selected to form a key because the value of the genes that represent these characters is equal to 1. The remaining characters, characters no.: 3,7,8,9,10,11,12,16,17,18,21,23, and 24, are ignored so the key suggest by the above chromosome is: (abdefmnostvyz).

2) rating the keys

Each key suggested by the genetic algorithm must be rated to find the best key appeared in the first generation. The key (chromosome) that offers a cipher text with a highest randomness

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

is the best key. Entropy test was used to measure the randomness of cipher text and evaluate keys which can be calculated according to the following formula:

$$\mathbf{H(entropy)= -\sum_{i=1}^{n} pi \log_2 pi} \qquad \mathbf{(2)}$$

The higher value of the entropy which mean cipher text contain higher randomness, in other words, the chromosome with highest entropy value will be the candidate key that vigenere using in encryption and decryption.

3) start evolution

At this stage, new generations are created through genetic manipulation: selection, mating, mutation. The designer of the algorithm must choose the methods for implementing the genetic processes, in this paper, we choose:

a) binary group for selection: the genetic algorithm has to choose two solutions for mating to produce new solutions. The first mate is chosen by nominating two solutions from the generation (randomly). By competing depending on the rate value of each candidate, the one that has the best rate is chosen, i.e. the candidate with the highest entropy. The second mate is selected using the same procedure.

b) one-point mating, figure 2: where the parents are divided into two pieces at a random point and then the pieces located after that point are exchanged, so that each child gets a piece from each parent. This process is applied with occurrence chance = 0.8.

Fig. 2: one-point mating

c) Swapping mutation, figure 3: which is the process of exchanging values between two randomly selected genes. The genetic algorithm applies with occurrence chance = 0.2.

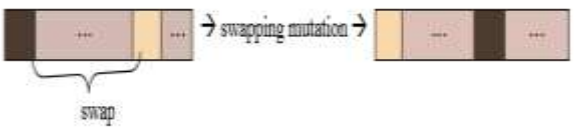

Fig. 3: swapping mutation

There are 20 members in the new generation and all of them are rated by calculating entropy**.** Each generation produced in this phase is an evolution of solutions in the previous

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

generation. By the end of this phase, we get 40 generations (a first generation in addition to 39 generations that are developed from genetic processing).

4) getting the final key

The best key is announced, the one that has the highest entropy. This key is used by vigenere to encode and decode data.

4. Results and discussion

Matlab R2018 was the language used to program the proposed system. Figure 4 displays the results of both vigenere alone and vigenere hybridized with genetic algorithm. Figure 5 shows the effect on the entropy when using the genetic algorithm for key selection compared to the entropy obtained with normal key selection.

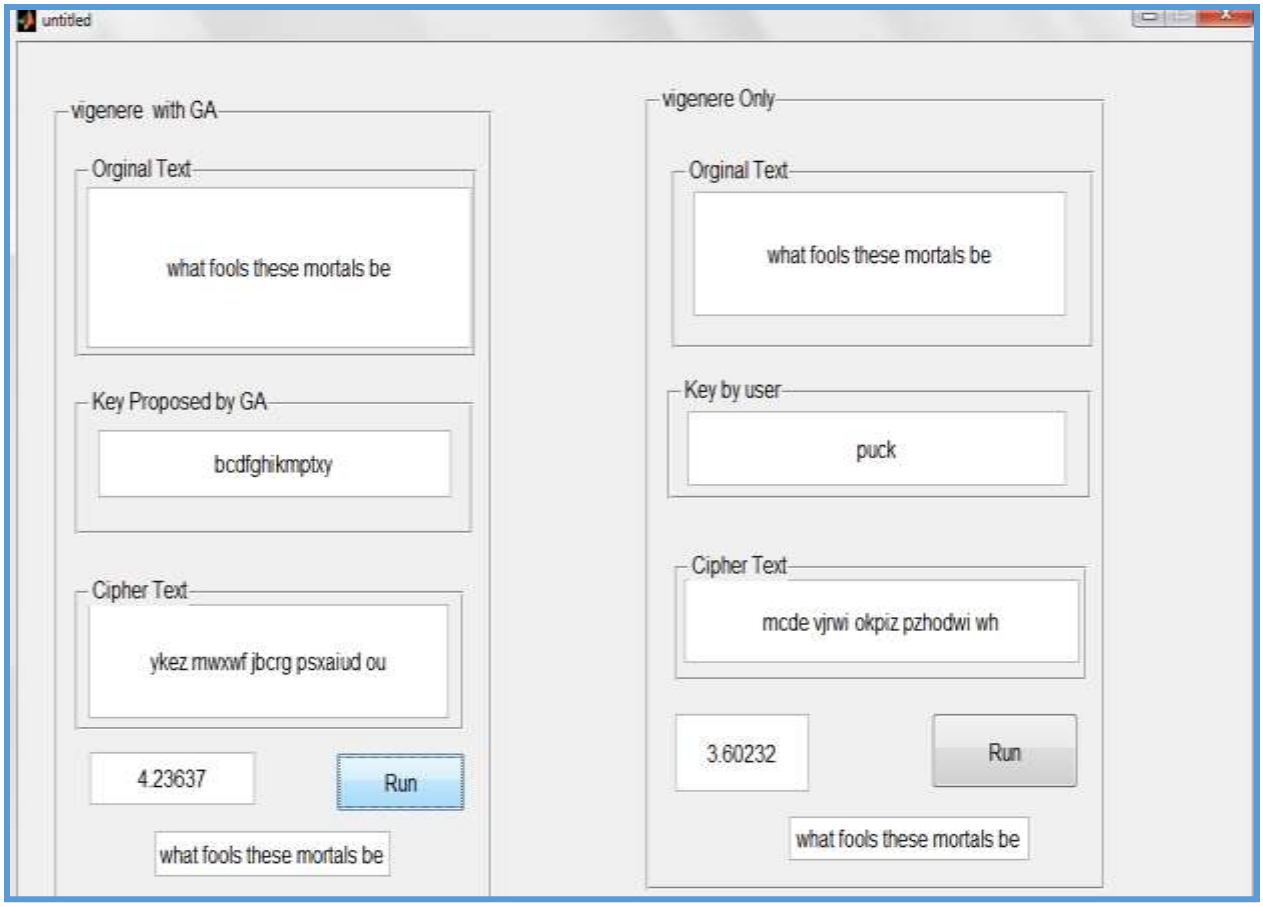

Fig. 4: sample of execution.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

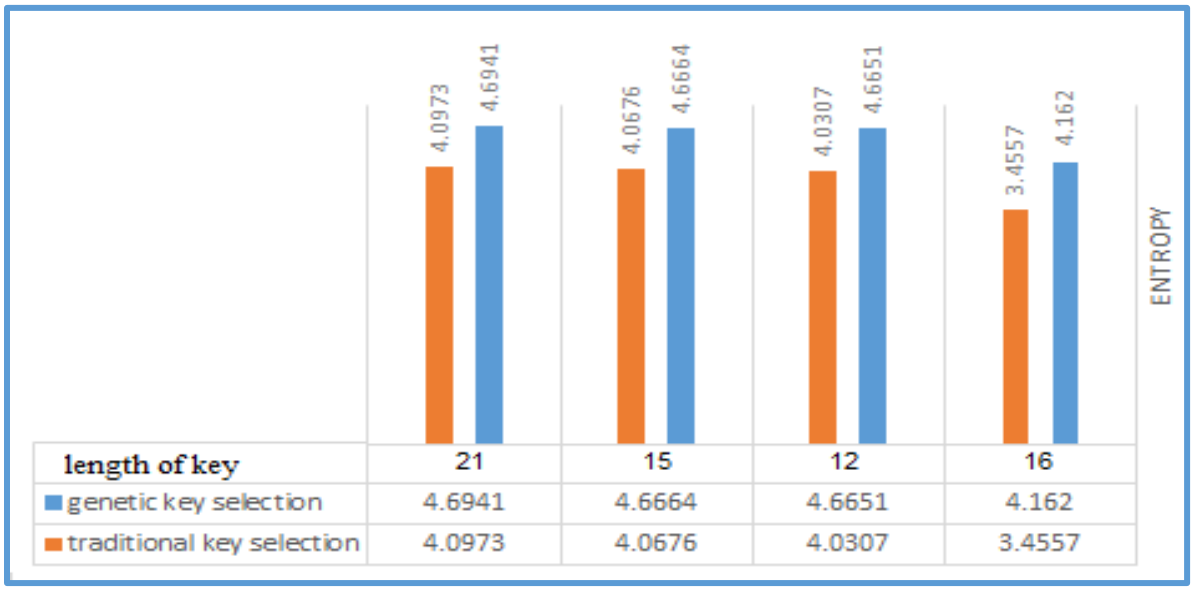

Fig. 5: the effect of GA on entropy

For comparison purpose, we select the closest researches to our work which are: [13] where the researchers presented the vigenere with Chaos method based on entropy as a scale to evaluation and [14] presented the vigener method with its advanced versions and measured the efficiency of each method using entropy. Figure 6 displays a comparison between the results of these studies and our proposed idea.

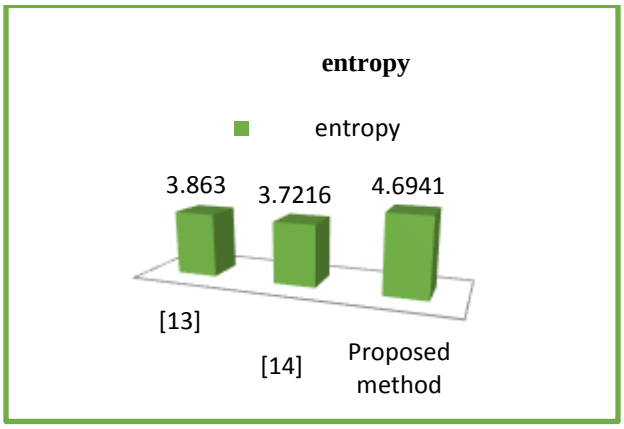

Fig. 6: the effect of GA on entropy

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Reference [9] used vigenere method with genetic algorithm by adopting a fitness function that includes calculating the sum of the difference of the occurrences of the plain text characters with the cipher text characters. In order to compare the results of the proposed method with the results of [9], their validity function was applied in the proposed system and the figure 7 shows that.

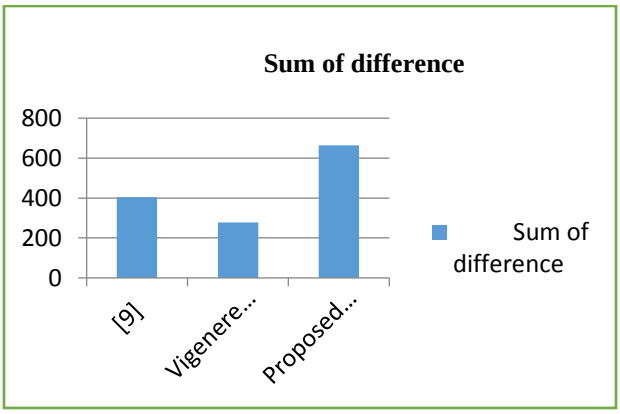

Fig. 7: compare proposed method with reference [9]

From the recorded results we noted that Shannon entropy was increased with the adoption of the genetic algorithm for generating a random key gave quite a good result compared to the vigenere with key selected in traditional way which means that the genetic algorithm strengthened vigenere. Choosing key classically causes the appearance of many repeated characters on cipher compared with cipher text produced by proposed method which is contain more randomness, this explains why Vigenere with a regular key are easy prey to break.

5. CONCLUSIONS

This research included the use of a genetic algorithm to increase the confidentiality of the vigenere cipher method by generating a random key and adopting it in the process of encoding and decoding. The results of the proposed system showed a significant increase in the value of Entropy test compared to the (vigenere cipher only), because this method is characterized by its ease compared to other methods of encryption such as RSA and DES thus this research considered as a gateway to experiment other strong encryption methods and compare the results with this research.

Conflict of interests.

There are non-conflicts of interest.

References

[1] Imam S., Mesran, Nelly Astuti Hasibuan and Robbi Rahim"Vigenere Cipher Algorithm with Grayscale Image Key Generator for Secure Text File",International Journal of Engineering Research & Technology (IJERT), Vol. 6 Issue 01, January-2017.

[2] Chukhu Chunka,Rajat Subhra Goswami and Subhasish Banerjee" An efficient mechanism to generate dynamic keys based on genetic algorithm", special issue article, onlinelibrary.wiley, 15 August 2018.

[3] Noor A. Ibraheem and Mokhtar M Hasan," Combining Several Substitution Cipher Algorithms using Circular Queue Data Structure", Baghdad Science Journal, 17(4) 2020.

[4] Sarita Kumari "A research Paper on Cryptography Encryption and Compression Techniques" ,International Journal Of Engineering And Computer Science , Volume 6 ,Issue 4 ,April 2017.

[5] A.A. M. Aliyu and Abdulrahman Olaniyan, "Vigenere Cipher: Trends, Review and Possible Modifications," Internasional Journal of Computer Application, vol. 135, no. 11, pp. 46-50, 2016.

[6] Noor K. A., Asraa A. H, Zainab F. H." Classificstion of Brain MRI Images using Classifier Techniques supported by Genetic and Fuzzy C-Means", Research Journal of Applied Sciences Vol.11, No.10, ,2016.

[7] Md. Shafiul Alam Forhad, Md. Sabir Hossain, Mohammad Obaidur Rahman, Md. Mostafizur Rahaman, Md. Mokammel Haque and Muhammad Kamrul Hossain Patwary" An improved fitness function for automated cryptanalysis using genetic algorithm", Indonesian Journal of Electrical Engineering and Computer Science Vol. 13, No. 2, February 2019.

[8] Sania Jawaid and Adeeba Jamal,"Generating the Best Fit Key in Cryptography using Genetic Algorithm " International Journal of Computer Applications (0975 – 8887) ,Volume 98 – No.20, July 2014.

[9] Ghusoon Salim Basheer," Application of Polyalphabetic Substitution Cipher using Genetic Algorithm", Raf. J. of Comp. & Math's. , Vol. 5, No.1, 2008.

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

[11] Quist-Aphetsi Kester," A cryptosystem based on Vigenère cipher with varying key", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 10, December 2012.

[12] Abdul Razzaq, Yasir Mahmood, Farooq Ahmed and Ali Hur" Strong Key Machanism Generated by LFSR based Vigenère Cipher", The 13th International Arab Conference on Information Technology ACIT'2012.

[13] Gede Arna Jude Saskara and Budi Rahardjo " Design and Implementation Randomized Cryptography Algorithm Detection Using Chaos Algorithm ", IEEE Catalog Number : CFP16B09-PRT, ISBN : 978-1-5090-1719-5,2015.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

[14] Al-Amin Mohammed Aliyu and Abdulrahman Olaniyan," Vigenere Cipher: Trends, Review and Possible Modifications", International Journal of Computer Applications , Volume 135 – No.11, February 2016.

[15]Taronisokhi Zebua and Eferoni Ndruru," How to Create and Randomize a Vegenere Cipher Key Based on the Key Procedure of RC4+ Algorithm", The IJICS (International Journal of Informatics and Computer Science), Vol 3 No 2, September 2019.

الخلاصة

التشفير هو علم حماية المعلومات. يتطلب التشفير مفاتيح قوية لتشفير أو فك تشفير البيانات و هذه المفاتيح يجب ان تكون غير متوقعة وليس من السهل كسرها. في هذا البحث نستخدم الخوارزمية الجينية لتوليد مفاتيح لخوارزمية التشفير vigenere. يتم استخدام أفضل مفتاح لإجراء التشفير و اختبار العشوائية للمفاتيح التي تم إنشاؤها بواسطة الخوارزمية الجينية تتم باستخدام اختبار entropy test. يُظهر حسابات entropy test أن عشوائية المفتاح الذي تم إنشاؤه بناءً على المعالجة الجينية أفضل من المفتاح المختار في تشفير vigenere الكلاسيكي.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com