International Journal of Advances in Applied Sciences (IJAAS)
Vol. 6, No.4, December 2017, pp. 313~318
ISSN: 2252-8814

Dynamic Key Matrix of Hill Cipher Using Genetic Algorithm

Andysah Putera Utama Siahaan
Faculty of Computer Science, Universitas Pembangunan Panca Budi, Medan, Sumatera Utara, Indonesia

Article Info

Article history:

Received Sep 16, 2017
Revised Nov 14, 2017
Accepted Nov 22, 2017

Keyword:

Cryptography
Genetic algorithm
Hill cipher

ABSTRACT

The matrix in Hill Cipher was designed to perform encryption and decryption. Every column and row must be inserted by integer numbers. But, not any key that can be given to the matrix used for the process. The wrong determinant result cannot be used in the process because it produces the incorrect plaintext when doing the decryption after the encryption. Genetic algorithms offer the optimized way to determine the key used for encryption and decryption on the Hill Cipher. By determining the evaluation function in the genetic algorithm, the key that fits the composition will be obtained. By implementing this algorithm, the search of the key on the Hill Cipher will be easily done without spending too much time. Genetic algorithms do well if it is combined with Hill Cipher.

Corresponding Author:

Andysah Putera Utama Siahaan,
Faculty of Computer Science,
Universitas Pembangunan Panca Budi,
Medan, Sumatera Utara, Indonesia.
Email: andiesiahaan@gmail.com

1. INTRODUCTION

Hill Cipher Encryption is a way employed to encrypt messages using the matrix as a key. In this key, there are nine pieces utilized random integers that set a matrix of 3x3. Each number will be associate with each other to generate the cipher text, but we cannot permanently use these numbers to restore the original messages. The numbers must have the exact value of the determinant. Before the numbers could be utilized, we should test these numbers whether we meet the true determinant. The test itself takes time meanwhile these numbers which make up the determinant correct is not necessarily obtained. If the result is wrong, the search of random integers has to be done again. So if we do repeatedly, it will cast a very long time. The problem that arises is an inefficient time if the key on Hill Cipher algorithm is performed manually. Generating keys on Hill Cipher algorithm by combining Genetic algorithms are supposed to speed up the search for the suitable key for the Hill Cipher encryption.Literature review that has been done author used in the chapter "Introduction" to explain the difference of the manuscript with other papers, that it is innovative, it are used in the chapter "Research Method" to describe the step of research and used in the chapter "Results and Discussion" to support the analysis of the results [2]. If the manuscript was written really have high originality, which proposed a new method or algorithm, the additional chapter after the "Introduction" chapter and before the "Research Method" chapter can be added to explain briefly the theory and/or the proposed method/algorithm [4].

2. THEORIES

Hill Cipher is the modulo arithmetic technique in cryptography [1]. Hill Cipher uses the symmetric key as the password to convert plaintext to ciphertext. The Symmetric key is one of the cryptography systems that have the same kind of keys in encryption and decryption. The key employed to encryption is actually distinct from decryption, but they are drawn from the same formula. We must inverse the key before it is

utilized to decrypt the ciphertext. This cryptographic technique has the matrix as the vessel of information exchanges either on encryption or decryption part [2]. The general theory of the matrix used in Hill Cipher is the multiplication between the matrix and the inverse of the matrix [3]. Without getting the right key, the process of encryption and decryption can be done. We can refer to the example of the encrypting message on Figure 1.

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} \quad Mod\ 26$$

Figure 1. Hill Cipher Encryption

C1 is dependent on the multiplication of K11, K12, K13 and P1, P2, P3. The result of the multiplication has modular expression of the total character. Genetic algorithm is a computational algorithm that inspired the theory of evolution which was later adopted into computational algorithms is used to solve a search value in an optimization problem [4]. This algorithm is built on the genetic processes in living organisms. This following figure explains the step of the Genetic algorithm process.

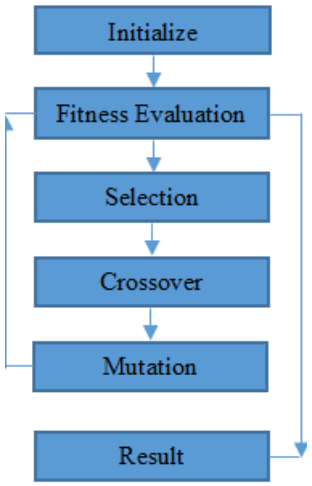

Figure 2. Step of Genetic Algorithm

There are three main steps in Genetic algorithms such as selection, crossover and mutation. Selection is used to recombine the population with the highest probability. The random number generated is combined with the cumulative probability. The nearest value is taken in order to replace the original value of the population.

Figure 3 shows one of the selection methods. The method above is Roulette Wheel Selection. Individuals are mapped into a line segment in a sequence such that each individual segment has the exact same size as its fitness. A random number generated. Individuals who have a segment within the segment in the region of the random numbers will be selected [6]. This process is repeated until the number of individuals is obtained. Crossover is a genetic algorithm operator to mix the chromosome with the extra chromosome chosen to produce child chromosome from one generation to the next. It usually selects some qualified parents [5]. The qualification is the crossover rate value. This value relates to select the parent chromosome.

Figure 4 indicates the crossover based on one cut point division. Mutation is a genetic operator employed to maintain genetic diversity from one generation of a population of genetic algorithm chromosomes to the next. This operator repositions the chromosomes by exchanging the value of the chromosome.

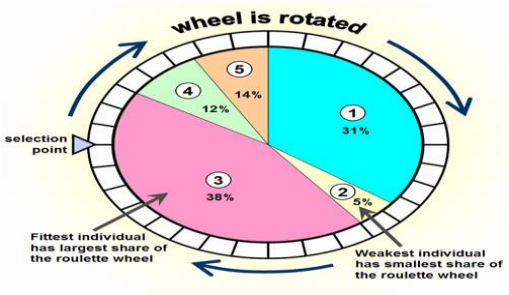

Figure 3. Roulette Wheel Selection

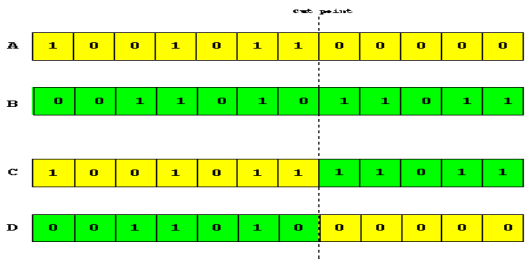

Figure 4. One Cut Point Crossover

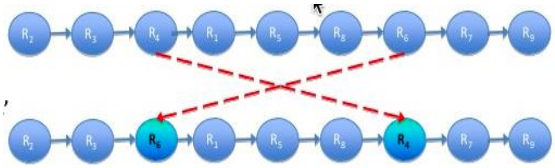

Figure 5. Swap Mutation

Figure 5 shows the example of swap mutation. We can note that the R4 and R6 are swapped. The chromosome index does not change, but the value of the index is substituted for the other value in the other index. Mutation causes movement in the search space and may produce stronger chromosome.

3. RELATED WORK

The previous researches of Hill Cipher do not say anything how to optimize the key search. The key is always obtained by trial and error method. This is not the clever idea to get the correct key in Hill Cipher. It needs the proper determinant to have the cipher text return its original message. It still requires more times to produce the key manually. Once the determinant gets wrong, it has to be recalculated from the beginning. There are nine numbers composed as the key for the encryption and decryption.

4. METHODOLOGY

Every chromosome in Hill Cipher consists of nine numbers. Each gene has a value between 0 to 255 which represents the number of a byte. Since the ASCII value does not exceed 255, we do not take an integer as the modular expression

K11	K12	K13
K21	K22	K23
K31	K32	K33

K11	K12	K13	K21	K22	K23	K31	K32	K33
x	x	x	x	x	x	x	x	x

Figure 6. Hill Cipher Chromosome

Figure 6 shows the form of the Hill Cipher chromosome. The matrix is transformed into one-dimension vector. Each cell is filled with a random integer number (x). The fitness function evaluates the determinant of the chromosome by this following formula.

$$F = D \quad (1)$$

Where
F = Fitness
D = Determinant

The genetic algorithm of this method is rather straightforward than the one of a usual method for scheduling because we do not have to search until the fitness value reaches no error. We just search the ideal fitness which does not contain equal value. Since the determinant is in the odd number, it is good for Hill Cipher. But, in this case, we still to find until the determinant reaches 1.

5. TESTING AND IMPLEMENTATION

Before knowing how the method works, first thing first is to prepare the data. The first step is to generate the initial population. Consider we set to Generation is 30 and Population Size is 20. The initial random population is given in table 1. Each number in the cell is a random integer.

Table 1. Initial Population

K11	K12	K13	K21	K22	K23	K31	K32	K33
108	85	165	89	69	185	97	185	54
65	86	135	47	224	116	213	112	6
227	29	41	44	1	141	101	218	32
174	195	136	196	220	37	108	144	43
31	232	46	145	120	234	196	242	63
190	54	140	128	118	179	151	108	43
90	64	85	24	242	106	154	178	244
172	124	121	98	194	81	73	183	135
215	82	163	62	103	13	79	165	164
147	47	77	212	45	112	249	18	80
213	7	244	195	246	197	244	25	119
196	92	115	59	55	190	44	191	27
134	29	216	72	200	78	196	7	131
168	214	80	10	104	177	1	114	177
145	20	91	221	73	79	149	137	73
50	89	246	142	168	108	85	116	244
119	141	61	167	254	239	66	77	65
40	187	243	193	58	195	14	154	172
190	168	210	137	178	63	5	146	173
139	26	47	226	179	242	187	137	228

Every chromosome is calculated to get the fitness, probability and cumulative probability. The next step is to conduct selection, crossover, and mutation. Table 2 shows the fitness, probability, and cumulative probability of population.

Table 2. Fitness, Probability, and Cumulative Probability

F	P	CP
1	0,0003652	0,0003652
64	0,0233747	0,02374
210	0,0766983	0,1004383
238	0,0869248	0,187363
9	0,0032871	0,1906501
76	0,0277575	0,2184076
150	0,0547845	0,2731921
186	0,0679328	0,3411249
203	0,0741417	0,4152666
174	0,06355	0,4788167
241	0,0880205	0,5668371
84	0,0306793	0,5975164
0	0	0,5975164
152	0,055515	0,6530314
172	0,0628196	0,715851
188	0,0686633	0,7845142
223	0,0814463	0,8659606
169	0,0617239	0,9276844
2	0,0007305	0,9284149
196	0,0715851	1

The process of selection, crossover and mutation has modified the population structure. The order of the key changes to the other number from other chromosomes. The new generation results from the latest updated chromosomes as saw in Table 2.

Table 3. Next Generation Population

K11	K12	K13	K21	K22	K23	K31	K32	K33
134	47	216	112	213	90	18	7	131
85	47	244	195	85	212	244	196	249
119	63	77	78	145	72	249	18	135
108	47	197	179	47	239	119	234	65
167	89	7	142	179	108	200	116	89
78	64	116	24	242	112	27	191	80
80	44	185	7	1	141	62	242	32
254	29	106	196	200	50	13	29	185
31	232	46	145	120	216	154	178	244
139	85	227	226	212	165	168	137	228
218	244	41	163	168	147	112	242	244
228	85	147	234	69	97	131	141	54
147	47	77	61	232	246	249	18	80
215	82	142	101	103	29	79	165	164
187	120	86	59	55	190	44	6	213
31	45	46	134	116	89	196	242	63
65	86	246	196	224	92	66	112	72
77	115	135	246	224	26	213	112	6
116	196	65	77	45	212	25	137	50
139	26	47	226	45	47	187	242	108

This process continues to the latest generation. It is the last result where the Hill Cipher keys are generated. At the end of the process to the total of the Hill Cipher key is variant. Table 4 shows the final result of the Genetic algorithm.

Table 4. The Final Result

K11	K12	K13	K21	K22	K23	K31	K32	K33
147	69	62	147	232	82	29	147	147
147	62	147	232	246	29	82	147	246
82	147	232	69	246	82	246	72	147

After getting the result, the key combination is now formed as showed in Table 5. The numbers showed in the blocks are generated by the genetic algorithm. We cannot do the same way manually because it takes much time to calculate the determinant manually.

Table 5. The Key Combination

Key 1			Key 2			Key 3		
147	69	62	147	62	147	82	147	232
147	232	82	232	246	29	69	246	82
29	147	147	82	147	246	246	72	147

The key in the block section is used to encrypt the plaintext in Hill Cipher algorithm.

Plaintext : ANDYSAHXX $\begin{pmatrix} 65 & 78 & 68 \\ 89 & 83 & 65 \\ 72 & 88 & 88 \end{pmatrix}$

Key : $\begin{pmatrix} 147 & 69 & 62 \\ 147 & 232 & 82 \\ 29 & 147 & 147 \end{pmatrix}$

Ciphertext : ÑË38%`H8 $\begin{pmatrix} 209 & 203 & 51 \\ 56 & 37 & 17 \\ 96 & 72 & 56 \end{pmatrix}$

Key Inverse : $\begin{pmatrix} 82 & 131 & 218 \\ 249 & 171 & 228 \\ 57 & 40 & 241 \end{pmatrix}$

Plaintext : ANDYSAHXX $\begin{pmatrix} 65 & 78 & 68 \\ 89 & 83 & 65 \\ 72 & 88 & 88 \end{pmatrix}$

We see there are three keys produced by the genetic algorithm. We have tested the first key above. The determinant has value 1 and it is suitable for the Hill Cipher process. When the determinant is 1, it can bring the ciphertext back to the plain text and vise versa.

6. CONCLUSION

On the Hill Cipher algorithm that uses a 3x3 matrix, searching key that has the proper determinant takes time. If we perform in that way, it slows down the process of cryptography. Genetic algorithms greatly assist the process of the encryption and the decryption on the Hill Cipher. It produces a series of numbers quickly. This technique generates several alternatives that can be used on the Hill Cipher algorithm. In this study, we conclude that the genetic algorithm has a valuable contribution when combined with the Hill Cipher method.

RERERENCES

[1] Abdullah, A. A., Khalaf, R., & Riza, M. (2015). A Realizable Quantum Three-Pass Protocol Authentication. Mathematical Problems in Engineering.
[2] Chase, J., & Davis, M. (2010). Extending the Hill Cipher.
[3] Chowdhury, S. I., Shohag, S. A., & Sahid, H. (2011). A Secured Message Transaction Approach by Dynamic Hill Cipher Generation and Digest Concatenation. International Journal of Computer Applications, 23(9), 25-31.
[4] Ghanbari, A. A., Broumandnia, A., Navidi, H., & Ahmadi, A. (2012). Brain Computer Interface with Genetic Algorithm. International Journal of Information and Communication Technology Research, 2(1), 79-86.
[5] Lin, C. H., Yu, J. L., Liu, J. C., Lai, W. S., & Ho, C. H. (2009). Genetic Algorithm for Shortest Driving Time in Intelligent Transportation Systems. International Journal of Hybrid Information Technology, 2(1), 21-30.
[6] Szénási, S., & Vámossy, Z. (2013). Implementation of a Distributed Genetic Algorithm for Parameter Optimization in a Cell Nuclei Detection Project. Acta Polytechnica Hungarica, 10(4), 59-86.