International Journal of Advanced Technology and Engineering Exploration, Vol 5(43)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2018.543001

Research Article

An efficient algorithm for polyalphabetic substitution using random tables

Ranju S Kartha[1*] and Varghese Paul[2]
Research Scholar, School of Computer Science, Mahatma Gandhi University, Kottayam, Kerala, India[1]
Professor, Department of Information Technology, RSET, Cochin, Kerala, India[2]

Abstract
There are many cryptographic techniques available for providing a secure communication. Encryption technique can be classified according to their encrypting process. They are substitution cipher and transposition cipher. Polyalphabetic cipher is based on substitution technique- the plaintext letters are encrypted differently depending upon their placement in the text and the keyword. Vigenere cipher is considered to be the most efficient and simplest Polyalphabetic substitution cipher. Due to its repeating nature of the keyword, it is vulnerable to attacks. To overcome this, here we are presenting a new cipher which uses multiple random Tables (26×26) for encryption. In this proposed cipher, for encrypting each plaintext letter we are generating a random table (26×26). Instead of using the same Vigenere table here we are using an infinite number of alphabetical tables depending on the length of the plaintext. Also, each table will be completely independent from the previous table. The repeating nature of the keyword does not help the crackers to break this code. So this proposed polyalphabetic cipher is unbreakable.

Keywords
Polyalphabetic cipher, Vigenere cipher, Vigenere table, Kasiski method, Index of coincidence (IC).

1.Introduction
Data is the heart of all organizations so data security remains the biggest concern for application customers [1]. Data needs to be handled securely at all stages, from transmission, computation and persistence. Nowadays the use of internet and mobile technology continues to rise in emerging markets. Thus there is a growing concern for privacy and data security. In today's world there are many applications such as secret message transferring in military systems, payments to private organisations, personal emails, data storage in personal devises and secure exchange of password, where the information security issues are more challenging and complex [2−6]. One of the methods for ensuring security is cryptography deals with hiding the real information [7, 8].

Cryptology is a science and the secret communication involves cryptography and cryptanalysis. Cryptography is the art and science of transforming messages to make them secure and immune to attack [3−9]. The word cryptography is a Greek term 'crypto' means secret and 'graphy' means writing. The basic goal of cryptography is to send messages so that no one, but the expected recipient can read.

Plaintext is the message in its original form and ciphertext is the message in coded form after the original message is encrypted.

Encryption is the process of transforming plaintext into ciphertext and decryption is the method of transforming ciphertext back into plaintext. To encrypt the message we need two main things: cipher and the key. Cipher means the set of rules that we are using to encode the message. And the key tells how to arrange those rules. Otherwise the rule will be same at all the time and anybody can decrypt the message very easily. To decrypt the message we need the cipher which we used and the key. Usually the attacker cracks the code by trying all possible combinations of key or analyzing the code by working backward from it. If it is not possible to determine the combination of cipher and the key we can say that the code is an unbreakable one. People keep coming up with new and better ciphers but it is hard to make them unbreakable.

Substitution cipher is a classical method of cryptography, it replace every plaintext letter by a corresponding ciphertext letter. One of the oldest and simplest ways of encrypting the message is Caesar cipher. In this case the key is just a number representing how many letters of the alphabet we shift it. For example, in a Caesar cipher of shift 3, A

*Author for correspondence

would become D, B would become E and so on. The shift is performed modulo 26 [9]. But it is easy to crack, even if we didn't know the key, we can do 25 tries to get the message. Caesar cipher is one simple type of monoalphabetic cipher. Monoalphabetic substitution cipher is a class of ciphers where the code is based on one letter of the alphabet standing in for another letter consistently throughout the message.

There are lot of ways to decrypt the message; the most commonly used method is Brute Force attack. In monoalphabetic cipher there are 26! possible keys, so Brute Force attack becomes infeasible. The most sophisticated technique for the cryptanalysis of monoalphabetic cipher is called frequency analysis. It is based on the language we used for encryption where certain letters and combination of letters occur with varying frequencies. In English, e is the most common letter, followed by t, then a and so on. The cracker can calculate the frequency of the letters appears in the ciphertext and relates them to the frequency of the language we used.

If the multiple occurrences are replaced by different cipher text characters, it is called polyalphabetic substitution ciphers [7]. There are many cryptographic algorithms in polyalphabetic substitution scheme. But as the number of ciphers increases, the crackers are trying to break to them. So it is very difficult to generate an unbreakable cipher.

2.Polyalphabetic substitution cipher

A polyalphabetic substitution cipher uses multiple monoalphabetic cipher substitutions so that same plaintext alphabet mapped into different alphabets. And the key is used to specify the mapping. For encryption the plaintext is divided into different groups, each group having *m* elements where *m* is equal to the length of the keyword. Here the keyword is repeated until it matches with the length of the plaintext. The elements in each group are encrypted using the corresponding letter in the keyword. If same letter is repeating in a group, it will be encrypted as a different element, depending upon the key letter. The Polyalphabetic ciphers have the advantage of hiding the letter frequency of the underlying language [8]. So the cracker cannot use single letter frequency analysis to break the code.

Let the plaintext P = P_1, P_2, P_3 …P_n is encrypting using the key stream K =K_1, K_2, K_3 …K_m, and then the encryption and decryption algorithm can be defined as:

Encryption: $C_i = P_i + K_{i\ (mod\ m)}$ *(mod26)*,
Decryption: $P_i = C_i + K_{i\ (mod\ m)}$ *(mod26)* (1)

Where $C_i = C_1, C_2, C_3 \ldots Cn$ is the ciphertext. The best well-known algorithm in polyalphabetic cipher is referred to as Vigenere cipher.

3.Existing cipher

Blaise de Vigenere a French cryptographer of the 1500s created this cipher called as Vigenere Cipher. Vigenere cipher has a similar method of encryption as Caesar cipher but uses a far more complex encryption key. In this method the alphabetic text is encrypted using a table of alphabets, termed as Vigenere table (*Figure 1*). It is a 26×26 table with A to Z as row heading and column heading. Each row comprises all the 26 alphabets of English. The first row has 26 letters in alphabetic order. From second row, each row has the letters shifted to left by one position in a cyclic way.

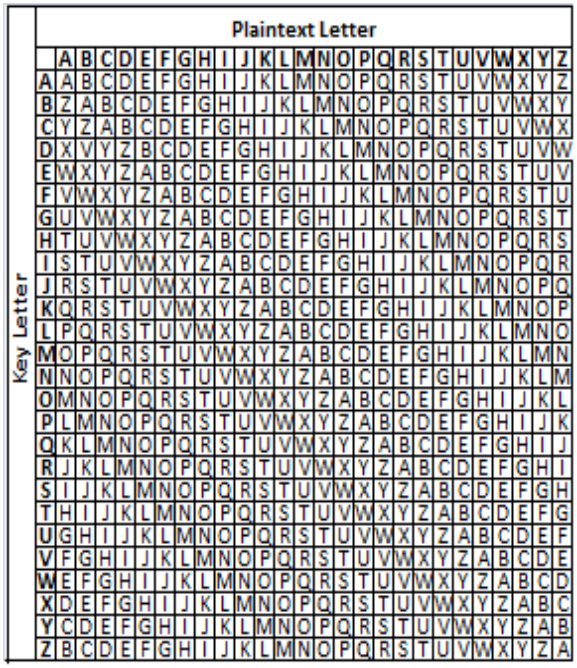

Figure 1 Vigenere table

In Vigenere cipher for encryption and decryption, the keyword is repeated until it matches with the plaintext [3]. For encryption, the letter in the plaintext selects the column index, and the corresponding letter in the keyword selects the row index and the entry at the corresponding row-column intersection is the letter in the ciphertext. Repeat this process until all the letters in the plaintext are processed.

Example: Suppose the message to be encrypted is cryptography (*Table 1*) and uses the keyword TIME, repeats the keyword until it is same length as the plaintext.

Table 1 Encryption process

Plaintext	C	R	Y	P	T	O	G	R	A	P	H	Y
Keyword	T	I	M	E	T	I	M	E	T	I	M	E
Ciphertext	V	Z	K	T	M	W	S	V	T	X	T	C

Decryption is performed by using the letter in the keyword to select the row index and find the position of the ciphertext letter in the corresponding row. The letter heading of the column that contains ciphertext letter is the needed plaintext letter. The number of possible solutions for this cryptosystem grows with the length of the text by a power of 26. The strength of this cipher is that the same letter in the plaintext can be encrypted in different ways. Vigenere cipher is one of the great breakthroughs in the world of cryptography; it was unbreakable for hundreds of years.

4.Cryptanalysis of Vigenere cipher

Cryptanalysis of classical ciphers is made possible because of the redundancy in the linguistic structure of natural languages [5]. In monoalphabetic substitution cipher, the most frequent letters in the ciphertext corresponds to the most frequent letters in the plaintext. So the cracker can easily break the code by performing frequency analysis on the letters in the ciphertext. The Vigenere cipher masks the frequency with which a character appears in a language, which makes the use of frequency analysis more difficult [4]. The frequency distribution of the ciphertext is much more flat.

The primary weakness of the Vigenere cipher is the repeating nature of its key [6]. Vigenere cipher is easily broken if the cracker discovers the length of the keyword. So the security of this cipher relies on having the key length unknown. Suppose the length of the keyword is *m*, once it is known to the cracker, he can split the ciphertext into different block of size *m*. Every m^{th} character of the ciphertext is encrypted using the same shift. So the cracker can write these blocks into a matrix in a way that each row is filled with letters of each block. Now the letters in the column have been encrypted using the same key. The cracker can easily break this code by performing frequency analysis on each column. This is possible if the keyword is repeated. Otherwise for short messages the Vigenere cipher is unbreakable.

Attacking a Vigenere cipher involves two steps, first one is to determine the length of the keyword and the second one is to find the letters of the keyword. There are two methods to find the length of the keyword. They are the Kasiski method - to find the keyword length using the repeated text sequence in the ciphertext and the IC to predict the number of alphabets used for substitution [4]. If the length of the ciphertext is too small or it does not include any repetitions of string, then these two methods cannot break the cipher.

The Kasiski Test was discovered independently by Charles Babbage and Friedrich Kasiski. This method is based on the following observation: if a string of characters of length three (trigrams) or more appears repeatedly in the ciphertext message, it is possible that the distance between the reoccurring characters is a multiple of the length of the keyword. This method follows the rule: if a message is encrypted with *m* alphabets (key length is *m* for Vigenere cipher), and if a particular word or letters group appears *d* times in the plaintext, then it should be encrypted approximately *d/m* times from the same alphabet [10]. So in this method first we find the repeated sequence of characters of length three or more in the ciphertext, and then find the distance between the successive repeated sequences. Next we have to determine the greatest common divisor of all these distances, and the keyword length should be one factor of that GCD.

$d \equiv 0 \pmod{m}$, where m is the key length (2)

Some of the repeated sequences in the ciphertext arise this way due to coincidence, but the probability of a repetition by chance is noticeably smaller. If the Kasiski test was successful, the cracker will get the keyword length. So he can divide the ciphertext into *m* different blocks and applies the methods which used to crack the monoalphabetic cipher, including frequency analysis.

William Friedman developed a statistical method that helps the cryptanalyst to guess the cryptosystem used in the ciphertext. Also it determines the length of the keyword if the cipher is Polyalphabetic. So the Friedman's test is used for cracking Vigenere cipher based on the value of Index of Coincidence (IC). If we pick two letters from the text at random, most of the time the letters will be different, but sometimes they will be same. The IC measures the probability that two randomly selected letters of the string are identical [4].

Suppose a particular letter appears *n* times among *N* letters, there are *N (N-1)* ways we can pick two letters at random. And there are *n (n-1)* ways we can pick the designated letter, so the probability that both letters we pick are designated letter will be *n (n-1)/N (N-1)*. It follows that the IC will be

$$IC = \frac{\sum_{i=1}^{c} n_i (n_i - 1)}{N(N-1)}, \quad (3)$$

Refer to "Equation (3)", where n_1 through n_c are the frequencies (as integers) of the *c* letters of the alphabet (*c* = 26 for English language).

In typical English text, because of the distribution of letters in this language about 6.8% of the randomly chosen pairs will consist of identical letters. While a text of randomly chosen letters will have IC as low as 3.8%. This feature is presented by a substitution cipher. In the case of monoalphabetic cipher, the frequency of letters in the ciphertext should be nearly the same as for English – but in a different order. So if the IC of the ciphertext is closer to 0.065, the more likely we have a monoalphabetic cipher. But in the case of Polyalphabetic substitution cipher, the frequencies of the letters would become more nearly uniform so the IC is closer to 0.038.

The IC can be used to estimate the length of the unknown keyword. The cracker can guess a keyword length *m* and divide the ciphertext into m strings. These substrings are referred as cosets. If the length of the keyword *m* is correct each coset would preserve the $IC_{English}$ to some degree. Therefore the average of IC's of these coset would still be high and close to $IC_{English}$ = 0.068. Otherwise the average of IC's would be low. Based on these observations the cracker can divide the ciphertext into 1 coset, 2coset etc according to the key length that he had guessed and compute the IC of each coset and its average. The length that yields the highest average IC value or close to $IC_{English}$ is likely to be the correct length of the keyword.

5.Proposed polyalphabetic cipher

The fundamental weakness of the Vigenere cipher is the repeating nature of its key. In the existing cipher, the keyword repeats until it is equal to the length of the plaintext during encryption.

Each time when the keyword repeats the Vigenere cipher uses the same Vigenere Table. In the proposed cipher also the keyword repeats until it is equal to the length of the plaintext. But here, for encrypting each plaintext letter this cipher will generate 26x26 random tables. That means multiple number of 26x26 tables is used for encrypting the plaintext depending on the length of plaintext, not depending on the keyword. The randomly generated 26*26 table will be exactly different from the previous tables. The table having 26 rows and 26 columns, each row and column have all the alphabetic characters without any repeat. The decryption of this cipher without knowing the key will be impossible. Here we can generate 26! = 4.032914611266056e+26 tables. Hence we can say that this cipher is unbreakable.

6.Result analysis

The program for new Polyalphabetic cipher was developed and used to encrypt a text message and the result was compared with the existing Vigenere cipher.

6.1Result from the existing Vigenere cipher

Example 1:
Enter the Text to be encrypted:
THEVI GENER ECIPH ERISA METHO DOFE NCRYP TINGA LPHAB ETICT EXTBY USING ASERI ESOFDI FFERE NTCAE SARCI PHERS BASED ONTH ELETT ERSOF AKEYW ORDI TISVE RYSIM PLE

Enter the Encryption Key: **TEXT**
Encrypted text:
MLBO BKBGX VBVBT EXKMP TFIQA HHLYX RZKRT QBGKX EILXU XXFVM IUMUC RLBRD TLIOB XWLYW MCYXV BGMGX XLEOVB TEXKWY TLIAHGX EXEIGM XVPHY EHXRA LKWMQ BLESX KCPBF TIX

The frequencies of occurrence of each letters in English alphabet in the above ciphertext are shown in *Figure 2*. Here we can see the frequency of occurrence of x is 18 and n and j are zero. Frequencies of occurrence of the remaining alphabets are varying between these two values. So it is almost identical to the letter frequencies in the ordinary English alphabets, but associated with different letters.

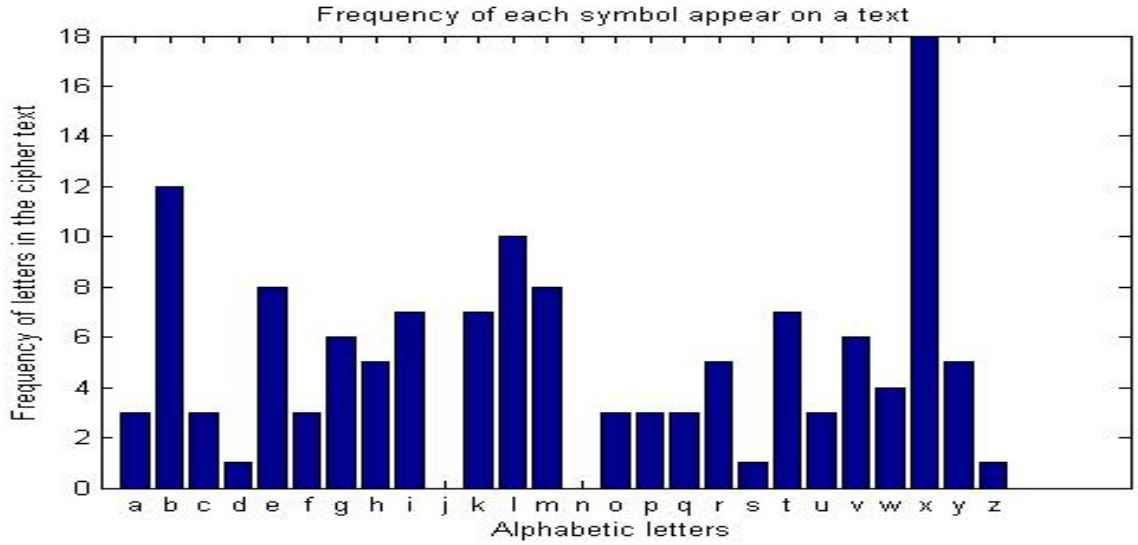

Figure 2 Frequency analysis of the traditional Vigenere cipher

6.2Result from the proposed algorithm

Enter the Text to be encrypted:

THEVI GENER ECIPH ERISA METHO DOFEN CRYPT INGAL PHABE TICTEX TBYUSI NGASE RIESOF DIFFER ENTCAE SARCIP HERSBA SEDONT HELETT ERSOFA KEYWOR DITISV ERYSIM PLE

Enter the Encryption Key: **TEXT**

The ciphertext is:

FXTI BNTWO ETSCH POBGE MHSKKJ IYIOA XAVHK DWNRL DMRQ OJHXGS ZFQZS YCANM ZSYCX DYPEG UPOLT WFURU ZQFUC JPRAD VMZSJ TWJP BLSKF BEETP QWRVB YAEGK LZQNO AZECI HQO

The frequency of occurrence of each alphabet in the above ciphertext is shown in *Figure 3*. In this graphical representation we can see that e, *o, s* and *z* having the highest frequency of occurrence and it is equal to 7. And letter *v* having least frequency of occurrence and that is 3. Frequency of occurrences of remaining letters in the alphabets is varying in between these two values. Here the frequency distribution of letters in the ciphertext is almost flat. It is impossible to break this cipher using frequency analysis.

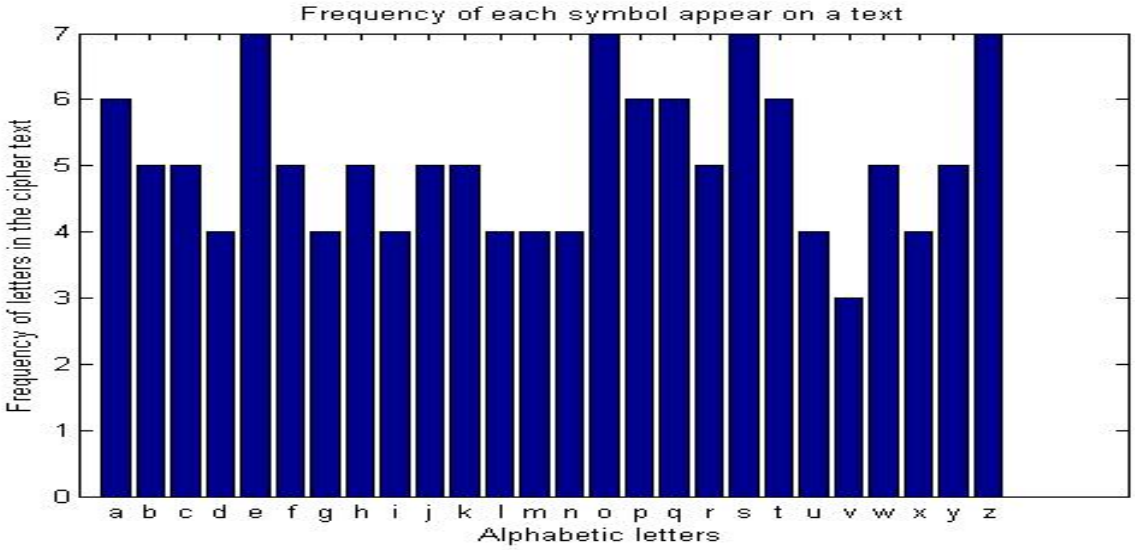

Figure 3 Frequency analysis of the proposed polyalphabetic cipher

6.3Analysis based on Kasiski method to determine the key length

The plaintext was encrypted using the program and analyzed below by applying the Kasiski attack. *Table 2* shows the analysis of trigram in the ciphertext.

Example 2:
Consider the existing Vigenere Cipher,
Plaintext:
THERE ARETW OWAYS OFCON STRUC TINGA SOFTW AREDE SIGNO NEWAY ISTOM AKEIT SOSIM PLETH ATTHE REARE OBVIO USLYN ODEFI CIENC IESAN DTHEO THERW AYIST OMAKE ITSOC OMPLI CATED THATT HEREA RENOO BVIOU SDEFI CIENC IESTH EFIRS TMETH ODISF ARMOR EDIFF ICULT
Encryption Key: **SYSTEM**
Ciphertext:
LFWKI MJCLP SISWK HJOGL KMVGU RAGKM KMXMA MJCVX WUYLG GIISW ALXAE YCXMF KMKBQ BDCLA EFLFW KIMJC GUZUG SKECZ GBWYM OACFV MQKYF WXTWM LAIDO YQBWF GKSDI ULQGV SYHJA VEFWB LAEFL FWKIM JCFHS NNGGN WPWDA VMQFA AXWFZ CXBVE LKWML AVGKY EDEMJ XHUXD AVYXL

Table 2 Analysis of trigram in the ciphertext

Positions	Distance	Plaintext	Keyword	Ciphertext
5	30	ARE	MSY	MJC
35		ARE	MSY	MJC
11	36	WAY	MSY	ISW
47		WAY	MSY	ISW
28	32	GAS	EMS	KMK
60		SOS	SYS	KMK
99	66	CIE	TEM	VMQ
165		CIE	TEM	VMQ
163	36	FIC	YST	DAV
199		FIC	YST	DAV

In the above ciphertext there are five repeating strings of length three. Its analysis can be shown in the following table. The repeating ciphertext **KMK** is encrypted from two plaintext portions **GAS** and **SOS** with keyword **EMS** and **SYS** respectively. This is a case of mere coincidence. So we are not considering this for finding the keyword length. *Table 3* shows the analysis of long repeated sequences in the ciphertext.

Table 3 Analysis of long repeated sequences in the ciphertext

Length	Distance	Factors
8	72	2, 3, 4, 6, 8, 9, 12, 18, 24, 36, 72
3	66	2, 3, 6, 11, 22, 33, 66
	36	2, 3, 4, 6, 9, 12, 18, 36
	30	2, 3, 5, 6, 10, 15

The next long string of length 4 is WMLA occurring at positions 108 and182. So the distance between these two positions is 74. At position 108 and 182 the plaintext is EOTH and the corresponding key is SIST. Next repeated long string LFWKIMJC of length 8. It occurs at three positions 0, 72 1nd144, so the distance between two occurrences is 72. At all these three positions the plaintext and the key is same.

Now we can compute the GCD of these distances and use it as the length of the keyword. Here GCD (30, 36, 66, 72) = 6, according to Kasiski test the length of the keyword is 6 and we know that this is correct.

Consider the proposed polyalphabetic cipher,
Plaintext:
THERE ARETW OWAYS OFCON STRUC TINGA SOFTW AREDE SIGNO NEWAY ISTOM AKEIT SOSIM PLETH ATTHE REARE OBVIO USLYN ODEFI CIENC IESAN DTHEO THERW AYIST OMAKE ITSOC OMPLI CATED THATT HEREA RENOO BVIOU SDEFI CIENC IESTH EFIRS TMETH ODISF ARMOR EDIFF ICULT

Encryption Key: **SYSTEM**
Ciphertext:

KKNNW MZZKZ BIPMR GVGSO RYINJ WUQOM RERGV HLBZB QKRGF QBZNX PODQK EEDKK UWRGA REHIB MVPWP MWONB ZBBTX GSEAF AOSVW JUVUB LGUZR DLEKK NGTVN LLDJR RWSSY YNGKF BGSQV WDGPW NVTQP WKMWW CBZBL PSIXI BNRTN ADGUZ CXDKR WYBNR ZDKSW HYEHI AJPLE GIKHG MADGQ

Table 4 shows the analysis of repeated sequences in the ciphertext.

Table 4 Analysis of repeated sequences in the ciphertext

Positions	Plaintext	Keyword	Ciphertext
1	THE	SYS	KKN
109	EOT	SYS	KKN
15	SOF	STE	RGV
33	FTW	STE	RGV
38	EDE	YST	BZB
152	ENO	TEM	BZB
99	BNR	TEM	BNR
178	IRS	MSY	BNR
163	CIE	MSY	ADG
202	CUL	MSY	ADG

In the above ciphertext there are five repeating strings of length three. The plaintext corresponding to the repeated sequence is different, so all the repeated sequence is by mere pure chance. If a match is due by pure chance, the factors of this distance may not be a factor of keyword length. So here it is impossible to find the keyword length.

6.4Index of coincidence

The IC for a given text can be computed using the following "Equation (3)", consider the plaintext from example 2 and encrypt using Vigenere cipher with the key COMPUTER we will get the following ciphertext:

VVQGY TVVVK ALURW FHQAC MMVLE HUCAT WFHHI PLXHV UWSCI GINCM UHNHQ RMSUI MHWZO DXTNA EKVVQ GYTVV QPHXI NWCAB ASYYM TKSZR CXWRP RFWYH XYGFI PSBWK QAMZY BXJQQ ABJEM TCHQS NAEKV VQGYT VVPCA QPBSL URQUC VMVPQ UTMML VHWDH NFIKJ CPXMY EIOCD TXBJW KQGAN

To determine the key length we have to guess the key length and divide the ciphertext into different coset according to the key length. By calculating the IC's of the cosets following observations are made.

Table 5 shows the largest average IC value 0.0731 corresponds to the keyword length 8, so m =8 is the most likely keyword length.
So in the case of Vigenere cipher we can find the correct keyword using this cryptanalysis.

Consider the above example encrypted using the proposed Polyalphabetic cipher we will get the ciphertext as:

TYWUR USHPO SLJNQ AYJLI FTMJY YZFPV EUZTS GAHTU WNSFW EEEVA MYFFD CZTMJ WSQEJ VWXTU QNANT MTIAW AOOJS HPPIN TYDDM VKQUF LGMLB XIXJU BQWXJ YQZJZ YMMZH DMFNQ VIAYE FLVZI ZQCSS AEEXV SFRDS DLBQT YDTFQ NIVKU ZPJFJ HUSLK LUBQV JULAB XYWCD IEOWH FTMXZ

Here the IC Value is calculated as 0.0399 which is less than the IC value obtained in the previous case and which is almost equal to the IC of random string (0.038).

Table 6 shows all the average IC values are low and almost equal to the IC of random string (0.038). So it is impossible to find the keyword length. Here the highest average IC value is 0.0581 which is lower than the $IC_{English}$ so m = 6 is the wrong guess.

Table 5 Average IC value for the possible keyword length from 1 to 10

Length	IC Value of the cosets	Average IC
1	0.0419	0.0419
2	0.0468. 0.0448	0.0458
3	0.04, 0.0456, 0.0465	0.044
4	0.0505, 0.047, 0.0572, 0.0423	0.0492
5	0.0451, 0.039, 0.0402, 0.0353, 0.039	0.0397
6	0.0521, 0.0623, 0.0588, 0.0392, 0.0516, 0.052	0.0481

Length	IC Value of the cosets	Average IC
7	0.039, 0.0413, 0.0418, 0.032, 0.0689, 0.0418, 0.0295	0.0421
8	0.0584, 0.0553, 0.0861, 0.04,0.1015,0.0633, 0.0966, 0.0833	0.0731
9	0.0316, 0.0434, 0.0750, 0.0474, 0.0434, 0.0434, 0.0237, 0.0432, 0.0432	0.0438
10	0.0666, 0.0190, 0.0285, 0.0333, 0.0428, 0.0631, 0.0315, 0.0473, 0.0315, 0.0263	0.039

Table 6 Average IC value for the possible keyword length from 1 to 10 in the case of proposed cipher

Length	IC Value of the cosets	Average IC
1	0.0399	0.0399
2	0.0438, 0.0412	0.0425
3	0.0403, 0.0456, 0.0465	0.0441
4	0.044, 0.0466, 0.0472, 0.0486	0.0466
5	0.0331, 0.037, 0.0402, 0.333, 0.037	0.0361
6	0.0621, 0.0633, 0.0588, 0.0492, 0.0536, 0.062	0.0581
7	0.039, 0.0313, 0.0418, 0.032, 0.0486, 0.0318, 0.0295	0.0362
8	0.0584, 0.0553, 0.0661, 0.04,0.0515,0.0433, 0.0366, 0.0343	0.0482
9	0.0336, 0.0434, 0.0750, 0.0474, 0.0542, 0.0433, 0.0435, 0.0441, 0.0432	0.0475
10	0.0666, 0.0190, 0.0285, 0.0333, 0.0428, 0.0631, 0.0315, 0.0473, 0.0315, 0.0283	0.0391

6.5Frequency analysis

The frequency analysis of the ciphertext (Example 1) does not help the cracker to decipher.

It has been stated in the *Table 7* with English letter frequency, Vigenere cipher frequency and proposed cipher frequency.

Table 7 Frequency analysis- proposed cipher with traditional vigenere cipher

English alphabet	Frequency of english letters %	Vigenere cipher %	Proposed cipher %
A	8.17	2.27	4.55
B	1.49	9.09	3.79
C	2.78	2.27	3.79
D	4.25	0.76	3.03
E	12.7	6.06	5.3
F	2.23	2.27	3.79
G	2.02	4.55	3.03
H	6.09	3.79	3.79
I	6.97	5.3	3.03
J	0.15	0	3.79
K	0.77	5.3	3.79
L	4.03	7.58	3.03
M	2.41	6.06	3.03
N	6.75	0	3.03
O	7.51	2.27	5.3
P	1.93	2.27	4.55
Q	0.10	2.27	4.55
R	5.99	3.79	3.79
S	6.33	0.76	5.3
T	9.06	5.3	4.55
U	2.76	2.27	3.03
V	0.98	4.55	2.27
W	2.36	3.03	3.79
X	0.15	13.64	3.03
Y	1.97	3.79	3.79
Z	0.07	0.76	5.3

Here we observe that the frequency of occurrence of alphabetic letters in the proposed cipher is almost flat. So it is hard to break proposed cipher compared to traditional Vigenere cipher. In the proposed cipher, for encryption we are using random number of 26×26 tables, leading to decrease in the effectiveness of Kasiski and IC attacks as shown in above tables.

The main weakness of proposed system is that, for encrypting long plaintext we have to generate an infinite number of 26×26 random table. But it increases the security and complexity of the cipher. So we can say that this proposed system is an unbreakable Polyalphabetic substitution cipher.

7.Conclusion

The proposed Polyalphabetic cipher overcomes the primary weakness of the Vigenere Cipher by using random number of 26×26 tables. In this cipher the keyword is repeating until it is equal to the length of the plaintext, but for encrypting each plaintext, it uses different 26×26 tables. So this system can generate infinite number of 26×26 random tables. This cipher does not depending on the length of the keyword. We can modify this cipher by constructing 68×68 matrix, consisting of alphabets (1 to 26), numbers (0 to 9) and all the symbols present on the keyboard (32). So we can encrypt all text and keyboard [1]. Proposed polyalphabetic cipher can provide security for many applications such as personal emails, web transactions, confidential information transmitted between public or private organization, military application etc. As cryptography grows without any boundary which in turn cause an increase in activities of the cryptanalyst to find new loopholes. Hence cryptology offers immense potential for research activities.

Acknowledgment

We like to thank the Director of School of Computer Science, Mahatma Gandhi University, PD Hills, Kottayam for providing all the facilities to complete the task.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Ravindra Babu K, Kumar DU, Babu V and kumar Shravan. A contemporary poly alphabetic cipher using comprehensive Vigenere table. World of Computer Science and Information Technology Journal. 2011; 1(4):167-71.

[2] Goel N, Maurya A, Kumar B. Information security: encryption and decryption with polyalphabetic substitution method. International Journal of Computer Science and Communication. 2011; 2(1):41-4.

[3] Mendrofa EH, Purba EY, Siahaan BY, Sembiring RW. Collaborative encryption algorithm between vigenere cipher, rotation of matrix (ROM), and one time pad (OTP) Algoritma. Advances in Science, Technology and Engineering Systems Journal.2017; 2(5):13-21.

[4] Kartha RS, Paul V. Survey: recent modifications in Vigenere Cipher. IOSR Journal of Computer Engineering. 2014; 16(2):49-53.

[5] Eskicioglu A, Litwin L. Cryptography. IEEE Potentials. 2001; 20(1):36-8.

[6] Razzaq A, Mahmood Y, Ahmed F, Hur A. Strong key machanism generated by LFSR based Vigenère cipher. International Arab conference on information technology 2012(pp.544-8).

[7] Benny A, Mathews M. An analysis into the efficiency of Ciphers. International Journal of Scientific & Engineering Research.2017; 8(7):1303-10.

[8] Forouzan BA. Cryptography & Network Security. McGraw-Hill; 2007.

[9] Katz J, Menezes AJ, Van Oorschot PC, Vanstone SA. Handbook of applied cryptography. CRC Press; 1996.

[10] http://www.scribd.com/doc/36928010/Classical-Cryptography. Accessed 26 March 2018.

Ms. Ranju S. Kartha completed B.Tech from M. G. University Kottayam in 2006 and M.Tech from Amrita University, Coimbatore in year 2008. She is currently pursuing Ph.D. and currently working as Assistant Professor in Department of Electronic and Communication Engineering, Sree Narayana Gurukulam College of Engineering, Kadayiruppu since 2008. She is a life member of Indian Society for Technichal Education (ISTE). She has published many research papers in reputed international journals. Her main research work focus on Cryptography Algorithms. She has 10 years of teaching experience.
Email: ranjuskartha@gmail.com

Dr. Varghese Paul completed B.Sc (Engg) from Kerala University and M.Tech from Cochin University of Science and Technology. He pursed Ph.D in Computer Science from Cochin University of Science and Technology and currently working as Professor in Department of IT, RSET, Cochin. He is a Research Supervisor of Cochin University of Science and Technology, M G University Kottayam, Anna Technical University Chennai, Bharathiar University Coimbatore, Bharathidasan University Trichy and Karpagam University Coimbatore. Under the guidance, 29 research scholars had already completed research studies and degree awarded. His research areas are Data Security using Cryptography, Data Compression, Data Mining, Image Processing and E-Governance. Developed TDMRC Coding System for character representation in computer systems and encryption system using this unique coding system. He has published many research papers in international as well as national journals and a text book also. He has 10 years of teaching experience and 19 years of industrial experience.