Volume 2, No. 3, May-June 2011
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info
ISSN No. 0976-5697

A Modern Hill Cipher Involving a Pair of Keys, XOR operation and Substitution

Aruna Varanasi*
Department of computer Science and Engineering,SNIST
Hyderabad, India,
varanasi.aruna2002@gmail.com

V.U.K.Sastry
Department of computer Science and Engineering,SNIST
Hyderabad, India,
vuksastry@rediffmail.com

S.Udaya Kumar
Department of Computer Science and Engineering, SNIST
Hyderabad, India
uksusarla@rediffmail.com

Abstract: In this investigation, we have developed a block cipher. This includes a pair of keys for strengthening the cipher. In the development of the cipher, we have used iteration process, and a pair of functions called mix() and substitute() in each round of the iteration process. These functions modify the plaintext in various ways before it takes the shape of the ciphertext. The avalanche effect and the cryptanalysis examined in this analysis clearly indicate that this cipher is a strong one.

Keywords: symmetric block cipher, cryptanalysis, avalanche effect, ciphertext, pair of keys, XOR operation, mixing, substitution.

I. INTRODUCTION

In the literature of the cryptography we have seen, in the recent past, some variants of the classical Hill cipher, called modern Hill cipher [1-2]. In a recent investigation [3], we have developed a block cipher which involves a pair of keys and modular arithmetic addition. The basic equations governing this cipher are

$$C = (KP + L) \bmod N, \qquad (1.1)$$
and
$$P = (K^{-1}(C - L)) \bmod N, \qquad (1.2)$$

where N is any positive integer and K^{-1} is the modular arithmetic inverse of K.

Here, the presence of K and L, one on the left side of the P and another on right side of P, preceded by addition operation, strengthens the cipher significantly. This cipher is thoroughly supported by iteration, mixing and substitution processes.

In the present paper, our objective is to develop a new cipher, which is quite similar to the earlier one put forth in [3]. In the development of this cipher we use XOR operation instead of modular arithmetic addition used in the earlier paper. Thus the fundamental equations describing this cipher are

$$C = (KP \oplus L) \bmod N, \qquad (1.1)$$
and
$$P = (K^{-1}(C \oplus L)) \bmod N, \qquad (1.2)$$

In this also we use iteration, mixing and substitution. However, in the development of the substitution table, we have placed the elements of the keys K and L in the first two columns of the table instead of the first two rows of the table utilized in the earlier analysis. We shall discuss the details of the substitution table, a little later, in section 2.

Here it is to be noted that the XOR in the present analysis is expected to play a very prominent role in mixing the binary bits of the keys (K and L) and the plaintext P.

Now, we mention the outlines of the paper. Section 2 is devoted to the development of the cipher and the algorithms concerned to encryption and decryption. In section 3, we have presented an illustration of the cipher by giving a suitable example. Further we have discussed the avalanche effect. Then in section 4, we have examined the cryptanalysis. Finally in section 5, we have mentioned about the computations carried out in this analysis and arrived at the conclusions obtained from this investigation.

II. DEVELOPMENT OF THE CIPHER

Let us consider a plaintext, P. On applying EBCDIC code, P can be represented in the form of a matrix given by

$P = [P_{ij}]$, i= 1 to n , j=1 to n, (2.1)

Let us have a pair of keys K and L, which can be written in the form

$K = [K_{ij}]$, i=1 to n, j=1 to n, (2.2)

and

$L = [L_{ij}]$, i=1 to n, j=1 to n. (2.3)

Here all the elements of the matrices P, K and L are decimal numbers, which lie in the interval [0 ,255].

On using the process of encryption, we get the ciphertext C, which can be written in the form

$C = [C_{ij}]$, i=1 to n, j=1 to n, (2.4)

in which all the elements of C are also lying in [0,255].

The process of encryption and the process of decryption are presented in terms of the flow charts given in Fig.1.

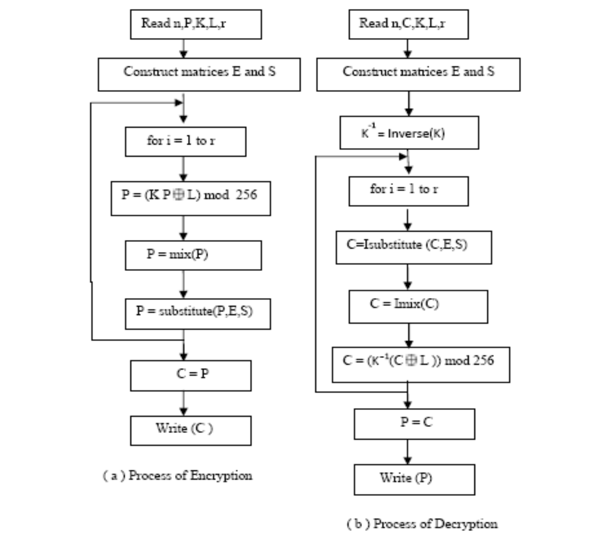

Fig.1 Schematic diagram of the Cipher

The algorithms for encryption and decryption are written below.

Algorithm for Encryption

1. Read n,P,K,L,r
2. for i = 1 to 16
   {
   for j = 1 to 16
   {
   E(i,j) =16(i-1)+(j-1)
   }
   }
3. S= Table(E,K,L)
4. for i = 1 to r
   {
   P = (K P ⊕ L) mod 256
   P = mix(P)
   P = substitute(P,E,S)
   }
   C = P
5. Write(C)

Algorithm for Decryption

1. Read n,C,K,L,r
2. for i = 1 to 16
   {
   for j = 1 to 16
   {
   E(i,j) =16(i-1)+(j-1)
   }
   }

3. S= Table(E,K,L)
4. K^{-1} = Inverse(K)

5. for i = 1 to r
   {

   C = Isubstitute(C,E,S)
   C = Imix(C)
   C = (K^{-1} (C ⊕ L))mod 256

   }
   P = C
6. Write (P)

Algorithm for inverse(K)

1. Read A, n, N
   // A is an n x n matrix. N is a positive integer with which modular arithmetic
   // is carried out. Here N= 256.
2. Find the determinant of A. Let it be denoted by Δ, where Δ ≠ 0.
3. Find the inverse of A. The inverse is given by $[A_{ji}]/\Delta$, i= 1 to n , j = 1 to n
   // $[A_{ij}]$ are the cofactors of a_{ij}, where a_{ij} are the elements of A
   for i = 1 to N
   {
   // Δ is relatively prime to N
   if((iΔ) mod N == 1) break;
   }
   d= i;
4. B = $[dA_{ji}]$ mod N. // B is the modular arithmetic inverse of A.

In the encryption algorithm, we have used the functions mix() and substitute().

In the function mix(), we adopt the following procedure. At each stage of the iteration process, the resulting plaintext matrix P, whose size is n^2, can be written in the form of a string of $8n^2$ binary bits, as each number can be represented in terms of 8 binary bits. This can be divided into four substrings, wherein each one is of size $2n^2$ binary bits. These strings can be written typically as shown below.

$$q_1 \quad q_2 \quad q_3 \quad q_4 \quad . \quad . \quad . \quad . \quad q_{2n^2},$$

$$r_1 \quad r_2 \quad r_3 \quad r_4 \quad . \quad . \quad . \quad . \quad r_{2n^2},$$

$$s_1 \quad s_2 \quad s_3 \quad s_4 \quad . \quad . \quad . \quad . \quad s_{2n^2},$$

$$t_1 \quad t_2 \quad t_3 \quad t_4 \quad . \quad . \quad . \quad . \quad t_{2n^2}.$$

The mixing is carried out by adopting the following arrangement:

$$q_1\, r_1\, s_1\, t_1\, q_2\, r_2\, s_2\, t_2\, q_3\, r_3\, s_3\, t_3\, q_4\, r_4\, s_4\, t_4 \ldots\ldots\ldots q_{2n^2}\, r_{2n^2}\, s_{2n^2}\, t_{2n^2}.$$

On decomposing this string into n^2 substrings and writing the binary bits in terms of decimal numbers, we get a square matrix of size n.

Let us now introduce the process of substitution. In the EBCDIC code, characters are represented by the numbers 0-255. These numbers can be written in the form of a matrix E given by

E (i , j) = 16(i-1)+(j-1), i=1 to 16 and j=1 to 16. (2.5)

In the development of the substitution table, having 16 rows and 16 columns, the first and second columns of the table are filled with the elements of the keys K and L (in order) respectively. The rest of the entries of the table are filled with the remaining elements of E (excluding the elements occurring in K and L) in a row wise manner in order. Thus we get the substitution table. This can be visualized as a matrix and it can be denoted by S(i,j), i=1 to 16, j= 1 to 16.

In order to have a clear insight into the substitution process, let us consider a plaintext. Let it be transformed (see encryption algorithm in section 2) by using the relations P = (KP $\oplus$ L) mod 256 and P = mix(P). Now the resulting plaintext contains a set of numbers. In the process of the substitution each number in the resulting plaintext is to be replaced by the corresponding number in the substitution matrix. If the number is E(i,j), it is to be replaced by S(i,j).

As it is seen in the algorithm, this substitution process is carried out in each round of the iteration process. For a detailed discussion of the substitution process, we may refer to (3).

It may be noted here that the function Imix() and Isubstitute(), in the process of decryption, are readily obtained by reversing the processes of mix() and substitute().

III. ILLUSTRATION OF THE CIPHER

Consider the plaintext mentioned below.

At present, though you have sleepless nights, work for your goal, your future will be with full of prosperity and happiness. We are bound to join very soon. (3.1)

Let us focus our attention on the first sixteen characters of the plaintext (3.1). This is given by

At present, thou (3.2)

On using the EBCDIC code, (3.2) can be written in the form

$$P = \begin{bmatrix} 193 & 163 & 64 & 151 \\ 153 & 133 & 162 & 133 \\ 149 & 163 & 107 & 64 \\ 163 & 136 & 150 & 164 \end{bmatrix} \quad (3.3)$$

Let us have the keys, K and L in the form

$$K = \begin{bmatrix} 123 & 25 & 9 & 67 \\ 134 & 17 & 20 & 11 \\ 48 & 199 & 209 & 75 \\ 39 & 55 & 85 & 92 \end{bmatrix} \quad (3.4)$$

and

$$L = \begin{bmatrix} 102 & 21 & 33 & 45 \\ 117 & 121 & 89 & 97 \\ 79 & 49 & 53 & 23 \\ 10 & 133 & 254 & 237 \end{bmatrix} \quad (3.5)$$

On using (2.5), (3.4), and (3.5), and applying the process mentioned in section 2, we get the following substitution table: This can be treated as matrix S(i,j), i=1 to 16, j=1 to 16.

On using (3.3) to (3.5), matrix S and the encryption algorithm (see section2) with r=16, we get the ciphertext C given by

$$C = \begin{bmatrix} 244 & 31 & 252 & 40 \\ 4 & 115 & 138 & 51 \\ 29 & 34 & 48 & 166 \\ 174 & 61 & 46 & 15 \end{bmatrix}. \quad (3.6)$$

On applying the decryption algorithm, we get back the original plaintext given by (3.3).

Let us now examine the avalanche effect, which shows the strength of the cipher in a qualitative manner.

To go ahead with the process, let us replace the fourth character ‘p’ of the plaintext (3.2) by ‘o’. The EBCDIC codes of ‘p’ and ‘o’ are 151 and 150. Readily we notice that these two numbers differ by one bit in their binary form. On using the modified plaintext, the keys K and L given by (3.4) and (3.5), the substitution matrix S, and the encryption algorithm, we get the ciphertext C in the form

$$C = \begin{bmatrix} 233 & 192 & 223 & 84 \\ 80 & 182 & 249 & 38 \\ 2 & 213 & 74 & 200 \\ 228 & 141 & 53 & 209 \end{bmatrix} \quad (3.7)$$

On converting (3.6) and (3.7) into their binary form, we find that the two ciphertexts differ by 72 bits (out of 128 bits). This shows that the cipher is expected to be a strong one.

Consider a one bit change in one of the keys, say key, L. To this end, we have replaced the first row third column element “33” of (3.5), by “ 32 ”. On performing the encryption with the modified key L, the corresponding substitution matrix S, and with the original plaintext (3.3), keeping the other key K intact, we get the ciphertext given by

.

$$C = \begin{bmatrix} 28 & 102 & 92 & 33 \\ 237 & 139 & 151 & 250 \\ 255 & 29 & 85 & 79 \\ 95 & 226 & 14 & 52 \end{bmatrix}. \quad (3.8)$$

Let us now compare the binary strings corresponding to (3.6) and (3.8). From this we find that the two ciphertexts differ by 68 bits (out of 128 bits). This also shows that the cipher is a potential one.

123	102	0	1	2	3	4	5	6	7	8	12	13	14	15	16
25	21	18	19	22	24	26	27	28	29	30	31	32	34	35	36
9	33	37	38	40	41	42	43	44	46	47	50	51	52	54	56
67	45	57	58	59	60	61	62	63	64	65	66	68	69	70	71
134	117	72	73	74	76	77	78	80	81	82	83	84	86	87	88
17	121	90	91	93	94	95	96	98	99	100	101	103	104	105	106
20	89	107	108	109	110	111	112	113	114	115	116	118	119	120	122
11	97	124	125	126	127	128	129	130	131	132	135	136	137	138	139
48	79	140	141	142	143	144	145	146	147	148	149	150	151	152	153
199	49	154	155	156	157	158	159	160	161	162	163	164	165	166	167
209	53	168	169	170	171	172	173	174	175	176	177	178	179	180	181
75	23	182	183	184	185	186	187	188	189	190	191	192	193	194	195
39	10	196	197	198	200	201	202	203	204	205	206	207	208	210	211
55	133	212	213	214	215	216	217	218	219	220	221	222	223	224	225
85	254	226	227	228	229	230	231	232	233	234	235	236	238	239	240
92	237	241	242	243	244	245	246	247	248	249	250	251	252	253	255

Table 1: Substitution Table.

IV. CRYPTANALYSIS

In the literature of cryptography, the general analytical methods for breaking the cipher, if possible, are

1. Ciphertext only attack (Brute force attack)
2. Known plaintext attack
3) Chosen plaintext attack and
4) Chosen ciphertext attack

In this cipher, as there are two keys K and L, where in each key is containing 16 numbers, the total length of the pair of keys is seen to be 256 binary bits. Thus the size of the key space is

$2^{256} = (2^{10})^{25.6} \approx (10^3)^{25.6} = 10^{76.8}$.

If the time required for breaking the cipher with one value of the key in the key space is taken as 10^{-7} seconds, then the time required for examining the breakability of the cipher with all possible values of the keys in the key space is

$$\frac{10^{76.8} \times 10^{-7}}{365 \times 24 \times 60 \times 60} = 31.71 \times 10^{60.8}\ years$$

As this number is very large, it is impossible to break the cipher by brute force attack.

244	31	252	40	4	115	138	51	29	34	48	166	174	61	46	15
14	14	178	132	130	56	59	91	37	249	218	192	107	244	224	174
205	180	174	1	70	181	51	181	17	1	84	163	185	129	105	124
92	229	176	242	188	81	109	154	222	29	215	221	30	89	98	81
91	206	13	52	218	93	88	72	88	208	114	140	223	61	210	35
196	201	141	210	139	87	132	177	55	46	166	139	100	26	79	224
165	3	96	45	4	201	150	124	30	73	13	152	109	162	67	11
233	111	56	108	113	178	96	230	54	224	88	31	82	200	72	233
61	103	5	50	66	3	204	99	160	191	19	64	11	70	26	144
199	133	37	93	166	60	186	117	12	184	105	38	249	197	239	194

Let us now consider the known plaintext attack. In this we know as many plaintext and ciphertext pairs as we require. On carrying out the encryption process which includes sixteen rounds of the iteration, the ciphertext C can be taken to be in the form

$$C = \Psi\, (M((K\Psi\, (M((\ldots\ldots\Psi\, (M((K\, \Psi\, (M((KP \oplus L) \bmod 256)) \oplus L\,) \bmod 256))\, \ldots\ldots \oplus L) \bmod 256)) \oplus L) \bmod 256)) \quad (4.1)$$

In writing (4.1), the functions mix() and substitute() are replaced by M() and Ψ () respectively. This replacement is done for the sake of elegance. Here we notice that (4.1) cannot be written in the form

$$C = F(K,L,M, \Psi)\, P$$

where F is a function, depending upon K,L,M and Ψ.

Thus, as (4.1) is a complicated relation, we cannot determine P or a function of P in terms of the other quantities. Hence, unlike in the case of classical Hill cipher, this cipher cannot be broken by the known plaintext attack.

Though it is worth examining the cryptanalysis in the case of the last two attacks (attacks 3 and 4), we restrain ourselves without further examination, as the cryptanalysis in these two cases is expected to be quite cumbersome [4-5].

In view of the above analysis, we conclude that this cipher cannot be broken by any easy means, and it is quite dependable.

V. COMPUTATIONS AND CONCLUSIONS

In this paper, we have developed a block cipher which involves iteration process, a pair of keys and XOR operation in each round of the iteration process. This cipher includes a pair of functions called mix () and substitute (), for achieving diffusion and confusion. The computations in this analysis are performed by writing programs for encryption and decryption in Java. The ciphertext corresponding to the entire plaintext given by (3.1) is obtained in the form

In obtaining the ciphertext we have divided the plaintext (3.1) into 10 blocks. As the last block is having 12 characters only, it is appended with 4 blank characters to make it a complete one.

From the avalanche effect and the cryptanalysis carried out in this investigation, we conclude that the cipher is fairly a strong one, and it can be used comfortably for the security of information.

VI. REFERENCES

[1] V.U.K.Sastry, Aruna Varanasi. S. Udaya Kumar, "A Modern Hill cipher Involving Permuted Key and Modular Arithmetic Addition Operation", International Journal of Advanced Research in Computer Science (IJARCS), Vol.2, No.1, pp.162-165, Jan-Feb 2011.

[2] V.U.K.Sastry, Aruna Varanasi, S.Udaya Kumar, "A Modern Hill Cipher Involving XOR operation and a Permuted Key", International journal of Advanced Research in Computer Science (IJARCS), Vol.2, No.1, pp.153-155, Jan-Feb 2011.

[3] Aruna Varanasi, V.U.K.Sastry, S.Udaya Kumar, A Modern Hill Cipher Involving a Pair of Keys, Modular Arithmetic Addition and Substitution, International journal of Advanced Research in Computer Science (IJARCS), sent for publication

[4] Biham, E., and Shamir, A. Differential Cryptanalysis of the Data Encryption Standard. New York: Springer-Verlag, 1993.

[5] Matsui, M. " Linear Cryptanalysis Method for DES Cipher", Proceedings, EUROCRYPT'93,1993: New York- Springer-Verlag.

