Volume 2, No. 5, May 2011

Journal of Global Research in Computer Science

ISSN-2229-371X

RESEARCH PAPER

Available Online at www.jgrcs.info

A MODERN ADVANCED HILL CIPHER INVOLVING A PAIR OF KEYS, XOR OPERATION AND SUBSTITUTION

Aruna Varanasi *1, V.U.K.Sastry 2 and S.Udaya Kumar3

1Department of computer Science and Engineering,SNIST
Hyderabad, India,
varanasi.aruna2002@gmail.com
*2Department of computer Science and Engineering,SNIST
Hyderabad, India,
vuksastry@rediffmail.com
3Department of computer Science and Engineering,SNIST
Hyderabad, India,
uksusarla@rediffmail.com

Abstract: In this paper, we have developed a block cipher, which involves a pair of keys, XOR operation, mixing and substitution. All these additional features are expected to strengthen the cipher as the plaintext undergoes several transformations which are causing confusion and diffusion. From the avalanche effect and the cryptanalysis carried out in this investigation, we have noticed that this cipher is a strong one, and it can be utilized effectively for the transmission of information in a secured manner.

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

INTRODUCTION

In a recent investigation [1], we have devoted our attention to the study of a modern advanced Hill cipher involving a pair of keys. In this, we have introduced modular arithmetic addition operation, mixing and substitution in each round of the iteration process. The basic equations governing this cipher are

$$C = (AP + B) \bmod N, \quad (1.1)$$

and

$$P = (A(C - B)) \bmod N, \quad (1.2)$$

where P is a plaintext matrix, A and B are square matrices of size n, N a positive integer, chosen appropriately, and C is the corresponding ciphertext matrix. In this analysis, matrices A and B are involutory matrices, which include the pair of keys K and L respectively.

Here it is to be noted that an involutory matrix is a matrix whose arithmetic inverse is the same as the matrix itself. The equations that are required for obtaining A are given by

$$A^{-1} = A, \quad (1.3)$$

$$(A A^{-1}) \bmod N = I, \quad (1.4)$$

$$A^2 \bmod N = I, \quad (1.5)$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad (1.6)$$

$$A_{11} = K, \quad (1.7)$$

$$A_{22} = -K, \quad (1.8)$$

$$A_{12} = [d(I - K)] \bmod N, \quad (1.9)$$

$$A_{21} = [\lambda(I + K)] \bmod N, \quad (1.10)$$

$$\text{where } (d\lambda) \bmod N = 1, \quad (1.11)$$

where A^{-1} is the arithmetic inverse of A, I the identity matrix, d a chosen positive integer and λ is determined from (1.11). Similar equations can be obtained for obtaining B (see [1]).

In order to have a detailed discussion concerned to the relations for obtaining an involutory matrix, we refer to [2].

In the present paper our objective is to develop a variant of the modern advanced Hill cipher, discussed in [1], by replacing the addition operation with XOR operation. The relations governing the block cipher that we are going to develop in this analysis are

$$C = (AP \oplus B) \bmod N, \quad (1.12)$$

and

$$P = (A(C \oplus B)) \bmod N. \quad (1.13)$$

In this analysis also we have included the iteration process, the functions mix() and substitute() in each round of the iteration. All these features together with the XOR operation are expected to strengthen the cipher significantly.

Let us now put forth the plan of the paper. In section 2, we have introduced the development of the cipher, and presented the flowcharts and algorithms for encryption and decryption. In section 3, we have illustrated the cipher and mentioned the avalanche effect. Section 4 is devoted to cryptanalysis. Finally in section 5, we have discussed the computations and drawn conclusions.

DEVELOPMENT OF THE CIPHER

In the development of this cipher, the plaintext P, the pair of keys K and L (basing upon which the involutory matrices A and B are found), and the ciphertext C are given by the relations

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

$K = [K_{ij}]$, i=1 to n/2, j=1 to n/2, (2.2)

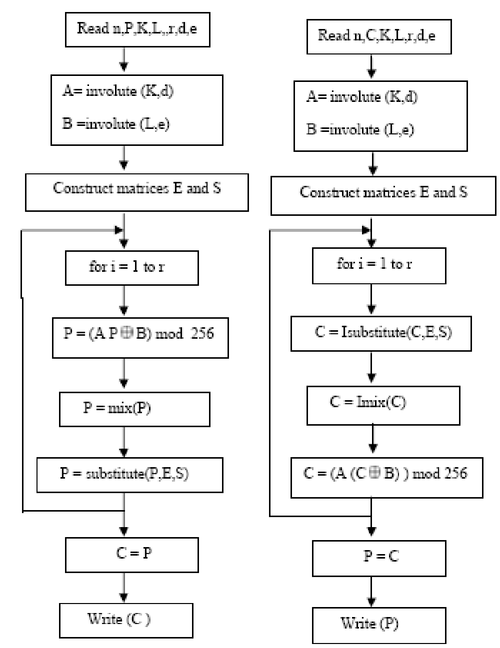
$L = [L_{ij}]$, i=1 to n/2, j=1 to n/2, (2.3)

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

Here n is an even positive integer and each element of P, K, L and C are decimal numbers, lying in [0, 255], as we have made use of EBCDIC code.

On using the keys K and L, and taking N=256, the involutory matrices A and B can readily be found by using the relations, mentioned in section 1 (see (1.3) to (1.11)).

As we have already pointed out in section 1, the relations governing the encryption and the decryption are (1.12) and (1.13). In what follows, we present the flowcharts and the algorithms.



(a) Process of Encryption   (b) Process of Decryption

Fig.1 Schematic diagram of the Cipher

Algorithm for Encryption

1. Read n,P,K,L,r,d,e
2. A = involute(K,d)
   B = involute(L,e)
3. Construct matrices E, S
4. for i = 1 to r
   {
   P = (A P $\oplus$ B) 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,d,e
2. A = involute(K,d)
   B = involute (L,e)
3. Construct matrices E,S
4. for i= 1 to r
   {
   C = Isubstitute(C,E,S)
   C = Imix(C)
   C= (A (C $\oplus$ B))mod 256
   }
   P = C
5. Write (P)

In this analysis, we have denoted the number of rounds as r, and it is taken as 16. The d and e are positive integers which are chosen in finding the involutory matrices A and B. The function involute() is used for obtaining the involutory matrix.

The functions mix() and substitute() used in the encryption algorithm can be mentioned as follows:

At each stage of the iteration process, the matrix P is of size nxn. It can be written in the form of four binary strings, wherein each string has $2n^2$ binary bits as shown below:

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

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

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

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

On mixing these strings, we get a single string given by

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

On taking 8 bits at a time, the above string, containing $8n^2$ binary bits can be written in the form of a square matrix of size n.

Let us now develop the process of substitution. We know that the EBCDIC code, requires the numbers 0-255 for the representation of the characters. 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)

Let us now see the development of the substitution table consisting of 16 rows and 16 columns. In order to achieve this one, let us firstly fill up the first two columns of the table with the elements of the keys K and L in order. Then rest of the table is filled with the remaining elements of E, in order in a row wise manner, excluding the numbers contained in K and L. This process yields the substitution table. This table can be represented in the form of a substitution matrix denoted by S(i,j).

For a detailed discussion of the process of substitution, we refer to [1].

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

ILLUSTRATION OF THE CIPHER

Consider the plaintext given below:

"Hello X! I am waiting for your email. I have already completed my B. Tech. examinations very well. My father is compelling me to do IAS, and to become a collector in this country. It is unfortunate! When are you completing your PhD program? I would like to come to you and finish there my MS. What about our marriage? I am waiting for your reply." (3.1)

Let us now consider the first sixty four characters of the plaintext given by (3.1). Thus we have

"Hello X! I am waiting for your email. I have already completed m" (3.2)

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

$$P = \begin{bmatrix} 200 & 133 & 147 & 147 & 150 & 64 & 231 & 79 \\ 64 & 201 & 64 & 129 & 148 & 64 & 166 & 129 \\ 137 & 163 & 137 & 149 & 135 & 64 & 134 & 150 \\ 153 & 64 & 168 & 150 & 164 & 153 & 64 & 133 \\ 148 & 129 & 137 & 147 & 75 & 64 & 201 & 64 \\ 136 & 129 & 165 & 133 & 64 & 129 & 147 & 153 \\ 133 & 129 & 132 & 168 & 64 & 131 & 150 & 148 \\ 151 & 147 & 133 & 163 & 133 & 132 & 64 & 148 \end{bmatrix} \quad (3.3)$$

Let us choose the keys K and L in the form

$$K = \begin{bmatrix} 69 & 124 & 27 & 167 \\ 135 & 79 & 99 & 111 \\ 248 & 199 & 209 & 75 \\ 239 & 45 & 255 & 92 \end{bmatrix}. \quad (3.4)$$

and

$$L = \begin{bmatrix} 215 & 113 & 19 & 147 \\ 223 & 109 & 254 & 12 \\ 56 & 1 & 127 & 174 \\ 59 & 146 & 189 & 81 \end{bmatrix}. \quad (3.5)$$

Let us now construct the involutory matrices A and B by adopting the process mentioned in section 1 (see (1.3) - (1.11). In obtaining A and B, we have taken d= 99 and e=189 respectively. Thus we get

$$A = \begin{bmatrix} 69 & 124 & 27 & 167 & 180 & 12 & 143 & 107 \\ 135 & 79 & 99 & 111 & 203 & 214 & 183 & 19 \\ 248 & 199 & 209 & 75 & 24 & 11 & 144 & 255 \\ 239 & 45 & 255 & 92 & 147 & 153 & 99 & 207 \\ 130 & 84 & 233 & 237 & 187 & 132 & 229 & 89 \\ 141 & 112 & 1 & 133 & 121 & 177 & 157 & 145 \\ 168 & 77 & 134 & 249 & 8 & 57 & 47 & 181 \\ 5 & 47 & 181 & 63 & 17 & 211 & 1 & 164 \end{bmatrix} \quad (3.6)$$

and

$$B = \begin{bmatrix} 215 & 113 & 19 & 147 & 2 & 147 & 249 & 121 \\ 223 & 109 & 254 & 12 & 93 & 68 & 122 & 36 \\ 56 & 1 & 127 & 174 & 168 & 67 & 250 & 138 \\ 59 & 146 & 189 & 81 & 113 & 54 & 119 & 240 \\ 184 & 197 & 15 & 143 & 41 & 143 & 237 & 109 \\ 203 & 6 & 214 & 252 & 33 & 147 & 2 & 244 \\ 152 & 149 & 128 & 70 & 200 & 255 & 129 & 82 \\ 87 & 250 & 1 & 186 & 197 & 110 & 67 & 175 \end{bmatrix} \quad (3.7)$$

As we have mentioned in section 2, the substitution matrix S can be written in terms of Table 1.

On using (3.3), (3.4), and (3.5), and the encryption algorithm (which uses the substitution process), we get

$$C = \begin{bmatrix} 9 & 204 & 21 & 245 & 209 & 19 & 10 & 192 \\ 202 & 15 & 30 & 64 & 115 & 112 & 75 & 180 \\ 128 & 157 & 223 & 223 & 114 & 195 & 241 & 185 \\ 152 & 12 & 38 & 108 & 70 & 94 & 145 & 233 \\ 208 & 153 & 64 & 199 & 251 & 56 & 53 & 27 \\ 40 & 143 & 184 & 154 & 226 & 19 & 152 & 41 \\ 84 & 198 & 231 & 32 & 157 & 102 & 102 & 137 \\ 126 & 144 & 115 & 68 & 74 & 90 & 176 & 70 \end{bmatrix} \quad (3.8)$$

On adopting the decryption algorithm, with the required inputs, it can be readily verified that we get back the original plaintext given by (3.3).

Let us now examine the avalanche effect, which gives an idea about the quality of the cipher.

To this end, in the plaintext (3.2), we replace the 18^{th} character 't' by 's',. As the EBCDIC codes of 't ' and 's ' are 163 and 162, they differ by one bit in their binary form. Now, on using the modified plaintext along with (3.4) and (3.5), and applying the encryption algorithm, we have the ciphertext C in the form

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

Table 1: Substitution Table.

$$C = \begin{bmatrix} 9 & 18 & 203 & 50 & 195 & 232 & 67 & 144 \\ 235 & 242 & 235 & 148 & 117 & 141 & 72 & 219 \\ 93 & 167 & 93 & 158 & 76 & 0 & 180 & 57 \\ 233 & 17 & 69 & 191 & 173 & 63 & 163 & 3 \\ 189 & 175 & 189 & 104 & 206 & 7 & 152 & 51 \\ 39 & 180 & 135 & 138 & 134 & 76 & 26 & 34 \\ 162 & 111 & 132 & 214 & 230 & 145 & 199 & 255 \\ 223 & 131 & 138 & 119 & 187 & 131 & 89 & 94 \end{bmatrix} \qquad (3.9)$$

On converting (3.8) and (3.9), in to their binary form, and comparing the corresponding strings, we notice that the two ciphertexts differ by 271 bits (out of 512 bits). This shows that the strength of the cipher is expected to be up to the mark.

Let us now focus our attention on one bit change in one of the keys, say key K. To achieve this one we change the 2nd row 1st column element of the key K, given by (3.4), from 135 to 134. On using the original plaintext (3.3), the modified key K, keeping the other key L intact, and using the encryption algorithm, we get

$$C = \begin{bmatrix} 64 & 149 & 123 & 14 & 122 & 220 & 136 & 86 \\ 92 & 48 & 64 & 65 & 52 & 38 & 97 & 124 \\ 122 & 42 & 193 & 20 & 165 & 158 & 150 & 241 \\ 186 & 233 & 224 & 199 & 72 & 98 & 53 & 149 \\ 34 & 213 & 182 & 72 & 31 & 70 & 219 & 126 \\ 111 & 34 & 134 & 47 & 50 & 155 & 137 & 225 \\ 1 & 188 & 232 & 137 & 83 & 28 & 134 & 214 \\ 25 & 202 & 28 & 121 & 209 & 17 & 222 & 234 \end{bmatrix} \qquad (3.10)$$

Now on comparing (3.8) and (3.10) in their binary form, we find that they differ by 278 bits (out of 512 bits). This also shows that the strength of the cipher is considerable.
In what follows, let us now consider the cryptanalysis, which exhibits more firmly about the strength of the cipher.

CRYPTANALYSIS

The different types of cryptanalytic attacks which are generally considered in the literature of Cryptography are

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

The key matrices K and L, involved in this analysis, contain 16 decimal numbers each. The constants d and e, which are chosen at our will in the construction of the involutory matrices A and B, are two more decimal numbers. In view of these facts, the total length of the keys is 34 decimal numbers, that is 272 binary bits. Hence the size of the key space is

$$2^{272} = (2^{10})^{27.2} \approx (10^3)^{27.2} = 10^{81.6}.$$

If the time required for obtaining the plaintext with one value of the key in the key space is 10^{-7} seconds, then the time required for the execution of the cipher with all the possible keys in the key space is

$$\frac{10^{81.6} \times 10^{-7}}{365 \times 24 \times 60 \times 60} = 3.171 \times 10^{66.6}\,years$$

As this number is very large, we can firmly say that this cipher cannot be broken by the brute force attack.

Let us now consider the known plaintext attack. In this we know as many pairs of the plaintext and the ciphertext as we desire. In the development of this cipher as we have an iterative process, which involves a pair of keys, functions mix() and substitute(), and XOR operation, at the end of the iteration process, the relation between the plaintext and the ciphertext can be viewed as shown below

$$C = \Psi\,(M((A\Psi\,(M((\ldots\ldots\ \Psi\,(M((A\,\Psi\,(M((AP \oplus B) \bmod 256)) \oplus B\,) \bmod 256))\ \ldots\ldots) \bmod 256)) \oplus B) \bmod 256)) \tag{4.1}$$

In writing (4.1), the function mix() and the function substitute() are represented as M and Ψ for simplicity and elegance. Here we notice that the equation (4.1) cannot be written in the form

$$C = F(A,B,M,\Psi)\,P \tag{4.2}$$

where F is a function, depending upon K,L,M and Ψ. This amounts to that we cannot find a direct relation between C and P as we could do in the case of the classical Hill cipher. Thus this cipher cannot be broken by the known plaintext attack.

The last two cases of cryptanalysis, namely chosen plaintext attack and chosen ciphertext attack are very complicated, and hence we leave them at the present stage.

In the light of the above discussion we conclude that this cipher is a strong one.

COMPUTATIONS AND CONCLUSIONS

In this investigation, we have developed a block cipher, called modern advanced Hill cipher, which includes a pair of keys, XOR operation and functions mix() and substitute(). In this cipher the computations are carried out by writing programs for encryption and decryption in Java.

The plaintext (3.1) is divided into 6 blocks, wherein each block is containing 64 characters. Nevertheless, as the last block is containing only 26 characters, it is supplemented with 38 blank characters so that it becomes a complete block. On using the encryption algorithm the ciphertext corresponding to the entire plaintext (3.1) is obtained in the form

9	204	21	245	209	19	10	192	202	15	30	64	115	112	75	180
128	157	223	223	114	195	241	185	152	12	38	108	70	94	145	233
208	153	64	199	251	56	53	27	40	143	184	154	226	19	152	41
84	198	231	32	157	102	102	137	126	144	115	68	74	90	176	70
130	44	62	203	105	71	89	28	77	162	107	5	69	166	138	152
213	215	97	20	61	189	128	4	11	231	55	114	134	67	204	252
195	179	217	112	114	95	8	38	122	41	245	53	80	18	105	28
222	239	81	51	11	110	88	0	207	170	206	189	65	243	248	146
44	29	213	27	171	154	244	212	103	160	86	88	243	243	132	68
139	152	24	63	49	47	29	101	184	160	159	118	25	111	107	135
219	108	148	89	253	130	4	53	13	148	65	243	104	26	27	177
165	105	79	87	216	146	34	97	144	9	111	119	22	71	87	35
68	227	105	191	188	44	106	237	191	26	180	191	188	11	58	196
9	127	213	222	118	65	84	61	186	61	175	45	24	119	238	50
175	27	111	73	75	89	14	126	33	218	96	142	145	137	154	174
199	213	152	47	197	236	194	94	133	220	67	21	71	227	246	77
39	102	178	249	227	56	102	160	97	199	58	188	153	37	131	31
106	133	137	44	137	134	92	202	227	175	160	173	120	107	64	70
232	122	71	211	88	104	101	205	45	52	191	32	209	107	17	79
232	245	166	167	83	214	76	104	179	171	247	167	30	90	223	87
77	13	179	143	153	221	59	80	29	212	5	40	93	198	138	254
190	131	4	39	27	112	224	147	180	202	238	200	212	207	53	163
117	154	41	245	215	129	160	99	128	230	60	221	1	77	3	33
155	99	125	199	64	5	230	44	31	215	225	181	184	173	39	59

The avalanche effect and the cryptanalysis, considered in sections 3 and 4, clearly indicate that the cipher is a strong one and it cannot be broken by any cryptanalytic attack. This generalization of the advanced Hill cipher is markedly an interesting one, and it can be applied comfortably for the transmission of information in a secured manner.

REFERENCES

[1] Aruna Varanasi, V.U.K.Sastry and S.Udaya Kumar, " A modern Advanced Hill cipher Involving a Pair of Keys, Modular Arithmetic Addition and Substitution", sent for publication.

[2] V.U.K.Sastry, Aruna Varanasi, S.Udaya Kumar, "Advanced Hill Cipher Involving Permutation and Iteration", International Journal of Advanced Research in Computer Science, Vol.1, No.4, pp. 141-145, Nov-Dec. 2010.

Aruna Varanasi is presently working as Associate Professor in the Department of Computer Science and Engineering (CSE), Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India. She was awarded "Suman Sharma" by Institute of Engineers (India), Calcutta for securing highest marks among women in India in AMIE course.

Dr. V. U. K. Sastry is presently working as Professor in the Dept. of Computer Science and Engineering (CSE), Director (SCSI), Dean (R & D), SreeNidhi Institute of Science and Technology (SNIST), Hyderabad, India. He was Formerly Professor in IIT, Kharagpur, India and worked in IIT, Kharagpur during 1963 – 1998. He guided 12 PhDs, and published more than 40 research papers in various international journals. His research interests are Network Security & Cryptography, Image Processing, Data Mining and Genetic Algorithms.

