

Model Estimation and Error Level for Okike's Merged Irregular Transposition

Cipher

Abstract

The researcher has developed a new encryption technique known as Merged Irregular Transposition Cipher. In this cipher method of encryption, a message to be encrypted is split into parts and each part encrypted separately. Before the encrypted message is transmitted to the recipient(s), the positions of the split in the encrypted messages could be swapped to ensure more security. This work seeks to develop a model by considering the split number, S and the average number of characters per split, L as the message under consideration is split from 2 through 10. Again, after developing the model, the error level in the model would be determined.

Keywords: Merged Irregular Transposition, Error Level, Model Estimation, Message

Splitting

1.0 Introduction

From time immemorial, the hiding of the real meaning of information has always been there. Julius Caesar's cipher gave him advantage over his opponents during his days. Caesar's cipher then employed simple substitution method to conceal the actual meaning of information to be exchanged.

Transposition cipher on the other hand rearranges the characters contained the messages, thereby hiding the real meaning of the message. Transposition may be Single Columnar, Double Columnar, Dynamic or Irregular Ciphers. The new transposition cipher developed by the researcher is called Merged Irregular Transposition Cipher which makes it possible for the message to be split into parts before the encryption process takes place. In this work, a sample message is deployed and split from 2 through 10 so that a model is estimated. As the message is split from 2 through 10, the average number of characters per split is taken. This is used to determine the error level in the estimated model.

2.0 Substitution Cipher

Substitution ciphers are one of the earliest types of ciphers. Examples of classic substitution ciphers include the well-known simple substitution and the less well-known homophonic substitution. Although simple substitution ciphers are indeed simple – both in terms of their use and attacks; the homophonic substitution ciphers are far more challenging to break. Even with modern computing technology, homophonic substitution ciphers remain a significant challenge.

The simple substitution cipher is a cipher that has been in use for many hundreds of years. It basically consists of substituting every plaintext character for a different ciphertext character. According to Christensen (2006), a simple substitution cipher is a method of concealment that replaces each letter of a plaintext message with another letter (Dhavare, 2011).

3.0 Transposition Cipher

Collberg (2012) defined transposition cipher as a method of encryption by which units of plaintext are rearranged to form the ciphertext. In a transposition cipher the original characters of the plaintext are not changed, but simply moved around in the ciphertext.

With transposition cipher, letters orders are rearranged without altering the actual letters. Row Transposition Ciphers allow letters in rows to be written out, reorder the columns according to the key before reading off (Jain, 2011).

4.0 Merged Irregular Transposition Cipher

A geometrical transposition, otherwise called the irregular transposition cipher is one which employs the use of keyword and has varied number of letters in its columns or rows. The number of letters in a given row or column is determined by the column or row number and the corresponding ordinal value

To improve on the Irregular Transposition Cipher, the researcher intends to develop a new type of cipher to be known as Merged Irregular Transposition Cipher. Unlike the Irregular Transposition Cipher already in existence, the Merged Irregular Transposition Cipher will utilize multiple tables and keywords to encrypt information. The first step toward the application of this model will be to divide the entire message into multiple equal or nearly equal parts. The number of parts may be determined by the length of the message to be encrypted. In this research work, the message to be encrypted will only be divided into ten parts for the sake of illustration and analysis carried on the behaviour of each of the part in relation to the overall behaviour of the cipher.

4.1 Structure of Merged Irregular Cipher

The structure of the table is depicted in table 4.1 below:

Col	1	2	3	-----			m
K/W	k	e	y	w	o	r	d
O/V	3	2	7	6	4	5	1
Row							
1							
2							
3							
:							
:							
:							
n							

Table 4.1 : Keyword written against Column

From the tables above, there are n rows and m columns. In table 4.1, the keyword is written against the columns. There would not be much difference if the keywords are written against the rows

To illustrate how the Merged Irregular Transposition Cipher works, assuming the information below originates from Anambra State Government House, Awka few days before the political impasse that engulfed the state took place:

ATTENTION: POLICE BOSS – HOODLUMS PLAN MAYHEM ON ANAMBRA RESIDENTS

To encipher the above information using the Merged Irregular Transposition Cipher, the following steps are involved:

1. Choose keywords to use, The number of keywords should depend on the length of the message.
2. Split the original message into multiple equal or nearly equal parts.
3. Encrypt each part of the split message using any of the keyword.
4. Combine the multiple encrypted messages into a single message.

Before encrypting the above message, there may be a need to define the variables to be used.

4.2 Definition of Variables

- i. Column, m is the number of columns contained in a table (matrix).
- ii. Row, n is the number of rows in a table.
- iii. Length of message, L is the total number of characters in the message to be encrypted for each split number, S .
- iv. Message Split number, S is the number of parts the entire message is split into.
- v. Message Characters, C is the individual characters in a message.
- vi. Available spaces, A refers to the number of spaces in a table that may contain any message character, C .
- vii. Empty character, X is used to fill up empty spaces when all the message characters, C have been entered and yet the ordinal value corresponding to the row has not been reached.
- viii. Empty Space, B is the number of space(s) in the table that either contains no message characters, C or empty character, X .
- ix. Non-empty space, Q is the number of space(s) that either contains a message character, C or empty character, X in a table.
- x. Swap Number, Z is the total number of positions that an encrypted message can be exchanged for each split, S .

At this point, an example may be used to illustrate how the Merged Irregular Transposition Cipher may be applied. The entire message, L to be encrypted may be split into a given number of parts, S , where $S \geq 2$.

$$S=2$$

Using UNIVERSE as a keyword and encrypting the first part of the message “ATTENTION: POLICE BOSS – HOODLUMS”

130

131

132	Col	1	2	3	4	5	6	7	8
133	K/W	U	N	I	V	E	R	S	E
134	O/V	7	4	3	8	1	5	6	2
135	Row								
136	1	A	T	T	E	N			
137	2	T	I	O	N	:	P	O	L
138	3	I	C	E					
139	4	B	O						
140	5	S	S	-	H	O	O		
141	6	D	L	U	M	S	X	X	

142

143

144 This will produce the ciphertext below:

145

146 N:OS L TOE-L TICOSL POX OX ATIBSD ENHM

147 Using SCIENCES as a keyword and encrypting the second part of the message” PLAN

148 MAYHEM ON ANAMBRA RESIDENTS.”

149	Col	1	2	3	4	5	6	7	8
150	K/W	S	C	I	E	N	C	E	S
151	O/V	7	1	5	3	6	2	4	8
152	Row								
153	1	P	L						
154	2	A	N	M	A	Y	H		
155	3	E	M	O	N				
156	4	A	N	A	M	B	R	A	
157	5	R	E	S					
158	6	I	D	E	N	T			
159	7	S							
160	8	.	X	X	X	X	X	X	X

161

162 This will produce the ciphertext below:

163

164 LNMNEDX HRX ANMNX AX MOASEX YBTX

165 PEAARIS. X

166

167 which will produce the combined ciphertext below:

168

169 N:OS L TOE-L TICOSL POX OX ATIBSD ENHM

170 LNMNEDX HRX ANMNX AX MOASEX YBTX

171 PEAARIS. X

172

173 At this point, the positions of the ciphertext information may be swapped, Since S=2, the
 174 first could take the second position and the second take the first position thereby
 175 increasing the complexity of the encryption algorithm.

176

177

178

S= 3

179

180 ATTENTION: POLICE BOS

19

181 S - HOODLUMS PLAN MAYHE

19

182 M ON ANAMBRA RESIDENTS.

20

183

184 Using the keyword FORSEE, the first part of the message “ATTENTION: POLICE
185 BOS”

186

187	Col	1	2	3	4	5	6
188	K/W	F	O	R	S	E	E
189	O/V	3	4	5	6	1	2
190	Row						
191	1	A	T	T	E	N	
192	2	T	I	O	N	:	P
193	3	O					
194	4	L	I				
195	5	C	E	B			
196	6	O	S	X	X		

197

198 This will yield the ciphertext below:

199

200 N: P ATOLCO THIES TOBX ENX

201 Using the keyword VISION to encrypt the second part of the message “S - HOODLUMS
202 PLAN MAYHE”

203

204	Col	1	2	3	4	5	6
205	K/W	V	I	S	I	O	N
206	O/V	6	1	5	2	4	3
207	Row						
208	1	S	–				
209	2	H	O	O	D		
210	3	L	U	M	S	P	L
211	4	A	N	M	A	Y	
212	5	H	E	X			

213

214 This will produce the ciphertext message below:

215

216 -OUNE DSA L PY OMMX SHLAH .

217

218

219 The keyword TONGUE when applied to encrypt the third part of the message “M ON
220 ANAMBRA RESIDENTS.”

221

222	Col	1	2	3	4	5	6
223	K/W	T	O	N	G	U	E
224	O/V	5	4	3	2	6	1
225	Row						
226	1	M	O	N	A	N	A
227	2	M	B	R	A		
228	3	R	E	S			

229 4 I D
 230 5 E
 231 6 N T S . X X

232

233 This produces the ciphertext message shown below:

234

235 AX AA NRSS OBEDT MMRIEN NX

236

237 which will produce the combined ciphertext below:

238

239 N: P ATOLCO TIIES TOBX ENX -OUNE DSA L PY

240 OMMX SHLAH . AX AA NRSS OBEDT MMRIEN NX

241

242 S= 4

243

244 ATTENTION: POLI 14

245 CE BOSS – HOODLUM 14

246 S PLAN MAYHEM ON AN 15

247 AMBRA RESIDENTS. 15

248

249 The first part of this message “ATTENTION: POLI” may be encrypted using the
 250 keyword BREAD as below:

251

252

253

254	Col	1	2	3	4	5
255	K/W	B	R	E	A	D
256	O/V	2	5	4	1	3
257	Row					
258	1	A	T	T	E	
259	2	N				
260	3	T	I	O	N	:
261	4	P	O	L		
262	5	I	X			

263

264 This will produce the ciphertext message below:

265

266 EN ANTPI : TOL TIOX

267

268 Applying the keyword MONEY to encrypt the second part of the message “CE BOSS –
 269 HOODLUM”

270	Col	1	2	3	4	5
271	K/W	M	O	N	E	Y
272	O/V	2	4	3	1	5
273	Row					
274	1	C	E	B	O	
275	2	S				
276	3	S	–	H		
277	4	O	O			
278	5	D	L	U	M	X

279

280 which produces the ciphertext message below:

281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332

OM CSSOD BHU E-OL X

The keyword PAPER when applied to encrypt the third part of the message with 15 characters” S PLAN MAYHEM ON AN”

Col	1	2	3	4	5
K/W	P	A	P	E	R
O/V	3	1	4	2	5
Row					
1	S	P			
2	L	A	N	M	
3	A				
4	Y	H	E		
5	M	O	N	A	N

will give rise to the ciphertext below:

PAHO MA SLAYM NEN N

The keyword WIPER if applied to the fourth part of the message “AMBRA RESIDENTS.”

Col	1	2	3	4	5
K/W	W	I	P	E	R
O/V	5	2	3	1	4
Row					
1	A	M	B	R	
2	A	R			
3	E	S	I		
4	D	E	N	T	S
5	.				

will produce the ciphertext message below:

RT MRSE BIN S AAED.

Which will produce the combined ciphertext message below:

EN ANTPI : TOL TIOX OM CSSOD BHU E-OL X
PAHO MA SLAYM NEN N RT MRSE BIN S AAED.

Further Splitting the message into 5 through 10 and choosing appropriate keywords would result in table 4.2 below

No. Of Split (S)	Average No. of Characters per split (L)
2	29
3	19.33333
4	14.5
5	11.6
6	9.666667
7	8.285714
8	7.25
9	6.444444
10	5.8

Table 4. 2: Merged Transposition Cipher Defined values of variables

From table 4. 2 above, the researcher intends to estimate a model and determine the level of error with respect to the number of split, S.

5.0 Model Estimation

In this section, the researcher intends to estimate a model and use the model to determine the level of error, ε with respect to the number of split, S of a given message.

Given a model:

$$L = \alpha + \beta S + \varepsilon, \text{ where}$$

L is the dependent variable, which is the average character No. in each split, S

S is the Split No. which is the independent variable.

α is point of interception of L at the point where S=0

β is the rate of change of L for each unit change in S.

ε is the error, the difference between the actual value of L and the estimated value, $\hat{L}$,

the model estimate $\hat{L}$ may be computed so as to find out error in the estimate with respect to the number of split., S.

From table 4.2 above, s, which is the difference between S and mean of S can be computed. Also l, the difference between L and mean of L are computed for each value of S (Okike, 2012). This is depicted in table 5.0 below:

S	L	s = S- Ś	l = L- Ĺ	sl	s ²
2	29	-4	16.56887	-66.2755	4
3	19.33333	-3	6.902202	-20.7066	9
4	14.5	-2	2.068872	-4.13774	16
5	11.6	-1	-0.83113	0.831128	25
6	9.666667	0	-2.76446	0	36
7	8.285714	1	-4.14541	-4.14541	49
8	7.25	2	-5.18113	-10.3623	64
9	6.444444	3	-5.98668	-17.9601	81
10	5.8	4	-6.63113	-26.5245	100
54	111.8802			-149.281	384

Table 5.0 : Values of S, L, l and s

Statistically,

$$\begin{aligned}\bar{S} &= \frac{\sum S}{n} \\ &= \frac{54}{9} \\ &= 6\end{aligned}$$

$$\begin{aligned}\bar{L} &= \frac{\sum L}{n} \\ &= \frac{111.8802}{9} \\ &= 12.43113\end{aligned}$$

$$\begin{aligned}\beta &= \frac{\sum sl}{\sum s^2} \\ &= \frac{-149.281}{384} \\ &= -0.38875\end{aligned}$$

$$\begin{aligned}\alpha &= \bar{L} - \beta \bar{S} \\ &= 12.43113 - (-0.38875 * 6) \\ &= 14.76363\end{aligned}$$

Therefore, the model is of the form:

$$L = 14.76363 - 0.38875S + \varepsilon, \text{ which means that } S \text{ is inversely proportional to } L.$$

$$\begin{aligned}L &= 14.76363 - 0.38875S + \varepsilon \quad \text{-----i} \\ \hat{L} &= 14.76363 - 0.38875S \quad \text{-----ii}\end{aligned}$$

then the error in the estimate can be computed from equation ii – i as shown in the table 5.1 below:

S	L	$\hat{l}$	ε
2	29	13.98613	15.01387
3	19.33333	13.59738	5.73595
4	14.5	13.20863	1.29137
5	11.6	12.81988	-1.21988
6	9.666667	12.43113	-2.76446
7	8.285714	12.04238	-3.75667
8	7.25	11.65363	-4.40363
9	6.444444	11.26488	-4.82044
10	5.8	10.87613	-5.07613

Table 5.1: Values of S, L, $\hat{l}$ and ε

From the table 5.1 above, it could be seen that as the split number, S of the message being encrypted increases, the error level in the model decreases.

Conclusion

The error level of the estimated model is inversely proportional to the number of split, S of the message.

References

- Christensen C. (2006), Simple Substitution Ciphers, www.nku.edu/~christensen/
- Collberg C. (2012), Cryptography — Introduction. Being Lecture delivered in the Department of Computer Science, University of Arizona, www.cs.arizona.edu/~collberg/Teaching/466-566/2012/
- Dhavare A. (2011), Efficient Attacks on Homophonic Substitution Ciphers, Being a Masters Project submitted to the Department of Computer Science, San Jose State University, <http://www.scholarworks.sjsu.edu/cgi/>
- Jain R. (2011), Classical Encryption Techniques, Washington University in Saint Louis, www.cse.wustl.edu/~jain/cse571-11/
- Okike B. (2012), Some New Encryption Techniques Using Firewalls and Random Generators, LAP Lambert Academic Publishing, Berlin – Germany, <https://www.lap-publishing.com/>