Diana Cipher

Virendra Sule

August 2020

1 Diana Cipher and its Variations

Diana cipher was developed by NSA during the Vietnam war and was a modification of the Vigenere cipher but essentially operating on the same principle as a periodic key pad. The main distinguishing feature of Diana was the symmetry in the three letters, the plaintext alphabet P, the key alphabet K and the ciphertext alphabet C. Diana pad had fixed trigrams, given any two of these three alphabets uniquely fixed the third. In Vigenere this property did not hold. Hence operation of the Vigenere cipher was lot more complicated compared to Diana since the trigrams could be easily memorised hence encryption and decryption was much simpler. In Vigenere encryption as well as decryption were complex and decryption resulted in many errors.

We shall first consider a Diana pad variant for only five alphabets A,B,C,D,E. the pad is constructed in a table as follows

1.1 Diana pad and its trigrams

The original configuration of the Diana pad is described in the table below. The key character K is located in the first column while the plaintext character P is in the first row. The Diana pad denotes the character as an operation K$\oplus$P in the intersection of the row and column defined by P and K.

$\oplus$	A	B	C	D	E
A	E	D	C	B	A
B	D	C	B	A	E
C	C	B	A	E	D
D	B	A	E	D	C
E	A	E	D	C	B

Table 1: Diana Pad with five alphabets

The trigrams are

AAE, ABD, ACC, BBC, BEE, CDE, DDD.

Plaintext	P	ADAEBCCDEAC
Key	K	DEACEDEACED
Ciphertext	P⊕K	BCEDEEDBDAE

Table 2: Encryption in Vigenere mode

Consider the encryption of the text ADAEBCCDEAC in these alphabets. By the key DEACE repeated to a stream as long as plaintext. Ciphertext is read out in the pad for key alphabet in the first column and plaintext alphabet in the first row of the table. Alternatively the ciphertext is the third alphabet of the trigram defined by plaintext and key letters.

This type of use of Diana pad can be called the Vigenere mode of operation.

1.2 Diana cipher as Vernam pad encryption by key stream

The encryption of a plaintext shown above uses Diana pad as Vigenere cipher by repetition of the key. Alternatively a text can be encrypted by a key stream of the same length by using the same Diana pad as a Vernam cipher. Hence a key stream needs to be generated using the symmetric key.

1. Generating a key stream using symmetric key and IV by padding, circulating and extending. Let the key as before be DEACE and IV be AACD. Then to construct the key stream of length 30 consider repetition of DEACEAACD by circulating and then extending as follows until length 40 is achieved

   DEACEAACD EACEAACDD ACEAACDDE CEAACDDEA EAAC

2. Self enciphering the key and IV. In this method the stream is generated by repeating the following process:

   (a) S(0)=key⊕IV

   (b) S(1)=S(0)⊕S(0)* where S* denotes left circulation of S by one character.

   (c) S(2)=S(1)⊕S(1)*

   (d) Keystream=S(0)||S(1)||S(2)||... until the plaintext length is achieved

   For above example of key and IV we get

   (a) S(0)=DEACEAACD, S(0)*=EACEAACDD

   (b) S(1)=CACDAECED, S(1)*=ACDAECEDC

   (c) S(2)=CCEBADDCE, S(2)*=CEBADDCEC

   (d) S(3)=ADEDBDEDD, S(3)*=DEDBDEDDA

   (e) S(4)=BCCAACCDB,

Hence the key stream of length 40 is

DEACEAACD CACDAECED CCEBADDCE ADEDBDEDD BCCA

Many such variations can be carried out. Diana cipher in Vigenere mode should as strong as Vigenere. In the key stream mode the Diana cipher is equivalent to a stream cipher. Cryptanalysis of these modes of operation of Diana shall be a future project.

1.3 Transformation of English text and numerals

Due to the limited alphabet size of 5 for our version of Diana we need to specify a scheme of translating English text along with numerals onto the alphabet. Consider the rule for translating decimal digits in the five alphabets as follows.

Digit	0	1	2	3	4
Alphabet	AA	AB	AC	AD	AE
Digit	5	6	7	8	9
Alphabet	BA	BB	BC	BD	BE

Table 3: Code for numerals in five alphabets

The code for numerals 0 to 9 is then given by double letters of the alphabet. In a text beginning and end of numerals may be indicated by letters XX which are not likely to occur in English text.

Next we construct the encoding of English text characters in the five alphabets. This is constructed in following table which assumes I=J rule. An English letter is located in the square and coded as the double alphabet Horizontal||Verticle in the first column and first row of the table respectively.

	A	B	C	D	E
A	A	B	C	D	E
B	F	G	H	J	K
C	L	M	N	O	P
D	Q	R	S	T	U
E	V	W	X	Y	Z

Table 4: Coding of English characters

1.4 Example

1. Plaintext: You Live only Twice, James Bond 007.
2. Coded text: EDCDDE CABDEAAE CDCCCAED DDEBBDACAE BDAACBAEDC ABCDCCAD ECECAAAABCECEC

3. Key: SIRMA

4. Key in alphabets: DCBDDBCBAA

5. Encryption: Vigenere mode words divided in 5 letter groups. See table 4.

	EDCDD	ECABD	EAAEC	DCCCA	EDDDE	BBDAC	
	DCBDD	BCBAA	DCBDD	BCBAA	DCBDD	BCBAA	
$\oplus$	CEBDD	EADDB	CCDCE	ABBCE	CEADC	CBAEC	
	AEBDA	ACBAE	DCABC	DCCAD	ECECA	AAABC	ECEC
	DCBDD	BCBAA	DCBDD	BCBAA	DCBDD	BCBAA	DCBD
$\oplus$	BDCDB	DACEA	DBDAE	ABCEB	CABEB	DCDDC	CBBE

Table 5: Example

1.5 Variations of the Diana pad

The Diana pad considered in the encryption above is not the only choice possible for the this kind of encryption. Since the pad used creates a fixed set of trigrams, the processes of encryption and decryption are decided by these trigrams. Hence any arrangement of pad which results into a set of triagrams is also a variant of the Diana pad considered above.

For simplicity of analysis consider Diana pad in just two alphabets A,B. Then the possible pads are

$\oplus$	A	B
A	A	B
B	B	A

$\oplus$	A	B
A	B	A
B	A	B

Similarly pads with other alphabet sets can be determined by filling up rows of the table.

1.5.1 Example of creating Diana pad for three alphabets

Consider alphabets A, B, C. First write the trigrams for two letters A$, where $ is chosen from the first row such that all alphabets appear in the first row of the table. Let these be chosen as AAA, the next enctry can be ABB or ABC. If ABB is chosen then the third entry must be ACC.

$\oplus$	A	B	C
A	A	B	C
B	B	A	C
C	C	?	B

Having filled the first row. The trigrams for the next row are BA$, BB$ and BC$. Hence these must be BAB (since ABB is already a trigram) or can be BAC. With the later choice we can fill next as BBA. But then the last column has to be BCA which repeats A. Hence if we start the second row with B to

get trigram BAB, then fill BBA. The last trigram has only option BCC. For the last row we have to consider trigrams CA$, CB$, CC$. Hence we have only options CAC, CBC and CCB. Hence the last row cannot be filled. Hence we take a different choice for first row as follows.

$\oplus$	A	B	C
A	B	A	C
B	A	B	C
C	C	B	A

This has trigrams AAB, ACC, BBB, BCC. Another choice of first row as given below shows that the Diana pad has no solution.

$\oplus$	A	B	C
A	C	B	A
B	B	A	C
C	A	C	?

1.5.2 Variation in coding of alphabets

Diana pad

1.6 Excercises

1. Construct more methods than given above, of generating a random looking stream of alphabets of a given length N dependent on a seed string of alphabet of small length n.

2. Construct all possible Diana pads of 3 alphabets, 4 alphabets and 5 alphabets.

3. Show that the Diana pad is as secure as Vigenere cipher of the same key length.

4. Construct a hash function and a MAC using the Diana pad.