Details Of Decryption Procedure

Steps performed in decryption are given as below.

Step-1 : **Finding Out Key Length**

Given cipher is taken as input and all the special characters are removed from it using the "**d_1_rmspch**" python program. Now the frequency analysis code which was written by me in python was employed on the cipher to find out the frequency of bigrams,trigrams and polygrams and checking for the spikes in the distances between those grams. From those results, i was not able to find any spikes which were related in terms of distances.

Eg :

Dif -Val	**Dif -Val**	**Dif -Val**	**Dif -Val**	**Dif -Val**	**Dif -Val**	**Dif -Val**	**Dif -Val**	**Dif -Val**
1 | 0	11 | 3	21 | 2	31 | 2	41 | 8	51 | 3	61 | 7	71 | 4	81 | 5
2 | 0	12 | 4	22 | 11	32 | 12	42 | 6	52 | 6	62 | 10	72 | 1	82 | 4
3 | 15	13 | 5	23 | 11	33 | 7	43 | 5	53 | 6	63 | 6	73 | 5	83 | 6
4 | 5	14 | 6	24 | 4	34 | 6	44 | 4	54 | 4	64 | 7	74 | 4	84 | 1
5 | 10	15 | 9	25 | 6	35 | 4	45 | 5	55 | 6	65 | 7	75 | 9	85 | 9
6 | 10	16 | 12	26 | 12	36 | 2	46 | 9	56 | 7	66 | 6	76 | 5	86 | 4
7 | 13	17 | 1	27 | 9	37 | 10	47 | 5	57 | 5	67 | 8	77 | 6	87 | 4
8 | 8	18 | 4	28 | 4	38 | 6	48 | 11	58 | 6	68 | 4	78 | 9	88 | 4
9 | 10	19 | 8	29 | 5	39 | 8	49 | 8	59 | 6	69 | 5	79 | 4	89 | 6
10 | 6	20 | 9	30 | 3	40 | 4	50 | 10	60 | 10	70 | 9	80 | 8	90 | 4

If you observe the above table there is no relation between any of the numbers in terms of the distances and its frequencies.

From this i deduced that it is not straight forward vigner cipher but its been tinkered with some shift cipher or substitution cipher as professor mentioned in the document.

So i started shifting the cipher by 0,1,2,3,.... and 2,4,6,8,... and 1,3,5,7,.... and so on and so forth using the "**d_2_Cipher_Shift**" python program with all the combinations possible. But i am not sure as to where should i apply this shift   to the cipher to form a relation between distances and its frequencies. As i don't know the length of keysize, i started applying shift changes in sequence, like (2,4,6,8,.... in lengths of 1,2,3,....) and (1,3,5,7... in lengths of 1,2,3,4).... and (0,1,2,3,4,.... in lenghts of 1,2,3,......)

After doing all the above steps in sequence i finally found a relation between lengths and distances at shifting the cipher by (0,1,2,3,... in length of 9) which is explained below.

Original Vignere Cipher	`Q d a r z e s Y l | s g l o l Z x p o | k x e c l f e w s`
After Shifting by 0,1,2,... in length of 9	`Q d a r z e s y l | s g l o l z x p o | k x e a j d c u q`

For this cipher, i ran the "**d_3_GramsFrequencyAnalysis**" python program and i was able to find the relation between distances and its frequencies. It is given below in the table.

Dif -Val	Dif -Val	Dif -Val	Dif -Val	Dif -Val	Dif -Val	Dif -Val	Dif -Val	Dif -Val
1 | 0	11 | 8	21 | 9	31 | 8	41 | 7	51 | 12	61 | 6	71 | 11	**81 | 29**
2 | 0	12 | 8	22 | 7	32 | 12	42 | 6	52 | 7	62 | 3	**72 | 20**	82 | 4
3 | 12	13 | 11	23 | 7	33 | 2	43 | 5	53 | 7	**63 | 48**	73 | 9	83 | 4
4 | 10	14 | 13	24 | 8	34 | 7	44 | 4	**54 | 53**	64 | 6	74 | 12	84 | 5
5 | 10	15 | 9	25 | 9	35 | 3	**45 | 35**	55 | 7	65 | 4	75 | 12	85 | 3
6 | 6	16 | 10	26 | 12	**36 | 35**	46 | 10	56 | 10	66 | 8	76 | 4	86 | 8
7 | 9	17 | 10	**27 | 78**	37 | 6	47 | 7	57 | 10	67 | 2	77 | 6	87 | 2
8 | 9	**18 | 49**	28 | 7	38 | 4	48 | 14	58 | 10	68 | 7	78 | 4	88 | 5
9 | 25	19 | 9	29 | 4	39 | 9	49 | 9	59 | 7	69 | 8	79 | 7	89 | 12
10 | 8	20 | 5	30 | 10	40 | 6	50 | 5	60 | 4	70 | 5	80 | 2	90 | 53

If you oberve the above table, you can see some relation between distances of multiples of 9 having higher values. From this , we can deduce that key can be of length 9 or its factors like 1 and 3. Key length cannot be 1 and 3 as they are too small.

So from this **KeyLength = 9**

Step-2 : Finding Out Key

Finding key is not straight forward as we have to check for the most frequently occurring characters in the vignere cipher text. For that i wrote a "**d_4_CharacterCount**" python program which will print the character and its frequency.

From this , i found out that letter **"H"** has occurred most number of times. Now the next step is to find out what are the patterns that are appearing most regularly like tri grams and poly grams as they are easy to debug compared to larger grams.

Frequency of Characters in Cipher text :

A | 76	G | 227	M | 98	S | 211	Y | 93
B | 76	H | 312	N | 111	T | 169	Z | 154
C | 133	I | 202	O | 217	U | 248	
D | 164	J | 145	P | 134	V | 291	
E | 200	K | 285	Q | 205	W | 211	
F | 192	L | 196	R | 287	X | 151	

Frequency of Characters in Engilsh Sentences or Paragraphs :

A | .082	G | .020	M | .024	S | .063	Y | .020
B | .015	H | .061	N | .067	T | .091	Z | .001
C | .028	I | .070	O | .075	U | .028	
D | .043	J | .002	P | .019	V | .010	
E | .127	K | .008	Q | .001	W | .023	
F | .022	L | .040	R | .060	X | .001	

If we observe the both tables, frequency of H is more in cipher and frequency of E is more in plain english words. So to convert **H** to **E** we have to use **D** from the Vignere Table.

Second most is V in cipher and S in plain . So we have to use **D.**

Current state till now is

Cipher Text : I V C A J D C U Q
Key : _ D _ _ _ _ _ _ _

Now let us take polygram DCUQ and analyze it.

D is 16th most frequent in Cipher and B is 16th most frequent in plain text. So for **D** to **B**, **C** is used.

Similarly C is 20th most in cipher and O is 20th most in plain. So for **C** to **O**, **O** is used.

Similarly U is 9th most in cipher and R is 9th most in plain. So for **U** to **R**, **D** is used.

Similarly Q is 10th most in cipher and N is 10th most in plain. So for **Q** to **N** , **D** is used.

Now final state is

Cipher Text : I V C A J D C U Q
Key : _ D _ _ _ C O D D

Now from permutations and some combinations in debugging the cipher text, i came to know it was a scientists named "**Edgar F Codd**" who is a Turing award receipient. And after decrypting the cipher text, i have observed that it is about "**Maurice Vincent Wilkes**" who is also a Turing award receipient.

Cipher Text : Q D A R Z E S Y L R F K N K Y W O N I V
Key : E D G A R C O D D E D G A R C O D D E D
Plain Text : M A U R I C E V I N C E N T W I L K E S

Key Found = EDGARCODD

Finally wrote a **"d_5_Decryption_Procedure_code"** python program which will do the decryption if key is provided.

List Of Programs Written By me :

1) d_1_rmspch.py
2) d_2_Cipher_Shift.py
3) d_3_GramsFrequencyAnalysis.py
4) d_4_CharacterCount.py
5) d_5_Decryption_Procedure_code.py