Exercise 3 in Cryptography
- Proposed Solution -

Prof. Dr. Rudolf Mathar, Henning Maier, Jose Angel Leon Calvo

2015-04-30

Solution of Problem 7

The first step is to perform a frequency analysis. The frequency analysis consists of counting the number of appearances of each letter and the number of ordered pairs as follows:

i	Character	k_i	p_i
0	A	1	0
1	B	4	6
2	C	12	66
3	D	4	6
4	E	2	1
5	F	1	0
6	G	6	15
7	H	2	1
8	I	7	21
9	J	2	1
10	K	14	91
11	L	7	21
12	M	1	0

i	Character	k_i	p_i
13	N	3	3
14	O	4	6
15	P	4	6
16	Q	4	6
17	R	10	45
18	S	4	6
19	T	0	0
20	U	0	0
21	V	9	36
22	W	6	15
23	X	8	28
24	Y	3	3
25	Z	0	0

- k_i the total number of ocurrences of each letter
- p_i the number of ordered pairs calculated as $p_i = \binom{k_i}{2}$

The next step is to calculate the total length of the ciphertext using that the text is divided in blocks:

$$n = 14 \cdot 8 + 6 = 118 \tag{1}$$

Now we have to use the formula of the index of coincidence as follows *(Def: 3.1, page 13, lecture notes)*:

$$I_C = \frac{|\{i,j\}|C_i = C_j, 1 \le i < j \le n\}|}{\binom{n}{2}} = \frac{\sum_{i=0}^{25}\binom{k_i}{2}}{\binom{n}{2}} \tag{2}$$

For 2, we calculate all the possible pair of combinations of n symbols as:

$$\binom{n}{2} = \frac{n!}{(n-2)!\,2!} = \frac{n(n-1)}{2!} = \frac{118 \cdot 117}{2} = 6903 \tag{3}$$

Therefore, substituting in 2, we have:

$$I_C = \frac{6 \cdot 0 + 3 \cdot 1 + 2 \cdot 3 + 6 \cdot 6 + 2 \cdot 15 + 2 \cdot 21 + 1 \cdot 28 + 1 \cdot 36 + 1 \cdot 45 + 1 \cdot 66 + 1 \cdot 91}{6093}$$

$$= \frac{383}{6903} = 0.055483 \tag{4}$$

Using the hint of the exercise, we can assume that the text is monoalphabetic and probably in English, because the index of coincidence obtained is close to its value.

Solution of Problem 8

a) We have the auto-key cryptosystem:

$$c_i = \begin{cases} m_i + k_i \pmod{26} & 0 \leq i \leq n-1 \\ m_i + c_{i-n} \pmod{26} & n \leq i \leq l-1 \end{cases}$$

Using a ciphertext only attack, we can compute the message as follows:

$$c_n = m_n + c_0 \iff m_n = c_n - c_0$$

$$c_{n+1} = m_{n+1} + c_1 \iff m_{n+1} = c_{n+1} - c_1$$

$$\implies m_{n+j} = c_{n+j} - c_j$$

Therefore, the next task is to determine n

b) Using the result from a) we decipher the following text, just shifting the ciphertext along itself:

For $n = 1$

D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A	
	D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
	I	V	P													

For $n = 2$

D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A		
		D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
		D	K	N													

For $n = 3$

D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A			
			D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
			S	I	S	T	H	E	A	U	T	O	K	E	Y			

Only the first characters are missing in the message. For these characters, we guess them.
Message: THIS IS THE AUTOKEY

c) Consider:

$$\hat{c}_i = \begin{cases} m_i + k_i \pmod{26} & 0 \leq i \leq n-1 \\ m_i + m_{i-n} \pmod{26} & n \leq i \leq l-1 \end{cases}$$

In this case, we know the keylength n and the message m. Therefore, we can obtain the message by:

$$c_i = m_i + m_{i-n} \tag{5}$$

With a Friedmann attack, using the most common characters in the English language, the message can be deciphered with a high probability

d)

Q	E	X	Y	I	R	V	E	S	I	U	X	X	K	Q	V	F	L	H	K	G
T		E		E		R		B		T		E		M		T		O		S
	H		R		A		E		E		T		R		E		H		D	

The plaintext is: THERE ARE BETTER METHODS

Solution of Problem 9

In this exercise, we have to apply the Kasiski-Babbage method as follows:

$$Y_{ij} = \begin{cases} 1 & \text{if} \quad c_i = c_j \\ 0 & \text{else} \end{cases}$$

then

$$\mathrm{E}[Y_{ij}] = \begin{cases} k_m & \text{if} \quad c_i = c_j \\ \frac{1}{m} & \text{else} \end{cases}$$

It follows for $m = 26$ (using English language):

$$k = \frac{0.028433n}{(n-1)I_C - 0.0385n + 0.066895} \tag{6}$$

In our case, the length of the message is $n = 3568$. Therefore, $k \approx 6.25643$. The length of the key has to be an integer, $k \approx 6$. We use the hint at the beginning of the exercise, getting $k \approx 5$.

Once we have the keylength, we perform a frequency analysis of the ciphertext. We create a frequency analysis for each of the 5 columns of the ciphertext. As we know, the most common characters in English language are: E, T, A, O, I, N.

The frequency analysis in detail is as follows:

Block	**Character**	**Frequency**	**Char**	**Frequency**	**Char**	**Frequency**
1	T	89	I	68	P	61
2	P	103	E	69	T	56
3	Y	94	N	63	C	58
4	X	101	B	59	G	53
5	S	85	H	68	B	58

Once this analysis is finished. We map the most common character to the character E, the second to T and we do the same with the following. Using this method, we obtain the key: Key = (T $\rightarrow$ E, P $\rightarrow$ E, Y $\rightarrow$ E,X $\rightarrow$E, S $\rightarrow$ E) = PLUTO

Using this key to decipher the ciphertext, the first sentence of the message is: THE BLACK CAT THE MOST WILD YET THE MOST HOMELY NARRATIVE WHICH I AM ABOUT