Prof. Dr. Rudolf Mathar, Dr. Michael Reyer

Tutorial 3
- Proposed Solution -

Friday, May 3, 2019

Solution of Problem 1

a) We have the autokey cryptosystem:

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

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

$$c_n = m_n + c_0 \mod 26 \iff m_n = c_n - c_0 \mod 26$$

$$c_{n+1} = m_{n+1} + c_1 \mod 26 \iff m_{n+1} = c_{n+1} - c_1 \mod 26$$

$$\implies m_{n+j} = c_{n+j} - c_j \mod 26$$

Now determine n by trying $n = 1, 2, \ldots$ until the ciphertext starting at position n sounds reasonable. You still need to guess the first part of the message.

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

For $n = 1$

c_k	D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A	
c_{k-n}		D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
m_k		I	V	P													

For $n = 2$

c_k	D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A		
c_{k-n}			D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
m_k			D	K	N													

For $n = 3$

c_k	D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A			
c_{k-n}				D	L	G	V	T	Y	O	A	C	O	U	V	C	E	Z	A
m_k				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

Now we also may calculate the key by calculating "DLG"-"THI"="KEY".

c) Consider:

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

In this case, we know the keylength n, and we know that the message $\mathbf{m}$ is used to generate $\hat{c}_i$. Therefore, we can obtain the message by frequency analysis on:

$$\hat{c}_i = m_i + m_{i-n}. \tag{1}$$

With a Friedmann attack, using the most common characters in the English language, we derive the most common $\hat{c}_i$'s. The message can be deciphered with a high probability then. Here, we can say 'I' is the most probable letter, if combining two english letters. Moreover, 'E'+'E'='I' is the most likely combination for getting the letter $\hat{c}_i =$ 'I'. Hence, we have a look at a positions $k \geq n$ in the cryptogram with $c_k =$'I' and now know that $m_k = m_{k-n} =$'E' holds true with high probability. Moreover, we know

$$m_{k-(j+1)n} = \hat{c}_{k-jn} - m_{k-jn} \mod 26 \; \forall j \in \mathbb{N} \text{ with } k - jn \geq n,$$
$$m_{k+jn} = \hat{c}_{k+jn} - m_{k+(j-1)n} \mod 26 \; \forall j \in \mathbb{N} \text{ with } k + jn < l.$$

d) In our case there are two positions with 'I'. The first occurence is used as described above to get two times 'E' and aftewards calculating each 2nd (n-th) letter of the message. The second occurence reveals the remaining text.

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

The key can be calculated by 'QE'-'TH'='XX'.

Solution of Problem 2

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

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

then

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

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

$$k = \frac{0.028433n}{(n-1)I_C - 0.0385n + 0.066895} \tag{2}$$

In our case, the length of the message is $n = 3568$. The index of coincidence is approximately $I_C \approx 0.043037$. 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 FOR THE MOST WILD YET MOST HOMELY NARRATIVE WHICH I AM ABOUT