Faris Halteh
Kim Feenstra Kuiper
Group 20

Cryptanalysis of the Vigenère cipher

(b) Breaking the cipher

How to run and use the code to crack the ciphertext of a Vigenère cipher

1. In order to start the program (that's written in Java), please open crackSweden.java and run this file using a compiler.
2. Once the file starts running, it will ask the user to enter the text to be decrypted. The user can paste the copied decrypted text here and press enter to let the program do the next step.
3. On pressing enter, the user will receive a table that represents the key length and the frequency corresponding to that key length. Then he or she will be asked to enter the number that corresponds to the most common spacing factor to identify the length of the keyword.
4. The user will then get a list of letters with the number of occurrence of each letter listed. All these letters have been encrypted with the same character from the key. The user will then receive the most frequent letter and will be asked to enter his or her suggestion if he or she thinks that another key is the most frequent one (this could be useful in situations where more than one letter have the same frequency).
    a. Press 1: to go on.
    b. Press 2: to suggest a different letter.
5. Step 4 will be repeated for all suggested key letters in order to amalgamate letters that could make the key.
6. The user will receive the key letters.
7. The user will then get the plain text that has been computed from the suggested key.
8. The user will finally be asked if he or she wants to decrypt another text or terminate the program.
    a. Press 1: to decrypt another text.
    b. Press 2: to terminate the program.

To decrypt the ciphertext we took several steps:
1) Check for frequent substrings (part of Kasiski test)
2) Check the distance between 2 of the same substrings (part of Kasiski test)
3) The distance modulus key length is 0, and a list of possible key lengths is provided
4) The user guesses the key length; based on the possible length list. (part of Kasiski test)
5) Divide the ciphertext up in substrings with the same length of the key. The text is now in a matrix
6) We apply frequency analysis on the 1st column, because they are encrypted with the same letter. (The Friedman Test)
7) We find the distance between the most common letter in the plain text (e) and in the ciphertext (letter most occurred in one column). From this we derive the shift and the letter which makes this shift. (The Friedman Test)

8) We repeat step 6 and 7 for each column.

9) All the letters result in a word; the key.

10) We decrypt the ciphertext with the found key.

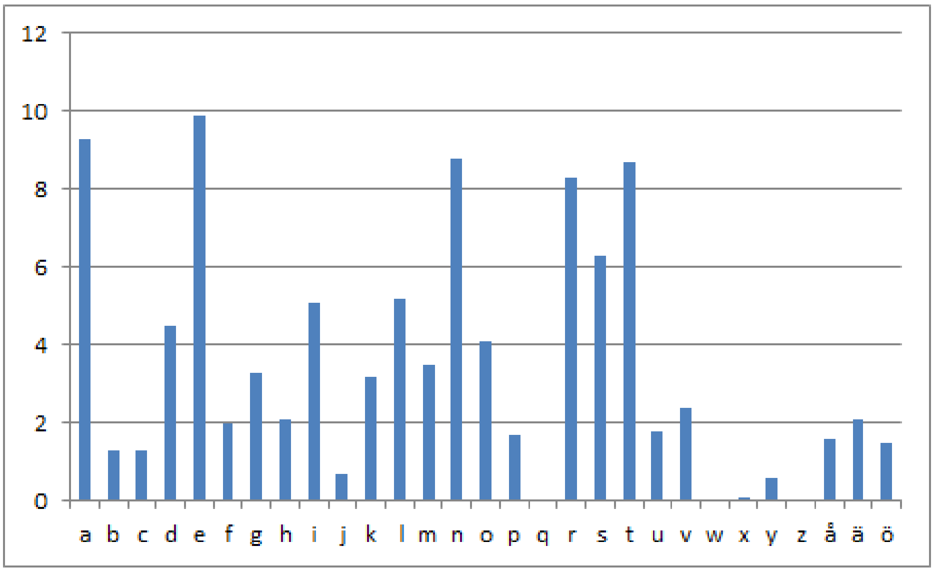


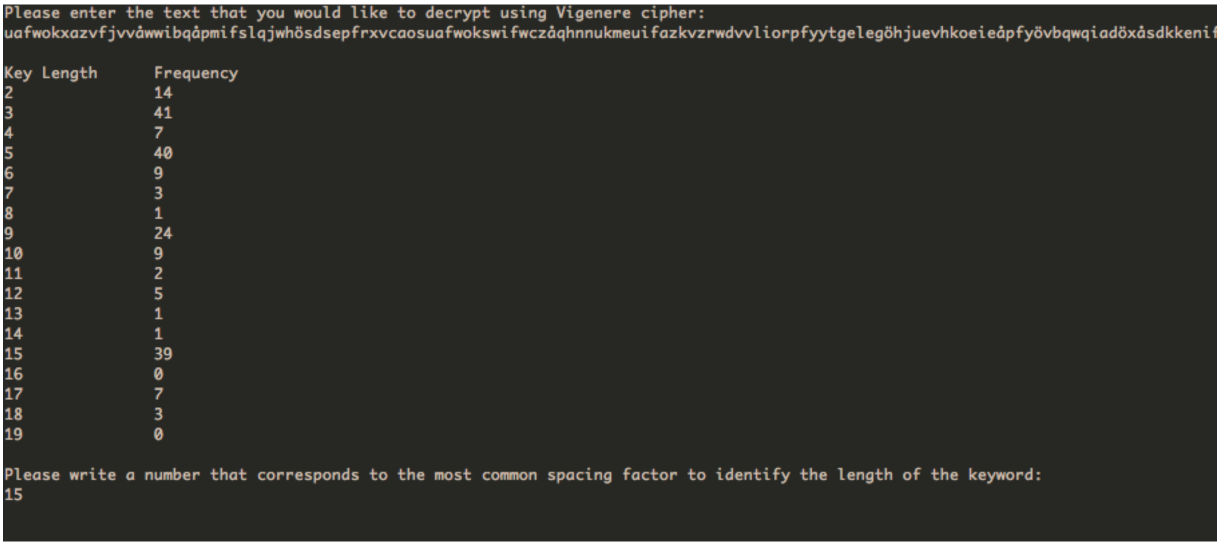
Table 1. Frequency table of Swedish letters[1]

Unfortunately we did not manage to decrypt any of the ciphertext; even not our own text. We think the problem is that the frequency analyses is not accurate enough to produce the right outcome. As we can see in table 1, is the frequency of *e, a, t, r* and *n* are all very high. On the other hand, if we would have english ciphertext, this would be much easier because there's a big difference in the frequency of occurrence between the letter *e* and all other letters. We did manage to guess the key length effectively using the Kasiski method. In addition, the program asks the user for a possible key length after printing out a list of the numbers corresponding the most common spacing factor to identify the length of the keyword. The highest number is not necessarily the length of the key, which makes it quite hard to compute this variable automatically without asking for any user input. The running time for group 8 is 12846000000 ns.

[1] Singh, Simon; Brogren Margareta (1999)
Kodboken: konsten att skapa sekretess - från det gamla Egypten till kvantkryptering Stockholm: Norstedt.
Libris 8345451. ISBN 91-1-300708-4

Faris Halteh
Kim Feenstra Kuiper
Group 20

Here you see one example:



```
Please enter the text that you would like to decrypt using Vigenere cipher:
uafwokxazvfjvvåwwibqåpmifslqjwhösdsepfrxvcaosuafwokswifwczåqhnnukmeuifazkvzrwdvvliorpfyytgelegöhjuevhkoeieåpfyövbqwqiadöxåsdkkenif

Key Length      Frequency
2               14
3               41
4               7
5               40
6               9
7               3
8               1
9               24
10              9
11              2
12              5
13              1
14              1
15              39
16              0
17              7
18              3
19              0

Please write a number that corresponds to the most common spacing factor to identify the length of the keyword:
15
```

Group Name	Suggested Length of the Key	Suggested Key (may not be accurate)
Group 1	15	dwabkglbrbdayre
Group 2	15	sakgisypllahöää
Group 3	12	khrsnöchlbml
Group 4	13	ltxnqaröeehlr
Group 5	7	fäbhden (färbaden-färboden)
Group 6	8	cåuöyonj
Group 7	15	öenröwöxääbwgow
Group 8	8	mjndwapt (mindwarp)
Group 9	15	åeååiugucäfbymc
Group 10	16	ikpöolhaöhdrpäqc
Group 11	15	datrngeguäwtöst
Group 12	12	rjäekuopamin
Group 16	11	äwmoråajdeö
Group 17	5	ctjox
Group 18	14	jntgbaöcajnhue

Group 19	11	päylhqhäobk
Group 20	9	gööerdpiw

Assist 1	34	jgfpkhkcmjcdönhnmiäieåcndhggmönels
Assist 2	30	deesjdfmnoehoyofkhnrkodcnpmweå
Assist 3	33	frjäataoöoäsöcnkcöpjbtcsåänpuhöwx
Assist 4	31	ixåqnleidnnpoböhjäjlaqggtdfmobc
Assist 5	23	ncwidööqnfpkäcbnpnhåefö

(c) Reasoning about ciphers

You have been asked to break a modified version of the Vigenère cipher, with 29 characters instead of 26. Do you think this version is easier to break, more difficult, or of the same level of difficulty?

On one side: the longer the alphabet, the more options and the longer your calculations will be. So it will be more difficult to solve it. But on the other side will there be more combinations of letter, substrings will be easier to recognize. In the end we think the swedish alphabet is slightly more difficult to solve than the english alphabet.

How does the length of the key affect the security of the Vigenère cipher? Are there other characteristics that can impact the security of a particular key?

The longer the key, the harder it gets to decrypt the ciphertext. If you have a very short key, the chance is high you will find a lot of repetitions, and you will find the key length easily. If one has a very long key, one will find less repetitions and it is harder to find the key length. If you do not know the key length it is impossible to solve the text (you have to guess for every possible key length). The Vigenère cipher is very safe if the key is as long or longer as the plaintext.

How could you break the Vigenère cipher if you don't know the language of the plain text?

We would look to buzz letters, like å, é, œ, Ø. Then we would organize the list of possible languages on the used letters. (If it includes a Ø it is more likely a Danish text then a Swedish text). And then you can start solving the text for each language.

1.

It is a cipher, which can not be cracked. Since the key is as long as the plaintext. This makes it quite challenging to extract the key from the decrypted text. As there is no relation between such text. Though it is hard to let Bob know the key (Cryptanalysis cannot be performed). If you know the key, you also immediately know the plain text. So this is not very useful.

2.

This cipher is very similar to vigenere, the only difference is the use of subtraction instead of addition. Because it is so similar to vigenere, it has the same security and it will have the same cost of a brute force attack. We would apply the same cryptanalysis on this technique (combination of Kasiski and Friedman tests in order to attempt to decrypt the key).

3.

This cipher is more secure than the vigenere, because it requires an extra step. It hard to brute force your solucion; the costs will be high. The technique we would try is to start in the end of the letters and work back. Though, it is very hard to figure out the letter before your letter.

4.

This cipher is more secure than a general Vigenere, because the whole key is longer than the text. And, it starts with a key which is different then the text. If you know the first key, it should be possible to solve the rest. But it is very hard to figure out the first key, because nothing gets repeated.

5.

This cipher has a fixed key length, though it is very hard to decrypt it because the letter constantly changes. It is more secure than the vigenere but not unbreakable. The initial key is embedded in the next key, which is embedded in the next key, and so on. This makes it very secure and very expensive to break without knowing the length of the first key.

Evaluation

Throughout this lab assignment we have faced numerous challenges that got in our way. However, it was quite enjoyable to figure out how to crack a vigenere cipher. We are slightly disapointed that we were not able to decrypt any of the ciphers. On the onther we are able to destinguish the keylength and analyse the frequency of the letters in the ciphertext. We learned a lot from bringing the theory into practice. The collaboration between us was very positive and efficient, we did pair programming and discussed all the details. The workload was equally devided.