



GAZI UNIVERSITY
FACULTY OF ENGINEERING – COMPUTER ENGINEERING

171180010 – Cansu AYTEN

CENG482 INTRODUCTION TO COMPUTER SECURITY

HOMEWORK 5

APRIL 2022

1. USING LETTER FREQUENCY FOR CRYPTOANALYSIS

The method called letter frequency is counting how often the letters appear in a text. Each language has a different number of letters and they appear at different frequencies in a text. For some languages, letters or groups of letters generally used in texts may be more or less than other languages. Therefore, the frequency of letters will be different for each language. For example, it can be said that while the letters E, T, A, and O are used frequently in English, the letters Z, Q, X, and J are less common. In addition, letter groups such as TH, ER, ON, and AN and the repetition of letters such as SS, EE, TT, and FF are frequently seen in English. This conclusion is valid for texts written in English that are long enough. The shortness of the text may lead to different letter distribution results. In cryptanalysis, the method based on how many letters or groups of letters occur in an encrypted message is called frequency analysis. In this method, a frequency table is created by making use of letter frequencies. Frequency analysis is a method used to decrypt texts encrypted with a substitution cipher. In a substitution cipher, the letters of a text are replaced with other letters. For example, if a letter a in the text is to be converted into an x, this process must be performed for each letter a. The receiver applies a reverse substitution to decipher this text. To explain the basic use of the frequency analysis method, first of all, the frequency of the letters in the ciphertext should be found. Then, using the letter frequency table of a language, the letters in the ciphertext should be associated with the letters of the predicted language. For example, the most common letter in the ciphertext may be the most common letter in the predicted language. Apart from this, similar results can be obtained by examining letter groups. [1,2,3].

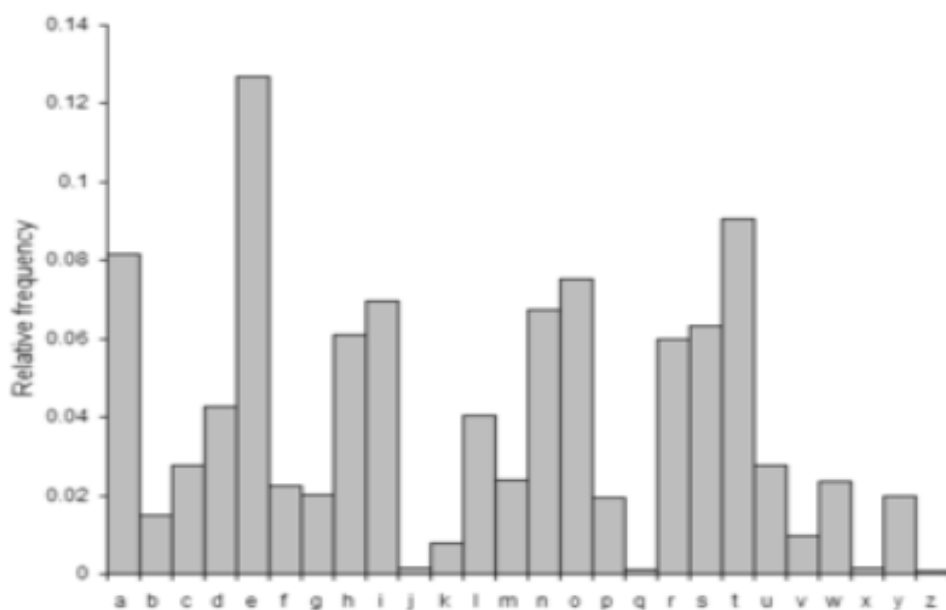


Figure 1: Letter Frequency in English [4]

In Figure 1, the letter frequency table for English is given. According to this table, it can be concluded that the letters e, t, a, o, i, n, s, h, r, d, l and u are the 12 most used letters.

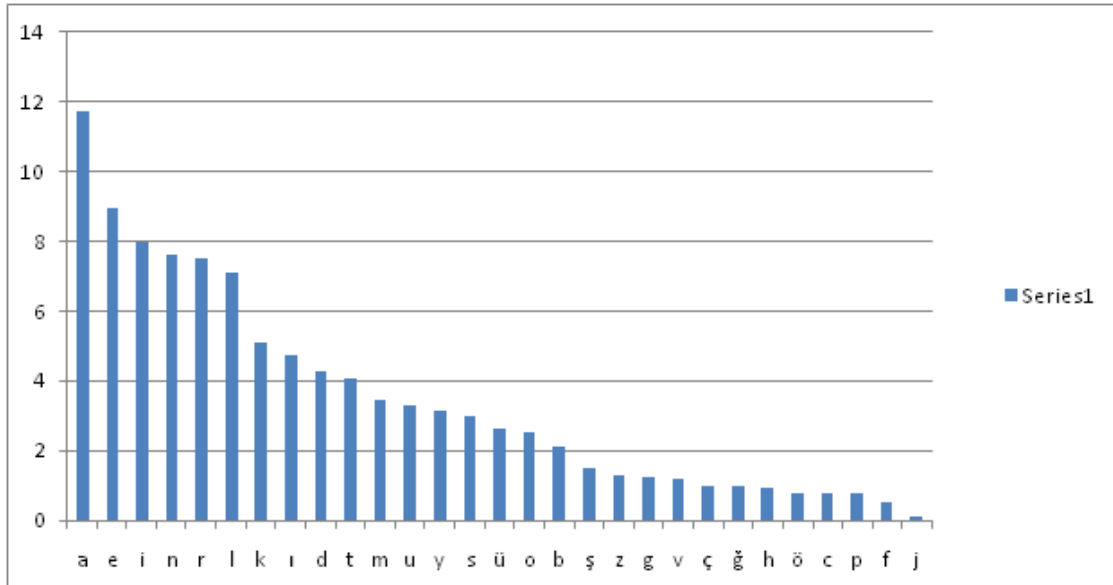


Figure 2: Letter Frequency in Turkish [5]

In Figure 2, the letter frequency table for Turkish is given. According to this table, it can be concluded that the letters a, e, i, n, r, l, ı, d, k, m, u and y are the 12 most used letters.

2. EXAMPLE FOR LETTER FREQUENCY

Story: Town Musicians of Bremen

Once upon a time an old and tired donkey ran away from his owner because he heard he was about to be sold to a butcher. The poor animal had to work carrying bags around a mill and it was no surprise he was out of strength. The owner gave him less and less food so he became weaker. In the end the owner thought it was best to get rid of him. The donkey was very unhappy about his owner's decision but he didn't want to surrender. He decided to abandon him before he ends up in the butcher shop. While he was wandering he stumbled across a sad, hunting dog whose owner made him guard the house. He breathed loudly and without any strength. He also decided to run away from his owner in order to avoid bad faith. The two sad animals started walking towards Bremen. They were lucky to have found each other because they could take care of each other. They had to get some food so they decided to become musicians. They also encountered an old cat and a rooster on their way and they also joined them. The cat's owner wanted to get rid of him because he wasn't good at catching mice anymore and the rooster's owner wanted to make lunch out of him. They walked and walked until sunset. They couldn't

reach Bremen but they found an old wooden shack where they spent the night. When they looked out the window they saw burglars preparing a tasteful dinner. They had a wonderful idea. They decided to sing under their window and earn their food. The dog climbed onto a donkey, cat onto the dog and the rooster on the cat. When they started singing it sounded like they were yelling. The burglars thought a wild animal is going to attack them so they got scared and ran away. When the burglars ran away the animals finally had something to eat. They enjoyed themselves after a very long journey. After dinner they went to bed to get some rest before they go to Bremen. They turned off the light and they made their beds. The donkey slept on the hay, the dog next of the door, the cat near a fireplace and the rooster sat on a high wooden beam. One of the robbers went back to the shack to see if there was a real animal there. As soon as he stepped in the cat scratched him. After that the donkey kicked him with his hoof, the dog bit him and the rooster crowed loudly. The burglar never came into the shack again. They could catch their breath after they defeated the burglar. They were very proud of themselves because they stood by each other's side when they needed to. They weren't ready to live under someone's reign again and that's why they got rid of the burglars. The burglar didn't even think about coming back to the shack again because he could get seriously injured. The animals lived in the shack until they died. They enjoyed good food and quality friendship [6].

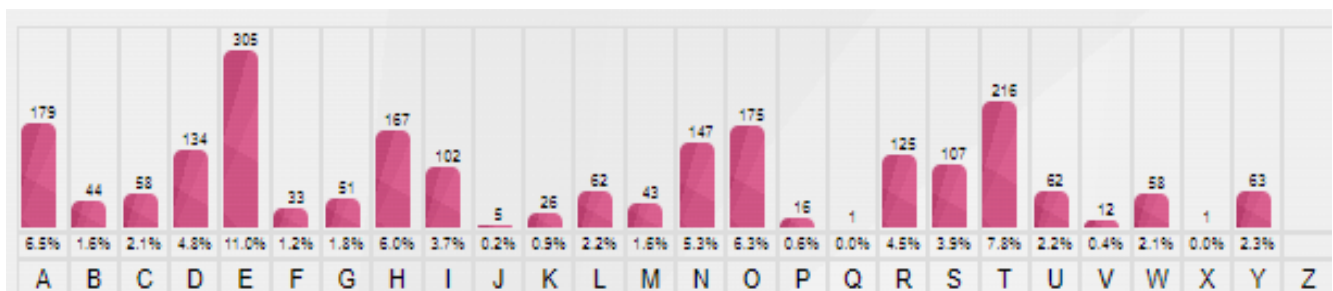


Figure 3: Letter Frequency For The Story

When Figure 3 is examined, it can be concluded that the letter E occurs 305 times, the letter T 216 times, the letter A 179 times, the letter O 175 times, and the letter H 167 times in a story of about one page. The letters that appear the least in the story are the letters J, Q, X, and Z. The letter J is used 5 times in total, the letter Q is used once, and the letter X is used once. The letter Z has never been mentioned in the story.

REFERENCES

1. Frequency analysis. (2021). Wikipedia
https://en.wikipedia.org/wiki/Frequency_analysis#Frequency_analysis_for_simple_substitution_ciphers
2. Frequency Analysis. (2019).
<https://www.101computing.net/frequency-analysis/>
3. Substitution cipher. (2022).
https://en.wikipedia.org/wiki/Substitution_cipher
4. The frequency of the letters of the alphabet in English. (n.d)
<https://www3.nd.edu/~busiforc/handouts/cryptography/letterfrequencies.html>
5. Gunasekara, M. (2010). Character Frequency Analysis for Turkish (Türkçe).
<http://melakatechnology.blogspot.com/2010/11/character-frequency-analysis-for.html>
6. Town Musicians of Bremen. (n.d)
<https://www.bookreports.info/town-musicians-of-bremen-summary/>