CS641A (Modern Cryptology)

Chapter-1 Report

Instructor: Prof. Manindra Agrawal

Group Name: CMEN

Names :	Pratyush	170500
	Pravar Deep Singh	160508
	Tushar Garg	160749

Date of submission of report : 23/01/2020

In Chapter 1, as we enter we see the following message:

> *You are standing at the base of a hill. There are big rocks and boulders strewn all over the hill. These make it difficult to estimate its height. However, it does appear to be very high. There is very little greenary on it. Someone seems to have carved out a trail -- seemingly leading to the top of the hill -- through the rocks. The trail appears very old. You are tempted to follow it.*

Now we exhaustively tried all the commands which we got by typing the keyword “list” which we got from the mail, and we get the following message by using the command “climb”:

> *You start to regret your decision. The climb has been very steep at places, and to add to your woes, the trail has been full of small rocks that have not been kind to your feet. Just when you decide to turn back, you notice a small opening where the trail seems to end. As you approach it, you notice a depression in the hill at the opening. From the distance, it appears as if someone has carved our a hole and fitted a door there. Coming closer, you realize that it is indeed a door! But it is closed and has no latches. How does one open it, you wonder? Then you notice that someone has carved some lines on the rocks on one side of the door. Looking closer, you realize that these are not random lines: it is some message written on the rocks!*

Since “some message” is written in the text we tried the command “read” and obtained the following message:

You clear up the dust from the rocks and read:
You are about to enter the Great Caves. There are many chambers inside. Doors to some chambers are open, and others closed. To open a closed door, you need to speak out a secret password. You will find these passwords and other messages written on screens near the doors. However, they would be coded and so you will have to first understand these codes to find out the passwords. But beware, once you enter the caves, you cannot come out unless you find the magic wand which is hidden somewhere inside. If you still want to go in, speak enter and the door in front of you will Open.

In the text itself it was specified to say “enter” so we entered the command “enter” and obtained the following message:

The door closes behind you with a loud noise. At first you cannot see a thing, but slowly, as your eyes adjust to the dim light coming through the small gaps between the door and walls, you see that this is a small chamber. The chamber is completely empty. You then notice another closed door in the chamber. Moving nearer, you see a glass panel near the door with something written on it. This must be the screen you think.

We then entered command “read” to read the message written on glass panel and obtained the following message:

Age qlmd dbvdhdt vqd nrhxv iqljsdh gn vqd ilmdx. Lx age ilb xdd vqdhd rx bgvqrbw gn rbvdhdxv rb vqd iqljsdh. Xgjd gn vqd olvdh iqljsdhx kroo sd jghd rbvdhdxvrbw vqlb vqrx gbd, r lj xdhrgex. Vqd igtd exdt ngh vqrx jdxxlwd rx l xrjpod xesxvrvevrgb irpqdh rb kqriq trwrvx qlmd sddb xqrnvdt sa 6 polidx. Ngh vqrx hgebt plxxkght rx wrmdb sdogk, krvqgev vqd uegvdx.

emTc88Qqjt

We realised that message is encrypted and we need to decrypt it.To do so we followed the below steps.

1. Finding the Encryption type:-

After looking at the text one can notice that the text contains proper punctuation, one letter words, two letter words and they seem to make a proper structure of a English sentence by which we thought that the encryption method uses may be **Substitute Cipher**. To find out what encryption method might be used, we first did the frequency analysis of the ciphertext. The result is shown in the graph below. The frequency distribution of normal English text is also shown below. We observed that the frequency distribution of ciphertext and normal English texts are very similar.

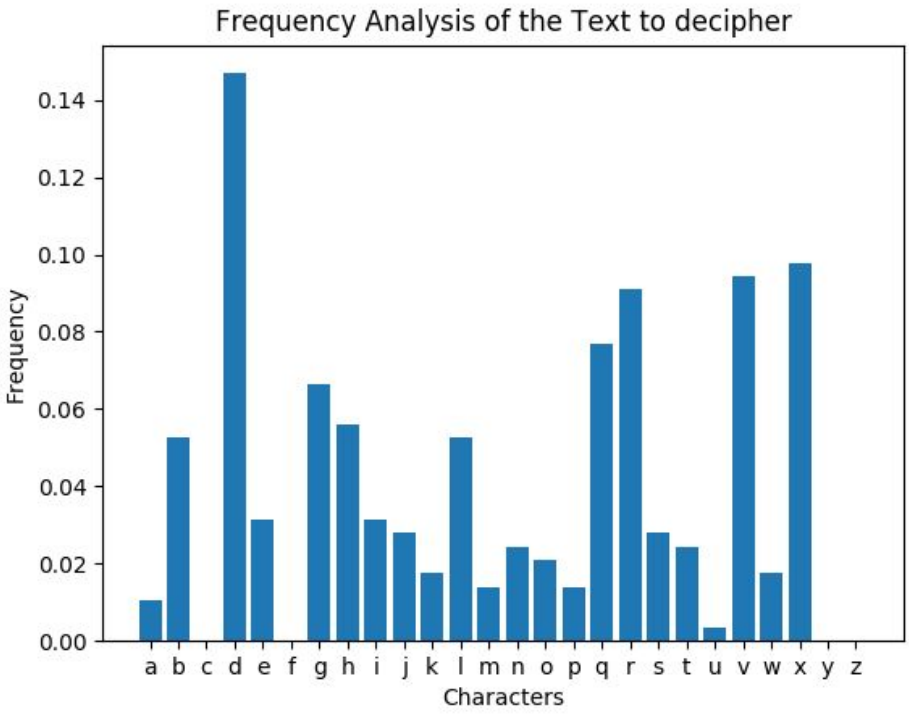

Fig. 1: Frequency Analysis of text to decrypt

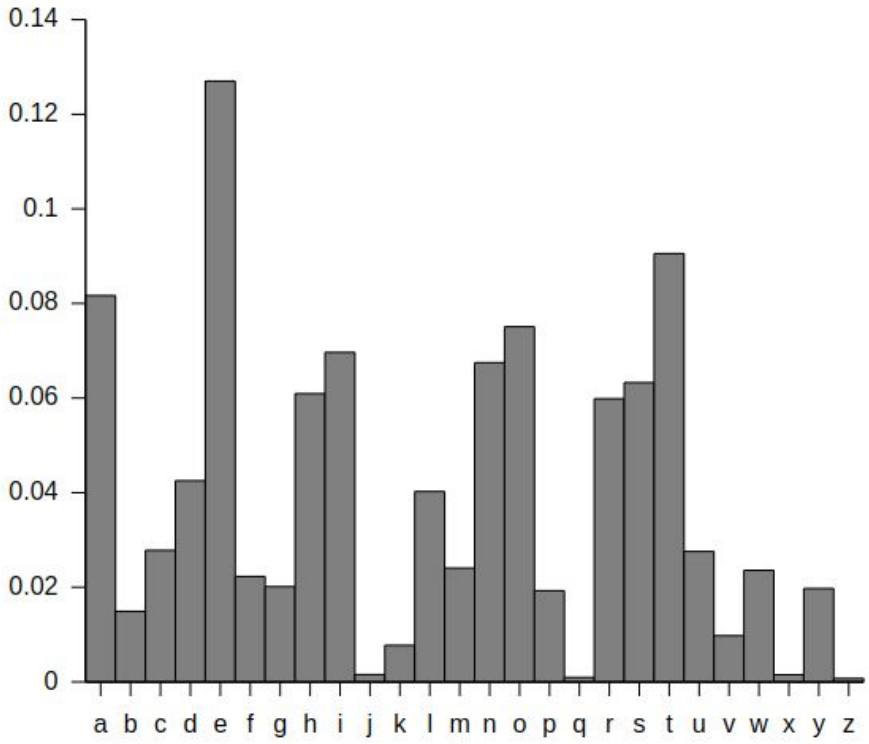

Fig. 2: Frequency Analysis of the Normal English Text

Now we calculated the Index of coincidence of the given ciphertext and found it to be **0.069857 ~ 0.07,** which is nearly the same as the standard English text's Index of coincidence value of **0.07**. Therefore, we could see high possibilities of **monoalphabetic substitution cipher**.

2. Breaking The Ciphertext:-

We started to analyse the word in the encrypted text and observed that the word "vqd" a three-letter word was present 7 times. Since most frequent three-letter word in English is "the", and also "d" was most frequent in ciphertext (in English most frequent is "e"), the correspondence of "vqd" with "the" seemed very probable.

Next, we observed the single letters "l" and "r" in the ciphertext. These could correspond only to "a" and "i" in the English text. Let us assume that "r" corresponds to "a" and "l" corresponds to "i". Now we have "r lj" in the ciphertext, it means "lj" is a two-letter word starting with "i" preceded by "a" (based on the above assumption). Now two letter words starting with "i" are "if", "is", "in" and "it" but none of them preceded by "a" makes sense. Hence, our assumption was wrong and we can conclude that "r" corresponds to "i" and "l" corresponds to "a". Also, in that case "r lj" corresponds to "i a_", where blank makes sense only when substituted with "m", hence "j" corresponds to "m".

Next, we have "sd" in the ciphertext which has only possibilities as "be" or "we" in English. But we notice that "sa" is also present in ciphertext, which rejects "s" being "w" as "w_" doesn't have two possibilities in English. Therefore "sd" corresponds to "be" and "sa" to "by". Now, we have "age" (appears multiple times) in ciphertext which corresponds to "y__" which has a high probability of being "you".

Next "qlmd", "gbd" and "xdd" in ciphertext gives after replacing currently found letters "ha_e" and "o_e" and "_ee" which seemed corresponding to "have", "one", and "see" in English. Next "dbvdhdt", "ilmdx", "uegvdx" and "wrmdb" in

ciphertext gives after replacing currently found letters “ente_e_”, “_aves”, “_uotes” and “_iven” respectively which seemed corresponding to “entered”, “caves”, “quotes” and “given” in English.

Next “kqriq”, “polidx” and “ngh” (occurring several times) in the ciphertext gives after replacing currently found elements “_hich”, “_ _ aces” and “_or” respectively, which seemed corresponding to “which”, “places” and “for” respectively in English.

After all these steps we found out following matches:

d	r	l	x	j	a	g	e	b	v	q	m	h	n	t	s	i	p	u	k	w	o
e	i	a	s	m	y	o	u	n	t	h	v	r	f	d	b	c	p	q	w	g	l

Table 1: Mapping of the words in the encrypted text

Here, the first row corresponds to the ciphertext and the second row corresponds to the English text. After replacing, we get the following English text.

> *You have entered the first chamber of the caves. As you can see there is nothing of interest in the chamber. Some of the later chambers will be more interesting than this one, i am serious.The code used for this message is a simple substitution cipher in which digits have been shifted by **6** places. For this round password is given below, without the quotes.*
>
> *uvDc88Hhmd*

Now the text says that while encryption the digits are moved by 6. Since the text is also an encryption where we have only decrypted the English alphabet,

hence the digit six is also after encrypting. Let us assume that in the original text the digit present was x then as the text says that we have to change the digit by x, which means the digit in the encrypted text will be $2x$.

Hence $2x = 6 \Rightarrow x = 3$, which finally gives

> *You have entered the first chamber of the caves. As you can see there is nothing of interest in the chamber. Some of the later chambers will be more interesting than this one, i am serious.The code used for this message is a simple substitution cipher in which digits have been shifted by **3** places. For this round password is given below, without the quotes.*
>
> *uvDc55Hhmd*

Now we find that we don't have a corresponding English letter for letter "c" in ciphertext. But only remaining english letters whos mapping was not found were "j", "k", "x" and "z". We tried them one by one to substitute "c" in password and entered the password. It turned out that the password "uvDz55Hhmd" worked and we entered into Chapter 2.

Explanation of codes:-

1. The file input.txt contains the message to be deciphered.
2. The file preprocess.py removes all the characters except alphabets from the input.txt and it also changes the upper case letters into lower case letters and saves the final text into the file level_1.py.
3. The file Frequency_Analysis.py creates the frequency histogram for the text to be deciphered.
4. The file Index_of_Coincidence.py calculates the index of coincidence of the text.

Reference:-

1. *https://en.wikipedia.org/wiki/Frequency_analysis*
2. *https://en.wikipedia.org/wiki/Index_of_coincidence*