ASSIGNMENT 2

CS-642

TEAM- Bullshot

Members –

Aman Tayal(180074)

Rohit Ranjan(180629)

Rishabh Kothary(180608)

For this assignment we needed to crack the following cipher :

“Lg ccud qh urg tgay ejbwdkt, wmgtf su bgud nkudnk lrd vjfbg. Yrhfm qvd vng sfuuxytj "vkj_ecwo_ogp_ej_rnfkukf" wt iq urtuwjm. Ocz iqa jdag vio uzthsivi pqx vkj pgyd encpggt. Uy hopg yjg fhkz arz hkscv ckoq pgfn vu wwygt nkioe zttft djkth.”

The first thing we noticed that none of the two letter and three letter were repeated which suggested to us this might not be a substitution cipher and thought there was a chance that this cipher might be a Vigenere Cipher. The further make our claim more sound we used the help of a statistical concept of ‘coincidence ratio’. Coincidence ratio is the probability that if we picked two random variables what is the chance that they are equal. For an English text it is found to be greater than 6.8% and a random text around 3.8%. So for every possible key size(i) we found the coincidence ratio of the subsequence of the original text of the form = a_{ik}

where k is any non- negative integer and ik < Total size of the text.

By doing this we concentrate on letters that are added (mod 26) with the same letter of the key.

The formula for finding the coincidence ratio we used was

$$\sum({}^{f(x_i)}C_2 \,/\, {}^{n}C_2)$$

where x_i are the letters in the cipher text, $f(x_i)$ are the total number of occurrences of that letter in the cipher text and n is the total number of character in the cipher text.

The results were as follows:

The coincidence ratio for keysize 1 is 4.236%

The coincidence ratio for keysize 2 is 4.652%

The coincidence ratio for keysize 3 is 3.966%

The coincidence ratio for keysize 4 is 5.180%

The coincidence ratio for keysize 5 is 3.153%

The coincidence ratio for keysize 6 is 5.161%

The coincidence ratio for keysize 7 is 3.989%

The coincidence ratio for keysize 8 is 4.710%

The coincidence ratio for keysize 9 is 6.316%

The coincidence ratio for keysize 10 is 2.924%

The coincidence ratio for keysize 11 is 4.412%

The coincidence ratio for keysize 12 is 8.333%

The coincidence ratio for keysize 13 is 5.714%

The coincidence ratio for keysize 14 is 2.198%

The coincidence ratio for keysize 15 is 1.282%

The coincidence ratio for keysize 16 is 3.030%

The coincidence ratio for keysize 17 is 1.818%

The coincidence ratio for keysize 18 is 12.727%

The coincidence ratio for keysize 19 is 2.222%

The coincidence ratio for keysize 20 is 0.000%

The coincidence ratio for keysize 21 is 2.778%

The coincidence ratio for keysize 22 is 2.778%

The coincidence ratio for keysize 23 is 0.000%

The coincidence ratio for keysize 24 is 7.143%

The coincidence ratio for keysize 25 is 0.000%

We observed a sharp peak at keysize 18 suggesting that cipher maybe a vigenere cipher as the subsequences defined earlier seemed to represent English text. And then we checked the coincidence ratio of factors of 18. We see that keysize 6 and 9 also had reasonably good coincidence ratio suggesting they could also be possible candidates for possible key sizes.

First we tried keysize 6. For each three letter words we replaced it with the word ‘the’ and found the key accordingly that would lead to such a word if encrypted. But every result turned out to be negative and it is almost certain that every English passage uses the word ‘the’. We even used letter frequency to find possible candidates for the key but every result turned out to be negative. Then we tried keysize 9 and again used the strategy of replacing every 3 letter words with ‘the’ and find the corresponding key. A certain try seemed promising and by using the approach of hangman we moved forward and found complete key which was ‘kcgcdfcc’. The decrypted text was:

‘Be wary of the next chamber, there is very little joy there. Speak out the password "the_cave_man_be_pleased" to go through. May you have the strength for the next chamber. To find the exit you first will need to utter magic words there..’

The code we used to crack the cipher is attached in the file.