SIL765 : Networks and System Security

Assignment – 1

Functions defined in decipher_text.py :-

1. content(filename) :
   It reads the content of the file and outputs it in a string format, this is the ciphertext which we want to decrypt.

2. modify_cipherCharacters(cipher_characters):
   Returns a map which contains the characters which will be used in the modified ciphertext to be passed in the hill climbing algorithm along with the list of the new cipher characters, i.e., we have a lot of special characters in the ciphertext but our hill climbing algorithm works on comparing the scores which we get by the quadrams present in the text, so we need to replace all the characters by english alphabets and hence a mapping from the given cipher characters to english alphabets is done.

3. modify(modified_ciphertext):
   It removes the spaces from the cipher text as the hill climbing algorithm inputs the text without spaces.

4. modifyCiphertext(text, cipher_characters, modified_cipher_characters):
   Modifies the ciphertext by replacing all the special characters with the new_cipher_characters which are the letters of the english alphabet.

5. pureciphertext(text,cipher_characters):
   This function gives the pure ciphertext which contains only the characters which are given to us as ciphertext characters. This pure ciphertext is used to count the frequency of the single letters.

6. frequency_singleLetters(new_text, cipher_characters):
   Returns the sorted frequency map of single letters in the pure ciphertext.

7. unmodifiedKey(Key, modified_cipher_characters_map):
   Returns the key made up of the actual cipher characters given, not the charcaters of the english alphabets which we substituted for running the algorithm. For eg. If “$” was substituted which “h” then it returns the modified key in which “h” is written with its original cipher character “$”.

8. allInDictionary(deciphered_text, dictionary, plain_characters):
   Chekcs whether all the words in the decrypted text are present in the dictionary or not

9. decryptText(ciphertext, key , plain_characters, cipher_characters):
   Decrypts the ciphertext when the key is passed as the input.

10. Ngram_score :
    This class is used to calculate the score of the text using the quadgrams in the text. The score of a quadgram is = log(No. of times the quadgrams appeared in the text / Total no. of quadgrams in the text). The score of the text is the summation of the scores of quadgrams.

11. mainFunction (ciphertext) :
    Implemented the hill climbing algorithm to find the best possible match of the plaintext with English.

Algorithm Used :-

1) Modify the given ciphertext by mapping the given cipher characters with the english alphabets.

2) Now in that ciphertext the frequency of all the cipher characters is found and then according the the frequency of the letters in English language the cipher characters are mapped to the english letters and in this way we got our first key.

3) Now we have to apply the hill climbing algorithm to approach towards the best matching plaintext which has the highest quadgram score with respect to the english language.

4) So, we start with the key obtained in step (2) and initialise the maximum score to be very negative. This key is the parentKey.

5) Now we randomly shuffle two characters in the key to obtain a child key and then we calculate the score of the text decrypted using this child key. If it comes out to be more than the parent score then we will update the parentkey and update the parent key.

6) We will continue calculating the child even if it not better than parent for 1000 iterations and then we will compare the maximum parent score found till that date with the maximum score at that point.

7) If the max parent score is more than the maximum score then we will update the maximum score.

8) This loop will continue till the timeout which is 4 minutes or the time when all the words of the decrypted text are present in the dictionary, whichever comes first.

Analysis of :

1. **ciphertext-1.txt :**

   Mapping :

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
y	5	n	8	@	p	7	q	1	r	w	u	0	9	$	3	4	2	v	o	s	6	#	t	x	z

Plaintext :

india, officially the republic of india, is a country in south asia. it is the seventh largest country by area, the second most populous country, and the most populous democracy in the world. bounded by the indian ocean on the south, the arabian sea on the southwest, and the bay of bengal on the southeast, it shares land borders with pakistan to the west; china, nepal, and bhutan to the north; and bangladesh and myanmar to the east. in the indian ocean, india is in the vicinity of sri lanka and the maldives; its andaman and nicobar islands share a maritime border with thailand, myanmar and indonesia. good, now turn for the second part of the question, good luck!

2. **ciphertext-2.txt :**

Mapping :

```
a b c d e f g h i j k l m n o p q r s t u v w x y z
8 o t 6 4 s p n r x z q w y $ 1 9 3 v u 2 0 5 @ # 7
```

Plaintext :

defeated and leaving his dinner untouched, he went to bed. that night he did not sleep well, having feverish dreams, having no rest. he was unsure whether he was asleep or dreaming. conscious, unconscious, all was a blur. he remembered crying, wishing, hoping, begging, even laughing. he floated through the universe, seeing stars, planets, seeing earth, all but himself. when he looked down, trying to see his body, there was nothing. it was just that he was there, but he could not feel anything for just his presence.

Resource used :-

http://practicalcryptography.com/cryptanalysis/stochastic-searching/cryptanalysis-simple-substitution-cipher/

Submitted By :-

--> Nikita Bhamu

--> 2018CS50413