Comp3260 Assignment 1

Jay Rovacsek and Cody Lewis

April 11, 2018

1 c1 Ciphertext

c1 was found to be a Viginere cipher with the key: 'remember'.

1.1 Analysis

First the frequency graph was observed, which expressed that the cipher was produced by substitution. Next the IC was calculated giving an approximate period, d = 3. Then a Kasiski of 17 characters shown that the string 'agxyvyzmffhjkefie' occurs in the text 3 times with the gaps 392 and 2024.

$$GCD(2024, 392) : 2024 = 5 \times 392 + 64$$

$$GCD(392, 64) : 392 = 6 \times 64 + 8$$

$$GCD(64, 8) : 64 = 8^2 + 0$$

$$\Rightarrow GCD(2024, 392) = 8$$

Hence the cipher definitely has 8 alphabets. Looking at the frequency graph with 8 alphabets shows that the cipher is Viginere as the letter frequency match the normal graph left to right, they only shift. The key is the letter a is shifted to in each alphabet, which means the key is 'remember'.

2 c2 Ciphertext

c2 was found to be a Beauford cipher with the key: 'triumph'.

2.1 Analysis

The frequency graphs shows that a substitution cipher was used, the IC give an approximate period, d = 10. A Kasiski at 10 gives the gaps of 287, 91.

$$GCD(287, 91) = 7$$

Hence there are definitely 7 alphabets in the cipher. Solving for the key to this cipher is very similar to the previous one, only the frequency graph is backwards meaning it is a Beauford cipher. Matching each letter to the letter in place of a in each alphabet, gives the key 'triumph'.

3 c3 Ciphertext

c3 was found to be a General substitution cipher with the key: 'mcafrgyjhkdplqiutvexbzwsno'

3.1 Analysis

General substitution was determined by the IC being analysed over 126 alphabets and returning an average of 0.06836808. Leading to the assumption that only a single alphabet had been used.
Furthermore, flags making c3 seem like a substitution cipher were the high counts of strings between 3 4 letters which stood out very abnormally from the remaining distribution.

3.2 Method of breaking

A boon to the breaking of the substitution cipher was Simon Singh's page [1] on cracking the substitution cipher and likely trigrams and digrams to occur within the English language. Moving forward with the assumption that the most common letters of the ciphertext were likely to mimic that of English, it was determined that:

$$the = xjr$$

$$and = hqy$$

However after a number of hours attempting to find words containing a mixture of the two strings, it was obvious that 'and' did not match the assumed ciphertext and instead another common three letter string that had occurred. It was still assumed 'the' matched the ciphertext proposed as the distribution was skewed such that it would be highly unlikely 't' could have been anything else.
Eventually a key was built based on comparison of ciphertext elements that were assumed to be certain words and if incorrect another key:value was attempted instead.

3.3 Second Analysis

On finding that

$$the = xjr$$

the cipher was checked as to whether it was an affine transformation. The following equations were formed:

$$e \mapsto r \Rightarrow (4k_0 + k_1) \pmod{26} = 17 \tag{1}$$

$$h \mapsto j \Rightarrow (7k_0 + k_1) \pmod{26} = 9 \tag{2}$$

$$t \mapsto x \Rightarrow (19k_0 + k_1) \pmod{26} = 23 \tag{3}$$

Using Gaussian elimination for (1) (2) brings the equation:

$$-3k_0 \pmod{26} = 8 \Rightarrow 23k_0 \pmod{26} = 8 \tag{4}$$

$$GCD(26, 23) : 26 = 1 \times 23 + 3$$
$$GCD(23, 3) : 23 = 7 \times 3 + 2$$
$$GCD(3, 2) : 3 = 2 \times 1 + 1$$
$$\Rightarrow 1 = 3 - 2$$
$$\Rightarrow 1 = 3 - (23 - 7 \times 3)$$
$$\Rightarrow 1 = 8 \times 3 - 23$$
$$\Rightarrow 1 = 8(26 - 23) - 23 = 8 \times 26 - 9 \times 23$$
$$\therefore k_0 = [-9] \times 8 \pmod{26} = 6 \tag{5}$$

Subbing (4) into (3) brings:

$$k_1 = 23 - (19 \times 6 \pmod{26}) = 13 \tag{6}$$

But these concluded values do not work for equations (1) and (2), hence c3 is not an affine transformation.

4 c4

c4's IC indicated that it is monoalphabetic, and its frequency graph matches that of normal English, hence c4 is a Transposition cipher. The first 3 letters of the cipher text is 'hte' according to digraphs is probably the message 'the', thus the cipher is a row Transposition. Through some anagramming further through the message the words 'enigma machine was' were found. After a bit of evaluation of the pattern against ciphertext and message the key '2,1,4,6,5,4' or 'bafced' was found to decipher c4 into the message.
The key was determined to have a length of x where:

$$x \pmod{6} = 0$$

This was determined via Kasiski analysis performed on the ciphertext resulting in strings 'ahcmni' and 'cyretp' occurring far more often than should be expected. Analysing the initial position and repeated position the key length was estimated with more accuracy given:

$$GCD(492, 414) = 6$$

References

[1] Simon Singh, *The Black Chamber Hints and tips*, `http://simonsingh.net/The_Black_Chamber/hintsandtips.html`

[2] Lawrie Brown based on code by Daryl Bossert, *jkypto* The program used for most of the statistical analysis, decryption and Kasiski calculations.