A history of cryptography and cryptanalysis

Christopher Colahan
Simpson College

February 15, 2017

Contents

All ciphers can be defined as two functions, one called the enciphering function that takes a plaintext as input and outputs a ciphertext, and one called a deciphering function that takes a ciphertext as input and outputs the plaintext.

For convenience, some common notation is used:

p is the plaintext.
c is the ciphertext.
k is the secret key.
$E(p)$ is the enciphering function.
$D(c)$ is the deciphering function.

1 Transposition Ciphers

A transposition cipher is a permutation of the plaintext.

2 Monoalphabetic Substitution Ciphers

2.1 Shift Ciphers

Shift ciphers work by shifting the symbols in the plaintext by an amount. For example, if we are using the English alphabet, then there are $n = 26$ possible symbols. We could then choose some k, $0 < k < n$ for our key. In our notation, this would look like

$$E(p_i, k) = p_i + k \pmod{n}.$$

(possibly cite abs alg textbook here?) To get the deciphering function, we shift backwards:

$$D(c_i, k) = c_i - k \pmod{n}.$$

2.2 Frequency Analysis

Shift ciphers are broken by frequency analysis. Below is a table that shows the letter frequency from a sample of English text.

Letter	Percentage
a	8.2
b	1.5
c	2.8
d	4.3
e	12.7
f	2.2
g	2.0
h	6.1

Letter	Percentage
i	7.0
j	0.2
k	0.8
l	4.0
m	2.4
n	6.7
o	7.5
p	1.9

Letter	Percentage
q	0.1
r	6.0
s	6.3
t	9.1
u	2.8
v	1.0
w	2.4
x	0.2
y	2.0
z	0.1

[1, pg. 19]

2.3 Homophonic Substitution Ciphers

A homophonic substitution cipher is a substitution cipher that maps each symbol to one of more symbols in order to prevent frequency analysis from being used.

For example, suppose we have 100 symbols $S = \{s_1, s_2, ...s_100\}$. The the letter e would map to approximately 12 or those symbols, but the letter a would only map to about 8 of those symbols.

Before enciphering, each letter is replaced at random with one of the symbols it maps to. This means that each symbol in the ciphertext only appears with a frequency of about 1%.

3 Polyalphabetic Substitution Ciphers

A polyalphabetic substitution cipher uses multiple monoalphabetic substitution ciphers to generate more possibilities for the ciphertext.

3.1 Vigenère Cipher

The vigenère cipher uses 26 alphabets to encrypt plaintext. A key is also used that consists of a string of symbols. Given a plaintext symbol p_i and a key symbol k_j, the ciphertext symbol c_i is the character in the i column and j row. (insert vigenère square here) (insert example here)

3.2 One Time Pad

The one time pad is a special vigenère cipher where

- the key is the same length as the plaintext,
- the key is random, and
- the same key is not used to encrypt two different plain texts.

References

[1] Simon Singh. *The Code Book*.