

# COMPUTER PROGRAMMING IN JAVA

COLUMBIA UNIVERSITY HIGH SCHOOL SCIENCE HONORS PROGRAM

Intro to Cryptology: Code Breaking

Lecture

2008 May 03 Sat

## Shift Ciphers

The shift cipher is arguably the simplest possible form of encryption. Our encoder/decoder is simple:

```
Plain:  ABCDEFGHIJKLMNOPQRSTUVWXYZ  
Cipher: DEFGHIJKLMNOPQRSTUVWXYZABC
```

The above shows a shift-4 cipher. When we want to encode a message, we simply find the index of the plaintext letter, and use as the letter at the same position in the cipher array as the ciphertext letter.

Legend has it that Caesar used a shift-3 cipher to communicate with his generals.

The cryptanalysis of shift ciphers is very easy, given that we know the language of the plaintext. Knowing the language is the key to breaking the code because the distribution of letters within a language is highly non-uniform. For example, in English, e is the most common letter by far, with t a strong second. Thus all we have to do is perform frequency analysis on the ciphertext, and make educated guesses about which letters map to which.

## Modified Shift Ciphers

There are numerous ways to extend the basic idea of the shift cipher. The method we will discuss here we call the “keyed” and “reversed” variant. Instead of picking a shift value  $k$  as our key, we instead pick a keyword. Starting at the beginning of the alphabet, we add each letter of the keyword (removing duplicates), followed the remaining letters of the alphabet, in reverse order. A figure illustrates this much better, where the keyword is “CANNON”:

```
Plain:  ABCDEFGHIJKLMNOPQRSTUVWXYZ  
Cipher: CANOZYXWVUTSRQPMLKJIHGFEDB
```

## Vigenère Ciphers

The Vigenère cipher is one of the simplest “polyalphabetic” cipher (guess what a shift cipher is then) that is immune to frequency analysis. The idea is simple: instead of using a single shift (or keyword) value, we generate a table of all (26) possible shift ciphers.

Let’s say we have a message, “I am bananas for SHP”, and a secret keyword, “false”. I

take my plaintext, and repeat as many instances of the keyword as necessary:

Plain: IAMBANANASFORSHP  
Cipher: FALSEFALSEFALSEF

Now to encrypt my message, I look up the row in the key, and the column of the plaintext letter (clearly to decrypt I do the reverse, i.e. I look up the row in the key, and look for the letter in row matching my ciphertext letter). The power of this technique is that the same plaintext letter is no longer mapped to the same ciphertext letter, rendering (direct) frequency analysis ineffectual.

Of course, this doesn't mean that the Vigenère cipher is unbreakable; the weakness is to note that it is still possible for there to be repeats in the ciphertext. Consider the following examples:

Key: ABCDAB CD ABCDA BCD ABCDABCDABCD  
Plaintext: **CRYPTO** IS SHORT FOR **CRYPTOGRAPHY**  
Ciphertext: **CSASTP** KV SIQUT GQU **CSASTPIUAQJB**

Ciphertext: **DYDUXRMHTVDVNQDQNWQDYDUXRMHARTJGWNQD**

(from [http://en.wikipedia.org/wiki/Vigenère\\_cipher](http://en.wikipedia.org/wiki/Vigenère_cipher))

This cryptanalysis technique is known as the Kasiski examination.