Birkbeck
UNIVERSITY OF LONDON

Information Security
Information & Network Security
Lecture 4

David Weston

Birkbeck, University of London
Autumn Term

Birkbeck
UNIVERSITY OF LONDON

Cryptography- Continued

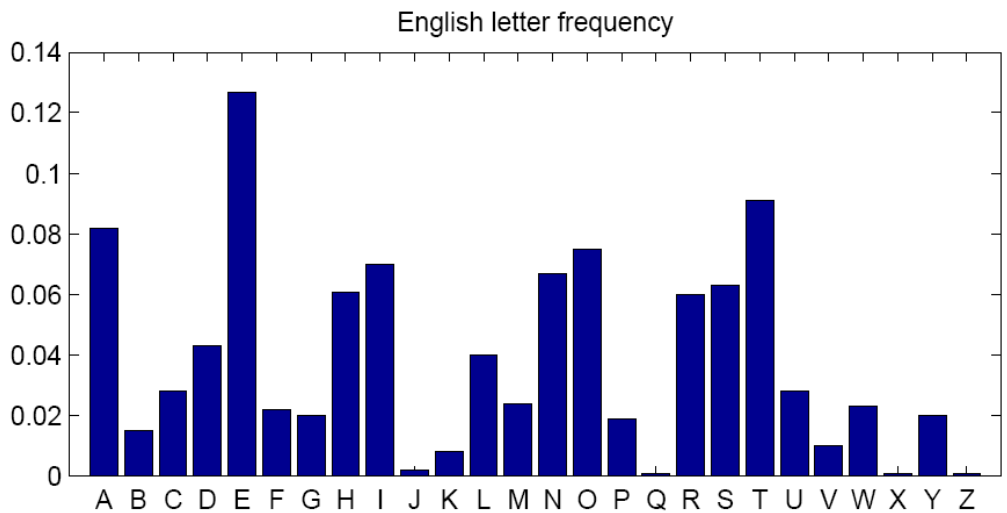
Monoalphabetic Substitution Ciphers

- Caesar cipher is a special case of a *simple substitution cipher* (or *monoalphabetic cipher*)
- Instead of just shifting characters, the alphabet can be mapped to a (random) permutation:

```
abcdefghijklmnopqrstuvwxyz
SECURITYABDFGHJKLMNOPQVWXZ
```

Monoalphabetic Substitution Ciphers (2)

- Monoalphabetic ciphers are not very secure and can be easily broken by statistical means:
  - Different characters have typical frequencies in languages



- There have been various attempts at making substitutions more secure

Homophonic Substitution Ciphers

- *Homophonic substitution ciphers* try to obscure the frequencies by mapping a character to more than one code
  - For example, “A” could correspond to 5, 13, 25, or 56; while for “B” this could be 7, 19, 32, or 42
- While this makes analysis a bit harder, it doesn’t hide all statistical properties
- With the help of a computer can usually be broken in a few seconds

Polygram Substitution Cipher

- Instead of encoding single characters, a *polygram substitution cipher* encrypts groups of letters
  - For example, “ABA” could correspond to “RTQ’, while “ABB” could correspond to “SLL”
  - The example above uses 3-grams (groups of 3 letters); can be generalized to n-grams
- Still not a very secure way of encrypting data:
  - This hides the frequencies of individual letters
  - However, natural languages also show typical frequencies for n-grams (although the curve is flattened)

Polyalphabetic Substitution Cipher

- A *polyalphabetic substitution cipher* uses multiple simple substitution ciphers
- The particular one used changes with the position of each character of the plaintext
  - There are multiple one-letter keys
  - The first key encrypts the first letter of the plaintext, the second key encrypts the second letter of the plaintext, and so on
  - After all keys are used, you start over with the first key
  - The number of keys determines the period of the cipher

Polyalphabetic Substitution Cipher (2)

- An early example is the Vigenère cipher
- Adds a key repeatedly into the plaintext using numerical codes: A=0, B=1, . . . , Z=25
- This is done modulo 26, i.e. if the result is greater than 26, then we divide the result by 26 and take the remainder

$$C = M + K \mod 26$$

- For example, P(15) + U(20) = 35 mod 26 = 9

Plain:	`tobeornottobethatisthequestion`
Key:	`runrunrunrunrunrunrunrunrunrun`
Cipher:	`KIOVIEEIGKIOVNURNVJNUVKHVMGZIA`

Polyalphabetic Substitution Cipher (3)

- Unfortunately, the Vigenère cipher is also not very secure
- Code can be broken by analyzing the period
  - Looking at the example from previous slide, we notice that KIOV is repeated after 9 letters, NU after 6 letters
  - As 3 is a common divisor of 6 and 9, this is a hint that the period could be 3
  - Knowing which letters were encoded with the same key allows the application of frequency methods again

Rotor Machines

- In the 1920s, mechanical encryption devices were developed (to automate the process)
- Basically, these machines implemented a complex Vigenère cipher
- A *rotor* is a mechanical wheel wired to perform a general substitution
- A *rotor machine* has a keyboard and a series or rotors, where the the output pins of one rotor are connected to the input of another
- For example, in a 4-rotor machine, the 1st rotor might substitute A $\rightarrow$ F, the 2nd F $\rightarrow$ Y, the 3rd Y $\rightarrow$ E, the 4th E $\rightarrow$ C; so the final output for A is C
- After each output, some of the rotors shift

Rotor Machines (2)

- The best-known rotor machine is the Enigma (pictured below)

Rotor Machines (3) - Enigma Machine

- Combination of several rotors and shifting of rotors leads to a period of 26^n (where n is the number of rotors)
- A long period makes it harder to break the code

Transposition Ciphers

- Substitution ciphers on their own are usually not very secure
- That is why they are combined with transposition ciphers
  - Transposition ciphers on their own are also not very secure
- In a *transposition cipher* the symbols of the plaintext remain the same, but their order is changed

Simple Columnar Transposition Cipher

- In a *simple columnar transposition cipher*
  - the plaintext is written horizontally onto a piece of graph paper of fixed width
  - while the ciphertext is read off vertically

W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C
E					

WIREEESEAACDTROFOEVLNDEEC

Double Columnar Transposition

- A simple columnar transposition can be broken by trying out different width/column lengths
- Putting the ciphertext through a second transposition enhances security
  - In contrast to a simple substitution cipher, where a second application does not increase security

Stream Ciphers/Block Ciphers

- In symmetric algorithms we also distinguish between stream ciphers and block ciphers
- *Stream ciphers* operate on the plaintext a single bit (or single character) at a time
  - Simple substitution cipher is an example
- *Block ciphers* operate on groups of bits (or groups of characters)
  - An example of an early block cipher is Playfair

Playfair

- Place the alphabet in a 5x5 grid, permuted by the keyword (omitting the letter J; J=I)

P	A	L	M	E
R	S	T	O	N
B	C	D	F	G
H	I	K	Q	U
V	W	X	Y	Z

- Then divide the plaintext into pairs of letters
  - preventing double letters in a pair (by separating them with an ‘x’)
  - adding a ‘z’ to the last pair (if necessary)

Playfair (2)

- For example, “Lord Granville’s letter” becomes “lo rd gr an vi lx le sl et te rz”
- Then encrypt text pair by pair
  - Replace two letters in the same row or column by succeeding letters, e.g. “am” $\rightarrow$ “LE”
  - Otherwise the two letters are in opposite corners of a rectangle, replace them with letters in the other corners, e.g. “lo” $\rightarrow$ “MT”

P	A	L	M	E
R	S	T	O	N
B	C	D	F	G
H	I	K	Q	U
V	W	X	Y	Z