System Security

Cryptography - Intro

Giovanni Russello

g.russello@auckland.ac.nz

http://www.cs.auckland.ac.nz/compsci725s2c/

Cryptography

- Security Engineering meets Mathematics
- Key enabling technology for protecting systems
- Very hard to do it right
  - Protecting the wrong assets
  - Protecting assets in the wrong way

Basic Terminology

- *Cryptography*: science and art of designing cypher algorithm
- *Cryptanalysis*: science and art of breaking cypher algorithm
- *Cryptology* (or *Crypto*): the study of both Cryptography and Cryptanalysis

Basic Terminology II

- *Cryptographic Primitives*: basic building blocks
  - *Block ciphers, stream ciphers, hash functions*
- *Symmetric, Shared-key, Secret-key Encryption*: one key for both encryption and decryption
- *Public-key*, *Asymmetric*: different keys for encryption and decryption
- *Cleartext, plaintext*: input to an encryption block
- *Ciphertext*: the output of an encryption block

What is the use of Crypto?

- Authorisation: control access to information
  - Secret-key and Private-key
- Authentication: verify the identity of an entity
  - Public-key
- Integrity: guarantee the message has not been tampered with
  - Hash function
- Authenticity of information: guarantee the information is not fake
  - Hash function

Historical Background

- Julius Caesar used to encrypt the dispatch to his legions: ‘D’ for ‘A’, ‘E’ for ‘B’

Plaintext: `abcdefgh…`

Ciphertext: `DEFGHIJK…`

- It can be generalised with the use of a keyword

Plaintext: `abcdefghijklmno…`

Ciphertext: `702ABCDEFGHIJKL…`

Monoalphabetic cypher

- Letters and combinations have a known frequencies for a given language
- It is not difficult to recover the original message once you have access to enough cypher text
  - In average, 600 letters of ciphertext are enough for breaking this type of code

Stream Cipher

- An improvement over Monoahphabetic
- The encryption rule depends on the actual position of a symbol in the plaintext
- The Vigenere:
  - Each letter correspond to a number (A=0, B=1,…)
  - Selecting a key and repeating it for the length of the plaintext
  - Adding the two strings in modulo 26
    - P(15) U (20) -> 15+20 = 35 -> 35 mod 26 = 9 -> J

Vigenere In Action

```
PT:    tobeornottobethatisthequestion
Key:   runrunrunrunrunrunrunrunrunrun
C:     KIOVIEEIGKIOVNURNVJNUVKHVMGZIA
```

Vigenere Issue

```
PT:    tobeornottobethatisthequestion
Key:   runrunrunrunrunrunrunrunrunrun
C:     KIOVIEEIGKIOVNURNVJNUVKHVMGZIA
```

With enough ciphertext, repeating patterns will appear

Vigenere Issue

```
PT:    tobeornottobethatisthequestion
Key:   runrunrunrunrunrunrunrunrunrun
C:     KIOVIEEIGKIOVNURNVJNUVKHVMGZIA
```

KIOV is repeated after 9 letters

Vigenere Issue

```
PT:    tobeornottobethatisthequestion
Key:   runrunrunrunrunrunrunrunrunrun
C:     KIOVIEEIGKIOVNURNVJNUVKHVMGZIA
```

NU is repeated after 6 letters

Vigenere Issue

```
PT:   tobeornottobethatisthequestion
Key:  runrunrunrunrunrunrunrunrunrun
C:    KIOVIEEIGKIOVNURNVJNUVKHVMGZIA
```

- We might guess the password length is 3 given it divides both 6 and 9.
- This also tells us that cyphertext symbols 1,4,7 and so on are encrypted under the same key symbol
- Again, using the frequency analysis we can guess which letters they represent
- Repeat the process for the second and third key symbols

One-time Pad

- To make stream ciphers more robust against patterns the one-time pad can be used
- Make the key sequence as long as the plaintext and never use it again
- Given a ciphertext and any plaintext of the same length there is always a key that decrypt the ciphertext to the plaintext

OTP in action

```
PT:    heilhitler
Key:   wclnbtdefj
C:     DGTYIBWPJA
```

The key material is destroyed after performing the encryption

OTP in action

Let us assume we intercept the cyphertext C

```
C:      DGTYIBWPJA
Key:    ??
PT:     ??
```

What is the plaintext?

OTP in action

The plaintext could be easily ‘Hang Hitler’

```
C:     DGTYIBWPJA
Key:   ??
PT:    hanghitler
```

OTP in action

By changing the key a new plausible plaintext can be generated from the ciphertext

```
C:     DGTYIBWPJA
Key:   wggsbtdefj (wclnbtdefj)
PT:    hanghitler
```

OTP in action

The ciphertext can also be changed in transmission

```
C:     DCYTIBWPJA
Key:   wclnbtdefj
PT:    hanghitler
```

Using the original key the plaintext message is changed.

One-Time Pad Properties

- The One-Time Pad offers ***perfect secrecy***
- According to Shannon:

"A cipher block has perfect secrecy if and only if there are as many possible keys as possible plaintexts, and every key is equally likely."

- CONS:
  - it does not offer message integrity
  - It is quite expensive as it uses the same amount of key material as plaintext

Strengthening Stream Ciphers

- Stream cyphers use pseudorandom generators to expand a short key into a long *keystream*
- The plaintext is then encrypted by performing a bit-wise xor with the keystream
- Stream ciphers are usually used in hardware because require less gates
  - GSM encryption using A5 algorithm

Block Ciphers

- Block ciphers operate on a fixed-length group of bits
- They operate by performing permutations on the bit blocks
- Usually better suited for software implementation

The Playfair BC

- The plaintext is prepared by
  - splitting in group of two letters
  - Replacing any ‘J’ with an ‘I’
  - Separating any double with a ‘x’
- Then the plaintext is permuted using the table below and applying two rules:

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

The Playfair Permutation Rule 1

- If a group of two letters is in the same row or column they are replaced by the succeeding letters.

For instance the group ‘am’ is encrypted to ‘LE”

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

The Playfair Permutation Rule 2

- Otherwise the two letters form a rectangle: we replace with the letters from the opposite corners

For instance the group ‘pf’ is encrypted to ‘MB”

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

Playfair issues

- The ciphertext looks random but changing one letter of the plaintext causes only one letter of the ciphertext to change as well
- Given enough ciphertexts the table can be reconstructed
- Also the size of the blocks is very small
  - It is possible to use the frequencies of the digraphs (letter pairs)

Pseudorandom and Random Oracles

- Building a cipher requires the **random** property to be satisfied under a given model
- Ciphers should be indistinguishable from a *random* function
- However, ciphers are algorithms built as a circuit or program: their output should look “random”
- A cipher is pseudorandom if its output is indistinguishable from that of a Random Oracle

Resources

- Security Engineering – Ross Anderson
- Chapter 5: http://www.cl.cam.ac.uk/~rja14/Papers/SEv2-c05.pdf