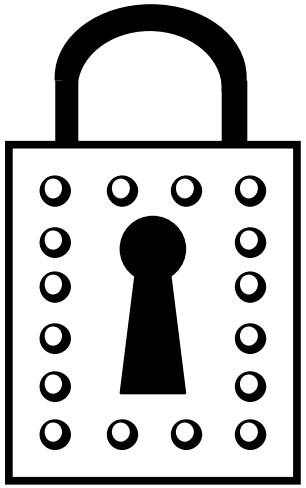
Introduction to Cryptographic Terminology



License

Cryptography

- The term *cryptography* has two parts: *crypt* and *-graphy*
- The word *crypt* comes from the Greek word *kryptos*
  - Means hidden or covered
  - A *crypt* is an underground burial place or a secret meeting place
  - The word *cryptic* means “difficult to understand”, that is, a “hidden meaning”
- The *-graphy* suffix denotes a process or science for drawing, writing, representing, describing and so on
  - Biography, choreography, geography, photography, typography, …
- So, cryptography is the science of hidden writing
  - To *encrypt:* to turn a plaintext message into a hidden message
  - To *decrypt:* to turn a hidden message back into a plaintext message

Cipher

- The Arabic number system had some important innovations:
  - Arithmetic is much simpler than arithmetic with Roman numbers
  - The concept of zero (*sifr* in Arabic) has two meanings:
    - It denotes “nothing”
    - It denotes an order of magnitude (10, 100, 1000, …)
- Initially, Europeans were confused by the concept of zero:
  - So *cipher* (*sifr*) was used to refer to something that was a mystery
  - The word *cipher* evolved to mean the deliberate hiding of meaning
- So, *cipher* and *cryptography* are almost synonyms
  - To *encipher* means to *encrypt*
  - To *decipher* means to *decrypt*

Cipher (cont')

- A *cipher* is an algorithm that enciphers and deciphers text
  - Many ciphers take a secret *key* that controls the algorithm
- Example of a (very simple to break) cipher:
  - Algorithm is: rotate each letter "N" places
  - "N" is the secret *key*
    - If "N" is 1 then A → B, B → C, …, Y → Z, Z → A
    - If "N" is 2 then A → C, B → D, …, Y → A, Z → B
- Knowing the cipher is not enough to decode a message
  - You also need to know the key that was used to encode the message

Plaintext and ciphertext

- The term *plaintext* means a readable message:
  - The message does *not* have to be text-based
  - It might be a graphic file or an audio file instead
- Conversely, *ciphertext* means an encrypted message

Security though obscurity

- Security through obscurity:
  - Develop your own cipher (probably with a hardcoded key)
  - Nobody else knows your cipher's algorithm (so it is obscure)
  - You mistakenly think your secrets are safe
- The flaw in is that there are always people smarter than you
  - Smarter people are likely to find flaws in your cipher
  - So they can decode all your secret messages

Well known ciphers

There is a better approach…

- When somebody invents a cipher, he publishes the details:
  - Mathematicians around the world test the cipher for flaws
  - If no flaws can be found then everybody has confidence in the cipher
- A *strong* (that is, good) cipher can be broken only by trying every single possible key value
  - If there, say, 10^{70} possible keys then this approach might take thousands of years of computer time
- All you need to do is:
  - Pick a key (at random) to use with the cipher
  - Keep the key secret

What about Moore's Law?

- In 1965, Gordon Moore (co-founder of Intel) made an observation:
  - Advances in technology mean you can put twice as many transistors onto a chip every 18 months
  - This observation has remained true for over 40 years
- Doubling the transistors usually means doubling the computational power
  - In 15 years time, computers will be 1000 faster than they are today
  - A cipher that takes 1000 years to crack today will take only one year to crack in 15 years' time
- There is no need to panic, because:
  - Most of today's secret messages will be worthless in 15 years' time (so it will not be worthwhile for somebody to crack them)
  - Better ciphers will be developed within 15 years

Key length

- Many ciphers consist of:
  - A well known algorithm, and…
  - A *key* (a number used to prime the algorithm)
- Example of an easy-to-break cipher:
  - Algorithm: rotate each letter "N" places
  - Key: a value in the range 1 to 26
- *Key length* is the number of bits used to represent a key
  - For example, a key length of 128 implies (at most) 2^{128} possible values
  - A value in the range 1 to 26 can be represented in 5 bits (2^5 is 32)
- There is a tradeoff. Longer keys:
  - Make the cipher more secure with only a little extra overhead (which is good)
  - Require more storage space and transmission bandwidth (which is bad)

Key length (cont')

- Some countries impose legal restrictions on key lengths
  - The term "export cipher" refers to a cipher used with a short key
- The intention is as follows:
  - Strong encryption can be used by the military
  - We do not want to allow strong encryption technologies to be used by foreign militaries (possible enemies)
  - So, we allow only weaker encryption to be exported
- Export restrictions on encryption make international e-commerce more difficult

Misplaced concern about public ciphers

- "Won't mathematicians working for the <such-and-such> government keep silent about flaws so they can decode your secret messages?"
- No, because:
  - A flaw is likely to be spotted by several people in different countries
    - So keeping silent about flaws just lets somebody else get the credit
  - E-commerce (which is *huge*) cannot work without reliable ciphers
    - E-commerce is not just buying books from Amazon
    - It is also online banking
    - And stock trading
    - And business-to-business transactions
  - E-commerce is driving advances in cryptography much more than the <such-and-such> government's attempts at spying

Checksums

- When data is being transmitted, it might become corrupted
  - For example, there might be noise on the transmission line
- A checksum is a way to detect accidental corruption. Example algorithm:

```
int calculateChecksum(char* data, int size)
{
    int result = 0;
    for (int i = 0; i < size; i++) {
        result += (unsigned int)data[i];
    }
    return result;
}
```

- Sender transmits data plus the checksum value
  - Receiver calculates checksum value for received data
  - Compares this to the transmitted checksum

Message Authentication Code (MAC)

- A checksum is not infallible, but it is a useful check
- A checksum can guard against *accidental* corruption
  - But is too simplistic to guard against deliberate corruption
- A *message authentication code* (MAC) is a kind of checksum:
  - The checksum value is encoded in, say, 20 bytes instead of just 4
  - And it is encrypted
  - These properties make a MAC unfeasibly difficult to deliberately fake
- A MAC is one of several ingredients used to ensure secure communication:
  - A cipher ensures nobody else can understand a secret message
  - A MAC ensures nobody can modify a secret message

8. Summary

Summary

- This chapter has introduced some basic terminology:
  - Cryptography, cipher
  - Encrypt = encipher; decrypt = decipher
  - Plaintext and ciphertext
  - Key length
  - Message authentication code (MAC) is an encrypted checksum
    - Used to detect tampering of messages
- Also explained:
  - Why security through obscurity is a bad idea
  - Well known ciphers tested by mathematicians worldwide are better
  - Lots of people can rely on the same cipher; but they use different (secret) keys