Topic 2

Introduction to Cryptography

Introduce the basic concepts of cryptography, and some classical techniques

Overview

- ❑ What is Cryptography?
- ❑ Terminology
- ❑ Classical encryption techniques
- ❑ Cryptanalytic attacks
- ❑ Conclusion

source: chapter 3 in Cryptography and Network Security

What is Cryptography?

- **Cryptography** is "the art of keeping messages secure" by Schneier.

Keep messages secure

Prevent unauthorised people from gaining access to the network

Scramble the message so that it can't be understood by unauthorised people.

Cryptography is used.

Terminology

- Cryptography: practice and theory of concealing text.
- Plaintext or cleartext: a message in its original form.
- Ciphertext: a message in an encrypted form.
- Encryption: code a message to hide its meaning.
- Decryption: convert an encrypted message back to its original form.
- Other terms: encode and encipher for encryption, and decode and decipher for decryption.
- Cipher/Cryptosystem: the system that performs encryption and decryption.
- Cryptanalysis: attempts to discover plaintext or key.

Classical Encryption Techniques

- Classical (historical) algorithms are based on substitution & permutation.
- Substitution -> Confusion
  - E.g. 'a' becomes 'b'
- Transposition/Permutation -> Diffusion
  - E.g. 'abcd' becomes 'dacb'
- XOR operator
- Simple/non-secure ciphers
  - Shift Cipher – Caesar Cipher,
  - Vigenere Cipher, etc ….
- Secure cipher
  - One-Time Pad

•Modern ciphers use substitution technique: take in N bits and output a different set of N bits using a lookup table, called S-Boxes.

•Modern ciphers use transposition technique: they permute N bits using a lookup table, called P-Boxes.

Classical Encryption Techniques - Caesar cipher

- **Caesar Cipher** (Shift cipher)
  - It uses simple substitution - each letter is translated to the letter a fixed number of letters after it in the alphabet.
  - The operation could be expressed using addition modulo 26.
    - The message must be a sequence of letters, each letter is identified with a number.
    - The key k is a number in the range 1 … 25.
    - Encryption/decryption involve $\pm$ k to each letter (mod 26).

$C_i = E_k(M_i) = E(k, M_i) = (M_i+k) \bmod 26.$

$M_i = D_k(C_i) = D(k, C_i) = (C_i-k) \bmod 26.$

mod n

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Classical Encryption Techniques - Caesar cipher

For example,

Plaintext:	treaty impossible
Key	± 3
Ciphertext:	wuhdwb lpsrvvleoh

That is, $C_i = E[3, M_i] = M_i + 3 \bmod 26$.

Classical Encryption Techniques - Caesar cipher

- Brute-force attack (or exhaustive key search) is by trying all possible keys
- The three characteristics which make brute-force attack practical:
  - The encryption and decryption algorithms are in public domain.
  - There are only 25 keys to try.
  - The language of the plaintext is easily recognisable (e.g. compressed text not).
- Given a small number of plaintext-ciphertext pairs encrypted under a key K, K can be recovered by exhaustive key search with 2^{n-1} processing complexity (where n is the bit-length of the key).
- With today's computing power, (symmetric) key length should be at least 128 bits.
- If the plaintexts are known to contain redundancy, then ciphertext-only exhaustive key search is possible with a relatively small number of ciphertexts.
- Also vulnerable to another form of attack - frequency distribution analysis of language letters.

Classical Encryption Techniques - Frequency Analysis

- **Letter Frequency Distribution in English (in percentage)**
  (of course this may vary depending on the content/size of the text)

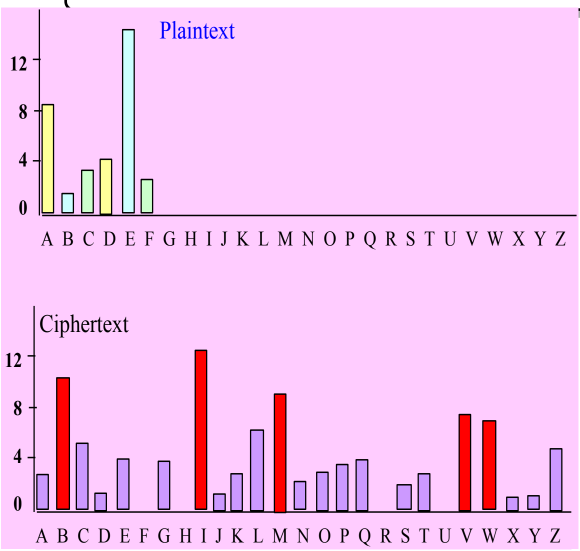
a	b	c	d	e	f	g	h	i
8.2	1.5	2.8	4.2	12.7	2.2	2.0	6.1	7.0
j	k	l	m	n	o	p	q	r
0.1	0.8	4.0	2.4	6.7	7.5	1.9	0.1	6.0
s	t	u	v	w	x	y	z	
6.3	9.0	2.8	1.0	2.4	2.0	0.1	0.1	

Classical Encryption Techniques - Frequency Analysis

- Take the following ciphertext as an example:

  bpmzm wvkm eia iv cotg lckstqvo eqbp nmibpmza itt abcjjg ivl jzwev ivl bpm wbpmz jqzla aiql qv aw uivg ewzla omb wcb wn bwev omb wcb, omb wcb, omb wcb wn bwev ivl pm emvb eqbp i yciks ivl i eilltm ivl i ……...

Frequency Analysis - how to do it



- Now let us to guess the key! K=????
- Look at 'E'
  - 4, 8, 17, 18 or 23
- Look at 'A'
  - 1, 8, 12, 21, or 22
- Ah…I guess **K = 8**?
- How can you solve this problem???

Classical Encryption Techniques - Frequency Analysis

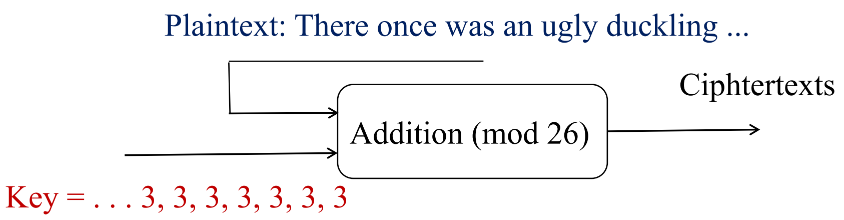
- **Now let us see how to break this ciphertext:**
  - it retains details about the **word lengths** of the underlying plaintext - this is valuable for cryptanalysis so in real-life, word breaks are removed prior to encryption.
  - Compute the frequencies of the letters in the ciphertext;
  - Compare them with the English letter frequencies; and
  - Try to deduce the plaintext by substituting letters by the most probable one……. , then you can work out the plaintext is:

  There once was an ugly duckling
  With feathers all stubby and brown…….. (you do the rest!)

A Quick Question

(a) Below is a diagram illustrating Caesar Cipher encryption operation. Could you propose a (simple) solution to hide letter frequency distributions in plaintexts, so that, from ciphertexts, the frequency distributions in plaintexts are not so obvious.

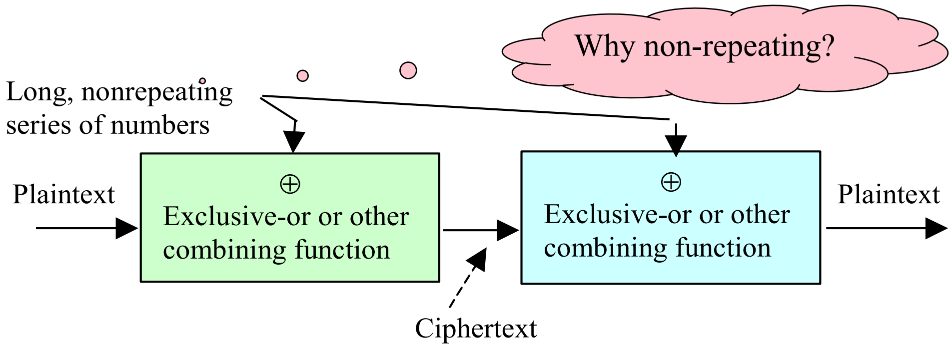
(b) How to choose the key stream to make the ciphertext the hardest to break?



Classical Encryption Techniques - One-time pad

- It was proposed by Gilbert Vernam during World War I.
- It is a special variant of the stream cipher.
- It is truly perfect cipher (perfect secrecy!):
  - **it uses a one-time random key that is as long as the plaintext with no repetitions (only used once).**
- If used properly, it is provably unbreakable. (Shannon, 1949)

Classical Encryption Techniques - One-time pad



- ❑ Problems
  - ❍It is too expensive for most applications - consumes as much key material as there is traffic;

$K = M \oplus C \;\; => \;\; M = K \oplus C$

Classical Encryption Techniques - One-time pad

- ❑ Problems (Cont.)
  - ❍Key management is hard!
    - ➢The need for non-repeating keys - problem with storing & distributing them, etc.
    - ➢Absolute synchronisation between sender and receiver. Otherwise, it fails completely to protect message integrity.

Plain: heilhitler
Key: wclnbtdefj
Cipher: DGTYIBWPJA
A spy's message

Cipher: DGTYIBWPJA
Key: wggsbtdefj
Plain: hanghitler
What the spy claimed he said

Stream Ciphers

- Basic idea: replace the random key in one-time pad by a pseudo-random sequence, generated by a cryptographic pseudo-random generator that is 'seeded' with the key.
- Ciphertext = plaintext XOR keystream

  M = m_1 m_2 m_3 … m_i …

  K = k_1 k_2 k_3 … k_i …

  C = c_1 c_2 c_3 … c_i …

  where $c_i = m_i \oplus k_i$, $i \geq 0$, and m_i is typically a byte (8 bits) or 1 bit.
- Same key used twice gives same keystream, as

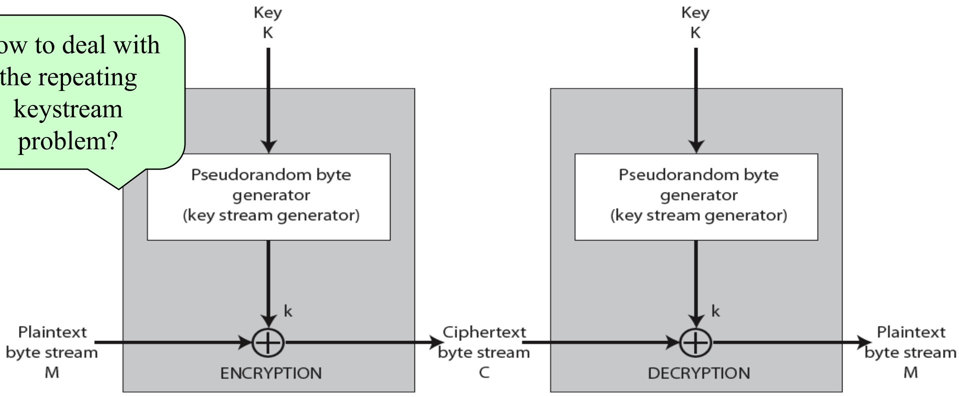
  $K = M \oplus C \;\; => \;\; M' = K \oplus C' = (M \oplus C) \oplus C'$

  This is a dangerous security property and we **must never *ever* reuse the same keystream** to encrypt two different messages.

Stream Ciphers - Structure

Generate a keystream from a short key that initializes the generator.

How to deal with the repeating keystream problem?



Classical Encryption Techniques - Transposition tech.

- Transposition technique (**permutation**)
  - To perform permutation on the plaintext. An example:

key	4	3	1	2	5	6	7
plaintext	a	t	t	a	c	k	p
	o	s	t	p	o	n	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

ciphertext ttnaaptmtsuoaodwcoixknlypetz

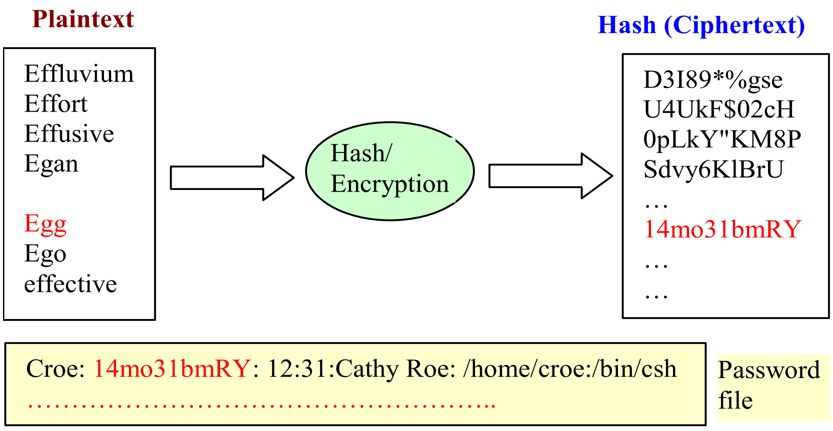
  - To write message in a rectangle, row by row, and read the message off, column by column, but permute the order of the column. Key = order of the columns.

What is the major difference between this cipher and the ciphers mentioned earlier?

Cryptanalytic Attacks

- The security of any (modern) cipher is based ***not*** on the secrecy of an algorithm, but on the security of the cryptographic keys!
- Common types of attacks
  - Try to break or 'crack' the algorithm by exploring any flaws in the algorithm, e.g. frequency analysis.
  - Try to decrypt the algorithm's ciphertext with every possible key until…, e.g. brute force attack (also exhaustive key search attack).
  - Run the algorithm on massive amounts of plaintext and find the one plaintext that encrypts to the ciphertext he is analysing, e.g. dictionary attack.

What is a dictionary attack? - an example



Exercise 2 – familiarise with CrypTool

- Download and install CrypTool 1.4.30 from http://www.cryptool.org/index.php/en/download-topmenu-63.html (or use the CrypTool already installed in the third year lab).
- This is a cryptographic e-learning software; it has a number of features which can make your learning interesting:
  - It is a freeware program with graphical user interface.
  - It visualises a number of algorithms.
  - It contains nearly all state-of-the-art cryptography functions.
  - It can be used to analyse cryptographic methods …
- Play with CryptTool and learn its capabilities.

Conclusions

- Explained a number of historical ciphers such as the Caesar cipher.
- Showed how these historical ciphers can be broken because they do not hide the underlying statistics of the plaintext.
- Introduced the concepts of substitution and transposition (permutation) as basic cipher components for classical cryptosystems.
- A good cryptosystem (cryptographic algorithms) must withstand all three sorts of attacks.
- Brute force and dictionary attacks can be thwart by using larger key space.