Topic 2: Introduction to Cryptography

Introduce the basic concepts of cryptography,
some classical techniques and cryptoanalysis attacks

Source: Stalling's book, chapter 3

Overview

- Part 1
  - What is cryptography and its applications
  - Terminology and definitions
- Part 2
  - Caesar cipher
- Part 3
  - More classical encryption techniques
  - More on cryptographic attacks
  - Conclusion

What is Cryptography?

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

Keep messages secure

Prevent unauthorised entities from gaining access to a network/system hosting messages (other resources).

Scramble messages so that they can't be understood by any unauthorised entities.

Cryptography is used.

Application of Cryptography

- Confidentiality (secrecy, privacy) of data in transmission & in storage
- Integrity of data (data authentication/authenticity) in transit & in storage
- Authentication of an identity (entity authentication)
- Credential systems (a proof of qualification or competence of a person)
- Digital signatures
- Electronic money (e.g. cryptocurrency, bit coins)

Application of Cryptography

- Threshold cryptosystems (a decryption key, or a signature signing key, is shared among a group of entities and a subset of these entities (more than some threshold number) have to collaborate to perform the decryption or signature signing).
- Secure multi-party computations (e.g. multiple parties compute a function jointly, the input is from the multiple parties, but no party should learn anything rather than its own input and the final result of the computation)
- Digital right management (e.g. activation of a software license by authorized users)
- Electronic voting
- …

Achieving Confidentiality using Encryption

- Using encryption to achieve secure communication over an insecure channel
- Ciphers (cryptosystems)
  - Symmetric-key based (conventional ciphers)
    - Same key is used for encryption and decryption
    - Historical ciphers
    - Modern ciphers
  - Asymmetric-key based (public-key ciphers)
    - Different keys are used for encryption and decryption
    - Modern ciphers

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 using an encryption key.
- Decryption: convert an encrypted message back to its original form using a decryption key.
- 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.

Security Definitions (in the context of cryptographic based solutions)

- Unconditionally secure
  - The system cannot be defeated, no matter how much power is available by the adversary.

- Conditionally secure based on computational complexity
  - The perceived level of computation required to defeat the system, which exceeds the computational resources of the hypothesized adversary.
  - e.g. given the computing power, it takes very long time to break a ciphertext.

Security Definitions (in the context of cryptographic based solutions)

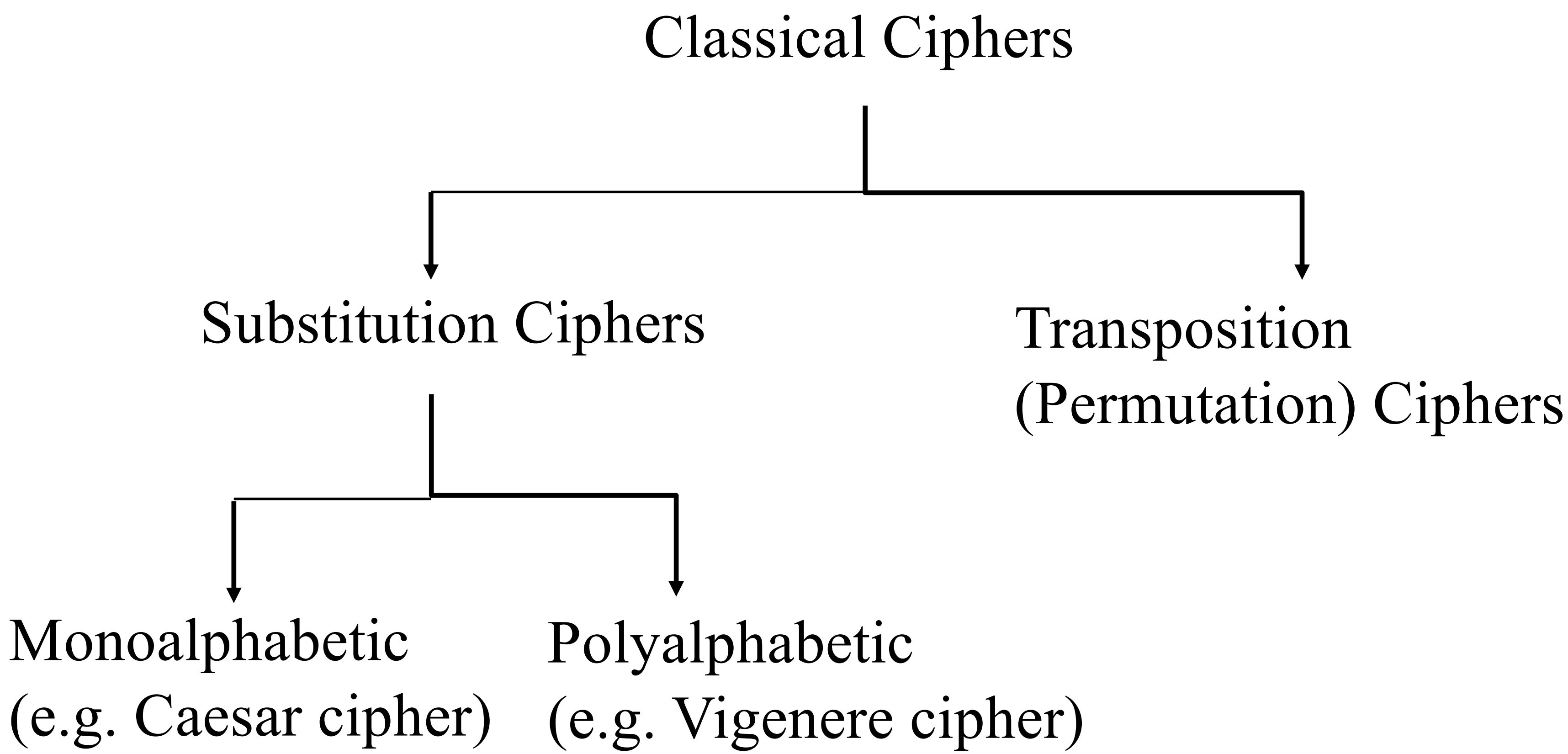
- Provably secure
  - Some solutions can be verified to be secure, e.g. in the case of security protocol designs.

- Ad hoc security
  - Claims of security generally remain questionable.
  - Unforeseen attacks remain a threat.

Part 2 Overview

- Caesar cipher

Classical Encryption Techniques



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, …
- 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 (Shift cipher)

- It uses simple substitution
- Encryption operation: each plaintext letter is translated to the letter a fixed number of letters further down the alphabet table (circular right shift).
- Decryption operation is the reverse of the encryption operation.
- The operation can 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 ± *k* to each letter (mod 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 – Attacks on 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 (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).
- If the plaintexts are known to contain redundancy, then ciphertext-only exhaustive key search is possible with a relatively small number of ciphertexts.
- With today's computing power, (symmetric) key length should be at least 128 bits.

- Also vulnerable to another form of attack - frequency analysis (also known as counting letters) attack.

Classical Encryption Techniques – Attacks on Caesar cipher

- Also vulnerable to frequency analysis (also known as counting letter) attack.
- Frequency analysis attack is a known-ciphertext attack based on the study of the frequency of letters or groups of letters in a ciphertext.

Part 3 Overview

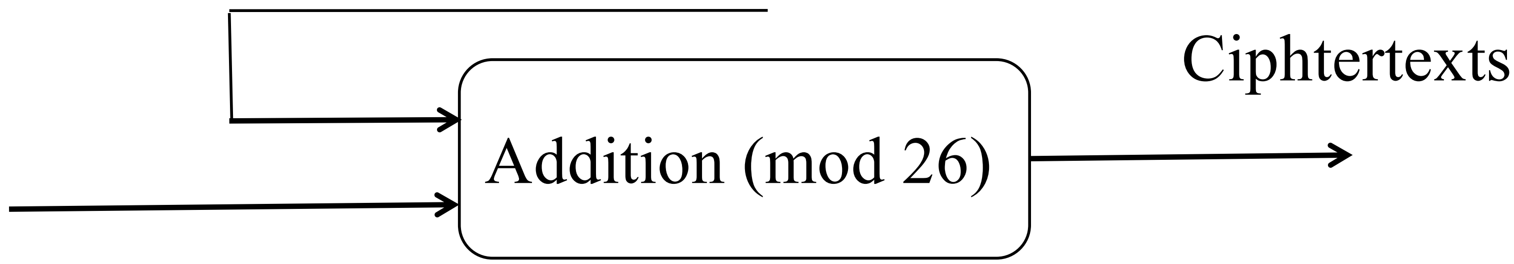
- More classical encryption techniques
  - Vigenere cipher
  - One-time pad and stream ciphers
  - Transposition cipher
- More on cryptographic attacks
- Conclusion

Let us start with a quick question on Caesar cipher

(a) This 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?

Plaintext: There once was an ugly duckling ...



Key = . . . 3, 3, 3, 3, 3, 3, 3

The University of Manchester

Letter Frequency Distribution in English (Literature)

- **This is in percentage term (**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	

Vigenere Cipher

- This cipher uses a keyword. For example, let's say the keyword (K) is 'bed', i.e. 143 (as b → 1, e → 4, d → 3). The plaintext (P) (*There once was an ugly duckling …*), and the corresponding ciphertext (C) are given below:

	P:	**T h**	**e**	**r**	**e**	**o**	**n**	**c**	**e**	**w**	**a**	**s**	**a**	**n**	**u**	**g**	**l**	**y**	**d**	**u**	**c**	**k**	**l**	**i**	**n**	**g**
	K:	1 4	3	1	4	3	1	4	3	1	4	3	1	4	3	1	4	3	1	4	3	1	4	3	1	4
	C	u l	h	s	i	r	o	i	h	x	e	v	b	r	x	h	p	b	e	y	f	l	p	l	0	k

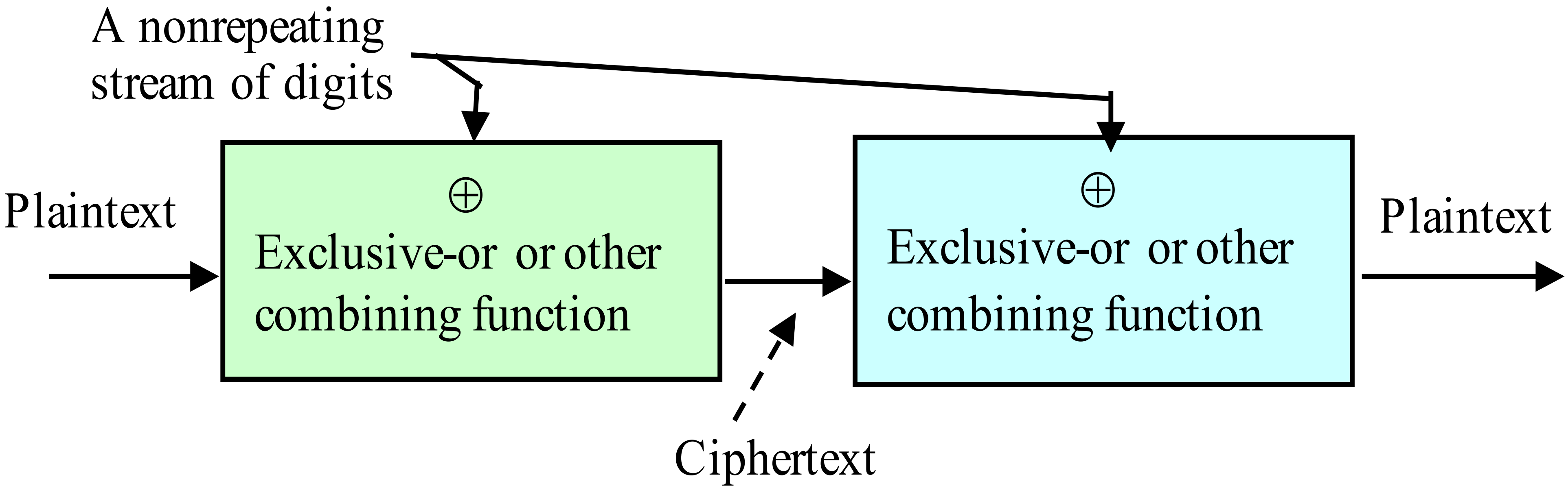
- Here, each plaintext letter has been shifted by a different amount that is determined by the key.
- This cipher is significantly more secure than a regular Caesar Cipher. Its security level is dependent on the keyword length.

One-time Pad

- If the keyword length is as long as the plaintext and is random, then we have a cipher with perfect secrecy.
- **One-time pad** is such a cipher.
  - **it uses an 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)
- It was proposed by Gilbert Vernam during World War I.
- It is a special variant of the stream cipher.
  - Typically a stream cipher uses (mod 2) (exclusive-or, i.e. XOR) function.

One-time Pad

M xor K = C;
C xor K = M



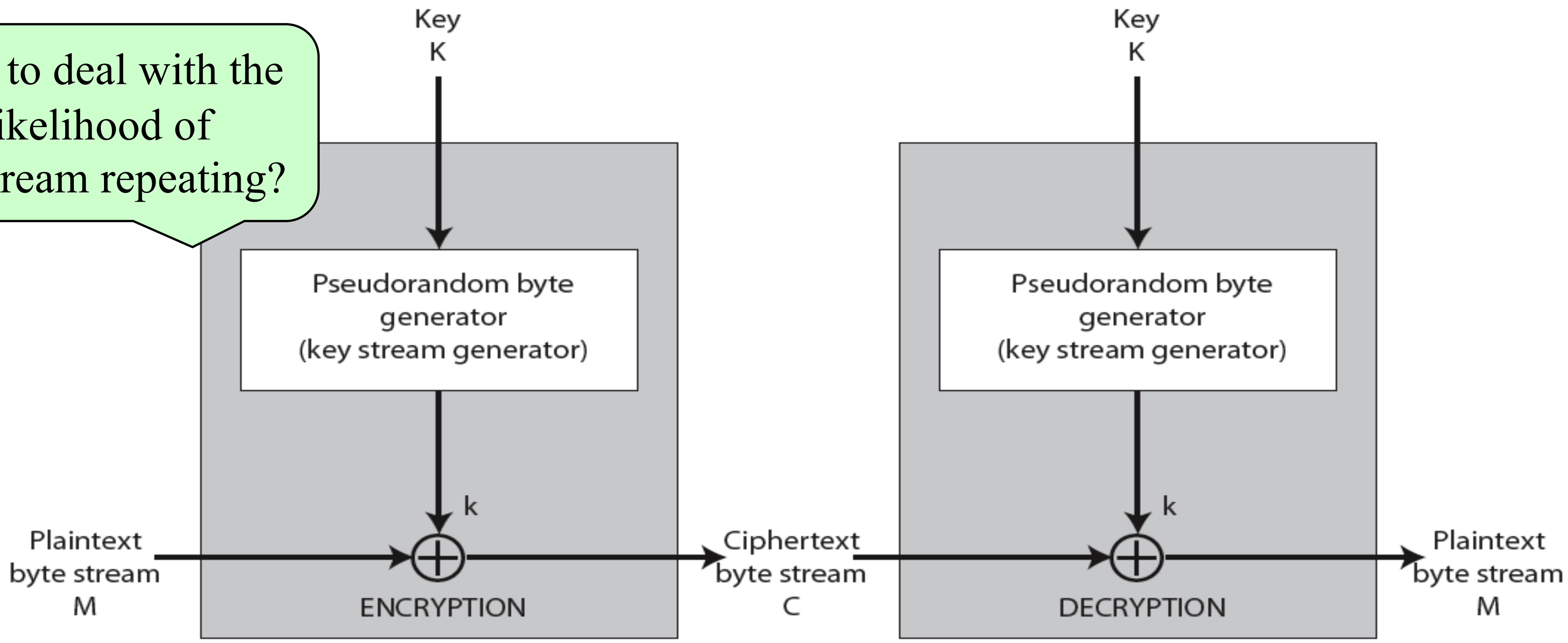
Stream Ciphers

- Stream ciphers encrypt individual bit or character streams.
- When encrypting individual bit streams, XOR is used, i.e.
  ciphertext (C) = plaintext (M) xor keystream (KS)
  $M = m_1 \; m_2 \; m_3 \; \ldots \; m_i \; \ldots$
  $KS = k_1 \; k_2 \; k_3 \; \ldots \; k_i \; \ldots$
  $C = c_1 \; c_2 \; c_3 \; \ldots \; c_i \; \ldots$
  where $c_i = m_i$ xor k_i, and m_i is typically a byte (8 bits) or 1 bit.
- 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.

Stream Ciphers

Generate a keystream using a key that initializes the generator.

How to deal with the likelihood of keystream repeating?



MANCHESTER 1824

The University of Manchester

Transposition Cipher

- The ciphertext is generated by performing permutation on the plaintext (i.e. changing the order of the alphabets in 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

  - Write the plaintext in a rectangle, row by row, and read the message off, column by column, but permuting the order of the columns, where Key = order of the columns to read.

Cryptographic 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 (frequency analysis).
  - Assume attackers can recognize a plaintext, try to decrypt a ciphertext with every possible key until a recognised plaintext is obtained (brute force attack or exhaustive key search attack).
  - Run the algorithm on massive amount of (probable) plaintexts until a plaintext that encrypts to the ciphertext he is analysing is found (dictionary attack).

What is a dictionary attack? - an example

Plaintext

Effluvium
Effort
Effusive
Egan

Egg
Ego
effective

Hash/
Encryption

Hash (Ciphertext)

D3I89*%gse
U4UkF$02cH
0pLkY"KM8P
Sdvy6KlBrU
…
14mo31bmRY
…
…

Croe: 14mo31bmRY: 12:31:Cathy Roe: /home/croe:/bin/csh
……………………………………………..

Password file

More on Cryptographic Attacks

- Ciphertext-only attack (e.g. frequency analysis)
  - Attacker knows ciphertexts of several messages encrypted with same key, plaintext is recognizable;
  - Goal: to find plaintext, possibly key

- Known-plaintext attack (e.g. dictionary attack)
  - Attacker observes <plaintext, ciphertext> pairs encrypted with same key;
  - goal: to find key

The University of Manchester

MANCHESTER 1824

More on Cryptographic Attacks

- Chosen-plaintext attack
  - Attacker can choose the plaintext and look at the paired ciphertext;
  - goal: to find the key

- Cryptographic attacks often exploit the redundancy of natural language
  - Lossless compression before encryption removes redundancy

Exercise Question – E2.1

- Given the Letter Frequency Distribution in English, as shown in the next slide, and the following ciphertext which has been generated using the Caesar cipher (but a different key), use the frequency analysis method to work out the encryption key and the corresponding plaintext.

Ciphertext:
bpmzm wvkm eia iv cotg lckstqvo eqbp nmibpmza itt abcjjg ivl jzwev …

Key: ??
Plaintext: ??

MANCHESTER 1824

The University of Manchester

Exercise Question – E2.1 continue

- **Letter Frequency Distribution in English (in percentage)** (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	

Exercise Question – E2.2

The stream cipher diagram given earlier shows that a key stream used for encryption/decryption is generated by using a pseudo-random generator that is 'seeded' with a (shorter) key, K. This key, K, is usually called encryption/decryption key.

(i) Comment on the benefit(s) of this approach, i.e. why is the key stream generated from K?

(ii) How to ensure (or to minimize the chances) that the output of the pseudo-random generator (i.e. the key stream) is non-repeating?

Familiarise with CrypTool

- Download and install CrypTool 1.4.30 from http://www.cryptool.org/index.php/en/download-topmenu-63.html (or use another on-line cryptotool).
- 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 the CryptTool and learn its capabilities.

Conclusions

- Explained a number of historical ciphers.
- Explained how some of the ciphers may be attacked.
- Introduced the concepts of substitution and transposition (permutation) as basic cipher operations for classical cryptosystems.
- Introduced some cryptographic attacks
- A secure cryptosystem must withstand all sorts of attacks.