### *Introduction to Cryptography*

François Koeune - Olivier Pereira

Slides 01





#### Goal of the course

#### Understand fundamental

- concepts,
- methods, and
- algorithms

used to secure information, with an emphasis on the algorithmic and mathematical aspects.

#### Related courses at UCL

# Option in *Cryptography and Information Security* (EPL – DATA/ELEC/INFO/MAP)

- ► LELEC2760 Secure electronic circuits and systems F -X. Standaert
- ► **LELEC2770** Privacy Enhancing Technologies O. Pereira, F.-X. Standaert
- ► LINGI2144 Secured systems engineering A. Legay
- ► LINGI2347 Computer System Security R. Sadre
- LING12348 Information theory and coding J. Louveaux, B. Macq, O. Pereira
- ▶ LMAT2440 Théorie des nombres O. Pereira, J.-P. Tignol
- ► **LMAT2450** Cryptography O. Pereira

#### Related courses at UCL

#### Other related courses

- ► **LELEC2870** Machine Learning J. Lee, M. Verleysen
- ▶ **LINGI1341** Computer networks O. Bonaventure
- ► LINMA2111 Discrete mathematics II : Algorithms and complexity J.-C. Delvenne
- ▶ LEPL2210 Ethics and ICT A. Gosseries, O. Pereira



#### Class Organisation

- ► Lectures/Exercises on Wednesday, 14:00 16:00 Exercises on Wednesday, 16:15 18:15
- ► TAs: Clément Hoffmann, Yaobin Shen
- ► We may offer homeworks:
  - ▶ make up to 20% of the January grade, if this helps you
  - do not count in August
- Examination: exercises. Slides and personal notes allowed
   Exam questions from past years often proposed as exercises

### Syllabus

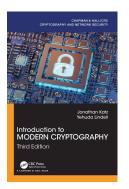
#### Expected distribution:

- ► Introduction (1 lecture)
- Symmetric cryptography (4 lectures)
- Asymmetric cryptography and algorithmic number theory (4 lectures)
- ► Protocols (2 lectures)



#### Support

Introduction to Modern Cryptography (2nd edition) by J. Katz and Y. Lindell – Chapman & Hall/CRC – 2020



http://www.cs.umd.edu/~jkatz/imc.html

#### Support

#### Other references (see also Moodle):

- ► W. Mao, *Modern Cryptography, Theory and Practice*, Prentice-Hall, PTR, 2003.
- ▶ D. Stinson, Cryptography, Theory and Practice, 3rd edition, Chapman & Hall/CRC, 2005.
- ► A.J. Menezes, P. van Oorschot, S. Vanstone, *Handbook of Applied Cryptography*, 1999. Free on http://www.cacr.math.uwaterloo.ca/hac/.
- ► D. Boneh, V. Shoup, A Graduate Course in Applied Cryptography, Free draft on http://toc.cryptobook.us/
- N. Koblitz, A Course in Number Theory and Cryptography, Graduate Texts in Math. No. 114, 2nd edition, Springer-Verlag, 1994.

#### Cryptography...

COD defines cryptography as: "the art of writing or solving codes."

- ► Certainly true until mid of 20<sup>th</sup> century
- ▶ Mostly used by armies and diplomats



#### Cryptography... today

#### Used every day!















### *Cryptography... today*

#### Much more than encryption:

- authentication
- ▶ key exchange
- ▶ identification
- elections
- ► Yao millionaire's problem
- ▶ ..

### *Cryptography... today*

#### Much more than encryption:

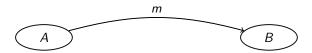
- authentication
- ▶ key exchange
- ▶ identification
- elections
- Yao millionaire's problem
- ▶ ..

From an art, cryptography became a science. . .

# The message encryption problem

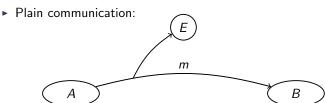
#### The setting:

► Plain communication:



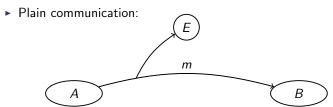
# The message encryption problem

#### The setting:

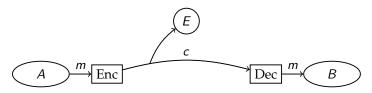


### The message encryption problem

#### The setting:



Encrypted communication:



What is an encryption scheme? A triple (Gen, Enc, Dec)



What is an encryption scheme? A triple  $\langle Gen, Enc, Dec \rangle$ 

- ▶ Gen probabilistically selects a key k
- ▶ Enc encrypts message m with key k, as  $c := \operatorname{Enc}_k(m)$
- ▶ Dec decrypts ciphertext c with key k as  $m := Dec_k(c)$

What is an encryption scheme? A triple (Gen, Enc, Dec)

- ▶ Gen probabilistically selects a key k
- ▶ Enc encrypts message m with key k, as  $c := \operatorname{Enc}_k(m)$
- ▶ Dec decrypts ciphertext c with key k as  $m := Dec_k(c)$

#### Remarks:

- same key is used for encryption and decryption: symmetric/private-key encryption
- ▶ correctness requirement:  $\forall k, m : m = Dec_k(Enc_k(m))$

An example: the Scytale (Greece, 7th century BC (?))



An example: the *Scytale* (Greece, 7<sup>th</sup> century BC (?))



- Gen defines the diameter of the cylinder (k := number of letters you can write on the circumference)
- ► Enc encrypts by transposing letters according to k
- Dec decrypts by performing the inverse transposition

Another example: the Caesar's cipher (Rome, 1st c. BC)

▶ Shift letters (D  $\rightarrow$  A, E  $\rightarrow$  B, F  $\rightarrow$  C, ..., C  $\rightarrow$  Z)

▶ Ex:  $\text{HELLO} \rightarrow \text{EBIIL}$ 

Another example: the Caesar's cipher (Rome, 1st c. BC)

- ▶ Shift letters (D  $\rightarrow$  A, E  $\rightarrow$  B, F  $\rightarrow$  C, ..., C  $\rightarrow$  Z)
- ► Ex: HELLO → EBIIL
- ▶ Gen defines the extent  $k \in [0, 25]$  of the shift
- Enc encrypts by substituting letters, applying the right shift
- Dec decrypts by performing the inverse shift

Another example: the Caesar's cipher (Rome, 1st c. BC)

- ▶ Shift letters (D  $\rightarrow$  A, E  $\rightarrow$  B, F  $\rightarrow$  C, ..., C  $\rightarrow$  Z)
- ► Ex: HELLO → EBIIL
- ▶ Gen defines the extent  $k \in [0, 25]$  of the shift
- ► Enc encrypts by substituting letters, applying the right shift
- Dec decrypts by performing the inverse shift

For more historical examples, see, e.g.,: http://www.apprendre-en-ligne.net/crypto/

Cryptanalysis: art of code breaking/cracking



Cryptanalysis: art of code breaking/cracking

#### Attacker model:

- 1. What should be considered as secret? Gen? Enc? Dec? k?
- 2. Which attack scenario? Eavesdropper? Chosen-plaintext? ...?

#### Attacker model:

1. What should be considered as secret? Gen? Enc? Dec? k?

Kerckhoffs' principle (1883): only the key should be secret



See: http://www.petitcolas.net/fabien/kerckhoffs/

Un grand nombre de combinaisons ingénieuses peuvent répondre au but qu'on veut atteindre dans le premier cas ; dans le second, il faut un système remplissant certaines conditions exceptionnelles, conditions que je résumerai sous les six chefs suivants :

1° Le système doit être matériellement, sinon mathématiquement, indéchiffrable ;

2° Il faut qu'il n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi ;

3° La clef doit pouvoir en être communiquée et retenue sans le secours de notes écrites, et être changée ou modifiée au gré des correspondants ;

4° Il faut qu'il soit applicable à la correspondance télégraphique;

5° Il faut qu'il soit portatif, et que son maniement ou son fonctionnement n'exige pas le concours de plusieurs personnes ;

6° Enfin, il est nécessaire, vu les circonstances qui en commandent l'application, que le système soit d'un usage facile, ne demandant ni tension d'esprit, ni la connaissance d'une longue série de règles à observer.



- 1. Keeping secrets is annoying:
  - ▶ A key is easier to exchange secretly than a full system
  - A key is easier to update in case of compromise
  - One encryption scheme per pair of users is not manageable
  - ► No need to kill the cryptographer

- 1. Keeping secrets is annoying:
  - ▶ A key is easier to exchange secretly than a full system
  - A key is easier to update in case of compromise
  - ▶ One encryption scheme per pair of users is not manageable
  - No need to kill the cryptographer
- 2. Public algorithms should be safer:
  - Public algorithms can be scrutinized by friends
  - Possibility to create standards
  - ► No risk of code reverse-engineering



- 1. Keeping secrets is annoying:
  - ▶ A key is easier to exchange secretly than a full system
  - A key is easier to update in case of compromise
  - ▶ One encryption scheme per pair of users is not manageable
  - No need to kill the cryptographer
- 2. Public algorithms should be safer:
  - Public algorithms can be scrutinized by friends
  - Possibility to create standards
  - ► No risk of code reverse-engineering
- 3. We can handle it...

#### Attacker model:

1. Which attack scenario?



#### Attacker model:

1. Which attack scenario?

#### Depending on the context:

- 1. Passive:
  - Ciphertext-only: you only see ciphertexts
  - ► Known-plaintext: you see some plaintext/ciphertext pairs

#### Attacker model:

1. Which attack scenario?

#### Depending on the context:

- 1. Passive:
  - Ciphertext-only: you only see ciphertexts
  - ► Known-plaintext: you see some plaintext/ciphertext pairs
- 2. Active:
  - Chosen-plaintext: you can ask for the encryption of some messages
  - Chosen-ciphertext: you can also ask for the decryption of some messages

Consider Caesar's cipher, with:

- ► Public algorithms
- Ciphertext only

How do we break it?



#### Consider Caesar's cipher, with:

- ▶ Public algorithms
- ► Ciphertext only

How do we break it?

▶ Just try the 26 possible keys!

# Cryptanalysis

#### Consider Caesar's cipher, with:

- ▶ Public algorithms
- Ciphertext only

#### How do we break it?

Just try the 26 possible keys!

#### Lesson:

### Improvement on Caesar's cipher:

- ▶ Not just a shift: take any permutation of the alphabet
- ▶ This is  $26! \approx 2^{88}$  keys

#### How do we break it?

Sherlock Holmes did it!



### Improvement on Caesar's cipher:

- Not just a shift: take any permutation of the alphabet
- ▶ This is  $26! \approx 2^{88}$  keys

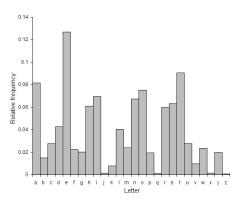
#### How do we break it?

► Sherlock Holmes did it!

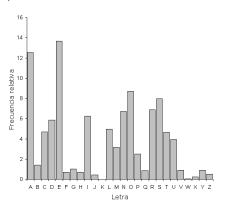


► If you know the plaintext language, frequency analysis is possible. . .

### Frequencies in English:



### Frequencies in Spanish:



### Improvement on Caesar's cipher:

- Not just a shift: take any permutation of alphabet
- ▶ This is  $26! \approx 2^{88}$  keys

#### How do we break it?

Just count the frequency of each symbol...

#### Lesson:

- Large key space is not enough!
- ▶ We need something secure independently of message distribution

# Vigenère cipher

#### Another improvement on Caesar's cipher:

- ▶ Instead of a constant shift, use different shifts according to position
- ▶ Key is a sequence of numbers in [0, 25]

#### Example:

- ► Suppose key is (2, 24, 5)
- Cryptography is great Attnvjetvnjt gu bpgvr

# Vigenère cipher

#### Another improvement on Caesar's cipher:

- ▶ Instead of a constant shift, use different shifts according to position
- Key is a sequence of numbers in [0, 25]

#### Example:

- ► Suppose key is (2, 24, 5)
- Cryptography is great Attnvjetvnjt gu bpgvr

#### How do we break it?

If you have enough ciphertext material... Make guesses on the key length, then make frequency analysis!

## Historic ciphers

#### Lessons:

We can keep playing like this for a long time...
 (See, e.g., D. Kahn, "The code-breakers" (Scribner) or J. Stern, "La science du secret" (Odile Jacob))

Can we do something else?

▶ In many cases: yes!

# Modern Cryptography

- "Modern cryptography"
  - 1. Definitions
  - 2. Assumptions
  - 3. Proofs



## Modern Cryptography: Definitions

### Definitions in cryptography:

- 1. What do we want to do?
- 2. Shall I use this scheme here?
- 3. Why choosing this scheme rather than that one?





What should the definition of security say for an encryption scheme?

1. Given any ciphertext, no adversary should be able to recover the key



- 1. Given any ciphertext, no adversary should be able to recover the key
- 2. Given any ciphertext, no adversary should be able to recover the plaintext



- 1. Given any ciphertext, no adversary should be able to recover the key
- 2. Given any ciphertext, no adversary should be able to recover the plaintext
- 3. Given any ciphertext, no adversary should be able to recover any character of the plaintext



- 1. Given any ciphertext, no adversary should be able to recover the key
- Given any ciphertext, no adversary should be able to recover the plaintext
- 3. Given any ciphertext, no adversary should be able to recover any character of the plaintext
- 4. Given any ciphertext, no adversary should be able to recover any function of the plaintext

What should the definition of security say for an encryption scheme?

- 1. Given any ciphertext, no adversary should be able to recover the key
- Given any ciphertext, no adversary should be able to recover the plaintext
- Given any ciphertext, no adversary should be able to recover any character of the plaintext
- 4. Given any ciphertext, no adversary should be able to recover any function of the plaintext

Still need to define the adversarial model...



# Modern Cryptography: Definitions

#### Limitations: Science vs. real world

- ▶ Check whether intuitive properties are guaranteed
- Compare with other definitions
- Compare with attack examples
- ▶ Use it during a few years...



# Modern Cryptography: Precise Assumptions

### Most schemes rely on computational assumptions

- ▶ Need to understand what we are trusting (challenges)
- Needed to write security proofs
- Useful for abstraction
- Useful for scheme comparison

# Modern Cryptography: Proof of Security

#### Relate schemes and definitions to assumptions

► Reductionist approach: if someone can break this scheme, (s)he is also able to falsify my assumption



We just broke encryption schemes... but



We just broke encryption schemes... but

Shannon (1949): perfect encryption is possible!



What is an encryption scheme? A triple (Gen, Enc, Dec)

- ▶ Gen probabilistically selects a key  $k \in \mathcal{K}$
- ▶ Enc encrypts message  $m \in \mathcal{M}$  with key k, as  $c \leftarrow \operatorname{Enc}_k(m)$
- ▶ Dec decrypts ciphertext  $c \in C$  with key k as  $m := Dec_k(c)$

#### Remarks:

- Enc may be probabilisitic
- $ightharpoonup \operatorname{Dec}_k(\operatorname{Enc}_k(m)) = m$ , always
  - ⇒ assume, wlog, Dec to be deterministic

What is an encryption scheme? A triple (Gen, Enc, Dec)

- ▶ Gen probabilistically selects a key  $k \in \mathcal{K}$
- ▶ Enc encrypts message  $m \in \mathcal{M}$  with key k, as  $c \leftarrow \operatorname{Enc}_k(m)$
- ▶ Dec decrypts ciphertext  $c \in C$  with key k as  $m := Dec_k(c)$

#### Remarks:

- Enc may be probabilisitic
- ▶  $Dec_k(Enc_k(m)) = m$ , always ⇒ assume, wlog, Dec to be deterministic
- Assume  $|\mathcal{M}| > 1$
- Assume  $\mathcal M$  and  $\mathcal C$  only contain messages and ciphertexts that may happen.

*Definition:*  $\langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$  over message space  $\mathcal{M}$  is *perfectly secret* if, for every probability distribution over  $\mathcal{M}$ , every message  $m \in \mathcal{M}$ , and every ciphertext  $c \in \mathcal{C}$ :

$$Pr[M = m | C = c] = Pr[M = m]$$

#### Remarks:

lacktriangleright Probability distribution over  ${\mathcal M}$  refers to distribution on messages





#### Equivalent definition:

 $\langle \operatorname{Gen}, \operatorname{Enc}, \operatorname{Dec} \rangle$  over message space  $\mathcal M$  is *perfectly secret* if, for every probability distribution over  $\mathcal M$ , every message  $m \in \mathcal M$ , and every ciphertext  $c \in \mathcal C$ :

$$Pr[C = c|M = m] = Pr[C = c]$$

#### Equivalent definition:

 $\langle \operatorname{Gen}, \operatorname{Enc}, \operatorname{Dec} \rangle$  over message space  $\mathcal M$  is *perfectly secret* if, for every probability distribution over  $\mathcal M$ , every message  $m \in \mathcal M$ , and every ciphertext  $c \in \mathcal C$ :

$$Pr[C = c|M = m] = Pr[C = c]$$

### Proof of equivalence:

#### Equivalent definition:

 $\langle \operatorname{Gen},\operatorname{Enc},\operatorname{Dec} \rangle$  over message space  $\mathcal M$  is *perfectly secret* if, for every probability distribution over  $\mathcal M$ , every message  $m\in \mathcal M$ , and every ciphertext  $c\in \mathcal C$ :

$$\Pr[C = c | M = m] = \Pr[C = c]$$

#### Proof of equivalence:

$$Pr[C = c | M = m] = Pr[C = c]$$

#### Equivalent definition:

 $\langle \operatorname{Gen},\operatorname{Enc},\operatorname{Dec} \rangle$  over message space  $\mathcal M$  is *perfectly secret* if, for every probability distribution over  $\mathcal M$ , every message  $m\in \mathcal M$ , and every ciphertext  $c\in \mathcal C$ :

$$\Pr[C = c | M = m] = \Pr[C = c]$$

#### Proof of equivalence:

$$\begin{array}{ll}
\operatorname{Pr}[C = c | M = m] &= \operatorname{Pr}[C = c] \\
\operatorname{(Bayes} \Rightarrow) & \frac{\operatorname{Pr}[M = m]C = c] \cdot \operatorname{Pr}[C = c]}{\operatorname{Pr}[M = m]} &= \operatorname{Pr}[C = c]
\end{array}$$

#### Equivalent definition:

 $\langle \operatorname{Gen}, \operatorname{Enc}, \operatorname{Dec} \rangle$  over message space  $\mathcal M$  is *perfectly secret* if, for every probability distribution over  $\mathcal M$ , every message  $m \in \mathcal M$ , and every ciphertext  $c \in \mathcal C$ :

$$\Pr[C = c | M = m] = \Pr[C = c]$$

#### Proof of equivalence:

$$Pr[C = c | M = m] = Pr[C = c]$$
(Bayes  $\Rightarrow$ ) 
$$\frac{Pr[M = m | C = c] \cdot Pr[C = c]}{Pr[M = m]} = Pr[C = c]$$
(Reorganize  $\Rightarrow$ ) 
$$Pr[M = m | C = c] = Pr[M = m]$$

#### Equivalent definition:

 $\langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$  over message space  $\mathcal{M}$  is *perfectly secret* if, for every  $m_0, m_1 \in \mathcal{M}$  and every ciphertext  $c \in \mathcal{C}$ :

$$Pr[C = c | M = m_0] = Pr[C = c | M = m_1]$$

#### Interpretation:

▶ It is impossible to distinguish the ciphertext corresponding to two plaintexts

### Equivalent definition. . .

Given  $\Pi := \langle Gen, Enc, Dec \rangle$ , and adversary  $\mathcal{A}$ , define the following experiment  $\mathsf{PrivK}_{\mathcal{A},\Pi}^{\mathsf{eav}}$ :

- 1.  $\mathcal{A}$  outputs  $m_0, m_1 \in \mathcal{M}$
- 2. Choose  $k \leftarrow \mathcal{K}$  and  $b \leftarrow \{0,1\}$ , and send  $\operatorname{Enc}_k(m_b)$  to  $\mathcal{A}$
- 3. A outputs b'
- 4. Define  $PrivK_{\mathcal{A},\Pi}^{eav} := 1$  iff b = b'

# *Perfect encryption (cont.)*

#### Equivalent definition:

 $\langle Gen, Enc, Dec \rangle$  over message space  ${\cal M}$  is *perfectly secret* if for every adversary  ${\cal A}$ :

$$\mathsf{Pr}[\mathsf{Priv}\mathsf{K}^{\mathsf{eav}}_{\mathcal{A},\mathsf{\Pi}}=1]=rac{1}{2}$$

### Interpretation:

Even if A chooses 2 messages, it cannot decide which of them has been encrypted

### One-time pad is perfectly secret!

- ▶ Fix I > 0.  $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0, 1\}^I$
- ightharpoonup Gen selects uniformly in  ${\cal K}$
- $ightharpoonup \operatorname{Enc}_k(m) := m \oplus k$
- ▶  $Dec_k(c) := c \oplus k$

#### Remarks:

- → denotes binary XOR (exclusive OR)
- ▶  $Dec_k(Enc_k(m)) = m \oplus k \oplus k = m$

One-time pad is perfectly secret!

#### Proof:

Fix any distribution over  $\mathcal{M}$ , any  $m \in \mathcal{M}$  and  $c \in \mathcal{C}$ .

$$\Pr[C = c | M = m] = \Pr[M \oplus K = c | M = m]$$

One-time pad is perfectly secret!

#### Proof:

Fix any distribution over  $\mathcal{M}$ , any  $m \in \mathcal{M}$  and  $c \in \mathcal{C}$ .

$$Pr[C = c|M = m] = Pr[M \oplus K = c|M = m]$$
  
=  $Pr[m \oplus K = c]$ 

One-time pad is perfectly secret!

#### Proof:

Fix any distribution over 
$$\mathcal{M}$$
, any  $m \in \mathcal{M}$  and  $c \in \mathcal{C}$ . 
$$\Pr[C = c | M = m] = \Pr[M \oplus K = c | M = m]$$
$$= \Pr[m \oplus K = c]$$
$$= \Pr[K = m \oplus c] = \frac{1}{2^l}$$

One-time pad is perfectly secret!

#### Proof:

Fix any distribution over  $\mathcal{M}$ , any  $m \in \mathcal{M}$  and  $c \in \mathcal{C}$ .  $Pr[C = c|M = m] = Pr[M \oplus K = c|M = m]$  $= \Pr[m \oplus K = c]$  $= \Pr[K = m \oplus c] = \frac{1}{2^l}$ = Pr[C = c|M = m'] for every m'

One-time pad is perfectly secret!



One-time pad is perfectly secret!

One-time pad is not convenient to use...

- key needs to be as long as message!
- suppose m, m' encrypted with k  $(m \oplus k) \oplus (m' \oplus k) = m \oplus m'$   $\mathcal{A}$  wins if it can play  $\mathsf{PrivK}_{\mathcal{A},\Pi}^{\mathsf{eav}}$  twice with same key!



# Limits of Perfect Secrecy

Suppose  $\Pi := \langle Gen, Enc, Dec \rangle$  is s.t.  $|\mathcal{K}| < |\mathcal{M}|$ . Then  $\Pi$  is not a perfectly secret encryption scheme.



## Limits of Perfect Secrecy

Suppose  $\Pi := \langle Gen, Enc, Dec \rangle$  is s.t.  $|\mathcal{K}| < |\mathcal{M}|$ . Then  $\Pi$  is not a perfectly secret encryption scheme.

#### Proof.

Consider uniform distribution on  $\mathcal{M}$  and any  $c \in \mathcal{C}$ . Define  $\mathcal{M}(c) := \{ \hat{m} : \hat{m} = \mathrm{Dec}_{\hat{k}}(c) \text{ for some } \hat{k} \in \mathcal{K} \}$ We must have  $|\mathcal{M}(c)| < |\mathcal{M}|$ Therefore,  $\exists m \in \mathcal{M} - \mathcal{M}(c)$ , and  $Pr[M = m | C = c] = 0 \neq Pr[M = m]$ 

Perfectly secret encryption schemes exist, but are difficult to use Can we do better?



Perfectly secret encryption schemes exist, but are difficult to use

Can we do better?

Shannon's theory says: "no!"



Perfectly secret encryption schemes exist, but are difficult to use

Can we do better?

Shannon's theory says: "no!"

Under which assumptions?

- A has perfect information
- $ightharpoonup \mathcal{A}$  has unbounded computational power



Perfectly secret encryption schemes exist, but are difficult to use

Can we do better?

Shannon's theory says: "no!"

Under which assumptions?

- A has perfect information
- A has unbounded computational power

Next week:

What about bounded computational power?

