Lezione 3

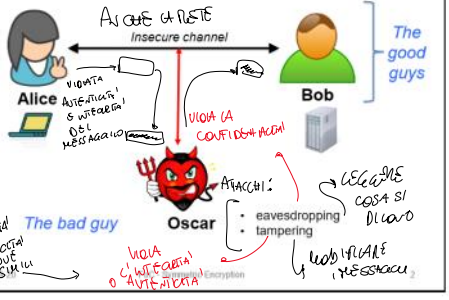
Monday, March 1, 2021 1:55 PM

Symmetric Encryption

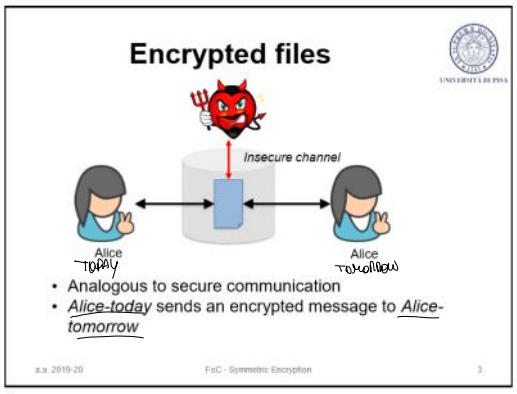
Gianluca Dini
Dept. of Ingegneria dell'Informazione
University of Pisa
gianluca.dini@unipi.it
Last version: 2021-03-01

1

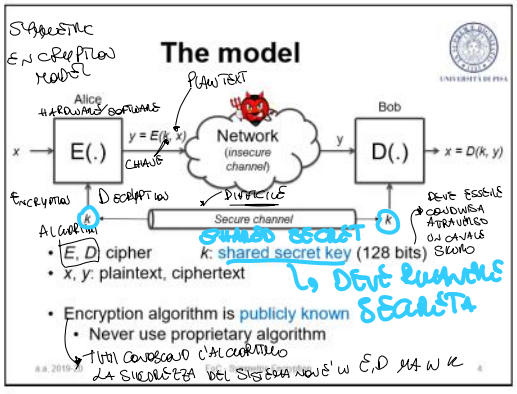
Main characters



2

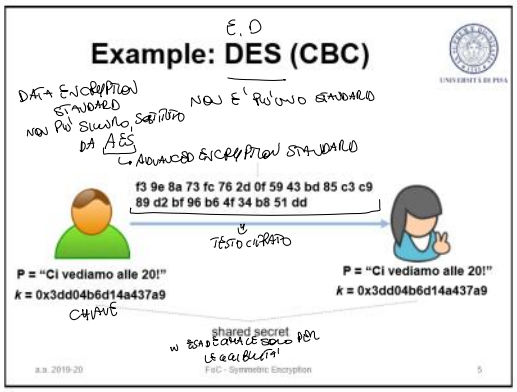

Encrypted files
Insecure channel
Alice
TODAY
Alice
TOMORROW
• Analogous to secure communication
• Alice-today sends an encrypted message to Alice-tomorrow
a.a. 2019-20
FoC - Symmetric Encryption
3

3


The model
Alice
Bob
Network
(insecure channel)
E(.)
D(.)
x
y = E(k, x)
y
x = D(k, y)
Secure channel
k
k
• E, D cipher
• x, y: plaintext, ciphertext
k: shared secret key (128 bits)
• Encryption algorithm is publicly known
• Never use proprietary algorithm
a.a. 2019-20

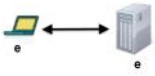
4

Example: DES (CBC)



5

Example: SSL



- **Handshake protocol**
  - establish a **shared secret key** by means of public key cryptography
    - 2nd part of the course
- **Record protocols**
  - use **shared secret key** to transmit data to ensure confidentiality and integrity
    - 1st part of the course

6

SYMMETRIC → CHIAVE PER CRIPTARE E DECRIPTARE SONO UGUALE

Cipher definition

SPAZIO DELLE CHIAVE (KEY SPACE) — PLAINTEXT — CYPHERTEXT SPACE

- **(DEF)** A cipher defined over $(\mathcal{K}, \mathcal{P}, \mathcal{C})$ is a pair of "efficient" algs (E, D) where

ALGORITMO EFFICIENTE ① POCO TEMPO ② COMPLESSITÀ POLINOMIALE

$$E: \mathcal{P} \times \mathcal{K} \rightarrow \mathcal{C} \qquad D: \mathcal{C} \times \mathcal{K} \rightarrow \mathcal{P}$$

- **Consistency Property** DEVO POTER RECUPERARE IL TESTO CIFRATO

$$\forall p \in \mathcal{P}, k \in \mathcal{K} : D(k, E(k, p)) = p$$

– E may be randomized; D is always deterministic — INVERSO

DECIDO SOLO UNO DEI DUE, L'ALTRO È DATO

7

POSSO DEFINIRE DIFFERENTI FUNZIONI, QUALI SONO SICURE?

Security of a cipher (informal)

SUPPOSIZIONI

① TUTTI I MESSAGGI CIFRATI POSSONO ESSERE LETTI

② TUTTI I MESSAGGI CHE POSSONO ESSERE LETTI VIOLANO LA CONFIDENZIALITÀ CIFRATO.

- A symmetric cipher is secure *iff* for each pair (p, c) then (IF AND ONLY IF)
  - given the ciphertext c, it is "difficult" to determine the corresponding plaintext p without knowing the key k, and vice versa — OSCAR SCOPRE TUTTI I MESSAGGI CIFRATI
  - given a pair of ciphertext c and plaintext p, it is "difficult" to determine the key k, unless it is used just once — CON DIFFICOLTÀ SI INTENDE L'OPPOSTO DI EFFICIENTE

↳ SUPPONIAMO FACCIA CADERE OSCAR POSSA CONOSCERE SIA IL TESTO CRIPTATO CHE QUELLO NON CIFRATO

NON VOGLIAMO IN QUESTO CASO PERCHÉ NON USIAMO MAI UNA CHIAVE!

8

An historical example

Mono-alphabetic substitution

Cleartext alphabet	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
Key	J	U	L	I	S	C	A	E	R	T	V	W	X	Y	Z	B	D	F	G	H	K	M	N	O	P	Q

P = "TWO HOUSEHOLDS, BOTH ALIKE IN DIGNITY,
IN FAIR VERONA, WHERE WE LAY OUR SCENE"
("Romeo and Juliet", Shakespeare)

P' = "TMOHO USEHO LDSBO THALI KEIND IGNIT
YINFA IRVER ONAWH EREWE LAYOU RSCEN E"

C = "HNZEZ KGSEI WIGUZ HEJWR VSRYI RAYRR
PRYCJ RFMSF ZYJNE SFSNS WJPZK FGLSY S"

9

First Attack

- Brute force attack (exhaustive key search)
  - Oscar has ciphertext (y) and some plaintext (x)
  - Oscar tries all possible keys
    - **for each** k in K
      if (y == E(k, x)) **return** k
- The attack is *always possible*
- The attack may be more complicated because of *false positives*

10

An historical example

Mono-alphabetic substitution

- The key is a permutation of the alphabet
- **Encryption algorithm**
  - every cleartext character having position *p* in the alphabet is substituted by the character having the same position *p* in the key
- **Decryption algorithm**
  - every ciphertext character having position *p* in the key is substituted by the character having the same position *p* in the cleartext
- **Number of keys** = 26! – 1 ≃ 4 ×10^{26} (number of seconds since universe birth!)

11

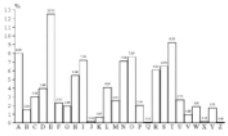
An historical example

- Brute force attack is practically infeasible given the enormous key space
- Brute force attack considers the cipher as a black box
- The monoalphabetic substitution algorithm by an analytical attack which analyze the internals of the algorithm

12

An historical example

- The monoalphabetic-substitution cipher maintains the redundancy that is present in the cleartext
- It can be "easily" crypto-analized with a ciphertext-only attack based on language statistics



Frequency of single characters in English text

13

An historical example

- The following properties of a language can be exploited
  - The frequency of letters
  - Generalize to pairs or triples of letters
  - Frequency of short words
    - If word separators (blaks) have been identified

14

Lesson learned

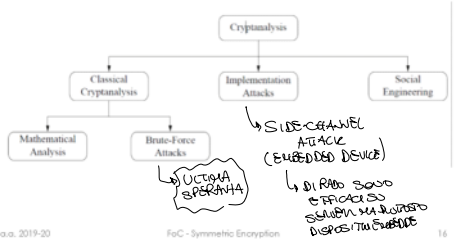
- Good ciphers should hide statistical properties of the encrypted plaintext
- The cyphertext symbols should appear to be random
- A large key space alone is not sufficient for strong encryption function (necessary condition)

CONDIZIONE NECESSARIA MA NON SUFFICIENTE

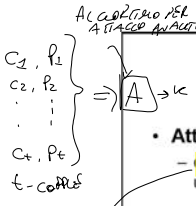
SE L'ALGORITMO È RESISTENTE AD ANALISI STATISTICHE ALLORA È VERO CHE PIÙ È GRANDE LA CHIAVE PIÙ È SICURO L'ALGORITMO

15

Crptanalysis



16



Attack Complexity

- **Attack complexity** is the dominant of:
  - **data complexity** — expected number of input data units required
    - Ex.: exhaustive data analysis is $O(2^n)$
  - **storage complexity** — expected number of storage units required
  - **processing complexity** — expected number of operations required to processing input data and/or fill storage with data
    - Ex.: exhaustive key search is $O(2^k)$

a.a. 2019-20 FoC - Symmetric Encryption 17

NUMERO DI COPPIE NECESSARIE PER OTTENERE LA CHIAVE: t

SE LE COPPIE SONO TUTTE UGUALI STESSO RAGIONAMENTO MA USANDO VARIABILI DIVERSE

NUMERO DI STEP DELL'ALGORITMO

STORAGE COMPLEXITY

DATA COMPLEXITY

STORAGE COMPLEXITY

SE UNA DELLE 3 È STRUTTURATE ALLORA IL CIFRARIO È SICURO E DIFFICILE DA DECIFRARE SENZA CHIAVE

Types of attacks

- Attacks are classified according to what information an adversary has access to
  1. **ciphertext-only attack** (the least strong)
  2. **known-plaintext attack**
  3. **chosen-plaintext attack** (the strongest)
- Fact.
  - A cipher secure against CPAs is also secure against COAs and KPAs
- Best practice.
  - It is customary to use ciphers resistant to a CPA even when mounting that attack is not practically feasible

a.a. 2019-20 FoC - Symmetric Encryption 18

18

1. $C_1, C_2 \dots C_t$ CONOSCIUTE
2. $(C_1, P_1), (C_2, P_2), \dots, (C_t, P_t)$ → A → k
   KNOWN PLAINTEXT ATTACK
3. CHIA È L'AVVERSARIO CHE SCEGLIE QUALE P_i FAR CONOSCERE

Kerchoff's principle
(19th century)

- Kerchoff's maxim
  - *A cryptosystem should be secure even if everything about the system, except the key, is public knowledge*
- Shannon's maxim
  - *The enemy knows the system*
- Pros
  - Maintaining security is easier
    - Keys are small secrets
      - Keeping small secrets it's easier than keeping large secrets
      - Replacing small secrets, once possibly compromised, is easier than replacing large secrets

BASTA CAMBIARE LA CHIAVE INVECE DI CAMBIARE L'ALGORITMO

a.a. 2019-20 FoC - Symmetric Encryption 19

19

Security through Obscurity

- Security through Obscurity
  - Attempt to use secrecy of design or implementation to provide security
- History shows that StO doesn't work
  - GSM/A1 disclosed by mistake
  - RC4 disclosed deliberately
  - Enigma disclosed by intelligence
  - ... many others…
- Defense in Depth
  - Solely relaying on StO is a poor design decision
  - StO is a valid secondary measure

a.a. 2019-20 FoC - Symmetric Encryption 20

20

Security through Obscurity

- Hiding security vulnerabilities in algorithms, software, and/or hardware decreases the likelihood they will be repaired and increases the likelihood that they can and will be exploited by evil-doers. Discouraging or outlawing discussion of weaknesses and vulnerabilities is extremely dangerous and deleterious to the security of computer systems, the network, and its citizens.
  [S. Bellovin, Steven, R. Bush, (February 2002), *Security Through Obscurity Considered Dangerous*, Internet Engineering Task Force (IETF), retrieved February 27, 2019]

21

Things to remember

- Cryptography is
  - a very useful tool
  - the basis for many mechanisms
- Cryptography is not
  - The solution to all security problems
    - Software bugs
    - Social engineering
  - Reliable if designed, implemented and used properly
    - WEP, Heartbleed,…
  - Something you should try to invent yourself

22

Symmetric Encryption

EXERCISES

23

Shift Cipher (Caesar Cipher)

- Shift every plaintext letter by a fixed number of positions (the key) in the alphabet with wrap around
- Ex.
  - PT = «ATTACK»
  - K = 17
  - CT = "RKKRTB"

24

Shift Cipher (Caesar Cipher)

- Letters are encoded as numbers
  - A: 0, B: 1, C: 2, …, Z: 25
- PT and CT are elements of the ring Z_{26}
  - Ecryption: $y = x + k \bmod 26$
  - Decryption: $x = y - k \bmod 26$
  - EX.
    - Pt = «ATTACK» => 0 19 19 0 2 10
    - K = 17
    - Ct = 17 10 10 17 19 1 => "RKKRTB"

25

Shift Cipher (Caesar Cipher)

- Possible attacks
  - Brute force attack
    - Small key space: 26 possible keys
  - Anlytical attack
    - Letter frequency analysis

26

Affine cipher

- Definition
  - Let a, b, x, y $\in \mathbb{Z}_{26}$
  - Encryption: $y = a \cdot x + b \bmod 26$
  - Decryption: $x = a^{-1} (y - b) \bmod 26$
  - With k (a, b) and gcd(a, 26) = 1
- Example
  - Plaintext: «ATTACK» => 0, 19, 19, 0, 2, 10
  - k = (9, 13)
  - Ciphertext: 13, 2, 2, 13, 5, 25 => «NCCNFZ»

27

Affine cipher

- Attacks
  - Brute force attack
    - Key space = (#values for a) × (#values for b) = 12 × 26 = 312
  - Analytical attack
    - Letter frequency analysis

28

Towards a secure cipher

- Attacker ability: cipher-text only
- Possible security requirements
  - Attacker cannot recover secret key
  - Attacker cannot recover plaintext
- Shannon's idea
  - Cipher-text should not reveal any information about plaint-text

a.a. 2019-20 FoC - Symmetric Encryption 29

29

Perfect secrecy
(Shannon, 1949)

- A cipher (E, D) defined over $(\mathcal{K}, \mathcal{P}, C)$ has **perfect secrecy** iff

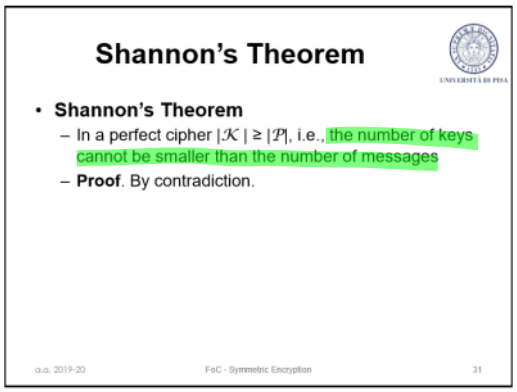
  $\forall p \in \mathcal{P}, c \in \mathcal{C} : \Pr(P = p \mid C = c) = \Pr(P = p)$

  where P is a random variable in $\mathcal{P}$ and C is a random variable in $\mathcal{C}$

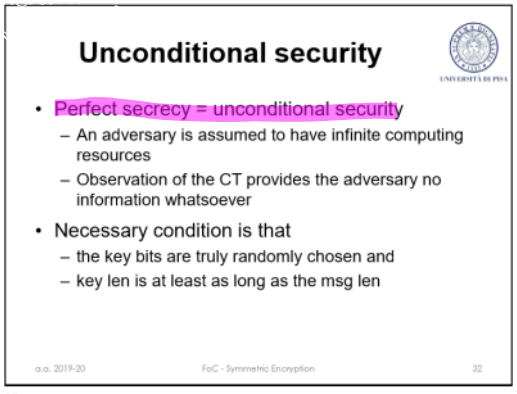
Information theoretical secure cipher
Unconditionally secure cipher

a.a. 2019-20 FoC - Symmetric Encryption 30

30



31



32

Shannon's theorem

Theorem 1. *In a perfect cipher $|\mathcal{K}| \geq |\mathcal{P}|$, i.e., the number of keys cannot be smaller than the number of messages.*

Proof. The proof is by contradiction.

First, let us assume that $|\mathcal{K}| < |\mathcal{P}|$. Then, let us observe that it had better be the case that $|\mathcal{C}| \geq |\mathcal{P}|$ or, otherwise, the cipher would not be an invertible (two plaintext messages would map into the same cipher-text message under the same key). It follows that

$$|\mathcal{C}| > |\mathcal{K}|. \tag{1}$$

Let us now look at the consequences of this inequality. Let us consider a p^* such that $Pr\{P = p^*\} \neq 0$. Let us encrypt p^* under every possible key. Since the number of keys is smaller than the number of ciphertexts because of inequality 1, then there must be a cipher-text, namely c^* that is not image of p^* under any key. If follows that $Pr\{P = p^*|C = c^*\} = 0$. It follows that there exists at least a pair (p^*, c^*), s.t. $Pr\{P = p^*|C = c^*\} \neq Pr\{P = p^*\}$. □

16

Perfect secrecy
(another definition)

- **Definition**. A cipher (E, D) over $(\mathcal{K}, \mathcal{P}, C)$ has perfect secrecy iff
  - $\forall m_1, m_2 \in \mathcal{P}, |m_1| = |m_2|, \forall c \in C,$ $Pr(E(k, m_1) = c) = Pr(E(k, m_2) = c)$, with $k \leftarrow$ random()

33