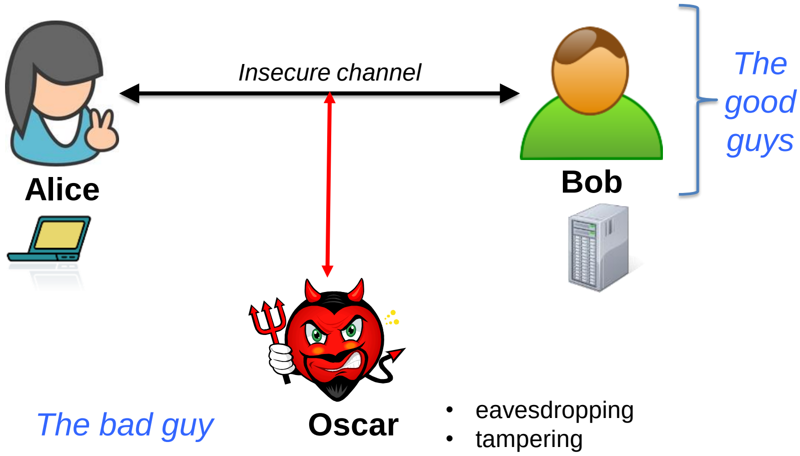
Symmetric Encryption

Gianluca Dini
Dept. of Ingegneria dell'Informazione
University of Pisa
gianluca.dini@unipi.it

Last version: 2021-03-01

1

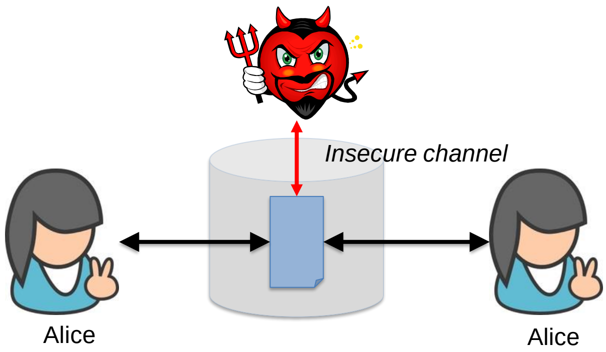
Main characters



- eavesdropping
- tampering

2

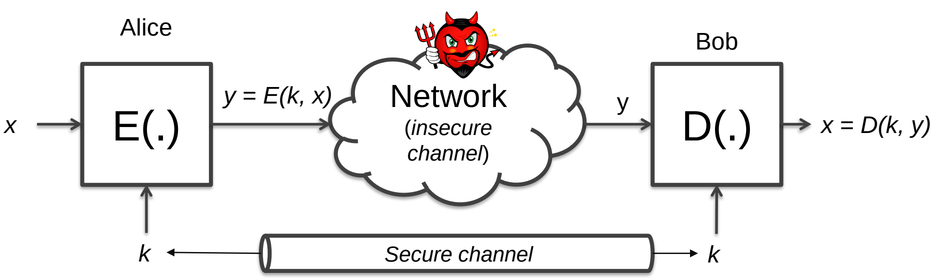
Encrypted files



- Analogous to secure communication
- *Alice-today* sends an encrypted message to *Alice-tomorrow*

3

The model



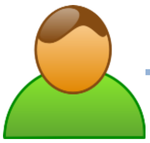
- *E*, *D*: cipher *k*: shared secret key (128 bits)
- *x*, *y*: plaintext, ciphertext

- Encryption algorithm is publicly known
  - Never use proprietary algorithm

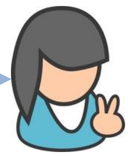
4

Example: DES (CBC)

UNIVERSITÀ DI PISA



f3 9e 8a 73 fc 76 2d 0f 59 43 bd 85 c3 c9
89 d2 bf 96 b6 4f 34 b8 51 dd



P = “Ci vediamo alle 20!”

***k* = 0x3dd04b6d14a437a9**

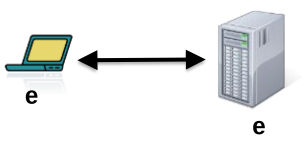
P = “Ci vediamo alle 20!”

***k* = 0x3dd04b6d14a437a9**

shared secret

Example: SSL

UNIVERSITÀ DI PISA



- **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

Cipher definition

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

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

- **Consistency Property**

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

  - *E* may be randomized; *D* is always deterministic

Security of a cipher (informal)

- A symmetric cipher is secure *iff* for each pair (p, c) then
  - given the ciphertext c, it is "difficult" to determine the corresponding plaintext p without knowing the key k, and vice versa
  - given a pair of ciphertext c and plaintext p, it is "difficult" to determine the key k, unless it is used just once

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

C = "HNZEZ KGSEZ WIGUZ HEJWR VSRYI RAYRH
PRYCJ RFMSF ZYJNE SFSNS WJPZK FGLSY S"

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*

An historical example

UNIVERSITÀ DI PISA

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** = $26! - 1 \backsimeq 4 \times 10^{26}$ (number of seconds since universe birth!)

An historical example

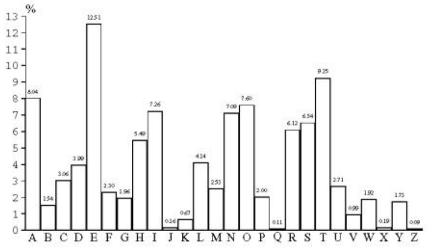
UNIVERSITÀ DI PISA

- 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

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



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

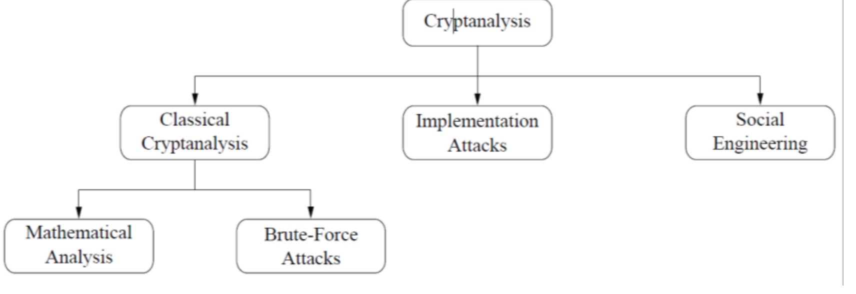
Lesson learned

UNIVERSITÀ DI PISA

- 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)

Crptanalysis

UNIVERSITÀ DI PISA



Attack Complexity

UNIVERSITÀ DI PISA

- **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 $O(2^k)$

Types of attacks

UNIVERSITÀ DI PISA

- Attacks are classified according to what information an adversary has access to
  - **ciphertext-only attack** (the least strong)
  - **known-plaintext attack**
  - **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

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

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

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]

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

Symmetric Encryption

EXERCISES

Shift Cipher (Caesar Cipher)

UNIVERSITÀ DI PISA

- 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"

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_6
  - 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”

Shift Cipher (Caesar Cipher)

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

Affine cipher

UNIVERSITÀ DI PISA

- Definition
  - Let a, b, x, y $\in$ Z26
  - Encryption: $y = a \cdot x + b \mod 26$
  - Decryption: $x = a^{-1} (y – b) \mod 26$
  - 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»

Affine cipher

UNIVERSITÀ DI PISA

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

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

Perfect secrecy
(Shannon, 1949)

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

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

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

Shannon's Theorem

UNIVERSITÀ DI PISA

- **Shannon's Theorem**
  - In a perfect cipher $|\mathcal{K}| \geq |\mathcal{P}|$, i.e., the number of keys cannot be smaller than the number of messages
  - **Proof**. By contradiction.

Unconditional security

UNIVERSITÀ DI PISA

- Perfect secrecy = unconditional security
  - An adversary is assumed to have infinite computing resources
  - Observation of the CT provides the adversary no information whatsoever
- Necessary condition is that
  - the key bits are truly randomly chosen and
  - key len is at least as long as the msg len

Perfect secrecy
(another definition)

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