UNIVERSITA' DEGLI STUDI DI PARMA
Dipartimento di Ingegneria dell'Informazione

Cryptography: Introduction

Luca Veltri

(mail.to: luca.veltri@unipr.it)

Corso di Sicurezza nelle reti di telecomunicazioni, a.a. 2004/2005

http://www.tlc.unipr.it/veltri

Cryptography

- cryptography <--> Greek: krupto+grafη (hidden/secret+writing)

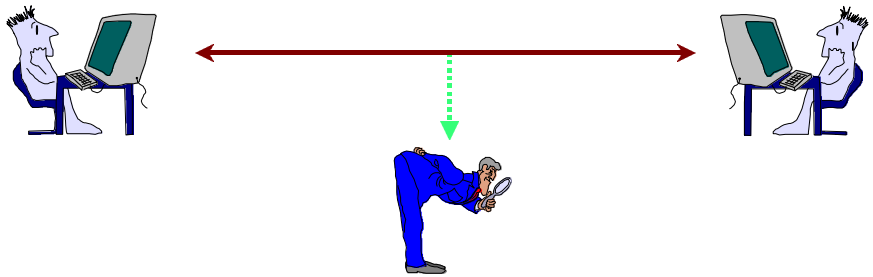
  the study of mathematical techniques related to information security that have the following objectives:
  - **Confidentiality**
    - ensuring information is accessible only by authorized persons
  - **Data integrity**
    - ensuring information is has not been altered by unauthorized or unknown means
  - **Authentication**
    - verification of the identity of an entity
  - **Non-repudiation**
    - preventing the denial of previous commitments or actions
- the most widely used tool for securing information and services
- it is one tool (not the only)

Crittografia e Crittoanalisi

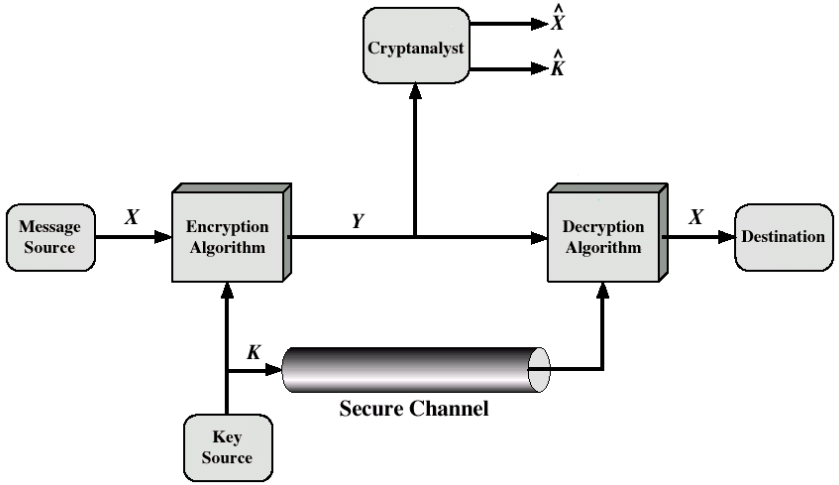
- Crittologia:
  - **scienza che ha lo scopo di studiare comunicazioni sicure**
- Crittografia:
  - **branca della crittologia che ha come scopo la progettazione di algoritmi di cifratura e decifratura, al fine di garantire la segretezza e o l'autenticità dei messaggi**
- Crittoanalisi:
  - **branca della crittologia che ha come scopo l'analisi di un cifrario per risalire all'informazione originaria, e/o la generazione di informazione cifrata contraffatta che possa essere accettata come autentica**

Cryptography

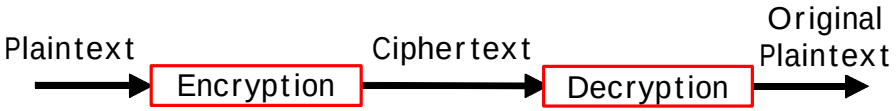
- Consiste nell'alterazione controllata di un messaggio (sequenza alfanumerica di caratteri) in maniera da renderlo non comprensibile a chi non dispone degli strumenti adeguati



Modello di sistema crittografico tradizionale



Cryptography: Basic Terminology



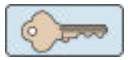
- **plaintext** - the original message
- **ciphertext** - the coded message
- **cipher** - algorithm for transforming plaintext to ciphertext
- **key** - info used in cipher known only to sender/receiver
- **encipher (encrypt)** - converting plaintext to ciphertext
- **decipher (decrypt)** - recovering ciphertext from plaintext

Semplice esempio di algoritmo crittografico (cifraro con shift)

Algoritmo: $X \leftarrow M+K \bmod 26$ $K \in \{0,1,\ldots,25\}$

Testo in chiaro: C A S A

Testo cifrato: R P H P

Chiave: K = 15 

Caesar cipher:

Chiave K=3

Cryptographic algorithms

- Funzioni matematiche usate per cifrare e decifrare un testo
- Possono essere caratterizzati da:
  - **type of encryption operations used**
    - substitution / transposition / product
  - **number of keys used**
    - single-key or private / two-key or public
  - **way in which plaintext is processed**
    - block / stream

Security of cryptographic algorithms

- The security of a cipher might rest in the secrecy of its *restricted* algorithm, however:
  - **whenever a users leaves a group, the algorithm must change**
  - **could be scrutinized by people smarter than you**

- Modern cryptography relies on *keys, a selected value from a large set (a keyspace), e.g., a 1024-bit number. 2^{1024} values!*
  - **Change of authorized participants requires only a change in key**

Kerckhoffs' principle

Security should be based on secrecy of the key,
not the details of the algorithm

(La sicurezza di un crittosistema, o cifrario,
deve dipendere solo dalla segretezza della chiave
e non dalla segretezza dell'algoritmo usato)

Jean Guillaume Hubert Victor Francois Alexandre Auguste Kerckhoffs von Nieuwenhof (1835-1903), filologo olandese, *"La Criptographie Militaire"* [1883]

Cryptography Attacks

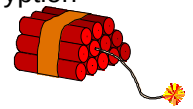
- Brute-force search
  - **si tenta ogni possibile chiave su un frammento di testo in chiaro**
  - **in media per avere successo occorre provare la metà delle chiavi possibili**
  - **deve essere possibile riconoscere quando si è trovato il testo in chiaro corretto**

- Cryptographic analysis
  - **Si basa sulla natura dell'algoritmo e sfrutta qualche conoscenza delle caratteristiche generali del testo in chiaro e/o qualche esempio di coppie testo in chiaro/testo cifrato**
    - non richiede necessariamente di scoprire la chiave

- The loss of a key without cryptoanalysis is called a "compromise"

Types of Cryptanalytic Attacks

- What do we mean with "a bad guy breaks an encryption scheme"?



- There are three basic attacks:
  - **Ciphertext-only attack**
    - The attacker has to recover the plaintext from only the ciphertext
  - **Known-plaintext attack**
    - Portions of the cipher are known as plaintext. The rest may be easier to recover
  - **Chosen-plaintext attack**
    - The attacker can choose what plaintext to encrypt, again making it easier to recover other ciphertext
  - **Chosen-ciphertext attack**
    - The attacker can choose ciphertext and obtain the plaintext

Ciphertext only - Attack

- The bad guy has seen (and presumably stored) some ciphertext that can be analyzed
- One possible strategy to figure out the plaintext is to try all keys
  - **it is essential that he/she is able to recognize when he/she has succeeded (often called *recognizable plaintext* attack)**
    - (attacco a parole probabili)
    - for example in case of normal text or known document formats (e.g. PostScript, etc.) with recognizable patterns
  - **it is necessary to have enough ciphertext**
- E' l'attacco più difficile da realizzare

Known plaintext - Attack

- The bad guy knows a <pliantext, ciphertext> pair
- How it is possible to obtain it...
  - **the secret data does not remain secret forever (e.g. the name of an attacked city)**
- From that pairs, the attacker can try to figure out the mapping of some fraction of the text
- Some cryptographic schemes might be good enough to be secure against *ciphertext only* attacks but not against to *known plaintext* attacks
  - **in these cases, it is important to minimize the possibility for a bad guy to obtain <pliantext, ciphertext> pairs**

Chosen plaintext (or ciphertext) - Attack

- The bad guy can choose any plaintext and get the corresponding ciphertext from the system (or the contrary)
  - **e.g. there is a telegraph service that encrypt and transmit messages; the bad guy can ask the telegraph company to transmit any plaintext he/she wants**
- Some cryptographic schemes might be good enough to be secure against *ciphertext only* attacks and *known plaintext* attacks but not against to *chosen plaintext* attacks

Computational and Unconditional Security

- **computational security**
  - **given limited computing resources (e.g. time needed for calculations is greater than age of universe, or the cost required for the attack is not affordable), the cipher cannot be broken**
- **unconditional security**
  - **no matter how much computer power is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext**

Computational Difficulty

- Cryptographic algorithms should be reasonably efficient to compute for good guys (who know the keys)
- Cryptographic algorithms are not impossible to break without the key: a bad guy can simply try all possible keys until one works
- The security depends on how much work is necessary to break it
  - **It is the *complexity* of launching the attack that secures us**
- Attack complexities:
  - **data complexity: a large number of expected inputs (e.g., ciphertext)**
  - **processing complexity: a large number of operations required**
  - **storage complexity: a large amount of storage units required**
- Often a scheme can be made more secure by making the key longer

Brute Force Search

- always possible to simply try every key
- most basic attack, proportional to key size
- assume either know / recognise plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/μs	Time required at 10^6 encryptions/μs
32	$2^{32} = 4.3 \times 10^9$	$2^{31}\ \mu s = 35.8$ minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55}\ \mu s = 1142$ years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127}\ \mu s = 5.4 \times 10^{24}$ years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	$2^{167}\ \mu s = 5.9 \times 10^{36}$ years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26}\ \mu s = 6.4 \times 10^{12}$ years	6.4×10^{6} years

Cryptanalysis: Example (Language Redundancy)

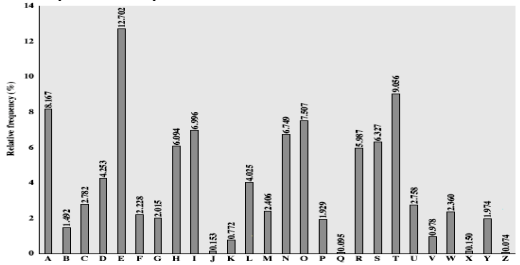
- Example of cryptoanalysis:
  - **cleartext: normal text**
  - **cryptography algorithm: monoalphabetic substitution**
- key concept:
  - **human languages are redundant: do not change relative letter frequencies**
  - **in English e is by far the most common letter**
  - **then T,R,N,I,O,A,S**
  - **other letters are fairly rare (e.g. Z,J,K,Q,X)**
- cryptoanalysis:
  - **based on lenguage redudancy**
  - **discovered by Arabian scientists in 9th century**
  - **calculate letter frequencies for ciphertext**
  - **compare counts/plots against known values**
  - **have tables of single, double & triple letter frequencies**

Cryptanalysis: Example

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- count relative letter frequencies (see text)



- guess P & Z are e and t
- guess ZW is th and hence ZWP is the
- proceeding with trial and error finally get:

```
it was disclosed yesterday that several informal but
direct contacts have been made with political
representatives of the viet cong in moscow
```

Principio fondamentale della crittografia

- Cryptographers invent clever secret codes
- Cryptoanalysts attempt to break these codes
- The two disciplines help each other!

Fundamental Tenet of Cryptography:

If lots of people have failed to solve a problem,
then it probably won't be solved (soon).

"Secondo principio fondamentale"

- Often breaking a cryptographic scheme is not the the only way of getting what you want!

You can get further with a kind word and a gun
than you can with a kind word alone.

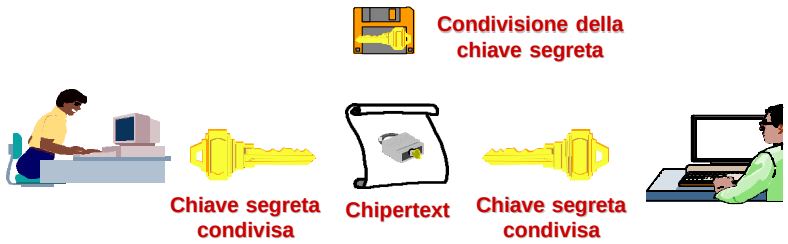
- Willy Sutton, bank robber

Differenti tipi di crittografia

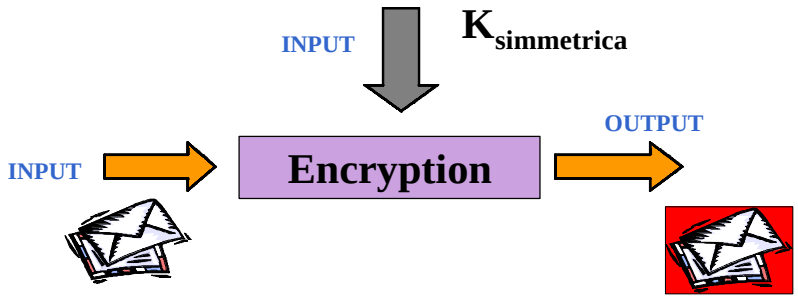
- A chiave segreta (simmetrica):
  - **le due parti che comunicano condividono un segreto (la chiave crittografica)**
- A chiave pubblica (asimmetrica):
  - **la chiave crittografica è composta da due parti, una che tutti conoscono (chiave pubblica) e una che solo l'interessato conosce (chiave privata)**
- Hash algorithm (message digest/one way transformation):
  - **una funzione hash è una trasformazione matematica in una sola direzione che a partire da un messaggio arbitrario (lunghezza variabile) genera messaggio/numero di dimensione fissata**

Crittografia a chiave segreta (simmetrica)

- La chiave utilizzata per cifrare è la medesima chiave utilizzata per decifrare (chiave simmetrica o Secret Key, Ks)
- Alcune volte detta: crittografia convenzionale o simmetrica



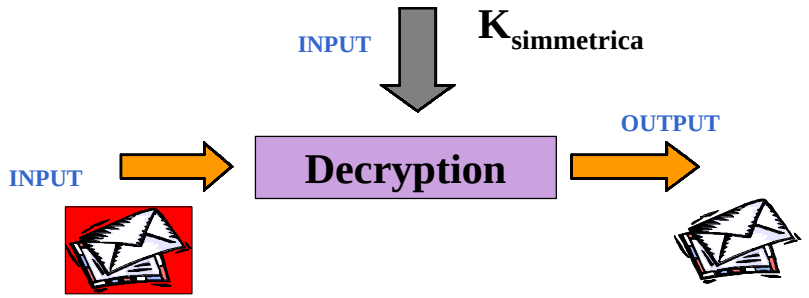
Schema di funzionamento: Cifratura



Messaggio in chiaro

Messaggio cifrato

Schema di funzionamento: Decifratura



Messaggio cifrato

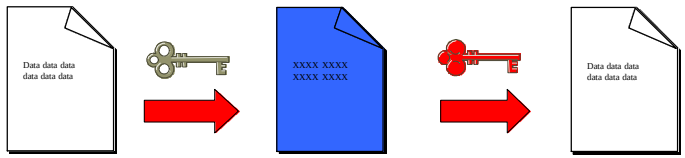
Messaggio in chiaro

Caratteristiche

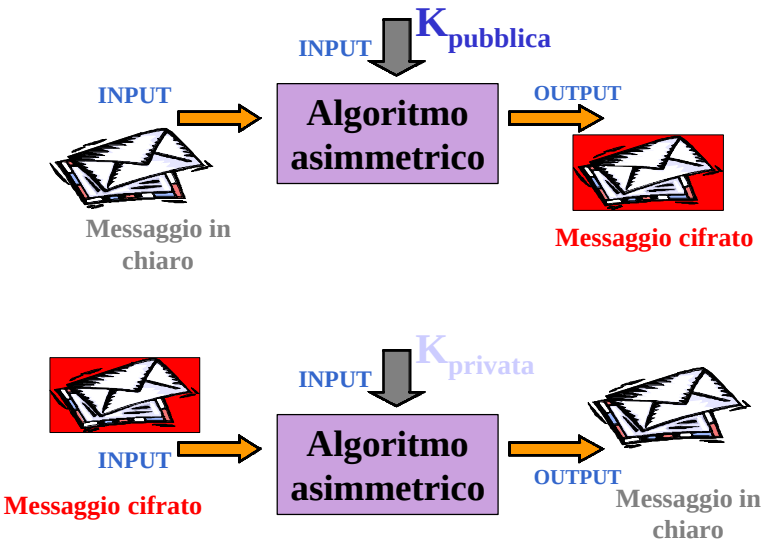
- Richiede una fase iniziale in cui ciascuna coppia di interlocutori si scambia la secret key in maniera sicura
- Il numero delle chiavi per realizzare una comunicazione reciproca tra N utenti (dispositivi) è pari a Nx(N-1)/2 (se le chiavi rimangono sempre le stesse)
- Viene generalmente utilizzato per proteggere, mediante codifica, informazioni (file) in un repository locale o trasmessi
- La robustezza dell'algoritmo è normalmente misurata dalla lunghezza delle chiavi: 40 bit (debole), 128 bit (forte)
- Algoritmi più diffusi: DES, 3DES, RC2, RC4, IDEA, AES

Crittografia a chiave pubblica (asimmetriche)

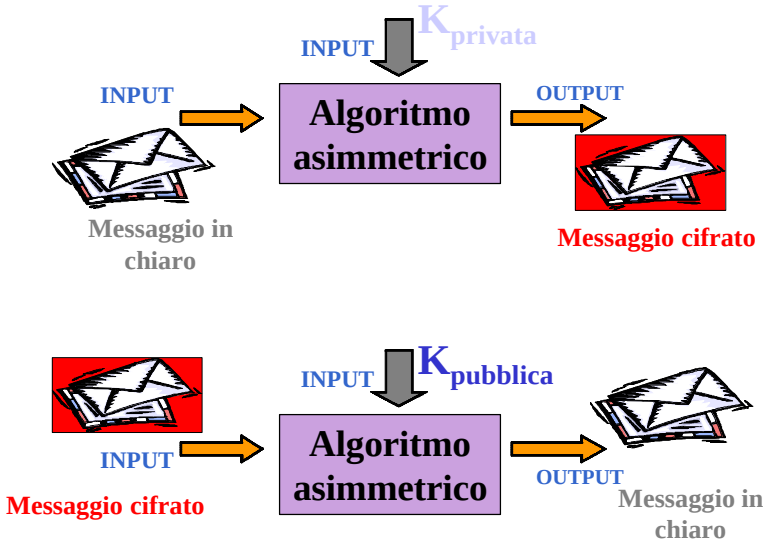
- Spesso detta crittografia asimmetrica
- La chiave utilizzata per cifrare (chiave di cifratura, Ke) è diversa dalla chiave utilizzata per decifrare (chiave di decifratura, Kd)
- Le chiavi si usano in coppie, di cui una detta privata (segreta) e l'altra detta pubblica (disponibile a tutti)
- Un clear-text cifrato con la chiave privata può essere decifrato solo con la chiave pubblica e viceversa
- Ciascuna coppia di chiavi è caratterizzata da proprietà peculiari



Schemi di funzionamento: Io caso



Schemi di funzionamento: IIo caso



Caratteristiche

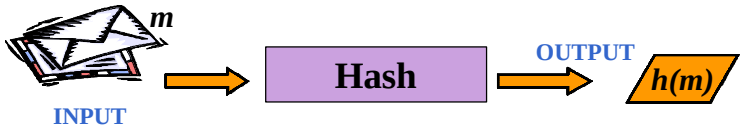
- Può richiedere una fase iniziale in cui gli interlocutori si scambiano le rispettive chiavi pubbliche
- Il numero delle chiavi è proporzionale a N per la comunicazione reciproca tra N utenti (dispositivi)
- Viene generalmente utilizzato per distribuire chiavi simmetriche in un ambiente distribuito
- Algoritmi più diffusi: RSA, Diffie-Hellman, DSA
- La robustezza del sistema dipende anche dal sistema di certificazione delle chiavi di cifratura (se esiste..)
  - **L'impiego della crittografia a chiavi asimmetriche comporta l'implementazione di una PKI (Public Key Infrastructure)**

algoritmi di cifratura simmetrici e assimmetrici

- Entrambi gli algoritmi, simmetrici o asimmetrici, consentono di cifrare un messaggio; gli algoritmi simmetrici sono preferiti nei casi in cui sia necessaria una certa velocità di esecuzione della cifratura
- Un buon sistema crittografico usa al meglio i due tipi di algoritmo:
  - **Uso di algoritmo asimmetrico per lo scambio sicuro di chiavi e per l'autenticazione**
  - **Uso di algoritmo simmetrico per lo scambio dei dati**

Hash Algorithms

- Also known as message digests or one way transformations
- Hash function is a mathematical transformation *m* that takes a message of arbitrary length and computes a fixed-length (short) number/string *h(m)*
- Properties:
  - **for any message *m*, it relatively easy to compute *h(m)***
  - **given *h(m)* there is no way to find *m* in a way easier than computing all possibilities**
  - **it is computationally infeasible to to find two values that hash to the same thing**



- Examples of hash algorithms: MD2, MD5, SHS

Dove la crittografia non ci aiuta

- Protezione di sistema:
  - **Accesso non autorizzato alle risorse HW e SW**
  - **Alterazione dei programmi SW (virus, ecc.)**
  - **Attacchi mirati a rendere inutilizzabili i dispositivi ed i programmi**
  - **...**
- Protezione dei servizi:
  - **Attacchi mirati a rendere indisponibili i servizi**
  - **Alterazione dei contenuti informativi di database**
  - **Utilizzo illecito di servizi ed informazioni**