CVMLab
Computer Vision
& Multimedia Lab

La crittografia

- Motivazioni
- Obiettivi
- Terminologia
- Storia
- Steganografia

Università
degli Studi
di Pavia

Storia della crittografia

- Sparta (Plutarco) scytala
- Cifrario di Cesare (Svetonio)
- Medioevo in Oriente
- Rinascimento in Italia
  - Cicco Simonetta (Sforza) → Primo trattato di decrittazione
  - Serenissima (sala dei segreti)
  - Roma (disco di Leon Battista Alberti) - De cifris
- XVII - XVIII secolo (Vienna, le camere nere)

Storia della crittografia

- Ottocento: Kasiski, Kerckhoffs e Babbage
  - Francia
- La seconda guerra mondiale
  - Codice Enigma (Tedesco)
  - A. M. Turing
- Inizio della criptografia moderna (1949): Claude Shannon pubblica Communication Theory of Secrecy Systems su Bell System Technical Journal
- Teoria dell'informazione e informatica
- Reti, algoritmi a chiave segreta e pubblica

Crittanalisi

- Le ipotesi fondamentali della crittanalisi sono due:
  - Eventuali attaccanti hanno una perfetta conoscenza dell'algoritmo utilizzato per cifrare il messaggio e di tutti i dettagli della sua realizzazione.
  - Eventuali attaccanti hanno completo accesso al canale di comunicazione e possono pertanto intercettare, interrompere, creare o modificare qualsiasi flusso di dati.
- I possibili attacchi vengono suddivisi nelle classi:
  - Ciphertext-only attack
  - Known-plaintext attack
  - Chosen-plaintext attack

Scopi della crittografia

- Si vuole garantire:
  - Confidenzialità: le informazioni sono accessibili solo da persone autorizzate
  - Autenticazione: l'identità dell'interlocutore è garantita
  - Integrità: garanzia della non alterazione dell'informazione
  - Non ripudiabilità: garanzia che nessun soggetto della comunicazione possa disconoscere di esserne l'autore

Criteri generali

- L'algoritmo di cifratura è generalmente noto (principio di Kerckhoffs)
- Nessun sistema è assolutamente sicuro
- Si deve rendere praticamente irrealizzabile l'attacco
  - Sistemi teoricamente sicuri (es. one-time pad) non praticabile come soluzione
  - Sistemi computazionalmente sicuri: è antieconomico tentare di aggirare le protezioni

Sistemi computazionalmente sicuri

- Il valore delle informazioni contenute nei messaggi cifrati non deve mai superare i costi stimati per violare l'algoritmo utilizzato
- Il periodo temporale durante il quale le informazioni cifrate devono essere mantenute confidenziali non deve superare il tempo stimato necessario per violare l'algoritmo

Terminologia

- Criptologia
  - scienza che studia i messaggi segreti
- Crittografia (κριπτοσγραφια): studio dei metodi per rendere un messaggio non intelleggibile a chiunque non sia il legittimo destinatario
  - Lo scopo NON è quello di nascondere un messaggio o di dissimularlo (steganografia)
- Crittanalisi: studio dei metodi per violare il segreto di un messaggio cifrato
  - Testo in chiaro vs. testo cifrato
  - Crittografo vs. crittanalista
  - Cifratura vs. decifratura oppure decrittazione

Steganografia

- Lo scopo è nascondere l'esistenza del messaggio
- Marcatura di caratteri: marcare caratteri con inchiostro speciale su un testo scritto o stampato su carta
- Inchiostro invisibile
- Perforazioni invisibili su carta

Steganografia: esempio

- Il formato Kodak Photo CD alla massima risoluzione visualizza 2048 X 3072 pixel a 24 bit.
  - Modificando a piacere il bit meno significativo posso nascondere 2.3 Mbyte di messaggio in una sola immagine
  - L'immagine però occupa 18Mbyte

Steganografia

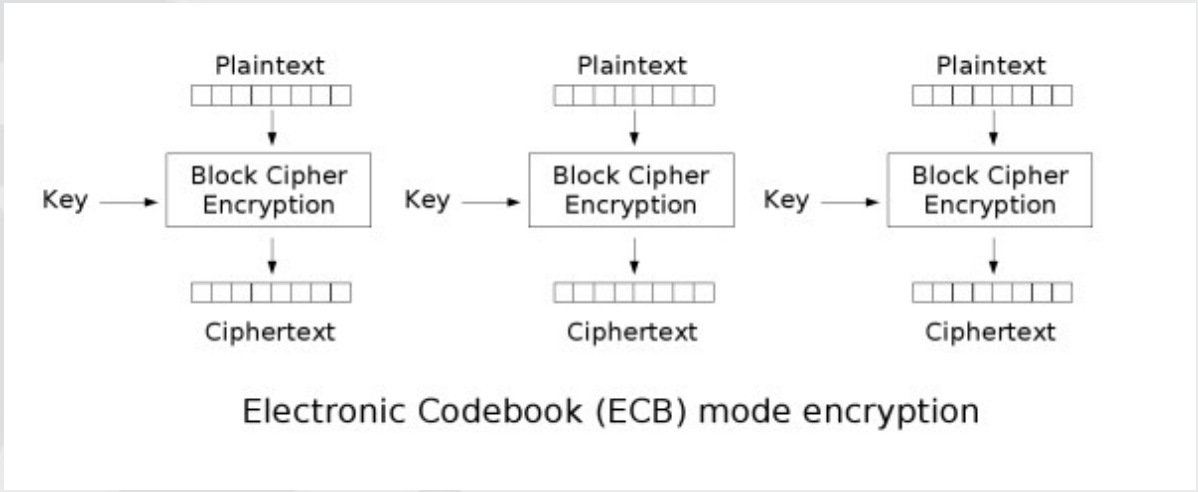
- Svantaggi:
  - richiede molti dati per nascondere pochi bit di informazione
  - una volta scoperto il meccanismo, è da buttare
  - può essere sfruttato se le due parti che comunicano devono nascondere la loro connessione, piuttosto che il messaggio stesso
  - applicazione nel copyright
  - watermarking nelle immagini

Watermarking

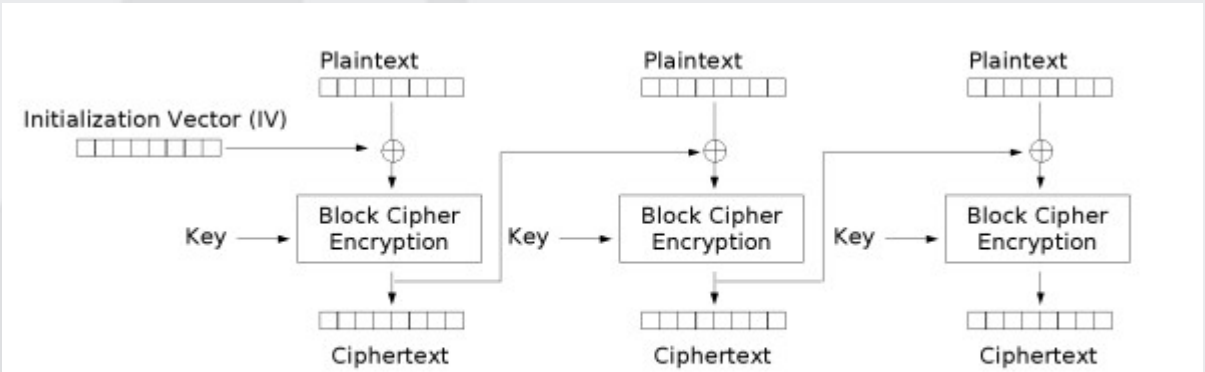
Crittografia

- **I sistemi crittografici sono generalmente classificati in base a tre criteri:**
  - **il tipo di operazioni per passare da testo in chiaro a testo cifrato (sostituzioni, trasposizioni, ecc.)**
  - **Il numero di chiavi usate (le funzioni di cifratura e di decifratura utilizzano una o più chiavi K per produrre il risultato).**
    - Algoritmi simmetrici o asimmetrici.
  - **Il modo in cui si elabora il testo in chiaro: a blocchi o a stream (sw, hw o real-time)**

Crittografia a blocchi



Electronic Codebook (ECB) mode encryption



Cipher Block Chaining (CBC) mode encryption

Crittografia

- Si distinguono due campi della crittografia: crittografia convenzionale e a chiave asimmetrica
- Utilizzando una notazione matematica ed indicando con M il testo in chiaro, con C il testo cifrato, con E( ) la funzione di cifratura e con D( ) quella di decifratura, un sistema convenzionale basato su una sola chiave k può essere descritto dalle equazioni:
  - $E_K(M) = C$
  - $D_K(C) = M$
- Con la proprietà che:
  - $D_K(E_K(M)) = M$
- Si suppone che un crittanalista conosca E e D, e cerchi di stimare M, K o entrambe.

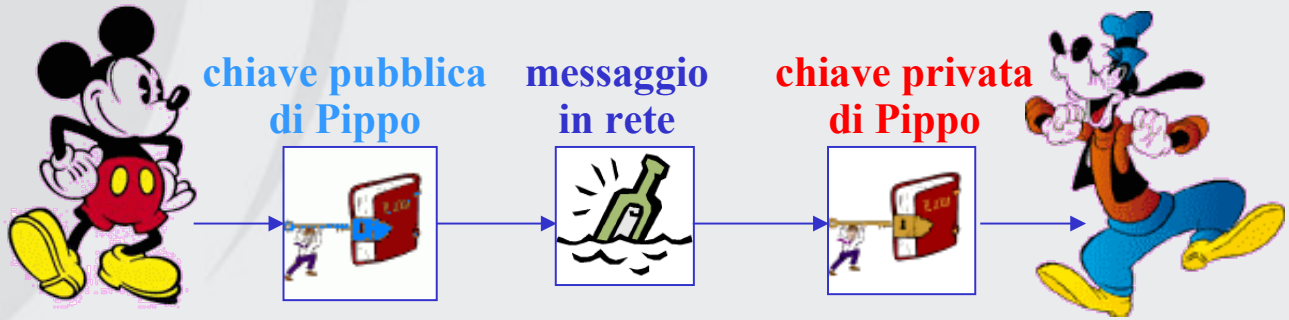
Cifratura a Chiave Simmetrica

Problema: come condividere la chiave

Il problema della scalabilità

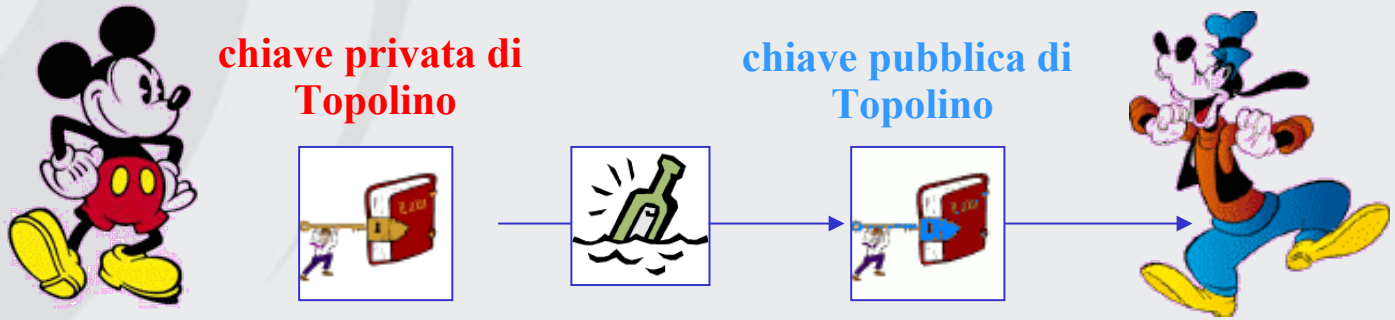
Cifratura a chiave pubblica

- Topolino vuole mandare un messaggio segreto a Pippo
  - Topolino usa la chiave pubblica di Pippo per cifrare il messaggio
  - Solo Pippo è in grado di decodificare il messaggio
  - Pippo tuttavia non può essere sicuro dell'identità di Topolino



Messaggi firmati

- Topolino vuole mandare un messaggio firmato a Pippo
  - Topolino usa la sua chiave privata per cifrare il messaggio
  - Pippo è in grado di decodificare il messaggio usando la chiave pubblica di Topolino
  - Solo Topolino poteva inviare quel messaggio
  - Il messaggio però non è segreto tutti lo possono leggere



Messaggi firmati

chiave pubblica
di Pippo

Hash( )

D(Hash( ))

chiave privata di Topolino

Messaggio in rete

chiave privata
di Pippo

Hash( )

D(Hash( ))

chiave pubblica di Topolino

Sono uguali?

Hash( )

Funzioni hash

- Dato un messaggio di lunghezza arbitraria produce una stringa di lunghezza predefinita
- Data una stringa hash è difficile trovare un messaggio compatibile
- Funzioni hash comuni:
  - MD5 (Message Digest Rivest 1992) 128 bit
  - SHA1 (Secure Hash Algorithm NIST 1995) 160 bit
    - Sha256 256 bit
    - Sha512 512 bit

Funzioni hash

```
$ md5sum <<< ciao
5f423b7772a80f77438407c8b78ff305 *-

$ md5sum <<< Ciao
Bba5159eba60f759a28b36834acf656c *-

$ sha1sum <<< ciao
953ed62a3246f2dbd96cdbfc0ec0d92b5cb2f5a8 *-

$ sha256sum <<< ciao
6f0378f21a495f5c13247317d158e9d51da45a5bf68fc2f366e450deafdc8302 *-

$ sha512sum <<< ciao
d380e3a08107af3a45bbe2539d9cc8d05a3eaf4a82a91bcc46bf8ca33fb72d37c2ec89893da
7ba76d9f2794155896760a23d5fe937de2e7a8cda52d0b8a0d62e *-
```

Problema: *uomo nel mezzo*

Tecniche convenzionali

- Cifrari a sostituzione: una lettera del testo in chiaro è sostituita da una o più lettere o numeri o simboli.
- Se il testo in chiaro è visto come una sequenza di bit, allora ciò implica la sostituzione di blocchi di bit (pattern) in chiaro con pattern di bit cifrati.
- Esempio storico: il cifrario di Cesare
  - chiaro:

    incontriamoci alle sette
  - ogni lettera è sostituita dalla lettera di tre posti successivi nell'alfabeto:

    LQFRQWULDPRFLDOOHVHWWH

Tecniche convenzionali

- Assumendo un valore numerico a ogni lettera, per ogni lettera del testo in chiaro p si sostituisce la lettera cifrata C tale che
  - C = E(p) = (p + 3) mod 26
  - C = E(p) = (p + k) mod 26 k assume valori da 1 a 25
  - La decifrazione è p = D(C) = (C - k) mod 26
- Possibile crittanalisi di tipo brute - force
  - gli algoritmi di E e D sono noti
  - la chiave k assume un numero di valori limitato
  - il linguaggio del testo in chiaro è noto

Cifrature monoalfabetiche

- Per aumentare lo spazio delle chiavi si esegue una sostituzione arbitraria

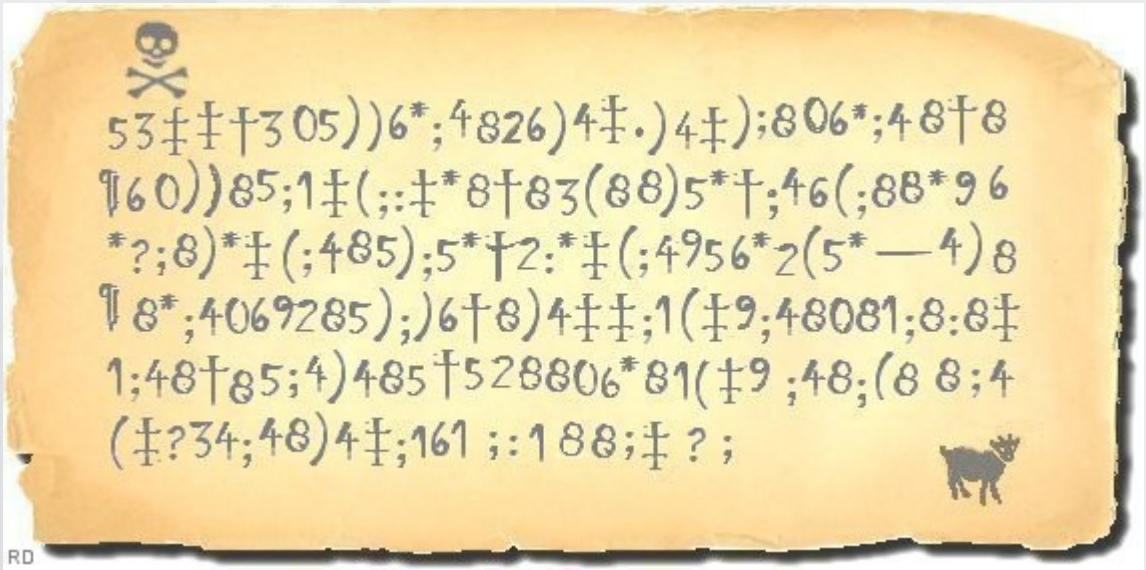
```
a b c d e f g h i l m n o p q r s t u v w x y z
h k l p o u t r e d f s w q a z c v b n m .....
```

- In questo caso il testo cifrato può essere ottenuto a una qualunque delle permutazioni di 26 caratteri, ovvero 26! = 4 X 10^{26} possibili chiavi.
- Non è ancora abbastanza sicuro perché si sfrutta la regolarità del linguaggio naturale

E 12.75	S 6.00	P 2.75	K 0.50
T 9.25	D 4.25	Y 2.75	X 0.50
R 8.50	H 3.50	G 2.00	Q 0.50
N 7.75	C 3.50	L 3.75	J 0.25
I 7.75	F 3.00	W 1.50	Z 0.25
O 7.50	U 3.00	V 1.50	
A 7.25	M 2.75	B 1.25	

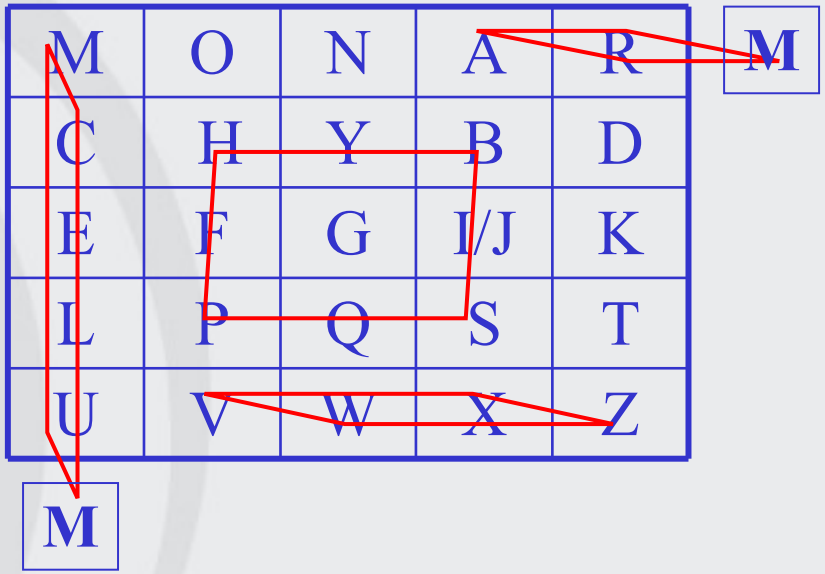
Cifrature monoalfabetiche

- Vedi Scarabeo d'oro - Poe



Cifrario Playfair

- Da una parola chiave si crea una matrice del tipo



Le doppie nel testo in chiaro sono separate da una lettera "filler"

cc → CZC

bp → HS

vx → WZ; ar → RM; mu → CM (righe e colonne si considerano periodiche)

Cifrari polialfabetici

- Partono da un insieme di cifrari monoalfabetici.
  - Una chiave determina quale cifrario usare
- Esempio: cifrario di Vigenére
  - Si tratta di una tabella di 26 cifrari di Cesare
  - Data una lettera chiave x e una lettera in chiaro y, la lettera cifrata corrispondente è quella corrispondente all'intersezione tra x e y

Cifrari polialfabetici

- Devo avere una chiave lunga quanto il testo da cifrare (soluzione: ripetizione)

```
chiave:           paviapaviapavia
testo chiaro:     dalleduealletre
testo cifrato:    sagtesuzilaeoze
```

```
A: ABCDEFGHIJKLMNOPQRSTUVWXYZ
B: BCDEFGHIJKLMNOPQRSTUVZWXYA
C: CDEFGHIJKLMNOPQRSTUVWXYZAB
...
I: IJKLMNOPQRSTUVWXYZABCDEFGH
...
P: PQRSTUVWXYZABCDEFGHIJKLMNO
...
V: VWXYZABCDEFGHIJKLMNOPQRSTU
```

Tecniche a trasposizione

- I cifrari a trasposizione non effettuano sostituzioni, ma una permutazione delle lettere del testo in chiaro (implementazione: macchine a rotori)

```
chiave:        4 3 1 2 5 6 7
chiaro:        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
cifrato: ttnaaptmtsuoaodwcoixknlypetz
```

Un cifrario perfetto

One-time pad, Gilbert Vernam, 1917

$$C = M \oplus K$$

0 1 1 0 1 0 0 1 0 1 1 1	messaggio
1 0 1 0 0 0 1 1 1 0 1 0	chiave (sequenza di bit casuale)
1 1 0 0 1 0 1 0 1 1 0 1	testo cifrato

M e C sono indipendenti (il testo cifrato non dà alcuna informazione utile sul messaggio)

Messaggio e chiave hanno la stessa lunghezza
La chiave si può usare una sola volta

Un cifrario perfetto

- Si pensi ad un cifrario polialfabetico con chiave lunga come il testo:
  - Testo criptato: wg ubsokwebalk a swqiu
  - Possibile testo: ci incontriamo a Pavia
  - Possibile testo: li incontriamo a Crema
- Qualunque testo della stessa lunghezza è lecito

DES (Data Encryption Standard)

- adottato nel 1977 come standar dal NIST (National Institute of Standards and Technology)
- utilizza una chiave simmetrica di 56 bit
- codifica blocchi di 64 bit

- Attacco brute-force
  - Un'operazione di cifratura DES per μs: 2^{55} μs = 1142 anni
  - 10^6 operazioni di cifratura DES per μs 10.01 h

- oggi è considerato obsoleto

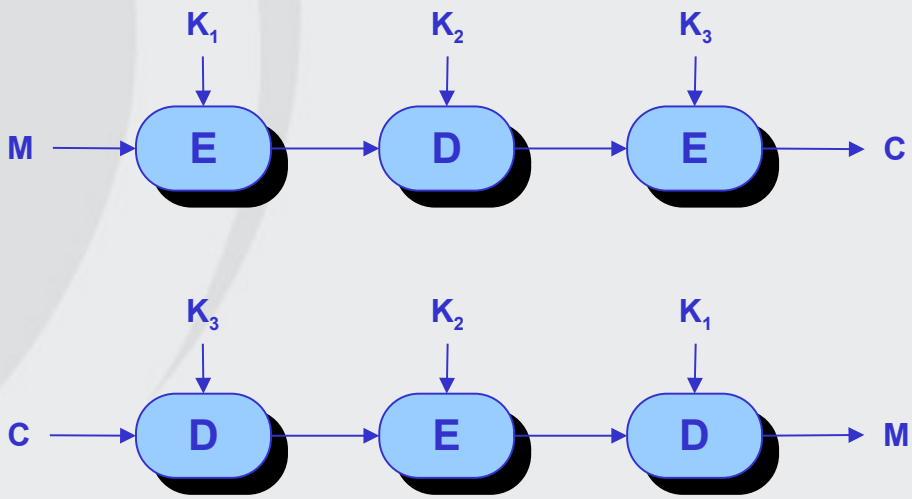
Tempo medio per una ricerca completa

Dimensione della chiave in bit	Numero di possibili chiavi	Tempo necessario	
		1 decifratura/μs	10^6 decifrature/μs
32	4.3×10^9	36 minuti	2.1 ms
56	7.2×10^{16}	1142 anni	10 ore
128	3.4×10^{38}	5.4×10^{24} anni	5.4×10^{18} anni
168	3.7×10^{50}	5.9×10^{36} anni	5.9×10^{30} anni

$$T_{medio} = ½\ 2^{nbit} / decifrature\_per\_secondo$$

Triplo DEA

- **Si usano tre chiavi e tre esecuzione dell'algoritmo DES (cifratura-decifratura-cifratura)**
  - $C = E_{K3}(D_{K2}(E_{K1}(M)))$
    - $E_K(X)$ = cifratura di X con la chiave K
    - $D_K(Y)$ = decifratura di Y con la chiave K
- **Lunghezza effettiva della chiave: 168 bit**



Cifratura a chiave pubblica

- L'algoritmo RSA prende il nome dai tre inventori: Ron Rivest, Adi Shamir, Len Adleman (MIT)
- La chiavi sono due coppie (D, N) e (E, N) dove N è il prodotto di due numeri primi p e q con
  - $ED \bmod (p-1)(q-1) = 1$
  - $C = m^E \bmod N$
  - $D = C^D \bmod N = m = m^{ED} \bmod N$
- La conoscenza dell'algoritmo, di una delle chiavi e di esempi di testo cifrato non è sufficiente per determinare l'altra chiave
  - Noti N e D è computazionalmente difficile ricavare E
  - Si sfrutta la funzione di Eulero $\Phi(N)$ (numero di interi positivi minori di N e primi rispetto a N, $\Phi(N)= (p-1)(q-1)$)

$m = (m^e \bmod n)^d \bmod n$

Dalla teoria dei numeri:

se p, q sono primi allora

$$x^y \bmod n = x^{y \bmod (p-1)(q-1)} \bmod n$$

$$\begin{aligned}
(m^e \bmod n)^d &= m^{ed} \bmod n \\
&= m^{(ed \bmod (p-1)(q-1))} \bmod n \\
&= m^{(1)} \bmod n \\
&= m
\end{aligned}$$

Si noti che i numeri e e d possono essere scambiati per cui le funzioni di codifica e decodifica possono essere applicate in qualsiasi ordine

M = C(D(M)) = D(C(M))

RSA: cifratura e decifratura

- Testo in chiaro: M < n
  - p=7,q=17,e=5, d=77, n=119, M=19
- Testo cifrato: $C=M^e \% n$
  - $19^5 \% 119 = 2476099 \% 119 = 66$
- Testo in chiaro: $M=C^d \% n$
  - $66^{77} \% 119 = 19$

Intermediario di fiducia

Problema per la crittografia a chiave simmetrica:

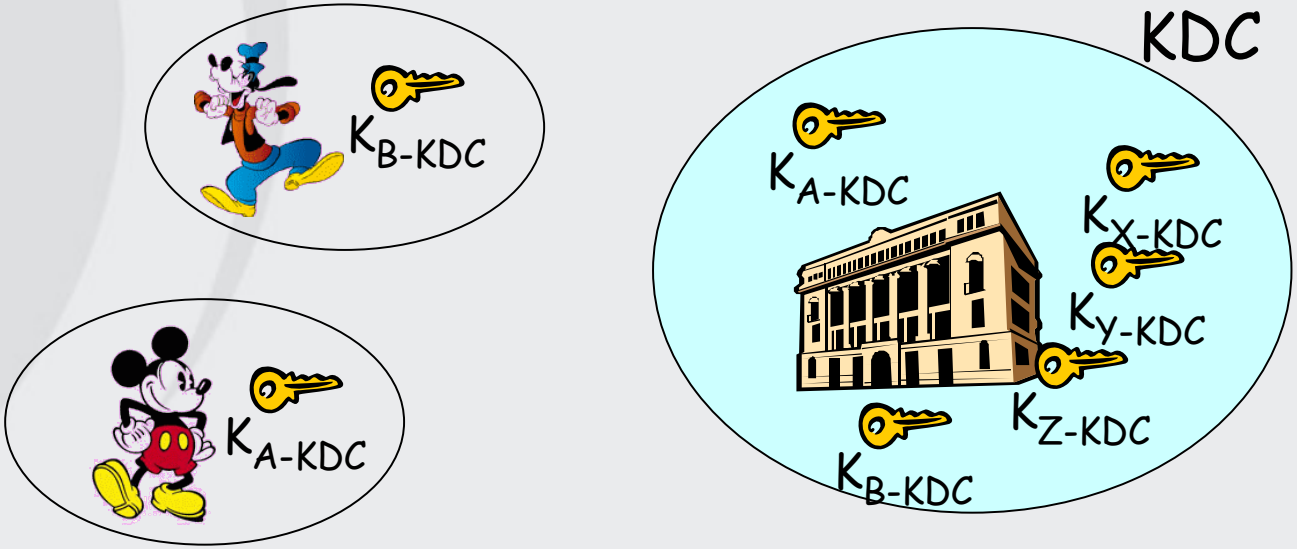
- Come possono le due parti concordare le chiavi prima di comunicare?
- Soluzione:
  - Un centro di distribuzione delle chiavi (KDC, key distribution center) di fiducia funge da intermediario tra le due entità

Problema per la crittografia a chiave pubblica:

- Quando Topolino riceve la chiave pubblica di Pippo (attraverso un CD, il sito web o via e-mail), come fa a sapere che è veramente la chiave pubblica di Pippo?
- Soluzione:
  - Autorità di certificazione (CA, certification authority)

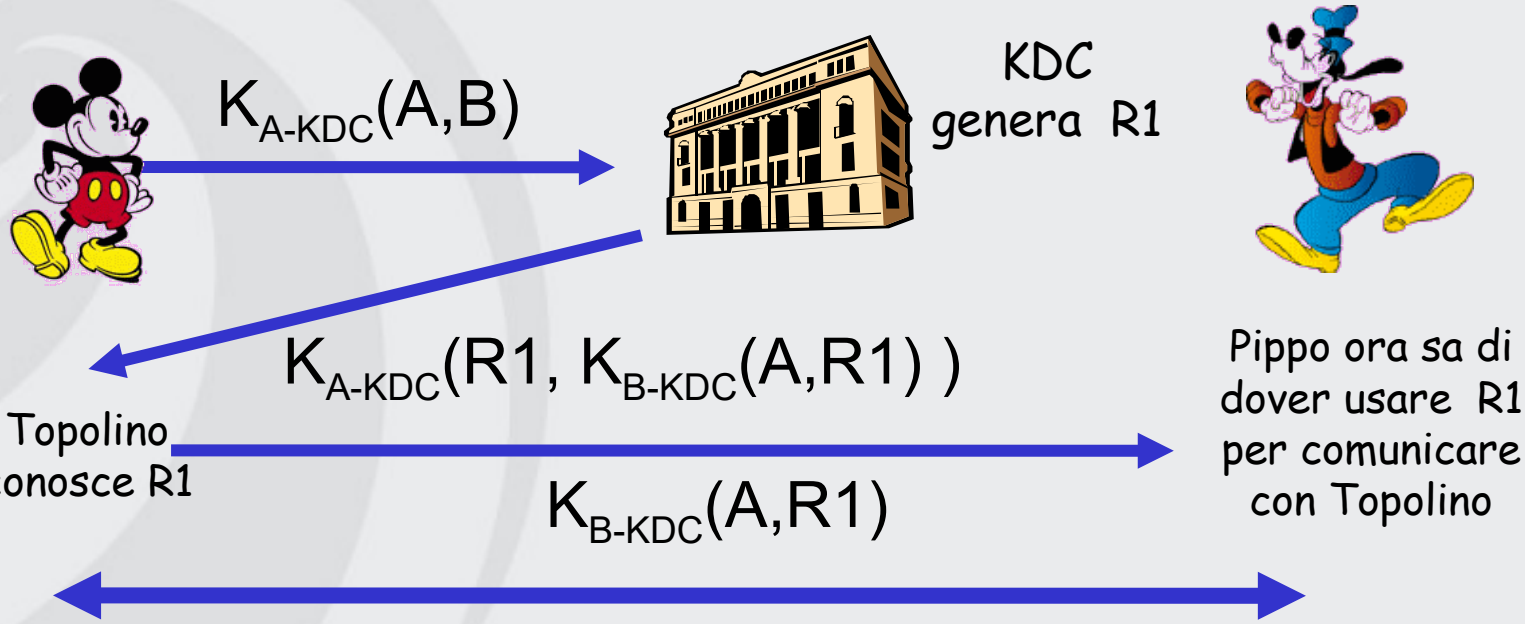
Centro di distribuzione delle chiavi (KDC)

- Topolino e Pippo vogliono comunicare protetti dalla crittografia a chiave simmetrica, ma non sono in possesso di una chiave segreta condivisa.
- KDC: è un server che condivide diverse chiavi segrete con ciascun utente registrato (molti utenti)
- Topolino e Pippo conoscono solo la propria chiave individuale, $K_{A\text{-}KDC}$ $K_{B\text{-}KDC}$, per comunicare con KDC



Centro di distribuzione delle chiavi (KDC)

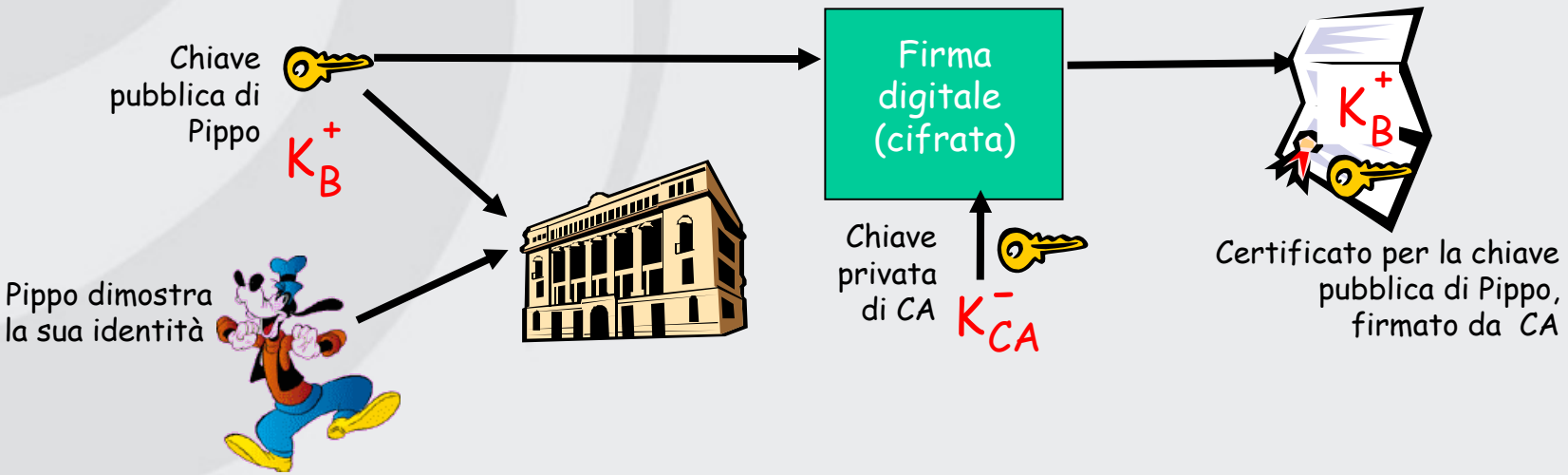
In che modo KDC consente a Topolino e Pippo di determinare la chiave segreta simmetrica condivisa per comunicare tra loro?



Topolino e Pippo comunicano usando R1 come chiave di sessione per la cifratura simmetrica condivisa

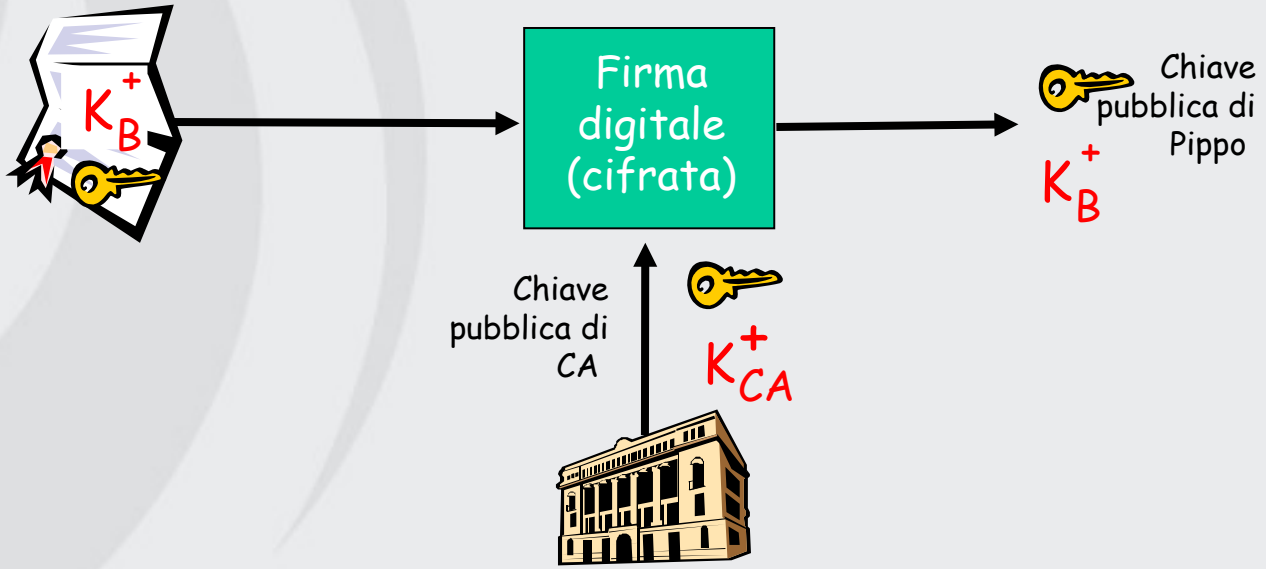
Autorità di certificazione

- Autorità di certificazione (CA): collega una chiave pubblica a una particolare entità, E
- E (persona fisica, router) registra la sua chiave pubblica con CA
  - E fornisce una "prova d'identità" a CA
  - CA crea un certificato che collega E alla sua chiave pubblica
  - Il certificato contiene la chiave pubblica di E con firma digitale di CA (CA dice "questa è la chiave pubblica di E")



Autorità di certificazione

- Quando Topolino vuole la chiave pubblica di Pippo:
  - prende il certificato di Pippo
  - applica la chiave pubblica di CA al certificato pubblico di Pippo e ottiene la chiave pubblica di Pippo

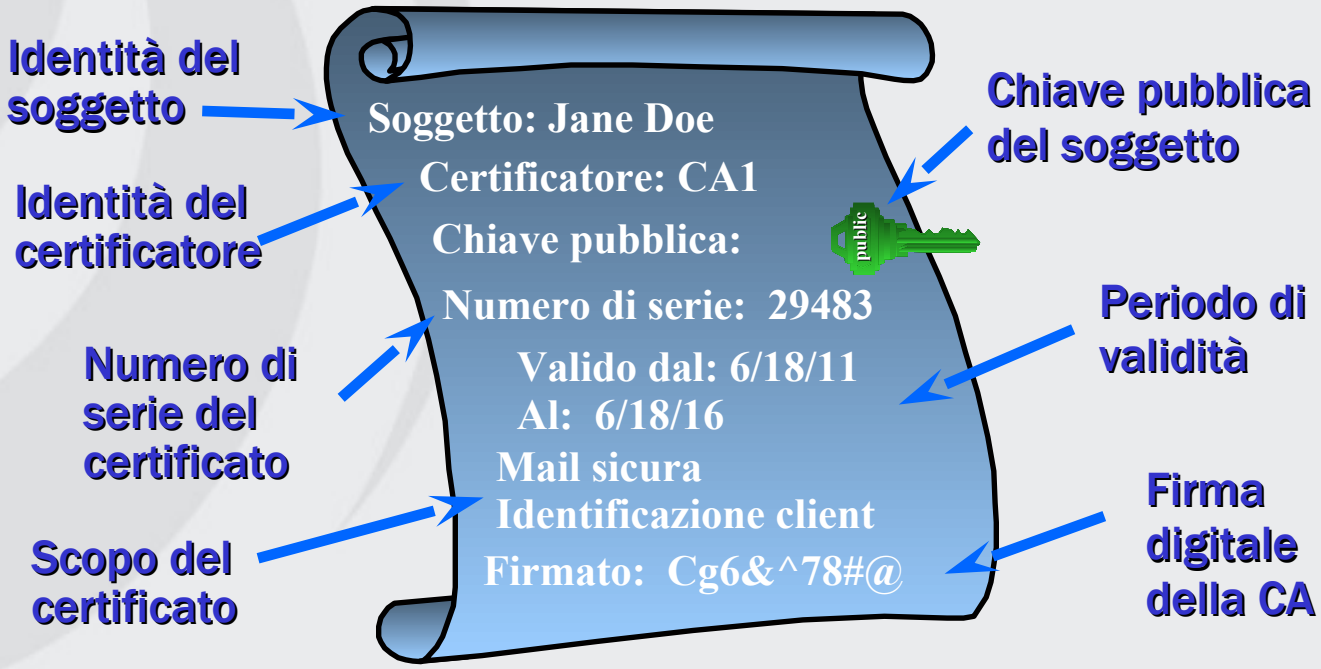


Certification Authority

- Una CA può erogare certificati a:
  - Se stessa (Root)
  - Un'altra CA (Subordinate)
  - Soggetti finali (utenti, computer)
- Una CA "fidata" deve fornire
  - Una prova della sua identità
  - Lista dei certificati revocati
  - Politiche di erogazione dei certificati

Certificati

- Legano l'identità di un soggetto ad una chiave pubblica
  - La chiave pubblica del soggetto è criptata con la chiave privata di CA (CA firma il certificato)



CVMLab
Computer Vision
& Multimedia Lab

Esercizi

Università
degli Studi
di Pavia

Cifrario Playfair

- Chiave: chiave
- Testo: dollaro
- Filler: Y

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

Dollaro → dolylaro

do → GM

ly → NW

la → NH

ro → TM

Cifrario Playfair

- Chiave: sistemioperativi
- Testo: Rossi Paolo
- Filler: Y

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

rossipaolo → ro sy si pa ol oy

ro → AP sy → EU si → IT

pa → RV ol → RH oy → AU