

#### Mensa Italia - The High "IQ" Society

in collaborazione con l'albergo "La Fenice"

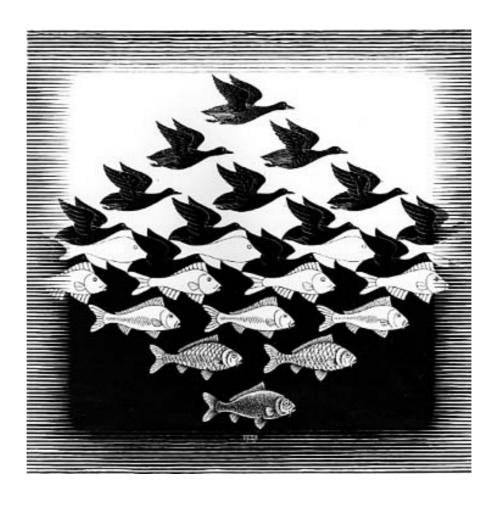


# Introduzione alla crittografia

Come affidare la privacy ad un'equazione matematica

di Enrico Zimuel

14 Luglio 2002 - Francavilla al Mare (CH)



#### **Sommario**

- Che cos'è la crittografia?
- Cenni storici
- Cifrari manuali: sostituzioni, trasposizioni
- Cifrari a sostituzione monoalfabetica, polialfabetica
- La crittoanalisi statistica
- Esiste il cifrario perfetto? Il cifrario di Vernam one-time pad
- La crittografia moderna: I cifrari simmetrici
- I cifrari Des, 3Des, Blowfish, Rijndael (AES)
- Il problema della trasmissione della chiave
- I cifrari asimmetrici o a chiave pubblica
- Il cifrario RSA
- I cifrari ibridi: Il metodo di Diffie-Hellman per lo scambio delle chiavi
- Il software PGP
- La firma digitale e le funzioni hash
- I keyserver e le certification authority

#### Note sul copyright:

Il presente documento può essere utilizzato liberamente a patto di citare la fonte e non stravolgerne il contenuto. CopyFree 2002 - Enrico Zimuel.

enrico@enricozimuel.net

# Che cos'è la crittografia?

- La crittografia (dal greco kryptos, nascosto, e graphein, scrivere) è la scienza che si occupa dello studio delle scritture "segrete".
- E' nata come branca della matematica e dell'informatica grazie all'utilizzo di tecniche di teoria dei numeri e di teoria dell'informazione.
- "Insieme delle tecniche che consentono di realizzare la cifratura di un testo e la decifrazione di un crittogramma" Dizionario Garzanti (1972)
- Alcune definizioni curiose:
   "Sistema segreto di scrittura in cifra o codice",
   "Gioco enigmistico consistente in una specie di
   rebus letterale particolarmente oscuro"
   Dizionario Zingarelli (1987)
- La crittografia fa parte della crittologia che racchiude in se un'altra scienza, la crittoanalisi (crittologia= crittografia + crittoanalisi)

#### Origini storiche

- La **crittografia** è una scienza antichissima utilizzata nell'antichità per nascondere messaggi tra regnanti, imperatori, nobili.
- La scitala lacedemonica è un antico esempio di un sistema per cifrare messaggi tramite l'utilizzo di un bastone cilindrico, cifrario a trasposizione (secondo gli scritti di Plutarco, in uso dai tempi di Licurgo, IX sec a.C.).



- Il periodo d'oro della crittologia è relativo alla seconda guerra mondiale quando Alan Turing, il padre dell'informatica teorica, insieme al gruppo di ricerca del Bletchley Park formalizzò la matematica necessaria per uno studio sistematico dei cifrari.
- Nasce nel 1943 in Inghilterra il primo elaboratore elettronico il **Colossus** utilizzato per decifrare le comunicazioni "segrete" dei nemici.

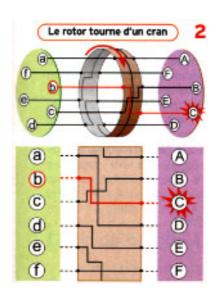


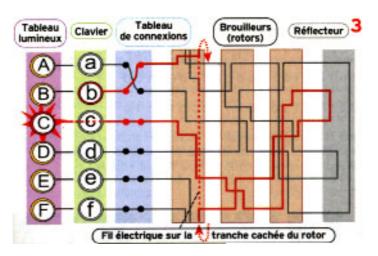
• Claude Shannon, l'ideatore della moderna teoria dell'informazione, che nel 1949 pubblicò un articolo rimasto nella storia "Communication theory of secrecy systems".

# La macchina cifrante Enigma

- **Enigma** è una delle macchine cifranti più famose della seconda guerra mondiale, ideata da Arthur Scherbius.
- La macchina Enigma consentiva di cifrare un testo scegliendo tra 17'576 x 6 x 100'391'791'500 = 10'000'000'000'000'000, 10 milioni di miliardi di combinazioni differenti.







## La crittografia moderna

- Le basi teoriche della moderna crittografia, quella attualmente utilizzata, sono ancora più giovani e risalgono a circa 30 anni fa a partire dal 1969 con le prime ricerche di **James Ellis** del quartier generale governativo delle comunicazioni britanniche (GCHQ).
- Sviluppata ed affinata nel 1976 in America grazie al contributo di Whitfield Diffie e Martin Hellman con la nascita del termine crittografia a chiave pubblica.
- Nasce nel 1977 il cifrario a chiave pubblica RSA da tre ricercatori del MIT (Massachusetts Institute of Technology), Ronald Rivest, Adi Shamir e Leonard Adleman.
- Con il cifrario RSA si inizia a parlare di strongencryption, crittografia forte.



#### Che cos'è un cifrario?

 Un cifrario è un sistema, di qualsiasi tipo, in grado di trasformare un testo in chiaro (messaggio) in un testo inintellegibile (testo cifrato o crittogramma).



- Per poter utilizzare un cifrario è necessario definire due operazioni: la cifratura del messaggio e la decifrazione del crittogramma.
- Definiamo con Msg "l'insieme di tutti i messaggi" e con Critto "l'insieme di tutti i crittogrammi".
- Cifratura: operazione con cui si trasforma un generico messaggio in chiaro m in un crittogramma c applicando una funzione
   C: Msg -> Critto.
- Decifrazione: operazione che permette di ricavare il messaggio in chiaro m a partire dal crittogramma c applicando una funzione D: Critto -> Msg.
- Matematicamente D(C(m))=m le funzioni C e D sono una inversa dell'altra e la funzione C deve essere iniettiva, ossia a messaggi diversi devono corrispondere crittogrammi diversi.

# Cifrari manuali (a sostituzione monoalfabetica)

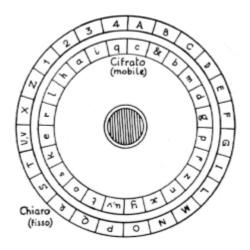
- Cifrari a sostituzione monoalfabetica: si utilizza un alfabeto sostitutivo con una permutazione della posizione delle lettere.
- **Cifrario di Cesare**: ogni lettera viene sostituita da quella che la segue di tre posizioni nell'ordinamento normale dell'alfabeto (in maniera circolare).

а	b	C	d	е	f	g	h	.—	I	m	n	0	р	q	r	S	t	u	V	Z
d	е	f	g	h	i		m	n	0	р	q	r	S	t	u	V	Z	а	b	С

- In termini matematici: C(x)= (x+3) mod 21, D(x)= (x-3) mod 21 (dove x rappresenta la posizione della lettera nell'alfabeto e l'operatore mod rappresenta il resto della divisione per 21).
- Esempio: "a"=1 per cui C(1)=(1+3) mod 21= 4 ossia la lettera "a" viene cifrata con la lettera "d", poiché "d"=4. Per decifrare utilizzo D(4)=(4-3) mod 21 = 1, ossia "a".
- Esempio: il testo in chiaro "prova di trasmissione" viene cifrato con il crittogramma "surbd gn zudvpnvvnrqh"

# Cifrari manuali (a sostituzione monoalfabetica)

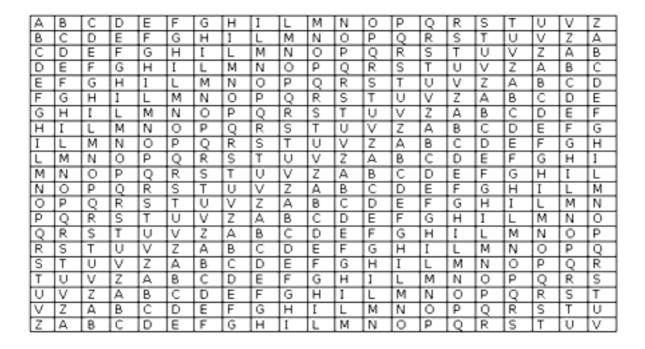
- Cifrario a rotazione o cifrario additivo: Si utilizzano tutte le 21 permutazioni circolari dell'alfabeto italiano.
- Leon Battista Alberti, nel suo Trattato della cifra, ha proposto un disco composto di due cerchi cifranti concentrici:



 Cifrario completo: si utilizzano tutte le permutazioni dell'alfabeto, per cui le possibili combinazioni risultano 21!-1 ≈ 5x10<sup>19</sup>.
 Vedremo che anche se il numero delle possibili combinazioni è molto grande la sicurezza di un cifrario completo a sostituzione monoalfabetica è molto bassa.

# Cifrari manuali (a sostituzione polialfabetica)

 Si utilizzano più alfabeti di sostituzione in cascata, il più famoso cifrario a sostituzione polialfabetica è il cifrario di Vigenère (Blaise de Vigenère, XVI sec.).



- Si utilizza una parola-chiave per identificare la riga corrispondente per la sostituzione con il relativo alfabeto individuato.
- Ad esempio con la parola-chiave MENSA, voglio cifrare la parola "INTELLIGENZA", ottengo il crittogramma: "URHZLVOTZNLE".
- Ogni lettera nel testo in chiaro non viene sempre sostituita con la stessa lettera nel crittogramma.

# Cifrari manuali (a trasposizione)

- Si utilizzano degli schemi di trasposizione della posizione dei caratteri in un testo in chiaro.
- Si fissa un numero intero P, detto periodo della traposizione, e si sceglie una permutazione degli interi da 1 a P, ad esempio per P=7:

1	2	3	4	5	6	7
4	6	3	5	7	1	2

 Il messaggio viene suddiviso in blocchi di lunghezza P e le lettere di ciascun blocco vengono "rimescolate" in base alla permutazione.

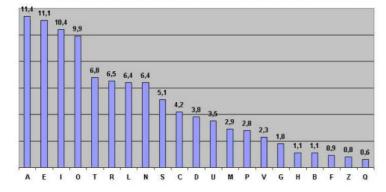
Ad esempio il crittogramma corrispondente a: "FUGGI PR/IMA DEL T/RAMONTO" è "GPGIRFU/DLAETIM/OTMNORA"

 Con questo cifrario le lettere vengono semplicemente "rimescolate", il contenuto informativo (entropia) associata ad uno schema del genere rimane invariata.

#### Crittoanalisi statistica del cifrario di Cesare

- Il cifrario di Cesare, come la maggior parte dei cifrari storici basati tu trasposizioni e traslazioni, può essere facilmente violato utilizzando tecniche statistiche (crittoanalisi statistica).
- Si analizzano le frequenze relative dei caratteri nel testo cifrato e le si confrontano con quelle di una lingua conosciuta, ad esempio l'italiano.
- Le frequenze relative al testo cifrato "surbd gn zudvpnvvnrqh" risultano s (1/19), u (2/19), r (2/19), b (1/19), d (2/19), g (2/19), n (3/19), z (1/19), v (3/19), p (1/19), h (1/19).
- Si confrontano tali frequenze con quelle della lingua italiana: a (0,114), e (0,111), i (0,104), o (0,099), t (0,068), r (0,065),...





 Con queste informazioni ottengo una prima approssimazione del testo in chiaro "sroba gi zravpivvioqh", procedo per tentativi ripetendo il procedimento.

# Il cifrario di Vernam one-time pad

- Esiste un cifrario teoricamente sicuro? Si, è il cifrario di Vernam, detto anche one-time pad, ideato nel 1926 da **G.S.Vernam**.
- Per ogni comunicazione si utilizza una sorgente perfettamente casuale di numeri, ogni volta differente (one-time pad), e l'operatore di somma XOR (⊕).

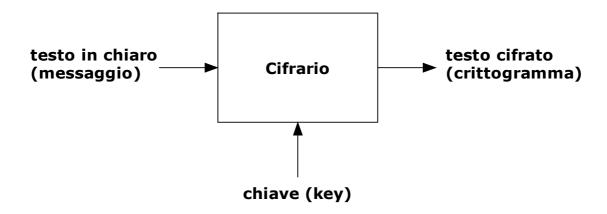
Sorgente casuale di numeri (one-time pad)



- La sicurezza teorica del sistema è garantita dalla perfetta casualità della sorgente (one-time pad) e dal fatto che per ogni operazione di cifratura si utilizzano dati differenti.
- Nella realtà un cifrario del genere non esiste poiché non esiste un algoritmo in grado di emulare una sorgente informativa perfettamente casuale.
- Svantaggio pratico, in una ipotetica implementazione: lunghezza della one-time pad identica a quella del messaggio.

#### La crittografia simmetrica

- Introduciamo un parametro chiamato k (key= chiave) all'interno delle funzioni di cifratura C(m,k) e decifrazione D(c,k).
- Si parla di crittografia simmetrica perchè si utilizza la stessa chiave k per le operazioni di cifratura e decifrazione.
- La robustezza del cifrario dipende, a differenza di prima, solo dalla segretezza della chiave k.

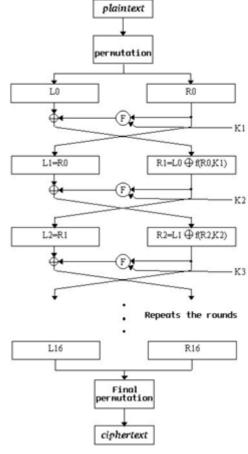


- Principio di Kerckhoffs (1883): "La sicurezza di un sistema crittografico è basata esclusivamente sulla conoscenza della chiave, in pratica si presuppone noto a priori l'algoritmo di cifratura e decifrazione."
- Purtroppo alcuni sistemi crittografici proprietari moderni non rispettano questo essenziale principio di sicurezza.

#### I cifrari simmetrici moderni:

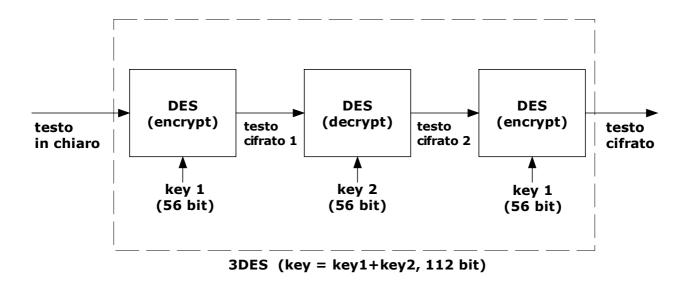
# **DES (Data Encryption Standard)**

- Sviluppato dall'IBM nel 1970 diventato standard nel 1976.
- Utilizza chiavi di 56 bit, divide il testo in chiaro in blocchi di 64 bit, effettua delle permutazioni iniziali e finali ed un ciclo di 16 iterazioni di permutazioni e xor (Feistel network, tecniche di confusione e diffusione).
- Il 17 Luglio 1998, l'**EFF**(Electronic Frontier
  Foundation) costruisce un
  sistema dedicato in grado di violare il DES in
  meno di 56 ore, tramite un attacco di tipo
  "brute-force".
- Morale della favola: non utilizzate sistemi di cifratura basati sul DES!



# I cifrari simmetrici moderni: 3DES (Triple Des)

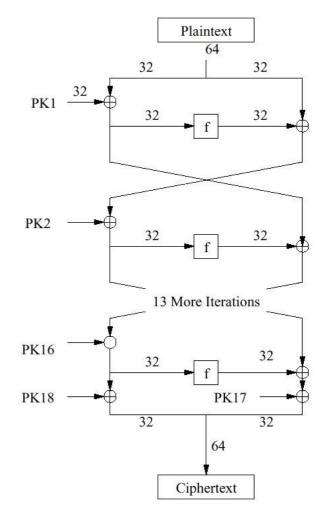
- Evoluzione del DES, è basato su un utilizzo del cifrario DES ripetuto, chiavi di 112 bit.
- Si utilizza la tecnica della codifica-decodificacodifica (EDE, Encrypt-Decrypt-Encrypt) utilizzando il cifrario DES.



#### I cifrari simmetrici moderni:

#### **Blowfish**

- Ideato nel 1993 da Bruce Schneier.
- E' stato sviluppato come algoritmo di encryption: veloce, compatto, semplice da implementare e sicuro con chiavi di dimensioni variabili fino a 448 bit.
- E' un cifrario a blocchi di 64 bit, basato sulle reti di Feistel.
- Non si conoscono attacchi efficaci.



 Può essere utilizzato liberamente, l'algoritmo non è brevettato, è utilizzato in molti sistemi open source (come ad esempio in OpenBSD).

#### I cifrari simmetrici moderni:

# Rijndael

- Sviluppato Joan
   Daemen e Vincent
   Rijmen.
- Questo algoritmo ha vinto la selezione per l'Advanced Encryption Standard (AES) il 2 Ottobre 2000. Ufficialmente il Rijndael è diventato lo standard per la cifratura del XXI secolo.
- Il cifrario utilizza
   chiavi di lunghezza
   variabile 128, 192,
   256 bit (gli autori
   hanno dimostrato
   come è possibile
   variare le dimensioni
   delle chiavi con
   multipli di 32 bit). Lo

Byte Sub Shift Row Mix Column

schema del Rijndael è stato influenzato dall'algoritmo SQUARE.

# Il problema della trasmissione della chiave

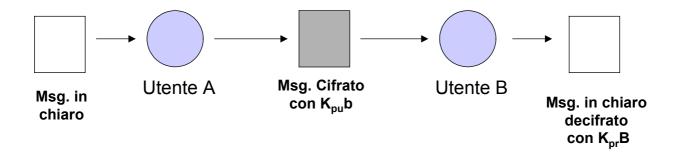
 Volendo utilizzare un cifrario simmetrico per proteggere le informazioni tra due interlocutori come posso scambiare la chiave segreta? Devo utilizzare una "canale sicuro" di comunicazione.



- Ma tale "canale sicuro" esiste nella realtà?
- Altro inconveniente: Per una comunicazione sicura tra n utenti si dovranno scambiare in tutto (n-1)\*n/2 chiavi, ad esempio con 100 utenti occorreranno 4950 chiavi, il tutto per ogni comunicazione!

#### La crittografia a chiave pubblica

- Utilizza una coppia di chiavi per le operazioni di cifratura (encryption) e decifrazione (decryption).
- Una chiave detta **pubblica** (**public key**) viene utilizzata per le operazioni di encryption.
- L'altra chiave, detta privata (private key),
   viene utilizzata per le operazioni di decryption.
- A differenza dei cifrari simmetrici non è più presente il problema della trasmissione delle chiavi.
- Sono intrinsecamente sicuri poiché utilizzano tecniche di tipo matematico basate sulla teoria dei numeri, sulla teoria delle curve ellittiche, etc.
  - Esempio di encryption (trasmissione sicura):

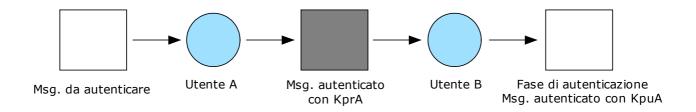


 $K_{pu}B$  = chiave pubblica dell'utente B

 $K_{pr}B$  = chiave privata dell'utente B

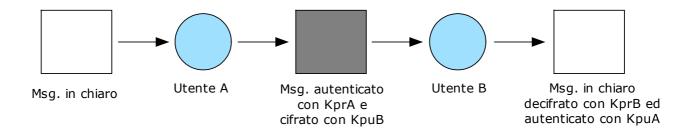
# La crittografia a chiave pubblica

# • Esempio di autenticazione:



KprA = chiave privata dell'utente A
KpuA = chiave pubblica dell'utente A

# • Esempio di encryption ed autenticazione:



KprA = chiave privata dell'utente A
KpuA = chiave pubblica dell'utente A
KprB = chiave privata dell'utente B
KpuB = chiave pubblica dell'utente B

#### Il cifrario RSA

- E' basato su tecniche di teoria dei numeri: prodotto di due numeri primi di dimensioni elevate (ad esempio con 300 cifre decimali), congruenza in modulo, funzione di Eulero.
- Definiamo alcuni concetti di teoria dei numeri per poter analizzare il funzionamento del cifrario RSA:
  - Un numero p>1 si dice primo se è divisibile solo per ±1 e ±p.
  - Dati tre interi a,b≥0 e n>0, si dice che a è congruo a b modulo n, e si scrive:

# $a \equiv b \pmod{n}$

se esiste un intero k per cui a = b + kn ( o equivalentemente se a mod n = b mod n, dove l'operatore mod indica il resto della divisione intera tra a e n, b e n).

 Per un intero n>1 si definisce la funzione di Eulero Φ(n) come il numero di interi minori di n e relativamente primi con esso. Se n è un numero primo si ha che Φ(n)=n-1.

#### Il cifrario RSA

- Le chiavi pubbliche e private vengono determinate con il seguente algoritmo:
  - Si scelgono due numeri primi p e q molto grandi (ad esempio con l'algoritmo di Solovay e Strassen, 1975);
  - Calcolo n= p\*q, e la funzione di Eulero
     Φ(n)= (p-1)\*(q-1);
  - Scelgo un intero e minore di Φ(n) e primo con esso;
  - Calcolo l'intero d, inverso di e modulo Φ(n) (ossia tale che e\*d = k\*Φ(n) + 1, con k numero intero);
  - La chiave pubblica è costituita dalla coppia di valori <e,n>, la chiave privata dai valori <d,n>.
- Le operazioni di encryption e decryption sono:

$$C = M^e \pmod{n}$$

$$\mathbf{M} = \mathbf{C}^{d} \pmod{\mathbf{n}} = (\mathbf{M}^{e} \pmod{\mathbf{n}})^{d} \pmod{\mathbf{n}}$$

dove M= blocco di testo in chiaro, C= crittogramma.

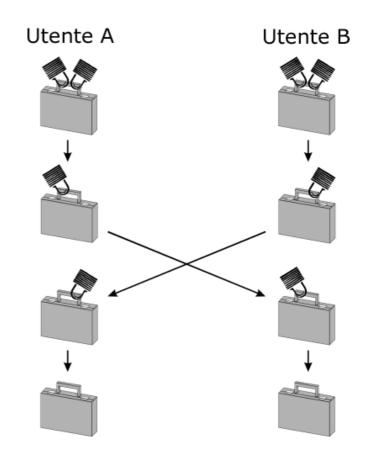
#### Il cifrario RSA

- La sicurezza del sistema è basata sul fatto che è difficile fattorizzare un prodotto di due numeri primi di dimensioni elevate (allo stato attuale), ossia l'algoritmo che risolve questo problema ha una complessità computazionale troppo elevata.
- La lunghezza delle chiavi è variabile: 512, 1024, 2048, 4096 bit ed oltre.
- Svantaggio: l'algoritmo RSA non è veloce, infatti viene utilizzato soprattutto nei sistemi crittografici ibridi che utilizzano contemporaneamente sia algoritmi simmetrici che algoritmi a chiave pubblica (come ad esempio nei software PGP e GNUPG).
- Il 6 Settembre 2000 l'algoritmo RSA è diventato di dominio pubblico, prima era di proprietà dell'omonima azienda RSA Security Inc (http://www.rsa.com).



# L'algoritmo Diffie-Hellman per lo scambio delle chiavi

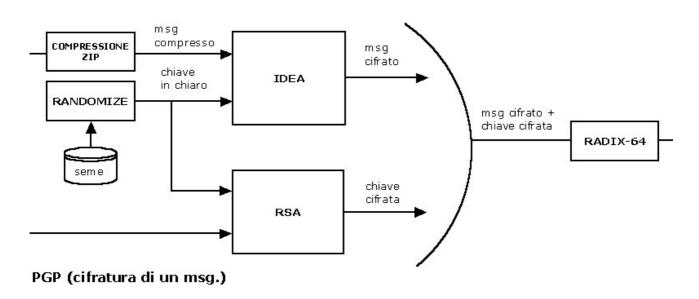
- Creato nel 1976
   dai ricercatori
   W.Diffie e
   M.Hellman è il
   primo algoritmo a
   chiave pubblica
   della storia.
- E' stato creato per eliminare il problema dello scambio delle chiavi di cifratura su di un canale insicuro di comunicazione.



- L'efficacia dell'algoritmo Diffie-Hellman dipende dalla difficoltà di calcolare logaritmi discreti.
- Il sistema lavora su delle strutture albebriche particolari, i campi di Galois con prodotti e potenze di numeri interi.
- Può essere utilizzato insieme ad un cifrario di tipo simmetrico per lo scambio della chiave simmetrica k di cifratura (**sistema ibrido**).

#### Il software PGP

- PGP (Pretty Good Privacy) è un software di pubblico dominio creato da Phil Zimmermann nel 1991.
- E' un software per la privacy personale: protezione delle email, dei files, firma digitale.
- Utilizza gli algoritmi di crittografia a chiave pubblica RSA, Diffie-Hellman, DSA e gli algoritmi simmetrici IDEA, CAST, 3-DES.
- E' basato su di un sistema di crittografia "ibrido" nel senso che utilizza crittografia simmetrica per le operazioni di encryption sui dati generando delle chiavi di sessione pseudocasuali cifrate con un algoritmo a chiave pubblica.



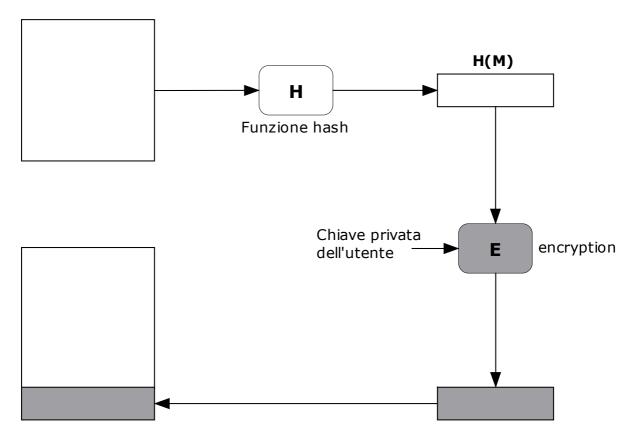
 Attualmente il progetto PGP è morto, l'ultima versione rilasciata dalla NAI è la 7.0.4, al suo posto è possibile utilizzare il GNUPG.

#### La firma digitale e le funzioni hash sicure

- Nasce come applicazione dei sistemi a chiave pubblica.
- Viene utilizzata per autenticare la paternità di un documento informatico e la sua integrità.
- Si utilizza un cifrario a chiave pubblica e si "cifra" un documento (file) con la propria chiave segreta. Chiunque può verificare la paternità del documento utilizzando la chiave pubblica dell'utente firmatario.
- *Problema*: per l'autenticazione di un documento di grandi dimensioni con un algoritmo a chiave pubblica occorre molto tempo.
- Soluzione: posso autenticare solo un "riassunto" del documento tramite l'utilizzo di una funzione hash sicura.
- Le funzioni hash sicure vengono utilizzate per generare un sorta di "riassunto" di un documento informatico (file).
- Una funzione hash accetta in ingresso un messaggio di lunghezza variabile M e produce in uscita un digest di messaggio H(M) di lunghezza fissa.

# Esempio di firma digitale di un documento

#### Documento da firmare M



#### **Documento firmato:**

Il ricevente può verificare la firma utilizzando la chiave pubblica dell'utente firmatario e riapplicando la funzione hash

#### Le funzioni hash sicure

- L'output di una funzione hash, il digest (impronta digitale, targa, riassunto), è strettamente legato al messaggio M, ogni messaggio M genera un H(M) univoco.
- Anche considerando due messaggi M
   ed M' differenti solo per un carattere le loro
   funzioni hash H(M) e H(M') saranno diverse.
- Requisiti di una funzione hash sicura:
  - H può essere applicata a un blocco di dati di qualsiasi dimensione;
  - H produce in uscita un risultato di lunghezza fissa (ad esempio 160 bit);
  - Per qualunque codice h il calcolo di x tale che H(x)=h deve avere una complessità computazionale improponibile;
  - Per qualunque blocco di dati x deve essere il calcolo di y≠x tale che H(x)=H(y) deve avere una complessità computazionale improponibile.
  - Ai fini pratici H(x) deve essere relativamente semplice da calcolare.

# Esempio di funzione hash

- Tutte le funzioni hash operano sulla base del seguente principio: i dati in ingresso sono considerati come una sequenza di blocchi di n bit, essi vengono elaborati un blocco alla volta iterativamente per produrre una funzione hash di n bit.
- Una delle più semplici funzioni hash è quella che esegue lo XOR (⊕) bit a bit di ciascun blocco, ossia:

$$C_i = b_{i1} \oplus b_{i2} \oplus ... \oplus b_{im}$$

- Dove C<sub>i</sub> rappresenta l'i-esimo bit del codice hash, m il numero di blocchi di n bit, b<sub>ij</sub> l'iesimo bit all'interno del j-esimo blocco e l'operatore ⊕ l'operazione di XOR.

# I Public Key Server e le Certification Authority

- Dove trovo le chiavi pubbliche dei miei destinatari?
- Creazione di "archivi di chiavi pubbliche", i Public Key Server.
- Ma chi mi garantisce la corrispondenza delle chiavi pubbliche con i legittimi proprietari?
- Nascita delle Certification Authority (CA): entità di certificazione in una rete informatica dell'identità elettronica delle chiavi pubbliche.
- A questo punto chi garantisce la validità delle certification authority?
- Atto di fede!
- In Italia esistono attualmente 14 entità di certificazione legalmente riconosciute dall'AIPA (Autorità per l'Informatica nella Pubblica Amministrazione) secondo l'articolo 27 comma 3 del DPR 28 dicembre 2000 n.445 specificato nel DPCM 8 febbraio 1999, esse sono: S.I.A. S.p.A., SSB S.p.A., BNL Multiservizi S.p.A., Infocamere SC.p.A., Finital S.p.A., Saritel S.p.A., Postecom S.p.A., Seceti S.p.A., Centro Tecnico per la RUPA, In.Te.S.A. S.p.A., ENEL.IT S.p.A., Trust Italia S.p.A., Cedacrinord S.p.A., Actalis S.p.A.

#### Bibliografia italiana essenziale

- "Crittografia" di Andrea Sgarro, Franco Muzzio Editore;
- "Crittografia Principi, Algoritmi, Applicazioni" di P. Ferragina e F. Luccio, Bollati Boringhieri Editore;
- "Crittologia" di L. Berardi, A.Beutelspacher, FrancoAngeli Editore;
- "Codici & Segreti" di Simon Singh, Rizzoli Editore;
- "La guerra dei codici" di Stephen Budiansky, Garzanti Editore;
- "Segreti, Spie e Codici Cifrati" di C.Giustozzi,
   A.Monti, E.Zimuel, Apogeo Editore;
- "Sicurezza delle reti Applicazioni e standard" di William Stallings, Addison-Wesley Editore;
- "Sicurezza dei sistemi informatici" di M.Fugini, F.Maio, P.Plebani, Apogeo Editore

# Su Internet per saperne di più

http://www.enricozimuel.net