La crittografia nell'era Internet

di Enrico Zimuel

Italian Web Awards 2002

Francavilla al Mare (CH) 21 Giugno 2002

Questa presentazione è stata creata con OpenOffice 1.0
www.openoffice.org

- Sommario:
  - Introduzione alla crittografia
  - La crittografia simmetrica o a chiave segreta
  - La crittografia asimmetrica o a chiave pubblica
  - La firma digitale e le funzioni hash sicure
  - I servizi di autenticazione: Kerberos, X.509
  - La sicurezza web: i protocolli SSL, TLS
  - Sicurezza IP e ESP
  - Sicurezza della posta elettronica: PGP/GNUPG

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)

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.

Origini storiche

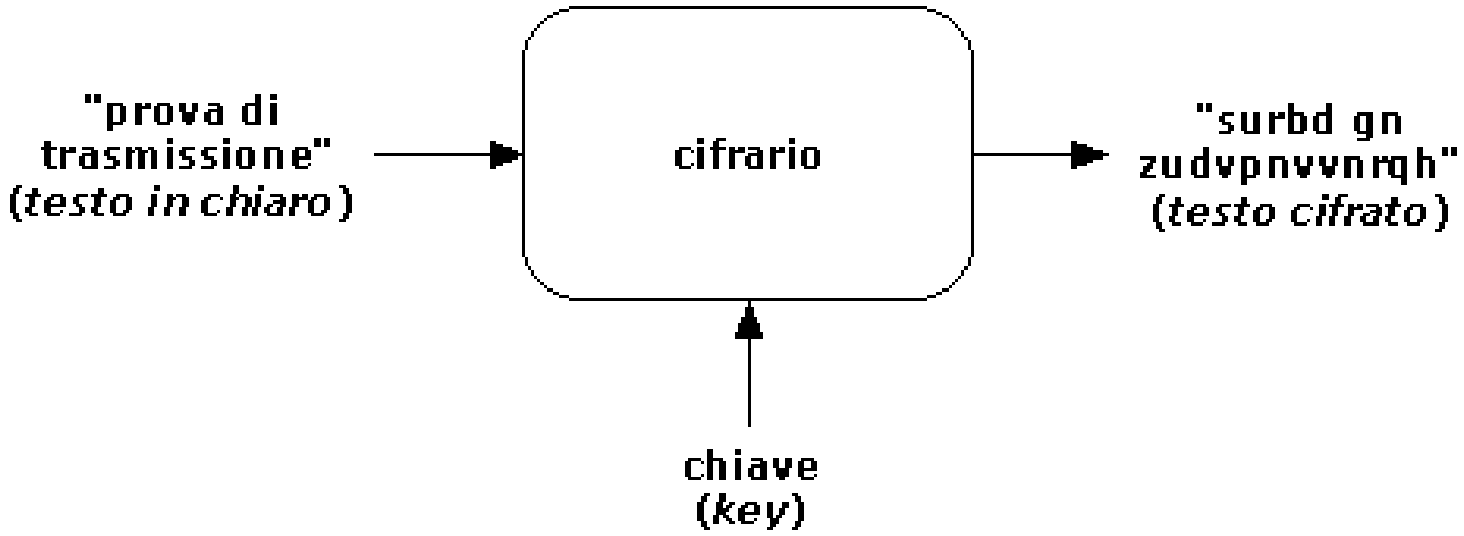
- **Enigma** è una delle macchine cifranti più famose della seconda guerra mondiale.
- **Claude Shannon**, l'ideatore della moderna teoria dell'informazione, che nel 1949 pubblicò un articolo rimasto nella storia "Communication theory of secrecy systems".
- Nasce nel 1943 in Inghilterra il primo elaboratore elettronico il **Colossus** utilizzato per decifrare le comunicazioni "segrete" dei nemici.

La moderna crittografia

- 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 *strong-encryption*, crittografia forte.

La crittografia simmetrica

- Si utilizza una sola chiave segreta (*key*) per cifrare e decifrare una comunicazione. Il cifrario "mischia" (encryption) i caratteri del testo in chiaro utilizzando come unica informazione la chiave segreta.



Esempio di cifrario simmetrico: il cifrario di Cesare

- Consideriamo l'alfabeto italiano, costruiamo un cifrario che sostituisce ad ogni lettera di questo alfabeto la lettera che si trova 3 posizioni in avanti.

- Il testo in chiaro “prova di trasmissione” viene cifrato nel testo “surbd gn zudvpnvvnrqh”.

Crittoanalisi 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 "s**ro**b**a** g**i** z**ra**vp**i**vv**io**qh", procedo per tentativi ripetendo il procedimento.

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?
- 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!

Alcuni cifrari simmetrici moderni

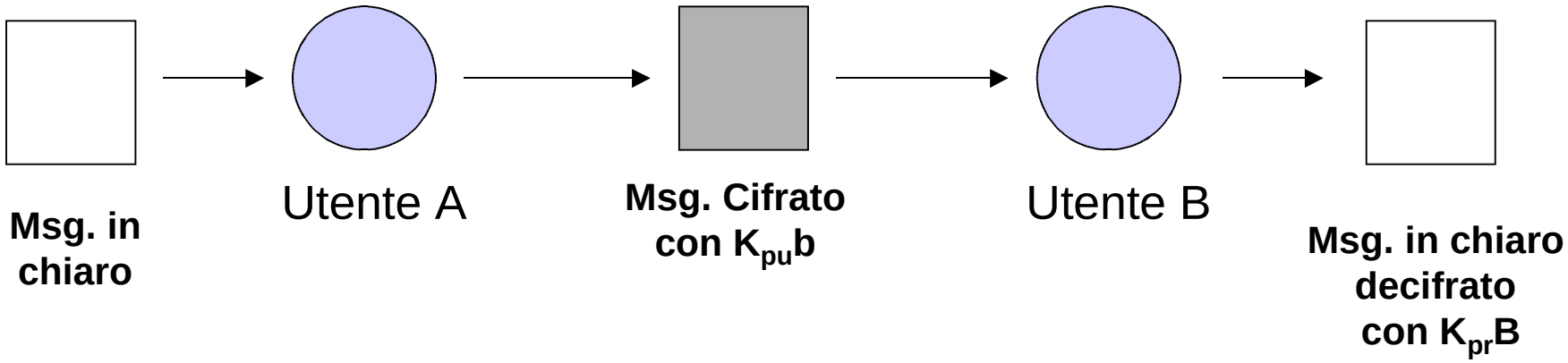
- Nella crittografia moderna vengono utilizzati molto nei *sistemi ibridi* per la loro velocità di elaborazione.
- I più conosciuti ed utilizzati cifrari simmetrici sono:
  - Feistel (1973)
  - DES (Data Encryption Standard, 1977), 3DES
  - IDEA (1991)
  - BlowFish (1993)
  - RC5 (1994)
  - CAST-128 (1997)
  - Rijndael (AES, Advanced Encryption Standard)

La crittografia a chiave pubblica

- Utilizzano una coppia di chiavi per le operazioni di encryption e 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.

La crittografia a chiave pubblica

- Esempio:



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

$K_{pr}B$ = chiave privata 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).
- Il cifrario RSA è basato sulle seguenti relazioni:

$$C=M^e \pmod n \quad , \quad M=C^d \pmod n = M^{ed} \pmod n$$

dove M= blocco di testo in chiaro, C= blocco di testo cifrato, la chiave pubblica è costituita dalla coppia (e,n) e la chiave privata dalla coppia (d,n).

- La sicurezza del sistema è basata sul fatto che è difficile fattorizzare un prodotto di due numeri primi di dimensioni elevate.

La nascita dei sistemi PKI

- 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**).
- A questo punto chi garantisce la validità delle certification authority?
- Atto di fede!

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.

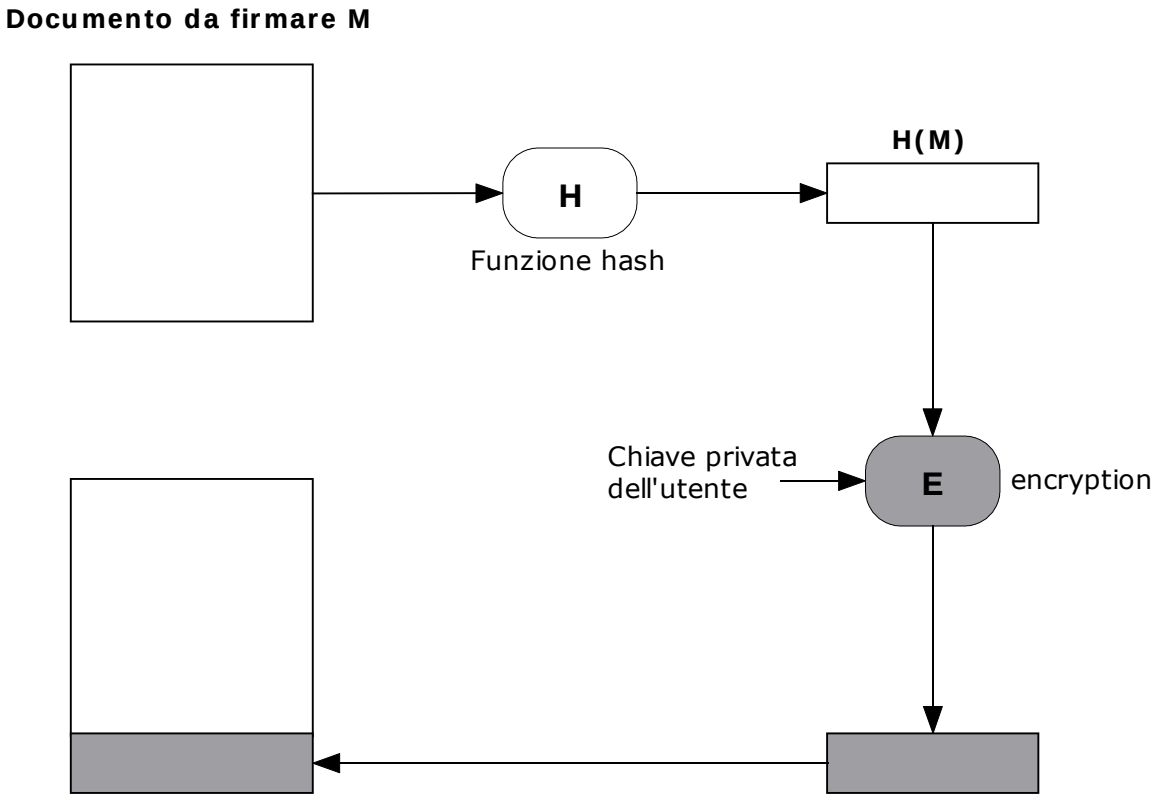
Le funzioni hash sicure

- 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.
- Questo 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(x):

- 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.
- per fini pratici H(x) deve essere relativamente semplice da calcolare.

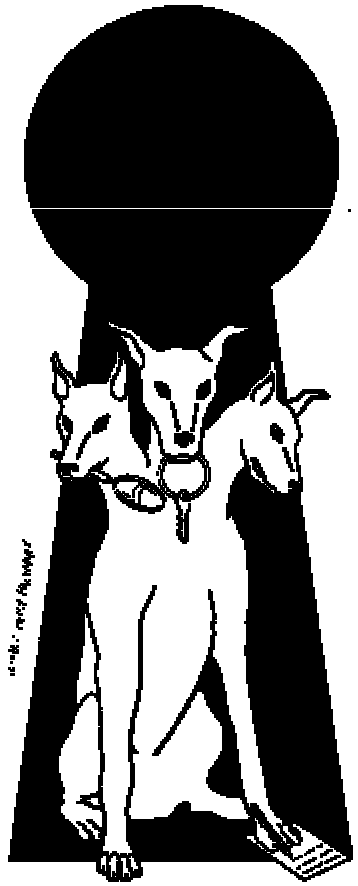
Esempio di firma digitale di un documento:



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

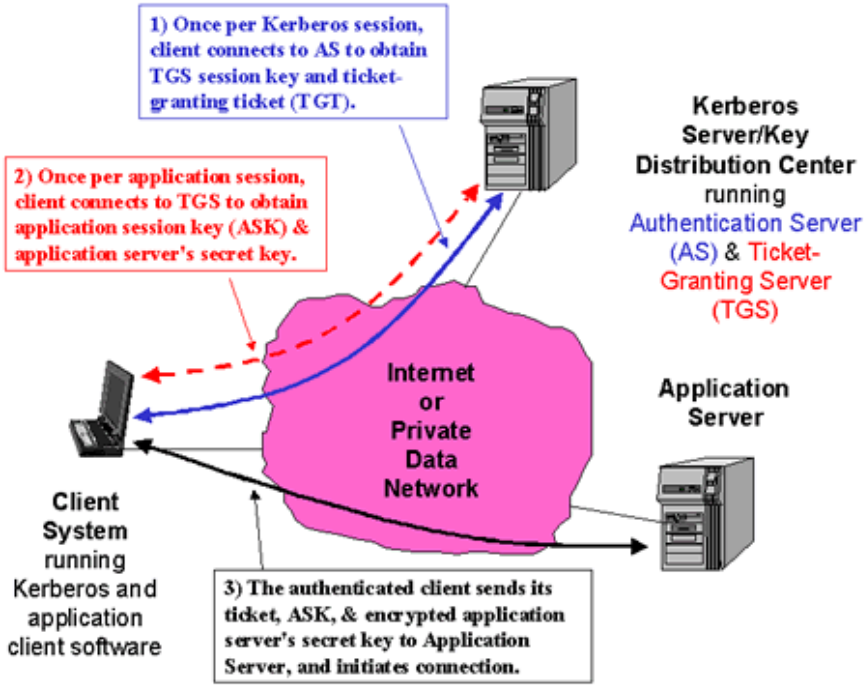
Il servizio di autenticazione Kerberos

- Kerberos è un servizio di autenticazione sviluppato nel 1991 al MIT nell'ambito del progetto Athena.
- Viene utilizzato in un ambiente distribuito in cui gli utenti, collegati a workstation, desiderano accedere a servizi forniti da server distribuiti sulla rete. I server devono essere in grado di consentire gli accessi solo a utenti autorizzati e di autenticare le richieste di servizi.
- Kerberos si basa su tecniche di crittografia simmetrica e non utilizza algoritmi a chiave pubblica.
- Le caratteristiche principali di Kerberos sono: sicurezza, affidabilità, trasparenza, scalabilità.



Il modello di fiducia di Kerberos

- L'autenticazione nel sistema Kerberos si basa su un nuovo modello di fiducia. A differenza del modello a due parti, in cui è prevista la presenza di due elementi aventi fiducia reciproca, nel sistema Kerberos le due parti vengono a trovarsi in una relazione di fiducia verso una terza parte avente funzione di garante dell'identità dell'uno verso l'altro.
- Essenzialmente, il funzionamento è basato sul modello di distribuzione delle chiavi di Needham e Schroeder modificato con l'aggiunta di un marcatore orario.

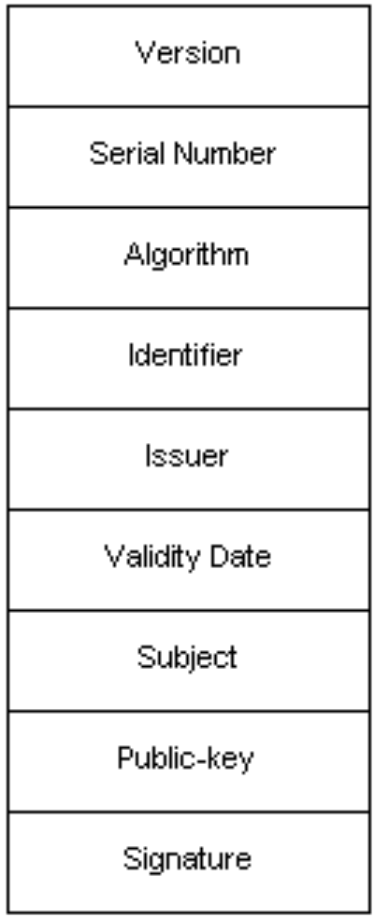


Il servizio di autenticazione X.509

- Fa parte della serie di raccomandazioni X.500 volte alla definizione di un servizio di directory (in generale una directory è un server o un insieme distribuito di server che mantengono una base di dati contenente informazioni sugli utenti).
- X.509 definisce un'architettura di riferimento per l'erogazione di servizi di autenticazione da parte delle directory X.500.
- X.509 utilizza crittografia a chiave pubblica, lo standard non impone l'utilizzo di un particolar algoritmo di cifratura ma raccomanda RSA.
- I servizi di autenticazione sono implementati con l'ausilio di certificati digitali contenenti le chiavi pubbliche degli utenti firmate con la chiave privata di un'autorità di certificazione (CA, Certification Authority).
- Il formato di certificato X.509 è utilizzato in molte applicazioni: S/MIME, sicurezza IP, SSL/TLS, SET.

I certificati X.509

- Il fulcro dello schema X.509 è costituito dai certificati a chiave pubblica associati a ciascun utente.
- Non è compito del directory server creare le chiavi pubbliche o effettuare la certificazione; il suo compito è univocamente quello di fornire un luogo facilmente accessibile da cui gli utenti possono ottenere i certificati.
- Un certificato X.509 è costituito essenzialmente dai seguenti elementi: versione, numero di serie, identificatore dell'algoritmo di firma, nome della CA, periodo di validità, nome del soggetto, informazioni sulla chiave pubblica del soggetto, identificatore univoco di chi emette il certificato, identificatore univoco del soggetto, estensioni, firma.

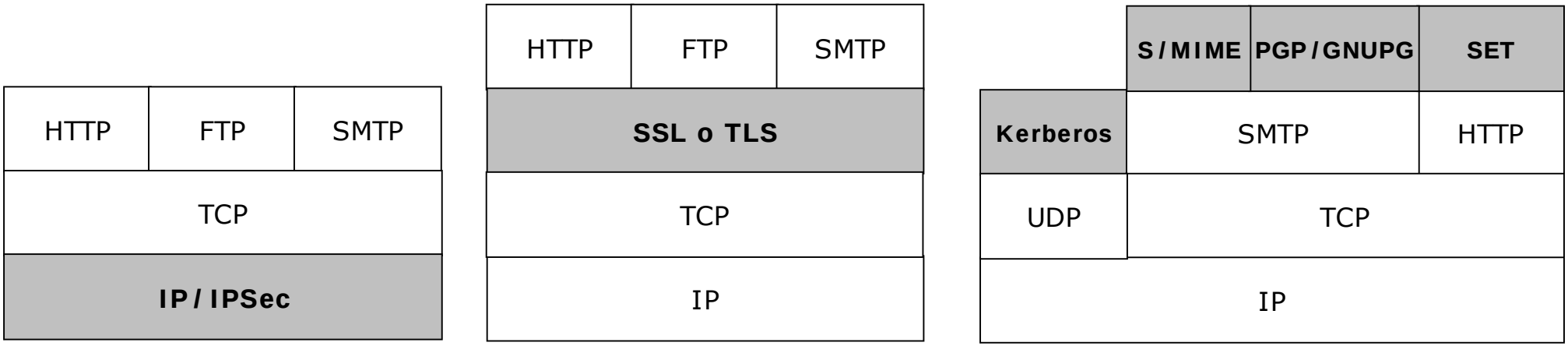


I protocolli SSL e TLS

- SSL (Secure Socket Layer) è stato creato dalla Netscape nel 1994, la versione 3.0 è del 1995.
- Successivamente il progetto è stato sottoposto al processo di standardizzazione Internet, all'interno di IETF () si costituì il gruppo di lavoro TLS (Transport Layer Security).
- La prima versione del TLS può essere considerata come un SSL 3.1, in pratica lo standard TLS è l'evoluzione del protocollo SSL.
- SSL è progettato per fare uso del protocollo TCP al fine di fornire un servizio di sicurezza end-to-end affidabile.
- Viene utilizzato per proteggere le transazioni via web di dati sensibili (https): acquisti legati all'e-commerce, numeri di carte di credito, informazioni aziendali, etc.

Sicurezza del traffico web

- Per fornire sicurezza del traffico web esistono diversi approcci simili dal punto di vista delle funzionalità ma differenti per quanto concerne l'ambito di applicabilità ed il loro posizionamento all'interno della pila del protocollo TCP/IP.

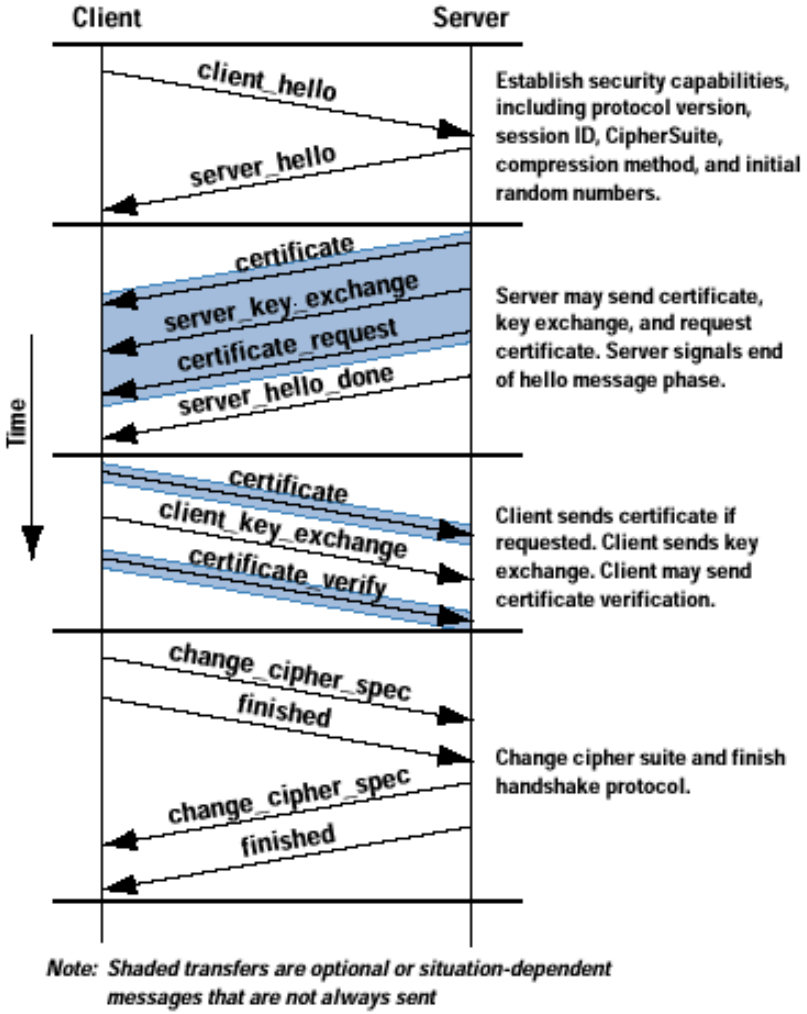


I protocolli SSL e TLS

- SSL è costituito da due livelli di protocolli: l'SSL **Record** e l'SSL **Handshake**; il primo viene utilizzato per il trasposto dei messaggi fornendo servizi di riservatezza ed integrità dei dati, il secondo fornisce il processo di autenticazione tra client e server per la creazione di un canale sicuro di comunicazione per l'utilizzo dell'SSL Record.
- La cifratura delle transazioni via web è variabile per ogni sessione di collegamento.
- Gli algoritmi di cifratura utilizzati dal protocollo SSL sono: IDEA, RC2-40, DES-40,DES, 3DES, Fortezza, RC4-40, RC4-128.
- Gli algoritmi di autenticazione per lo scambio delle chiavi temporali utilizzati dal protocollo SSL sono: RSA,DSS, MD5, SHA-1, Diffie-Hellman.

Il protocollo di autenticazione SSL: l'Handshake

- I passi principali dell'SSL Handshake possono essere così schematizzati:
  - Utente: client hello
  - Server: server hello
  - Server: invio del certificato
  - Server: server hello done
  - Utente: autenticazione del sistema
  - Utente: invio del pre-master secret e costruzione del master secret
  - Server: ricezione del pre-master secret e costruzione del master secret.
  - Utente: invio del certificato (opzionale)
  - Utente/Server: messaggio finished



Sicurezza IP

- Il protocollo di comunicazione attualmente utilizzato su Internet Ipv4 non prevede la cifratura dei messaggi dati (nell'IPv6 si), chiunque può “sniffare” ossia intercettare dalla rete un pacchetto dati e ricostruire il messaggio originario.
- La sicurezza IP (IPSec, *IP Security*) nasce da una precisa esigenza: proteggere le comunicazioni IP da attacchi di tipo IP spoofing (falsificazione degli indirizzi IP) e di tipo IP sniffing (intercettazione dei pacchetti dati).
- La protezione avviene tramite encryption dei pacchetti IP.
- IPSec fornisce la possibilità di rendere sicure le comunicazioni su LAN, su WAN private e pubbliche e su Internet.
- IPSec fornisce un insieme di servizi di sicurezza a livello IP: controllo dell'accesso, integrità in assenza di connessione, autenticazione della sorgente dati, rifiuto di pacchetti originati da un attacco di replay, riservatezza (cifratura), parziale riservatezza del flusso di traffico.

Sicurezza IP

- Alcuni esempi d'utilizzo di IPSec:
  - Connettività sicura di filiali su Internet;
  - Accessi remoti sicuri su Internet;
  - Possibilità di stabilire connettività extranet e intranet con i partner;
  - Miglioramento della sicurezza del commercio elettronico;
- La caratteristica più importante di IPSec è che può cifrare e/o autenticare tutto il traffico a livello IP, adattandosi quindi a una notevole gamma di applicazioni.

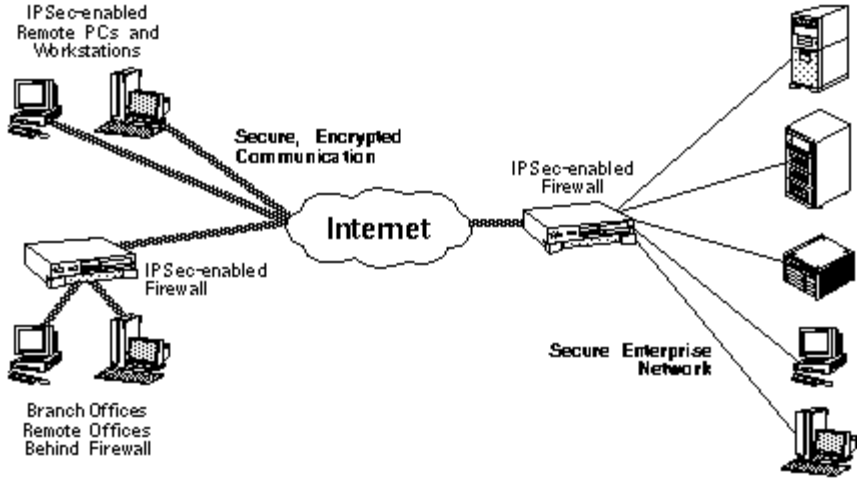
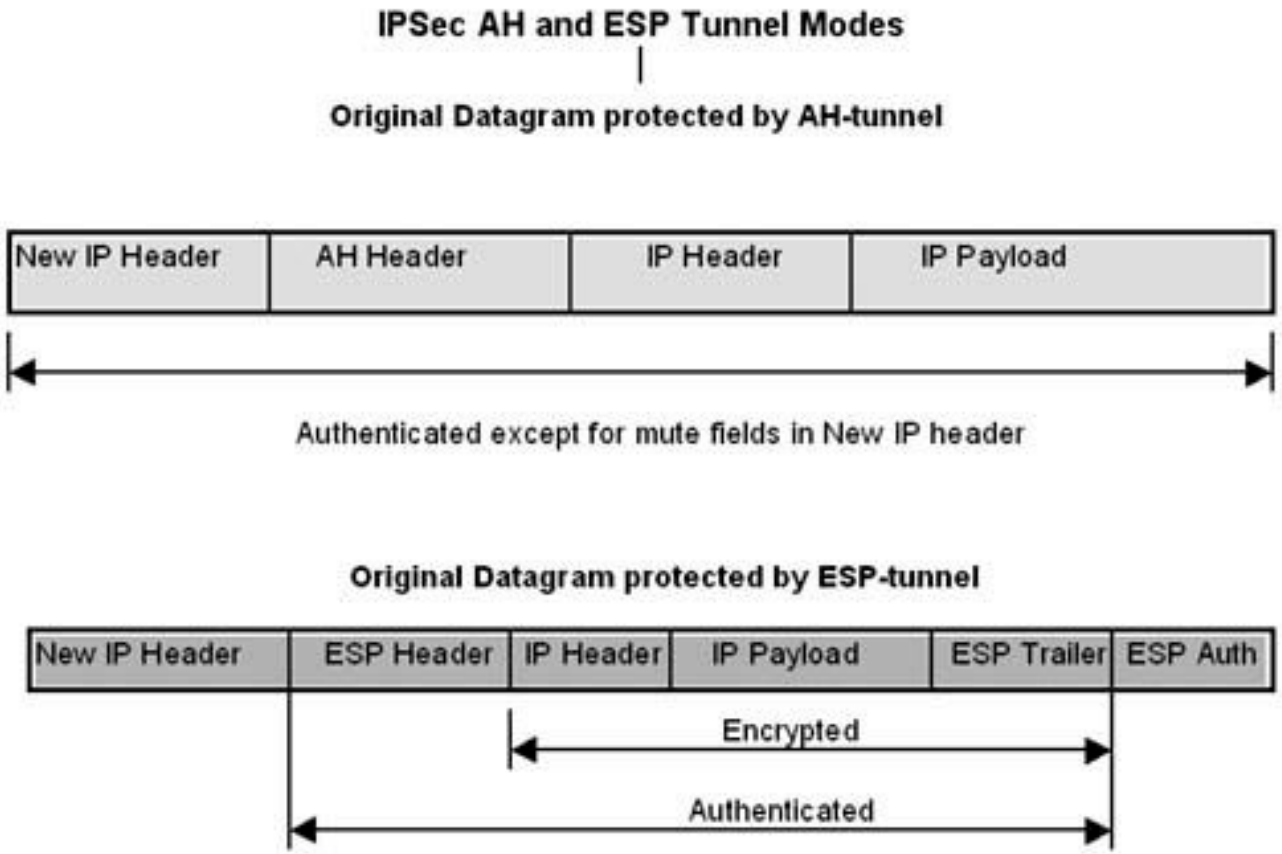


Diagramma pacchetti IPSec

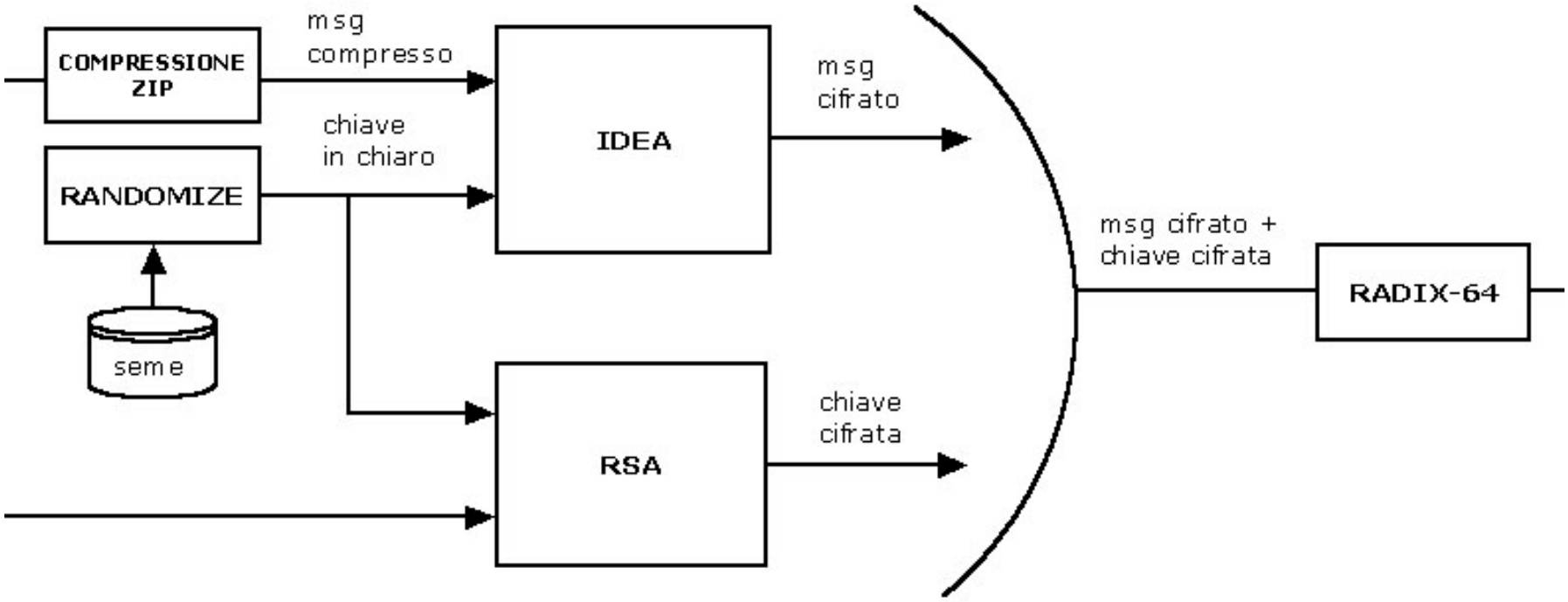
•



Sicurezza della posta elettronica: 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 pseudo-casuali cifrate con un algoritmo a chiave pubblica.
- Attualmente il progetto PGP è morto, l'ultima versione rilasciata dalla NAI è la 7.0.4.

Il funzionamento del PGP



PGP (cifratura di un msg.)

Sicurezza della posta elettronica: GNUPG

- Il progetto tedesco GnuPG (GNU Privacy Guard) nasce nel 1997 per opera di Werner Koch, sviluppatore indipendente interessato alla crittografia OpenSource.
- L'obiettivo del progetto è la realizzazione di un engine crittografico, alternativo al Pgp, totalmente open source basato su algoritmi crittografici standard e non proprietari.
- Disponibile in più versioni: Gnu/Linux, Ms Windows, FreeBSD, OpenBSD, AIX, Sun Os, BSDI, IRIX, etc.
- Disponbili vari front-end per sistemi GUI: Gnome, KDE, Ms Windows, etc.
- Supporto algoritmi crittografici ElGamal, DSA, RSA, AES, 3DES, Blowfish, Twofish, CAST5, MD5, SHA-1, RIPE-MD-160 e TIGER
- La versione attuale è la 1.0.7

Bibliografia italiana essenziale

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