Sicurezza dei dati

Premessa

Il tema della sicurezza dei dati (di lavoro o personali) è ormai uno dei punti critici da affrontare in ogni sistema.

In alcuni casi è importante proteggere sia l'integrità dei contenuti in uno scambio tra mittente e destinatario sia l'identità stessa del mittente e del destinatario. In altri casi è sufficiente proteggere soltanto la visibilità delle informazioni e dei contenuti.

Autenticazione e crittografia

Due sono le politiche che si implementano in un'ottica di sicurezza:

- l'**autenticazione** identifica con certezza un interlocutore (tipico è il caso di user name e password per l'accesso a siti o a risorse);
- la **crittografia** mira a trasformare le informazioni per renderle irriconoscibili a chiunque, tranne a chi ha una chiave per poterle interpretare.

Tecniche di intrusione

Gli **hacker** (o, con un termine più evocativo, pirati informatici) hanno ideato, e continuano a farlo, numerose tecniche di intrusione nei sistemi.

Si possono individuare due grosse famiglie:

- lo **sniffing** (fiutare) avviene quando i dati vengono ricopiati mentre sono in transito e senza modificarli, in modo che chi li invia e chi li riceve non si accorgano di nulla;
- lo **spoofing** (raggirare) consiste nel modificare gli indirizzi dei dati in transito, in modo che i dati passino dal pirata senza che mittente e destinatario originario lo rilevino.

Ci si rende conto che in entrambi i casi l'hacker deve essere molto esperto del sistema con cui ha a che fare e della struttura dei dati da intercettare. Non è da escludere, quindi, che molto spesso siano le stesse software house a "monitorare" i propri prodotti per dedurre informazioni sull'utilizzo degli stessi. I produttori di software, infatti, conoscono in partenza le caratteristiche del computer (dato che sono fornite dagli stessi utenti in fase di installazione) e le caratteristiche dei dati (ideate da loro stessi).

Casi tipici di sniffing o, ancora peggio, di spoofing sono familiari a tutti i lettori:

- il **phishing** (pescare) consiste nel "travestimento" dell'hacker in una forma che l'utente conosce o che gli è familiare, in modo da indurlo a fornire informazioni confidenziali o password;
- il **pharming** (falsare) consiste in falsi siti web, del tutto identici a quelli originali di banche o enti pubblici, dove l'utente – pensando di essere al sicuro – inserisce password o altre informazioni sensibili;
- l'**identity theft** (furto d'identità) consiste nella sostituzione di una persona o di una società da parte dell'hacker, il quale – finché nessuno lo scopre – può fare quello che vuole accreditandosi con un falso nome.

mistersnike@gmail.com

Fasi critiche nello scambio di contenuti

Gli elementi delicati in un qualsiasi scambio di dati sono essenzialmente tre:

- l'**autenticazione**, che avviene tipicamente con user name e password;
- l'**integrità dei contenuti**, che ha come obiettivo la consistenza delle informazioni;
- la **riservatezza dei collegamenti**, che avviene proteggendo i dati con sistemi di crittografia (sia per i dati che per le password).

Crittografia

La **crittologia** (dal greco "nascondere") è la scienza che studia la **crittografia**, ovvero quell'insieme di tecniche di trasformazione dei messaggi. In genere si tenta di trasformare ogni lettera dell'alfabeto in un'altra ed è un'attività che risale fino a Giulio Cesare, tanto per capirci. Il processo di ricostruzione è la **decifratura** o **decrittografia**.

Le moderne tecniche di crittografia si basano su concetti di tipo statistico: nella lingua italiana le lettere obbediscono a probabilità di ricorrenza ben precise. Decifrare un messaggio in cui ogni lettera è stata sostituita con un'altra, quindi, risulta molto semplice per gli attuali calcolatori.

Risultano più robusti algoritmi di crittografia che aumentano la ridondanza aggiungendo molti più simboli di quelli che servono e quindi, in pratica, aumentando le possibili combinazioni da provare con un attacco di forza bruta.

Un cifrario di questo tipo viene detto **monoalfabetico** se c'è una sola regola di trasformazione fra l'alfabeto originale e l'alfabeto criptato e, quindi, per ogni lettera criptata c'è un'unica corrispondenza con una sola lettera originale e viceversa. Un cifrario monoalfabetico si costruisce definendo una parola chiave che, sovrapposta all'alfabeto originale, consente di dedurre l'alfabeto criptato per trasposizione in sequenza successiva di tutte le altre sue lettere.

Tale tecnica è chiamata **Data Encryption Standard** (**DES**), inventata da IBM una trentina di anni fa e ancora valida oggigiorno. In particolare, in questo standard, ogni lettera è codificata su 64 bit, mentre la chiave è codificata su 56 bit. Purtroppo, in alcune condizioni sono le chiavi ad essere il punto debole del DES. In crittografia una **chiave** è un informazione usata come parametro in un algoritmo crittografico. Le chiavi sono utilizzate in molte applicazioni crittografiche e, secondo il principio di Kerckhoffs, sono l'unico dato che è davvero necessario tenere segreto. La dimensione della chiave, generalmente misurata in bit, dipende dal particolare algoritmo usato. Alcuni algoritmi possono utilizzare chiavi di lunghezze diverse e in questo caso più lunga è la chiave tanto più difficile sarà forzare il messaggio cifrato. È anche vero, però, che all'aumentare della lunghezza della chiave aumenta anche la complessità computazionale nell'eseguire l'algoritmo secondo una legge che può essere lineare, quadrata, cubica o anche peggiore. Nel caso del DES, macchine specializzate sono in grado di esaminare tutte le possibili chiavi e decifrare il messaggio in meno di 24 ore. Per rendere sicuro il DES si sono sviluppate delle evoluzioni come il **Triple DES** (**3DES**). Questo algoritmo espande la chiave impedendo per il momento un attacco a forza bruta, sebbene renda l'algoritmo in teoria vulnerabile a degli attacchi che però non sono attuabili nella pratica. Negli ultimi anni è stato sostituito – come si vedrà – dall'**Advanced Encryption Standard** (**AES**), un nuovo algoritmo che elimina molti dei problemi del DES.

Il DES è l'archetipo della **cifratura a blocchi**, un algoritmo che prende in ingresso una stringa di lunghezza fissa di testo in chiaro e la trasforma – con una serie di operazioni complesse – in un'altra stringa

di testo cifrato della stessa lunghezza. Nel caso del DES la dimensione del blocco è di 64 bit. Il DES usa inoltre una chiave per modificare la trasformazione in modo che l'operazione di decifratura possa essere effettuata solo conoscendo la chiave stessa. La chiave è lunga 64 bit ma solo 56 di questi sono effettivamente utilizzati dall'algoritmo. Otto bit sono utilizzati solo per il controllo di parità e poi scartati, per questo la lunghezza effettiva è riportata come di 56 bit. Il DES prevede quindi 16 cifrature successive (trasposizioni e sostituzioni di bit). In pratica, il testo chiaro viene suddiviso in blocchi da 64 bit (equivalenti a 8 caratteri); ogni blocco è sottoposto a una trasposizione data in base ad una chiave di 56 bit; si applica quindi per 16 volte una funzione cifrante e alla fine la trasposizione inversa di quella iniziale. Viene definito un sistema simmetrico perchè sia il mittente del messaggio sia il ricevente devono conoscere la stessa chiave segreta. In un ambiente di poli-utenza si era parlato di una diffusione della chiave fino a renderla di pubblico dominio; ma poi ci si è chiesti se ne valesse davvero la pena. Se usato singolarmente, il DES può essere un ottimo sistema per inserire files in un disco fisso, nella forma cifrata.

Più affidabili sono i cifrari polialfabetici, nei quali le lettere sono codificate usando più regole di trasformazione che – quindi – riportano ad altrettanti alfabeti criptati, tutti diversi. Ciò significa che per ogni lettera originale possono esserci più lettere criptate e, viceversa, ogni lettera criptata ha la sua regola di ricostruzione diversa dalle regole della lettere adiacenti.

Classificazione dei sistemi di crittografia

Solitamente una classificazione è basata discriminando la chiave che si utilizza. Si hanno così sistemi a chiave non riutilizzabile, simmetrici e asimmetrici.

I sistemi a chiave non riutilizzabile, o monouso, sono detti **One Time Password** (**OTP**): la chiave è memorizzata su una comune chiave USB e viene costruita sostituendo a ogni lettera dell'alfabeto un numero casuale che viene definito una sola volta per ogni trasformazione. La tecnica è senza dubbio indecifrabile, ma vale per un solo messaggio perché essa può essere usata una sola volta. C'è però il problema di difendere la chiave almeno quanto il contenuto dei messaggi.

La crittografia si definisce **simmetrica** quando usa una sola chiave sia per criptare sia per decriptare i messaggi. Essa è anche chiamata crittografia a chiave segreta, perché per la sua importanza la chiave deve essere comunicata dal mittente al destinatario in maniera "segreta". I sistemi simmetrici più famosi sono il già analizzato DES e l'**AES** (**Advanced Encryption Technology**). Esso utilizza una chiave di 128, 192 o 256 bit, che lo rende più affidabile del predecessore DES.

La crittografia **asimmetrica** è anche detta a chiave pubblica (**Public Key Infrastructure**, **PKI**) perché usa due chiavi: una chiave di codifica, detta pubblica perché è distribuita a chiunque dal destinatario, il quale – però – è l'unico a possedere l'altra chiave, detta asimmetrica o privata, che serve a riconoscere i contenuti dei messaggi: infatti il destinatario vuole essere l'unico a sapere leggere i messaggi dai mittenti a cui ha fornito la prima chiave. Questo concetto è alla base delle cosiddette firme digitali o certificati digitali, che i produttori leader di software incorporano nei loro prodotti per verificare se sono effettivamente utilizzati da chi li acquista. Naturalmente le due chiavi sono legate matematicamente da una regola di trasformazione che conosce solo il destinatario e, quindi, è assolutamente impossibile per un pirata decifrare un messaggio criptato senza la chiave privata, qualora conoscesse la chiave pubblica. Il metodo più famoso di crittografia asimmetrica è l'**RSA**, dalle iniziali dei tre ricercatori del MIT (Rivest, Shamir e Adleman) che lo idearono

vent'anni fa. Per semplificare il funzionamento, immaginiamo che A debba spedire un messaggio segreto a B. Occorrono i seguenti passaggi:

1. B sceglie due numeri primi molto grandi (per esempio da 300 cifre) e li moltiplica con il suo computer (impiegando meno di un secondo);
2. B invia il numero che ha ottenuto ad A. Chiunque può vedere questo numero;
3. A usa questo numero per crittografare il messaggio;
4. A manda il messaggio a B, che chiunque può vedere ma non leggere;
5. B riceve il messaggio e utilizzando i due fattori primi, che solo lui conosceva decifra il messaggio.

A e B hanno impiegato pochi secondi a cifrare e decifrare, ma chiunque avesse intercettato le loro comunicazioni impiegherebbe milioni di anni per scoprire i due fattori primi, con cui si può decifrare il messaggio. In realtà questo sistema non è così semplice e per trasmettere grandi quantità di dati occorre tanto tempo, quindi A e B si scambieranno con questo sistema una chiave segreta (che non occupa molto spazio), che poi useranno per comunicare tra loro usando un sistema a crittografia simmetrica, più semplice, sicuro e veloce.

Esistono, quindi, forme di crittografia ibride tra la simmetrica e l'asimmetrica: ad esempio, la struttura è simmetrica, mentre la chiave è criptata in modo asimmetrico.

Conclusioni

Qualunque sia il metodo scelto, permane il problema di dover adeguatamente proteggere le chiavi e ancora oggi la tecnica consigliata è quella di cambiare le chiavi periodicamente; per far ciò, esistono anche algoritmi molto efficaci a disposizione.

Inoltre, si capisce che l'aumento della protezione e della sicurezza è proporzionale all'aumento della complessità e dei costi.

Negli ultimi tempi si sente parlare di autenticazione forte (**strong authentication**): something you have, something you know. In pratica, chi vuole farsi identificare deve avere una chiave, intesa come memoria (USB, smart card, certificato digitale) e poi deve anche conoscere un'informazione chiave (codice identificativo, password, PIN). Le due chiavi possono usarsi solo insieme.

Riferimenti

"Elettronica Oggi", novembre 2007, numero 372, articolo "Autenticazione e cifratura" di Lucio Pellizzari.

http://it.wikipedia.org