La crittografia

In questa breve clip vogliamo introdurvi al mondo della crittografia. Trattandosi di un argomento molto vasto, si vuole creare interesse per il documento più approfondito associato a questo modulo.

La crittografia si riferisce ai metodi di protezione delle informazioni e delle comunicazioni in modo tale che solo le persone a cui sono destinate possano leggerle.

In informatica questa protezione si ottiene utilizzando applicazioni che implementano concetti matematici complessi. Questi concetti sono codificati mediante algoritmi per la generazione di chiavi crittografiche, la firma digitale, la protezione della privacy dei dati, la navigazione sul Web e le comunicazioni riservate come le e-mail e le transazioni con carta di credito.

La cifratura è il processo che rende illeggibili le informazioni e la decifratura recupera le informazioni nascoste. Il processo di cifratura/decifratura prevede solitamente un algoritmo e una *chiave. La* chiave è solo un'altra informazione, di solito un numero che specifica come l'algoritmo viene applicato a un *testo in chiaro* (o *cleartext*) per cifrarlo in quello che viene chiamato *cyphertext*. In un sistema crittografico sicuro anche se si conosce l'algoritmo utilizzato per cifrare il messaggio dovrebbe essere difficile o quasi impossibile recuperare il testo in chiaro dal testo cifrato senza la chiave.

Un po' di storia

La crittografia risale al 2000 a.C., quando nell'antico Egitto si usavano geroglifici "segreti" non standard. Nell'antica Grecia, un "dispositivo" chiamato Scytale era usato in particolare dagli spartani per comunicare durante le campagne militari; consisteva in un'asticella con una striscia di pergamena avvolta intorno ad essa su cui era scritto un messaggio. Il destinatario utilizzava un'asticella dello stesso diametro, su questa era avvolta la pergamena per leggere il messaggio.

Un metodo un po' più sofisticato che anticipa i metodi moderni è noto come *cifrario di Cesare*, utilizzato da Giulio Cesare per la sua corrispondenza riservata. In sostanza, ogni lettera viene sostituita da un'altra lettera successiva nell'alfabeto, più

precisamente 3 lettere dopo. Qui possiamo riconoscere il concetto di chiave: quante lettere bisogna andare più avanti nell'alfabeto per creare il testo cifrato. Per quanto semplice e presente persino nei giocattoli per bambini, Wikipedia riporta che è stato utilizzato ancora di recente.

Questo tipo di crittografia è semplice e può essere utilizzata per qualsiasi messaggio. A differenza di un sistema di frasi in codice in cui, ad esempio, "Compra il libro" significa "Manda avanti le legioni". In questo caso le due estremità della comunicazione devono condividere un libro comune di frasi in codice e non c'è modo di codificare altre frasi.

Il cifrario di Cesare è un *cifrario a sostituzione*, in quanto ogni lettera viene cambiata con un'altra; altre varianti sostituiscono blocchi di lettere o intere parole. Fino alla metà del secolo scorso la crittografia consisteva in cifrari di sostituzione più o meno complessi utilizzati per mantenere sicure le comunicazioni governative e militari. Nel frattempo i crittografi trovarono modi sempre più sofisticati per recuperare il testo originale da un testo cifrato, ma prima dell'avvento dei computer era difficile eseguire trasformazioni matematiche abbastanza rapidamente da rendere facili le operazioni di cifratura e decifratura. Durante la Seconda guerra mondiale, i tedeschi usarono la macchina elettromeccanica Enigma per criptare i messaggi e Alan Turing guidò un team britannico che sviluppò una macchina simile per decifrare il codice, fornendo anche le basi teriche per i primi computer moderni. La crittografia divenne decisamente più complessa con l'avvento dei computer.

La crittografia ha iniziato a essere interessante per gli usi civili quando i computer sono stati collegati per mezzo di reti non esclusive in quanto era facile spiare i dati che passavano sulla rete. I servizi finanziari sono stati i primi utenti delle reti e per essi era essenziale trovare un modo per mantenere segrete le informazioni.

All'inizio degli anni '70 la IBM propose l'algoritmo Data Encryption Standard (DES) per la protezione di dati elettronici sensibili e non classificati del governo. Nel 1977 fu pubblicato come standard ufficiale FIPS (Federal Information Processing Standard) e largamente adottato dall'industria. L'algoritmo era pubblicamente disponibile.

Oggi il fatto che gli algoritmi crittografici siano pubblicamente disponibili è visto come necessario, invece di essere considerato un male inevitabile. In questo modo, molti crittografi esperti possono studiare gli algoritmi e trovare le falle prima che vengano sfruttate.

La crittografia moderna

La crittografia moderna ha quattro obiettivi:

Riservatezza. Le informazioni non devono essere comprese da nessuno altro se non da chi è il legittimo destinatario.

Integrità. Le informazioni non possono essere alterate senza che l'alterazione venga rilevata.

Non ripudio. Il creatore dell'informazione non può negare di essere l'origine dell'informazione.

Autenticazione. Il mittente e il destinatario possono confermare l'identità l'uno dell'altro.

Nella crittografia moderna vige un principio importante: il principio di Kerckhoffs. Esso dice che "*un sistema crittografico dovrebbe essere sicuro anche se tutto ciò che riguarda il sistema, tranne la chiave, è di dominio pubblico*". In altre parole, si vuole progettare un algoritmo che non abbia bisogno di essere segreto per nascondere con successo le informazioni.

Tipi di algoritmi crittografici

Gli algoritmi crittografici in uso sono numerosi, ma rientrano essenzialmente in tre categorie: *crittografia simmetrica, crittografia asimmetrica* e *funzioni hash.*

Crittografia simmetrica

La crittografia simmetrica è la famiglia più tipica e antica di algoritmi crittografici, quella a cui si pensa tipicamente quando si pensa alla crittografia dei messaggi. Il loro nome deriva dalla simmetria delle operazioni di crittografia e decrittografia. Questi algoritmi utilizzano un'unica chiave che deve essere nota sia al mittente sia al destinatario; essa viene utilizzata per cifrare un messaggio e poi per decifrarlo.

Il cifrario di Cesare è un esempio di crittografia simmetrica. Ciò che Cesare e uno dei suoi centurioni avrebbero dovuto condividere è il numero di lettere da considerare in avanti o indietro nell'alfabeto per trasformare il testo in chiaro in cifrato o viceversa. Il numero è lo stesso sia per la cifratura sia per la decifratura: è questo che la rende simmetrica.

La crittografia simmetrica è ampiamente utilizzata per mantenere la riservatezza dei dati. Può essere utile per mantenere privato un disco rigido locale, poiché è lo stesso utente che generalmente cifra e decifra i dati protetti; in questo caso la chiave segreta non deve essere condivisa con nessun altro.

Crittografia asimmetrica

Gli algoritmi asimmetrici non utilizzano una sola chiave, ma due: una chiave viene utilizzata per cifrare (chiave privata) e l'altra per decifrare (chiave pubblica). Pertanto nello scambio di messaggi il mittente e il destinatario non utilizzano la stessa chiave. Queste chiavi sono generate e gestite a coppie: ciò che una chiave (qualsiasi della coppia) cifra, può essere decifrato solo dall'altra chiave della coppia.

Il meccanismo crittografico su cui si basa la crittografia asimmetrica è quello delle *funzioni unidirezionali*: operazioni matematiche molto difficili da invertire. Ad esempio, la moltiplicazione di due numeri primi molto grandi (la chiave pubblica e quella privata, ad esempio) è un calcolo semplice e veloce da eseguire, ma sarebbe molto difficile (in un tempo accettabile) scoprire i due numeri primi originali dal solo risultato.

Il motivo dell'utilizzo di algoritmi asimmetrici deriva dai problemi che presenta la crittografia simmetrica:

1. Le comunicazioni riservate tra 2 entità richiedono la condivisione di un'unica chiave, se le entità sono n allora sono necessarie $n \cdot (n-1)/2$ chiavi diverse per consentire la comunicazione riservata tra ogni coppia di parti.

2. Sia il mittente che il destinatario conoscono la loro chiave condivisa, quindi il non ripudio non è possibile (il ripudio è la negazione di essere l'autore di un certo messaggio).

Il primo problema non riguarda il numero di chiavi che ogni soggetto deve gestire, ma lo scambio delle stesse. Ogni soggetto deve concordare una chiave con ciascuno degli altri n-1 soggetti, un'operazione complessa che rende la crittografia simmetrica poco pratica per molti usi.

Il secondo problema, quello del non ripudio, risiede nel fatto che spesso si vuole essere sicuri che un certo messaggio sia stato effettivamente inviato da un certo soggetto. È chiaro che se un soggetto riceve un messaggio da un altro soggetto e i due soggetti sono gli unici a conoscere la chiave usata per cifrarlo, ogni soggetto è sicuro che l'altro soggetto è l'autore del messaggio. Viceversa, poiché solo loro conoscono la chiave, ogni soggetto sa che solo l'altro soggetto può leggere il messaggio. Il problema si presenta quando c'è la necessità di dimostrarlo a una terza parte per un qualche tipo di controversia, allora questo meccanismo non funziona: ogni soggetto è in grado di creare un messaggio, cifrarlo e poi dire che l'autore è l'altro soggetto. Il problema del non ripudio non si risolve utilizzando la crittografia simmetrica.

I calcoli necessari per la crittografia asimmetrica sono molto più complessi e richiedono più risorse di quelli utilizzati dagli algoritmi simmetrici. Per ovviare a questo problema, i messaggi inviati in rete (che di solito sono lunghi) sono tipicamente cifrati con un algoritmo simmetrico e la chiave simmetrica utilizzata (che è breve, alcune centinaia di byte) è cifrata con un algoritmo asimmetrico. In questo modo la riservatezza è assicurata.

Per garantire l'identificazione del mittente e stabilire il non ripudio è necessario un meccanismo che certifichi l'associazione di una chiave pubblica a un soggetto. Un'infrastruttura a chiave pubblica (PKI) offre la possibilità di essere certi che una determinata chiave pubblica sia associata a una specifica persona o istituzione.

Funzioni di hash

Gli algoritmi crittografici simmetrici e asimmetrici trasformano entrambi un testo in chiaro in un testo cifrato e poi di nuovo in un testo in chiaro. Una funzione hash è un algoritmo di crittografia unidirezionale: questa funzione applicata a un testo in chiaro produce un testo cifrato di dimensioni fisse (ad esempio 256 byte) chiamato *hash*, dal quale non è possibile recuperare il testo in chiaro. La loro utilità sta nel fatto che è estremamente difficile trovare un testo in chiaro con un determinato hash. Gli algoritmi di hashing sono utilizzati per garantire l'integrità dei dati, ad esempio inviando un messaggio insieme al suo hash. Quando si riceve il messaggio lo stesso algoritmo di hashing applicato al testo del messaggio deve dare lo stesso valore di hash, altrimenti è chiaro che il messaggio è stato modificato durante il trasporto. Per motivi di sicurezza, l'hash può essere crittografato con la chiave privata del mittente: questa è la base del protocollo di firma digitale.