Sicurezza ed informatica

Leonardo Miliani | 5 ottobre 2006

Da un po' di tempo mi sto appassionando alla sicurezza intesa come protezione dei dati in campo informatico: in una parola, alla **crittografia**. Questa scienza, diversamente da come si potrebbe pensare, non è nata in epoca moderna ma esiste da migliaia di anni: il suo nome deriva dalle parole greche *kryptos*, che significa nascosto, e *graphein*, che significa scrivere. E' una scienza molto interessante perché unisce matematica ed informatica in un *coctail* davvero gustoso. Oggi torna molto alla ribalta a causa della grande diffusione di internet, dove è utilizzata per rendere sicure le transazioni economiche che avvengono in rete: spesso, però, se ne abusa o se ne parla a sproposito, dando alla crittografia, considerata una specie di bacchetta magica, "poteri" che non possiede e colpe di cose non commesse. Cerchiamo quindi di fare chiarezza e vediamo di capire cos'è la crittografia, come ci aiuta nel mondo moderno e quali sono le leggende metropolitane a cui dovremmo evitare di credere. Ah, premessa: non sono ne' un crittografo (quindi non scrivo algoritmi di cifratura) ne' un crittanalista (non studio sistemi per violare questi algoritmi) per cui cercherò di affrontare l'argomento non addentrandomi nel campo matematico della questione ma solo analizzando le sue problematiche a livello pratico.

Come detto, la crittografia è una scienza antica: pensate che già qualche centinaio d'anni prima di Cristo gli Ebrei avevano inventato un sistema per "scrivere nascosto" denominato Atbash e che consisteva in un semplice scambio della prima lettera del loro alfabeto con l'ultima, della seconda con la penultima, e così via.

Ma è ad un grande personaggio dell'antichità che si deve l'utilizzo del primo vero cifrario, o sistema di cifratura, che da lui prese il nome: **Cesare**.
Per le comunicazioni segrete con i suoi generali, Cesare adottava un cifrario a sostituzione in cui ogni lettera dell'alfabeto era sostituita da quella che si trovava ad un certo numero di posizioni avanti. Cesare utilizzava 3 spostamenti (la "chiave" del cifrario) per cui le lettere ABC venivano cifrate come DEF. L'utilizzo di un tecnica di protezione di dati sensibili portò allo studio dei sistemi per poter decifrare i messaggi segreti: nacque così anche la **crittanalisi**, che è la scienza che studia i cifrari e cerca di trovarne i punti deboli per poter riportare i dati alla forma originale e leggibile.
Nel corso dei secoli le tecniche di protezione dei dati si sono fatte sempre più sofisticate via via che la conoscenza matematica apportava nuovi sistemi ai crittografi. Il 1500 è il secolo che dà i maggiori contributi a questa scienza: sono di quegli anni alcuni dei più noti cifrari antichi, come **il disco dell'Alberti** o il famoso **cifrario di Vigenère** (in realtà questo cifrario è di Giovan Batista Belaso ma è stato attribuito erroneamente allo scienziato francese), considerato all'epoca della sua pubblicazione inviolabile e rimasto in uso fino agli inizi del XX secolo, nonostante nella seconda metà dell'800 ne fosse stata pubblicata la crittoanalisi.

Nel 1883 Kerckhoffs capì il reale problema della crittografia:
"La sicurezza di un crittosistema non deve dipendere dal tener celato il crittoalgoritmo. La sicurezza dipenderà solo dal tener celata la chiave."

Nel 1918 **Gilbert Vernam** perfezionò l'idea alla base del cifrario di Vigenère, e cioè l'uso di una parola chiave da interpolare con il testo da cifrare, proponendo l'uso di chiavi lunghe almeno quanto il messaggio. **Claude Shannon**, padre della Teoria dell'informazione, ha dato ragione a Vernam dimostrando che questo "è l'unico metodo crittografico totalmente sicuro possibile". Questo assioma è valido anche oggi: tutti gli algoritmi di cifratura, infatti, sono solo dei sistemi per poter ovviare al fatto che non è possibile utilizzare chiavi lunghe migliaia di caratteri.

Il più famoso sistema di cifratura del XX secolo è la macchina **Enigma**, sfruttata durante la II Guerra mondiale dalle truppe naziste per cifrare i loro messaggi militari. Enigma era una specie di macchina da scrivere elettrica in cui, impostando un codice su delle apposite ruote e battendo il testo da cifrare o decifrare sulla tastiera si aveva il corrispondente testo decifrato o cifrato visualizzato dall'accensione in sequenza di una serie di lettere luminose.
Grazie alla nascita dei computer ed alla possibilità di eseguire calcoli sempre più complessi in sempre minor tempo la crittografia ha sviluppato algoritmi di cifratura sempre più sofisticati e complessi. Grazie a questo si sono potute adottare chiavi di lunghezza più contenuta, mantenendo comunque elevati standard di sicurezza.

E siamo così arrivati ai giorni nostri. Gli algoritmi di cifratura si sono notevolmente affinati parallelamente allo sviluppo della tecnica: più i computer divenivano potenti, più gli algoritmi venivano perfezionati e resi sempre più complessi. Esaminando il codice di un algoritmo si può benissimo capire su che genere di macchine deve girare. Uno dei più noti algoritmi di cifratura è stato sicuramente il **DES**, inventato negli anni ‘70 e rimasto in uso fino ai giorni nostri. Il DES fu adottato dal Governo americano come standard, il che significa che tutti i documenti segreti erano cifrati con questo algoritmo. Nonostante questa investitura ufficiale, molti crittanalisti erano scettici nei confronti del DES, principalmente perché la sua chiave è lunga soltanto **56 bit** (vale a dire solo 7 caratteri!). Molti critici affermano che ciò era dovuto al fatto che il Governo americano aveva la tecnologia per violare un testo cifrato a 56 bit in tempi ragionevoli ma non per una cifratura a 64 bit. Nonostante questo, il DES è rimasto in uso fino ai giorni nostri: solo nel 2002 è stato scelto il suo successore, l'**AES**, un algoritmo di cifratura che può operare con chiavi a 128 e 256 bit, garantendo quindi uno standard di sicurezza estremamente più elevato.

Abbiamo citato 2 algoritmi molto noti: entrambi appartengono alla classe degli **algoritmi di cifratura “a blocchi”** (block cipher): il testo da cifrare viene suddiviso in blocchi di lunghezza fissa che vengono poi cifrati uno ad uno. Un'altra classe di algoritmi che opera in maniera differente è quella dei **cifrari “a flusso”** (stream cipher), fino a poco tempo fa considerati meno sicuri di quelli a blocchi. Come per tutte le cose, se non si investe in ricerca e studio in un campo di qualsivoglia natura, poco si conoscerà e pochi saranno gli sviluppi ad esso correlati. E così è stato anche per gli algoritmi a flusso, che venivano snobbati dai crittografi a favore dei più noti cugini. A differenza di un algoritmo a blocchi, uno a flusso codifica il testo carattere per carattere: per questo motivo cominciano ad essere molto usati in quei campi in cui si devono gestire grosse moli di dati di lunghezza ignota, come ad esempio una connessione *wireless*. Un altro vantaggio di alcuni algoritmi di flusso è dato dal fatto che spesso il codice di cifratura e quello di decifratura sono identici: questo comporta una minore dimensione e semplicità del programma che implementa l'uso di questi algoritmi. Uno dei più noti e diffusi algoritmi a flusso è l'**RC4**, del 1987, utilizzato ancora oggi in alcuni protocolli informatici, quali il WEP e l'SSL.

Tutti questi algoritmi hanno però un problema di base, quello fatto notare dal Kerckhoffs: nessuno di essi è sicuro se non si tiene segreta la chiave. E' inutile investire su algoritmi capaci di lavorare con chiavi a 256 bit o più se poi le chiavi stesse sono trasmesse su canali non sicure. E per ovviare a questo problema che nel 1976 furono gettate le basi del sistema di negoziazione di una chiave segreta comune su un canale insicuro noto come **Diffie-Hellman**, dal nome dei suoi ideatori. Questo sistema è alla base degli algoritmi di **cifratura “a chiave pubblica”**, che si basa su una chiave liberamente distribuibile (da qui il nome di “chiave pubblica”) e grazie alla quale chiunque può poi stabilire una connessione sicura con chi l'ha distribuito. Generalizzando, il sistema si basa sul principio matematico della fattorizzazione: dato un grosso numero primo risultante dal prodotto di 2 numeri primi più piccoli, è matematicamente impossibile risalire a questi. Grazie a questo fatto, 2 utenti su internet che non si possono incontrare per scambiarsi di persona una chiave sicura possono farlo senza problemi. Il primo utente genererà, infatti, una coppia di chiavi, una pubblica ed una privata: invierà quindi la prima all'altro utente e terrà per sé quella privata. Il secondo utente utilizzerà questa chiave per cifrare il suo messaggio e poi lo spedirà al primo utente: questo potrà

decifrarlo grazie alla sua chiave privata, che sarà l'unica in grado di aprire il messaggio cifrato. Se un altro utente si inserisse sul canale di comunicazione ed intercettasse la chiave pubblica prima ed il messaggio cifrato poi non sarebbe lo stesso in grado di decifrare il testo che i due utenti si scambiano perché la chiave pubblica serve soltanto per la cifratura dei dati.

Questo sistema è alla base del più noto algoritmo di cifratura a chiave pubblica, l'**RSA**, inventato nel 1977 ed ancora largamente utilizzato. In generale, poi, molti degli algoritmi di sicurezza per internet si basano sulla cifratura a chiave pubblica: non potendo sapere chi si collegherà e non avendo un altro canale per contattare il visitatore, un server che riceve una richiesta di collegamento sicuro da un qualsiasi utente negozierà una coppia di chiavi pubbliche con il computer del visitatore: una chiave servirà all'utente per cifrare i messaggi spediti al server, l'altra servirà per cifrare i messaggi spediti all'utente. L'importante, in questo tipo di collegamenti, è evitare il cosiddetto attacco del “man-in-the-middle”, vale a dire dell'intruso che si interpone fra il server ed il visitatore e negozia al loro posto le chiavi. In questo caso sia il server che l'utente credono di negoziare le chiavi con la controparte ma in realtà il pirata informatico sta ri-negoziando le chiavi in modo da far sembrare la negoziazione regolare ma in realtà sta iniziando l'attacco… Esempio:

Alex conosce Marco su internet e decide di scambiare con lui dei dati personali per cui decide di utilizzare un algoritmo di cifratura a chiave pubblica, però non si è accorto che Luca sta intercettando le sue comunicazioni con Marco. Alex genera una coppia di chiavi ed invia quella pubblica a Marco: Luca, che è in ascolto sul canale, intercetta la chiave pubblica di Alex e la tiene per sé. Genera poi un'altra coppia di chiavi ed invia quella pubblica a Marco. Marco, a questo punto, riceve una chiave e pensa che sia quella inviatagli da Alex ma non sa che in realtà è quella di Luca. Adesso Marco cifra un messaggio con la chiave di Luca e lo invia ad Alex. Luca lo intercetta e lo può decifrare tranquillamente con l'appropriata chiave segreta. Decifrato il messaggio, lo cifra nuovamente ma questa volta con la chiave pubblica di Alex e glielo rispedisce. Alex riceve il messaggio e lo decifra in tranquillità pensando che sia arrivato da Marco ma in realtà gli è stato spedito da Luca. Nella migliore delle ipotesi, Luca si potrebbe essere limitato ad una lettura di esso, così da non far capire né ad Alex né a Marco che il canale è sotto controllo; nella peggiore delle ipotesi, Luca potrebbe aver alterato un documento importante così da far ricevere ad Alex dati ingannevoli. Considerate ad esempio il caso in cui Alex chieda a Marco il fatturato della sua ditta in previsione di un possibile acquisto. Se Luca ha degli interessi affinché la vendita avvenga, potrebbe alterare questi dati facendo credere ad Alex che la ditta di Marco fatturi molto più del reale oppure anche molto meno, a seconda dei casi. Insomma, la trasmissione dei dati sembra sia avvenuta su un canale sicuro ma in realtà i dati sono stati intercettati e modificati da un terzo soggetto!

Tutto questo è possibile perché molti pensano che un messaggio cifrato sia inviolabile. La realtà è che la possibilità che un messaggio cifrato venga violato è molto più elevata del previsto! Questo a causa di diversi fattori:

- si è utilizzato un algoritmo che è stato dimostrato essere facilmente violabile: ce ne sono molti e, purtroppo, continuano ad essere utilizzati!
- si è utilizzata una chiave di una lunghezza non sufficientemente elevata: con la potenza di calcolo degli attuali computer, scegliere un una chiave di lunghezza inferiore ai 128 bit è oggi assolutamente un azzardo!
- si è utilizzata la stessa chiave per cifrare diversi messaggi: con gli algoritmi a flusso e con quelli a blocchi utilizzati in determinate modalità, l'utilizzo della stessa chiave su uno stesso testo in chiaro genera lo stesso testo cifrato! Questo è un aiuto enorme che si dà ad un potenziale malintenzionato!
- si è scambiata la chiave su un canale non sicuro: spesso gli utenti cifrano dei dati e poi si scambiano la chiave via e-mail oppure via IRC!

Quello che si può fare per evitare che un messaggio possa essere alterato od intercettato da un soggetto esterno a chi se lo scambia è utilizzare un algoritmo di autenticazione, una specie di firma digitale che certifica che il messaggio è stato inviato proprio dall'utente con cui stavamo

intrattenendo una corrispondenza. L'autenticazione non cifra il messaggio ma serve a **“certificare”** che i dati ricevuti siano effettivamente stati spediti dalla persona da cui li stavamo aspettando e che sono giunti integri, vale a dire non manipolati. Ma come è possibile? Si può utilizzare un certificato di autenticità rilasciato da un servizio quale VeriSign o Kerberos, certificato che non è ricostruibile da terzi e che è costruito sulla base della chiave pubblica che si distribuirà: questo certificato viene poi utilizzato per “firmare digitalmente” i messaggi che saranno inviati.

Prendiamo il caso precedente di Alex che comunica con Marco e con Luca che spia il loro canale di comunicazione. Alex genera le chiavi ed invia la chiave pubblica ad un servizio di generazione di certificati per avere un certificato di autenticità per la chiave che sta per utilizzare. Ottenuto il certificato, lo spedisce a Marco in allegato alla chiave pubblica. Luca intercetta questa spedizione e decide di agire come prima: prende la chiave pubblica di Alex e la mette da parte e poi genera un'altra coppia di chiavi. Luca però non può utilizzare il suo certificato di autenticità per firmare la nuova chiave altrimenti Marco saprebbe chi ha intercettato il suo messaggio. Luca decide quindi di allegare il certificato di Alex e spedisce i dati a Marco. Questa volta Marco non cadrà nell'inganno perché l'esame del certificato dimostrerà che quella allegata non era la chiave pubblica per la quale era stato generato ed avvertirà Alex di non inviare più nessun dato perché il canale è sotto attacco.

Spero che queste note abbiano aiutato qualcuno a comprendere che la sicurezza non è mai garantita al 100% e che in ambito informatico vale il principio della paranoia: non si è mai sicuri abbastanza di essere… sicuri. Controllate sempre, prima di inviare dati importanti su internet, di aver stabilito una connessione sicura con il server a cui vi state connettendo: la presenza del “lucchetto” nella barra di stato del browser è un segnale importante di presenza di una connessione cifrata. Se non vedete questo simbolo NON inviate nessun dato, anche se il testo sullo schermo sembra affermare il contrario! Potreste essere in presenza di uno di quei numerosi siti illegali che presentano un'interfaccia di un sito noto ma che poi raccolgono le vostre informazioni in maniera fraudolenta!