

crittografia a chiave pubblica



Whitfield Diffie



Martin Hellman

New Directions in Cryptography

We stand today on the brink of a revolution in cryptography. The development of cheap digital hardware . . . has brought the cost of high grade cryptographic devices down to where it can be used in . . . remote cash dispensers and computer terminals. At the same time, theoretical developments in information theory and computer science show promise of providing provably secure cryptosystems, changing this ancient art into a science.

Diffie e Hellmann, IEEE IT **22** (1976)

- nel 1976, Diffie e Hellman pubblicano “New directions in cryptography”
- viene considerato l’inizio della crittografia a chiave pubblica (PKC)
- propongono uno schema di **scambio della chiave** basato sul logaritmo discreto
- introducono l’idea di un **crittosistema a chiave pubblica**
- la prima realizzazione di un crittosistema di questo tipo si ha nel 1978 con l’RSA

crittografia simmetrica

- Alice e Bob condividono la stessa chiave k – scelta e scambiata fra loro prima di cominciare a comunicare
- la chiave dà luogo a una funzione di cifratura e_k e una funzione di decifratura d_k
- è facile ricavare d_k da e_k (nel DES, l’unica cosa che cambia è l’ordine delle chiavi)
- se si sa cifrare, si sa anche decifrare

idea della crittografia a chiave pubblica

- sviluppare un crittosistema in cui data la funzione di cifratura e_k sia **computazionalmente difficile** determinare d_k
- Bob rende pubblica la **sua** funzione di cifratura e_k
- Alice (e chiunque altro) può scrivere a Bob, cifrando il messaggio con la e_k senza bisogno di accordi preliminari
- Bob è l'unico che può decifrare il messaggio
- analogia con un lucchetto, che chiunque può usare, ma di cui solo Bob ha la chiave

funzioni unidirezionali

- bisogna che la funzione di cifratura e sia una **funzione unidirezionale** (one-way function)
- informalmente, una funzione invertibile $f : \mathcal{P} \rightarrow \mathcal{C}$ si dice unidirezionale se
 - dato $x \in \mathcal{P}$, il calcolo di $f(x)$ è **facile**
 - per **quasi tutti** gli $y \in \mathcal{C}$ il calcolo di $f^{-1}(y)$ è **difficile**
 - dato $x \in \mathcal{P}$, il calcolo di $f(x)$ è realizzabile con una **complessità polinomiale**
 - per **quasi tutti** gli $y \in \mathcal{C}$ il calcolo di $f^{-1}(y)$ **non** è realizzabile con una complessità polinomiale (è **NP-completo**?)
- **Esempio** una funzione ritenuta unidirezionale: sia $n = pq$, p e q numeri primi “abbastanza grandi”, b un intero coprimo con $\phi(n)$; sia $f : \mathbb{Z}_n \rightarrow \mathbb{Z}_n$ t.c.

$$f(x) = x^b \pmod{n}$$

trapdoor

- se la funzione è unidirezionale, anche per Bob è impossibile decifrare il messaggio
- una trapdoor (botola) one-way function è una funzione unidirezionale che diventa facile da invertire, se si conosce un'informazione supplementare
- l'informazione supplementare viene tenuta segreta da Bob, e usata per decifrare
 - ci sarà una **chiave pubblica** nota a tutti – che serve per cifrare
 - e una **chiave privata** nota solo a Bob – che serve per decifrare

- un crittosistema a chiave pubblica non sarà mai a segretezza perfetta
- il testo cifrato $y = e_k(x)$ ci dà informazioni su x
- se conosco e_k - posso provare tutti i plaintext fino a trovare quello che, cifrato, ci dà y
- questo deve essere computazionalmente infattibile

cenni sulla complessità computazionale

- come si misura la complessità computazionale di un algoritmo?
- si parla di algoritmi che operano su numeri interi
- la “grandezza” di un input è pari al numero di bit nella sua rappresentazione binaria ($\approx \log_2 n$)
- la complessità di un algoritmo si misura in termini di **operazioni bit**
- sono operazioni bit:
 - addizione fra due cifre binarie (es. $0 + 1$);
 - sottrazione fra due cifre binarie (es. $1 - 0$);
 - moltiplicazione fra due cifre binarie (es. $1 \cdot 1$);
 - divisione di un intero a *due* cifre binarie per *una* cifra binaria (es. 10 diviso 1);
 - traslazione a *sinistra* di un posto, cioè moltiplicazione per 2, ad esempio 11 diventa 110, e traslazione a *destra* di un posto, cioè divisione per 2.

complessità polinomiale e esponenziale

Definizione

La complessità computazionale di un algoritmo che opera sugli interi è data dal numero di operazioni bit occorrenti per eseguirlo.

è una funzione (della lunghezza dell'input)

Definizione

Un algoritmo A per eseguire un calcolo su numeri interi si dice polinomiale se esiste un intero positivo d tale che il numero di operazioni bit necessarie per eseguire l'algoritmo su interi di lunghezza binaria al più k è $\mathcal{O}(k^d)$.

Definizione

Un algoritmo si dice esponenziale se il numero di operazioni bit necessarie per eseguire l'algoritmo su interi di lunghezza binaria al più k è dello stesso ordine di 2^{ck} , per una costante $c > 0$

esempi

- siano n, m interi, $L(n) = k$ (lunghezza binaria), $L(m) = h$, $k \geq h$
- per calcolare la somma $n + m$ servono al più $2k$ operazioni bit quindi $\mathcal{O}(k)$ l'algoritmo che calcola la somma è polinomiale
- per calcolare il prodotto (usando l'algoritmo "tradizionale") servono al più $\mathcal{O}(k^2)$ operazioni bit
- per calcolare il MCD(n, m) usando l'algoritmo di Euclide servono al più $\mathcal{O}(k^3)$ operazioni bit
- per trovare la fattorizzazione di n usando il metodo delle divisioni successive servono $\mathcal{O}(2^{ck})$ operazioni bit

- polinomiale $\leftrightarrow$ (computazionalmente) facile
- esponenziale $\leftrightarrow$ (computazionalmente) difficile
- mostrare che esiste un algoritmo polinomiale che risolve un dato problema è semplice - basta descrivere l'algoritmo
- ma come si fa a mostrare che **non** esiste un algoritmo polinomiale che risolve un dato problema?

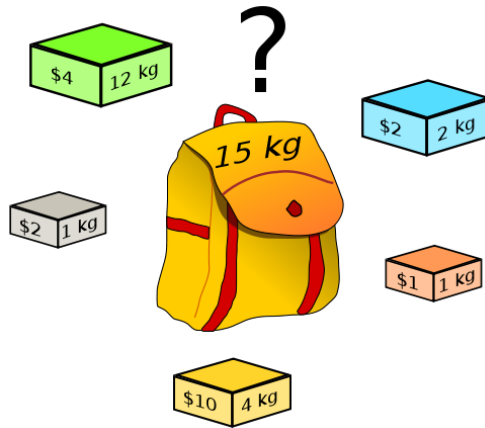
- sia **P** la classe dei problemi P per cui esiste un algoritmo che risolve P in tempo polinomiale (esempio trovare il massimo comun divisore di due numeri interi)
- il problema P appartiene alla classe **NP** se esistono algoritmi (non necessariamente polinomiali) che risolvono il problema ma è possibile verificare, con un algoritmo polinomiale, se un dato assegnato è soluzione o meno del problema (esempio: fattorizzare un intero n)
- si ha $\mathbf{P} \subseteq \mathbf{NP}$
- il problema centrale della teoria della complessità è stabilire se se

$$\mathbf{P} \neq \mathbf{NP}$$

- un problema di **NP** si dice **NP-completo** se per ogni problema Q in **NP** esiste un algoritmo polinomiale che riduce la soluzione di Q a quella di P
- quindi se esiste un algoritmo polinomiale che risolve un problema **NP-completo** allora ogni problema di tipo **NP** sarebbe anche in **P**
- quindi si avrebbe $\mathbf{P} = \mathbf{NP}$
- i problemi **NP-completi** sono i problemi *computazionalmente più difficili* tra quelli di tipo **NP**.

problema dello zaino (knapsack problem)

- un problema **NP**–completo è il **problema dello zaino**
- dato uno zaino che può contenere b unità
- e dati k oggetti di volume $a_1, \dots, a_k$
- il problema è riempire lo zaino usando alcuni oggetti della lista



problema dello zaino (knapsack problem)

- dobbiamo trovare una k -pla $(e_1, \dots, e_k)$, $e_i = 0, 1$
- tale che $b = \sum_i e_i a_i$
- se possibile – oppure mostrare che una tale k -pla non esiste
- il problema è in **NP** – data una k -pla, è facile verificare se è una soluzione
- ma trovare una soluzione è difficile – procedendo per tentativi ho 2^k possibilità da controllare
- e non si conosce un algoritmo veloce per risolverlo – il problema dello zaino è **NP**–completo

esempio

- data la lista (323, 412, 33, 389, 544, 297, 360, 486)
- e dato $b = 1228$
- un metodo possibile è un procedimento **greedy** (avido, goloso)
- a ogni passo, prendo l'oggetto più grande che entra ancora nello zaino
- in questo esempio, ho
$$1228 = 544 + 684 = 544 + 486 + 198 = 544 + 486 + 33 + 165$$
- e qui ci fermiamo
- invece si ha che $1228 = 412 + 33 + 297 + 486$
- se avessimo una lista di 100 numeri di 50 cifre ognuno il problema diventa inavvicinabile

un crittosistema dal problema dello zaino

- la chiave pubblica di Bob è una lista $(a_1, \dots, a_k)$ di interi positivi
- per cifrare, il messaggio dev'essere una stringa binaria di lunghezza k
- $x = (e_1, \dots, e_k)$, $e_i = 0, 1$ – Alice calcola $\sum_i e_i a_i = b$
- e trasmette b a Bob
- nessun attaccante può decifrare – bisogna risolvere un problema dello zaino!
- in questo momento, non può decifrare neanche Bob...

sequenze supercrescenti

- ci sono particolare k -ple per cui il problema dello zaino è facile
- la sequenza $(a_1, \dots, a_k)$ si dice supercrescente se ogni termine è maggiore della somma dei termini precedenti
- $\sum_{j=1}^{i-1} a_j < a_i$
- se ho una lista $(a_1, \dots, a_k)$ supercrescente
- è facile vedere che il procedimento greedy visto prima risolve il problema dello zaino
- e che - se esiste - la soluzione è unica
- per esempio la successione $1, 2, 4, 8, 16, \dots$ delle potenze di due è supercrescente
- il procedimento greedy è quello che si usa per scrivere un intero in base 2

il crittosistema di Merkle–Hellman

- non possiamo usare una sequenza supercrescente come chiave pubblica!
- bisogna “mascherare” questa proprietà
- Bob sceglie una lista $(a_1, \dots, a_k)$ supercrescente
- sceglie poi un intero $n > \sum a_i$ e un intero u con $(u, n) = 1$
- costruisce una nuova lista $(a_1^*, \dots, a_k^*)$
- dove $a_i^* = ua_i \pmod{n}$
- $(a_1^*, \dots, a_k^*)$ è la chiave pubblica
- $(a_1, \dots, a_k)$, n e u sono la chiave privata

- la cifratura funziona come già detto: Alice deve cifrare $x = (e_1, \dots, e_k)$ – calcola $\sum_i e_i a_i^* = b^*$ e trasmette b^* a Bob
- per decifrare, Bob calcola v , l'inverso di u modulo n
- poi calcola $vb^* \equiv b \pmod{n}$
- $b \equiv vb^* = v \sum e_i a_i^* \equiv v \sum e_i u a_i \pmod{n}$
- $v \sum e_i u a_i = uv \sum e_i a_i \equiv \sum e_i a_i \pmod{n}$
- ora sia b che $\sum e_i a_i$ sono minori di $n = \sum a_i$
- quindi $b = \sum e_i a_i$
- Bob deve risolvere un caso facile del problema dello zaino per decifrare

esempio

- data la lista supercrescente $(1, 3, 7, 15, 31, 63, 127, 255)$
- Bob sceglie $n = 557$ maggiore della somma dei termini della lista e $u = 323$ coprimo con 557
- moltiplicando per 323 e riducendo mod 557 ottiene $(323, 412, 33, 389, 544, 297, 360, 486)$
- la stringa binaria 01100101 viene codificata da Alice come $412 + 33 + 297 + 486 = 1228$
- Bob riceve 1228 – sa che l'inverso di 323 mod 557 è 169
- calcola $1228 \cdot 169 \pmod{557}$ e ottiene 328
- $328 = 255 + 73 = 255 + 63 + 10 = 255 + 63 + 7 + 3$
- riottiene quindi la stringa 01100101

il crittosistema di Merkle–Hellman

- il crittosistema ora descritto è dovuto a Merkle e a Hellman (1978)
- è elegante – e molto più semplice dell’RSA
- sfortunatamente non è molto sicuro
- A. Shamir, A polynomial-time algorithm for breaking the basic Merkle - Hellman cryptosystem, IEEE Transactions on Information Theory (1984)
- il “mascheramento” non funziona molto bene