

La crittografia

Elia Alessandro Calderan

Alice e Bob sono figli di due famiglie in inimicizia da tempo immemorabile. Alice proviene da una famiglia di artisti, mentre Bob da una famiglia di matematici. I nostri si conoscono a una mostra presso il National Cryptologic Museum (<http://www.nsa.gov/museum/>) ed è amore a prima vista. I loro incontri sono però impediti da Charles, il padre di Bob; l'unico modo che hanno per comunicare è spedirsi messaggi senza che Charles lo venga a sapere e diseredi senza indugio Bob.

La prima tattica che Alice e Bob possono utilizzare per comunicare è quella di mascherare il messaggio in modo che non sembri una comunicazione di grande rilevanza o non abbia nemmeno la parvenza di essere un messaggio d'amore. La prima volta Alice maschera il messaggio in un dipinto attraverso ridipinture successive, in questo modo Charles quando vede Bob con un quadro in mano non può che sospettare un interessamento all'arte da parte del figlio; Bob per decifrare il messaggio deve esporlo ai raggi-X. Successivamente, per comodità, Alice decide di scriverlo con inchiostri chimici particolari (acido citrico o latte); e, se Bob vuole leggere il messaggio deve semplicemente scaldare il foglio.

In teoria il messaggio può essere nascosto in un microfilm, così come avviene in una tradizione ormai decennale nei film di spionaggio. La cosa non è del tutto assurda, basti pensare che alcune tecniche moderne prevedono addirittura di scrivere un messaggio utilizzando le quattro basi che formano il DNA.

Le tecniche che abbiamo qui descritto fanno parte delle tecniche steganografiche. Tra queste tecniche rientra anche un gioco molto diffuso sulle riviste enigmistiche, il rebus; qui il messaggio è steganografato con delle immagini. La steganografia ha i suoi pro e i suoi contro, il pro risiede nel fatto che il messaggio non è identificabile in quanto tale in modo facile, il contro principale è il fatto che molte volte per la steganografia sono necessari macchinari speciali (i raggi-X, il microfilm...).

Dopo un certo tempo, Bob decide di proporre ad Alice metodi più pratici e meno appariscenti dal punto di vista tecnologico per scambiarsi informazioni. L'idea sta non nel nascondere il messaggio in sé, camuffandolo, ma nel rendere il messaggio incomprensibile a terzi, ovvero di camuffare il significato del messaggio. Si parla in questo caso di crittografia.

Quello che Alice vuole inviare a Bob si chiama, banalmente, messaggio (o chiaro); prima di inviarlo, però lo cifra (o crittografa o critta) attraverso una procedura (l'algoritmo) basata su una chiave (mentre l'algoritmo è lo stesso, il risultato fornito potrebbe variare al variare della chiave utilizzata) affinché il messaggio venga trasformato e risulti a prima vista del tutto incomprensibile per Charles; la comunicazione che può

intercettare Charles si chiama cifrato (o crittato), mentre Bob, una volta ottenuto il cifrato, per sapere quello che Alice gli vuole veramente comunicare, deve decifrarlo (o decrittare), e, per far ciò, ha bisogno di un'altra procedura e di un'altra chiave da applicare al messaggio ricevuto. Le tecniche che Charles utilizza per cercare di carpire le informazioni che Alice invia a Bob si chiamano crittanalitiche.

Partiamo subito con un esempio per chiarire i concetti appena introdotti. Una tecnica che Alice può usare per crittografare il messaggio è la cosiddetta tecnica di shifting (o traslazione); il metodo consiste nel traslare l'alfabeto di un certo numero di lettere, a scelta del mittente. Un caso potrebbe essere il seguente:

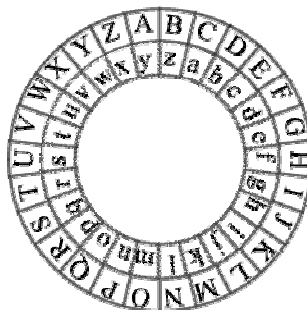
A	B	C	D	E	F	G	..	X	Y	Z
Y	Z	A	B	C	D	E	..	V	W	X

Con questa procedura se Alice vuole inviare a Bob il messaggio "FUGGI", spedisce il crittato "DSEEG", poichè nel crittato ogni lettera è ottenuta andando indietro di due lettere rispetto al messaggio originale. Infatti da G si passa ad E, e da C si passa ad A; un problema è costituito dalla lettera B, dalla quale non è possibile fare due passi indietro, si decide così (vedremo che la scelta non è casuale) di ricominciare da Z, quindi B viene crittata con la lettera Z; la stessa cosa avviene per la lettera A.

Vediamo ora di rendere la cosa leggermente più matematica. Associando a ogni lettera un numero in un modo piuttosto naturale A=0, B=1 e così via, otteniamo la seguente tabella:

A	B	C	D	E	F	G	..	X	Y	Z
0	1	2	3	4	5	6	..	23	24	25
24	25	0	1	2	3	4	..	21	22	23
Y	Z	A	B	C	D	E	..	V	W	X

Come vengono tradotti i ragionamenti fatti poco fa? Allora, la G andava in E e ora il 6 va in 4; la C andava in A e il 2 va in 0. Sembra che sia sufficiente sottrarre 2 dal numero associato alla lettera del messaggio per ottenere il numero del messaggio crittato. Anche qui qualche problema si ha con la B e con la A, fortunatamente questo problema è destinato a sparire se cambiamo rappresentazione. Invece di scrivere in riga i numeri li scriviamo su un cerchio, in questo modo la sottrazione viene a essere uno spostamento antiorario sul cerchio



mentre l'addizione è uno spostamento in senso orario sulla ruota. Questa accoppiata di ruote si chiama disco cifrante di Leon Battista Alberti. Con questo disco Alice è in grado di crittare in maniera rapida qualsiasi messaggio, basta che decida di quanto ruotare il disco interno e legga come devono venire cambiate le lettere del

chiaro per ottenere il cifrato. Noi con questo disco possiamo studiare meglio la matematica che sta alla base delle crittografie, chiamata aritmetica modulare.

Abbiamo visto che sul disco le operazioni possono dare i risultati usuali come $5+2=7$ o $7-2=5$, ma anche dare risultati un pò strani come $1-2=25$. Questi risultati sembrano strani se non si nota il seguente fatto: se noi aggiungiamo, o togliamo, 26 torniamo al punto di partenza, poichè questo equivale a fare un giro completo attorno alla ruota! Quindi $2+26=2$, o anche $2-26=2$. In questa maniera vediamo che $1-2=25$ perchè $1-2=-1=-1+26=25$; possiamo così anche dire che $25=25+26=51$. Quello che stiamo facendo si chiama aritmetica modulo 26, poichè addizione e sottrazione sono insensibili ai multipli di 26. Vedremo in seguito cosa succede con la moltiplicazione.

Dopo questa digressione torniamo ad Alice che ha scritto il messaggio "CIAO BOB, FUGGI" e deve adesso scegliere una chiave, che in uno shifting è un numero da 0 a 25 e rappresenta la distanza tra una lettera e la sua crittata (ovvero di quanto deve essere ruotato il disco interno del disco di Alberti), nell'esempio precedente la chiave è la distanza tra A ed Y ovvero 24 (che è uguale a -2 modulo 26). Il messaggio cifrato che Charles può intercettare è "AGYM ZMZ, DSEEG". Dal momento che anche Bob possiede un disco cifrante, la cosa che deve fare per capire il messaggio che Alice vuole comunicargli è posizionare la Y del disco esterno con la A del disco interno e decrittare il messaggio con lo stesso algoritmo adottato da Alice (che è esattamente il procedimento inverso rispetto a quello adottato da Alice).

Purtoppo anche Charles possiede un disco cifrante, ed è riuscito ad intercettare il messaggio e a capirne il contenuto. Il suo lavoro crittanalitico non è stato molto faticoso; non sapendo la posizione del disco nel momento del crittaggio ha provato tutte le chiavi possibili che Bob potrebbe utilizzare (ovvero tutte le posizioni del disco, male che vada ne deve provare, al massimo, 26) ed ha ottenuto il chiaro di Alice, attraverso i seguenti tentativi:

il messaggio intercettato:

AGYM ZMZ, DSEEG

spostando di una posizione tutte le lettere:

BHZN ANA, ETFFH

spostando di un'ulteriore posizione tutte le lettere:

CIAO BOB, FUGGI

e ha impedito l'uscita di Bob di casa per incontrare Alice. Il problema della sicurezza della comunicazione tra Alice e Bob è un problema serio, allora Alice decide di fare un'operazione più complicata sulle lettere dell'alfabeto: non solo traslarle, ma anche di permutarle. Bob ritiene sia un metodo sufficientemente sicuro dal momento che la forza bruta utilizzata da Charles precedentemente non serve a nulla. Vediamo perchè. Se l'alfabeto fosse costituito da due sole lettere le permutazioni possibili sono 2:

A B | A B

A B | B A

con tre lettere saliamo a 6:

A B C | A B C | A B C | A B C | A B C | A B C

A B C | A C B | B A C | B C A | C A B | C B A

con quattro il numero incomincia a crescere sensibilmente 24, per poi passare a 120 con 5 lettere. Le permutazioni possibili con 26 lettere sono circa 403 milioni di miliardi di miliardi. Quindi è molto improbabile che Charles provi a forzare il codice provando tutte le chiavi; ma le capacità linguistico-matematiche di Charles gli permettono di portare a buon fine la crittanalisi comunque senza fare tutte le prove. Infatti in ogni lingua per ciascuna lettera è possibile associare una frequenza con la quale viene usata in un testo standard, non solo, ma anche ogni sillaba ha una frequenza associata; per esempio, in un testo in lingua italiana, la lettera A compare molto più spesso della lettera V, e la sillaba CHE è molto più usata della sillaba SCI. Ricorrendo a queste statistiche, Charles è in grado di ricostruire il testo originale, calcolando la frequenza con la quale compaiono le lettere e le sillabe nel cifrato e poi tentando di indovinare le rimanenti.

Per chi fosse interessato ad altri metodi di crittografia o alla crittanalisi si consigliano i seguenti siti: <http://www.tonymcrypt.com/Crittografia.htm> per la crittanalisi e altre tecniche crittografiche, e <http://www.liceofoscarini.it/studenti> per una breve storia della crittografia.

Alice e Bob devono trovare tecniche più sicure per comunicare, altrimenti Charles riuscirà sempre a intercettare i loro messaggi.

Decidono quindi che scrivere il messaggio con le lettere non è sicuro poichè sono troppo semplici da maneggiare, e convertono quindi il messaggio in un numero. A ogni carattere viene assegnato un numero, e 'incollandò i vari numeri si ottiene un unico numero che rappresenta il messaggio.

Operando delle manipolazioni aritmetiche su questo numero si può ottenere il crittato; ovviamente, per riuscire poi a decrittare il messaggio, le varie operazioni aritmetiche devono poter essere invertite. Dal momento che stiamo operando in aritmetica modulare dobbiamo vedere quando è possibile invertire un'operazione. Abbiamo già verificato precedentemente che la sottrazione (ciò l'inversa dell'addizione) esiste; dobbiamo ora chiederci se esiste la divisione, ovvero indaghiamo l'esistenza dell'operazione inversa della moltiplicazione; per far questo osserviamo come si comporta la moltiplicazione costruendo delle tabelle moltiplicative nelle aritmetiche modulo 3, 4, 5, 6 e 7.

Modulo 3	Modulo 4	Modulo 5
0 1 2	0 1 2 3	0 1 2 3 4
0 0 0 0	0 0 0 0 0	0 0 0 0 0 0
1 0 1 2	1 0 1 2 3	1 0 1 2 3 4
2 0 2 1	2 0 2 0 2	2 0 2 4 1 3
	3 0 3 2 1	3 0 3 1 4 2
		4 0 4 3 2 1
Modulo 6	Modulo 7	
0 1 2 3 4 5	0 1 2 3 4 5 6	
0 0 0 0 0 0	0 0 0 0 0 0 0	
1 0 1 2 3 4 5	1 0 1 2 3 4 5 6	
2 0 2 4 0 2 4	2 0 2 4 6 1 3 5	
3 0 3 0 3 0 3	3 0 3 6 2 5 1 4	
4 0 4 2 0 4 2	4 0 4 1 5 2 6 3	
5 0 5 4 3 2 1	5 0 5 3 1 6 4 2	
	6 0 6 5 4 3 2 1	

Vediamo che, come ci aspettiamo dalle normali regole dell'aritmetica, un numero moltiplicato per zero dà zero. Ma entra in gioco una novità: non è sempre possibile effettuare la divisione! Questo perchè non è sempre

assicurata l'esistenza di un reciproco (il reciproco di un numero n è $1/n$, ovvero il numero che moltiplicato per n restituisce 1), infatti dire 3 diviso 2, ovvero $3/2$ è come dire tre volte un mezzo, ovvero $3 \cdot 1/2$. Vediamo cosa succede nell'aritmetica modulare cercando di calcolare $2/3$, che è $2 \cdot 1/3$ ovvero due volte il numero che moltiplicato per 3 dà 1:

- Modulo 3
Non fattibile poichè $3=0$ e quindi stiamo dividendo per 0
- Modulo 4
 $2/3 = 2 \cdot 1/3$
Cercando nella tabella qual è il numero che moltiplicato per 3 dà 1 abbiamo che $1/3=3$
 $2/3 = 2 \cdot 3 = 6 = 2$
- Modulo 5
Cercando nella tabella qual è il numero che moltiplicato per 3 dà 1 abbiamo che $1/3=2$
 $2/3 = 2 \cdot 2 = 4$
- Modulo 6
Non esiste numero che moltiplicato per 3 dia 1, infatti danno tutti 0 o 3. Quindi non si può fare $2/3$!
- Modulo 7
Cercando nella tabella qual è il numero che moltiplicato per 3 dà 1 abbiamo che $1/3=5$
 $2/3 = 2 \cdot 5 = 10 = 3$

Si provi a vedere in quali aritmetiche modulari proposte si può calcolare $3/8$ (la risposta dovrebbe essere 0 in modulo 3, 1 in modulo 5, 3 in modulo 7, mentre in modulo 4 non è fattibile perchè $8=0$ e in modulo 6 non è fattibile poichè non esiste $1/2$). Allora quando la divisione è sempre fattibile? Vediamo che se in una riga compare più di uno zero non compare un uno, quindi la presenza di più di un singolo zero sembra implicare la non esistenza del reciproco. Che caratteristiche hanno i numeri che hanno nella riga più di uno zero? In modulo 3, 5 e 7 questo non accade, modulo 4 abbiamo il 2, modulo 6 abbiamo il 2 e il 3. Si osserva che il 2 divide il 4 e che sia il 2 che il 3 dividono il 6, mentre 3, 5, e 7 non hanno divisori, ovvero sono primi. È vero che il modulo in cui lavoriamo è primo allora possiamo tranquillamente fare la divisione? Supponiamo di lavorare modulo n e che n abbia un divisore, diciamo sia p , allora esiste un numero q tale che $p \cdot q = n$. Questo significa che $q \cdot p = 0$ poichè lavoriamo modulo n ma sia q che p non sono 0! Ed ecco quindi che nella riga di p compare un altro 0. C'è infatti un teorema che afferma che solo le aritmetiche modulari in modulo un numero primo ammettono reciproco per ogni numero (che non sia uguale a zero).

Per qualche ulteriore ragguglio più approfondito sulle operazioni modulari (infatti si possono definire anche l'esponenziazione e il logaritmo) si può consultare il sito <http://www.woodrow.org/teachers/math> che offre anche applet Java interattive, mentre per la teoria si può consultare il sito <http://www.amagri.it/Crittologia/Crittografia/> (tutto il sito <http://www.amagri.it/index.htm> è interessante, tant'è che verrà citato anche tra poco, quando parleremo dell'RSA).

Adesso che abbiamo indagato un poco la struttura delle aritmetiche modulari possiamo parlare di crittografia moderna: i metodi a chiave pubblica, più precisamente parleremo dell'RSA. Questo metodo, inventato nel 1978 è così chiamato dai nomi dei suoi tre ideatori: Ron Rivest, Adi Shamir, and Leonard

Adleman. Se nel caso dello shifting l'operazione che effettuava la codifica era l'addizione (e la sottrazione effettuava la decodifica), nella crittografia a chiave pubblica come l'RSA l'operazione che realizza la codifica è la moltiplicazione (e l'esponenziazione). Conseguentemente, per decrittare un messaggio abbiamo bisogno della "divisione" (e del logaritmo).

Dal momento che di matematica dietro questo metodo ce ne è un bel pò, la spiegazione si baserà più sul concetto di firma digitale che sulla tecnica in sè; ma, per chi fosse interessato sin da subito a un'introduzione un poco più formale e dettagliata ci sono i siti <http://mathcircle.berkeley.edu/> e l'articolo sul già citato sito <http://www.amagri.it/Crittologia/Crittografia>, più semplice e discorsivo.

Avevamo lasciato Alice e Bob con il messaggio trasformato in un numero da manipolare; la manipolazione numerica del messaggio si basa su quattro numeri, che sono le chiavi dell' algoritmo di crittaggio e che chiameremo Ca, Da, Cb, Db, sigle che significano "cifraggio Alice", "decrittaggio Alice", "cifraggio Bob", "decrittaggio Bob". Il sistema funziona, a grandi linee, come segue: Alice scrive il messaggio e lo cifra con la chiave Ca, il messaggio arriva a Bob, che lo critta nuovamente con Cb e lo rispedisce ad Alice che lo decritta con Da. adesso il messaggio viene rispedito a Bob, che lo può leggere decrittandolo con la chiave Db. Nonostante questo possa sembrare un'inutile perdita di tempo, nonchè un modo neanche tanto originale per congestionare la rete, è uno tra i metodi più sicuri per lo scambio di informazioni. Osserviamo attentamente cosa è successo nella comunicazione tra Alice e Bob:

$$\text{Msg} \rightarrow \text{Ca}(\text{Msg}) \rightarrow \text{Cb}(\text{Ca}(\text{Msg})) \rightarrow \text{Da}(\text{Cb}(\text{Ca}(\text{Msg}))) \rightarrow \text{Db}(\text{Da}(\text{Cb}(\text{Ca}(\text{Msg})))) \rightarrow \text{Msg} \\ A \rightarrow B \rightarrow A \rightarrow B \rightarrow B \rightarrow A \rightarrow B$$

Tutto questo accade grazie a due caratteristiche della procedura, che prende di volta in volta in input un numero (ovvero una delle chiavi): 1) Applicando la procedura con Ca e poi applicandola sul risultato con Da si ottiene nuovamente il messaggio iniziale, la stessa cosa avviene con Cb e Db. 2) Applicare a un messaggio Cb e poi Da è come applicare prima Da e poi Cb, ovvero i due procedimenti commutano.

Alla luce di queste nuove rivelazioni, con spirito più analitico:

$$\text{Msg} \rightarrow \text{Ca}(\text{Msg}) \rightarrow \text{Cb}(\text{Ca}(\text{Msg})) \rightarrow \text{Da}(\text{Cb}(\text{Ca}))$$

$$A \rightarrow B \rightarrow A \rightarrow B$$

a questo punto le chiavi commutano:

$$\text{Da}(\text{Cb}(\text{Ca}(\text{Msg}))) = \text{Cb}(\text{Da}(\text{Ca}(\text{Msg})))$$

e Da e Ca si cancellano a vicenda:

$$\text{Cb}(\text{Da}(\text{Ca}(\text{Msg}))) = \text{Cb}(\text{Msg}).$$

Adesso Bob applica la procedura con input Db e ottiene il messaggio iniziale poichè Cb e Db si cancellano a vicenda. Quindi alla fine Bob avrà tra le mani il messaggio scritto da Alice!

Veniamo ora al dunque: perchè questo metodo è molto sicuro? Inoltre perchè si chiama a chiave pubblica? Rispondiamo dapprima alla seconda domanda. Nel caso analizzato avevamo una comunicazione tra Alice e Bob; ma, nella maggior parte dei casi anche David ed Emma vorranno spedire messaggi a Bob e Alice. Ovviamente sarebbe poco pratico avere quattro chiavi per ogni corrispondente, allora ognuno espone una "chiave pubblica", che è una coppia di numeri in stretta relazione con le chiavi in possesso di ogni utente, quindi

per i vari processi ognuno userà la propria chiave pubblica e la chiave pubblica del destinatario, senza dover cambiare la propria chiave al variare di questo.

Possiamo così rispondere alla seconda domanda. Il metodo per creare la chiave pubblica si basa sulla fattorizzazione in numeri primi che, per numeri molto grandi, come quelli utilizzati nella tecnica RSA (sono dell'ordine di alcune centinaia di cifre), richiede un tempo computazionale spropositato, infatti per fattorizzare un numero di circa 1000 cifre coinvolto nell'RSA senza utilizzare computer appositamente progettati serve poco meno di un secolo; quindi la comunicazione è praticamente incorruttibile in un tempo ragionevole e, nel caso in cui venga demolito il codice di una chiave pubblica, basta creare una nuova chiave.

La fattorizzazione degli "RSA-numbers", così come vengono chiamati, richiede anni; l'RSA-200, un numero composto da 200 cifre, è stato fattorizzato solo nel maggio di quest'anno (circa 80 Opteron -una famiglia di computer molto potenti-, collegati da una rete che supportava trasmissioni da 1 Gigabit, hanno lavorato ininterrottamente dal Dicembre 2004 al Maggio 2005; se ne fosse stato utilizzato solo uno, la fattorizzazione avrebbe richiesto 55 anni) mentre rimangono ancora aperte le sfide per i numeri tra l'RSA-640 e l'RSA-2048.

Charles, non riuscendo a decrittare i messaggi che Alice e Bob si scambiano, rimane profondamente colpito dalle capacità matematiche di Alice ed esaltato si convince che Alice è più matematica che artista, decide quindi di darle qualche chance di entrare a far parte della famiglia lasciando a Bob la possibilità di frequentarla.

Sitografia

The Numeroscope: an Interactive Laboratory of Numbers and Cryptography, Woodrow Wilson National Fellowship Foundation. <http://ulisse.sissa.it/rec_16set05_1.jsp>

La Crittografia da Atbash a RSA, Liceo Foscarini - Venezia. <http://ulisse.sissa.it/rec_16set05_2.jsp>

Tonycrypt. <http://ulisse.sissa.it/rec_16set05_3.jsp>

Amagri.it, Antonio Amagri. <http://ulisse.sissa.it/rec_16set05_4.jsp>

Icosaedro.it, Umberto Salsi. <http://ulisse.sissa.it/rec_16set05_8.jsp>

National Cryptologic Museum, National Security Agency. <http://ulisse.sissa.it/rec_16set05_5.jsp>

Codes and Ciphers in the Second World War, Tony Sale. <http://ulisse.sissa.it/rec_16set05_6.jsp>

RSA Encryption, Tom Davis. <http://ulisse.sissa.it/rec_16set05_7.jsp>