Ratio Math. 2
(1991), 13 - 28

APPLICAZIONI DELLA MATEMATICA IN PROBLEMI DI COMUNICAZIONE

Luigia Berardi
Dipartimento di Ingegneria Elettrica, Università, L'Aquila.

1. INTRODUZIONE

Sono molto lieta di poter intervenire in questo Convegno, che del resto ho contribuito ad organizzare,e sono lieta di poter parlare a tutti gli intervenuti di questo argomento riassunto nel titolo del mio intervento: le applicazioni della Matematica nei problemi di comunicazioni. Questo argomento è alla base della Crittologia, che è uno degli argomenti del nostro gruppo di ricerca. Inoltre la crittologia e i problemi di comunicazione in genere, sono di interesse non solo nell'ambiente degli Ingegneri ma anche in ambiente Economico. Ciò naturalmente è conseguenza della profonda rivoluzione che l'Elettronica e l'Informatica hanno provocato nel trattamento dell'Informazione e quindi in tutte le transazioni finanziarie del mondo di oggi.

Si pensi che soltanto un centinaio di anni fa avvenne una prima significativa rivoluzione in relazione al nostro problema.

Il primo francobollo postale è del 1839 (Inghilterra), il servizio "Pony Express", ormai mitico e leggendario, è del 1860. Nello stesso periodo nasce il telegrafo via cavo (Morse-1850) e nascono le prime macchine da scrivere (Sholes-1870), il telefono (Bell e Meucci -1876), il cinema (F.lli Lumiere-1895) e il fonografo (Edison-1899). L'aereo è successivo, è del 1918 l'inizio di un costosissimo servizio postale aereo.

(*) Lavoro eseguito nell'ambito del gruppo 40% MPI-Comitato 1.3: Modelli economici, dinamici, discreti, ottimizzazione, crittologia.Il lavoro è stato anche oggetto di una Conferenza durante il Convegno Nazionale della "Mathesis" tenutosi a Gioia del Colle (BA) nel Maggio 1989.

Fin dal 1895 Guglielmo Marconi riusciva a trasmettere. telegraficamente, senza cavo a distanze sempre maggiori. Con l'invenzione della valvola termoionica (Fleming- 1907) nasce la vera possibilità tecnica di una "radio-comunicazione". Nel 1930 Marconi, dal suo panfilo "Elettra", ancorato nel porto di Genova, inaugura l'Esposizione mondiale di "Sidney" parlando agli Elettrotecnici riuniti a congresso e accendendo luci augurali mediante impulsi radio.

La televisione,da una intuizione dell'abate Caselli (1860) risale al 1930, quando fu presentato il primo brevetto da Farnsworth.Essa viene inaugurata in Italia sulla fine degli anni '50.

L'elaboratore elettronico ha enormemente facilitato lo sviluppo delle comunicazioni. Il primo vero elaboratore ENIAC venne realizzato attorno al 1945, esso era dotato di ben diciottomila valvole.La sempre maggiore necessità di migliori comunicazioni condusse nel 1964 alla costituzione di un Consorzio INTELSAT (International Telecommunication Satellite Consortium) che si occupasse della realizzazione e gestione di satelliti artificiali che consentissero collegamenti telefonici e televisivi in tempo reale. Satelliti che percorrono in un giorno tutta l'orbita terrestre ad una distanza di 36.000 Km dalla Terra, tre di essi sono in grado di gestire le comunicazioni di tutto il nostro pianeta.

L'avvento dell'elettronica, la scoperta dei microprocessori che sono stati realizzati negli Stati Uniti,grazie alla corsa allo "spazio", durante gli anni '60 sono il fenomeno forse più significativo in questo contesto. Oggi in quasi ogni casa del mondo, ad alto tenore economico, esiste un Personal Computer, che per potenza e capacità elaborative è di gran lunga superiore a un gigantesco computer che ieri occupava un palazzo di più piani.

Siamo di fatto in grado e ciò sarà realizzato certamente in pochi anni, di metterci in rete con il resto del mondo. Oggi infatti un gran numero di messaggi viaggiano su canali elettronici: posta elettronica, telex, moneta elettronica, ordini e transazioni bancarie. Un grande pericolo incombe: una massa di informazioni che cada in mano di personaggi non autorizzati può dar luogo ad un sottobosco di illeciti di vario genere quale ad esempio lo spionaggio industriale con il relativo commercio delle informazioni ovvero la contraffazione di transazioni ovvero della posta elettronica ovvero l'appropriazione del denaro elettronico. Nasce quindi la necessità di *proteggere l'informazione*.

Si pensi soltanto ai rischi che si possono correre se un "bandito elettronico" acceda a reti o archivi quali quelli di Banche, Enti assicurativi, di Società ovvero di Enti governativi. Ci si potrebbe appropriare quindi non solo di denaro ma anche di informazioni su movimenti contabili, situazioni fiscali, piani di produzione e sviluppo, informazioni militari etc.etc., un operatore finanziario potrebbe disconoscere l'ordine elettronico di un cliente e viceversa.

Un rimedio è certamente quello di disporre di buoni prodotti crittografici. La crittografia, che oggi si presenta ben più sofisticata che non ieri quando il suo uso era riservato al solo ambito militare, è oggi una disciplina nella quale

si osserva un grosso fermento. Convegni internazionali, europei, nazionali, gruppi di ricerca, conferenze, tesi di Laurea e Corsi Universitari.

Lo scopo di questa Conferenza è quello di accostare alla problematica della Crittografia ed alla Matematica che in essa si adopera. Attraverso una panoramica di tipo anche storico ci si vuole accostare a Modelli Matematici concreti, nei quali è utilizzata una Matematica di estrema semplicità. Gli argomenti sono presentati in forma problematica ed euristica.Si parte da situazioni concrete, anche sotto forma di gioco, ovvero dal racconto di fatti storici dai quali è facile evincere la visione di una Crittografia interagente con la Matematica in una continua e stimolante evoluzione .

Tratteremo essenzialmente i seguenti tre problemi:

I. Come "viaggia" un messaggio attraverso un "filo conduttore" o, meglio "via cavo"? Come stabiliamo se il messaggio ricevuto è quello trasmesso?

II. Come si costruiscono le sequenze di numeri e/o di simboli che quotidianamente vediamo su prodotti commerciali e a cosa servono?

III. Come si può trasmettere un messaggio segreto sotto gli occhi di tutti?

Naturalmente approfondendo in altro ambito queste problematiche si incontra anche matematica più sofisticata. Nei problemi di autenticazione ad esempio si fa grande uso della Matematica Discreta ,che nella Crittografia ha trovato una delle sedi più naturali per le Applicazioni.

Per chi volesse approfondire i problemi connessi con l'autenticazione di un messaggio elettronico, ovvero con la firma elettronica consiglio i miei articoli indicati in bibliografia con i numeri [1] e [2]. Un interessante libro, esauriente e vivace in pari tempo e quello di Andrea Sgarro [8]. Un classico con una bibliografia poderosa è il Manuale scritto nel 1948 dal Generale Sacco [7]. Infine un articolo che indichi problematiche complementari a quelle che andremo a trattare ora, scritto da me e dall'amico Prof.Bruno Rizzi, è apparso sul Periodico di Matematiche n.4 (1987), 1-22.

2. UNA APPLICAZIONE DEL SISTEMA DI NUMERAZIONE BINARIO

Dopo aver trattato i vari sistemi di numerazione, ed in particolare quello in base 2, si può porre il problema della trasmissione di un messaggio attraverso un filo conduttore. Costruiamo

"UN CODICE NATURALE"

I PASSO

Ad ogni lettera dell'alfabeto associamo il suo numero d'ordine, rappresentato in base 2. Si ha:

A → 1 → 2^0	→	1
B → 2 → 2^1	→	10
C → 3 → $2^1 + 2^0$	→	11
.....	.....	
Z → 26 → $2^4+2^3+2^1$	→	11010

Se ai simboli 1 e 0 associamo due "segnali" diversi, ad esempio ad 1 una tensione positiva ed a 0 una tensione negativa, oppure ad 1 corrisponde il passaggio di corrente ed a 0 il non passaggio, la trasmissione è possibile.

Ma, all'arrivo, decodificare la sequenza di 0 ed 1 non è possibile perché ogni lettera non è rappresentata da una sequenza di lunghezza costante. Quindi se ad esempio arriva 11010, non si può stabilire se è Z oppure FS (essendo F=110 e S=10).

II PASSO

Rappresentiamo ogni lettera con una sequenza di 5 simboli (tanti quanti ne occorrono per l'ultima lettera dell'alfabeto), anteponendo in ogni sequenza tanti zeri quanti ne servono.

A → 1 → 2^0	→	00001
B → 2 → 2^1	→	00010
C → 3 → $2^1 + 2^0$	→	00011
.....	.....	
Z → 26 → $2^4+2^3+2^1$	→	11010

Ulteriori sequenze possono essere usate, volendo, per codificare anche i simboli di interpunzione (ad esempio, = 11011 etc.)

A questo punto è semplice trasmettere e ricevere, se ogni segnale trasmesso arriva in uscita senza "ombre".

Ma ciò in genere non accade nei canali reali di comunicazione, nei quali sono presenti dei disturbi, che sono causa di *errori*, nel senso che qualche "segnale" ricevuto può essere diverso da quello corrispondente trasmesso.Cosi, ad esempio,se trasmettiamo C e l'ultimo 1 diventa O durante la trasmissione, riceviamo B, ed il nostro codice non *rivela* l'errore, nel senso che non ci possiamo accorgere che era stata trasmessa C e non S.

I canali di trasmissione sono buoni abbastanza per poter supporre che in ogni sequenza non entri più di un errore.

III PASSO

Costruiamo un codice rivelatore di un errore.

In ogni sequenza rappresentativa di una lettera, aggiungiamo un simbolo

scelto tra 0 ed 1 in modo che nella sequenza stessa compaia un numero pari di 1 (Codice di controllo di parità), (ovvero un numero dispari di 1 (Codice di controllo di disparità)).

A → 00001 → 000011
B → 00010 → 000101
C → 00011 → 000110
.....

L'ultimo simbolo, quello sottolineato, è detto *simbolo di controllo di parità*. Ogni blocco di questo nuovo codice è lungo 6.Se in un blocco in uscita compare un numero dispari di 1, vuol dire che in quel blocco c'è un errore. Dunque abbiamo costruito un codice binario *rivelatore di un errore*.

Il problema della correzione di un errore è più complicato dal punto di vista tecnico, ma vale la pena di introdurne almeno la problematica, per completezza.

IV PASSO

Esistono codici che, oltre a rivelare errori, sono in grado di correggere un certo numero di errori, essi prendono il nome di *codici correttori* di errori. Relativamente ad essi diciamo soltanto che se vogliamo un codice correttore di 1 errore, i blocchi rappresentativi delle lettere devono avere a due a due almeno tre simboli diversi.

Infatti, se due qualsiasi blocchi hanno almeno tre simboli diversi, ed entra un solo errore in un blocco, allora nel codice c'e un solo blocco che ha un solo simbolo diverso dal blocco in uscita, mentre tutti gli altri differiscono dal blocco in uscita per almeno 2 simboli.

ESEMPIO.- Supponiamo che un codice sia costituito dalle due parole:

0000 1110

(Le due sequenze potrebbero significare NO e SI, oppure FALSO e VERO oppure NERO e BIANCO, etc.). Trasmettiamo 0000 e supponiamo che entri 1 errore, cioè uno dei quattro simboli diventa 1, allora possiamo avere una delle sequenze:

(*) 1000 0100 0010 0001

Poiché l'altra parola del codice,cioè 1110,ha almeno due simboli diversi da ogni sequenza (*), è possibile correggere la sequenza in uscita sostituendola con la parola del codice che differisce da essa per un solo simbolo.Il principio applicato si chiama *principio di massima somiglianza*.

3. UNA APPLICAZIONE DELLA DIVISIBILITA' PER 11

Nella vita di tutti i giorni capita di vedere impresse sulle confezioni di prodotti commerciali delle sequenze di numeri e/o simboli, che caratterizzano

il prodotto.

Tali sequenze sono costruite mediante *codici rivelatori di* errore.Esempi sono dati dal codice EAN (**E**uropean **A**rticle **N**umber), che identifica i prodotti alimentari (vedasi l'articolo di Ippoliti et altri su Ratio 1), dal codice ISBN (**I**nternational **S**tandard **B**ook **N**umber) che identifica i libri, dal Codice Fiscale Italiano, che identifica ogni persona Italiana o che ha rapporti di lavoro con l'Italia. Il significato dei vari simboli che compaiono in una sequenza costruita tramite uno di questi codici scaturiscono da un preciso procedimento matematico. Inoltre è interessante vedere come viene costruito il simbolo di controllo che permette di rilevare un errore. Allo scopo si presta molto bene il codice ISBN, nel quale il simbolo di controllo viene stabilito in base alla divisibilità di un numero per 11.

Leggiamo su un libro :

ISBN 88 08 03858 0

10 numeri in fila,ognuno con un significato.Vogliamo controllare se il "numero del libro" è giusto, indipendentemente dal suo significato. Moltiplichiamo da sinistra verso destra ogni simbolo per i numeri 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 e sommiamo. Si ottiene:

$$10*8+9*8+8*0+...+2*8+1*0 = 286$$

ed essendo la somma pesata calcolata, divisibile per 11, il numero ISBN è giusto! Vogliamo ora capire.

Un numero ISBN è una sequenza di 10 simboli, anche ripetuti, presi nell'insieme $\{0, 1, 2,...,9, X\}$. La prima parte, costituita da 1 o 2 simboli, indica la Nazione della Casa Editrice. La seconda parte è il numero della Casa Editrice, la terza individua il libro, l'ultima parte, costituita da un solo simbolo è un simbolo di controllo. Il simbolo X, che sta per 10, può comparire solo come simbolo di controllo. Vediamo come si calcola il simbolo di controllo. Sia a_{10} a_9 a_8...a_2 a_1 un numero ISBN nel quale tutti i numeri sono noti tranne a_1. Calcoliamo la somma

$$S = 10\ a_{10}+9\ a_9+8\ a_8+...+2\ a_2$$

Il simbolo a_1, per definizione, è il più piccolo intero positivo tale che $S + a_1$ sia un multiplo di 11. E' chiaro che a_1 può prendere un valore tra 0 ed X=10.

Tale codice rivela un errore. Infatti supponiamo che il numero ISBN

$$a_{10}\ a_9\ a_8...a_2\ a_1$$

presenti un errore, ad esempio nel primo simbolo. Allora esso diventa

$$a'_{10}\ a_9\ a_8...a_2\ a_1$$

La somma $S = 10\,a_{10}+9\,a_9+8\,a_8...+2\,a_2+a_1$ è divisibile per 11.
Supponiamo che anche la somma

$$S' = 10\,a'_{10}+9\,a_9+8\,a_8+...+2\,a_2+a_1$$

sia divisibile per 11. Allora anche la loro differenza è divisibile per 11,cioè

$$S - S' = 10\,a_{10} - 10\,a'_{10} = 10\,(a_{10} - a'_{10})$$

è divisibile per 11. Essendo 11 un primo, segue che $a_{10} - a'_{10}$ è divisibile per 11. Poiché $a_{10} \leq 9$ ed $a'_{10} \leq 9$ segue che:

$$-9 \leq a_{10} - a'_{10} \leq 9$$

L'unica possibilità è allora data da $a_{10} = a'_{10}$
Quindi se vi è un errore il numero ISBN (errato) non è divisibile per 11. Naturalmente non è faticoso inventare dei "codici ISBN" con 4, 6, 12 cifre legati ai numeri primi 5, 7, 13 ed ai rispettivi criteri di divisibilità.

4. LA SEMPLICITA' DELLA MATEMATICA NEI CODICI SEGRETI: LORO EVOLUZIONE STORICA

Il problema trattato nel paragrafo 2 è quello della trasmissione di un messaggio usando un codice rivelatore di almeno un errore.

Non abbiamo trasmesso messaggi segreti, e quindi tutti potevano decodificare e conoscere i messaggi trasmessi.

In questo paragrafo il problema è ben diverso. Il problema è quello di trasmettere *messaggi segreti*, cioè trasmettere in un canale pubblico un messaggio in modo che esso sia *comprensibile solo al mittente ed al destinatario.* Questo problema è molto antico .Un metodo per raggiungere lo scopo, secondo Plutarco, è stato usato già nel IX secolo a.C. Tale metodo è noto come "Scitala Lacedemonica" e consiste in un cilindro su cui viene avvolta ad elica una striscia (a quei tempi di cuoio), sulla quale viene scritto un messaggio per righe longitudinali. Tolto il cilindro,... è realizzato il caos delle lettere. Il destinatario riavvolge la striscia su un cilindro identico al precedente e ... il messaggio è chiaro. La chiave segreta, cioè il segreto, è il diametro del cilindro.

Vogliamo ora "nascondere" un messaggio letterale, sostituendo ad ogni lettera che in esso compare un'altra lettera. L'arte di "nascondere" i messaggi si chiama *Crittografia.*

Scriviamo due alfabeti su due righe successive in modo che le lettere del secondo alfabeto siano slittate di un certo numero di posti rispetto a quelle del primo. Chiamiamo il primo alfabeto *alfabeto in chiaro* ed il secondo *alfabeto*

cifrante.

ESEMPIO

CHIARO : A B C D E F G H I L M N O P Q R S T U V Z
cifrante : f g h i l m n o p q r s t u v z a b c d e

Sostituiamo ogni lettera del messaggio da trasmettere con la lettera corrispondente ad essa sull'alfabeto cifrante. Ad esempio si ha:

TUTTI ASPETTIAMO LA PRIMAVERA
bcbbp faulbbpfrt qf uzprfdlzf

E' un gioco "tradurre il messaggio" cioè cifrare. Se il destinatario conosce il numero dei posti di cui è stato slittato l'alfabeto (tale numero è la chiave segreta del codice), è ancora un gioco "ritradurre" cioè decifrare il messaggio ricevuto.

Una persona sa che un messaggio in italiano è stato cifrato (o *criptato*) con questo sistema,ma non conosce la chiave segreta usata, vede il messaggio e ne vuole conoscere il contenuto, anche se non è autorizzato, vuole cioè *forzare* il codice.

Tale arte si chiama *Crittoanalisi*.

Prima di tutto per forzare il codice, il messaggio deve essere "abbastanza lungo". In tale caso la crittoanalisi di un testo è basata sul fatto che in ogni lingua ogni lettera si presenta con una frequenza sua propria. Esistono tabelle delle frequenze nelle lingue più usate.

Frequenza delle lettere di una lingua.

Tedesco		Inglese		Francese		Italiano		Spagnolo		Portoghese	
A	5	A	7.81	A	9.42	A	11.74	A	12.69	A	13.5
B	2.5	B	1.28	B	1.02	B	.92	B	1.41	B	.5
C	1.5	C	2.93	C	2.64	C	4.50	C	3.93	C	3.5
D	5	D	4.11	D	3.38	D	3.73	D	5.58	D	5
E	18.5	E	13.05	E	15.87	E	11.79	E	13.15	E	13
F	1.5	F	2.88	F	.95	F	.95	F	.46	F	1
G	4	G	1.39	G	1.04	G	1.64	G	1.12	G	1
H	4	H	5.85	H	.77	H	1.54	H	1.24	H	1
I	8	I	6.77	I	8.41	I	11.28	I	6.25	I	6
J	..	J	.23	J	.89	J	..	J	.56	J	.5
K	1	K	.42	K	..	K	..	K	..	K	..
L	3	L	3.60	L	5.34	L	6.51	L	5.94	L	3.5
M	2.5	M	2.62	M	3.24	M	2.51	M	2.65	M	4.5
N	11.5	N	7.28	N	7.15	N	6.88	N	6.95	N	5.5

O	3.5	O	8.21	O	5.14	O	9.83	O	9.49	O	11.5
P	.5	P	2.15	P	2.86	P	3.05	P	2.43	P	3
Q	..	Q	.14	Q	1.06	Q	.61	Q	1.16	Q	1.5
R	7	R	6.64	R	6.46	R	6.37	R	6.25	R	7.5
S	7	S	6.46	S	7.90	S	4.98	S	7.60	S	7.5
T	5	T	9.02	T	7.26	T	5.62	T	3.91	T	4.5
U	5	U	2.77	U	6.24	U	3.01	U	4.36	U	4
V	1	V	1.00	V	2.15	V	2.10	V	1.07	V	1.5
W	1.5	W	1.49	W	..	W	..	W	..	W	..
X	..	X	.30	X	.30	X	..	X	.13	X	.2
Y	..	Y	1.51	Y	.24	Y	..	Y	1.06	Y	..
Z	1.5	Z	.09	Z	.32	Z	.49	Z	.35	Z	.3

Si studia la frequenza di tutte le lettere che compaiono nel testo criptato. Probabilmente la lettera che ha una maggiore frequenza corrisponde nell'alfabeto in chiaro alla lettera che ha una maggiore frequenza in quella lingua.

E' chiaro che quanto più il testo è lungo, tanto più lo studio delle frequenze porta al testo reale.

Il codice descritto è molto antico ed è noto come Codice di Cesare, in quanto usato da Giulio Cesare nella guerra in Gallia e nella corrispondenza epistolare con i familiari. Esso è un esempio di codice monoalfabetico. (Si chiamano monoalfabetici i codici in cui per criptare viene usato un solo alfabeto).

Si potrebbe ragionare diversamente per criptare un messaggio.Dividiamo il testo del messaggio in parti di una lunghezza fissata l, ogni parte sarà chiamata *blocco*.

In ogni blocco "riordiniamo" le lettere in un modo fissato. Come si può fissare un "qualche modo"? Usiamo ancora una permutazione di l oggetti. Ad esempio se l è 4 e se fissiamo la permutazione che porta 1 2 3 4 in 3 1 4 2 si ha:

TESTO IN CHIARO : **ARRI VERE MOMA RTED IORE OTTO** 1 2 3 4
testo cifrato : **rair rvee mmao erdt rieo toot** 3 1 4 2

cioè in ogni blocco scambiamo le lettere usando la permutazione fissata. In effetti facciamo un anagramma di ogni blocco nel "modo" fissato.

I codici segreti di questo tipo si chiamano codici a trasposizione.

Da un punto di vista storico, possiamo dire che nel Medio Evo, per le condizioni politiche, i codici segreti vengono usati poco ed in genere solo per nascondere nomi di personaggi importanti. Verso la fine del Medio Evo, con l'inizio delle relazioni diplomatiche tra i vari stati, i codici segreti diventano una necessità.

Secondo ricerche storiche dovute a Meister (1902) l'uso sistematico dei

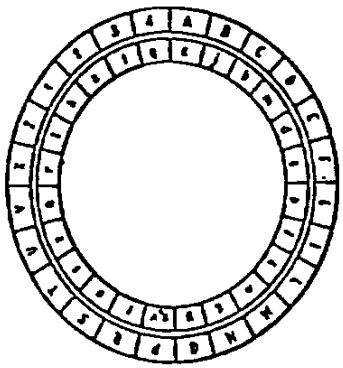
codici segreti ebbe inizio nella Corte Papale, nelle Repubbliche e Signorie Italiane, a partire dal 1300.

E' in questo periodo che la *Crittografia* ha una grossa evoluzione. Troviamo un cambiamento radicale, infatti nascono i primi Codici segreti che non usano un solo alfabeto cifrante, ma molti alfabeti cifranti. Tali codici si chiamano *codici polialfabetici.*

Il primo codice polialfabetico e dovuto ad un nome italiano illustre: Leon Battista Alberti. Egli, su commissione di un segretario pontificio, Leonardo Dato, compilò un codice polialfabetico che fu messo a punto intorno al 1466.

Descriviamo tale codice. Esso è realizzato tramite una coppia di cerchi concentrici. Nel disco più esterno, in cui c'e l'alfabeto in chiaro troviamo 24 caselle, 20 delle quali contengono lettere (mancano per ragioni di sicurezza crittografica, le lettere che si presentano con minore frequenza, cioè J, K , Y, W, Q, H) e le rimanenti quattro i numeri 1 , 2 , 3 ,4.

chiaro



Il disco interno contiene 24 lettere (manca la lettera w ed è u=v, inoltre tali lettere sono in disordine) formanti l'alfabeto cifrante. Esso può ruotare rispetto al primo disco.

Si fissa, prima di cominciare a criptare il messaggio, una lettera dell'alfabeto in chiaro detta *indice* del codice. Poiché c'è una corrispondenza biunivoca tra le caselle dei due dischi, allora alla lettera scelta come indice del codice, sia ad esempio D, corrisponde una ed una sola lettera del disco interno. Come prima lettera del testo cifrato si scrive la lettera corrispondente a D e poi ogni lettera del messaggio viene sostituita con la corrispondente sempre del disco più interno. Supponiamo ora che dopo un certo numero di lettere (anche uno solo) si desideri cambiare alfabeto per rendere più difficile una possibile decrittazione. Si scelga allora uno dei numeri 1, 2, 3, 4 che sono a disposizione, ad esempio il numero 2, che va pensato come inserito nel testo in chiaro.A questo numero si sostituisce la lettera che gli corrisponde nella corrispondenza

dato dai due dischi. Fatto ciò, si ruota il disco finché la lettera corrispondente al numero scelto non si vada a situare esattamente sotto l'indice del codice, la lettera D nel nostro caso. Questa operazione cambia la biezione tra l'alfabeto in chiaro e quello cifrante, quindi otteniamo un altro alfabeto.fino a che non decidiamo di fissare un nuovo numero e cosi via.

Il codice dell'Alberti non ha avuto molto successo.La tendenza dominante nel periodo e quella di semplificare il modo di criptare. Troviamo vari altri codici polialfabetici (ricordiamo,ad esempio, quelli di G.Cardano e di Bellaso) e viene apportata un'altra innovazione: l'introduzione della *parola chiave*. Anche questo è dovuto ad un italiano Giambattista Della Porta (1563), l'inventore della camera oscura.

Vediamo come si usa la parola chiave in un codice polialfabetico.

1. Si fissa una parola del tutto arbitraria ma contenente lettere tutte distinte (ciò perché ad ogni parola corrisponderà un diverso alfabeto).

2. Si scrive tale parola il numero di volte necessario per avere una lunghezza pari a quella del messaggio.

ESEMPIO. Sia PORTA la parola chiave e si voglia inviare un messaggio:

testo in chiaro : **NOI ABBIAMO UN SEGRETO**
parola chiave : **por taporta po rtaport**

Cosa ne facciamo della parola chiave? Ogni lettera di essa, ad esempio p, ci dice quale alfabeto dobbiamo usare per criptare la lettera del messaggio corrispondente. Nel caso generale quindi si deve dare una permutazione dell'alfabeto per ogni lettera della parola chiave. La costruzione di queste permutazioni costituisce il codice stesso. Come esempio vediamo il *sistema Porta* che si assegna mediante le seguenti *tavole*,dette di Della Porta.

Tavola Della Porta

ab	A B C D E F G H I J K L M	N O P Q R S T U V W X Y Z
cd	A B C D E F G H I J K L M	Z N O P Q R S T U W V X Y
ef	A B C D E F G H I J K L M	Y Z N Q P Q R S T U V W X
gh	A B C D E F G H I J K L M	X Y Z N O P Q R S T U V W
ij	A B C D E F G H I J K L M	W X Y Z N O P Q R S T U V
kl	A B C D E F G H I J K L M	V W X Y Z N O P Q R S T U
mn	A B C D E F G H I J K L M	U V W X Y Z N O P Q R S T
op	A B C D E F G H I J K L M	T U V W X Y Z N O P Q R S
qr	A B C D E F G H I J K L M	S T U V W X Y Z N O P Q R
st	A B C D E F G H I J K L M	R S T U V W X Y Z N O P Q
uv	A B C D E F G H I J K L M	Q R S T U V W X Y Z N O P
wx	A B C D E F G H I J K L M	P Q R S T U V W X Y Z N O
yz	A B C D E F G H I J K L M	O P Q R S T U V W X Y Z N

In questo esempio le lettere minuscole scritte in testa danno il nome all'alfabeto di quella riga, che è ottenuto dividendo l'alfabeto in due parti di 13 lettere ognuna e stabilendo una biezione tra i due insiemi di 13 elementi. Allora se la lettera della parola chiave che stiamo considerando e una a oppure una b,si cripta la lettera del testo corrispondente con l'alfabeto di nome ab, e così via.

parola chiave : **por taporta po rtaport**
testo in chiaro : **NOI ABBIAMO UN SEGRETO**
testo cifrato : **hin rouosqb bh avtlxbk**

Un codice polialfabetico, che ha raggiunto maggiore notorietà, è quello dovuto al francese Blaise de Vigénère. Egli nel 1586 pubblica il suo codice nel quale fa uso di una tavola quadrata, già introdotta dall'abate Tritemio e nota come *tavola di Vigénère*,sulla quale vi e veramente molto da dire.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Ogni riga della tavola e un alfabeto di Cesare, a cui diamo il nome della lettera posta a sinistra. Per criptare un messaggio si usano tanti alfabeti di Cesare quante sono le lettere della parola chiave.

ESEMPIO

parola chiave	:	**por taporta po rtaport**
testo in chiaro	:	**NOI ABBIAMO UN SEGRETO**
testo cifrato	:	**ddc tbqzrgomc**. . . . .

La tavola è un quadrato latino cioè una matrice in ogni riga e colonna della quale vi è una permutazione della prima riga o colonna. Si tratta anche della tabella additiva delle classi resto modulo 26, quindi un esempio di gruppo finito. Se prendo 24 ho la matematica dell'orologio e se prendo la moltiplicazione modulo 26 ho un caso di divisori dello zero. Si può anche scrivere 00,01,....,25 al posto di A, B, ..., Z e allora il passaggio dal testo in chiaro a quello cifrato è la somma modulo 26. Insomma gli spunti sono tanti.

Il segreto di questo codice è tutto nella parola chiave. Quindi, il destinatario del messaggio (e si spera solo lui, oltre il mittente) conosce la parola chiave. Allora è in grado di decifrare il messaggio in arrivo usando il procedimento al contrario (ovvero la sottrazione modulo 26).

Il problemaè molto più complicato che nei codici di Cesare o comunque monoalfabetici. Si pensi che il codice di Vigénère è stato usato da eserciti di diversa nazionalità ed ha resistito a tutti gli "attacchi" dei crittoanalisti per circa tre secoli.

Doveva arrivare un ufficiale prussiano di nome Kasiski (1863) per forzare il codice di Vigénère.Il metodo da lui usato è noto come *test di Kasiski* e si può applicare se il testo cifrato è abbastanza lungo. Il test di Kasiski è basato sui seguenti due punti.

1. Trovare la lunghezza della parola chiave.
2. Trovare le lettere della parola chiave.

Trattiamo il punto 1. Si cercano nel messaggio cifrato tutte le sequenze (cioè i gruppi) di tre o più lettere consecutive che si ripetono. *E' molto probabile* che a sequenze uguali del cifrato corrispondano sequenze uguali del testo in chiaro. Se cosi è vuol dire che le prime lettere di sequenze eguali sono state criptate con lo stesso alfabeto della tavola di Vigénère, analogamente le seconde, le terze, ... Da ciò segue che alle prime lettere delle sequenze uguali corrisponde la stessa lettera della parola chiave, analogamente alle seconde lettere e cosi via. Ma allora la distanza (= numero delle lettere) tra due sequenze uguali è un multiplo della lunghezza della parola chiave.

Quanto è lunga la parola chiave? *Molto probabilmente* la sua lunghezza è pari al M.C.D delle distanze tra le sequenze uguali tra loro, che si ripetono.

Chiaramente una distanza "strana" tra due sequenze uguali, nel senso che non ha divisori comuni con le altre distanze, deve essere scartata; questo è il motivo per cui abbiamo detto "molto probabilmente la sua lunghezza ...". Ciò accade quando due sequenze uguali non corrispondono a due sequenze uguali del testo in chiaro.

Il problema al punto 1 è risolto e la lunghezza l è nota.

Trattiamo il punto 2. Conosciamo la lunghezza l della parola chiave. Il problema è scoprire le lettere della parola chiave. Notiamo che, nel testo cifrato, alla prima lettera, alla (l+1) - ma lettera, alla (2l+1) - ma lettera ... corrisponde la stessa lettera della parola chiave e quindi tutte queste lettere sono state criptate con uno stesso codice di Cesare (cioè con una stessa riga del quadrato di Vigénère).

Ripetendo il ragionamento, si ha che la seconda, (l+2) - ma riga, (2l+2) - ma riga, ..., sono state criptate con lo stesso codice di Cesare. In definitiva si ha:

{ I, (l+1) - ma, (2l+1) - ma, ... }
{ II, (l+2) - ma, (2l+2) - ma, ... }
{ III, (l+3) - ma, (2l+3) - ma, ... }
.....
{l - ma, 2l - ma, 3l - ma, ... }

Le lettere di ognuno di questi insiemi sono state criptate con lo stesso codice di Cesare.Allora studiamo la frequenza delle lettere in ognuno di essi,con lo stesso metodo usato nei codici di Cesare.Scoperta una lettera, è noto il codice di Cesare usato, e quindi anche il nome dell'alfabeto (cioè la lettera che compare in testa), che ci dà la lettera della parola chiave. Allora il codice di Vigénère *è forzato* e quindi perde il suo interesse.

E' possibile "complicare" il codice di Vigénère per avere un codice sicuro? (sicuro significa capace di resistere agli attacchi).

SI

Un codice completamente sicuro è il *codice di Vernam* (1926). Questo è un codice di Vigénère (cioè si usa la tavola di Vigénère) nel quale si usa una parola chiave avente una lunghezza pari alla lunghezza del messaggio.

Sembra che il *telefono rosso*, esistente tra la Casa Bianca e il Cremlino, usi un codice di Vernam per comunicare.

Questo codice di Vernam è in un certo qual senso il padre dei codici costruiti mediante massaggi scritti con sequenze binarie e chiavi con sequenze pseudocasuali; codifica e decodifica sono la somma modulo 2. Siamo nel pieno della Teoria di Shannon ed il poco spazio ci impedisce di approfondire. Notiamo solo che la parola chiave, in un codice di Vernam, non è detto che debba essere di senso compiuto (come nel codice di Vigénère, poiché ai tempi di Vigénère il messaggero aveva necessità di ricordare a memoria la parola chiave), ed essa può essere una sequenza, lunga come il messaggio, di simboli

qualsiasi,ad esempio di O ed 1 casuali ovvero pseudocasuali. Vi è oggi una grande richiesta di sequenze pseudocasuali che, dal punto di intuitivo, sono sequenze scritte con una legge e quindi facilmente riproducibili in un secondo computer, ma con proprietà statistiche molto vicine a quelle della casualità.

Riprendiamo i codici a sostituzione ed a trasposizione. Se mescoliamo questi codici applicando sia gli uni che gli altri otteniamo un codice composto.

Un codice molto usato oggi per trasmettere un messaggio segreto attraverso un canale pubblico è un codice costruito dalla IBM il cui algoritmo (ma non le chiavi che ognuno fissa come vuole) è stato addirittura pubblicato al Federal Register of Information Processing Standard con data 1 Agosto 1975. Tale codice si chiama Data Encription Standard, in sigla DES.

Il DES è un codice binario. Per trasmettere un messaggio segreto attraverso un canale pubblico con il DES si opera nel modo seguente:

1. Si trasforma il messaggio in una sequenza di simboli binari.
2. Si suddivide tutta la sequenza in blocchi, ciascuno costituito da 64 simboli.
3. Ad ogni blocco si applicano per 16 volte codici a sostituzione e a trasposizione.

La chiave segreta K del DES è data da tutte le permutazioni scelte nei 16 passi.

Il DES è oggi uno dei codici segreti più usati. Ad esempio nella realizzazione di una carta a banda magnetica tipo BANCOMAT, la Banca di emissione usa il DES. Indaghiamo sull'argomento.Una carta Bancomat ha una zona a banda magnetica che può memorizzare alcuni dati. Quando una Banca consegna ad un suo utente una carta del Bancomat, insieme ad essa comunica al cliente un numero segreto che va sotto il nome di PIN (Personal Identification Number), composto da 5 cifre della numerazione decimale.

Come si calcola il PIN?

Il conto corrente dell'utente ha un certo numero N. La Banca cripta il numero N tramite il DES, cioè scrive N in forma binaria e poi applica a questa sequenza il DES, usando una chiave segreta K.

Sulla striscia magnetica viene memorizzato il numero di codice della Banca ed il numero N del conto dell'utente. Quando l'utente usa la carta del Bancomat in un terminale, il terminale legge sulla carta sia N che il numero di codice della Banca che ha emesso la carta.Dal codice della Banca è facile per il terminale desumere da un archivio in suo possesso la chiave K usata dalla Banca,quindi applicando il DES ad N e a K è in grado di ricalcolare il PIN della carta. A questo punto chiede il PIN all'utente che lo deve digitare e solo se i due numeri coincidono il terminale fornisce l'assenso all'operazione. Tutto dura pochi istanti, naturalmente. Se una persona vuole scoprire il PIN di una carta, deve fare mediamente 50.000 tentativi, cioè 200 ore di sportello. Naturalmente ci sono varie altre possibilità, non ultimo leggere il PIN con un binocolo!

L'esempio del PIN, abbastanza semplice, fa capire come può essere usato

un codice segreto per proteggere un messaggio. Oggi con il DES si riescono a realizzare procedimenti molto più sofisticati del semplice calcolo di un PIN,quali ad esempio l'autentifica di un documento inviato elettronicamente o addirittura la sua "firma elettronica", cfr. [1].

BIBLIOGRAFIA

1. L.BERARDI, *Some remarks about an electronic signature derived from a generalized RSA-code*, J.of Information & Opti.Scien. 1 (1990), 189-194.
2. L.BERARDI-A.BEUTELSPACHER, *I buoni angeli custodi, ovvero i protettori di un messaggio*, Archimede 2-3 (1988), 129-140.
3. L.BERARDI-M.DI FONZO, *Protezione delle informazioni su personal computer*, Atti del Primo Simposio Nazionale su "Stato e prospettive della Ricerca Crittografica in Italia", a cura della Fondazione "U.Bordoni", Roma, Ottobre 1987, 167-173.
4. L.BERARDI-F.EUGENI, *Bloking sets e teoria dei Giochi: origini e problematiche*, (dedicato al Prof.Renato Nardini per il suo 70-mo compleanno), Atti Sem.Mat.Fis.Univ.Modena,34 (1988),165-196.
5. A.BEUTELSPACHER, *La scuola elementare della teoria dei Codici*, Quaderno n.1, suppl.didatt., Semin. Geom. Combin. Univ. L'Aquila, 1985.
6. M.CERASOLI-F.EUGENI-M.PROTASI, *Elementi di Matematica Discreta*, Zanichelli, Bologna, 1988.
7. L.SACCO, *Manuale di Crittografia*, Roma, Litografia Covi, 1947.
8. A.SGARRO, *Crittografia*, Muzzio Editore, 1986.
9. C.E.SHANNON, *Communication theory of secrecy system*, BSTJ 28 (1949), 666-715.

I lettori interessati a proseguire le ricerche possono consultare i vari volumi della serie Springer Verlag "grigia", dedicata all'Informatica con titoli Crypto 19.., Eurocrypto 19..