Crittografia

L. Giuzzi A. Sonnino

6 dicembre 2004

Indice

Introduzione

Il criptosistema più antico di cui si abbia notizia è il *cifrario di Atbash,* utilizzato nel Libro di Geremia del Vecchio Testamento. Il principio su cui si basava era estremamente semplice: si invertiva l'ordine delle lettere dell'alfabeto (ebraico) e per cifrare si usava la seguente sostituzione:

ת	ש	ר	ק	···	ד	ג	ב	א
↓	↓	↓	↓		↓	↓	↓	↓
א	ב	ג	ד	···	ק	ר	ש	ת

mentre per decifrare bastava compiere sul testo cifrato la stessa operazione. Ad esempio, utilizzando il cifrario di Atbash la parola **אלהים** viene mutata in **תמצנך**. Questo dimostra quanto antica sia la necessità di proteggere una informazione riservata e di particolare importanza contro possibili intrusioni.

In realtà, la crittografia moderna inizia con l'introduzione, nel 1976, del concetto di *chiave pubblica* ad opera di W. Diffie & M. Hellman [DH76] SEGUE...

Capitolo 1

Preliminari

In questo paragrafo introdurremo alcuni concetti di base per stabilire la terminologia e le notazioni che verranno utilizzate in seguito.

1.1 Rappresentazione dei messaggi

Un *alfabeto* è un insieme finito $A = \{a_0, a_1, \ldots, a_{n-1}\}$, che d'ora in poi supporremo sempre in corrispondenza biunivoca con $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, per qualche $n \in \mathbb{N}$, oppure con un campo finito $\mathbb{F}_n$, con $n = p^h$, $h \in \mathbb{N}$ e p primo. Una sequenza composta da k elementi di A si dice $r-$*blocco,* o blocco di lunghezza r, e chiameremo *lettere,* o *simboli,* gli elementi che lo compongono. Normalmente denoteremo un r-blocco con $\mathbf{a} = (a_0, a_1, \ldots, a_{r-1})$.

Ad esempio, l'alfabeto latino `A-Z` può essere rappresentato mediante l'insieme $\mathbb{Z}_{26}$ degli interi modulo 26. Volendo invece considerare 2-blocchi possiamo etichettare le coppie ordinate di simboli (x, y) mediante la corrispondenza

$$\begin{array}{ccc} \mathbb{Z}_{26} \times \mathbb{Z}_{26} & \longrightarrow & \mathbb{Z}_{676} \\ (x, y) & \longmapsto & 26x + y. \end{array}$$

Così facendo, otteniamo un'espressione delle singole lettere che compongono ciascun blocco come cifre di un elemento di $\mathbb{Z}_{676}$ espresso in base 26, mentre i 2-blocchi sono rappresentati da numeri di due cifre nella stessa base. Similmente, volendo considerare 3-blocchi, si può usare la corrispondenza

$$\begin{array}{ccc} \mathbb{Z}_{26} \times \mathbb{Z}_{26} \times \mathbb{Z}_{26} & \longrightarrow & \mathbb{Z}_{17576} \\ (x, y, z) & \longmapsto & 676x + 26y + z. \end{array}$$

In generale, gli r-blocchi composti da simboli di un alfabeto di lunghezza N possono essere rappresentati da sequenze di interi compresi fra 0 e $N^r - 1$, facendo corrispondere ad ogni blocco un elemento di $\mathbb{Z}_{N^r}$ espresso in base N.

Un *testo* è un elemento o una sequenza concatenata di elementi dell'insieme

$$\mathcal{T} := \bigcup_{r \geq 0} \mathbb{Z}_n^r,$$

dove nel caso di un campo finito l'unione viene fatta su $\mathbb{F}_n^r$. Un *linguaggio* è un sottoinsieme di $\mathcal{T}$. Nel caso di un linguaggio di programmazione, il sottoinsieme è definito in modo molto preciso per mezzo di un certo numero di regole ricorsive, mentre per una lingua parlata tali regole sono perlopiù molto vaghe. In quest'ultimo caso è conveniente adottare un approccio probabilistico per descrivere le possibili occorrenze delle parole e delle singole lettere che le compongono.

Uno *spazio dei messaggi* è una sequenza di variabili casuali $\{X_0.X_1, \ldots, X_{r-1}\}$ con cui a ciascun evento $(m_0, m_1, \ldots, m_{r-1})$ è associata una probabilità

$$Pr(X_j = m_0, X_{j+1} = m_1, \ldots X_{j+r} = m_{r-1})$$

dove gli indici sono ridotti modulo r. Ponendo $j = 0$ possiamo scrivere, più semplicemente, $Pr(m_0, m_1, \ldots, m_{r-1})$. Dalla teoria della probabilità seguono le seguenti relazioni:

i) $Pr(m_0, m_1, \ldots, m_{r-1}) \geq 0$ per ogni sequenza $(m_0, m_1, \ldots, m_{r-1})$;

ii) $$\sum_{(m_0, m_1, \ldots, m_{r-1}) \in \mathcal{T}} Pr(m_0, m_1, \ldots, m_{r-1}) = 1;$$

iii) $$\sum_{(m_0, m_1, \ldots, m_{r-1}) \in \mathcal{T}} Pr(m_0, m_1, \ldots, m_{r-1}) = Pr(m_0, m_1, \ldots, m_{r-1}).$$

Se nello studio di un linguaggio ci limitiamo a valutare blocchi a di lunghezza 1, osserviamo che la loro distribuzione $p(a)$ è indipendente dalla posizione che assumono all'interno di un blocco di lunghezza maggiore di 1, cosicché

$$Pr(m_0, m_1, \ldots, m_{r-1}) = p(m_0)p(m_1) \ldots p(m_{r-1}).$$

Ad esempio, utilizzando i valori riportati nella tabella 1.1, possiamo calcolare la probabilità di occorrenza della parola "ten" nella lingua Inglese come segue

$$Pr(\texttt{TEN}) = p(\texttt{T})p(\texttt{E})p(\texttt{N}) = Pr(\texttt{NET}) \approx 8{,}2 \cdot 10^{-4}.$$

Osserviamo che in questo modello alle parole `TEN` e `NET` resta associata la stessa probabilità.

Un modello più accurato può essere definito considerando il modo in cui, in una certa lingua, una lettera segue un'altra all'interno di una parola. In tal caso lo spazio dei messaggi genera una catena di Markov finita, ossia, una sequenza di

A	0,0804	H	0,0549	O	0,0760	V	0,0099
B	0,0154	I	0,0276	P	0,0200	W	0,0192
C	0,0306	J	0,0016	Q	0,0011	X	0,0019
D	0,0399	K	0,0067	R	0,0612	Y	0,0173
E	0,1251	L	0,0414	S	0,0654	Z	0,0009
F	0,0230	M	0,0253	T	0,0925		
G	0,0196	N	0,0709	U	0,0271		

Tabella 1.1: Distribuzione di probabilità delle lettere in Inglese

variabili casuali $\{X_0, X_1, \ldots, X_{r-1}\}$ tali che la distribuzione di probabilità condizionata X_{s+1}, con $0 \leq s < r$, dipende solo dalla distribuzione di probabilità X_s ed è indipendente dalle distribuzioni di probabilità precedenti. Le probabilità di transizione $Pr(X_{s+1} = m_i | Pr(X_s = m_j)) = p_{ij}$ definiscono una matrice di transizione $P = (Pr(X_i|X_j))$ alla quale resta associato il vettore delle probabilità stazionarie $\mathbf{p} = (p(0), p(1), \ldots, p(r-1))$ che si può calcolare risolvendo il sistema

$$\begin{cases} \mathbf{p}P = \mathbf{p} & \\ p(i) \geq 0 & \text{per ogni } 0 \leq i < r \\ \sum_{i=0}^{r-1} p(i) = 1 & \text{per ogni } 0 \leq i < r. \end{cases}$$

Quindi, per un evento $(m_0, m_1, \ldots, m_{r-1})$ si ha

$$Pr(m_0, m_1, \ldots, m_{r-1}) = p(0)Pr(X_1|X_0)Pr(X_2|X_1) \cdots Pr(X_{r-1}|X_{r-2}).$$

Una descrizione dettagliata delle catene di Markov esula dallo scopo di queste note; per un ulteriore approfondimento al riguardo si rimanda il lettore ad uno dei molti testi di teoria della probabilità che trattano l'argomento.

La matrice di transizione ed il vettore delle probabilità stazionarie per la lingua Inglese, come per altre lingue, sono stati calcolati e sono ben noti. Per l'Inglese si veda, ad esempio, [vT89, tabelle 1.2 e 1.3]. Utilizzando queste tabelle troviamo $P(\texttt{T}|\texttt{E}) = 0{,}1417$, $P(\texttt{E}|\texttt{T}) = 0{,}0404$, $P(\texttt{E}|\texttt{N}) = 0{,}1381$, $P(\texttt{N}|\texttt{T}) = 0{,}1641$, $P(\texttt{N}|\texttt{E}) = 0{,}1212$, e le componenti del vettore delle probabilità stazionarie relative alle lettere `E`, `N` e `T` che sono, rispettivamente, 0,1566, 0,0814 e 0,0773. Per mezzo di tali valori possiamo calcolare le seguenti probabilità:

$$\begin{aligned} Pr(\texttt{TEN}) &= 0{,}0773 \cdot 0{,}1417 \cdot 0{,}1381 \approx 1{,}51 \cdot 10^{-3}, \\ Pr(\texttt{NET}) &= 0{,}0814 \cdot 0{,}1212 \cdot 0{,}0404 \approx 3{,}98 \cdot 10^{-4}, \\ Pr(\texttt{TNE}) &= 0{,}1566 \cdot 0{,}0015 \cdot 0{,}1212 \approx 2{,}85 \cdot 10^{-5}. \end{aligned}$$

Occorre comunque osservare che nel modello appena descritto la probabilità $Pr(X_j = m_0, X_{j+1} = m_1, \ldots, X_{j+r-1} = m_{r-1})$ è indipendente da j, ossia dalla posizione di una certa parola all'interno di un testo. Questo in una lingua comune non sempre è vero. Ad esempio, è quasi certo che il testo di una lettera abbia inizio con la parola `CARO` o `CARA`, mentre è assai difficile che inizi con `MARE`, anche se "mare" è una parola molto comune nella lingua Italiana.

1.2 I criptosistemi classici

Una *trasformazione crittografica* è un'applicazione iniettiva $\varphi : \mathbb{Z}_n \to \mathbb{Z}_m$ oppure $\varphi : \mathbb{F}_n \to \mathbb{F}_m$. In pratica, nella maggior parte dei casi si supporrà $n = m$. Sia m un pezzo di testo—o una qualsiasi sequenza di dati in espressa in forma numerica—che chiameremo *testo in chiaro.* Allora diremo che $\varphi(m)$ è un *testo cifrato,* e definiremo un *criptosistema* come un insieme di trasformazioni crittografiche

$$\Phi := \{\, \varphi_k \mid k \in K \,\}.$$

dove l'insieme degli indici K, i cui elementi si dicono *chiavi,* viene detto *spazio delle chiavi,* Definiremo inoltre *spazio dei messaggi in chiaro* l'insieme M di tutti i possibili testi in chiaro, e *spazio dei messaggi cifrati* l'insieme C dei testi cifrati, cosicché per una trasformazione crittografica con chiave $k \in K$ scriveremo

$$\varphi_k : M \longrightarrow C.$$

Dato che φ_k, per ogni $k \in \Phi$, è una funzione iniettiva, la sua inversa ϑ_k esiste, e cioè $\vartheta_k(c) = \vartheta_k(\varphi_k(m)) = m$ per ogni $c \in C$. Naturalmente ciò che ci si aspetta da un buon criptosistema è che il calcolo di $m = \vartheta_k(c)$ non sia "fattibile" a partire dalla sola conoscenza di c, a meno di conoscere anche la chiave k. Secondo l'idea di Shannon [Sha49] un siffatto criptosistema, detto convenzionale o a chiave privata, può essere rappresentato schematicamente come in fig. 1.1.

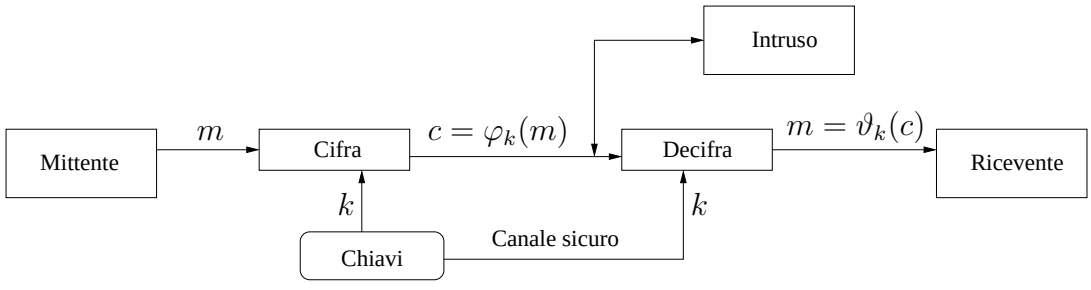

Figura 1.1: Un criptosistema convenzionale

1.2.1 Criptosistema di Giulio Cesare

Uno dei più antichi criptosistemi di cui abbiamo notizia è quello detto *di Giulio Cesare.* Si considerino blocchi di lunghezza 1 e si rappresentino le lettere dell'alfabeto latino, compreso lo spazio fra le parole, con gli elementi di $\mathbb{Z}_{26}$. Per ogni simbolo m si applichi la trasformazione

$$\begin{array}{rccl} \varphi_k : & \mathbb{Z}_{26} & \longrightarrow & \mathbb{Z}_{26} \\ & m & \longmapsto & m+k \pmod{26} \end{array}$$

cosicché, con la chiave $k = 3$, il testo in chiaro `CAESAR` viene trasformato nel testo cifrato `FDHVDU`. In questo caso lo spazio delle chiavi è costituito dall'insieme $K = \{0, 1, 2, \ldots, 24, 25\}$, e si ha $\vartheta_k = \varphi_k^{-1} = \varphi_{26-k}$.

Per un criptanalista è abbastanza facile rompere il criptosistema di Giulio Cesare. Infatti, dato che lo spazio delle chiavi è molto piccolo, la chiave utilizzata può essere determinata per mezzo di una ricerca esaustiva in un tempo ragionevole. Ad esempio nella tabella 1.2 è riportata la criptanalisi del testo cifrato `FDHVDU`.

0	FDHVDU	7	MKOCKB	14	TRVJRI	21	AYCQYP
1	GEIWEV	8	NLPDLC	15	USWKSJ	22	BZDRZQ
2	HFJXFW	9	OMQEMD	16	VTXLTK	23	CAESAR
3	IGKYGX	10	PNRFNE	17	WUYMUL	24	DBFTBS
4	JHLZHY	11	QOSGOF	18	XVZNVM	25	ECGUCT
5	KIMAIZ	12	RPTHPG	19	YWAOWN		
6	LJNBJA	13	SQUIQH	20	ZXBPXO		

Tabella 1.2: Criptanalisi di `FDHVDU` nel criptosistema di Giulio Cesare

Si potrebbe pensare che alla scarsa sicurezza del criptosistema di Giulio Cesare sia possibile ovviare adottando un criptosistema con uno spazio delle chiavi molto grande. Il prossimo esempio dimostra che in generale questo non è vero.

1.2.2 Sostituzione semplice

Nel *criptosistema a sostituzione semplice* si pone $K = \mathrm{Sym}(26)$, ossia, come spazio delle chiavi si prende l'insieme costituito da tutte le permutazioni definite sopra $\mathbb{Z}_{26} = \{0, 1, 2, \ldots, 25\}$. Ovviamente, il criptosistema di Giulio Cesare è un caso particolare di criptosistema a sostituzione semplice, nel senso che le chiavi sono scelte comunque in un sottoinsieme di $\mathrm{Sym}(26)$.

L'insieme delle sostituzioni crittografiche di un criptosistema a sostituzione semplice è definito come

$$\Phi = \{\, \varphi_\pi \mid \pi \in \mathrm{Sym}(26) \},$$

mentre per ogni chiave $\pi \in \mathrm{Sym}(_{26})$ la funzione per decifrare un testo cifrato con φ_π è data da $\theta_\pi = \varphi_{\pi^{-1}}$. Dato che $|K| = 26! = 403291461126605635584 \cdot 10^6$, in questo caso certamente il problema di uno spazio delle chiavi troppo piccolo non si pone. D'altra parte, per forzare questo criptosistema si può tentare con un attacco probabilistico, confrontando la frequenza delle lettere in un testo cifrato con quelle di una tabella di distribuzione simile alla 1.1. Oppure si può utilizzare qualche informazione aggiuntiva, tipo l'argomento trattato, su di un pezzetto di testo cifrato che si è riusciti ad intercettare. In realtà, la quantità di informazione necessaria per rompere un siffatto criptosistema è sorprendentemente poca. Nell'esempio che segue supponiamo di stare analizzando il testo cifrato, composto di 5-blocchi, della tabella 1.3, di cui sappiamo che riguarda la "teoria della comunicazione".

WSPBP	ZFKYL	BQXKW	SBQYU	BXSFP	FICFH	FIZIP
PKFZZ	FIFBS	MXIKY	UBXSF	LYWSH	WSZXY	LWSYE
ZIXBE	IBPQG	BXHIB	SQBHY	EFFQG	FBEKE	PPYAA
BXQGF	PBCWX	EFZIY	PKFZZ	FIFCF	SAYYL	YEZFI
YIPBS	FEEYQ	XKWSB	QYUBX	SFQBX	HWXLB	HFSLF
IFPBY	LYIYA	BXSBB	SZIBS	PFQGF	YEKFQ	QYSBP
KXLBZ	IYPKB	PPBXN	FWZBE	BUUYZ	XPBYL	YLBMF
ZZBSF	EPBPZ	FKYBE	BSAWY	AABSY	ZWIYE	BHIFI
FSZYS	XWSYI	FPBPZ	FSUYY	HBQQX	EBFII	XIBLB
ZIYPK	BPPBX	SFBSO	WYSZX	ZFSLX	SXYQX	LBMBQ
YIFOW	YSZXC	BFSFQ	XKWSB	QYZXB	SKXLX	IBLXS
LYSZF	...					

Tabella 1.3: Testo cifrato con una permutazione di Sym(26)

Possiamo assumere che la parola `COMUNICAZIONE` sia contenuta in un testo simile, cosicché dobbiamo semplicemente cercare una sequenza composta da tredici lettere consecutive in cui la prima lettera è uguale alla settima, la seconda all'undicesima, la quinta alla dodicesima e la sesta alla decima. Scopriamo che una siffatta sequenza compare due volte nel testo ed è `QXKWSBQYUBXSF`. Così facendo ricaviamo la seguente informazione sulla chiave $\pi \in \mathrm{Sym}(26)$:

C	O	M	U	N	I	A	Z	E
↓	↓	↓	↓	↓	↓	↓	↓	↓
Q	X	K	W	S	B	Y	U	F.

Possiamo anche assumere che la parola `TRASMISSIONE` sia contenuta nel testo; quindi dobbiamo cercare sequenze della forma `**Y*KB**BXSF`. Nel testo compare

due volte la sequenza ZIYPKBPPBXSF da cui ricaviamo:

```
T  R  S
↓  ↓  ↓
Z  I  P.
```

Ora sappiamo che all'inizio del testo deve esserci una frase del tipo UN SISTEMA *I COMUNI... da cui, ponendo D al posto di *, ricaviamo la corrispondenza D → L. Procedendo in questo modo alla fine siamo in grado di ricostruire completamente la corrispondenza π:

```
A  B  C  D  E  F  G  H  I  J  K  L  M
↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓
Y  T  Q  L  F  M  A  G  B  D  R  E  K

N  O  P  Q  R  S  T  U  V  W  X  Y  Z
↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓  ↓
S  X  H  O  I  P  Z  W  C  V  J  N  U
```

e quindi di decifrare il testo.

1.2.3 Criptosistema di Vigenère

Il criptosistema che tratteremo in questo paragrafo è stato introdotto nel 1586 da B. de Vigenère [dV86]. Il *criptosistema di Vigenère* è essenzialmente una sequenza di criptosistemi di Giulio Cesare applicati periodicamente. Più precisamente, l'insieme delle trasformazioni crittografiche del criptosistema di Vigenère è definito come

$$\Phi = \{\, \varphi_{(k_0,k_1,\dots,k_{s-1})} \mid (k_0, k_1, \dots, k_{s-1}) \in \mathbb{Z}_n^s \,\},$$

ponendo

$$\varphi_{(k_0,k_1,\dots,k_{s-1})}(m_0, m_1, m_2, \dots) = (c_0, c_1, c_2, \dots)$$

con

$$c_t = m_t + k_u \pmod{n}$$

e gli indici u ridotti modulo s. In altri termini, in un criptosistema di Vigenère il testo viene cifrato mediante una *sostituzione polialfabetica.*. Ad esempio, usando la consueta rappresentazione dell'alfabeto latino $A = \{A, \dots, Z\}$ per mezzo degli elementi di $\mathbb{Z}_{26}$, prendendo come parola chiave la sequenza CHIAVE e sommando lettera per lettera modulo 26, otteniamo la seguente trasformazione:

```
testo in chiaro:   UNCRI PTOSI STEMA COMEQ UESTO NONSE MBRAM OLTOS ...
chiave:            CHIAV ECHIA VECHI AVECH IAVEC HIAVE CHIAV ECHIA ...
testo cifrato:     WUKRD TVWSD NXGTI CJQGX CENXQ UWNNI OIZAH SNAWS ...
```

	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
0	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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

Tabella 1.4: Tavola di Vigenère

Un utile strumento per cifrare e decifrare con il criptosistema di Vigenère è costituito dalla *tavola di Vigenère* riportata in tabella 1.4. Per cifrare, basta scrivere ripetutamente, sotto il testo in chiaro, la parola chiave tante volte quante è necessario ed eventualmente troncando alla fine. Quindi si cerca nella tavola di Vigenère la riga di ciascuna lettera del testo in chiaro e la colonna della relativa lettera della parola chiave. A questo punto la lettera corrispondente nel testo cifrato è quella che si trova nell'intersezione fra riga e colonna. Vice versa, per decifrare si cerca la colonna di ciascuna lettera della parola chiave e si scende fino a trovare la relativa lettera del testo cifrato. La lettera che etichetta la riga trovata è la lettera corrispondente nel testo in chiaro.

Va osservato che una scelta poco oculata della chiave (ad esempio, il nome di un parente, una data di nascita, etc.) potrebbe diminuire notevolmente, a priori, la sicurezza del sistema.

Dopo essere stato considerato inviolabile per quasi tre secoli, il criptosistema di Vigenère è stato rotto nel 1863 dall'ufficiale prussiano F. W. Kasiski con l'in-

troduzione di una tecnica detta dell'*incidenza delle coincidenze.* Questa tecnica viene utilizzata per determinare la lunghezza s della chiave, cosicché per rompere il criptosistema basta scomporre il testo cifrato in una sequenza concatenata di s frammenti di testo, ciascuno dei quali risulta essere cifrato, con una chiave distinta, mediante un criptosistema di Giulio Cesare.

Siano $\mathbf{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,r-1})$, con $i \in \{1, 2\}$, due sequenze di variabili casuali sopra $\mathbb{Z}_n$ indipendenti fra di loro ma identicamente distribuite, cioè:

$$Pr(X_{i,j} = m) = p(m), \quad i \in \{1, 2\}, \quad 0 \leq j < r.$$

Supponiamo di voler calcolare il numero di coincidenze, definito come

$$\sigma[X_1, X_2] = |\{\, j \mid 0 \leq j < r,\ X_{1,j} = X_{2,j}\}|.$$

Dal calcolo delle probabilità si ha

$$Pr(X_{1,j} = X_{2,j}) = \sum_{m \in \mathbb{Z}_n} Pr(X_{1,j} = X_{2,j} = m) = \sum_{m \in \mathbb{Z}_n} p^2(m).$$

Sia G un sottoinsieme di $\mathrm{Sym}(n)$ e si consideri una variabile casuale Π sopra G con distribuzione

$$Pr(\Pi = \pi) = q(\pi).$$

Se X_1 ed X_2 sono cifrati, rispettivamente, con le sostituzioni semplici π_1 e π_2 di G, denotiamo le loro rispettive immagini con $Y_1 = (Y_{1,0}, Y_{1,1}, \ldots, Y_{1,r-1})$ e $Y_2 = (Y_{2,0}, Y_{2,1}, \ldots, Y_{2,r-1})$. Per $0 \leq j < n$ si ha

$$Pr(Y_{1,j} = Y_{2,j} = c) = \sum_{\pi \in G} q(\pi) p(\pi^{-1}(c)).$$

Restano da considerare due possibilità:

Δ_0 : X_1 ed X_2 sono stati cifrati mediante la stessa sostituzione semplice π con probabilità $q(\pi)$;

Δ_1 : X_1 ed X_2 sono stati cifrati, rispettivamente, mediante due sostituzioni semplici π_1 e π_2, scelte in modo indipendente e con rispettive probabilità $q(\pi_1)$ e $q(\pi_2)$.

Ne consegue

$$Pr(Y_{1,j} = Y_{2,j} \mid \Delta_0) = \sum_{c \in \mathbb{Z}_n} Pr(Y_{1j} = Y_{2,j} = c \mid \Delta_0) = \sum_{\pi \in G} \sum_{c \in \mathbb{Z}_n} q(\pi) p^2(\pi^{-1}(c)) = \sum_{\pi \in G} \sum_{m \in \mathbb{Z}_n} q(\pi) p^2(m) = \sum_{m \in \mathbb{Z}_n} p^2(m), \quad (1.1)$$

mentre

$$Pr(Y_{1,j} = Y_{2,j} \mid \Delta_1) = \sum_{c \in \mathbb{Z}_n} Pr(Y_{1j} = Y_{2,j} = c \mid \Delta_1) =$$
$$\sum_{\pi_1, \pi_2 \in G} \sum_{c \in \mathbb{Z}_n} q(\pi_1) q(\pi_2) p(\pi_1^{-1}(c)) p(\pi_2^{-1}(c)) =$$
$$\sum_{c \in \mathbb{Z}_n} \left(\sum_{\pi_1 \in G} q(\pi_1 p(\pi_1^{-1}(c))) \right) \left(\sum_{\pi_2 \in G} q(\pi_2 p(\pi_2^{-1}(c))) \right) =$$
$$\sum_{c \in \mathbb{Z}_n} \left(\sum_{\pi \in G} q(\pi) p(\pi^{-1}(c))) \right)^2 = \sum_{c \in \mathbb{Z}_n} Pr^2(Y = c). \quad (1.2)$$

Utilizzando la matrice di transizione di una certa lingua, la (1.1) fornisce un certo valore δ_0, mentre se prendiamo G come il gruppo composto dalle 26 chiavi del criptosistema di Giulio Cesare, e supponiamo che ciascuna di esse ha la stessa probabilità di $1/26$, allora la (1.2) restituisce il valore $\delta_1 \approx 0{,}03846$. Ne consegue che il valore atteso per $\sigma[X_1, X_2]$ è $\delta_0 r$ sotto l'ipotesi Δ_0 e $\delta_1 r$ sotto l'ipotesi Δ_1.

Ora, per determinare la lunghezza della chiave usata per cifrare un testo con il criptosistema di Vigenère, poniamo

testo in chiaro:	$(m_0, m_1, \ldots, m_{r-1})$
chiave:	$(k_0, ki_1, \ldots, k_{s-1})$
testo cifrato:	$(c_0, c_1, \ldots, c_{r-1})$

cosicché per ogni $0 \leq i < r$ si ha

$$c_i = m_i + k_i$$

con gli indici ridotti modulo s. Supponiamo che le componenti $m_i \in \mathbb{Z}_n$ siano valori indipendenti che la variabile casuale X può assumere con probabilità $p(m_i)$, e che le componenti k_i siano valori indipendenti scelti in G con probabilità $q(k_i)$. Definiamo, inoltre,

$$\mathbf{c}^{(v)} = (c_0, c_1. \ldots, c_{r-v-1}),$$
$${}^{(v)}\mathbf{c} = (c_v, c_{v+1}, \ldots, c_{r-1}).$$

Fatto ciò, il valore atteso per $\sigma[{}^{(v)}\mathbf{c}, \mathbf{c}^{(v)}]$ dovrebbe fornire un'indicazione per determinare la lunghezza della chiave. Invero, se s è un divisore di v e $0 \leq i < r - v$, allora

$$Pr({}^{(v)}\mathbf{c} = \mathbf{c}^{(v)}) = \sum_{c \in \mathbb{Z}_n} Pr(c_i = c_{i+v} = c) =$$
$$\sum_{\pi \in G} \sum_{c \in \mathbf{Z}_n} q(\pi) p^2(\pi^{-1}(c)) = \sum_{\pi \in G} \sum_{m \in \mathbf{Z}_n} q(\pi) p^2(m) = \sum_{m \in \mathbb{Z}_n} p^2(m),$$

Da quanto appena visto deriva il seguente

Teorema 1.1 Il valore più probabile di $\sigma[{}^{(v)}\mathbf{c}, \mathbf{c}^{(v)}]$ è dato da

$$\sigma[{}^{(v)}\mathbf{c}, \mathbf{c}^{(v)}] = \begin{cases} (r-v)\sum_{m\in\mathbb{Z}_n} p^2(m) & \text{se } s \text{ divide } v \\ (r-v)\sum_{c\in\mathbb{Z}_n} Pr^2(Y=c) & \text{se } s \text{ non divide } v. \end{cases}$$

Utilizzando il teorema 1.1 siamo in grado di determinare la lunghezza della parola chiave utilizzata per cifrare un testo, purché esso sia abbastanza lungo da rendere attendibile un'analisi probabilistica. Fissato un certo valore τ, ad esempio $\tau = 26$ potrebbe essere una scelta ragionevole, si elencano tutti i valori $\sigma[{}^{(v)}\mathbf{c}, \mathbf{c}^{(v)}]/(r-v)$ che si ottengono per $0 < v < \tau$ e si determina il periodo s con cui i valori più vicini a quello di δ_0 calcolato con la (1.1) compaiono nell'elenco. A questo punto sappiamo che s è la lunghezza più probabile per la chiave.

L'esempio che segue, ancorché breve, fornisce una semplice schematizzazione dell'idea che è alla base della tecnica di Kasiski. Supponiamo di avere intercettato il seguente frammento di testo cifrato, che sappiamo essere stato cifrato con il criptosistema di Vigenère e con una chiave la cui lunghezza s ci accingiamo a determinare.

```
ECQTC GJWDZ GDBOH TCWGA STKGK HDTCW EWHVB ...
```

Cerchiamo tutti i gruppi di due lettere che compaiono più volte nel testo. In questo caso abbiamo tre occorrenze di `TC` alle posizioni 4–5, 16–17, 28–29 e due occorrenze di `CW` alle posizioni 17–18, 29–30. Osserviamo che le coppie `TC` distano fra di loro 12 posizioni, come anche le coppie `CW`. Ne deduciamo che la lunghezza della chiave deve essere 2, 3, 4, 6 oppure 12, cosicché otteniamo una notevole limitazione della scelta.

Ovvialmente il testo analizzato è troppo breve per affermare con certezza che il calcolo è accurato. Infatti, non è detto che tutte le coppie di caratteri che si incontrano vengano dalla stessa duplice sostituzione; potrebbero esserci, casualmente, coppie uguali che vengono da una cifratura diversa. Comunque, in un testo abbastanza lungo, molte coppie coincidenti vengono dalla stessa cifratura e quindi forniscono un'indicazione sufficientemente precisa per dererminare la lunghezza della chiave. In particolare, è molto probabile che questa lunghezza sia il prodotto dei primi che compaiono più frequentemente nelle fattorizzazioni delle distanze fra coppie coincidenti. A questo punto, determinata la lunghezza s della chiave, il criptanalista deve solo scomporre il testo ciclicamente in s sequenze distinte, una lettera ogni s in ciascuna sequenza, e quindi effettuare un'analisi del criptosistema di Giulio Cesare su ognuna di esse. Nel nostro caso la chiave, di lunghezza 6, è `CODICE`; pertanto sono sei le sequenze da analizzare per forzare il criptosistema.

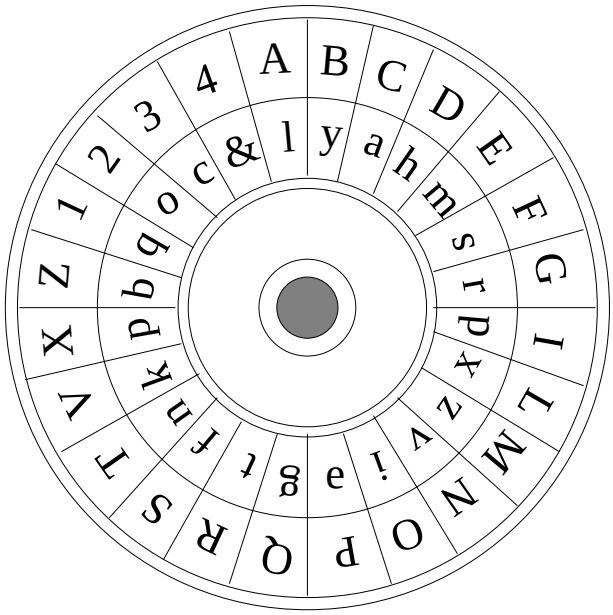

Figura 1.2: Il disco cifrante di Leon Battista Alberti

Una variante più recente del criptosistema di Vigenère è costituita dal *criptosistema di Playfair,* che prende il nome da un dipendente della A. T. & T. che lo propose nel 1917. Si tratta essenzialmente di un criptosistema di Vigenère in cui la chiave—generata in modo completamente casuale per cifrare una volta sola—ha la stessa lunghezza del testo da cifrare. Anche se un criptosistema così congegnato è intrinsecamente sicuro, nella maggior parte dei casi, proprio per la scarsa maneggevolezza di una chiave tanto lunga, non ha alcuna utilità pratica. È interessante ricordare che, negli anni della guerra fredda, la famosa linea rossa fra la Casa Bianca ed il Cremlino utilizzava questo sistema di cifratura. Una chiave come quella del criptosistema di Playfair è detta anche *chiave monouso.*

1.2.4 Il disco cifrante di Leon Battista Alberti

Il *disco cifrante* fu introdotto nel 1470 da Leon Battista Alberti, filosofo, architetto, musicista, pittore e scultore, nel suo "Modus scribendi in ziferas". In effetti, l'Alberti può essere considerato più un criptonalista ante litteram che un crittografo, dato che a lui sono dovuti i primi studi noti sulla distribuzione delle lettere in un testo in lingua latina.

Il criptosistema ideato da Alberti è basato su di un marchingegno come quello illustrato in figura 1.2, composto da due dischi di rame in grado di ruotare indipendentemente intorno allo stesso asse. Il disco esterno, detto stabile, è per il testo

in chiaro. Esso ha 24 caselle contenenti 20 lettere latine maiuscole, con l'esclusione di H, K, W, Y, e con U=V ed I=J. Le rimanenti quattro posizioni sono occupate dai numeri 1, 2, 3 e 4. Il disco interno, detto mobile, contiene le 24 lettere latine minuscole per il testo cifrato. Inoltre, mentre le 20 lettere maiuscole sono disposte in ordine alfabetico, le 24 minuscole sono in ordine casuale.

Un tipico utilizzo del disco cifrante di Alberti può essere schenatizzato come segue. Mittente e destinatario sono entrambi dotati dello stesso disco, con le lettere disposte allo stesso modo, e concordano una lettera iniziale come chiave comune. Per crittare un messaggio, il mittente inizia ruotando il disco interno in modo arbitrario; quindi scrive il testo cifrato riportando per prima la lettera sul disco piccolo corrispondente alla chiave concordata che si trova sul disco grande. Quindi esegue la sostituzione del testo prelevando i caratteri sul disco piccolo in corrispondenza dei caratteri da cifrare sul disco più grande. Dopo aver cifrato la prima parola, il mittente ruota di nuovo, in maniera casuale, il disco interno ed inizia a scrivere un'altra parola facendola precedere dalla nuova lettera che corrisponde, sul disco piccolo, alla chiave concordata sul disco grande. Ad esempio, supponiamo di voler cifrare il testo `NEL MEZZO DEL CAMMIN...`, avendo concordato con il destinatario la lettera `Q` come chiave di partenza, e con il disco nella posizione iniziale illustrata in figura 1.2. All'inizio riportiamo la lettera `g` corrispondente alla chiave iniziale, quindi scriviamo, per la prima parola, `GVMX`. Poi ruotiamo a caso il disco e supponiamo che ora la lettera del disco piccolo corrispondente con la chiave sia la `h`. Ora possiamo scrivere, per la seconda parola del testo cifrato, `H&PZZY`. Ruotiamo di nuovo e, se ora la lettera corrispondente alla chiave è la `r`, scriviamo `RQOY`. Qunidi, se dopo un'altra rotazione abbiamo la lettera `f` in corrispondenza con la chiave, scriviamo `FMAIIZE` cosicché, alla fine, il testo cifrato è `GVMX H&PZZY RQOY FMAIIZE...`

Oltre a quanto visto, Leon Battista Alberti aveva ideato un codice formato da 336 valori, combinando 1, 2, 3 e 4 in gruppi di 2, 3 e 4 cifre. Grazie ai quattro numeri riportati nel disco piİ grande, era possibile cifrare tale codice rendendolo ancora più sicuro anche se, per l'epoca, garantisse già di per sé una certa sicurezza. Per cifrare questii numeri si utilizzava il disco cifrante con la stessa tecnica già descritta.

Un aspetto particolarmente interessante è che, contrariamente al criptosistema di Vigenère, il criptosistema di Alberti è resistente ad un attacco portato per mezzo di un'analisi probabilistica sulla frequenza delle lettere, che l'Alberti stesso aveva precedentemente studiato. È interessante inoltre notare che il criptosistema basato sul disco cifrante di Alberti non un ebbe successo immediato, per la decisione dell'Alberti stesso di tenerlo segreto. In effetti, il suo trattato di crittografia fu pubblicato a Venezia solo un secolo più tardi.

1.2.5 Il criptosistema di Playfair

Il *criptosistema di Playfair* fu introdotto dal fisico C. Wheatstone, ma prende il nome da L. Playfair che lo divulgò nel 1854. Esso fu utilizzato dagli Inglesi nel corso della prima guerra mondiale. Nel criptosistema di Playfair il testo viene diviso in 2-blocchi, e non si fa distinzione fra le lettere `I` e `J`, cosicché l'alfabeto utilizzato consta solamente di 25 lettere. Queste lettere vengono poi disposte riga per riga in una tabella 5×5, i cui primi elementi sono occupati dalle lettere di una parola chiave, in modo tale che una lettera che vi compare più di una volta viene inserita una volta sola; le rimanenti lettere, invece, sono disposte secondo l'ordine alfabetico naturale. Ad esempio, la parola chiave `LUIGI` dà origine alla seguente tabella per cifrare:

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

Sia $(x, y) = (a_{ij}, a_{mn})$ il 2-blocco composto da a_{ij}, cioè la lettera all'intersezione fra la i-esima riga e la j-esima colonna della tabella per cifrare, ed a_{hl}, la lettera all'intersezione fra l'm-esima riga e l'm-esima colonna. Tale 2 blocco si cifra come segue:

$$\begin{aligned}
(a_{ij}, a_{mn}) &\longmapsto (a_{in}, a_{mj}) && \text{se } i \neq m \text{ e } j \neq n,\\
(a_{ij}, a_{mn}) &\longmapsto (a_{i,j+1}, a_{i,n+1}) && \text{se } i = m \text{ e } j \neq n,\\
(a_{ij}, a_{mn}) &\longmapsto (a_{i+1,j}, a_{m+1,j}) && \text{se } i \neq m \text{ e } j = n,
\end{aligned}$$

in cui gli indici sono ridotti modulo 5. Inoltre, se $x = y$, si inserisce la lettera `Q` fra x ed y e quindi si cifra il nuovo testo contenente $\ldots x\mathtt{Q}y \ldots$ Inoltre, se alla fine si rimane con un numero dispari di lettere nel testo da cifrare, si può aggiungere una qualsiasi lettera diversa dall'ultima in fondo alla sequenza per completare l'ultimo 2-blocco. Ad esempio, con la parola chiave `LUIGI` il testo in chiaro

```
CO MP RE SS O ...
```

viene dapprima trasformato in

```
CO MP RE SQ SO ...
```

e quindi cifrato come

```
FK HR SR TR TN ...
```

Conoscendo la parola chiave, il testo in chiaro si può ricavare immediatamente invertendo il procedimento sugli indici.

1.2.6 Trasposizioni

Un approccio completamente differente è rappresentato dall'uso della cifratura mediante *trasposizione*. In un siffatto criptosistema il testo viene spezzato in k-blocchi e quindi cifrato, blocco per blocco, con una permutazione di $\mathrm{Sym}(k)$. Ad esempio, ponendo $k = 7$ e scegliendo la permutazione $(1\,3\,5)(2\,4\,7) \in \mathrm{Sym}(7)$, il testo in chiaro

```
LACRITT OGRAFIA EUNASCI ENZAPOC OESATTA ...
```

viene mutato nel testo cifrato

```
ITLACTR FAOGRIA SIEUNCA PCENZOA TAOESTA ...
```

Si può anche usare una permutazione scelta in un modo particolare. Un esempio è fornito dalla cosiddetta *trasposizione per colonne,* la quale è di natura essenzialmente geometrica. In questo caso il testo viene scritto riga per riga in una matrice $k \times k$, ma poi viene letto colonna per colonna secondo un ordine dipendente da una certa parola chiave. Ad esempio, scegliendo di cifrare mediante una matrice 6×6 con parola chiave `CHIAVE` intendiamo che la prima colonna da trasmettere è la quarta perché la prima lettera in ordine alfabetico che compare in `CHIAVE` è la `A` che si trova in quarta posizione. In questo caso, il testo "la crittografia è una scienza poco esatta perché ci sono molti modi di interpretarla" viene ordinato come segue:

416235	416235	416235
LACRIT	ERCHEC	ETARLA
TOGRAF	ISONOM	...
IAEUNA	OLTIMO	
SCIENZ	DIDIVE	
APOCOE	RSIDII	
SATTAP	NTERPR	

e quindi trasmesso come

```
AOACPA  RRUECT  IANNOA  LTISAS  TFAZEP  CGEIOT
RSLIST  HNIIDR  EOMVIP  EIODRN  CMOEIR  COTDIE
  ...     ...     ...     ...     ...     ...
```

Osserviamo che la trasposizione non cambia la frequenza delle lettere in un testo. D'altra parte, contrariamente al criptosistema di Vigenère, in questo caso viene distrutta la logica con cui una certa lettera segue un'altra all'interno di una parola. Per questo motivo la trasposizione è stata utilizzata quasi sempre in associazione con qualche altro criptosistema (ad esempio, quello di Vigenère).

http://www.impan.gov.pl/Great/Rejewski/article.html

Figura 1.3: La macchina enigma

1.2.7 Macchine a rotore

Molti ricorderanno che nei film di spionaggio degli anni '60 uno dgli tipici degli agenti segreti era quello di impossessarsi di una certa "valigetta nera". Quella valigetta non era altro che la chiave di un criptosistema usato per proteggere dati strategici da cui dipendeva la sicurezza (o l'insicurezza, a seconda dei punti di vista...) del mondo civile. Più precisamente, nella valigetta c'era un rotore come quello schematizzato in figura 1.4. Esistono diverse varianti di criptosistemi basati su rotori. La più famosa era probabilmente quella basata sulla macchina enigma, inventata originariamente nel 1918 a Berlino da Arthur Scherbius ed utilizzata, dopo notevoli perfezionamenti, dai Tedeschi durante la seconda guerra mondiale.

Una macchina enigma è composta essenzialmente da una valigetta contenente una tastiera, due ruote fisse (gli statori), tre ruote mobili (i rotori) comprese fra i due statori, ed una serie di 26 lampadine corrispondenti alle lettere dell'alfabeto latino. I rotori e gli statori hanno vari punti di ingresso e di uscita della corrente, ed in corrispondenza di questi hanno dei contatti elettrici sistemati in modo tale che in ogni momento ciascuno dei contatti di una ruota sia collegato con uno ed un solo contatto di ogni ruota contigua. Inoltre le ruote sono cablate internamente in modo irregolare, cosicché una corrente che entra in una di esse ne esce "altrove" in modo pressoché casuale, vedi figura 1.4. Oltre ai rotori, per estendere quanto più possibile la scelta delle chiavi di cifatura, la valigetta contiene un circuito elettrico

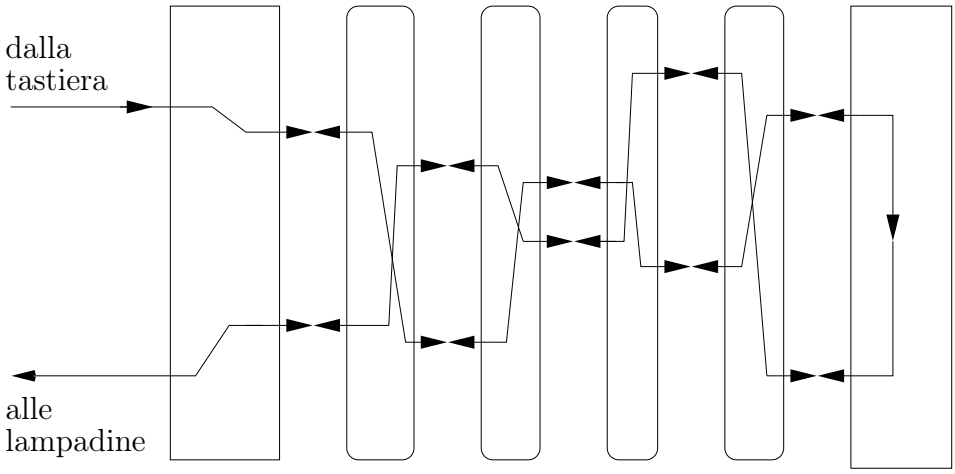

Figura 1.4: Schema di funzionamento della macchina enigma

il quale, mediante un sistema di cavetti mobili, permette di scambiare un certo numero di lettere prima di avviarle ai rotori.

Ora, premendo un tasto sulla tastiera, una corrente passa verso lo statore di sinistra. Da questo statore la corrente passa al primo rotore, al secondo, al terzo, quindi arriva allo statore di destra, detto anche ruota di riflessione. Da qui la corrente torna indietro, percorrendo di nuovo i tre rotori, verso il primo statore e quindi verso il pannello con le lampadine. L'effetto della pressione di un tasto sulla tastiera consiste nell'accensione di una lampadina corrispondente ad una lettera dell'alfabeto: questa è la lettera del testo cifrato corrispondente alla lettera del testo in chiaro.

Il movimento dei rotori all'interno della macchina è fondamentale: ad ogni pressione di tasto, il primo rotore scatta di una posizione. Questo significa che premendo di nuovo la stessa lettera, la cifratura non è più la stessa. Il primo rotore funge quindi da chiave di cifratura con lunghezza 26 caratteri, che si ripete ciclicamente per tutto il messaggio. Al completamento di un giro del primo rotore, il secondo rotore scatta di una posizione e dopo un giro completo del secondo rotore, segue uno scatto del terzo, dando così luogo ad una chiave di crittazione lunga 17576 caratteri, sufficienti per rendere la cifratura simile all'utilizzo della chiave monouso con tutti i testi non più lunghi di questo valore. In realtà, la scelta delle chiavi è molto più estesa. Il pannello elettrico, ad esempio, permette di scambiare fra loro sei, otto o anche dieci coppie di lettere e, tenuto conto del numero di possibili accoppiamenti di ventisei lettere, estende notevolmente il numero di scelte per la chiave. In pratica, una chiave di enigma viene definita per mezzo dei seguenti parametri:

- la scelta e la disposizione dei tre rotori fra i due statori;

- la posizione iniziale dei rotori;
- una certa permutazione iniziale dell'alfabeto ottenuta per mezzo del circuito con i cavetti mobili.

Considerando che i rotori, ciascuno con una propria circuiteria interna, venivano generalmente scelti in un gruppo di sei, e che il circuito a cavetti mobili poteva essere utilizzato per effettuare fino a dieci sostituzioni alfabetiche preliminari, si ottengono, per la scelta della chiave, un numero di combinazioni di ordine superiore a 10^{20},

L'elemento strutturale che forniva grande flessibilità d'uso a Enigma, ma che allo stesso tempo costituiva un elemento di debolezza, era costituito dal fatto che i due testi, quello in chiaro e quello cifrato, erano ottenibili l'uno dall'altro allo stesso modo; cioè, con due applicazioni la macchina enigma forniva l'identità. L'utilità di un siffatto sistema nell'uso pratico è evidente: per leggere il messaggio, il legittimo destinatario, che aveva predisposto una macchina uguale a quella cifrante e nelle stesse condizioni iniziali, non doveva far altro che digitare il testo cifrato per ricostruire il testo in chiaro sul visore.

Nonostante l'apparente inviolabilità, alla fine il criptosistema enigma fu rotto. Varie furono le ragioni che condussero gli Alleati a questo risultato. Innanzitutto, i Tedeschi si abbandonarono incautamente ad un certo senso di sicurezza, probabilmente utilizzando molte volte le stesse combinazioni, e quindi fornendo informazioni preziose ai criptanalisti. Ad esempio, la posizione iniziale dei rotori veniva cambiata ogni ventiquattro ore secondo una regola prefissata, cosicché questa regola finiva per essere la vera chiave; questo introduceva già di per sé un notevole elemento di debolezza. Un altro elemento di debolezza strutturale di enigma era, come già detto, insito nella sua struttura simmetrica: dal punto di vista della teoria dei gruppi, ad esempio, ciò significa che la sostituzione complessiva ottenuta è semplicemente un prodotto di scambi.

Il primi a rompere il criptosistema enigma nella sua versione più semplice furono, negli anni '30, i Polacchi grazie al lavoro di un gruppo di matematici guidato da Marian Rejewski [Rej80]. Successivamente, anche gli Inglesi riuscirono nell'intento di rompere il criptosistema enigma basato sulla macchina più avanzata allora disponibile. Tale risultato fu raggiunto, oltre che con le informazioni passate dai Polacchi, grazie al matematico Alan Touring ed all'uso dei Colossi, i precursori dei moderni computer. Molti dettagli sulle tecniche di decifrazione utilizzate all'epoca sono ancora tenuti segreti. Per aver un'idea un po' più precisa su come il criptosistema enigma fu rotto si può consultare [Kon81, cap. 5].

Una variante un po' più sofisticata di enigma è la macchina ideata dallo svedese Boris Caesar Wilhelm Hagelin alla fine degli anni '30, detta CSP-1500, ed adottata dall'U.S. Army con la denominazione M-209 durante gli anni '40.

Nella macchina di Hagelin sono presenti 6 rotori aventi, rispettivamente, 26, 25, 23, 21, 19 e 17 contatti. Una volta che una lettera è stata cifrata—a seconda della mutua posizione dei contatti dei rotori—i sei rotori si muovono di una posizione. Ne consegue che il primo rotore torna alla posizione iniziale dopo 26 lettere cifrate. Ora, dato che ciascuno dei rotori ha un numero di contatti primo con gli altri, in pratica il criptosistema basato sulla macchina di Hagelin può venire riguardato come un criptosistema di Vigenère di periodo $26 \cdot 25 \cdot 23 \cdot 21 \cdot 17 = 101405850$. Anche il criptosistema basato sulla macchina di Hagelin è stato rotto. Il lettore interessato può consultare [BP82, §23] al riguardo.

Capitolo 2

Sovracifratura

Inserire DES & i suoi fratelli.

Bibliografia

[ARS78] L. M. Adleman, R. L. Rivest, and A. Shamir, A method for obtaining digital signatures and public-key cryptosystems, *Communications of the ACM* **21** (1978), 120–126.

[BB96] L. Berardi and A. Beutelspacher, *Crittologia*, Collana Quaderni di Informatica, Angeli, 1996.

[Beu94] A. Beutelspacher, *Cryptology*, The Mathematical Association of America, 1994.

[BL96] D. Boneh and R. J. Lipton, Algorithms for black-box fields and their application to cryptography, *Advances in cryptology—CRYPTO '96 (Santa Barbara, CA)*, Lecture Notes in Comput. Sci., vol. 1109, Springer, Berlin, 1996, pp. 283–297.

[BLS+83] J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckermann, and S. S. Wagstaff Jr., *Factorization of* $b^n \pm 1$*,* $b = 2, 3, 5, 6, 7, 10, 11, 12$ *up to high powers*, AMS, Providence RI, 1983.

[BN58] J. Bretagnolle Nathan, Cubiques définies sur un corps de caractéristique quelconque, *Ann. Fac. Sci. Toulouse* **22** (1958), 175–234.

[BP82] H. Beker and F. Piper, *Cypher Systems, The Protection of Communications*, Northwood Books, London, 1982.

[BS66] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Pure and Applied Mathematics, vol. 20, Academic Press, New York, London, 1966.

[Can87] D. Cantor, Computing in the jacobian of a hyperelliptic curve, *Math. Comp.* **48** (1987), 95–101.

[Del74] P. Deligne, La conjecture de Weil, *Inst. Hautes Etudes Sci. Publ. Math.* **43** (1974), 273–307.

[DH76] W. Diffie and M. Hellman, New directions in cryptography, *IEEE Trans. Inform. Theory* **22** (1976), 644–654.

[dV86] B. de Vigenère, *Traicté des Chiffres, ou Secrètes Manières d'Escrire*, A. L'Angelier, Paris, 1586.

[ElG85] T. ElGamal, A public key cryptosystem and a signature scheme based on discrete logarithms, *IEEE Trans. Inform. Theory* (1985), 469–472.

[Ful69] W. Fulton, *Algebraic Curves*, Benjamin, New York, 1969.

[GM86] R. Gupta and M. R. P. Murty, Primitive points on elliptic curves, *Compositio Math.* **58** (1986), 13–44.

[Her82] I. N. Herstein, *Algebra*, Editori Riuniti, 1982.

[Hir80] J. W. P. Hirschfeld, Sulle varietà algebriche negli spazi proiettivi finiti, 1980, Quaderni Seminario di Geometrie Combinatorie N. 27, Dip. Mat. Univ. di Roma "La Sapienza".

[Hir83] J. W. P. Hirschfeld, The Weil conjecture in finite geometry, *Proc. of Australian Combinatorial Conference, Adelaide*, Lecture Notes in Mathematics, vol. 1036, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1983.

[Hir98] J. W. P. Hirschfeld, *Projective Geometries over Finite Fields, 2nd edition*, Oxford University Press, Oxford, 1998.

[Hus87] D. Husemöller, *Elliptic Curves*, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1987.

[HW60] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon, Oxford, 1960.

[Kob84] N. Koblitz, *An Introduction to Elliptic Curves and Modular Forms*, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1984.

[Kob87a] N. Koblitz, *A Course in Number Theory and Cryptography*, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1987.

[Kob87b] N. Koblitz, Elliptic curve cryptosystems, *Math. Comp.* **48** (1987), 203–209.

[Kob88] N. Koblitz, Primality of the number of points on an elliptic curve, *Pacific J. Math.* **131** (1988), 157–165.

[Kob89] N. Koblitz, Hyperelliptic cryptosystems, *J. of Cryptology* **1** (1989), 139–150.

[Kob98] N. Koblitz, *Algebraic Aspects of Cryptography*, Algorithms and Computation in Mathematics, vol. 3, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1998.

[Kon81] A. G. Konheim, *Cryptography, a primer*, John Wiley & Sons, New York, Chichester, Brisbane, 1981.

[Lan58] S. Lang, *Introduction to Algebraic Geometry*, Interscience, New York, 1958.

[Lan78] S. Lang, *Elliptic Curves: Diophantine Analysis*, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1978.

[Le65] S. Lang and J. T. Tate (eds.), *The Collected Papers of Emil Artin*, Addison Wesley, Reading MA, 1965.

[Len86] H. W. Lenstra Jr., Elliptic curves and number-theoretic algorithms, Report 86-19 Math. Inst. Universiteit Amsterdam, 1986.

[Len87] H. W. Lenstra Jr., Factoring integers with elliptic curves, *Ann. Math.* **126** (1987), 649–673.

[LN83] R. Lidl and H. Niederreiter, *Finite Fields*, Addison Wesley, Reading MA, 1983.

[LN86] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and Their Applications*, Cambridge University Press, Cambridge, 1986.

[LT87] S. Lang and H. Trotter, Primitive points on elliptic curves, *Bull. Amer. Math. Soc.* **83** (1987), 289–292.

[Mas83] J. L. Massey, Logarithms in finite cyclic groups—cryptographic issues, *Proc. 4th Benelux Symposium on Information Theory*, 1983, pp. 17–25.

[Mil85] V. Miller, Use of elliptic curves in cryptography, *Advances in Cryptology—Crypto '85*, Lecture Notes in Computer Science, vol. 218, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1985, pp. 417–426.

[MLB78] S. Mac Lane and G. Birkhoff, *Algebra*, Mursia, 1978.

[Mur83] M. R. P. Murty, On Artin's conjecture, *J. Number Theory* **16** (1983), 147–168.

[MV90a] A. Menezes and S. Vanstone, The implementation of elliptic curve cryptosystems, *Advances in cryptology—AUSCRYPT '90 (Sydney, 1990)*, vol. 453, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1990.

[MV90b] A. Menezes and S. Vanstone, Isomorphism classes of elliptic curves over finite fields of characteristic 2, *Utilitas Math.* **38** (1990), 135–153.

[MV90c] A. Menezes and S. Vanstone, Isomorphism classes of elliptic curves over finite fields, Research Report CORR 90–1, Department of Combinatorics and Optimization, University of Waterloo, January 1990.

[MV93] A. Menezes and S. Vanstone, Elliptic curve cryptosystems and their implementation, *J. Cryptology* **6** (1993), 209–224.

[MVZ93] A. Menezes, S. Vanstone, and R. Zuccherato, Counting points on elliptic curves over $\mathbf{f}_{2^m}$, *Math. Comp.* **60** (1993), 407–420.

[NZ80] I. Niven and H. S. Zuckerman, *An introduction to the theory of numbers, fourth edition*, John Wiley & Sons, New York, Chichester, Brisbane, 1980.

[Odl85] A. M. Odlyzko, Discrete logarithms and their cryptographic significance, *Advances in Cryptography: Proc. of Eurocrypt '84*, Lecture Notes in Computer Science, vol. 209, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1985, pp. 224–314.

[Pol74] J. M. Pollard, Theorems on factorization and primality testing, *Proc. Cambridge Phil. Soc.* **76** (1974), 521–528.

[Rej80] M. Rejewski, An application of the theory of permutations in breaking the enigma cipher, *Applicationes Mathematicae* **16** (1980), no. 4.

[Riv85] R. L. Rivest, Advances in cryptology, *Advances in Cryptography: Proc. of Eurocrypt '84*, Lecture Notes in Computer Science, vol. 209, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1985, pp. 159–165.

[Sch85] R. Schoof, Elliptic curves over finite fields and the computation of square roots mod. p, *Math. Comp.* **44** (1985), 483–494.

[Sha48] C. E. Shannon, A mathematical theory of communication, *Bell System Tech. J.* **27** (1948), 379–423, 623–656.

[Sha49] C. E. Shannon, Communication theory of secrecy systems, *Bell System Tech. J.* **28** (1949), 656–715.

[Sil86] J. Silverman, *The Arithmetic of Elliptic Curves*, Springer Verlag, Berlin, Heidelberg, New York, Tokio, 1986.

[SW79] E. Seah and H. C. Williams, Some primes of the form $(a^n - 1)/(a - 1)$, *Math Comp.* **33** (1979), 1337–1342.

[vT89] H. C. A. van Tilborg, *An Introduction to Cryptology*, Kluwer Academic Publishers, Boston, 1989.

[WW84] P. K. S. Wah and M. Z. Wang, Realization and application of the massey-omura lock, *Proc. International Zürich Seminar*, 1984, pp. 175–182.

Indice analitico