Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea in Matematica

Storia della crittoanalisi dei cifrari monoalfabetici e polialfabetici

Tesi di Laurea in storia della crittoanalisi

Relatore:
Chiar.mo Prof.
DAVIDE ALIFFI

Presentata da:
MATTEO GRILLINI

II Sessione
Anno Accademico 2016/2017

Dedicato a tutti i folli che alla parola impossibile associano il pensiero: "posso essere il primo a farlo"

Introduzione

La crittografia è la scienza che studia la comunicazione nascosta, ovvero come nascondere in bella vista ciò che si vuole celare. Un buon sistema crittografico non dipende in maniera vincolante dal metodo col quale viene trasmesso ma unicamente dalla sua formulazione. Nella storia dell'umanità la possibilità di inviare messaggi non comprensibili o di comprendere quelli ritenuti indecifrabili dall' avversario fornisce uno strumento fondamentale. Da ciò nasce la lotta fra crittoanalisti e crittografi che si è costantemente intrecciata con la storia del genere umano. La storia presenta un' alternanza fra periodi dove vengono creati codici molto resistenti e periodi in cui la sicurezza sembra inesistente.
In questo elaborato verranno prese in esame i modelli che si sono affermati nella crittografia monoalfabetica e polialfabetica e come sono stati violati.

Indice

Capitolo 1

Cifrari moalfabetici

1.1 Struttura della cifratura

L' umanità comprese fin dai tempi dell' antica Roma quanto in guerra la possibilità di fornire ordini senza temere che essi fossero trafugati dal nemico fosse fondamentale. I primi sistemi di cifratura furono *monoalfabetici*.

Definizione 1.1. Un cifrario è detto monoalfabetico se utilizza un alfabeto per il testo in chiaro e una permutazione dello stesso per il testo cifrato.

Le cifrature monoalfabetiche si differenziano per la *chiave segreta*, che è l' unica informazione necessaria per decifrare ogni messaggio. Uno dei cifrari monoalfabetici più celebri è quello di Cesare; in esso la *chiave segreta* è un numero che indica di quanti caratteri viene traslato l'alfabeto dal messaggio in chiaro a quello cifrato. Quindi prendendo in considerazione l' alfabeto italiano e una traslazione di un fattore 3, tipicamente utilizzata da Cesare, si otterrebbe il risultato seguente:

Esempio 1.1. l' alfabeto traslato di 3 unità diventa:
A B C D E F G H I L M N O P Q R S T U V Z
D E F G H I L M N O P Q R S T U V Z A B C
Ora è possibile cifrare qualsiasi messaggio tipo:
C I F R A R I O D I C E S A R E
F N I U D U N R G N F H V D U H

La semplicità della chiave, in quanto corrispondente a un numero, rappresenta uno dei punti di forza del cifrario, poichè ricordarla e comunicarla risulta semplice. La debolezza del cifrario di Cesare è intrinseca nella ristrettezza del numero delle possibili chiavi. Infatti il sistema possiede nel caso dell' alfabeto italiano solo 21 possibili *chiavi*, in quanto indicando con k la traslazione, se $k > 21$ è sufficiente applicare una semplice riduzione modulo 21. I crittografi dell' epoca, per cercare di aumentare la sicurezza pensarono di ricifrare il messaggio cifrato, ciò chiaramente non aumenta la sicurezza poichè semplicemente da un chiave k si passa a $2k$, e quindi due cifrature sono equivalenti a una sola.

Esempio 1.2. Da una prima cifratura di chiave 3 si ottiene:
A B C D E F G H I L M N O P Q R S T U V Z
D E F G H I L M N O P Q R S T U V Z A B C
Dalla seconda cifratura si ottiene:
D E F G H I L M N O P Q R S T U V Z A B C
G H I L M N O P Q R S T U V Z A B C D E F
Si può facilmente scivere una cifratura di chiave 6:
A B C D E F G H I L M N O P Q R S T U V Z
G H I L M N O P Q R S T U V Z A B C D E F
Come si può osservare le due cifrature coincidono perfettamente.
Invece considerando un fattore 24 si otterrebbe:
A B C D E F G H I L M N O P Q R S T U V Z
D E F G H I L M N O P Q R S T U V Z A B C
Che coincide anch' essa con una cifratura di fattore 3, infatti 24 *mod* 21= 3

I cifrari monoalfabetici oltre, a possedere come *chiave* un fattore di traslazione, possono fondarsi su una *parola* o una $frase\ chiave$.
Per cifrare un messaggio, data la $frase\ chiave$ si eliminano da essa gli spazi e le lettere ripetute; la sequenza ottenuta rappresenta l' inizio dell' alfabeto cifrato che verrà poi completato con il resto dell' alfabeto mantenendone l' ordine. Procedimento del tutto analogo si compie se la *chiave* è una parola,

ovviamente omettendo il passaggio dell' eliminazione degli spazi poichè non presenti.

Esempio 1.3. Data la frase chiave: "dalla vita alla morte".
Riscrivendola eliminando spazzi e ripetizioni si ottiene (dalvitmore), ponendo questa sequenza come inizio dell' alfabeto si ottiene:
A B C D E F G H I L M N O P Q R S T U V Z
D A L V I T M O R E B C F G H N P Q S U Z
Quindi cifrando il testo "dalla vita alla morte" si ottiene "VDEE URQD DEED BFNQI":
D A L L A V I T A A L L A M O R T E
V D E E D U R Q D D E E D B F N Q I

1.2 Analisi delle frequenze

I cifrari monoalfabetici possiedono la debolezza intrinseca di far corrispondere una lettera in chiaro sempre la stessa lettera nei messaggi cifrati. Ciò, e la conoscenza della lingua in cui è scritto il messaggio sono i fondamenti dell' analisi delle frequenze, che rende i cifrari presentati molto fragili. Immaginando di possedere un testo cifrato scritto in lingua italiana, si procede analizzando la frequenza con la quale compaiono le lettere.

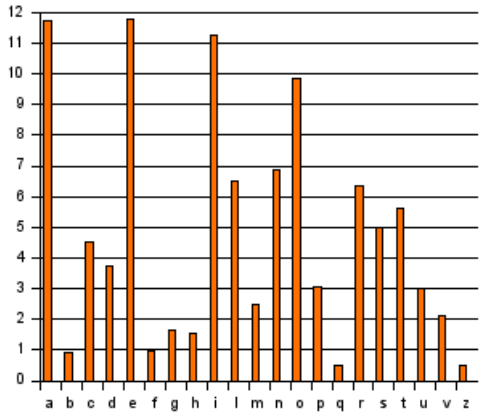

Figura 1.1: *Tabella delle frequenze nella lingua italiana*

Definizione 1.2. f = frequenza con cui compare una lettera nel testo, in forma percentuale.

Definizione 1.3. n = numero di volte che compare un carattere nel testo.

Definizione 1.4. N = numero di caratteri nel testo.

Per ottenere la frequenza sfruttiamo la seguente formula:

$$f = 100 * n/N \tag{1.1}$$

Ogni lingua possiede una propria tabella delle frequenze.

	TEDESCO	INGLESE	FRANCESE	ITALIANO	SPAGNOLO
A	5	7.81	9.42	11.74	12.69
B	2.5	1.28	1.02	0.92	1.41
C	1.5	2.93	2.64	4.5	3.93
D	5	4.11	3.38	3.73	5.58
E	18.5	13.05	15.87	11.79	13.15
F	1.5	2.88	0.95	0.95	0.46
G	4	1.39	1.04	1.64	1.12
H	4	5.85	0.77	1.54	1.24
I	8	6.77	8.41	11.28	6.25
J	-	0.23	0.85	-	0.56
K	1	0.42	-	-	-
L	3	3.60	5.34	6.51	5.94
M	2.5	2.62	3.24	2.51	2.65
N	11.25	7.28	7.45	6.88	6.95
O	3.5	8.21	5.14	9.83	9.49
P	0.5	2.15	2.85	3.05	2.43
Q	-	0.14	1.06	0.61	1.16
R	7	5.64	6.46	6.37	6.25
S	7	6.46	7.90	4.98	7.60
T	5	9.02	7.26	5.62	3.91
U	5	2.77	6.24	3.01	4.36
V	1	1	2.15	2.10	1.07
W	1.5	1.49	-	-	-
X	-	0 30	0 30	-	0 13
Y	-	1.51	0.24	-	1.06
Z	1.5	0.09	0.32	0.49	0.35

Figura 1.2: *Confronto fra le 5 principali lingue europee*

Posseduti questi dati, è possibile utilizzare l' analisi delle frequenze. Inizialmente si confronta la frequenza delle lettere cifrate con quella della lingua

del testo in chiaro, supponendo che le lettere ad alta frequenza nel testo cifrato coincidano con quelle ad alta frequenza nella lingua, come ad esempio la lettera e che ha $f = 11,79$. Da osservare che non è necessario comprendere la lingua, basta possedere una tabella di frequenza delle lettere. Oltre ad un' analisi sul singolo carattere si prendono in considerazione le combinazioni di lettere, la ripetizione consecutiva di un carattere fornisce informazioni importanti; infatti le doppie formano casi molto particolari. In italiano le doppie ad alta frequenza sono ll, tt, ss, nn, pp, cc, rr e gg che danno un' ottima indicazione sui caratteri doppi. Ulteriori ipotesi sulle coppie di caratteri derivano dalla lingua, nell' Italiano per esempio la lettera q è seguita dalla u; questo binomio si identifica con facilità perchè il carattere q ha una frequenza molto bassa e lo troviamo praticamente sempre seguito con uno stesso carattere, ovvero la cifratura di u. Ogni lingua possiede casi similari che quindi forniscono appigli ai crittoanalisti.
Successivamente ad un' analisi sui singoli caratteri e sulle combinazioni di essi, si pone l' attenzione sulle parole. Le parole particolarmente brevi sono molto poche e quindi forniscono indicazioni precise; le uniche parole formate da una lettera nella lingua italiana sono a, e, i, o e l. Inoltre bisogna sfruttare il contesto del messaggio, nei vari ambiti le tabelle di frequenza possono essere differenti.
Il metodo delle frequenze rende obsoleti i cifrari monoalfabetici.

Capitolo 2

Cifrari polialfabetici

2.1 Struttura dei cifrari

Nel XV secolo la sicurezza dei cifrari monoalfabetici era resa praticamente nulla dall'analisi delle frequenze, ma nella decade fra il 1460 e il 1470 l' architetto Leon Battista Alberti elaborò un' idea che era destinata a rivoluzionare la crittografia globale, un cifrario polialfabetico.

Definizione 2.1. Un cifrario è detto polialfabetico se alterna più alfabeti cifranti che sono permutazioni dell' alfabeto in chiaro.

L'idea consiste nell' introduzione di almeno un secondo alfabeto cifrante da alternare col primo, codificando una lettera con uno e poi con l' altro.

Esempio 2.1. Dati:
alfabeto in chiaro: A B C D E F G H I L M N O P Q R S T U V Z
alfabeto cifrato 1: E U F A V O D N P H S G T M I L B R Z C Q
alfabeto cifrato 2: C M U N B I P L O V A T G S D R H Q F Z E
Per cifrare la parola "CASSA" si codifica la "C" con l' alfabeto 1 e la prima "A" con l' alfabeto 2, a questo punto si ricifra la prima "S" con l' alfabeto cifrato 1. Così proseguendo si ottiene:
C A S S A
F C B H E
Da osservare che le due "A" sono cifrate in modo differente.

Alberti non riuscì mai a sfruttare a pieno il potenziale della sua idea. Ma nonostante ciò elaborò la prima forma di meccanizzazione della crittografia, il disco cifrato, che era composta da un alfabeto esterno e da due alfabeti più interni che fornivano la cifratura.

Figura 2.1: *Disco cifrato*

Le idee di Alberti furono la base del cifrario di Vigenère che rapprentò fino al XIX secolo un sistema crittografico inviolabile, e la rivincita dei crittografi sui crittoanalisti.
La struttura del cifrario si fonda su 26 alfabeti e una parola chiave.

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

Figura 2.2: *Tabella di Vigenère*

Le righe sono semplicemente ottenute tramite cifrari di Cesare dell' alfabeto inglese con traslazione pari al numero della riga. Per cifrare un messaggio si utilizza per ogni lettera una diversa riga e quindi un diverso alfabeto cifrante. Le righe da usare e il loro ordine sono determinate dalla parola chiave. Data una parola chiave, si riscrive questa tante volte quante sono necessarie per raggiungere la lunghezza del messaggio. Identificando ogni riga con la propria prima lettera, si cifra ogni lettera del messaggio in chiaro con la rispettiva riga della tabella di Vigenère.

Esempio 2.2. Data la parola chiave "MONTE" e il messaggio in chiaro: "spostare truppe su cima est". Si ottiene:

Chiave segreta: M O N T E M O N T E M O N T E M O N T E M O N
Testo in chiaro: S P O S T A R E T R U P P E S U C I M A E S T
Testo cifrato: E D B L X M F R M V G D C X W G Q V F E Q G G

Le lettere sottostanti la lettera M della chiave segreta vengono cifrate dalla dodicesima riga della tabella mentre quelle sottostanti la O dalla quattordicesima e così via.

Il cifrario di Vigenère si rivela incredibilmente resistente all'analisi delle frequenze, in quanto con il continuo cambiamento degli alfabeti non si ha più una corrispondenza univoca fra il carattere in chiaro e quello cifrato. Così le frequenze delle diverse lettere tendono a diventare costanti.

2.2 Breccia in Vigenère

La forza del cifrario di Vigenère consiste nel fatto che una stessa lettera può essere cifrata in n modi dove n è la lunghezza della chiave. Per cui se una parola viene riscritta $n + 1$ volte una delle possibili cifrature è ripetuta. Questo è il primo passo per la decifrazione del cifrario. Se si possiede un messaggio cifrato col sistema di Vigenère, per poterlo decifrare si compie una ricerca sulle ripetizioni di sequenze dei caratteri; trovate due sequenze identiche di caratteri ci sono due possibili modi con i quali possono essere state generate. Il caso più improbabile è che lettere differenti abbiano generato

sequenze del tutto identiche, ma questo risulta statisticamente trascurabile. Altrimenti ci si è imbattuti nella cifratura dello stesso gruppo di lettere con la chiave nella medesima posizione. La distanza fra le ripetizioni che verrà denotata con u, restringe l' insieme delle possibili lunghezze della chiave, se tale lunghezza si indicasse con l allora l deve essere un divisore di u.

Esempio 2.3. Dato il messaggio cifrato di cui un estratto è:
D F G H *U W R* L M Z W R T Q *U W R* R T T W
La distanza fra la prima U della sequenza $U\ W\ R$ e la seconda U della medesima sequenza è di 10 caratteri quindi la chiave del cifrario può avere come lunghezza 2, 5 ,10 e 1.

Per poter determinare quale divisore identifica la lunghezza corretta, si ipotizza che un qualsiasi divisore sia corretto, escludendo il caso che $l = 1$ perchè si avrebbe semplicemente un cifrario di Cesare. Si può notare che ogni l lettere il messaggio è cifrato tramite la medesima riga della tabella di Vigenère, ma questa non è altro che un cifrario di Cesare sul quale si può compiere l' analisi delle frequenze. Se il risultato è coerente con il grafico delle frequenze della lingua del messaggio l' ipotesi risulta corretta, altrimenti si ripete la medesima operazione su un' altra l. Possedendo le frequenze dei 26 cifrari di Cesare, e confrontandole con quelle ottenute nel metodo appena esposto è possibile determinare non solo la lunghezza ma anche la chiave stessa del cifrario di Vigenère.

Capitolo 3

Enigma

Nella prima guerra mondiale i sistemi crittografici di maggior successo erano quelli fondati sulla lingua. Ovvero il codice tramite il quale si crittava non era altro che una lingua sconosciuta al nemico. Il più celebre di questi sistemi era quello americano fondato sulla lingua Navajo, che ebbe un ruolo fondamentale anche nel secondo conflitto mondiale. In contrapposizione con questo metodo crittografico si sviluppò in Germania la macchina Enigma che ideata da Artur Scherbius nel 1918, rappresenta un netto balzo in avanti per i crittografi. Infatti Scherbius era convinto che fondarsi su semplici algoritmi di codifica e riponendo di fatto la resistenza del codice praticamente solo nella mancanza di conoscenza linguistica del nemico, non fosse affatto sufficiente. Da notare che il primo modello di Enigma era stato commercializzato nel 1918, diversi anni prima dello scoppio della guerra, ma tale macchina rimase a lungo non utilizzata per colpa del suo elevato costo. Dopo la prima guerra mondiale non si pensava a un possibile secondo conflitto e ai privati sembrava inutile spendere grandi cifre per un sistema di crittografia. Ma con l' avvicinarsi dello scoppio della seconda guerra mondiale il costo non apparve più sproporzionato, e gli addetti alle comunicazioni del Terzo Reich non ebbero alcun dubbio sul vantaggio di sfruttare il macchinario. Per completezza è necessario puntualizzare che gli alleati giunsero in possesso del primo schema parziale di enigma nel 1931, ma decisero che la Germania non era un pericolo

sufficiente per investire tempo e denaro in quella macchina L' unico governo a mostrare un reale interesse era quello polacco che si mise subito al lavoro.

3.1 La macchina enigma

Figura 3.1: *Enigma*

Il funzionamento di una macchina Enigma risulta apparentemente molto semplice. L' operatore riceveva le impostazioni della macchina e le inseriva; a quel punto per ogni carattere digitato sulla tastiera appartenente al testo in chiaro si illuminava sul display soprastante la corrispondente lettera cifrata. Quando il messaggio giungeva a destinazione l' addetto configurava la macchina e digitando il testo cifrato vedeva illuminarsi il rispettivo testo in

chiaro.
La struttura interna di Enigma e il suo sistema di crifratura si fonda su 3 parti meccaniche e alcune accortezze prese dai cifratori tedeschi; ciò vale per la prima versione della macchina.

1. 3 rotori
2. Prese multiple e riflettore
3. Accorgimenti dei crittoanalisti

3.1.1 I rotori

Definizione 3.1.
Il rotore è un disco suddiviso in 26 parti ognuna corrispondente a una lettera dell'alfabeto.

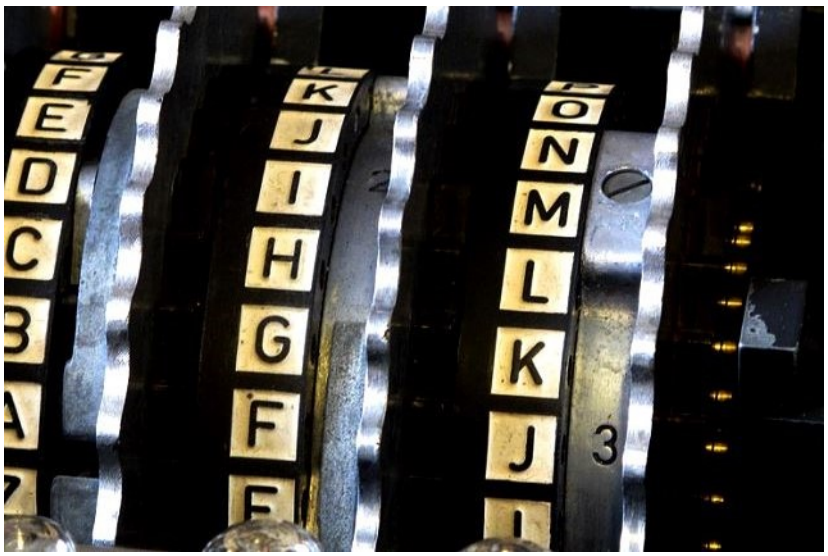

Figura 3.2: *Rotori di enigma*

Come si osserva in figura la struttura appare discretamente semplice. A una prima osservazione si potrebbe ipotizzare una semplice cifratura monoalfabetica ripetuta 3 volte, ma come è noto, ciò non aumenterebbe la sicurezza

in quanto equivarrebbe a una singola cifratura eseguita con il prodotto delle 3 permutazioni. Ma il sistema è più complesso. Infatti ogni lettera inserita causa una rotazione di 1/26 del primo rotore, e nessun cambiamento negli altri due; ogni 26 lettere il primo rotore compie una rotazione completa e di conseguenza il secondo rotore compie una rotazione di 1/26 e così via. Ciò trasforma la cifratura da monoalfabetica in polialfabetica.

Esempio 3.1. Si immaginino i tre rotori che dopo 25 inserimenti siano nelle seguenti posizioni:
$1 = (abcdefghijklmnopqrstuvwxyz)$
$2 = (bcdefghijklmonpqrstyuwxyza)$
$3 = (cdefghijklmnopsrtuvwxyzab)$
Ora si osservi cosa accade cifrando 2 volte la lettera a
tramite i tre rotori:
Il primo $a = a$
Il secondo $a = b$
Il terzo: $b = d$
Perciò la prima cifratura termina con $a = d$
Ma l' avere cifrato una lettera ha effetto sui rotori:
Il primo diventa $1 = (bcdefghijklmnopqrstuvwxyza)$
Poichè si era supposto che fosse la ventiseiesima lettera si è riottenuta la posizione iniziale e quindi viene modificato anche il secondo rotore
$2 = (cdefghijklmnopqrstuvwxyzab)$
Mentre il terzo rimane invariato. Ricifrando a si ottiene $a = f$ che risulta profondamente diverso dal risultato precedente.

Nell' esempio si è supposto che ogni rotore corrisponda semplicemente a una traslazione dell'alfabeto; in realtà ogni rotore corrisponde a una permutazione del tutto generica.
Il sistema dei rotori possiede 26^3 possibili impostazioni iniziali, inoltre i 3 rotori sono costruiti in modo tale da essere interscambiabili, ed essendo diversi fra loro incrementano le possibili impostazioni di un fattore 3! portandole a 105456.

3.1.2 Pannello a prese multiple e riflettore

Il sistema dei 3 rotori è la base della macchina Enigma, ma esso fornisce solo un numero "limitato" di impostazioni possibili. Se si cercasse una decifrazione a forza bruta controllando tutte le possibili impostazioni sarebbero necessarie a una persona circa 2 settimane di lavoro. Ipotizzando una dozzina di persone che si suddividono il lavoro, sarebbe possibile trovare una soluzione in poche ore. È chiaramente necessario aumentare il numero delle impostazioni possibili. L' aumento delle possibili impostazioni è dovuto al pannello a prese multiple.

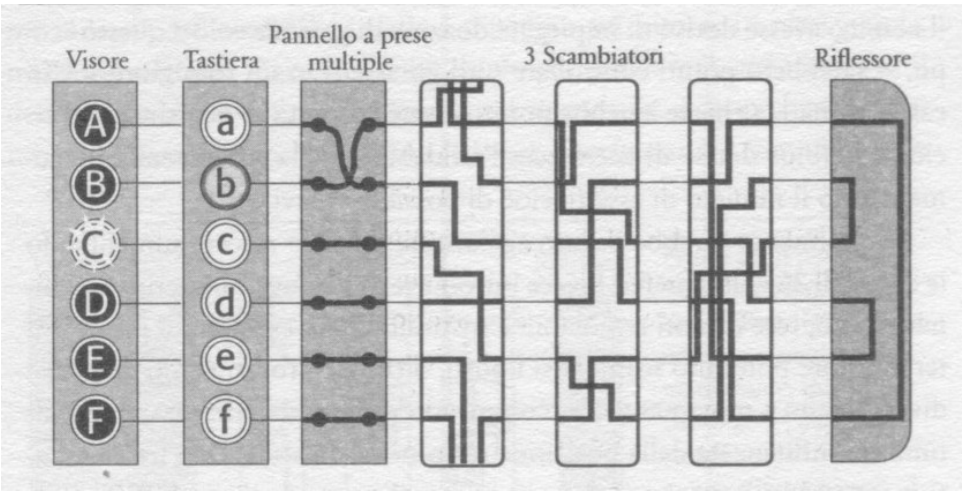

Figura 3.3: *Schema di enigma*

Come si può osservare in figura 3.3 il pannello a prese multiple è posto tra la tastiera di inserimento e il primo rotore.
Esso consiste in un pannello con 26 ingressi corrispondenti alle 26 lettere, o per maggior correttezza alle 26 parti del primo rotore. I 26 ingressi possono essere collegati fra loro tramite cavi elettrici,questo comporterà che le lettere corrispondenti agli ingressi collegati verranno scambiate. Ovvero se collego la lettera a alla lettere k la macchina cifra la lettera a come se fosse k e la k come se fosse a.

Esempio 3.2.
Sfruttando i dati dell'esempio precedente si otterrebbe che collegando a a k:
$a = l$
$k = d$

Ogni macchina Enigma in media possedeva 6 cavi di collegamento, quindi poteva collegare 12 lettere. Questi collegamenti fornivano 100 miliardi 391 milioni 791 mila 500 possibili assetti dovuti solo al pannello a prese multiple. Una domanda che nasce spontanea è: se il sistema degli scambiatori genera un numero tanto grande di impostazioni qual è l' utilità dei rotori? La risposta è che utilizzato da solo, il pannello equivale a un sistema crittografico banale, una semplice sostituzione monoalfabetica che come già visto è vulnerabile al metodo dell' analisi delle frequenze. Tale sistema inoltre è limitato a solo 12 caratteri. Combinando il pannello con il sistema dei rotori invece si giunge a circa 10 milioni di miliardi di impostazioni possibili. Il tempo per testare tutti i possibili setting da parte di un crittoanalista risulta maggiore del tempo dell' universo stesso, per cui si può affermare che il sistema non è violabile con un attacco a forza bruta.
A questo punto sono state analizzate tutte le parti crittograficamente rilevanti di Enigma; da sottolineare rimane il compito del riflettore. Benché quest'ultima parte non fornisca nessun vantaggio per la cifratura, infatti non aumenta il numero delle possibili impostazioni, risulta assolutamente fondamentale, é infatti un pannello posto dopo l' ultimo rotore che permette l' illuminazione della lettera cifrata o decifrata e permette di usare lo stesso procedimento sia per cifrare che per decifrare.
In questa sezione è stato descritto il primo modello di macchina Enigma militare che rimarrà immutato fino al 1939, quando i crittografi tedeschi applicarono alcune migliorie. Ma questo sarà trattato in seguito

3.1.3 Accorgimenti dei crittoanalisti

I funzionari tedeschi che si occupavano della sicurezza delle trasmissioni benché molto fiduciosi nella macchina Enigma,utilizzarono ulteriori accorgi-

menti per rendere i messaggi assolutamente sicuri.
Le ulteriori accortezze prese sono principalmente due: il sistema a doppia chiave e la mancanza di cifrature identiche.
Il sistema a doppia chiave consisteva in un doppio approccio di sicurezza, peraltro in uso comune nella crittografia contemporanea. Ogni giorno Enigma è configurato su una delle possibili impostazioni chiamate chiave giornaliera; per quanto la sicurezza di tale chiave sia data dal gran numero di possibili impostazioni, risulta chiaro come cifrare ogni messaggio con la stessa chiave fornisca al nemico molto materiale da analizzare, da ricordare che la macchina per ogni messaggio veniva riportata alla configurazione iniziale. La soluzione adottata dai crittografi tedeschi è l' introduzione di una nuova chiave detta chiave di messaggio. Come lascia intendere il nome tale chiave è propria di ogni messaggio, è composta da tre lettere cifrate con la chiave giornaliera. Le prime tre lettere di ogni messaggio sono quindi cifrate con la chiave giornaliera, forniscono la chiave di messaggio che sarà utilizzata per codificare il resto del testo, da notare che la chiave di messaggio differisce da quella giornaliera solo per la posizione iniziale dei 3 rotori. Così facendo si risolve il rischio di fornire troppo materiale cifrato nel medesimo modo al nemico, ma ci si espone al rischio di errori umani. Infatti un semplice errore di battitura dei primi tre caratteri renderebbe il messaggio illeggibile. Per limitare questa eventualità, la chiave di messaggio è scritta due volte consecutive così da poter immediatamente individuare l'errore umano.
L' altro accorgimento preso riguarda le impostazioni delle singole lettere, i tedeschi pensavano che cifrare una lettera con se stessa, semplificasse il compito del decrittatore; per cui si scartano a priori tutte le impostazioni che cifrano una lettera con se stessa. Quest' ultima analisi dei crittoanalisti tedeschi è incredibilmente ingenua poichè così facendo si escludono un numero enorme di possibili assetti della macchina, semplificando in realtà notevolmente il compito di chi la voglia violare.
Nonostante questa ingenuità è da sottolineare come Enigma rimanga comunque inattaccabile con un metodo a forza bruta.

Riassumendo: ogni giorno veniva usata la stessa chiave per cifrare i primi 6 caratteri di ogni messaggio, quelle 6 lettere rappresentavano la chiave di messaggio ripetuta 2 volte; tale chiave forniva la posizione dei rotori. Questa chiave viene usata per cifrare il singolo messaggio.

3.2 Decifrazione

La decifrazione di Enigma avvenne in due momenti differenti poichè la macchina venne prima forzata dai crittoanalisti polacchi e successivamente dopo che fu modificata, si rivelò necessaria l' intervento del controspionaggio inglese.

3.2.1 Scuola polacca

La Polonia iniziò il proprio lavoro sulla macchina Enigma nel 1931, con largo anticipo rispetto agli altri stati europei. Alla guida del gruppo dei crittoanalisti polacchi c' era Rejewski, i loro risultati furono straordinari in funzione delle risorse a loro disposizione.
Come già detto il sistema a doppia chiave innalza la sicurezza del messaggio cifrato ma in esso è intrinseca una debolezza di base, la riscrittura della chiave di messaggio. Nonostante la riscrittura limiti notevolmente errori umani dovuti alla digitazione, crea anche una ripetizione, la quale è destinata a diventare uno dei punti deboli della cifratura.
Per la decodifica della chiave di messaggio è necessario concentrarsi sulla stessa, quindi solo sulle prime 6 lettere di ogni messaggio, tralasciando il resto.
Ricordiamo che queste 6 lettere, che hanno la struttura $xyzxyz$ sono cifrate sempre con la stessa chiave giornaliera. Indichiamo con $ABCDEF$ i 6 caratteri iniziali di un messaggio.
Partendo da un qualsiasi messaggio è possibile, possedendo sufficienti messaggi, creare un algoritmo che fornisca sotto forma di ciclo la codifica della prima lettera come composizione della permutazione che porta x in A e x in D. Per ottenere tale ciclo si prende un messaggio di partenza dato, succes-

sivamente si ricerca un messaggio con A uguale alla D del primo messaggio, è necessario iterare questo procedimento fino a quando non si ha un' uguaglianza fra la D del messaggio attuale e la A di quello di partenza. Ripetendo il procedimento con messaggi di partenza le cui A non appartengono ai cicli già determinati, otteniamo la permutazione $A - D$ come prodotto di cicli disgiunti. È chiaro che il processo può essere ripetuto per ottenere anche la permutazione $B - E$ e $C - F$ come prodotto di cicli disgiunti.
Per fare chiarezza viene presentato l' esempio seguente:

Esempio 3.3.

Si prendano in considerazione le seguenti chiavi di messaggio:
1) dmx vbj 2)vbj pmk 3)puk fny 4)fny ifw 5) ifw xgx
6) xga gqc 7) gqc zdb 8) zdb yzd 9) yzd ojf 10)ojf dua
È evidente dal primo messaggio che d vada in v e dal secondo che v venga mandato in p, ora si possiede un ciclo parziale dato da $(dvp...)$. Per completare il ciclo si prosegue in modo del tutto analogo fino al messaggio 10, grazie al quale si chiude il ciclo ottenendo: $(dvpfixgzyo)$.
Iterando l' algoritmo su sufficienti messaggi, qui non riportati, si ottiene un prodotto di cicli disgiunti che forniscono la permutazione $A - D$.
$A - D = (dvpfixgzyo)(eijmunglht)(bc)(rw)(a)(s)$
Proseguendo in maniera del tutto analoga si ricavano:
$B - E = (unfgqdzj)(axywv)(elot)(cki)(mb)(r)(s)$
$C - F = (giloqztmersuvnp)(xjkyw)(acbdf)$

Nello studio matematico sulle permutazioni si sono ottenuti molteplici risultati, che risultano fondamentali per la crittoanalisi. Ne sono un esempio i seguenti due teoremi, fondamentali per decifrare Enigma

Teorema 3.2.1 (Primo teorema sulle permutazioni)**.** *Date due permutazioni X e Y, se una trasposizione appartiene a entrambe allora appartiene a due cicli diversi nel prodotto XY.*

Teorema 3.2.2 (Secondo teorema sulle permutazioni)**.** *Se in una permutazione compare un numero pari di cicli diversi ma della stessa lunghezza allora quella permutazione puó essere considerata come il prodotto di due permutazioni X e Y ottenute da trasposizioni disgiunte.*

Sono trattati in modo approfondito nell'articolo di Rejewski "How polish Mathematicians Broke the Enigma Cipher". Sfruttandoli si giunge a una conclusione assai interessanti:

Osservazione 1. Se due lettere appartengono a due diversi cicli della stessa lunghezza di XY allora appartengono alla stessa permutazione alla quale appartengono anche i loro vicini; ove con vicini si intendono la lettera antecedente e precedente a quella presa in esame.

Queste proposizioni forniscono la base matematica per ottenere informazioni fondamentali sulla chiave giornaliera. Infatti avere diverse informazioni sulla chiave di messaggio fornisce informazioni sulla chiave giornaliera, perchè la chiave di messaggio è cifrata proprio con quella giornaliera. Rimane da comprendere come queste poche informazioni possano fornire la permutazione dovuta sia ai rotori che al pannello a prese multiple.

La genialità di Rejewski ebbe il suo culmine nel comprendere come si potessero separare gli effetti dei rotori e del pannello a prese multiple. Sfruttando i teoremi e l' osservazione sopracitati si può evincere che la lunghezza del ciclo di ogni carattere, nella permutazione dovuta alla chiave giornaliera, risulta essere costante pe i rotori quindi si otteneva una base per l' analisi di questi.
Considerando solo i rotori si è già visto che si hanno 17.576x6=105.456 possibili assetti. Provando tutti questi possibili assetti su una copia trafugata di Enigma e annotando la lunghezza dei cicli si giunge a possedere un elenco delle possibili cifrature, differenziate per la lunghezza dei cicli. Perciò ogni giorno Rejewski compiva lo studio della lunghezza dei cicli della chiave giornaliera sfruttando quella di messaggio, e successivamente confrontava i dati con l' elenco delle possibili 105.456 impostazioni e risaliva a quella del giorno. Nonostante l'incredibile successo la scuola crittoanalitica polacca era lontana

da decifrare Enigma, infatti il panello a prese multiple rappresentava ancora un ostacolo apparentemente insormontabile. Il processo per la sua decodifica è un esempio perfetto della complessità della crittoanalisi, per cui esistono metodi ottimali ma l' intelletto del crittoanalista mantiene un ruolo centrale. Infatti il processo per la decodifica del pannello a prese multiple avviene tramite la rimozione dei cavi di collegamento del pannello a prese multiple da una macchina Enigma; il successivo tentativo tramite la giusta posizione dei rotori della decodifica del messaggio. Chiaramente il messaggio per larghi tratti risulta incomprensibile ma alcuni frammenti, quelli con molte lettere che non vengono scambiate, risulteranno quasi comprensibili.

Esempio 3.4. Si immagini di aver decodificato la posizione dei rotori e sia ottenuto un messaggio di cui un frammento è:
$ALLACCAREATTATBA$
Si può suppore che il messaggio sia in realtà "attaccare all' alba". Quindi è avvenuto uno scambio fra la lettera L e la lettera T

Sfruttando situazioni analoghe a quelle dell' esempio si è in grado di determinare i 6 scambi dovuti al pannello a prese multiple. Il controspionaggio polacco poteva dedurre le impostazioni di Enigma in poche ore.
Le incredibili capacità di Rejewsk e del suo team si scontrarono nel 1939 con due modifiche sostanziali della macchina. Furono introdotti 2 nuovi rotori da poter inserire nei 3 box predisposti, ciò aumentava notevolmente il numero di chiavi, in quanto dalle 6 possibili configurazioni dei rotori si passava a 60 e di conseguenza le possibili permutazioni dovute ai rotori a 1.054.560. Inoltre furono introdotti altri 6 collegamenti nel pannello a prese multiple fornendo ai tedeschi 159 miliardi di miliardi di possibili chiavi. La complessità computazionale era divenuta troppo grande per i mezzi polacchi.

3.2.2 Bletchley Park

Allo scoppio del conflitto il governo inglese decise di spostare la *Stanza* 40, ovvero il centro della crittoanalisi governativa, nel più ampio sito di Blet-

chley Park che poteva ospitare nettamente più personale.
I crittoanalisti di Bletchley Park assimilarono il metodo polacco e cercarono di elaborare strategie per ridurre i tempi di decifrazione, che dopo le modifiche tedesche si erano dilatati. Il personale e i mezzi maggiori rispetto ai crittoanalisti polacchi permisero agli inglesi di possedere sufficiente materiale per osservare che una debolezza risiedeva negli operatori. Infatti nonostante le indicazioni fossero quelle di inserire chiavi casuali per ogni messaggio, molti addetti inserivano terne di lettere consecutive sulla tastiera come ASD o ZXC; queste chiavi banali venivano chiamate *cillies*. Questi non erano punti deboli del sistema crittografico in sè ma di come gli addetti lo utilizzavano.
Inoltre i responsabili della scelta della chiave giornaliera decisero di non riposizionare per due giorni consecutivi un rotore nella medesima posizione. Ovvero nominando i 5 rotori con N, M, P, Q e L se un giorno l' assetto è $N - M - L$ il giorno seguente sicuramente N non può essere in prima posizione, M non può essere in seconda e L non può essere in terza; quindi si passa da 60 possibili assetti a 31.
Il pannello a prese multiple viene configurato con un ragionamento simile. Infatti sono vietati gli scambi fra lettere consecutive; ciò non fornisce nessun vantaggio alla resistenza del cifrario, ma al contrario esclude molteplici possibili assetti del pannello ai decifratori.
Sfruttando i cillies e questi accorgimenti si riusciva a decifrare alcuni messaggi di Enigma, però la comunicazione usando la macchina nonostante ciò rimaneva sufficientemente sicura.
Come nel caso polacco si rivelò necessario affidarsi a una mente più acuta delle altre. Ad Alan Turing e agli analisti di Bletchley Park furono forniti fondi illimitati. Turing osservò che i messaggi trasmessi dall' esercito tedesco avevano una struttura prevedibile. Ad esempio intorno alle 6:00 di ogni mattina veniva trasmesso un bollettino meteo con una struttura molto rigida, per cui era logico supporre che la parola *wetter* (tempo atmosferico) comparisse in una determinata posizione; l' esperienza accumulata a Bletchley Park for-

niva diversi indizi di questo tipo; queste supposizioni furono chiamate *cribs*. Questo rivoluzionò il mondo della crittografia, per la prima volta un attacco a un sistema crittografico viene condotto non solo fondandosi sul testo cifrato perchè non più sufficiente ma fondandosi sul testo cifrato e testi precedenti di cui si possiede testo cifrato e testo in chiaro. Nella crittoanalisi moderna questo tipo di attacco viene chiamato "known plaintext".
Turing pensò che i cribs rappresentassero la chiave per la risoluzione di Enigma e la supposizione si rivelò esatta. Il procedimento di Turing si rivelò assolutamente geniale: individuato un potenziale crib si chiama e una configurazione di Enigma che fornisce la cifratura della prima lettera in chiaro a con A. Si ricerca nel crib la lettera in chiaro coincidente con A e questa dista k lettere da a quindi ha assetto $e + k$. Si procede in modo analogo fino trovare una lettera cifrata coincidente con a e si indica con n il numero di lettere necessarie a tornare ad a così da formare un ciclo. Collegando n macchine Enigma in modo che esse abbiano collegamenti tra lettere cifrate e le lettere del crib che si sono supposte correlate, si esclude il pannello a prese multiple in modo non banale. Nella pratica Turing collegò 3 macchine Enigma, trovato un potenziale crib del tipo descritto e di lunghezza 3, osservò che a subisce una modifica dovuta al pannello a prese multiple che sarà indicata con L_1 prima dei rotori. Sucessivamente subirà una codifica sempre dal pannello prima dell' illuminazione sul display che si denoterà con L_2, questa sarà la stessa subita dalla seconda lettera in chiaro. Quindi la seconda lettera subirà anch' essa L_2 in ingresso e una nuova codifica dovuta al pannello indicata con L_3 prima di essere visualizzata. La terza e ultima lettera del ciclo in entrata subisce L_3 ma poichè deve risultare uguale ad a in uscita deve subire L_1; per la precisione tutte le L_{n} in uscita sarebbero le inverse di quelle in entrata ma essendo trasposizioni coincidono. Così facendo si ottiene che la cifratura dipende solo dai rotori ed è sufficiente controllare le 17.576 possibili impostazioni compatibili.

Esempio 3.5. Si prenda il messaggio cifrato: ETJPXW, supponendo che per la struttura del messaggio esso sia uguale a *wetter* e si indichi con e una

configurazione di Enigma che codifichi w con E.
Si ottiene:

Assetto ipotizzato:	e	e+1	e+2	e+3	e+4	e+5
Testo in chiaro:	w	e	t	t	e	r
Testo cifrato:	E	R	J	P	X	W

La configurazione e manda w in E, mentre e+1 manda e in R e per concludere $e+5$ manda r in W. Concatenando 3 macchine Enigma con impostazione e, $e+1$, $e+5$ e facendole cifrare contemporaneamente è necessario controllare quale delle 17.576 impostazioni che in maniera compatibili con le ipotesi fornisse come risultato il crib utilizzato.

La concatenazione di 3 macchine Enigma che ruotavano ogni secondo la propria posizione nel peggiore dei casi avrebbe impiegato 5 ore per trovare l' impostazione corretta. Quindi la macchina detta *bomba* creata da Turing testava tutte le possibili impostazioni facendo ruotare diversi rotori che si fermavano nel caso di un assetto corretto.

Figura 3.4: *Immagine della bomba*

Superata la complessità di creare una macchina che controllasse in un secondo un' impostazione, si sarebbe comunque dovuto ricavare la configurazione del pannello a prese multiple e c' era sempre il rischio che la macchina

lavorasse su una combinazione errata delle 60 possibili dei rotori. Non fu trovato un modo per meccanizzare la risoluzione del pannello a prese multiple, la cui decodifica rimase completamente nelle mani dei crittoanalisti. Quindi per la decodifica di Enigma nonostante i tanti passi tecnologici fatti rimase per tutto il conflitto necessaria la genialità dei crittoanalisti di Bletchley Park.

Capitolo 4

Conclusioni

Nell' elaborato si è vista l' evoluzione delle tecniche crittografiche e in particolare quella della crittoanalisi. Partendo dai cifrari monoalfabetici si è osservato come semplici conoscenze linguistiche e intuitive siano sufficienti per decodificare un messaggio cifrato sfruttando l' analisi delle frequenze. Poi si è affrontata la struttura polialfabetica e in particola quella di Vigenère, vedendo come per secoli questa sia stata ritenuta indecifrabile, poichè è necessaria una buona capacità di astrazione per poter riconoscere la presenza di cifrari monoalfabetici nella struttura del testo, riconducendosi al caso monoalfabetico per la decodifica. Inoltre con i cifrari polialfabetici nascono le prime forme di meccanizzazione della crittografia, di cui si è osservata in particolare Enigma. La macchina tedesca è il massimo risultato ottenuto prima dell' introduzione dei computer, ed è anche il primo esempio dell' impossibilità della decifrazione conoscendo solo testi cifrati. Inoltre con Enigma nasce l' idea della necessità di una macchina crittoanalitica per decodificare una macchina crittografica, il solo ingegno dei crittoanalisti non è più sufficiente anche se ancora necessario. Il risultato straordinario di Alan Turing si accompagna ad alcune curiosità storiche. Come detto nel capitolo precedente, i crittoananlisti inglesi si fondarono sul lavoro polacco, che come visto aveva il suo maggior esponente in Rejewski; la curiosità è che lo stesso Rejewski era in Inghilterra mentre Turing compiva la sua opera di decifrazione, ma non

venne in alcun modo coinvolto. Il crittografo polacco venne a sapere dell' utilità delle proprie ricerche solo nel 1974, poichè la violazione di Enigma rimase per diversi decenni un segreto di stato. Alan Turing morirà suicida nel 1954 senza vedersi riconosciuto pubblicamente il suo enorme contributo alla vittoria alleata.

Bibliografia

1 [1] Simon Singh, *Codici e segreti*. Rizzoli libri (1996)

2 [2] Marian Rejewski, *How polish mathematicians deciphered Enigma*. Annals of the History of Computing, volume 3, number 3 (1981)

3 [3] David Kahn, *The codebreakers*. Scribner (1996)