Il cifrario di Vigenère

Bizzoni Stefano
De Persiis Angela
Freddi Giordana

Cifrari monoalfabetico e polialfabetico

- mono: cifrari a *sostituzione* o a *trasposizione*, associano ad ogni lettera dell'alfabeto sempre lo stesso carattere cifrato;
- poli: associano ad ogni lettera dell'alfabeto caratteri diversi; cade la decrittazione basata sulla frequenza delle lettere.

Cifrario di Vigenère (1586)

- violato nel 1863 da Kasiski;
- generalizzazione del cifrario di Cesare;
- utilizza 26 alfabeti cifranti per cifrare un solo messaggio;
- INPUT:
  - testo in chiaro/cifrato
  - chiave (infinite soluzioni)
  - tavola di Vigenère

Tavola di Vigenère

Testo

chiave

	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

alfabeto in chiaro

Osservazioni

- testo senza ridondanza;
- chiave lunga e casuale → diminuisce la corrispondenza tra lettere in chiaro e cifrate;
- cambiare spesso la chiave;
- condizioni difficili da rispettare.

Algoritmo di cifratura

- considera la prima lettera del testo in chiaro e della chiave;
- le usa come coordinate cartesiane nella tavola;
- l'intersezione fornisce il carattere da sostituire nel testo cifrato;
- itera per tutta la lunghezza del testo.

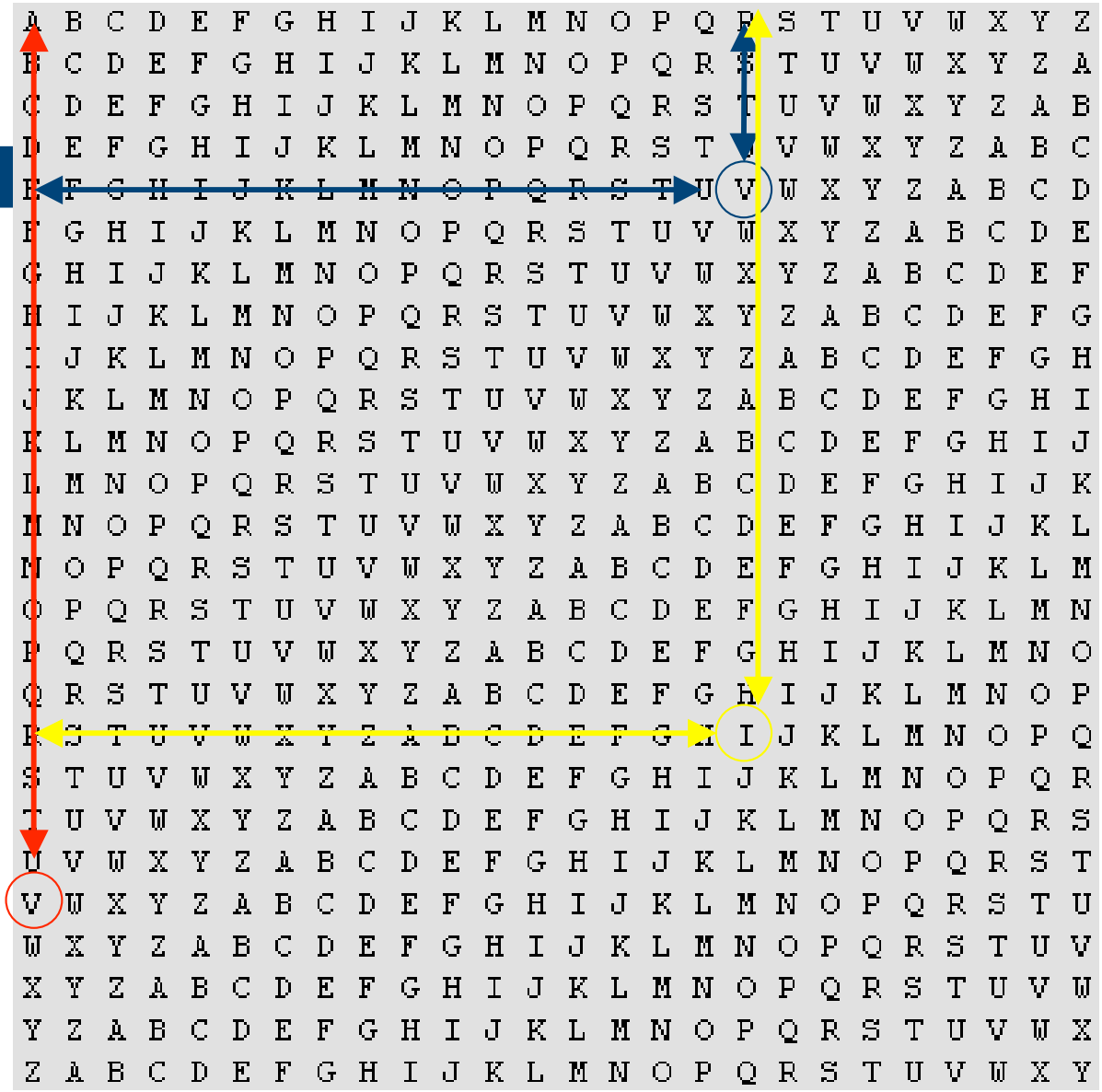
Cifratura

Testo in chiaro: ARRIVANO
Chiave: VERME

A R R I V A N O

V E R M E V E R

V V I U Z V R F



Algoritmo di decifrazione

- considera la prima lettera della chiave;
- nella riga corrispondente alla lettera della chiave individua la lettera del testo cifrato;
- il carattere che contrassegna tale colonna è la lettera in chiaro;
- itera per tutta la lunghezza del testo.

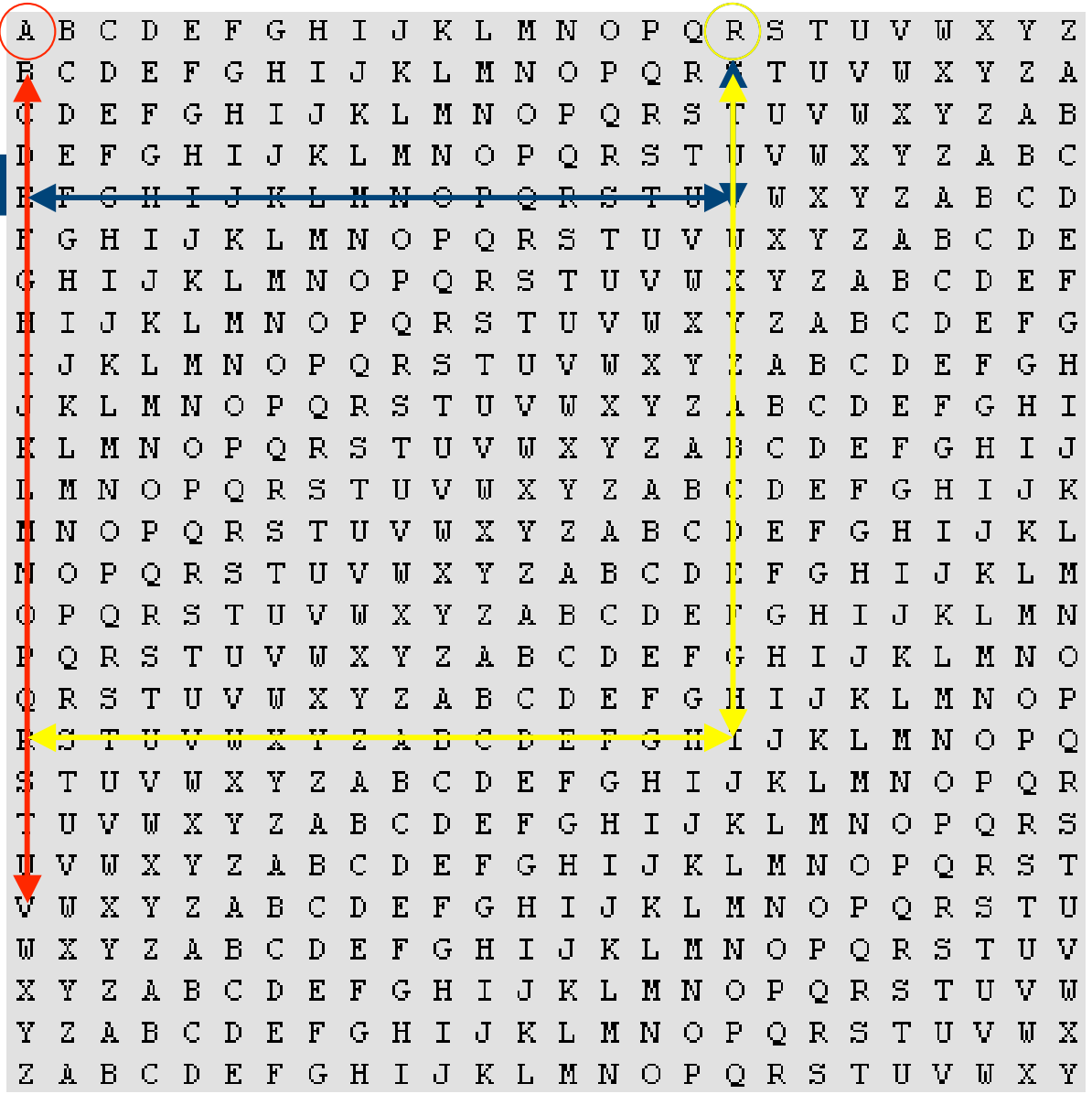
Decifrazione

Testo cifrato: VVIUZVRF
Chiave: VERME

V V I U Z V R F

V E R M E V E R

A R R I V A N O



Modello matematico

- altra rappresentazione per cifrare il testo;
- corrispondenza tra le lettere dell'alfabeto e i numeri naturali
  - a $\rightarrow$ 0, b $\rightarrow$ 1, c $\rightarrow$ 2,…, z $\rightarrow$ 25
- si somma al numero del simbolo del messaggio in chiaro il numero d'ordine del simbolo della chiave;
- si applica l'operatore *mod26* per ottenere il carattere cifrato;
- per decifrare si procede allo stesso modo ma sottraendo gli indici.

Cifratura

- Testo in chiaro: ARRIVANO
- Chiave: VERME

a=0, v=21 a+v=0+21=21=V
r=17, e=4 r+e=17+4=21=V
r=17, r=17 r+r=17+17=34_8 mod26=I
….
Si ottiene, come prima, VVIUZVRF

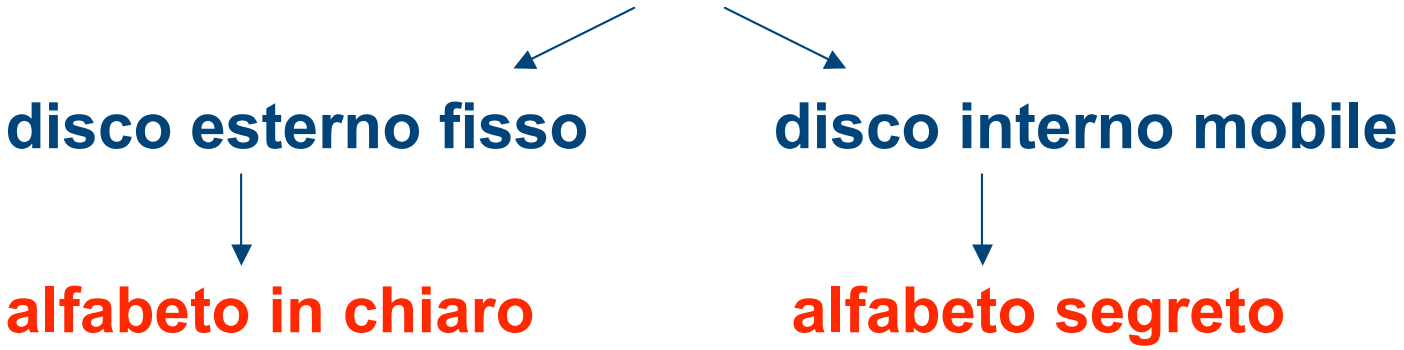
Decifrazione

- Testo cifrato: VVIUZVRF
- Chiave: VERME

v=21, v=21 v-v=21-21=0=A
v=21, e=4 r-e=21-4=17=R
i=8, r=17 i-r=8-17=-9_17 mod26=R
….
Si ottiene, come prima, ARRIVANO

Cifrario di Alberti (1466)

- primo cifrario polialfabetico
- dati di Input:
  - **Testo in chiaro**
  - ***Disco di Alberti***: coppia di cerchi concentrici



Algoritmo di Cifratura . . .

- **fisso una lettera dell'alfabeto cifrante:** *Indice del Cifrario*
- **ruoto il disco interno per portare l'indice del cifrario in corrispondenza di una lettera dell'alfabeto in chiaro scelta**
- **ogni lettera del testo in chiaro viene cifrata con la corrispondente lettera del disco interno (corrispondenza biunivoca)**

…Algoritmo di Cifratura

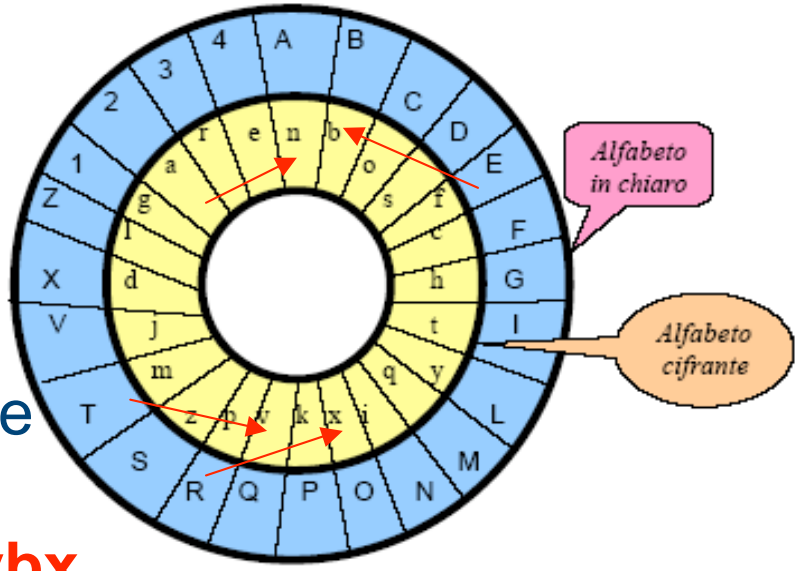
- **per cambiare alfabeto**:
  - **scelgo uno dei quattro numeri del disco esterno**
  - **inserisco la lettera del disco interno in corrispondenza del numero**
  - **ruoto il disco finché l'indice del cifrario sia in corrispondenza del numero scelto**
  - **l'alfabeto ottenuto con questa rotazione viene impiegato fino ad un ulteriore cambio**

Cifrare…..

Messaggio reale: Teramo è una città

Messaggio inviato: TERAMOEUNACITTA

Indice del cifrario: *a*



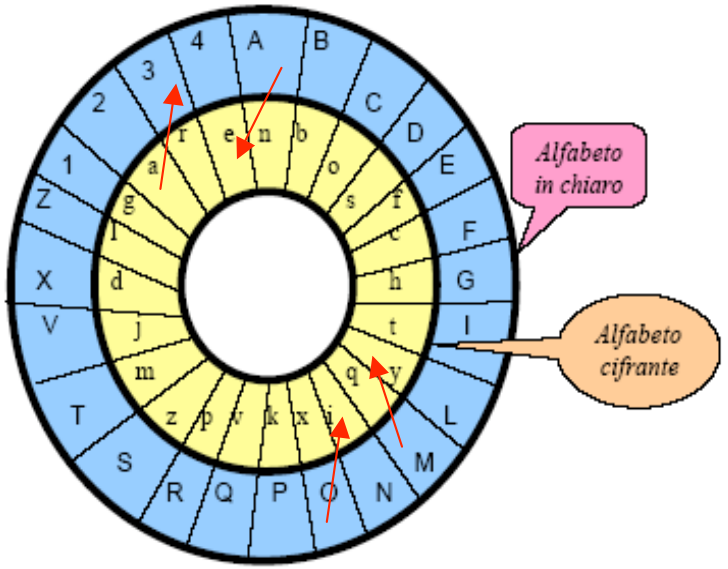
Passo 1:
- Pongo la a sotto la A
- cifro secondo le corrispondenze

T→v , E→b, R→ x ⇒ TER→ vbx

…..Cifrare

- **Passo 2:**
  - Scelgo di inserire il 3:**3** → **l**
  - Porto la a sotto il 3
  - cifro con il nuovo alfabeto

A → e
M → y ⇨ **AMO → eyi**
O → i



TERAMOEUNACITTA ⇨ TER**3**AMOEUNACITTA

…Cifrare

- **Passo 3:**

  - scelgo di inserire **1:** 1→ I

  - porto la a sotto 1

  E → c, U=V → d, N → x, A → b ⇨ EUNA → cdxb

Messaggio in chiaro: T E R A M O E U N A

Messaggio cifrato: vbx I eyi I cdxb

Confronto Alberti e Vigenère . . .

Vantaggi del cifrario di Vigenère:

- modifica dell'alfabeto ad ogni lettera (con Alberti il cambio avviene con l'inserimento saltuario dei numeri)
- cambio degli alfabeti in modo *"coperto"*: la chiave è un oggetto di cui non si ha traccia nel messaggio

…Confronto Alberti e Vigenère

- Vantaggi del cifrario di Alberti:

  - alfabeto sul disco più piccolo (alfabeto cifrante) disposto con un ordine casuale

  - si possono avere 24!-1 modi di combinare il disco interno

  - 1/24!: probabilità di successo se si attacca il cifrario di Alberti

Ideale…

- Avere una chiave che cambia ad ogni messaggio
- Utilizzare un ordinamento casuale come quello di Alberti

Problema del Cifrario di Vigenère…

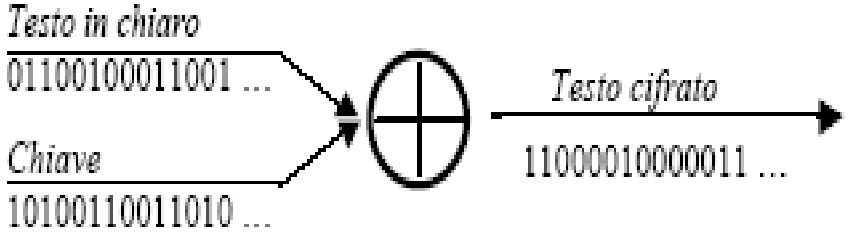
- è un insieme di cifrari di Cesare intercalati a distanza fissa
- la crittoanalisi è tanto più facile quanto più la parola chiave è breve
- diversa è la situazione se la chiave ha lunghezza uguale al testo in chiaro

…soluzione: Cifrario di Vernam (1917)

- L'One-Time-Pad prevede:
  - **Chiave: sequenza di bit del tutto casuale**
  - **Algoritmo: XOR** di ogni bit del messaggio in chiaro e bit della chiave

Messaggio-in-Chiaro XOR Chiave → Mess-Cifrato

Messaggio-Cifrato XOR Chiave → Mess-in-Chiaro



Esempio

Messaggio in-Chiaro: 1 0 1 1 0 1 0 0 0 1

Chiave: 0 1 1 0 1 0 0 0 1 1

XOR

Messaggio cifrato: 1 1 0 1 1 1 0 0 1 0

- È immediato verificare che l'operatore XOR serve sia per cifrare che per decifrare

Cifrario di Vernam: cifrario perfetto…

- **Shannon ha dimostrato che ogni cifrario "teoricamente sicuro"è un cifrario di Vernam se:**
  - la chiave segreta è una stringa casuale di bit lunga come il testo (distrugge proprietà statistiche messaggio)
  - la chiave non deve essere riutilizzata

- **Allora il testo cifrato non contiene alcuna informazione sul testo in chiaro**

…Cifrario perfetto

- Provare a decifrare senza la chiave:
  - **provare tutte le chiavi e ottenere tutti i possibili messaggi di quella lunghezza**
  - **numero delle chiavi è pari a 2^N con N numero di bit del messaggio**

Utilizzo Pratico

- **Non semplice perché:**
  - bisogna avere chiavi casuali e lunghe
  - difficoltà di trasmissione:chiave e testo cifrato possono essere letti con stessa probabilità se strasmessi sullo stesso canale
  - soluzione: scambiarsi in anticipo le chiavi
- **Questa tecnica è tuttora usata dai servizi segreti per comunicazioni di estrema importanza**

Attacchi al cifrario di Vigenère

Attacchi: Kasiski - Friedman

- Obiettivo:determinare la lunghezza della parola chiave
- Decrittare i messaggi implica considerazioni statistiche sulle caratteristiche di ciascuna lingua.
- Il lavoro di decrittazione consiste in successive induzioni e deduzioni in merito al presumibile significato

Kasiski (1863)

Principio base dell' attacco di Kasiski

- Porzioni ripetute di messaggio cifrate con la stessa porzione di chiave risultano segmenti di testo cifrato identici

P R O V A D I C I F R A T U R A

H T M L H T M L H T M L H T M L

W K A G H W U N P Y D L A N D L

• Stessa lettera viene cifrata in modo diverso nelle sue varie occorrenze (lettera R)

• Se due lettere sono poste ad una distanza pari alla lunghezza della chiave vengano cifrate nello stesso modo (lettera A)

Procedimento

- Dobbiamo spostare la nostra attenzione non su lettere ma su sequenze di lettere uguali.
- Si individuano tutte le sequenze ripetute nel testo cifrato
- Il numero di lettere comprese tra gli intervalli dei poligrammi è multiplo del numero di lettere della chiave
- Il massimo comune divisore tra le distanze tra sequenze identiche è la lunghezza della chiave.

Sequenze di lettere formate da stesse lettere

EVFTS IXRES CEKHF
XVVGH IBSDZ SICYO
XVVOP LIRES IMTNL
SICINF XVVIS UDIMT
KXVVM MFIIG VMNFL

SEQUENZE	DISTANZE	SCOMPOSIZIONE
RES	25	5*5
SIC	15	5*3
XVV	10,15	5*2,5*3

MCD =5

Riduzione al codice di Cesare

- Determinazione della lunghezza della parola chiave equivale alla determinazione del numero degli alfabeti
- Il messaggio si riduce a messaggi intercalati, tutti cifrati con un codice di Cesare ed è allora molto facile completarne la decifratura.
- Le risultanti porzioni monoalfabetiche possono essere risolte individualmente

Considerazioni

- Segmenti ripetuti di testo cifrato di lunghezza 4 o maggiore sono più utili, poiché le ripetizioni accidentali sono meno probabili
- Il calcolo del MCD deve essere fatto considerando solo le sequenze sospette
- E’ possibile che la lunghezza non sia esattamente il MCD ma un suo multiplo

Friedman (1925)

Friedman

- Nel 1925 Friedman trovò un nuovo metodo più efficiente per risalire alla lunghezza della chiave L
- Si basa sul numero delle coppie uguali: probabilità che prese due lettere nel testo cifrato queste siano uguali
- Si calcola L in funzione di tale probabilità denominata indice di coincidenza

Numero di coppie uguali nel testo cifrato (Z)

- Consideriamo una sequenza di n lettere:n_1 è il numero di occorrenze della lettera “A” ….n_{26} numero di occorrenze della lettera “Z”.
- Calcoliamo il numero delle coppie formate dalla stessa lettera e indichiamolo con Z

Numero di coppie formate dalla lettera A

$$\frac{n_1(n_1-1)}{2}$$

Possibilità che la prima lettera sia A

Possibilità che la seconda lettera sia A

Z

$$\sum_{i=1}^{26} \frac{n_i(n_i-1)}{2}$$

Indice di coincidenza(I)

- Abbiamo calcolato il numero dei casi favorevoli
- Il numero di coppie possibili è $\frac{n\ (n-1)}{2}$

I = probabilità di prendere una coppia uguale in un messaggio cifrato

$$\frac{\sum_{i=1}^{26} \frac{n_i (n_i - 1)}{2}}{\frac{n\ (n-1)}{2}}$$

Calcolo delle probabilità

- Definiamo P come la probabilità che presa una coppia di simboli in un testo in chiaro questi siano uguali

$$\sum_{i=1}^{26} p_i^2 = 0.075$$

Probabilità relativa alla lettera iesima nella lingua italiana

- Definiamo Q come la probabilità che preso una coppia di simboli in un testo casuale questi siano uguali

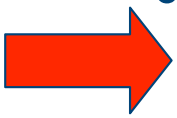
$$\sum_{i=1}^{26} \left(\frac{1}{26}\right)^2 = 0.038$$

Valori di I in funzione di L

- La probabilità di prendere una coppia uguale in testo cifrato soddisferà la seguente relazione $Q < I < P$
- $I = P$ se L(lunghezza della chiave)=1
- $I = Q$ se $L = N$(lunghezza del messaggio)

Esempio

Frasysidhxlakahg
Hjhqwaapqqsvsja
Adagsiuqywisbajl
Ajhajsuscqywoeb
ahjkabsswqomjab



L=1

$I(c_0,c_1..c_n)=$ { 0.075 se L=1 ; 0.038 se L≠1

L=2

$I(c_1,c_3\ldots)=$

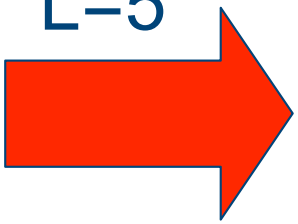
$=I(c_0.c_2\ldots.)$

{ 0.075 se L=2 ; 0.038 se L≠2

Esempio (2)

- L=1 ? $I(C_0,C_1\ldots..C_n)=0.045$
- L=2 ? $I(C_0.C_2\ldots)=0.0463$
  $I(C1.C3\ldots.)=0.0468$
- L=3 ? $I(C_0.C3\ldots)=0.0431$
  $I(C1.C4\ldots.)=0.0459$
  $I(C2,C5\ldots.)=0.0456$
- L=4? $I(C_0.C4\ldots)=0.0421$
  $I(C1.C5\ldots.)=0.0495$
  $I(C2,C6\ldots.)=0.0437$
  $I(C3,C7\ldots)=0.0444$

L=5



I(C0.C5…)=0.07221
I(C1.C6….)=0.0715
I(C2,C7….)=0.0810
I(C3,C8…)=0.0684
I(C4,C9…)=0.0759

Tutti vicini a 0.075

L=5

Suddivisione del testo

- Dividiamo il testo cifrato in L colonne
- Probabilità di prendere due lettere uguali appartenenti alla stessa colonna = P
- Probabilità di prendere due lettere uguali appartenenti a colonne diverse = Q

c_1	c_2	c_3	….	c_L
c_{L+1}	c_{L+2}	c_{L+3}	….	c_{2L}
…	…	…	…	…

Numero di coppie di lettere appartenenti alla stessa colonna N_C

- In ogni colonna ci stanno n/L simboli
- Scelto un simbolo ne rimangono (n/L)-1

Nc

$$\frac{n\left(\frac{n}{L}-1\right)}{2}$$

Numero delle coppie di lettere appartenenti a colonne diverse N_d

- n- n/L rappresenta il numero delle lettere presenti fuori dalla colonna in cui si trova la prima lettera

Nd

$$\frac{n\left(n-\frac{n}{L}\right)}{2}$$

Calcolo di L(coppie uguali)

- **Valore atteso delle coppie di lettere uguali**

$$Z = Nc * P + Nd * Q$$

$$Z = \frac{n\left(\frac{n}{L} - 1\right)}{2} 0.075 + \frac{n\left(n - \frac{n}{L}\right)}{2} 0.038$$

$$I = \frac{Z}{n(n-1)/2} \quad \Rightarrow \quad L = \frac{0.037n}{I(n-1) - 0.038 + 0.075}$$

Esempio Applicazione formula

- N=514

$$I = \sum_{i=1}^{26} \frac{n_i (n_i - 1)}{2}$$

$$\frac{n\ (n-1)}{2}$$

Contiamo le lettere e le frequenze

I = 0.045

L=5