CRITTOANALISI STATISTICA

La sicurezza di un cifrario non dipende solo dalla grandezza dello spazio delle chiavi

Spesso la vulnerabilità non è nell'algoritmo ma nelle applicazioni

(Chiavi usate male, conoscenza del formato del messaggio)

Gli attacchi di crittoanalisi statistica prevedono di forzare un cifrario con metodi statistici usando perlopiù attacchi di tipo known cypher text

Ipotesi: Il crittoanalista conosce:

- linguaggio naturale in cui è scritto il messaggio;
- modello impiegato per cifrare e decifrare;

Si ammette che il messaggio sia sufficientemente lungo

Sapendo che un cifrario è a sostituzione monoalfabetica allora

$$y \text{ nel crittogramma} \rightarrow x \text{ nel messaggio}$$

vale che

$$\text{frequenza}(y) = \text{frequenza}(x)$$

Nel caso di frequenze simili si tentano alcune permutazioni

Nel caso del cifrario di Cesare:

svelare la coppia (x, y) permette di svelare l'intero messaggio

Affini:

Servono due coppie di lettere per impostare il sistema per calcolare le incognite a, b

Completo:

si va per somiglianza tra frequenze

Sapendo che un cifrario è a sostituzione polialfabetica:

- l'istogramma delle frequenze risulta appiattito

Nel caso del cifrario di de Vigenere:

y dipende dalla coppia (x, k)

la debolezza è la chiave ripetuta più volte

Se $|k| = h$ allora le stesse lettere distanti un multiplo di h vengono codificate nella stessa lettera (è come se fosse un cifrario monoalfabetico)

$\forall i$: intero $\leq h$:

- Costruisco un sottomessaggio $m[i]$ formato da lettere in posizioni $i, i+h, i+2h \dots$
- tutte le lettere sono allineate con la

stessa lettera nella chiave

- Il messaggio è decomposto in h sottomessaggi cifrati con un cifrario monoalfabetico

Come conoscere h?

Si cercano sequenze di lettere ripetute ("q-grammi" frequenti nella lingua)

La distanza tra le coppie di q-grammi è un multiplo di h

Nel caso di cifrari a trasposizione:

si usano i q-grammi

permutazione semplice:

- si divide il criptogramma in porzioni lunghe h
- si cercano i gruppi di q lettere che formano i q grammi più diffusi
- se un gruppo deriva da un q-gramma allora si trova parte della permutazione

ESERCIZI (da Es1.pdf)

1) $|k| = 46$ bit

128 istruzioni per ogni bit di m

10^{-8} s per istruzione

Quanti anni per decifrare un messaggio da 1000 bit?

→ $2^{46} \cdot 10^{3} \cdot 10^{-8} \cdot 128 \;/\; 365 \cdot 24 \cdot 60 \cdot 60 \approx 2500$ anni

2) BBS, $x(i) = x(i-1)^2 \bmod n$, $b(i) = 1 \Leftrightarrow x(m-i)$ dispari

Scelgo $n = 11 \cdot 23$.

11 e 23 sono valori legittimi?

$M = 123456$

$y = M \bmod 100$, $x(0) = y^2 \bmod n$

indicare una sequenza di bit generati

→ $11 \bmod 4 = 23 \bmod 4 = 3$

⇒ legittimi

$2\left\lfloor \frac{11}{4} \right\rfloor + 1$, $2\left\lfloor \frac{23}{4} \right\rfloor + 1$ primi

$y = M \bmod 100 = 123456 \bmod 100 = 56$

$x(0) = 56^2 \bmod 100$

$x(1) = 36$

$x(2) = 96$

$x(3) = 16$

$x(4) = 56$ ...

3) C sequenza binaria

Calcola $37^C \bmod 100$

$34_2 \rightarrow 011\,100 \rightarrow 1011100 = C$

$\overset{64}{1}0\overset{16+8+4}{1}1\overset{0}{1}00_{10} = 64+16+8+4 = 92$

$37^{92} = 37^{64+16+8+4}$

$37^2 \bmod 100 = 69$

$37^4 = 61$

$37^8 = 21$

$37^{16} = 41$

$37^{32} = 81$

$37^{64} = 61$

$37^{92} \ (100) = 61 \cdot 41 \cdot 21 \cdot 61 \ (100) =$

$= 21 \cdot 21 \cdot 41 \ (100) =$

$= 41 \cdot 41 \ (100) =$

$81 \ (100).$