# Crittografia classica

Nessun algoritmo di cifratura è incondizionatamente sicuro.

la crittografia è la scrittura segreta basata su di un codice condiviso tra gli interlocutori, il livello di segretezza di un testo cifrato dipende da due fattori:

* Algoritmo utilizzato
* Complessità della chiave di cifratura che determina il modo in cui un messaggio viene cifrato

#### Algoritmi simmetrici

Nei sistemi simmetrici viene utilizzata una sola chiave, detta **segreta,** utilizzata come parametro unidirezionale e invertibile.

Poiché si utilizza una funziona invertibile, il destinatario dovrà soltanto elaborare nuovamente il crittogramma, **richiamando l'inverso della funzione di cifratura avente come parametro la medesima chiave utilizzata dal mittente**.

Tipi di algoritmi simmetrici classici:

* **Codifica** (simboli diversi)
* Cifrari per **sostituzione** (simboli uguali)
* Cifrari per **trasposizione**
* Cifrari **prodotto**: cifrari che integrano sostituzione e trasposizione

Sostituzione

L'alfabeto sostitutivo per il messaggio originale è costituito da stessi simboli mediante: **Sostituzione**, **Trasposizione.**

#### Algoritmi di shifting

L'alfabeto utilizzato usa lettere spostate in avanti (shiftate) di **n** posizioni. La **chiave** è il numero di posizioni di shift n.

Esempio: A->E, B->F, C->G ... “MANTOVA” -> “QERXSZE”

#### Cifrario basato su alfabeto sostitutivo (BASE)

Si modifica l'alfabeto con un alfabeto sostitutivo con sostituzione casuale delle lettere. La chiave è l'intero alfabeto sostitutivo.

Esempio:

*Alfabeto*: **abcdefghijklmnopqrstuvz**

*Alf. Sost*:  **DKVQFIBJWPESCXHTMYAUOLRGZN**

*Testo*: **ifwewishtoreplaceletters**

*Testo cifrato*: **WIRFRWAJUHYFTSDVFSFUUFYA**

#### Cifrario basato su alfabeto sostitutivo (AVANZATO)

La modifica è riassunta da una "parola chiave" o "frase" che **non deve avere lettere ripetute** (se ci sono, si eliminano quelle dopo la prima)

Esempio:

*Chiave*: **XRAYFILMS** (lunghezza = 9)

*Alfabeto*: **abcdefghijklmnopqrstuvwxyz**

*Alf. Sost*:  **XRAYFILMSBCDEGHJKNOPQTUVWZ**

#### Crittoanalisi statistica

Per la crittoanalisi:

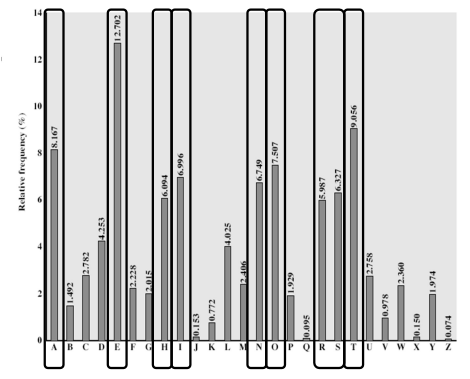
* Si calcolano le frequenze delle lettere nel testo cifrato
* Si confrontano con i valori noti di quella lingua
* Si provano a sostituire le lettere sulla base delle frequenze corrispondenti: molte funzionano, alcune no

E’ possibile utilizzare la conoscenza della **lingua del messaggio** per semplificare il problema della crittoanalisi. La conoscenza della lingua del messaggio consente di determinare anche:

* Diverse frequenze delle combinazioni di lettere: bigrammi e trigrammi più comuni o impossibili
* Lettere doppie
* *Affinità*: in italiano la lettera Q è sempre seguita dalla U, le parole tendono a terminare per vocale, prima e dopo una doppia c'è una vocale
* *Repulsione*: in italiano, la lettera N non precede mai la P, le vocali non sono mai doppie
* Spazi significativi nelle frasi (anche se questo problema può essere eliminato rimuovendo gli spazi o mettendoli a distanza fissa per disorientare il crittoanalista)

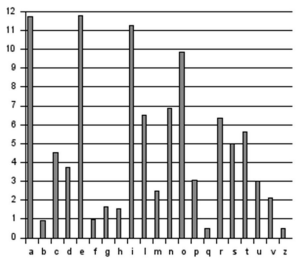
*Alfabeto inglese* (esempio):

* e = 12,7%
* t = 9,1%
* a = 8,2%
* o = 7,5%
* Altre lettere (es., j, q, x, z) sono poco frequenti



*Alfabeto Italiano* (esempio):

* Le vocali **E, A, O, I** sono le lettere più frequenti (arrivano a circa il 45% del testo)
* Seguono le consonanti **L, N, R, S, T, C**
* Tra i *bigrammi* più frequenti vi è **QU** seguito da vocale, ma anche **MP, ST**, e le doppie **CC, RR, TT, SS, LL**
* *Trigrammi* più frequenti sono: **CHE, CHI, GHE, GHI, STR**
* Due lettere uguali sono necessariamente consonanti doppie, che sono sicuramente precedute e seguite da vocali.



Pattern d'azione

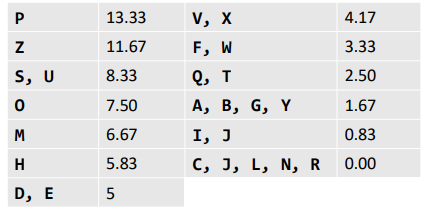
1. Concentrarsi sulle 4 vocali (e, a, o, i) e 5 consonanti più frequenti (t, n, s, h, r) consente di scoprire circa il 70% del contenuto di un messaggio.
2. Analogamente, concentrarsi sulle 4 lettere meno frequenti (j, q, x, z) consente di eliminare molte combinazioni
3. **CAVEAT**: per testi brevi, l'analisi di frequenza può non funzionare perfettamente e qualche iterazione è necessaria; per testi molto brevi non funziona.

Esempio pratico:

Decifrare il seguente testo cifrato mediante "crittoanalisi delle frequenze":

“UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZVUEPHZ HMDZSHZOWSFPAPPDTSVPQUWYMXUZUHSXEPYEPOPDZSZUFPOMBZ WPFUPZHMDJUDTMOHMQ”

Frequenza delle lettere:



possiamo ipotizzare che:

* **P** -> **e** (13.33 vs 12.70)
* **Z** -> **t** (11.67 vs 9.05)
* Le lettere {S, U, O, M, H} hanno frequenza relativamente alta e potrebbero appartenere all'insieme {a, h, i, n, o, r, s}
* Le lettere {A, B, G, Y, I, J} hanno frequenza relativamente bassa e potrebbero appartenere all'insieme {b, j, k, q, v, x, z}

A questo punto possiamo procedere verificando i bigrammi e i trigrammi, per tentativi:

* si deduce che **W** -> **h**, in quanto **ZW** è il bigramma più frequente del nel testo cifrato
* Verificando la sequenza **ZWSZ** non si sa se forma una parola, ma se lo fosse e avendo "th\*t" si può ipotizzare che **S** -> **a**

# Cifrari misti e poli-alfabetici

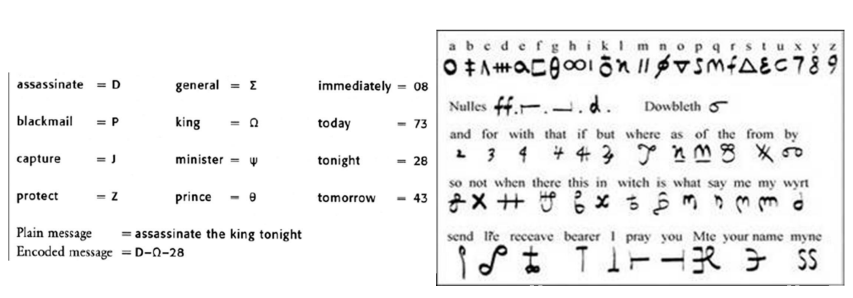
Soluzioni alla crittoanalisi per frequenza:

* **Nomenclatore**
* **Cifrari omofonici**
* **Cifrari polialfabetici**
* **Cifrari poligrafici**

Nomenclatore

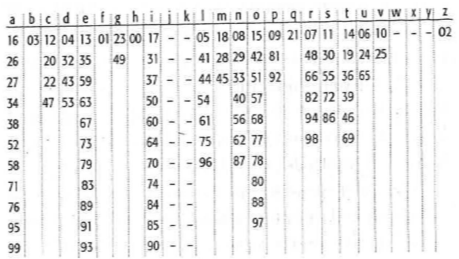
Le sue lettere erano cifrate secondo un algoritmo che combinava cifratura e codifica: non solo le lettere venivano permutate, ma venivano usati dei simboli runici per riferirsi a certe parole d'uso comune.

I nomenclatori furono usati per secoli con un numero di simboli sempre maggiore.



#### Cifratore per sostituzione omofonica

Si utilizza per la stessa lettera un numero di valori proporzionale alla sua frequenza. Usati per annullare la debolezza della frequenza e dell'identità delle lettere.



Algoritmi polialfabetici

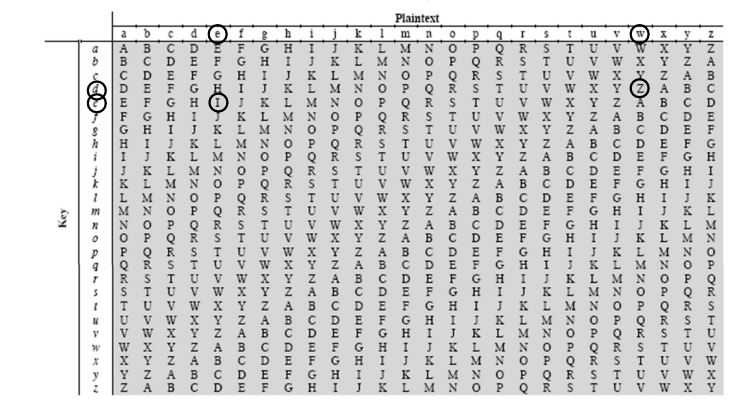
Utilizzano più alfabetici di sostituzione.

Utilizzo di 4 alfabeti: A1, A2, A3, A4:

* La prima lettera viene cifrata con la prima lettera di A1, la seconda con la prima lettera di A2 e così via. La quinta, di nuovo con A1.
* Riesce ad eliminare l'evidenza delle frequenze caratteristiche delle lettere più popolari.

#### Algoritmi polialfabetici: Vigenère

* Si utilizza una chiave di d lettere: K = k1, k2, …, kd, dove l'ima lettera specifica l'i-mo alfabeto da utilizzare
* Si utilizza un alfabeto per ogni lettera da cifrare ripetendo la chiave se (come tipicamente accade) il testo è più lungo di d lettere
* Per semplificare le operazioni di cifratura/decifratura si può utilizzare la cosiddetta **Tavola di Vigenère**



Esempio:

*Testo*: “uccidereilrediscoziadomaninotte”

*Chiave*: “mantovamantovamantovamantovamantova”

*Vigenère*: “GCPBRZRQ IY KS YI ECBSWV DAMNGW BJTFE”

*Mono-alfabetico*: “AIIOJKXK OR XK JO YIUFOG JUSGTO TUZZK”

#### Passi di crittoanalisi Vigenère

1. Identificare la lunghezza della chiave L
2. Suddividere il testo cifrato in L sottotesti
3. Applicare un algoritmo di crittoanalisi per frequenza ad ogni sottotesto

*Calcolare la lunghezza della chiave:*

1. Ricercare sequenze di caratteri identici nel testo cifrato
2. Calcolare la distanza t tra gli inizi di queste sequenze
3. Identificare tutte le sequenze identiche nel testo cifrato e calcolare le varie distanze **t1 , t2 , …, tk**
4. Lunghezza chiave: **L = MCD(t1 , t2 , …, tk )**

Esempio:

“KIOVIEEIGKIOVNURNVJNUVKHVMGZIANBRKBRICF”

L = MCD(9, 6, 3) = 3 è un multiplo della chiave

#### XOR Alfabetico

Il simbolo di operatore XOR in crittografia viene applicato anche in contesti non binari come **somma modulare**

*Esempio*:

S ⨁ I -> A equivale a: (18+8) mod 26 = 0

E ⨁ H -> L equivale a: (5+7) mod 26 = 11

# One Time Pad

Uso del "blocco usa e getta" (one-time pad): chiave lunga quanto il messaggio quindi senza alcuna ricorrenza. Idea originaria:

* Ogni pagina del blocco contiene una lista di numeri casuali (es., chiavi di 30 caratteri)
* Per cifrare un messaggio, il mittente deve staccare tanti fogli quanti ne sono necessari

Mittente e destinatario devono avere una copia identica del blocco e usarlo in modo perfettamente sincronizzato.

Ogni sistema basato su ripetizione è critto-analizzabile; **One Time Pad non usa ripetizione**.

Ogni lettera del messaggio originale è accoppiata ad un numero casuale differente, il testo cifrato risultante è incondizionatamente sicuro.

*Esempio*:

Un libro è una fonte di "simboli casuali" sufficientemente lunga per implementare nella pratica la tecnica del blocco "usa e getta" Mittente e destinatario devono mettersi d'accordo su:

* Libro
* Pagine
* Righe
* Frequenza lettere

La sicurezza dell'OTP si basa sulla effettiva casualità della chiave: se la sequenza è veramente casuale, il crittogramma non è soggetto ad attacchi.

**Mai riutilizzare il PAD**

Si ipotizzi di utilizzare due volte lo stesso pad K

**C1 = M1** ⨁ **K**

**C2 = M2** ⨁ **K**

Dove C1 e C2 sono i testi cifrati, M1 e M2 sono i testi in chiaro e K la chiave.

Se il crittoanalista intercetta i messaggi crittografati con lo stesso pad, può sfruttare la seguente proprietà: **C1** ⨁ **C2 = (M1** ⨁ **K)** ⨁ **(M2** ⨁ **K) = M1** ⨁ **M2**

Se i messaggi sono sufficientemente lunghi è possibile ricavare M1 e M2.

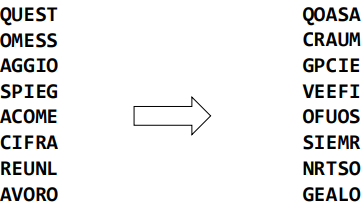
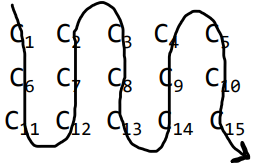
# Cifrari per trasposizione

L'obiettivo della sostituzione è la confusione (nascondere le lettere originali) per rendere difficile il compito di determinare come un messaggio viene trasformato in un testo cifrato.

Con la trasposizione, la crittografia mira alla diffusione, spargendo le informazioni del messaggio nel testo cifrato.

**Trasposizione colonnare**

Si scrivono i caratteri di un testo in chiaro su righe di cinque caratteri, una sotto l'altra. Il testo cifrato viene ottenuto leggendo le colonne dall'altro verso il basso



# Cifrari prodotto

La **sostituzione** e la **trasposizione** sono le pietre angolari della crittografia. Tuttavia, la crittografia degli ultimi decenni si basa su una combinazione di più tecniche. Una combinazione di due o più cifrature è detta **cifratura prodotto**.

**Algoritmo ADFGVX**

Algoritmo prodotto che combina sostituzione e trasposizione. Fase di sostituzione: a ciascun elemento della Tabella si fa corrispondere una coppia di lettere

Esempio:

“**MANTOVA CENTRO**” -> “**XX GV GX XV VD GD GV DG GA GX XV AG VD**”

