



cenni di crittografia

Alberto Ferrari – Analisi dei Dati

steganografia

- La steganografia è una tecnica che si prefigge di **nascondere** la comunicazione tra due interlocutori. Il termine è composto dalle parole greche στεγανός (coperto) e γραφία (scrittura).
- Erodoto scrive che intorno al 500 a.C., il re spartano Demarato utilizzò un geniale stratagemma per nascondere messaggi di grande importanza. Per le comunicazioni scritte venivano utilizzate delle tavolette ricoperte di cera. La cera permetteva di scrivere e cancellare un messaggio, garantendo il riutilizzo della tavoletta. L'idea del sovrano spartano consistette nello scrivere il messaggio sotto la cera, direttamente sulla tavoletta. Successivamente, per depistare ogni sospetto, sulla cera veniva scritto un altro messaggio. In questo modo, l'informazione importante veniva nascosta e riusciva a passare “sotto il naso” di un eventuale controllo nemico.



steganografia

- Nascondere un messaggio di testo all'interno di un'immagine
- <https://stylesuxx.github.io/steganography/>
- <https://www.cryptool.org/en/cto/bacon/>



Alberto Ferrari – Analisi dei Dati

crittografia

- ***crittografia***
 - κρυπτός (**nascosto**) e γραφία (**scrittura**)
 - metodo per rendere un messaggio non comprensibile a persone non autorizzate a leggerlo
 - il messaggio è definito ***crittogramma***
- ***crittoanalisi***
 - studio dei metodi per ottenere il significato di informazioni cifrate senza avere accesso all'informazione segreta
 - è la "controparte" della crittografia
- ***crittologia***
 - studio della crittografia e della crittoanalisi

chiave

- *chiave*
 - informazione usata come parametro in un algoritmo crittografico
- unico ***dato*** che è necessario mantenere ***segreto***
- la dimensione della chiave (misurata in bit) dipende dall'algoritmo usato
- ***brute force***
 - attacco a forza bruta (ricerca esaustiva)
 - provare tutte le chiavi possibili
 - una chiave di n bit ha 2^n chiavi distinte

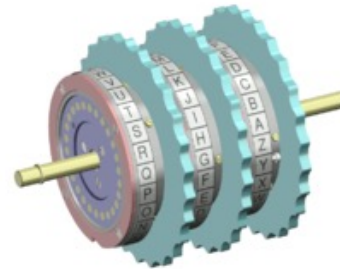
crittografia classica ***(dall'antichità al 1975)***

- ***metodi antichi***
 - scitala spartana
 - scacchiera di Polibio
 - codice atbash
 - codice di Cesare
- ***rinascimento***
 - Blaise Vigenère
- ***XX secolo***
 - macchina Enigma
 - usata dai tedeschi durante la Seconda Guerra Mondiale
 - DES (Data Encryption Standard)

crittografia moderna

- ***crittografia moderna***
 - nasce nel 1975 con un articolo di Diffie & Hellman
 - si propone un **nuovo protocollo per lo scambio delle chiavi** (tallone d'Achille della crittografia classica)
- rende possibile la trasmissione sicura di dati fra entità che non hanno concordato preventivamente le chiavi
- esempio:
 - RSA





crittografia classica

Alberto Ferrari – Analisi dei Dati

cifrari monoalfabetici

- cifrario a *sostituzione monoalfabetica*
 - utilizza un alfabeto per il testo in chiaro e una *permutazione* dello stesso per il testo cifrato
 - la permutazione costituisce la chiave del sistema
 - ad ogni lettera del testo in chiaro viene associata la corrispondente lettera dell'alfabeto permutato

500-600 a.c. *cifrario atbash*

- ***atbash***
 - cifrario a sostituzione ***monoalfabetica***
 - la prima lettera dell'alfabeto è sostituita con l'ultima, la seconda con la penultima, e così via
- *testo in chiaro:*
 - **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
- *testo cifrato:*
 - **ZYXWVUTSRQPONMLKJIHGFEDCBA**
 - esempio:
 - LEZIONI DI CRITTOGRAFIA
 - OVARLMR WR XIRGGLTIZURZ

<https://www.cryptool.org/en/cto/atbash/>



python – atbash crittare un testo

```
alfabeto = "abcdefghijklmnopqrstuvwxyz"
cifrato = "zyxwvutsrqponmlkjihgfedcba"
testo = "big data"
testo_cifrato = ""
for c in testo:
    pos = alfabeto.find(c)
    if pos == -1:
        testo_cifrato = testo_cifrato + c
    else:
        testo_cifrato = testo_cifrato + cifrato[pos]
print('il testo:          ', testo)
print('viene crittato come:', testo_cifrato)
```

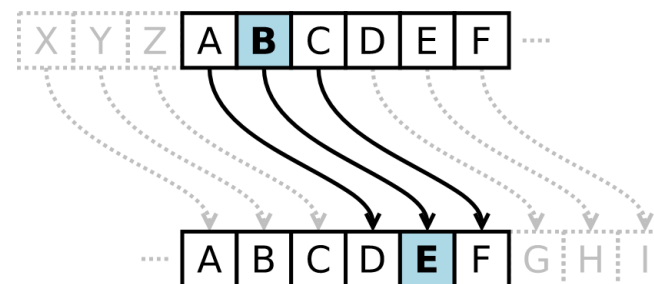

python – atbash decrittare un testo

```
testo_decifrato = ""
for c in testo_cifrato:
    pos = alfabeto.find(c)
    if pos == -1:
        testo_decifrato = testo_decifrato + c
    else:
        testo_decifrato = testo_decifrato + cifrato[pos]
print('il testo crittato:      ',testo_cifrato)
print('viene decrittato come:',testo_decifrato)
```


50-60 a.c.

metodo di Cesare

- cifrario di **Cesare**
 - cifrario a sostituzione monoalfabetica in cui ogni lettera del testo in chiaro è sostituita nel testo cifrato dalla lettera che si trova un certo numero di **posizioni successive** nell'alfabeto
- Cesare utilizzava uno spostamento di **3** posizioni (CHIAVE 3)
- testo in chiaro:
 - ABCDEFGHIJKLMNOPQRSTUVWXYZ
- testo cifrato:
 - DEFGHIJKLMNOPQRSTUVWXYZABC
- esempio:
 - LEZIONI DI CRITTOGRAFIA
 - OHCLRQL GL FULWWRJUDILD

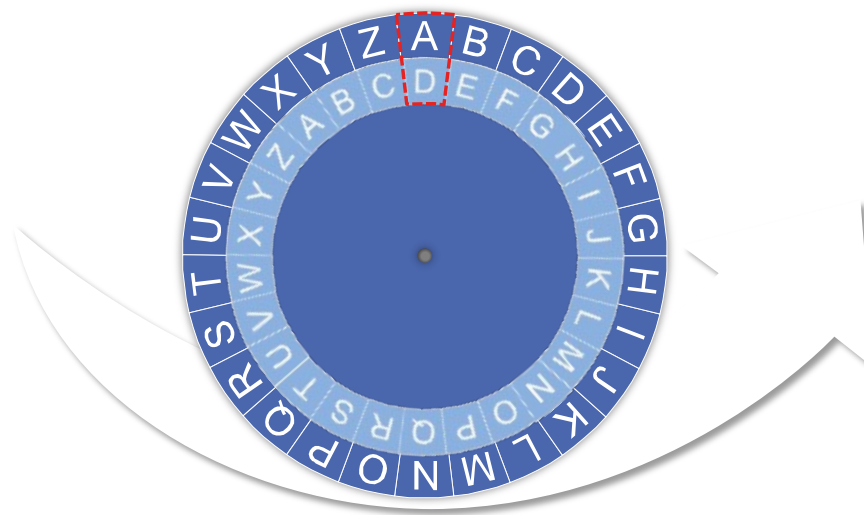


<https://www.cryptool.org/en/cto/caesar/>

Cesare

- P (alfabeto testo in chiaro [plaintext])
 - C (alfabeto testo crittato)
 - K_E (chiave di cifratura, parametro per f)
 - K_D (chiave di decifratura, parametro per f^{-1})
 - $f()$ (funzione di trasformazione crittografica)
-
- $P = C = \{A, B, C, \dots, X, Y, Z\}$
 $= \{0, 1, 2, \dots, 23, 24, 25\}$
 - $K_E = k \in P, k \neq 0$
 - $f(x_i) = (x_i + k) \bmod 26$
 - $K_D = k_d \in P, k_d = 26 - k$
 - $f^{-1}(x_i) = (x_i + k_d) \bmod 26$

cifrari a scorrimento



python – Cesare crittare un testo

```
alfabeto = "abcdefghijklmnopqrstuvwxyz"
cifrato = "defghijklmnopqrstuvwxyzabc"
testo = "big data"
testo_cifrato = ""
for c in testo:
    pos = alfabeto.find(c)
    if pos == -1:
        testo_cifrato = testo_cifrato + c
    else:
        testo_cifrato = testo_cifrato + cifrato[pos]
print('il testo: ', testo)
print('viene crittato come:', testo_cifrato)
```


python – Cesare decrittare un testo

```
decifrato = "xyzabcdefghijklmnopqrstuvw"  
testo_decifrato = ""  
for c in testo_cifrato:  
    pos = alfabeto.find(c)  
    if pos == -1:  
        testo_decifrato = testo_decifrato + c  
    else:  
        testo_decifrato = testo_decifrato + decipherato[pos]  
print('il testo crittato:      ',testo_cifrato)  
print('viene decrittato come:',testo_decifrato)
```

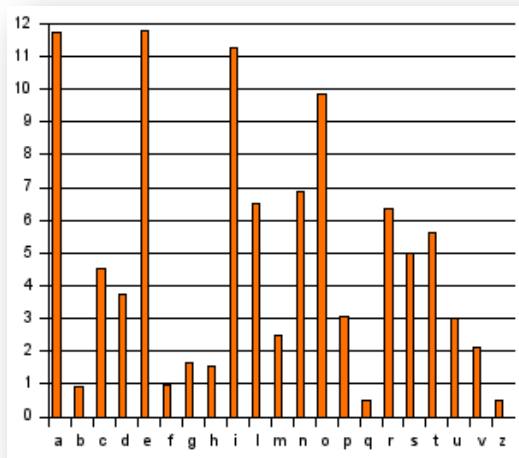

debolezza del metodo di Cesare

- il metodo di Cesare ha due principali ***debolezze***:
 - è sensibile all'**analisi di frequenza**
 - sono possibili solo **poche chiavi** diverse ($n - 1$) se n è il numero di caratteri dell'alfabeto
- chi intercetta un messaggio cifrato con il metodo di Cesare può limitarsi a provare successivamente tutte le possibili chiavi di cifratura e trovare il testo in chiaro in un tempo ragionevolmente breve (attacco a **forza bruta**)

analisi di frequenze

- *analisi delle frequenze*

- studio della frequenza di utilizzo delle lettere o gruppi di lettere in un testo cifrato
- in ogni lingua la **frequenza** di uso di ogni lettera è piuttosto determinata



Alberto Ferrari – Analisi dei Dati

Lettera	Frequenza
a	11.74%
b	0.92%
c	4.50%
d	3.73%
e	11.79%
f	0.95%
g	1.64%
h	1.54%
i	11.28%
l	6.51%
m	2.51%
n	6.88%
o	9.83%
p	3.05%
q	0.51%
r	6.37%
s	4.98%
t	5.62%
u	3.01%
v	2.10%
z	0.49%

Cifrario Kāma Sūtra

- Il cifrario del Kamasutra è uno dei primi metodi di sostituzione conosciuti. È descritto nel Kamasutra intorno al 400 a.C. Lo scopo era quello di insegnare alle donne come nascondere messaggi segreti da occhi indiscreti.
- La chiave viene creata mescolando l'alfabeto in un ordine casuale. Viene poi divisa in due metà e scritta una sotto l'altra.

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

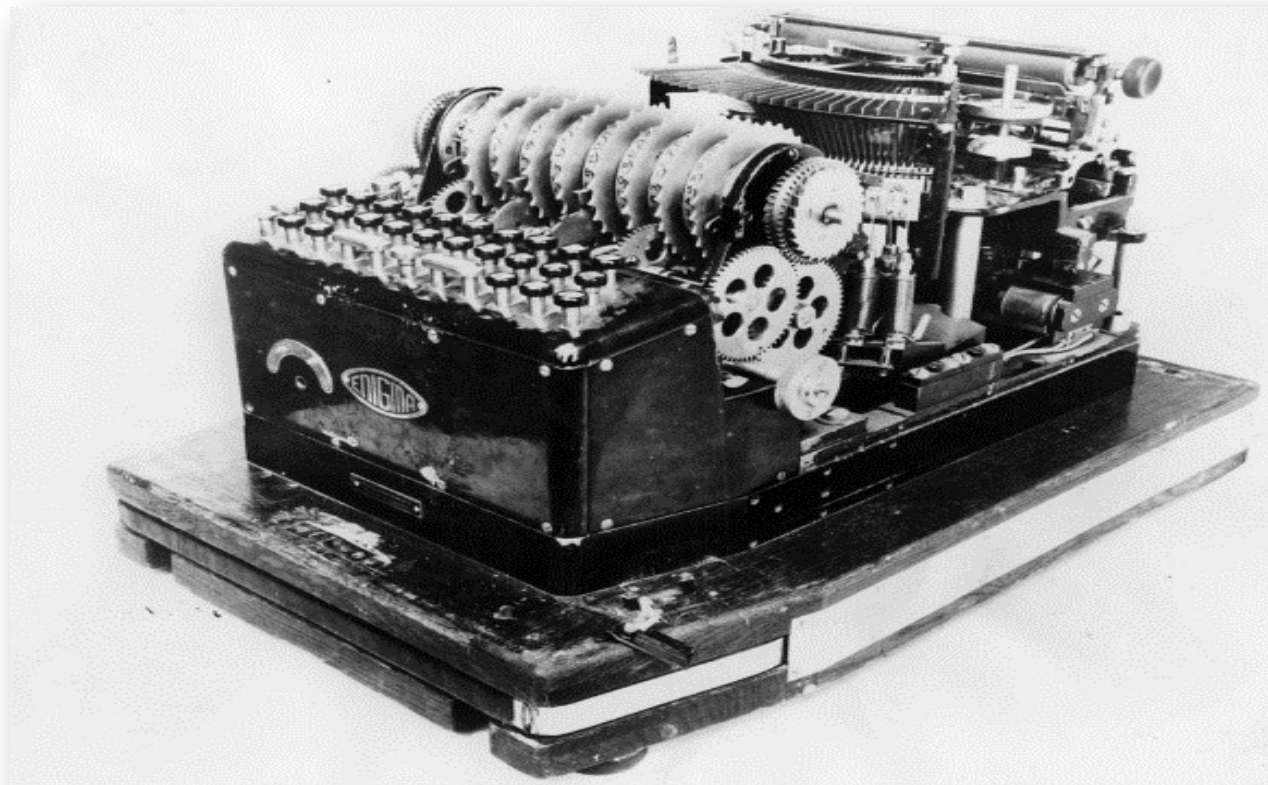
<https://www.cryptool.org/en/cto/kamasutra/>



la macchina Enigma

- macchina elettromeccanica usata dai tedeschi nella seconda guerra mondiale
- una serie di rotori effettuano la trasformazione di un carattere dell'alfabeto in un altro che viene a sua volta trasformato dal rotore successivo
- dopo la digitazione di ogni carattere il primo rotore effettua una rotazione che può comportare la rotazione eventuale del successivo
- la scelta della posizione iniziale dei rotori (e di altri meccanismi di traslazione) costituisce la chiave
- ritenuta per molto tempo inattaccabile
- il matematico polacco Marin Rejewsky con il suo lavoro riuscì a decifrare numerosi messaggi militari tedeschi, un fattore che probabilmente contribuì alla vittoria finale degli alleati

Enigma



Alberto Ferrari – Analisi dei Dati

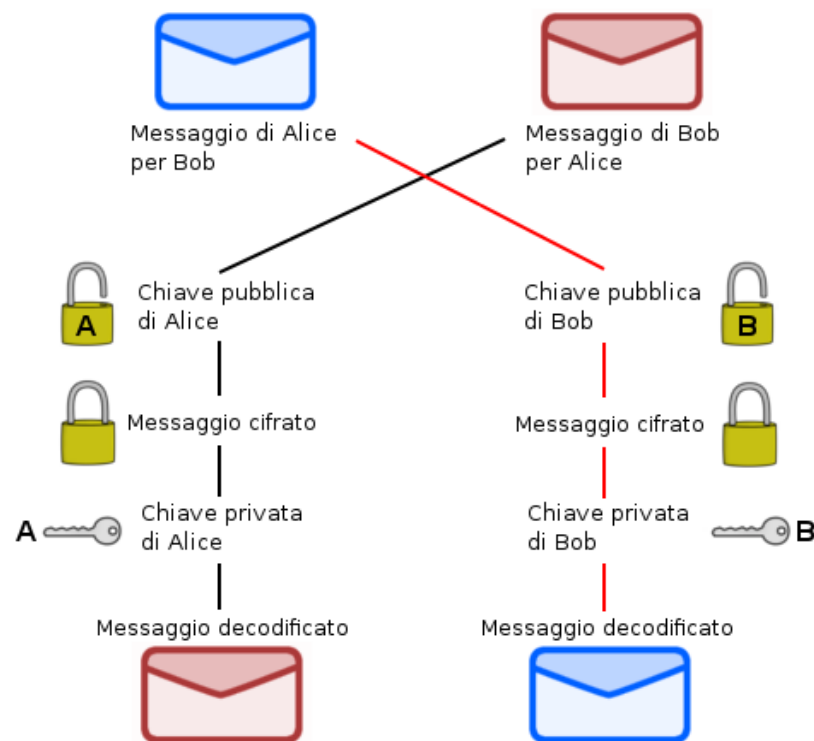
Whatsapp - crittografia a chiave asimmetrica

- si basa sull'utilizzo di ***due chiavi***
 - chiave ***pubblica***
 - chiave ***privata***
- le chiavi sono generate in modo automatico tra i due contatti
- da una chiave ***non è possibile risalire*** all'altra
- nel caso di Whatsapp
 - la ***cifratura*** avviene all'interno del proprio smartphone
 - poi il messaggio viene consegnato al server di Whatsapp
 - il server lo invia al dispositivo del destinatario
 - lo smartphone del destinatario ***decifra*** il messaggio tramite la sua chiave privata

crittografia end to end su WhatsApp

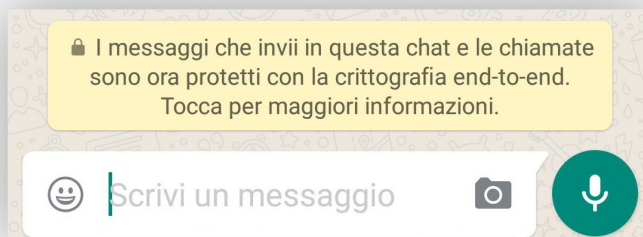
- nel momento in cui viene aggiunto un contatto le app dei due dispositivi si connettono e creano due ***coppie di chiavi interdipendenti***
- le ***chiavi private*** sono memorizzate sui rispettivi ***dispositivi*** e sono ***invisibili*** anche a WhatsApp stessa
- quando viene inviato un messaggio il server dell'azienda lo riceve e lo indirizza al destinatario
 - ***non è in grado di decifrarlo***
- un eventuale ***attacco hacker*** ai server di WhatsApp non può in ogni caso scoprire le chiavi private né accedere ai messaggi
- WhatsApp ***non registra*** le conversazioni e non può condividerle con altre organizzazioni
 - ***neanche con le forze dell'ordine***

crittografia end-to-end (End-to-End Encryption (E2EE))



Whatsapp – Telegram ... assoluta privacy

- ***nessuno*** oltre al mittente e al destinatario può leggere il messaggio
- ***neanche la polizia*** di nessun paese al mondo potrà chiedere al gestore di vedere i messaggi inviati e ricevuti
- ***neanche la società*** stessa li può leggere



MD5 - Message Digest 5

- algoritmo di crittografia ***a senso unico***
- non prevede la decrittazione
- Ronald Rivest 1991
- è una ***funzione di compressione***
 - ***input*** una stringa di lunghezza arbitraria
 - ***output*** (firma digitale) una stringa da 128 bit
 - si presuppone che l'output restituito dalla funzione sia ***univoco***
 - più precisamente che sia molto improbabile ottenere due output identici da due input diversi

python – esempio MD5

```
import hashlib
testo = "big data"
testo_cifrato = hashlib.md5(testo.encode())

print("testo cifrato in caratteri esadecimali : ")
print(testo_cifrato.hexdigest())
```


Whatsapp - crittografia

- Informazioni sulla crittografia end-to-end
- https://faq.whatsapp.com/820124435853543/?locale=it_IT

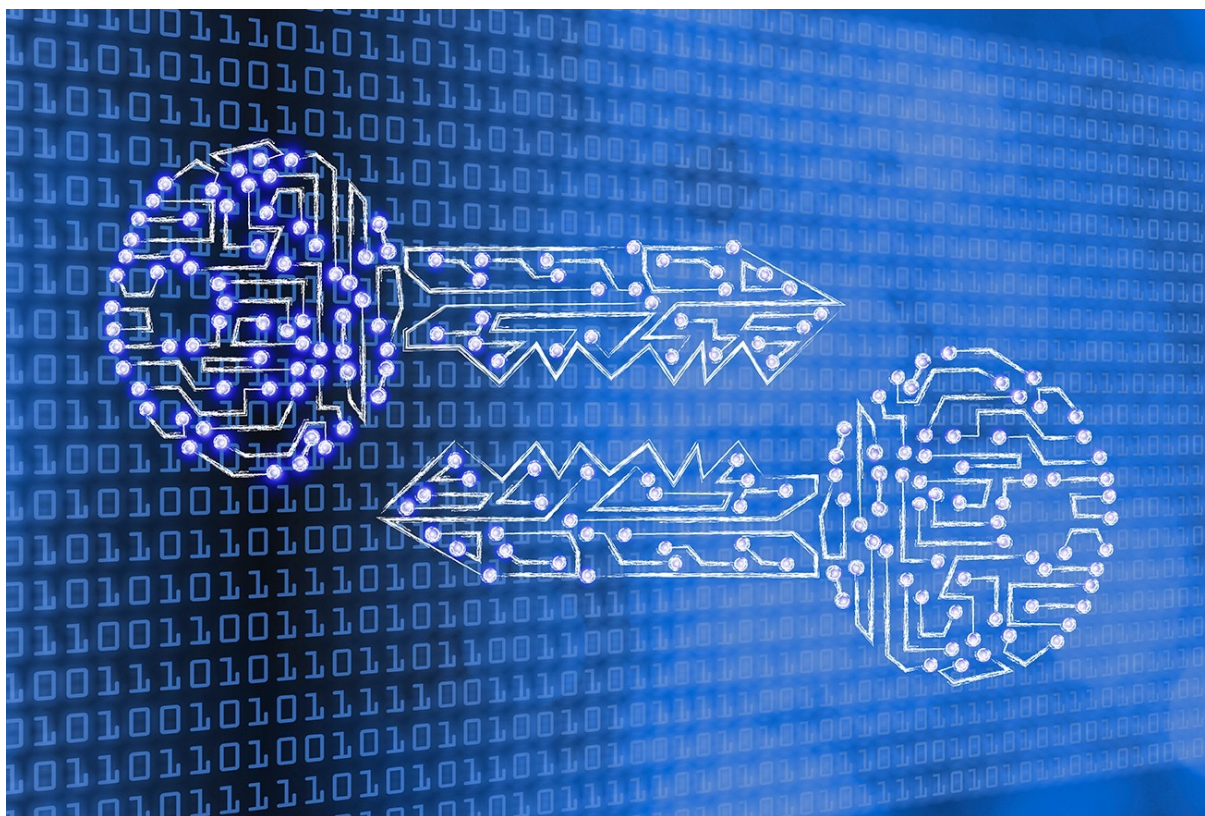


Telegram - sicurezza

- Quindi come crittografate i dati?
- <https://telegram.org/faq#sicurezza>



crittografia moderna RSA



Alberto Ferrari – Analisi dei Dati

RSA

(dal nome dei suoi inventori Rivest, Shamir e Adleman)

- RSA è uno dei metodi più usati per la crittografia asimmetrica, cioè con chiave pubblica e chiave privata
- si basa su un principio matematico semplice ma potente:
 - è facile moltiplicare due numeri primi grandi, ma è difficile fare l'operazione inversa (trovare i fattori primi del prodotto)
- questo principio garantisce la sicurezza dell'algoritmo

1. Scegli due numeri primi grandi, p e q .

Esempio:

$$p = 61, q = 53$$

2. Calcola il loro prodotto:

$$n = p \times q = 3233$$

👉 Questo sarà parte della chiave pubblica e privata.

3. Calcola la funzione di Eulero:

$$\varphi(n) = (p - 1)(q - 1) = 60 \times 52 = 3120$$

4. Scegli un numero e tale che:

- $1 < e < \varphi(n)$
- e sia coprimo con $\varphi(n)$

(cioè non abbiano divisori comuni)

Esempio: $e = 17$

5. Calcola d , l'inverso moltiplicativo di e mod $\varphi(n)$:

$$d \times e \equiv 1 \pmod{\varphi(n)}$$

👉 Qui $d = 2753$

chiavi risultanti

- chiave pubblica: $(e,n)=(17,3233)$
- chiave privata: $(d,n)=(2753,3233)$

cifratura (encryption)

- per cifrare un messaggio M (convertito in numero):
- $C = M^e \bmod n$



- esempio:
- $M=123 \Rightarrow C=123^{17} \bmod 3233 = 855$
- il messaggio cifrato è 855

decifratura (decryption)

- per decifrare:
- $M = C^d \bmod n$
- $M = 855^{2753} \bmod 3233 = 123$
- il messaggio originale è stato recuperato



procedimento

Passaggio	Descrizione
Generazione chiavi	si scelgono due numeri primi e si calcola (n, e, d)
Cifratura	si eleva il messaggio a $(e) \bmod (n)$
Decifratura	si eleva il messaggio cifrato a $(d) \bmod (n)$
Sicurezza	dipende dalla difficoltà di fattorizzare (n) in (p) e (q)

dove viene utilizzato

- HTTPS / SSL per connessioni sicure
- Firma digitale (per autenticare messaggi)
- Scambio di chiavi in protocolli sicuri