DOI: 10.13137/2039-8646/35612

La crittografia classica come occasione di ragionamento matematico[*]

FABIO PASTICCI
Dipartimento di filosofia,
scienze sociali, umane e della formazione
Università di Perugia
fabio.pasticci@unipg.it

ABSTRACT

Cryptography can be a tool that, enables students to consolidate knowledge and develop mathematical skills if used as a recreational or playful activity. In this paper we introduce Caesar's cryptosystem, which is one of the simplest encryption schemes. It is very easy to crack the encryption of this system: a brute force attack allows easy recovery of the key and the plaintext. We made a simple modification that allows us to have a greater number of keys, in order to make such a brute force attack infeasible. Finally, we show a statistical method that allows us to force this cryptosystem.

PAROLE CHIAVE

CRITTOSISTEMA DI CESARE / CAESAR CRYPTOSYSTEM; CRITTOGRAFIA / CRYPTOGRAPHY; CRITTOANALISI / CRYPTOANALYSIS; ANALISI DELLE FREQUENZE / FREQUENCY ANALYSIS; DIDATTICA DELLA MATEMATICA / MATHEMATICS EDUCATION.

1. INTRODUZIONE

La *crittografia*, ovvero l'arte e la scienza di trasmettere messaggi facendo in modo che solo il legittimo destinatario sia in grado di leggerli, può essere utilizzata in classe come strumento utile per il recupero di abilità operative e il consolidamento di conoscenze matematiche. Al fine di rendere meno "astratto" lo studio della matematica è possibile trovare schemi crittografici alla cui base ci sono elementi matematici che sono collegati con i bisogni formativi delle singole classi: ad esempio già in una classe terza della Scuola primaria è possibile presentare un *cifrario di Cesare* e far sperimentare

[*] *Title: Classical cryptography as an opportunity for mathematical reasoning.*

ai ragazzi la tecnica per cifrare e decifrare i messaggi, mentre già in una classe quarta, sempre della Scuola primaria, è possibile mostrare alcuni collegamenti tra il cifrario di Cesare e l'aritmetica modulare.

Nello specifico si può evidenziare il fatto che spostare di alcuni posti le lettere dell'alfabeto significa addizionare numeri naturali con l'accortezza che quando si arriva all'ultima lettera dell'alfabeto è necessario ripartire dalla prima. Questo, in termini numerici, si traduce nel seguente modo: se si ha un alfabeto con 26 lettere l'addizione 26 più 1 avrà come risultato 1.

Inoltre la crittografia

> *consente di creare collegamenti con altre discipline, quali la storia (dai metodi antichi di cifratura e decifratura di codici segreti fino al suo utilizzo durante la Seconda Guerra Mondiale), la letteratura (utilizzando testi famosi come esempi di cifratura e decifratura) e l'educazione civica (ad esempio ragionando sulle frodi digitali e sulla necessità di limitare la condivisione di dati personali).*[1]

Ovviamente esistono anche altri metodi per tramettere messaggi segreti: scrivere il testo con l'inchiostro simpatico, affidarsi a un messaggero fidato, dividere la comunicazione in più parti per poi inviarle utilizzando canali di trasmissione diversi, solo per citarne alcuni. Quindi, perché si preferisce la crittografia?

> *La risposta è semplice [...] la matematica fornisce, almeno all'ingrosso, la giustificazione teorica per la forza di un particolare algoritmo o protocollo.*[2]

Un'altra motivazione molto rilevante dal punto di vista didattico è che

> *cryptography is a charming and rewarding way to introduce into the classroom subjects of traditional or less traditional mathematics.*[3]

In questo lavoro si illustra il *crittosistema* (o *cifrario*) di Cesare e l'attacco esaustivo che è facilmente realizzabile e che permette di risalire alla chiave e al testo in chiaro in breve tempo. Si illustra quindi una modifica al suddetto schema di cifratura che rende inutile l'attacco a forza bruta. Poi viene presentata una ulteriore tipologia di attacco che permette di forzare questo crittosistema derivato da quello di Cesare.

[1] Cfr. CAZZOLA, GRAZIAN 2021.
[2] Cfr. BERARDI, BEUTELSPACHER 1996.
[3] Cfr. BORRELLI, FIORETTO, SGARRO, ZUCCHERI 2002.

In questo percorso viene dato spazio alla concretizzazione del concetto di *funzione* e all'introduzione di nozioni di base di *aritmetica modulare* che, opportunamente presentata, come è stato già visto negli obiettivi collaterali dell'attività di cifrazione e di decifrazione descritti da Zuccheri[4], amplia quella che può essere definita la *matematica dell'orologio*.

Occorre mettere in atto tutte le strategie possibili, in modo da far scaturire il tutto da attività ludiche da presentare ai ragazzi nell'ottica di far

> *acquistare confidenza con la matematica, riuscire ad unire un argomento all'altro, in maniera continua, senza interruzioni, in modo da costruire una specie di «filo di Arianna» che permetta di percorrere il labirinto delle proprie conoscenze matematiche con una certa dimestichezza, senza bruschi salti nel buio.*[5]

Per poter illustrare i vari aspetti della crittografia si precisa qui la terminologia con alcune definizioni.

Si chiama *schema crittografico*, *cifrario* o *crittosistema* il raggruppamento dell'insieme dei possibili *testi in chiaro*, dei corrispondenti *testi cifrati*, degli *algoritmi* (di *cifratura* e di *decifratura*) e delle *chiavi*. Il *testo in chiaro* è la sequenza di lettere, numeri o simboli che si vogliono trasmettere. La *cifratura* è l'operazione che trasforma il *testo in chiaro* in *testo cifrato*. Questa azione viene posta in essere per mezzo di una *chiave segreta* cioè di una informazione che deve essere nota solamente al mittente e al destinatario del messaggio. L'operazione inversa della *cifratura* si chiama *decifratura* se eseguita dal legittimo destinatario per mezzo della chiave e si chiama invece *decrittazione* o *crittoanalisi* se è svolta da un intruso che tenta di risalire al testo in chiaro senza conoscere la chiave. Da quanto appena detto appare evidente che la crittografia permette di creare situazioni didattiche in cui il docente «cattura l'attenzione e rende piacevole fare matematica»[6]. Inoltre

> *l'osservazione di processi e competenze permette di riconoscere molteplici gradualità, dando all'insegnante la possibilità di apprezzare i progressi e le potenzialità di ogni allievo.*[7]

[4] Cfr. ZUCCHERI, 1992.
[5] Cfr. PERES 1986.
[6] Cfr. D'AMORE 1999.
[7] Cfr. DI MARTINO, ZAN 2020.

2. Il cifrario di Cesare

Il *cifrario di Cesare* è uno dei più semplici schemi crittografici: può essere presentato utilizzando pochi strumenti teorici e soprattutto permette di vedere la «matematica come strumento per indagare e descrivere la realtà»[8].

Si tratta di uno dei crittosistemi che vengono definiti *simmetrici*, o anche *classici*, poiché la chiave utilizzata dal mittente per cifrare il messaggio è la stessa che permette al destinatario di risalire al testo in chiaro partendo dal testo cifrato che ha ricevuto.

Svetonio, in *Vite dei Cesari* scrive:

> *si qua occultius perferenda erant, per notas scripsit, id est sic structo litterarum ordine, ut nullum uerbum effici posset: quae si qui inuestigare et persequi uelit, quartam elementorum litteram, id est* D *pro* A *et perinde reliquas commutet.*[9]

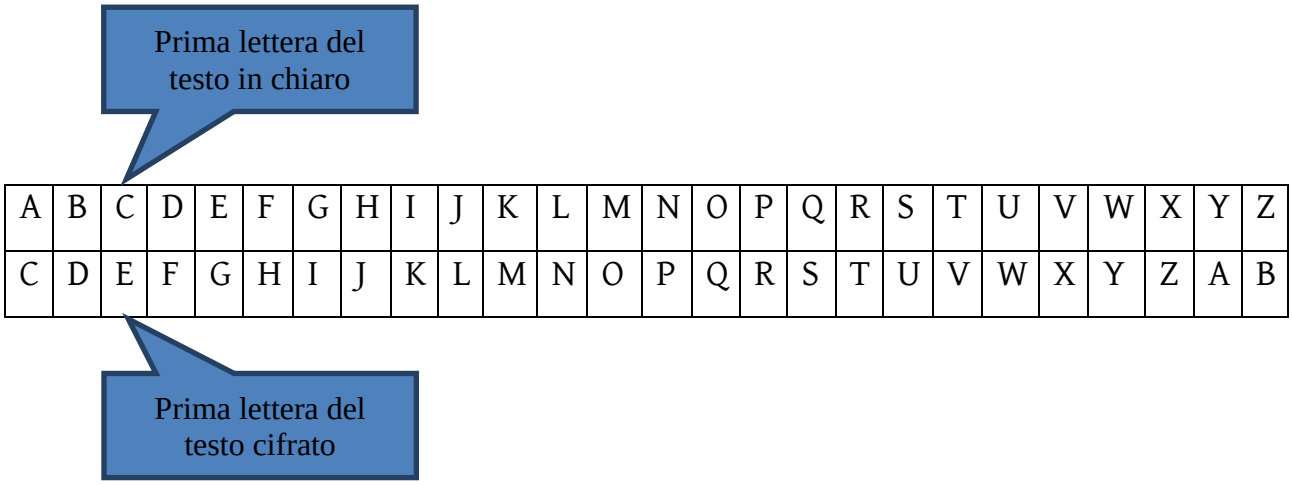
La cifratura di un messaggio, mettendo in pratica quanto appena descritto, avviene nel modo seguente:

- si prende una tabella con due righe e tante colonne quante sono le lettere dell'alfabeto che si considera (l'esempio nella Tabella 1 utilizza l'alfabeto italiano con 26 lettere);
- si scrive sulla prima riga l'intero alfabeto e sulla seconda lo si va a riscrivere iniziando da una lettera diversa dalla lettera A. Quando si arriva alla lettera Z, si continua ripartendo dalla lettera A e proseguendo in ordine fino ad arrivare all'ultima casella disponibile. Questo si può anche definire uno scorrimento circolare dell'alfabeto.

La cifratura avviene utilizzando la seguente procedura. Ogni lettera del testo in chiaro viene cercata sulla prima riga della tabella. La lettera che si trova sulla seconda riga, sotto la lettera trovata, costituisce la corrispondente lettera cifrata. Ad esempio se si deve cifrare la parola "CIAO" si inizia cercando la lettera C sulla prima riga.

[8] Cfr. Sabena, Ferri, Martignone, Robotti 2019.

[9] Cfr. Suetonius Tranquillus 2016. «Se doveva comunicare informazioni in modo riservato, le scriveva in linguaggio cifrato, cioè seguendo una sequenza alfabetica disposta in modo tale che non se ne potesse ricavare nessuna parola di senso compiuto: se si desidera comprendere il testo bisogna sostituire la quarta lettera dell'alfabeto alla prima, cioè la D al posto della A e così tutte le altre».



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
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

Tabella 1. Crittosistema di Cesare con chiave A --> C.

Quindi la prima lettera del testo cifrato sarà la lettera E.

Poi si cerca la lettera I sempre sulla prima riga e quindi la lettera K sarà la seconda lettera del testo cifrato. Si procede allo stesso modo per la lettera A e per la lettera O. Quindi "EKCQ" è la parola cifrata che corrisponde alla parola in chiaro "CIAO".

La chiave di cifratura è il numero di posizioni di cui viene spostata la lettera A dell'alfabeto nella seconda riga. Il destinatario del messaggio, che è a conoscenza della chiave, riceve la parola cifrata "EKCQ".

Come affermato da Zuccheri[10], questa procedura «fornisce l'occasione di far lavorare i bambini con un'applicazione biunivoca (*cifratura*) e con la sua inversa (*decifrazione*)».

La decifratura del messaggio avviene nel modo seguente: si prende la stessa Tabella 1 utilizzata per la cifratura del messaggio e ogni lettera del testo cifrato viene cercata sulla seconda riga e sostituita nel corrispondente testo in chiaro dalla lettera che si trova sopra la lettera trovata.

Nel caso in questione si cerca la lettera E sulla seconda riga e si sostituisce con la lettera C, poi si cerca la K sempre sulla seconda riga e si sostituisce con la lettera I poi analogamente per la lettera C e per la lettera Q. Quindi si ottiene la parola "CIAO" che è il testo in chiaro di partenza.

[10] Cfr. ZUCCHERI 1992.

Dopo aver presentato agli alunni – ad esempio, di una classe terza della Scuola primaria - il cifrario di Cesare, si possono organizzare giochi in aula, anche a squadre, in cui vince chi riesce a decifrare un messaggio prima degli altri fornendo la Tabella 1 tutta compilata.
Prima di passare alla descrizione delle attività legate ai procedimenti che un intruso può attuare per forzare il crittosistema, possiamo notare che il cifrario di Cesare è di tipo monoalfabetico: «ogni lettera dell'alfabeto in chiaro è cifrata sempre con la stessa lettera dell'alfabeto segreto»[11]. In altre parole, a lettere uguali tra loro nel testo in chiaro corrispondono lettere uguali tra loro nel testo cifrato.
Per chiarire quanto detto basta osservare che se si cifra la parola "LIBRI" con il cifrario di Cesare in Tabella 1 si ottiene "NKDTK" da cui si evince che le due lettere I del testo in chiaro vengono cifrate con due lettere K. In realtà questa caratteristica appena descritta è uno degli elementi che favorisce il lavoro dei crittoanalisti. È una delle vulnerabilità propria di tutti i cifrari monoalfabetici, compreso lo schema crittografico di Cesare.

3. CRITTOANALISI

Successivamente, dopo aver lavorato su cifratura e decifratura dei messaggi, si propone quindi ai ragazzi un'ulteriore sfida: si mostra loro un testo cifrato, rendendo noto il fatto che è stato criptato utilizzando un cifrario di Cesare, senza comunicare la seconda riga della tabella (cioè si mantiene segreta solamente la *chiave*). A questo punto si invitano gli studenti a decrittare il messaggio.
Partendo dalle soluzioni proposte dai ragazzi si illustra il lavoro del crittoanalista. Occorre una premessa. In generale i crittoanalisti, cioè coloro che vogliono risalire al testo in chiaro senza essere i legittimi destinatari del messaggio, non sono sprovveduti. Spesso possiedono strumenti e capacità di un certo rilievo. Pertanto è opportuno chiedersi quali siano i componenti dello schema crittografico che devono rimanere segreti. Gli elementi candidati a non essere divulgati sono l'*algoritmo* e la *chiave*.
Il crittografo olandese Auguste Kerckhoffs ha evidenziato sei "desiderata" riguardanti

[11] Cfr. BERARDI, BEUTELSPACHER 1996.

gli schemi crittografici. In particolare il secondo di questi afferma che «Il faut qu'il [le système] n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi»[12], cioè la sicurezza di uno schema crittografico deve dipendere solamente dalla segretezza della chiave e non dalla segretezza dell'algoritmo di cifratura. Quindi, tornando al gioco proposto ai ragazzi, si lavora sulle strategie proposte da loro per scoprire il testo in chiaro e quindi la chiave del crittosistema, rendendo noto in anticipo l'algoritmo di cifratura utilizzato e mantenendo segreta la chiave.
Dopo un brainstorming si tabulano le risposte e si raggruppano quelle che utilizzano tecniche simili tra loro. Uno dei metodi che generalmente gli studenti propongono è il cosiddetto "attacco a forza bruta".

3.1 Attacco a forza bruta

Si definisce "attacco a forza bruta" quella procedura, utilizzata dal crittoanalista per risalire al testo in chiaro senza conoscere la chiave, che consiste nel provare tutte le possibili chiavi fino ad arrivare a quella che permette di decifrare il messaggio. Questo tipo di attacco viene chiamato anche *ricerca esaustiva.*
Questa strategia di crittoanalisi, applicata su messaggio che è stato cifrato con il crittosistema di Cesare, si attua mediante la seguente procedura. Si inizia a decifrare il messaggio supponendo che la chiave sia costituita dallo spostamento dell'alfabeto di una sola posizione cioè partendo dalla lettera B nella compilazione della seconda riga della tabella utilizzata per cifrare.

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	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

Tabella 2. Crittosistema di Cesare con chiave A --> B.

In caso di insuccesso si parte con le lettere successive cioè si prova a iniziare con la

[12] Cfr. Kerckhoffs 1883, p. 12. «È necessario che [il sistema] non richieda di essere segreto, e che possa cadere senza pericoli nelle mani del nemico».

lettera C, poi D, fino a ottenere la tabella di cifratura giusta che permette di risalire al teso in chiaro. Si prenda ad esempio il testo cifrato dell'esempio precedente, cioè si vuole risalire al testo in chiaro corrispondente alla parola EKCQ.

Partendo dalla Tabella 2, si cerca la lettera E sulla seconda riga e si nota che sopra la E si trova la lettera D. Tale lettera D è la prima lettera del testo in chiaro che si sta ricostruendo. Poi si cerca la lettera K sempre sulla seconda riga e si nota che sopra c'è la lettera J quindi il testo in chiaro inizia con DJ. Già da queste prime lettere ci si accorge che DJ non è una possibile sillaba con cui può iniziare una parola italiana.

Poiché si ipotizza che il testo in chiaro sia una parola italiana, si intuisce che la chiave appena utilizzata è da scartare. Ovviamente, se si volesse proseguire con questa chiave, si otterrebbe un testo in chiaro formato da una parola che non ha significato. Per quanto detto, quindi, si cambia la chiave prendendo in considerazione uno spostamento di due posizioni, cioè si fa partire l'alfabeto della cifratura con la lettera C (cfr. Tabella 1).

Si ripete la stessa procedura e, nel caso in cui non fosse nemmeno questa la chiave, si continua facendo partire l'alfabeto cifrante dello schema crittografico (cfr. Tabella 1) dalla lettera D. In caso di insuccesso si sposta ulteriormente di una lettera l'avvio della procedura fino ad arrivare alla chiave corretta.

Si chiede quindi ai ragazzi di stabilire il numero massimo di tentativi che possono essere necessari per scoprire la chiave. Si coglie l'occasione per ricordare loro che in generale la bontà di un *algoritmo di crittoanalisi* si misura su quello che viene definito "il caso pessimo", cioè si considera il numero massimo di tentativi che un cracker potrebbe essere costretto a compiere per scoprire il messaggio segreto. Ovviamente se nell'esempio appena visto fosse stata utilizzata la Tabella 2 come schema di cifratura, il crittoanalista potrebbe scoprire la chiave con un solo tentativo.

Quindi, qual è questo numero massimo? La risposta corretta è n-1 dove n è il numero delle lettere dell'alfabeto utilizzato. Si sottrae 1 perché non si considera la chiave che cifra ogni lettera con se stessa.

4. Il cifrario di Cesare "reso più sicuro"

Come è stato visto il cifrario di Cesare è uno schema di cifratura che può essere forzato molto velocemente, anche senza l'ausilio di strumenti tecnologici, mediante l'attacco a forza bruta. Si chiede quindi agli studenti se e come è possibile incrementare il livello di sicurezza di questo crittosistema.

Una strada che può essere percorsa, per rendere più difficile il lavoro del crittoanalista, è quella di cercare metodi che proteggano dall'attacco a forza bruta che è stato visto in precedenza. Quindi, per vanificare questo tipo di lavoro svolto dall'intruso, si possono "mescolare" le lettere che si scrivono sulla seconda riga della tabella utilizzata per la cifratura.

In questo modo, per conoscere la chiave non è più sufficiente sapere il numero di caselle dopo le quali si inizia a scrivere l'alfabeto partendo dalla lettera A nella seconda riga. Per poter cifrare i messaggi è necessario conoscere tutta la stringa di lettere che forma la seconda riga della tabella.

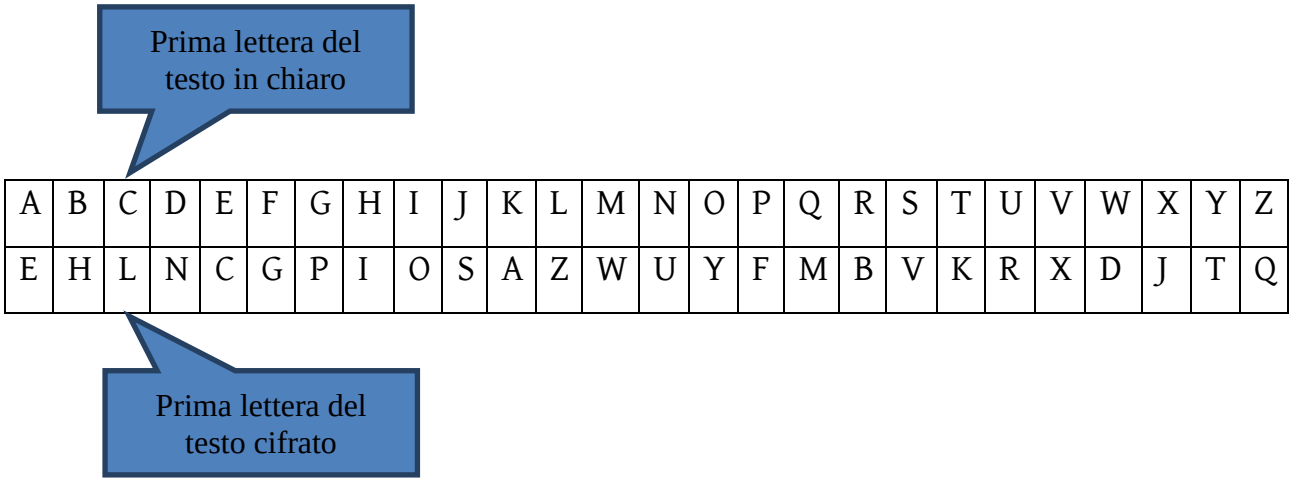


Tabella 3. Crittosistema di Cesare "reso più sicuro".

Le procedure per cifrare e decifrare i messaggi rimangono le stesse del cifrario di Cesare descritto in precedenza: ad esempio se si deve cifrare la parola "CIAO" si inizia cercando la lettera C sulla prima riga. La prima lettera del testo cifrato sarà quindi la L. Poi si cerca

la lettera I sempre sulla prima riga e quindi la lettera O sarà la seconda lettera del testo cifrato. Si prosegue in questo modo fino ad arrivare alla fine del testo in chiaro. Quindi al testo in chiaro "CIAO" corrisponde il testo cifrato "LOEY".
Il destinatario del messaggio riceve la parola cifrata e per risalire al testo in chiaro procede, come nel cifrario di Cesare "tradizionale", cercando ogni singola lettera sulla seconda riga e sostituendola, nel testo in chiaro ricostruito, con la lettera che si trova sopra la lettera appena cercata.

4.1 Attacco "a forza bruta"

L'attacco a forza bruta, come è stato visto, non richiede particolari conoscenze teoriche, basta provare tutte le chiavi possibili. Nel caso del crittosistema di Cesare "tradizionale", come abbiamo già visto, il numero di tentativi da compiere per risalire al testo in chiaro non preoccupa più di tanto il crittoanalista.
Si chiede quindi agli studenti di formulare ipotesi su quale possa essere il numero massimo di chiavi di un crittosistema di Cesare "reso più sicuro". In base a quanto detto sulla costruzione della chiave, tale numero equivale al numero di sequenze di 26 lettere che si possono ottenere utilizzando tutte le lettere dell'alfabeto senza ripetere la stessa lettera.
La domanda può essere fatta precedere da un gioco sugli *anagrammi*: ad esempio si può far calcolare agli studenti il numero di tutti i possibili anagrammi di alcune parole.
Per rispondere alla richiesta di calcolare il numero degli anagrammi dell'alfabeto si può partire ponendo agli studenti il seguente quesito: quanti sono gli anagrammi della parola APE? Dopo aver lavorato sulle strategie proposte dagli alunni si può concludere che APE, AEP, PEA, PAE, EAP, EPA sono tutti gli anagrammi richiesti.
Considerando l'ordine in cui sono scritti, peraltro ininfluente sul risultato che si sta cercando, si può notare che occorre scrivere sequenze costituite da tre lettere: per la lettera iniziale ci sono tre possibilità, per la seconda ce ne sono due e, infine, una sola per la terza lettera. Quindi tutti gli anagrammi della parola APE possono essere contati moltiplicando tra loro i numeri delle possibili scelte delle singole lettere quindi

in totale si hanno $3 \cdot 2 \cdot 1 = 6$ anagrammi. Se si volessero contare i possibili anagrammi della parola CIAO allora si dovrebbe calcolare $4 \cdot 3 \cdot 2 \cdot 1 = 24$.

In generale il numero degli anagrammi di una parola formata da n lettere, tutte diverse tra loro, è di $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$. Questo numero si indica con $n!$ e si legge "n fattoriale".

Quindi, tornando al crittosistema "reso più sicuro", gli studenti arriveranno ad affermare che le possibili chiavi sono $26! = 26 \cdot 25 \cdot 24 \cdot \ldots \cdot 2 \cdot 1$ (togliere l'unica chiave che non permuta nessuna lettera dell'alfabeto non influisce sull'ordine di grandezza del numero). Il risultato di questo prodotto, almeno in apparenza, garantisce una certa sicurezza del crittosistema. Cercare di indovinare una chiave, tra tutte queste, "sembra" impossibile.

È interessante notare che è sufficiente un semplice "rimescolamento" delle lettere che si trovano sulla seconda riga della tabella del cifrario di Cesare, come è stato fatto nella Tabella 3, per rinvigorire la sicurezza del crittosistema.

Il numero di tentativi da compiere per forzare il crittosistema rende inapplicabile l'attacco a forza bruta realizzato, come in precedenza, con il solo ausilio di un foglio di carta con cui i ragazzi provano a decifrare il messaggio. Anzi, «anche procedendo al ritmo di una chiave al secondo, per completare il controllo occorrerebbe un tempo di gran lunga superiore all'età dell'universo».[13]

Gli studenti incontreranno nel loro percorso scolastico i possibili modi per stabilire l'ordine di grandezza di questi numeri, ma, anche senza effettuare molti calcoli, si è portati a concludere che il numero che si ottiene rende impossibile la ricerca esaustiva della chiave. Occorre ricordare che in qualsiasi lavoro di crittoanalisi il *tempo* impiegato per ottenere il testo in chiaro è un elemento fondamentale. Basti pensare che le transazioni bancarie, anche di capitali ingenti, avvengono in tempi molto rapidi, quindi riuscire a forzare un crittosistema svolgendo calcoli che durano ore o giorni non ha utilità pratica.

[13] Cfr. SINGH 2001.

4.2 Attacco "statistico"

L'attacco a forza bruta, come è stato visto, pur non richiedendo particolari conoscenze teoriche, spesso non permette di ottenere la chiave in tempi ragionevoli.

Un metodo di attacco più efficiente è quello che si basa sul fatto che in generale in un testo alcune lettere compaiono più frequentemente di altre. Per questo si considera la *frequenza percentuale* (che si ottiene moltiplicando per 100 la *frequenza relativa*[14]) che è caratteristica della lingua in cui è scritto il testo in chiaro.

I numeri della Tabella 4 «rappresentano valori medi, che corrispondono solo in modo approssimativo a quelli riscontrabili in un brano specifico»[15].

Lettera	Freq	Lettera	Freq	Lettera	Freq	Lettera	Freq
A	11,74%	G	1,64%	O	9,83%	U	3,01%
B	0,92%	H	1,54%	P	3,05%	V	2,10%
C	4,50%	I	11,28%	Q	0,51%	Z	0,49%
D	3,73%	L	6,51%	R	6,37%		
E	11,79%	M	2,51%	S	4,98%		
F	0,95%	N	6,88%	T	5,62%		

Tabella 4. Frequenza percentuale delle diverse lettere in italiano (Fonte: Singh 2001).

Per quanto detto, ovviamente, esistono singole frasi in cui la percentuale delle lettere si allontana notevolmente dai valori indicati nella Tabella 1.

Quindi la crittoanalisi, oltre all'applicazione meccanica della sostituzione di lettere in base a frequenze più o meno simili, necessita anche di una buona dose di intuito.

Il lavoro del crittoanalista è, nel caso di cifrari come quello di Cesare (sia quello originale sia quello "reso più sicuro"), agevolato dal fatto che si tratta di cifrari monoalfabetici.

4.3 Esempio

Si chiede ai ragazzi di immaginare di aver intercettato il seguente messaggio segreto (di cui si sa che è scritto in lingua italiana) di cui si vuole comprendere il significato[16]:

SGTI P ZQPTBI NPMGUGTP PRQAELPG BEUUP TIAMLP VPDGETDP VEL ZQPTMI MEOVI PTNILP NIBEAMI MQI SQLILE NG RESSELP P ZQPUE UGOGME AG AVGTHELP UP MQP ASLETPMP PQBPNGP

[14] Si chiama *frequenza relativa* il numero che si ottiene dividendo la frequenza assoluta per il numero di unità statistiche.
[15] Cfr. Singh 2001.
[16] Per abbreviare il lavoro, senza comunque perdere di generalità, in questo esempio utilizziamo l'alfabeto italiano con 21 lettere.

Per fare questo serve risalire alla *chiave*. In base al *principio di Kerckhoffs*, secondo cui è segreta solamente la chiave e non la procedura utilizzata, si sa che il messaggio è stato cifrato con il crittosistema di Cesare "reso più sicuro".

Si può notare che, con un attacco a forza bruta, ci si può accorgere che non è stato utilizzato un cifrario di Cesare "convenzionale", cioè ottenuto facendo scorrere di alcune posizioni l'inizio dell'alfabeto e poi procedendo secondo l'ordine alfabetico. Quindi, sapendo che nella seconda riga della tabella cifrante le lettere sono scritte in ordine casuale, si prova con l'attacco "statistico".

Il primo passaggio da compiere è quello di contare il numero totale delle lettere che compaiono nel messaggio (nel caso dell'esempio sono 133) e il numero di occorrenze di ogni lettera per costruire una tabella delle frequenze, indicando accanto a ogni lettera la frequenza percentuale con cui tale lettera compare nel testo cifrato.

Lettera	Freq	Lettera	Freq	Lettera	Freq	Lettera	Freq
A	4,51%	G	8,27%	O	1,5%	U	4,51%
B	3%	H	0,75%	P	18,05%	V	3%
C	0%	I	7,52%	Q	6,02%	Z	2,26%
D	1,5%	L	6,77%	R	1,5%		
E	9,78%	M	6,77%	S	3,76%		
F	0%	N	3,76%	T	6,77%		

Tabella 5. Frequenze percentuali delle singole lettere nel testo intercettato.

Confrontando i valori ottenuti con quelli della Tabella 4, si può notare che la lettera del testo cifrato che ha la frequenza maggiore è la P e quindi a tale lettera dovrebbe corrispondere nel testo in chiaro una delle lettere dell'alfabeto che ha frequenza maggiore nella Tabella 1.

Non è automatico però che alla lettera P corrisponda la lettera 'e', potrebbe succedere, come vedremo in questo caso, che alla lettera P corrisponda la lettera 'a' (utilizziamo lettere maiuscole per il testo cifrato e lettere minuscole per il testo in chiaro). Nell'esempio in esame, come sarà descritto in seguito, ci si accorgerà di questo osservando le prime parole che saranno completamente in chiaro.

In generale, si procede per tentativi ed errori. Se il messaggio cifrato, come in questo caso, è formato da parole divise dagli spazi (cioè non è formato da una stringa di caratteri senza interruzione) il compito potrebbe risultare più immediato perché si possono prendere in considerazione le lettere finali di ciascuna parola cifrata sostituendole con le vocali, procedendo in ordine con le frequenze (anche se in italiano vi sono molte parole che non finiscono per vocale e il testo potrebbe contenere parole tronche). In questo caso la P è la più frequente tra le lettere finali delle parole cifrate e questo supporta l'ipotesi che alla P corrisponda una vocale, che in prima istanza dobbiamo supporre sia la 'e'. Tuttavia ciò non è necessario.

Vediamo come procedere in generale. Innanzitutto, ordiniamo le tabelle 4 e 5 in modo decrescente rispetto alla frequenza e mettiamole a confronto.

Tabella 4 ordinata

E	11,79
A	11,74
I	11,28
O	9,83
N	6,88
L	6,51
R	6,37
T	5,62
S	4,98
C	4,5
D	3,73
P	3,05
U	3,01
M	2,51
V	2,1
G	1,64
H	1,54
F	0,95
B	0,92
Q	0,51
Z	0,49

Tabella 5 ordinata

P	18,05
E	9,78
G	8,27
I	7,52
L	6,77
M	6,77
T	6,77
Q	6,02
A	4,51
U	4,51
N	3,76
S	3,76
B	3
V	3
Z	2,26
D	1,5
O	1,5
R	1,5
H	0,75
C	0
F	0

Proviamo a eseguire, l'una di seguito all'altra, le seguenti sostituzioni nel testo cifrato (nella prima si sostituisce la lettera P con la lettera 'e'), ottenendo testi parzialmente modificati (ricordiamo che nel testo da decifrare la parte modificata è scritta con lettere minuscole mentre la parte restante è scritta con lettere maiuscole).

- P con e:

SGTI e ZQeTBI NeMGUGTe eRQAELeG BEUUe TIAMLe VeDGETDe VEL ZQeTMI MEOVI eTNILe NIBEAMI MQI SQLILE NG RESSELe e ZQeUE UGOGME AG AVGTHELe Ue MQe ASLETeMe eQBeNGe

- E con a:

SGTI e ZQeTBI NeMGUGTe eRQAaLeG BaUUe TIAMLe VeDGaTDe VaL ZQeTMI MaOVI eTNILe NIBaAMI MQI SQLILa NG RaSSaLe e ZQeUa UGOGMa AG AVGTHaLe Ue MQe ASLaTeMe eQBeNGe

- G con i:

SiTI e ZQeTBI NeMiUiTe eRQAaLei BaUUe TIAMLe VeDiaTDe VaL ZQeTMI MaOVI eTNILe NIBaAMI MQI SQLILa Ni RaSSaLe e ZQeUa UiOiMa Ai AViTHaLe Ue MQe ASLaTeMe eQBeNie

- I con o:

SiTo e ZQeTBo NeMiUiTe eRQAaLei BaUUe ToAMLe VeDiaTDe VaL ZQeTMo MaOVo eTNoLe NoBaAMo MQo SQLoLa Ni RaSSaLe e ZQeUa UiOiMa Ai AViTHaLe Ue MQe ASLaTeMe eQBeNie

- L con n

SiTo e ZQeTBo NeMiUiTe eRQAanei BaUUe ToAMne VeDiaTDe Van ZQeTMo MaOVo eTNone NoBaAMo MQo SQnona Ni RaSSane e ZQeUa UiOiMa Ai AViTHane Ue MQe ASnaTeMe eQBeNie

- M con l:

SiTo e ZQeTBo NeliUiTe eRQAanei BaUUe ToAlne VeDiaTDe Van ZQeTlo laOVo eTNone NoBaAlo lQo SQnona Ni RaSSane e ZQeUa UiOila Ai AViTHane Ue lQe ASnaTele eQBeNie

- T con r:

Siro e ZQerBo NeliUire eRQAanei BaUUe roAlne VeDiarDe Van ZQerlo laOVo erNone NoBaAlo lQo SQnona Ni RaSSane e ZQeUa UiOila Ai AVirHane Ue lQe ASnarele eQBeNie

- Q con t:

Siro e ZterBo NeliUire eRtAanei BaUUe roAlne VeDiarDe Van Zterlo laOVo erNone NoBaAlo lto Stnona Ni RaSSane e ZteUa UiOila Ai AVirHane Ue lte ASnarele etBeNie

- A con s:

Siro e ZterBo NeliUire eRtsanei BaUUe roslne VeDiarDe Van Zterlo laOVo erNone NoBaslo lto Stnona Ni RaSSane e ZteUa UiOila si sVirHane Ue lte sSnarele etBeNie

- U con c:

Siro e ZterBo Nelicire eRtsanei Bacce roslne VeDiarDe Van Zterlo laOVo erNone NoBaslo lto Stnona Ni RaSSane e Zteca ciOila si sVirHane ce lte sSnarele etBeNie

- N con d:

Siro e ZterBo delicire eRtsanei Bacce roslne VeDiarDe Van Zterlo laOVo erdone doBaslo lto Stnona di RaSSane e Zteca ciOila si sVirHane ce lte sSnarele etBedie

A questo punto, però, si può notare che alcune parole sono formate da tutte lettere modificate, quali ad esempio "delicire". Purtroppo quest'ultima non è una parola appartenente al vocabolario italiano, quindi occorre rivedere le sostituzioni effettuate.
Dalla Tabella 4 si evince che la differenza tra le frequenze delle lettere 'a' ed 'e' è 0,05% quindi nel prossimo tentativo di decrittazione si prova a sostituire la P del testo cifrato con la 'a' e la E del testo cifrato con la 'e'.
Si procede quindi con nuove sostituzioni.

- P con a:

SGTI a ZQaTBI NaMGUGTa aRQAELaG BEUUa TIAMLa VaDGETDa VEL ZQaTMI MEOVI aTNILa NIBEAMI MQI SQLILE NG RESSELa a ZQaUE UGOGME AG AVGTHELa Ua MQa ASLETaMa aQBaNGa

- E con e:

SGTI a ZQaTBI NaMGUGTa aRQAeLaG BeUUa TIAMLa VaDGeTDa VeL ZQaTMI MeOVI aTNILa NIBeAMI MQI SQLILe NG ReSSeLa a ZQaUe UGOGMe AG AVGTHeLa Ua MQa ASLeTaMa aQBaNGa

Poi come già fatto in precedenza, si procede con le seguenti sostituzioni.

- G con i:

SiTI a ZQaTBI NaMiUiTa aRQAeLai BeUUa TIAMLa VaDieTDa VeL ZQaTMI MeOVI aTNILa NIBeAMI MQI SQLILe Ni ReSSeLa a ZQaUe UiOiMe Ai AViTHeLa Ua MQa ASLeTaMa aQBaNia

- I con o:

SiTo a ZQaTBo NaMiUiTa aRQAeLai BeUUa ToAMLa VaDieTDa VeL ZQaTMo MeOVo aTNoLa NoBeAMo MQo SQLoLe Ni ReSSeLa a ZQaUe UiOiMe Ai AViTHeLa Ua MQa ASLeTaMa aQBaNia

A questo punto si può notare che le lettere L, M e T hanno le stesse frequenze nel testo cifrato.

Si prova a sostituire L con 'n' e poi si procede con le sostituzioni successive.

- L con n:

SiTo a ZQaTBo NaMiUiTa aRQAenai BeUUa ToAMna VaDieTDa Ven ZQaTMo MeOVo aTNona NoBeAMo MQo SQnone Ni ReSSena a ZQaUe UiOiMe Ai AViTHena Ua MQa ASneTaMa aQBaNia

- M con l:

SiTo a ZQaTBo NaliUiTa aRQAenai BeUUa ToAlna VaDieTDa Ven ZQaTlo leOVo aTNona NoBeAlo lQo SQnone Ni ReSSena a ZQaUe UiOile Ai AViTHena Ua lQa ASneTala aQBaNia

- T con r:

Siro a ZQarBo NaliUira aRQAenai BeUUa roAlna VaDierDa Ven ZQarlo leOVo arNona NoBeAlo lQo SQnone Ni ReSSena a ZQaUe UiOile Ai AVirHena Ua lQa ASnerala aQBaNia

- Q con t:

Siro a ZtarBo NaliUira aRtAenai BeUUa roAlna VaDierDa Ven Ztarlo leOVo arNona NoBeAlo lto Stnone Ni ReSSena a ZtaUe UiOile Ai AVirHena Ua lta ASnerala atBaNia

- A con s:

Siro a ZtarBo NaliUira aRtsenai BeUUa roslna VaDierDa Ven Ztarlo leOVo arNona NoBeslo lto Stnone Ni ReSSena a ZtaUe UiOile si sVirHena Ua lta sSnerala atBaNia

- U con c:

Siro a ZtarBo Nalicira aRtsenai Becca roslna VaDierDa Ven Ztarlo leOVo arNona NoBeslo lto Stnone Ni ReSSena a Ztace ciOile si sVirHena ca lta sSnerala atBaNia

Anche qui si vede che ci sono parole che, sebbene non ancora completamente decifrate, non sembrano appartenere al vocabolario italiano.

Quindi si riparte dal punto in cui è stato osservato che le lettere L, M, T hanno la stessa frequenza e si procede con le seguenti sostituzioni: L con 'r' e T con 'n'.

- T con n:

Sino a ZQanBo NaMiUina aRQAeLai BeUUa noAMLa VaDienDa VeL ZQanMo MeOVo anNoLa NoBeAMo MQo SQLoLe Ni ReSSeLa a ZQaUe UiOiMe Ai AVinHeLa Ua MQa ASLenaMa aQBaNia

- M con l:

Sino a ZQanBo NaliUina aRQAeLai BeUUa noAlLa VaDienDa VeL ZQanlo leOVo anNoLa NoBeAlo lQo SQLoLe Ni ReSSeLa a ZQaUe UiOile Ai AVinHeLa Ua lQa ASLenala aQBaNia

- L con r:

Sino a ZQanBo NaliUina aRQAerai BeUUa noAlra VaDienDa Ver ZQanlo leOVo anNora NoBeAlo lQo SQrore Ni ReSSera a ZQaUe UiOile Ai AVinHera Ua lQa ASrenala aQBaNia

- Q con t:

Sino a ZtanBo NaliUina aRtAerai BeUUa noAlra VaDienDa Ver Ztanlo leOVo anNora NoBeAlo lto Strore Ni ReSSera a ZtaUe UiOile Ai AVinHera Ua lta ASrenala atBaNia

- A con s:

Sino a ZtanBo NaliUina aRtserai BeUUa noslra VaDienDa Ver Ztanlo leOVo anNora NoBeslo lto Strore Ni ReSSera a ZtaUe UiOile si sVinHera Ua lta sSrenala atBaNia

Qui l'intuizione può aiutarci osservando, ad esempio, la settima parola. Si ipotizza sia "nostra" e non "noslra" e quindi si sostituisce M con 't' e non più con 'l'.

Sostituendo M con 't' si ottiene la seguente frase:

Sino a ZQanBo NatiUina aRQAeLai BeUUa noAtLa VaDienDa VeL ZQanto teOVo anNoLa NoBeAto tQo SQLoLe Ni ReSSeLa a ZQaUe UiOite Ai AVinHeLa Ua tQa ASLenata aQBaNia

Poi, si può procede con altre sostituzioni.

- L con r

Sino a ZQanBo NatiUina aRQAerai BeUUa noAtra VaDienDa Ver ZQanto teOVo anNora NoBeAto tQo SQrore Ni ReSSera a ZQaUe UiOite Ai AVinHera Ua tQa ASrenata aQBaNia

Adesso Q non si può più sostituire con 't' poiché abbiamo già sostituito M con 't'. Riprendiamo il testo ottenuto dopo la sostituzione di L con 'r' e sostituiamo A con 's' e otteniamo la frase seguente:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra VaDienDa Ver ZQanto teOVo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si sVinHera Ua tQa sSrenata aQBaNia

Adesso l'intuizione diventa sempre più efficace e tale da sostituire, almeno parzialmente, l'analisi delle frequenze. Le parole VaDienDa e Ver suggeriscono di sostituire V con 'p' e D con 'z'.

- V con p:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra paDienDa per ZQanto teOpo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si spinHera Ua tQa sSrenata aQBaNia

- D con z:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra pazienza per ZQanto teOpo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si spinHera Ua tQa sSrenata aQBaNia

Un ulteriore suggerimento, ad esempio, è fornito dall'undicesima parola: teOpo ci fa pensare a sostituire O con 'm'.

- O con m

Sino a ZQanBo NatiUina aRQserai BeUUa nostra pazienza per ZQanto tempo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe Uimite si spinHera Ua tQa sSrenata aQBaNia

La dodicesima parola suggerisce di sostituire N con 'c':

Sino a ZQanBo catiUina aRQserai BeUUa nostra pazienza per ZQanto tempo ancora coBesto tQo SQrore ci ReSSera a ZQaUe Uimite si spinHera Ua tQa sSrenata aQBacia

La sesta parola suggerisce di sostituire B con 'd' e U con 'l':

Sino a ZQando catilina aRQserai della nostra pazienza per ZQanto tempo ancora codesto tQo SQrore ci ReSSera a ZQale limite si spinHera la tQa sSrenata aQdacia

Poi, per completare, dalla terza parola si intuisce che si possono effettuare le seguenti sostituzioni: Q con 'u' e Z con 'q' e si ottiene la frase seguente:

Sino a quando catilina aRuserai della nostra pazienza per quanto tempo ancora codesto tuo Surore ci ReSSera a quale limite si spinHera la tua sSrenata audacia.

Adesso la frase è praticamente in chiaro. Solo per completare lo schema di cifratura appena scoperto possiamo sostituire S con 'f', R con 'b' e H con 'g'. Si ottiene quindi il testo in chiaro:

Fino a quando Catilina abuserai della nostra pazienza per quanto tempo ancora codesto tuo furore ci beffera a quale limite si spingera la tua sfrenata audacia.[17]

5. CONCLUSIONE

Il crittosistema di Cesare viene di solito presentato come la prima applicazione della matematica alla crittografia. Le successive modifiche che abbiamo visto hanno permesso di ottenere schemi crittografici più sicuri e contemporaneamente hanno dimostrato che serve qualche conoscenza ulteriore in ambito matematico per essere in grado di forzare il crittosistema in tempi "ragionevoli".

Nel secondo crittosistema che abbiamo incontrato, oltre ad aver descritto una possibile strategia di decrittazione facilmente inseribile in un percorso didattico, è stato mostrato che sono sufficienti anche pochi concetti base di statistica descrittiva per agevolare il lavoro del crittoanalista, permettendogli di risalire al testo in chiaro in tempi brevi.

Questa necessità di forzare rapidamente un crittosistema è il compito principale di chi vuole conoscere un segreto senza essere in possesso della chiave: tale attività, come abbiamo viso è agevolata dalla conoscenza della matematica. Quindi, con l'inserimento opportuno della crittografia in percorsi di apprendimento si permette anche di superare quella visione della matematica secondo cui essa è qualcosa di astratto in cui è difficile poter vedere un'applicazione pratica.

BIBLIOGRAFIA

BERARDI L., BEUTELSPACHER A.
1996, *Crittologia. Come proteggere le informazioni riservate*, Milano, Franco Angeli.

[17] Cfr. CICERONE 1962 (frase tradotta e considerata senza punteggiatura né lettere accentate per agevolare il lavoro di decrittazione).

BORRELLI M., FIORETTO A., SGARRO A., ZUCCHERI L.
2002, *Cryptography and Statistics: A Didactical Project*, Proceedings of the 2nd International Conference on the Teaching of Mathematics (at the undergraduate level), Hernissos, Crete. 1-6 luglio 2002, IRAKLIO, CRETE. John Wiley & Sons Inc., pp. 1-6.

CICERONE M. T.
1962, *Orationes in Catilinam*, Roma, A. Signorelli.

CAZZOLA M., GRAZIAN V.
2021, *Giochi di Crittografia elementare per la scuola primaria*, in: R. BONINO, D. MAROCCHI, M. RINAUDO, M. SERIO (a cura di), *Apprendimento laboratoriale in Matematica e Fisica in presenza e a distanza* (Torino, 11-12-13 ottobre 2021 – online), Torino, Università degli Studi di Torino, pp. 350-357.

D'AMORE B.
1999, *Elementi di Didattica della Matematica*, Bologna, Pitagora Editrice.

DI MARTINO P., ZAN R.
2020, *Problemi per crescere Matematica senza paura*, Firenze, Giunti Scuola.

KERCKHOFFS A.
1883, «La cryptographie militaire», *Journal des sciences militaires*, Vol IX, pp. 5-38.

MARCEDDU M. C.
2002, *Il gioco dell'agente segreto - II Parte*, in: «Matematica dei ragazzi: Scambi di esperienze tra coetanei – antologia 2000 – 2002», Trieste, EUT – Edizioni Università Trieste, pp. 36-38.

PERES E.
1986, *Giochi matematici*, Roma, Editori Riuniti.

SABENA C., FERRI F., MARTIGNONE F., ROBOTTI E.
2019, *Insegnare e apprendere matematica*, Milano, Mondadori.

SINGH S.
2001, *Codici e segreti*, Milano, BUR.

SUETONIUS TRANQUILLUS G.
2016, *De vita Cæsarum*, Libro 1, Capitolo LVI, Chieti, Vestigium.

ZUCCHERI L.
1992, «Crittografia e statistica nella scuola elementare», *Insegnamento della matematica e delle scienze integrate*, 15, n. 1, pp. 19-38.