CODICI SEGRETI: UN VIAGGIO NELLA CRITTOGRAFIA

MARIA CONCETTA BROCATO, AGOSTINO DOVIER

1.1 Introduzione

Il laboratorio PLS *"Codici segreti: un viaggio nella crittografia"* è nato con due finalità: da un lato si desiderava introdurre gli studenti partecipanti all'affascinante mondo della crittografia, dall'altro si intendeva suggerire l'utilizzo degli algoritmi di cifrazione/decifrazione/decrittazione come strumento veicolare per la pratica della codifica di algoritmi che necessitano di leggere/scrivere su file e operare con i vettori. Avendo l'obiettivo (più apparente che reale) di essere in grado di forzare gli attuali codici segreti (sulla scia dello scandalo *WikiLeaks*, accaduto nei mesi precedenti alla prima edizione del laboratorio), di indovinare la password per l'accesso alle reti wireless oppure (più concretamente) di indovinare un testo cifrato dai propri compagni con una chiave ignota, i ragazzi hanno avuto la possibilità di affrontare attività talvolta considerate *noiose* con rinnovato entusiasmo.

Nella fase iniziale di organizzazione del laboratorio si pensava di porre come prerequisito la conoscenza dei principi di base della programmazione e dei costrutti principali di almeno un linguaggio di programmazione; in un secondo momento questa richiesta è stata allentata per permettere la partecipazione anche a studenti di istituti in cui lo studio dell'informatica non viene affrontato. In tal caso ci si è concentrati sugli aspetti fondanti della teoria dei numeri e dei campi finiti in particolare, utili per affrontare il problema della fattorizzazione, la cui risoluzione efficiente permetterebbe, di fatto, di aprire nuovi orizzonti crittografici.

1.2 Inquadramento storico

Il desiderio e la necessità di trasferire l'informazione da mittente a destinatario in modo tale che un eventuale malintenzionato che ne fosse venuto in possesso non potesse comprenderla permea la storia delle comunicazioni umane. Nella Bibbia si legge di come il profeta Daniele fosse in grado di decrittare i messaggi inviati da Dio a Baldassarre, nonché si trovano diverse istanze di semplici codici per nascondere alcuni nomi (per esempio Babel veniva scritta come *Scheschach*, utilizzando un codice noto come *AtBash*). L'impiego dei cifrari in campo politico e militare fu probabilmente introdotto da Giulio Cesare (101–44 a.C.), il quale codificava l'informazione sostituendo ad ogni lettera la lettera che la seguiva di tre posizioni nell'alfabeto (per esempio, la A diviene D, in breve $A \rightsquigarrow D$), ricominciando dall'inizio quando l'alfabeto termina (si veda Figura 1.1). Conoscendo il tipo di codifica, ma non la chiave (in questo caso la D) una spia poteva decrittare il messaggio provando 26 possibili chiavi (anzi 25: la A co-

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
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

P I A N O L A U R E E S C I E N T I F I C H E
↓ ↓
S L D Q R O D X U H H V F L H Q W L I L F K H

Figura 1.1: Il codice monoalfabetico di Cesare: la tabella che descrive la sostituzione ciclica ed un esempio di codifica (per semplicità abbiamo usato il moderno alfabeto inglese a 26 lettere).

P I A N O L A U R E E S C I E N T I F I C H E
B A C C A B A C C A B A C C A B A C C A B A C
↓ ↓
Q I C P O M A W T E F S E K E O T K H I D H G

Figura 1.2: Il codice di Vigènere: esempio di codifica con la parola chiave *BACCA* (la sostituzione è determinata sempre utilizzando l'alfabeto inglese).

me chiave non è molto interessante). Una estensione apparentemente più complessa è quella che si ottiene sostituendo le lettere sulla base di una qualunque permutazione. Il numero di possibilità ora cresce notevolmente ($26! \approx 4 \cdot 10^{26}$), ma mediante l'impiego di semplici tecniche di statistica linguistica risulta possibile forzare il codice anche utilizzando carta e penna (si legga il bellissimo racconto [88]).

Un enorme balzo in avanti nella storia della crittografia fu la definizione del cifrario polialfabetico o di Vigènere (Blaise de Vigènere, 1523–1596), che venne definito *l'indecifrabile* (e, per un po', lo rimase). Esso usa una parola chiave per commutare tra diversi cifrari "alla Cesare". Ad esempio, se la parola chiave è *BACCA* la prima lettera del testo viene sostituita dalla seguente ($A \rightsquigarrow B$), la seconda e la quinta vengono lasciate immutate ($A \rightsquigarrow A$), la terza e la quarta vengono sostituite dalla lettera che la segue di due posizioni ($A \rightsquigarrow C$), dalla sesta si ricomincia come per la prima e si procede ciclicamente (si veda Figura 1.2). Mentre un cifrario monoalfabetico può essere forzato tentando 26 casi (usiamo l'alfabeto inglese), il cifrario alla Vigènere richiede un numero di tentativi pari a 26^n dove n è la lunghezza della chiave (anch'essa ignota al malintenzionato). Se la chiave è "corta" rispetto al testo vi sono delle tecniche algoritmico-statistiche escogitate da Babbage e formalizzate da Kasiski nel diciannovesimo secolo (si legga il recente racconto [3]) che permettono, con alcuni tentativi, di forzare tale codice. Tuttavia tali tecniche sono inefficaci qualora la chiave sia piuttosto lunga rispetto al testo.

Nei primi anni del ventesimo secolo l'ingegnere tedesco Arthur

Figura 1.3: L'Enigma (Fonte: www.copyright-free-images.com/)

Scherbius brevettò una macchina elettromeccanica per la cifratura dei messaggi: l'*ENIGMA* (Figura 1.3). Essa, mediante l'utilizzo di 3 o 4 *rotori* (interscambiabili) e del modo odometrico (come nei contachilometri meccanici) riusciva a simulare un codice alla Vigènere con chiave lunga 26^3 o 26^4 (a seconda del numero di rotori), rendendo vano un attacco basato sulla statistica. L'ENIGMA fu impiegato massivamente nella Seconda guerra mondiale: solo il gran lavoro del controspionaggio inglese, nel quale spiccava la figura di Alan Mathison Turing (1912–1954), padre dell'Informatica intesa come scienza, riuscì a forzarlo, sfruttando dei risultati combinatorici di valenti matematici polacchi e la costruzione di simulatori elettromeccaninci (le bombe) prima e del primo calcolatore elettronico (colossus) poi, contribuendo così ad invertire in modo deciso le sorti della guerra (e in ogni caso a salvare migliaia di vite umane).

Portando la situazione al limite, ovvero *utilizzando* una chiave

- lunga quanto il testo
- generata casualmente,

un codice di questo tipo è dimostrabilmente indecifrabile (*one time pad* o cifrario perfetto: Gilbert Vernam 1890–1960; Figura 1.4). Questa tecnica, apparentemente impraticabile a causa del problema della condivisione della chiave, fu usata a lungo durante la Guerra fredda (per la famosa linea rossa Washington-Mosca) e ancor più recentemente dal gruppo di spie cubane denominato *Wasp (Red Avispa)*, i cui messaggi, numerici, inviati usando una normale radio, furono intercettati (ma non decrittati) dagli Stati Uniti nel 1998.

Negli anni '70 del secolo scorso, l'inizio della diffusione dei calcolatori e delle prime reti che li interconnettevano richiese la progettazione di una nuova modalità di crittografia. I messaggi erano dei files (in ultima analisi, binari) e gli algoritmi di cifratura e decifratura dovevano essere implementati da un programma al calcolatore. Ovviamente non si può assumere che la spia non venga a prima o poi

Alfabeto:	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 ♣ ♢	♡ ♠ ♭ ♯
Codifica digitale:	0 1 2 3	4 5 6 7	8 9 10 11	12 13 14 15	16 17 18 19	20 21 22 23	24 25 26 27	28 29 30 31

Testo in chiaro:	I	N	F	O	R	M	A	T	I	C	A
Digitalizzazione:	01000	01101	00101	01110	10001	01100	00000	10011	01000	00010	00000
Chiave:	00111	00101	00101	00011	11111	00111	00000	00111	11001	00110	00100
Somma in $\mathbb{Z}_2$:	01111	01000	00000	01101	01110	01011	00000	10100	10001	00100	00100
Testo cifrato:	P	I	A	N	O	L	A	U	R	E	E

Figura 1.4: Crittografia digitale: ogni lettera viene rappresentata da una sequenza di bits (numeri da 0 a 31 espressi in binario). In questo esempio usiamo 5 bits per l'alfabeto inglese esteso con gli altri 6 simboli utilizzati. La chiave è una sequenza di 0 e 1 e viene sommata bit per bit al testo in chiaro (ove $0+0=1+1=0$ e $0+1=1+0=1$). Con la sequenza riportata, il testo in chiaro "INFORMATICA" viene cifrato nel testo in cifra "PIANOLAUREE". Se la chiave è una sequenza veramente casuale (ad esempio generata da una moneta) e della stessa lunghezza del testo, il codice è dimostrabilmente indecifrabile.

a conoscenza di tale programma. Si pensò pertanto di progettare un algoritmo noto a tutti, eliminando questo tipo di incertezza: solo la chiave non dev'essere nota alla spia. L'algoritmo di cifrazione deve mascherare l'informazione al punto tale che la spia non abbia altre soluzioni che provare tutte le possibili chiavi. Fu quindi approvato il cifrario DES (*Data Encryption Standard*—Horst Feistel e il suo team all'IBM), che si basava su una chiave a 56 bits. In effetti, nei 20 anni circa in cui fu impiegato nessuno trovò delle scorciatoie per forzarlo. Fu forzato nel 1996 per la prima volta grazie ad un algoritmo che sfruttava internet per distribuire le chiavi da testare in tutti i PC nella rete che aderivano al progetto. Una chiave di 56 bit necessita, nel caso peggiore, di $2^{56} \approx 10^{17}$ tentativi; se riuscissimo a fare un tentativo in 100 ns (10^{-7}s), saremmo in grado di provare tutte le chiavi in 10^{10} secondi, ovvero in circa 115.000 giorni. Se 100.000 PC partecipano al progetto, in poco più di un giorno saremmo sicuri di forzare il DES.

Il DES fu sostituito da altri algoritmi ed, in particolare, il suo successore "ufficiale" è l'*Advanced Encryption Standard* (AES—Vincent Rijmen e Joan Daemen) che può lavorare con chiavi da 128, 192, e 256 bits e, allo stato attuale, non è attaccabile con metodi esaustivi.

DES e AES sono entrambi algoritmi a *chiave privata*, ovvero è prevista la condivisione (da farsi ogni tanto, quando ci si vede di persona) di una chiave, cosí come in tutta la storia della crittografia accennata sopra. Alla fine degli anni '70 nasce dall'informatica un'idea innovativa: la *crittografia a chiave pubblica* (Whitfield Diffie e Martin Hellman). L'idea è che ognuno si genera una chiave divisa in due pezzi: una parte privata che non viene mai condivisa con altri e una parte pubblica che ognuno può vedere (andando sulla pagina web o sul caro vecchio elenco del telefono). Ad esempio per inviare un messaggio ad un'amico lo cifriamo usando la sua chiave pubblica. Qui c'è l'aspetto geniale. La decifrazione di quel messaggio possedendo la chiave privata è un'impresa algoritmicamente semplice. Per contro,

```
read(n);
i = 2;
while (i ≤ √n){
   if (n mod i = 0)
      then return i;
      else i = i + 1;
}
```

Figura 1.5: Un semplice algoritmo che permette (avendo molto tempo a disposizione) la decrittazione dell'RSA. Dato un numero intero $n \geq 2$ restituisce il più piccolo divisore non unitario di n. La funzione *mod* restituisce il resto della divisione tra numeri interi. Nel caso peggiore sono necessarie circa $\sqrt{n}$ divisioni per trovare il numero cercato (o stabilire se il numero è primo). Si osservi che se il numero n consta di 100 cifre allora il numero di operazioni è dell'ordine di $\sqrt{10^{100}} = 10^{50}$. Se ogni operazione si potesse fare in 10^{-10} secondi, sarebbero necessari 10^{40} secondi, ovvero $3 \cdot 10^{32}$ anni.

la decrittazione di quel messaggio pur possedendo la chiave pubblica del destinatario, è un'impresa teoricamente possibile, ma che richiede un numero di tentativi e di conseguenza tempi di esecuzione non praticabili. L'idea fu da principio messa da parte, in quanto non pareva esistere una sua implementazione, finché Ronald Rivest, Adi Shamir e Leonard Adleman proposero una tecnica di un'eleganza disarmante che la rendeva possibile. Il cifrario RSA (dai cognomi degli autori) permette di realizzare un esempio di crittografia a chiave pubblica. La cosa interessante è che per poterlo forzare, sarebbe sufficiente escogitare una tecnica efficiente per trovare i due numeri (entrambi primi) che moltiplicati tra loro forniscono un numero (di qualche centinaio di cifre) che costituisce la chiave pubblica. Finora nessuno ha dimostrato che un tale algoritmo non esiste, anche se nessuno è riuscito ancora a realizzarlo.

Le idee della crittografia informatica a chiave privata (AES) e a chiave pubblica (RSA) confluiscono nel protocollo *Pretty Good Privacy* (PGP): la crittografia a chiave pubblica (che necessita di tempi maggiori per la codifica) viene impiegata per passarsi la chiave privata per l'AES, prendendo il meglio dai due sistemi (P. H. Zimmemann).

Non entreremo qui nell'affascinante mondo della crittografia quantistica, che potrebbe costituire il futuro di quest'arte; in ogni modo, per un trattato esauriente (e divertente) su tutto questo materiale, si suggerisce la lettura di [100, 102].

1.3 Descrizione

Il laboratorio si è svolto per 3 anni e c'è stata qualche leggera variazione tra scuola e scuola. L'organizzazione generale prevedeva dapprima due lezioni di due ore (talvolta nelle scuole, altre volte presso l'Università di Udine), in cui il materiale storico è stato pre-

sentato utilizzando illustrazioni provenienti da diversi testi e dalla rete. Nella presentazione ci si è soffermati a lungo sulle tecniche per la decrittazione del Vigènere, illustrando i risultati dell'esecuzione di alcuni programmini che, in seguito, gli studenti avrebbero dovuto comprendere e re-implementare in tutto o in parte. Nella presentazione dell'Enigma è stato illustrato un simulatore (ce ne sono diversi disponibili on-line, sia per PC che per dispositivi mobili), che ha catturato l'attenzione dei ragazzi. In qualche caso è stato anche citato il film "Enigma", tratto dal libro [**Enigma**], che poi è stato proiettato nelle scuole.

I docenti delle scuole hanno successivamente dedicato 10 ore per aiutare i ragazzi a lavorare sul materiale. L'obiettivo principale era di scrivere i programmi per forzare il Vigènere; tali programmi erano stati anche forniti prima ai docenti con alcune spiegazioni tecniche circa il loro funzionamento e la correttezza delle ipotesi statistiche su cui si basavano. I programmi sono stati scritti nei linguaggi Prolog, Visual Basic, Pascal e C. In questa fase, tuttavia, è stata data ampia libertà e di conseguenza sono state realizzate attività diverse.

In ogni caso, dopo l'attività "interna" alle scuole si è organizzato un incontro finale per la relazione del lavoro svolto e qualche considerazione di chiusura.

Entriamo ora più in dettaglio su come è stata realizzata l'attività all'interno dell'ISIS A. Malignani di Udine.

1.4 La voce della scuola

Lo scopo principale del Piano Lauree Scientifiche è quello di sperimentare azioni che rafforzino i rapporti tra la Scuola Secondaria e l'Università in ottica orientante per gli studenti; in questo contesto di prospettiva verso il futuro e di scoperta, la Scuola Secondaria può trovare spazi per svolgere azioni didattiche innovative nel metodo, negli strumenti e negli scenari. L'attività nei laboratori PLS permette a studenti e docenti una flessibilità di ruoli; il docente della classe può assumere un ruolo di guida, di "tutor" del percorso di esplorazione e orientamento, può affiancare, supportare ed accompagnare con efficacia le attività dei propri studenti, in quanto agisce in stretta collaborazione con il docente referente dell'Università e con il suo supporto. In questo modo l'insegnante può scendere dalla cattedra e sperimentare in prima persona.

In quest'ottica di sperimentazione sinergica tra docente di classe e referente si è svolta l'attività del laboratorio "Codici Segreti: Un viaggio nella crittografia" all'interno dell'ISIS A. Malignani di Udine, precisamente nelle classi III TELA dell'anno 2012/2013 e 2013/2014 unite ad alcuni studenti della prima delle due, diventata IV TELA nel successivo anno 2013/2014. Il percorso di esplorazione biennale ha permesso di maturare e svolgere, nel secondo anno, oltre alle attività pianificate l'anno precedente, anche alcune attività verticali e trasversali tra le due classi, con azioni di "peer education" tra studenti:

- in presenza (in orario extra-scolastico),
- virtualmente, sfruttando le moderne possibilità offerte dalla tecnologia informatica: l'utilizzo della piattaforma didattica Moodle e di ambienti cooperativi, quali wiki, forum e bacheche virtuali.

Le attività seminariali di introduzione alla crittografia, svolte nel 2012/2013 dal prof. A. Dovier sono state precedute dalla visione del film "Enigma"ed affiancate da attività di "*live twitting*" con hashtag identificati in maniera collaborativa dagli studenti (#malignaniUd #PLScrittografia). Tale attività ha permesso il coinvolgimento degli studenti, che da passivi ascoltatori sono diventati parte attiva. Nello stesso anno, per svolgere la successiva attività di laboratorio interna con il docente della scuola, i gruppi classe sono stati suddivisi in team per aree di interesse:

- studio e codifica di algoritmi crittografici in C,
- progettazione e pubblicazione di un sito HTML/CSS che descrivesse le attività del laboratorio,
- creazione della documentazione per la relazione finale,
- editing della presentazione delle attività utilizzando modalità collaborative (utilizzando il software Prezi).

Tutti i tweet raccolti dagli studenti durante le conferenze e le attività, i contenuti, gli algoritmi, le immagini, le slide ed i concetti ritenuti "chiave" dagli studenti sono stati raccolti e pubblicati nel sito descritto che è stato fruito, nel secondo anno, dalla successiva classe terza per svolgere l'attività didattica interna alla scuola. Nelle Figure 1.6–1.8 riportiamo alcuni estratti dal contenuto in rete e i riferimenti per accedervi.

Nell'annualità successiva gli studenti della classe quarta che avevano già svolto il modulo PLS hanno avuto la possibilità di essere dei *referenti esperti* per la nuova classe che doveva iniziare il percorso; nel contempo, essi hanno approfondito il programma disciplinare dell'anno in corso attraverso il miglioramento del foglio di stile e della veste grafica del sito, l'introduzione di pagine dinamiche e l'interazione con una base dati normalizzata. La classe terza del 2013/2014, oltre a svolgere le attività di *live twitting* e in team di interesse (similmente all'anno precedente) ha sperimentato la realizzazione di uno streaming video dei due incontri seminariali con il docente referente dell'Università, con la contestuale pubblicazione del video su un canale YouTube. Durante le attività interne alla scuola gli allievi sono stati il motore;nil compito dell'insegnante, in questo contesto, è stato quello di guida per far emergere le potenzialità e gli interessi degli studenti, in stretta collaborazione con il docente referente dell'Università.

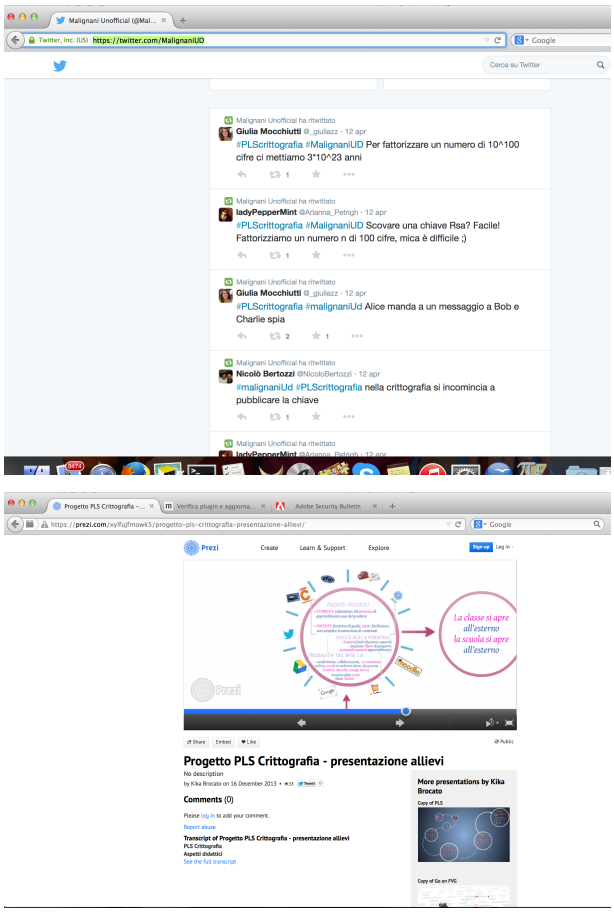

Figura 1.6: Esempi di tweet scritti dagli studenti durante la presentazione e loro visualizzazione web: `https://twitter.com/MalignaniUD` (in alto), una pagina della presentazione dell'attività da parte degli allievi usando Prezi: `https://prezi.com/xylfujfmowk5/progetto-pls-crittografia-presentazione-allievi/` (in basso)

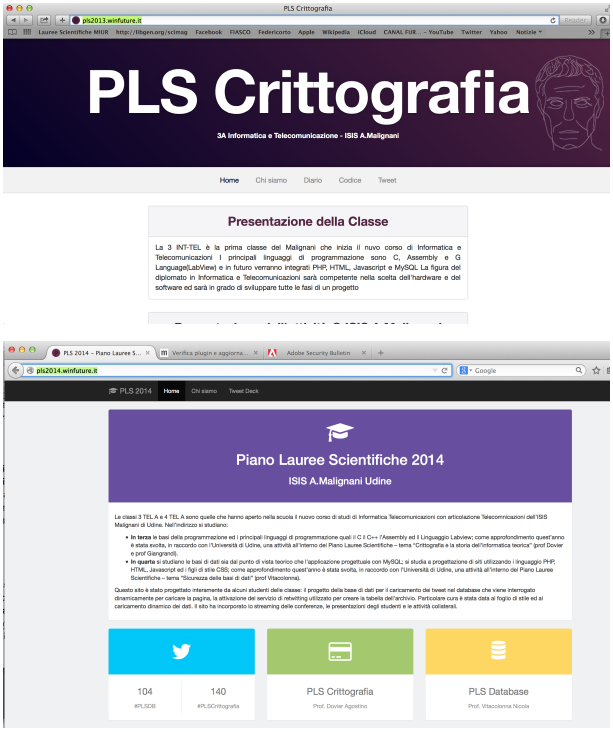

Figura 1.7: Sito web progettato nel 2012/2013 (`http://pls2013.winfuture.it/`), dove si possono anche trovare codici Java (in alto) e sito web progettato nel 2013/2014 (`http://pls2014.winfuture.it/`) nel corso del progetto con il dr. Vitacolonna (in basso)

1.5 Conclusioni

Come si evince dalla sezione precedente, i risultati del laboratorio sono andati oltre le aspettative. Anche in istituti con minori competenze di tipo informatico sono emersi spunti estremamente interessanti da parte degli studenti. Inoltre diversi allievi si sono appassionati a tal punto da iscriversi al corso di laurea in Matematica ed addirittura in Informatica, di cui prima nemmeno conoscevano l'esistenza. Il lavoro svolto ha permesso anche la presentazione dell'attività al GO On FVG del 5 maggio 2014 da parte della prof.ssa Brocato assieme agli allievi S. Cragnolini, I. Manfredi, R. Nobile, A. Roccaforte della IV TELA 13/14: https://prezi.com/e7rg_oqffetd/go-on-fvg/. Per ulteriore materiale e approfondimenti si rimanda al sito https://www.dimi.uniud.it/scuole/pls/moduli/codici-segreti/.

Figura 1.8: Materiale a cura degli allievi III TELA 13/14 contenente lo streaming video delle lezioni in presenza organizzate con Prezi `https://prezi.com/ijfmnkgeycoo/pls/`

1.6 Ringraziamenti

Desideriamo ringraziare: la dr.ssa Anna Barbieri per la collaborazione nell'edizione 2011/2012 nella quale ha tenuto dei seminari sul problema della fattorizzazione di numeri interi, il dr. Marco Peressotti per il porting in C, Visual Basic e Pascal degli algoritmi di cifrazione, decifrazione e decrittazione del codice Vigénere, le proff. Clara Veronese e Stefania Pividori per il loro lavoro svolto con passione e originalità presso l'istituto Zanon, nonché le professoresse Sabrina Capobianco, Eva Caluzzi e Alessandra Maniglio del Liceo Scientifico "Le Filandiere" di San Vito al Tagliamento e i proff. Laura Candotti e Massimo Bove del Liceo Scientifico "Magrini" di Gemona, il prof. Davide Fattori dell'ISIS di Tarvisio e i proff. Giorgio Tuan, Luca Peresson e Maria Fontana dell'ISIS Malignani. Infine si ringraziano tutti gli studenti partecipanti ed in particolare gli allievi delle classi III TELA A.S. 12/13, III TELA A.S. 13/14, IV TELA A.S. 13/14 dell'ISIS A. Malignani di Udine.

Riferimenti bibliografici

[3] Swanston Andrew. *Il codice del traditore*. IBS, 2012.

[63] Robert Harris. *Enigma*. Arnoldo Mondadori Editore, 1995.

[88] Edgar Allan Poe. *Lo Scarabeo D'Oro*. 1843.

[100] Andrea Sgarro. *Codici Segreti*. Arnoldo Mondadori Editore, 1989.

[102] Simon Singh. *Codici & Segreti*. BUR Biblioteca Univ. Rizzoli, 2001.

[105] Alan M. Turing. «Computability and λ-Definability». In: *J. Symb. Log.* 2.4 (1937), pp. 153–163. DOI: 10.2307/2268280. URL: http://dx.doi.org/10.2307/2268280.