UNIVERSITA' DEGLI STUDI DI TERAMO
Facoltà di Scienze della Comunicazione

Franco Eugeni

ELEMENTI DI CRITTOGRAFIA

.

Edizioni Telematiche APAV - Teramo

1° Edizione 13 Giugno 2007

Indice

PREMESSA

Questo volume telematico è una opera aperta (nel senso che può essere integrato successivamente).

La Crittografia è una disciplina matematica nella quale i processi di calcolo, che poi diventeranno tipici dell'informatica, sono sempre presenti per questo basta ricordare che quando a Leon Battista Alberti fu posto dal segretario papale Leonardo Dati il problema di capire come mai gli avversari intercettavano, Alberti rispose inventando il metodo della statistica dei simboli ecc. Sostanzialmente uno strumento di calcolo che anche se non era automatizzato in futuro sarebbe stato facile farlo.
D'altra parte possiamo dire che la rottura di Enigma, macchina tedesca degli anni '30-'40, avvenne grazie al calcolatore bomba dei polacchi, e successivamente ai Colossi inglesi dando luogo ad un processo che è un naturale passaggio all'informatica come mezzo indispensabile per i crittoanalisti. Del resto anche il metodo per trasferire i messaggi in binario, cifrarli, decifrarli e quando leggerli in linguaggio reale è di fatto possibile grazie a tecniche automatiche.
Andando avanti nel mondo attuale l'attacco alla decomposizione di un prodotto di due primi di almeno 100 cifre l'uno per ottenere la decomposizione, è possibile solo per via informatica.
Lo studio della Crittografia ci condurrà lentamente, in un esame attraverso i secoli che sempre più ci conduce alla necessità dell'utilizzo di strutture informatiche e di grandi potenze di calcolo che ormai sono possibili solo attraverso sistemi cooperativi mastodontici anzi mostruosi, che nascono dall'organizzazione migliore di PC connessi tra loro attraverso la rete.

LA CRITTOGRAFIA

1. Introduzione

Una visione storica del problema è sempre di grande utilità per la comprensione dello stesso. In questa introduzione, pertanto, intendiamo dare una breve ed efficace panoramica sulla *crittologia*.
La crittologia può essere divisa in tre grandi branche.

- La *crittografia* o arte delle scritture segrete: in questo ambito il problema che ci si pone consiste nel nascondere un messaggio con procedimenti noti solo al mittente e al destinatario, al fine di impedire che un personaggio estraneo alla comunicazione possa, senza essere autorizzato, comprendere il messaggio stesso. I procedimenti per "cifrare" e "decifrare" i messaggi rappresentano il “segreto” del codice.
- L'*autenticazione* è un ulteriore procedimento mediante il quale il ricevente ha una garanzia che il messaggio sia esattamente quello spedito dal mittente e che il testo sia autentico. I metodi di autenticazione più sofisticati sono i procedimenti di firma numerica, grazie ai quali il mittente non può disconoscere di aver inviato quel dato messaggio ed il destinatario è in grado di provare ad un terzo l'identità dei mittente.
- La *crittoanalisi* è la metodologia di ricerca che mira alla ricostruzione, parziale e/o totale, dei sistemi di cifratura usati senza essere in possesso né della loro architettura, né delle norme d'impiego, né delle chiavi usate. Dunque sono definiti crittoanalisti coloro che si occupano di “rompere” i messaggi senza conoscerne la chiave segreta.

Risulta interessante ora riflettere sulla frase, di seguito riportata:

"*L'ingegno umano non riuscirà mai a concepire un cifrario*
di cui l'ingegno stesso non possa scoprire la chiave"

(Edgar Allan Poe)

Tale frase si può riassumere nel modo seguente: non appena un uomo è in grado di costruire un cifrario, si riesce a trovare sempre un altro uomo pronto ad abbatterlo.

2. La storia

L'uso delle scritture segrete risale a tempi antichissimi. Lo stesso Erodoto (VII, 139) ci narra che un tale Demarato riuscì ad informare i Lacedemoni del progetto di Serse di invadere la Grecia facendo pervenire loro un messaggio inciso su di una tavoletta, che era stata ricoperta con cera. Inoltre, durante le guerre persiane, si usava rapare uno schiavo, incidere il messaggio sopra la sua testa, fargli poi ricrescere i capelli ed infine inviare il messaggio. Una volta giunto a destinazione, lo schiavo veniva rapato di nuovo e ciò consentiva al destinatario di leggere il messaggio.
Aulo Gellio (*Noct. att.*, XVII, 9) parla di un sistema simile, usato dai cartaginesi, e descrive la *scytala lacedemonica*, della quale parla anche Plutarco: il messaggio veniva inciso su un nastro di cuoio avvolto attorno ad un tubo di legno e

poi, sfilando il nastro, si spediva il messaggio che poteva essere letto solo da chi possedeva un tubo dello stesso diametro. Si introdusse, così, il primo strumento che possedeva un codice. Svetonio (*Caes.*, LVI) parla di un alfabeto convenzionale usato da Giulio Cesare: era sufficiente sostituire ogni lettera con quella che la seguiva dopo aver effettuato lo slittamento di tutte le lettere dell'alfabeto normale, di un certo numero di posti prestabilito.

Esempio

ABCDEFGHIJKLMNOPQRSTUVWXYZ (*alfabeto in chiaro*)

XYZABCDEFGHIJKLMNOPQRSTUVW (*alfabeto cifrato*)

Testo chiaro: ALEA IACTA EST

Testo cifrato: XIBX FXZQX BPQ

Quindi, il testo cifrato è stato ottenuto slittando di tre posti tutte le lettere dell'alfabeto in chiaro.
A parte il cifrario di Giulio Cesare sopra analizzato, va ricordato anche l'antico sistema persiano di trasmettere segnali accendendo fuochi su località elevate; simili sistemi di comunicazione furono usati dai greci, cartaginesi e romani, addirittura fino ai tempi della guerra anglo-boera. Alcuni scrivevano messaggi sulle foglie che servivano a bendare le piaghe purulente dello schiavo-corriere mentre, ad esempio, Scipione l'Africano concepì un sistema di comunicazione, recentemente molto usato dall'Unione Sovietica, che consisteva nell'inviare in altri paesi spie esperte travestite da domestici degli ambasciatori.
Nel VI secolo, la rete informativa dell'impero bizantino era divenuta uno dei fondamenti dello stato: agenti mercanti erano sparsi in tutto il paese attraverso un sistema di import-export.
Comunque l'invenzione di alcuni dei più importanti sistemi di comunicazione viene fatta risalire ad Alessandro il Grande: ad una certa distanza dal suo quartier generale, il macedone aveva infatti costituito un vero e proprio ufficio di raccolta dati, basato sulle informazioni ottenute da agenti inviati nei paesi nemici. Costoro avevano anche il compito di spargere notizie false sulle intenzioni e le mosse di Alessandro, in modo da disorientare il nemico o indurlo a svelarsi.
Nel Medioevo non si ebbe una sostanziale evoluzione dei sistemi crittografici, mentre nel Rinascimento la crittografia ebbe notevole impulso, grazie alla scoperta di nuovi sistemi di cifratura ideati da:

- Leon Battista Alberti, sul cui codice torneremo più avanti
- Giovan Battista Della Porta, celebre fisico napoletano (1540–1615), inventore della camera oscura ed autore, fra l'altro, del trattato *De furtivis literarum notis* (Napoli 1563)
- Gerolamo Cardano, medico e matematico (1501–1576), che, durante la sua sfortunata esistenza, affrontò anche problemi crittografici nel *De subtilitate* (Lione 1554).

Fuori Italia occorre menzionare il tedesco Tritemio (Johannes da Trittenheim, 1462–1516), autore della *Polygrafia* (Francoforte 1550) e *della Steganographia, hoc est ars per occultam scripturam animi sui voluntatem absentibus aperiendi* (Francoforte 1606–1622). Alcuni metodi di Tritemio sono riportati anche nel libro di Umberto Eco, *Il pendolo di Focault*. Il francese Blaise de Vigenère (1522–1596) autore di un *Traicté des chiffres ou secrètes nianières d'escrire* (Parigi 1586), invece, merita una descrizione a parte. Il suo codice, infatti, fu considerato "sicuro" per più di 200 anni, fin quando, nel 1863, Kasiski, ufficiale prussiano, ideò un test statistico (test di Kasiski) che spezzo il codice.

Nel secolo XVII fu attribuita sempre maggiore importanza alle scritture in cifra e le varie Nazioni adottarono sistemi di cifratura sempre più complessi.

Di conseguenza durante il Rinascimento una rete fittissima di agenti copriva l'Europa tant'è che neanche il Papa poteva stare tranquillo: si dice, infatti, che il segretario di Adriano VI fosse una spia dell'Imperatore Carlo V, al corrente di tutti i segreti del Papa e della sua corte[1].

Il miglior servizio di spionaggio dell'epoca era probabilmente quello spagnolo: l'agente principale di Filippo II era nientemeno che l'ambasciatore inglese a Parigi, Sir Edward Stafford, il quale fra l'altro riuscì a fornire agli spagnoli notizie sulla flotta di Sir Francis Drake, pronta a salpare contro l'Invincibile Armata. A questo punto però si inserì nel gioco Sir Francis Walsingham, che, raccolte le prove del tradimento di Stafford, decise di sfruttarlo a suo vantaggio, così tramite Stafford si ottennero molteplici informazioni, ad esempio l'elenco preciso di tutte le spie spagnole in Inghilterra.

Padre Giuseppe, religioso al servizio del Cardinale Richelieu, fu famoso per la sua scuola di informatori. Essi trovarono pane per i loro denti solo con il servizio segreto inglese, guidato allora da John Thurloe, abilissimo collaboratore di Oliver Cromwell.

Verso la fine del XVIII secolo il primato dell'organizzazione spionistica passa dall'Inghilterra alla Francia: Napoleone Bonaparte domina l'Europa, non solo con la potenza degli eserciti e con il genio strategico, ma anche con l'efficienza della sua rete di informatori. Tra questi il più grande di tutti, Karl Schulmeister.

Presentato da René Savary a Napoleone con le parole "Ecco, Sire, un uomo tutto cervello e senza cuore, ai Vostri ordini", Schulmeister si accinse a quella che resta forse la più incredibile azione di spionaggio della storia: diventare capo del servizio di informazioni militari della coalizione avversa a Napoleone. Schulmeister si trasferì a Vienna ed offrì informazioni strategiche di grande importanza e assolutamente vere.

In meno di un anno, Schulmeister riuscì a farsi nominare capo del servizio di informazioni austriaco: da quel momento fu come se Napoleone stesso potesse esaminare i piani strategici del nemico.

Le vittorie di Ulm e di Austerlitz furono in gran parte dovute a Schulmeister, che non solo informava Napoleone delle mosse nemiche, ma forniva agli alleati false indicazioni.

Nel 1900 la tecnica fa notevoli passi avanti ed i mezzi di comunicazione subiscono una vera e propria rivoluzione: la fotografia, il telegrafo, la radio, il telefono, l'aereo. Nascono i servizi militari organizzati. A volte forse anche più di uno per nazione. In Italia durante la seconda guerra vi erano ad esempio i seguenti:

- Servizio Informazioni Militari (SIM) dell'Esercito
- Servizio Informazioni Militari (SIS) della Marina
- Servizio Informazioni Aeronautica (SIA) dell'Aeronautica
- Centro di Controspionaggio Militare e Servizi Speciali (CCMSS) alle dipendenze del Ministro della Guerra
- Organizzazione di Vigilanza e Repressione Antifascista (OVRA)

Chiusa l'era delle spie romantiche e degli avventurieri di genio, i nuovi personaggi saranno soprattutto colonnelli inglesi a riposo con l'hobby della decifrazione; giornalisti mondani affiliati da anni da qualche servizio segreto; taciturni camerieri turchi e scienziati atomici convinti che una potenza diversa dalla loro patria avrebbe fatto un uso più giusto dei terribili segreti di cui erano a conoscenza; ufficiali della Marina particolarmente dotati nel campo dell'intercettazione e delle trasmissioni radio.

[1] Con la Riforma nasce così la "spia ideologica": uomini e donne di ogni ceto sociale divennero traditori della loro patria per servire gli interessi di una delle due fazioni cristiane in lotta.

Accanto a loro, spesso alle loro dipendenze, tutta una galleria di personaggi minori, mossi da motivi sia nobili che ignobili, ideologici o bassamente economici, ma sostanzialmente non diversa dagli informatori dei faraoni d'Egitto o dei consoli romani del passato.

Durante la prima guerra mondiale (1915-1918) va ricordato l'episodio del telegramma Zimmerman nell'aprile 1917, che è, di fatto, il primo caso di messaggio segreto che decriptato al momento giusto creò un evento assai importante: l'entrata in guerra degli Stati Uniti.

Come premessa va ricordato che nel 1915 un U-boot tedesco in immersione aveva silurato il transatlantico *Lusitania* causando la morte di 1200 civili tra cui 150 americani. L'incidente avrebbe causato l'entrata in guerra degli Stati Uniti ma un accordo risolse l'incidente. Del resto lo stesso presidente Woodrow Wilson era contrario ad un intervento poiché sperava di poter pesare nelle trattative e condurre l'Europa ad una pace risparmiando vite americane. Nel novembre 1916 la nomina di Arthur Zimmerman a ministro degli esteri tedesco di notoria indole liberale, sembrò facilitare l'ipotesi di Wilson.

In realtà Zimmerman riteneva che si dovesse scatenare una guerra sottomarina senza avvisaglie e senza quartiere e consigliò il Kaiser in questa direzione. L'unica preoccupazione era l'eventuale intervento USA. Così Zimmerman tentò un accordo con il presidente del Messico spingendolo ad invadere gli USA al fine di chiedere la restituzione del Nuovo Messico, del Texas e dell'Arizona. Il presidente messicano, inoltre, avrebbe tentato di coinvolgere anche il Giappone a dare manforte in questa operazione. Gli Stati Uniti impegnati su questi fronti non avrebbero prestato molta attenzione all'Europa. L'inizio dell'operazione avvenne con un telegramma cifrato inviato all'ambasciatore tedesco in Messico, che doveva informare il presidente messicano del da farsi. Di seguito riportiamo una traccia di quel famoso telegramma:

*"**Segretissimo** – Per conoscenza personale di Vostra Eccellenza e per l'ulteriore inoltro al Ministro Imperiale in (Messico?) con telegramma No. 1 (...) per via sicura.*

Intendiamo iniziare dal 1° febbraio la guerra sottomarina senza restrizioni. Ci sforzeremo, ciononostante, di mantenere neutrali gli Stati Uniti d'America. (?)

Se non dovessimo (riuscirci) proponiamo al (? Messico) un'alleanza su queste basi:

[condurre una comune] condotta di guerra, [pervenire ad comune] conclusione della pace (...) con nostro sostegno finanziario e nostro accordo a ceder a Messico i perditi territori del Nuovo Messico, del Texas e dell'Arizona. A lei definire dettagli.

Vostra Eccellenza informi il presidente [del Messico] segretamente (? che noi prevediamo) guerra con gli Stati Uniti (forse) e nello stesso tempo d'intavolare trattative fra noi ed il Giappone.

(La preghiamo dire al Presidente) che (...) i sommergibili costringeranno l'Inghilterra alla pace entro pochi mesi. Accusi ricevuta" – Zimmerman"

Il telegramma prese, per un incidente su alcuni cavi sottomarini, la via della Svezia. Intercettato dallo spionaggio inglese venne decrittato nella famosa stanza 40 dal crittoanalista, nonché reverendo, William Montgomery. Così nonostante la lunga riluttanza del Presidente Wilson e dei non interventisti, quando il congresso ne venne a conoscenza, ed insieme al congresso la stampa e i media, la situazione cambiò globalmente dalla sera alla mattina. Una semplice decrittazione era valsa tre anni di fallite trattative e l'operazione bellica che ne derivò fu di proporzioni colossali e con il ben noto esito.

Durante la seconda guerra mondiale fu molto in auge una macchina di decrittazione costruita da Alan Turing il cui nome era Enigma.

Oggi lo scenario è cambiato: oltre al mondo politico–militare (statale in genere) è coinvolto anche il mondo privato (civile). Il problema crittografico nuovo è quello di proteggere qualsiasi informazione. Si pensi a quanti dati attualmente occorre mantenere riservati, contenuti, ad esempio, presso:

- archivi di banche
- archivi di grosse e medie industrie
- archivi di enti di stato dai quali possono trarsi informazioni sullo stato patrimoniale anche di singoli soggetti
- archivi di compagnie aeree dai quali sapere in anticipo spostamenti di singoli soggetti
- archivi di ospedali

Inoltre, si pensi che oggi nessuno è in grado di garantirci la riservatezza non solo dei dati contenuti negli archivi sopraindicati ma anche la segretezza di una telefonata o l'autenticità di un ordine di pagamento fatto in uno dei canali disponibili (telefonico o telefax). Inoltre un nostro agente potrebbe disconoscere l'ordine ricevuto ovvero un malintenzionato potrebbe modificare l'ordine dall'esterno a suo favore e a nostro danno.
Quindi, come è facilmente intuibile anche per i non addetti ai lavori, la crittologia nasce e prolifera come supporto per le comunicazioni militari. Con il passare dei secoli, però, i metodi sono cambiati ed i crittografi naturalmente sono sempre a caccia di nuovi codici che i critttoanalisti sistematicamente scoprono.

3. Il problema dei crittografi: fare un caos ordinato

Il problema più delicato nell'organizzazione di una rete di informazioni è quello di stabilire comunicazioni sicure tra gli agenti e il centro da cui essi dipendono. I diversi sistemi escogitati nel passato si può dire siano stati sufficientemente simili tra loro.
Spesso le informazioni venivano spedite per mezzo di lettere redatte secondo un codice convenzionale precedentemente stabilito: ad esempio comuni lettere commerciali. Cifre e nomi di particolare interesse erano "travestite" con vocaboli innocenti: cliente, denaro e ordinazioni indicavano effettivamente truppe, dati bellici o altre informazioni. Ecco un esempio di codice molto elementare.

Cliente = Alleato	Acquistare = Attaccare
Merce = Esercito nemico	Ditta = Divisione in soccorso
Interpellare = Comunicare a ...	Inviare = Far saltare

Per messaggi di questo tipo era necessario avere fantasia e tempo a disposizione. A causa della loro ingenuità essi risultavano immediatamente cifrabili, non erano quindi così adatti per trasmettere informazioni di una certa importanza[2]. Messaggi convenzionali venivano alle volte inviati anche per mezzo di inserzioni nei giornali.
Ad esempio: "Vendesi terreni per 10 ettari, 2 fabbricati, 823 bovini" può essere benissimo un messaggio in cui 102823 è la chiave da usare per decifrare un successivo o precedente messaggio.
L'abilità consiste nel pensare il testo in modo da non destare il sospetto e da non richiamare l'attenzione dell'avversario.

[2] É bene che i messaggi cifrati sembrino veri

Il mezzo più usato da quando esiste è stata la radio. Con essa è possibile inviare tempestivamente qualsiasi messaggio al proprio centro, da questo è possibile ricevere in ogni momento ordini o informazioni. Le trasmissioni avvenivano ad ore prestabilite ed iniziavano sempre con la sigla di una delle parti. Il dispaccio veniva trasmesso cifrato, con un sistema la cui chiave era conosciuta quasi sempre da un solo trasmettitore che, una volta raccolte le informazioni, le cifrava per affidarle all'agente incaricato di effettuare la trasmissione. Così questi inviava al centro una serie di numeri o lettere a lui stesso incomprensibili. Nel corso dell'emissione l'operatore inseriva il cosiddetto *parity check* che segnalava eventuali errori dovuti al canale e il *security check*, cioè un errore, una parola sbagliata sempre uguale, in modo che il centro avesse la certezza che l'agente stava trasmettendo liberamente. La radio presentava il grave inconveniente di poter essere individuata dalle stazioni di ascolto e di intercettazione avversari.

Come si è detto, i messaggi, prima di essere trasmessi, venivano tradotti "in cifra" secondo particolari metodi appositamente studiati.

I vari sistemi di cifratura possono a grandi linee raggrupparsi in tre categorie: sistemi *a trasposizione*, sistemi *a sostituzione* e sistemi *misti*. Nei primi la traduzione del linguaggio chiaro in linguaggio segreto ha luogo mediante una specie di anagramma degli elementi dei testi chiari; nei secondi mediante una sostituzione degli elementi stessi con cifre crittografiche, cioè con segni convenzionali, o con gruppi di tali segni; nei terzi mediante entrambe le operazioni eseguite successivamente l'una dopo l'altra in un certo ordine.

I sistemi a sostituzione, più precisamente, consistono nel mettere al posto di ogni lettera (o gruppo di lettere, o parole o frasi del testo) un'altra lettera (o numero, o gruppo di lettere o di numeri). La sostituzione letterale può essere *monoalfabetica o polialfabetica* a secondo se avviene in base ad un solo alfabeto cifrante o a più alfabeti cifranti, da adoperare in blocco, ma con una certa legge, di volta in volta, prestabilita.

Ad esempio il seguente messaggio: "Appuntamento ora X…" può essere così trasmesso:

BQQVOUBNFOUPPSBY ...

Osserviamo innanzitutto che lo spazio tra le parole è stato abolito (cosa che accade anche per l'immagine crittografata) e che l'alfabeto usato in questo caso è un alfabeto slittato di un posto, nel quale la B ha preso il posto della A, la C della B, la D della C e così via. Naturalmente l'alfabeto cifrante si può ottenere spostando di due, tre, quattro, o anche di ventisei posti l'alfabeto reale. In questo caso ci si trova di fronte ad un codice monoalfabetico poiché vi è una corrispondenza biunivoca (uno ad uno) tra l'alfabeto chiaro e quello cifrante. Per problematiche statistiche, come si vedrà nel seguito, questi codici presentano una debolezza estrema.

Un altro messaggio: "Giorno X ora Y in arrivo navi ed aerei alleati ..." cifrato per trasposizione senza chiave diviene:

GIIE INEA OADT RRAI NREA OIRB XVEC OOID RNAE AALF YVLG

Esso si ottiene scrivendo su di un rettangolo il messaggio da inviare nel modo seguente:

GIORNOXORAY
INARRIVONAV
IEDAEREIALL
EATIABCDEFG

La cifratura può aver luogo per *lettere*, per *sillabe*, o per *gruppi* di un numero fisso di lettere (*poligrammi*), o per *frasi*, o in modo promiscuo.

La maggior parte degli autori ripartisce i sistemi di cifratura nelle due categorie di sistemi letterali o di sistemi a repertorio. I primi consistono nella trasposizione e/o sostituzione delle lettere o di poligrammi. I secondi consistono invece nella sostituzione delle parole e delle frasi. Questa operazione può essere seguita da una seconda cifratura, o sopracifratura, per trasposizione e/o per sostituzione delle relative cifre crittografiche. I segni convenzionali che rappresentano gli elementi del linguaggio chiaro possono essere di qualsiasi genere, ma nei tempi moderni è frequente l'uso di segni adoperati nella corrispondenza telegrafica. Sono usati i segni della normale scrittura, e per lo più si usano o sole cifre arabe o sole lettere e si formano gruppi di un numero fisso di elementi.
Molto comuni sono i sistemi a gruppi di cinque lettere o di cinque cifre arabe, ovvero di dieci lettere costituenti un insieme pronunciabile, essendo tali i massimi tassabili per una parola nelle comunicazioni telegrafiche internazionali.
La convenzione in base alla quale si eseguono le operazioni di sostituzione e di trasposizione è sovente rappresentato da una *chiave*, cioè da una serie di numeri o di lettere, il cui uso può spiegarsi mediante esempi.

Esempio

Alfabeto chiaro	A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T	U	V	W	Z
Alfabeto cifrato	g	i	n	e	p	r	o	a	b	c	d	f	h	l	m	q	s	t	u	v	w	z

Si fissa l'attenzione sulla parola chiave, "ginepro", formata da lettere tutte distinte, che rappresenta la vera e propria chiave. Il metodo di cifratura è ovvio: nell'alfabeto cifrato sono assenti tutte quelle lettere comprese nella parola chiave. Chiaramente se la parola chiave fosse stata costituita anche da doppie allora bisognava contarle una sola volta, ossia esse andavano contratte. Da notare che, in questo caso, potrebbe essere conveniente cambiare la parola chiave per il fatto che da s a z ci sono troppe lettere fisse.
La cifratura può anche essere eseguita mediante l'uso di griglie, cioè di poligoni di cartone o di altra materia ripartiti in caselle, delle quali un certo numero forate. Il tipo originario è la griglia quadrata, ideata da Girolamo Cardano (1501-1576), con la quale la cifratura ha luogo sovrapponendo quadrati di cartone forati in un certo modo convenuto. Il messaggio viene scritto nei fori della griglia appoggiata su di un foglio di carta quadrettata, ruotando il cartoncino in un modo stabilito quando tutte le fessure sono state riempite.
Dopo aver tolto la griglia, si rilevano le lettere o colonna per colonna, sia in senso verticale che in senso orizzontale, oppure mediante l'aiuto di una "chiave" come nei casi precedenti.

Esempio

PRIMO QUADRATO

	1				
2					
	3		5		
		4		6	
					7
				8	

SECONDO QUADRATO

				7	
			6		8
		5			
			4		
1		3			
	2				

Si osservi che il secondo quadrato è stato ruotato, rispetto al primo, di 90° (in senso antiorario).

Testo chiaro LA CRITTOGRAFIA E' INTERESSANTE

	L									E			N				
A									A		I	T					
	C		I					I					E		E		
		R		T					F					R		S	
					T	G		A									S
				O			R									A	

Gli spazi vuoti vengono riempiti ad esempio con lettere casuali.

Per eliminare in parte le ripetizioni che compaiono necessariamente in ogni crittogramma, specialmente nei testi molto lunghi, si ricorre alla sopracifratura, operazione che consiste nel cifrare di nuovo il testo ottenuto con la prima cifratura, magari cifrando parzialmente il testo o usando un cifrario tipo quello che vedremo più avanti di Leon Battista Alberti, nel quale cambiare alfabeto è facile.

Naturalmente il destinatario per poter decifrare un crittogramma doveva compiere operazioni inverse a quelle eseguite dal mittente.

Da un punto di vista storico, possiamo dire che nel Medio Evo, per le condizioni politiche, i codici segreti vengono usati poco ed in genere solo per nascondere nomi di personaggi importanti.

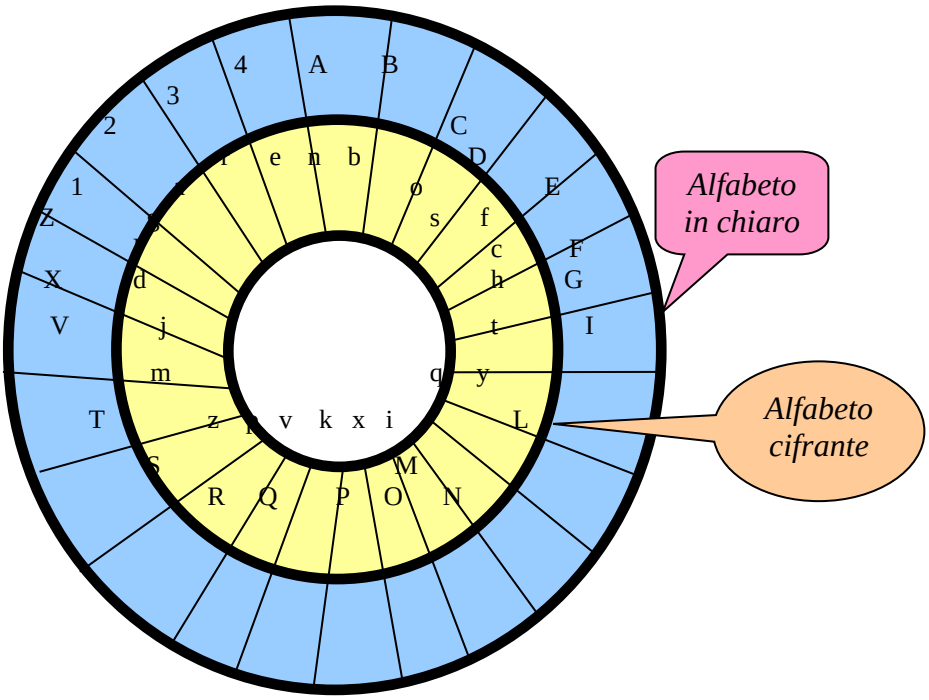
Verso la fine del Medio Evo, con l'inizio delle relazioni diplomatiche tra i vari stati i codici segreti diventano una necessità.

Secondo ricerche storiche dovute a Meister (1902) l'uso sistematico dei codici segreti ebbe inizio nella corte papale, nelle repubbliche e signorie italiane, a partire dal 1300.

E' in questo periodo che la Crittografia ha una grossa evoluzione. Troviamo un cambiamento radicale, infatti nascono i primi codici segreti che non usano un solo alfabeto cifrante, ma molti alfabeti cifranti. Tali codici si chiamano codici polialfabetici.

Il primo codice polialfabetico è dovuto ad un illustre pensatore italiano: Leon Battista Alberti, architetto, urbanista, pedagogo, matematico e crittografo. É suo, su commissione del segretario pontificio, Leonardo Dato, il primo codice polialfabetico della storia, messo a punto intorno al 1466.

Esso era costituito da due cerchi o dischi concentrici: in quello più esterno compare l'alfabeto in chiaro formato da 24 caselle, 20 delle quali contenenti lettere (mancano per ragioni di sicurezza crittografica, le lettere che si presentano con minore frequenza, cioè J, K, Y, W, Q, H) e le rimanenti i numeri 1, 2, 3, 4.



Osserviamo che il cerchio più esterno contiene lettere tutte maiuscole e ben ordinate (sono scritte, infatti, dalla A alla Z, anche se alcune lettere sono state omesse per timore di un'eventuale confusione: manca la U perché identificata con la V così come la J è stata associata alla I). Tale cerchio, però, contiene ben quattro numeri, dall'1 al 4.
Leon Battista Alberti, inoltre, utilizzava vari dischi interni, in base, ad esempio, al giorno in cui veniva inviato il messaggio. Il disco interno (quello riportato in figura è uno dei tanti possibili dischi interni) contiene una permutazione di altre 24 lettere (manca la lettera w ed è u = v) formanti l'alfabeto cifrante ed esso può ruotare rispetto al primo disco. Trattandosi di 24 lettere dell'alfabeto, il numero dei riordinamenti possibili, ovvero il numero dei dischi che bisognava costruire, era pari esattamente a 24! (si legge 24 fattoriale), essendo:

$$24! = 24\times23\times22\times21\times \ldots \times4\times3\times2\times1$$

Poiché il numero che si ottiene è eccessivamente grande è possibile costruire solo 365 dischi interni, tanti quanti sono i giorni dell'anno.

Esempio

Messaggio esatto: TERAMO E' UNA CITTA'
Messaggio da inviare: TERAMOEUNACITTA

Da notare che sono stati eliminati gli spazi e gli accenti.

Prima di cominciare a codificare il messaggio, va scelto il disco da utilizzare in accordo con il destinatario (sia esso quello riportato in figura) e poi va inizializzata la macchina fissando, a piacere, una lettera dell'alfabeto cifrante, ad esempio la a, da collocare sotto la A dell'alfabeto in chiaro, attraverso una rotazione del disco interno. Poiché c'è una corrispondenza biunivoca tra le caselle dei due dischi, allora alla lettera scelta come indice del codice, cioè la a, corrisponde una ed una sola lettera del disco esterno. Dopo aver posto la a sotto la A si inizia a codificare il messaggio tenendo conto della corrispondenza tra le lettere dell'alfabeto in chiaro e quelle dell'alfabeto cifrante.

I° PASSO) T→v, E→b, R→x ⇒ TER→**vbx**

Se continuiamo per questa strada, cioè usando un codice monoalfabetico, la probabilità che esso sia scoperta è estremamente alta per cui occorre cambiare alfabeto utilizzando uno dei quattro numeri che compaiono nell'alfabeto in chiaro. Scegliamo, ad esempio, **3**, da inserire tra la R e la A del messaggio da cifrare:

TERAMOEUNACITTA→TER**3**AMOEUNACITTA

II° PASSO) 3→l

Cambiamo così l'inizializzazione portando la a, scelta all'inizio, sotto il numero 3 e procediamo come in precedenza:

III° PASSO) A→e, M→y, O→i ⇒ AMO→**eyi**

Cambiamo di nuovo alfabeto, scegliendo questa volta il numero **1**, da inserire tra la O e la E del messaggio da decifrare:

TERAMOEUNACITTA→TER**3**AMO**1**EUNACITTA

IV° PASSO) 1→l,

Inizializziamo nuovamente la macchina, portando la a sotto il numero 1 e proseguendo come al solito:

V° PASSO) E→c, U=V→d, N→x, A→b ⇒ EUNA→**cdxb**

Si può ora cambiare di nuovo l'inizializzazione della macchina scegliendo un altro numero a piacere e ripetendo il discorso precedente, ottenendo così:

Messaggio in chiaro: TERAMOEUNA ……

Messaggio cifrato: vbxleyilcdxb ……

Per decifrare il messaggio chiaramente si utilizza lo stesso procedimento al contrario, conoscendo a priori il disco cifrante da utilizzare e l'inizializzazione della macchina.

Anche se la macchina dell'Alberti non ebbe molta fortuna lo stesso autore ne riconobbe i seguenti vantaggi:

i) nessuna cifra è più rapida;
ii) nessuna cifra è più facile a leggersi;
iii) nessuna cifra, se si ignorano gli indici convenuti tra due persone, si può pensare più segreta.

La tendenza del periodo medioevale, quindi, era proprio quella di costruire dei semplici algoritmi per cifrare messaggi. Nascono così altri vari codici polialfabetici, quali, ad esempio, quelli di Cardano e di Giovan Battista Bellaso.

Successivamente si avvertì la necessità di utilizzare la parola chiave introdotta, per la prima volta, dall'italiano Giovan Battista Della Porta (1563), inventore della camera oscura, rivale di Galileo Galilei al quale fu attribuito, dopo una lunga disputa, il merito di avere scoperto il cannocchiale, il commediografo, il crittografo, etc.

Con riferimento alla tavola di Della Porta vediamo l'uso della parola chiave in quel codice.

Si comincia con il fissare una parola del tutto arbitraria ma contenente lettere tutte distinte (ciò perché ad ogni parola corrisponderà un diverso alfabeto); sia AMBRISEMLO (abbiamo tolto tre I). Si scrive tale parola sotto il messaggio un numero di volte tale da "coprire" il messaggio stesso, indi si usano le tavole come nel seguente esempio.

Esempio

Ogni lettera di essa, ad esempio a, ci dice quale alfabeto dobbiamo usare per criptare la lettera del messaggio corrispondente. Nel caso generale quindi si deve dare una permutazione dell'alfabeto per ogni lettera della parola chiave. La costruzione di queste permutazioni costituisce il codice stesso. Come esempio vediamo, di seguito, il sistema di Della Porta:

ab	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
cd	A	B	C	D	E	F	G	H	I	J	K	L	M
	Z	N	O	P	Q	R	S	T	U	V	W	X	Y
ef	A	B	C	D	E	F	G	H	I	J	K	L	M
	Y	Z	N	O	P	Q	R	S	T	U	V	W	X
gh	A	B	C	D	E	F	G	H	I	J	K	L	M
	X	Y	Z	N	O	P	Q	R	S	T	U	V	W
ij	A	B	C	D	E	F	G	H	I	J	K	L	M
	W	X	Y	Z	N	O	P	Q	R	S	T	U	V
kl	A	B	C	D	E	F	G	H	I	J	K	L	M
	V	W	X	Y	Z	N	O	P	Q	R	S	T	U
mn	A	B	C	D	E	F	G	H	I	J	K	L	M
	U	V	W	X	Y	Z	N	O	P	Q	R	S	T
op	A	B	C	D	E	F	G	H	I	J	K	L	M
	T	U	V	W	X	Y	Z	N	O	P	Q	R	S
qr	A	B	C	D	E	F	G	H	I	J	K	L	M
	S	T	U	V	W	X	Y	Z	N	O	P	Q	R
st	A	B	C	D	E	F	G	H	I	J	K	L	M
	R	S	T	U	V	W	X	Y	Z	N	O	P	Q
uv	A	B	C	D	E	F	G	H	I	J	K	L	M
	Q	R	S	T	U	V	W	X	Y	Z	N	O	P
wx	A	B	C	D	E	F	G	H	I	J	K	L	M
	P	Q	R	S	T	U	V	W	X	Y	Z	N	O
yz	A	B	C	D	E	F	G	H	I	J	K	L	M
	O	P	Q	R	S	T	U	V	W	X	Y	Z	N

Nelle tavole di della Porta le lettere minuscole scritte in testa danno il nome all'alfabeto di quella riga, che è ottenuto dividendo l'alfabeto in due parti di 13 lettere ognuna e stabilendo una bidirezione tra i due insiemi di 13 elementi.

Allora se la lettera della parola chiave che stiamo considerando è una a oppure una b, si cripta la lettera del testo corrispondente con l'alfabeto di nome *ab*, e così via.

Esempio

Testo in chiaro: TERAMOEUNACITTA
Parola chiave: PORTAPORTAPORTA
Testo cifrato: ???????????????

In primo luogo osserviamo che la parola chiave (porta) è ripetuta un numero di volte sufficienti a raggiungere la medesima lunghezza del testo in chiaro che, nel caso specifico, è costituito da 15 lettere.

Si procede ora nel seguente modo: stabiliamo in primo luogo una corrispondenza tra il testo in chiaro e la parola chiave.

T	E	R	A	M	O	E	U	N	A	C	I	T	T	A
P	O	R	T	A	P	O	R	T	A	P	O	R	T	A

Utilizziamo poi la tavola di Della Porta: poiché alla lettera T (di Teramo) è associata la lettera P (di porta) dobbiamo usare l'alfabeto *op*, cioè a T corrisponde A; analogamente alla lettera E (di Teramo) è associata la lettera O (di porta), per cui anche in questo caso dobbiamo utilizzare l'alfabeto *op*, cioè E $\rightarrow$ X; proseguendo, alla lettera R (di Teramo) corrisponde la lettera R (di porta), quindi, questa volta, dobbiamo cambiare alfabeto, cioè usare *qr* (R $\rightarrow$ M); continuando, si ottiene che: A $\rightarrow$ R; M $\rightarrow$ Z; O $\rightarrow$ I; E $\rightarrow$ X; U $\rightarrow$ C; N $\rightarrow$ J; A $\rightarrow$ N; C $\rightarrow$ V; I $\rightarrow$ O; T $\rightarrow$ B; T $\rightarrow$ C; A $\rightarrow$ N.
Dunque:
Testo cifrato: axmrzixcjnvobcn

Per essere in grado di compilare i messaggi secondo i sistemi esposti, gli operatori devono conoscere sia le chiavi sia i sistemi di cifratura. Alle volte, invece, si sono usati codici cifranti, cioè fascicoli contenenti liste di cifre o di lettere da sostituire alle rispettive voci. Queste liste possono essere costituite da parole, frasi o periodi, più frequentemente usati nei messaggi, con a fianco un gruppo cifrante, numerico o letterale, che può trasmettersi via telegrafo.
I sistemi monoalfabetici sono i più antichi; l'alfabeto cifrante è stabilito in un qualsiasi modo convenzionale e si può fare uso anche di segni nulli, nonché di omofoni, cioè di più segni rappresentanti la stessa lettera dell'alfabeto normale, usabili indifferentemente.
I sistemi polialfabetici derivano tutti dalle tabelle ideate dall'Alberti, da Della Porta e da Giovanni Tritemio.
Tritemio, citato ampiamente da Umberto Eco nel pendolo di Focault, usava rivestire i suoi molteplici codici, molti basati su peculiarità del latino, di misticismo. L'idea di base fu quella di aggiungere molte parole in modo che il messaggio finale avesse senso compiuto. In Tritemio appare anche il quadrato latino 26×26 usato come tavola per cifrare e decifrare. Questo quadrato può essere visto come la tavola di addizione di Z_{26} qualora che sia

$$A = 0,\ B = 1, \ldots, Z = 25$$

L'introduzione di un codice polialfabetico, che ha raggiunto maggiore notorietà rispetto a quello dell'Alberti, va attribuito al francese Blaise de Vigenère.

Egli nel 1586 pubblica il suo codice nel quale fa uso di una tavola quadrata, già introdotta dall'abate Tritemio e nota come tavola di Vigenère, sulla quale vi è veramente molto da dire. Il quadrato è passato alla storia come quadrato o *chiffre carré* del Vigenère. Esso è stato molto in voga, sino a epoca relativamente recente, per scopi militari e diplomatici.

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

La prima riga contiene l'alfabeto di Cesare. La seconda riga contiene l'alfabeto di Cesare slittato di un posto. La terza riga contiene l'alfabeto di Cesare slittato di due posti e così via.

Ogni riga della tavola, dunque, è un alfabeto di Cesare, a cui diamo il nome della lettera posta a sinistra. La tavola, cioè, contiene l'elenco di tutti i possibili codici di Cesare.

Per criptare un messaggio si usano tanti alfabeti di Cesare quante sono le lettere della parola chiave.

La tavola è un quadrato latino cioè una matrice in ogni riga e colona della quale vi è una permutazione della prima riga o colonna[3]. Il segreto di questo codice è tutto nella parola chiave. Quindi, il destinatario del messaggio conosce la parola chiave ed è quindi in grado di decifrare il messaggio in arrivo usando il procedimento al contrario[4].

Esempio

Testo chiaro: VIGENERE
Chiave: **APEAPEAP**
Testo cifrato: VXKECIRT
Dopo aver scritto il testo chiaro, ovvero il messaggio, bisogna ripetere la parola chiave (nel nostro caso **ape**) un numero di volte sufficienti per coprire il messaggio.
Per l'uso della suddetta tabella si adotta generalmente una chiave letterale consistente per lo più in una parola o in una frase, e la cifratura ha luogo, sostituendo ogni lettera del testo chiaro con quella della colonna verticale, nel punto d'intersezione delle colonne comincianti rispettivamente con la lettera da cifrare e con la corrispondente lettera della chiave, o viceversa. Pertanto risulta:

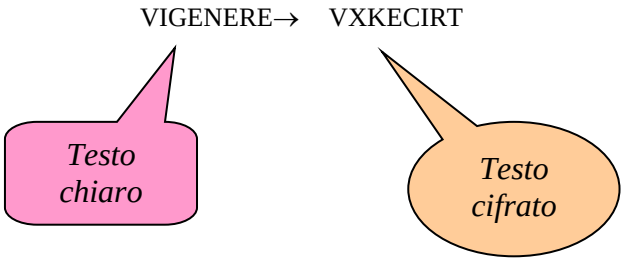
- V si codifica usando A
- I si codifica usando P
- G si codifica usando E
- E si codifica usando A
- N si codifica usando P
- E si codifica usando E
- R si codifica usando A
- E si codifica usando P

Ne segue che:

- l'intersezione tra V ed A è V;
- l'intersezione tra I e P è X;
- l'intersezione tra G ed E è K;
- l'intersezione tra E ed A è E;
- l'intersezione tra N e P è C;
- l'intersezione tra E ed E è I;
- l'intersezione tra R ed A è R;
- l'intersezione tra E ed P è T.

Dunque:

[3] Si tratta anche della tabella additiva delle classi resto modulo 26, quindi un esempio di gruppo finito.
[4] Ovvero la sottrazione modulo 26



Possono pure compilarsi tabelle per cifratura polialfabetica aventi per alfabeto base un alfabeto invertito, che sia cioè una permutazione dell'alfabeto ordinario.
Sono stati ideati anche sistemi a rappresentazione numerica, nei quali gli alfabeti cifrati sono costituiti da numeri, ma essi non differiscono in maniera sostanziale, agli effetti del segreto crittografico, da quelli letterali.
Altro esempio di codice polialfabetico è rappresentato dal cosiddetto "codice dello spillo", basato esclusivamente sulle prime facciate di un giornale. Questo tipo di codice si rivelò molto efficace durante il periodo della guerra.

Esempio

Testo chiaro: Alea iacta est

Ogni volta che troviamo una lettera del testo chiaro nella prima pagina del giornale dobbiamo apportare un buco con lo spillo su tale lettera. Ultimata questa operazione dobbiamo piegare il giornale e lasciarlo in un posto precedentemente convenuto con il destinatario che provvederà a raccogliere il giornale e a leggerlo.

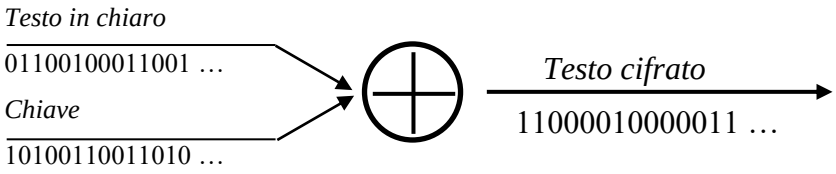
I sistemi polialfabetici possono essere a *chiave fissa*, a *chiave variabile*, a *interruzione della chiave* e *autocifranti*. Nei sistemi a chiave variabile e in quelli a interruzione della chiave occorre stabilire per convinzione il modo in cui si deve segnalare al destinatario del messaggio il cambio o l'interruzione nell'uso della chiave. Nei sistemi autocifranti si adopera la chiave per le prime lettere del testo e si usa poi, come chiave, o lo stesso testo chiaro o il testo segreto ottenuto.
Si può ricorrere per la sostituzione polialfabetica all'uso di "*macchine cifranti*". Ciò consente di eseguire rapidamente le operazioni di cifratura pur adoperando chiavi diverse e di notevole lunghezza. La cifratura per sostituzione può anche aver luogo per poligrammi, cioè per gruppi di un numero fisso di lettere, e per frazioni di lettere, cioè sostituendo le singole lettere del testo chiaro con gruppi di lettere o di cifre o di altri segni convenzionali che si sottopongono poi a seconda cifratura. La cifratura potrebbe anche aver luogo per sillabe, ma i sistemi di questo tipo sono rarissimi e invero poco pratici. Le macchine cifranti moderne sono naturalmente i computer, ove tutto è veloce e dove la grande mole di lavoro non conta.
I sistemi a *repertorio*, che si ritengono ideati nel sec. XVII, sono di larghissimo uso nei tempi moderni, in quanto consentono maggior garanzia di sicurezza in confronto dei sistemi letterali e producono, d'altra parte, una notevole economia di spese telegrafiche. I repertori, altrimenti chiamati codici, vocabolari telegrafici, dizionari cifrati, ecc., sono libri o anche programmi di computer contenenti un certo numero di voci, a ciascuna delle quali corrisponde un gruppo cifrante, composto generalmente di quattro o cinque lettere dell'alfabeto o di cifre arabe.

Subito dopo la seconda guerra mondiale si assistette al caso di Pearl Harbor, base militare aereo−navale statunitense dal 1887. Attaccata il 7 dicembre 1941 dalle forze giapponesi diede il via alle ostilità nippo-statuntensi. Ma i messaggi erano stati decrittati ed i crittoanalisti sapevano già dell'attacco.

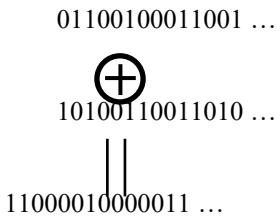
Inoltre si capì che un messaggio, attraverso opportuni codici, poteva essere scritto come una sequenza di 0 e di 1, utilizzando il codice ASCII o qualsiasi sua variante. Del resto, se si pensa ad una codifica tipo Vigenère si può dimostrare che se la chiave ha la stessa lunghezza del testo e la sequenza $k_1, k_2, \ldots, k_n$ è casuale (cioè prodotta dal meccanismo di testa e croce), allora il codice non si può rompere indipendentemente dalle operazioni fatte (teorema di Turing). Per cui si cercò di cambiare sistema considerando sia il testo chiaro che la chiave come una sequenza di 0 e di 1. L'addizione è quella di Boole, cioè:

$$0 + 0 = 0,\ 1 + 1 = 0,\ 0 + 1 = 1,\ 1 + 0 = 1$$

IL SISTEMA ONE-TIME PAD

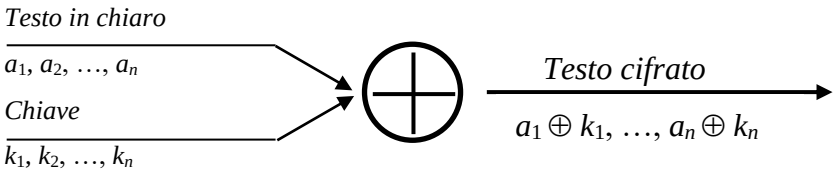


Infatti risulta:



poiché 0 (del testo in chiaro) + 1 (della chiave) = 1 (del testo cifrato), 1 + 0 = 1, 1 + 1 = 0, 0 + 0 = 0, e così via.

IL SISTEMA VERNAM



I metodi sopra illustrati sicuramente sono tra i più sicuri ma non consentono di trasmettere chiavi in tempo reale; attualmente, quindi, non è ancora stata trovata una soluzione del problema ma sono state scoperte delle chiavi cosiddette pseudocasuali.

4. Il problema statistico

Una simpatica storiella crittografica mostra come, a volte, non bisogna fidarsi di piccole statistiche per generalizzare un problema.

Esempio

Un intruso, volendo carpire la parola d'ordine di un accampamento militare, si nasconde dietro un cespuglio, in prossimità dell'ingresso dell'accampamento, ed ascolta quanto segue:

la sentinella domanda: dodici?
il 1° soldato risponde: sei! e passa;

la sentinella domanda: dieci?
il 2° soldato risponde: cinque! e passa;

la sentinella domanda: otto?
il 3° soldato risponde: quattro! e passa;

la sentinella domanda: sei?
il 4° soldato risponde: tre! e passa;

la sentinella domanda: quattro?
l'intruso risponde: due!

e così viene fulminato da un colpo di fucile.
Per passare senza alcun problema, infatti, l'intruso doveva pronunciare il numero delle lettere componenti il numero della domanda, quindi, nel caso specifico, sette invece di due!
Dunque, non bisogna mai accontentarsi di una piccola statistica ma occorrono numeri sufficientemente grandi per poter effettuare in modo corretto le statistiche.
La decrittazione dei crittogrammi é la traduzione di essi in linguaggio chiaro, eseguita da chi non sia a conoscenza dei cifrari e delle chiavi costituenti la base del segreto.
I metodi usati per decrittare i messaggi si basano su considerazioni statistiche, in particolare sulle caratteristiche di ciascuna lingua, cioè sul fatto che in ogni lingua le singole lettere, alcuni bigrammi e trigrammi e certe parole si ripetono più frequentemente di altre. Si chiama logoscopico il calcolo statistico delle parole, frasi e periodi che più frequentemente si riscontrano nel linguaggio. Il lavoro di decrittazione consiste in successive induzioni e deduzioni in merito al presumibile significato dei testi presi in esame e può essere agevolato da alcune circostanze favorevoli, quali il possesso di più testi cifrati relativi allo stesso testo chiaro, ma ottenuti con cifrari diversi, la conoscenza anche vaga del sistema di cifratura adottato, la conoscenza parziale o totale del testo chiaro corrispondente a qualche testo cifrato del quale si sia in possesso. Quindi cifrando un testo chiaro, se ad ogni lettera**,** o gruppo di lettere o parole si sostituisce un

determinato segno, questo si ripeterà con la stessa frequenza del testo chiaro, cosicché alla lettera o alla cifra più frequentemente ripetuta nel testo cifrato, con ogni probabilità corrisponderà la lettera o la parola più frequente nel linguaggio chiaro, tenendo conto chiaramente del fatto che la frequenza cambia in base alla lingua utilizzata, ovvero è costante solo all'interno di precisi margini presi in considerazione.

Le basi linguistiche della decrittazione consistono nelle caratteristiche particolari di ciascuna lingua e cioè nelle sequenze percentuali delle lettere, dei bigrammi e trigrammi più comuni e di alcune parole, nelle sequenze percentuali delle lettere e di alcune parole tra di loro, nelle sequenze obbligate o molto probabili, in quelle escluse o assai poco probabili, nelle terminazioni più frequenti delle parole, ecc. I dati caratteristici delle principali lingue sono contenuti in varie opere di crittografia.

Ad esempio nella lingua italiana il 45% delle lettere sono vocali e le sequenze con le quali si ripetono le varie lettere sono le seguenti:

A = 10, 41%	B = 0, 95%	C = 4,28%
D = 3, 82%	E = 12, 63%	F = 0, 75%
G = 2, 01%	H = 1, 10%	I = 11, 62%
L = 6, 61%	M = 2, 58%	N = 6, 49%
O = 8, 71%	P = 3, 26%	Q = 0, 57%
R = 6, 70%	S = 6, 04%	T = 6, 06%
U = 3, 04%	V = 1, 51%	**Z =** 0, 93%

Se, quindi, volessimo scrivere in sequenza le lettere dell'alfabeto italiano in base alla frequenza decrescente, avremmo:

E = 12, 63%	T = 6, 06%	G = 2, 01%
I = 11, 62%	S = 6, 04%	V = 1, 51%
A = 10, 41%	C = 4,28%	H = 1, 10%
O = 8, 71%	D = 3, 82%	B = 0, 95%
R = 6, 70%	P = 3, 26%	**Z =** 0, 93%
L = 6, 61%	U = 3, 04%	F = 0, 75%
N = 6, 49%	M = 2, 58%	Q = 0, 57%

Esempio

Nel seguente messaggio cifrato:

VSNHQ HPNFM HVHLA RQRQR VZUND LHQZN

si contano 5 H, 4 N, 4 Q, 3 R e 3 V alcune delle quali con ogni probabilità saranno vocali. Perciò quasi sicuramente la H corrisponderà alla "E". Risulta probabile, quindi, che, dato l'argomento in questione, la prima parola sarà "spie" e quindi V = S ed S = P. Perciò il crittogramma è stato ottenuto con un alfabeto spostato di tre posti. Un'accurata verifica ci darà il seguente testo chiaro: "Spie nemiche seguono nostri agenti".

Bisogna però precisare che i metodi di decrittazione sopra riportati sono i più semplici da usare. Ove non si conosca o non si presuma a quale sistema crittografico appartiene il metodo di cifratura, si deve anzitutto eseguire questa indagine, la quale spesso non presenta grandi difficoltà, possedendo i sistemi più usuali delle caratteristiche particolari. Non è raro tuttavia, e ciò avviene generalmente per i sistemi misti, che nel testo non si rilevino sufficienti indizi per la determinazione del sistema di cifratura; in questi casi il lavoro di crittoanalisi è, naturalmente, più difficile e può anche non portare, in mancanza di circostanze favorevoli, a un pratico risultato.

La decrittazione dei sistemi letterali a trasposizione si tenta mediante successive disposizioni e spostamenti delle lettere del testo cifrato, sino a quando non si riscontra qualche combinazione di lettere, parte di parola o parola, che valga a mettere sulla buona via.

Per i sistemi letterali a costituzione monoalfabetica la decrittazione consiste nella ricostruzione dell'alfabeto cifrante.

Sia dato ad esempio il seguente crittogramma, che si presume riferirsi a un testo chiaro scritto in lingua italiana:

KSAJOHTQSVTKHXHJKTQPZJHQMTZFJATXS
ZQTJNNZTXJDTKHSXHFTPHMHEYSNZTDTE
TKKTETQEHKHJOHTQSILAASZHDTETQMSZS
QOJNJQJFSZHEJQJ

Il calcolo delle frequenze letterali dà il seguente risultato: 17 T; 13 H; 12 J; 10 Q; 10 S; 8 Z; 7 K; 6 E; 4 A; 4 N; 4 X; 3 D; 3 F; 3 M; 3 O; 2 P; 1 I; 1 L; 1 V; 1 Y.

Il confronto delle suddette frequenze con quelle medie della lingua italiana (e = 12, 6%; i = 11,6%; a = 10, 3%; o = 8, 7% ecc.) fa presumere che quattro delle lettere T, H, J, Q, S rappresentino le 4 vocali a, e, i, o; nel testo si riscontrano i bigrammi HT, HJ, HS, dei quali i primi due ricorrono due volte ciascuno, e poiché i dittonghi più frequenti in italiano sono io, ia, ie, può ritenersi H = i. I trigrammi HTQ e JQJ, che ricorrono due volte l'uno, e JHQ inducono a presumere, in considerazione della scarsa frequenza dei dittonghi, che Q rappresenti una consonante e, dato ciò, può presumersi T = o e Q = N. Il gruppo HTQS (ione) che ricorre due volte, fa ritenere S = e e conseguentemente J = a, dopo di che la posizione e la frequenza della Z fanno presumere Z = R. A questo punto il crittogramma presenta l'aspetto seguente:

K S A J O H T Q S V T K H X H J K T Q P Z J H Q M T Z F J A T X S
  e   a   i o n e   o   i   i a   o n   r a i n   o r   a   o   e

Z Q T J N N Z T X J D T K H S X H F T P H M H E Y S N Z T D T E
r n o a   r o   a   o   i e   i   o   i   i   e r o   o

T K K T E T Q E H K H J O H T Q S I L A A S Z H D T E T Q M S Z S
o   o   o n   i   i a   i o n e   e r i   o o n   e r e

Q O J N J Q J F S Z H E J Q J
n   a   a n a   e r i   a n a

Riesce ora facile completare il lavoro di decrittazione, dal quale il testo chiaro e l'alfabeto convenzionale risultano ricostituiti così:

Testo chiaro: "Legazione Bolivia Londra informa Governo approvato lievi modifiche protocollo conciliazione suggerito conferenza panamericana".
Alfabeto chiaro: 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
Alfabeto cifrante: JVEPSMAYH KFQT N ZIDLX O

La decrittazione dei sistemi letterali a sostituzione polialfabetica si fonda sui seguenti principi, rilevati dal Kasiski e dal Kerckhoffs:

a. in qualunque testo cifrato due poligrammi simili sono, salvo casi eccezionali, il prodotto di due gruppi di lettere uguali cifrati con alfabeti uguali;
b. il numero delle cifre compreso nell'intervallo dei due poligrammi simili è un multiplo del numero di lettere della chiave.

Determinata la lunghezza della chiave, ciò che equivale a determinare il numero degli alfabeti cifranti, si divide il testo in gruppi di lunghezza pari a quella della chiave e si calcolano le frequenze delle lettere che occupano nei gruppi lo stesso posto; confrontando le frequenze suddette con quelle medie della lingua è possibile quasi sempre intuire il significato delle lettere più frequenti e formare in tal modo, in successivi tentativi, gruppi di lettere, parti di parole e parole, tenendo conto del presumibile contenuto del messaggio.
Se gli alfabeti convenzionali sono disposti normalmente (tabella del Vigenère e simili) la conoscenza del significato di una lettera in un alfabeto ha per conseguenza quella del valore di tutte le altre lettere nello stesso alfabeto. Ciò non si verifica per i sistemi ad alfabeti intervertiti, nei quali però, se l'interversione è regolare, si ha l'equidistanza relativa delle varie lettere in tutti gli alfabeti e conseguentemente si deduce, dal significato di una lettera in più alfabeti cifranti, il valore che hanno in tutti questi alfabeti le lettere di cui si conosca il significato soltanto per uno di essi.
Per i sistemi a rappresentazione numerica, tanto monoalfabetici quanto polialfabetici, i metodi di decrittazione non sono diversi da quelli innanzi indicati per le rispettive categorie di sistemi a rappresentazione letterale.
La decrittazione dei sistemi a repertorio si tenta in base al confronto della frequenza dei gruppi cifranti con quella delle parole di uso più frequente nella lingua ovvero delle cosiddette parole vuote più importanti (articoli, verbi ausiliari, preposizioni, congiunzioni). La posizione di questi gruppi nel testo, il presumibile contenuto del messaggio, ed eventualmente la conoscenza, parziale o totale, di qualche testo chiaro corrispondente a un testo cifrato di cui si disponga, danno modo di intuire il significato di altri gruppi e di procedere nel lavoro di decrittazione.
La decrittazione dei repertori ordinati e non sopracifrati è in linea di massima possibile, ove si disponga di testi aventi complessivamente una lunghezza opportuna, dato che la conoscenza o presunzione del significato di alcuni gruppi cifranti costituisce una guida di grande importanza.
Circostanza particolarmente favorevole è l'esistenza di serie di gruppi che facciano pensare ad una parola, generalmente nome proprio, cifrata mediante cifratura delle lettere o delle sillabe. Per esempio, se in un testo ottenuto con un repertorio di 10000 voci che si ritenga in lingua italiana e in ordine normale, si riscontra una serie di gruppi come la seguente, si deve presumere che essa rappresenti un nome proprio e che i singoli gruppi corrispondano ad una vocale o ad una consonante secondo queste indicazioni:

3733	3409	7462	7462	0002	7462	4517
conson	vocale	conson	conson	vocale	conson	vocale

Il gruppo 0002 è senza dubbio uguale ad a; i gruppi 3409 e 4517, data la distanza dall'inizio del cifrario e la loro distanza rispettiva equivalgono presumibilmente ad e ed i; la consonante rappresentata da 3733 è molto probabilmente f. Il gruppo 7462 sarà presumibilmente corrispondente a p, r, s, ma, poiché delle tre ipotesi "Feppapi", "Ferrari", "Fessasi" la seconda è evidentemente la più probabile, può concludersi ritenendo 7462 = r. In tal modo si pongono dei punti di riferimento abbastanza attendibili per la ricostruzione del cifrario ed è agevole proseguire il lavoro di decrittazione, deducendo anzitutto, in base al calcolo delle frequenze, il significato delle parole vuote più importanti.

Per i repertori paginati e sempre non sopracifrati si procede in maniera analoga, tenendo conto non soltanto delle frequenze dei singoli gruppi cifranti ma anche di quelle complessive dei gruppi contenuti in ogni pagina.

Molto più difficile è la decrittazione dei repertori intervertiti e non sopracifrati per la mancanza di qualsiasi ordine nella disposizione dei gruppi cifranti corrispondenti alle voci chiare; il lavoro, generalmente, può avere pratico risultato soltanto ove ricorra qualche circostanza favorevole o soccorra un'intuizione particolarmente felice, e comunque disponendo di un adeguato numero di messaggi cifrati.

Il problema della decrittazione dei sistemi a codice (intervertito o non) sopracifrato è di estrema complessità ed è anche piuttosto lungo descrivere casi particolari per farsi un'idea. Comunque, va detto che un tale problema è di fatto teoricamente impossibile se non vi sono sovrapposizioni nell'impiego del verme di sopracifratura. Se sovrapposizioni vi sono, si riesce ad eliminare la complicazione dovuta al verme, lavorando sulle "differenze" fra i gruppi costituenti la differenza prima dei messaggi. Il "procedimento" usato è piuttosto complesso ed è praticamente impossibile descriverlo in questo lavoro.

Il problema è molto più complicato che nei codici di Cesare o comunque monoalfabetici. Si pensi che il codice di Vigenère è stato usato da vari eserciti resistendo ai vari "attacchi" che i crittoanalisti gli hanno dato per ben tre secoli. Nel 1863 un ufficiale prussiano, Kasiski, forza il codice di Vigenère con un metodo noto come test di Kasiski. Il test di Kasiski è basato sui seguenti punti:

- il messaggio è ”sufficientemente lungo”;
- Il primo passo è "trovare la lunghezza della parola chiave";
- Il secondo passo è "trovare le lettere della parola chiave".

Circa il primo passo iniziamo con il ricercare nel messaggio cifrato tutte le sequenze (cioè i gruppi) di tre o più lettere consecutive che si ripetono. E' molto probabile che a sequenze uguali del cifrato corrispondano sequenze uguali del testo in chiaro. In ogni caso è questa una ipotesi di lavoro compatibile con la lunghezza del messaggio. Se è così vuol dire che le prime lettere di sequenze uguali sono state criptate con lo stesso alfabeto della tavola di Vigenère, analogamente le seconde, le terze, ecc. Da ciò segue che alle prime lettere delle sequenze uguali corrisponde la stessa lettera della parola chiave, analogamente alle seconde lettere e così via. Ma allora, se la distanza (= numero delle lettere) tra due sequenze uguali è un multiplo della lunghezza della parola chiave, quanto è lunga la parola chiave? Molto probabilmente la sua lunghezza è pari al M.C.D. delle distanze tra le sequenze uguali tra loro, che si ripetono.

Chiaramente una distanza "strana" tra due sequenze uguali, nel senso che non ha divisori comuni con le altre distanze, deve essere scartata; questo è il motivo per cui abbiamo detto "molto probabilmente la sua lunghezza …" Ciò accade quando due sequenze uguali non corrispondono a due sequenze uguali del testo in chiaro.

Trovata la lunghezza d della parola chiave cerchiamo le lettere che la compongono. Notiamo che, nel testo cifrato, alla prima lettera, alla (d + 1**)**-ma lettera, alla (2d +1)-ma lettera, ... corrisponde la stessa lettera della parola chiave e quindi

tutte queste lettere sono state criptate con uno stesso codice di Cesare (cioè con una stessa riga del quadrato di Vigenère). Segue, ripetendo il ragionamento, che la seconda, la (d + 2)-ma, la (2d + 2**)-**ma riga sono anche loro crittografate usando uno stesso codice di Cesare. In definitiva ciascuna delle righe seguenti sono crittografate con una stessa lettera della parola chiave:

{I, (*d* + 1)–ma, (2*d* + 1)–ma, …}
{II, (*d* + 2)–ma, (2*d* + 2)–ma, …}
{III, (*d* + 3)–ma, (2*d* + 3)–ma, …}
……..
……..
{*d*–ma, 2*d*–ma, 3*d*–ma, …}

Le lettere di ognuno di questi insiemi sono state criptate con lo stesso codice di Cesare. Allora studiamo la frequenza delle lettere in ognuno di essi, con lo stesso metodo usato nei codici di Cesare. Così scoperta una lettera, è noto il codice di Cesare usato, e quindi anche il nome dell'alfabeto (cioè, la lettera che compare in testa), che ci dà la lettera della parola chiave. In questo modo il codice di Vigenère è forzato e quindi perde il suo interesse.
E proprio dall'idea di Vigenère nacque un codice completamente sicuro, precisamente il codice di Vernam (1926), sopra analizzato.

BIBLIOGRAFIA

Ambrisi -Eugeni,

Ambrisi E.-.Eugeni F., Il problema della protezione dell'Informazione, Ratio Math.,(1990) pp.15,37.

AA. VV., *L'identité. Séminaire interdisciplinaire dirigé par Claude Lévi-Strauss*, Paris 1977

AA.VV., *Ratio Math. (fasciolo dedicato alla Crittografia*), n.6 (1993) *Contiene articoli di Berardi L-, Dass B.K, Donini L. (Ammiraglio), Elia M., Eugeni F., Innamorati S., Manara C.F., Maturo A., Zuanni F. Moro G. (ammiraglio).*

Berardi L., *Algebra e teoria dei codici correttori*, (1994), Franco Angeli, Milano.

Berardi L., Beutelspacher A., *Crittologia* , (1996), Franco Angeli, Milano.

Berardi L., Beutelspacher A., *Matematica Discreta*, (2003), Franco Angeli, Milano.

De Finetti B., *Macchine "che pensano" (e che fanno pensare)*, "Tecnica ed Organizzazione", n. 2, 1952, p. 31/70.

Eugeni -D.- Eugeni F., *Il codice di Leon Battista Alberti*, Ratio Math.(1994), pp179-186

Eugeni F. e Mascella R., *Società e fondamenti dell'informatica*, Zikkurat, Teramo (in preparazione).

Ginzburg, C., "*Spie. Radici di un paradigma indiziario"*, in A. Gargani, a cura, *Crisi della ragione*, Einaudi, Torino 1979

Kuhn, T.S., *La struttura delle rivoluzioni scientifiche*, Einaudi, Torino 1969, (*The Structure of Scientific Revolution*s, Chicago 1962)

Tallini L., *Sui codici unidirezionali*, Ratio Math., n.4 (1992).