

## TEMA 5: INTRODUCERE ÎN CRIPTOGRAFIE

### Obiectivele temei:

- Familizarizarea cu obiectivele și noțiunile de bază ale criptografiei.
- Analiza cifrurilor clasice.
- Studiarea mecanismelor de criptare.
- Introducere în atacurile criptografice.

### Cuvinte-cheie:

cifru,	cifru polialfabetic,
sistem de criptare,	mașini rotor,
text clar,	criptanaliză,
text criptat,	atacuri criptografice,
substituția,	forță brută ( metoda
transpunerea,	exhaustivă).
cifru monoalfabetic,	

Există dovezi că criptografia ca metodă de protecție a textului s-a născut odată cu apariția scrisului, iar metodele de secretizare a mesajelor erau deja cunoscute de civilizațiile antice din India, Egipt și Mesopotamia. Cele mai vechi texte indiene printre cele 64 de arte menționau și metodele de modificare a textului, unele dintre care pot fi considerate metode criptografice. Pentru a ascunde ce a scris, autorul unei plăci din Mesopotamia ce conținea o rețetă pentru fabricarea glazurii pentru ceramică utiliza semne necunoscute, sărea peste unele litere, iar numele le înlocuia cu cifre. De atunci au fost găsite mai multe referiri la criptografie, cea mai mare parte a având o utilizare militară.

În zilele noastre Criptografia oferă unele dintre cele mai bune metode (dacă nu chiar cele mei bune) de protecție a informației. Metodele criptografice sunt aplicate în rețele și pe WEB, ele sunt implementate în aplicații software și pe dispozitive hardware, devenind deja un accesoriu indispensabil al tehnologiilor moderne.

### 5.1. Obiective și noțiuni de bază ale criptografiei

#### 5.1.1. Noțiune de criptografie

*Criptografia* reprezintă o ramură a matematicii care se ocupă cu securizarea informației, asigurând de asemenea autentificarea și restricționarea accesului într-un sistem informatic. În realizarea acestora se utilizează în mare parte metode matematice, profitând de unele probleme cu complexitate de rezolvare suficient de înaltă. Termenul „criptografie” este compus din cuvintele de

origine greacă  $\kappa\rho\upsilon\pi\tau\acute{o}\varsigma$ , *kryptós* (ascuns) și  $\gamma\rho\acute{\alpha}\phi\epsilon\iota\nu$ , *gráfein* (a scrie). Criptografia modernă urmărește realizarea următoarelor obiective:

- *confidențialitatea (privacy)* – proprietatea de a păstra secretul informației, pentru ca aceasta să fie folosită numai de persoanele autorizate;
- *integritatea datelor (integrity)* – proprietatea de a evita orice modificare (inserare, ștergere, substituție) neautorizată a informației;
- *autenticitatea (authenticity)* – proprietatea de a identifica o entitate conform anumitor standarde, care include autenticitatea unei entități sau a sursei informației;
- *non-repudierea (nonrepudiation)* – proprietatea care previne negarea unor evenimente anterioare.

Celelalte obiective legate de securitatea informației (*autentificarea mesajelor, semnături, autorizare, validare, controlul accesului, certificare, timestamping, confirmarea recepției, anonimitate, revocare*) pot fi derivate din aceste patru.

Împreună cu Criptografia se dezvoltă *Criptanaliza* – (din greacă, *kryptós*, „ascuns”, și *analýein*, „a dezlega”) studiul metodelor de obținere a înțelesului informațiilor criptate, fără a avea acces la informația secretă necesară în mod normal pentru aceasta. De regulă, aceasta implică găsirea unei chei secrete. Criptografia și Criptanaliza împreună constituie *Criptologia* (din greacă, *kryptós*, „ascuns”, și  $\lambda\acute{o}\gamma\omicron\varsigma$ , „cuvânt”) – știința care se ocupă cu metodele de criptare și decriptare.

În continuare sunt date unele noțiunile fundamentale cu care se operează în criptologie.

O mulțime nevidă  $T$  se numește *alfabet*.

Elementele alfabetului  $T$  se numesc *litere*.

O consecutivitate finită de elemente din alfabetul  $T$  se numește *cuvânt*. Una și aceeași literă poate intra într-un cuvânt de mai multe ori.

Numărul de elemente ale alfabetului se numește *lungimea alfabetului*.

Un cuvânt ce nu conține nici o literă se numește *cuvânt nul*.

*Lungimea cuvântului*, notată deseori cu  $w$ , este numărul de litere în acest cuvânt, unde fiecare literă se consideră de câte ori se întâlnește în el.

Vom nota cu  $T^*$  mulțimea tuturor cuvintelor alfabetului  $T$ . Submulțimile mulțimii  $T^*$  le vom numi *limbaje* (formale) peste  $T$ .

Un mesaj în forma sa originală se numește *text clar* (uneori *text în clar* sau *text plan*, în engleză *plaintext*) și îl vom nota cu  $pt$  sau cu  $m$  sau  $M$  (de la „*message*” - mesaj).

Rescrierea textului clar, folosind o metodă cunoscută numai de expeditor (eventual și de destinatar), se numește *criptare* (sau *cifrare*) a mesajului. Scopul rescrierii textului clar are ca scop camuflarea *textului clar* în așa fel încât substanța să nu sufere modificări semantice pentru ca mesajul să poată fi restabilit de destinatar în versiunea inițială a acestuia.

*Text criptat* sau *text cifrat* (în engleză *ciphertext*) se numește textul obținut în rezultatul operației de criptare a textului plan. Textul criptat îl vom nota cu  $ct$ , cu  $c$  sau cu  $C$  (de la „*cipher*” - cifrul). Textul cifrat se mai numește *criptogramă*.

Procesul retransformării criptogramei în textul original este numit *decriptare* sau *descifrare*.

În procesul schimbului de informație criptată destinatarul primește printr-un canal textul criptat de la expeditor și îl decriptează, știind metoda folosită pentru criptare, iar în rezultat el obține mesajul inițial. În literatura de specialitate expeditorul de obicei este numit *Alice* iar destinatarul este numit *Bob*. Deci, Alice și Bob trebuie să stabilească în prealabil detaliile modalității de criptare și de decriptare.

Criptarea se folosește pentru a fi siguri că informația este inaccesibilă oricărei persoane care nu deține instrumentul necesar decriptării. Chiar dacă oricine poate vedea datele în formă criptografică, oricum nu va înțelege nimic, care să conducă spre descifrarea textului original. Persoana care interceptează criptograma și încearcă să obțină textul clar aplicând diverse metode, însă fără a avea cheia de decriptare, este numită *criptanalist*.

### 5.1.2. Sisteme de criptare

Un sistem care realizează operațiile de criptare și decriptare se numește *sistem de criptare* (sau *sistem criptografic*, sau *criptosistem*). În criptografia modernă un sistem de criptare este definit ca o structură cu cinci componente ( $P, C, K, E, D$ ):

- $P = \{m : m \in T^*\}$  – spațiul (mulțimea) textelor în clar, scrise pentru un alfabet nevid  $T$  (în mod obișnuit  $T = \{0,1\}$ );
- $K$  – spațiul (mulțimea) cheilor de criptare  $k, k \in K$ ;
- Familia funcțiilor de criptare dependentă de chei și de un algoritm de criptare  $E$

$$E_k : P \rightarrow C, E_k = \{e_k : e_k(m) = c \text{ și } e_k \text{ este injectivă}\};$$

- Familia funcțiilor de decriptare dependentă de chei și de un algoritm de decriptare  $D$

$$D_k : C \rightarrow P, D_k = \{d_k : d_k(e_k(m)) = m \text{ pentru orice } m \in P\};$$

- $C$  spațiul (mulțimea) mesajelor cu text criptat unde:

$$C = \{c : \text{există } k \in K, m \in P, c = E_k(m)\}.$$

Schema aplicării unui sistem de criptare este prezentată în figura 5.1.

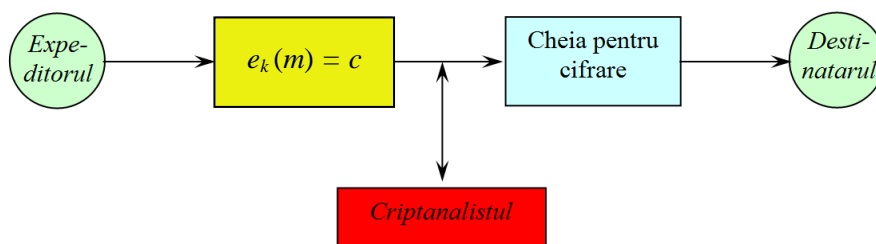


Figura 5.1. Schema aplicării unui sistem de criptare

Un sistem de criptare este realizat în baza a trei algoritmi: algoritmul de criptare  $e_k$ , algoritmul de decriptare  $d_k$  și algoritmul de generare a cheilor.

Pentru ca un sistem de criptare să fie considerat bun, el trebuie să îndeplinească trei criterii (enunțate de Francis Bacon în sec. XVII):

1. Fiind date  $e_k$  și  $m \in \mathbf{P}$  să fie ușor de calculat  $e_k(m)$ .
2. Fiind date  $d_k$  și  $c \in \mathbf{C}$  să fie ușor de determinat  $d_k(c)$ .
3. Să fie imposibil de determinat  $m$  din  $c$ , fără a cunoaște  $k$ .

Criteriile 1 și 2 implică faptul că pentru utilizatorii legali algoritmi de criptare și decriptare nu trebuie să fie prea complicați (se presupune că utilizatorii au un timp acceptabil pentru calcule). În criteriul 3 „imposibilitatea” e înlocuită în prezent cu „dificultatea de a calcula”. Se presupune că un interceptor de asemenea are acces la tehnica de calcul. Ultimul criteriu definește (sub o formă vagă) ideea de ”securitate” a sistemului. La aceste criterii, Bacon adăuga și o a patra regulă:

4. Textul criptat trebuie să fie un text banal, fără suspiciuni.

Această condiție nu mai poate fi considerată importantă și este utilizată astăzi doar de un subdomeniu strict al criptografiei, numit *steganografie* – știința despre transmiterea secretă a informației prin păstrarea secretului a însuși faptului transmiterii acestei informații.

### 5.1.3. Tipuri de cifruri

Metodele de criptare pot fi divizate pe categorii în felul următor (figura 5.2):

a) *în funcție de tipul operațiilor folosite:*

- bazate pe substituții;
- bazate pe transpuneri;

b) *în funcție de tipul de chei folosite:*

- metode simetrice (cu cheie secretă);
- metode asimetrice (cu cheie publică);

c) *metoda prin care datele sunt procesate:*

- cu cifruri bloc;
- cu cifruri fluide (flux, șir).

Cifrurile clasice foloseau substituția sau transpoziția. Printre metodele moderne de criptare deosebim două direcții principale: *cifruri cu cheie secretă* (sau *cifruri simetrice*) în care  $e_k$  este bijectivă și *cifruri cu chei publice* (sau *cifruri asimetrice*) – cifruri în care  $e_k$  nu este bijectivă.

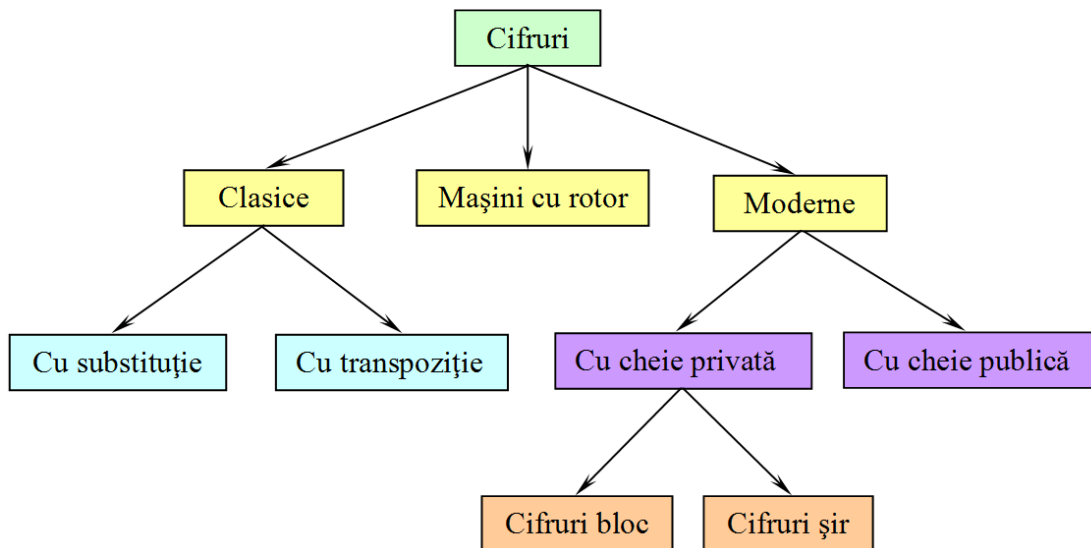


Figura 5.2. Clasificarea cifrurilor

Cifrurile simetrice pot fi de *tip bloc* sau de *tip șir* (sau *flux*, sau *fluide*, în engleză *stream cipher*). Algoritmii de tip bloc acționează asupra blocurilor de text clar și text cifrat. Algoritmii de tip șir se aplică șirurilor de text clar și text cifrat, la nivel de bit sau octet.

Algoritmii moderni de tip bloc criptează mesajul în blocuri de 64 – 265 biți (lungimea de 64 biți este deja considerată insuficientă, standardul curent de cifrare simetrică – *AES* – lucrează cu blocuri de lungime 128, 192 sau 256 biți). Pentru acesta se aplică o funcție matematică între un bloc de biți ai mesajului în clar și cheie (care poate varia ca mărime), rezultând același număr de biți pentru mesajul criptat. Funcția de criptare este realizată astfel încât să îndeplinească următoarele cerințe:

- cunoscând un bloc de biți ai textului clar și cheia de criptare, algoritmul să poată genera rapid un bloc al textului criptat;
- cunoscând un bloc de biți ai textului criptat și cheia de criptare/decriptare, sistemul să poată genera rapid un bloc al textului clar;
- cunoscând blocurile textului clar și ale textului criptat, să fie dificil de generat cheia.

Cifrurile șir la fel formează o clasă importantă de algoritmi de criptare. Ceea ce le caracterizează și le diferențiază față de cifrurile bloc este faptul că cifrurile șir procesează informația în unități oricât de mici, chiar bit cu bit, aplicând funcția XOR între biții cheii și biții de cifrat, iar funcția de criptare se poate modifica în cursul criptării. Cifrurile șir sunt algoritmi cu memorie, în sensul că procesul de criptare nu depinde doar de cheie și de textul clar, ci și de starea curentă. În cazul în care probabilitatea erorilor de transmisie este mare, folosirea cifrurilor șir este avantajoasă deoarece au proprietatea de a nu propaga erorile. Ele se folosesc și în cazurile în care datele trebuie procesate una câte una, datorită lipsei de spațiu de memorie.

Cifruri *asimetrice* utilizează o pereche de chei: o *cheie publică* și o *cheie privată*. Un utilizator care deține o astfel de pereche își publică o cheie (cheia publică) astfel încât oricine

dorește să o poată folosi pentru a-i transmite un mesaj criptat. Numai deținătorul cheii secrete (private) este cel care poate decripta mesajul astfel criptat.

Cele două chei sunt legate matematic, însă cheia privată nu poate fi obținută din cheia publică. În caz contrar, oricine ar putea decripta mesajele destinate unui alt utilizator, fiindcă oricine are acces la cheia publică a acestuia. O analogie foarte potrivită pentru proces este folosirea cutiei poștale. Oricine poate pune în cutia poștală a cuiva un plic, dar la plic nu are acces decât posesorul cheii de la cutie.

Criptografia asimetrică se mai numește criptografie cu chei publice și e compusă din două mari ramuri:

- criptarea cu cheie publică – un mesaj criptat cu o cheie publică nu poate fi decodificat decât folosind cheia privată corespunzătoare. Metoda este folosită pentru a asigura confidențialitatea;
- semnături digitale – un mesaj semnat cu cheia privată a emițătorului poate fi verificat de către oricine, prin acces la cheia publică corespunzătoare, astfel asigurându-se autenticitatea mesajului.

## 5.2. Cifruri clasice

Criptografia clasică se încadrează în clasa criptografiei cu chei simetrice și este criptografia dinaintea calculatorului, de unde și denumirea de criptografie pre-computațională. În criptografia clasică, algoritmi erau bazați pe caracter și constau dintr-o serie de transformări elementare (substituții și transpoziții) ale caracterelor textului clar. Unii algoritmi aplicau aceste transformări în mod repetat, îmbunătățind în acest mod securitatea algoritmului. În criptografia modernă bazată pe calculator (criptografie computațională), lucrurile s-au complicat, dar multe dintre ideile criptografiei clasice au rămas nemodificate.

### 5.2.1. Cifrul de substituție

Cifrul de substituție (*substitution cipher*) este cifrul la care fiecare caracter sau grup de caractere ale textului clar  $m$  este substituit cu un alt caracter sau grup de caractere în textul cifrat  $c$ , descifrarea făcându-se prin aplicarea substituției inverse asupra textului cifrat. În criptografia clasică există patru tipuri de cifruri de substituție: monoalfabetică, polialfabetică, poligramică și homofonică.

*Cifruri de substituție monoalfabetică (monoalphabetic ciphers)* sunt cifrurile în care fiecare caracter al textului în clar  $m$  este înlocuit cu un caracter corespunzător în textul cifrat  $c$ . Reprezentanți ai acestei clase sunt: cifrul Cezar, cifrul Afin, cifrul Polibios, etc.

Cifrul lui *Cesar* (sau *Cezar*). În acest cifru fiecare literă a textului clar este înlocuită cu o nouă literă obținută printr-o deplasare alfabetică. Cheia secretă  $k$ , care este aceeași la criptare cât și la decriptare, constă în numărul care indică deplasarea alfabetică, adică  $k \in \{1, 2, 3, \dots, n-1\}$ , unde  $n$

este lungimea alfabetului. Criptarea și decriptarea mesajului cu cifrul Cezar poate fi definită de formulele

$$c = e_k(x) = x + k \pmod{n},$$

$$m = d_k(y) = y - k \pmod{n},$$

unde  $x$  și  $y$  sunt reprezentarea numerică a caracterului respectiv al textului clar. Funcția numită *Modulo* ( $a \pmod{b}$ ) returnează restul împărțirii numărului întreg  $a$  la numărul întreg  $b$ . Această metodă de criptare este numită așa după Iulius Cezar, care o folosea pentru a comunica cu generalii săi, folosind cheia  $k = 3$  (tabelul 5.1).

De exemplu, pentru  $k = 3$  avem

$$e_k(S) = 18 + 3 \pmod{26} = 21 = V$$

$$d_k(V) = 21 - 3 \pmod{26} = 18 = S$$

În acest caz pentru  $m = „cifrul\ cezard”$ , obținem  $c = „fliuxo\ fhcdud”$ .

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
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
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	0	1	2
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

Tabelul 5.1. Cifrul Cezar cu cheia  $k=3$

Cifrul lui Cezar este foarte ușor de spart, deci este un cifru foarte slab. Astfel, un criptanalist poate obține textul clar prin încercarea tuturor celor 25 de chei. Nu se știe cât de util era cifrul Cezar în timpul când era folosit de către cel de la care îi provine numele, dar este probabil ca el să fi fost destul de sigur, atât timp cât numai câțiva dintre inamicii lui Cezar erau în stare să scrie și să citească, dar mai ales să cunoască concepte de criptanaliză.

*Cifrul afin* este o generalizare a cifrului Cezar, având cheia

$$k = \{(a, b) \mid a, b \in \mathbb{Z}_{26} = \{0, 1, 2, \dots, 25\}, \text{cmmdc}(a, 26) = 1\},$$

iar funcțiile de criptare și decriptare pentru o cheie  $k = (a, b)$  sunt

$$e_k(x) = ax + b \pmod{26}$$

și

$$d_k(y) = a^{-1}(y + 26 - b) \pmod{26},$$

unde  $a \cdot a^{-1} = 1 \pmod{n}$ .

Condiția ca  $a$  să fie prim cu 26 asigură existența lui  $a^{-1}$  în  $\mathbb{Z}_{26}$ . Pentru calcularea lui  $a^{-1}$  poate fi aplicat *algoritmul Euclid extins* (vezi compartimentul 7), sau se verifică pe rând toate numerele reciproce prime cu  $a$  în clasa modulo  $n$  până când găsim acel număr pentru care  $a \cdot a^{-1} = 1 \pmod{n}$ , acesta fiind și unicul cu o astfel de proprietate.

De exemplu, pentru  $a = 7$ ,  $b = 16$  funcția de criptare este  $e_k(x) = 7x + 16$ . În acest caz pentru  $m = „cifrul\ afin”$ , obținem  $c = „euzfap\ qzud”$ .

Deoarece  $7 \cdot 15 = 1 \pmod{26}$ , avem  $7^{-1} = 15 \pmod{26}$ , iar decriptarea se realizează matematic folosind funcția

$$d_k(y) = 15 \cdot (y+26-15) \pmod{26} = 15y + 15 \cdot 11 \pmod{26} = 15y + 9 \pmod{26}.$$

Condiția  $\text{cmmdc}(a, 26) = 1$  asigură de asemenea injectivitatea funcției  $e_k$ , necesară pentru un proces corect de criptare-decriptare. De exemplu, pentru  $e_k(x) = 10x + 1$ ,  $A$  și  $N$  se transformă ambele în  $B$ , iar  $O$  nu apare ca imagine în alfabetul substituției. În acest caz avem:

$$\text{cmmdc}(a, 26) = \text{cmmdc}(10, 26) = 2 \neq 1.$$

Astfel de cazuri au loc întotdeauna atunci când  $a$  nu respectă condiția de mai sus.

Orice cheie  $k$  a cifrului afin este determinată complet de valorile întregi  $(a, b)$  pentru care  $\text{cmmdc}(a, 26) = 1$ . Sunt posibile 12 valori pentru  $a$ : 1, 3, 5, 7, 9, 11, 15, 19, 21, 23, 25 și 26 valori pentru  $b$ , care se iau independent de  $a$ , cu o singură excepție  $a = 1, b = 0$  (care se exclude deoarece nu conduce la nici o criptare). Așadar mulțimea cheilor în acest caz este alcătuită din  $12 \cdot 26 - 1 = 311$  chei diferite, număr suficient de mic pentru atacul prin forța brută, adică prin verificarea tuturor cheilor posibile.

Trebuie de avut în vedere că în cazul în care este necesar de a cripta mesaje în limba română sau rusă (sau oricare altă limbă care are un alfabet) procedul rămâne același, doar că trebuie să ținem cont de faptul că alfabetele respective au 31 și 33 de litere, adică  $n=31$  și respectiv  $n=33$ .

**Cifrul Polibios.** Cifrul Cezar nu este cel mai vechi algoritm de criptare. Se pare că primul astfel de algoritm a fost utilizat de Polybios (istoric grec, care a decedat cu 30 ani înaintea nașterii lui Cezar). Inițial acesta a fost doar un sistem maritim de semnalizare cu torțe iar ulterior i s-a dat o semnificație criptografică.

În cifrul Polibios pentru fiecare alfabet se construiește un careu aparte de cifrare, de cele mai dese ori cu numărul de coloane și linii egal (însă nu e o condiție necesară). Dimensiunile careului depind de lungimea  $n$  a alfabetului. Pentru a crea careul se iau două numere întregi, produsul cărora e cel mai aproape de  $n$ . Liniile și coloanele se numerotează. După aceasta literele alfabetului se înscriu în acest careu în ordinea lor naturală. Dacă nu sunt suficiente celule pentru literele alfabetului se pot înscrie într-o celulă 2 litere sau se pot omite câteva litere (de obicei de frecvența cât mai redusă). Pentru alfabetul latin, putem avea careuri Polibios  $5 \times 5$ , după cum este reprezentat în tabelul 5.2 (în al treilea careu a fost omisă litera Q care este una cu frecvență redusă):

	<i>m</i>	<i>o</i>	<i>u</i>	<i>s</i>	<i>e</i>
<i>m</i>	A	B	C	D	E
<i>u</i>	F	G	H	I/J	K
<i>s</i>	L	M	N	O	P
<i>c</i>	Q	R	S	T	U
<i>a</i>	V	W	X	Y	Z

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
<i>a</i>	A	B	C	D	E
<i>b</i>	F	G	H	I	J
<i>c</i>	K	L	M	N	O
<i>d</i>	P	R	S	T	U
<i>e</i>	V	W	X	Y	Z

Tabelul 5.2. Exemple de tabele ale cifrului Polibios



În operația de criptare, fiecare caracter  $x$  va fi reprezentat printr-o pereche de litere  $(x, y)$ , unde  $x \in \{m, o, u, s, e\}$  iar  $y \in \{m, u, s, c, a\}$  pentru primul exemplu din figură, sau  $x, y \in \{1, 2, 3, 4, 5\}$  și  $x, y \in \{a, b, c, d, e\}$  pentru exemplele doi și trei.

Astfel, pentru primul exemplu de tabel, textul clar VENI VIDI VICI este criptat în

*ma em us su ma su sm su ma su um su.*

Cifrul Polibius, prezentat în exemplul din mijloc al tabelului 5.2, a fost folosit de condamnații din penitenciarele rusești și de către prizonierii americani din Vietnam. Cifrul este foarte simplu de învățat și poate fi aplicat folosind diverse semne drept coordonate-chei (cifre, puncte, figuri, etc). Cifrul Polibios a fost utilizat de asemenea în cadrul altor sisteme de criptare, cum ar fi sistemul nihilist, cifrul ADFGVX (utilizat de armata germană în primul război mondial) sau sistemul Bifid.

Punctul slab al sistemelor de criptare monoalfabetice constă în frecvența de apariție a caracterelor în text. Dacă un text criptat este suficient de lung și se cunoaște limba în care este scris textul clar, sistemul poate fi spart printr-un atac bazat pe *frecvența apariției literelor* într-o limbă (atacul prin analiza frecvenței), această frecvență fiind o problemă studiată intens (nu neapărat în scopuri criptografice) iar în rezultat au fost construite diverse structuri de ordine relativ la frecvența apariției literelor în fiecare limbă europeană și în alte limbi.

De obicei, cu cât un text criptat este mai lung, cu atât frecvența literelor folosite se apropie de această ordonare generală. O comparație între cele două relații de ordine (cea a caracterelor din textul criptat și cea a literelor din alfabetul limbii curente) conduce la realizarea câtorva corespondențe (literă text clar – literă text criptat), ceea ce stabilește în mod univoc cheia de criptare. Pentru algoritmul Cezar este suficientă stabilirea unei singure perechi, pentru cel afin trebuie două perechi etc. Pentru limba română frecvența literelor este prezentată în tabelul 5.3 și figura 5.3.

A	Ă	Â	B	C	D	E	F	G	H	I	Î	J	K	L	M
9,95	4,06	0,91	1,07	5,28	3,45	11,47	1,18	0,99	0,47	9,96	1,40	0,24	0,11	4,48	3,10
N	O	P	Q	R	S	Ș	T	Ț	U	V	W	X	Y	Z	
6,47	4,07	3,18	0,00	6,82	4,40	1,55	6,04	1,00	6,20	1,23	0,03	0,11	0,07	0,71	

Tabelul 5.3. *Frecvența literelor limbii române*

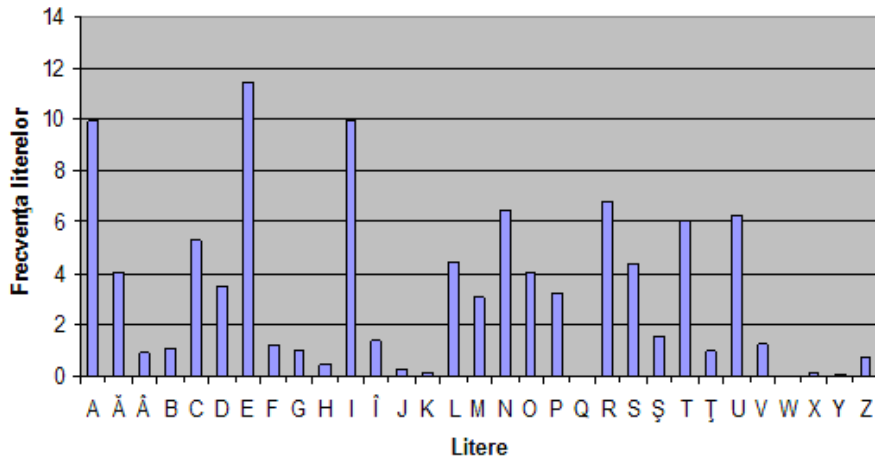


Figura 5.3. Frecvența literelor limbii române

*Cifruri de substituție polialfabetică (polyalphabetic ciphers).* Slăbiciunea cifrurilor monoalfabetice este definită de faptul că distribuția lor de frecvență reflectă distribuția alfabetului folosit. Un cifru este mai sigur din punct de vedere criptografic dacă prezintă o distribuție cât mai regulată, care să nu ofere informații criptanalistului.

O cale de a aplatiza distribuția este combinarea distribuțiilor ridicate cu cele scăzute. Dacă litera  $T$  este criptată câteodată ca  $a$  și altă dată ca  $b$ , și dacă litera  $X$  este de asemenea câteodată criptată ca  $a$  și altă dată ca  $b$ , frecvența ridicată a lui  $T$  se combină cu frecvența scăzută a lui  $X$  producând o distribuție mai moderată pentru  $a$  și pentru  $b$ . Două distribuții se pot combina prin folosirea a doua alfabete separate de criptare, primul pentru caracterele aflate pe poziții pare în textul clar, al doilea pentru caracterele aflate pe poziții impare, rezultând necesitatea de a folosi alternativ doua tabele de translatere, de exemplu permutările

$$p_1(a)=(3 \cdot a) \bmod 26 \text{ și } p_2(a)=(7 \cdot a + 13) \bmod 26.$$

Diferența dintre cifrurile polialfabetice și cele monoalfabetice constă în faptul că substituția unui caracter variază în text, în funcție de diverși parametri (poziție, context etc.). Aceasta conduce bineînțeles la un număr mult mai mare de chei posibile. Se consideră că primul sistem de criptare polialfabetic a fost creat de Leon Battista în 1568. Unele aplicații actuale folosesc încă pentru anumite secțiuni astfel de sisteme de criptare.

*Cifrul Vigenere.* La fel ca cifrul Cezar, cifrul Vigenere deplasează literele, dar, spre deosebire de acesta nu se poate sparge ușor în 26 combinații. Cifrul Vigenere folosește o deplasare multiplă. Cheia nu este constituită de o singură deplasare, ci de mai multe, fiind generate de câțiva întregi  $k_i$ , unde  $0 \leq k_i \leq 25$ , dacă luăm ca reper alfabetul latin cu 26 de litere. Criptarea se face în felul următor:

$$c_i = m_i + k_i \pmod{26}.$$

Cheia poate fi, de exemplu,  $k = (5, 20, 17, 10, 20, 13)$  și ar provoca deplasarea primei litere cu 5,  $c_1 = m_1 + 5 \pmod{26}$ , a celei de a doua cu 20,  $c_2 = m_2 + 20 \pmod{26}$ , ș.a.m.d. până la sfârșitul cheii

și apoi de la început, din nou. Cheia este de obicei un cuvânt, pentru a fi mai ușor de memorat – cheia de mai sus corespunde cuvântului „*furtun*”. Metoda cu deplasare multiplă oferă protecție suplimentară din două motive:

- primul motiv este că ceilalți nu cunosc lungimea cheii;
- cel de al doilea motiv este că numărul de soluții posibile crește; de exemplu, pentru lungimea cheii egală cu 5, numărul de combinații care ar fi necesare la căutarea exhaustivă ar fi  $26^5 = 11\,881\,376$ .

Decriptarea pentru cifrul Vigenere este asemănătoare criptării. Diferența constă în faptul că se scade cheia din textul cifrat,

$$m_i = c_i - k_i \pmod{26}.$$

Pentru simplificarea procesului de cifrare se poate utiliza următorul tabel, numit *Tabula Recta* (tabelul 5.4). Aici toate cele 26 cifruri sunt situate pe orizontală și fiecărui cifru îi corespunde o anumită literă din cheie, reprezentată în colana din stânga tabelului. Alfabetul corespunzător literelor textului clar se află în prima linie de sus a tabelului. Procesul de cifrare este simplu – este necesar ca având litera *x* din cheie și litera *y* din textul clar să găsim litera textului cifrat care se află la intersecția liniei *x* și coloanei *y*. Se poate de procedat și în conformitate cu ecuațiile ce definesc cifrul  $c_i = m_i + k_i \pmod{26}$  și  $m_i = c_i - k_i \pmod{26}$ , așa cum este arătat în exemplul ce urmează.

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>	<i>m</i>	<i>n</i>	<i>o</i>	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>u</i>	<i>v</i>	<i>w</i>	<i>x</i>	<i>y</i>	<i>z</i>
<i>a</i>	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
<i>b</i>	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
<i>c</i>	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
<i>d</i>	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
<i>e</i>	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
<i>f</i>	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
<i>g</i>	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
<i>h</i>	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</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
<i>j</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	I
<i>k</i>	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
<i>l</i>	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
<i>m</i>	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
<i>n</i>	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
<i>o</i>	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
<i>p</i>	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
<i>q</i>	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
<i>r</i>	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
<i>s</i>	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
<i>t</i>	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
<i>u</i>	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
<i>v</i>	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
<i>w</i>	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
<i>x</i>	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
<i>y</i>	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
<i>z</i>	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

Tabelul 5.4. *Tabula Recta* pentru cifrul Vigenere

**Problemă.**

De cifrat, utilizând cifrul Vigenere, mesajul „*Per aspera ad astra*” folosind cheia  $K = \text{SUPER}$ .

**Soluție.** Pentru a cifra sau descifra mai întâi facem corespondența următoare:

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
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

Apoi alcătuim și completăm tabelul:

Textul clar $M$	P	E	R	A	S	P	E	R	A	A	D	A	S	T	R	A
Cheia $K$	S	U	P	E	R	S	U	P	E	R	S	U	P	E	R	S
Textul clar $M$	15	4	17	0	18	15	4	17	0	0	3	0	18	19	17	0
Cheia $K$	18	20	15	4	17	18	20	15	4	17	18	20	15	4	17	18
$M+K \pmod{26}$	7	24	6	4	9	7	24	6	4	17	21	20	7	23	8	18
Textul cifrat $C$	H	Y	G	E	J	H	Y	G	E	R	V	U	H	X	I	S

Criptograma este  $C = \text{HYGEJHYGERVUHXIS}$ .

Pentru decriptare procedăm la fel, cu excepția  $m_i = c_i - k_i \pmod{26}$ .

Pentru acesta la fel completăm următorul tabel:

Textul cifrat $C$	H	Y	G	E	J	H	Y	G	E	R	V	U	H	X	I	S
Cheia $K$	S	U	P	E	R	S	U	P	E	R	S	U	P	E	R	S
Textul cifrat $C$	7	24	6	4	9	7	24	6	4	17	21	20	7	23	8	18
Cheia $K$	18	20	15	4	17	18	20	15	4	17	18	20	15	4	17	18
$M-K \pmod{26}$	15	4	17	0	18	15	4	17	0	0	3	0	18	19	17	0
Textul clar $M$	P	E	R	A	S	P	E	R	A	A	D	A	S	T	R	A

Textul clar este  $M = \text{PERASPERAADA STRA}$ .

Criptanaliza sistemului Vigenere constă în următoarele:

fie  $c = c_0 c_1 \dots c_{n-1}$  textul criptat cu cheia  $k = k_0 k_1 \dots k_{p-1}$ ; putem aranja acest text sub forma unei matrice cu  $p$  linii și  $\lceil n/p \rceil^1$  coloane, astfel:

$$\begin{array}{cccc}
 c_0 & c_p & c_{2p} & \dots \\
 c_1 & c_{p+1} & c_{2p+1} & \dots \\
 \vdots & \vdots & \vdots & \vdots \\
 c_{p-1} & c_{2p-1} & c_{3p-1} & \dots
 \end{array}$$

Elementele de pe prima linie au fost criptate după formula

$$c_{pr} = a_{pr} + k_0 \pmod{26}, k \geq 0,$$

adică cu un sistem Cezar ( $k_0$  fiind o valoare fixată din  $\mathbf{Z}_{26}$ ). În mod similar și celelalte linii. Deci, dacă s-ar cunoaște lungimea  $p$  a cheii, problema s-ar reduce la criptanaliza a  $p$  texte criptate cu Cezar – sistem de criptare monoalfabetic. Sunt cunoscute două metode pentru aflarea lungimii cheii: testul lui Kasiski și indexul de coincidență.

<sup>1</sup>  $\lceil x \rceil$  - funcția plafon, sau partea întreagă prin exces, este cel mai mic număr întreg mai mare decât numărul real  $x$  dat.

Prima metodă constă în studiul textului criptat și aflarea de perechi de segmente de cel puțin 3 caractere identice (această lungime este propusă de Kasiski). Pentru fiecare astfel de pereche, se determină distanța dintre segmente. După ce s-au găsit mai multe astfel de distanțe, valoarea lui  $p$  va fi cel mai mare divizor comun al lor (sau – eventual un divizor al acestuia).

A doua metodă de aflare a lungimii cheii de criptare într-un sistem Vigenere se bazează pe un concept definit în 1920 de Wolfe Friedman - *indexul de coincidență*. Dacă  $c = c_1c_2\dots c_n$  este o secvență de  $n$  caractere alfabetice, probabilitatea ca două caractere din  $c$ , alese aleator, să fie identice se numește "*index de coincidență*"  $I_c(x)$  al lui  $c$ .

*Cifrul omofonic* (homophonic ciphers) este un cifru de substituție în care un caracter al alfabetului mesajului clar (alfabet primar) poate să aibă mai multe reprezentări. Ideea utilizată în aceste cifruri este uniformizarea frecvențelor de apariție a caracterelor alfabetului textului cifrat (alfabet secundar), pentru a îngreuna atacurile criptanalitice. Astfel, litera  $A$  - cu cea mai mare frecvență de apariție în alfabetul primar – poate fi înlocuită de exemplu cu  $H$ ,  $\#$  sau  $m$ .

Cifrul omofonic este un cifru intermediar între sistemele mono și cele polialfabetice. Principalul lui scop este de a evita atacul prin frecvența de apariție a caracterelor. Se presupune că a fost utilizat prima oară în 1401 de către ducele de Mantua. În cifrul omofonic fiecărui caracter  $a \in m$  i se asociază o mulțime  $H(a) \subset c$  astfel încât:

- $H(a) \cap H(b) = \emptyset \Leftrightarrow a \neq b$ ;
- Dacă  $a$  apare mai frecvent decât  $b$  în textele clare, atunci  $\text{card}(H(a)) \geq \text{card}(H(b))$ .

Criptarea unui caracter  $x \in m$  se face cu un element ales aleator din  $H(x)$ . Pentru decriptarea lui  $y \in c$  se caută o mulțime  $H(x)$  astfel încât  $y \in H(x)$ .

*Exemplu.* Pentru limba engleză poate fi utilizat cifrul definit de tabelul 5.5, în primele două linii ale căruia sunt așezate literele alfabetului latin și frecvențele rotunjite ale acestora. În coloanele de sub litera  $x$  este situat  $H(x)$ . De exemplu

$$H(n) = \{18, 58, 59, 66, 71, 91\}.$$

Pentru criptarea textului „ $ab$ ” se poate folosi oricare din secvențele 0948, 1248, 3348, 4748, 5348, 6748, 7848, 9248, 0981, 1281, 3381, 4781, 5381, 6781, 7881, 9281.

Tabelul 5.5 a fost completat cu numerele 00, 01, ..., 99, ele fiind în cantitate de 100 pentru a simplifica calculele legate de frecvența literelor (adică pentru a corespunde cu 100% a frecvențelor). Însă elementele acestui tabel pot fi oricare simboluri, iar cantitatea lor va determina coeficientul de proporționalitate pentru frecvențele fiecărei litere din alfabet.

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
8	2	3	4	12	2	2	6	6	1	1	4	2	6	7	2	1	6	6	9	3	1	2	1	2	1
09	48	13	01	14	10	06	23	32	15	04	26	22	18	00	38	94	29	11	17	08	34	60	28	21	02
12	81	41	03	16	31	25	39	70			37	27	58	05	95		35	19	20	61		89		52	
33		62	45	24			50	73			51		59	07			40	36	30	63					

47			79	44			56	83			84		66	54			42	76	43								
53				46			65	88					71	72			77	86	49								
67				55			68	93					91	90			80	96	69								
78				57										99					75								
92				64															85								
				74															97								
				82																							
				87																							
				98																							

Tabelul 5.5. Exemplu de cifru omofonic pentru limba engleză

Deși mai greu de spart decât cifrurile de substituție simple (monoalfabetice), cifrul omofonic nu maschează total proprietățile statistice ale textului clar. În cazul unui atac cu text clar cunoscut, cifrul se sparge extrem de ușor. Atacul cu text cifrat este mai dificil, dar unui calculator îi va lua doar câteva secunde pentru a-l sparge.

*Substituție poligramică* realizează substituția unor blocuri de caractere (poligrame) din textul clar, distrugând astfel semnificația, atât de utilă în criptanaliză, a frecvențelor diferitelor caractere. Vom considera un mesaj  $m=m_1m_2\dots m_dm_{d+1}\dots$  și un cifru care prelucrează poligrame de lungime  $d$ . Criptograma rezultată este  $c=c_1c_2\dots c_dc_{d+1}\dots c_{d+d}\dots$ , fiecare poligramă  $m_{i,d+1}\dots m_{i,d+d}$  fiind transformată în poligrama  $c_{i,d+1}\dots c_{i,d+d}$  prin funcția de substituție  $f_j$  astfel:

$$c_{i,d+j} = f_j(m_{i,d+1}\dots m_{i,d+d})$$

În cazul cifrării literelor singulare frecvența de apariție a literelor în textul cifrat este egală cu frecvența de apariție a literelor corespunzătoare din textul clar. Această invarianță a frecvențelor furnizează o cantitate de informație suficientă criptanalistului pentru spargerea cifrului. Pentru minimizarea informației colaterale furnizate de frecvența de apariție a literelor s-a utilizat cifrarea grupurilor de  $d$  litere ( $d$ -grame). În cazul când un grup de  $d$  litere este substituit printr-un alt grup de  $d$  litere, substituția se numește poligramică. Substituția poligramică cea mai simplă se obține pentru  $d=2$  când digrama  $m_1m_2$  din textul clar se substituie cu digrama  $c_1c_2$  din textul cifrat.

Un exemplu clasic pentru substituția diagramelor este cifrul lui *Playfair* (tabelul 5.6).

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

Tabelul 5.6. Exemplu de cifru Playfair

Primele litere din pătrat (în caz general – un dreptunghi) reprezintă un cuvânt cheie  $k$  (literele care se repetă se scriu o singură dată, în acest exemplu cheia fiind  $k=PLAYFAIR$ ), după care pătratul se completează cu literele alfabetului în ordinea lor naturală, fără repetare. Cifrarea se execută după următoarele reguli:

- Pentru separarea literelor identice alăturate se introduc niște caractere de separare care, de regula, au frecvența de apariție redusă, cum sunt de exemplu literele X, Q în limba română. În cazul în care numărul de caractere în textul clar este impar se procedează la fel. La descifrare aceste litere introduse se omit.
- Dacă  $m_1m_2$  sunt dispuse în vârfurile opuse ale unui dreptunghi, atunci  $c_1c_2$  sunt caracterele din celelalte vârfuri ale dreptunghiului,  $c_1$  fiind în aceeași linie cu  $m_1$ . De exemplu RK devine CG, deci  $RK \rightarrow CG$ ;
- Dacă  $m_1$  și  $m_2$  se găsesc într-o linie, atunci  $c_1$  și  $c_2$  se obțin printr-o deplasare ciclică spre dreapta a literelor  $m_1$  și  $m_2$ . De exemplu  $PA \rightarrow LY$  iar  $NO \rightarrow OQ$ ;
- Dacă  $m_1$  și  $m_2$  se află în aceeași coloană atunci  $c_1$  și  $c_2$  se obțin printr-o deplasare ciclică a lui  $m_1, m_2$  de sus în jos. De exemplu  $AH \rightarrow BQ$  iar  $OV \rightarrow VL$ ;

Descifrarea se execută după reguli asemănătoare cu cele de cifrare, inversând direcția.

**Exemplu.** Folosind exemplul de mai sus ( $k = \text{PLAYFAIR}$ ) pentru textul clar  $m = \text{„VINE IARNA”}$  obținem textul cifrat  $c = \text{„UR UN BP IO YW”}$ . Aici am introdus la sfârșitul mesajului litera X iar  $AX \rightarrow YW$ . La descifrare după sensul mesajului se omite această literă.

Cifrul Playfair se folosea în scopuri tactice de către forțele militare britanice în timpul celui de-al doilea război al Burilor (1899-1902) dar și în primul război mondial. La fel a fost utilizat de către australieni și germani în timpul celui de-al doilea război mondial. El era utilizat deoarece era suficient de rapid în aplicare și nu necesita nici un utilaj special. Scopul principal al utilizării lui era protecția informației importante (însă nu și secrete) pe parcursul unei lupte. La momentul când criptanalizii inamici spargeau cifrul, informația deja nu mai era utilă pentru inamic.

Utilizarea cifrului Playfair în prezent nu are sens deoarece laptop-urile moderne pot sparge cu ușurință cifrul în câteva secunde. Primul algoritm de spargere pentru Playfair a fost descris în anul 1914 de către locotenentul Iosif O. Moubornom într-o broșură de 19 pagini.

### 5.2.2. Cifrul de transpoziții

Spre deosebire de cifrurile cu substituție, care păstrează ordinea literelor din textul sursă, însă le transformă, cifrurile cu transpoziție (*transposition ciphers*) reordonează literele, fără a le „deghiza”. Criptarea prin metoda transpoziției este o tehnică mai eficientă decât criptarea prin substituție, dar are, la rândul ei, o mulțime de dezavantaje. Textul criptat prin metoda transpoziției păstrează toate caracterele textului inițial, dar în altă ordine obținută prin aplicarea algoritmului ce va fi prezentat în continuare.

Criptarea prin transpoziție constă în scrierea textului inițial din care s-au eliminat spațiile și semnele de punctuație într-o matrice de dimensiune  $M \times N$  și interschimbarea anumitor linii (sau coloane) între ele. Textul criptat se obține prin scrierea caracterelor din noua matrice de pe fiecare coloană în parte, începând cu colțul din stânga-sus. Dacă lungimea textului inițial este mai mică

decât numărul de elemente ce pot fi scrise în matrice, atunci textul se completează cu elemente aleatoare, până ajunge la dimensiunea  $M \cdot N$ .

Pentru textul „*Securitatea este asigurată*”, care are lungimea de 24 de caractere, se pot alege mai multe matrice de dimensiune  $M \times N$ , o posibilitate ar fi ca matricea să aibă 4 linii și 6 coloane, dar pentru ca textul să fie mai greu de decodificat trebuie să conțină și caractere alese aleator, sau într-un mod mai inteligent, care să îngreuneze munca celui care dorește să afle conținutul secret din mesajul criptat. Fie am ales o matrice care are 5 linii și 6 coloane. Textului inițial i se adaugă 6 caractere aleatoare și se obține textul *Securi tateae steasi gurată xyzwtwu* care se scrie în matricea din partea stângă, așa cum e arătat în tabelul 5.7:

	1	2	3	4	5	6
1	S	e	c	u	r	i
2	t	a	t	e	a	e
3	s	t	e	a	s	i
4	g	u	r	a	t	ă
5	x	y	z	t	w	u

	1	2	3	4	5	6
5	x	y	z	t	w	u
3	s	t	e	a	s	i
4	g	u	r	a	t	ă
1	S	e	c	u	r	i
2	t	a	t	e	a	e

Tabelul 5.7. Exemplu de cifru cu transpoziție

Prin scrierea liniilor 1, 2, 3, 4, 5 în ordinea 5, 3, 4, 1, 2, această ordine fiind cheia cifrului, adică  $k=5,3,4,1,2$ , se obține matricea din partea dreaptă. Textul criptat care se obține este

$$c = xsgSt ytuea zerct taauw wstra uiăie.$$

Pentru ca procesul de decriptare să fie mai simplu și să nu mai fie nevoie de ordinea în care au fost puse liniile din matricea creată, se folosește o versiune a criptării prin transpoziție care se bazează pe un cuvânt-cheie. Pentru a cripta un text folosind o cheie și metoda transpoziției, se alege o cheie ale cărei litere determină ordinea în care se vor scrie coloanele din matricea aleasă. Pentru a afla ordinea în care vor fi scrise coloanele din textul inițial, se ordonează alfabetic literele din cheie, și fiecărei litere  $i$  se asociază numărul de ordine din șirul ordonat. Lungimea cheii trebuie să fie egală cu numărul de coloane din matrice.

Considerăm textul anterior, scris într-o matrice de dimensiuni  $5 \times 6$ , și cheia „*butuc*” (tabelul 5.8.). Literele din cheie se ordonează alfabetic și se obține șirul:  $b, c, t, u, u$ . Indicele 1 este asociat cu litera  $b$ , indicele 2 - cu litera  $c$ , indicele 3 - cu litera  $t$ , indicele 4 - cu prima literă  $u$ , iar indicele 5 este asociat cu a doua literă  $u$ . Pentru a rescrie liniile, pentru fiecare indice  $i$  din șirul ordonat se caută indicele  $j$ , care reprezintă poziția literei cu indicele  $i$  din cheie și se scrie linia  $j$ , astfel:

	1	2	3	4	5	6	
b	1	S	e	c	u	r	i
u	4	t	a	t	e	a	e
t	3	s	t	e	a	s	i
u	5	g	u	r	a	t	ă
c	2	x	y	z	t	w	u

	1	2	3	4	5	6
1	S	e	c	u	r	i
2	x	y	z	t	w	u
3	s	t	e	a	s	i
4	t	a	t	e	a	e
5	x	y	z	t	w	u



### Tabelul 5.8. *Exemplu de transpoziție cu cheie*

Textul cifrat care se obține în final este  $c = Sxstx\ eytay\ czetz\ utaet\ rwsaw\ iui eu$ .

Pentru a decifra un mesaj criptat cu această metodă, criptograma se scrie în matrice pe coloane, începând cu colțul stânga-sus, și apoi se realizează operația inversă, adică pentru fiecare indice  $j$  al literelor din cheie, se caută indicele  $i$  asociat literei din șirul sortat și se scrie linia cu indicele  $i$ . Din noua matrice astfel obținută se scriu literele de pe fiecare linie, în ordinea naturală.

Spargerea unui cifru cu transpoziție începe cu verificarea dacă acesta este într-adevăr de acest tip prin calcularea frecvențelor literelor și compararea acestora cu statisticile cunoscute. Dacă aceste valori coincid, se deduce că fiecare literă este „ea însăși”, deci este vorba de un cifru cu transpoziție. Următorul pas este emiterea unei presupunerii în legătură cu numărul de coloane. Acesta se poate deduce pe baza unui cuvânt sau expresii ghicite ca făcând parte din text. Considerând spre exemplu sintagma „*săprezinte*”, cu grupurile de litere (luate pe coloane) „*si*”, „*an*”, „*pt*”, „*re*”, se poate deduce numărul de litere care le separă, deci numărul de coloane. Notăm în continuare cu  $m$  acest număr de coloane.

Pentru a descoperi modul de ordonare a coloanelor, dacă  $m$  este mic, se pot considera toate posibilitățile de grupare a câte două coloane (în număr de  $m \cdot (m - 1)$ ). Se verifică dacă ele formează împreună un text corect numărând frecvențele literelor și comparându-le cu cele statistice. Perechea cu cea mai bună potrivire se consideră corect poziționată. Apoi se încearcă, după același principiu, determinarea coloanei succesoare perechii din coloanele rămase iar apoi - a coloanei predecesoare. În urma acestor operații, există șanse mari ca textul să devină recognoscibil.

Unele proceduri de criptare acceptă blocuri de lungime fixă la intrare și generează tot un bloc de lungime fixă. Aceste cifruri pot fi descrise complet prin lista care definește ordinea în care caracterele vor fi trimise la ieșire (șirul pozițiilor din textul de intrare pentru fiecare caracter din succesiunea generată).

De la apariția cifrurilor cu substituție și a celor cu transpoziție anii au trecut și tehnicile de criptare au evoluat foarte mult. Problema construirii unui cifru imposibil de spart i-a preocupat îndelung pe criptanaliști; ei au dat o rezolvare teoretică simplă încă de acum câteva decenii dar metoda nu s-a dovedit fiabilă din punct de vedere practic, după cum se va vedea în continuare.

Tehnica propusă pentru un cifru perfect presupune alegerea unui șir aleator de biți pe post de cheie și aducerea textului sursă în forma unei succesiuni de biți prin înlocuirea fiecărui caracter cu codul său ASCII. Apoi se aplică o operație logică - de tip SAU exclusiv (operația inversă echivalentei:  $0 \text{ xor } 0 = 0$ ,  $0 \text{ xor } 1 = 1$ ,  $1 \text{ xor } 0 = 1$ ,  $1 \text{ xor } 1 = 0$ ) - între cele două șiruri de biți. Textul cifrat rezultat nu poate fi spart pentru că nu există indicii asupra textului sursă și nici textul cifrat nu oferă criptanalistului informații. Pentru un eșantion de text cifrat suficient de mare, orice literă sau grup de litere (diftong, triftong) va apărea la fel de des.

Acest procedeu este cunoscut sub numele de metoda cheilor acoperitoare. Deși este perfectă din punct de vedere teoretic, metoda are, din păcate, câteva dezavantaje practice:

- cheia nu poate fi memorată, astfel încât transmiiătorul și receptorul să poarte câte o copie scrisă a ei fiindcă în caz că ar fi „capturați”, adversarul ar obține cheia;
- cantitatea totală de date care poate fi transmisă este determinată de dimensiunea cheii disponibile;
- o nesincronizare a transmiiătorului și receptorului care generează o pierdere sau o inserare de caractere poate compromite întreaga transmisie fiindcă toate datele ulterioare incidentului vor apărea ca eronate.

### 5.3. Mecanisme de criptare

Sistemele de criptare pot fi aduse la un grad mai mare de securitate dacă se folosesc mijloace mecanice de criptare. Astfel de mecanisme special construite vor ușura operațiile de criptare/decriptare și în același timp vor fi capabile să creeze un număr mult mai mare de chei posibile.

#### 5.3.1. Mijloace mecanice de criptare

Primele astfel de mecanisme au apărut încă în antichitate. În secolul V î.e.n. pentru criptarea datelor se folosea un baston, numit *Schitala* (figura 5.4), în jurul căruia se înfășura spiră lângă spiră o panglică foarte îngustă de piele, papirus sau pergament pe care, pe generatoare se scriau literele mesajelor. După ce textul era scris panglica se desfășura, mesajul devenea indescifrabil, deoarece literele erau dezamblate. Mesajul se putea descifra numai de o persoană care dispunea de un baston de grosime și lungime identice cu bastonul inițial pe care se înfășura din nou panglica primită de receptor. Astfel *Schitala* realiza o transpoziție, aceasta fiind o primă formă a acestei metode de criptare. Conform istoricilor greci, acest mod de comunicare era folosit de spartani în timpul campaniilor militare. El avea avantajul de a fi rapid și nu genera erori de transmitere. Dezavantajul însă era acela că putea fi ușor de spart.

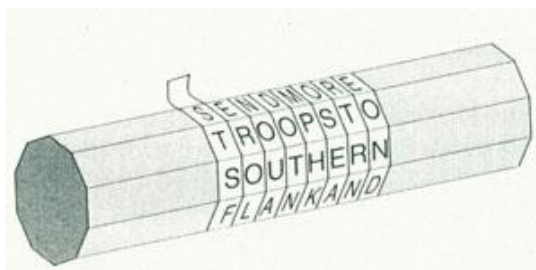


Figura 5.4. *Schitala*

*Leon Battista Alberti* (14.02.1404 – 25.04.1472) – scriitor, arhitect, pictor, sculptor, matematician, criptograf, filozof italian și umanist al Renașterii a inventat „*Criptograful lui Alberti*” (figura 5.5), care era alcătuit din două discuri concentrice cu diametre diferite, suprapuse. Fiecare disc era împărțit în 24 sectoare pe care erau înscrise litere și cifre. Pe discul exterior, care rămânea

static, erau scrise 20 de litere ale alfabetului italian (alfabetul italian nu avea literele *H, J, K, W, X, Y*) în ordinea lor firească, iar apoi cifrele 1, 2, 3, 4. Pe discul interior care se rotea, erau scrise 23 de litere ale alfabetului latin (fără *J, K, Y*) și conjuncția *ET*. Ordinea lor era arbitrară. Pentru cifrare se stabilea o cheie, de exemplu  $G=a$ . Aceasta însemna că pentru cifrare litera *a* de pe discul mic se așeza în dreptul literei *G* de pe discul mare și apoi începea cifrarea. Alberti recomanda schimbarea cheii după un număr de cuvinte.

Criptograful lui Alberti a fost perfecționat de către Silvester, Argenti și alții, constituind un element de bază pentru criptografele de tip disc apărute ulterior. Silvester Porta a împărțit discurile în 26 sectoare (Figura 5.5) utilizând astfel toate cele 26 litere ale alfabetului latin (nu numai italian), criptograful său realizând astfel o substituție simplă complet literală.

Criptograful avea două particularități care fac ca invenția să fie un mare eveniment în criptografie. În primul rând acest mecanism nu era altceva decât un algoritm de criptare polialfabetică. În rândul al doilea discul respectiv permitea utilizarea așa numitelor coduri cu recifrare, care au apărut abia la sfârșitul secolului XIX, adică peste patru secole după invenția lui Alberti. În acest scop pe discul exterior erau scrise cifrele 1, 2, 3, 4. Alberti a compus un cod care consta din 336 grupuri de coduri numerotate de la 11 la 4444. Fiecărui cod îi corespundea o oarecare frază terminată. Când fraza se întâlnea în mesaj ea se înlocuia cu codul respectiv, iar cu ajutorul discului cifrele erau criptate ca niște semne ordinare ale mesajului, fiind transformate în litere.

Leon Alberti poate fi considerat un criptograf ilustru și din motivul că este autorul primei lucrări de criptologie din Europa („*De cifris*”) publicată în 1469. În această lucrare erau prezentate exemple de versiuni posibile de cifrare dar și se argumenta necesitatea aplicării criptografie în practică ca un instrument ieftin și sigur de protecție a informației.

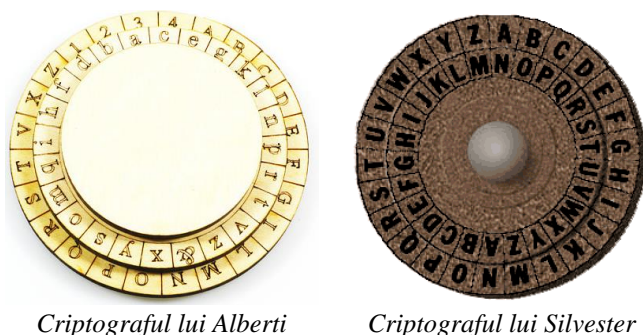


Figura 5.5. Versiuni de Criptografe

Ideea de mașină de criptare apare clar prima dată la Thomas Jefferson, primul secretar de Stat al Statelor Unite (președinte era George Washington), care a inventat un aparat de criptat numit roată de criptare, folosit pentru securitatea corespondenței cu aliații – în special cei francezi. Un cilindru Jefferson (figura 5.6) este format din  $n$  discuri de dimensiuni egale (inițial  $n = 26$  sau  $n =$

36) așezate pe un ax. Discurile se pot roti independent pe ax, iar pe muchia fiecăruia sunt înscrise cele 26 litere ale alfabetului, într-o ordine aleatoare (dar diferită pentru fiecare disc).

La criptare, textul clar se împarte în blocuri de  $n$  caractere. Fiecare astfel de bloc se scrie pe o linie (generatoare) a cilindrului, rotind corespunzător fiecare disc pentru a aduce pe linie caracterul căutat. Oricare din celelalte  $n-1$  linii rămase poate constitui blocul de text criptat.



Figura 5.6. Cilindre Jefferson

Pentru decriptare este necesar un cilindru identic, în care se scrie pe o linie textul criptat (de  $n$  caractere) și apoi se caută printre celelalte 25 linii un text cu semnificație semantică. Probabilitatea de a avea un singur astfel de text crește cu numărul de discuri din cilindru.

O mică diferență apare dacă textul clar nu are nici o semnificație semantică (s-a folosit o dublă criptare). Atunci trebuie convenită dinainte o anumită distanță de criptare  $s$  ( $1 \leq s \leq 25$ ).

Ordinea discurilor poate fi de asemenea schimbată. De exemplu, un cilindru cu  $n = 20$  discuri poate realiza  $20! = 2\,432\,902\,008\,176\,640\,000$  texte criptate diferite pentru același text clar. Cilindrul Jefferson realizează o substituție polialfabetică de perioadă  $n$ . Dacă ar fi privit ca un sistem de criptare Vigenere, lungimea cheii este enormă (de multe ori  $n^n$ , în funcție de modalitățile de aranjare a alfabetelor pe discuri). Cilindrul Jefferson a fost reinventat ulterior de mai multe ori, cea mai celebră fiind se pare mașina de criptat „Enigma”.

Thomas Jefferson a folosit acest aparat în perioada 1790 – 1802, după care se pare că ideea s-a pierdut. Devenit președinte, Jefferson a fost atras de sistemul Vigenere, pe care îl consideră mai sigur și-l recomandă secretarului său de stat James Madison ca înlocuitor al sistemului pe care îl inventase anterior.

### 5.3.2. Mașini rotor

O mașină rotor (rotor machine, figura 5.7) are o tastatură și o serie de rotoare ce permit implementarea unei versiuni a cifrului Vigénère. Fiecare rotor face o permutare arbitrară a alfabetului, are 26 de poziții și realizează o substituție simplă. Deoarece rotoarele se mișcă cu viteze de rotație diferite, perioada unei mașini cu  $n$  rotoare este  $n \cdot 26!$ .

Aplicarea practică a acestor mașini a început numai la începutul secolului XX. Una dintre primele mașini rotor a fost mașina germană „Enigma”, elaborată în anul 1917 de către Eduard Hebern și perfectată mai târziu de mai multe ori. Din punct de vedere comercial ea a fost

disponibilă pe piață încă din anul 1920, însă importanța ei a fost dată de utilizarea mașinii de către diverse guverne, în mod special de către Germania nazistă înainte și în timpul celui de-al doilea război mondial.

Dintre toate dispozitivele criptografice create de-a lungul timpului mașina Enigma a fost un echipament mai special din 2 puncte de vedere: criptografic și istoric.



Figura 5.7. Modelul militar german numit *Wehrmacht Enigma*

Importanța din punct de vedere criptografic este dată de faptul că echipe de criptanaliști (matematicieni la origine) de toate naționalitățile, în efort combinat, au încercat pe de o parte perfecționarea mașinii, pe de altă parte spargerea cifrurilor. Printre cei care au participat la spargerea cifrului au făcut parte și polonezul Rajewski și britanicul Turing (inventatorul mașinilor Turing).

Importanța istorică rezidă din rolul mare jucat de aceste mașini în timpul celui de-al doilea război mondial, mai precis faptul că descifrarea de către aliați a codului (nume de proiect ULTRA) a dus, după unii istorici, la scurtarea războiului cu aproximativ un an.

*Construcția mașinii.* Mașina Enigma era o combinație de părți mecanice și electrice. Principalele ei componente erau, după cum urmează:

- *Tastatura:* o tastatură obișnuită similară cu cea pentru mașinile de scris.
- *Placa cu lămpi:* asemănătoare unei tastaturi cu lămpi în loc de taste. Pe lămpi erau tipărite literele alfabetului ce deveneau vizibile prin aprinderea lămpii corespunzătoare.
- *Placa cu comutatoare:* mufe (prize) câte una pentru fiecare literă, ce se conectau prin fire în 6 perechi (Această componentă fusese adăugată de germani pentru a crește securitatea mașinii).
- *Trei roți:* se mai numeau *rotoare (roți detașabile)* fiecare dintre ele având câte un set de 26 de contacte, câte unul pentru fiecare literă a alfabetului (figura 5.8).
- *Roata reflectoare:* roată fixă identică cu celelalte 3, având un set de 26 de contacte grupate în perechi.
- *Cabluri:* asigurau conexiunile între taste și lămpi precum și între lămpi și primul rotor, între primul rotor și al doilea, al doilea și al treilea, al treilea și roata reflectoare.
- *Baterie:* pentru alimentarea circuitelor electrice.





Figura 5.8. Rotor Enigma

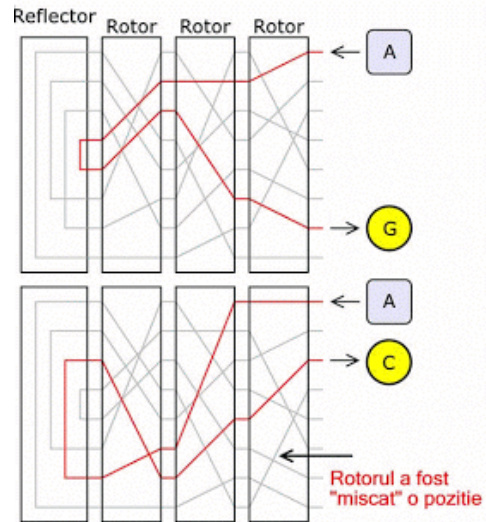


Figura 5.9. Principiul de funcționare al mașinii (Enigma)

Principiul de funcționare al mașinii Enigma se prezenta conform schemei din Figura 5.9. Prin apăsarea tastei "A" curentul era trecut prin setul de rotoare până la reflector de unde se "întorcea" înapoi aprinzându-se becul "G". Litera "A" se criptează diferit ("G" și "C") doar printr-o simplă rotire a primului rotor care face ca semnalul să circule pe o rută complet diferită.

Pentru *operarea mașinii* în primul rând toți operatorii aveau mașini identice (pentru asigurarea inter-operabilității). Inițierea criptării unui mesaj se făcea în 2 pași:

- *Pasul 1*: setarea mașinii – operație ce consta în fixarea ordinei și poziției fiecărui rotor precum și alegerea celor 6 perechi de conectori prin placa cu comutatoare (switch board).
- *Pasul 2*: scrierea propriu-zisă a mesajului – pentru criptarea mesajului operatorul apăsa pe tasta corespunzătoare primei litere din textul necodat (să zicem "N"). În acest moment se aprindea o lampă (să zicem "T") corespunzătoare codificării. Repetând și pentru celelalte litere, rezulta textul codat.

Trebuie de menționat că toate setările din *Pasul 1* erau înscrise în manuale de operare (code books), setări ce se schimbau de regulă zilnic. Fiecare operator avea câte un exemplar. De fapt, aceste setări constituiau cheia criptosistemului Enigma. Un atribut extrem de important al mașinii *Enigma* era că cheile de cifrare și cele de decifrare erau aceleași. Cu alte cuvinte dacă la "transmitere" "N" se transforma în "T", la "destinație" "T" se transforma în "N" (folosind bine-nțelesele aceleași setări ale mașinii).

Utilizarea intensivă colaborată cu posibilitatea transmiterii informației folosind aceleași *day key* la care se adăugau intensele activități de contraspionaj i-au condus pe germani la teama că mașina ar putea fi compromisă. Efectul, a fost introducerea unui *protocol*. Acesta spunea: „*fiecare operator va transmite suplimentar, înaintea mesajului propriu zis, o cheie a mesajului (message*

key)”. Aceste chei erau cuvinte (nu neapărat cu sens) formate din 3 litere alese în mod aleator de operatorul mașinii. Cu alte cuvinte, operatorul trebuia să seteze mașina conform instrucțiunilor zilnice din manualul de operare (*code book*), după care trimitea cheia din cele trei litere alese aleator. În așa fel mașina se seta într-un mod complet aleator. Condițiile radio proaste, lucrul sub presiune, precum și alte condiții de lucru nefavorabile puteau conduce la transmiterea (sau recepționarea) greșită (alterată) a cheii, fapt ce ar fi făcut inutilă transmiterea mesajului propriu-zis (evident, datorită faptului că mașinile de la transmițător și cea de la receptor ar fi fost setate diferit). Pentru a minimiza astfel de incidente, operatorilor li s-a cerut să transmită cheia de 2 ori. De exemplu cheia *hot* se transmitea *hothot* și se recepționa *dugraz*. Însă în mod ironic, ceea ce se dorea o măsură de securitate în plus, de fapt a compromis mașina.

Matematic vorbind, mulțimea cheilor posibile era atât de mare încât nici nu se punea problema ”atacării” mașinii, cel puțin nu la aceea vreme, prin metoda exhaustivă („*brute-force*”). Enigma a fost elaborată astfel încât securitatea să fie păstrată chiar dacă inamicul cunoaște schemele rotoarelor, cu toate că în practică setările erau secrete. Cu o schemă secretă de setare cantitatea totală a configurărilor posibile era de ordinul  $10^{114}$  (circa 380 biți) iar dacă schema și alte setări operaționale erau cunoscute acest număr se reducea la  $10^{23}$  (76 biți). Germanii credeau că mașina Enigma este una infailibilă datorită imensității setărilor posibile ce i se puteau aplica. Era ireal să începi măcar să alegi o configurare posibilă.

Din punct matematic de vedere transformarea Enigmei pentru fiecare literă este rezultatul matematic a permutărilor. Pentru un aparat cu trei rotoare fie  $P$  transformarea pe tabela de prize,  $U$  - reflectorul, și  $L$ ,  $M$ ,  $R$  - acțiunea rotorului din stânga, din mijloc, dreapta respectiv. Atunci criptarea  $E$  poate fi notată cu:

$$E = PRMLUL^{-1}M^{-1}R^{-1}P^{-1}$$

După fiecare apăsare de tastă rotoarele se rotesc, schimbând transformarea. De exemplu dacă rotorul de dreapta  $R$  e rotit cu  $i$  poziții, transformarea devine:  $\rho^i R \rho^{-i}$ , unde  $\rho$  este permutarea ciclică. Similar, rotorul din mijloc și cel din stânga pot fi reprezentate ca  $j$  și  $k$  rotații respectiv a lui  $M$  și  $L$ . Funcția de criptare poate fi descrisă astfel:

$$E = P(\rho^i R \rho^{-i})(\rho^j M \rho^{-j})(\rho^k L \rho^{-k})U(\rho^k L^{-1} \rho^{-k})(\rho^j M^{-1} \rho^{-j})(\rho^i R^{-1} \rho^{-i})P^{-1}$$

Pentru elucidarea funcționării mașinii Enigma este sugestivă simularea (în flash) de la [www.enigmaco.de/enigma](http://www.enigmaco.de/enigma) (figura 5.10).

Primele spargerii ale mașinii Enigma au avut loc la începutul anilor 30 de către matematicienii polonezi *Alicen Rejewski*, *Jerzy Rozycki* și *Henryk Zygalsk*. Cu noroc și intuiție Rejewski și echipa lui au reușit să compromită mașina, totul fiind posibil nu datorită vreunei ”scăpări” în proiectarea mașinii ci deciziei nemților de a transmite repetitiv (de 2 ori) cheia.

Ulterior Enigma a fost perfecționată, spargerea ei devenind practic imposibilă pentru acele timpuri. Un aport considerabil în direcția spargerii acestei mașini a avut Alan Turing, care proiectase o mașină electromecanică (denumită „*Bombe*” după modelul original polonez) ce putea ajuta la spargerea mașinii Enigma mai rapid decât „bomba” din 1932 a lui *Rejewski*, din care s-a și inspirat. „*Bombe*” (figura 5.11), cu o îmbunătățire sugerată de matematicianul Gordon Welchman, a devenit una din principalele unelte automate utilizate pentru a ataca traficul de mesaje protejat de Enigma.

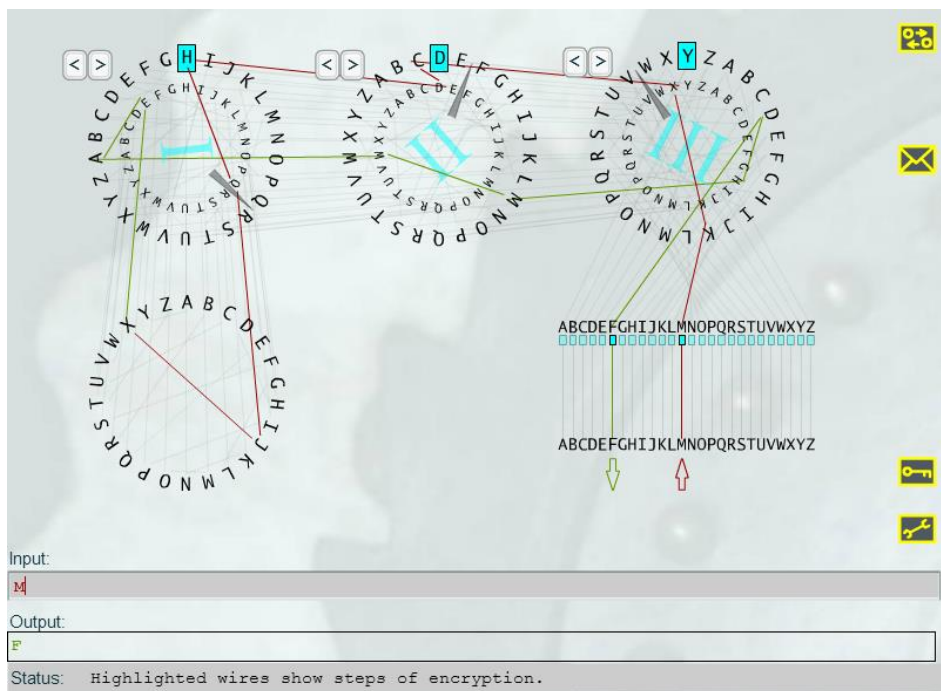


Figura 5.10. *Simulator mașina Enigma*

Mașina „*Bombe*” căuta setări potențial corecte pentru un mesaj Enigma (adică, ordinea rotoarelor, setările rotoarelor, etc.), folosind un fragment de text clar probabil. Pentru fiecare setare posibilă a rotoarelor (numărul maxim posibil fiind de ordinul a 1019 stări, sau 1022 pentru mașinile Enigma de la U-boat, care aveau patru rotoare, față de mașina Enigma standard care avea doar trei). Aceasta efectua un lanț de deducții logice pe baza fragmentului probabil, deducții implementate electric. „*Bombe*” detecta când avea loc o contradicție, și elimina setarea, trecând la următoarea. Peste două sute de astfel de mașini create de Alan Turing au fost în funcțiune până la sfârșitul războiului.

Mașinile cu rotor au fost folosite activ pe parcursul războiului II mondial. Pe lângă mașina germană Enigma au fost folosite și Sigaba (SUA), Typex (Marea Britanie), Red, Orange, Purple (Japonia). Mașinile cu rotor au fost vârful criptografiei formale deoarece realizau cifruri suficient de rezistente într-un mod relativ simplu.



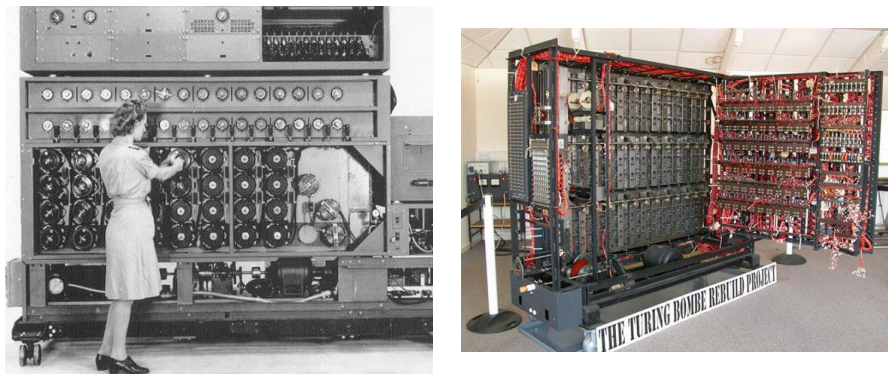


Figura 5.11. Mașina BOMBE (Alan Turing)

Atacurile încununare de succes asupra mașinilor cu rotor au fost posibile numai la începutul anilor 40 odată cu apariția mașinilor electronice de calcul. Tot în această perioadă criptografia devine științific ramură aparte a matematicii odată cu publicarea (anul 1949) articolului lui Claude Elwood Shannon<sup>2</sup> „Communication Theory of Secrecy Systems”, care a pus bazele științifice ale sistemelor de criptare cu cheie secretă (sistemelor simetrice).

## 5.4. Atacuri criptografice

### 5.4.1. Noțiune de criptanaliză și atac criptografic

Atacul asupra securității unui sistem criptografic definește orice acțiune ce compromite securitatea aceluia sistem. Atacurile criptografice pot fi îndreptate împotriva:

- algoritmilor criptografici;
- tehnicilor utilizate pentru implementarea algoritmilor protocoalelor;
- protocoalelor.

După modelul de atacare al unui atacator, aceste atacuri se pot clasifica după cum urmează:

- a) atacuri pasive (de interceptie):
  - de înregistrare a conținutului mesajelor;
  - de analiză de trafic;
- b) atacuri active:
  - de întrerupere (atac la disponibilitate);
  - de modificare (atac la integritate);
  - de fabricare(atac la autenticitate).

Atacurile pasive sunt atacuri în care intrusul (persoană, calculator, program) doar ascultă, monitorizează transmisia, deci sunt atacuri de interceptie. Ele pot fi de două feluri:

1. de înregistrare a conținutului mesajelor;
2. de analiză a traficului.

Caracteristicile atacurilor pasive:

<sup>2</sup> Claude Elwood Shannon (30 aprilie 1916 - 24 februarie 2001) a fost un matematician american, inginer electrotehnist și criptograf cunoscut sub numele de „tatăl teoriei informației”.

- sunt greu de detectat pentru că datele nu sunt alterate;
- măsurile ce pot fi luate pentru evitarea acestor atacuri sunt acelea care fac criptanaliza extrem de grea, dacă nu imposibilă;
- este necesară prevenirea și nu detecția lor.

Atacurile active sunt atacuri în care intrusul are o intervenție activă atât în desfășurarea normală a traficului, cât și în configurarea datelor (modificarea datelor, crearea unor date false).

Dintre atacurile active se regăsesc:

- întreruperea serviciului;
- modificarea;
- modificări în program astfel încât acesta va lucra diferit;
- modificarea conținutului mesajelor transmise în rețea;
- fabricarea: un neavizat inserează informații false în sistem.

Atacurilor active pot fi caracterizate prin faptul că deși pot fi detectate, prevenirea lor este foarte grea, deoarece ar însemna protecție fizică permanentă a întregului sistem.

Atacurile amintite mai sus sunt, în principiu, atacuri mai generale, nu neapărat specifice exclusiv algoritmilor criptografici, iar o analiză mai detaliată a unora dintre ele va fi făcută în compartimentul 9.

Atacurile specifice algoritmilor criptografice sunt atacurile de criptanaliză.

*Criptanaliza* este știința spargerii cifrurilor și se ocupă de obținerea valorii inițiale a informației criptate fără a avea acces la informația secretă, adică la cheia necesară pentru acest lucru. Persoana care se ocupă cu criptanaliza se numește *criptanalist*. Atacul criptografic bazat pe metode de criptanaliză se mai numește atac criptanalitic.

Analiza criptografică studiază metode de atac pornind de la informații minimale despre cheile de criptare, algoritmi utilizați, protocoalele de autentificare, segmente de text clar și segmentele corespondente din textul criptat, sau doar pe baza unuia sau a unui set de texte criptate utilizând același algoritm.

Scopul metodelor de criptanaliză este descoperirea mesajelor în clar și/sau a cheii din mesajul criptat. Orice cifru este creat în scopul de a asigura confidențialitatea informației protejate și, reieșind din acest concept, întotdeauna se vor găsi oameni care doresc să obțină accesul la informația dată. În esență, se încearcă determinarea unui punct vulnerabil al algoritmului, care să poată fi exploatat folosind metode pentru care timpul de căutare să fie considerabil mai mic decât timpul necesar verificării tuturor combinațiilor de chei posibile (atac cu forța brută).

Nu există încă un sistem criptografic despre care să se poată afirma că este pe deplin sigur, dar pot fi considerate sigure acele criptosisteme pentru care atacurile cunoscute necesită un timp mult prea îndelungat pentru a putea fi considerate practice.

#### 5.4.2. Tipuri de atacuri criptografice

Se cunosc mai multe tipuri de atacuri criptografice. O categorie aparte sunt atacurile ce se realizează prin forță brută (*brute force*), adică aplică o metodă exhaustivă de căutare prin încercarea tuturor combinațiilor posibile fie de chei de criptare, fie de simboluri din text pentru deducerea textului în clar (de exemplu, la metodele de criptare prin substituția sau transpoziția literelor din mesaje de tip text). Complexitatea acestui atac este în funcție de cantitatea tuturor variantelor posibile.

Acest tip de atac este unul general, adică poate fi aplicat la orice algoritm de criptare. Din acest motiv în elaborarea sistemelor de criptare autorii încearcă să obțină ca acest atac să fie cel mai eficient, comparativ cu celelalte metode de spargere. Sistemul se proiectează astfel încât forța brută să aibă un spațiu al soluțiilor suficient de voluminos pentru ca rezultatul aplicării forței brute să nu fie obținut pe parcursul a câtorva ani, sau uneori și secole.

În baza complexității aplicării forței brute se face evaluarea securității sistemului. În particular, cifrul se consideră sigur dacă nu există o metodă de spargere semnificativ mai rapidă decât forța brută.

Atacurile criptografice bazate pe forța brută sunt cele mai universale, dar și cele mai îndelungate. În legătură cu aceasta există deja și se mai elaborează încontinuu posibilități de optimizare a metodei forței brute. Printre aceste metode optimizate se numără următoarele:

- metoda ramifică și mărginește (*branch and bound method*);
- metoda calculelor paralele (*parallel calculation method*).

Metoda *ramifică și mărginește* reprezintă un algoritm de căutare a soluțiilor optime pentru diverse probleme. Esența lui constă în separarea submulțimii de soluții admisibile care nu conține soluții optime. Metoda a fost propusă pentru prima dată de către A. H. Land și A. G. Doig în 1960 pentru programarea discretă.

*Calculul paralel* este execuția în paralel pe mai multe procesoare a acelorași instrucțiuni, sau și a unor instrucțiuni diferite, cu scopul rezolvării mai rapide a unei probleme, de obicei special adaptată sau subdivizată. Ideea de bază aici constă în divizarea mulțimii soluțiilor în  $N$  submulțimi, fiecare din ele fiind mult mai „mică” decât originalul, astfel realizarea „forței brute” va necesita de  $n$  ori mai puțin timp, în funcție de numărul de submulțimi ale partiției făcute. Problema fundamentală aici constă în determinarea și divizarea mulțimii soluțiilor. „Forța brută” se aplică până când un procesor nu a găsit soluția (cheia) necesară.

De rând cu forța brută se mai aplică și *metodele statistice* de atac. Aceste metode se divizează în două subcategorii:

- metode de criptanaliză a proprietăților statistice ale gamei de criptare;
- metode de criptanaliza a complexității șirului.

Prima subcategorie studiază șirurile la ieșirea algoritmilor de criptare. În acest caz criptanalistul cu ajutorul diverselor teste statistice încearcă să găsească valoarea următorului bit al șirului cu o probabilitate mai mare decât probabilitatea alegerii aleatoare.

În cazul al doilea criptanalistul încearcă să genereze șiruri analogice cu gama însă aplicând metode mult mai simple.

Din diversitatea de tipuri și metode de atac criptografic, demonstrate, verificate și aplicate de matematicieni, informaticieni și criptanaliști trebuie menționate atacurile cu text clar sau text cifrat:

- *Atac cu text cifrat (ciphertext-only attack)* interceptat, prin analiza căruia se încearcă găsirea textului original sau a cheii de criptare. Acest atac se bazează pe informații referitoare la secvențe de text cifrat și este una dintre cele mai dificile metode criptografice din cauza informației sumare pe baza căreia trebuie să se deducă informații referitoare la textul clar sau la chei.
- *Atac cu text clar cunoscut (known plaintext attack)*, este un atac în care criptanalistul are acces nu doar la textul cifrat, ci și la textul clar corespunzător. El încearcă să descopere o corelație între cele două pentru a găsi cheia de criptare sau pentru a crea un algoritm care îi va permite să descifreze orice mesaje criptate cu această cheie. Textele în clar necesare pentru acest atac pot fi obținute prin diverse metode, de exemplu dacă se cunoaște că se trimite un fișier cifrat cu un nume știut, atunci din extensia fișierului se pot face concluzii despre conținutul anumitor fragmente ale fișierelor, de exemplu a header-ului. Acest atac este mai puternic decât atacul cu text cifrat.
- *Atac cu text cifrat ales (chosen ciphertext attack)*, este un atac în care criptanalistul alege un text cifrat și încearcă să găsească textul clar potrivit. Acest lucru se poate face cu o decriptare oracul (o mașină care decriptează fără a demasca cheia). Atacul este aplicat la criptarea cu cheie publică – se începe cu un text cifrat și căutări de potrivire de date ale textului clar postate public.
- *Atac cu text clar ales (chosen plaintext attack)*, este un atac în care criptanalistul poate cripta un text clar la alegerea sa și studia textul cifrat rezultat. Scopul criptanalistului este același ca la atacul cu text clar cunoscut: de a afla cheia de criptare sau de a găsi o altă metodă pentru descifrarea mesajelor cifrate cu aceeași cheie. Însă criptanalistul trebuie să mai aibă posibilitatea de a alege câteva texte în clar și să obțină rezultatul cifrării lor. Obținerea textului cifrat respectiv pentru textul clar dat uneori se poate face prin crearea și transmiterea unui mesaj necifrat în numele unuia din utilizatorii care folosesc criptarea. În cazul coincidenții unor factori acest mesaj poate fi cifrat și retransmis înapoi. Acest atac este cel mai des utilizat în criptografia asimetrică, în cazul în care criptanalistul are acces la o cheie publică.

- *Atac cu text clar ales adaptiv (adaptive chosen plaintext attack)*, este un caz particular, mai confortabil, al atacului cu text clar ales. Confortul atacului constă în faptul că pe lângă posibilitatea alegerii textului clar, criptanalistul poate lua decizia de a cifra un oarecare text clar în baza operațiilor de cifrare deja efectuate. Cu alte cuvinte la realizarea atacului cu text clar ales criptanalistul alege doar un bloc mare al textului clar pentru a fi cifrat și apoi, în baza datelor obținute începe spargerea sistemului. În cazul organizării atacului cu text clar ales adaptiv criptanalistul poate obține rezultatul cifrării oricărui bloc al textului clar pentru a acumula datele care îl interesează și care vor fi luate în seamă la alegerea ulterioară a blocurilor textului clar etc. Anume adaptivitatea (posibilitatea feedback-ului) îi dă un avantaj atacului cu text clar ales adaptiv.
- *Atac cu text cifrat ales adaptiv (Adaptive Chosen Ciphertext Attack)*, este un atac analogic atacului cu text clar ales adaptiv.

În continuare vom menționa alte tipuri și metode de atac utilizate în prezent.

*Atacul de tip dicționar (dictionary attack)* este o metodă criptanalitică în care atacatorul pregătește și memorează un tabel cu corespondențe text clar - text criptat de tipul perechilor  $(P_i C_i = E_{K_i}(P), K_i)$  sortate după  $C_i$ . Ulterior, atacatorul monitorizează comunicația și în momentul în care va găsi un text criptat  $C_j$  care se regăsește în tabelul său va găsi imediat cheia de criptare  $K_j$ .

*Atacul zilei de naștere (Birthday attack)*, se bazează pe cunoscutul paradox al „zilei de naștere” și a variantelor sale (într-un grup de 23 de persoane, probabilitatea să existe două dintre ele născute în aceeași zi din an este peste 0,5; în general, într-un grup de  $\sqrt{k}$  numere aleatoare având  $k$  valori posibile, probabilitatea ca cel puțin două să fie egale este în jur de 0,5). Problema poate fi astfel generalizată: dacă o funcție  $f: A \rightarrow B$  poate lua oricare din cele  $n$  valori din mulțimea  $B$  cu probabilități egale, atunci după calculul funcției pentru  $\sqrt{n}$  valori diferite este foarte posibil să găsim o pereche de valori  $x_1$  și  $x_2$  astfel încât  $f(x_1) = f(x_2)$ . Evenimentul reprezintă o coliziune, iar pentru funcții cu distribuție impară, coliziunea poate apărea și mai devreme. Semnătura digitală este susceptibilă de a fi supusă unui astfel de atac.

*Atac cu întâlnire la mijloc (Meet-in-the-middle attack)* este similar cu atacul zilei de naștere, cu excepția faptului că în acest caz analistul are o flexibilitate mai mare. În loc să aștepte coincidența a două valori într-o singură mulțime de date, analistul poate căuta o intersecție a două mulțimi.

Presupunem că atacatorul cunoaște o mulțime de texte în clar  $P$  și texte criptate  $C$  cu cheile  $k_1$  și  $k_2$ . Atunci el poate calcula  $E_K(P)$  pentru toate cheile posibile  $K$  și să memoreze rezultatele, apoi poate calcula  $D_K(C)$  pentru fiecare  $K$  și să compare cu rezultatele memorate - dacă va găsi o coincidență este ca și cum ar fi găsit cele două chei și poate verifica direct pe textul în clar și cel

criptat. Dacă dimensiunea cheii este  $n$ , atacul va folosi doar  $2^{n+1}$  criptări în contrast cu un atac clasic, care ar avea nevoie de  $2^{2n}$  criptări.

*Atacul omului din mijloc (Man-in-the-middle attack)* descrie situația când un atacator are posibilitatea să citească și să modifice mesajele schimbate între doi corespondenți fără ca cele două părți să sesizeze faptul că metoda de comunicare între ei a fost compromisă.

Atacul începe de obicei cu ascultarea canalului și se termină cu încercarea criptanalistului de a înlocui mesajul interceptat, extragerea informațiilor utile din el, redirectionarea mesajului la unele resurse externe. Fie că subiectul  $A$  planifică transmiterea spre subiectul  $B$  a unei informații oarecare. Subiectul  $C$  posedă cunoștințe despre structura și proprietățile metodei de transmitere a datelor, precum și a însuși faptului transmiterii acelei informații pe care intenționează să o intercepteze (de exemplu cheia privată). Pentru săvârșirea atacului  $C$  se „prezintă” lui  $A$  drept  $B$ , iar lui  $B$  drept  $A$ . Subiectul  $A$  consideră eronat că transmite informația lui  $B$  și o trimite lui  $C$ , care la rândul său efectuează unele operații cu ea (o copiază sau o modifică în scopuri personale) și o transmite lui  $B$ . Ultimul consideră că informația a fost primită direct de la  $A$ .

Posibilitatea unui astfel de atac rămâne o problemă serioasă pentru sistemele bazate pe chei publice.

*Atacul în reluare (replay attack)* este un atac în care atacatorul memorează o sesiune de comunicare în ambele sensuri (mesajele schimbate de ambii corespondenți) sau bucăți din sesiune. Ideea atacului nu este de a decripta o sesiune de comunicare, ci de a crea confuzii și mesaje false.

*Atacul cu chei relaționate (related keys attack)*. În acest caz atacatorul descoperă o relație între un set de chei și are acces la funcțiile de criptare cu astfel de chei relaționate. Scopul declarat este de a găsi chiar cheile de criptare. Algoritmi ca IDEA, GOST, RC2 și TEA au prezentat slăbiciuni când au fost supuse atacului.

*Atacul prin alunecare (slide attack)* poate fi văzut ca o variantă a atacului cu chei relaționate în care relațiile sunt definite pe aceeași cheie. Atacul este eficient în cazul unor procese iterative sau recursive (algoritmi simetrici de tip șir sau bloc) care prezintă grade de similitudine între cicluri succesive ale procesului iterativ. Complexitatea atacului este independentă de numărul de cicluri ai algoritmului. Slăbiciuni în cazul acestui atac au fost relevate în algoritmul Feistel și chiar în cazul SHA-1.

*Atacul de corelație (correlation attack)* se efectuează asupra generatorului de filtrare din cifrurile șir bazate pe generatoare de tip LFSR (Linear Feedback Shift Register, mai detaliat despre LFSR vezi în compartimentul 6), în două faze: întâi se determină o funcție între șirul de biți cheie generat și biții registrului de deplasare, după care șirul de chei este interpretat ca o versiune afectată de zgomot a șirului generat de LFSR.

*Atacul de corelație rapidă (fast correlation attack)* se aplică generatoarelor de chei bazate pe LFSR, ca și atacul de corelație, dar sunt mai rapide și exploatează existența unei corelații între șirul de chei și ieșirea unui LFSR, numit LFSR țintă, a cărui stare inițială depinde de anumiți biți ai cheii secrete.

Atacurile de corelație rapidă evită examinarea tuturor inițializărilor posibile ale LFSR-ului țintă folosind anumite tehnici eficiente de corectare a erorii. Astfel, descoperirea stării inițiale a LFSR-ului constă în decodarea subșirului de chei relativ la codul FSR-ului.

*Atacul prin interpolare (interpolation attack)* este o tehnică de atac asupra cifrurilor simetrice bloc construite din funcții algebrice simple. Dacă textul criptat este scris ca un polinom, funcție de elementele textului clar și gradul polinomului este suficient de mic, atunci un număr limitat de perechi de text clar/criptat sunt suficiente pentru determinarea funcției de criptare. Acest lucru permite atacatorului să crijteze sau să decrijteze blocuri de date fără a recupera propriu zis cheia de criptare. Atacul a fost introdus în 1997 și aplicat prima dată pe o variantă a algoritmului SHARK, un predecesor al algoritmului Rijndael, care stă la baza standardului AES. Acest tip de atac poate fi generalizat, astfel că ideea interpolării poate fi aplicată și în cazul unor polinoame probabilistice.

*Atacul „divide și cucerește” (divide and conquer attack)*. Atacurile din această categorie încearcă diviziunea cheilor în bucăți mai mici, pentru a face posibilă căutarea exhaustivă. Acest tip de atac este eficient în măsura în care este posibil să determinăm bucăți separate din chei. Problema constă în validarea sau invalidarea unui segment de cheie, fără a avea informații despre restul cheii.

*Atacul temporal (timing attack)*. Durata de execuție a unui echipament hardware de criptare poate furniza informații despre parametrii implicați astfel încât analiza atentă și măsurarea timpului de execuție poate duce, în anumite condiții, la recuperarea cheilor secrete. Pentru a putea realiza un astfel de atac, atacatorul are nevoie de un set de mesaje împreună cu durata lor de procesare pe echipamentul criptografic. Metodele de măsurare a timpului sunt diverse: monitorizarea activității procesorului, măsurarea timpului într-o secvență de interogare/răspuns etc. Atacul a fost aplicat algoritmilor RSA și RC5, dar și unor protocoale de Internet de tipul SSL.

#### 5.4.3. Criptanaliza liniară și diferențială

În continuare sunt expuse două tehnici de criptanaliză - *criptanaliza liniară* și *criptanaliza diferențială* - care la momentul actual sunt unele dintre cele mai răspândite metode de spargere a cifrurilor bloc.

*Criptanaliza liniară (linear cryptanalysis)* este o tehnică introdusă de Matsui și Yamagishi în 1991 (*A new Method for known plain text attack of FEAL cipher*) care încearcă să exploateze aparițiile cu probabilitate mare ale expresiilor liniare ce implică biți de text clar, biți de text criptat și biți ai subcheilor. În acest caz se presupune că atacatorul cunoaște un set aleator de texte clare,

precum și textele criptate corespunzătoare. Se aplică algoritmilor simetrici, de tip bloc și a fost utilizată cu succes în criptanaliza algoritmului DES – precedentul standard decriptare simetrică în SUA. Sunt elaborate atacuri pentru cifrurile bloc și cele de tip flux. Descoperirea criptanalizei liniare a constituit un imbold pentru elaborarea noilor scheme de criptare.

Criptanaliza liniară este o tehnică prin care se urmărește construcția unui sistem de ecuații liniare între biții textului clar, ai textului cifrat și ai cheii. Rezolvarea acestui sistem de ecuații duce la aflarea cheii de cifrare. Sistemul de ecuații ce se construiește poate fi chiar un sistem probabilist, în sensul că o ecuație este verificată cu o anumită probabilitate.

Criptanaliza se face în două etape. Prima - construirea relațiilor dintre textul clar, textul cifrat și cheie, care sunt adevărate cu o probabilitate mare. A doua - utilizarea acestor relații, împreună cu perechile cunoscute de text clar și text cifrat pentru obținerea biților cheii.

La momentul actual pentru orice cifru nou este necesar de demonstrat că el este rezistent la criptanaliza liniară.

*Criptanaliza diferențială (differential cryptanalysis)* este o tehnică introdusă de Biham și Schamir prima dată în 1991 și concretizată în „Differential Cryptanalysis of the Data Encryption Standard” publicat în 1993. Tehnica face parte din categoria atacurilor cu text clar ales. În general fiind dată perechea  $(D, D')$ , numită caracteristică, tehnica constă în generarea de texte clare  $(P_1, P_2)$  cu  $D = P_1 - P_2$  astfel încât  $D' = C_1 - C_2$  și aflarea unei părți a cheii, restul fiind căutată exhaustiv. Operația „-” este operația de grup care, spre exemplu în cazul cifrurilor bloc, constă în adunarea cheii de runda.

Criptanaliza diferențială exploatează aparițiile cu mare probabilitate a diferențelor din textele clare, precum și a diferențelor apărute în ultimul ciclu al criptorului în care atacatorul poate selecta intrările și examina ieșirile în încercarea de a deduce cheia.

În final trebuie de subliniat că la elaborarea sistemelor de criptare trebuie să fie luate în considerare performanțele criptanalizei pentru a diminua riscurile posibile.

### **Întrebări și subiecte pentru aprofundarea cunoștințelor și lucrul individual**

1. Care este locul criptografiei în procesul de asigurare a securității informației?
2. Numiți și comentați cele cinci componente ale unui sistem de criptare a informației.
3. Faceți o clasificare a tipurilor de algoritmi de criptare. Care sunt avantajele și dezavantajele fiecărui tip de cifru?
4. Fie mesajul  $M = \text{NumePrenume}$ . Utilizând algoritmii de criptare Cezar, Afin, Polibios, Playfair, Vigenere și de transpoziție realizați criptarea și decriptarea mesajului  $M$ , alegând cheia de criptare potrivită (înlocuiți *NumePrenume* cu cele personale).
5. Realizați un studiu pentru a comenta argumentat efectele utilizării mașinii Enigma în cel de-al Doilea Război Mondial.
6. Descrieți etapele realizării unui atac criptografic bazat pe analiza frecvenței.