

**ACADEMIA TEHNICĂ MILITARĂ**

**Prof. univ. dr. ing.  
CIPRIAN RĂCUCIU**

**Lect. univ. dr. ing.  
MARIUS ROGOBETE**

**Conf. univ. dr.  
MADLENA NEN**

**CRIPTOGRAFIE  
ȘI  
SECURITATEA INFORMAȚIEI**

**CURS PENTRU ÎNVĂȚĂMÂNT LA DISTANȚĂ**

**BUCUREȘTI - 2016**

**Descrierea CIP a Bibliotecii Naționale a României**

**RĂCUCIU, CIPRIAN**

**Criptografie și securitatea informației : curs pentru învățământ la distanță** / prof. univ. dr. ing. Ciprian RĂCUCIU, lect. univ. dr. ing. Marius ROGOBETE, conf. univ. dr. Madlena NEN. - București : Editura Academiei Tehnice Militare, 2016

Conține bibliografie

ISBN 978-973-640-255-5

I. Rogobete, Marius

II. Nen, Madlena

004.056(075.8)

004.056.55(075.8)

Editat: *Academia Tehnică Militară*

Șef Editură: *Cpt. Răzvan CHIRIȚĂ*

Tehnoredactare: *Ciprian RĂCUCIU, Marius ROGOBETE,  
Madlena NEN, Camelia COMAN*

Corectură: *Magdalena Corina MAZILU*

Operațiuni tipografice: *Alina ANDREESCU, Viorica TOMA, Adrian STĂNICĂ*

Bun de tipar: **14.12.2016**

Hârtie ofset: **61 × 86**

Coli tipar: **17.5**

Format: **32 / 61 × 86**

Coli editură: **8.75**

Tiparul: *Academia Tehnică Militară*

Lucrarea conține **136** pagini

**0208**

C.C.-21

# CUPRINS

<b>CAPITOLUL 1 BAZELE MATEMATICE ALE SISTEMELOR DE SECRETIZARE ..</b>	<b>5</b>
1.1 Noțiuni de teoria numerelor .....	5
1.1.1 Numere prime.....	5
1.2 Congruențe .....	8
1.2.1 Noțiunea de congruențe .....	8
1.2.2 Congruența $ax \equiv b \pmod{m}$ .....	10
1.2.3 Congruența de gradul al II-lea.....	14
1.2.4 Logaritmi discreți .....	16
1.3 Inele de polinoame .....	19
1.3.1 Proprietăți generale .....	19
1.3.2 Factorialitatea inelelor de polinoame.....	22
1.3.3 Criterii de ireductibilitate .....	25
1.3.4 Polinoame ciclotomice .....	27
<b>Test de autoevaluare</b> .....	<b>30</b>
<b>Tema de autoinstruire</b> .....	<b>30</b>
<b>Lucrare de verificare</b> .....	<b>30</b>
<b>CAPITOLUL 2 BAZELE TEORETICE ALE SISTEMELOR SECRETE .....</b>	<b>31</b>
2.1 Introducere.....	31
2.2 Modelul matematic al sistemului secret .....	31
2.3 Reprezentarea sistemelor secrete .....	33
2.4 Compunerea sistemelor secrete .....	35
<b>Test de autoevaluare</b> .....	<b>48</b>
<b>Tema de autoinstruire</b> .....	<b>48</b>
<b>Lucrare de verificare</b> .....	<b>48</b>
<b>CAPITOLUL 3 SUCEȘIUNI PSEUDOALEATOARE ÎN SECRETIZAREA</b>	
<b>INFORMAȚIEI .....</b>	<b>49</b>
3.1 Sucesiuni de numere aleatoare.....	49
3.2 Teste de aleatorism.....	52
3.2.1 Conceptul de aleatorism.....	52
3.2.2 Teste de aleatorism .....	54
3.3 Scheme liniare și neliniare pentru generarea succesiunilor	
pseudoaleatoare .....	57
<b>Test de autoevaluare</b> .....	<b>67</b>
<b>Tema de autoinstruire</b> .....	<b>67</b>
<b>Lucrare de verificare</b> .....	<b>67</b>

<b>CAPITOLUL 4 METODE DE CIFRARE .....</b>	<b>68</b>
4.1 Corpuri Galois. Calcule.....	68
4.1.1 Câmpuri Galois .....	68
4.2 Funcții polinomiale .....	71
4.3 Metode de cifrare bazate pe funcții de permutare .....	72
4.3.1 Metoda utilizată în cazul $N=P^n$ .....	72
4.3.2 Metoda utilizată în cazul $M=p^n+1$ .....	73
4.3.3 Metoda de cifrare a mai multor câmpuri și funcții de permutare ....	75
4.3.4 Metoda de cifrare prin folosirea unui singur câmp .....	79
<b>Test de autoevaluare .....</b>	<b>81</b>
<b>Tema de autoinstruire .....</b>	<b>81</b>
<b>Lucrare de verificare .....</b>	<b>81</b>
<b>CAPITOLUL 5 DISPOZITIVE ȘI MAȘINI CRIPTOGRAFICE.....</b>	<b>84</b>
5.1 Dispozitive și mașini criptografice .....	84
5.1.1 Evoluția dispozitivelor criptografice.....	84
5.1.2 Utilizarea mijloacelor criptografice .....	86
5.2 Funcționarea mașinilor criptografice.....	87
5.2.1 Mașina „Enigma” .....	91
5.2.2 Mașini cu cifrare poligramică.....	95
5.2.3 Mașini tomogramice.....	98
<b>Test de autoevaluare .....</b>	<b>101</b>
<b>Tema de autoinstruire .....</b>	<b>101</b>
<b>Lucrare de verificare .....</b>	<b>101</b>
<b>CAPITOLUL 6 ELEMENTE DE CRIPTANALIZĂ .....</b>	<b>103</b>
6.1 Caracteristicile statistice ale limbajelor naturale .....	103
6.2 Metode de decriptare .....	113
6.3 Spargerea sistemelor criptografice .....	119
6.4 Spargerea sistemelor poligrafice.....	128
<b>Test de autoevaluare .....</b>	<b>132</b>
<b>Tema de autoinstruire .....</b>	<b>132</b>
<b>Lucrare de verificare .....</b>	<b>133</b>

## CAPITOLUL 1

### BAZELE MATEMATICE ALE SISTEMELOR DE SECRETIZARE

*Cuvinte cheie: divizibilitate, număr prim, congruență, inele.*

#### OBIECTIVE

Cunoașterea și înțelegerea următoarelor concepte din teoria numerelor:

- numere prime, numere prime Fermat și Mersenne
- congruențe, congruența modulo  $n$ , congruența de gradul al II-lea, logaritmi discreți în sisteme de secretizare cu chei publice
- inele de polinoame, proprietățile lor și polinoame ciclotomice; funcția lui Möbius

### 1.1 Noțiuni de teoria numerelor

#### 1.1.1 Numere prime

Având două numere naturale  $m$  și  $n$ , spunem că  $m$  divide pe  $n$ , sau că  $n$  este multiplu al lui  $m$ , dacă există un număr natural  $k$ , astfel încât:  $n = m \cdot k$ . În acest caz se scrie  $m/n$  sau  $m:n$ . Relația de divizibilitate pe  $N$  o vom nota cu  $|$ .

Pentru un număr  $n \in N$ , un număr  $m \in N$  se numește *divizor* al lui  $n$  dacă  $m/n$ . Deoarece  $n = 1 \cdot n$  și  $n = n \cdot 1$ , avem  $1/n$  și  $n/n$ , deci 1 și  $n$  sunt divizori ai lui  $n$ , pentru  $\forall n \in N$ . Numerele 1 și  $n$  se numesc *divizori improprii* ai lui  $n$ , iar orice alt divizor al lui  $n$  se va numi *divizor propriu*. Orice număr natural  $m$  este divizor al lui 0, deci  $0 = m \cdot 0$  și  $0 = 0|m$ .

Cu excepția numărului 1, care are un singur divizor, orice număr natural  $n > 1$  are cel puțin doi divizori distincți, aceștia fiind 1 și  $n$ . Un număr natural  $n > 1$  care are numai doi divizori se numește număr *prim*.

#### **Proprietăți**

- 1) Relația de divizibilitate este o relație de ordine pe  $N$ .

2) Pentru  $\forall n \in N$  avem  $n|n$ , deci relația este *reflexivă*.

3) Dacă  $m$  și  $n$  sunt două numere naturale și avem  $m|n$  și  $n|m$ , putem scrie:

$$n = k_1 m, \quad m = k_2 n, \quad \text{unde } k_1, k_2 \in N,$$

deci  $n = k_1(k_2 n) = k_1 k_2 n = n(k_1 k_2)$ , ceea ce arată că  $n = 0$  sau  $k_1 k_2 = 1$ .

Dacă  $n = 0$ , atunci  $m = k_2 n = 0 = n$ . Dacă  $k_1 k_2 = 1$ , atunci se demonstrează că  $k_1 = k_2 = 1$ , deci  $m = n$ . Putem spune că relația este *antisimetrică*.

4) Dacă  $m, n$  și  $p$  sunt numere naturale și avem:  $m|n$  și  $n|p$ , atunci:

$$n = k_1 m \text{ și } p = k_2 n = k_1 k_2 m.$$

Deci  $m|p$  și relația  $|$  este *tranzitivă*.

**Definiție.** Un număr natural  $p > 1$  este număr prim dacă și numai dacă pentru orice două numere naturale  $m$  și  $n$  avem:

$$p = m \cdot n \Rightarrow m = 1 \text{ sau } n = 1.$$

Numerele prime sunt foarte importante, în primul rând, datorită faptului că *orice număr natural nenul se scrie în mod unic ca un produs de numere prime*. Acest rezultat, cunoscut sub numele de *teorema fundamentală a aritmeticii*, a devenit, prin generalizările care i s-au dat, instrumentul de bază în multe capitole ale teoriei algebrice a numerelor și ale algebrei abstracte. Numerele prime sunt, de asemenea, importante, deoarece multe teoreme despre numere prime sunt ușor de formulat, dar foarte greu de demonstrat. Unele dintre aceste „teoreme” se dovedesc adevărate în toate cazurile accesibile calculului, prin mijloacele cunoscute până în prezent, dar aceste mijloace se dovedesc insuficiente pentru a se verifica valabilitatea generală a „teoremei” respective.

Una dintre primele probleme care s-a pus este dacă mulțimea numerelor prime este infinită sau nu. Răspunsul este dat de teorema lui Euclid.

**Teoremă.** Mulțimea numerelor prime este infinită.

Demonstrația se face foarte simplu dacă, prin reducere la absurd, presupunând că mulțimea numerelor prime este finită. Fie  $P = \{p_1, p_2, \dots, p_n\}$  această mulțime: considerăm numărul natural:  $N = p_1 \cdot p_2 \dots p_n + 1$ . Deoarece  $N > 1$ , există un număr  $p$  astfel încât  $p|N$ . Deoarece  $p \in P$ , rezultă  $p | p_1 p_2 \dots p_n$  și deci  $p | p_1 p_2 \dots p_n p | [N - (p_1 p_2 \dots p_n)] = 1$ . Așadar  $p|1$ , ceea ce contrazice faptul că  $p$  este număr prim.

Legat de faptul că mulțimea numerelor prime este infinită, s-a pus problema distribuției acestor numere. Notând cu  $\prod(x)$  numărul numerelor

prime mai mici decât  $x$ , se pune problema găsirii unei formule de calcul pentru  $\prod(x)$ . Mai mulți matematicieni au găsit experimental că:  $\prod(x) \cong \frac{x}{\ln x}$ , însă abia la sfârșitul secolului trecut *J. Hadamard* și *Ch. J. de la Valle Poussin* au demonstrat că:

$$\lim_{x \rightarrow \infty} \frac{\prod(x)}{\frac{x}{\ln x}} = 1.$$

S-a pus și problema dacă anumite mulțimi de numere prime, cu anumite proprietăți sunt infinite sau nu. Astfel, numerele prime de forma  $2^{2^n} + 1$  se numesc *numere prime Fermat*, iar numerele prime de forma  $2^p - 1$ , unde  $p$  este număr prim, se numesc *numere prime Mersenne*. Nu se cunoaște dacă mulțimea numerelor prime Fermat sau Mersenne este finită sau nu.

**Teoremă.** Dacă  $a$  și  $n$  sunt numere naturale,  $a \geq 1$  și  $n \geq 2$ , astfel încât  $a^n - 1$  este număr prim, atunci  $a = 2$  și  $n$  este număr prim.

Într-adevăr

$$a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \dots + a + 1)$$

și în ipoteza că  $a^n - 1$  este prim, rezultă că  $a - 1 = 1$ , deci  $a = 2$ .

Dacă  $m|n$ , deci  $n = m \cdot k$ , avem:

$$2^n - 1 = 2^{mk} - 1 = (2^m)^k - 1 = (2^m - 1) \left[ 2^{m(k-1)} + \dots + 2^m + 1 \right].$$

Rezultă că:  $2^{mk} - 1 = 2^m$  sau  $mk = m$ , deci  $m = n$  sau  $k = 1$ . Deci  $n$  este număr prim.

Se demonstrează că nu numai  $n$  este prim dacă și numai dacă:

$$\sum_{1 \leq m \leq n} \left( \left[ \frac{n}{m} \right] - \left[ \frac{n-1}{m} \right] \right) = 2,$$

unde  $[x]$  – partea întreagă a lui  $x$ .

Se întâlnesc două situații:

a)  $m|n$ , deci  $n = km$  și atunci

$$\left[ \frac{km}{m} \right] - \left[ \frac{km-1}{m} \right] = k - k + 1 = 1;$$

b)  $m$  nu divide pe  $n$ , deci  $n = km + s$ ;  $1 \leq s \leq m$

$$\left\lfloor \frac{km+s}{m} \right\rfloor - \left\lfloor \frac{km+s-1}{m} \right\rfloor = k - k = 0.$$

Deci dacă  $n$  este un număr prim, atunci el are doar cei doi divizori improprii, iar dacă  $n$  are  $\tau(n)$  divizori atunci:

$$\sum_{1 \leq m \leq n} \left( \left\lfloor \frac{n}{m} \right\rfloor - \left\lfloor \frac{n-1}{m} \right\rfloor \right) = \tau(n).$$

## 1.2 Congruențe

### 1.2.1 Noțiunea de congruențe

**Definiție.** Fie  $n$  un număr întreg pozitiv. Pe inelul  $\mathbb{Z}$  al numerelor întregi, definim o relație liniară, numită *congruența modulo  $n$* . Dacă  $a, b \in \mathbb{Z}$ , spunem că  $a$  este congruent cu  $b$  modulo  $n$  și scriem  $a \equiv b \pmod{n}$ , dacă  $n|(a-b)$ .

Relația de congruență este o relație de echivalență. Într-adevăr:

- 1) Dacă  $a \in \mathbb{Z}$ , atunci  $n|(a-a)$ , deci  $a \equiv a \pmod{n}$ . Deci relația este reflexivă.
- 2) Fie  $a, b \in \mathbb{Z}$ , atunci  $n|(a-b)$  și deci  $n|-(a-b)$  sau  $n|(b-a)$ , adică  $b \equiv a \pmod{n}$ . Așadar relația este *simetrică*.
- 3) Fie  $a, b, c \in \mathbb{Z}$ , astfel încât  $a \equiv b \pmod{n}$  și  $b \equiv c \pmod{n}$ . Atunci  $n|(a-b)$  și  $n|(b-c)$ , deci  $n|((a-b) + (b-c))$  sau  $n|(a-c)$ , de unde rezultă  $a \equiv c \pmod{n}$ . Deci relația este *tranzitivă*.

Pentru  $a \in \mathbb{Z}$  notăm cu:  $d = \{b \in \mathbb{Z} \mid a \equiv b \pmod{n}\}$  clasa de echivalență a lui  $a$ , numită clasa de resturi modulo  $n$ . Mulțimea claselor de resturi modulo  $n$  o vom nota  $Z_n$ :

$$Z_n = \{\bar{0}, \bar{1}, \dots, \bar{(n-1)}\}.$$

Pe mulțimea claselor de resturi modulo  $n$  se definesc operațiile algebrice de adunare și înmulțire a claselor de resturi modulo  $n$  în modul următor:

$$\begin{cases} \bar{a} + \bar{b} = \overline{a+b} \\ \bar{a} \cdot \bar{b} = \overline{ab}, \forall \bar{a}, \bar{b} \in Z_n. \end{cases}$$



Mulțimea înzestrată cu operațiile de adunare și înmulțire a claselor de resturi formează un inel unitar și comutativ. Aceasta se verifică simplu prin proprietățile:

$$1. (\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C});$$

$$2. \mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A};$$

$$3. \mathcal{A} + \mathcal{O} = \mathcal{O} + \mathcal{A};$$

$$4. \mathcal{A} + (-\mathcal{A}) = (-\mathcal{A}) + \mathcal{A} = \mathcal{O};$$

$$5. (\mathcal{A} \cdot \mathcal{B}) \cdot \mathcal{C} = \mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C});$$

$$\mathcal{A} \cdot (\mathcal{B} + \mathcal{C}) = \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{C},$$

$$6. (\mathcal{B} + \mathcal{C}) \mathcal{A} = \mathcal{B} \cdot \mathcal{A} + \mathcal{C} \cdot \mathcal{A},$$

$$7. \hat{a} \cdot \hat{1} = \hat{1} \cdot \hat{a} = \hat{a}, \quad \forall \hat{a}, \hat{b}, \hat{c} \in Z_n.$$

O clasă de resturi este ireversibilă în inelul  $Z_n$  dacă și numai dacă  $(a, n) = 1$  (cel mai mare divizor comun).

Ne propunem să descriem subnivelele și idealele inelului  $Z_n$ . Orice ideal și orice subinel al lui  $Z_n$  este în particular un subgrup al grupului aditiv  $(Z_n, +)$  al nivelului. Grupul  $(Z_n, +)$  este ciclic, fiind generat, de exemplu, de elementul  $\hat{1}$ . Atunci orice subgrup al său este, de asemenea, ciclic și deci este de forma:

$$\langle \hat{d} \rangle = \{ \hat{d}a / a \in Z \}, \text{ unde } \hat{d} \in Z_n.$$

Un astfel de subgrup este în același timp și subnivel și ideal al inelului  $Z_n$ . Deci subinelele și idealele lui  $Z_n$  coincid cu subgrupurile grupului  $(Z_n, +)$  al inelului.

De exemplu, să determinăm idealele inelului  $Z_6$ :

$$Z_6 = \{ \hat{0}, \hat{1}, \hat{2}, \hat{3}, \hat{4}, \hat{5} \}.$$

Cum doar  $\mathcal{P}$  și  $\mathcal{S}$  sunt inversabile în  $Z_6$ , rezultă că  $\langle \mathcal{P} \rangle = \langle \mathcal{S} \rangle$ . Considerând celelalte elemente din  $Z_6$ , obținem:

$$\begin{aligned}\langle \emptyset \rangle &= \{ \emptyset \}; \\ \langle \mathbb{Z} \rangle &= \langle \mathbb{Z} \rangle = \{ \emptyset, \mathbb{Z}, \mathbb{Z} \}; \\ \langle \mathbb{Z} \rangle &= \{ \emptyset, \mathbb{Z} \}.\end{aligned}$$

Deci  $Z_6$  are patru ideale, care sunt în același timp și subinelele sale:

$$\{ \emptyset \}, \{ \emptyset, \mathbb{Z} \}, \{ \emptyset, \mathbb{Z}, \mathbb{Z} \}, \{ \emptyset, \mathbb{Z}, \mathbb{Z} \} Z_6.$$

Mulțimea  $Z_n$  are  $n$  elemente. În general, numerele întregi  $a_1, a_2, \dots, a_n$  formează un sistem complet de resturi modulo  $n$  dacă:

$$Z_n = \{ \hat{a}_1, \hat{a}_2, \hat{a}_3, \dots, \hat{a}_n \}.$$

De exemplu, mulțimea:  $\{0, 1, 2, 3, 4, 5\}$  formează un sistem complet de resturi pentru  $Z_6$ , dar și  $\{-6, 7, 14, 9, 16, -1\}$  formează tot un sistem de resturi modulo 6.

*Mulțimea claselor de resturi care sunt relativ prime cu modulul formează sistemul redus de resturi împreună cu numărul 1. În cazul  $Z_6$ ,  $\{1, 5\}$  este sistem redus de resturi.*

Mulțimea  $Z_n$ , înzestrată cu operațiile de adunare și înmulțire, definite prin:

$$\begin{aligned}\hat{a} + \hat{b} &= a + b; \\ \hat{a} \cdot \hat{b} &= a \cdot b,\end{aligned}$$

este un inel unitar și comutativ, iar aplicația  $p: Z \rightarrow Z_n$ , definită prin  $p(a) = \hat{a}$ ,  $a \in Z$  este un morfism de inele surjective.

Într-adevăr:

$$\begin{aligned}p(a + b) &= a + b = \hat{a} + \hat{b} = p(a) + p(b); \\ p(a \cdot b) &= a \cdot b = \hat{a} \cdot \hat{b} = p(a) \cdot p(b).\end{aligned}$$

### 1.2.2 Congruența $ax \equiv b \pmod{m}$

**Teoremă.** *Congruența  $ax \equiv b \pmod{m}$  are soluții dacă și numai dacă  $d$ , cel mai mare divizor comun al numerelor  $a$  și  $m$  ( $d = (a, m)$ ), divide pe  $b$ .*

**Demonstrație.** Fie  $\hat{x}_0 \in Z_m, x_0 \in Z$  o soluție a ecuației date. Atunci  $ax_0 \equiv b \pmod{m}$ , deci  $m \mid (ax_0 - b)$ , adică există  $y_0 \in Z$  astfel încât  $ax_0 - b = my_0$ . Dar există  $d \mid a$ ;  $d \mid m$ , rezultă  $d \mid (ax_0 - my_0)$ , deci  $d \mid b$ . Reciproc, să presupunem că  $d \mid b$ . Se știe că există  $x'_0$  și  $y'_0$  numere întregi, astfel încât

$ax'_0 + my'_0 = d$ . Luând  $b' = \frac{b}{d} \in Z$ , avem  $b = db' = a(x'_0b') + m(y'_0b')$ , deci  $a(x'_0b') \equiv b \pmod{m}$ , ceea ce arată că  $x'_0b' \in Z_m$  este o soluție a congruenței date. Fie  $d = (a, m)$  și să presupunem că  $d|b$ . Atunci, dacă  $m' = \frac{m}{d} \in Z$ , iar  $x_0$  este o soluție a congruenței  $ax \equiv b \pmod{m}$ , rezultă că  $x_0, x_0 + m', x_0 + 2m', \dots, x_0 + (d-1)m'$  sunt toate soluțiile congruenței date.

Evident că dacă  $(a, m)=1$ , atunci congruența  $ax \equiv b \pmod{m}$  are soluție unică. Aceasta înseamnă că dacă  $m$  și  $a$  sunt relativ prime, atunci există totdeauna  $a^{-1} \pmod{m}$ .

**Definiție.** Un element  $a \in Z_m$  este unitate în  $Z_m$  dacă există  $b \in Z_m$  astfel încât  $ab \equiv 1 \pmod{m}$ .

Se poate astfel descrie grupul  $U(Z_m)$  care conține unitățile inelului  $Z_m$ :

$$U(Z_m) = \{a_{i_1}, a_{i_2}, \dots, a_{i_n}\}.$$

În acest caz numărul  $n$  se numește indicatorul Euler al lui  $m$  și este egal cu numărul elementelor relativ prime cu  $m$  plus 1. Aceasta se notează  $\varphi(m)$ .

De exemplu  $U(Z_{12}) = \{1, 5, 7, 11\}$ , deci  $\varphi(12) = 4$ .

**Teorema lui Euler.** Pentru orice număr întreg  $a$  relativ prim cu  $m$  avem  $a^{\varphi(m)} \equiv 1 \pmod{m}$ .

**Demonstrație:** Fie  $\{a_1, a_2, \dots, a_{\varphi(m)}\}$  un sistem redus de resturi modulo  $m$ . Atunci și  $aa_1, aa_2, \dots, aa_{\varphi(m)}$  este de asemenea un sistem redus de resturi. Produsul  $a_1a_2 \dots a_{\varphi(m)} \equiv aa_1 \cdot aa_2 \dots aa_{\varphi(m)}$  este comutativ, deci  $a_1a_2 \dots a_{\varphi(m)} \equiv a^{\varphi(m)} \cdot a_1 \cdot a_2 \dots a_{\varphi(m)}$ . De aici rezultă  $a^{\varphi(m)} \equiv 1 \pmod{m}$ .

Dacă  $m = p$ ,  $p$  fiind un număr prim, atunci orice număr  $a$  cu  $p|a$  avem  $(a, p)=1$  și prin urmare congruența  $ax \equiv b \pmod{m}$  are soluție unică pentru orice  $b \in Z$ . În acest caz  $U(Z_p) = Z_p - \{0\}$  este corp.

Este evident că în cazul  $p$  este număr prim,  $\varphi(p) = p - 1$ .

O consecință imediată a teoremei lui Euler o constituie *teorema lui Fermat*.

Pentru orice  $p$  număr prim și  $a \in \mathbb{Z}$  astfel încât  $p|a$ , avem  $a^{p-1} \equiv 1 \pmod{p}$ .

Indicatorul Euler al unui număr întreg  $m$  este o funcție numerică multiplicativă. Deci dacă  $m = m_1 \cdot m_2$ , atunci  $\varphi(m) = \varphi(m_1) \cdot \varphi(m_2)$ .

Dacă  $m = p^\alpha$ ,  $p$  fiind număr prim, atunci în șirul  $1, 2, \dots, p, \dots, p^\alpha$  sunt  $p^\alpha - p^{\alpha-1}$  termeni relativ primi cu  $p^\alpha$ . Deci  $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$ .

Pe această bază se poate calcula indicatorul Euler când  $m = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \dots p_k^{\alpha_k}$

$$\varphi(m) = m \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_k}\right).$$

De exemplu, fie  $m = 120 = 2^3 \cdot 3 \cdot 5$

$$\varphi(120) = 120 \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{5}\right) = 120 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 32.$$

Se demonstrează că  $\sum_{d|n} \varphi(d) = n$ . De exemplu, pentru  $n = 12$ , avem

$$\begin{aligned} \sum_{d|n} \varphi(d) &= \varphi(1) + \varphi(2) + \varphi(3) + \varphi(4) + \varphi(6) + \varphi(12) = \\ &= 1 + 1 + 2 + 2 + 2 + 4 = 12. \end{aligned}$$

Dacă notăm cu  $\tau(n)$  numărul divizorilor lui  $n$  (inclusiv divizorii improprii) și  $n$  este dat de relația:  $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \dots p_k^{\alpha_k}$ , atunci  $\tau(n)$  se determină cu ajutorul relației  $\tau(n) = (\alpha_1 + 1) \cdot (\alpha_2 + 1) \dots (\alpha_k + 1)$  și este egal cu numărul termenilor produsului

$$P = (1 + p_1 + p_1^2 + \dots + p_1^{\alpha_1}) (1 + p_2 + p_2^2 + \dots + p_2^{\alpha_2}) \dots (1 + p_k + p_k^2 + \dots + p_k^{\alpha_k}).$$

**Teorema chineză a resturilor.** Fie  $m_1, m_2, \dots, m_s$  numere întregi cu  $(m_i, m_j) = 1$ , pentru orice  $i, j \in \{1, 2, \dots, s\}$ ,  $i \neq j$  și  $b_1, b_2, \dots, b_s$  numere întregi oarecare. Atunci există un număr întreg  $x$ , soluție a sistemului de congruențe:

$$\begin{aligned} x &\equiv b_1 \pmod{m_1} \\ x &\equiv b_2 \pmod{m_2} \\ &\dots\dots\dots \\ x &\equiv b_s \pmod{m_s} \end{aligned}$$

și, pentru orice altă soluție  $y$  a aceluiași sistem, avem  $x \equiv y \pmod{m}$ , unde

$$m = m_1 \cdot m_2 \dots m_s.$$

**Demonstrație.** Pentru fiecare  $i \in \{1, 2, \dots, s\}$  notăm  $n_i = \frac{m}{m_i}$ , cu  $(n_i, m_i) = 1$ .

Există numerele întregi  $u_i, v_i$ , astfel încât  $u_i m_i + v_i n_i = 1$ . Luăm  $e_i = v_i n_i$  și atunci este evident că  $e_i \equiv 1 \pmod{m_i}$  și  $e_i \equiv 0 \pmod{m_j}, i \neq j$ . Luând

$x \equiv \sum_{i=1}^s b_i e_i$  avem  $x \equiv b_i e_i \pmod{m_j}$ , deci  $x \equiv b_i \pmod{m_i}$  știind că  $e_i \equiv 1 \pmod{m_i}$ .

*Exemplu*

$$x \equiv 1 \pmod{7}; \quad x \equiv 4 \pmod{9}; \quad x \equiv 3 \pmod{5}.$$

Avem:

$$m_1 = 7; b_1 = 1; \quad m_2 = 9; b_2 = 4; \quad m_3 = 5; b_3 = 3.$$

Deci

$$m = m_1 \cdot m_2 \cdot m_3 = 315; \quad n_1 = \frac{m}{m_1} = 45; \quad n_2 = \frac{m}{m_2} = 35; \quad n_3 = \frac{m}{m_3} = 63.$$

Soluția generală a sistemului de congruențe va fi  $x \equiv \sum_{i=1}^3 b_i v_i n_i \pmod{m}$ .

Se calculează  $v_i$  ( $i = 1, 2, 3$ ), astfel:  $n_i \cdot v_i \equiv 1 \pmod{m_i}$ . Deci

$$n_1 v_1 \equiv 1 \pmod{m_1};$$

$$3v_1 \equiv 1 \pmod{7}; \quad 45 \pmod{7} = 3 \pmod{7};$$

$$v_1 \equiv 3^{\varphi(7)-1}; \quad \varphi(7) = 6;$$

$$v_1 \equiv 3^5 \pmod{7} = 5 \pmod{7};$$

$$n_2 v_2 \equiv 1 \pmod{m_2}; \quad 35 \pmod{9} = 8 \pmod{9};$$

$$8v_2 \equiv 1 \pmod{9}; \quad \varphi(9) = 6;$$

$$v_2 \equiv 8^5 \pmod{9} = 8 \cdot 8 \cdot 8 \cdot 8 \pmod{9};$$

$$v_2 \equiv 8 \pmod{9};$$

$$n_3 v_3 \equiv 1 \pmod{m_3}; \quad 63 \pmod{5} = 3 \pmod{5};$$

$$3v_3 \equiv 1 \pmod{5};$$

$$v_3 \equiv 3^3 \pmod{5} = 2 \pmod{5};$$

$$x = 1 \cdot 5 \cdot 45 + 4 \cdot 8 \cdot 35 + 3 \cdot 2 \cdot 63 = 148 \pmod{315}.$$

### 1.2.3 Congruența de gradul al II-lea

Fie congruența:  $ay^2 + by + c \equiv 0 \pmod{s}$ ,  $a \neq 0$ .

Prin prelucrări succesive, această congruență poate fi adusă la forma  $x^2 \equiv n \pmod{m}$ :

$$\begin{aligned} ay^2 + by + c &\equiv 0 \pmod{s} / 4a \\ 4a^2y^2 + 4aby + 4ac &\equiv 0 \pmod{4as} \\ (2ay + b)^2 + 4ac - b^2 &\equiv 0 \pmod{4as}. \end{aligned}$$

Notăm

$$\begin{cases} 2ay + b \equiv x \\ b^2 - 4ac \equiv n \text{ și ecuația devine } x^2 \equiv n \pmod{m} \\ 4as \equiv m \end{cases}$$

Numărul relativ prim  $m$  se numește *rest pătratic* modulo  $m$ , dacă congruența  $x^2 \equiv n \pmod{m}$  are soluții, și *nerest pătratic* în caz contrar.

În cazul  $m=2$ , congruența  $x^2 \equiv n \pmod{2}$  are totdeauna soluții, și anume:

$$\begin{aligned} x &= 1, \text{ pentru } n = 1; \\ x &= 0, \text{ pentru } n = 0. \end{aligned}$$

Pentru cazul  $m=p$  număr prim,  $m \geq 3$ , într-un sistem de resturi modulo  $p$  există  $\frac{p-1}{2}$  resturi pătratice și  $\frac{p-1}{2}$  neresturi pătratice.

Stabilirea faptului că  $n$  este sau nu rest pătratic modulo  $p$  se poate face cu ajutorul simbolului lui Legendre  $\left(\frac{n}{p}\right)$ , care se definește astfel:

$$\left(\frac{n}{p}\right) = \begin{cases} 1n, \text{ rest pătratic mod } p \\ -1n, \text{ nerest pătratic mod } p \end{cases}$$

O serie de proprietăți ale simbolurilor lui Legendre permit cu ușurință calculul acestora:

$$\text{a) } \left(\frac{n_1 n_2 \dots n_s}{p}\right) = \left(\frac{n_1}{p}\right) \cdot \left(\frac{n_2}{p}\right) \dots \left(\frac{n_s}{p}\right);$$

$$b) \left( \frac{n_1 n_2^2}{p} \right) = \left( \frac{n_1}{p} \right);$$

$$c) \left( \frac{-1}{p} \right) = (-1)^{\frac{p-1}{2}};$$

$$d) \left( \frac{2}{p} \right) = (-1)^{\frac{p^2-1}{8}};$$

$$e) \left( \frac{p}{q} \right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} \left( \frac{q}{p} \right).$$

Simbolul lui Legendre se utilizează în cazul în care  $p$  este număr prim. O generalizare a acesteia o reprezintă simbolul lui Jacobi care este valabil și pentru  $m = p_1 \cdot p_2 \dots p_r$ ,  $p_i =$  numere prime,  $p_i > 2$

$$\left( \frac{a}{m} \right) = \left( \frac{a}{p_1} \right) \cdot \left( \frac{a}{p_2} \right) \dots \left( \frac{a}{p_r} \right).$$

$$\text{Pentru orice } a \equiv a_1 \pmod{m} \text{ avem } \left( \frac{a}{m} \right) = \left( \frac{a_1}{m} \right).$$

Simbolul lui Legendre se poate calcula și cu ajutorul relației lui Euler:

$$\left( \frac{n}{p} \right) \equiv n^{\frac{p-1}{2}} \pmod{p}.$$

Pentru această relație există și o generalizare pentru resturi de ordin  $n$ , și anume: numărul  $a$  este rest de ordin  $n$  pentru modulul  $p$  atunci și numai atunci când:

$$a^{\frac{p-1}{d}} \equiv 1 \pmod{p},$$

unde  $d = (n, p-1)$ .

Există în acest caz  $\frac{p-1}{d}$  resturi de ordin  $n$  față de modulul  $p$ .

Numărul soluțiilor congruenței:  $x^2 \equiv n \pmod{m}$  îl notăm cu  $N(n, m)$  și se calculează astfel:

a) cazul  $m = p^\alpha$ ,  $\alpha \geq 1$ ,  $p$  număr prim:

$$N(n, p^\alpha) = \begin{cases} 0, & \left(\frac{n}{p}\right) = -1 \\ 2, & \left(\frac{n}{p}\right) = 1 \end{cases}$$

b) cazul  $m = p^\alpha; n = p^\beta \cdot n_1; \alpha > \beta$ :

$$N(p^\beta \cdot n_1, p^\alpha) = 0 \text{ dacă } \beta \text{ este impar;}$$

$$N(p^\beta \cdot n_1, p^\alpha) = p^{\frac{\beta}{2}} \cdot N(n_1, p^{\alpha-\beta}) \text{ pentru } \beta \text{ număr par.}$$

### 1.2.4 Logaritmi discreți

Fie un sistem de secretizare cu chei publice de tip exponențial, R.S.A., în care modulul  $m$  este făcut public, iar exponentul de criptare  $e$  este păstrat secret. Presupunem că un criptanalist a interceptat cel puțin o pereche  $(M, M^e)$  și încearcă să spargă sistemul, adică să găsească exponentul de decriptare  $d$ . Deci criptanalistul se găsește în fața problemei de a găsi  $\log_M M^e \pmod{m}$ . Acesta este un caz special de calculare a logaritmilor discreți. Tăria metodei constă tocmai în dificultatea calculării acestui logaritm. Dacă considerăm ecuația  $a^x = y$  pentru numere reale pozitive, problema găsirii lui  $x$  cunoscând pe  $a$  și  $y$  este aproape aceeași cu a găsi pe  $y$  știindu-i pe  $a$  și  $x$ . Găsirea soluției presupune efectuarea unor calcule mai mult sau mai puțin complicate și căutarea în tabele de logaritmi. Problema este complet diferită dacă se lucrează cu logaritmi discreți. Noțiunea generală de *logaritm discret* poate fi formulată după cum urmează. Fie  $g$  un element al unui grup finit  $G$  și fie  $y$  un alt element din  $G$ . Atunci orice număr întreg pozitiv  $x$ , astfel încât  $g^x = y$  se numește *logaritm discret* al lui  $y$  în baza  $g$ . Evident, orice element  $y$  al lui  $G$  are un logaritm discret în baza  $g$  dacă și numai dacă  $G$  este ciclic cu generatorul  $g$ . De exemplu, în grupul multiplicativ al numerelor întregi modulat, numai 1,2,4 au un logaritm discret în baza doi, dar toate numerele au un logaritm discret în baza trei:

$y$	1	2	3	4	5	6	$(3^x = y)$
$x$	6	2	1	4	5	3	



Dacă  $x \in GF(3^2)$ , atunci cu ajutorul ecuației  $x^2 + 2x + 2 = 0$  se pot genera toate elementele câmpului  $GF(3^2)$  ridicând la putere un element  $x$  numit element generator:

$i$	1	2	3	4	5	6	7	8
$\alpha^i$	$\alpha$	$\alpha + 1$	$2\alpha + 1$	2	$2\alpha$	$2\alpha + 2$	$\alpha + 2$	1

și atunci:

$y$	1	2	$\alpha$	$\alpha + 1$	$\alpha + 2$	$2\alpha$	$2\alpha + 1$	$2\alpha + 2$
$\log_{\alpha} y$	8	4	1	2	7	5	3	6

Grupurile cu cardinalitate mică nu prezintă dificultăți de calcul. Există și algoritmi eficienți de calcul al logaritmilor discreți pentru unele cazuri speciale. În general însă calculul logaritmilor discreți sunt considerați de aceeași dificultate ca și algoritmi pentru factorizarea lui  $m$ .

Algoritmul propus de Silver, Pohlig și Hellman, probabil cel mai bun algoritm general de acest tip, este exemplificat după cum urmează și el lucrează atunci când factorii primi ai lui  $m$  sunt mici.

Fie  $GF(q)$ ,  $q = p^n$  un câmp finit. Considerăm logaritmii discreți în baza  $g$ , unde  $g$  este generator pentru  $GF(q)$ . Pentru fiecare divizor  $d$  prim al lui  $q - 1$ , calculăm numerele:

$$a(i, d) = g^{\left(\frac{q-1}{d}\right)} (\text{mod } q), 0 \leq i < p.$$

Dacă fiecare  $d$  care îl divide pe  $q - 1$  este mai mic, atunci mărimea tabelului precalculate conținând numerele auxiliare  $a(i, d)$  este acceptabilă, destul de ușor de realizat.

$\alpha$ $i$	2	3	5
0	1	1	1
1	180	48	59
2		132	42
3			125
4			135

De exemplu,  $GF(181)$  și  $q=2$  (2 este într-adevăr un generator). Acum  $180 = 2^2 \cdot 3^2 \cdot 5$  și tabela numerelor  $a(i, d)$  arată ca mai sus.

Să calculăm acum logaritmul discret al lui 62 în baza 2. În general, dacă  $q-1 = \pi d^\alpha$ , atunci pentru a găsi logaritmul discret  $x$  al lui  $y$  în baza  $g$  este suficient să găsim  $(x, \text{mod } d^\alpha)$  pentru fiecare  $d$  din factorizarea lui  $q-1$ . Folosind teorema chineză a resturilor  $x$  este ușor de calculat din valorile  $(x, \text{mod } d^\alpha)$ . Pentru a calcula  $(x, \text{mod } d^\alpha)$ , considerăm reprezentarea acestui număr în baza  $d$ :

$$(x, \text{mod } d^\alpha) = x_0 + x_1 d + \dots + x_{\alpha+1} d^{\alpha-1}; 0 \leq x_i \leq d-1.$$

În exemplul dat considerăm factorul  $d^\alpha = 3^2$  și scriem  $(z, \text{mod } 181) = x_0 + 3x_1$ .

Pentru a afla  $x_0$  calculăm  $\left( y^{\frac{q-1}{d}}, \text{mod } q \right)$ , care este egal cu  $a(i, d)$  pentru un număr  $i$ . Alegem  $x_0 = i$ . În exemplul dat  $(62^{60}, \text{mod } 181) = 48$ , deci  $x_0 = 1$ . Acest lucru este valabil, în general, deoarece  $(g^{q-1}, \text{mod } q) = 1$  deci

$$y^{\frac{q-1}{d}} \equiv g^{\frac{\alpha(q-1)}{d}} = g^{\frac{x_0(q-1)}{d}} \equiv a(x_0, d) (\text{mod } q).$$

Pentru a-l obține pe  $x_1$ , calculăm mai întâi inversul  $g^{-x_0}$  al lui  $g^{x_0} (\text{mod } q)$  și considerăm  $y_1 = yg^{-x_0}$ . Dacă acum  $\left( y_1^{\frac{q-1}{d^2}}, \text{mod } q \right) = a(i, d)$ , atunci  $x_1 = i$ . Pentru a-l obține pe  $x_2$ , considerăm numărul  $y_2 = yg^{-x_0 - x_1 d}$  și calculăm  $\left( y_2^{\frac{q-1}{d^2}}, \text{mod } q \right)$ .

Procedura se repetă până când este găsit  $(x, \text{mod } d^\alpha)$ .

În exemplul dat găsim  $y_1 = 31$ . Aceasta implică faptul că  $\left( y_1^{\frac{180}{3^2}}, \text{mod } 181 \right) = (y_1^{20}, \text{mod } 181) = 1$ , deci  $x_1 = 0$  și  $z \equiv 1 (\text{mod } 9)$ .

Să considerăm acum factorul  $d^\alpha = 2^2$ . Trebuie să determinăm  $x_0 + 2x_1$ . Deoarece  $(62^{90}, \text{mod} 181) = 1$ , rezultă că  $x_0 = 0$ . Acum  $y_1 = y = 62$  și  $(62^{45}, \text{mod} 181) = 1$ , de unde  $x_1 = 0$  și  $z \equiv 0 \pmod{4}$ .

Acum considerăm  $d^\alpha = 5^1$ . Trebuie determinat doar  $x_2$ . Deoarece  $(62^{36}, \text{mod} 181) = 1$ , rezultă  $x_0 = 0$  și  $z \equiv 0 \pmod{5}$ .

Deci avem de rezolvat congruențele:

$$\begin{aligned} z &\equiv 0 \pmod{4} \\ z &\equiv 0 \pmod{5} \Rightarrow z \equiv 100 \pmod{180} \\ z &\equiv 0 \pmod{9} \quad \log 62 = 100 \end{aligned}$$

## 1.3 Inele de polinoame

### 1.3.1 Proprietăți generale

Fie  $R$  un inel comutativ și unitar și fie  $R^{(N)}$  mulțimea șirurilor  $f = (a_0, a_1, \dots, a_n, \dots)$ ,  $a_i \in R$  care au numai un număr finit de termeni  $a_i$  nenuli. Există deci un număr natural  $m$  astfel încât  $a_i = 0$ , pentru orice  $i > m$ .

1) Șirurile  $f = (a_0, a_1, \dots, a_n, \dots)$  și  $g = (b_0, b_1, \dots, b_n, \dots)$  sunt egale dacă și numai dacă:  $a_i = b_i$  pentru orice  $i$ . Pe mulțimea  $R^{(N)}$  se definesc două operații algebrice, adunarea și înmulțirea, în raport cu care  $R^{(N)}$  devine un inel comutativ și unitar.

2) Dacă  $f, g \in R^{(N)}$ , atunci  $k > \max\{m, n\}$   $f + g = (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n, \dots)$  aparține, de asemenea, mulțimii  $R^{(N)}$ . Într-adevăr, dacă  $m$  și  $n$  sunt două numere naturale astfel încât  $a_i = 0$ , pentru orice  $i > m$  și  $b_j = 0$ , pentru orice  $j > n$ , atunci  $a_k + b_k = 0$ , pentru orice  $k > \max\{m, n\}$ .

3) Se verifică ușor că  $(R^N, +)$  formează un grup abelian, elementul nul fiind  $(0, 0, 0, \dots)$ .

4) Pentru orice element  $f = (a_0, a_1, \dots, a_n, \dots)$  opusul său va fi  $-f = (-a_0, -a_1, \dots, -a_n, \dots)$ .

5) Înmulțirea pe  $R^{(N)}$  se definește astfel:

$$f \cdot g = (c_0, c_1, \dots, c_k, \dots), \text{ unde } c_k = \sum_{i+j=k} a_i b_j, \text{ pentru orice } k = 0, 1, 2, \dots$$

6) Să arătăm că  $f \cdot g \in R^{(N)}$ . Într-adevăr  $c_k = 0$ , pentru orice  $k > m + n$ .

7) Înmulțirea pe  $R^{(N)}$  este asociativă, comutativă și are element unitatea,  $(1, 0, 0, \dots)$ , proprietăți care se verifică cu ușurință.

8) În plus, operația de înmulțire este distributivă față de adunare, adică:

$$f(g + h) = fg + fh$$

$$(f + g)h = fh + gh$$

În concluzie, s-a demonstrat că  $(R^{(N)}, +)$  formează un inel unitar comutativ. Elementele acestui inel se numesc polinoame cu coeficienți în  $R$ .

Gradul unui polinom se notează cu  $\text{grad}(f)$  și este dat de coeficientul dominant  $a_n$  al polinomului  $f$ , deci:  $\text{grad}(f) = n$  dacă  $a_i = 0$ , pentru orice  $i > n$  și  $a_n \neq 0$ .

Notăm prin  $X$  polinomul  $(0, 1, 0, \dots)$  care se numește nedeterminata  $x$ .

Înmulțirea polinoamelor ne dă  $X^n = \begin{pmatrix} 0, 0, \dots, 1, 0, 3 \dots \\ \text{nori} \end{pmatrix}$ , cu 1 aflat pe poziția  $(n + 1)$ .

Fie  $f$  un polinom de grad  $n$  ai cărui coeficienți sunt  $a_0, a_1, \dots, a_n$ , adică

$$f = (a_0, a_1, \dots, a_n).$$

Folosind adunarea și înmulțirea definite pe  $R^{(N)}$ , se obține:

$$\begin{aligned} f &= (a_0, 0, 0, \dots, 0) + (0, a_1, 0, \dots, 0) + \dots + (0, 0, \dots, a_n, 0) = \\ &= (a_0, 0, 0, \dots) + (a_1, 0, 0, \dots) \cdot (0, 1, 0, \dots) + \dots + (a_n, 0, 0, \dots) \cdot \begin{pmatrix} 0, 0, \dots, 1, 0, 3 \dots \\ \text{nori} \end{pmatrix}, \end{aligned}$$

$$\text{deci } f = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n = \sum_{i=0}^n a_i x^i.$$

Mulțimea  $R^{(N)}$  se mai poate nota  $R[x]$  și se numește inelul polinoamelor în nedeterminata  $X$  cu coeficienți în inelul  $R$ .

Fie un inel și  $f, g$  două polinoame din  $R[x]$ , atunci:

$$1. \text{ grad}(f + g) \leq \max[\text{grad}(f), \text{grad}(g)];$$

$$2. \text{ grad}(fg) \leq \text{grad}(f) + \text{grad}(g).$$

Avem egalitate dacă  $f$  și  $g$  sunt nenule, iar cel puțin unul dintre coeficienții dominanți ai lui  $f$  și  $g$  nu este divizor al lui zero.

Fie  $R$  un inel comutativ și unitar și inelul polinoamelor  $R[x]$ . Atunci au loc afirmațiile:

- un element  $a \in R$  este inversabil în  $R$  dacă și numai dacă  $a$  este inversabil în  $R[x]$ ;
- dacă  $R$  este domeniu de integritate, atunci  $R[x]$  este de asemenea domeniu de integritate și  $U(R) = U(R[x])$ , unde  $u: R \rightarrow R[x]$  este un morfism al lui  $R$  pe  $R[x]$  și  $u(a) = (a, 0, 0, \dots)$ .

**Teoremă.** Fie  $R$  un inel comutativ și unitar și  $R[x]$  inelul polinoamelor de o nedeterminată cu coeficienți în  $R$  și  $u: R \rightarrow R[x]$  morfismul canonic. Atunci oricare ar fi inelul comutativ unitar  $S$ , morfismul de inele  $v: R \rightarrow S$  și  $x \in S$ , există un unic morfism de inele  $\varphi: R[x] \rightarrow S$  astfel încât  $u(x) = x$  și diagrama să fie comutativă.

*Demonstrație.* Să definim mai întâi morfismul  $\varphi$ . Dacă  $f \in R[x]$ ,

$$f = \sum_{i=0}^m a_i x^i, \text{ atunci } \varphi(f) = \sum_{i=0}^m v(a_i) x^i.$$

Arătăm că  $\varphi$  are proprietățile din enunț. Fie  $g = \sum_{i=0}^n b_i x^i$  un alt polinom din  $R[x]$  și să presupunem că  $n \geq m$ . Completând eventual polinomul  $f$  cu termeni ai căror coeficienți sunt zero, putem scrie:

$$f = \sum_{i=0}^n a_i x^i,$$

unde  $a_{m+1} = a_{m+2} = \dots = a_n = 0$ .

Atunci

$$\begin{aligned} \varphi(f + g) &= \varphi\left(\sum_{i=0}^n a_i x^i + \sum_{i=0}^n b_i x^i\right) = \sum_{i=0}^n v(a_i + b_i) x^i = \\ &= \sum_{i=0}^n (v(a_i) + v(b_i)) x^i = \sum_{i=0}^n v(a_i) x^i + \sum_{i=0}^n v(b_i) x^i = \varphi(f) + \varphi(g). \end{aligned}$$

Dacă notăm cu  $c_k$  coeficienții produsului  $fg$ , avem  $c_k = \sum_{i+j=k} a_i b_j$ , și cum

$v$  este un morfism de inele, obținem  $v(c_k) = \sum_{i+j=k} v(a_i) \cdot v(b_j)$ .

Ținând seama de acest lucru, se verifică imediat că:

$$v(fg) = v(f) \cdot v(g).$$

Comutativitatea diagramei se verifică ușor. Într-adevăr, dacă  $a \in \mathfrak{p}$ , avem:

$$(\varphi \cdot u)(a) = \varphi(u(a)) = \varphi(a) = \varphi(ax^0) = v(a)x^0 = v(a).$$

Deci  $\varphi \cdot u = v$ .

### 1.3.2 Factorialitatea inelelor de polinoame

Un inel  $R$  se numește inel factorial sau cu descompunere unică în factori primi (ireductibili) dacă orice element neinversabil și nenul din  $R$  se descompune într-un produs finit de elemente prime. Orice inel de polinoame de o nedeterminată este un inel factorial.

**Teoremă.** Fie  $R$  un inel factorial. Atunci inelul de polinoame  $R[x]$  este factorial. Pentru demonstrație avem nevoie de o serie de rezultate preliminare.

**Lema 1.** Fie  $a \in R$  și  $f = a_0 + a_1x + \dots + a_nx^n \in R[x]$ . Dacă  $a|f$ , atunci  $a|a_i$ , oricare ar fi  $i = 0, 1, \dots, n$ .

*Demonstrație.* Cum  $a|f$  rezultă că:

$$f = a(b_0 + b_1x + \dots + b_nx^n) = ab_0 + ab_1x + \dots + ab_nx^n,$$

deci  $a|a_i$ .

**Lema 2.** Fie  $R$  un domeniu de integritate. Dacă  $p \in R$  este un element prim din  $R$  atunci  $p$  este un element prim în  $R[x]$ .

*Demonstrație.* Fie  $f, g \in R[x]$  astfel încât  $p|fg$ . Presupunem că  $f = a_0 + a_1x + \dots + a_nx^n$  și  $g = b_0 + b_1x + \dots + b_mx^m$  și că  $p|f$  și  $p|g$ . Conform lemei anterioare, rezultă că există un  $a_k$  astfel încât  $p|a_k$ . Analog, din  $p|g$  rezultă că există  $b_l$  astfel încât  $p|b_l$ . Deci  $p|a_0, p|a_1, \dots, p|a_{k-1}$  și  $p|b_0, p|b_1, \dots, p|b_{l-1}$ .

Coeficientul lui  $x^{k+l}$  este

$$c_{k+l} = \sum_{i+j=k+l} a_i b_j = a_0 b_{k+l} + a_1 b_{k+l-1} + \dots + a_{k-1} b_{l+1} + a_k b_l + a_{k+1} b_{l-1} + \dots + a_{k+l} b_0.$$

Deoarece  $p \nmid a_i b_i$  cu  $i \neq k$  și  $j \neq l$ , iar  $p \mid a_k b_l$ , rezultă că  $p \mid c_{k+l}$ , deci  $p \mid fg$ , contradicție, deci trebuie ca  $p \mid f$  sau  $p \mid g$ .

Presupunem acum că inelul  $R$  este factorial și fie  $f = a_0 + a_1 x + \dots + a_n x^n$  un polinom din  $R[x]$ . Vom nota cu  $c(f)$  cel mai mare divizor comun al elementelor  $a_0, a_1, \dots, a_n$ .  $c(f)$  se numește conținutul polinomului  $f$ . Dacă  $c(f) = 1$ , atunci polinomul  $f$  se numește primitiv. Se observă că  $f = c(f) \cdot f'$ , unde  $f'$  este un polinom primitiv.

**Lema lui Gauss (3).** Dacă  $R$  este un inel factorial și  $f, g \in R[x]$ , atunci  $c(fg) = c(f) \cdot c(g)$ .

*Demonstrație.* Cum  $f = c(f) \cdot f'$  și  $g = c(g) \cdot g'$ , rezultă  $fg = c(f) \cdot f' \cdot c(g) \cdot g' = c(fg) f'g'$ . Deci  $c(fg) = c(f) \cdot c(g) \cdot c(f'g')$ .

Trebuie arătat că  $c(f'g') = 1$ . Presupunem că  $c(f'g') \neq 1$ , deci există  $p \in \mathfrak{p}$  element prim, astfel încât  $p \mid c(f'g')$ . Deci  $p \mid f'g'$ , deci  $p \mid f'$  sau  $p \mid g'$  rezultă că  $f'$  sau  $g'$  nu sunt primitive, rezultă o *contradicție*.

**Lema 4.** Fie  $R$  un inel factorial și  $f, g \in \mathfrak{p}[x]$ , unde  $g$  este un polinom primitiv. Dacă  $a \in \mathfrak{p}$ ,  $a \neq 0$ , și  $g \mid af$ , atunci  $g \mid f$ .

*Demonstrație.* Avem  $af = gh$ ,  $h \in R[x]$   $ac(f) = c(g) \cdot c(h) = c(h)$  deoarece  $c(g) = 1$ . Cum  $h = c(h) \cdot h'$ , rezultă că  $af = g \cdot c(h) \cdot h'$  sau  $af = g \cdot a \cdot c(f) \cdot h' : a \Rightarrow f = g \cdot c(f) \cdot h' = c(f) \cdot g \cdot h'$ , deci  $g \mid f$ .

**Lema 5.** Fie  $R$  un inel factorial cu corpul de fracții  $K$ , iar  $f, g \in R[x]$  două polinoame primitive. Atunci  $f$  și  $g$  sunt asociate în  $R[x]$  dacă și numai dacă sunt asociate în inelul  $K[x]$ .

*Demonstrație.* Evident că dacă  $f$  și  $g$  sunt asociate în  $R[x]$  sunt asociate și în  $K[x]$ . Invers, presupunem că  $f$  și  $g$  sunt asociate în  $K[x]$ . Deci există  $u \in K[x]$ , element inversabil, astfel încât  $g = fu$ . Cum  $u \in K$ , atunci putem scrie  $u = \frac{a}{b}$  cu  $a \in K$  și  $b \in K$ ,  $a \neq 0, b \neq 0$ . Deci  $bg = af$  și  $f/g$  și  $g/f$ , adică  $f$  și  $g$  sunt asociate în inelul  $R[x]$ .

**Lema 6.** Fie  $R$  un inel factorial cu  $K$  corpul său de fracții și  $f \in R[x]$  un polinom primitiv cu  $\text{grad}(f) \geq 1$ . Atunci  $f$  este ireductibil în  $R[x]$  dacă și numai dacă  $f$  este ireductibil în  $K[x]$ .

*Demonstrație:* Presupunem că  $f$  este ireductibil în  $R[x]$  și fie  $f = gh$ , cu  $g \in K[x]$ ,  $h \in K[x]$  și  $\text{grad}(g) \geq 1$ ,  $\text{grad}(h) \geq 1$ . Evident putem scrie  $g = \frac{a}{b}g_1$ ,  $a, b \in K$ ,  $(a, b) = 1$  și  $g_1 \in R[x]$ . Analog  $h = \frac{c}{d}h_1$ ,  $c, d \in K$ ,  $(c, d) = 1$ ,  $h_1 \in R[x]$ . În plus,  $\text{grad}(g) = \text{grad}(g_1)$  și  $\text{grad}(h) = \text{grad}(h_1)$ .

Deci  $f = \frac{a}{b}g_1 \cdot \frac{c}{d}h_1 = \frac{ac}{bd}g_1h_1$ ,  $g_1 = c(g_1) \cdot g'_1$ ;  $h_1 = c(h_1) \cdot h'_1$  ( $g'_1$  și  $h'_1$  sunt polinoame primitive).

Obținem  $f = u \cdot g'_1 \cdot h'_1$ ;  $u = \frac{ac}{bd}c(g_1) \cdot c(h_1)$  – element inversabil din  $R$ . Deci  $f$  și  $g'_1 \cdot h'_1$  sunt asociate în  $K[x]$ . Deci  $f$  și  $g'_1 \cdot h'_1$  sunt asociate și în  $R[x]$ . Deci  $f = v \cdot g'_1 \cdot h'_1$ ,  $v \in U[R]$ . Cum  $\text{grad } g_1 \geq 1$  și  $\text{grad } h_1 \geq 1$ , aceasta implică că  $f$  nu este ireductibil în  $R[x]$ , rezultând deci o *contradicție*.

Invers, presupunem că  $f$  este ireductibil în  $R[x]$  și că  $f = gh$ ,  $g, h \in R[x]$ . Cum  $f$  este ireductibil în  $K[x]$ , rezultă că  $g$  este inversabil sau  $h$  este inversabil în  $K[x]$ . Dacă  $g$  este inversabil în  $K[x]$ , rezultă că  $g \in K$ , adică  $g = a \in K$ . Deci  $f = ah$ . Cum  $f$  este primitiv, rezultă că  $a$  este inversabil în  $R$ . Deci  $f$  este ireductibil în  $R[x]$ .

**Teoremă.** Fie  $R$  un inel factorial. Atunci inelul de polinoame  $R[x]$  este factorial.

*Demonstrație:* Fie  $f \in R[x]$ . Putem scrie  $f = c(f)f_0$ , unde  $f_0$  este un polinom primitiv. Cum  $f_0 \in K[x]$ , iar  $K[x]$  este un inel factorial, rezultă că  $f_0 = f_1f_2 \dots f_n$ , unde  $f_1, f_2, \dots, f_n \in K[x]$  și sunt polinoame ireductibile. Putem scrie pentru  $f_i$ ,  $f_i = \frac{a_i}{b_i}g_i$ , unde  $a_i, b_i \in R$  și  $g_i \in R[x]$  este un polinom primitiv, deci  $g_i$  este ireductibil în  $R[x]$ . Rezultă că  $f_0$  se scrie sub forma  $f_0 = \frac{a}{b}g_1g_2 \dots g_n$ , unde  $a, b \in R$ . Cum  $f_0$  este primitiv, și produsul  $g_1g_2 \dots g_n$  este



primitiv. Din lema 6 rezultă  $f_0 = ug_1g_2\ldots g_n, u \in U[R]$ . Cum  $c(f)$  este un produs finit de elemente prime în  $R$ , care sunt elemente prime și în  $R[x]$ , rezultă conform lemei 3 că  $f$  este un produs de elemente ireductibile în  $R[x]$ . Scrierea lui în produs de elemente ireductibile în  $R[x]$  este unică. Să presupunem că  $f = f_1f_2\ldots f_n = g_1g_2\ldots g_m$ , unde  $f_i, g_i \in R[x]$  sunt elemente ireductibile în  $R[x]$ . Dacă  $\text{grad } f_i \geq 1$ , atunci evident  $c(f_i) = 1$ . Deci putem scrie:

$$f = f_1f_2\ldots f_sf_{s+1}\ldots f_n = g_1g_2\ldots g_rg_{r+1}\ldots g_m,$$

unde  $f_{s+1}, \ldots, f_n, g_{r+1}, g_m$  au gradul  $\geq 1$ . Din lema lui Gauss rezultă că  $f_1, f_2, \ldots, f_s$  și  $g_1, g_2, \ldots, g_r$  sunt asociate în divizibilitate în  $R$ . Cum  $R$  este factorial, rezultă că  $r = s$  și abstractie făcând de o renumerotare avem  $f_i \sim g_i$  \*) oricare ar fi  $1 \leq i \leq s$ . Rezultă:  $f_{s+1}\ldots f_n = g_{r+1}\ldots g_m$ .

Din lema 7 rezultă  $m = n$  și  $g_k \sim f_n$  în  $K[x]$ , oricare ar fi  $s+1 \leq k \leq n$ , deci  $g_k \sim f_n$  în  $R[x]$ . Deci descompunerea este unică.

### 1.3.3 Criterii de ireductibilitate

1) **Criteriul lui Eisenstein.** Fie  $R$  un inel factorial și  $K$  corpul funcțiilor sale, iar  $f = a_0 + a_1x + \ldots + a_nx^n$  din  $R[x]$ . Presupunem că există un element prim  $g_k \sim f_n$ ,  $p \in R$  cu proprietățile:

i)  $p \mid a_0, p \mid a_1, \ldots, p \mid a_{n-1}$ ;

ii)  $p \nmid a_n$ ;

iii)  $p^2 \nmid a_0$ .

Atunci polinomul este ireductibil în  $K[x]$ .

*Demonstrație.* Putem presupune că polinomul  $f$  este primitiv. Prin reducere la absurd vom presupune că  $f$  este reductibil în  $K[x]$ . Deoarece  $f$  este primitiv, rezultă că el este reductibil și în  $R[x]$ . Fie atunci  $f = gh$ , cu  $g, h \in R[x]$ , unde

$$g = b_0 + b_1x + \ldots + b_mx^m, \quad b_m \neq 0;$$

$$h = c_0 + c_1x + \ldots + c_rx^r, \quad c_r \neq 0.$$

Prin identificare rezultă  $a_0 = b_0c_0$ . Din  $p \mid a_0$  rezultă  $p \mid b_0$  sau  $p \mid c_0$  și cum  $p^2 \nmid a_0$  avem că  $p$  nu divide în același timp pe  $b_0$  și  $c_0$ . Să presupunem că

\*) Relația „ $\sim$ ” înseamnă asociat în divizibilitate. Dacă  $f \sim g$ , rezultă  $f \mid g$  și  $g \mid f$ .

$p \mid b_0$  și  $p \mid c_0$ . Deoarece  $p \mid a_n$  este clar că polinomul  $g$  are coeficienți care nu se divid cu  $p$ . Fie  $i$  minim astfel încât  $p \mid b_i$  și să considerăm

$$a_i = \sum_{k+l=i} b_k c_l = b_i c_0 + \sum_{j=1}^{i-1} b_j c_{i-j}, \text{ cu } i \leq m < n.$$

Deoarece  $p \mid b_0, p \mid b_i, \dots, p \mid b_{i+1}$ , rezultă că  $p \mid \sum_{j=1}^{i-1} b_j c_{i-j}$ ,  $i \leq m < n$ , și

cum  $p \mid b_i c_0$ , avem evident că  $p \mid a_i$ , contradicție.

*Aplicație.* Fie  $p$  un număr prim. Polinomul  $f = x^{p-1} + \dots + x + 1$  cu coeficienți întregi este ireductibil în  $\mathbb{Q}[x]$ . Într-adevăr, este suficient să demonstrăm că polinomul  $f(x+1)$  este ireductibil în  $\mathbb{Q}[x]$ . Avem:

$$\begin{aligned} f(x+1) &= \frac{(x+1)^{p-1} + (x+1)^{p-2} + \dots + (x+1) + 1}{1} = \frac{(x+1)^p - 1}{x+1-1} = \frac{(x+1)^p - 1}{x} = \\ &= \frac{x^p + c_p^1 x^{p-1} + \dots + c_p^{p-1} x + 1 - 1}{x} = x^{p-1} + C_p^1 x^{p-2} + \dots + C_p^{p-1} \end{aligned}$$

Observăm că  $p \mid C_p^k, 1 \leq k \leq p-1$ ,  $p \mid 1$  și  $p^2 \mid C_p^{p-1}$ . Conform criteriului lui Eisenstein,  $f(x+1)$  este ireductibil în  $\mathbb{Q}[x]$  și deci  $f(x)$  este de asemenea ireductibil în  $\mathbb{Q}$ .

**2) Criteriul de reducere:** Fie  $R$  și  $S$  două inele factoriale,  $\varphi: R \rightarrow S$  un morfism (unitar) de inele,  $\bar{\varphi}: R[x] \rightarrow S[x]$  morfismul de inele care extinde pe  $\varphi$ , adică dacă  $f = a_0 + a_1 x + \dots + a_n x^n \in R[x]$ , atunci  $\bar{\varphi}(f) = \varphi(a_0) + \dots + \varphi(a_n) x^n$ . Presupunem că  $f$  este un polinom primitiv în  $R[x]$ , astfel încât  $\text{grad}(f) = \text{grad}(\bar{\varphi}(f))$  și  $\bar{\varphi}(f)$  este ireductibil în  $S[x]$ . Atunci polinomul  $f$  este ireductibil.

*Demonstrație.* Prin absurd presupunem că  $f$  este reductibil în  $R[x]$ , adică  $f = gh$ , cu  $g, h \in R[x]$ , iar  $g$  și  $h$  nu sunt inversabile în  $R[x]$ . Cum  $f$  este primitiv, atunci trebuie ca  $\text{grad}(g) \geq 1$  și  $\text{grad}(h) \geq 1$ . Dar din egalitatea  $f = gh$  rezultă că  $\bar{\varphi}(f) = \bar{\varphi}(g) \cdot \bar{\varphi}(h)$ . Cum  $\text{grad}(\bar{\varphi}(g)) \leq \text{grad}(g)$  și  $\text{grad}(\bar{\varphi}(h)) \leq \text{grad}(h)$ , iar  $\text{grad}(\bar{\varphi}(f)) = \text{grad}(f)$  rezultă că:

$$\text{grad}(\bar{\varphi}(g)) = \text{grad}(g) \geq 1;$$

$$\text{grad}(\bar{\varphi}(h)) = \text{grad}(h) \geq 1.$$

Deci  $\overline{\varphi}(f)$  nu este ireductibil în  $S[x]$ , contradicție.

*Caz particular:* Fie  $p$  un număr prim,  $f = a_0 + a_1x + \dots + a_nx^n$  un polinom primitiv din  $Z[x]$  și  $\overline{f} = \overline{a}_0 + \overline{a}_1x + \dots + \overline{a}_nx^n$  polinomul din  $Z_p[x]$ , astfel încât  $\overline{a}_i$  este clasa de resturi modulo  $p$  a lui  $a_i$ , pentru orice  $i = 0, 1, \dots, n$ . Dacă  $\text{grad}(f) = \text{grad}(\overline{f})$  și  $\overline{f}$  este ireductibil, atunci și  $f$  este ireductibil.

*Aplicație.* Polinomul  $f = x^5 - 5x^2 + 1$  are coeficienți întregi și este ireductibil în  $Z[x]$ .

Într-adevăr, aplicăm criteriul de reducere pentru  $p = 2$ . Avem  $\overline{f} = x^5 + x^2 + 1 \in Z_2[x]$ ,  $\text{grad}(f) = \text{grad}(\overline{f})$ . Demonstrăm că  $f$  este ireductibil în  $Z_2[x]$ . Deoarece  $\overline{f}(\overline{0}) = \overline{1}$  și  $\overline{f}(\overline{1}) = \overline{1} \neq \overline{0}$ , rezultă că  $f$  nu are divizori de gradul întâi.

Deci  $x^5 + x^2 + 1 = (ax^2 + bx + c)(\alpha x^3 + \beta x^2 + \gamma x + \delta)$ ,  $a, b, c, \alpha, \beta, \gamma, \delta \in \mathbb{F}_2$ .

Prin identificare rezultă:

$$\begin{cases} a\alpha = 1 \\ \alpha b + a\beta = 0 \\ a\gamma + b\beta + c\alpha = 0 \\ a\delta + b\gamma + c\beta = 1 \\ b\delta + c\gamma = 0 \\ c\delta = 1 \end{cases}$$

Acest sistem nu are soluție, deci  $x^5 + x^2 + 1$  este ireductibil.

### 1.3.4 Polinoame ciclotomice

Pentru orice număr natural  $n \geq 1$ , rădăcinile complexe ale polinomului  $f_n(x) = x^n - 1$  se numesc rădăcini de ordin  $n$  ale unității. Vom nota cu  $U_n$  mulțimea acestora, adică

$$U_n = \{x \in \mathbb{C} / x^n = 1\}.$$

Mulțimea  $U_n$  conține exact  $n$  elemente, și anume:

$$x_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}; k = 0, 1, 2, \dots, n-1.$$

Mulțimea  $U_n$  înzestrată cu operația de înmulțire este un „grup ciclic”, adică un grup în care toate elementele sunt puteri ale unui anumit element. Dacă  $\alpha = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ , se demonstrează că dacă  $\alpha^p = 1$ , atunci  $p$  este multiplu de  $n$ .

Un element  $x_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$  se numește generator al grupului  $U_n$  (dacă prin ridicări succesive la putere se generează toate elementele mulțimii  $U_n$ ) dacă și numai dacă  $(k, n) = 1$ .

Grupul ciclic  $U_n$  are  $\varphi(n)$  generatori care sunt rădăcinile primitive de ordinul  $n$  ale unității.

Pentru  $n$  număr prim, mulțimea rădăcinilor primitive de ordin  $n$  este

$$P_n = \{\alpha, \alpha^2, \dots, \alpha^{n-1}\},$$

unde

$$\alpha = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, 0 \leq k \leq n-1.$$

Pentru orice  $n \in \mathbb{N}^*$ , polinomul

$$f_n = \prod_{\alpha \in P_n} (x - \alpha)$$

se numește al  $n$ -lea *polinom ciclotomic* (al  $n$ -lea polinom de diviziune circulară).

*Relația lui Dedekind.* Pentru orice  $n \geq 1$  are loc egalitatea

$$x^n - 1 = \prod_{d|n} f_d,$$

unde produsul se face după divizorii naturali  $d$  ai numărului  $n$ . Această relație, a cărei demonstrație este relativ simplă, poate fi utilizată ca o relație de recurență cu ajutorul căreia se pot determina din aproape în aproape polinoamele ciclotonice.

De exemplu, ne propunem să calculăm polinoamele ciclotonice  $f_1, f_2, f_3, f_4, f_5$  și  $f_6$ .

Evident,  $f_1 = x - 1$  și apoi  $x^2 - 1 = f_1 \cdot f_2$ . Deci  $f_2 = x + 1$ ,  $x^3 - 1 = f_1 \cdot f_3$ , rezultă  $f_3 = x^2 + x + 1$ ;

$$x^4 - 1 = f_1 \cdot f_2 \cdot f_4, \text{ deci } f_4 = \frac{x^4 - 1}{(x-1)(x+1)} = x^2 + 1;$$

$$x^5 - 1 = f_1 \cdot f_5, \text{ deci } f_5 = \frac{x^5 - 1}{x - 1} = x^4 + x^3 + x^2 + 1;$$

$$x^6 - 1 = f_1 \cdot f_2 \cdot f_3 \cdot f_6, \text{ deci } f_6 = \frac{x^6 - 1}{(x - 1)(x + 1)(x^2 - x + 1)} = x^2 - x + 1.$$

Pentru  $n = p$  număr prim:

$$f_p = \sum_{i=0}^{p-1} x^i.$$

Polinomul  $f_p$  este în acest caz ireductibil. Mai mult, se demonstrează (teorema lui Dedekind) că orice polinom  $f_n, n \geq 1, n \in \mathbb{N}$  este ireductibil în inelul  $\mathbb{Z}[x]$  și în inelul  $\mathbb{Q}[x]$ .

Polinoamele ciclotonice pot fi exprimate cu ajutorul funcției lui Möbius, care este definită astfel:

$$\mu(n) = \begin{cases} 1, & \text{dacă } n = 1; \\ (-1)^k, & n = P_1 P_2 \dots P_k, n \text{ liber de pătrate}; \\ 0, & \text{dacă există } p^2 \mid n, p \text{ număr prim.} \end{cases}$$

Pentru orice  $n \geq 1, n \in \mathbb{N}$  avem egalitatea

$$f_n = \prod_{d \mid n} (x^d - 1)^{\mu\left(\frac{n}{d}\right)}.$$

De exemplu

$$f_6 = (x - 1)^{\mu(6)} \cdot (x - 1)^{\mu(3)} \cdot (x^3 - 1)^{\mu(2)} \cdot (x^6 - 1)^{\mu(1)} = \frac{(x - 1)(x^6 - 1)}{(x^2 - 1)(x^3 - 1)} = x^2 - x + 1,$$

deoarece

$$\begin{aligned} \mu(1) &= 1; \\ \mu(2) &= -1; \\ \mu(3) &= -1; \\ \mu(6) &= 1. \end{aligned}$$

Pentru demonstrarea acestei relații se utilizează teorema de inversiune a lui Möbius.

Vom prezenta formulele de inversiune ale lui Möbius fără demonstrație:

a) formula aditivă

Dacă  $f(n) = \sum_{\substack{d \\ d|n}} g(d)$ , atunci  $g(n) = \sum_{\substack{d \\ d|n}} \mu(d) f\left(\frac{n}{d}\right)$ , unde

$$\mu(d) = \begin{cases} 1; & d=1 \\ (-1)^k; & d = P_1 P_2 \dots P_k \\ 0; & d = P_1 P_2 \dots P_j^\alpha, \quad \alpha > 1 \end{cases}$$

b) formula multiplicativă

Dacă  $f(n) = \sum_{\substack{d \\ d|n}} g(d)$ , atunci  $g(m) = \sum_{\substack{n \\ n|m}} f(n)^{\mu\left(\frac{m}{n}\right)}$ .

### Test de autoevaluare

1. Care este corolarul definiției numerelor prime, referitor la modul de scriere al acestora?
2. Dacă  $a$  și  $n$  sunt numere naturale,  $a \geq 1$  și  $n \geq 2$ , astfel încât  $a^n - 1$  este număr prim, ce valori pot avea  $a$  și  $n$ ? Demonstrați.

### Temă de autoinstruire

1. Propuneți o schemă logică de implementare a algoritmului Silver, Pohlig și Hellman (independentă de limbajul de programare).

### Lucrare de verificare

1. Enunțați teorema chineză a resturilor.
2. Ce este un număr prim?
3. Definiți noțiunea de congruență.
4. Definiți criteriul de ireductibilitate al lui Eisenstein.
5. Arătați că polinomul  $f = x^5 + 5x^2 + 1$ , care are coeficienți întregi, este ireductibil în  $Z[x]$ .
6. Enunțați teorema lui Euler.
7. Definiți noțiunea generală de logaritm discret.

## CAPITOLUL 2

### BAZELE TEORETICE ALE SISTEMELOR SECRETE

*Cuvinte cheie: cifrator, receptor, interceptor, mesaj, criptogramă, chei, algoritm  
cifrare, sistem de cifrare secret.*

#### OBIECTIVE

- Înțelegerea modelului matematic al unui sistem de cifrare
- Definiția sistemului de cifrare dată de Shannon
- Estimarea numărului de mesaje false, a numărului de decriptări cu cheie falsă și a numărului de mesaje false
- Sisteme de criptare compuse, proprietățile cifrurilor obținute și aplicarea lor în practică

#### 2.1 Introducere

Primele informații referitoare la criptografie datează de acum circa 4000 de ani și provin din Egiptul antic, dar abordarea teoretică a acestora și descrierea modelului matematic al unui sistem secret sunt de dată mult mai apropiată de zilele noastre.

O contribuție deosebită a avut-o, pe această linie, C.E. Shannon, care a introdus o serie de concepte și idei originale, cum ar fi noțiunile de secret într-un sistem de cifrare și redundanța unui anumit limbaj.

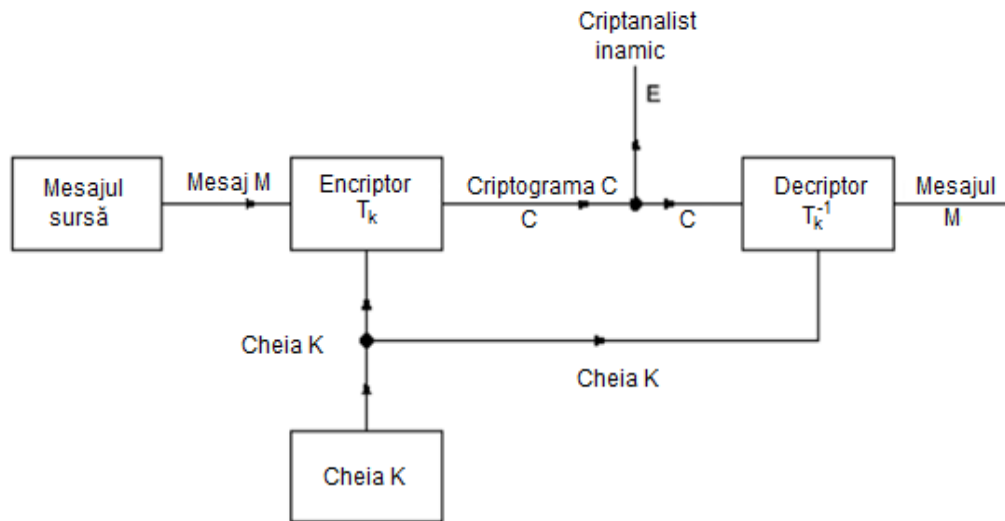
#### 2.2 Modelul matematic al sistemului secret

Figura 2.1 ilustrează un sistem de cifrare sau, cum este deseori numit, un sistem de secretizare.

Interceptorul este introdus în schemă pentru a arăta unde este cel mai probabil să aibă loc interceptarea.

Înainte ca un mesaj să fie transmis, cele două părți implicate, numite cifrator și receptor, își stabilesc o cheie particulară din mulțimea disponibilă, cheie care este păstrată secretă.

Folosind mesajul ce urmează a fi transmis și cheia aleasă, cifratorul îl cifrează cu ajutorul algoritmului utilizat înaintea transmiterii.



**Fig. 2.1 Un sistem de cifrare**

Mulțimea tuturor mesajelor posibile pe care cifratorul le poate transmite este numită „spațiul mesajelor” și o vom nota cu  $M$ . Mulțimea tuturor criptogramelor posibile este numită „spațiul criptogramelor” și o vom nota cu  $C$ , iar mulțimea tuturor cheilor o vom nota cu  $K$ . Mulțimile  $M$ ,  $C$  și  $K$  sunt mulțimi finite, dar, de obicei, prea mari pentru a le enumera explicit elementele. Procedura de cifrare poate fi redefinită după cum urmează:

- cifratorul și receptorul își aleg o cheie;
- cifratorul își alege și apoi folosește algoritmul de cifrare  $f$  pentru a determina în mod unic  $c \in C$ .

Fiecare cheie  $k_i \in K$  împreună cu algoritmul de cifrare determină o transformare  $t_{k_i} : M \rightarrow C$ .

Deci putem privi sistemul de cifrare ca un triplet  $(M, T, C)$ , unde  $T$  cuprinde totalitatea transformărilor posibile. Este foarte posibil ca utilizând două chei diferite să obținem aceeași transformare. Pot fi uneori mai multe chei decât transformări. Acest lucru poate determina ca la probabilități egale ale cheilor să obținem probabilități inegale ale transformărilor. Pe de altă parte, o anumită cheie determină în mod unic o transformare. Pentru simplificare vom considera că transformările și cheile sunt aceleași.

O cerință de bază a unui sistem de cifrare este următoarea: cunoașterea criptogramei, cheii și a algoritmului trebuie să facă receptorul să determine în mod unic mesajul.



Deci dacă  $C = t(m)$ , atunci  $m = t^{-1}(c)$ .

Putem da acum definiția lui Shannon pentru un sistem de cifrare: „un sistem de cifrare este o familie de transformări ireversibile dintr-o mulțime de mesaje  $M$  într-o mulțime de criptograme. Cele trei mulțimi  $T$ ,  $M$  și  $C$  se consideră, de obicei, finite”.

Diferența între cunoștințele receptorului și interceptorului constă în aceea că receptorul cunoaște care  $k$  a fost folosită, în timp ce interceptorul poate cunoaște doar probabilitățile „apriori” ale diferitelor transformări. Dacă toate cheile sunt egal probabile, atunci interceptorul are numai lista tuturor transformărilor posibile.

Definiția unui sistem de cifrare permite ca transformările și cheile să aibă probabilități diferite și este important de observat că acest lucru se poate întâmpla, și chiar se întâmplă.

## 2.3 Reprezentarea sistemelor secrete

Sistemele secrete pot fi reprezentate în diferite moduri. Unul dintre acestea folosește schemele liniare arătate în figura 2.2.

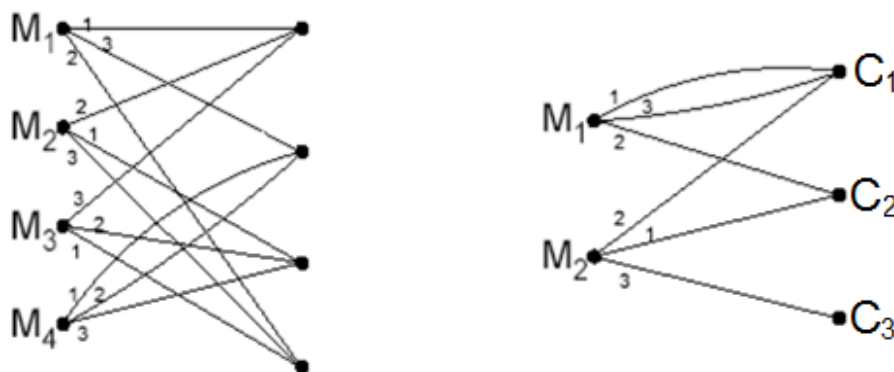


Fig. 2.2

Mesajele posibile reprezintă punctele din stânga, iar criptogramele posibile, punctele din dreapta. Dacă o anumită cheie (de exemplu 3) transformă mesajul  $M_2$  în criptograma  $C_4$ , punctele  $M_2$  și  $C_4$  se unesc cu o linie pe care se trece cifra 3. Pentru fiecare cheie din fiecare mesaj trebuie să rezulte numai o linie. Dacă acest text este adevărat și pentru fiecare criptogramă, sistemul secret se numește sistem închis. O metodă mai generală de descriere a sistemului secret constă în prezentarea operației cu ajutorul căreia, operând asupra mesajului cu o cheie oarecare, se poate obține criptograma.

De asemenea, se poate determina probabilitatea diferitelor chei, fie indicând metoda de alegere a cheilor, fie descriind modul cum, de regulă, adversarul alege cheile. Probabilitățile mesajelor se determină din datele apriori care se cunosc despre conținutul probabil al mesajului sau din orice informație specială care se referă la criptogramă.

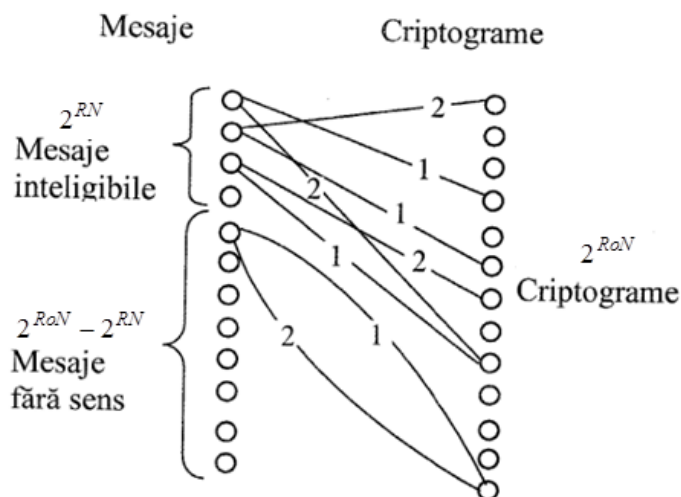
Dacă fiecare mesaj sau criptogramă conține  $N$  caractere dintr-un alfabet finit de  $L$  simboluri și dacă se notează cu  $R_c = \log_2 L$  viteza absolută de generare a limbajului, caracteristicile de bază ale sistemului arată ca în figura 2.3, unde:  $L^n = 2^{RoN}$  este numărul de mesaje posibile și egal cu numărul de criptograme:

$2^{RN}$  - mulțimea mesajelor cu sens, cu probabilitatea apriori  $a2^{-RN}$  ( $R$  viteza de generare a limbajului);

$2^{RoN} - 2^{RN}$  - mulțimea mesajelor fără sens în limbaj, atribuindu-li-se probabilitatea apriorică egală cu 0;

$2^{H(k)}$  - mulțimea cheilor, toate egal prealabile și apriori independente de mesaj:  $H(R)$  – entropia cheii.

Se observă că o linie cu numărul  $i$  indică cifrarea mesajului din stânga cu criptograma din dreapta, când s-a folosit cheia  $i$ .



**Fig. 2.3**

În figura 2.3 sunt prezentate 12 mesaje și două chei, deci  $H(k) = 1$ .

Se observă că aceeași criptogramă poate rezulta prin cifrarea a două sau mai multe mesaje diferite dacă se folosesc chei diferite.

Mesajele obținute prin descrierea unei criptograme folosindu-se o altă cheie decât cea utilizată la cifrare se numesc mesaje false. Numărul mesajelor

false  $n_m$  este o variabilă aleatoare, fiind determinată de cifru și de criptogramă. Dacă  $n_m$  ia valori mari cu probabilități apropiate de 1, sistemul este sigur chiar dacă criptanalistul are posibilități nelimitate de calcul.

Poate exista și situația când o criptogramă este descifrată în același mesaj cu mai multe chei. Astfel, criptanalistul știe care mesaj a fost transmis, dar nu știe care cheie a fost folosită. Această situație se numește descriere cu cheie falsă. Dacă notăm cu  $n_k$  numărul de decriptări cu cheie falsă, atunci în cazul unui cifru cu o bună folosire a cheii trebuie ca  $n_k = n_m$ .

În timp ce  $n_m$  prezintă interes mai mare, este mai simplu de calculat  $n_k$ . Astfel, dacă  $l(c)$  reprezintă numărul de linii care se termină în criptograma  $C$ , atunci  $n_k(c)$ , numărul descifrărilor corecte cu cheie falsă când este recepționată criptograma  $C$ , este:

$$n_k(c) = \max \{ \lfloor l(c) - 1 \rfloor, 0 \}.$$

## 2.4 Compunerea sistemelor secrete

Fiind date două sisteme secrete  $T$  și  $R$ , ele pot fi combinate în diferite feluri pentru obținerea unui nou sistem secret  $S$ .

Dacă  $R$  și  $T$  au același spațiu al mesajelor, se poate forma o sumă ponderată:

$$S = pT + qR;$$

$$q + p = 1.$$

Dacă  $T$  constă din prezentările:  $T_1, T_2, \dots, T_m$  cu probabilitățile  $p_1, p_2, \dots, p_m$ , iar  $R$  din reprezentările:  $R_1, R_2, \dots, R_k$ , cu probabilitățile  $q_1, q_2, \dots, q_k$ , sistemul rezultat,  $S$ , are reprezentările:  $T_1, T_2, \dots, T_m, R_1, R_2, \dots, R_k$ , cu probabilitățile:

$$p p_1, p p_2, \dots, p p_m, q q_1, q q_2, \dots, q q_k.$$

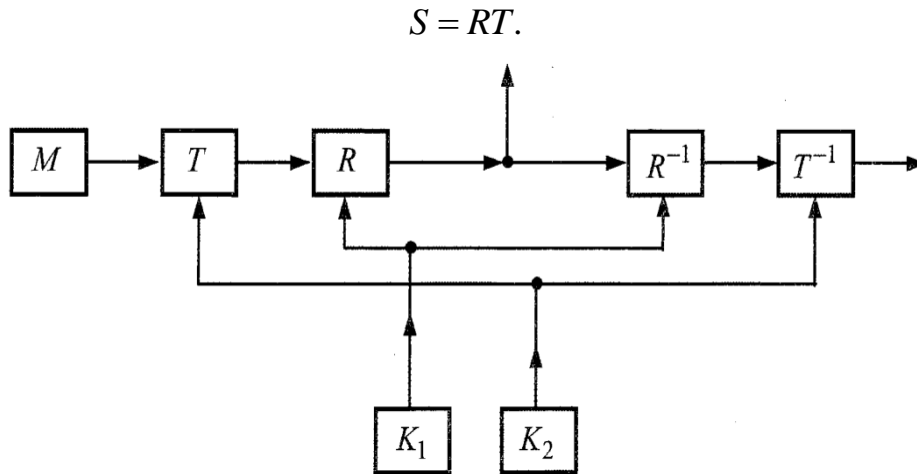
Generalizând, se poate forma suma mai multor sisteme:

$$S = P_1 T + P_2 R + \dots + P_m U; \quad \sum_{i=1}^m P_i = 1.$$

O altă metodă de combinare a două sisteme secrete constă în formarea produsului lor, ca în figura 2.4.

Să presupunem că  $T$  și  $R$  sunt două sisteme astfel încât spațiul mesajelor sistemului  $R$  poate fi identificat cu spațiul criptogramelor sistemului  $T$ . În acest caz se poate folosi sistemul  $T$  la mesajele inițiale, iar apoi sistemul  $R$ , la

rezultatul primei operații, ceea ce dă o operație rezultată  $S$ , care se poate scrie sub forma:



**Fig. 2.4**

Cheia sistemul  $S$  constă atât din cheia sistemului  $T$ , cât și din cea a sistemului  $R$ , aceste chei alegându-se independent și cu probabilitățile lor inițiale. Dacă sistemul  $T$  are  $m$  chei cu probabilitățile  $P_1, P_2, \dots, P_m$ , iar sistemul  $R$  are  $n$  chei cu probabilitățile  $P_j \cdot q_j, i = 1, m$  și  $j = 1, n$ .

Cifrurile de tip produs se folosesc în mod frecvent, dar trebuie remarcat că produsul, în general, nu este comutativ, adică nu totdeauna  $RT = TR$ . În schimb produsul este asociativ:

$$R(ST) = (RS)T = RST.$$

Se verifică de asemenea:

- legea asociativă ponderată pentru adunare:

$$p(p'T + q'R) + qS = pp'T + Pq'$$

- legea distributivă la stânga și la dreapta:

$$T(pR + qS) = pTR + qTS.$$

$$(pR + qS)T = pRT + qST.$$

Sistemele la care spațiile  $M$  și  $C$  coincid (caz frecvent când succesiunea la litere se transformă tot în succesiune de litere) se numesc endomorfe.

Sistemul endomorf  $T$  poate fi ridicat la putere:  $T^n$  există deci. Dacă  $T \cdot T = T$  sistemul se numește independent. Din aceste combinații rezultă metode pentru construirea unor noi tipuri de sisteme secrete din sisteme secrete date.

Proprietățile prezentate mai sus se pot folosi pentru descrierea situațiilor pe care le poate întâlni criptanalistul când încearcă să descifreze o anumită criptogramă. În mod real adversarul descifrează sistemul secret de tipul:

$$T = p_1A + p_2B + \dots + p_nS + p'X; \sum_j p_j = 1,$$

unde  $A, B, \dots, S$  sunt cifruri cunoscute cu probabilitățile lor apriori  $p_i$ , iar  $p'x$  reprezintă probabilitatea folosirii unui cifru nou necunoscut.

Sistemul în cazul căruia pentru fiecare din transformările  $T_i, T_j, T_k$  există o transformare  $T_S$ , astfel încât cheile să fie echiprobabile și să aibă loc relația  $T_i T_j - 1_{Tk} = T_S$  se numește sistem omogen.

În caz contrar sistemul se numește neomogen.

Sistemele omogene au o serie de proprietăți:

- în cifrul omogen operația care transformă spațiul mesajelor în sine însuși,  $T_i^{-1}T_j$ , formează un grup de ordinul  $m$ , unde  $m$  este numărul de chei diferite;
- produsul a două cifruri omogene comutative este un cifru omogen;
- mesajele se pot împărți în clase de resturi  $R'_1 R'_2 \dots R'_S$ .

Aceste clase de resturi sunt disjuncte, conțin toate mesajele posibile, dacă se cifrează un mesaj din clasa  $R_i$  se obține o criptogramă din clasa  $R'_i$ , fiecare mesaj din clasa de resturi  $R_i$  poate fi cifrat în fiecare criptogramă din  $R'_i$  cu ajutorul a exact  $K / \varphi_i$  chei diferite ( $K$  reprezintă numărul de chei, iar  $\varphi_i$  numărul de mesaje din clasa  $R_i$ ). Se spune despre două sisteme că sunt asemenea dacă există o transformare  $A$  care are o transformare  $A^{-1}m$ , astfel încât:

$$R = AS.$$

Aceasta înseamnă că cifrarea cu  $R$  dă același rezultat ca și cifrarea cu  $S$  dacă aceasta din urmă este aplicată ulterior transformării. Dacă notăm  $R \approx S$  ( $R$  asemenea cu  $S$ ), atunci se poate scrie:

$$\begin{aligned} R &\approx R(\text{reflexivitate}) \\ \left\{ \begin{array}{l} R \approx S \Rightarrow R \approx T (\text{tranzitivitatea}) \\ S \approx T \end{array} \right. \end{aligned}$$

Dacă  $R \approx S \Rightarrow S \approx R$  (simetrie). Deci dacă  $R \approx S$ , atunci cele două sisteme sunt echivalente din punctul de vedere al descifrării. Într-adevăr, dacă

criptanalistul advers interceptează o criptogramă din sistemul  $S$ , el o trece într-o criptogramă din sistemul  $R$ , aplicând o transformare  $A$ . Invers, criptograma, deci sistemul  $R$ , poate fi trecută într-o criptogramă din sistemul  $S$  prin transformarea  $A^{-1}$ . Din punct de vedere teoretic, se pun o serie de probleme în ceea ce privește asigurarea secretului. Se pune întrebarea cât de rezistent este un sistem secret, dacă criptanalistul advers nu este limitat de timp și are la dispoziție specialiști și mijloace tehnice pentru analizarea criptogramelor interceptate. De asemenea, se pune întrebarea dacă o criptogramă oarecare are una sau mai multe soluții și ce volum de text cifrat într-un anumit sistem trebuie să fie interceptat pentru a se obține o rezolvare unică. Trebuie dat răspunsul la întrebarea dacă există sisteme secrete la care nu se poate găsi rezolvarea unică indiferent de volumul textului cifrat interceptat. Pentru a putea defini un sistem secret perfect, să presupunem că există un număr finit de mesaje  $M_1, M_2, \dots, M_n$  cu probabilitățile lor apriorice  $p_1, p_2, \dots, p_n$ . Aceste mesaje sunt transformate în criptogramele  $C_1, C_2, \dots, C_n$ , astfel încât să se păstreze regula de cifrare  $C = T_i m$ . După interceptarea unei criptograme, se pot calcula probabilitățile aposteriori ale diferitelor mesaje,  $P(M / C)$ . Din rezolvarea acestor calcule rezultă că *sistemul secret perfect* se definește ca un sistem la care pentru toate criptogramele  $C$ , probabilitățile sunt egale cu probabilitățile apriori ale mesajelor, fără a depinde de acestea. În caz contrar, când probabilitățile aposteriori se deosebesc de cele apriori pentru anumite mesaje de chei, adversarul poate să-și corecteze acțiunile sale privind alegerea cheilor. Cu alte cuvinte, dacă un sistem secret este perfect, adversarul nu primește nicio informație suplimentară la recepționarea unei noi criptograme. Dacă notăm:  $P(M)$  – probabilitatea apriori a mesajului  $M$ ;  $P(C / M)$  – probabilitatea condiționată a criptogramei  $C$  cu condiția că s-a ales mesajul  $M$ ;  $P(C)$  – probabilitatea că se obține criptograma  $C$ ;  $P(M / C)$  – probabilitatea aposterioară a mesajului  $M$ , în condiția că s-a recepționat criptograma  $C$ , atunci se poate scrie condiția necesară și suficientă pentru ca un sistem secret să fie perfect;

$$P(M / C) = \frac{P(M) \cdot P(C / M)}{P(C)}.$$

Probabilitatea  $P(C/M)$  nu trebuie să depindă de mesaj, deci  $P(C / M) = P(C)$  și  $P(M / C) = P(M)$ .

Altfel spus, probabilitatea totală a tuturor cheilor care transformă mesajul  $M_i$  în criptograma  $C$  este egală cu probabilitatea totală a cheilor care transformă mesajul  $M_j$  în aceeași criptogramă  $C$  pentru orice  $M_i, M_j$  și  $C$ .

În cazul sistemului secret perfect trebuie ca numărul mesajelor  $M$  să fie egal cu numărul criptogramelor  $C$ , deoarece, pentru orice  $i$  dat, transformarea  $T_i$  dă o corespondență univocă între toate mesajele  $M$  și unele criptograme din  $C$ . Mai mult, există cel puțin o cheie care transformă pe  $M$  dat în orice criptogramă  $C$ .

Toate cheile care transformă pe  $M$  dat în diferite criptograme  $C$  trebuie să fie diferite, deci numărul de chei nu trebuie să fie mai mic decât numărul mesajelor  $M$ .

Un sistem secret perfect este reprezentat în figura de mai jos:

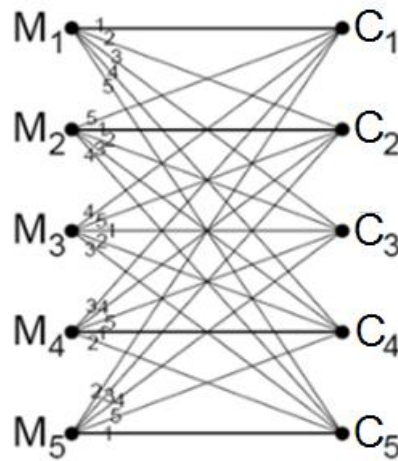


Fig. 2.5

După cum se observă, fiecare mesaj se leagă cu fiecare criptogramă cu câte o linie, toate cheile fiind echiprobabile. Cantitatea de informație obținută la alegerea mesajului sau cheii este:

$$H(M) = -\sum_{i=1}^m P(M_i) \log P(M_i),$$

$$H(C) = -\sum_{j=1}^m P(C_j) \log P(C_j).$$

Aceasta este maximă când mesajele sunt echiprobabile, deci  $P(M_i) = \frac{1}{n}$

și rezultă:

$$H(M)_{\max} \leq \log n.$$

Această informație este complet ascunsă când nedeterminarea cheii este maximă, adică:

$$H(C) = \log n.$$

Deci cantitatea de nedeterminare care poate fi introdusă într-un sistem secret nu poate fi mai mare decât nedeterminarea cheii, deci decât  $\log n$ .

În practică sistemele perfecte se folosesc pentru transmiterea unor documente de importanță deosebită sau când numărul de mesaje este mic. Neajunsul principal îl constituie faptul că aceste sisteme necesită chei cu volum mare.

Pentru depistarea unui text este necesar să se intercepteze un număr suficient de litere  $N$ . Dacă  $N$  este suficient de mare, este posibilă rezolvarea univocă a decriptării. Probabilitățile apriori ale literelor pentru diferite mesaje și cifruri se pot calcula din timp, iar cele aposteriori se pot determina după interceptarea criptogramelor. Pe măsură ce  $N$  crește, probabilitatea ca literele respective să reprezinte anumite mesaje crește, iar pentru alte mesaje scade. Procesul continuă până când rămâne un singur mesaj cu probabilitatea tinzând către 1. Această situație este similară cu cea din teoria informației, unde semnalul transmis este deformat de perturbații, ceea ce a determinat introducerea noțiunii de entropie condiționată ca măsură a nedeterminării semnalului transmis când se cunoaște semnalul recepționat.

Deci este normal să se folosească ca măsură a secretului entropia condiționată a cheii și a mesajului

$$H(K/C) = - \sum_{K,C} P(K,C) \log P(K/C),$$

$$H(M/C) = - \sum_{C,M} P(C,M) \log P(M/C),$$

unde  $P(K,C)$  este probabilitatea cheii  $K$  și criptogramei  $C$ ,  $P(K/C)$  – probabilitatea aposterioară a cheii  $K$  dacă s-a interceptat criptograma  $C$ , iar  $P(C,M)$ ,  $P(M/C)$  probabilitățile analoge pentru mesaje.

Din aceste relații se observă că rezistența sistemului secret este egală cu zero când cheia are probabilitatea egală cu 1, iar la toate celelalte probabilități este 0.

Entropia depinde de numărul de  $N$  interceptate, ea scăzând când  $N$  crește.

Rezistența sistemului secret are câteva proprietăți interesante, dintre care menționăm:



- rezistența cheii  $H(K/C)$  este o funcție necrescătoare cu  $N$ . Dacă s-au recepționat  $N$  litere, rezistența primelor  $N$  litere ale mesajului este mai mică sau egală cu rezistența cheii, deci:

$$H(K/C, S) \geq H(K/C, N), \quad S \leq N$$

$$H(M/C, S) \geq H(M/C, N)$$

$$H(M/C, N) \geq H(K/C, N)$$

- rezistența secretului în cazul unui sistem produs  $S = T R$  nu este mai mică decât rezistența unui sistem  $R$ . Dacă  $M_1, C_1$  și  $C_2$  sunt mesajul și respectiv criptogramele corespunzătoare celor două cazuri, atunci:

$$H(M/C_2) \geq H(M/C_1)$$

- pentru o sumă ponderată de sisteme, rezistența  $H$  este limitată de inegalitatea:

$$\sum_i P_i H_i \geq H > \sum_i P_i H_i - \sum_i P_i \log P_i,$$

unde  $H_i$  putând fi rezistența cheii sau rezistența mesajului.

Un sistem secret este ideal dacă  $H(K/C)$  și  $H(M/C)$  nu tind spre zero când  $N$  tinde spre infinit. Sistemul este strict ideal dacă:

$$H(X/C) = H(K).$$

Un exemplu de sistem secret ideal este simpla substituție folosită la un limbaj artificial în care toate literele sunt egal probabile.

În cazul limbajelor naturale, sistemele secrete pot în general să se apropie de caracteristica ideală, dar aceasta duce la creșterea rapidă a complexității sistemului.

Sistemele secrete reale au o serie de neajunsuri:

- sistemul trebuie să se găsească în strictă concordanță cu limbajul, deci cel care realizează cifrul trebuie să cunoască profund structura limbii;
- structura limbajelor naturale fiind complexă, înlăturarea reductanței cere transformări complicate, iar dispozitivul ce ar realiza această transformare ar trebui să fie foarte complicat;
- transformările folosite pot duce la multiplicarea erorilor. Eroarea produsă în transmiterea unei litere duce la erori într-un sector mare, definit de lungimea legăturilor statistice în limbajul inițial.

Orice sistem de cifrare cu un număr finit de mesaje poate fi ilustrat în felul următor: în stânga se scrie o coloană de puncte reprezentând mulțimea mesajelor și în dreapta o mulțime de puncte reprezentând mulțimea

criptogramelor. Dacă un mesaj  $m_i$  se transformă cu o cheie  $K_j$  într-o criptogramă  $C_k$ , atunci o săgeată etichetată  $K_j$  va uni  $m_i$  cu  $C_k$ .

O dată definit sistemul secret, se pot analiza cele cinci principii care stau la baza sistemelor secrete, principii elaborate de Shannon în 1940. Analiza va trebui să țină seama de evoluția tehnologiei în perioada care a trecut de atunci, mai ales în ceea ce privește tehnica de calcul. Cele cinci principii propuse de Shannon erau:

1. Cantitatea de secretizare oferită;
2. Mărimea „cheii”;
3. Simplitatea operațiunilor de cifrare și descifrare;
4. Propagarea erorii;
5. Extensia mesajului.

Importanța primului principiu este evidentă, complexitatea și tăria sistemelor secrete fiind direct proporțională cu importanța mesajelor transmise. În ceea ce privește „cheia”, aceasta trebuie ținută secretă, periodic ea trebuie schimbată și transmisă corespondenților de către cel care gestionează sistemul de „chei”. Deci cheia ar trebui să fie cât mai scurtă posibil. Pe de altă parte, „lungimea cheii” poate determina numărul de chei, deci mulțimea transformărilor făcute, care trebuie să fie cât mai mare. De aici rezultă că mărimea „cheii” trebuie să satisfacă două cerințe contradictorii.

Uneori se poate împărți sistemul de chei în clase, cum ar fi:

- chei de sistem;
- chei de mesaj.

Operațiile de cifrare și descifrare pot fi făcute manual sau automat. Dacă se lucrează manual, complexitatea poate duce la erori sau la creșterea timpului necesar efectuării operațiilor respective. Dacă se lucrează automat, complexitatea poate duce la necesitatea unor mașini sofisticate și scumpe.

În anumite sisteme secrete apariția unor erori de transmitere se poate propaga afectând porțiuni întregi din mesaj sau chiar întregul mesaj. De aceea este bine ca erorile de propagare să fie minimizate.

În multe cazuri, în urma operațiunii de cifrare, lungimea textului crește. O astfel de extensie a mesajului este de nedorit pentru majoritatea sistemelor de comunicație.

Din cele arătate mai sus rezultă că există o anumită incompatibilitate între cerințele celor cinci principii și este practic imposibil să fie satisfăcute.

Problema realizării unor sisteme de secretizare „sigure” a preocupat dintotdeauna proiectanții unor asemenea sisteme. Pentru a putea realiza sisteme cât mai rezistente la atacuri criptanalitice, trebuie cunoscute amenințările la adresa sistemelor secrete.

Trebuie presupus întotdeauna că un criptanalist are cunoștințe despre sistemul de secretizare și dispune de mijloacele și forțele necesare pentru a încerca să „spargă” sistemul secret analizat.

Proiectanții trebuie să fie capabili să evalueze „tăria” sistemului realizat sau timpul de „acoperire” (timpul necesar unui criptanalist pentru a-l „sparge”).

În scopul evaluării securității unui sistem, se fac următoarele presupuneri, care se consideră a fi condiții fundamentale:

- C1. Criptanalistul are cunoștințe complete despre sistemul de cifrare;
- C2. Criptanalistul a acumulat un volum considerabil de text cifrat;
- C3. Criptanalistul are la dispoziție un anumit volum de text clar și echivalentul său cifrat.

Condiția C1 implică faptul că nu există siguranță în sistemul de cifrare însuși și securitatea trebuie să fie dată de cheia utilizată. Firește, sarcina criptanalistului este mult mai grea dacă nu cunoaște metoda folosită. Dar este posibil ca algoritmul de cifrare să fie în posesia criptanalistului (uneori acesta este public).

În ceea ce privește celelalte două condiții, este posibil ca prin acumularea unui volum mare de text cifrat și prin faptul că toate comunicațiile între două surse încep cu un anumit antet, criptanalistul să aibă la dispoziție text clar și echivalentul său cifrat. O întrebare pe care trebuie să și-o pună criptograful (realizatorul unui sistem de cifrare): este dificil să determini parțial sau în totalitate mesajul cunoscând criptograma plus o mică parte din textul clar echivalent? Manual, răspunsul depinde de anumiți factori, printre care și lungimea textului cunoscut.

Desigur, cele trei condiții sunt destul de pesimiste din punctul de vedere al criptografului, dar dacă un sistem de secretizare nu ține seama de ele, acel sistem poate fi considerat necorespunzător.

Dezvoltarea sistemelor de cifrare a fost mult influențată de demonstrația lui Shannon, că sistemul cu șir cu unică folosință este practic imbatabil. Sistemul de chei este un șir aleator de lungime mult mai mare decât mesajul și care se utilizează o singură dată. De exemplu, dacă mesajul este un șir de biți:

$$m = 01001011101...,$$

iar cheia, un alt șir de biți:

$$k = 101100010111000....,$$

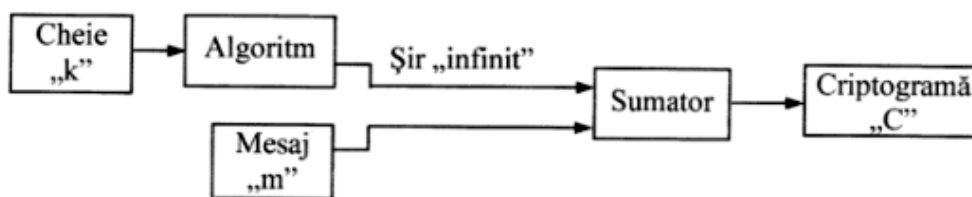
atunci criptograma  $C$  se poate obține „însumând” cele două șiruri (anticoincidență):

$$m = 01001011101...$$

$$k = 10110001011...$$

$$C = 11111010110.$$

Mulți criptografi și-au dat seama că dacă realizează un sistem mai bun decât acesta, sistemul realizat ar asigura un înalt grad de securitate. Un astfel de cifru este schematizat în figura 2.6.



**Fig. 2.6**

Pe baza unui sistem de chei și a unui algoritm, se obține un șir „infinit” aleator. Shannon a sugerat criptografilor utilizarea a două tehnici de cifrare, numite de el „difuzie” și „confuzie”. Ideea de „difuzie” se referă la împrăștierea caracteristicilor statistice ale spațiului mesajelor într-o structură statistică ce implică lungi combinații de litere din criptogramă. De exemplu, se urmărește modificarea frecvenței de apariție a caracterelor din spațiul mesajelor, în sensul micșorării dispersiilor.

Efectul este că un criptanalist trebuie să intercepteze o criptogramă mult mai lungă pentru a putea încerca o descifrare statistică. Prin tehnica „confuziei” se urmărește corelarea cât mai complexă între criptogramă și cheie, aceasta implicând dependența fiecărui caracter cifrat de întreaga structură de „cheie”. Sistemul secret cu acoperire unică prezentat mai sus se bazează pe tehnica „confuziei”. În plus, acest sistem nu prezintă fenomenul de propagare a erorii, necesitând, în schimb, o sincronizare perfectă între transmițător și receptor.

Spargerea acestui sistem implică cunoașterea șirului aleator în întregime, cunoașterea unor secvențe mai scurte sau mai lungi nu compromite întregul sistem. În general, s-ar putea asocia „țări” unui sistem secret cu entropia cheii, care este cu atât mai mare cu cât numărul lor este mai mare și cu cât au probabilități mai apropiate

$$H(k) = - \sum_{i=1}^n P_i \log_2 P_i.$$

Dacă  $P_i = \frac{1}{n}$ , atunci  $H(k) = \log_2 n$ .

Se poate considera că sunt două concepte diferite în ceea ce privește securitatea unui sistem: „securitate teoretică” și „securitate practică”.

Din punctul de vedere al criptografului, mai important este conceptul de „securitate teoretică”, el fiind interesat de șansele obținerii unei soluții unice prin încercarea fiecărei chei asupra criptogramei interceptate, precum și din determinarea cantității de criptograma necesară pentru obținerea soluției unice.

În sistemul secret se poate recurge la alegeri statistice: acelea ale mesajului clar și ale cheii. Entropia mesajului  $H(M)$  este dată de relația:

$$H(M) = - \sum_{m \in M} P(m) \cdot \log_2 P(m)$$

și reflectă siguranța cu care putem aprecia că un mesaj particular va fi transmis. Criptanalizii încearcă să determine ce mesaj a fost transmis și ce cheie a fost utilizată. Acest lucru ne conduce la ideea de entropie condiționată, numită de Shannon, *echivocitate*. Dacă pentru mulțimea criptogramelor  $C$  și întregul pozitiv  $n$ , notăm cu  $C_n$  mulțimea criptogramelor de lungime  $n$ , putem defini echivocitățile:

$$H_C(M, n) = - \sum_{\substack{m \in n \\ c \in C_n}} P(c, m) \cdot \log P_c(m)$$

$$H_C(K, n) = - \sum_{\substack{k \in K \\ c \in C_n}} P(c, k) \cdot \log P_c(k)$$

$P(c, m)$  reprezintă probabilitatea cu care se recepționează criptograma  $C$  fiind transmis mesajul  $m$ , iar  $P_c(m)$  este probabilitatea a posteriori cu care mesajul  $m$  a fost transmis știind că s-a recepționat criptograma  $C$ . Probabilitatea de a fi transmis oricare mesaj este egală cu probabilitatea de a fi aleasă cheia corespunzătoare și deci:

$$H_C(m, n) = H_C(k, n).$$

Dacă se consideră numărul literelor alfabetului egal cu 26 (deci mulțimea mesajelor are 26 de elemente), atunci, folosind logaritmul zecimal, obținem:

$$H_C(M, 1) = 1,26$$

$$H_C(M, 2) = 0,948$$

$$H_C(M, 3) = 0,394$$

$$H_C(M, 4) = 0,206$$

Echivocitatea dă o oarecare măsură cantitativă incertitudinii pe care o avem în prezicerea mesajului transmis (altfel spus, cheia utilizată). Creșterea lui  $n$  nu garantează ușurarea decriptării.

Dacă alfabetul sursei de mesaje are  $L$  caractere, iar sistemul folosește mesaje de lungime  $n$ , atunci vor exista  $L^n$  mesaje posibile.

Dacă notăm  $r_0 = \log_2 L$ , atunci numărul de mesaje de lungime  $n$  posibile va fi:  $2^{r_0 \cdot n}$ . Cele  $2^{r_0 \cdot n}$  mesaje posibile pot fi împărțite în două grupe:

- mesaje cu semnificație care au probabilități aproximativ egale  $2^{r_n \cdot n}$ ,

$$\text{unde } r_n = \frac{H_c(m, n)}{n};$$

- mesaje fără semnificație, care au probabilități neglijabile:  $2^{r_0 n} - 2^{r_n n}$ .

Dacă sistemul are  $h$  chei egal probabile, atunci  $h = 2^{H(k)}$ .

Probabilitatea ca un mesaj ales la întâmplare să aibă semnificație este:  $2^{-r_0 n + r_n n} = 2^{-n(r_0 - r_n)}$ . Astfel, dacă un mesaj  $m$  se codifică într-o criptogramă  $C$ , atunci probabilitatea de a găsi un mesaj cu semnificație prin descifrarea lui  $C$  este:  $2^{-n(r_0 - r_n)}$ . Dacă  $r_0 = r_m$ , atunci această probabilitate este egală cu 1. Dacă însă  $r_n \ll r_0$ , atunci această probabilitate este mică.

Diferența  $d_n = r_0 - r_n$  se numește *redundanță* a sistemului secret.

Deci dacă avem redundanță  $d_n$ , probabilitatea de a obține o decodificare semnificativă este  $2^{-nd_n}$ . Dacă încercăm toate cheile posibile ( $2^{H(k)}$ ), atunci numărul decodificărilor cu semnificație va fi:  $2^{H(k)} \cdot 2^{-nd_n} = 2^{H(k) - nd_n}$ . Dacă  $H(k)$  este mult mai mare decât  $nd_n$ , atunci probabilitatea de a obține o decodificare semnificativă este mai mare, deci probabilitatea de a determina mesajul corect este mică.

În practică mesajele nu au totdeauna aceeași lungime. Dacă  $n$  devine oricât de mare, s-a constatat că redundanța tinde asimptotic spre o anumită valoare  $\underline{d}$ .

Dacă interceptăm primele  $q$  litere ale unei criptograme, numărul așteptat de decodificări semnificative pentru aceste  $q$  poziții este:

Aceasta sugerează să determinăm cel mai mic întreg  $n_0$  care dă decodificarea unică așteptată cu relația:

$$n_0 = \frac{H(k)}{d}.$$

Vom încerca să determinăm  $n_0$  pentru un cod monoalfabetic. Având 26! chei egal probabile, atunci  $H(k) = \log_2 26! \cong 88,4$ . Alfabetul are 26 de litere, deci  $L = 26$  și deci  $r_0 = \log_2 26 \cong 4,7$ . Pentru calculul lui  $d$  trebuie determinat debitul real al limbii  $r_n$ . Nu există nicio cale de a furniza o evaluare precisă a probabilității oricărui cuvânt al limbii, dar diferite studii statistice oferă anumite aproximări ale probabilităților diferitelor cuvinte. De exemplu, pentru limba engleză, Hellmann estimează  $r_n \cong 1,5$ , iar Deavours,  $r_n ; 1$ . Deci  $d = 4,7 - 1,5 = 3,2$  (redundanța mai arată cât la sută din caracterele unui text sunt redundante. De exemplu, dacă  $d = 3,2$ , iar  $r_0 = 4,7$ , atunci  $\frac{3,2}{4,7}$ ; 68% din caractere sunt redundante.).

Este de așteptat că pentru orice criptogramă de lungime mai mare ca 28 să avem soluție unică. În practică însă e nevoie de mai mult de 28 de litere pentru a determina soluția unică. Dar cu cât se interceptează mai multe criptoграme, cu atât redundanța limbajului crește și deci unicitatea soluției se obține mai ușor. De asemenea, numărul de caractere interceptate trebuie să facă astfel încât probabilitatea de utilizare a tuturor literelor să fie suficient de mare.

Una dintre cerințele unui sistem secret aleator este ca pentru orice criptogramă dată  $C$ , decodificarea folosind toate cheile conduce la o selecție aleatoare a tuturor mesajelor, în particular la o selecție aleatoare atât a mesajelor semnificative, cât și nesemnificative. Aceasta implică în mod clar că numărul decodificărilor este mai mic decât putem garanta, dacă am știut că decodificarea unei criptoграme care a provenit de la un mesaj semnificativ trebuie să dea un mesaj semnificativ. Această ultimă condiție impune ca și spațiul criptoграмelor să fie împărțit în două grupe:

- criptoграme care provin de la codificarea mesajelor semnificative;
- criptoграme care nu provin din codificarea mesajelor semnificative.

Dacă este posibilă o astfel de împărțire a spațiului criptoграмelor, atunci se poate reduce spațiul mesajelor astfel ca toate mesajele să fie semnificative.

Dar îndepărtarea mesajelor nesemnificative este echivalentă cu îndepărtarea redundanței sistemului și în acest caz orice sistem de cod ar

satisfacă din punctul de vedere al securității. Acest lucru este practic imposibil. Structura tuturor limbajelor naturale este mult mai complexă. Un sistem secret se poate considera ideal dacă creșterea volumului de text interceptat nu este în mod necesar de ajutor, deci dacă  $H_c(k, n)$  nu tinde spre zero când  $n$  crește oricât de mult. Un exemplu de sistem ideal ar fi un cifru monoalfabetic pe un limbaj artificial în care toate literele sunt echiprobabile, iar literele succesive sunt alese independent.

### **Test de autoevaluare**

1. Ce este redundanța unui sistem secret și care este utilizarea sa practică?
2. Care sunt tipurile de tehnici de cifrare propuse de Shannon?

### **Tema de autoinstruire**

1. Propuneți o schemă logică de combinare a mai multor sisteme secrete prin formarea produsului lor, pornind de la figura 2.4.

### **Lucrare de verificare**

1. Estimarea numărului de mesaje false, a numărului de decriptări cu cheie falsă și a numărului de mesaje false.
2. Sisteme de criptare compuse, proprietățile cifrurilor obținute și aplicarea lor în practică.
3. Enunțați definiția lui Shannon pentru un sistem de cifrare.
4. Enunțați cele cinci principii care stau la baza evaluării sistemelor secrete (principii elaborate de Shannon în 1940).
5. Definiți procedura de cifrare.
6. Explicitați modul de construire a unui sistem secret  $S$  din două sisteme secrete  $R$  și  $T$ .
7. Enunțați proprietățile care se referă la rezistența sistemului secret.
8. Care sunt neajunsurile sistemelor secrete reale?



## CAPITOLUL 3

# SUCCESIUNI PSEUDOALEATOARE ÎN SECRETIZAREA INFORMAȚIEI

*Cuvinte cheie:* numere aleatoare, conceptul de aleator, generarea numerelor aleatoare.

### OBIECTIVE

Înțelegerea și cunoașterea următoarelor noțiuni:

- numere aleatoare, secvențe de numere aleatoare
- aleatorismul și teste de aleatorism: frecvență, serial, packer, autocorelare și spectral
- generarea succesiunilor pseudoaleatoare; scheme și aplicații

### 3.1 Succesiuni de numere aleatoare

O succesiune finită de numere sau evenimente de orice fel este aleatoare dacă ea a fost obținută într-un mod care să nu permită prevederea apariției diferitelor elemente ale succesiunii.

Noțiunea de *aleatoriu* se referă la condițiile apriorice pentru formarea succesiunii, și nu la stabilirea aposteriorică a caracterului și proprietăților succesiunii.

De cele mai multe ori succesiunile aleatoare se obțin în mod determinist dintr-o secvență scurtă, cu ajutorul unor generatoare de succesiuni, cum ar fi registrul de deplasare cu reacții liniare.

Cunoașterea mai multor secvențe din șir nu permite stabilirea legii de generare și de aceea se urmărește generarea unor succesiuni cu perioade lungi de repetiție, care sunt apoi testate asupra proprietăților lor aleatoare.

Un șir de numere reale

$$\{u_n\} = u_0, u_1, \dots, u_n, \dots, \quad 0 \leq u_i \leq 1$$

se numește succesiune de numere aleatoare dacă ele sunt alese la întâmplare.

Aceste succesiuni se dovedesc utile în multe tipuri de aplicații:

- simularea fenomenelor naturale;
- selectarea unui eșantion aleator pentru obținerea de informații despre ceea ce poate constitui comportare tipică;
- analiza numerică;
- luarea deciziilor în criptografie.

Șirul  $\{x_n\} = x_0, x_1, x_2, \dots$  se numește șir  $b$ -nar dacă oricare termen al șirului este unul dintre numerele întregi  $0, 1, 2, \dots, (b-1)$ . Un șir binar (2-nar) este format din 0 și 1.

La început, cei ce voiau să obțină numere aleatoare în diferite lucrări științifice le realizau prin simularea unor evenimente aleatoare (aruncarea unei monede sau extragerea unei bile dintr-o urnă etc.). După introducerea calculatoarelor electronice, obținerea numerelor aleatoare se face prin intermediul unor programe. Asupra acestor metode s-a ridicat obiecția cu privire la caracterul aleator al șirului generat, deoarece fiecare număr obținut este complet determinat de predecesorii săi. Șirurile de acest tip, generate în mod determinist, sunt numite *succesiuni pseudoaleatoare*. Generarea succesiunilor aleatoare lungi s-a dovedit o operație dificilă. Pericolul constă în faptul că șirul degenerază și tinde să se stabilizeze la anumite cicluri de elemente. De aceea s-au elaborat metode adecvate care să garanteze obținerea unor șiruri lungi de numere, cu proprietăți aleatoare.

O clasă de metode de generare o constituie metodele matematice. Una dintre aceste metode este și metoda *congruențial-liniară*.

În conformitate cu această metodă, șirul  $\{x_n\}$  se obține pe baza unei relații de recurență

$$x_{n+1} = ax_n + c \pmod{m},$$

în care:  $m$  este modulul și  $m > 0$ ;  $a$  se numește multiplicator,  $0 < a < m$ ;  $c$  se numește increment,  $0 \leq c < m$ ;  $x_0$  este termenul inițial,  $0 \leq x_0 < m$ .

Acești parametri, care se mai numesc și numere magice, trebuie să ia astfel de valori încât să permită obținerea unui șir cât mai mare.

Trebuie ca șirul de numere să aibă o distribuție uniformă pe o mulțime finită (deci toate numerele să aibă aceeași probabilitate de obținere).

Dacă se ia  $m=10$  și  $x_0 = a = c = 7$ , atunci șirul obținut va fi 6,9,0,7, după care se repetă. Șirul obținut nu este aleator. În general, șirurile care se obțin prin relații de recurență  $x_{n+1} = f(x_n)$ ; au această proprietate de a se închide într-un ciclu

(bucă), care apoi se repetă. Ciclul care se repetă se numește *perioadă*. În exemplul de mai sus, perioada este 4, cu valorile alese, dar *perioada maximă* este 10.

Pentru rezultate bune trebuie ca parametrii  $m, a, c, x_0$  să îndeplinească anumite condiții:

- $m$  să fie cât mai mare, având în vedere că perioada maximă este egală cu  $m$ ;
- $(c, m) = 1$ ;  $c$  și  $m$  să fie relativ prime între ele;
- $a$  să fie de forma  $a = p + 1$ , unde  $p$  este un divizor prim al lui  $m$ .

Dacă  $c = 0$ , generatorul se numește congruențial multiplicativ, relațiile devenind:

$$x_{n+1} = ax_n \pmod{m};$$

$$x_0 \neq 0 \text{ și } (x_0, m) = 1.$$

Numărul  $a$  trebuie să fie un element primitiv modulo  $m$ . Deci, dacă  $(a, m) = 1$ , iar  $\lambda$  este cel mai mic număr întreg pentru care  $a^\lambda = 1 \pmod{m}$ , numărul  $\lambda$  este numit *ordinul lui a modulo m*. Orice număr  $a$  pentru care  $\lambda$  este maxim posibil (în cazul nostru  $m - 1$ ) se numește *element primitiv modulo m*.

Von Neumann a lansat ideea folosirii procedeeleor aritmetice pentru generarea algoritmică de numere cu proprietăți aleatoare. Astfel, acesta propune metoda așa-numită a părții de la mijlocul pătratului, metodă care constă în următoarele:

- să presupunem că avem o reprezentare în baza  $b$  a numărului cu care lucrăm ( $b = 2$  sau  $b = 10$ );
- dacă numărul are  $2a$  cifre, atunci pătratul său va avea  $4a$  cifre (dacă numărul nu are  $2a$  cifre, se adaugă un 0 în față);
- se extrage din pătratul numărului un număr format din  $2a$  cifre aflate la mijloc. Deci:

$$x_{n+1} = \left[ \frac{x_n^2}{b^a} \right] - \left[ \frac{x_n^2}{b^{3a}} \right] \cdot b^{2a},$$

unde prin  $[x]$  se înțelege partea întreagă a lui  $x$ . Dar, de exemplu, dacă

$$x_0 = 3792$$

$$a = 2$$

$$b = 10$$

$$x_0^2 = 14379264,$$

rezultă  $x_1 = x_0$ .

Notând  $\left[ \frac{x^2}{b^{3a}} \right] = k$  și  $\left[ \frac{x^2}{b^a} \right] = \frac{x^2 - r}{b^a}$ , atunci găsirea numărului care se repetă revine la studierea ecuației diofantice:

$$x^2 - b^a x = r + b^{3a} k,$$

unde  $x \in [0, b^{2a}]$ ,  $r, k \in [0, b^a]$ .

A fost găsit ulterior un algoritm care permite determinarea numerelor repetabile în funcție de  $a$ ,  $b$  și  $r$ .

Altă metodă este metoda *congruențial aditivă*, adică:

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_k x_{n-k} \quad (\text{modulo } m).$$

Dacă  $m = p$  număr prim, atunci din teoria corpurilor finite se cunoaște că șirul definit de  $\{x_n\}$  cu relația de mai sus are *perioada maximă de lungime*  $p^k - 1$ . Dacă polinomul  $f(x) = x^k - a_1 x^{k-1} - \dots - a_k$  este primitiv modulo  $p$ , adică are o rădăcină care este element primitiv în corpul cu  $p^k$  elemente, constantele  $a_1, a_2, \dots, a_k$  se pot determina astfel încât perioada maximă să fie  $(p^k - 1)$ .

Pe lângă metodele matematice de generare a succesiunilor aleatoare pot fi utilizate unele procedee care se bazează pe fenomene fizice, de exemplu zgomotul electronic sau radioactiv.

De asemenea, s-au realizat tabele cu numere întâmplătoare. Aceste tabele conțin, de obicei, numere întregi uniform distribuite pe un interval. Cea mai mare tabelă de numere aleatoare este „A Million Random Digits and 100.000 Normal Deviates”, publicată în 1955 de Rand Corporation.

## 3.2 Teste de aleatorism

### 3.2.1 Conceptul de aleatorism

Se pune problema de a decide dacă o secvență periodică are sau nu proprietăți care să o apropie de o secvență aleatoare. Evident nicio secvență periodică nu este aleatoare în întregime. În criptografie, ceea ce se dorește de la o secvență periodică este să fie nedeterminată, astfel încât un criptanalist să nu poată prevedea ce urmează în continuare din secvența interceptată. În primul rând trebuie ca textul cifrat să fie mult mai scurt decât secvența periodică.

Așa cum s-a văzut, secvențele utilizate sunt pseudoaleatoare, generându-se cu ajutorul unor registre de deplasare cu reacție. Dar textul cifrat prin utilizarea unei astfel de secvențe trebuie să capete proprietățile specifice unui text aleator.

Pentru a putea da o definiție aleatorismului, trebuie definită o serie de termeni specifici șirurilor aleatoare binare. Vom utiliza astfel termenul de *serie* ca fiind un număr de biți identici. De exemplu, în secvența 1011000101100 sunt 8 serii:

- 3 serii de lungime 2;
- 1 serie de lungime 3.

Dacă o anumită secvență de lungime  $m$  are perioada  $p$ , atunci simbolurile de pe pozițiile  $m$  și  $m + p$  satisfac relația:

$$S_{m+p} = S_m.$$

Pentru fiecare  $\tau$  fixat putem compara biții secvenței date cu biții secvențelor translatate cu  $\tau$  biți.

Dacă notăm cu  $A$  numărul de coincidențe și cu  $D$  numărul  $(p - A)$ , atunci se poate defini funcția de autocorelație  $C(\tau)$ :

$$C(\tau) = \frac{A - D}{p}.$$

Evident,  $C(p + \tau) = C(\tau)$ , pentru orice  $\tau$ , astfel că putem considera că  $\tau$  verifică relația  $0 \leq \tau < p$ . Dacă  $\tau = 0$ , atunci avem autocorelație în fază, caz în care  $A = p$  și  $D = 0$ , iar  $C(0) = 1$ . Pentru  $\tau \neq 0$  vom avea autocorelație în afara fazei.

Golomb a propus trei criterii de aleatorism pentru secvențele binare:

C1. Dacă  $p$  este par, atunci secvența de lungime  $p$  va avea  $p/2$  biți de valoare 0 și  $p/2$  biți de valoare 1. Dacă  $p$  este impar, atunci numărul de zerouri și numărul de cifre unu diferă printr-o unitate;

C2. Într-o secvență de lungime  $p$ , jumătate din numărul seriilor va avea lungimea 1,  $1/4$  va avea lungimea 2, ...,  $1/2^k$  din numărul seriilor va avea lungimea  $k$ ;

C3. Autocorelația defazată este constantă.

Vom considera în continuare aceste trei criterii ca o necesitate acceptată pentru a putea spune că o secvență este „bună”.

### 3.2.2 Teste de aleatorism

Pe lângă criteriile de aleatorism propuse de Gaulomb, care sunt necesare dar nu și suficiente, se folosesc o serie de teste statistice pentru a determina proprietățile aleatoare ale unei secvențe. Vom prezenta, în continuare, cinci asemenea teste statistice, care pot fi realizate pentru a demonstra o măsură cantitativă a aleatorismului.

De asemenea, stabilim anumite niveluri de încredere pentru fiecare test, astfel încât să putem decide dacă o secvență a trecut sau nu testul respectiv. Vom nota în continuare cu  $n_0$  numărul de zerouri și cu  $n_1$  numărul de „1”.

#### Testul 1: Testul de frecvență

Acesta este testul cel mai evident și se aplică pentru a ne asigura că  $n_0 \cong n_1$ . Pentru aceasta se calculează

$$\chi^2 = \frac{(n_0 - n_1)^2}{n},$$

unde  $n = n_0 + n_1$ .

Evident că dacă  $n_0 = n_1$ , rezultă  $\chi^2 = 0$ .  $\chi^2$  este cu atât mai mare cu cât este mai mare discrepanța dintre frecvențele observate și cele așteptate. Pentru a decide dacă valoarea obținută este destul de bună astfel ca secvența analizată să treacă testul, trebuie doar să comparăm valoarea obținută cu un tabel al distribuției  $\chi^2$  (tabele care există și în care găsim valoarea lui  $\chi^2$  pentru anumite niveluri de semnificație). Dacă valoarea obținută este mai mică decât cea aflată în tabel, atunci secvența trece testul, în caz contrar ea este respinsă. De asemenea, secvența este respinsă dacă  $\chi^2 = 0$ , deoarece fiind prea bună poate fi suspectă.

#### Testul 2: Testul serial

Testul serial este folosit pentru a asigura că probabilitățile de tranziții sunt egale sau foarte apropiate, ceea ce ar demonstra că fiecare bit este independent de predecesorul său. Să presupunem că perechea 01 apare de  $n_{01}$  ori, perechea 10 de  $n_{10}$  ori, 00 de  $n_{00}$  ori și 11 de  $n_{11}$  ori. Evident că:

$$\begin{cases} n_{00} + n_{01} = n_0 \text{ sau } n_0 - 1 \\ n_{10} + n_{11} = n_1 \text{ sau } n_1 - 1 \end{cases}$$

(Apare  $n_0 - 1$  sau  $n_1 - 1$ , deoarece într-o secvență de lungime  $n$  sunt doar  $n - 1$  tranziții.)

Ideal ar fi ca  $n_{00} = n_{01} = n_{10} = n_{11} = \frac{n-1}{4}$ .

Pentru valorile  $n_{00}, n_{01}, n_{10}, n_{11}, n_1$  și  $n_0$  determinate, se calculează valoarea:

$$\frac{4}{n-1} \sum_{i=0}^l \sum_{j=0}^l (n_{ij})^2 - \frac{2}{n} \sum_{i=0}^l (n_i)^2 + 1,$$

unde  $l$  este numărul de caractere distincte.

Această valoare are o distribuție  $\chi^2$  cu două grade de libertate. Există tabele corespunzătoare cu valorile lui  $\chi^2$ . Se compară valoarea obținută cu cea din tabele și se decide dacă secvența trece testul sau nu.

### Testul 3: Testul packer

Pentru orice număr întreg  $m$  sunt  $2^m$  posibilități distincte pentru o secțiune de lungime  $m$ . În acest test împărțim secvența noastră în blocuri de lungime  $m$  și apoi numărăm fiecare tip de secțiune de lungime  $m$  din secvența analizată.

Dacă frecvențele sunt  $f_0, f_1, \dots, f_{2^m-1}$ , atunci  $\sum_{i=0}^{2^m-1} f_i = F = \left\lceil \frac{n}{m} \right\rceil$ , unde cu  $\left\lceil \frac{n}{m} \right\rceil$  am notat cel mai mare întreg care nu este mai mare decât  $\frac{n}{m}$  (partea întreagă),  $n$  – lungimea totală a secvenței. Atunci evaluăm:

$$\chi^2 = \frac{2^m}{F} \sum_{i=0}^{2^m-1} (f_i)^2 - F.$$

Se compară, în final, valoarea obținută cu cea găsită în tabele ale variabilei  $\chi^2$  cu  $2^m - 1$  grade de libertate și se decide dacă secvența trece sau nu testul.

O variantă de aplicare a acestui test ar fi următoarea:

- se împarte secvența în blocuri de  $m$  biți (de exemplu,  $m = 4$ );
- se transformă în baza 10 numărul scris cu patru biți, obținându-se numerele de la 0 la 15;
- șirul de numere astfel obținut se împarte la rândul său în grupe de câte 5 a căror structură poate fi:
  - cinci numere identice;
  - patru plus unu;

- trei plus două;
- două plus două, plus unu;
- trei plus unu, plus unu;
- două plus unu, plus unu, plus unu;
- cinci numere diferite;
- se determină frecvențele de apariție a acestor grupuri și se compară cu valorile calculate pentru un șir de numere aleatoare;
- pe baza unor diferențe admise, se decide dacă secvența analizată trece sau nu testul.

**Testul 4: Testul de autocorelare**

Dacă secvența de  $n$  biți pe care o testăm este  $a_1, a_2, \dots, a_n$ , atunci se calculează:

$$A(d) = \sum_{i=1}^{n-d} a_i a_{i+d}; \quad 0 \leq d \leq n-1.$$

Dacă  $d = 0$ , atunci  $A(0) = \sum_{i=1}^n a_i^2 = \sum_{i=1}^n a_i = n_1$ .

Dacă secvența are  $n_0$  cifre 0 și  $n_1$  cifre 1, care sunt distribuite aleatoriu, atunci valoarea estimată a lui  $A(d)$ ,  $d \neq 0$  este

$$\mu = \frac{n_1^2 (n-d)}{n^2}$$

Pentru o secvență de  $n$  biți se poate determina corelația care există între secvența dată și orice deplasare circulară a acesteia cu  $q$  biți ( $0 < q < n$ ). De exemplu, pentru  $q = 1$ , se calculează coeficientul de corelație serială  $c$ :

$$c = \frac{n(a_1 a_2 + a_2 a_3 + a_3 a_4 + \dots + a_{n-1} a_n) - (a_1 + a_2 + \dots + a_n)^2}{n(a_1^2 + a_2^2 + \dots + a_n^2) - (a_1 + a_2 + \dots + a_n)^2}.$$

Acest coeficient apare frecvent în statistică și are valori cuprinse între  $-1$  și  $1$ . Dacă  $c$  este 0, sau foarte apropiat de 0, indică o slabă corelație între biții secvenței, iar dacă  $c$  tinde spre 1 indică o dependență liniară totală.

**Testul 5: Testul spectral**

Se datorează lui COVEYOU și Mc. PHERSON și s-a impus datorită faptului că s-a dovedit că toate secvențele care au trecut toate celelalte teste de

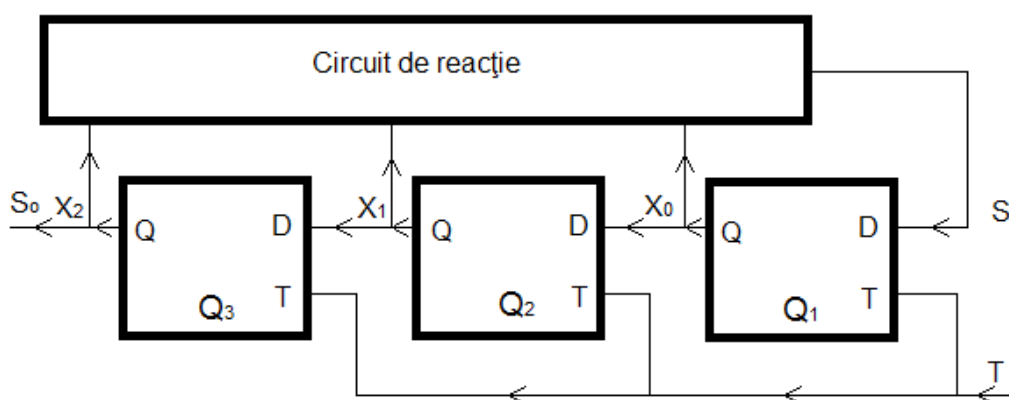


aleatorism sunt acceptate și de acest test, iar cele care au fost respinse de unul dintre teste sunt respinse și de acesta.

Testul se bazează pe tehnica transformărilor Fourier aplicate funcțiilor de variabile întregi, transformări care permit scoaterea în evidență a caracterului aleator al unei secvențe de numere întregi.

### 3.3 Scheme liniare și neliniare pentru generarea succesiunilor pseudoaleatoare

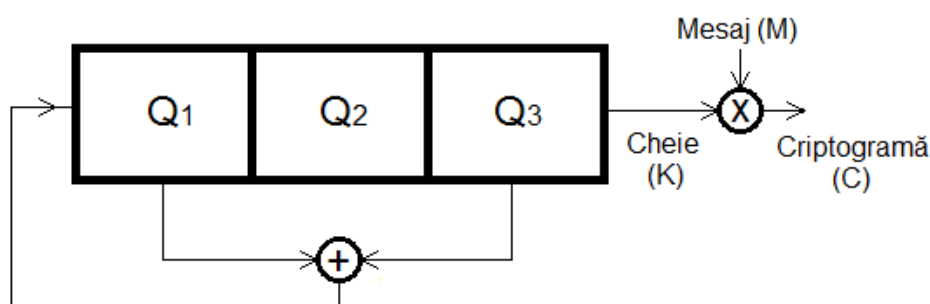
Pentru a produce secvențe aleatoare lungi dintr-un cod scurt, lucru dorit în criptologie, se utilizează în mod frecvent un generator format din registre de deplasare cu reacție. Caracteristica principală este reprezentată de adăugarea unui circuit logic combinațional cu funcția de reacție la registrul de deplasare. Prin efectul de reacție se înțelege introducerea unei variabile care se aplică intrării seriale a registrului, care va influența următoarea stare obținută.



**Fig. 3.1**

Un registru de deplasare cu trei celule de memorie tip CBB tip D și un circuit de reacție este dat în figura 3.1.

Un exemplu de registru de deplasare cu trei celule de memorie și un numărător modulo 2 este dat în figura 3.2



**Fig. 3.2**

Dacă inițial conținutul registrului este 111, prin deplasări succesive combinațiile formate din stările ieșirilor celor trei celule  $Q_1, Q_2, Q_3$  vor fi:

mod 10	Q3,Q2,Q1	Si
7	111	1
3	011	1
5	101	0
2	010	1
1	001	0
4	100	0
6	110	1
7	111	1

La ieșire se obține șirul 1110100, care se repetă și care se adună cu mesajul, obținându-se criptograma  $C$ . Perioada secvenței depinde de numărul de celule care compun registrul, precum și de alegerea conexiunilor la sumatorul modulo 2.

Pentru tratarea teoretică a succesiunilor și a schemelor generatoare, acestora li se atașează polinoame ale căror coeficienți  $(a_i)$  sunt elemente din câmpul  $GF(2)$ .

Succesiunea obținută la ieșire se poate de asemenea exprima polinomial, astfel:

$$a(x) = a_0 + a_1x + \dots + a_nx^n,$$

unde  $a_0, a_1, \dots, a_n$  constituie șirul de la ieșirea generatorului.

În exemplul de mai sus,  $a(x) = 1 + x + x^2 + x^4$ .

Polinoamele care sunt atașate schemelor semnifică, prin coeficienții acestora, conexiunile sumatoarelor de la diferite celule ale registrului.

De exemplu, în schema de mai sus,  $g(x) = 1 + x^2$ .

Conținutul informațional al registrelor se poate exprima, de asemenea, printr-un polinom de grad egal cu numărul de celule ale registrului de deplasare, cu puterile mici spre rangurile inferioare de la intrarea registrului și cu puterile mari spre rangurile de la ieșire.

În figura 3.2 se prezintă un registru de deplasare cu conexiune inversă care calculează relațiile de recurență liniară:

$$a_{i+k} = - \sum_{j=0}^{k-1} h_j \cdot a_{i+j}.$$

Se poate determina  $a_k$  și următoarele valori, pe baza celor  $k$  valori anterioare ale succesiunii.

Structura schemei este determinată de polinomul:

$$h(x) = h_0 + h_1x + \dots + h_kx^k,$$

unde  $h_0 \neq 0$  și  $h_k = 1$ .

Conținutul inițial al registrului este secvența  $a_0, a_1, \dots, a_k$ .

Prin deplasări succesive, la ieșirea schemei se obține la început conținutul registrului, apoi elementele formate pe baza combinațiilor liniare, până când se va găsi un moment  $n$  după care succesiunea se repetă.

Perioada de repetiție a succesiunii este dată de cel mai mic număr natural  $n$  pentru care polinomul  $x^n - 1$  se împarte exact la  $h(x)$ :

$$\frac{x^n - 1}{h(x)} = g(x).$$

Polinomul  $h(x)$  se numește *polinom caracteristic* al succesiunii  $\{a_n\}$  și al registrului de deplasare care o generează. Dacă gradul polinomului  $h(x)$  este  $k$ , gradul polinomului  $g(x)$  este  $(n - k)$ .

O succesiune generată de un registru de deplasare cu reacție cu  $k$  ranguri (celule) are lungimea maximă dacă perioada sa este  $2^k - 1$ .

Dacă secvența  $\{a_n\}$  are lungimea maximă, atunci polinomul  $h(x)$  este ireductibil peste câmpul coeficienților. Reciproca acestei afirmații nu este adevărată, deci pot exista polinoame caracteristice ireductibile care nu dau succesiuni de lungime maximă.

De exemplu, polinomul:  $h(x) = 1 + x + x^2 + x^3 + x^4$  este ireductibil peste  $GF(2)$ , dar nu este primitiv. El divide polinomul  $x^5 - 1$ , iar schema respectivă generează o succesiune de lungime 5, care nu este de perioadă maximă  $n = 2^4 - 1 = 15$ .

Dacă  $h(x)$  este reductibil,  $h(x) = s(x)t(x)$ , atunci exponentul lui  $h(x)$  este cel mai mic multiplu comun al exponenților lui  $s(x)$  și  $t(x)$ .

Când polinomul  $h(x)$  este ireductibil și primitiv, succesiunea obținută la ieșirea registrului de deplasare nu depinde de condițiile inițiale (cu excepția conținutului zero peste tot care trebuie evitată).

Dacă  $2^k - 1 = p$  este un număr prim, atunci fiecare polinom primitiv de grad  $p$  corespunde unei succesiuni de lungime maximă.

În cazul în care polinomul este un trinom de forma:  $x^p + x^q + 1$ , se demonstrează că rezultate mai bune se obțin dacă gradul  $p$  este ales astfel încât însuși  $2^p - 1$  să fie un număr prim (astfel de numere prime se numesc *numere prime Mersenne*).

Funcționarea registrului poate fi descrisă matriceal. Fiecare stare a unui registru de deplasare cu reacție având  $k$  ranguri poate fi considerată ca un vector  $k$ -dimensional. În acest caz, registrul de deplasare este un operator liniar care schimbă fiecare stare în starea următoare. Matricea unui registru de deplasare cu reacție este o matrice pătrată  $k \times k$  având forma:

$$M = \begin{bmatrix} h_1 & 1 & 0 & L & 0 \\ h_2 & 0 & 1 & K & 0 \\ K & K & K & K & K \\ h_{k-1} & 0 & 0 & K & 1 \\ h_k & 0 & 0 & K & 0 \end{bmatrix},$$

unde  $h_i$  sunt coeficienții polinomului caracteristic.

Dacă notăm cu  $S(i)$  vectorul coloană, având  $k$  elemente din  $GF(2)$  care dau starea celulelor registrului la momentul  $i$ :

$$S(i) = \begin{pmatrix} a_0^i \\ a_1^i \\ M \\ a_{k-1}^i \end{pmatrix},$$

unde cu  $a_j^i$  am notat starea celulei  $j$  la momentul  $i$ .

Se poate ușor observa că:

$$a_0^{i+1} = h_1 a_0^i + h_2 a_1^i + K + h_k a_{k-1}^i$$

$$a_1^{i+1} = a_0^i$$

$M$

$$a_{k-1}^{i+1} = a_{k-2}^i$$

Deci starea registrului la momentul  $i + 1$  va fi:

$$S(i+1) = M^T \cdot S(i)$$

și

$$S(i+j) = (M^T)^j \cdot S(i).$$

Dacă se cunosc  $2k$  biți ai succesiunii de ieșire se poate stabili atât conținutul inițial al registrului, cât și conexiunile la sumatoare prin rezolvarea unui sistem de  $2k$  ecuații liniare, sau matriceal.

Dacă notăm cu  $S(1)$  un vector coloană format din primii  $k$  biți ai șirului de ieșire,  $S(2)$  un alt vector coloană format din  $k$  biți începând cu al doilea etc., se pot forma două matrice pătrate de dimensiune  $k \times k$ :

$$X(1) = [S(1), S(2), \dots, S(k)];$$

$$X(2) = [S(2), S(3), \dots, S(2k)].$$

Între cele două matrice există relația:

$$X(2) = T \cdot X(1).$$

Matricea  $T$  este matrice de trecere și are forma:

$$T = \begin{pmatrix} 0 & 1 & 0 & K & 0 \\ 0 & 0 & 1 & K & 0 \\ K & K & K & K & K \\ 0 & 0 & 0 & K & 1 \\ h_1 & h_2 & h_3 & K & h_k \end{pmatrix}.$$

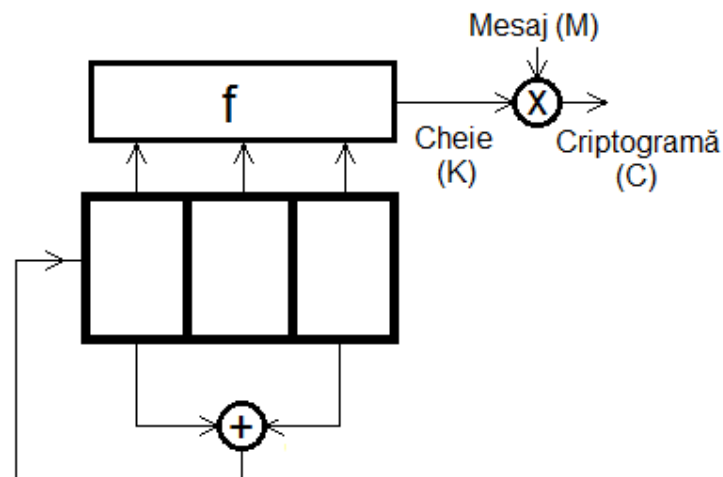
Se observă cu ușurință legătura dintre matricele  $M$  și  $T$ .

Deci  $T$  se poate calcula cu relația:

$$T = X(2) \cdot (X(1))^{-1}.$$

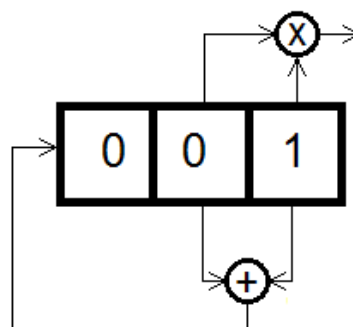
Inversarea matricei  $X(1)$  necesită cel mult operații de ordinul  $k^3$  și se pot efectua ușor cu ajutorul calculatorului pentru orice  $k$ .

Pentru funcționarea mai sigură a schemelor secvențiale, registrul de deplasare se completează cu o logică neliniară. În figura 3.3 se prezintă un registru de deplasare cu reacție pentru generarea unei succesiuni, care apoi este filtrată neliniar cu o funcție  $f$  pentru a realiza o distribuție uniformă între biții de 0 și cei de 1, secvență care se poate utiliza drept cheie pentru cifrarea mesajelor.



**Fig. 3.3**

Există o clasă de generatori neliniari realizată pe bază de polinoame caracteristice primitive și cu logică neliniară la două ranguri pentru a asigura ieșirea. Schema unui astfel de generator neliniar este dată mai jos:



**Fig. 3.4**

Succesiunea la ieșire poate fi privită ca o soluție generală a ecuației caracteristice a unui generator liniar echivalent, determinat de polinomul  $X^6 + X^5 + X^4 + X^3 + X^2 + X + 1$ .

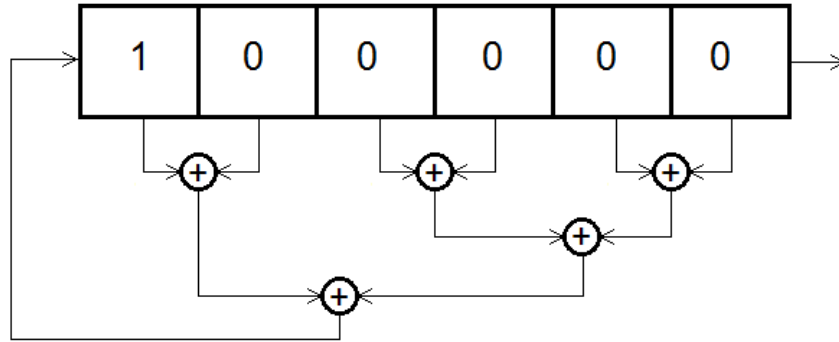


Fig. 3.5

Analiza se poate extinde la generatoare având mai mult de două ranguri combinate neliniar, așa cum se arată în figurile de mai jos.

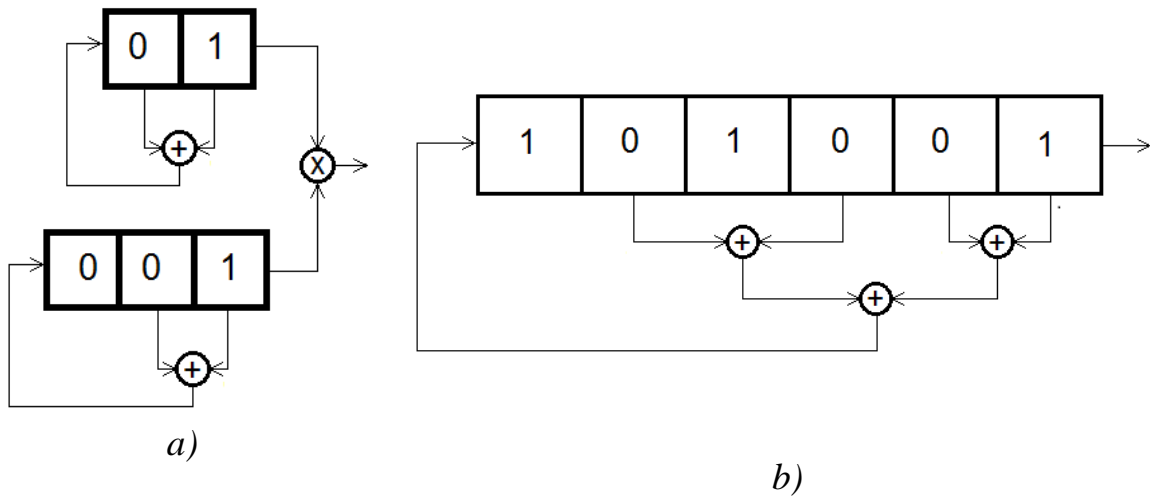


Fig. 3.6

Schema din dreapta reprezintă echivalentul liniar al schemei din stânga.

Procesul neliniar analizat poate fi extins la sinteza unor generatoare mai complexe. Un astfel de generator este reprezentat mai jos.

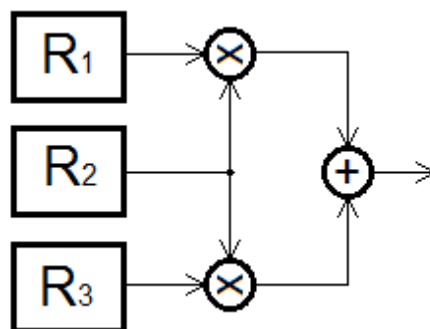
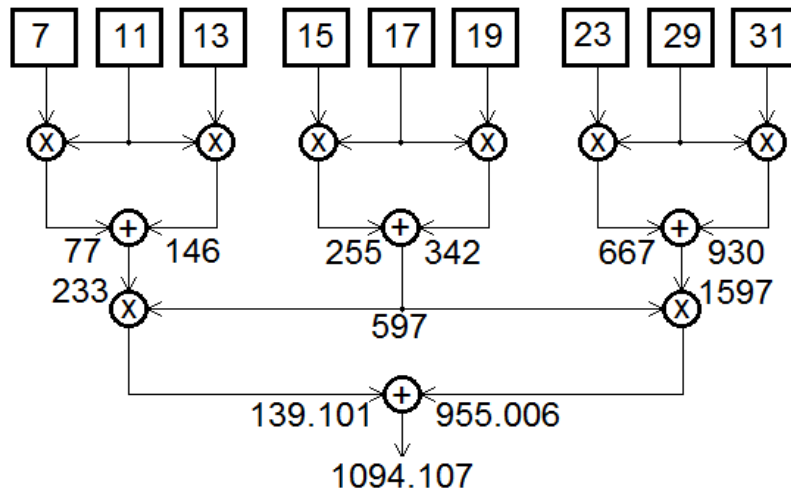


Fig. 3.7

Dacă cele trei registre  $R_1, R_2$  și  $R_3$  au polinoame caracteristice primitive de grade respectiv  $r, s$  și  $t$ , atunci se demonstrează că generatorul rezultat va avea complexitatea echivalentă cu:  $rs + (s + 1)t$ . Perioada de repetiție va fi egală cu cel mai mic multiplu comun al numerelor:  $2^r - 1; 2^s - 1; 2^t - 1$ .

Sucesiunea obținută poate fi combinată cu alte succesiuni generate în mod analog într-un dispozitiv asemănător celui de mai jos.

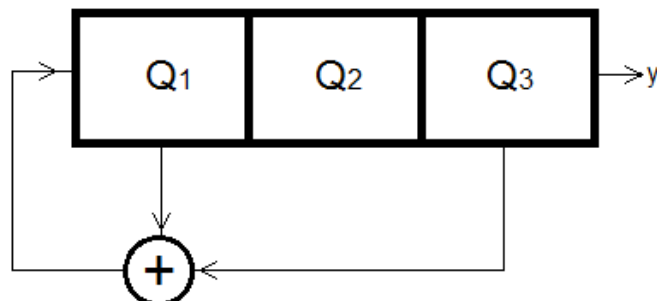


**Fig. 3.8**

La aceste scheme se poate încerca maximizarea complexității succesiunii în funcție de numărul total de ranguri conținute în registre. Așadar, combinând registrele de deplasare cu reacție liniară cu un număr mic de ranguri cu ajutorul operațiilor neliniare, se pot obține generatoare foarte complexe cu perioade mari de repetiție, care pot fi folosite cu succes în criptografie.

### *Aplicație*

Să analizăm în continuare generatorul din figura 3.9.



**Fig. 3.9**

Dacă se încarcă inițial registrul cu starea: 1,0,0, atunci tabela de adevăr care conține funcționarea acestuia este:



$Q_1$	$Q_2$	$Q_3$	$y$
1	0	0	0
1	1	0	0
1	1	1	1
0	1	1	1
1	0	1	1
0	1	0	0
0	0	1	1

Se definește funcția de autocorelație  $R(i)$ , astfel:

$$R(i) = \sum_{k=1}^{2^n-1} Q(k) \cdot Q(k+i) = \begin{cases} 2^n - 1, & \text{pentru } i = 0 \\ 0, & \text{în rest} \end{cases}$$

$R(i)$  este o funcție periodică cu perioada  $2^n - 1$ , deci  $R(i) = R(i + 2^n - 1)$ .

$Q(k + i)$  reprezintă decalarea ciclică cu  $i$  biți a secvenței obținute la ieșirea registrului.

Dacă îi asociem lui 0 valoarea numerică  $-1$ , iar lui 1 valoarea numerică 1, atunci pentru ieșirea:  $Y = 0011101$  se obține:

$Q(k)$ :        -1 -1 1 1 1 -1 1

$Q(k + 0)$ :   -1 -1 1 1 1 -1 1

$Q(k + 1)$ :    1 -1 -1 1 1 1 -1

$Q(k + 2)$ :   -1 1 -1 -1 1 1 1

$Q(k + 3)$ :    1 -1 1 -1 -1 1 1

$Q(k + 4)$ :    1 1 -1 1 -1 -1 1

$Q(k + 5)$ :    1 1 1 -1 1 -1 -1

$Q(k + 6)$ :   -1 1 1 1 -1 1 -1

$R(0) = (-1)(-1) + (-1)(-1) + 1 + 1 + 1 + (-1)(-1) + 1 = 7$

$R(1) = (-1) + (-1)(-1) + (-1) + 1 + 1 + (-1) + (-1) = -1$

În mod analog:  $R(2) = -1$

$R(3) = -1$

$R(4) = -1$

$R(5) = -1$

$R(6) = -1$

Existența maximului pronunțat în graficul funcției de autocorelație face posibilă utilizarea acestei secvențe pseudoaleatoare ca secvență de sincronizare în transmisiunile de date. Dacă se intercalează în mod periodic la transmitere o astfel de succesiune și se face corelația dintre secvența recepționată cu o secvență generată local, atunci la ieșirea corelatorului va apărea un maxim ori de câte ori va apărea secvența respectivă. Avantajul utilizării acestei secvențe într-o astfel de aplicație constă în faptul că sincronizarea se poate face chiar dacă „ $e$ ” biți ai acestuia au fost eronați, cu condiția ca  $e < \frac{2^n - 1}{2}$  (deci  $e < 3$ ).

În registrele din schemele de generare a succesiunilor trebuie evitată situația în care toate celulele sunt în starea 0, deoarece nu se mai poate ieși din această stare. Un exemplu de evitare a acestei stări este dat în schema de mai jos.

Presupunând că se cunosc primii 6 biți ai succesiunii, să determinăm acum structura și starea inițială a registrului. Pe baza celor 6 biți cunoscuți: 001110 se construiesc matricele  $X(1)$  și  $X(2)$

$$X(1) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad X(2) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Din relația:  $X(2) = T \cdot X(1)$  se obține  $T = X(2) \cdot X(1)^{-1}$ .

Dar

$$X(1)^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

deci:

$$T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Rezultă că  $h_1 = 1; h_2 = 0; h_3 = 1$ .

Primii trei biți ai șirului ne arată starea inițială a celor trei celule:

$$Q_3 = 0; Q_2 = 0 \text{ și } Q_1 = 1.$$

### Test de autoevaluare

1. Ce reprezintă numerele aleatoare și succesiunile pseudoaleatoare?
2. De ce sunt folosite testele de aleatorism și care sunt tipurile de teste?
3. Care este utilizarea numerelor aleatoare?

### Tema de autoinstruire

1. Propuneți o schemă logică pentru aplicația prezentată la generarea succesiunilor pseudoaleatoare.

### Lucrare de verificare

1. Definiți metoda congruențial-liniară de generare a succesiunilor pseudoaleatoare.
2. Explicați metoda Von Neumann de generare algoritmică a numerelor cu proprietăți aleatoare.
3. Definiți metoda congruențială aditivă de generare a succesiunilor pseudoaleatoare.
4. Care sunt criteriile de aleatorism ale lui Golomb pentru secvențele binare?
5. Explicați metoda de testare a unui șir pseudoaleator prin aplicarea testului de frecvență.
6. Explicați metoda de testare a unui șir pseudoaleator prin aplicarea testului serial.
7. Explicați metoda de testare a unui șir pseudoaleator prin aplicarea testului poker.
8. Utilizând polinomul  $f = x^5 + x^2 + 1$  cu coeficienți în GF2, să se genereze o secvență pseudoaleatoare plecând din starea 00001.

## CAPITOLUL 4

### METODE DE CIFRARE

*Cuvinte cheie: metode de cifrare, polinoame primitive, polinoame de permutare.*

#### OBIECTIVE

Înțelegerea și cunoașterea următoarelor noțiuni:

- câmpuri Galois, funcții de permutare;
- metode de cifrare, exemple de cifruri;
- simularea și implementarea unor exemple practice de cifrare.

#### 4.1 Corpuri Galois. Calcule

##### 4.1.1 Câmpuri Galois

Fie câmpul  $GF(q)$  în care  $q = p^n$  și care se notează cu  $F$ . Elementele  $x \in F$  pot fi reprezentate printr-un set de  $q$  polinoame:

$$x = C_0 t^{n-1} + C_1 t^{n-2} + \dots + C_{n-2} t + C_{n-1},$$

unde  $0 \leq c_i < p$ , iar  $t$  este un element primitiv al câmpului  $F$ , adică orice  $x \neq 0$  poate fi exprimat prin  $x = t^k$ ,  $1 \leq k \leq q-1$ .

În plus,  $t^{q-1} = 1$ , unde  $q-1$  este cel mai mic întreg pozitiv cu această proprietate. Deci se poate scrie:  $F = \{0, t, t^2, \dots, t^{q-1}\}$ .

Elementul  $t$  satisface o ecuație ireductibilă de forma:

$$f(t) = t^n + k_1 t^{n-1} + k_2 t^{n-2} + \dots + k_{n-1} t + k_n = 0,$$

unde  $0 \leq k_i < p$ . Polinomul  $f(t)$  se numește polinom primitiv peste câmpul  $F$ , iar ecuația  $f(t) = 0$  se poate folosi pentru a converti orice formă de puteri

$(x = t^k)$  într-o formă polinomială  $\left( x = \sum_{i=0}^{n-1} c_i t^{n-1-i} \right)$ , exprimând  $t^n = -\sum_{i=1}^n k_i t^{n-i}$ .

De asemenea, fiecărui element  $x \in F$  i se poate atribui o valoare numerică cu ajutorul relației:

$$f = \sum_{i=0}^{n-1} C_i p^{n-1-i}, \quad 0 \leq f \leq q-1.$$

Dacă

$$x = (C_0, C_1, \dots, C_{n-1}), \text{ iar } y = (d_0, d_1, \dots, d_{n-1}),$$

atunci

$$x + y = (C_0 + d_0, C_1 + d_1, \dots, C_{n-1} + d_{n-1}) \pmod{p}.$$

Dacă  $x = t^k$ , iar  $q = t^m$ , atunci  $xy = t^{k+m}$ , unde  $k + m$  se calculează modulo  $(q - 1)$ .

Așadar orice element al câmpului  $F$  poate fi exprimat în două moduri:

- prin componentele sale  $(C_0, C_1, \dots, C_{n-1})$ ;
- sub formă de puteri  $(t^k)$ , în afara elementului  $x = 0$ .

Vom ilustra remarcile de mai sus, considerând câmpul  $GF(3^3)$ , căruia îi asociem polinomul primitiv  $f(t) = t^3 + 2t + 1$ .

Elementele câmpului vor fi:  $F = (0, t, t^2, \dots, t^{26})$

$$0 = (0, 0, 0) \quad \text{cu} \quad f_0 = 0$$

$$t = (0, 1, 0) \quad f_1 = 3$$

$$t^2 = (1, 0, 0) \quad f_2 = 9$$

$$t^3 = (0, 1, 2)$$

se obține cu ajutorul polinomului primitiv

$$f(t), \quad t^3 + 2t + 1 = 0.$$

Calculul acesta se bazează pe faptul că valorile termenilor vor fi întotdeauna pozitive și vor reprezenta valorile calculate  $\text{mod}(p)$ , adică

pentru  $-2 \Rightarrow +1$ , iar pentru  $-1 \Rightarrow +2$ . Pentru valoarea  $t=3$ , rezultă  $t^3 = -2t - 1 = t + 2 = 3 + 2 = 5$ , adică  $f_3 = 5$ .

Valorile următoare se calculează pornind de la această relație

$$\begin{aligned}
 t^4 &= (1, 2, 0) & f_4 &= 15 \\
 t^5 &= (2, 1, 2) & f_5 &= 23 \\
 t^6 &= (1, 1, 1) & f_6 &= 13 \\
 t^7 &= (1, 2, 2) & f_7 &= 17 \\
 t^8 &= (2, 0, 2) & f_8 &= 20 \\
 t^9 &= (0, 1, 1) & f_9 &= 4 \\
 t^{10} &= (1, 1, 0) & f_{10} &= 12 \\
 t^{11} &= (1, 1, 2) & f_{11} &= 14 \\
 t^{12} &= (1, 0, 2) & f_{12} &= 11 \\
 t^{13} &= (0, 0, 2) & f_{13} &= 2 \\
 t^{14} &= (0, 2, 0) & f_{14} &= 6 \\
 t^{15} &= (2, 0, 0) & f_{15} &= 18 \\
 t^{16} &= (0, 2, 1) & f_{16} &= 7 \\
 t^{17} &= (2, 1, 0) & f_{17} &= 21 \\
 t^{18} &= (1, 2, 1) & f_{18} &= 16 \\
 t^{19} &= (2, 2, 2) & f_{19} &= 26 \\
 t^{20} &= (2, 1, 1) & f_{20} &= 22 \\
 t^{21} &= (1, 0, 1) & f_{21} &= 10 \\
 t^{22} &= (0, 2, 2) & f_{22} &= 8 \\
 t^{23} &= (2, 2, 0) & f_{23} &= 24 \\
 t^{24} &= (2, 2, 1) & f_{24} &= 25 \\
 t^{25} &= (2, 0, 1) & f_{25} &= 19 \\
 t^{26} &= (0, 0, 1) & f_{26} &= 1
 \end{aligned}$$

Se observă că fiecare element al câmpului are două reprezentări și i se asociază o valoare numerică. Pentru efectuarea operațiilor de înmulțire și împărțire se folosește reprezentarea exponențială, iar pentru operațiile de adunare și scădere, reprezentarea polinomială.

De exemplu:

$$t^5 \cdot t^{25} = t^{30} = t^4 = (1, 2, 0)$$

$$t^5 + t^{25} = (2, 1, 2) + (2, 0, 1) = (1, 1, 0) = t^{10}$$

$$t^3 : t^{14} = t^{29} : t^{14} = t^{15} = (2, 0, 0)$$

## 4.2 Funcții polinomiale

Un polinom  $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$  cu coeficienți  $a_i \in GF(q)$  este numit polinom de permutare dacă pentru cele  $q$  elemente ale câmpului  $(x_1, x_2, \dots, x_q)$  imaginile  $(f(x_1), f(x_2), \dots, f(x_q))$  reprezintă o permutare a elementelor câmpului.

Fie, de exemplu, elementele câmpului  $GF(7)$  și polinomul  $f(x) = 2x + 3$ . Acest polinom realizează următoarea permutare:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 0 & 2 & 4 & 6 & 1 \end{pmatrix}$$

Aceste funcții de permutare pot fi utilizate în criptografie dacă literelor unui alfabet  $L$  li se asociază valori numerice corespunzătoare, iar cifrarea are loc după următoarea schemă:

$$\alpha \rightarrow x \rightarrow f(x) \rightarrow \alpha'.$$

Deci unei litere  $\alpha$  i se asociază valoarea numerică  $x$ . Se calculează  $f(x)$  care corespunde literei  $\alpha'$ .

Această procedură realizează o substituție simplă. În scopul de a da o complexitate mai mare se folosește o permutare polinomială cu mai mulți parametri care să fie modificați după o anumită regulă.

De exemplu, polinomul:  $f(x) = ax + b$ ,  $a \neq 0$ ,  $a, b \in GF(q)$  asigură o substituție simplă  $a$  și  $b$  putând fi considerate ca elemente ale cheii de cifrare.

În general dacă permutarea este reprezentată prin:  $y_i = f(x_i, a_1, a_2, \dots, a_r)$ ,  $a_i \in GF(q)$ , unde  $a_1, a_2, \dots, a_r$  pot fi supuși unor anumite restricții impuse de

funcția utilizată, atunci descifrarea va impune utilizarea funcției de permutare inverse  $f^{-1}(x)$ .

Deci  $x_i = f^{-1}(y_i, a_1, a_2, \dots, a_r)$ .

### 4.3 Metode de cifrare bazate pe funcții de permutare

Vom prezenta mai multe metode bazate pe funcțiile de permutare care pot fi aplicate în raport cu numărul de litere al alfabetului  $N$  și de câmpul asociat.

#### 4.3.1 Metoda utilizată în cazul $N=P^n$

Vom ilustra această metodă printr-un exemplu.

Fie câmpul  $GF(3^3)$ , deci  $N=3^3=27$  litere (cele 26 de litere ale alfabetului plus cuvântul - spațiu), și fie polinomul de permutare:  $y_i = a_i x_i + b_i$ , cu  $a_i \neq 0$ . Între parametrii  $a_i$  și  $b_i$  vom stabili o anumită regulă de recurență, de exemplu:

$$\begin{aligned} a_{i+2} &= a_{i+1} + a_i; & i &= 1, 2, 3, \dots \\ b_{i+2} &= b_{i+1} \cdot b_i; & i &= 1, 2, 3, \dots \end{aligned}$$

și valorile inițiale:

$$\begin{aligned} a_1 &= (0, 2, 1) = t^{16} \\ a_2 &= (1, 1, 1) = t^6 \\ b_1 &= (0, 0, 2) = t^{13} \\ b_2 &= (1, 1, 0) = t^{10} \end{aligned}$$

Cifrarea cuvântului SECRET poate fi realizată astfel:

(a)	S	E	C	R	E	T
(b)	19	5	3	18	5	20
(c)	(2,0,1)	(0,1,2)	(0,1,0)	(2,0,0)	(0,1,2)	(2,0,2)
(d)	25	3	1	15	25	8
(e)	(0,2,1)	(1,1,1)	(1,0,2)	(2,1,0)	(0,1,2)	(2,2,2)
(f)	16	6	12	17	3	19
(g)	(0,0,2)	(1,1,0)	(2,2,0)	(1,2,2)	(1,2,0)	(1,1,2)
(h)	20	16	26	18	24	17
(i)	T	P	Z	R	S	Q



- (a) textul în clar  $(x_i)$ ;
- (b) valorile numerice asociate literelor;
- (c) reprezentare polinomială;
- (d) reprezentare exponențială;
- (e)  $a_i$  reprezentate polinomial;
- (f)  $a_i$  reprezentate ca puteri;
- (g)  $b_i$  reprezentate polinomial;
- (h) valori numerice obținute pentru reprezentările cifrate;
- (i) criptograma  $(y_i)$ .

Pentru descifrare se folosește funcția inversă:  $x_i = a_i^{-1}(y_i - b_i)$ .

### 4.3.2 Metoda utilizată în cazul $M=p^n+1$

Carmichael a prezentat o metodă de cifrare bazată pe utilizarea funcțiilor de permutare, introducând simbolul  $\infty$  la elementele câmpului  $GF(p^n)$ , metodă care poate fi utilizată atunci când numărul literelor din alfabet poate fi scris sub forma:  $M = p^n + 1$ .

Vom ilustra această metodă utilizând ca funcție de permutare o funcție rațională  $R(x)$ :

$$R(x) = y_i = \frac{a_i x_i + b_i}{c_i x_i + d_i}, a_i, b_i, c_i \in GF(p^n) \text{ și } a_i d_i - b_i c_i \neq 0.$$

Simbolul  $\infty$  este reprezentat sub forma  $\frac{x}{0}$  ( $x \neq 0$ ), unde  $x \in GF(p^n)$ .

Astfel  $R\left(-\frac{d_i}{c_i}\right) = \infty$ ,  $c_i \neq 0$  și  $R(\infty) = \frac{a_i}{c_i}$ , iar dacă  $c_i = 0$ , atunci  $R(\infty) = \infty$ .

Vom utiliza elementele câmpului  $GF(5^2)$  cu polinomul primitiv  $f(t) = t^2 + 4t + 2$ . Alfabetul este cel al limbii engleze cu  $26 = 5^2 + 1$  litere, iar corespondența dintre litere și elementele câmpului  $GF(5^2)$  este dată mai jos:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
(a)	0	1	2	3	4	5	6	7	8	9	10	11	12	13
(b)	(0,0)	(1,0)	(1,3)	(4,3)	(2,2)	(4,1)	(0,2)	(2,0)	(2,1)	(3,1)	(4,4)	(3,2)	(0,4)	(4,0)
(c)		(1,2)	(2,3)											

- (d) (0,1) (3,3)  
 (e) (3,4) (1,2)  
 (f) (2,0) (1,0)

	O	P	Q	R	S	T	U	V	W	X	Y	Z
(a)	14	15	16	17	18	19	20	21	22	23	24	
(b)	(4,2)	(1,2)	(3,3)	(1,4)	(0,3)	(3,0)	(3,4)	(2,4)	(1,1)	(2,3)	(0,1)	
(c)	...											
(d)	...											
(e)	...											
(f)	...											

Linia (a) cuprinde reprezentarea exponențială, iar linia (b), reprezentarea polinomială a elementelor de câmp, iar (c), (d), (e) și (f) reprezintă parametrii  $a_i, b_i, c_i$  și respectiv  $d_i$ .

Linia (a) folosește valorile obținute din:

$$t^2 = -4t - 2 = t + 3$$

$$t^3 = t * t^2 = t * (t + 3) = t^2 + 3t = t + 3 + 3t = 4t + 3$$

$$t^4 = t * t^3 = t * (4t + 3) = 4t^2 + 3t = 7t + 12 = 2t + 2$$

.....

Vom folosi ca regulă de variație a parametrilor  $a_i, b_i, c_i$  și  $d_i$  selectând două matrice pătrate de ordinul 2 cu elemente ale câmpului  $GF(5^2)$  și stabilind:

$$M_{i+2} = M_i \times M_{i+1}, \text{ deci } M_i \neq 0 \text{ și } M_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$$

Fie de exemplu  $M_1 = \begin{bmatrix} (1,2) & (0,1) \\ (3,4) & (2,0) \end{bmatrix}; M_2 = \begin{bmatrix} (2,3) & (3,3) \\ (1,2) & (1,0) \end{bmatrix}$  și folosind

regula stabilită calculăm  $M_3$ :

$$M_3 = \begin{bmatrix} (0,4) & (3,0) \\ (4,1) & (2,0) \end{bmatrix}$$

$$a_3 = (1,2) \cdot (2,3) + (0,1) \cdot (1,2) = t^{15} \cdot t^{23} + t^{24} \cdot t^{15} = t^{14} + t^{15} = (4,2) + (1,2) = (0,4)$$

$$b_3 = (1,2) \cdot (3,3) + (0,1) \cdot (1,0) = t^{15} \cdot t^{16} + t^{24} \cdot t = t^7 + t = (2,0) + (1,0) = (3,0)$$

$$c_3 = (3,4) \cdot (2,3) + (2,0) \cdot (1,2) = t^{20} \cdot t^{23} + t^7 \cdot t^{15} = t^{19} + t^{22} = (3,0) + (1,1) = (4,1)$$

$$d_3 = (3,4) \cdot (3,3) + (2,0) \cdot (1,0) = t^{20} \cdot t^{16} + t^7 \cdot t = t^{12} + t^8 = (0,4) + (2,1) = (2,0)$$

Pentru fiecare matrice calculată trebuie verificată condiția de nenulitate.

Se observă că:

$$a_1d_1 - b_1c_1 = (3,2) \neq 0$$

$$a_2d_2 - b_2c_2 = (3,1) \neq 0$$

$$a_3d_3 - b_3c_3 = (3,4) \neq 0$$

Cifrarea cuvântului COD se va face astfel:

	C	O	D
(a)	2	14	3
(b)	(1,3)	(4,2)	(4,3)
(c)	(3,0)	(2,3)	(3,1)
(d)	Q	X	J

$$Y_1 = \frac{a_1x_1 + b_1}{c_1x_1 + d_1} = \frac{t^2 \cdot t^{15} + (0,1)}{t^{20} \cdot t^{15} + (2,0)} = \frac{(1,4) + (0,1)}{(3,2) + (2,0)} = \frac{(1,0)}{(0,2)} = \frac{t}{t^6} = t^{19} = (3,0)$$

$$Y_2 = \frac{a_2x_2 + b_2}{c_2x_2 + d_2} = \frac{t^{23} \cdot t^{14} + (3,3)}{t^{15} \cdot t^{14} + (1,0)} = \frac{(4,0) + (3,3)}{(4,1) + (1,0)} = \frac{(2,3)}{(0,1)} = t^{23} = (2,3)$$

$$Y_3 = \frac{a_3x_3 + b_3}{c_3x_3 + d_3} = \frac{t^{12} \cdot t^3 + (3,0)}{t^5 \cdot t^3 + (2,0)} = \frac{(1,2) + (3,0)}{(2,1) + (2,0)} = \frac{(4,2)}{(4,1)} = \frac{t^{14}}{t^5} = t^9 = (3,1)$$

Descifrarea se face folosind transformarea inversă:

$$x_i = -\frac{d_i y_i - b_i}{c_i y_i - a_i}.$$

### **4.3.3 Metoda de cifrare a mai multor câmpuri și funcții de permutare**

Metoda care va fi prezentată în continuare este echivalentă cu cifrarea independentă și simultană a mai multor mesaje, fiecare cifrat printr-o metodă diferită dintre tipurile prezentate anterior.

Fie un alfabet  $L$ , având  $l$  litere. Dacă se vor cifra câte  $n$  litere simultan, cifrare poligrafică, se poate spune că avem de-a face cu un alt alfabet  $A$  care are  $N = l^n$  caractere ( $n$ -grame) de forma  $\lambda = \overline{l_1 l_2 \dots l_n}$ ,  $l_i \in L$ .

La cifrare, textul clar compus din literele alfabetului  $L$  se împarte în blocuri de lungime  $n$ . Numărul  $N$  se poate scrie ca o sumă de  $k$  numere, astfel:

$$N = N_1 + N_2 + \dots + N_k,$$

în care  $M_i = P_i^{m_i}$  sau  $N_i = P_i^{m_i} + 1$ .

Alfabetul  $A$  este partiționat în  $K$  subalfabeturi, astfel încât  $A = \bigcup_{i=1}^k A_i$ , subalfabetele fiind disjuncte două câte două. Fiecărui subalfabet îi este destinat un câmp  $GF(P_i^{m_i}) = F_i$  și o funcție de permutare  $f_i(x)$ , care poate fi polinomială dacă  $N_i = P_i^{m_i}$  sau rațională dacă  $M_i = P_i^{m_i} + 1$ .

Modul în care cele  $N$   $n$ -grame sunt împărțite în cele  $k$  subalfabete este următorul:

- se repartizează fiecărei litere ale alfabetului inițial  $L$  o valoare numerică din secvență:  $0, 1, 2, \dots, (l-1)$ . Această repartiție poate fi diferită pentru mesaje diferite;
- fiecare  $n$ -gramă se scrie sub formă numerică  $\lambda = (v_1, v_2, \dots, v_n)$ , unde  $v_i$  este valoarea numerică a literei  $l_i$  din  $n$ -gramă;
- fiecărei  $n$ -grame  $i$  se atașează un număr  $\sigma = v_1 l^{n-1} + v_2 l^{n-2} + \dots + v_{n-1} l + v_n$  care reprezintă valoarea numerică asociată acesteia;
- în subalfabetul  $A_1$  vor fi incluse toate  $n$ -gramele pentru care  $0 \leq \sigma \leq N_1 - 1$ ;
- în subalfabetul  $A_2$  vor fi incluse toate  $n$ -gramele pentru care  $N_1 \leq \sigma \leq M_1 + M_2 - 1$ ;
- în subalfabetul  $A_k$  vor fi incluse toate  $n$ -gramele pentru care  $N_1 + N_2 + \dots + N_{k-1} \leq \sigma < N - 1$ .

Se definește, în continuare, pentru fiecare  $n$ -gramă, numărul  $\rho$  astfel  $\rho = \sigma - (N_1 + N_2 + \dots + N_{i-1})$ , astfel încât  $0 \leq \rho \leq N_i - 1$ ,  $n$ -gramă fiind inclusă în subalfabetul  $A_i$ .

Se scrie  $\rho$  sub formă polinomială:

$$\begin{aligned} \rho &= C_0 \rho_i^{m_i-1} + C_1 \rho_i^{m_i-2} + \dots + C_{m_i-2} \rho_i + C_{m_i-1} \\ \rho &= (C_0, C_1, C_2, \dots, C_{m_i-1}) \end{aligned}$$

Procedura de mai sus definește o bijecție între  $n$ -gramele fiecărui subalfabet  $A_i$  și elementele câmpului  $F_i$ . Cu ajutorul funcțiilor  $f_i(x)$  se face apoi cifrarea după metodele descrise anterior.

Vom ilustra această procedură printr-un exemplu. Să considerăm cazul  $n = 2$  și  $l = 26$ , iar repartizarea valorilor numerice pentru literele alfabetului normală:

$A = 0$	$H = 7$	$O = 14$	$V = 21$
$B = 1$	$I = 8$	$P = 15$	$W = 22$
$C = 2$	$J = 9$	$Q = 16$	$X = 23$
$D = 3$	$K = 10$	$R = 17$	$Y = 24$
$E = 4$	$L = 11$	$S = 18$	$Z = 25$
$F = 5$	$M = 12$	$T = 19$	
$G = 6$	$N = 13$	$U = 20$	

Numărul de bi-grame este  $N = 26^2 = 676$ .

Acesta se poate scrie astfel:

$$26^2 = 7^3 + 13^2 + (3^4 + 1) + (3^4 + 1)$$

Deci:

$$\begin{aligned} N_1 &= 7^3 = 343 & N_1 + N_2 &= 512 \\ N_2 &= 13^2 = 169 & N_1 + N_2 + N_3 &= 594 \\ N_3 &= 3^4 + 1 = 82 \\ N_4 &= 3^4 + 1 = 82 \end{aligned}$$

Vom avea deci pentru subalfabete  $A_1, A_2, A_3$  și  $A_4$ .

O bigramă  $\lambda$  va aparține:

- subalfabetului  $A_1$  dacă  $0 \leq \sigma \leq 342$ ;
- subalfabetului  $A_2$  dacă  $343 \leq \sigma \leq 511$ ;
- subalfabetului  $A_3$  dacă  $512 \leq \sigma \leq 593$ ;
- subalfabetului  $A_4$  dacă  $594 \leq \sigma \leq 675$ .

Fie diagrama:  $\lambda = RM = (17, 12)$ .

Deci  $\sigma = 17 \times 26 + 12 = 454$  și de aici rezultă că această diagramă aparține subalfabetului  $A_2$  ( $RM \in A_2$ ). Se calculează  $\rho$ :

$$\rho = \sigma - N_1 = 454 - 343 = 111$$

Subalfabetului  $A_2$  îi este asociat câmpul  $GF(13^2)$  și valorii 111 îi corespunde elementul (8,7) din  $GF(13^2)$ , deoarece  $(8 \times 13 + 7 = 111)$ .

Câmpului  $GF(7^3)$  corespunzător primului subalfabet îi atașăm polinomul primitiv  $f(t) = t^3 - t - 5$ . Celui de-al doilea subalfabet îi asociem câmpul  $GF(13^2)$  și polinomul primitiv  $f(t) = t^2 + t + 2$ , celui de-al treilea subalfabet îi asociem câmpul  $GF(13^2)$  și polinomul primitiv  $f(t) = t^2 + t + 2$ , la fel și celui de-al patrulea subalfabet ( $A_4$ ).

Cele patru funcții de permutare utilizate vor fi:

- pentru  $A_1$   $y_i = x_i^5 + a_i$ ;
- pentru  $A_2$   $y_i = x_i^5 + b_i$ ;
- pentru  $A_3$   $y_i = c_i + \frac{1}{x_i^3 + 1}$ ;
- pentru  $A_4$   $y_i = d_i + \frac{1}{x_i^3 + 1}$ .

Modul de variație a parametrilor este următorul:

$$\begin{aligned} a_{i+1} &= (1,0,2)a_i + (2,3,6); & a_1 &= (2,4,5) \\ b_{i+1} &= (7,8)b_i; & b_1 &= (7,10) \\ c_{i+1} &= (1,0,0,0)c_i; & c_1 &= (1,1,0,1) \\ d_{i+1} &= (1,0,0,0)d_i; & d_1 &= (1,2,2,0) \end{aligned}$$

În cazul nostru:  $x_1 = (8,7)$ .

Se calculează  $y_1 = x_1^5 + b_1 = (8,7)^5 + (7,10)$ .

Pentru descifrare se vor aplica funcțiile inverse care pentru  $A_2$  este:

$$\begin{aligned} x_i^5 &= y_i - b_i \\ x_i &= \sqrt[5]{y_i - b_i} \end{aligned}$$

Dacă  $n=3$ , atunci vom avea  $N = 26^3 = 17576$ , care se poate scrie:  $N = 26^3 = 7^5 + 3^6 + 3^3 + 13$  și se vor putea cifra trigrame.

#### 4.3.4 Metoda de cifrare prin folosirea unui singur câmp

Vom prezenta o metodă de cifrare în care numărul de litere ale alfabetului  $N$  este mai mic decât elementele câmpului asociat.

Fie  $A = (\alpha_1, \alpha_2, \dots, \alpha_N)$  alfabetul utilizat și  $F = GF(p^n)$  câmpul asociat,  $(8^n > N)$ . Se face partiționarea elementelor câmpului în două seturi:

$$S = (x_1, x_2, \dots, x_N)$$

și

$$S' = F - S = (x'_1, x'_2, \dots, x'_{q-1}); \quad q = p^n.$$

Se stabilește o bijecție între cele  $N$  litere ale alfabetului  $A$  și cele  $N$  elemente ale câmpului  $S$ .

Folosind o funcție polinomială de permutare de forma  $y_i = f(x_i, a_i)$ , să cifrăm textul clar:  $P_1 P_2 P_3 \dots$ . Se stabilește bijecția între literele textului clar și valorile numerice  $\lambda_1, \lambda_2, \lambda_3, \dots$ , unde  $\lambda_i \in S$  și se evaluează apoi  $f(\lambda_i a_i) = g_i$ .

Acum dacă  $g_i \in S$  și  $g_i$  corespunde literei  $C_i$ , atunci  $P_i$  este cifrat în  $C_i$ . Dacă  $g_i \in S'$ , se repetă procedura. Deoarece  $S'$  este finit, și dacă prin alegerea făcută se evită situația:

$$f(x_i, a_i) = x_i,$$

atunci după cel mult  $(q - N)$  pași se va obține un element din  $S$ .

Pentru un parametru  $\alpha_i$  fixat, dacă  $x_j \neq x_k$ , se va obține  $y_i \neq y_k$ , însă prin variația parametrilor  $\alpha_i$  este posibil ca pentru  $x_i \neq x_k$  să se obțină  $y_i = y_k$ .

Vom ilustra această procedură prin două exemple.

**Exemplul 1:** Fie câmpul  $GF(u_1)$  și polinomul de permutare  $y_i = a_i x_i + b_i$ , unde  $a_i = 6^i \pmod{41}$ , iar  $b_{i+2} = b_{i+1} + b_i \pmod{41}$  și  $b_1 = 1$ , iar  $b_2 = 5$ .

În setul  $S$  vor intra elementele câmpului care au valori între 0 și 25, deci:

$$S = (0, 1, 2, \dots, 25),$$

iar

$$S = (0, 1, 2, \dots, 25), \quad S' = (26, 27, \dots, 40).$$

Să cifrăm textul PERMUTARE.

	P	E	R	M	U	T	A	R	E
(a)	15	4	17	12	20	19	0	17	4
(b)	6	36	11	25	27	39	29	10	19
(c)	1	5	6	11	17	28	4	32	36
(d)	9	15*	14*	24	24	1*	4	2*	13*
(e)	J	P	O	Y	Y	H	E	C	N

(a) valorile numerice asociate textului clar

(b)  $a_i = 6^i \quad i = 1, 2, 3, 4, 5, 6, 7, 8, 9$

(c)  $b_i \quad i = \overline{1, 9}$

(d) valorile numerice ale literelor criptogramei

(e) criptograma.

Valorile marcate cu " \* " s-au obținut după mai multe iterații.

Pentru descifrare se folosește funcția inversă:

$$v_i = 6^{40-i} (y_i - b_i),$$

aplicându-se aceleași iterații în cazul obținerii unor elemente din  $S'$ .

**Exemplul 2:** Fie un alfabet cu  $N = 1000$  caractere utilizate, format din 26 de litere,  $26 \times 26 = 676$  diagrame și 298 dintre cele mai frecvente trigrame:  $1000 = 26 + 26^2 \times 26 + 298$ . Cifrarea se face în câmpul  $GF(11^3)$  care conține 1331 elemente și care pot fi generate cu ajutorul polinomului primitiv:  $f(t) = t^3 + 10t + 4$ . Se utilizează funcția de generare polinomială:

$$y_i = a_i x_i^3 + b_i$$

Elementele câmpului  $GF(11^3)$  sunt partiționate în două seturi: setul  $S$  de 1000 de elemente și setul  $S'$  cu 331 elemente:

$$S = \{0, 1, 2, \dots, 999\} \text{ și } S' = \{1000, 1001, \dots, 1331\}$$

Se face corespondența între cele 1000 de caractere ale alfabetului și elementele setului  $S$ , sub forma:

$$A = 000$$

$$TE = 789$$

$$STR = 672 \text{ etc.}$$



Pentru cifrare se împarte textul în clar în secvențe de lungime 1,2 sau 3 elemente conținute în  $A$ ; fie acestea  $\lambda_1\lambda_2\lambda_3\dots$ .

Se înlocuiește secvența  $\lambda_i$  prin elemente de câmp  $x_i$  cu valoarea stabilită prin corespondență. Parametrii  $a_i$  și  $b_i$  pot fi variați în modul următor:

$$a_i = t^{3i+1}$$

$$b_i = t^{n^2+n+2}$$

iar reprezentările cifrate se obțin cu relația:  $y_i = a_i x_i^3 + b_i$  și dacă  $y_i \in S$  atunci el corespunde caracterului  $\lambda_k$ . Deci  $\lambda_i$  s-a cifrat prin  $\lambda_k$ .

Dacă însă  $y_i \in S'$ , atunci se repetă procedeul cu aceiași parametri până când  $y \in S$ .

Ca text cifrat poate fi folosit și cel obținut din înșiruirea valorilor numerice ale elementelor  $y_i \in S$  obținute în urma procesului iterativ de cifrare.

Pentru descifrare se utilizează funcția inversă:  $x_i = [(y_i - b_i)/a_i]^{\frac{1}{3}}$ , cu specificația că  $t^{\frac{1}{3}} = t^{887}$ .

### **Test de autoevaluare**

1. Ce reprezintă o metodă de cifrare și de câte tipuri sunt?
2. Care sunt diferențele dintre metodele de cifrare bazate pe funcții de permutare?
3. Care este utilizarea numerelor aleatoare?

### **Tema de autoinstruire**

1. Încercați crearea unei scheme logice care să algoritmeze metoda Carmichael, în speță exemplul atașat acesteia.

### **Lucrare de verificare**

1. Fie câmpul  $GF(26)$ , căruia îi asociem polinomul de permutare  $f(x) = 6x + 7$ .  
Să se cifreze următoarea expresie:

**CODE**

2. Fie câmpul  $GF(3^3)$  căruia îi asociem polinomul de permutare  $y_i = a_i * x_i + b_i$ .  
Între parametrii  $a_i$  și  $b_i$  există următoarele relații de recurență:

$$\begin{array}{lll} a_{i+2} = a_{i+1} + a_i & & a_1 = t^1 \\ b_{i+2} = b_{i+1} * b_i & \text{cu valorile inițiale} & a_2 = t^3 \\ i = 1, 2, 3, \dots & & b_1 = t^2 \\ & & b_2 = t^{15} \end{array}$$

Să se cifreze cuvântul înscris în tabelul următor, cu completarea acestuia în mod corespunzător.

(a)	<b>C</b>	<b>O</b>	<b>D</b>	<b>E</b>
(b)				
(c)				
(d)				
(e)				
(f)				
(g)				
(h)				
(i)				

unde:

- (a) textul în clar ( $x_i$ )
- (b) valorile numerice asociate literelor;
- (c) reprezentare polinomială;
- (d) reprezentare exponențială;
- (e)  $a_i$  reprezentate polinomial;
- (f)  $a_i$  reprezentate ca puteri;
- (g)  $b_i$  reprezentate polinomial;
- (h) valori numerice obținute pentru reprezentările cifrate;
- (i) criptograma  $y_i$

Nota 1: Valorile numerice asociate literelor alfabetului sunt prezentate în tabelul următor:

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
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	0

3. Fie câmpul  $GF(3^3)$  căruia îi asociem polinomul de permutare  $y_i = a_i * x_i + b_i$ . Între parametrii  $a_i$  și  $b_i$  există următoarele relații de recurență:

$$\begin{aligned} a_{i+2} &= a_{i+1} + a_i & a_1 &= t^1 \\ b_{i+2} &= b_{i+1} * b_i & a_2 &= t^3 \\ i &= 1, 2, 3, \dots & b_1 &= t^2 \\ & & b_2 &= t^{15} \end{aligned}$$

cu valorile inițiale

Să se cifreze cuvântul înscris în tabelul următor cu completarea acestuia în mod corespunzător

(a)	S	A	L	T
(b)				
(c)				
(d)				
(e)				
(f)				
(g)				
(h)				
(i)				

unde:

- (a) textul în clar ( $x_i$ )
- (b) valorile numerice asociate literelor;
- (c) reprezentare polinomială;
- (d) reprezentare exponențială;
- (e)  $a_i$  reprezentate polinomial;
- (f)  $a_i$  reprezentate ca puteri;
- (g)  $b_i$  reprezentate polinomial;
- (h) valori numerice obținute pentru reprezentările cifrate;
- (i) criptograma  $y_i$

Nota 1: Valorile numerice asociate literelor alfabetului sunt prezentate în tabelul următor:

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
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	0

3. Fie câmpul  $GF(26)$  căruia îi asociem polinomul de permutare  $f(x) = 6x + 7$ . Să se cifreze următoarea expresie:

GEST

## CAPITOLUL 5

### DISPOZITIVE ȘI MAȘINI CRİPTOGRAFICE

*Cuvinte cheie: mașini de criptare, mașina Enigma, mașini poligramice și tomogramice.*

#### OBIECTIVE

Înțelegerea și cunoașterea următoarelor noțiuni:

- care sunt și cum funcționează dispozitivele de criptare

Explicarea și exemplificarea funcționării mașinilor criptografice

Criptografia a jucat un rol foarte important în istorie, iar producătorii de cifruri au dorit realizarea unor sisteme de cifrare cât mai rezistente. Dacă la început și-au pus mai puțin problema operativității și a productivității, odată cu creșterea volumului de corespondență ce trebuia cifrată, aceste probleme au devenit foarte importante. De asemenea, crescând complexitatea metodelor de cifrare a crescut și riscul de a greși. Toate acestea au impus necesitatea realizării unor dispozitive care să facă mai sigură și mai rapidă activitatea de cifrare. Au apărut mai întâi simple rigle, discuri, abace, apoi acestea au fost mereu perfecționate ajungându-se la adevărate mașini la început mecanice, apoi electromecanice ajungându-se astăzi la realizarea unor sisteme de cifrare bazate pe utilizarea calculatoarelor și a microprocesoarelor specializate.

#### 5.1 Dispozitive și mașini criptografice

##### 5.1.1 Evoluția dispozitivelor criptografice

Dispozitivele și mașinile criptografice pot fi grupate în șase generații care urmăresc evoluția dezvoltării tehnicii și tehnologiilor de realizare precum și a metodelor și principiilor de cifrare.

*În prima generație* pot fi incluse acele dispozitive simple bazate pe principiul riglei și care au apărut începând se pare cu anul 475 î.e.n.

Dintre acestea cel mai vechi dispozitiv de cifrat a fost SKYTALA spartanilor care va fi descrisă ulterior.

**A doua generație** este generația mașinilor mecanice bazate pe cilindri, tamburi, roți dințate și pârghii. Acestea au apărut la sfârșitul secolului al XIX-lea, s-au dezvoltat apoi rapid atingând apogeul în timpul primului război mondial. Aceste mașini utilizează, pentru asigurarea secretului, cheile în număr foarte mare și practic nerepetabile. Deci algoritmul de cifrare se complică dar se automatizează, iar cele mai perfecționate mașini din această generație asigură și tipărirea textelor clare și a criptogramelor.

Acest lucru elimină greșelile de transcriere, mărește siguranța și operativitatea cifrării.

**A treia generație** apare la mijlocul secolului al XIX-lea, fiind condiționată de apariția telegrafului. Ea se dezvoltă în paralel cu a doua generație împrumutându-i multe din principiile și modalitățile practice de realizare. Această generație impune metode de cifrare și criptanaliză, mașinile fiind în general electromecanice.

**A patra generație** este generația aparaturii electrice și începe să se impună puțin înainte de izbucnirea primului război mondial. Primele realizări au utilizat sisteme cu relee sau se bazau pe transformări ale mașinilor de scris electrice și pe utilizarea roților de cod.

Ele au fost considerate mult timp deosebit de rezistente chiar dacă în 1929 Friedman a reușit să spargă un asemenea cod. Mai târziu prof. ABRAHAM SINKOV de la universitatea din ARIZONA a demonstrat eficacitatea utilității teoriei grupurilor la spargerea acestor coduri.

**A cincea generație** a debutat puțin înaintea izbucnirii celui de-al doilea război mondial, odată cu accentuarea dezvoltării electronicii și apariția calculatoarelor cu memorii magnetice.

Alături de telegraf, comunicațiile prin radio influențează tot mai mult criptologia impunând o dezvoltare rapidă a criptoanalizei.

**A șasea generație** apărută după deceniul al VI-lea al secolului XX se impune ca fiind generația microelectronicii, a informaticii, a procesoarelor specializate. Se bazează pe algoritmi complecși, iterativi, pe utilizarea cheilor aleatoare și unice, pe trecerea de la folosirea ca suport de informație a benzilor sau colilor de hârtie la dischete sau benzi magnetice.

Dar evoluția acestor mijloace este departe de a fi încheiată ea continuând și astăzi într-un ritm și mai alert.

Înainte de a trece la prezentarea unor dispozitive și mașini criptografice, considerăm util să facem o prezentare a principalilor beneficiari ai acestora.

### ***5.1.2 Utilizarea mijloacelor criptografice***

Beneficiarii de mijloace criptoanalitice pot fi clasificați în cinci categorii mai importante:

- militari;
- diplomați;
- servicii de informații;
- oameni de afaceri;
- corespondenți particulari.

Din aceste categorii militarii sunt de departe principalii utilizatori ai mijloacelor criptologice.

Un exemplu este edificator pe această linie. Astfel, la foarte scurt timp după ce inginerul suedez Boris Hagelin a inventat mașina de codificat C-36, uzinele Smith-Carana au construit pentru armata SUA circa 14000 de exemplare ale acestei mașini, cu denumirea „CONVERTER M-209”. Se disting două tipuri de mașini criptografice militare:

- modele „strategice” care funcționează la eșaloanele superioare ale armatelor;
- modele „tactice” care servesc la transmisiuni de campanie.

Dintre acestea, mașinile „strategice” trebuie să asigure un grad superior de securitate și de aceea sunt mai costisitoare. Mesajele prelucrate cu aceste mașini pot păstra valabilitatea și importanța perioade mai îndelungate de timp, de ordinul anilor sau chiar deceniilor.

Mașinile „tactice” nu trebuie să fie atât de complexe pentru că informațiile de la acest nivel au o valoare limitată în timp. Astfel, poziția unei subunități, să zicem, nu mai prezintă niciun interes pentru inamic la un interval de câteva zile, chiar ore, după transmiterea mesajului. Aceste mașini vor utiliza coduri de campanie care nu necesită în practică mult timp și efort pentru a fi decodificate.

Ar fi greșit să se creadă că serviciile diplomatice nu reprezintă decât o piață restrânsă pentru dispozitivele de codificare. Să admitem că, de exemplu, 120 de state fac schimb de reprezentanțe; se obține un total de peste 28000 de mașini criptografice dacă se consideră doar câte două mașini pentru ambasadă sau consulat și alte două la Ministerul de Externe al fiecărui stat.

Serviciile de informații sunt la rândul lor mari „consumatoare” de mașini de cifrat dacă avem în vedere aria largă a activităților pe care le desfășoară: contraspionaj, securitate națională și în multe situații spionaj (economic, militar etc.).

Dispozitivele și mașinile criptografice folosite de industriași și oameni de afaceri reprezintă mai puțin de 5% din totalul acestora. Mai trebuie luate în considerare aici organisme financiare (marile bănci, instituții de credit, companii de asigurări) care utilizează în mod frecvent mașinile criptografice.

Costul relativ mare al aparaturii de cifrare face ca particularii care le utilizează să fie extrem de puțini. În sfârșit, organizațiile criminale, în special cele ce se ocupă cu traficul de orice natură, recurg deseori la criptografie.

De exemplu, în perioada prohibiției, navele care soseau în SUA cu alcool primeau prin radio instrucțiuni codificate de la căpeteniile traficantilor.

Numai între anii 1927-1929 o echipă condusă de o femeie Elisabeth Smith Friedman a descifrat peste 12000 din aceste mesaje, iar poliția dispunea, la rândul ei, de păzitoare de coastă dotate cu laboratoare de decodificare.

### 5.2 Funcționarea mașinilor criptografice

Facem în continuare o prezentare a dispozitivelor și mașinilor criptografice, respectând pe cât posibil ordinea cronologică a apariției lor, dar făcând în același timp distincție între cele două categorii.

**SKYTALA** apărută în secolul V î.e.n. era un baston 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 scrierea textului, panglica era dublată, mesajul devenea indescifrabil, întrucât literele erau dezasamblate. Mesajul se putea descifra numai de persoana care dispunea de un baston de lungime și grosime identice cu dimensiunile inițiale pe care să fie înfășurate din nou panglica primită. Ea realiza o transpoziție, fiind o primă formă a acestei metode de cifrare.

**Criptograful lui Alberti.** Era alcătuit din două discuri concentrice, cu diametre diferite, suprapuse. Fiecare disc era împărțit în 24 de sectoare, pe care erau înscrise literele și cifrele. Pe discul mic erau înscrise 20 de litere (fără H, J, K, U, W) și cifrele de la 1 la 4, iar pe al doilea 23 de litere (fără H, K, Y) și conjuncția „ET”. Ordinea lor era arbitrară. Pentru cifrare se stabilea o cheie, de exemplu  $D = A$ . Aceasta însemna că pentru cifrare litera  $D$  de pe discul mic se așează în dreptul literei  $A$  de pe discul mare și apoi începea cifrarea. Alberti

recomanda pentru mărirea rezistenței schimbarea cheii după un număr de cuvinte. Criptograful lui Alberti a fost perfecționat de Silvester, Argenti și alții constituind un element de bază pentru criptoграfele de tip disc apărute ulterior. Silvester Porta a împărțit discurile în 26 de sectoare, utilizând apoi toate cele 26 de litere, criptograful său permițând o substituție simplă, dar complet literală.

**Criptograful lui Wheatstone.** A fost inventat în 1867, este realizat tot cu două discuri și realizează o substituție dublă cu cheie finită. Cele două discuri erau fixe, pe unul dintre ele cele 26 de litere fiind dispuse în mod normal, iar pe cel de-al doilea literele erau așezate aleatoriu, conform unei chei. Peste cele două discuri se mișcau două ace într-un raport 27 la 26 de rotiri, producându-se astfel un decalaj continuu și ciclic între cele două alfabete.

Aparent, se produce o cifrare cu autocheie, ceea ce i-a făcut pe mulți specialiști să-l declare indescifrabil.

Mai târziu însă *Kerchoffer* a demonstrat că numărul alfabetelor este limitat și că de fapt el realizează o substituție cu dublă cheie.

Cifrarea se face pornind de la o cheie, de exemplu  $A = C$ , adică acul mare se afla la A, iar acul mic se afla la C. În continuare se duce acul mare la prima literă a textului clar și se ia drept literă cifrată litera pe care se oprește acul mic. Operația se repetă pentru a doua literă ș.a.m.d, deplasarea făcându-se în sensul acelor de ceasornic.

La descifrare se procedează invers, plecând din aceleași puncte determinate de cheie și se deplasează acele până când acul mic ajunge la poziția primei litere din criptogramă, prima literă a textului în clar fiind litera la care a ajuns acul mare.

**Criptograful Bazeries.** A fost inventat în 1891, fiind considerat o mare realizare din punct de vedere criptologic la vremea respectivă.

Era realizat dintr-un ax și 20 de inele care aveau înscrise pe marginea exterioară un număr, o literă de identificare și un alfabet de 25 de litere (fără W).

Cheia sistemului se formează din litere sau cifre și reprezintă ordinea de montare a inelelor pe ax, ea fiind valabilă atât la cifrare, cât și la descifrare. Cifrarea se face împărțind textul clar în grupe de 20 litere și apoi aranjând inelele, astfel încât pe o generatoare să apară înscrise cele 20 de litere ale primei grupe. Textul cifrat poate fi considerat oricare alt grup de litere de pe o altă generatoare.

Descifrarea se face în mod identic.

Rezistența acestui criptograf constă în numărul mare al cheilor: 20!



## Dispozitive și mașini criptografice

S-a stabilit că cifrarea cu ajutorul cilindrului lui Bazaries a asigurat secretul mesajelor în funcție de pregătirea criptanaliștilor, de la 6 ore până la 3 zile. De aceea, ca măsură de protecție a secretului, cheia era schimbată la fiecare 6 ore.

**Criptograful lui Soudart.** A fost inventat în anul 1914 și realizează un cifru bazat pe dubla transpoziție. Principiul este asemănător celui al lui Bazaries, având tot 20 de inele montate pe un ax. Marginea exterioară este împărțită în două, pe partea stângă erau trecute cele 25 de litere ale alfabetului (fără W), iar pe marginea din dreapta, în dreptul literelor erau trecute literele textului clar.

Cheia de cifrare era literală sau numerică, putând fi formată din cel mult 20 de caractere (litere sau cifre de la 1 la 20), într-o ordine aleatoare. De exemplu, cheia DISCRET va deveni în numeric 2461537, ceea ce determină montarea inelelor pe cilindru, în ordinea dată de cheie. După montare se rotesc inelele până apare pe poziția inițială cuvântul cheie.

2	4	6	1	5	3	7
DM	IE	ST	CO	RD	EE	TC
ER	JI	TP	DT	SO	FG	UR
FA	KF	UI	EC	TE	GC	VL
GA	LS	VI	FC	UE	HP	XT

În locurile din dreapta rămase libere se înscriu literele textului, dar pe linii de la stânga la dreapta și de sus în jos. Fie textul clar: METODE CRİPTOGRAFICE CLASICE.

Punctul s-a folosit cu PT. Se scot inelele și se montează în ordine numerică normală.

1	2	3	4	5	6	7
CO	DM	EE	IE	RD	ST	TC
DT	ER	FG	JI	SO	TP	UR
EC	FA	GC	KF	TE	UI	VL
FC	GA	HP	LS	UE	VI	XT

Se realizează a doua transpoziție aliniind inelele după alfabetele înscrise la stânga. Criptograma va fi: **OTCCM/ RAAEG/ CPEIF/ SDOEE/ TPIIC/ RLT.**

Dintre criptografele de tip riglă exemplificăm: *rigla Saint-Cyr* – realizează o substituție polialfabetică; se compune dintr-un suport fix pe care este înscris câte un alfabet într-o anumită ordine și un cursor.

Cifrarea se execută simplu, plecând de la o cheie literală care realizează corespondența dintre litera de pe cursor și litera respectivă de pe suport.

Litera cifrată se citește pe suport în dreptul literei clare de pe cursor. Cheia determină poziția începutului alfabetului de pe cursor față de alfabetul de pe suport. De exemplu, dacă prima litera a cheii este *D*, atunci cifrarea literei *C* va fi *F*.

Riglele au fost perfecționate în permanență și se folosesc și azi. S-au realizat rigle la care alfabetul s-a completat cu cifre și semne de punctuație. Cea mai reprezentativă este *rigla universală*. Aceasta este formată din 20 de riglete care se deplasează orizontal, astfel încât pe coloana orificiilor să apară cheia numeric convenientă. Această riglă oferă foarte multe posibilități, cum ar fi: cifrarea numerică, duplicarea transpozițiilor, utilizarea de chei multiple literale sau numerice.

Oricât de complicate ar fi, criptografele nu sunt dispozitive care ajută pe cel ce execută cifrarea să elimine greșeli de transcriere, să scurteze timpul de lucru. Spre deosebire de ele, mașinile de cifrat automatizează procesul cifrării, eliminând în întregime greșelile de transmisie, rămânând doar greșelile de operare dependente de instruirea și antrenamentul cifratorului. Transformările realizate cu mașinile de cifrat nu sunt simple transpoziții de elemente, existând posibilitatea realizării unor funcții de transformare, bazate pe calcule complexe.

În ceea ce privește productivitatea, aceasta este mult superioară procedeelelor de cifrare utilizate anterior. Perfecționarea mașinilor de cifrat este indiscutabil legată de nivelul de dezvoltare al științei și tehnicii din epoca respectivă.

Concepute pentru mecanizarea operațiunilor de cifrare și descifrare, mașinile clasice, tradiționale constituie sisteme criptografice complexe, îndeplinind atât cerința transparenței algoritmilor de cifrare-descifrare pentru utilizatori, cât și mutarea ponderii secretului pe sistemul de chei utilizat. În linii mari, o mașină de cifrat, respectiv descifrat, trebuie să cuprindă următoarele elemente:

- mulțimea (algoritmul) transformărilor;
- sistemul de chei;
- dispozitive auxiliare care să permită introducerea datelor și cheilor, adaptarea acestora la canalul de legătură etc.

Din punctul de vedere al metodei utilizate, mașinile de cifrat polialfabetice pot fi împărțite în trei categorii:

- mașini polialfabetice simple, care realizează substituții diferite la fiecare literă a textului în clar, în funcție de numărul de alfabet utilizate. Nu pot exista mai mult de 26 de alfabet diferite și acest impediment duce la apariția repetărilor, factor deosebit de decisiv în bătălia descifrărilor;
- mașini polialfabetice bazate pe principiul autocifrării. Substituțiile realizate depind atât de numărul alfabetelor, cât și de caracteristicile textului cifrat. Aceasta duce la reducerea numărului de repetări, deci îngreunează lucrul criptanalizatorilor, dar complică în același timp descifrarea;
- mașini polialfabetice complexe, la care există posibilitatea schimbării aleatoare a rangurilor diferitelor alfabet utilizate.

### 5.2.1 Mașina „Enigma”

Dintre mașinile din această categorie, vom prezenta mașina „Enigma”.

Este o mașină polialfabetică complexă. A fost fabricată în Germania și larg utilizată în timpul celui de-Al Doilea Război Mondial. De altfel, enigma „Enigmei” încă nu a fost elucidată pe deplin. Specialiștii oscilează de la a-i atribui în întregime eșecul Germaniei, până la a o trata doar ca un factor conjunctural, complementar, în desfășurarea celei mai mari conflagrații mondiale din istoria omenirii și deci fără contribuție semnificativă la destinul beligeranților. Înainte de a încerca să tratăm și noi acest subiect, să vedem ce a însemnat din punct de vedere criptografic mașina Enigma.

Algoritmul criptografic al mașinii avea la bază o substituție polialfabetică complexă, realizată cu ajutorul a trei discuri mobile, prevăzute fiecare cu câte 26 de contacte pe o față și 26 de ace pe cealaltă.

Contactele și acele sunt legate electric între ele; două câte două, dar nu direct, ci într-o anumită ordine.

Există și un al patrulea disc, fix, prevăzut cu 26 de contacte, legate două câte două. Discurile sunt coaxiale și montate în așa fel încât acele unui disc presează contactele celui alt, realizându-se astfel 13 linii paralele, formate din câte 3 segmente cu variație independentă.

Circuitul electric se realizează de la sursă la tastă, care marchează prin apăsare litera clară (la cifrare) sau pe cea cifrată (la descifrare), prin contactul

acesteia la o contraplață cu 20 de contacte fixe, apoi prin discurile I, II și III, după circuite realizate din segmentele cu variație independentă, la discul al patrulea, fix, și prin una dintre cele 13 perechi de contacte ale acestuia se înapoiază prin alt circuit, format din segmente de către cele trei discuri mobile, la un bec ce iluminează litera cifrată (la cifrare) sau pe cea clară (la descifrare).

Complexitatea mașinii constă în numărul de segmente, care realizează variația circuitelor, datorită rotirii celor trei discuri în funcție de tastele apăsată și de caracteristicile componenteii textului clar.

La fiecare apăsare de tastă, în funcție de cheia adoptată, cele trei discuri se deplasează unul față de celălalt, cu un anumit număr de poziții, astfel încât se realizează de fiecare dată un anumit circuit electric și deci o anumită corespondență între litera clară apăsată și litera cifrată iluminată.

Cele trei discuri, cu diferite scheme de legături între contactele și acele acestora, se pot schimba între ele, realizând practic, în acest fel, o altă variantă a mașinii. De altfel, în afara celor trei discuri montate, fiecare mașină dispunea și de un alt set de discuri cu legături complet diferite.

Dacă se consideră că fiecare dintre cele cinci discuri poate fi realizat într-un alt mod, se obține o succesiune de patru alfabetete diferite, dintre care trei, fiind realizate de discurile mobile, sunt permutabile. Dar datorită legăturilor întrețesute, se poate considera că la fiecare transformare de literă clară în litera cifrată contribuie șapte alfabetete (șase realizate de traseele – dus și întors – din discurile mobile și unul în cel de-al patrulea disc).

Fără îndoială, aceste posibilități de schimbare a alfabetetelor determină un număr foarte mare de transformări, greu de descifrat, chiar și în cazul în care se posedă o variantă a mașinii de cifrat.

Un alt element caracteristic al mașinii îl constituie cheia; aceasta definește ordinea discurilor, poziția inițială și deplasările relative ale discurilor în timpul cifrării unei anumite litere.

Deși robustă și comodă, datorită faptului că nu avea posibilitatea de a imprima textul cifrat sau clar (acestea se culegeau și se notau sau se scriau la altă mașină), greșelile de transcriere erau destul de frecvente. La acestea se mai adăugau și câteva necazuri tehnice produse de imperfecțiunea unor contacte, arderea becurilor și scăderea capacității sursei de alimentare.

Pentru a vedea cum se lucra cu această mașină (din care armata noastră a achiziționat și exploatat circa 100 de bucăți), vom reproduce câteva pasaje (cu eliminarea unor elemente caracteristice de limbaj) din „Instrucțiunile de

întrebuințare a mașinii de cifrat Enigma”, editate la 12 ianuarie 1937 de către Marele Stat Major, reactualizate cu unele precizări organizatorice la 13 ianuarie 1940.

„Extinderea întrebuințării mașinii de cifrat Enigma la toate Secțiile Armatei se ordonă de către Comandamentul Suprem al Armatei”.

Lungimea minimă a unei radiograme cifrate cu mașina de cifrat Enigma nu este limitată, dar cea maximă nu va depăși 250 de litere.

Cheile se vor schimba zilnic la orele 00.00. Tabelele cu cheile zilnice se distribuie lunar.

Cheile zilnice pentru cifrare impun schimbarea:

- poziției discurilor (în cifre romane);
- poziției inelelor (în cifre arabe sau litere);
- legăturilor fișelor (în litere).

Correspondența dintre literele și numerele cheii este normală:  $A = 01$ ,  $B = 02, \dots, Y = 25$  și  $Z = 26$ .

Poziția discurilor (II, I, III etc.) ne arată ordinea în care acestea sunt introduse în mașina de cifrat, de la stânga la dreapta.

Poziția inelelor (13 08 11) ne arată aranjarea lor inițială. Correspondența dintre poziția discurilor și ordinea numerelor ce reprezintă poziția inelelor se realizează pentru fiecare caz în parte. Pentru exemplu dat: II--13, I--08 și III--11.

Legăturile fișelor se realizează conform succesiunii indicate de perechile de litere din cheie.

Dacă cheia legăturilor fișelor este:

AO BI DV EH GZ KW LX MU RY QT

atunci se realizează următoarele legături:

A cu O, B cu I etc.

Cheia unei radiograme se va determina printr-o grupă cu cinci litere, din care primele două sunt fără semnificație (se vor schimba la fiecare radiogramă), iar următoarele trei indică cheia din „Tabelul cu chei”.

La telegramele compuse din mai multe fragmente, fiecare fragment va avea litere diferite ca grupă de recunoaștere și alte două litere fără semnificație. Grupa de recunoaștere se transmite prima și nu se cifrează.

Cele trei litere care formează poziția inelelor vor fi alese din combinații de la AAA la ZZZ, excluzându-se cele cu semnificație, prescurtări -RGT-, indicative de apel, comenzi de trafic, litere în ordine alfabetică normală sau inversă - ABC și CBA etc.

Cifratorul așează tamburii, literele în fereștruci și cablurile de legătură conform cheii zilnice.

LA A DOUA RADIOGRAMĂ POZIȚIA INELELOR SE MODIFICĂ CU CHEIA DE RADIOGRAMĂ.

După așezarea cheii de radiogramă, se bate textul în clar, iar literele luminate – cifrate – se culeg pe formularul de radiogramă, în grupe de câte cinci.

La descifrare, cifratorul va aranja mașina conform cheii zilnice. Va extrage apoi cheia de radiogramă (cele trei litere cifrate din preambulul radiogramei, adăugate după cele trei litere de bază), și o va descifra, apăsând tastele corespunzătoare.

După descifrarea cheii de radiogramă se vor stabili pozițiile celor trei inele, în concordanță cu pozițiile discurilor indicate în cheia zilnică.

În tabel, o cheie zilnică se prezintă astfel: 4;I II III;16 11 13; BN KE VZ CO DI FR HU JW LS TX; adq nuz opw vxz, unde:

- 4 - era data (ziua);
- I - poziția discurilor;
- 16 - poziția inelelor;
- BN - legăturile fișelor;
- nuz - grupele indicative.

Pentru exemplificare, o radiogramă va arăta astfel:

1755-139-wep hfi -  
ulznu sgexu nfopr salmc  
ydr jd.....

unde:

- 1755 - ora expedierii;
- 139 - numărul literelor;
- wep - poziția de bază;
- hfi - cheia cifrată a telegramei sau radiogramei;
- ul - litere fără semnificație;
- znu - grupa de recunoaștere;
- sgexu - textul cifrat.

Pentru exercițiu se vor folosi un anumit număr de combinații de chei, care însă nu se vor întrebuiți și pentru transmiterea radiogramelor. După analiza acestor instrucțiuni, ne putem da seama că pentru a evita greșelile și mai ales încurcăturile în fixarea cheilor, operatorii trebuiau pregătiți și antrenați. De

asemenea, o mare parte dintre posibilitățile mașinii erau puse în valoare sau estompate de sisteme de realizare a cheilor și de frecvența lor de schimbare.

Greșelile de operare, de schimbare a cheilor, repetarea radiogramelor care aveau elemente inițiale greșite și lungimea excesivă a textelor erau elemente ce puteau compromite mașina, lucru care de altfel s-a și întâmplat.

### 5.2.2 Mașini cu cifrare poligramică

Dacă mașinile polialfabetice realizează o cifrare caracter cu caracter, există și tipuri de mașini care fac o cifrare a  $n$ -gramelor, prin  $n$ -grame înțelegând grupuri de  $n$  caractere.

Principiul cifrării poligramice a fost imaginat în deceniul trei al acestui secol de Lester S. Hill și publicat în „The American Mathematical Monthly”, în 1929, și se bazează pe calcule matematice.

Se alocă literelor alfabetului valori numerice de la 0 la 25 în mod aleator. Se pot imagina diverse combinații între numerele asociate literelor, operații care se pot materializa practic cu ajutorul unei mașini de cifrat. Toate operațiile se efectuează modulo 26. De exemplu, fie alfabetul:

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

Dorim să cifrăm trigrame ale textului clar:

*C R I P T O G R A F I E*

Aceste trigrame vor fi: CRI-PTO-GRA-FIE.

Alegând cheia de cifrare: **ENIGMATIC**, forma numerică fiind 25 7 0/ 24 22 5/ 23 0 21, pentru prima criptogramă se pot scrie ecuațiile:

$$25 \times C + 7 \times R + 0 \times I = 25 \times 21 + 7 \times 12 + 0 \times 0 = 11 \rightarrow J$$

$$24 \times C + 22 \times R + 5 \times I = 24 \times 21 + 22 \times 12 + 5 \times 0 = 14 \rightarrow F$$

$$23 \times C + 0 \times R + 21 \times I = 23 \times 21 + 0 \times 12 + 21 \times 0 = 15 \rightarrow O$$

Deci trigramei *CRI* îi corespunde trigrama *JFO*.

Descifrarea se face în mod asemănător. Se pleacă de la trigrama cifrată și folosind alți coeficienți se determină mesajul clar.

Practic, se rezolvă sistemul:

$$\begin{cases} K_1x + K_2y + K_3z = C_1 \\ K_4x + K_5y + K_6z = C_2 \\ K_7x + K_8y + K_9z = C_1 \end{cases}$$

unde:  $x, y, z$  sunt valorile numerice ale literelor textului clar;

$K_1, K_2, \dots, K_9$  sunt valorile numerice ale literelor care constituie cheia;

$C_1, C_2, C_3$  sunt valorile numerice ale literelor criptogramei.

În cazul nostru vom avea:

$$\begin{cases} 25x + 7y + 0z = 11 \\ 24x + 22y + 5z = 14 \\ 23x + 0y + 21z = 15 \end{cases} \pmod{26}$$

Soluția sistemului este  $x = 21, y = 12, z = 0$ .

Această soluție se obține în mod practic din următoarele calcule:

$$\begin{cases} x = K'_1 C_1 + K'_2 C_2 + K'_3 C_3 \\ y = K'_4 C_1 + K'_5 C_2 + K'_6 C_3 \\ z = K'_7 C_1 + K'_8 C_2 + K'_9 C_3 \end{cases}$$

În lucrarea „*Some Cryptographic Applications of Permutation Polynomials*”, Jack Levine și J.V. Brawley prezintă o altă metodă de cifrare poligrafică prin utilizarea mai multor câmpuri și permutarea polinomială.

Fie un alfabet  $P$  cu  $L$  litere. Cu cele  $L$  litere pot fi construite  $N = L^n$  poligrame (fiecare având  $n$  litere), descrise de succesiunile:

$$\lambda = L_1 L_2 \dots L_n; L_i \in P; i = \overline{1, n}.$$

Considerăm acum cele  $N$  poligrame ca fiind litere ale unui alfabet  $A$ . La cifrare textul clar compus din literele alfabetului  $P$  se divide în grupuri de  $n$  litere, obținând astfel un număr de poligrame. Numărul  $N$  se poate descompune sub forma:  $N = \sum_i N_i$ , unde:  $N_i = P_i^{m_i}$  sau  $N_i = P_i^{m_i} + 1$ , iar

$P_i$  – numere prime.

Se poate realiza astfel o partiționare a alfabetului  $A$  în  $K$  subalfabete, astfel că  $A = \bigcup_{i=1}^k A_i$ .

Subalfabetului  $A_i$ ,  $i$  se repartizează  $N_i$  poligrame, un câmp Galois extins  $GF(P_i^{m_i})$  și o permutare polinomială  $P_i(x_i)$ , dacă  $N_i = P_i^{m_i}$ , sau o funcție rațională  $R_i(x_i)$ , dacă  $N_i = P_i^{m_i} + 1$ .



Modul în care se face împărțirea poligramelor este următoarea:

1. Se asociază fiecărei litere ale alfabetului  $P$  o valoare numerică din secvența  $0, 1, 2, \dots, (L - 1)$ .

2. Schimbăm orice poligramă  $\lambda = L_{i_1} L_{i_2} K L_{i_n}$  cu o formă numerică  $\beta = V_{i_1} V_{i_2} K V_{i_n}$ , unde  $V_i$  este valoarea numerică a literei  $L_i$ .

3. Fie  $\sigma = V_{i_1} L^{n-1} + V_{i_2} L^{n-2} + K + V_{i_{n-1}} L + V_{i_n}$  pe care o definim ca fiind valoarea numerică repartizată poligramei  $\lambda$  ( $0 \leq \sigma \leq L^n - 1$ ).

4. În subalfabetul  $A_1$  vor fi repartizate poligramele pentru care  $0 \leq \sigma \leq N_1 - 1$ , în subalfabetul  $A_2$  acele poligrame pentru care  $N_1 \leq \sigma \leq N_2 - 1$  ș.a.m.d.

5. Definim un număr  $\rho$  astfel:

$\rho = \sigma - (N_1 + N_2 + \dots + N_{i-1})$  asociat poligramelor subalfabetului  $A_i$ . Se observă că  $0 \leq \rho \leq N_{i-1}$ .

6. Se exprimă numărul  $\rho$  în baza  $P_i$ , astfel:

$$\rho = C_0 P_i^{m_i} + C_1 P_i^{m_i-2} + K + C_{m_{i-1}} = (C_0, C_1, \dots, C_{m_{i-1}}),$$

unde  $C_0, C_1, \dots, C_{m_{i-1}}$  sunt elementele câmpului  $GF(P_i^{m_i})$ .

7. Printr-o procedură oarecare stabilim o corespondență între poligramele fiecărui subalfabet  $A_i$  și elementele câmpului  $GF(P_i^{m_i})$  care sunt cifrate cu folosirea unui polinom de permutare sau a unei funcții de permutare, după cum  $N_i = P_i^{m_i}$  sau  $N_i = P_i^{m_i} + 1$ . Vom ilustra diferitele etape ale procedurii printr-un exemplu. Se consideră cazul cifrării diagramelor, deci  $n = 2$ ,  $L = 26$  având repartizate literelor următoarele valori numerice:  $A = 0, b = 1, C = 2, \dots, Z = 25$ .

Vor fi deci  $N = 26 \times 26 = 676$ , pe care îl partiționăm astfel:

$$676 = 7^3 + 13^2 + (3^4 + 1) + (3^4 + 1)$$

Deci  $N_1 = 7^3 = 343$  cu câmpul  $GF(7^3)$ ;

$N_2 = 13^2 = 169$  cu câmpul  $GF(13^2)$ ;

$N_3 = N_4 = 3^4 + 1 = 82$  cu câmpul  $GF(3^4)$ .

Vom împărți cele 676 diagrame în patru subalfabete, astfel:

subalfabetul  $A_1$ , dacă  $0 \leq \sigma \leq 342$ ;

subalfabetul  $A_2$ , dacă  $343 \leq \sigma \leq 511$ ;

subalfabetul  $A_3$ , dacă  $512 \leq \sigma \leq 592$ ;

subalfabetul  $A_4$ , dacă  $593 \leq \sigma \leq 675$ .

Având de exemplu diagrama  $RM = (17,12)$  calculăm:

$$\sigma = 17 \times 26 + 12 = 454.$$

Deci diagrama  $RM$  face parte din subalfabetul  $A_2$ .

Calculăm  $\rho = \sigma - N_1 = 454 - 343 = 111$ , valoare pe care o exprimăm în baza 13.

$$111 = (8,7) \text{ (baza 13), adică } 111 = 8 \times 13 + 7.$$

Rezultă că  $RM = (8,7)$ .

Dacă folosim polinomul de permutare de forma  $y_i = x_i^5 + b_i$ , având  $b_1 = (12,8)$ , se obține  $y_i = 88$ , deoarece  $x_1 = 111$ , iar  $b_i = 164$ . Toate aceste operații se fac modulo 169, câmpul utilizat fiind  $GF(13^2)$ .

$$\sigma = 88 + 343 = 431 = (16,15) = QP \text{ (adică } 16 \times 26 + 15).$$

Deci diagramei  $RM$  îi corespunde criptograma  $QP$ .

Această metodă, necesitând o serie de calcule complicate, impune utilizarea calculatorului pentru executarea operației de cifrare și descifrare.

### **5.2.3 Mașini tomogramice**

O altă categorie de mașini de cifrat o constituie *mașinile de cifrat tomogramice*.

Fracționarea literelor (tomogramia) în scopul cifrării a fost și continuă să fie, chiar și cu mijloace moderne, o temă de studiu deosebit de atractivă. De la substituții cu numere fracționare, până la accesul pe bit al criptologiei moderne, iată gama procedeele tomogramice utilizate în acest tip de cifrare.

Alfabetul lui Collon, trifidul lui Delastelle, alfabetul Morse și codul Baudot, iată numai câteva dintre cifrurile tomogramice.

Collon realiza un alfabet de substituție cu grupe formate din două simboluri, bazate pe utilizarea combinatorie a 5 sau 6 simboluri diferite.

Trifidul lui Delastelle realiza substituția prin grupe de câte trei cifre, rezultate din combinarea cifrelor 1, 2 și 3.

Alfabetul Morse, bazat pe combinații de puncte, linii și spații, și mai ales codul Baudot sunt larg cunoscute și utilizate și în telegrafia modernă. Ele nu mai reprezintă de mult un secret pentru un mare număr de oameni și, deci, din punct de vedere criptografic nu mai au demult nicio valoare.

Sunt considerate, acum, la fel ca și alfabetele limbilor de circulație mondială și utilizate ca elemente de reprezentare a textelor clare și ca atare supuse unor transformări comparatorii, bazate pe diferite combinații.

*Sistemul Vernam* a fost inventat în decembrie 1917 și apoi larg răspândit până la sfârșitul Primului Război Mondial. El are marele merit că a realizat un dispozitiv practic de mixare a textelor clare transpuse pe banda perforată (în cod Baudot), cu o succesiune de grupe, în același cod, transpuse pe altă bandă.

Din punct de vedere criptografic, procedeul de mixare reprezintă algoritmul sistemului, iar cheile, benzile suport ale grupelor utilizate pentru mixaj.

Se cunoaște că în cod Baudot literele sunt reprezentate prin grupe de câte cinci simboluri (1 și 0, perforații pentru 1 sau pentru 0 pe banda de hârtie, impulsuri electrice pozitive și respectiv negative etc.).

De exemplu:  $A = 11000$  și  $I = 01100$  etc.

Metoda Vernam se bazează pe o relație de tip modulo 2 între aceste simboluri, conform următoarei convenții:

$$1 + 1 = 0$$

$$1 + 0 = 1$$

$$0 + 1 = 1$$

$$0 + 0 = 0$$

Dacă se consideră că textul clar este  $A$  și cheia  $B$  se obține ca text cifrat  $G$ , astfel:

$$A = 11000$$

$$B = 10011$$

$$G = 01011$$

Ca să mixeze impulsurile electrice, Vernam a inventat un dispozitiv realizat cu magneți, relee și bobine. Cum cifrarea și descifrarea erau reciproce, același aparat era folosit la ambele operații.

Dacă cineva interceptează aceste mesaje, obține doar secvențe de 0 și 1 sau șiruri de caractere fără semnificație. De asemenea, a crescut mult operativitatea cifrării și s-au eliminat, aproape în întregime, erorile de transcriere.

Prin sistemul Vernam s-a impus în lumea criptografiei schimbarea accentului preocupării principale de la algoritmul de criptare la strategia cheilor.

În primele zile, cheile pentru acest aparat se obțineau prin perforarea benzilor cu simboluri aleatoare.

Apoi s-au căutat și s-au găsit și alte metode de elaborare a cheilor.

Cu ajutorul simbolurilor utilizate în codul Baudot, se poate realiza un tabel de  $32 \times 32$ , în care, pe rând, coloanele puteau să fie socotite chei.

Întrucât secretul sistemului Vernam constă tocmai în cheia folosită, s-a renunțat la utilizarea unor chei repetate în favoarea unor foarte lungi. Aceste chei foarte lungi prezentau însă inconveniente de manipulare. Dificultatea a fost rapid escaladată, datorită ideii inginerului Morshouse de a combina două chei scurte, ca și cum una ar fi servit la cifrarea celeilalte.

Rezultatul obținut era o cheie foarte lungă, denumită apoi cheie secundară. Cheia secundară se obține din două chei primare scurte, datorită diferenței numerelor de perforații (semne) în cele două cazuri. Astfel, dacă prima cheie primară avea 1000 de semne, iar a doua 999, datorită diferenței de un semn, prin combinații succesive, se putea obține o cheie cu 999000 de semne. Acest lucru înseamnă că în loc de a utiliza o cheie de 4000 m lungime se utilizau două chei de aproximativ 4 m.

Acest procedeu însă nu exclude complet repetițiile, mai ales pe anumite sectoare de bandă și impune cert marcarea sfârșitului cheii, neexcluzând în totalitate nici sensul, inconveniențe prompt „speculate” de criptanaliști, ceea ce făcea ca rezistența sistemului Vernam să scadă pe măsura ce el era mai cunoscut și mai des utilizat.

Saltul următor l-a făcut Mauborgne prin introducerea „cheii cu o singură utilizare”, rezultată din aplicarea pentru realizarea unei chei atât a principiului aleatorismului, cât și al non-repetiției.

*Sistemul Jammet* realizat de Jammet Onde Electrique, în august 1926, este analog sistemului Vernam, bazându-se însă pe alfabetul Morse.

Problema cea mai grea pentru acest sistem a fost sincronizarea manipulatorilor și, bineînțeles, extragerea semnalului clar la recepție din cel secretizat. Sistemul a reprezentat un interesant experiment tehnic, dar s-a dovedit fără valoare criptografică.

*Sistemul Belin* este un prim sistem de secretizare a transmisiilor imaginilor. Principiul de funcționare se baza pe exploatarea optică a imaginii și transformarea impulsurilor luminoase în impulsuri electrice. Succesul acestui tip de transmisii era asigurat de sincronizarea perfectă a discurilor transmițătoare și receptoare.

Pentru a evita recepția unor astfel de imagini, Belin a realizat un sistem de variație controlată a vitezelor celor două discuri, astfel că orice interceptare și încercare de reproducere produceau o imagine deformată, neinteligibilă.

### Test de autoevaluare

1. Cum au evoluat mașinile criptografice și care sunt principalii beneficiari?
2. Care sunt principalele clase de mașini de criptare clasificate după modul de funcționare?

### Tema de autoinstruire

1. Realizați schema logică a metodei Vernam, bazată pe exemplul din curs.

### Lucrare de verificare

1. Descrieți funcționarea mașinii de cifrat ENIGMA.
2. Să se descrie funcționarea următoarelor sisteme criptografice:
  - a. Sistemul Vernam
  - b. Criptograful Bazaries.

3. Considerăm un cifru de permutare la care se fixează numerele naturale  $p, q$ . Textul clar se împarte în blocuri de câte  $p \cdot q$  caractere. Fiecare astfel de bloc se scrie pe liniile unei matrice de  $p$  linii și  $q$  coloane. Criptarea blocului se realizează scriind aceste matrice pe coloane. De exemplu, pentru  $p = 3, q = 4$ , textul clar MAINI CURATE se scrie:

M A I N  
I C U R  
A T E X

(textul s-a completat cu litera X). Textul criptat va fi MIAACTIUENRX.

Decriptați următorul text DJNOUDNAINPAPANONZ criptat într-un mod similar.

4. Folosind atacul prin forță brută, decriptați mesajul WYPTBSJBYZ criptat cu un sistem Cezar.

5. Presupunem că Alice dorește să transmită mesajul "TRIMIT AJUTOR" folosind un canal de comunicație nesigur. Atât ea, cât și Bob se presupune că au obținut printr-o metodă necunoscută un set de chei generate absolut aleator și că una dintre aceste chei este "AFVKGHOPERTS". Folosind sistemul de criptare Vernam, cifrați textul clar enunțat.

6. În condițiile enunțate de exercițiul de la punctul 5, descifrați următorul text cifrat: T W D W O A O Y O K H J.

7. Să se creeze textul clar INAINTE SI LA DREAPTA folosind sistemul de criptare Hill cu matricea:  $\begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 22 & 19 & 1 \end{pmatrix}$

## CAPITOLUL 6

### ELEMENTE DE CRIPTANALIZĂ

*Cuvinte cheie: limbaje naturale, decriptarea, spargerea sistemelor cripto.*

#### OBIECTIVE

Înțelegerea și cunoașterea următoarelor noțiuni:

- limbaje naturale și caracteristici
- rolul *sistemului* și al *cheii de criptare* în metodele de decriptare
- elementele ce stau la baza spargerii sistemelor criptografice și poligrafice

#### 6.1 Caracteristicile statistice ale limbajelor naturale

Din punctul de vedere al teoriei informației, limba poate fi socotită ca o sursă de comunicări, care creează o oarecare cantitate de informație, măsurată în biți. În orice limbă s-ar face exprimarea, comunicările se compun din combinații și șiruri de litere și cuvinte care nu sunt cu totul întâmplătoare, cum ar părea la prima vedere. Șirurile de litere formează cuvinte și fraze cu o structură specifică unei anumite limbi. Studiarea unei astfel de structuri specifice fiecărei limbi permite să se obțină o economie de timp în transmiterea textelor cu ajutorul liniilor de telecomunicații, prin codificarea optimă a comunicărilor și literelor.

În scopuri ilustrative, în studiul limbilor se folosesc uneori limbaje abstracte sau artificiale. În acest caz se consideră fie că toate literele au aceeași probabilitate și se succed independent una de alta (*aproximație de ordin 0*), fie că literele se succed independent, dar cu probabilitățile limbilor reale (*aproximație de ordin 1*). Se poate impune, de asemenea, să se țină seama de structura diagramelor, adică după fixarea unei anumite litere, litera următoare se alege în funcție de frecvențele cu care diferitele litere se succed după litera fixată (*aproximație de ordin 2*). Se poate ține seama, de asemenea, de structura trigramei, adică literele se aleg în funcție de probabilitățile de apariție care depind de primele două litere ale trigramei (*aproximație de ordin 3*).

Așadar teoretic se poate calcula entropia medie,  $H_0$ , care revine unei litere de text (în care literele au aceeași frecvență) dacă se cunoaște numărul  $n$  de litere al alfabetului. De exemplu, dacă alfabetul are 8 litere: (a, b, c, d, e, f, g, h) iar probabilitățile lor de apariție sunt  $p = 0,125$ , pentru fiecare literă, atunci entropia medie pe literă este dată de mărimea:

$$H_0 = \log_2 8 = 3 \text{ biți/literă}$$

$$H_0 = \log_2 26 = 4,7 \text{ biți/literă.}$$

În cazul alfabetului latin care conține 26 de litere, entropia medie pe literă  $H_0$  dă o indicație destul de vagă asupra entropiei unei limbi (de fapt  $H_0$  reprezintă entropia alfabetului).

În realitate, frecvențele de apariție a literelor nu sunt aceleași și ținând seama de probabilitățile de apariție a diferitelor litere se obține *entropia de ordin 1* cu relația:

$$H_1 = - \sum_{i=1}^n p_i \log_2 p_i < H_0$$

În tabelul de mai jos sunt prezentate entropiile de ordin 0 și 1 ale câtorva limbi de circulație mondială, precum și ale limbii române:

Limba	$H_0$	$H_1$
Engleză	4,7	4,03
Rusă	5	4,34
Germană	4,75	4,08
Franceză	4,7	3,94
Română	4,7	4,47

Dar nici această mărime nu este precisă, deoarece în calculul entropiei  $H_1$  s-au neglijat dependențele ce există între diferitele litere.

La limbile reale se observă că unele litere apar de cele mai multe ori împreună. Pentru a ține seama de aceasta, se va lua în considerație și probabilitatea de tranziție  $P(j/i)$ : probabilitatea ca după litera  $\alpha_i$  să urmeze litera  $\alpha_j$ . O altă metodă echivalentă pentru redarea acestei structuri constă în stabilirea probabilităților diagramelor  $P(i, j)$ :

$$P(i, j) = P(i) \cdot P(j/i).$$



Calculul entropiei medii pe litera  $H_2$  va fi dată de relația:

$$H_2 = - \sum_{i,j} P(i,j) \cdot \log_2 P(j/i) \text{ biți/literă}$$

Ilustrarea modului de calcul al probabilităților diagramelor poate fi făcută cu exemplu de mai jos.

Presupunem că alfabetul conține trei litere A, B și C, cu următoarele probabilități:

$\sum P(i) = 1$		$P(j/i)$			
I	P(I)	I	J		
A	0,33	A	A	B	C
B	0,60	A	0	0,8	0,2
C	0,07	B	0,5	0,5	0
		C	0,5	0,4	0,1

$$P(A,A) = P(A) \cdot P(A/A) = 0,33 \cdot 0 = 0$$

$$P(A,B) = P(A) \cdot P(B/A) = 0,33 \cdot 0,8 = 0,264$$

$$P(A,C) = P(A) \cdot P(C/A) = 0,33 \cdot 0,2 = 0,066$$

$$P(B,A) = P(B) \cdot P(A/B) = 0,6 \cdot 0,5 = 0,3$$

$$P(B,B) = P(B) \cdot P(B/B) = 0,6 \cdot 0,5 = 0,3$$

$$P(B,C) = P(B) \cdot P(C/B) = 0,6 \cdot 0 = 0$$

$$P(C,A) = P(C) \cdot P(A/C) = 0,07 \cdot 0,5 = 0,035$$

$$P(C,B) = P(C) \cdot P(B/C) = 0,07 \cdot 0,4 = 0,028$$

$$P(C,C) = P(C) \cdot P(C/C) = 0,07 \cdot 0,1 = 0,007$$

$$H_2 = -0,264 \log_2 0,8 - 0,066 \log_2 0,2 - 2 \cdot 0,3 \log_2 0,5 - 0,035 \log_2 0,5 - 0,021 \log_2 0,4 - 0,007 \log_2 0,1 = 3,3 \text{ biți}$$

A fost calculată  $H_2$  pentru diferite limbi; astfel, pentru limba engleză, s-a găsit  $H_2 = 3,3$  biți/literă, iar pentru limba rusă,  $H_2 = 3,51$  biți/literă.

Calculul frecvențelor de apariție a combinațiilor de trei litere și mai multe este necesară pentru descrierea statistică a unei limbi. Deși foarte complexe, calculele referitoare la stabilirea frecvenței trigramelor au fost efectuate pentru mai multe limbi europene, calculându-se entropia de ordinul 3,  $H_3$ . Aceasta este 2,98 biți/literă pentru limba engleză, 3 biți/literă pentru limba rusă.

Calculul frecvențelor de apariție a combinațiilor de 4, 5 sau 6 litere este foarte complicat și până acum nu s-a efectuat analitic pentru nicio limbă. S-au făcut însă unele determinări experimentale, care arată că pentru texte mai lungi care cuprind 100 de litere și mai mult entropia scade în jurul valorii de 1 bit/literă.

Studierea limbii ca mesaj conduce la necesitatea calculării cantității de informație pe care el o conține, iar măsura cantității de informație o constituie entropia limbii. O valoare mare a acesteia denotă o anumită claritate a limbii. Din considerentele de mai sus se vede că eficacitatea cu care limba (ca sursă) produce informație depinde de distribuția probabilităților de apariție a literelor care compun comunicarea; cu cât distribuția este mai uniformă, cu atât cantitatea de informație adusă de literă este mai mare. Raportul dintre entropia medie pe literă și entropia maximă,  $H_{\max} = \log_2 n$  ( $n$  fiind numărul de litere utilizat), se numește entropie relativă și exprimă gradul de compresie al comunicării în funcție de structura ei statistică.

Cantitatea  $R = 1 - \frac{H}{H_{\max}}$  se numește *redundanță* și reprezintă un indicator util în cazul mesajelor lungi.

Până nu demult frecvența de apariție a literelor se determina manual, solicitând o muncă laborioasă în special pentru texte lungi.

Dacă toate combinațiile de litere ar forma cuvinte, atunci într-un alfabet de 30 de litere ar exista:

30 de cuvinte de o literă;  
 $30^2$  cuvinte de două litere;  
 .....  
 $30^n$  cuvinte formate din  $n$  litere.

În realitate, în fiecare limbă vorbită, numărul de cuvinte este mai mic de 50000, deci numai o parte foarte mică dintre combinații formează cuvinte, restul fiind combinații fără sens.

Fie un mesaj  $M$  de lungime  $N$  și să presupunem că numărul literelor distincte din  $M$  este egal cu  $n$  ( $n \leq N$ ). Numărul tuturor literelor din alfabetul limbii în care este redactat textul este  $L$ . Dacă la mesajul  $M$  atașăm pe rând câte o literă din alfabet, obținem un număr de  $L$  texte de lungime  $(N+1)$  din care:  $(L-n)$  vor avea  $(n+1)$  litere distincte, restul de  $n$  texte vor rămâne cu  $n$  litere distincte.

Deci, în medie, numărul literelor distincte pentru fiecare din cele  $L$  texte este:

$$\frac{(L-n)(n+1)+n^2}{L} = \frac{n(L-1)+L}{L}.$$

Dacă notăm cu  $\Delta_n$  variația numărului de litere distincte corespunzătoare unei variații  $\Delta N$  a lungimii textului, avem:

$$\Delta_n = \left[ \frac{n(L-1)+L}{L} - n \right] \cdot \Delta N = \frac{L-n}{L} \cdot \Delta n.$$

Această relație poate fi asimilată cu o ecuație diferențială cu variabile separabile, adică:

$$dn = \frac{L-n}{L} dN$$

Prin integrare, se obține:

$$n = L \left( 1 - e^{-\frac{N}{L}} \right)$$

$$0 = L - e^{-k}; \quad k = -\ln L.$$

Constanta  $k$  se determină din condițiile la limită: pentru  $N=0$  rezultă  $n=0$ .

Corelând această relație cu rezultatele obținute printr-o cercetare statistică, rezultă că relația finală este de forma:

$$n = L_0 \left( 1 - e^{-\frac{N}{L}} \right)$$

În această relație,  $L_0$  reprezintă numărul literelor din alfabet cu o frecvență de apariție mai mare decât 1%. În tabelul următor se dau valorile indicatorului „numărul literelor distincte” ( $n$ ) pentru texte de diferite lungimi calculate cu relația de mai sus pentru diferite limbi:

Limba	$L_0$	$N = 30$	$N = 40$	$N = 50$	$N = 100$
Italiană	18	12,69	13,98	15,57	17,96
Română	18	12,69	13,98	15,57	17,96
Germană	21	14,8	16,31	18,16	20,95
Engleză	21	14,8	16,31	18,16	20,95
Franceză	19	13,39	14,76	16,43	18,98

Din analiza acestor date rezultă că pentru texte având mai mult de 100 de litere intervine o anumită stabilitate, în sensul că în astfel de texte vor apărea, în medie, toate cele  $L_0$  litere din alfabet și deci nu mai are rost calculul numărului de litere distincte cu această relație.

În tabelul de mai jos se prezintă frecvențele relative de apariție a literelor în câteva limbi europene, determinate la un text de 1000 de litere.

Engleză		Rusă		Română		
litera	frecvența	litera	frecvența	litera	text ziar	text literă
e	0,131		0,090	a	0,105	0,093
t	0,105		0,072	b	0,0092	0,0092
a	0,086		0,062	c	0,053	0,054
o	0,08		0,053	d	0,04	0,039
n	0,071		0,045	e	0,12	0,1
r	0,068		0,040	f	0,008	0,009
i	0,063		0,038	g	0,006	0,008
s	0,061		0,035	h	0,007	0,005
h	0,053		0,028	i	0,13	0,11
d	0,038		0,024	j	0,0014	0,002
l	0,034		0,025	k	—	—
f	0,029		0,023	l	0,046	0,046
c	0,028		0,021	m	0,039	0,04
m, u	0,025		0,018	n	0,056	0,064
p, y, q	0,020		0,016	o	0,025	0,04
w	0,015		0,014	p	0,026	0,031
b	0,014		0,013	q	—	—
v	0,009		0,012	r	0,062	0,07
k	0,004		0,010	s	0,024	0,046
x	0,002		0,009	t	0,065	0,053
J, g, z	0,001		0,007	u	0,074	0,06
			0,006	v	0,012	0,013
			0,006	w	—	—
			0,003	x	—	—
			0,0002	y	—	—
				z	0,008	0,008
				ț	0,006	0,001
				ș	0,17	0,02

### Elemente de criptanaliză

Un alt indicator util în criptanaliză este raportul dintre numărul vocalelor și cel al consoanelor. S-au făcut cercetări statistice pe această temă și în tabelul următor sunt sintetizate unele rezultate:

Limba %	Română	Franceză	Italiană	Germană	Engleză
vocale	49,4	43,36	47,74	38,86	37,4
consoane	50,6	56,64	52,26	61,14	62,6

Rangul pe care îl ocupă literele în cadrul cuvântului constituie un indicator prețios în criptanaliză. Cercetarea literelor inițiale și finale ale cuvintelor oferă o serie de rezultate specifice dat fiind că parte inițială din cuvânt deține un loc principal în privința cantității de informație conținută în cuvântul respectiv. În tabelul de mai jos sunt date, în ordine descrescătoare a frecvențelor de apariție, cele mai des întâlnite litere inițiale și finale:

Limba	Rangul literei	Litere
română	inițială	SCPAIDMNTUVL...
	finală	EAITRUMD...
germană	inițială	DASIEWV...
	finală	NERSTHDI...
engleză	inițială	TASIWCFRDB...
	finală	ESTNDYOR
italiană	inițială	ADOVRCUSP...
	finală	EAOILN
franceză	inițială	LEDASCPV
	finală	?

Cuvintele scurte cele mai frecvente constituie, de asemenea, un indicator prețios pentru criptanaliști. În tabelul de mai jos sunt prezentate cele mai frecvente cuvinte scurte din câteva limbi europene:

Limba	Cuvinte scurte
Franceză	A, DE, DU, EST, ET, LA, LE, LES, PAR, QUE...
Germană	AN, DASS, DEN, DER, DIE, ES, IN, IST, UND...
Engleză	A, AND, AT, FOR, HAS, IN, IT, IS, OF, THE...
Italiană	A, CHE, CON, DI, E, HA, HO, IN, PER

Cercetările statistice efectuate asupra cuvintelor au condus la rezultate deosebit de interesante atunci când s-a adoptat drept criteriu numărul silabelor componente. S-a stabilit că, indiferent de limba folosită, distribuția de probabilitate a cuvintelor după numărul silabelor componente este o distribuție de tip Poisson.

Astfel, dacă  $P(n)$  este probabilitatea de apariție a unui cuvânt format din  $n$  silabe și cu  $I$  lungimea medie a cuvintelor, în silabe, într-o anumită limbă avem:

$$I = 1 \cdot P(1) + 2 \cdot P + \dots + n \cdot P(n)$$

$$P(n) = \frac{a^{n-1} \cdot e^{-a}}{(n-1)!}, \quad \text{unde } a = I - 1$$

Luând în considerare apropierea limbă-cod (limba fiind considerată ca un cod limitat probabilistic), se poate calcula cantitatea de energie informațională conținută într-un mesaj oarecare.

Energia informațională este o noțiune introdusă de matematicianul român Octav Onicescu și se definește astfel: fie  $A$  un sistem oarecare și  $A_1, A_2, \dots, A_n$

diferitele sale stări. Notând cu  $P_1, P_2, \dots, P_n$  probabilitățile stărilor  $\left( \sum_{i=1}^n P_i \right)$ ,

mărimea  $I(A)$  calculată cu relația:  $I(A) = P_1^2 + P_2^2 + \dots + P_n^2 = \sum_{i=1}^n P_i^2$  se numește

energie informațională. Acest indicator reflectă starea de organizare a sistemului.

În cazul în care  $P_1 = P_2 = \dots = P_n = \frac{1}{n}$ , energia informațională este minimă:

$$I(A) = \frac{1}{n}.$$

Deoarece probabilitățile  $P_i$  nu se cunosc, ele vor fi estimate cu ajutorul frecvențelor relative, astfel că expresia energiei informaționale devine:

$$I(A) = \sum_{i=1}^n \left( \frac{n_i}{N} \right)^2 = \frac{1}{N^2} \sum_{i=1}^n n_i^2,$$

unde  $n_i$  reprezintă frecvența absolută a stării  $A_i$ , iar  $N = \sum_{i=1}^n n_i$ .

Cea mai bună aproximație pentru  $I(A)$  se obține folosind pentru probabilitățile  $P_i$  o estimare eficientă și absolut corectă și, în acest caz, expresia energiei informaționale devine:

$$I(A) = \sum_{i=1}^n \frac{n_i(n_i-1)}{N(n-1)}.$$

Cercetările statistice au determinat probabilitățile de apariție a cuvintelor în funcție de lungimea lor. Pentru limba română, aceste probabilități sunt date în tabelul de mai jos:

Lungimea	Probabilitățile de apariție
1	0,0206
2	0,2723
3	0,1066
4	0,1396
5	0,1146
6	0,0966
7	0,0990
8	0,0603
9	0,0505
10	0,0173
11	0,0152
12	0,0063
13	0,0033
14	0,0006

Se observă că, având probabilitatea de apariție de 0,2723, cuvintele de două litere sunt cele mai frecvente. Lungimea medie a cuvântului se poate calcula cu relația:

$$l_{med} = \sum_{i=1}^n P_i l_i,$$

$n$  fiind cea mai mare lungime a unui cuvânt din limba română. S-a obținut lungimea medie  $l_{med} = 4,6$  litere/cuvânt.

Dacă se ia  $n = 14$  și dacă toate cuvintele ar fi egal probabile, atunci:

$$l_{med} = \frac{1}{14} \sum_{i=1}^{14} l_i = 7,5 \text{ litere.}$$

Pentru un limbaj natural  $L$  și un text  $T$  de lungime  $n$ , al cărui vocabular conține  $N$  cuvinte, avem:

$$F_{\gamma} \cdot \gamma = k(L, n); \quad \gamma = 1, 2, \dots, N,$$

lege cunoscută sub numele de legea ESTOUP-ZIPP și în care  $F_{\gamma}$  reprezintă frecvența cuvântului  $\gamma$ , iar  $k$  este o constantă care diferă în funcție de limbă și de lungimea textului  $n$ . Există și o extindere naturală a acestei legi:

$$F_{\gamma} \cdot \gamma^a = k a_0,$$

unde  $a_0$  este o constantă reală pozitivă caracteristică limbajului. Pentru calculul lui  $a_0$  și a constantei  $k_{a_0}$ , se poate folosi relația:

$$k_a(L) = \frac{1}{N} \sum_{\gamma=1}^N F_{\gamma} \cdot \gamma^a.$$

Pentru limba română s-a calculat  $a_0 = 0,8$  și  $k_{a_0} = 655$ .

O altă lege utilă criptanaliștilor este legea rang-frecvență: într-o listă de frecvență, produsul dintre rangul unui cuvânt și frecvența corespunzătoare este constantă.

Prin *rang* se înțelege numărul de ordine al unui cuvânt în cadrul unei liste de frecvențe. *Lista de frecvențe* este un vocabular special, în care cuvintele sunt așezate în ordinea descrescătoare a frecvențelor cu care apar în text. *Vocabularul* unui text cuprinde numărul total de cuvinte diferite din text.

Există și legea MANDELBROT:

$$F_{\gamma} = P \cdot L \cdot (\gamma + a)^{-b},$$

în care:  $L$  este lungimea textului în cuvinte;  $a, b$  – doi parametri caracteristici ai limbii;  $p$  este o constantă care se calculează cu relația:

$$\frac{1}{P} = \frac{1}{(1-a)^b} + \frac{1}{(2-a)^b} + \dots + \frac{1}{(N-a)^b}.$$

Un indicator interesant pentru limbajele naturale legat de numărul și lungimea șirurilor repetate (secvențe) care pot să apară și care este implicit legat de transformările la care sunt supuse textele prin aplicarea unor prelucrări criptografice a fost introdus de ELLIOT FISCHER. Fie  $A$  un alfabet finit și fie  $L(s)$  lungimea unui șir  $S$  din  $A_1$ , unde  $A_1$  reprezintă mulțimea tuturor șirurilor de lungime finită din  $A$  care se repetă.



Notăm cu  $S(i, j)$  un subșir al lui  $S$  care începe cu poziția  $i$  și se termină cu poziția  $j$ . Un șir  $R$  obținut prin concatenarea a două șiruri  $S$  și  $Q$  îl notăm:  $R = SQ$ .

Spunem că  $R$  este reproductibil din  $S$  dacă  $Q$  este un subșir al lui  $S$ . Spunem că  $R$  este productibil din  $S$  dacă șirul  $RQ_1$  format prin ștergerea ultimului simbol din  $Q$  este reproductibil din  $S$ . Deci reproductibilitatea presupune *copierea*, iar productibilitatea permite apariția unui simbol nou la sfârșitul operației de copiere.

Dându-se un șir  $S$  din  $A$ , îl putem considera fie ca rezultat al unei concatenări, fie ca rezultat al unui proces de producere în care fiecare nouă secvență este productibilă din secvența deja existentă.

## 6.2 Metode de decriptare

Prin decriptarea unui sistem de cifrare  $S$  se înțelege evidențierea textelor clare pe baza analizelor criptogramelor realizate prin intermediul lui  $S$ , fără a avea cunoștințe despre *sistemul utilizat* și *cheia folosită*. Cifrarea are ca efect perturbarea unor caracteristici statistice ale limbajelor de redactare a mesajelor, perturbări mai accentuate sau mai puțin accentuate în funcție de complexitatea sistemului respectiv.

După ce volumul textului interceptat a trecut de *punctul de unicitate*, criptograma are o rezolvare unică. Textul cifrat poate fi decriptat prin încercarea succesivă a tuturor cheilor posibile. Această metodă, numită *metoda sistemului complet de încercări*, nu este însă practică. De exemplu, dacă o cheie are  $26!$  posibilități ( $\cong 2 \cdot 10^{26}$ ) care apar în cazul unei simple substituții și care se consideră a avea un volum mic.

Dacă adversarul (criptanalistul) are un calculator pentru încercarea cheilor cu o viteză de 1 cheie/microsecundă și dacă el va găsi cheia adevărată făcând jumătate din încercările posibile, ar avea nevoie de

$$\frac{2 \cdot 10^{26}}{2 \cdot 60^2 \cdot 24 \cdot 365 \cdot 10^6} \cong 3 \cdot 10^2 \text{ ani}.$$

Deci metoda sistemului complet de încercări nu poate fi făcută în practică. În analiza criptografică se folosește metoda încercărilor de o factură specială, încercările făcându-se în ordinea descrescătoare a probabilităților de utilizare a cheilor; mai mult decât atât, încercările se referă la o grupă cât mai mare de chei și nu la o cheie anume.

În acest caz mulțimea cheilor se împarte în submulțimi care conțin aproximativ același număr de chei. Cu ajutorul unor încercări se determină submulțimea cheilor din care face parte cheia utilizată.

În cazul de mai sus numărul de încercări s-ar reduce la  $26 \times 5 = 130$  de încercări.

Deci în timp ce metoda sistemului complet de încercări face ca numărul acestora să fie egal cu numărul cheilor, metoda împărțirii în subgrupe reduce numărul încercărilor și-l face să fie egal cu volumul cheii exprimat în biți.

Deci procesul de reducere a numărului de încercări constă în împărțirea mulțimii cheilor în submulțimi echiprobabile, iar numărul mediu de încercări va fi:

$$h = \frac{H(c)}{\log 2}.$$

Dacă fiecare încercare are  $S$  rezultate posibile, iar fiecare rezultat corespunde posibilității găsirii cheii într-una dintre  $S$  grupe echiprobabile, atunci numărul mediu de încercări va fi:

$$h = \frac{H(c)}{\log s}.$$

Dacă  $S = 2$ , fiecare încercare aduce 1 bit de informație față de  $\frac{1}{1,25}$  biți în cazul sistemului complet de încercări.

În analiza criptografică repetările din textul criptat care pot fi cauzale sau accidentale au un rol însemnat.

Repetările cauzale reprezintă repetările cifrării textului clar care a trecut prin prelucrarea criptografică.

Repetările accidentale sunt cele care prin circumstanțe neprevăzute provin din cifrările diferitelor elemente ale textului clar. Găsirea repetărilor de diferite lungimi (4, 5 sau mai multe caractere) duce, de obicei, la rezolvarea criptogramelor.

Alteori, dacă repetările de diferite lungimi sunt insuficiente, se pune problema dacă ele sunt cauzale sau accidentale, evaluându-se semnificația criptografică a elementelor cifrate repetate.

În analiza criptografică prin text aleator se înțelege textul în care elementele cifrate vor apărea aproximativ cu aceeași probabilitate.

Pentru exemplificare, presupunem că avem un text aleator de  $N$  elemente ale unui sistem secret cu  $n$  elemente diferite (de exemplu,  $N = 50$  litere ale unui

alfabet care conține  $n = 26$  litere); probabilitatea de apariție a unui anume element va fi  $p_i = \frac{1}{n} = \frac{1}{26}$ .

În mod similar, dacă dispunem de un text cu  $N = 376$  diagrame din cele 676 posibile, probabilitatea de apariție a unei anumite diagrame este  $p_d = \frac{1}{676}$ .

Dar nu toate cele  $n$  elemente posibile vor apărea în textul de  $N$  elemente. Numărul de elemente care nu apar este un indicator semnificativ în criptanaliză. Pentru a putea fi utilizat, este necesar să se cunoască distribuția teoretică a elementelor care nu apar (blancuri). Probabilitatea că există un număr de exact  $r$  blancuri este:

$$P(r) = \frac{n!}{r!} \sum_{i=0}^{n-r} (-1)^i \frac{1}{n!(n-r-i)!} \left(1 - \frac{r+i}{n}\right)^N.$$

S-au tabulat valorile lui  $P(r)$  pentru  $N = n = 10$  (deci cazul cifrelor).

$r$	$P(r)$
0	0,00036288
1	0,0163296
2	0,13608
3	0,3556224
4	0,34514424
5	0,1285956
6	0,01718892
7	0,00067176
8	0,000004549
9	0,000000001

Numărul probabil al blancurilor dintr-un text aleator cu  $N$  elemente, într-un sistem în care există  $n$  caractere diferite, este dat de relația:

$$B_N = n \left(1 - \frac{1}{n}\right)^N.$$

S-a determinat numărul mediu de blancuri pe mulțimi de  $n$  elemente ale unui text aleator format din  $N$  elemente. Rezultatele sunt prezentate în tabelul de mai jos, pentru cazul  $N = 200$  și  $n = 10$ .

$r$	$200 P(r)$	Frecvența ( $f$ )	$r \times f$
0	0,08	0	0
1	3,26	8	8
2	27,22	22	44
3	71,12	72	216
4	69,02	72	288
5	25,72	21	105
6	3,44	4	24
7	0,14	1	7
8	0	0	0
9	0	0	0
	200	200	692

Deci numărul mediu al blancurilor pe submulțimi de 10 simboluri distincte este:

$$\frac{692}{200} = 3,46$$

Pentru un text aleator de  $N = 100$  diagrame din cele  $n = 676$  posibile, se poate calcula numărul estimat de blancuri:

$$B = 676 \left( 1 - \frac{1}{676} \right)^{100} \cong 582$$

Deci textul va conține 94 de diagrame distincte.

Prin *text nealeator* se înțelege un text în care elementele sunt bine stabilite în concordanță cu prelucrarea lor criptografică. Deci din punct de vedere statistic, textul clar și textul nealeator sunt identice. Dacă  $n$  elemente posibile ale unui sistem au probabilitățile de apariție  $P_1, P_2, \dots, P_n$ , atunci media blancurilor într-un text compus din  $N$  elemente:

$$B_N = (1 - P_1)^N + (1 - P_2)^N + \dots + (1 - P_N)^N,$$

sau cu o bună aproximație:

$$B_N = \sum_{i=1}^n e^{-NP_i}.$$

Pentru un text în limba engleză cifrat monoalfabetic, numărul blancurilor are valoarea din tabelul de mai jos:

## Elemente de criptanaliză

$n$	Numărul blancurilor	$n$	Numărul blancurilor
10	18,50	110	5,64
20	14,13	120	5,46
30	11,55	130	5,21
40	10,03	140	5,04
50	8,84	150	4,88
60	7,98	160	4,78
70	7,33	170	4,67
80	6,74	180	4,56
90	6,29	190	4,44
100	5,83	200	4,4

Pentru texte cu caracter nealeator, în analizele criptografice se utilizează așa-numitul test.

Să considerăm un text având  $N$  caractere dintr-un sistem cu  $n$  elemente posibile. Presupunem că din cele  $n$  elemente sunt posibile  $f_1, f_2, \dots, f_N$  elemente, astfel încât  $f_1 + f_2 + \dots + f_N = N$ . Deci elementul  $\alpha_1$  apare de  $f_1$  ori, elementul  $\alpha_2$  apare de  $f_2$  ori și așa mai departe.

Dacă definim mărimea  $\Phi$  cu ajutorul relației:

$$\Phi = f_1(f_1 - 1) + f_2(f_2 - 1) + \dots + f_n(f_n - 1),$$

atunci se poate calcula valoarea medie a acestuia ca fiind  $E(\Phi)$ :

$$E(\Phi) = S_2 N(N - 1),$$

unde  $S_2$  reprezintă suma pătratelor probabilităților de apariție a fiecărui element din cele  $n$  elemente posibile din sistem. Pentru un sistem aleator:

$$S_2 = \sum_{i=1}^n \frac{1}{n^2} = \frac{1}{n}.$$

Pentru un text aleator în care gruparea literelor a fost monografică (caracter cu caracter), digrafică (grupuri de două caractere cifrate odată) și trigrafice (grupuri de trei caractere cifrate odată) s-a calculat  $S_2$ , obținându-se valorile următoare:

- grupare monografică,  $S_2 = 0,038$ ;
- grupare digrafică  $S_2 = 0,0015$ ;

- grupare trigrafică  $S_2 = 0,000057$ .

Pentru un text clar și pentru mai multe limbi s-au obținut următoarele valori:

Limba	$E(\Phi)$	
	text monografic	text digrafic
Engleză	$0,0661 N(N-1)$	$0,0069 N(N-1)$
Franceză	$0,0778 N(N-1)$	$0,0093 N(N-1)$
Rusă	$0,0529 N(N-1)$	$0,0058 N(N-1)$
Germană	$0,0762 N(N-1)$	$0,0112 N(N-1)$
Română	$0,08595 N(N-1)$	

Având la dispoziție aceste date statistice, se poate stabili din ce limbă face parte următorul text și ce fel de cifrare s-a folosit:

IBMQO PBIUO MBBGA JCZOF MUUQB

Textul conținând 25 de caractere are următoarea distribuție:

A B C D E F G H I J K L M N O P Q U Z R S T V W X Y  
 1 5 1 0 0 1 1 0 2 1 0 0 3 0 3 1 2 3 1 0 0 0 0 0 0 0

Valoarea testului  $\Phi$  va fi:

$$\Phi = 1(1-1) + 5(5-1) + 1(1-1) + 1(1-1) + 1(1-1) + 2(2-1) + 1(1-1) + 3(3-1) + 3(3-1) + 1(1-1) + 2(2-1) + 3(3-1) + 1(1-1) = 42$$

În limba engleză  $E(\Phi) = 0,0661 \times 25 \times 24 = 39,6$  pentru cifrare monografică. Pentru un text aleator:  $E(\Phi) = 0,038 \times 25 \times 24 = 22,8$ .

Deoarece 39,6 este mult mai apropiat de 42 decât 22,8, se poate spune că mai curând este un text în engleză decât un text întâmplător.

Dacă se iau în considerare pătratele numerelor care reprezintă frecvențele de apariție ale literelor:

$$\Psi = f_1^2 + f_2^2 + \dots + f_n^2,$$

atunci  $\Psi = S_2 N^2 + (1 - S_2) N$ , care reprezintă testul  $\Psi$  pentru diagrame.

Multe tipuri de cifruri pot fi sparte cu ajutorul analizei statistice, mai ales când textul interceptat are mai mult de 200 de caractere.

O metodă care poate fi utilizată în descifrare este *metoda cuvintelor probabile*. Cuvintele probabile sunt acele cuvinte sau expresii care pot fi întâlnite mai ales într-un mesaj ca urmare a faptului că sunt caracteristice pentru sursa de mesaje respectivă.

Cuvintele probabile pot fi considerate comune sau silabele care se întâlnesc mai frecvent într-o limbă; de exemplu: are, cu, și, ile, lor etc. pentru limba română.

Metoda se poate utiliza astfel: presupunând că o anumită parte a criptogramei reprezintă un anumit cuvânt al mesajului, se găsește o parte a cheii. Această parte se folosește pentru descifrarea celorlalte părți ale criptogramei. Metoda cuvintelor probabile s-a dovedit destul de eficace, în special, pentru mesaje destul de lungi. Sunt puține cifruri clasice care având un volum mic al cheii rămân mult timp nedecryptate prin metoda cuvintelor probabile. De aceea, această metodă se folosește pentru verificarea calității cifrurilor. În practică, metoda se utilizează împreună cu alte metode.

### 6.3 Spargerea sistemelor criptografice

Orice sistem criptografic conține două elemente esențiale:

- procedeul general de cifrare;
- cheia specifică utilizată.

De exemplu, cifrul lui Caesar este un procedeu general monoalfabetic, care operează asupra unui alfabet standard și în care cheia specifică este cheia 3. Deci relația de cifrare ar fi:

$$c_i = m_i + 3,$$

unde  $c_i$  este valoarea numerică a literei criptogramei, nu valoarea numerică a literei din mesajul clar  $m_i$ .

Presupunând că nu este cunoscută cheia:

$$c_i = m_i + a \text{ (modulo 26)},$$

unde  $m_i \in \{A, B, C, \dots, Z\}$ , iar  $a$  trebuie determinată. Pentru un criptanalist această problemă nu este dificilă, el neavând de efectuat decât 25 de încercări. Este posibil ca numai prin efectuarea deplasărilor succesive ale literelor unui singur cuvânt din textul cifrat să se afle cheia specifică.

De exemplu, fie cuvântul din criptograma:

**K I T K C T C T C Q**

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

Deci după 8 deplasări ale literelor s-a obținut un cuvânt cu sens în limba română. Așadar cheia de cifrare este  $-8$ , iar relația de cifrare este  $c_i = m_i + 8$  (modulo 26).

În acest caz s-a presupus că se cunoaște procedeul general de cifrare și s-a reușit stabilirea cheii specifice. Mergând de la simplu la complex, se presupune că există situația când trebuie determinat și procedeul de cifrare prin metode criptanalitice. Determinarea procedeului general de cifrare se poate face cu ajutorul frecvenței relative de apariție a diferitelor litere din alfabet. În acest scop se consideră un eșantion de text clar și se determină frecvența de apariție a literelor, folosind în acest sens relația:

$$f_i = \frac{n_i}{N},$$

unde  $N$  este numărul de litere pe care îl conține textul clar, iar  $n_i$  reprezintă numărul de apariții a literei  $\alpha_i$ . Se determină apoi frecvențele de apariție a literelor din textul criptat și, prin comparație, se pot stabili anumite corespondențe între literele din textul criptat și literele alfabetului. Pentru exemplificare, să analizăm o criptogramă obținută prin transformări de tipul:  $c = am + b$ , unde perechea  $a, b$  reprezintă cheia specifică. Această transformare nu mai realizează o simplă translație a literelor din alfabet, ci o substituție oarecare, iar constanta  $a$  nu trebuie să fie divizor al lui 26.

Fie criptograma:

**JAHGZTGTH JFJMTRTOXG NGFAMXZGHWENT TJMT X  
XATGHMFT QFWFNFOH NHGT JXOFNFMH RPOMH MTUHNFMHMT**



### Elemente de criptanaliză

Lungimea criptogramei este 80, iar frecvențele de apariție a literelor sunt:

Litera	Frecvența	Litera	Frecvența
A	3	P	1
F	10	Q	1
G	7	R	2
H	10	T	12
J	5	U	1
M	9	W	2
N	6	X	5
O	4	Z	2

Comparând aceste frecvențe cu frecvențele literelor din limba română, se pot face unele corespondențe:

- $T_c \rightarrow E_m$  ( $T$  din criptogramă corespunde lui  $E$ );
- $F_c \rightarrow J_m$  ( $F$  din criptogramă corespunde lui  $I$ ).

Pe baza acestor corespondențe se poate scrie sistemul:

$$\begin{cases} 19 \equiv 4a + b \\ 5 \equiv 8a + b \end{cases} \quad (\text{modulo } 26)$$

Scăzând cele două ecuații între ele, se obține:

$$-4a \equiv 14 \quad (\text{modulo } 26)$$

$$22a \equiv 14 \quad (\text{modulo } 26)$$

$$11a_0 \equiv 7 \quad (\text{modulo } 13)$$

$$a_0 \equiv 7 \cdot 11^{11} \quad (\text{modulo } 13)$$

$$a_0 = 3 \text{ și } a_1 = 16$$

Pentru  $a_0 = 3$  și  $a_1 = 16$  se obține:

$$b = 5 - 11 = 5 + 15 = 20$$

Deci s-au obținut două soluții:

$$c = 3m + 7 \quad (\text{modulo } 26)$$

$$c = 16m + 20 \quad (\text{modulo } 26)$$

Aplicând, pe rând, cele două transformări, numai prima soluție duce la obținerea unui text cu înțeles în limba română:

**SPARGEREA SISTEMELOR CRİPTOGRAFICE ESTE O OPERAȚIE DIFICILĂ CARE SOLICITĂ MULTĂ TENACITATE**

Pentru descifrare s-a folosit relația:

$$m = 3^{-1}(c - 7) \pmod{26}$$

Dar

$$3^{-1} \pmod{26} = 9 \pmod{26}$$

$$-7 \pmod{26} = 19 \pmod{26}$$

Deci  $m = 9c + 15 \pmod{26}$ .

A doua soluție nu este posibilă, deoarece  $(16, 26) = 2$  și deci 16 nu are invers în mulțimea claselor de resturi modulo 26.

Să analizăm un alt exemplu în care s-au folosit substituții simple cu reprezentări unice și uniforme.

Fie criptograma:

33. 15. 44. 35. 14. 15. 32 .15./ 13. 42. 24. 41. 44. 35. 22. 42. 11. 21. 24.  
13. 15./ 44. 42. 15. 12. 45. 24. 15./ 43. 44. 45. 14. 24. 11. 44. 15./ 13. 11./ 45.  
34./ 13. 11. 41. 24. 44. 35. 32./ 43. 41. 15. 13. 24. 11. 31./ 11. 32./ 43. 13. 42. 24.  
43. 45. 32. 45. 24./ 32. 11./ 34. 35. 24./

Pentru început, facem următoarele ipoteze:

- mesajul clar este redactat în limba română;
- criptograma respectă împărțirea pe cuvinte a mesajului, fiecare literă având ca reprezentare cifrantă un grup de două cifre;
- nu sunt folosite semne de punctuație.

Vom căuta acum să vedem în ce măsură ipotezele de mai sus se verifică.

În primul rând, lungimea medie a cuvintelor este:

$$l_{med} = \frac{2 \cdot 8 + 1 \cdot 13 + 3 \cdot 7 + 4 \cdot 2 + 1 \cdot 3 + 1 \cdot 9}{13} = 5,83 \text{ litere}$$

Valoarea de 5,83 este foarte apropiată de lungimea medie a cuvintelor din limba română.

Frecvențele de apariție a reprezentărilor cifrante sunt prezentate în tabelul următor:

### Elemente de criptanaliză

Reprezentare cifrantă	11	12	13	14	15	21	22	23	24	25
Frecvență	7	1	6	2	8	1	1	–	9	–
Reprezentare cifrantă	31	32	33	34	35	41	42	43	44	45
Frecvență	–	6	1	2	4	3	4	4	6	5

Criptograma conține în total 70 de litere, din care 17 distincte. Pentru un text în limba română, de lungime  $N = 70$ , se obține:

$$l_0 = 26 \left( 1 - e^{-\frac{70}{26}} \right) = 16,80.$$

Deci ipoteza că avem de-a face cu o substituție de litere ca reprezentări unice de lungime 2 devine plauzibilă. Reprezentarea cifrantă 24 este cea mai frecventă, apărând de 9 ori și s-ar putea face corespondența  $24 \rightarrow A$  ( $I$  sau  $E$ ).

Așezând reprezentările cifrante cele mai frecvente în ordinea descrescătoare a numărului de apariție, s-ar putea face următoarele corespondențe:

Reprezentare cifrantă	24	15	11	13	44	32
literă	A	E	I	R	T	N

Înainte de a înlocui aceste posibile echivalențe în criptogramă, facem următoarele observații:

- cea mai frecventă reprezentare din finalul cuvintelor este 15, de aici corespondența  $15 \rightarrow E$ ;

- diagramele 11.32 și 32.11 cele mai plauzibile ar fi: AL-LA sau NU-UN, dar în corelație cu frecvențele de apariție a literelor decidem  $11 \rightarrow A$ ;  $32 \rightarrow L$ .

Deci în tabelul de mai sus operăm două modificări:

Reprezentare cifrantă	24	15	11	13	44	32
literă	A	E	I	R	T	N

- reprezentările cele mai frecvente ca literă inițială sunt 43 și 13, iar în limba română C și S, deci putem avea  $S \rightarrow 13$  ( $43$ )  $C \rightarrow 43$  ( $13$ ).

Dar există un cuvânt care începe cu 43.13 care nu poate fi CS... ci SC... Așadar  $43 \rightarrow S$  și  $13 \rightarrow C$ . Răspândind în criptogramă echivalențele probabile stabilite, se obține în final mesajul:

**METODELE CRİPTOGRAFICE TREBUIE STUDIMATE CA UN  
CAPITOL SPECIAL AL SCRISULUI LA NOI**

iar tabelul de substituție este:

V	1	2	3	4	5
0					
1	A	B	C	D	E
2	F	G	H	I	J
3	K	L	M	N	O
4	P	R	S	T	U
5	V	W	X	Y	Z

Literele care nu apar în text s-au trecut în tabel pe baza observației că primă linie formată din litere care apar în criptogramă a fost scrisă în ordine alfabetică normală.

În exemplul de mai sus am folosit faptul că criptograma a respectat împărțirea reală pe cuvinte a mesajului. Acest lucru nu este însă esențial și o criptogramă poate fi descifrată și în cazul în care nu sunt puse în evidență spațiile dintre cuvinte.

Vom analiza, în continuare, tehnicile folosite pentru spargerea sistemelor criptografice polialfabetice, la care lungimea cheii este mai mare decât 1.

Decriptarea acestor sisteme comportă două etape:

- determinarea lungimii cheii de cifrare;
- determinarea efectivă a alfabetelor folosite și apoi descoperirea textului clar.

În cifrarea polialfabetică, cu cât numărul de alfabete este mai mare, cu atât distribuția reprezentărilor cifrante este mai uniformă.

Dacă mesajul este scurt sau dacă numărul alfabetelor folosite la cifrare este mic, atunci examinarea distribuției frecvențelor poate fi total neconcludentă. Neuniformitatea distribuției variabilei aleatoare implicată de un mesaj se caracterizează prin coeficientul  $M_R$  (măsura neuniformității), care se definește astfel:

$$M_R = \sum_{i=1}^N \left( P_i - \frac{1}{N} \right)^2,$$

unde  $P_i$  este probabilitatea de apariție a literei  $\alpha_i$ .

Se observă că dacă:

$$P_1 = P_2 = \dots = P_N = \frac{1}{N},$$

atunci  $M_R = 0$ .

Având în vedere că  $\sum_{i=1}^N P_i = 1$ , atunci  $M_R$  se mai poate scrie:

$$M_R = \sum_{i=1}^N \left( P_i^2 - \frac{2P_i}{N} + \frac{1}{N} \right) = \sum_{i=1}^N P_i^2 - \frac{2}{N} N + N \frac{1}{N^2} = \sum_{i=1}^N P_i^2 - \frac{1}{N},$$

iar dacă  $N = 26$ , atunci  $M_R = \sum_{i=1}^{26} P_i^2 - \frac{1}{26}$ .

Pentru text clar, având 1000 de litere, coeficientul  $M_R$  are valoarea determinată experimental de 0,047592. Așadar  $M_R$  poate lua valori cuprinse între 0 și 0,047492, și permite stabilirea cu o anumită probabilitate dacă cifrarea a fost monoalfabetică sau polialfabetică.

Dificultatea constă în calcularea sumei pătratelor probabilităților de apariție a literelor. Având în vedere că  $P_i^2$  se poate interpreta ca probabilitatea evenimentului ca doua litere selectate la întâmplare din textul considerat să fie identice, apare posibilitatea de a estima  $\sum_{i=1}^N P_i^2$  fără a cunoaște  $P_i$ , cu probabilitatea ca două litere luate la întâmplare să fie identice, ceea ce este echivalent cu calculul raportului dintre numărul dubletelor și numărul total de diagrame din textul considerat.

Dacă notăm cu  $f_A$  frecvența absolută a literei  $A$  din textul considerat, atunci numărul dubletelor  $AA$  va fi:

$$C_{f_A}^2 = \frac{1}{2} f_A (f_A - 1).$$

Dacă numărul total de litere din text este  $N$ , atunci probabilitatea unui dublet este:

$$P_i^2 = \frac{f_i (f_i - 1)}{N(N - 1)},$$

iar  $\sum_{i=1}^N P_i^2 = \frac{\sum_{i=1}^N f_i(f_i - 1)}{N(N - 1)}$  și semnifică șansa ca două litere dintr-o distribuție să

fie identice; acest număr se numește indice de coincidență ( $I_c$ ), putând lua valori cuprinse între 0,038 și 0,086 la un alfabet format din 26 de litere. Limita inferioară corespunde unei distribuții uniforme, iar cea superioară unui cifru monoalfabetic.

O măsură a modului în care numărul de alfabete folosite la cifrare influențează indicele de coincidență se poate determina statistic și  $I_c$  se poate exprima cu ajutorul relației:

$$I_c = \frac{1}{m} \cdot \frac{N - m}{N - 1} \cdot 0,086 + \frac{m - 1}{m} \cdot \frac{N}{N - 1} \cdot 0,038,$$

în care:  $N$  – lungimea textului cifrat;  $m$  – numărul alfabetelor folosite.

Dacă  $N = m$ , atunci  $I_c$  tinde spre valoarea 0,038, iar dacă  $m = 1$ , atunci  $I_c = 0,086$ .

S-au determinat valorile lui  $I_c$  pentru texte mai lungi de 1000 de litere și pentru diferite valori ale lui  $m$  (numărul de alfabete utilizat), rezultatele fiind prezentate în tabelul de mai jos:

$m$	$I_c$
1	0,086
2	0,068
5	0,047
7	0,044
10	0,042
> 10	0,038

În cazul utilizării indicelui de coincidență pentru determinarea numărului de alfabete utilizate la cifrare, trebuie să se țină seama de caracterul statistic al lui  $I_c$ ; deci pentru mesaje scurte valoarea calculată poate fi foarte diferită față de valoarea așteptată. De asemenea, trebuie avut în vedere că valoarea scontată a lui  $I_c$  presupune folosirea alfabetelor de cifrare de același număr de ori. Deci cuvântul cheie nu trebuie să conțină litere identice.

Un alt indicator pentru stabilirea numărului de alfabete utilizat se poate obține datorită proprietăților algoritmului de cifrare polialfabetic, și anume: dacă

se scrie textul clar sub cuvântul cheie, atunci literele textului clar de pe o coloană se cifrează monoalfabetic, prin corespondența stabilită de litera cheii din coloana respectivă.

Ca urmare, dacă se repetă un cuvânt sau un șir de litere în textul clar și dacă se întâmplă ca literele identice să fie în aceeași coloană, atunci și în textul cifrat apar secvențe de litere identice. Observația este prezentată mai jos și este esențială în spargerea sistemelor.

Să presupunem că o cheie de cifrare de lungime  $K$  și  $d$  este multiplu de  $K$ , deci  $d = K \times m$ ; atunci este evident că porțiunile  $[AB]$  și  $[CD]$  vor fi cifrate cu aceeași cheie și dacă ele cuprind porțiuni identice, atunci vor apărea în criptogramă ca secvențe identice.

Deci calculând distanțele dintre secvențele identice din criptogramă și luând factor comun cel mai frecvent al acestora, vom putea găsi lungimea cheii.

Apoi se împarte textul cifrat în coloane (în număr de  $K$ ), fiecare reprezentând cifrări monoalfabetice.

Fie criptograma:

WBHNS/ PGHRF/ SGGGH/ EFHVW/ CCFHK/ ASNSP/ GBWCH/  
HRSFC/ ZFXZD/ VCDTS/ SGFVQ/ BDNSV/ KWQTS/ WGZH.

Se observă că secvențele NSPG se repetă la o distanță de 24 de litere; grupul NS apare odată la distanța de 30 de litere, iar grupul TS mai apare odată la 15 caractere. Putem trage concluzia că lungimea cheii este 3. Deci vom avea trei criptograme cifrate monoalfabetic:

- 1) W N G F G E V C K N G C R C X V T G Q N K T G;
- 2) B S H S G F W F A S B H S Z Z C S F B S W S Z;
- 3) H P R G H H C H S P W H F F D D S V D V Q W H.

Aplicând acum observațiile de la cifrurile monoalfabetice și rezolvând unele congruențe modulo 26, se găsește pentru prima criptogramă cheia 2, pentru a doua cheia 14, iar pentru a treia, cheia 3.

Cu acestea, soluția apare simplu:

UNELE METODE DE SECRETIZARE IMPLEMENTATE PE  
CACULATOARELE PERSONALE ȘI ÎN REȚELE.

Dacă sistemele de cifrare au folosit reprezentări multiple proporționale, vom urmări reducerea acestora la sisteme bazate pe substituții simple cu reprezentări unice uniforme.

În cazul reprezentărilor multiple pentru aceeași literă a mesajului clar, rezultă că numărul elementelor cifrante este mai mare decât numărul de litere din alfabet și trebuie găsite acele reprezentări cifrante care sunt echivalente în sensul că reprezintă aceeași literă.

Dacă, de exemplu, în criptogramă apar două secvențe asemănătoare, cum ar fi:

13 11 42 70 53  
13 11 42 70 53

este aproape sigur că 42 și 18 sunt reprezentări cifrante ale aceleiași litere.

Toate echivalențele care se fac trebuie verificate căutând alte secvențe care le cuprind.

Metoda dă rezultate, însă necesită un volum mare de text interceptat.

## **6.4 Spargerea sistemelor poligrafice**

Pentru o mai bună mascare a caracteristicilor statistice ale variabilei aleatoare implicate de textul cifrat, se recurge adeseori la un algoritm de cifrare în care unitatea de prelucrare este un șir de litere (două sau mai multe).

Procedura de decriptare a unui astfel de cifru se exemplifică în cele ce urmează:

1	FFTVT	SPUCM	BAUKA	BVCAA	PMYIK	KWEXO
2	UKABY	ECUXJ	XXYCY	PETKE	HOYRX	WCMZY
3	XIOGIP	ZDAZI	FFZVP	JNNCH	GIOUL	WGAZC
4	IMSXR	ZPBTV	LPDAN	FFFZI	BLYLP	UDUKP
5	CMERG	IXGZV	VYPMD	FEGVC	SYDQZ	VVYPM
6	DFDPU	RPUPZ	EBTSO	UVQTV	TSKIX	XUSFL
7	XWAST	YKGNB	GKZRA	DYCLW	HKHW	CXXFL
8	IMLPJ	FLCYE	CUZJX	XOGFR	MORVK	XXZZ
9	JRFFA	BNANP	XXUUA	CQUNU	RMDFD	FQUJM
10	SOYLX	XUFRI	JHZRN	FRFIM	ZZXNF	ITAZI
11	VAJEF	IGYXF				
12	HQDFN	NGYFF	DHXJO	AGIWN	XOABW	EKSVU
13	WWDUL	HOUKG	NBRYL	PCHNF	GAKGO	UFBET
14	PUAAN	LRMWX	OUQUG	AKGOU	BANQD	FNN

Sarcina primordială a criptanalistului este de a obține anumite date privind sistemul criptografic folosit la cifrare. Pentru aceasta se stabilește distribuția monografică a variabilei aleatoare implicate de textul cifrat și se calculează indicele de coincidență:



$$I.C. = \frac{\sum_{i=A}^Z f_i(f_i - 1)}{N(N - 1)}$$

În cazul criptogramei date,  $N = 418$  caractere, rezultă  $I.C. = 0,036$ , ceea ce arată că cifrarea nu este monoalfabetică. Pentru a verifica dacă cifrarea este sau nu polialfabetică, se examinează șirurile de litere care se repetă în criptogramă. În criptograma recepționată, astfel de șiruri de litere cu lungimea minimă egală cu 4 sunt următoarele:

Șiruri de litere care se repetă	Poziția primei litere		$l$ – lungimea intervalului	Divizorii primi ai lui $l$
ZVVYPMOF	129	145	16	2
HQDFNN	330	412	62	2; 31
GAKGOU	380	404	24	2; 3
TVTS	3	169	166	2; 83
KIXX	173	235	82	2; 41

Este foarte puțin probabil că repetițiile formate din 8, respectiv din 6 litere să fie accidentale. Singurul factor prim comun al intervalelor de repetiție fiind 2, se verifică dacă s-a executat sau nu cifrarea cu ajutorul a două alfabete. Totuși 2 fiind factor pentru fiecare din intervalele de repetiție, se pune întrebarea dacă sistemul de cifrare este digrafic. Pentru a răspunde la această întrebare, se întocmește tabelul de frecvențe de apariție a digramelor din textul cifrat.

Într-un fel sistemul digrafic se poate considera ca un sistem monoalfabetic, însă pe un alfabet de  $26^2 = 676$  caractere. Fiecare dintre aceste caractere (diagrame) are o frecvență caracteristică în textul clar. În acest caz, se poate aplica raționamentul de la decriptarea monoalfabetică.

Pentru un eșantion de text clar (2500 de diagrame), calculând suma pătratelor frecvențelor caracteristice ale celor 676 de diagrame, rezultă:

$$\sum_{i=AA}^{ZZ} P_i^2 = 0,0116,$$

ceea ce arată gradul de neuniformitate a distribuției (generate) digrafice, ținând seama că pentru o distribuție uniformă acest coeficient este egal cu  $\frac{1}{676} = 0,0015$ .

Cu ajutorul frecvențelor de apariție a digramelor se determină indicele de coincidență:

$$I.C. = \frac{\sum_{i=AA}^{ZZ} f_i(f_i - 1)}{N(N - 1)} = 0,00989 \approx 0,001$$

unde  $N = 209$  reprezintă numărul de digrame din criptogramă.

Având în vedere că I.C este o aproximare liniară a sumei  $\sum P_i^2 = 0,0116$ , rezultă că sistemul folosit este digrafic; în cadrul acestui sistem, ori de câte ori apare o digramă din textul clar, ea este întotdeauna înlocuită cu aceeași digramă de cifru. În continuare, se pune problema stabilirii tipului sistemului digrafic de cifrare, precum și a cheii specifice de cifrare. Există mai multe tipuri de sisteme digrafice. De exemplu, transformarea liniară ilustrată de ecuația matriceală:

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

Pentru a verifica acest lucru se realizează o serie de texte speciale. Astfel, dacă s-ar putea identifica două sau mai multe corespondențe de digrame de forma  $C_1C_2 \rightarrow M_1M_2$ , atunci cu ajutorul echivalențelor numerice ale literelor  $C_1, C_2, M_1$  și  $M_2$  se poate obține un sistem de congruențe de forma:

$$\begin{cases} M_1 = C_1b_{11} + C_2b_{12} & (\text{mod } 26) \\ M_2 = C_1b_{21} + C_2b_{22} & (\text{mod } 26) \end{cases}$$

Matricea sistemului fiind inversa matricei din transformarea liniară de mai sus, rezultă:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Căutarea digramelor componente se face în mod similar cu căutarea literelor pentru soluționarea unui sistem de cifrare monoalfabetică. Dacă s-ar putea stabili două corespondențe de forma  $C_1C_2 \rightarrow M_1M_2$ , atunci se pot scrie patru congruențe, suficiente pentru calcularea necunoscutelor  $b_{ij} (i, j = \overline{1, 2})$ . Dacă apar ambiguități pentru valorile coeficienților  $b_{ij}$ , atunci este necesar să se examineze un text cifrat suplimentar pentru obținerea soluției corecte. Desigur, se presupune că echivalențele numerice ale literelor sunt cunoscute.

În căutarea digramelor corespunzătoare se pleacă de la frecvențele de apariție a digramelor și a șirurilor de litere din textul cifrat și din textul în clar în eșantion.

Digrame în ordine descrescătoare (la un text clar):

TI	AR	EN
AT	TA	MI
DE	ST	NT
IN	CA	PE
ER	TO	SI
TE	IC	UN
UL	OK	EC
RE	RA	IE
		IT

Din exemplul:

<i>OU</i>	<i>LP</i>
<i>XX</i>	<i>QU</i>
<i>DF</i>	<i>TV</i>
<i>FG</i>	<i>ZI</i>
<i>PU</i>	<i>ZV</i>
<i>AB</i>	<i>FF</i>
<i>GA</i>	

se încearcă 3 comparații:

$$\begin{aligned} OU_c &\rightarrow TI_m \\ XX_c &\rightarrow AT_m \\ DF_c &\rightarrow DE_m \end{aligned}$$

Acestea conduc la un sistem de congruențe:

$$\begin{aligned} 19 &\equiv 14b_{11} + 20b_{12} \\ 8 &\equiv 14b_{21} + 20b_{22} \\ 0 &\equiv 23b_{11} + 23b_{12} \pmod{26} \\ 19 &\equiv 23b_{21} + 23b_{22} \\ 3 &\equiv 3b_{11} + 5b_{12} \\ 4 &\equiv 3b_{21} + 4b_{22} \end{aligned}$$

iar din rezolvare rezultă un sistem incompatibil

$$\begin{aligned} OU &\rightarrow IE \\ XX &\rightarrow TI \\ DF &\rightarrow TA \end{aligned}$$

$$\begin{cases} 8 \equiv 14b_{11} + 20b_{12} \\ 4 \equiv 14b_{22} + 20b_{22} \\ 19 \equiv 23b_{11} + 23b_{12} \pmod{26} \\ 8 \equiv 23b_{21} + 23b_{22} \\ 19 \equiv 3b_{11} + 5b_{12} \\ 0 \equiv 3b_{22} + 5b_{22} \end{cases}$$

Se rezolvă primele 4, grupându-le astfel:

$$\begin{cases} 8 \equiv 14b_{11} + 20b_{12} \mid \cdot 23 \\ 19 \equiv 23b_{11} + 23b_{12} \mid \cdot 20 \end{cases} \quad \begin{cases} 4 \equiv 14b_{21} + 20b_{22} \\ 8 \equiv 23b_{21} + 23b_{22} \end{cases}$$

$$184 \equiv 322b_{11} + 460b_{12}$$

$$380 \equiv 460b_{11} + 460b_{22}$$

Soluțiile 1, 4, 5 și 8 verifică și congruențele din sistemul dat mai sus (cu 6 ecuații).

Având în vedere că soluția 1 duce la matricea neinvertibilă:

$$P = \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix},$$

aceasta nu poate fi o soluție a transformării liniare:

$$P_8 = \begin{bmatrix} 18 & 19 \\ 15 & 17 \end{bmatrix}.$$

## **Test de autoevaluare**

1. Enumerați principalele statistici ale limbajelor naturale.
2. Ce reprezintă un text nealeator?
3. Ce formă are măsura neuniformității și care este rolul ei în analiza cifrării monoalfabetice sau polialfabetice?

## **Tema de autoinstruire**

1. Realizați schema bloc a spargerii unei criptograme realizate prin utilizarea metodei cifrului lui Caesar, bazată pe exemplul din curs.

### Lucrare de verificare

1. Să considerăm că s-a interceptat următorul text, criptat cu un sistem monoalfabetic (nu se știe exact ce sistem a fost utilizat). Prin metode de criptanaliză, să se descifreze textul de mai jos.

lqakc sp gcxk aca pcmgqb kq kxc pkersmpqsb vk vsmgxkbc mkacpc tcacpbqlqs  
vk cgele cmtxq ms nocxgsb mbxcsp vk exsgk oxcbsbcbk texbslk spclbk gcxk  
cmgqpvkq bxkgcbexslk gqxbslk xktxknkpbq tkpbxq mbxcsp q  
cfkxbsmakpb mqtexcbex vcx lsatkvk pq bxkrqscq mc zsk txkc gqxsems ps  
mk cmbktbk mc czlk acxk lqgxq vk lc gkl gq gcxk fkpkcq sp gepbcgb

## BIBLIOGRAFIE

1. \*\*\*\*\* – *Applications of Digital Signal Processing*, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1978.
2. \*\*\*\*\* – *The Current State of DES*, Dr. Dobb's Journal, comunicare Internet, 2001.
3. \*\*\*\*\* – *Cracking DES: Secrets of Encryption Research, Wiretap Politics & Chip Design*, Electronic Frontier Foundation, 1999.
4. **G. Agnew** – *Random Sources for Cryptographic Systems*, Advances in Cryptology, EUROCRYPT '87 (LNCS 304), 77-81, 1988.
5. **N. Ahmed, K.R. Rao** – *Orthogonal Transforms for Digital Signal Processing*. Springer-Verlag, New York, 1975.
6. **I. Angheloiu, V.V. Patriciu** – *Securitatea și protecția informației în sistemele electronice de calcul*, Ed. Militară, București, 1986.
7. **I. Angheloiu, E. Gyorfi** – *Elemente de teoria transmiterii informației*, Editura Academia Militară, București, 1976.
8. **A. Angot** – *Complemente de matematici*, Editura Tehnică, București, 1965.
9. **W. Barker** – *Cryptanalysis of the Hagelin Cryptograph*, Aegean Park Press, Laguna Hills, California, 1977.
10. **U. Baum, S. Blackburn** – *Clock-controlled Pseudorandom Generators on Finite Groups*, B. Preneel (Editor), Fast Software Encryption, Second International Workshop (LNCS 1008), p. 6-21, Springer-Verlag, 1995.
11. **H. Beker, F. Piper** – *Cipher Systems: The Protection of Communications*, John Wiley & Sons, New York, 1982.
12. **M. Bellare, J. Kilian, P. Rog-Away** – *The Security of Cipher Block Chaining*, Advances in Cryptology, CRYPTO '94 (LNCS 839), p. 341-358, 1994.
13. **I. Ben-Aroya, E. Biham** – *Differential Cyptanalysis of Lucifer*, Advances in Cryptology, CRYPTO '93 (LNCS 773), p. 187- 199, 1994.
14. **E. Biham, A. Shamir** – *Differential Cryptanalysis of DES-like Cryptosystems*, Advances in Cryptology – CRYPTO'90 Proceedings, Springer-Verlag, 1991.

15. **E. Biham** – *New Types of Cryptanalytic Attacks Using Related Keys*, Advances in Cryptology, EUROCRYPT '93 (LNCS 765), p. 398-409, 1994.
16. **L. Blum, M. Blum, M. Shub** – *Comparison of Two Pseudo Random Number Generators*, Advances in Cryptology, Proc. of Crypto 82, p. 61-78, 1983.
17. **L. Coculescu, V. Cristea, I. Finta, V.V. Patriciu, F. Pilat** – *Proiectarea sistemelor teletinformatice*, Editura Militară, București, 1988.
18. **I. Constantinescu, S. Condrea, E. Nicolau** – *Teoria informației*, Editura Tehnică, București, 1958.
19. **R. Fano** – *Transmission of Information. A Statistical Theory of Communications*, J. Wiley & Sons, Inc., New York, 1961
20. **F. Mac Williams, N.J.A. Sloane** – *The Theory of Error Correcting Codes*, Murray Hill, New York, 1977.
21. **A. Menezes, P. van Oorschot, S. Vanstone** – *Handbook of Applied Cryptography*, CRC Press, Canada, 1997.
22. **A.-T. Murgan** – *Principiile teoriei informației în ingineria informației și a comunicațiilor*, Editura Academiei Române, București, 1998.
23. **V.V. Patriciu** – *Criptografia și securitatea rețelelor de calculatoare cu aplicații în C și Pascal*, Editura Tehnică, București, 1994.
24. **V.V. Patriciu, M. Pietroșanu-Ene, I. Bica, C. Cristea** – *Securitatea informatică în UNIX și INTERNET*, Editura Tehnică, București, 1994.
25. **C. Răuciu** – *Aspecte privind secretizarea semnalelor video*, A XXVI-a Sesiune de Comunicări Științifice cu Participare Internațională, Academia Tehnică Militară, București, 1995.
26. **C. Răuciu, E.L. Jinga** – *Secretizarea imaginilor statice*, A XXIX-a Sesiune de Comunicări Științifice cu Participare Internațională, Academia Tehnică Militară, București, 2001.
27. **C. Răuciu, Ș.C. Pobereznenco** – *Security of Steganography Systems. Concealing Information in Image Files*, Communications 2002, Academia Tehnică Militară, București, 2002.
28. **C. Răuciu** – *Considerații asupra algoritmilor de criptanaliză statistică*, A XXX-a Sesiune de Comunicări Științifice cu Participare Internațională, Academia Tehnică Militară, București, 2003.
29. **C. Răuciu, V.V. Patriciu, I. Bica, I. Rîncu** – *Studiu privind condițiile ce trebuie îndeplinite de rețelele de comunicații utilizate de către o autoritate de certificare de tranzit*, Contract nr.:A-2887/2002, Beneficiar: Serviciul de Telecomunicații Speciale București, București, 2002.

30. **C. Răcuciu** – *Aspecte ale criptării imaginilor prin metode simetrice*, A XXX-a Sesiune de Comunicări Științifice cu Participare Internațională, Academia Tehnică Militară, București, 2003.
31. **D.R. Stinson** – *Cryptography: Theory and Practice*, CRC Press, Florida, 1995.
32. **R. Rivest, A. Shamir, L.M. Adleman** – *Cryptographic Communications System and Method*, U.S. Patent # 4,405,829, 20 Sep. 1983.
33. **A. Salomma** – *Criptografia cu chei publice*, Editura Militară, București, 1993.
34. **B. Schneier** – *Applied Cryptography*, John Wiley & Sons, 1996.