Criptarea lui Cezar

Când trimitea mesaje oficiale către subalternii săi, Cezar folosea o tehnică foarte simplă de codificare a acestora. Fie două cercuri pe care sunt scrise cele 26 de litere ale alfabetului; dacă rotim cercul mic cu un număr de litere față de celălalt în sensul acelor de ceasornic, atunci, pentru a cripta un mesaj, se caută fiecare literă din mesaj pe cercul mare şi se scrie litera corespunzătoare de pe cercul mic.

Figura 1 Figura 2

Dacă avem textul IULIUS CAESAR şi deplasăm cercul din interior din figura 1 cu două litere în sensul acelor de ceasornic după cum se observă în figura 2, atunci textul criptat va fi GSJGSQ AYCQYP.

Procedeul de rotire a cercului mic cu un anumit număr de litere poartă numele de deplasare. Pentru această tehnică de criptare, o parte din cheia (parola) de criptare o constituie numărul de deplasări (deplasament) de câte o literă în sensul acelor de ceasornic. Cealaltă parte din cheia de criptare o constituie alfabetul ales. Alfabetul folosit poate să fie altul decât cel prezentat în exemplul anterior. Acesta poate să conțină pe lângă majuscule şi alte caractere printre care pot apărea cifre, caractere speciale, alte litere ale alfabetelor diferitelor limbaje utilizate sau poate să fie întreg setul de caractere ASCII folosit de calculator.

Pentru a decripta un mesaj criptat cu această metodă, trebuie să se cunoască numărul de deplasări utilizat la criptarea textului inițial. Procesul de obținere a mesajului inițial din mesajul criptat când se cunoaşte deplasamentul este foarte simplu. Se deplasează cercul din interior cu atâtea litere cu câte indică deplasamentul şi fiecare literă din textul criptat se caută pe cercul din interior şi se scrie litera corespunzătoare de pe cercul din exterior. Din textul FRIFRP ZXBPXO criptat în cazul alfabetului englezesc, folosind un deplasament de trei caractere, prin decriptare se obține IULIUS CAESAR. Dacă notăm cu L numărul de caractere al alfabetului utilizat, atunci textul criptat folosind un deplasament D mai mare sau egal cu L este acelaşi cu cazul în care deplasamentul este restul împărțirii lui D la L.

Această metodă de criptare nu este eficientă, deoarece tex- tul criptat poate fi uşor decriptat dacă se cunoaşte alfabetul utilizat. La decriptare se folosesc L deplasamente şi astfel se generează L texte dintre care numai unul este cel bun şi poate
fi uşor identificat.

Variante ale criptării lui Cezar

O variantă a criptării lui Cezar este tehnica folosită de Ovidius. Acesta nu folosea deplasarea cercului din interior, în schimb literele de pe cercul din interior erau scrise în ordine inversă (de la ultima literă din alfabet la prima, în sensul acelor de ceasornic). Acest lucru este echivalent cu scrierea literelor din alfabet pe un rând în ordine şi pe celălalt rând în ordine inversă, după cum se poate observa în continuare:

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

Pentru a cripta cuvântul OVIDIUS se caută pentru fiecare literă din cuvânt litera de pe primul rând şi se scrie litera de pe al doilea rând de sub litera găsită. Aplicând acest procedeu obținem cuvântul LERWRFH.

Pentru decriptare se foloseşte acelaşi procedeu ca în cazul criptării deoarece, dacă unei litere i de pe primul rând îi corespunde a (L + 1 - i)- a literă din al doilea rând, atunci celei de-a (L + 1 - i)- a litere din primul rând îi corespunde a (L + 1 -(L + 1 - i))- a literă din al doilea rând.
Având aceste componente, procesul de decriptare este similar celui prezentat în secţiunea anterioară.

Dacă se foloseşte alfabetul englezesc, ca posibilă modalitate de aranjare a caracterelor pe cele două cercuri ar putea fi aleasă aranjarea lor de pe tastatură,şi anume:

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

În cazul în care se cunoaşte numai tehnica de criptare utilizată, decriptarea mesajului este mai complicată, deoarece se bazează pe statistică matematică şi în principal pe faptul că într-un text necriptat cu o dimensiune suficient de mare vocalele au o frecvenţă foarte mare şi nu există două vocale consecutive din text care să fie la o distanţă mai mare de şase caractere (pentru limba română).

Metoda substituţiei

Prin substituţie se înţelege înlocuirea fiecărui caracter din alfabet cu un altul şi nu există două caractere care să fie înlocuite cu acelaşi caracter. O altă posibilitate ar fi aceea de
a înlocui fiecare caracter din alfabetul ales cu un simbol sau un număr.

Tehnicile de criptare prezentate până acum sunt cazuri particulare ale criptării prin substituţie. Considerând alfabetul englezesc, există 26! posibilităţi de substituire a caracterelor cu altele din care, pentru o mai mare securitate a datelor se scade numărul de posibilităţi în care un caracter este substituit cu el însuşi (acest număr se obţine rezolvând problema concordanţelor).
Criptarea unui text folosind această metodă se face ca în cazul criptării lui Ovidius, adică pentru fiecare caracter din text se scrie litera care s-a ales pentru substituţie.
Fie următoarea aranjare a caracterelor alfabetului englezesc:

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

Pentru textul PAROLA ESTE ASDFG, mesajul criptat folosind aranjarea de mai sus, este: NZFEDZ JOXJ ZOASB.

Procesul de decriptare se face în sens invers, adică pentru fiecare literă din textul criptat se scrie litera pe care aceasta a substituit-o în procesul de criptare.

A. Criptosisteme monografice sau caracter.

Abordarea matematică pentru Cifrul Cezar

Vom considera în toate exemplele că mesajul este alcătuit din litere ale unui alfabet cu 26 de litere (cel al limbii engleze, de exemplu). Atribuim fiecărei litere din alfabet un echivalent numeric de la 0 la 25. Astfel realizăm corespondenţele A - 0, B – 1,…, Z – 25 ca în tabelul de mai jos:

a	b	c	D	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	Q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Aici, unitatea de mesaj este formată dintr-o singură literă. Un prim exemplu simplu este un criptosistem care se presupune că a fost inventat şi folosit de Iulius Cezar.

Exemplu. În textul de bază fiecare literă, pe care o notăm cu P, este înlocuită cu litera aflată la 3 poziţii la dreapta faţă de aceasta pe care o notăm cu C. Astfel, litera “a” devine “d”; litera “x” se transformă în “a” ; “y” devine “b” iar “z” devine “c” . Transformarea se poate scrie ţinând cont de echivalenţii numerici ai literelor astfel:

$$C \equiv P + 3 (\mathrm{mod}\, 26),\ 0 \le C \le 25.$$

Presupunem că textul de bază este : THIS MESSAGE IS TOP SECRET

Pentru început, textul de bază se împarte în blocuri de 5 litere

THISM ESSAG EISTO PSECR ET

care apoi sunt înlocuite cu echivalenţii lor numerici:

19 7 8 18 12 4 18 18 0 6 4 8 18 19 14 15 18 4 3 17 4 10.

Fiecare echivalent numeric este transformat după regula precizată şi rezultă:

22 10 11 21 15 7 21 21 3 9 7 11 21 22 17 18 21 7 6 20 7 13.

Blocurile nou construite sunt transformate în litere corespunzătoare echivalenţilor numerici:

WKLVP HVVDJ HLVWR SVHGU HW

care formează textul cifrat.

Pentru decifrare, se foloseşte transformarea inversă $P \equiv C - 3 (\mathrm{mod}\, 26)$ cu $0 \le \mathrm{P} \le 25$ şi se aplică acelaşi procedeu după care se refac din blocuri cuvintele iniţiale.

Acest exemplu este un caz particular al criptosistemelor descrise prin transformările de forma $C \equiv P + k \,(\mathrm{mod}\ 26)$, $0 \le \mathrm{C} \le 25$. Acestea se numesc *transformări de deplasare.* Cheia de cifrare este k .

Fie acum transformarea definită prin $C \equiv aP + b (\mathrm{mod}\ 26)$, $0 \le \mathrm{C} \le 25$, a,b întregi cu $(a,26) = 1$. Se observă că pentru a există $\varphi(26) = 12$ posibilităţi de atribuire şi pentru b avem 26. Deci există $12 \cdot 26 = 312$ astfel de transformări (printre care şi transformarea identică) numite *transformări afine*. Cheia de cifrare este dată de a şi b.

Se remarcă faptul că transformările de deplasare sunt cazuri particulare ale celor afine obţinute pentru $a = 1$. Procesul de criptare se desfăşoară la fel ca în exemplul dat, numai că de această dată echivalenţii numerici se modifică după noua relaţie. Pentru decriptare se foloseşte transformarea $P \equiv \alpha(C + b) (\mathrm{mod}\, 26)$, $0 \le \mathrm{P} \le 25$ unde α este inversul lui a modulo 26.

Dacă se doreşte spargerea unui cifru presupus a fi de tip monografic trebuie făcută o analiză o frecvenţei apariţiei literelor din textul cifrat şi făcută o comparare cu frecvenţa literelor dintr-un text obişnuit. Se cunoaşte că în limba engleză cele mai frecvente litere din cadrul unui text sunt E, T, N, R, I, O, A (pentru limba română ele ar fi I, E, A, B). Astfel, punând în corespondenţă cea mai des întâlnită literă din textul cifrat (de preferinţă mai lung pentru o mai corectă realizare a corespondenţei între literele de frecvenţă maximă) cu cea care apare de cele mai multe ori într-un text arbitrar se pot dobândi informaţii legate de transformarea folosită la criptare.

Exemplu. Presupunem că un text a fost cifrat printr-o transformare de deplasare şi observăm că litera care are frecvenţa cea mai mare în textul cifrat este “p”. Putem atunci presupune că ea corespunde literei “e” din textul de bază, cum aceasta are cea mai mare frecvenţă într-un text scris în limba engleză. Ţinând cont de echivalenţii numerici corespunzători şi înlocuind în relaţie obţinem $15 \equiv 4 + k \,(\mathrm{mod}\, 26)$, de unde cheia de cifrare posibil folosită este $k = 11$. Pentru transformările de deplasare, determinarea cheii nu presupune necesitatea ca textul cifrat să fie lung pentru că, de fapt, există doar 25 de posibilităţi pentru k ceea ce nu presupune un volum atât de mare de muncă. În concluzie, acest tip de criptosistem nu este prea bun.

Exemplu. Să presupunem acum că un text a fost criptat printr-o transformare afină. Din analiza frecvenţei literelor din textul cifrat, vedem că cel mai des apar literele “l” şi “u” . Bănuim atunci că “l” corespunde lui “e” iar “u” lui “ t”. Obţinem atunci relaţiile $11 \equiv 4a + b (\mathrm{mod}\ 26)$ şi $20 \equiv 19a + b (\mathrm{mod}\ 26)$. Din rezolvarea acestui sistem de congruenţe obţinem $a \equiv 11 (\mathrm{mod}\, 26)$ şi $b \equiv 19 (\mathrm{mod}\, 26)$. Dacă presupunea noastră este corectă, transformarea afină folosită este dată de $C \equiv 11P + 19 (\mathrm{mod}\ 26)$ iar pentru decriptare se foloseşte $P \equiv 19(C - 19) \equiv 19C + 3 (\mathrm{mod}\, 26)$ unde am ţinut cont că $a \equiv 19 (\mathrm{mod}\, 26)$.

Criptosisteme poligrafice sau bloc (Hill)

Pentru a evita faptul că primele criptosisteme sunt vulnerabile, criptanaliza realizându- se folosind frecvența literelor în text, s-a preferat împărțirea textului de bază în blocuri de o anumită lungime şi transformarea acestora în blocuri cu aceeaşi dimensiune. Aceste *criptosisteme* se numesc *poligrafice* sau *bloc*.

Studiem întâi cazul *cifrului diagrafic* pe un exemplu concret. Aici, blocurile sunt formate din două litere. Considerăm ca text de bază THE GOLD IS BURIED IN ORONO.

Întâi se împarte textul de bază în blocuri de două litere. Dacă numărul literelor este impar, ultimul bloc este completat cu a literă, de exemplu "x".

TH EG Ol DI SB UR IE DI NO RO NO

Fiecare literă din bloc este înlocuită cu echivalentul sau numeric

19 7 4 6 14 11 3 8 18 1 20 17 8 4 3 8 13 14 17 14 13 14

Fiecare bloc de numere din textul de bază P_1P_2 este înlocuit cu blocul C_1C_2 după transformarea $C_1 \equiv 5P_1 + 17P_2 (\mathrm{mod}\, 26)$. Obținem acum blocurile $C_2 \equiv 4P_1 + 15P_2 (\mathrm{mod}\, 26)$.

6 25 18 2 23 13 21 2 3 9 25 23 4 14 21 2 17 2 11 18 17 2

care transformate în litere formează textul cifrat

GZ SC XN VC DJ YX EO VC RC LS RC

Procesul de decifrare se face după regula $P_1 \equiv 17C_1 + 5C_2 (\mathrm{mod}\, 26)$, $P_2 \equiv 18C_1 + 23C_2 (\mathrm{mod}\, 26)$

Acest criptosistem este mult mai uşor de descris matriceal, şi anume:

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \equiv \begin{pmatrix} 5 & 17 \\ 4 & 15 \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} (\mathrm{mod}\, 26)$$

Se observă că matricea care intervine are invers modulo 26, matricea inversă intervenind în procesul de decriptare.

Trecem acum la cazul general, în care blocurile în care este împărțit textul de bază conțin fiecare n litere. Procesul de cifrare urmăreşte aceeaşi cale ca pentru cifrul diagrafic doar că, transformarea acum este dată de $C \equiv A \cdot P (\mathrm{mod}\ 26)$ unde $A \in M_n(Z)$ cu $(\det A, 26)$

$$C = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}, \quad P = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix}.$$

Păstrând notațiile făcute în primul capitol, A^t este inversa modulo 26 a matricei A, adică $A * A^t \equiv I_n (\mathrm{mod}\, 26)$. De aici, pentru decriptare se foloseşte relația: $P \equiv A^t \cdot C (\mathrm{mod}\, 26)$.

Pentru o mai bună exemplificare, considerăm cazul $n = 3$. Textul de bază este

STOP PAYMENT

iar transformarea este dată de

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} \equiv A \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} (\mathrm{mod}\, 26) \text{ unde } A = \begin{pmatrix} 11 & 2 & 19 \\ 5 & 23 & 25 \\ 20 & 7 & 1 \end{pmatrix}.$$

$\det(A) \equiv 5 (\mathrm{mod}\, 26)$ deci, $(\det A, 26) = 1$.

Cum la împărțirea textului în blocuri de câte 3 litere ultimul bloc este format doar din două litere, mai adăugăm un "x" în final:

STO PPA YME NTX

Transformăm blocurile în numere folosind echivalenții numerici

18 19 14 15 15 0 24 12 4 13 19 23

Pentru primul bloc avem:

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} \equiv \begin{pmatrix} 11 & 2 & 19 \\ 5 & 23 & 25 \\ 20 & 7 & 1 \end{pmatrix} \begin{pmatrix} 18 \\ 19 \\ 4 \end{pmatrix} \equiv \begin{pmatrix} 8 \\ 19 \\ 13 \end{pmatrix} (\mathrm{mod}\, 26).$$

Repetând calculul pentru fiecare bloc, obținem

8 19 13 13 4 15 0 2 22 20 11 0

care convertite în litere dau textul cifrat

ITN NEP ACW ULA

Pentru decifrare folosim matricea

$$A^t = \begin{pmatrix} 6 & -5 & 11 \\ -5 & -1 & -10 \\ -7 & 3 & 7 \end{pmatrix}.$$

Aceste criptosisteme sunt şi ele vulnerabile privind analiza frecvenței blocurilor de litere. De exemplu, cele mai des întâlnite grupuri de două litere în limba engleză sunt TH, HE iar pentru blocuri de trei litere apar cel mai frecvent THE, AND, THA. Făcând analiza corespunzătoare, putem găsi matricea de cifrare.

Exemplu. Dacă într-un text cifrat cu cifru diagrafic cele mai frecvente perechi de litere sunt KX şi VZ putem bănui că acestea corespund în textul de bază lui TH respectiv, HE.
Atunci, (19 7) şi (7 4) sunt trimise în (10 23) respectiv în (21 25). Obținem astfel

$$\begin{pmatrix} 10 & 21 \\ 23 & 25 \end{pmatrix} \equiv A \begin{pmatrix} 19 & 7 \\ 7 & 4 \end{pmatrix} (\mathrm{mod}\, 26)$$

Inversa modulo 26 a matricei $\begin{pmatrix} 19 & 7 \\ 7 & 4 \end{pmatrix}$ este $\begin{pmatrix} 4 & 19 \\ 19 & 19 \end{pmatrix}$ de unde,

$A \equiv \begin{pmatrix} 10 & 21 \\ 23 & 25 \end{pmatrix} \begin{pmatrix} 4 & 19 \\ 19 & 19 \end{pmatrix} \equiv \begin{pmatrix} 23 & 17 \\ 21 & 2 \end{pmatrix} (\mathrm{mod}\, 26)$ dă o posibilă cheie.

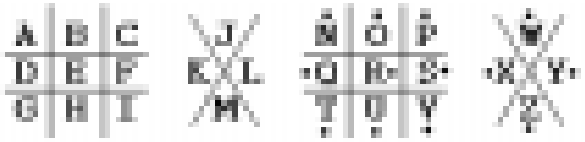
Criptarea cu alfabet aleator

O altă variantă a criptării lui Cezar este folosirea unui alfabet şi aranjarea caracterelor acestuia, pe cele două cercuri, într-o ordine aleatoare, dar caracterele de pe ambele cercuri trebuie să fie aranjate în aceeaşi ordine.

Pentru decriptarea mesajelor criptate cu ajutorul unui alfabet ales aleator, trebuie cunoscute alfabetul, ordinea caracterelor din alfabet precum şi numărul de deplasări.
substitui această literă. În procesul de criptare, pentru fiecare literă din textul inițial se alege aleator un număr din mulțimea care i-a fost atribuită.

De exemplu, dacă pentru caracterul A se alege mulți mea {1, ..., 10}, pentru litera B se alege mulțimea {11, ..., 20} etc., există mai multe posibilități pentru a cripta textul de mai sus; una dintre ele poate fi: 154, 7, 172, 149, 120, 156, 48, 185, 194, 44, 4, 184, 33, 51, 63. Nu este obligatoriu ca mulțimile alese pentru fiecare 31 caracter să aibă același cardinal; de fapt, este indicat ca mulțimile asociate literelor care apar frecvent să aibă mai multe elemente.

Pentru transmiterea de texte criptate se mai pot folosi imagini în care caracterele din textul inițial sunt înlocuite cu simboluri predefinite. Una dintre cele mai cunoscute metode de înlocuire a caracterelor cu simboluri este prezentată în figura următoare.



Pentru fiecare caracter din figură se desenează conturul regiunii în care acesta se află. De exemplu, pentru litera N se desenează laturile regiunii din stânga-sus ale diagramei în care se află împreună cu un punct între laturi.

Metoda transpoziției

Criptarea prin metoda transpoziției este o tehnică mai eficientă decât criptarea prin substituție, dar are, la rândul ei,

o mulțime de dezavantaje.

Textul criptat prin metoda transpoziției păstrează toate caracterele textului inițial, dar în altă ordine obținută prin aplicarea algoritmului ce va fi prezentat în continuare. Criptarea prin transpoziție constă din scrierea textului inițial din care s-au eliminat spațiile şi semnele de punctuație, într-o matrice de dimensiune M × N, interschimbarea anumitor linii (sau coloane) între ele şi textul criptat se obține prin scrierea caracterelor din noua matrice de pe fiecare coloană în parte, începând cu colțul din stânga-sus. Dacă lungimea textului inițial este mai mică decât numărul de elemente ce pot fi scrise în matrice, atunci textul se completează cu elemente aleatoare, până ajunge la dimensiunea M*N. Pentru textul Misiunea a fost îndeplinită, care are lungimea de 24 de caractere, se pot alege mai multe matrice de dimensiune M x N, o posibilitate ar fi ca matricea să aibă 4 linii şi 6 coloane, dar pentru ca textul să fie mai greu de decodificat trebuie să conțină şi caractere alese aleator, sau într-un mod mai inteligent, care să îngreuneze munca celui care doreşte să afle conținutul secret din mesajul criptat. Fie o matrice care are 5 linii şi 6 coloane. Textului inițial i se adaugă 6 caractere aleatoare şi se obține textul Misiu neaaf ostîn depli nită x yztwu şi se scrie în matricea din partea stângă astfel:

	1	2	3	4	5	6
1	M	i	s	i	u	n
2	e	a	a	f	o	s
3	t	î	n	d	e	p
4	l	i	n	i	t	ă
5	x	y	z	t	w	u

	1	2	3	4	5	6
5	x	y	z	t	w	u
3	t	î	n	d	e	p
4	l	i	n	i	t	ă
1	M	i	s	i	u	n
2	e	a	a	f	o	s

Prin scrierea liniilor 1, 2, 3, 4, 5 în ordinea 5, 3, 4, 1, 2,2 se obține matricea din partea dreaptă. Textul criptat care se obține este: xtlMe yîiia znnsa tdiif wetuo upăns.

Transpoziție cu parolă

Pentru ca procesul de decriptare să fie mai simplu şi să nu mai fie nevoie ca ordinea în care au fost puse liniile din matricea creată, se foloseşte o variantă a criptării prin transpozițe care se bazează pe o parolă.

Pentru a cripta un text folosind o parolă şi metoda transpoziției, se alege o parolă ale cărei litere determină ordinea în care se vor scrie coloanele din matricea aleasă. Pentru a afla ordinea în care vor fi scrise coloanele din textul inițial, se ordonează alfabetic literele din parolă, şi fiecărei litere i se asociază numărul de ordine din şirul or donat. Lungimea parolei trebuie să fie egală cu numărul de coloane din matrice.

Considerăm textul anterior, scris într-o matrice de dimensiuni 5 x 6, şi parola vultur. Literele din parolă se ordonează alfabetic şi se obține şirul: l, r, t, u, u, v. Indicele 1 este asociat cu litera l, indicele 2 cu litera r, indicele 3 cu litera t, indicele 4 cu prima literă u din parolă, indicele 5 cu a doua literă u din parolă, iar indicele 6 este asociat cu litera v. Pentru a scrie coloanele, pentru fiecare indice i din şirul ordonat se caută indicele j, care reprezintă poziția literei cu indicele i, din parolă şi se scrie coloana j, astfel:

v u l t u r	
6 4 1 3 5 2	1 2 3 4 5 6
1 M i s i u n	5 s n i i u M
2 e a a f o s	3 a s f a o e
3 t î n d e p	4 n p d î e t
4 l i n i t ă	1 n ă i i t l
5 x y z t w u	2 z u t y w x

Textul care se obține în final este: sannz nspău ifdit iaîiy uoetw Metlx.

Pentru a decripta un mesaj criptat folosind această metodă, mesajul se scrie în matrice pe coloane, începând cu colțul stânga-sus, şi apoi se realizează operația inversă, adică pentru fiecare indice j al literelor din parolă, se caută indicele i asociat literei din şirul sortat şi se scrie coloana cu indicele i. Din noua matrice astfel obținută se scriu literele de pe fiecare linie, în ordine.