Lucrare de laborator nr. 3. Cifruri polialfabetice

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

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

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

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

3.1. Cifrul Vigenère

Metoda de criptare cunoscută sub numele de „cifrul Vigenère" a fost atribuită greșit lui *Blaise de Vigenère* în secolul al XIX-lea și, de fapt, a fost descrisă pentru prima dată de *Giovan Battista Bellaso* în cartea sa din 1553 *La cifra del. Sig*. Vigenère a creat un cifru asemănător, dar totuși diferit și mai puternic în 1586.

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

$$c_i = (m_i + k_i) \bmod 26.$$

Cheia poate fi, de exemplu, $k = (5, 20, 17, 10, 20, 13)$ și ar provoca deplasarea primei litere cu 5, $c_1=m_1 + 5 \pmod{26}$, a celei de a doua cu 20, $c_2 = m_2 + 20 \pmod{26}$, ș.a.m.d. până la sfârșitul cheii și apoi de la început, din nou. Cheia este de obicei un cuvânt, pentru a fi mai ușor de memorat – cheia de mai sus corespunde cuvântului „*furtun*". Metoda cu deplasare multiplă oferă protecție suplimentară din două motive:

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

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

$$m_i = (c_i - k_i) \bmod 26.$$

Pentru simplificarea procesului de cifrare se poate utiliza următorul tabel, numit *Tabula Recta* (tabelul 3.1), care se utiliza de către Vigenere. Aici toate cele 26 cifruri sunt situate pe orizontală și fiecărui cifru îi corespunde o anumită literă din cheie, reprezentată în colana din stânga tabelului. Alfabetul corespunzător literelor textului clar se află în prima linie de sus a tabelului. Procesul de cifrare este simplu – este necesar ca având litera $\boldsymbol{m_i}$ din mesaj și litera $\boldsymbol{k_i}$ din

cheie să găsim litera textului cifrat c_i, care se află la intersecția liniei m_i și coloanei k_i. În exemplul din tabelul 3.1 este prezentat cazul $m_i = M$ și $k_i = H$, iar în rezultat se obține $c_i = T$.

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

Tabelul 3.1. *Tabula Recta pentru cifrul Vigenere*

Se poate de procedat și în conformitate cu ecuațiile ce definesc modelul matematic al cifrului:

$$c_i = m_i + k_i \pmod{26} \text{ și } m_i = c_i - k_i \pmod{26},$$

așa cum este arătat în exemplul ce urmează.

Exemplu.
De cifrat, utilizând cifrul Vigenere, mesajul „*Per aspera ad astra*” cu cheia ***K***= SUPER.
Soluție. Pentru a cifra sau descifra mai întâi facem corespondența următoare (codificăm 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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Apoi alcătuim și completăm tabelul:

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

Criptograma este ***C***= *HYGEJHYGERVUHXIS*.

Pentru decriptare procedăm la fel, aplicând formula reciprocă: $m_i = c_i - k_i$ (mod 26).

Textul cifrat C	H	Y	G	E	J	H	Y	G	E	R	V	U	H	X	I	S
Cheia K	S	U	P	E	R	S	U	P	E	R	S	U	P	E	R	S
Textul cifrat C	7	24	6	4	9	7	24	6	4	17	21	20	7	23	8	18
Cheia K	18	20	15	4	17	18	20	15	4	17	18	20	15	4	17	18
M–K (mod 26)	15	4	17	0	18	15	4	17	0	0	3	0	18	19	17	0
Textul clar M	P	E	R	A	S	P	E	R	A	A	D	A	S	T	R	A

Textul clar este ***M*** = PERASPERAADASTRA.

3.2. Algoritmul de criptare *Playfair*

3.2.1. Scurt istoric

Cu toate că poartă numele baronului Lyon Playfair, algoritmul a fost inventat de prietenul acestuia, Charles Wheatstone și descris pentru întâia dată într-un document la 26 martie 1854. La început a fost respins de "British Foreign Office" deoarece a fost considerat foarte greu de înțeles. Atunci când Wheatstone s-a oferit să demonstreze că în 15 minute va învăța să folosească algoritmul 3 băieți din 4 din școala aflată în apropiere, secretarul biroului de externe i-a răspuns: „Da, este foarte posibil, însă nu îi vei putea învăța să fie buni diplomați".

După crearea algoritmului, baronul Playfair a convins guvernul britanic să adopte acest algoritm pentru uz oficial și de aceea poartă numele său și nu al creatorului, Wheatstone. Algoritmul a fost utilizat de către armata britanică în războiul cu burii din Africa de Sud, iar versiuni modificate au fost folosite tot de britanici în primul război mondial cât și de armata australiană în cel de-al doilea război mondial.

Din punct de vedere al criptografiei moderne, algoritmul de criptare Playfair este unul învechit, chiar primitiv. Orice calculator personal modern poate găsi (sparge) cheia și descifra mesajul într-un interval de timp de câteva secunde sau chiar sutimi de secunde, folosind software-ul potrivit. Unii dintre cei mai iscusiți criptanaliști sau chiar unii experți în cuvinte încrucișate pot sparge mesajul criptat în câteva minute folosind doar un creion și o foaie de hârtie.

Cu toate că este un algoritm depășit din toate punctele de vedere, algoritmul Playfair este unul dintre primii algoritmi ce folosește principiile moderne ale cifrurilor bloc. Studierea acestui algoritm vă poate oferi o mai bună înțelegere intuitivă a criptografiei moderne fără a folosi cunoștințe complexe de matematică sau teoria numerelor.

3.2.2. Descrierea generală a algoritmului Playfair

Criptarea Playfair implică parcurgerea următorilor pași:

a) pregătirea textului ce urmează a fi criptat;
b) construirea matricei de criptare;
c) construirea mesajului criptat.

3.2.3. Exemplu de criptare Playfair

a) Pregătirea textului ce urmează a fi criptat

Acest prim pas implică scrierea tuturor literelor cu majuscule, în perechi, fără spații și punctuație. Toate literele '***J***' din text vor fi înlocuite de '***I***' (în exemplul de mai jos, nu există litera '***J***').

Fie mesajul care trebuie criptat este:

m = *Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.*

Acesta va deveni mai întâi:

CONGRESS SHALL MAKE NO LAW RESPECTING AN ESTABLISHMENT OF RELIGION OR PROHIBITING THE FREE EXERCISE THEREOF OR ABRIDGING THE FREEDOM OF SPEECH OR OF THE PRESS OR THE RIGHT OF THE PEOPLE PEACEABLY TO ASSEMBLE AND TO PETITION THE GOVERNMENT FOR A REDRESS OF GRIEVANCES

Apoi, după împărțirea lui în grupuri de două litere, obținem:

CO NG RE SS SH AL LM AK EN OL AW RE SP EC TI NG AN ES TA BL IS M EN TO FR EL IG IO NO RP RO HI BI TI NG TH EF RE EE EX ER CI SE TH ER EO FO RA BR ID GI NG TH EF RE ED OM OF SP EE CH OR OF TH EP RE SS OR TH ER IG HT OF TH EP EO PL EP EA CE AB LY TO AS SE MB LE AN DT OP ET IT IO NT HG OV ER NM EN TF OR AR ED RE SS OF GR IE VA NC ES

Pasul următor în pregătirea textului pentru criptare este inserarea unei litere '***Q***', '***X***' sau '***Z***' (care sunt literele cel mai rar întâlnite în vocabularul limbii engleze) între fiecare cuplu dublură de litere. De exemplu, cuvântul "*FR **EE** DO M*" din exemplu de mai sus va deveni "*FR **EX E**D OM*". Din cauza repetării de trei ori a literei *S* între primele 2 cuvinte ale exemplului ("*CO NG RE **SS S**H AL*") acestea vor fi rescrise ca și "*CO NG RE **SX SZ S**H AL L*".

Această regulă a literelor duble a fost introdusă din două motive:

1. deoarece literele duble sunt foarte des întâlnite în limba engleză, iar aceasta poate ajuta un criptanalist;
2. pentru a reduce numărul de cuvinte ușor de intuit la o primă vedere în mesajul ce trebuie criptat (de ex "*miss*", "*missing*", etc).

Pasul final al pregătirii textului de criptat este adăugarea unei litere adiționale aleasă de persoana care criptează mesajul în cazul în care există un număr impar de litere la pasul anterior. Textul final, pregătit de criptare, al exemplului nostru va fi:

CO NG RE SX SZ SH AL LM MA KE NO LA WR ES PE CT IN GA NE ST AB LI SH ME NT OF RE LI GI ON OR PR OH IB IT IN GT FR EX EZ EX ER CI SE TH ER EO FO RA BR RI DG IN GT HE FR EX ED OM OF SP EX EC HO RO FT HE PR ES SO RT HE RI GH TO FT HE PE OP LE PE AC EA BL YT OA SX SE MB LE AN DT OP ET IT IO NT HE GO VE RN ME NT FO RA RE DR ES SO FG RI EV AN CE SB

Observație: în cazul grupului de litere SS din grupul *ES SO* nu a fost folosite literele *Q*, *X* sau *Z* deoarece cei doi *S* nu fac parte din același grup de două litere.

b) Construirea matricei de criptare

La acest pas mai întâi se vor scoate din cheia de criptare literele care se repetă, începând de la a doua apariție a acestora, Spre exemplu, dacă cheia *k = dublura*, ea se va transforma în *dublra*.

Pentru criptarea mesajului nostru vom folosi drept cheie *k = First Amendment*, care va deveni după prelucrare *FIRST AMEND*.

După aceasta se construiește matricea de criptare, care, pentru alfabetul englez cu 26 de litere, poate fi una de dimensiunile 5x5. În caz general, dacă alfabetul limbii, în care este scris mesajul ce trebuie criptat, are un alt număr de litere, matricea poate fi alta, spre exemplu 6x5, 5x6, 6x4, 4x6, etc., în care pot fi plasate toate (sau aproape toate) literele alfabetului. În cazul în care nu sunt suficiente spații pentru a plasa toate literele alfabetului – se poate proceda așa cum arătat mai sus, adică să înlocuim litera *J* cu *I*, sau se poate de eliminat litera (literele) cel mai rar întâlnită

în limba respectivă, iar la decriptare, intuitiv ea va fi restabilită. În cazul în care matricea are mai multe celule decât litere, celulele libere pot fi completate cu orice simboluri.

Apoi matricea va fi completată după algoritmul:

- începând cu colțul din stânga sus, se va completa pe linie cheia de la pasul precedent;
- se va completa matricea cu celelalte litere ale alfabetului, cu excepția literei *J*, luate în ordine alfabetică.

Pentru exemplul nostru vom obține următoare matrice de criptare:

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

Observație: Cu cât cheia utilizată este mai lungă, cu atât textul cifrat va fi mai greu de criptanalizat. Metoda cea mai des folosită pentru utilizarea unei chei lungi era memorarea unor propoziții scurte (3-5 cuvinte) ușor de reținut.

c) Construirea mesajului criptat

Perechile de litere din textul inițial vor fi criptate după algoritmul:

1. dacă cele două litere sunt în linii și coloane diferite, fiecare literă va fi înlocuită de litera aflată pe aceeași linie dar pe coloana celeilalte litere din cuplul curent. De exemplu, cuplul ***NP*** va fi criptat ca și ***EQ***;
2. dacă cele două litere sunt pe aceeași linie a matricei, fiecare va fi înlocuită de următoarea de pe linia curentă; ultima literă din linie se va înlocui cu prima literă din aceeași linie. De exemplu, cuplul **IT** va fi codificat ca și **RF**;
3. în mod similar, dacă literele sunt pe aceeași coloană, vor fi înlocuite fiecare de cea aflată imediat pe aceeași coloană dar cu o linie mai jos; ultima literă din coloană se va înlocui cu prima literă din aceeași coloană. De exemplu cuplul **CW** va fi codificat **OI** (deoarece *W* este ultima de pe coloană și nu are altă literă sub ea, va fi criptată în prima literă din aceeași coloană).

Folosind mesajul inițial și cheia prelucrate la pașii anteriori precum și matricea de la pasul *b)*, vom obține următorul text criptat:

OWEHEGRYTYNQBVOAEMGDMQVBXINRXGKISMBEDNTFBLOF
NQENDSLIEGOFCRQMPIXEQCFCRFSMKRISGRDXGRGEOMRNSK
GEMPILFEGFSREKSMKRGNISGRNAWCLIRQGRMGCQIPIFGNXE
NRIQSFGNSRHKIUIFGNXGPQPAXGMBNMLVZSLMRYRNACPAM
DKDPQDRRFMWDSGNCPXASEENDSILFEEGETNRIQRBSRAXMDG
MFH

3.2.4. Decriptarea mesajului

Pentru a decripta un mesaj folosind algoritmul Playfair, vom inversa toți pașii urmați la criptare. Textul criptat îl împărțim în perechi (nu avem nevoie de pașii pregătitori din etapa de criptare):

OW EH EG RY TY NQ BV OAEM GD MQ VB XI NR XG KI SM BE DN TF BL OF NQ EN DS LI EG OF CR QM PI XE QC FC RF SM KR IS GR DX GR GE OM RN SK GE MP IL FE GF SR EK SM KR GN IS GR NA WC LI RQ GR MG CQ IP IF GN XE NR IQ SF GN SR HK IU IF GN XG PQ PA XG MB NM LV ZS LM RY RN AC PA MD KD PQ DR RF MW DS GN CP XA SE EN DS IL FE EG ET NR IQ RB SR AX MD GM FH

În mod analogic, având aceeași cheie ca și la criptare, vom obține aceeași matrice:

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

Pentru decriptare vom transforma perechile de litere folosind aceleași reguli ca și la criptare, cu următoarele precizări:

1. dacă cele două litere sunt în linii și coloane diferite, decriptarea se face exact la fel ca și criptarea;
2. dacă cele două litere sunt pe aceeași linie a matricei, fiecare va fi înlocuită de *precedenta* de pe linia curentă; prima literă din linie se va înlocui cu ultima literă din aceeași linie;
3. dacă literele sunt pe aceeași coloană, vor fi înlocuite fiecare de cea aflată imediat pe aceeași coloană dar cu o linie mai sus; prima literă din coloană se va înlocui cu ultima literă din aceeași coloană.

CO NG RE SX SZ SH AL LM MA KE NO LA WR ES PE CT IN GA NE ST AB LI SH ME NT OF RE LI GI ON OR PR OH IB IT IN GT FR EX EZ EX ER CI SE TH ER EO FO RA BR RI DG IN GT HE FR EX ED OM OF SP EX EC HO RO FT HE PR ES SO RT HE RI GH TO FT HE PE OP LE PE AC EA BL YT OA SX SE MB LE AN DT OP ET IT IO NT HE GO VE RN ME NT FO RA RE DR ES SO FG RI EV AN CE SB

Mesajul poate fi acum citit dacă scoatem spațiile dintre cuplurile de litere și adăugăm spații noi, în funcție de limba folosită și de logica mesajului:

CONGRESS SHALL MAKE NO LAW RESPECTING AN ESTABLISHMENT OF RELIGION OR PROHIBITING THE FREE EXERCISE THEREOF OR ABRIDGING THE FREEDOM OF SPEECH OR OF THE PRESS OR THE RIGHT OF THE PEOPLE PEACEABLY TO ASSEMBLE AND TO PETITION THE GOVERNMENT FOR A REDRESS OF GRIEVANCES

Sarcină de laborator:

Sarcină 3.1. De implementat algoritmul Playfair în unul din limbajele de programare pentru mesaje în limba română (31 de litere). Valorile caracterelor textului sunt cuprinse între 'A' și 'Z', 'a' și 'z' și nu sunt premise alte valori. În cazul în care utilizatorul introduce alte valori - i se va sugera diapazonul corect al caracterelor. Lungimea cheii nu trebuie să fie mai mică de 7. Utilizatorul va putea alege operația - *criptare* sau *decriptare*, va putea introduce *cheia*, *mesajul* sau *criptograma* și va obține *criptograma* sau *mesajul decriptat*. Faza finală de adăugare a spațiilor noi, în funcție de limba folosită și de logica mesajului – se va face manual.

Sarcină 3.2. De implementat algoritmul Vigenere în unul din limbajele de programare pentru mesaje în limba română (31 de litere), acestea fiind codificate cu numerele 0, 1, ... 30. Valorile caracterelor textului sunt cuprinse între 'A' și 'Z', 'a' și 'z' și nu sunt premise alte valori. În cazul în care utilizatorul introduce alte valori - i se va sugera diapazonul corect al caracterelor. Lungimea cheii nu trebuie să fie mai mică de 7. Criptarea și decriptarea se va realiza în conformitate cu formulele din modelul matematic prezentat mai sus. În mesaj mai întâi trebuie eliminate spațiile, apoi toate literele se vor transforma în majuscule. Utilizatorul va putea alege operația - *criptare* sau *decriptare*, va putea introduce *cheia*, *mesajul* sau *criptograma* și va obține *criptograma* sau *mesajul decriptat*.

Notă: Termen de realizare - 2 săptămâni.