A. Noţiuni de bază ale criptografiei

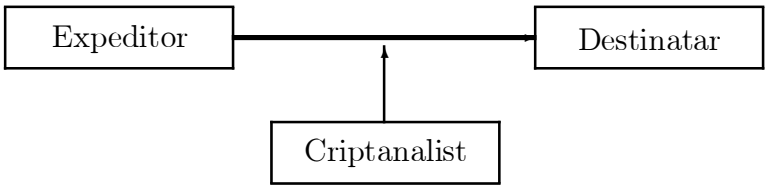
Cuprins

1. Definiţii şi notaţii preliminare.
2. Sisteme simetrice de criptare:
   (a) Cifruri de substituţie;
   (b) Cifruri de permutare.
3. Criptare cu cheie publică.
   (a) Consideraţii generale;
   (b) Funcţii neinversabile;
   (c) Trapa secretă.

A1. Definiţii şi notaţii preliminare

Termenul general folosit pentru scrierea secretă este *criptografie* (format din cuvintele greceşti *cryptos* – ascuns şi *grafie* – scriere). Domeniul cuprinde atât operaţia de cifrare a unui text, cât şi eventualele încercări de descifrare şi de aflare a cheii de criptare. În unele lucrări, cadrul general de lucru este numit *criptologie*, termenul de *criptografie* desemnând numai operaţia de cifrare şi descifrare legală.

Situaţia generală de care se ocupă criptografia este următoarea:



Expeditorul (personalizat în majoritatea lucrărilor cu apelativul *Alice*) doreşte să trimită destinatarului (numit *Bob*) un mesaj printr-un canal de comunicaţie, canal cu un grad ridicat de nesiguranţă. Această insecuritate o prezintă un adversar criptanalist (*Oscar*) care doreşte – din diverse motive – să cunoască conţinutul mesajului, deşi acesta nu îi este destinat.

Hackerul *Oscar* poate avea două tipuri de comportament:

- Pasiv: el se mulţumeşte să intercepteze mesajele şi să le citească, folosindu-le în scop personal;

- Activ: doreşte să modifice mesajul sau să introducă propriile sale mesaje, în intenţia de a trece faţă de *Bob* drept *Alice*. În acest caz, mesajul va trebui să verifice – înafară de condiţia de confidenţialitate – şi pe cea de autenticitate: *Bob* trebuie să fie sigur că mesajul primit a fost de la *Alice*.

  În unele cazuri, problema se poate complica prin faptul că există anumite mesaje pe care *Alice* neagă că îi aparţin, deşi le-a trimis chiar ea. În acest caz trebuie prevăzute anumite protocoale care să întărească proprietăţile de autentificare; proprietăţi care să o silească pe *Alice* să îşi recunoască propriile mesaje.

În toate aceste scenarii nu există personaje pozitive sau negative. Orice serviciu de criptare/decriptare are şi o secţie de criptanaliză. Se pot da numeroase exemple din istorie care să arate rolul pozitiv al lui *Oscar* în anumite situaţii. În general, într-un astfel de scenariu (expeditor, destinatar, criptanalist) nimeni nu are încredere în nimeni. Variantele studiate în care *Alice* are încredere în *Bob* sau invers, sunt mult mai simple şi, de aceea, extrem de rare.

Să stabilim întâi terminologia folosită uzual:

Un mesaj în forma sa originară este numit *text clar*. Expeditorul rescrie acest mesaj folosind o metodă cunoscută numai de el (eventual şi de destinatar); spunem că el *criptează* (sau *cifrează*) mesajul, obţinând un *text criptat*.

Algoritmul care realizează această transformare se numeste *sistem de criptare*. Formal,

Definiţia 1 *Un sistem de criptare este o structură* $(\mathcal{C}, \mathcal{K}, \mathcal{O})$*, unde:*

- $\mathcal{C}= \{w \mid w \in V^*\}$ *este mulţimea textelor clare, scrise peste un alfabet nevid* V *(uzual* $V = \{0, 1\}$*).*
- $\mathcal{K}$ *este o mulţime de elemente numite chei; Fiecare cheie* $K \in \mathcal{K}$ *determină o metodă de criptare* e_K *şi o metodă de decriptare* d_K *cu proprietatea* $d_K(e_K(w)) = w \ \forall w \in \mathcal{C}$*.*
- $\mathcal{O}= \{w \mid \exists k \in \mathcal{K},\ \alpha \in \mathcal{C}, w = e_K(\alpha)\}$*.*

  Elementele lui $\mathcal{O}$ *se numesc texte criptate*

Dacă funcţia e_K este bijectivă, sistemul de criptare este *simetric*.

Pentru ca un sistem de criptare să fie considerat **bun**, trebuie îndeplinite două criterii (enunţate de Francis Bacon, sec. 17):

1. Fiind date e_K şi $\alpha \in \mathcal{C}$, este uşor de determinat $e_K(\alpha)$;
2. Fiind date d_K şi $w \in \mathcal{O}$, este uşor de determinat $d_K(w)$;
3. α este imposibil de determinat din w, fără a cunoaşte d_K.

La acestea, Bacon adăuga şi o a patra regulă:

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

Această condiţie este utilizată astăzi doar de un subdomeniu strict al criptografiei, numit *steganografie*.

În termeni de complexitate, prin "uşor" se înţelege folosirea unui algoritm polinomial de grad mic – preferabil algoritm liniar; o problemă se consideră "imposibilă" dacă pentru rezolvarea ei nu se cunosc decât algoritmi de complexitate exponenţială.

Exemplul 1 *Unul din primele sisteme de criptare cunoscute este sistemul de criptare Cezar. Conform istoricului Suetoniu, el a fost folosit de Cezar în corespondenţa sa.*

Să considerăm alfabetul latin scris, în ordine

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

Fie k un număr întreg din intervalul $[0, 25]$. El se va numi "cheie de criptare". Rescriem alfabetul latin permutat ciclic, începând însă cu litera având numărul de ordine k (litera A are numărul de ordine 0). Această nouă scriere o aşezăm sub prima scriere, astfel (am presupus $k = 2$):

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

Dacă Alice are un text clar pe care vrea să-l cripteze cu sistemul Cezar, ea va proceda astfel:

Să presupunem că acest text clar este NIMIC NOU. Alice va aşeza sub fiecare literă a acestui text, litera aflată pe linia a doua din tabelul de sus, astfel:

$$N\ I\ M\ I\ C\ N\ O\ U$$
$$P\ K\ O\ K\ E\ P\ Q\ W$$

Textul criptat obţinut este PKOKEPQW (din motive suplimentare de securitate, spaţiile dintre cuvinte se ignoră de obicei).

La primirea textului, Bob – care ştie că este vorba de sistemul de criptare Cezar – va proceda astfel: el cunoaşte (fiind destinatar legal) cheia de criptare e_k. Cheia sa de decriptare este e_{26-k}. Pe baza ei Bob va putea construi cele doua linii ale tabelului, după care va proceda ca Alice: scrie textul criptat pe prima linie, iar pe a doua linie determină literele corespunzătoare, conform tabelului.

În cazul $k = 2$, Bob va folosi drept cheie numarul $e_{26-2} = e_{24}$, iar tabelul (litera 24 corespunde lui Y) este

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

Literele PKOKEPQW determină pe a doua linie textul NIMICNOU.

Să studiem puţin poziţia unui criptanalist (*Oscar*). Se presupune că el dispune de facilităţi de calcul excelente, adesea superioare celor de care dispun cei doi parteneri *Alice* şi *Bob*.

Mai mult, se poate considera că *Oscar* cunoaşte sistemul de criptare. S-a constatat că – practic – acest lucru se întâmplă totdeauna. Ce nu cunoaşte însă criptanalistul este cheia $K \in \mathcal{K}$.

În cazul când numărul cheilor posibile este mic (în Exemplul **??** sunt numai 26 chei), aceasta se poate afla foarte uşor după un număr finit de încercări. De aceea, în practică sunt evitate sistemele de criptare cu $card(\mathcal{K})$ mic.

În general, un criptanalist este pus în faţa următoarelor situaţii, care îi solicită strategii diverse de urmat:

1. Ştie numai textul criptat w; în acest caz atacurile sunt direct legate de lungimea textului.

   În sisteme simple – cum este *Cezar* – sunt suficiente texte scurte, pentru că o singură potrivire produce textul - clar. Pentru sistemele mai evoluate, este necesară o lungime mai mare a textului criptat. Metodele de criptanaliză folosesc diverse informaţii statistice relative la alfabetul textului - clar, la frecvenţa apariţiei caracterelor etc.

2. Ştie textul - clar; aici criptanalistul încearcă să determine câteva perechi $(x, e_K(x))$cu $x \in V$; cunoaşterea lor poate ajuta efectiv la decriptarea întregului text criptat. De exemplu, la *Cezar*, dacă se ştie o singură astfel de pereche, ea dezvăluie imediat şi cheia.

   Ca un alt exemplu istoric, aceasta a fost metoda folosită de orientalistul francez Champollion pentru descifrarea hieroglifelor, folosind piatra de la Rosetta.

3. Criptanalistul cunoaşte criptarea unor texte clare alese de el; este atacul cu text clar ales, considerat în majoritatea studiilor de criptanaliză. Această situaţie este adesea superioară celei din cazul precedent; să exemplificăm acest lucru:

   Exemplul 2 *Fie sistemul de criptare Hill, care a jucat un rol destul de important în istorie. Aici textele clare şi cele criptate folosesc alfabetul latin. Vom numerota literele în modul următor*

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

   Toate calculele sunt efectuate modulo 26. *Fie $d \geq 2$ un număr întreg; se defineşte o matrice $M_{d,d}$ inversabilă, cu elemente din Z_{26}.*

   Textul clar w se împarte în blocuri de lungime d : $w = \alpha_1\alpha_2 \ldots \alpha_n$, $|\alpha_i| = d$ (ultimul bloc se completează eventual până ajunge la lungimea d). Textul criptat va fi $x = \beta_1\beta_2 \ldots \beta_n$ unde $\beta_i = M\alpha_i$ $(1 \leq i \leq n)$.

   Să luăm de exemplu $d = 2$ şi $M = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}$, cu inversa $M^{-1} = \begin{pmatrix} 15 & 17 \\ 20 & 9 \end{pmatrix}$.

   Dacă textul clar este $w = HELP$, vom avea

$$\alpha_1 = \begin{pmatrix} H \\ E \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \end{pmatrix} \quad şi \quad \alpha_2 = \begin{pmatrix} L \\ P \end{pmatrix} = \begin{pmatrix} 11 \\ 15 \end{pmatrix}.$$

   Din relaţiile $M\alpha_1 = \beta_1$, $M\alpha_2 = \beta_2$ se obţine textul criptat $x = HIAT$.

   Să ne situăm acum pe poziţia lui Oscar: presupunem că am găsit dimensiunea $d = 2$ şi trebuie să aflăm matricea M (sau – echivalent – M^{-1}). Alegând textul clar $HELP$ vom primi textul criptat $HIAT$.

   Deci Oscar se află acum în faţa următoarei probleme: care este matricea $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ cu $a, b, c, d \in \{0, 1, \ldots, 25\}$ astfel ca $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} 7 & 11 \\ 4 & 15 \end{pmatrix} = \begin{pmatrix} 7 & 0 \\ 8 & 19 \end{pmatrix}$.
   Vom avea

$$M = \begin{pmatrix} 7 & 0 \\ 8 & 19 \end{pmatrix} \times \begin{pmatrix} 7 & 11 \\ 4 & 15 \end{pmatrix}^{-1} = \begin{pmatrix} 7 & 0 \\ 8 & 19 \end{pmatrix} \times \begin{pmatrix} 19 & 19 \\ 14 & 21 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}$$

   Să presupunem acum că Oscar lucrează în ipoteza (2)*; a aflat perechea (text - clar, text - criptat)= $(SAHARA, CKVOZI)$. Scriind ecuaţia matricială*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} 18 & 7 & 17 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 21 & 25 \\ 10 & 14 & 8 \end{pmatrix},$$

va afla $a = 3$, $c = 2$, dar b şi d vor fi nederminate. Deci, deşi textul aflat este mai lung, el nu permite aflarea matricii de criptare M. Va şti doar că orice matrice inversabilă de forma $M = \begin{pmatrix} 3 & b \\ 2 & d \end{pmatrix}$ poate fi baza sistemului de criptare Hill.

4. Ştie cheia de criptare; acum *Oscar* va cunoaşte cheia e_K şi încearcă să determine d_K înainte de interceptarea mesajelor criptate.

   Aceasta este situaţia tipică sistemelor de criptare cu cheie publică: cheia de criptare e_K este cunoscută public cu mult înainte de a fi folosită pentru criptare. Deci criptanalistul are la dispoziţie destul de mult timp pentru prelucrarea ei; după ce se primesc mesaje criptate, timpul devine *scump*, şi el trebuie să fie scurtat cât mai mult.

A2. Sisteme simetrice de criptare

În general, sistemele de criptare clasice se numesc şi *sisteme simetrice*. Motivul este acela că odată cu aflarea cheii de criptare e_K, cheia de decriptare se obţine imediat, fiind funcţia inversă. Folosind Exemplul **??**, la sistemul Cezar e_{26-k} este cheia de decriptare pentru cheia de criptare e_K.

Sistemele de criptare simetrice se împart în două clase mari: cifruri de permutare şi cifruri cu substituţie.

A.2.1. Cifruri de permutare

La aceste sisteme de criptare, textul clar se împarte în blocuri de n ($n \geq 2$) caractere, după care fiecărui bloc i se aplică o permutare fixată.

Exemplul 3 *Să presupunem că vom lua drept cheie de criptare permutarea* $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. *Atunci un text clar, de exemplu FLOARE ALBASTRA se împarte în grupuri de câte trei caractere (s-a considerat şi spaţiul drept caracter, notat _)*

FLO ARE _AL BAS TRA

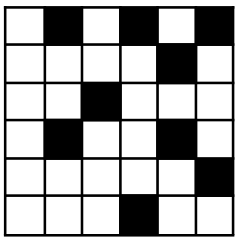
Textul criptat va fi

LFO RAE A_L ABS RTA

sau – eliminând grupările, LFORAEA LABSRTA.

Un sistem de criptare cu permutări celebru este sistemul *Richelieu*, prezentat şi în literatură de Jules Verne, în romanul *Mathias Sandorf*. Dăm un exemplu de utilizare a unui astfel de sistem.

Fie cartonul 6×6, în care zonele haşurate constituie găuri.



Vrem să criptăm textul

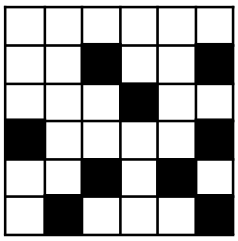
EMINESCU A FOST UN MARE POET NATIONAL

Vom scrie acest text sub forma unui tabel cu şase linii şi şase coloane, astfel:

E	M	I	N	E	S
C	U		A		F
O	S	T		U	N
M	A	R	E		P
O	E	T		N	A
T	I	O	N	A	L

Aplicând cartonul peste acest text, vor rămâne vizibile 9 caractere: MNS TA AN (citite de la stânga la dreapta şi de sus în jos).

Vom roti acum cartonul cu 90^o în sensul acelor de ceasornic. El va arăta



Aşezând acum peste text, rămân vizibile caracterele _F MPTNIL (primul caracter a fost un spaţiu şi l-am marcat cu _ pentru a-l face vizibil).

La a treia rotire a cartonului se obţine similar textul ICSUEETOA, iar la a patra – EEUAOURO_

Deci textul criptat este

MNS TA AN F MPTNILICSUEETOAEEUAOURO

Operaţia de decriptare se realizează similar.

A.2.2. Cifruri de substituţie

Cifrurile de substituţie sunt cele mai utilizate sisteme de criptare simetrice; ele se întâlnesc şi azi, exemple fiind sistemele DES şi AES.

Un astfel de cifru constă în înlocuirea fiecărui caracter cu alt caracter. Există două clase mari de cifruri de substituţie: sisteme monoalfabetice şi polialfabetice.

Sisteme de criptare monoalfabetice

Un astfel de sistem substituie fiecare caracter cu alt caracter – totdeauna acelaşi, indiferent de poziţie. Sistemul de criptare Cezar este un sistem monoalfabetic: odată stabilită cheia de criptare e_K, fiecare caracter cod x se înlocuieşte prin caracterul cod $x + k \ (mod\ 26)$.

Exemplul 4 *Un sistem de criptare afin este determinat de două numare întregi* a, b $(0 \leq a, b \leq 25)$ *cu* $(a, 26) = 1$. *Fiecare literă de cod* x *se înlocuieşte cu* $f(x) = ax + b \ (mod\ 26)$.

De exemplu, pentru $f(x) = 3x + 5$, *codificarea numerică arată astfel:*

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

sau – scris direct pentru caractere

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

Astfel, un text clar PRIMAVARA TARZIE se criptează în YEDPFQFEF KDECDR.

Condiţia $(a, 26) = 1$ *asigură injectivitatea aplicaţiei* $f(x) = ax+b$. *De exemplu, pentru* $f(x) = 10x+1$, A *şi* N *se transformă ambele în* B, *iar* O *nu apare ca imagine în alfabetul substituţiei.*

Decriptarea se realizează tot cu o aplicaţie afină, $g(x) = a^{-1}x + a^{-1}(26 - b)\ (mod\ 26)$. *Astfel, pentru funcţia de criptare* $f(x) = 3x+5$, *decriptarea se realizează cu* $g(x) = 9x+7$.

Să studiem spaţiul cheilor $\mathcal{K}$. *Orice cheie* $K \in \mathcal{K}$ *este determinată complet de valorile întregi* a *şi* b *cu* $(a, 26) = 1$. *Sunt posibile* 12 *valori pentru* a : $1, 3, 5, 7, 9, 11, 15, 19, 21, 23, 25$. *Pentru* b *sunt posibile* 26 *valori, care se iau independent de* a, *cu singura excepţie* $a = 1,\ b = 0$ *(care se exclude deoarece nu conduce la nici o criptare). Deci* $card(\mathcal{K}) = 311$.

Punctul slab al sistemelor de criptare monoalfabetice constă în frecvenţa literelor. Dacă un text criptat este suficient de lung şi se cunoaşte limba în care este scris textul clar, sistemul poate fi spart printr-un atac bazat pe frecvenţa apariţiei literelor într-o limbă.

Sunt construite diverse tabele pentru a da informaţia relativă la frecvenţa apariţiei literelor în fiecare limbă europeană. De obicei, cu cât un text criptat este mai lung, cu atât frecvenţa literelor folosite se apropie de această listă. Aceasta conduce la realizarea câtorva corespondenţe (literă text clar – literă text criptat), ceea ce stabileşte univoc cheia de criptare. Pentru sistemul Cezar este suficientă stabilirea unei singure perechi; pentru sistemul afin trebueisc două perechi etc.

Pentru limba română, un tabel al literelor cele mai frecvent întâlnite este

Literă	Frecvenţă
A	13.04 %
I	12.89 %
E	11.75 %
R	7.39 %
T	6.62 %
N	6.44 %
U	6.44 %
S	5.50 %
C	5.47 %

Literă	Frecvenţă
L	4.58 %
O	3.85 %
D	3.68 %
M	3.33 %
P	2.91 %
F	1.50 %
V	1.26%

(restul caracterelor au o frecvenţă sub 1 %).

Sisteme de criptare polialfabetice

Diferenţa dintre aceste sisteme de criptare şi cele monoalfabetice constă în faptul că substituţia unui caracter variază în text, în funcţie de diverşi parametri (poziţie, context etc.). Aceasta conduce bineînţeles la un număr mult mai mare de chei posibile.

Vom da două exemple de sisteme de criptare polialfabetice, extrem de mult folosite de-a lungul timpului.

Exemplul 5 Sistemul de criptare Playfair*:*

A fost definit de baronul Palyfayr de St. Andrews. Ideea de bază este următoarea;

Din cele 26 *litere ale alfabetului se elimină cea de frecvenţă minimă; să spunem* Q. *Restul se aranjează arbitrar sub forma unui pătrat* 5×5, *cum ar fi*

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

Acest tabel va forma atât cheia de criptare cât şi cea de decriptare. Regulile de criptare (decriptare) sunt:

- *Textul clar este separat în blocuri de câte două caractere, restricţia este ca nici un bloc să nu conţină aceiaşi literă, iar textul să fie de lungime pară. Aceste deziderate se realizează uşor modificând puţin textul clar.*

- *Fiecare bloc se criptează astfel: dacă cele două litere nu sunt plasate în tabel pe aceiaşi linie sau coloană (de exemplu A şi E), se cercetează colţurile dreptunghiului determinat de cele două litere (în cazul nostru A, F, O, E). Perechea AE este criptată în FO. Ordinea este determinată de ordinea liniilor pe care se află literele din textul clar. Astfel, EA se criptează în OF, SF în ZB etc.*

  Dacă cele două litere se găsesc pe aceaşi linie (coloană), se merge cu o poziţie la dreapta (jos) în mod ciclic. Deci CA se criptează în AX, WX în UG, CA în AX etc.

De exemplu, textul clar AFARA PLOUA se criptează în XHHPPDPEPX.

O modificare ciclică a liniilor şi coloanelor tabloului nu modifică criptarea.

Regulile de bază pot fi modificate sau completate după necesităţi. Astfel, se poate adăuga din loc în loc câte o literă falsă (cu frecvenţă foarte redusă, cum ar fi X, Y) care să modifice textul criptat. Pătratul 5×5 poate fi înlocuit cu un dreptunghi 4×6 sau 3×8, cu schimbările corespunzătoare în alegerea literelor care se elimină.

Pentru a păstra cheia în siguranţă, se recomandă memorarea acesteia. Cum o astfel de cheie este extrem de greu de memorat, se foloseşte un cuvânt cheie sau o propoziţie cu toate literele distincte. Acesta cuvânt este scris la începutul tabloului. Spaţiile rămase sunt completate cu restul literelor alfabetului, scrise în ordinea apariţiei lor.

Astfel, în preajma primului război mondial, armata română folosea un dreptunghi 3×8 din care lipseau literele Q şi K. Cuvântul cheie era $ROMANESC$. Un astfel de tablou avea forma

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

Exemplul 6 Sistemul de criptare Vigenere*:*

Numele vine de la baronul francez Blaise de Vigenere ($1523 - 1596$). Construcţia matematica a algoritmului este următoarea:

Considerăm cele 26 litere ale alfabetului, numerotate de la 0 (pentru A) până la 25 (pentru Z), conform tabelului:

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

Cheia este formată dintr-un cuvânt a cărui codificare numerică este $\mathbf{c} = c_0 c_1 \dots c_{k-1}$.

Fie $\mathbf{w} = w_1 w_2 \dots w_n$ *codificarea textului clar care trebuie trnasmis. Textul criptat va fi* $\mathbf{x} = x_1 x_2 \dots x_n$, *unde*

$$x_i = w_i + c_{i \bmod k} \pmod{26} \qquad (*)$$

Să considerăm de exemplu cuvântul cheie FOCAR; deci $k = 5$ *şi* $\mathbf{c} = 5\ 14\ 2\ 0\ 17$.

Dacă vrem să criptăm cu această cheie textul clar NU POT VENI AZI, vom proceda astfel:

Codificarea textului este $\mathbf{w} = 13\ 20\ 15\ 14\ 19\ 21\ 4\ 13\ 8\ 0\ 25\ 8$.

Sub fiecare număr din $\mathbf{w}$ *se aşează câte un număr din* $\mathbf{c}$*; când cheia se termină, ea se reia ciclic, până se termină* $\mathbf{w}$. *Deci vom avea*

13	20	15	14	19	21	4	13	8	0	25	8
5	14	2	0	17	5	14	2	0	17	5	14
18	8	17	14	10	0	18	15	8	17	4	22
S	*I*	*R*	*O*	*K*	*A*	*S*	*P*	*I*	*R*	*E*	*W*

Linia a treia conţine suma modulo 26 *a numerelor de pe primele două linii, iar pe ultima linie s-a scris textul criptat rezultat.*

Decriptarea se realizează similar, scăzând (modulo 26*) din codul caracterului criptat, codul caracterului corespunzător din cheie.*

O variantă a sistemul Vigenere este sistemul Beaufort (amiral englez, autorul şi a unei scale a vânturilor care îi poartă numele); aici relaţia de criptare $(*)$ *este înlocuită cu*

$$x_i = c_{i \bmod k} - w_i \pmod{26}$$

Avantajul sistemului Beaufort constă în faptul că ecuaţia de criptare se aplică şi la decriptare ($w_i = c_{i \bmod k} - x_i$).

Sistemul Vigenere a fost utilizat secole de-a rândul, fiind considerat ca fiind unul din cele mai sigure sisteme de criptare. În 1917 *de exemplu, prestigioasa revistă "Scientific American" îl considera imposibil de atacat. Numai că acest sistem a fost spart de Kasiski încă din* 1860.

Nu prezentăm modalitatea de criptanaliză a lui Kasiski; ea se bazează pe construcţia unor conjecturi asupra lui k (lungimea cheii). După aceea, totul se transformă în atacul a k sisteme de criptare Cezar, problemă aproape banală. Astăzi, orice student poate decripta un text criptat suficient de lung cu Vigenere, în cadrul unei teme de seminar.

A3. Sisteme de criptare cu cheie publică

A.3.1. Consideraţii generale

În sistemele de criptare clasice, *Alice* şi *Bob* îşi aleg o cheie secretă K care defineşte regulile de criptare (e_K) şi decriptare (d_K). În aproape toate cazurile d_K şi e_K coincideau sau se puteau deduce imediat una din alta. Astfel de sisteme sunt numite *sisteme cu cheie privată* deoarece publicarea lui e_K face sistemul extrem de vulnerabil (cazul DES - ului este o excepţie).

Un punct slab al sistemelor cu cheie privată este acela că necesită o comunicare prealabilă a cheii între *Alice* şi *Bob* printr-un canal sigur, înainte de transmiterea mesajului criptat. Practic, acest lucru este tot mai dificil de realizat.

Obiectivul sistemelor de criptare cu cheie publică este acela de a face "imposibil" (asupra acestui termen vom reveni) de obţinut cheia d_K plecând de la e_K. Astfel, regula de criptare e_K poate fi publicată într-un registru public (de unde şi numele sistemelor).

Avantajul constă în faptul că *Alice* (sau oricare altă persoană) poate trimite lui *Bob* un mesaj criptat cu e_K fără a intra în prealabil în contact. *Bob* este singura persoană capabilă să decripteze textul, utilizând cheia sa secretă d_K.

Ideea de sistem de criptare cu cheie publică apare în 1976 şi este prezentat de Diffie şi Hellman. De atunci au apărut diverse astfel de sisteme, a căror securitate este bazată pe probleme calculatorii. Cele mai cunoscute sunt:

- Sistemul RSA: se bazează pe dificultatea descompunerii în factori primi a numerelor mari (de sute de cifre). Este sistemul cel mai larg utilizat în acest moment.

- *Sistemul Merkle - Hellman*: primul sistem definit cu cheie publică, cunoscut şi sub numele de *problema rucsacului*, problemă cunoscută ca fiind NP - completă. Fiind spart din 1984, este adesea utilizat doar pentru exemplificări.

- *Sistemul McEliece*: este bazat pe teoria algebrică a codurilor, decodificarea unui cod liniar fiind de asemenea NP - completă.

- *Sistemul El Gamal*: se bazează pe dificultatea calculului logaritmului discret într-un corp finit.

- *Curbe eliptice*: sunt modificări ale altor sisteme, care utilizează curbe eliptice în loc de corpuri. Operând cu chei de lungime mică, sunt extrem de utile în scrierea datelor de autentificare pe smart - carduri.

A.3.2. Funcţii neinversabile

O observaţie importantă este aceea că că un sistem cu cheie publică nu este necondiţionat sigur: orice persoană – putând să efectueze criptări, nu este exclus să găsească anumite puncte slabe care să îi permită să şi decripteze mesajele. Ideea de bază folosită este aceea de *funcţie neinversabilă*. Să clarificăm puţin acest aspect.

Exemplul 7 *Ne putem imagina uşor străzile cu sens unic dintr-un oraş. Astfel, este uşor ca mergând pe astfel de străzi să ajungi de la A la B, dar este imposibil să ajungi de la B la A. În acest mod, criptarea este privită ca direcţia $A \rightarrow B$; deşi este foarte uşor de parcurs drumul în această direcţie, nu te poţi întoarce înapoi spre A (adică să decriptezi mesajul).*

Exemplul 8 *Să considerăm cartea de telefon a unui oraş mare; cu ajutorul ei este foarte uşor să găsim numărul de telefon al unei anumite persoane. În schimb, este extrem de greu – practic imposibil – să afli persoana care are un anumit număr de telefon. Te afli în situaţia parcurgerii secvenţiale a (cel puţin) unui volum gros, ceea ce conduce la o creştere exagerată a timpului.*

Aceasta dă o sugestie de construcţie a unui sistem de criptare cu cheie publică. Criptarea se face independent de context, literă cu literă. Pentru fiecare literă a textului clar se alege un nume care începe cu acest caracter şi numărul de telefon al persoanei respective va constitui criptarea. Sistemul este polialfabetic; două apariţii diferite ale aceleiaşi litere vor fi codificate foarte probabil cu numere diferite. De exemplu, textul clar SOLIST se poate cripta astfel:

S	*Simion Pavel*	6394502
O	*Olaru Ştefan*	7781594
L	*Lambru Stelian*	6300037
I	*Ilie Romeo*	3134971
S	*Solovean Raluca*	6281142
T	*Tecuceanu Paul*	3359962

Deci, textul criptat va fi

639450 277815 946300 037313 497162 811423 359962.

De remarcat că metoda este nedeterministă; din acelaşi ext clar se pot obţine enorm de multe texte criptate. Pe de-altă parte, orice text criptat conduce la un text clar unic.

Bob va avea la dispoziţie pentru decriptare o carte de telefon scrisă în ordinea crescătoare a numerelor. Aceasta îi va permite să decripteze mesajele cu un algoritm de complexitate logaritmică.

Deci, o funcţie neinversabilă f trebuie să verifice două condiţii:

- Fiind dat x, $f(x)$ este uşor de calculat;
- Calculul lui x din $f(x)$ este imposibil.

De remarcat că, din punct de vedere strict matematic, nu se cunosc astfel de funcţii. A demonstra că există funcţii neinversabile este echivalent cu a demonstra relaţia $\mathcal{P} \neq \mathcal{NP}$, conjectură care stă la baza întregii teorii criptografice. De aceea, termenii folosiţi sunt relativi la complexitatea calculatorie. Astfel, o problemă este:

1. *uşoară* dacă se poate rezolva cu un algoritm liniar;
2. *grea* dacă se poate rezolva cu un algoritm polinomial neliniar;
3. *imposibilă* dacă este $\mathcal{NP}$ - completă.

Am listat anterior o serie de probleme $\mathcal{NP}$ - complete care stau la baza unor sisteme de criptare cu cheie publică.

Exemplul 9 *Să considerăm "problema rucsacului". Ea constă dintr-un vector de dimensiune n $A = (a_1, a_2, \ldots, a_n)$ cu elemente numere întregi pozitive distincte, şi un număr întreg pozitiv k. Trebuiesc aflaţi acei a_i din A (dacă există) a căror sumă este k. Numele intuitiv dat problemei este evident. De exemplu, fie*

$A = (43, 129, 215, 473, 903, 302, 561, 1165, 696, 1523)$ *şi* $k = 3231$.

Se determină $3231 = 129 + 473 + 903 + 561 + 1165$*, care este o astfel de soluţie.*

În principiu o soluţie se poate găsi parcurgând sistematic toate submulţimile lui A şi verificând dacă suma elementelor lor este k. În cazul de sus, aceasta înseamnă $2^{10} - 1 = 1023$ *submulţimi (fără mulţimea vidă), dimensiune acceptabilă ca timp de lucru.*

Ce se întâmplă însă dacă A are câteva sute de componente ? În acest caz se cunoaşte faptul că problema rucsacului este $\mathcal{NP}$ - completă.

Cu ajutorul lui A se poate defini o funcţie f astfel:

Fie $x \in [0, 2^n - 1]$; x poate fi reprezentat în binar ca un cuvânt de lungime n (adăugând eventual 0 *- uri în faţă). $f(x)$ va fi numărul obţinut din A prin însumarea tuturor numerelor a_i aflate pe poziţiile marcate cu* 1 *în reprezentarea binară a lui x. Formal,*

$$f(x) = A \cdot B_x$$

unde B_x este reprezentarea binară a lui x, scrisă ca un vector coloană.

Să definim acum un sistem de criptare bazat pe problema rucsacului. Textul clar este codificat inițial în binar şi segmentat apoi în blocuri de câte n biţi (eventual ultimul bloc este completat la sfârşit cu zerouri). Fiecare bloc rezultat este apoi criptat calculând valoarea corespunzătoare a funcţiei f.

Pentru alfabetul latin sunt suficienţi 5 biţi pentru codificarea binară a literelor şi a spaţiului. Mai exact, dacă asociem literelor A - Z reprezentările binare ale numerelor $1-26$, vom avea:

	–	00000	A	–	00001	B	–	00010
C	–	00010	D	–	00011	E	–	00101
F	–	00110	G	–	00111	H	–	01000
I	–	01001	J	–	01010	K	–	01011
L	–	01100	M	–	01101	N	–	01110
O	–	01111	P	–	10000	Q	–	10001
R	–	10010	S	–	10011	T	–	10100
U	–	10101	V	–	10110	W	–	10111
X	–	11000	Y	–	11001	Z	–	11010

Să considerăm un text clar, FLOARE DE COLT de exemplu. Cum fiecare caracter se codifică în 5 biţi, în fiecare bloc intră două caractere: FL OA RE _D E_ CO LT. Codificând,se obţin şapte blocuri de câte 10 biţi:
0011001100 0111100001 1001000101 0000000100 0000000101 0001101111 0110010100
care conduc la textul criptat: $(1814, 3243, 3204, 1165, 1118, 5321, 1811)$.

Să considerăm sistemul de criptare definit în Exemplul **??**. Dacă îl privim ca un sistem clasic, criptanalistul trebuie să afle vectorul de bază A şi apoi să rezolve problema rucsacului.

Dacă el foloseşte metoda textelor clare alese, îl va afla uşor pe A: este suficient să trimită n texte clare cu cîte un singur 1 iar restul 0. Problema apare în momentul rezolvării problemei rucsacului; aici atât *Bob* cât şi *Oscar* sunt puşi în faţa aceleiaşi probleme $\mathcal{NP}$ - complete. Ori, practic, doar *Oscar* trebuie să rezolve o problemă dificilă, nu şi *Bob*.

A.3.3. Trapa secretă

Pentru ca *Bob* să nu fie pus în aceaşi situaţie ca şi *Oscar*, el tebuie să dispună de un procedeu care să îi permită să transforme problema $\mathcal{NP}$ - completă publică, într-o problemă uşoară. Acest procedeu este numit *trapă secretă*. În primul exemplu, acest procedeu era cartea de telefon ordonată după numerele de telefon, nu după abonaţi. Să vedem care este trapa secretă în sistemul de criptare din Exemplul **??**:

Exemplul 10 *Sunt clase de probleme ale rucsacului uşor de rezolvat; una din ele o formează vectorii cu creştere mare.*

Spunem că vectorul rucsac $A = (a_1, a_2, \ldots, a_n)$ este cu creştere mare dacă

$$\forall\, j \geq 2, \quad a_j \geq \sum_{i=1}^{j-1} a_i.$$

În acest caz, este suficient să parcurgem vectorul de la dreapta la stânga. Cunoscând k, cercetăm dacă $a_n \leq k$. Un răspuns negativ asigură eliminarea lui a_n din suma căutată.

Dacă răspunsul este afirmativ, a_n trebuie să fie în sumă (pentru suma celorltor elemente a_i rămase nu poate depăşi k). Definim apoi

$$k_1 = k - a_n \text{ dacă } a_n \leq k, \quad k_1 = k \text{ altfel}$$

şi reluăm precedeul pentru valorile k_1 şi a_{n-1}. Algoritmul este liniar şi se opreşte la a_1.

Dacă se face public un vector A cu creştere mare ca bază a sistemului de criptare, atunci decriptarea va fi la fel de simplă atât pentru Bob cât şi pentru Oscar. Pentru a evita acest lucru, vom modifica A în aşa fel încât vectorul rezultat B să nu mai fie cu creştere mare, ci să arate ca un vector obişnuit.

Alegem un număr întreg $m > \sum_{i=1}^{n} a_i$, pe care îl numim modul. Fie t cu $(m,t) = 1$ un întreg numit multiplicator; deci există un număr p cu $tp \equiv 1 \pmod m$.

Pe baza lor construim vectorul $B = (b_1, b_2, \ldots, b_n)$ unde $b_i \equiv ta_i \pmod m$. El este oferit public pentru criptare, iar m şi p sunt cunoscute numai de Bob, formând trapa sa secretă.

De exemplu, dacă se ia $m = 1590$ şi $t = 43$ (deci $p = 37$), atunci vectorul ruscac cu creştere mare $A = (1, 3, 5, 11, 21, 44, 87, 175, 701$ se transformă în $B = (43, 129, 215, 473, 903, 302, 561, 1165, 697, 1523)$.

Orice expeditor va folosi pentru criptarea mesajului acest vector B. Ce va face Bob la primirea unui text criptat c' de n biţi ? În transformă în zecimal, calculează $c = c'p \pmod m$ şi va efectua decriptarea folosind vectorul A. Soluţia este o secvenţă unică de n biţi, care formează textul clar.

Principiile generale de cosntrucţie a sistemelor de criptare cu cheie publică sunt:

1. Se alege o problemă dificilă $\mathcal{P}$, a cărei rezolvare este imposibilă în conformitate cu teoria complexităţii.

2. Se selectează o subproblemă P' a lui P, rezolvabilă în timp polinomial (preferabil liniar).

3. Se aplică o transformare problemei P' astfel ca să se obţină o problemă P_1 care să nu semene deloc cu P' şi să fie foarte apropiată de problema P.

4. Se face publică problema P_1 şi se descrie modalitatea de criptare. Informaţia referitoare la modul de obţinere a lui P' din P constituie trapa secretă.

5. Se construiesc detaliile sistemului de criptare, în care principiile de lucru să difere esenţial pentru destinatar în comparaţie cu criptanalistul: în timp ce primul va folosi trapa secretă şi va rezolva problema P', al doilea va trebui să rezolve problema P_1 – care prin asemănarea cu P este imposibilă algoritmic.