(19) OFICIUL DE STAT
PENTRU INVENŢII ŞI MĂRCI
Bucureşti

ROMÂNIA

(11) **RO 129751 A2**
(51) **Int.Cl.**
H04L 9/28 (2006.01);
G06N 3/12 (2006.01)

(12) # CERERE DE BREVET DE INVENŢIE

(21) Nr. cerere: **a 2013 00149**

(22) Data de depozit: **14.02.2013**

(41) Data publicării cererii:
29.08.2014 BOPI nr. **8/2014**

(71) Solicitant:
• **UNIVERSITATEA TEHNICĂ DIN CLUJ-NAPOCA**, *STR.MEMORANDUMULUI NR.28, CLUJ-NAPOCA, CJ, RO*

(72) Inventatori:
• **BORDA MONICA ELENA**, *STR. TARNITA NR.8, AP.4, CLUJ-NAPOCA, CJ, RO;*
• **TORNEA OLGA**, *STR. ONISIFOR GHIBU NR. 39, CLUJ-NAPOCA, CJ, RO;*
• **TEREBES ROMULUS**, *ALEEA GODEANU NR. 5, AP. 19, CLUJ-NAPOCA, CJ, RO;*
• **MALUTAN EMIL RAUL**, *STR. MINTIULUI NR. 6, AP. 33, GHERLA, CJ, RO*

(74) Mandatar:
CABINET DE PROPRIETATE INDUSTRIALĂ CIUPAN EMILIA, *STR.MESTECENILOR NR.6, BL.E, AP.2, CLUJ NAPOCA, JUDEŢUL CLUJ*

(54) # METODĂ ŞI SISTEM DE CRIPTARE DE TIP OTP BAZATE PE SECVENŢE ALEATOARE DETERMINATE DIN STRUCTURI ADN

(57) **Rezumat:**

Invenţia se referă la o metodă şi la un sistem de criptare de tip OTP (One-Time-Pad), bazate pe secvenţe aleatoare determinate din structuri ADN. Metoda de criptare, conform invenţiei, constă din transmiterea, de la o parte emiţătoare la o parte receptoare, a unei chei secrete, odată cu un mesaj criptat, cheia secretă fiind formată dintr-un antet compus dintr-un cod (k_1) pe 2 biţi, ce reprezintă modul de formare a cheii de criptare, dintr-un cod (k_2) pe 3 biţi, reprezentând numărul de ID-uri alestructurilor ADN utilizate la obţinerea cheii de criptare, şi dintr-o înşiruire (k_3) a ID-urilor structurilor ADN utilizate la obţinerea cheii de criptare (K_{ADN}). Sistemul de criptare, conform invenţiei, este alcătuit din două părţi: o parte emiţătoare de mesaje şi o parte receptoare, fiecare dintre cele două părţi fiind compusă dintr-o bază de date ADN (**BD ADN**), publică sau privată, identică la ambele părţi, un bloc al datelor de intrare (**DI**), un generator de cheie ADN (**Gen K_{ADN}**), un convertor al cheii ADN în cheie binară (**Conv ADN-B**), un sumator modulo 2 (**S**), un bloc de criptare (**E***) a datelor de intrare cu care se generează cheia secretă, şi un bloc de decriptare (**D***) a acesteia, mesajul criptat şi cheia secretă fiind transmise de la emiţător la receptor, acesta din urmă obţinând datele de intrare prin decriptarea cheii secrete, şi continuând cu generarea cheii ADN pe care o utilizează la decriptarea mesajului recepţionat.

Revendicări: 6
Figuri: 9

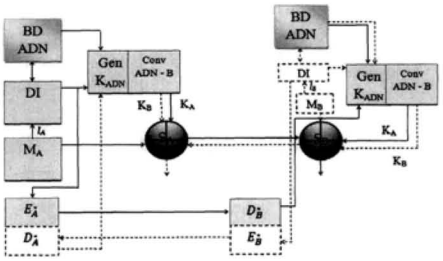

Fig. 7

RO 129751 A2

Metodă şi sistem de criptare de tip OTP bazate pe secvenţe aleatoare determinate din structuri ADN

Invenţia se referă la o metodă şi la un sistem de criptare de tip OTP (One-Time-Pad) bazate pe utilizarea secvenţelor aleatoare (SA) de orice lungime, utilizând structuri ADN (acid dezoxiribonucleic), sisteme demonstrate ca fiind de nespart [Sch-96], [Sta-99].

La ora actuală există generatoare pseudo-aleatoare ADN de tip congruenţial [PP-06], [RG-97], [PTD-06], care ca orice generator pseudoaleator generează secvenţe de lungime finită, iar parametrii generatorului (seed) trebuie transmişi celor 2 părţi. Există şi propuneri de secvenţe aleatoare de tip ADN [CG-08] cu proprietăţi foarte apropiate de ale unei secvenţe aleatoare pure, obţinuţi prin utilizarea mai multor valori iniţiale (seed-uri), dar a căror utilizare în aplicaţii criptografice este practic imposibilă.

Un sistem de criptare OTP, singurul sistem de nespart, are marele dezavantaj că, cheia (aleatoare), trebuie transmisă utilizatorilor.

Problema tehnică pe care o rezolvă invenţia constă în elaborarea unei metode de criptare şi a unui sistem de comunicare sigur, de nespart, bazate pe generarea unor chei de criptare a mesajelor de transmis între două părţi sau utilizate într-un sistem de stocare de date, chei cu o lungime suficient de mare astfel încât să asigure codificarea întregului mesaj, indiferent de tipul acestuia (text, imagine, sunet, video) şi care să nu necesite transmiterea cheii de criptare între participanţii la comunicaţie.

Metoda de criptare, conform invenţiei urmăreşte obţinerea de chei unice de criptare constând în secvenţe aleatoare bazate pe structuri ADN din baze de date biologice existente (publice sau private) sau realizate prin sinteză. Secvenţele aleatoare propuse se obţin din secvenţe ADN (cromozomi, gene etc.) stocate in baze de date biologice existente, cele mai mari fiind: [DDBJ], [ENA], [GenBank]. Structura unei secvenţe ADN dintr-un organism viu (om, animal, plantă) are un caracter aleator probat prin faptul că nu poate fi comprimată sau factorul de compresie este extrem de mic [NW-99], [CLLX-09], [FLCB-11], [DRCX+10], [RA-11].

Sistemul de criptare de tip OTP conform invenţiei este destinat utilizării în transmisiuni de tip duplex sau în sisteme de stocare, acesta asigurând protecţia informaţiei printr-un proces de criptare - decriptare în care cheile de criptare constau în secvenţe aleatoare obţinute din structuri ADN.

Sistemul, conform invenţiei, utilizat în transmisiuni de tip duplex este alcătuit din două părţi, una emiţătoare de mesaje şi cealaltă receptoare, fiecare dintre cele două părţi fiind compusă dintr-o bază de date ADN (BD ADN), publică sau privată , identică la ambele părţi, un bloc al datelor de intrare (DI), un generator de cheie ADN (Gen K_{ADN}), un convertor al cheii ADN în cheie binară (Conv ADN-B), un sumator modulo 2 (S) cu rol de criptare a mesajului, un bloc de criptare (E*) a datelor de intrare cu care se generează cheia secretă şi un bloc de decriptare a acesteia (D*), mesajul criptat şi cheia secretă fiind transmise de la emiţător către receptor, acesta din urmă (receptorul) obţinând datele de intrare prin decriptarea cheii secrete şi continuând cu generarea cheii ADN pe care o utilizează la decriptarea mesajului recepţionat.

Se dă în continuare un exemplu de realizare a metodei de criptare bazate pe secvenţe aleatoare de structuri ADN în legătură cu figurile 1-9 care reprezintă:

- figura 1 – secvenţă de cromozomi din structura ADN a pisicii domestice (Felis catus)
- figura 2 – secvenţă de cromozomi din structura ADN a porumbului (Zea mays)
- figura 3 – secvenţă de cromozomi din structura ADN a omului (Homo sapiens)
- figura 4 - formatul cheii secrete
- figura 5 – exemple de secvenţe ADN în format GenBank şi ID-urile lor
- figura 6 – formatul cheii secrete pentru exemplul prezentat
- figura 7 – schema bloc pentru criptosistemul OTP bazat pe secvenţe aleatoare de tip ADN utilizat în transmisiuni de tip duplex
- figura 8 – schema bloc pentru criptosistemul OTP utilizat pentru stocare de date bazat pe secvenţe aleatoare de tip ADN - partea de criptare a unităţii de înscriere
- figura 9 – schema bloc pentru criptosistemul OTP utilizat pentru stocare de date bazat pe secvenţe aleatoare de tip ADN - partea de decriptare a unităţii de citire.

Cheia de criptare/decriptare a unui criptosistem OTP bazat pe secvenţe ADN se generează pornind de la informaţii cunoscute, validate. Se prezintă în continuare fundamentele pe care se construiesc metoda şi sistemul de criptare de tip OTP bazate pe secvenţe aleatoare determinate din structuri AND, conform invenţiei.

- Codul genetic este format din 4 baze [CDLT04]: Adenina – A, Citozina – C, Guanina – G, Timina - T care vor fi substituite cu o codare binară uniformă conform tabelului 1.

A	00
C	01
G	10
T	11

Tabel 1. Tabel de conversie ADN → binar

Aceste substituţii sunt uşor de realizat prin structura de selecţie switch-case a unui limbaj de programare.

- Lungimile secvenţelor ADN aflate în bazele de date genetice sunt variabile: de la zeci de baze (o genă) până la sute de milioane de baze (un cromozom). Figurile 1-3 corelate cu tabele 2-4, demonstrează faptul că cromozomii din structura ADN a unor organisme vii precum pisica domestică (figura 1 – tabelul 2), porumbul (figura 2 – tabelul 3) sau organismul uman (figura 3 – tabelul 4) au lungimi de ordinul zecilor sau sutelor de milioane de perechi de baze (bp).

Cromozom	Lungimea în bp
A1	239,302,903
A2	169,043,629
A3	142,459,683
B1	205,241,052
B2	154,261,789
B3	148,491,654
B4	144,259,557
C1	221,441,202
C2	157,659,299
D1	116,869,131
D2	89,822,065
D3	95,741,729
D4	96,020,406
E1	63,002,102
E2	64,039,838
E3	43,024,555
F1	68,669,167
F2	82,763,536
X	126,427,096

2.

Cromozom	Lungimea în bp
1	301,354,135
2	237,068,873
3	232,140,174
4	241,473,504
5	217,872,852
6	169,174,353
7	176,764,762
8	175,793,759
9	156,750,706
10	150,189,435

3.

Cromozom	Lungimea în bp
1	249,250,621
2	243,199,373
3	198,022,430
4	191,154,276
5	180,915,260
6	171,115,067
7	159,138,663
8	146,364,022
9	141,213,431
10	135,534,747
11	135,006,516
12	133,851,895
13	115,169,878
14	107,349,540
15	102,531,392
16	90,354,753
17	81,195,210
18	78,077,248
19	59,128,983
20	63,025,520
21	48,129,895
22	51,304,566
X	155,270,560
Y	59,373,566

4.

Tabele 2, 3, 4. Lungimea secvenţelor ADN din cromozomii organismelor vii:
2. pisică, 3. porumb, 4. om

Pentru ca o secvenţă aleatoare (SA) să poată fi utilizată în aplicaţii criptografice de tip OTP, acea secvenţă trebuie să aibă lungimea cel puţin egală cu a mesajului în clar şi să fie utilizată o singură dată (cheie de sesiune). Acest deziderat se obţine realizând cheile de sesiune de lungimi corespunzătoare mesajului de criptat.

Lungimea necesară a secvenţei aleatoare folosite ca şi cheie de criptare depinde de tipul mesajului de transmis (text, imagine, sunet(cu debit specificat), video (cu debit specificat)), precum şi de tipul de transmitere pentru sunet şi video. Aceste date vor determina lungimea necesară a secvenţei aleatoare folosite ca şi cheie.

Se observă că în bazele de date genomice există secvenţe ADN (cele din cromozomi) de lungime foarte mare ce pot asigura lungimi cerute de diverse aplicaţii. Tabelul 5 prezintă tipuri de mesaje şi mărimea lor, datele fiind obţinute prin măsurători pe fişiere reale.

Text	Imagine	Sunet	Video
15KB/1p	402KB	5.34MB/278s (160kbps)	113 MB/1140s (audio 111 kbps video 704 kbps)
639KB/23p	573KB	4.74MB/248s (160kbps)	329MB/3359s (audio 112 kbps video 695 kbps)
1.14MB/65p	1.31MB	8.48MB/222s (320kbps)	349MB/2640s (audio 153 kbps video 934 kbps)
207MB/924p	3.44MB	3.16MB/195s (128kbps)	699MB/5708s (audio 99 kbps video 909 kbps)

Tabel 5. Diferite tipuri de mesaje şi mărimea lor

În tabelul 6 sunt exemplificate diferite mesaje în clar şi lungimea secvenţei ADN (cheia de criptare) necesare pentru criptarea lor, datele fiind obţinute prin măsurători pe fişiere reale. Fiecare secvenţă are un număr de identificare unic (ID) care va fi utilizat la datele de intrare în transmisia cheilor.

Mesajul în clar	Lungimea secvenţei ADN (cheii)	Dimensiunea datelor de intrare
Text 15KB/1 pagină	61,440 bp	ID a unei secvenţe ADN
Imagine 402KB	1,646,592 bp	ID a unei secvenţe ADN
Sunet (160kbps) 5.34MB/278s	22,397,583 bp	ID a unei secvenţe ADN
Video 329MB/3359s (audio 112 kbps, video 695 kbps)	1,384,120,320 bp	~ 5 ID-uri ale secvenţelor ADN din cromozomi

Tabel 6. Diferite tipuri de mesaje în clar şi lungimea cheilor de criptare

Pentru obţinerea unui număr ridicat de secvenţe aleatoare de lungime mare se propun mai multe modalităţi:

a) Un simplu cromozom obţinut prin simplă citire (Tabelele 2 - 4)
b) Multiplexarea, deplasarea ciclică şi concatenarea mai multor secvenţe obţinute din acelaşi cromozom.
   Pentru exemplificare se consideră o secvenţă originală divizată aleatoriu în trei subsecvenţe:

AATAGCACAATAA | TCACATTCTTG | GCTTCTACTCATCT

Prin deplasare ciclică şi concatenare se obţine secvenţa modificată:

GCTTCTACTCATCT | AATAGCACAATAA | TCACATTCTTG

c) Multiplexarea de secvenţe obţinute de la cromozomi distincţi ale aceleiaşi specii.
Se exemplifică cu secvenţe ADN conţinute în cromozomii 4, 7 şi 9 ai porumbului:

Zea mays Cr. 4:	A	A	G	CTTCTACTCATCTCCCGGCAAACAGATAT...
Zea mays Cr. 7:	G	G	A	ATAGCACAATAAGTGCGCAAAATCGAAG...
Zea mays Cr. 9:	G	A	T	CACATTCTTGGATTTTTGGTGGAGACCAT...

Multiplexarea, în acest caz, se face prin parcurgerea nucleotidelor de pe aceeaşi poziţie din secvenţele cromozomilor 4, 7 şi 9, de la stânga spre dreapta. Secvenţa rezultată este:

MUX(Zea mays{Cr. 4, Cr. 7, Cr.9}) = AGGAGAGATCACTTATAC...

d) Multiplexarea de secvenţe de la specii diferite după regula c)

Pentru exemplificare se consideră cromozomii din tabelul 7, cu lungimile lor exprimate în perechi de baze (bp) şi în biţi.

Cromozom	Lungimea secevnţei ADN (bp)	Lungimea secevnţei ADN (biţi)	Notaţie lungime
Homo Sapiens Cr. 5	180,915,260	361,830,520	l_1
Zea mays Cr. 8	175,793,759	351,587,518	l_2
Felis catus Cr. C1	221,441,202	442,882,404	l_3

Tabel 7. Selecţie de cromozomi utilizaţi în multiplexare

În urma multiplexării celor trei secvenţe după regula c) se obţine o secvenţă nouă de lungime

$$l_{mux} = 3*\min(l_1, l_2, l_3) = 1{,}054{,}762{,}554 \text{ biţi}$$

Cu o astfel de cheie se pot cripta 125,7 MB de date.

E important de reţinut faptul că părţile implicate în comunicare nu îşi transmit cheia de criptare obţinută din secvenţe ADN, ci o cheie secretă al cărei mod de realizare se prezintă în continuare.

Cheia secretă (figura 4), ce trebuie transmisă utilizatorului, va fi formată dintr-un antet şi ID-urile secvenţelor ADN utilizate la generarea cheii de criptare ADN.

Antetul se obţine prin concatenarea a două secvenţe binare pe 2 şi respectiv pe 3 biţi, notate $\mathbf{k_1}$ şi $\mathbf{k_2}$ unde:

- $\mathbf{k_1}$ reprezintă codificarea, pe 2 biţi, a modului de realizare a cheii de criptare K_{ADN} din structuri ADN, o posibilă codificare a acestuia putând fi următoarea:
  a) – 00, b) – 01, c) – 10, d) – 11
- $\mathbf{k_2}$ reprezintă codificarea, pe 3 biţi, a numărului **n** de ID-uri de structuri ADN din baza de date ADN, structuri utilizate în crearea cheii de criptare K_{ADN}, printr-unul din modurile a) – d).

Partea a treia, $\mathbf{k_3}$, din structura cheii secrete se obţine concatenând ID-urile pe 8 biţi (ID_1, ID_2, ... , ID_n) ale structurilor ADN folosite la generarea cheii de criptare K_{ADN}.
Se exemplifică, în continuare, modul de creare a cheii secrete.

Exemplu:
Pentru codarea unui fişier video de dimensiune 125 MB (Tabel 6) este nevoie de o secvenţă ADN de lungimea 500,000,000 bp, care se poate obţine prin multiplexarea a 3 secvenţe diferite, fiecare de minim 167,000,000 bp. Prin urmare în cheia secretă se vor transmite 3 ID-uri. În figura 2 sunt exemplificate secvenţele în formatul GenBank şi ID-ul fiecărei secvenţe.
În figura 6 este exemplificat formatul unei astfel de chei secrete.

Metoda de criptare, conform invenţiei, pentru sistemul de transmisiune duplex, se descrie prin următoarele etape:

1. preluarea mesajului M_A, respectiv M_B la partea emiţătoare şi stabilirea datelor de intrare DI (lungimea mesajului, modul de generare a cheii de criptare (unul din modurile a) – d) şi ID-urile structurilor ADN utilizate la generarea cheii)
2. generarea cheii de criptare K_{ADN} constând într-o secvenţă aleatoare de structuri ADN preluate din baza de date biologică BD ADN (publică sau privată)
3. conversia cheii ADN în format binar obţinându-se cheia OTP binară K_A
4. criptarea mesajului în clar M_A, respectiv M_B cu ajutorul cheii binare K_A, respectiv K_B
5. crearea cheii secrete din datele de intrare DI în blocul de criptare E^*_A, respectiv E^*_B
6. transmiterea mesajului criptat şi a cheii secrete părţii receptoare
7. decriptarea cheii secrete în blocul D^*_B, respectiv D^*_A la receptor şi obţinerea datelor de intrare DI din cheia secretă
8. generarea cheii ADN la receptor folosind datele de intrare DI şi baza de date BD ADN, identică cu baza de date de la partea emiţătoare
9. conversia cheii K_{ADN} la formatul binar K_A, respectiv K_B
10. decriptarea mesajului la receptor cu ajutorul cheii binare K_A, respectiv K_B.

Metoda de determinare de secvenţe aleatoare bazate pe structuri ADN prezintă următoarele avantaje:

- Asigură generarea de secvenţe aleatoare binare de orice lungime utilizând secvenţe de structuri ADN aflate în baze de date publice sau private
- Numărul de secvenţe aleatoare distincte este practic nelimitat datorita modurilor de obţinere versatile (a se vedea regulile a), b), c), d) prezentate anterior). Aceasta constituie premisa creării unui sistem de criptare cu cheie unică pe sesiune (OTP).

Se dă în continuare un exemplu de realizare a criptosistemului OTP conform invenţiei utilizat în transmisiuni de tip duplex, bazat se secvenţe aleatoare de tip ADN (figura 7) obţinute prin regulile a) – d) prezentate mai sus. Pentru o mai uşoară descriere a sistemului conform invenţiei, în figura 7 se utilizează următoarele notaţii:

BD ADN - Baza de date cu secvenţe ADN (identică celor 2 corespondenţi A şi B), care poate fi publică sau privată

DI – bloc al datelor de intrare: lungimea necesară a secvenţei după tipul mesajului în clar, numerele de identificare a secvenţelor utilizate la generarea cheii, modul de realizare a cheii.

M_A – mesajul în clar al utilizatorului A

M_B – mesajul în clar al utilizatorului B

Gen K_{ADN} – bloc de generare a secvenţei ADN care va fi utilizată drept cheie de criptare (K_{ADN})

Conv ADN – B – bloc de transformare a cheii K_{ADN} în format binar (K_A, K_B)

$E_{A,B}^{\bullet}$ – bloc de criptare a informaţiei de intrare (DI) folosind un algoritm simetric sau public pentru utilizatorii A, B

$D_{A,B}^{\bullet}$ – bloc de decriptare a informaţiei de intrare (DI) utilizând algoritmul ales la $E_{A,B}^{\bullet}$

S_A – sumator modulo 2 utilizat pentru criptare, respectiv pentru decriptare OTP pentru utilizatorul A

S_B – sumator modulo 2 utilizat pentru criptare, respectiv pentru decriptare OTP pentru utilizatorul B.

Se descrie în continuare **protocolul de utilizare a criptosistemului OTP** conform invenţiei, utilizat în transmisiuni duplex:

I. Transmisia A → B

a. *Criptarea mesajului de transmis, la partea emiţătoare A:*

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvenţelor ADN, modul de generare a cheii

(2) Generarea cheii ADN din datele de intrare şi baze de date genetice: K_{ADN} a utilizatorului A

(3) Generarea cheii OTP (K_A) prin conversia cheii K_{ADN} în format binar

(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul $E_A^{\bullet}$) şi transmiterea către partea receptoare B

(5) Criptarea OTP a datelor M_A: $C_A = K_A \oplus M_A$ şi transmiterea criptogramei C_A la B.

b. Decriptarea mesajului la partea receptoare B

(6) Decriptarea datelor de intrare, realizată în blocul $D_B^{\bullet}$

(7) Generarea cheii K_A din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la A şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)

(8) Decriptarea OTP a criptogramei C_A utilizând cheia K_A obţinută la (7): $C_A \oplus K_A = M_A$

II. Transmisia B → A

a. Criptarea mesajului de transmis, la partea emiţătoare B

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvenţelor ADN, modul de generare a cheii.

(2) Generarea cheii ADN din datele de intrare si baze de date genetice: K_{ADN} a utilizatorului B

(3) Generarea cheii OTP (K_B) prin conversia cheii ADN la format binar binar
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul E_B^*) şi transmiterea către partea receptoare A
(5) Criptarea OTP a datelor M_B: $C_B = K_B \oplus M_B$ şi transmiterea criptogramei C_B la A

b. Decriptarea mesajului la partea receptoare A

(6) Decriptarea datelor de intrare, realizată în blocul D_A^*
(7) Generarea cheii K_B din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la B şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)
(8) Decriptarea OTP a criptogramei C_B utilizând cheia K_B obţinută la (7): $C_B \oplus K_B = M_B$

Criptosistemul OTP utilizat în transmisiuni duplex, bazat pe secvenţe aleatoare de tip ADN are următoarele avantaje:

- Sistemul de tip OTP (o cheie secretă utilizată o singură dată şi având lungimea cel puţin egală cu a mesajului în clar) este demonstrat matematic ca de nespart [Ver-26], [Sha-49]
- Cheia OTP nu trebuie transmisă integral părţii de recepţie, ea putând fi generată uşor de aceasta prin transmiterea criptată (E_A^*, E_B^*) a datelor de intrare şi utilizarea aceleiaşi baze de date ADN publice sau private, ceea ce asigură un management extrem de uşor al cheilor, înlăturând astfel principalul neajuns al criptografiei simetrice (managementul extrem de greu al cheilor cu cât numărul de utilizatori creşte)
- Securitatea criptosistemului este dată de securitatea algoritmului de criptare a datelor de intrare (E*, D*).

Utilizând aceleaşi convenţii de notare ca în cazul criptosistemului OTP utilizat pentru transmisiuni duplex (figura 7), se descrie în continuare **protocolul de utilizare a criptosistemului OTP utilizat pentru stocare de date** conform invenţiei (figurile 8 şi 9):

a. Criptarea (la unitatea de scriere)

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvenţelor ADN, modul de generare a cheii de criptare
(2) Generarea cheii de criptare ADN din datele de intrare şi baze de date genetice: K_{ADN}
(3) Generarea cheii OTP (K) prin conversia cheii ADN la formatul binar
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul E^*) şi memorarea pe mediul de stocare
(5) Criptarea OTP a datelor M: $C = K \oplus M$ şi memorarea criptogramei C pe mediul de stocare (CD, DVD, etc.)

b. Decriptarea (la unitatea de citire)

(6) Decriptarea datelor de intrare, realizată în blocul D^*
(7) Generarea cheii K din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la înscriere şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)
(8) Decriptarea OTP a criptogramei C utilizând cheia K obţinută la (7): $C \oplus K = M$.

Criptosistemul OTP utilizat pentru stocare de date bazat pe secvenţe aleatoare de tip ADN prezintă avantajul că stocarea cheilor de criptare este mult mai uşoară decât în sistem clasic datorită faptului că sunt formate doar din datele de intrare (DI), deci au o lungime mult mai scurtă decât a cheii de criptare ADN folosite (K).

Bibliografie:

[Sch-96] B. Schneier, "Applied Cryptography: Protocols, Algorithms, and Source Code in C", *John Wiley & Sons*, Inc, 1996.

[Sta-99] W. Stallings, "Cryptography and Network Security: Principles and Practice", (5th Edition), *Prentice Hall*, 2011.

[PP-06] F. Piva, G. Principato, "RANDNA: a random DNA sequence generator", *In Silico Biology* , Volume 6 (3) IOS Press, Jan 1, 2006.

[CG-08] V. S. Chandra, G. Gopakumar, A.S. Nair, "Biolets: Statistical Approach to Biological Random Sequence Generation", *Malaysian Journal of Computer Science*, Vol. 21(2), 2008.

[RG-97] A. Rambaut, N.C. Grassly, "Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees", *CABIOS*, Vol. 13 no. 3, pp. 235-238, 1997.

[PTD-06] Y. Ponty, M. Termier and A. Denise, "GenRGenS: Software for Generating Random Genomic Sequences and Structures", *Bioinformatics*, Vol. 22, pp. 1534-1535, 2006.

[DDBJ] http://www.ddbj.nig.ac.jp/Welcome-e.html

[ENA] http://www.ebi.ac.uk/ena/

[GenBank] http://www.ncbi.nlm.nih.gov/pubmed/21071399

[NCBI] http://www.ncbi.nlm.nih.gov/genome

[NW-99] C. G. Nevill-Manning, I.H. Witten, "Protein is incompressible", In Proceedings of the Conference on Data Compression (DCC '99), pp. 257-, 1999.

[CLLX-09] Christley S, Lu Y, Li C, Xie X. "Human genomes as email attachments", Bioinformatics Vol. 25, pp. 274–275, 2009.

[FLCB-11] M. H.Y. Fritz, R. Leinonen, G. Cochrane and E. Birney, "Efficient storage of high throughput DNA sequencing data using reference-based compression", Genome Res. Vol. 21, pp. 734-740, 2011.

[DRCX+10] Daily K, Rigor P, Christley S, Xie X, Baldi P. "Data structures and compression algorithms for high-throughput sequencing technologies", BMC Bioinformatics 11: 514, 2010.

[RA-11] P. Rajarajeswari, A. Apparao, "DNABIT Compress – Genome compression algorithm", Bioinformation, Vol. 5(8), pp. 350–360, 2011.

[CDLT04] C. R. Calladine, H. R. Drew, B. F. Luisi, A. A. Travers, "Understanding DNA The Molecule & How It Works", *Academic Press*, April 2004.

[Ver-26] G. S. Vernam, "Cipher Printing Telegraph Systems", *Journal of the American Institute of Electrical Engineers*, Vol. XI.V, pp. 109-115, 1926.

[Sha-49] C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, Vol. 28, No. 4, pp. 656-715, 1949.

REVENDICĂRI

1. Metodă de criptare bazată pe secvenţe aleatoare determinate din structuri ADN, utilizată la transmisiuni de tip duplex, **caracterizată prin aceea că**, transmiterea securizată a mesajelor se realizează printr-un proces emisie - recepţie constând în următorii paşi:

   i) citirea mesajului în clar (M_A, respectiv M_B) la partea emiţătoare şi stabilirea datelor de intrare (DI) formate din:
   - lungimea mesajului în clar (l_A, respectiv l_B)
   - codul corespunzător modului de formare a cheii de criptare, acesta fiind unul dintre modurile a), b), c) sau d)
   - numărul de structuri ADN din baza de date biologică (BD ADN) care se vor utiliza la generarea cheii de criptare (K_{ADN})

   ii) generarea cheii de criptare (K_{ADN}) constând într-o secvenţă aleatoare de structuri ADN preluate din baza de date biologică (BD ADN), publică sau privată

   iii) conversia cheii ADN la formatul binar, obţinând cheia OTP (K_A)

   iv) criptarea mesajului în clar (M_A), respectiv M_B) cu ajutorul cheii OTP (K_A, respectiv K_B)

   v) crearea cheii secrete din datele de intrare (DI) în blocul de criptare (E^*_A, respectiv E^*_B)

   vi) transmiterea mesajului criptat (C_A, respectiv C_B) şi a cheii secrete părţii receptoare

   vii) decriptarea cheii secrete în blocul de decriptare (D^*_B, respectiv D^*_A) la receptor şi obţinerea datelor de intrare (DI)

   viii) generarea cheii ADN la receptor folosind datele de intrare (DI) şi baza de date biologică (BD ADN)

   ix) conversia cheii ADN (K_{ADN}) la formatul binar (K_A, respectiv K_B)

   x) decriptarea mesajului la receptor cu ajutorul cheii binare (K_A, respectiv K_B).

2. Metodă de criptare bazată pe secvenţe aleatoare determinate din structuri ADN, conform revendicării 1, **caracterizată prin aceea că**, cheia secretă care se transmite odată cu mesajul criptat, de la partea emiţătoare către partea receptoare, este formată dintr-un antet care are structura următoare:
   - un cod (k_1) pe 2 biţi care reprezintă modul de formare a cheii de criptare, adică unul din modurile a) - d)

- un cod (k_2) pe 3 biţi reprezentând numărul de ID-uri ale structurilor ADN utilizate la obţinerea cheii de criptare

şi dintr-o înşiruire (k_3) a ID-urilor structurilor ADN utilizate la obţinerea cheii de criptare (K_{ADN}).

3. Sistem de criptare de tip OTP bazat pe secvenţe aleatoare determinate din structuri ADN, alcătuit din două părţi, una emiţătoare de mesaje şi cealaltă receptoare, utilizat în transmisiuni de tip duplex **caracterizată prin aceea că**, fiecare dintre cele două părţi este compusă dintr-o bază de date ADN (BD ADN), publică sau privată , identică la ambele părţi, un bloc al datelor de intrare (DI), un generator de cheie ADN (Gen K_{ADN}), un convertor al cheii ADN în cheie binară (Conv ADN-B), un sumator modulo 2 (S) cu rol de criptare a mesajului, un bloc de criptare (E*) a datelor de intrare cu care se generează cheia secretă şi un bloc de decriptare a acesteia (D*), mesajul criptat şi cheia secretă fiind transmise de la emiţător către receptor, acesta din urmă (receptorul) obţinând datele de intrare prin decriptarea cheii secrete şi continuând cu generarea cheii ADN pe care o utilizează la decriptarea mesajului recepţionat.

4. Sistem de criptare de tip OTP bazat pe secvenţe aleatoare determinate din structuri ADN, conform revendicării 3, **caracterizat prin aceea că** utilizează următorul protocol:

I. Transmisia A → B

a. Criptarea mesajului de transmis, la partea emiţătoare A:

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, modul de generare a cheii şi alegerea secvenţelor ADN

(2) Generarea cheii ADN din datele de intrare şi baze de date genetice: K_{ADN} a utilizatorului A

(3) Generarea cheii OTP (K_A) prin conversia cheii ADN în format binar

(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul E_B^*) şi transmiterea către partea receptoare B

(5) Criptarea OTP a datelor M_A: $C_A = K_A \oplus M_A$ şi transmiterea criptogramei C_A la B.

b. Decriptarea mesajului la partea receptoare B

(6) Decriptarea datelor de intrare, realizată în blocul D_B^*

(7) Generarea cheii K_A din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la A şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)

(8) Decriptarea OTP a criptogramei C_A utilizând cheia K_A obţinută la (7): $C_A \oplus K_A = M_A$

II. Transmisia B → A

a. Criptarea mesajului de transmis, la partea emiţătoare B

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvenţelor ADN, modul de generare a cheii.

(2) Generarea cheii ADN din datele de intrare si baze de date genetice: K_{ADN} a utilizatorului B

(3) Generarea cheii OTP (K_B) prin conversia cheii ADN la format binar binar

(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul E_B^*) şi transmiterea către partea receptoare A

(5) Criptarea OTP a datelor M_B: $C_B = K_B \oplus M_B$ şi transmiterea criptogramei C_B la A

b. Decriptarea mesajului la partea receptoare A

(6) Decriptarea datelor de intrare, realizată în blocul D_A^*

(7) Generarea cheii K_B din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la B şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)

(8) Decriptarea OTP a criptogramei C_B utilizând cheia K_B obţinută la (7): $C_B \oplus K_B = M_B$

5. Sistem de criptare de tip OTP bazat pe secvenţe aleatoare determinate din structuri ADN, alcătuit din două unităţi, una de înscriere (de criptare) şi alta de citire (de decriptare) utilizat în sisteme de stocare, **caracterizat prin aceea că**, unitatea de înscriere este compusă dintr-o bază de date ADN (BD ADN), publică sau privată , un bloc al datelor de intrare (DI), un generator de cheie ADN (Gen K_{ADN}), un convertor al cheii ADN în cheie binară (Conv ADN-B), un sumator modulo 2 (S) cu rol de criptare a datelor, un bloc de criptare (E*) a datelor de intrare, iar unitatea de citire este compusă dintr-o bază de date ADN (BD ADN), identică cu baza de date ADN a unităţii de scriere, un bloc de decriptare a datelor de intrare (D*), un generator de cheie ADN (Gen K_{ADN}), un convertor al cheii ADN în cheie binară (Conv ADN-B) şi un sumator modulo 2 (S) cu rol de decriptare a datelor.

6. Sistem de criptare de tip OTP bazat pe secvenţe aleatoare determinate din structuri ADN, conform revendicării 5, **caracterizat prin aceea că** utilizează următorul protocol:

a. Criptarea (la unitatea de înscriere)

(1) Stabilirea datelor de intrare: lungimea mesajului în clar, modul de generare a cheii şi alegerea secvenţelor ADN

(2) Generarea cheii ADN din datele de intrare şi baze de date genetice: K_{ADN}

(3) Generarea cheii OTP (K) prin conversia cheii ADN la formatul binar

(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul E^*) şi memorarea pe mediul de stocare

(5) Criptarea OTP a datelor M: $C = K \oplus M$ şi memorarea criptogramei C pe mediul de stocare (CD, DVD, etc.)

b. Decriptarea (la unitatea de citire)

(6) Decriptarea datelor de intrare, realizată în blocul D^*

(7) Generarea cheii K din datele de intrare obţinute la (6) utilizând baza de date ADN identică cu cea de la scriere şi acelaşi generator de cheie (Gen K_{ADN} şi Conv ADN - B)

(8) Decriptarea OTP a criptogramei C utilizând cheia K obţinută la (7): $C \oplus K = M$.

Figura 1

Figura 2

Figura 3

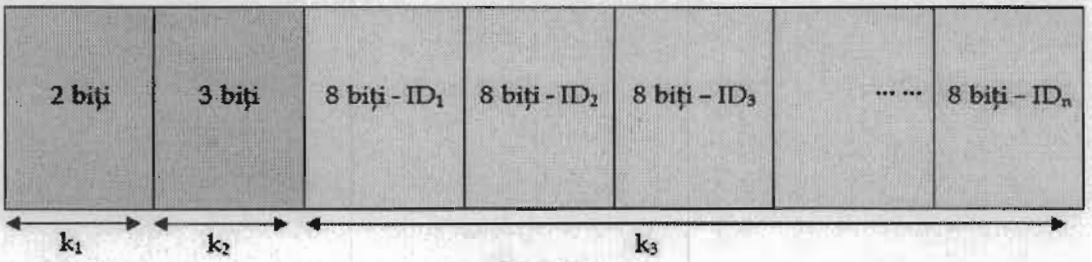

Figura 4

```
LOCUS       CM000663            249250621 bp    DNA     linear   CON 29-JUN-2009
DEFINITION  Homo sapiens chromosome 1, GRCh37 primary reference assembly.
ACCESSION   CM000663
VERSION     CM000663.1  GI:224384768
DBLINK      BioProject: PRJNA31257
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

a) ID_1 – CM000663

```
LOCUS       CM000664            243199373 bp    DNA     linear   CON 29-JUN-2009
DEFINITION  Homo sapiens chromosome 2, GRCh37 primary reference assembly.
ACCESSION   CM000664
VERSION     CM000664.1  GI:224384767
DBLINK      BioProject: PRJNA31257
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

b) ID_2 – CM000664

```
LOCUS       CM001378            239302903 bp    DNA     linear   CON 14-DEC-2011
DEFINITION  Felis catus breed Abyssinian chromosome A1, whole genome shotgun
            sequence.
ACCESSION   CM001378
VERSION     CM001378.1  GI:362110686
DBLINK      BioProject: PRJNA16726
KEYWORDS    WGS.
SOURCE      Felis catus (domestic cat)
  ORGANISM  Felis catus
```

c) ID_3 – CM001378

Figura 5

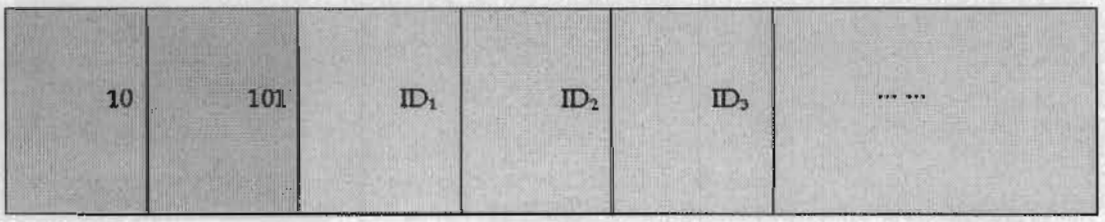

Figura 6

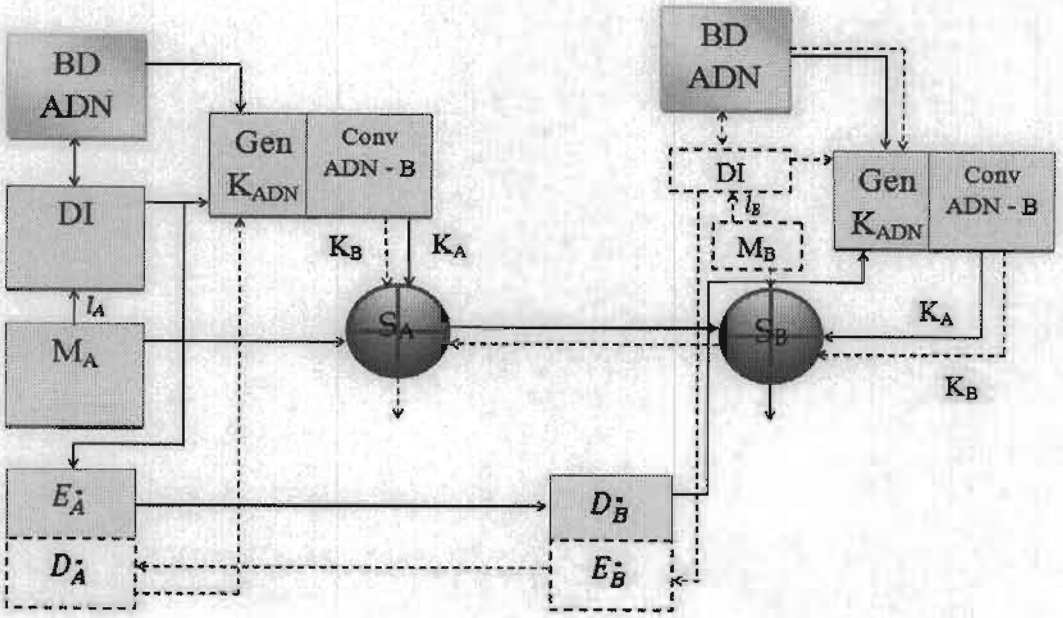

Figura 7

Unitatea de scriere (partea de criptare)

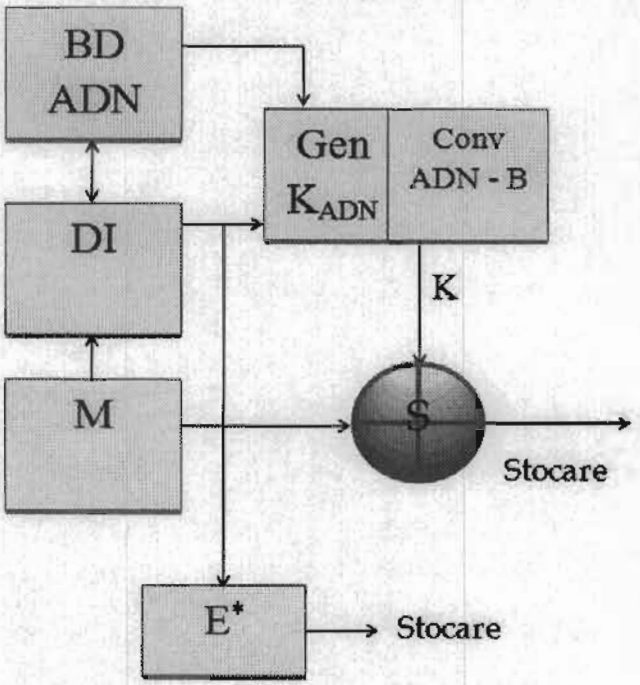

Figura 8

Unitatea de citire (partea de decriptare)

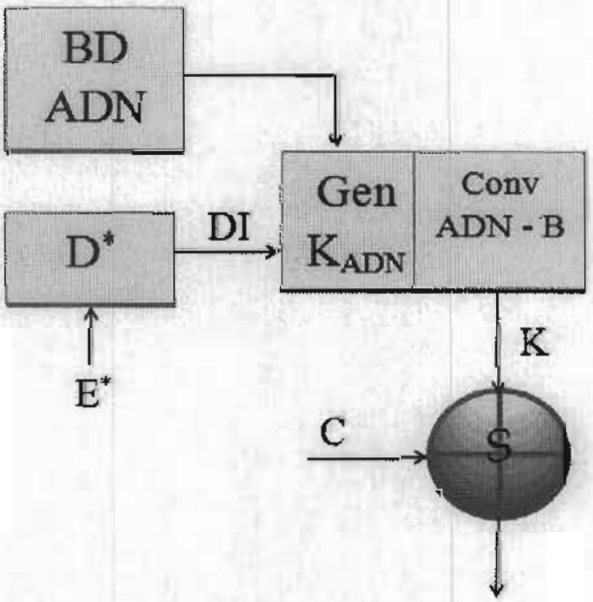

Figura 9