



(12)

## BREVET DE INVENȚIE

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

(22) Data de depozit: **14/02/2013**

(45) Data publicării mențiunii acordării brevetului: **30/12/2020** BOPI nr. **12/2020**

(41) Data publicării cererii:  
**29/08/2014** BOPI nr. **8/2014**

(73) Titular:  
• **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.E9, SC.1,  
AP.2, CLUJ NAPOCA, CJ**

(56) Documente din stadiul tehnicii:  
**Olga Tornea, Monica Borda, Tatiana  
Hodorogea, Mircea-Florin Vaida,  
"Encryption System with Indexing DNA  
Chromosomes Cryptographic Algorithm",  
IASTED, International Conference on  
Biomedical Engineering (BioMed 2010),  
Innsbruck, Austria, paper 680-099,  
pp.12-15, ISBN: 978-0-88986-825-0,  
975-0-88986-826-7, Acta Press, 15-18  
February 2010; M. Borda, O. Tornea, "DNA  
Secret Writing Techniques", Proceedings  
of 8th International Conference on  
Communications, invited paper, Vol.2,  
IEEE Explore, pp. 451-460, Bucharest,  
2010; CN102025482**

(54) **METODĂ ȘI SISTEM DE CRIPTARE DE TIP OTP BAZATE  
PE SECVENȚE ALEATOARE DETERMINATE  
DIN STRUCTURI ADN**



# RO 129751 B1

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

5           La ora actuală există generatoare pseudo-aleatoare ADN de tip congruențial [PP-06],  
[RG-97], [PTD-06], care ca orice generator pseudoaleator generează secvențe de lungime  
7           finită, iar parametrii generatorului (seed) trebuie transmiși celor 2 părți.

9           Există și propuneri de secvențe aleatoare de tip ADN [CG-08] cu proprietăți foarte  
apropiate de ale unei secvențe aleatoare pure, obținuți prin utilizarea mai multor valori inițiale  
(seed-uri), dar a căror utilizare în aplicații criptografice este practic imposibilă.

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

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

19           Metoda de criptare, conform invenției urmărește obținerea de chei unice de criptare  
constând în secvențe aleatoare bazate pe structuri ADN din baze de date biologice existente  
(publice sau private) sau realizate prin sinteză. Secvențele aleatoare propuse se obțin din  
21           secvențe ADN (cromozomi, gene etc.) stocate în baze de date biologice existente, cele mai mari  
fiind: [DDBJ], [ENA], [GenBank]. Structura unei secvențe ADN dintr-un organism viu (om,  
23           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].

25           Sistemul de criptare de tip OTP conform invenției este destinat utilizării în transmisiuni  
de tip duplex sau în sisteme de stocare, acesta asigurând protecția informației printr-un proces  
27           de criptare - decriptare în care cheile de criptare constau în secvențe aleatoare obținute din  
structuri ADN.

29           Sistemul, conform invenției, utilizat în transmisiuni de tip duplex este alcătuit din două  
părți, una emițătoare de mesaje și cealaltă receptoare, fiecare dintre cele două părți fiind  
31           compusă dintr-o bază de date ADN (BD ADN), publică sau privată, identică la ambele părți, un  
bloc al datelor de intrare (DI), un generator de cheie ADN ( $Gen K_{ADN}$ ), un convertor al cheii ADN  
33           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 și un bloc de decriptare  
35           a acesteia ( $D^*$ ), mesajul criptat și cheia secretă fiind transmise de la emițător către receptor,  
acesta din urmă (receptorul) obținând datele de intrare prin decriptarea cheii secrete și  
37           continuând cu generarea cheii ADN pe care o utilizează la decriptarea mesajului recepționat.

39           Se da în continuare un exemplu de realizare a metodei de criptare bazate pe secvențe  
aleatoare de structuri ADN în legătură cu fig. 1...9 care reprezintă:

41           - fig. 1, secvență de cromozomi din structura ADN a pisicii domestice (Felis cactus);

43           - fig. 2, secvență de cromozomi din structura ADN a porumbelului (Zea mays);

45           - fig. 3, secvență de cromozomi din structura de ADN a omului (Homo sapiens);

47           - fig. 4, formatul cheii secrete;

49           - fig. 5, exemple de secvențe ADN în format GenBank și ID-urile lor;

            - fig. 6, formatul cheii secrete pentru exemplul prezentat;

            - fig. 7, schema bloc pentru criptosistemul OTP bazat pe secvențe aleatoare de tip AND  
utilizat în transmisiuni de tip duplex;

            - fig. 8, schema bloc pentru criptosistemul OTP utilizat pentru stocarea de date bazat  
pe secvențe aleatoare de tip ADN - partea de criptare a unității de înscriere;

# RO 129751 B1

- fig. 9, schema bloc pentru criptosistemul OTP utilizat pentru stocarea de date bazat pe secvențe aleatoare de tip ADN - partea de decriptare a unității de citire.

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

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.

Tabel de conversie ADN → binar

Tabelul 1

A	00
C	01
G	10
T	11

Aceste substituții sunt ușor de realizat prin structura de selecție switch - case a unui limbaj de programare.

Lungimile secvențelor 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). Fig. 1...3 corelate cu tabelele 2...4, demonstrează faptul că cromozomii din structura ADN a unor organisme vii precum pisica domestică (fig. 1 - tabelul 2), porumbul (fig. 2 - tabelul 3) sau organismul uman (fig. 3 - tabelul 4) au lungimi de ordinul zecilor sau sutelor de milioane de perechi de baze (bp).

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

Tabele 2, 3, 4

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.

# RO 129751 B1

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

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

Se observă că în bazele de date genomice există secvențe ADN (cele din cromozomi) de lungime foarte mare ce pot asigura lungimi cerute de diverse aplicații.

Tabelul 5 prezintă tipuri de mesaje și mărimea lor, datele fiind obținute prin măsurători pe fișiere reale.

*Diferite tipuri de mesaje și mărimea lor*

*Tabelul 5*

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)

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

*Diferite tipuri de mesaje în clar și lungimea cheilor de criptare*

*Tabelul 6*

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

Pentru obținerea unui număr ridicat de secvențe aleatoare de lungime mare se propun mai multe modalități:

a) Un simplu cromozom obținut prin simpla citire (Tabelele 2-4).

b) Multiplexarea, deplasarea ciclică și concatenarea mai multor secvențe obținute din același cromozom.

# RO 129751 B1

Pentru exemplificare se consideră o secvență originală divizată aleatoriu în trei subsecvențe:

AATAGCACAATAA TCACATTCTTG GCTTCTACTCATCT

Prin deplasare ciclică și conctenare se obține secvența modificată:

GCTTCTACTCATCT AATAGCACAATAA TCACATTCTTG

c) Multiplexarea de secvențe obținute de la cromozomi distincți ale aceleiași specii. Se exemplifică cu secvențe ADN conținute în cromozomii 4, 7 și 9 ai porumbului:

Zea mays Cr. 4:	A	A	G	CTTCTACTCATCTCCCGGCAAACAGATAT...
Zea mays Cr. 7:	G	G	A	ATAGCACAATAAGTGCGCAAATCGAAG...
Zea mays Cr. 9:	G	A	T	CACATTCTTGGATTTTGGTGGAGACCAT...

Multiplexarea, în acest caz, se face prin parcurgerea nucleotidelor de pe aceeași poziție din secvențele cromozomilor 4, 7 și 9, de la stânga spre dreapta. Secvența rezultată este:

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

d) Multiplexarea de secvențe de la specii diferite după regula c).

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

## Selecție de cromozomi utilizați în multiplexare

Tabelul 7

Cromozom	Lungimea secevntei ADN (bp)	Lungimea secevntei ADN (biți)	Notăție 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	$l_3$

În urma multiplexării celor trei secvențe după regula c) se obține o secvență nouă de lungime  $1_{\text{mux}} = 3 \cdot \min(1_1, 1_2, 1_3) = 1,0547,762,554$  biți.

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

E important de reținut faptul că părțile implicate în comunicare nu își transmit cheia de criptare obținută din secvențe ADN, ci o cheie secretă al cărei mod de realizare se prezintă în continuare.

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

Antetul se obține prin concatenarea a două secvențe binare pe 2 și respectiv pe 3 biți, notate  $k_1$  și  $k_2$  unde:

-  $k_1$  reprezintă codificarea, pe 2 biți, a modului de realizare a cheii de criptare  $K_{\text{ADN}}$  din structuri ADN, o posibilă codificare a acestuia putând fi următoarea:

a) - 00, b) - 01, c) - 10, d) - 11;

-  $k_2$  reprezintă codificarea, pe 3 biți, a numărului  $n$  de ID-uri de structuri ADN din baza de date ADN, structuri utilizate în crearea cheii de criptare  $K_{\text{ADN}}$ , printr-unul din modurile a) - d).

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

# RO 129751 B1

## 1 Exemplu:

2 Pentru codarea unui fișier video de dimensiune 125 MB (Tabel 6) este nevoie de o  
3 secvență ADN de lungimea 500,000,000 bp, care se poate obține prin multiplexarea a 3 sec-  
4 vențe diferite, fiecare de minim 167,000,000 bp. Prin urmare în cheia secretă se vor transmite  
5 3 ID-uri. În fig. 2 sunt exemplificate secvențele în formatul GenBank și ID-ul fiecărei secvențe,  
6 în fig. 6 este exemplificat formatul unei astfel de chei secrete.

7 Metoda de criptare, conform invenției, pentru sistemul de transmisiune duplex, se  
8 descrie prin următoarele etape:

9 1. preluarea mesajului  $M_A$ , respectiv  $M_B$  la partea emițătoare și stabilirea datelor de  
10 intrare DI (lungimea mesajului, modul de generare a cheii de criptare (unul din modurile a) - d)  
11 și ID-urile structurilor ADN utilizate la generarea cheii);

12 2. generarea cheii de criptare  $K_{ADN}$  constând într-o secvență aleatoare de structuri ADN  
13 preluate din baza de date biologică BD ADN (publică sau privată);

14 3. conversia cheii ADN în format binar obținându-se cheia OTP binară  $K_A$ ;

15 4. criptarea mesajului în clar  $M_A$ , respectiv  $M_B$  cu ajutorul cheii binare  $K_A$ , respectiv  $K_B$ ;

16 5. crearea cheii secrete din datele de intrare DI în blocul de criptare  $E^*_A$ , respectiv  $E^*_B$ ;

17 6. transmiterea mesajului criptat și a cheii secrete părții receptoare

18 7. decriptarea cheii secrete în blocul  $D^*_B$ , respectiv  $D^*_A$  la receptor și obținerea datelor  
19 de intrare DI din cheia secretă;

20 8. generarea cheii ADN la receptor folosind datele de intrare DI și baza de date BD ADN,  
21 identică cu baza de date de la partea emițătoare;

22 9. conversia cheii  $K_{ADN}$  la formatul binar  $K_A$ , respectiv  $K_B$ ;

23 10. decriptarea mesajului la receptor cu ajutorul cheii binare  $K_A$ , respectiv  $K_B$ .

24 Metoda de determinare de secvențe aleatoare bazate pe structuri ADN prezintă  
25 următoarele avantaje:

26 - asigură generarea de secvențe aleatoare binare de orice lungime utilizând secvențe  
27 de structuri ADN aflate în baze de date publice sau private;

28 - numărul de secvențe aleatoare distincte este practic nelimitat datorită modurilor de  
29 obținere versatile (a se vedea regulile a), b), c), d) prezentate anterior). Aceasta constituie  
30 premisa creării unui sistem de criptare cu cheie unică pe sesiune (OTP).

31 Se dă în continuare un exemplu de realizare a criptosistemului OTP conform invenției  
32 utilizat în transmisiuni de tip duplex, bazat pe secvențe aleatoare de tip ADN (fig. 7) obținute prin  
33 regulile a) - d) prezentate mai sus. Pentru o mai ușoară descriere a sistemului conform invenției,  
34 în fig. 7 se utilizează următoarele notații:

35 BD ADN - Baza de date cu secvențe ADN (identică celor 2 corespondenți A și B), care  
36 poate fi publică sau privată;

37 D1 - bloc al datelor de intrare: lungimea necesară a secvenței după tipul mesajului în  
38 clar, numerele de identificare a secvențelor utilizate la generarea cheii, modul de realizare a  
39 cheii;

40  $M_A$  - mesajul în clar al utilizatorului A;

41  $M_B$  - mesajul în clar al utilizatorului B;

42 Gen  $K_{ADN}$  - bloc de generare a secvenței ADN care va fi utilizată drept cheie de  
43 criptare( $K_{ADN}$ );

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

45  $E^*_{AB}$  - bloc de criptare a informației de intrare (DI) folosind un algoritm simetric sau public  
46 pentru utilizatorii A, B;

47  $D^*_{AB}$  - bloc de decriptare a informației de intrare (DI) utilizând algoritmul ales la  $E^*_{AB}$

# RO 129751 B1

$S_A$ - sumator modulo 2 utilizat pentru criptare, respectiv pentru decriptare OTP pentru utilizatorul A;	1
$S_B$ - sumator modulo 2 utilizat pentru criptare, respectiv pentru decriptare OTP pentru utilizatorul B.	3
Se descrie în continuare protocolul de utilizare a criptosistemului OTP conform invenției, utilizat în transmisiuni duplex:	5
1. Transmisia $A \rightarrow B$	7
a) Criptarea mesajului de transmis, la partea emițătoare A:	
(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvențelor ADN, modul de generare a cheii;	9
(2) Generarea cheii ADN din datele de intrare și baze de date genetice: $K_{ADN}$ a utilizatorului A;	11
(3) Generarea cheii OTP ( $K_A$ ) prin conversia cheii $K_{ADN}$ în format binar;	13
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul $E^*_A$ ) și transmiterea către partea receptoare B;	15
(5) Criptarea OTP a datelor $M_A$ : $C_A = K_A \oplus M_A$ și transmiterea criptogramei $C_A$ la B;	
b) Decriptarea mesajului la partea receptoare B	17
(6) Decriptarea datelor de intrare, realizată în blocul $D^*_B$ ;	
(7) Generarea cheii $K_A$ din datele de intrare obținute la (6) utilizând baza de date ADN identică cu cea de la A și același generator de cheie (Gen $K_{ADN}$ și Conv ADN - B)	19
(8) Decriptarea OTP a criptogramei $C_A$ utilizând cheia $K_A$ obținută la (7): $C_A \oplus K_A = M_A$ .	21
2. Transmisia $B \rightarrow A$	23
a) Criptarea mesajului de transmis, la partea emițătoare B	
(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvențelor ADN, modul de generare a cheii;	25
(2) Generarea cheii ADN din datele de intrare și baze de date genetice: $K_{ADN}$ a utilizatorului B;	27
(3) Generarea cheii OTP ( $K_B$ ) prin conversia cheii ADN la format binar binar	29
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public (blocul $E^*_B$ ) și transmiterea către partea receptoare A;	31
(5) Criptarea OTP a datelor $M_B$ : $C_B = K_B \oplus M_B$ și transmiterea criptogramei $C_B$ la A;	
b) Decriptarea mesajului la partea receptoare A	33
(6) Decriptarea datelor de intrare, realizată în blocul $D^*_A$ ;	
(7) Generarea cheii $K_B$ din datele de intrare obținute la (6) utilizând baza de date ADN identică cu cea de la B și același generator de cheie (Gen $K_{ADN}$ și Conv ADN - B)	35
(8) Decriptarea OTP a criptogramei $C_B$ utilizând cheia $K_B$ obținută la (7): $C_B \oplus K_B = M_B$ .	37
Criptosistemul OTP utilizat în transmisiuni duplex, bazat pe secvențe aleatoare de tip ADN are următoarele avantaje:	39
- sistemul de tip OTP (o cheie secretă utilizată o singură dată și având lungimea cel puțin egală cu a mesajului în clar) este demonstrat matematic ca de nespart [Ver-26], [Sha-49];	41
- cheia OTP nu trebuie transmisă integral părții de recepție, ea putând fi generată ușor de aceasta prin transmiterea criptată ( $E^*_A, E^*_B$ ) a datelor de intrare și utilizarea aceleiași baze de date ADN publice sau private, ceea ce asigură un management extrem de ușor al cheilor, înlăturând astfel principalul neajuns al criptografiei simetrice (managementul extrem de greu al cheilor cu cât numărul de utilizatori crește);	43 45 47
- securitatea criptosistemului este dată de securitatea algoritmului de criptare a datelor de intrare ( $E^*, D^*$ ).	49

# RO 129751 B1

1 Utilizând aceleași convenții de notare ca în cazul criptosistemului OTP utilizat pentru  
transmisiuni duplex (fig. 7), se descrie în continuare protocolul de utilizare a criptosistemului  
3 OTP utilizat pentru stocare de date conform invenției (fig. 8 și 9):

a) Criptarea (la unitatea de scriere)

5 (1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvențelor ADN,  
modul de generare a cheii de criptare;

7 (2) Generarea cheii de criptare ADN din datele de intrare și baze de date genetice:  $K_{ADN}$ ;

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

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

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

b) Decriptarea (la unitatea de citire)

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

15 (7) Generarea cheii K din datele de intrare obținute la (6) utilizând baza de date ADN  
identică cu cea de la înscriere și același generator de cheie (Gen  $K_{ADN}$  și Conv ADN - B);

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

19 (9) Criptosistemul OTP utilizat pentru stocare de date bazat pe secvențe aleatoare de  
tip ADN prezintă avantajul ca stocarea cheilor de criptare este mult mai ușoară decât în sistem  
21 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).

## 23 Bibliografie:

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

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

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

31 [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),  
33 2008.

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

37 [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.

39 [DDBJ] <http://www.ddbi.nig.ac.jp/Welcome-c.html> [ENA] <http://www.ebi.ac.uk/ena/>

[GenBank] <http://www.ncbi.nlm.nih.gov/pubmed/21071399> [NCBI]  
41 <http://www.ncbi.nlm.nih.gov/gcnomc>

[NW-99] C. G. Nevill-Manning, I.H. Witten, "Protein is incompressible", In Proceedings  
43 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",  
45 Bioinformatics Vol. 25, pp. 274-275, 2009.



# RO 129751 B1

- [FLCB-11] M. H.Y. Fritz, R. Leinonen, G. Cochrane and E. Bimey, "Efficient storage of high throughput DNA sequencing data using reference-based compression", *Genome Res.* Vol. 21, pp.734-740, 2011. 1  
3
- [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. 5
- [RA-11] P. Rajarajeswari, A. Apparao, "DNABIT Compress - Genome compression algorithm", *Bioinformation*, Vol. 5(8), pp. 350-360, 2011. 7
- [CDLT04] C. R. Calladine, H. R. Drew, B. F. Luisi, A. A. Travers, "Understanding DNA The Molecule & How It Works", Academic Press, April 2004. 9
- [Ver-26] G. S. Vemam, "Cipher Printing Telegraph Systems", *Journal of the American Institute of Electrical Engineers*, Vol. XLV, pp. 109-115, 1926. 11
- [Sha-49] C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, Vol. 28, No. 4, pp. 656-715, 1949. 13

# RO 129751 B1

## Revendicări

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

7           i) citirea mesajului în clar la partea emițătoare și stabilirea datelor de intrare formate  
din:

          - lungimea mesajului în clar,

9           - un cod corespunzător modului de formare a cheii de criptare,

          - numărul de structuri ADN din baza de date (**BD ADN**) biologică care se vor  
11 utiliza la generarea cheii ( $K_{ADN}$ ) de criptare;

          ii) generarea cheii de criptare ( $K_{ADN}$ ) constând într-o secvență aleatoare de structuri ADN  
13 preluate dintr-o bază de date (**BD ADN**) biologică, publică sau privată;

          iii) conversia cheii ADN la formatul binar, obținând cheia OTP ( $K_A$ );

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

17           v) crearea cheii secrete din datele de intrare (**DI**) într-un bloc de criptare ( $E_A$ , respectiv  
 $E_B$ );

19           vi) transmiterea mesajului criptat ( $C_A$ , respectiv  $C_B$ ) și a cheii secrete părții receptoare;

          vii) decriptarea cheii secrete într-un bloc de decriptare ( $D_B$ , respectiv  $D_A$ ) la receptor și  
21 obținerea datelor (**DI**) de intrare aferente;

          viii) generarea cheii ADN la receptor folosind datele (**DI**) de intrare noi și baza de date  
23 (**BD ADN**) biologică;

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

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

27           2. Metodă de criptare bazată pe secvențe aleatoare determinate din structuri ADN,  
conform revendicării 1, **caracterizată prin aceea că**, cheia secretă care se transmite odată cu  
29 mesajul criptat, de la partea emițătoare către partea receptoare, este formată dintr-un antet care  
are structura următoare:

          - un cod ( $k_1$ ) pe 2 biți care reprezintă modul de formare a cheii de criptare;

31           - un cod ( $k_2$ ) pe 3 biți reprezentând numărul de ID-uri ale structurilor ADN utilizate la  
obținerea cheii de criptare;

33           - și dintr-o înșiruire ( $k_3$ ) a ID-urilor structurilor ADN utilizate la obținerea cheii de criptare  
( $K_{ADN}$ ).

35           3. Sistem de criptare de tip OTP bazat pe secvențe aleatoare determinate din structuri  
ADN, ce folosește metoda de la revendicarea 1, utilizat în transmisiuni de tip duplex, alcătuit  
37 din două părți, una emițătoare de mesaje și cealaltă receptoare, **caracterizat prin aceea că**,  
fiecare dintre cele două părți este compusă dintr-o bază de date ADN (**BD ADN**), publică sau  
39 privată, identică la ambele părți, un bloc al datelor de intrare (**DI**), un generator (**Gen**  $K_{ADN}$ ) de  
cheie ADN, un convertor (**Conv ADN-B**) al cheii ADN în cheie binară, un sumator (**S**) modulo  
41 2 cu rol de criptare a mesajului, un bloc (**E**), de criptare a datelor de intrare cu care se gene-  
rează cheia secretă și un bloc de decriptare a acesteia **D**, mesajul criptat și cheia secretă fiind  
43 transmise de la emițător către receptor, receptorul obținând datele de intrare prin decriptarea  
cheii secrete și continuând cu generarea cheii ADN pe care o utilizează la decriptarea mesajului  
45 recepționat.

# RO 129751 B1

4. Sistem de criptare de tip OTP bazat pe secvențe aleatoare determinate din structuri ADN, conform revendicării 3, <b>caracterizat prin aceea că utilizează următorul protocol:</b>	1
1. Transmisia $A \rightarrow B$	3
a. Criptarea mesajului de transmis, la partea emițătoare A:	
(1) Stabilirea datelor de intrare: lungimea mesajului în clar, modul de generare a cheii și alegerea secvențelor ADN;	5
(2) Generarea cheii ADN ( $K_{ADN}$ ) din datele de intrare și baze de date genetice a utilizatorului ( <b>A</b> );	7
(3) Generarea cheii OTP ( $K_a$ ) prin conversia cheii ADN în format binar;	9
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public și transmiterea către partea receptoare ( <b>B</b> );	11
(5) Criptarea OTP a datelor $M_A$ : $C_A = K_a \oplus M_A$ și transmiterea criptogramei ( $C_A$ ) la <b>B</b> ;	
b. Decriptarea mesajului la partea receptoare B	13
(6) Decriptarea datelor de intrare, realizată în blocul ( $D_B$ );	
(7) Generarea cheii ( $K_A$ ) din datele de intrare obținute la (6) utilizând baza de date ADN identică cu cea de la ( <b>A</b> ) și același generator de cheie ( <b>Gen <math>K_{ADN}</math></b> ) și ( <b>Conv ADN - B</b> );	15
(8) Decriptarea OTP a criptogramei ( $C_a$ ) utilizând cheia ( $K_a$ ) obținută la (7): $C_a \oplus K_a = M_A$	17
2. Transmisia $B \rightarrow A$	19
a. Criptarea mesajului de transmis, la partea emițătoare ( <b>B</b> )	
(1) Stabilirea datelor de intrare: lungimea mesajului în clar, alegerea secvențelor ADN, modul de generare a cheii;	21
(2) Generarea cheii ADN( $K_{ADN}$ ) din datele de intrare și baze de date genetice a utilizatorului ( <b>B</b> );	23
(3) Generarea cheii OTP ( $K_B$ ) prin conversia cheii ADN la format binar binar;	25
(4) Criptarea datelor de intrare utilizând un algoritm simetric sau public și transmiterea către partea receptoare ( <b>A</b> )	27
(5) Criptarea OTP a datelor $M_B$ : $C_B = K_B \oplus M_B$ și transmiterea criptogramei ( $C_B$ ) la ( <b>A</b> );	
b. Decriptarea mesajului la partea receptoare ( <b>A</b> )	29
(6) Decriptarea datelor de intrare, realizata in blocul ( $D_A$ )	
(7) Generarea cheii ( $K_B$ ) din datele de intrare obținute la (6) utilizând baza de date ADN identică cu cea de la ( <b>B</b> ) și același generator ( <b>Gen <math>K_{ADN}</math></b> și <b>Conv ADN - B</b> ) de cheie	31
(8) Decriptarea OTP a criptogramei ( $C_B$ ) utilizând cheia ( $K_B$ ) obținută la (7): $C_B \oplus K_B = M_B$ .	33
5. Sistem de criptare de tip OTP bazat pe secvențe aleatoare determinate din structuri ADN, ce folosește metoda de la revendicarea 1, alcătuit din două unități, una de înscriere și alta de citire utilizat în sisteme de stocare, <b>caracterizat prin aceea că</b> , unitatea de înscriere este compusă dintr-o bază de date ADN ( <b>BD ADN</b> ), publică sau privată , un bloc al datelor de intrare ( <b>DI</b> ), un generator ( <b>Gen <math>K_{ADN}</math></b> ) de cheie ADN, un convertor ( <b>Conv ADN-B</b> ) al cheii ADN în cheie binară, un sumator ( <b>S</b> ) modulo 2 cu rol de criptare a datelor, un bloc ( <b>E</b> ) de criptare a datelor de intrare, iar unitatea de citire este compusă dintr-o bază de date ADN ( <b>BD ADN</b> ), identică cu baza de date ADN a unității de scriere, un bloc de decriptare a datelor de intrare ( <b>D</b> ), un generator ( <b>Gen <math>K_{ADN}</math></b> ) de cheie ADN, un convertor ( <b>Conv ADN-B</b> ) al cheii ADN în cheie binară și un sumator ( <b>S</b> ) modulo 2 cu rol de decriptare a datelor.	35
	37
	39
	41
	43

# RO 129751 B1

- 1           6. Sistem de criptare de tip OTP bazat pe secvențe aleatoare determinate din structuri  
ADN, conform revendicării 5, **caracterizat prin aceea că** utilizează următorul protocol:
- 3           a. Criptarea - la unitatea de înscriere
- 5           (1) Stabilirea datelor de intrare: lungimea mesajului în clar, modul de generare a cheii  
și alegerea secvențelor ADN;
- 7           (2) Generarea cheii ADN ( $K_{ADN}$ ) din datele de intrare și baze de date genetice;
- 9           (3) Generarea cheii OTP ( $K$ ) prin conversia cheii ADN la formatul binar;
- 11          (4) Criptarea datelor de intrare realizată în blocul ( $E$ ) utilizând un algoritm simetric sau  
public și memorarea pe mediul de stocare;
- 13          (5) Criptarea OTP a datelor  $M$ :  $C = K \oplus M$  și memorarea criptogramei ( $C$ ) pe mediul de  
stocare;
- 15          b. Decriptarea - la unitatea de citire
- (6) Decriptarea datelor de intrare, realizată în blocul ( $D$ );
- (7) Generarea cheii ( $K$ ) din datele de intrare obținute la (6) utilizând baza de date ADN  
identică cu cea de la scriere și același generator (**Gen  $K_{ADN}$  și Conv ADN**) de cheie;
- (8) Decriptarea OTP a criptogramei ( $C$ ) utilizând cheia ( $K$ ) obținută la (7).

(51) Int.Cl.

H04L 9/28 (2006.01);

G06N 3/12 (2006.01)

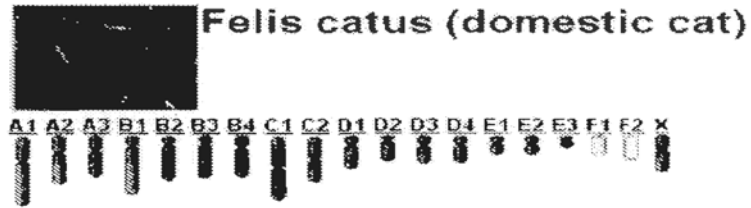


Fig. 1



Fig. 2

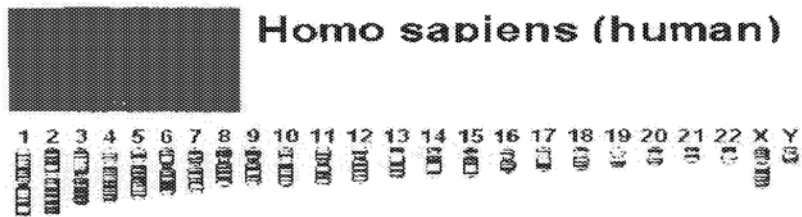


Fig. 3

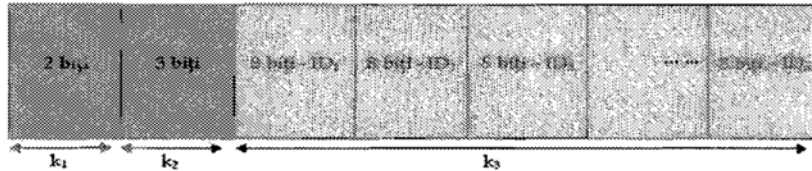


Fig. 4

```

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:224304768
DELINK      BioProject: EBI:NA12878
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo_sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

a) ID1 - CM000663

LOCUS       CM000664             243199371 bp    DNA     linear   CON 29-JUN-2009
DEFINITION  Homo sapiens chromosome 2, GRCh37 primary reference assembly.
ACCESSION   CM000664
VERSION     CM000664.1   GI:224304767
DELINK      BioProject: EBI:NA12878
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo_sapiens

b) ID2 - 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
DELINK      BioProject: EBI:NA12878
KEYWORDS    WGS.
SOURCE      Felis catus (domestic cat)
ORGANISM    Felis_catus

c) ID3 - CM001378
    
```

Fig. 5



Fig. 6

(51) Int.Cl.

H04L 9/28 (2006.01);

G06N 3/12 (2006.01)

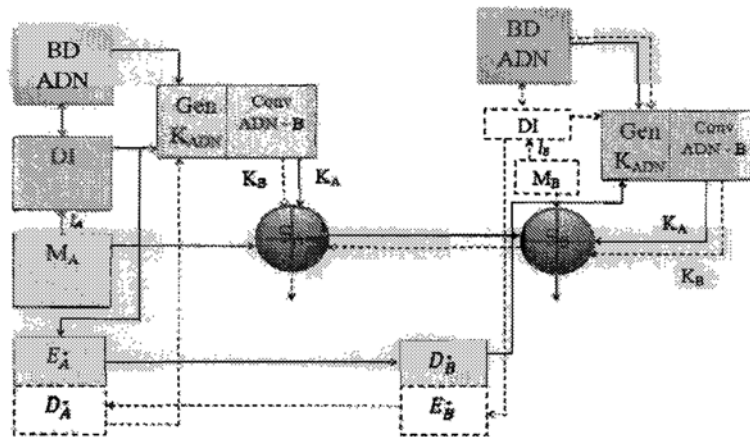


Fig. 7

(51) Int.Cl.

H04L 9/28 (2006.01);

G06N 3/12 (2006.01)

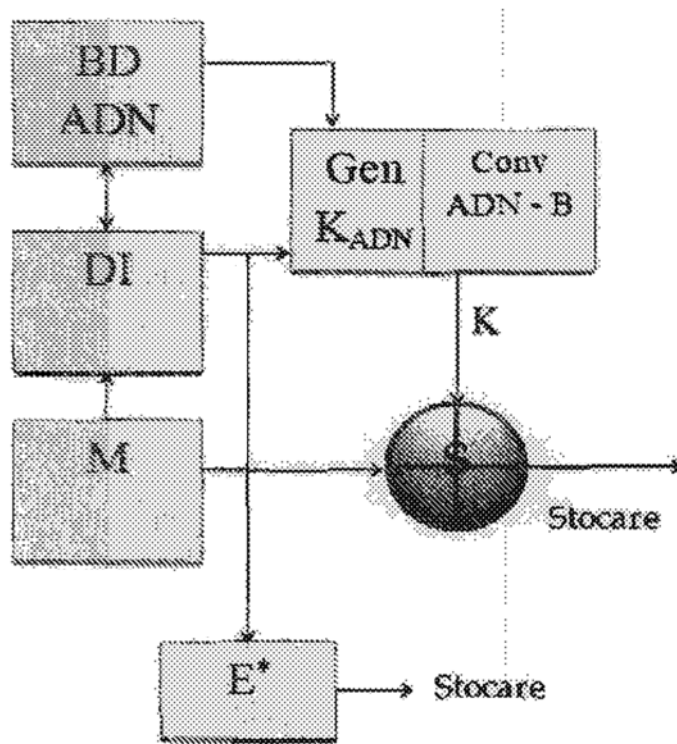


Fig. 8



(51) Int.Cl.

H04L 9/28 (2006.01);

G06N 3/12 (2006.01)

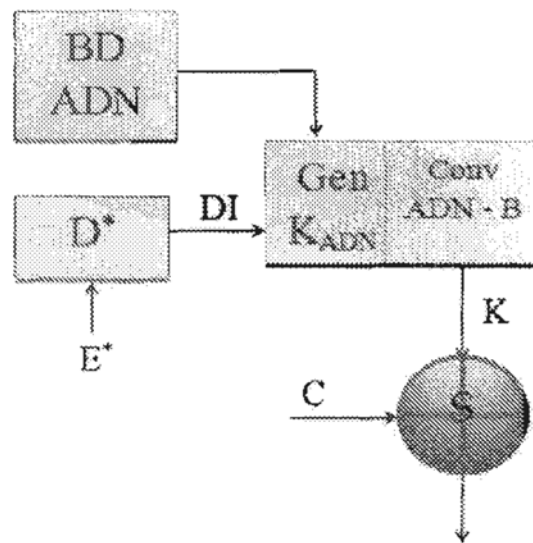


Fig. 9

