

# **CRIPTANALIZA. REZULTATE ȘI TEHNICI MATEMATICE**

Ediția I apărută la: Ed. Univ. Buc, 2004, ISBN 973575975-6.

Vasile PREDA, Emil SIMION și Adrian POPESCU

Ediția a doua 2011



# Cuprins

<b>1</b>	<b>INTRODUCERE</b>	<b>15</b>
<b>2</b>	<b>NOȚIUNI GENERALE</b>	<b>19</b>
2.1.	Obiectul criptanalizei . . . . .	19
2.2.	Criteria și standarde . . . . .	20
2.2.1.	Beneficii ale standardelor . . . . .	20
2.2.2.	Organisme de standardizare . . . . .	21
2.2.3.	Standardele ISO 15408 și FIPS 140-2 . . . . .	22
2.3.	Modelul OSI (Open System Interconectation) . . . . .	22
2.3.1.	Definirea nivelurilor rețelei . . . . .	22
2.3.2.	Nivelul fizic . . . . .	23
2.3.3.	Nivelul legătură date . . . . .	23
2.3.4.	Nivelul rețea . . . . .	24
2.3.5.	Nivelul transport . . . . .	24
2.3.6.	Nivelul sesiune . . . . .	24
2.3.7.	Nivelul prezentare . . . . .	24
2.3.8.	Nivelul aplicație . . . . .	25
2.3.9.	Protocolul TCP/IP . . . . .	25
2.4.	Testarea sistemelor criptografice . . . . .	25
2.4.1.	Introducere . . . . .	25
2.4.2.	Programul de validare a modulelor criptografice . . . . .	26
2.4.3.	Proceduri de certificare și autorizare . . . . .	28
2.4.4.	Autoritate de certificare . . . . .	28
2.5.	Procesul de selectare a modulelor criptografice . . . . .	29
2.5.1.	Faza de planificare . . . . .	29
2.5.2.	Faza de definire a cerințelor și specificațiilor de securitate . . . . .	31
2.5.3.	Faza de achiziție . . . . .	31
2.5.4.	Faza de operare . . . . .	32
2.6.	Operații în criptanaliză . . . . .	32

2.6.1.	Principii criptanalitice . . . . .	32
2.6.2.	Criterii de evaluare . . . . .	33
2.6.3.	Patru operații de bază ale criptanalizei . . . . .	33
2.6.4.	Evaluare și spargere . . . . .	34
2.7.	Clasificări ale atacurilor criptanalitice . . . . .	38
2.7.1.	Tipuri de atac asupra algoritmilor de cifrare . . . . .	38
2.7.2.	Tipuri de atac asupra cheilor . . . . .	40
2.7.3.	Tipuri de atac asupra protocoalelor de autentificare . . . . .	41
2.7.4.	Tipuri de atac asupra sistemului . . . . .	42
2.7.5.	Atacuri hardware asupra modulelor criptografice . . . . .	42
2.8.	Aplicații . . . . .	43
<b>3</b>	<b>TEORIA COMPLEXITĂȚII ALGORITMILOR</b>	<b>45</b>
3.1.	Introducere . . . . .	45
3.2.	Algoritmi și mașini Turing . . . . .	45
3.3.	Teoria problemelor $NP$ - complete . . . . .	46
3.4.	Exemple de probleme $NP$ - complete . . . . .	48
3.5.	Limite actuale ale calculatoarelor . . . . .	51
3.6.	Aplicații . . . . .	52
<b>4</b>	<b>ANALIZA STATISTICO-INFORMAȚIONALĂ</b>	<b>53</b>
4.1.	Noțiuni teoretice . . . . .	53
4.2.	Generatoare și teste statistice . . . . .	54
4.2.1.	Generatoare uniforme . . . . .	54
4.2.2.	Conceptul de test statistic . . . . .	54
4.2.3.	Modele statistice pentru generatoare . . . . .	56
4.2.4.	Teste elementare de aleatorism statistic . . . . .	57
4.2.5.	Interpretarea rezultatelor testelor statistice . . . . .	58
4.3.	Entropia variabilelor aleatoare discrete . . . . .	59
4.4.	Surse de aleatorism de numere întregi . . . . .	62
4.4.1.	Formula analitică a probabilității ideale a unei surse de aleatorism de numere întregi . . . . .	62
4.4.2.	Metoda de calcul efectiv al lui $p$ respectiv $q$ . . . . .	63
4.5.	Metode de decorelare . . . . .	64
4.5.1.	Metode deterministe . . . . .	64
4.5.2.	Metode nedeterministe . . . . .	64
4.6.	Teste statistice de aleatorism . . . . .	65
4.6.1.	Algoritmul de implementare al testului frecvenței . . . . .	65
4.6.2.	Algoritmul de implementare al testului serial . . . . .	66
4.6.3.	Algoritmul de implementare al testului succesiunilor . . . . .	67

4.6.4.	Algoritmul de implementare al testului autocorelației temporale	68
4.6.5.	Algoritmul de implementare al testului autocorelațiilor temporale . . . . .	68
4.6.6.	Algoritmul de implementare al testului autocorelației circulare	69
4.6.7.	Algoritmul de implementare al testului autocorelațiilor circulare	70
4.6.8.	Algoritmul de implementare al testului poker . . . . .	70
4.6.9.	Algoritmul de implementare al testului <i>CUSUM</i> (sumelor cumulate) . . . . .	72
4.6.10.	Algoritmul de implementare al testului de aproximare a entropiei	75
4.6.11.	Algoritmul de implementare al testului lui Maurer (1992) și testul entropiei . . . . .	76
4.6.12.	Algoritmul de implementare al testului $\chi^2$ . . . . .	78
4.6.13.	Algoritmul de implementare al testului Kolmogorov-Smirnov	80
4.6.14.	Testul spectral (transformarea Fourier discretă) . . . . .	81
4.6.15.	Teste de corelație . . . . .	83
4.6.16.	Algoritmul de implementare al testului corelațiilor temporale și circulare . . . . .	83
4.6.17.	Creșterea sensibilității algoritmilor de testare statistică . . . . .	83
4.7.	Teste de aleatorism algoritmic . . . . .	85
4.7.1.	Scurt istoric . . . . .	85
4.7.2.	Măsurarea complexității . . . . .	86
4.7.3.	Complexitatea segmentului . . . . .	89
4.7.4.	Complexitatea segmentului ca măsură a aleatorismului . . . . .	90
4.8.	Teste de necorelare algoritmică . . . . .	92
4.8.1.	Formularea problemei . . . . .	92
4.8.2.	Principii de test . . . . .	92
4.9.	Teste de verificare a jocurilor de tip Casino . . . . .	93
4.9.1.	Metoda $3-\sigma$ pentru ruletă . . . . .	94
4.9.2.	Metoda $3-\sigma$ pentru diferențe la ruletă . . . . .	94
4.9.3.	Metoda $X^2$ pentru ruletă . . . . .	94
4.9.4.	Metoda $X^2$ aplicată diferențelor pentru ruletă . . . . .	95
4.9.5.	Metoda $X^2$ pentru jocurile de tip loto . . . . .	95
4.10.	Aplicații . . . . .	95
<b>5</b>	<b>CODIFICAREA IN ABSENȚA PERTURBAȚIEI</b>	<b>99</b>
5.1.	Introducere . . . . .	99
5.2.	Codificarea în absența perturbației . . . . .	99
5.3.	Codurile Fano și Huffman . . . . .	102
5.3.1.	Algoritmul de implementare a codului Fano . . . . .	102
5.3.2.	Algoritmul de implementare a codului Huffman . . . . .	102

5.4.	Coduri optime . . . . .	103
5.5.	Aplicații . . . . .	104
<b>6</b>	<b>CRIPTANALIZA CIFRURILOR CLASICE</b>	<b>109</b>
6.1.	Substituția simplă și multiplă . . . . .	109
6.1.1.	Substituția simplă . . . . .	109
6.1.2.	Substituția multiplă . . . . .	111
6.2.	Substituția polialfabetică . . . . .	113
6.2.1.	Caracteristicile și identificarea sistemelor de substituție polialfabetică . . . . .	113
6.2.2.	Atacul sistemelor polialfabetice . . . . .	114
6.3.	Soluția unui cifru de substituție . . . . .	114
6.4.	Transpoziția . . . . .	115
6.5.	Sisteme mixte . . . . .	115
6.6.	Proceduri de identificare a sistemului . . . . .	115
6.6.1.	Funcția Kappa . . . . .	116
6.6.2.	Funcția Chi . . . . .	117
6.6.3.	Funcția Psi . . . . .	118
6.6.4.	Funcția Phi . . . . .	120
6.7.	Funcții simetrice de frecvență a caracterelor . . . . .	121
6.8.	Atac stereotip asupra cifrurilor de substituție . . . . .	122
6.9.	Atac de tip frecvență maximă a cifrurilor de substituție . . . . .	122
6.10.	Concluzii . . . . .	123
6.11.	Aplicații . . . . .	124
<b>7</b>	<b>CRIPTANALIZA CIFRURILOR FLUX</b>	<b>127</b>
7.1.	Atacul generatoarelor pseudoaleatoare . . . . .	127
7.2.	Criptanaliza liniară . . . . .	128
7.2.1.	Complexitatea liniară . . . . .	128
7.2.2.	Algoritmul Berlekamp-Massey. Rezultate teoretice . . . . .	134
7.2.3.	Implementarea algoritmului Berlekamp-Massey . . . . .	134
7.2.4.	Testul Berlekamp ca test statistico-informațional . . . . .	135
7.3.	Metoda corelației . . . . .	137
7.4.	Metoda corelației rapide . . . . .	137
7.4.1.	Transformata Walsh-Hadamard . . . . .	137
7.4.2.	Testul statistic Walsh-Hadamard . . . . .	141
7.4.3.	Caracterizarea proprietăților criptografice . . . . .	144
7.5.	Atacul Siegenthaler . . . . .	148
7.6.	Atacul consistenței liniare . . . . .	148
7.7.	Metoda sindromului linear . . . . .	149

7.7.1.	Formularea problemei . . . . .	149
7.7.2.	Preliminarii teoretice . . . . .	149
7.7.3.	Algoritmul Sindromului Linear . . . . .	150
7.7.4.	Numere critice și numere de rafinare . . . . .	150
7.8.	Corecția iterativă a erorii . . . . .	154
7.8.1.	Prezentare generală . . . . .	154
7.8.2.	Prezentarea algoritmilor de corecție iterativă . . . . .	155
7.8.3.	Rezultate experimentale . . . . .	157
7.8.4.	Concluzii . . . . .	157
7.9.	Algoritm de criptanaliză diferențială . . . . .	159
7.10.	Câteva tehnici de proiectare . . . . .	160
7.10.1.	Transformarea neliniară feed-forward . . . . .	160
7.10.2.	Generatorul Geffe . . . . .	160
7.10.3.	Generatorul Jennings . . . . .	161
7.10.4.	Generatorare cu tact controlat . . . . .	162
7.10.5.	Generatoare cu ceasuri multiple . . . . .	165
7.10.6.	Generatoare autodecimate . . . . .	166
7.11.	Exemplu de atac criptanalitic . . . . .	166
7.12.	Aplicații . . . . .	168
<b>8</b>	<b>CRIPTANALIZA CIFRURILOR BLOC</b>	<b>173</b>
8.1.	Introducere și concepte generale . . . . .	173
8.2.	Securitatea și complexitatea atacurilor . . . . .	174
8.3.	Criterii de evaluare a cifrurilor bloc . . . . .	174
8.4.	Moduri de operare . . . . .	174
8.4.1.	Modul ECB (electronic code-book) . . . . .	175
8.4.2.	Modul CBC (cipher-block chaining) . . . . .	176
8.4.3.	Modul CFB (cipher feedback) . . . . .	179
8.4.4.	Modul OFB (output feedback) . . . . .	180
8.4.5.	Modul BC (block chaining) . . . . .	182
8.4.6.	Modul BC cu sumă de control (BC-checksum) . . . . .	182
8.4.7.	Modul OFBNLF (output feedback block with a nonlinear function) . . . . .	182
8.4.8.	Cascade de cifruri și cifrări multiple . . . . .	183
8.5.	Generarea tabelor de substituție . . . . .	186
8.6.	Criptanaliza diferențială . . . . .	187
8.7.	Criptanaliza liniară . . . . .	187
8.8.	Alte metode . . . . .	187
8.9.	Implementări și rezultate experimentale . . . . .	188
8.9.1.	Implementarea standardului de cifrare A.E.S. . . . .	188

8.9.2.	Testarea algoritmului AES . . . . .	189
8.9.3.	Rezultate experimentale . . . . .	189
8.9.4.	Interpretarea rezultatelor . . . . .	192
8.10.	Concluzii . . . . .	192
8.11.	Aplicații . . . . .	193
<b>9</b>	<b>CRIPTANALIZA CIFRURILOR CU CHEI PUBLICE</b>	<b>197</b>
9.1.	Principii de bază . . . . .	197
9.1.1.	Introducere . . . . .	197
9.1.2.	Securitatea algoritmilor cu cheie publică . . . . .	198
9.1.3.	Comparații ale algoritmilor asimetrice și a algoritmilor simetrici	198
9.2.	Algoritmi de tip rucsac . . . . .	199
9.2.1.	Algoritmi rucsac supercrescător . . . . .	199
9.2.2.	Crearea cheii publice din cheia privată . . . . .	199
9.2.3.	Cifrarea . . . . .	200
9.2.4.	Descifrarea . . . . .	200
9.2.5.	Implementarea efectivă . . . . .	200
9.3.	Algoritmul RSA . . . . .	200
9.3.1.	Descrierea principiilor de cifrare și descifrare . . . . .	200
9.3.2.	Viteza algoritmilor tip RSA . . . . .	201
9.3.3.	Securitatea RSA-ului . . . . .	203
9.3.4.	Tipuri de atacuri asupra algoritmilor RSA . . . . .	204
9.3.5.	Trape în generarea cheilor RSA . . . . .	209
9.4.	Algoritmul Pohlig-Hellman . . . . .	209
9.5.	Algoritmul Rabin . . . . .	210
9.6.	Algoritmul ElGamal . . . . .	211
9.7.	Curbe eliptice . . . . .	211
9.8.	Aplicații practice . . . . .	213
9.9.	Teste de primalitate și metode de factorizare . . . . .	214
9.9.1.	Teste de primalitate . . . . .	214
9.9.2.	Metode de factorizare . . . . .	217
9.9.3.	Metode de generare a numerelor prime . . . . .	217
9.10.	Infrastructura Cheilor Publice (PKI) . . . . .	218
9.10.1.	Elementele PKI . . . . .	218
9.10.2.	Ghid de folosire a tehnologiei PKI în rețele deschise . . . . .	219
9.10.3.	Riscuri ale utilizării tehnologiei PKI . . . . .	220
9.10.4.	Standarde ale PKI . . . . .	221
9.11.	Aplicații . . . . .	221



<b>10</b>	<b>CRIPTANALIZA SEMNĂTURILOR DIGITALE</b>	<b>225</b>
10.1.	Prezentare generală . . . . .	225
10.2.	Noțiuni preliminare . . . . .	226
10.3.	Funcții hash . . . . .	228
10.3.1.	Generalități . . . . .	228
10.3.2.	Algoritmi hash . . . . .	229
10.3.3.	Funcții hash bazate pe cifruri bloc . . . . .	230
10.3.4.	Funcții hash nebazate pe cifruri bloc . . . . .	230
10.4.	Modalități de realizare a semnăturilor digitale . . . . .	231
10.4.1.	Aplicarea criptosistemelor simetrice . . . . .	231
10.4.2.	Aplicarea criptosistemelor asimetrice . . . . .	232
10.4.3.	Apelarea la funcții hash unidirecționale . . . . .	232
10.4.4.	Semnături digitale convenționale (normale) . . . . .	233
10.5.	Alte tipuri de semnături digitale . . . . .	235
10.5.1.	Semnătura invizibilă . . . . .	235
10.5.2.	Semnături fail-stop . . . . .	235
10.6.	Legislația în domeniu . . . . .	235
10.7.	Aplicații . . . . .	236
<b>11</b>	<b>CRIPTANALIZA PROTOCOALELOR CRIPTOGRAFICE</b>	<b>237</b>
11.1.	Protocoale elementare . . . . .	237
11.1.1.	Protocoale de schimb de chei . . . . .	237
11.1.2.	Protocoale de autentificare . . . . .	241
11.1.3.	Autentificarea și schimbul de chei . . . . .	244
11.1.4.	Protocoale de transfer orb . . . . .	248
11.1.5.	Analiza formală a protocoalelor de autentificare și a protocoalelor de schimb de chei . . . . .	248
11.1.6.	Divizarea secretului . . . . .	249
11.1.7.	Partajarea secretului . . . . .	249
11.2.	Protocoale avansate . . . . .	249
11.2.1.	Protocol de tip demonstrație convingătoare fără detalii . . . . .	249
11.2.2.	Protocol de tip dezvăluire minimă . . . . .	250
11.2.3.	Protocol de tip dezvăluire zero . . . . .	250
11.2.4.	Protocoale de tip transfer bit și aplicații . . . . .	250
11.2.5.	Alte protocoale avansate . . . . .	254
11.3.	Divizarea și partajarea secretelor . . . . .	254
11.3.1.	Protocol de divizare a secretului . . . . .	255
11.3.2.	Protocolul de partajare LaGrange . . . . .	255
11.3.3.	Protocolul de partajare vectorial . . . . .	256
11.3.4.	Protocolul de partajare Asmuth-Bloom . . . . .	256

11.3.5. Protocolul de partajare Karnin-Greene-Hellman . . . . .	256
11.3.6. Atacuri asupra protocoalelor de partajare (divizare) a secretului . . . . .	257
11.4. Exemple de implementare . . . . .	257
11.4.1. Scheme de autentificare . . . . .	257
11.4.2. Algoritmi de schimb al cheilor . . . . .	260
11.5. Aplicații . . . . .	264
<b>12 CRIPTANALIZA SISTEMELOR DE CIFRARE ANALOGICE</b>	<b>265</b>
12.1. Formularea problemei . . . . .	265
12.2. Calcul Operațional . . . . .	265
12.2.1. Transformata Laplace . . . . .	266
12.2.2. Transformata Fourier . . . . .	267
12.2.3. Transformata Fourier Discretă . . . . .	269
12.2.4. Transformata Cosinus Discretă . . . . .	271
12.2.5. Transformata Walsh . . . . .	271
12.2.6. Transformata z . . . . .	272
12.3. Caracterizarea variabilelor aleatoare . . . . .	273
12.4. Conversia Analogic/Digital . . . . .	274
12.4.1. Modulația în puls . . . . .	274
12.5. Cifrarea Analogică . . . . .	275
12.5.1. Inversarea spectrului . . . . .	275
12.5.2. Rotirea spectrului inversat . . . . .	276
12.5.3. Amestecarea spectrului . . . . .	277
12.5.4. Multiplexarea în timp . . . . .	279
12.6. Aplicații . . . . .	279
<b>13 MANAGEMENTUL CHEILOR CRIPTOGRAFICE</b>	<b>281</b>
13.1. Managementul cheilor . . . . .	281
13.2. Generarea cheilor . . . . .	283
13.3. Protecția cheilor criptografice . . . . .	284
13.3.1. Cheie utilizator . . . . .	284
13.3.2. Arhivarea cheilor . . . . .	285
13.3.3. Distrugerea cheilor . . . . .	285
13.4. Lungimea cheilor criptografice . . . . .	286
13.5. Aplicații . . . . .	287
<b>A METODELE ȘI TEHNICILE DE PROGRAMARE</b>	<b>289</b>
A.1. Structuri de date . . . . .	289
A.2. Alocarea memoriei . . . . .	290

A.3. Recursivitate . . . . .	290
A.4. Metoda backtracking . . . . .	290
A.5. Tehnica Divide et Impera . . . . .	291
A.6. Tehnica branch and bound . . . . .	291
A.7. Programarea dinamică . . . . .	292
A.8. Tehnica greedy . . . . .	292
A.9. Aplicații . . . . .	293
<b>B ELEMENTE DE TEORIA PROBABILITĂȚILOR</b>	<b>295</b>
B.1. Caracteristici ale variabilelor aleatoare . . . . .	295
<b>C STATISTICĂ DESCRIPTIVĂ</b>	<b>297</b>
C.1. Momentele unei variabile aleatoare . . . . .	297
C.2. Teoria selecției . . . . .	298
C.3. Aplicații . . . . .	299
<b>D TEORIA ESTIMAȚIEI</b>	<b>301</b>
D.1. Tipuri de estimatori . . . . .	301
D.2. Margini inferioare ale estimatorilor . . . . .	302
D.3. Estimația de verosimilitate maximă . . . . .	304
D.4. Estimația Bayesiană . . . . .	304
<b>E REPATIȚII STATISTICE</b>	<b>307</b>
E.1. Repartiții continue . . . . .	307
E.1.1. Repartiția normală . . . . .	307
E.1.2. Repartiția lognormală . . . . .	308
E.1.3. Repartiția uniformă . . . . .	308
E.1.4. Repartiția exponențială . . . . .	309
E.1.5. Repartiția gama . . . . .	310
E.1.6. Repartiția beta . . . . .	312
E.1.7. Repartiția Cauchy . . . . .	313
E.2. Distribuții discrete . . . . .	313
E.2.1. Distribuția Bernoulli . . . . .	313
E.2.2. Distribuția binomială . . . . .	313
E.2.3. Distribuția Poisson . . . . .	314
E.2.4. Distribuția hipergeometrică . . . . .	314
E.2.5. Distribuția geometrică . . . . .	314
E.3. Calculul numeric al cuantilelor . . . . .	314
E.3.1. Cuantila repartiției normale . . . . .	315
E.3.2. Cuantilele repartiției chi-pătrat . . . . .	315

<b>F</b>	<b>SERII DINAMICE STAȚIONARE</b>	<b>317</b>
F.1.	Exemple de serii dinamice . . . . .	317
F.2.	Procese stochastice . . . . .	317
F.3.	Staționaritate și strict staționaritate . . . . .	319
F.3.1.	Relația dintre Staționaritate și Strict Staționaritate . . . . .	320
F.4.	Estimarea și eliminarea componentelor tendință și sezoniere . . . . .	322
F.4.1.	Eliminarea tendinței în absența componenetei sezoniere . . . . .	323
F.4.2.	Eliminarea simultană a componentelor tendință și sezoniere . . . . .	325
F.5.	Funcția de autocovarianță a unui proces staționar . . . . .	327
F.5.1.	Funcția de autocovarianță de selecție . . . . .	329
<b>G</b>	<b>MODELUL AUTOREGRESIV-MEDIE MOBILĂ</b>	<b>331</b>
G.1.	Modelul autoregresiv AR(p) . . . . .	331
G.2.	Modelul medie mobilă MA(q) . . . . .	332
G.3.	Modelul ARMA(p,q) . . . . .	332
G.4.	Modelul ARIMA(p,d,q) . . . . .	333
G.5.	Probleme puse proceselor ARIMA(p,d,q) . . . . .	333
<b>H</b>	<b>SIMULAREA VARIABILELOR ALEATOARE</b>	<b>335</b>
H.1.	Tehnici de simulare . . . . .	335
H.2.	Legea tare a numerelor mari . . . . .	336
<b>I</b>	<b>ELEMENTE DE TEORIA NUMERELOR</b>	<b>339</b>
I.1.	Teste de primalitate . . . . .	339
I.2.	Lema chinezescă a resturilor . . . . .	340
I.3.	Numărul de numere prime . . . . .	341
I.4.	Simbolurile lui Legendre și Jacobi . . . . .	341
	<b>BIBLIOGRAFIE</b>	<b>343</b>



## Capitolul 6

# CRIPTANALIZA CIFRURILOR CLASICE

*Deciphering is, in my opinion, one of the most fascinating of arts, and I fear I have wasted upon it more time than it deserves.*

*Charles Babbage, 1864*

### 6.1. Substituția simplă și multiplă

#### 6.1.1. Substituția simplă

Operația de cifrare se bazează pe o *corespondență* biunivocă între *alfabetul clar* notat prin  $\mathcal{A}$  și *alfabetul cifrat* notat prin  $\mathcal{C}$ . Pentru exemplificarea ideilor vom presupune că alfabetul clar este format din cele 26 de litere ale *limbii române* (fără diacritice) plus *delimitatorul de cuvânt* spațiul. Alfabetul cifrat poate fi format din aceleași caractere sau doar din cele 26 de litere ale limbii române caz în care spațiul se va înlocui cu cea mai puțin frecventă literă ( $Q$ ) sau se va ignora pur și simplu.

Corespondența dintre cele două alfabetice (presupunem că delimitatorul de cuvânt este înlocuit cu litera  $Q$ ) poate fi:

-aleatoare;

-pseudoaleatoare: plecând de la o parolă se construiește alfabetul cifrat.

Dacă în cazul corespondenței aleatoare lucrurile sunt cât se poate de clare, vom prezenta pe scurt o metodă de construcție a corespondenței în cel de-al doilea caz. Pornind de la o parolă, alfabetul cifrat este construit după următorul algoritm:

- se scriu, o singură dată, în ordinea apariției, literele din parolă;
- se scriu literele alfabetului ce nu apar în parolă.

*Corespondența* între cele două alfabet se realizează după regula alfabet în alfabet după o permutare fixă  $\sigma$  (aceasta poate fi chiar permutarea identică iar la decriptare se aplică același procedeu dar cu inversa permutării  $\sigma$ ).

În funcție de forma permutării substituția se numește:

-*directă* (alfabetul cifrat are același sens lexicografic cu alfabetul clar, sunt în total 26 astfel de substituții). Exemplu de substituție directă:

A	B	C	D	E	F	G	H	I	J	K	L	M
G	H	I	J	K	L	M	N	O	P	Q	R	S

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
T	U	V	W	X	Y	Z	A	B	C	D	E	F

-*inversă* (alfabetul cifrat are sens invers lexicografic cu alfabetul clar, sunt în total 26 de astfel de substituții). Exemplu de substituție inversă:

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

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

Reamintim aici trei exemple celebre (vechile coduri ebraice) de substituții reciproce (dacă litera  $\mathcal{X}$  se substituie cu litera  $\mathcal{Y}$  atunci  $\mathcal{Y}$  se va substitui cu  $\mathcal{X}$ ) și anume:

-*atbash* (prima jumătate a literelor alfabetului se mapează în cea de-a două jumătate în ordine invers lexicografică):

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

-*albam* (prima jumătate a literelor alfabetului se mapează în cea de-a două jumătate în ordine lexicografică):

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

-*atbah*:

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

**Definiția 6.1.1.** Un *cifru de substituție liniar* de la  $\mathbf{Z}_m$  la  $\mathbf{Z}_m$  ( $m$  fiind numărul de caractere al alfabetului sursă) poate fi descris prin funcția  $f : \mathbf{Z}_m \rightarrow \mathbf{Z}_m$  definită prin  $f(x) = \alpha x + \beta$  cu  $\gcd(\alpha, m) = 1$ , funcția de descifrare fiind  $f^{-1}(x) = \alpha^{-1}(x - \beta)$ . Cheia de cifrare sunt numerele  $\alpha$  și  $\beta$ .

**Observația 6.1.1.** Primul manual de analiză criptografică a cifrurilor de substituție a fost scris de *Al-Kindi* [2] în anul 850 AD.

**Observația 6.1.2.** Cifrul de substituție are *proprietatea de confuzie* (ascunderea legăturii dintre textul clar și textul cifrat).

### 6.1.2. Substituția multiplă

În cazul substituției multiple (*substituție poligrafică*)  $M$  caractere din textul clar sunt substituite în  $N$  caractere de text cifrat. Evident pentru ca operația de cifrare să fie reversibilă trebuie ca  $M \geq N$ . Dacă  $M = N = 2$  atunci cifrul se numește cifru de *substituție digrafică*.

**Definiția 6.1.2.** Un cifru de *substituție poligrafică liniar* de la  $\mathbf{Z}_m^k$  la  $\mathbf{Z}_m^k$  ( $m$  fiind numărul de caractere al alfabetului sursă și  $k$  dimensiunea  $k$ -gramei) poate fi descris prin funcția  $f : \mathbf{Z}_m^k \rightarrow \mathbf{Z}_m^k$  definită prin  $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \boldsymbol{\beta}$  cu matricea  $\mathbf{A}$  inversabilă, funcția de descifrare fiind  $f^{-1}(\mathbf{y}) = \mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\beta})$ . Cheia de cifrare este matricea  $\mathbf{A}$  și vectorul  $\boldsymbol{\beta}$ .

#### Identificarea cifrului de substituție poligrafică

Un cifru de substituție poligrafică se poate detecta foarte ușor după frecvența  $N$ -gramelor. Distribuția acestora este departe de a fi uniformă. Evident, o condiție necesară de identificare a acestui tip de cifru este aceea ca dimensiunea mesajului cifrat să fie suficient de mare. O procedură de detectare a lui  $N$  se bazează, de exemplu, pe minimul entropiei  $H$  calculate din  $R$ -grame:

$$H(m; N) = \min_R H(m; R).$$

Alte proceduri de identificare corectă a parametrului  $n$  sunt prezentate în secțiunea *Proceduri de identificare a sistemului*. Pentru a se asigura biunivocitatea parametrul  $M$  trebuie să fie egal cu  $N$ . Construcția tabelii de substituție poligrafică se realizează cu ajutorul frecvenței  $N$ -gramelor din textul cifrat.



### Cifrul celor 4(2) tabele rectangulare

În continuare vom prezenta două sisteme de cifrare de tip substituție digrafică. Literele alfabetelor (clar și cifrat) sunt trecute într-un careu de  $5 \times 5$  (litera  $I$  fiind asimilată literei  $J$ ). Modul de aranjare al literelor alfabetelor clar este fixat apriori (de exemplu ordine lexicografică) iar al alfabetelor cifrate în funcție de parolă. Textul clar este preprocesat astfel încât acesta să fie compatibil cu matricea de cifrare (delimitatorul de cuvânt este ignorat sau este înlocuit cu cea mai puțin frecventă literă, litera  $I$  este asimilată cu litera  $J$ , și în fine dacă este cazul mai adjuncționăm o literă la text pentru a avea un număr par de digrame). Pentru cifrul celor 4 tabele regula de cifrare este dată de regula dreptunghiului: o digramă din textul clar se cifrează în digrama corespunzătoare diagonalei secundare a dreptunghiului determinat de cele două caractere ale digramei clare. Modul de aranjare al celor patru alfabete ( $P1$  și  $P2$  sunt alfabetele clare iar  $C1$  și  $C2$  sunt alfabetele cifrate în funcție de parolă) este prezentat mai jos:

P1	C1
C2	P2

Algoritmul prezentat mai sus degenerază în algoritmul celor două tabele (verticale sau orizontale) după cum urmează:  $P1 \equiv C1$  și  $P2 \equiv C2$  respectiv  $P1 \equiv C2$  și  $P2 \equiv C1$ . Dacă literele ce formează digrama se află pe aceeași coloană, respectiv linie, atunci reprezentarea cifrată a acestora sunt ele însele, respectiv acestea scrise în ordine inversă. Cele două tabele de cifrare sunt:

P1[C1]
C2[P2]

respectiv

P1[C2]	P2[C1]
--------	--------

### Cifrul Playfair

*Cifrul Playfair* (numele acestui sistem de cifrare provine de la Lordul englez Playfair) este unul dintre cele mai cunoscute sisteme de cifrare digrafice (transformă un grup de 2 litere într-un grup de alte două litere). Acest sistem de cifrare este foarte simplu de folosit dar mult mai sigur decât sistemele de substituție monoalfabetice. Descriem în continuare modul de utilizare al acestui sistem de cifrare. Literele alfabetului sunt trecute într-un careu de  $5 \times 5$  (litera  $I$  fiind asimilată literei  $J$ ). Textul clar este preprocesat astfel încât acesta să fie compatibil cu matricea de cifrare (delimitatorul de cuvânt este ignorat sau este înlocuit cu cea mai puțin frecventă literă,

litera  $I$  este asimilată cu litera  $J$ , și în fine dacă este cazul mai adjuncționăm o literă la text pentru a avea un număr par de digrame). Regula de cifrare este următoarea:

i) Dacă digrama ce se dorește cifrată nu are literele pe aceeași linie sau coloană, atunci regula de cifrare este *regula dreptunghiului*, traseul fiind pe verticală de la cea de-a doua literă a digramei către prima literă.

ii) Dacă digrama ce se dorește cifrată are literele pe aceeași linie, atunci se aplică regula: *cifrează la dreapta, descifrează la stânga*.

iii) Dacă digrama ce se dorește cifrată are literele pe aceeași coloană, atunci se aplică regula: *cifrează în jos, descifrează în sus*.

**Observația 6.1.3.** Dacă o digramă apare în textul clar în ordine inversă atunci același lucru se va întâmpla și în textul cifrat.

**Observația 6.1.4.** Cifrul Playfair nu are regulă pentru cifrarea literelor duble: digramele ce conțin două litere identice sunt sparte prin introducerea artificială a unei alte litere.

**Observația 6.1.5.** Cifrul Playfair apare ca o extindere, în sensul reducerii numărului de tabele rectangulare folosite (de la două la unul), al cifrului cu 2 tabele.

Metoda cea mai frecventă de atac a acestui tip de cifru constă în analiza frecvenței digramelor de text clar combinată cu metoda comparației patternurilor din textul cifrat cu patternuri din dicționar. Spectaculos este faptul că în *Manualul Teroristului* [104], în capitolul 13 intitulat *Scrieri Secrete, Coduri și Cifruri*, se prezintă tehnici steganografice, tehnici de codificare și o serie de metode de cifrare clasice cum ar fi substituția simplă și substituția multiplă.

## 6.2. Substituția polialfabetică

### 6.2.1. Caracteristicile și identificarea sistemelor de substituție polialfabetică

Un sistem de cifrare de tip substituție polialfabetică este generalizarea sistemului de cifrare de substituție monoalfabetică. Fie un sistem de cifrare polialfabetic compus dintr-un număr  $N$  de alfabete. Fiecare alfabet reprezintă o permutare (stabilită în funcție de parolă) a alfabetului de intrare. Algoritmul de cifrare constă în substituția celei de a  $i$ -a litere  $m$  din textul clar cu litera corespunzătoare din cel de al  $i \bmod N$  alfabet. Sistemele polialfabetice sunt ușor de identificat prin aplicarea tehnicii frecvențelor secvențelor decimate din textul cifrat sau a valorii maxime a funcției *Kappa* (v. paragraful Proceduri de identificare a sistemului).

### 6.2.2. Atacul sistemelor polialfabetice

Atacul sistemelor polialfabetice este similar cu atacul a  $N$  sisteme de substituție monoalfabetică. Deci, o procedură de *tip divide et impera* are o complexitate de  $O(N)$ . Procedura este descrisă în continuare:

**Intrare:** Textul cifrat de lungime  $M$  suficient de mare.

**Ieșire:** Textul clar corespunzător sistemului de cifrare polialfabetic.

**Pas 1.** Determină numărul de alfabete  $N$ .

**Pas 2.** Pentru  $j = 0$  to  $4$  execută:

pentru  $i = 1$  to  $N - j$  execută:

aplică procedura de reconstrucție parțială (pe baza frecvențelor  $(j + 1)$ -gramelor) a alfabetelor  $i, \dots, i + j$ .

**Pas 3.** Conform celor  $N$  alfabete reconstruiește textul clar.

**Observația 6.2.1.** Procedura descrisă mai sus are ca parametru implicit de analiză numărul maxim de legături 4 : astfel, 1-gramele sunt caracterele, 2-gramele sunt dubletii, etc.

## 6.3. Soluția unui cifru de substituție

În criptanaliză, soluția unui cifru de substituție parcurge următoarele etape:

a) *analiza criptogramelor:*

- a1) pregătirea unui tabel de frecvențe;
- a2) căutarea repetițiilor;
- a3) determinarea tipului de sistem utilizat;
- a4) pregătirea unei foi de lucru;
- a5) pregătirea unui alfabet individual;
- a6) tabelarea repetițiilor lungi.

b) *Clasificarea vocalelor și consoanelor prin studierea:*

- b1) frecvențelor;
- b2) spațiilor;
- b3) combinațiilor de litere;
- b4) repetițiilor.

c) *Identificarea literelor:*

- c1) partiționarea literelor în clase de probabilitate;
- c2) verificarea presupunerilor;
- c3) înlocuirea valorilor corecte în criptogramă;
- c4) descoperirea altor valori pentru a avea soluția completă.

d) *Reconstrucția sistemului:*

- d1) reconstrucția tabelii de cifrare;

- d2) reconstrucția cheilor folosite în operația de cifrare;
- d3) reconstrucția cheilor sau a cuvintelor cheie ce au fost folosite pentru construcția șirurilor de alfabete.

## 6.4. Transpoziția

Dacă substituția schimbă valoarea unui caracter din textul cifrat, permutarea schimbă poziția pe care acest caracter apare în textul cifrat și nu valoarea sa. Operația de cifrare prin permutarea  $\sigma \in S_N$  se realizează prin permutarea caracterelor din blocurile adiacente, de lungime  $N$ , ale textului clar. Un caz particular al permutării este acela al transpoziției de lungime  $N$ . Textul este scris într-o tabelă (completă sau nu) cu  $N$  coloane, literele fiind scrise linie cu linie. Pornind de la o parolă literală se construiește o parolă numerică (spre exemplu se asociază fiecărei litere din parolă poziția sa în scrierea lexicografică a acesteia). Textul este apoi citit coloană cu coloană într-o ordine stabilită apriori (crescător, descrescător, etc.). Pentru descifrare se aplică același algoritm dar cu parola numerică  $\sigma^{-1}$ .

**Observația 6.4.1.** Cifrul de transpoziție (mai general cifrul de permutare) are *proprietatea de difuzie* (disiparea redundanței textului clar de-a lungul textului cifrat).

## 6.5. Sisteme mixte

Sistemele mixte de cifrare au la bază o cifrare succesivă a mesajului prin metoda substituției iar apoi prin metoda transpoziției sau viceversa. Tot ceea ce trebuie să facem acum este să atacăm sistemul de cifrare de la ultima sa componentă către prima. Remarcăm faptul că în cazul substituției simple aceasta este comutativă cu operația de transpoziție deci se poate aborda mai întâi substituția iar apoi transpoziția. În cazul utilizării unui sistem polialfabetic, cu număr necunoscut de alfabete, recomandarea este ca după stabilirea, prin metode statistice, a numărului de alfabete, să se abordeze concomitent identificarea efectivă a alfabetelor și al transpoziției utilizate. În cazul utilizării unui sistem poligrafic (tabele de cifrare) și o transpoziție este recomandabilă o tehnică de tip backtracking.

## 6.6. Proceduri de identificare a sistemului

Procedurile de criptanaliză prezentate în cadrul acestui paragraf sunt bazate pe calculul unor estimatori pentru o serie de funcții de test (vezi *Friedman* [20], *Bauer* [6], *Preda* și *Simion* [58]). Textul clar trebuie să fie omogen din punct de vedere

statistic. Dacă textul nu este omogen, atunci, cu ajutorul unei proceduri, acesta se poate diviza în părți omogene. Procedurile ce urmează ne permit să identificăm modelul de cifrare, structura statistică a textului clar și în cele din urmă soluția problemei.

Să notăm prin  $T = (t_1, \dots, t_M)$  și  $T' = (t'_1, \dots, t'_M)$  două șiruri de lungime  $M$  din același vocabular  $Z_N$  (de lungime  $N$ ). Notăm prin  $m_i$  și  $m'_i$  frecvențele de apariție ale celei de a  $i - a$  litere din alfabetul sursă din șirul  $T$  respectiv  $T'$ . Avem deci:

$$\sum_{i=1}^N m_i = M$$

și

$$\sum_{i=1}^N m'_i = M.$$

În cele ce urmează vom descrie funcțiile de decizie *Kappa*, *Chi*, *Psi* și *Phi*.

### 6.6.1. Funcția Kappa

*Funcția de test Kappa* este definită prin:

$$Kappa(T, T') = \frac{\sum_{i=1}^M \delta(t_i, t'_i)}{M},$$

unde  $\delta$  este simbolul lui *Kronecker*.

Avem inegalitățile:

$$0 \leq Kappa(T, T') \leq 1,$$

cu egalitate în partea stângă pentru  $t_i \neq t'_i$  (pentru orice  $i$ ) respectiv  $T \equiv T'$ . Testul *Kappa* este similar cu testul de corelație. Câteodată *Kappa* se va nota mai simplu prin  $K$ .

Avem următoarele teoreme de invarianță (vezi *Bauer* [6]):

**Teorema 6.6.1.** *Pentru toate sistemele de cifrare de tip substituție polialfabetică valoarea lui Kappa a două texte de lungimi egale, cifrate cu aceeași cheie, este invariantă.*

**Teorema 6.6.2.** *Pentru toate sistemele de cifrare de tip transpoziție valoarea lui Kappa a două texte de lungimi egale, cifrate cu aceeași cheie, este invariantă.*

Valoarea medie a lui  $Kappa(T, T')$ , pentru  $T$  și  $T'$  definite de două surse  $Q$  respectiv  $Q'$  (cu probabilitatea simbolurilor  $p_i$  respectiv  $p'_i$  pentru  $i = 1, \dots, N$ ) este:

$$E[Kappa(T, T')] = \sum_{i=1}^N p'_i p_i.$$

Dacă sursele  $Q$  și  $Q'$  sunt identice (acest lucru se notează prin  $Q \equiv Q'$ ), adică  $p_i = p'_i$ , atunci:

$$E[Kappa(T, T')] = \sum_{i=1}^N p_i^2.$$

**Teorema 6.6.3.** Pentru două surse identice  $Q$  și  $Q'$  are loc inegalitatea:

$$\frac{1}{N} \leq E[Kappa(T, T')] \leq 1,$$

cu egalitate în partea stângă pentru distribuția uniformă, iar în partea dreaptă pentru sursa deterministă (adică există un indice  $i$  pentru care  $p_i = 1$ ).

**Observația 6.6.1.** Dispersia funcției Kappa se calculează după formula:

$$D^2(X) = E((E(X) - X)^2).$$

**Observația 6.6.2.** Funcția Kappa poate fi folosită la atacul sistemelor polialfabetice: diferența dintre valorile maxime ale valorilor  $Kappa(T^{(i)}, T)$  (am notat prin  $T^{(i)}$  textul  $T$  deplasat ciclic cu  $i$  poziții la dreapta) este divizibil cu numărul de alfabet utilizate. Pentru a găsi alfabetele utilizate se poate aplica o tehnică de tip divide et impera și o procedură bazată pe maximul frecvenței sau atac de tip stereotip.

### 6.6.2. Funcția Chi

Funcția Chi este definită prin formula:

$$Chi(T, T') = \frac{\sum_{i=1}^N m_i m'_i}{M^2}.$$

Vom nota această funcție și prin litera grecească  $\chi$ . Au loc următoarele teoreme de invarianță (vezi Bauer [6]):

**Teorema 6.6.4.** Pentru toate sistemele de cifrare de tip substituție monoalfabetică valoarea lui  $Chi$  a două texte de lungimi egale, cifrate cu aceeași cheie, este invariantă.

**Teorema 6.6.5.** Pentru toate sistemele de cifrare de tip transpoziție valoarea lui  $Chi$  a două texte de lungimi egale, cifrate cu aceeași cheie, este invariantă.

Avem inegalitatea:

$$Chi(T, T') \leq 1,$$

cu egalitate pentru  $T \equiv T'$ . Pentru un text cu distribuție uniformă  $T$  ( $m_i = \frac{M}{N}$ ) și un text arbitrar  $T'$  avem:

$$Chi(T, T') = \frac{1}{N}.$$

Valoarea medie a lui  $Chi$  a două texte  $T$  și  $T'$  de lungime egală  $M$  peste același vocabular  $Z_N$  se calculează din probabilitățile  $p_i$  și  $p'_i$  ale frecvenței de apariție ale celui de al  $i$ -lea caracter în sursele stochastice  $Q$  și  $Q'$  ale textelor:

$$E[Chi(T, T')] = \sum_{i=1}^N p_i p'_i.$$

Dacă sursele  $Q$  și  $Q'$  sunt identice atunci  $p_i = p'_i$  pentru orice  $i$  și

$$E[Chi(T, T')] = \sum_{i=1}^N p_i^2.$$

Avem inegalitatea:

$$\frac{1}{N} \leq E[Chi(T, T')] \leq 1,$$

cu egalitate în partea stângă pentru distribuția uniformă iar în partea dreaptă pentru sursa deterministă.

### 6.6.3. Funcția Psi

Funcția  $Psi$  este legată de funcția  $Chi$  prin formula:

$$Psi(T) = Chi(T, T).$$

Vom nota această funcție prin litera grecească  $\psi$ . În mod similar ca pentru funcția  $\chi$  avem următoarele teoreme de invarianță (vezi Bauer [6]).

**Teorema 6.6.6.** Pentru toate sistemele de cifrare de tip substituție monoalfabetică valoarea lui  $Psi$  este invariantă.

**Teorema 6.6.7.** Pentru toate sistemele de cifrare de tip transpoziție valoarea lui  $Psi$  este invariantă.

Avem inegalitatea:

$$\frac{1}{N} \leq Psi(T) \leq 1,$$

cu egalitate în partea stângă pentru un text cu distribuție uniformă ( $m_i = \frac{M}{N}$ ), și în partea dreaptă dacă  $T$  este construit dintr-un singur caracter.

Valoarea medie a lui  $Psi$  al unui text de lungime  $M$  din alfabetul  $Z_N$  se calculează din probabilitățile  $p_i$  ale frecvenței de apariție a celui de al  $i$ -lea caracter din sursa stochastică  $Q$  ce produce textul:

$$E[Psi(T)] = \sum_{i=1}^N p_i^2.$$

Are loc inegalitatea:

$$\frac{1}{N} \leq E[Psi(T)] \leq 1,$$

cu egalitate în partea stângă pentru un text cu distribuție uniformă și în partea dreaptă pentru o sursă deterministă.

Avem teorema *Kappa – Chi* (vezi *Bauer* [6]):

**Teorema 6.6.8.** Pentru două texte  $T$  și  $T'$  peste același vocabular și de aceeași lungime, valoarea  $Chi(T, T')$  este media aritmetică a valorilor  $Kappa(T^{(i)}, T')$  ( $T^{(i)}$  este textul  $T$  rotit ciclic cu  $i$  poziții la dreapta obținut prin formula  $t_j^* = t_{(j-i-1) \bmod M+1}$  pentru  $j = 1, \dots, M$ ):

$$Chi(T, T') = \frac{1}{M} \sum_{i=0}^{M-1} Kappa(T^{(i)}, T').$$

Corolarul următor este cunoscut sub numele de teorema *Kappa – Psi*:

**Corolar 6.6.1.** Pentru un text  $T$  de lungime  $M$  peste vocabularul  $Z_N$  avem:

$$Psi(T) = \frac{1}{M} \sum_{i=0}^{M-1} Kappa(T^{(i)}, T).$$



#### 6.6.4. Funcția Phi

Vom defini acum o nouă funcție denumită *Phi* care este dată prin formula:

$$Phi(T) = \frac{\sum_{i=1}^N m_i(m_i - 1)}{M(M - 1)}.$$

Funcția *Phi* va fi notată prin litera grecească  $\varphi$ . În cazul în care  $T \equiv T'$  avem  $Kappa(T^{(0)}, T) = 1$ , în timp ce pentru  $i \neq 0$  valorile lui  $Kappa(T^{(i)}, T)$  sunt relativ mici. Deci cazul  $i = 0$  este atipic pentru medie și deci este mult mai natural să extindem media numai peste cele  $M - 1$  cazuri:

$$\frac{1}{M - 1} \sum_{i=1}^{M-1} Kappa(T^{(i)}, T).$$

Avem teorema *Kappa - Phi*:

**Teorema 6.6.9.** (*Bauer [6]*) Pentru un text  $T$  de lungime  $M$  peste vocabularul  $Z_N$  valoarea lui  $Phi(T)$  este media valorilor  $Kappa(T^{(i)}, T)$  ( $T^{(i)}$  este textul  $T$  rotit ciclic cu  $i$  poziții la dreapta):

$$Phi(T) = \frac{1}{M - 1} \sum_{i=1}^{M-1} Kappa(T^{(i)}, T).$$

Avem de asemenea următoarele teoreme de invarianță (vezi *Bauer [6]*) pentru funcția *Phi*.

**Teorema 6.6.10.** Pentru toate sistemele de cifrare de tip substituție monoalfabetică valoarea lui *Phi* este invariantă.

**Teorema 6.6.11.** Pentru toate sistemele de cifrare de tip transpoziție, valoarea lui *Phi* este invariantă.

**Teorema 6.6.12.** Dacă  $T$  este o secvență distribuită uniform atunci

$$Phi(T) = \frac{M - N}{(M - 1)N}$$

unde  $M$  este lungimea secvenței iar  $N$  numărul de simboluri (caractere).

**Teorema 6.6.13.** *Dacă notăm prin  $\Phi_c(T)$  valoarea funcției  $\Phi$  calculată pe  $c$ -grame atunci:*

$$\lim_{c \rightarrow \infty} \Phi_c(T) = 0.$$

*Valoarea  $c$  pentru care  $\Phi_c(T) = \max_r \Phi_r(T)$  este dimensiunea blocului pe care se realizează codificarea datelor după procesul de cifrare.*

**Observația 6.6.3.** i) Valoarea medie a lui  $\Phi(T)$  este egală cu valoarea medie a lui  $\Psi(T)$  deci este egală cu  $\sum_{i=1}^n p_i^2$ .

ii) Funcția  $\Phi$  este utilă deoarece literele rare ( $m_i = 1$ ) sunt eliminate.

iii)  $\Phi(T) = 0$  dacă și numai dacă  $0 \leq m_i \leq 1$ .

iv) Funcția  $\Phi$  este utilă în identificarea limbii folosite în textul clar, cu condiția ca sistemul de cifrare folosit să fie o tranpoziție și/sau o substituție monoalfabetică.

v) Funcția  $\Phi$  a fost propusă inițial de *Solomon Kullback* din considerente stochastice.

## 6.7. Funcții simetrice de frecvență a caracterelor

Teoremele de invarianță pentru  $\Psi$  au loc pentru toate funcțiile de frecvență  $m_i$  a caracterelor care sunt simetrice. Cea mai simplă funcție polinomială netrivială este  $\sum_{i=1}^N m_i^a$ . Aceasta aparține familiei de funcții:

$$\Psi_a(T) = \begin{cases} \left( \sum_{i=1}^N \left( \frac{m_i}{M} \right)^a \right)^{\frac{1}{a-1}} & \text{dacă } 1 < a < \infty \\ \exp\left( \sum_{i=1}^N \left( \frac{m_i}{M} \right) \ln\left( \frac{m_i}{M} \right) \right) & \text{dacă } a = 1 \\ \max_{1, \dots, N} \frac{m_i}{M} & \text{dacă } a = \infty \end{cases}$$

cu  $\sum_{i=1}^N \frac{m_i}{M} = 1$ . Observăm că  $\Psi_2$  este  $\Psi$ . Cantitatea logaritmică  $-\ln \Psi_a(T)$  se numește entropia Rényi de ordinul  $a$  a lui  $T$  (Alfred Rényi 1921-1970, matematician ungar). Familia are următoarea reprezentare:

$$-\ln \Psi_a(T) = \begin{cases} -\frac{1}{a-1} \ln\left( \sum_{i=1}^N \left( \frac{m_i}{M} \right)^a \right) & \text{if } 1 < a < \infty \\ -\sum_{i=1}^N \left( \frac{m_i}{M} \right) \ln\left( \frac{m_i}{M} \right) & \text{if } a = 1 \\ -\max_{1, \dots, N} \ln\left( \frac{m_i}{M} \right) & \text{if } a = \infty. \end{cases}$$

Entropia Rényi de ordinul 1 se numește *entropie Shannon* (Claude Shannon 1916-2000, inginer American, părintele teoriei informației) iar entropia Rényi de ordinul 2 este *entropia Kullback*. Toate proprietățile lui  $Psi$  se generalizează la proprietăți ale lui  $Psi_a$ .

## 6.8. Atac stereotip asupra cifrurilor de substituție

Fie  $t_1, \dots, t_M$  mesajul cifrat de lungime  $M$  (suficient de mare). Notăm prin  $c_1, \dots, c_N$  literele care apar în mesajul cifrat și prin  $p_1, \dots, p_N$  frecvența de apariție a acestora. Presupunem că sistemul de cifrare folosit este o substituție monoalfabetică. Notăm prin  $c_1^*, \dots, c_N^*$  alfabetul textului clar și prin  $p_1^*, \dots, p_N^*$  probabilitatea apariției literelor alfabetului. O metodă de rezolvare a acestui sistem de cifrare este atacul brut care constă în generarea tuturor permutărilor alfabetului. Numărul total de permutări ce trebuie generate este de  $N!$  care este imens și deci trebuie găsită o altă cale de atac. Tehnica este de a diviza mulțimea literelor din textul cifrat și din alfabetul clar în  $K$  clase astfel încât dispersia din fiecare clasă este mică. Notăm prin  $m_i$  numărul de litere din cea de a  $i$ -a clasă corespunzătoare textului cifrat și prin  $n_i$  numărul de litere din cea de a  $i$ -a clasă corespunzătoare alfabetului clar. Notăm cu  $k_i = \max(m_i; n_i)$  și  $l_i = \min(m_i; n_i)$ . Astfel trebuie generate numai  $\prod_{i=1}^K A_{k_i}^{l_i}$  (avem  $\sum_{i=1}^K n_i = \sum_{i=1}^K m_i = N$ ) număr relativ mic în raport cu  $N!$ . Pentru a realiza această clasificare vom proceda de exemplu pentru literele  $c_1, \dots, c_N$  după cum urmează:

1) Sortăm literele  $c_1, \dots, c_N$  în ordinea crescătoare a probabilităților  $p_1, \dots, p_N$ . Notăm prin  $\bar{c}_1, \dots, \bar{c}_N$  și prin  $\bar{p}_1, \dots, \bar{p}_N$  literele sortate, respectiv probabilitățile sortate.

2) Calculăm  $\delta_i = \bar{p}_i - \bar{p}_{i-1}$  ( $i = 2, \dots, n$ ). Cele mai mari  $K - 1$  valori ale lui  $\delta_i$  sunt puncte delimitatoare ale claselor.

**Observația 6.8.1.** Pentru texte de lungime mică, în loc de probabilitatea  $p_i$  se folosește  $f_i$  frecvența absolută a literei  $i$ .

## 6.9. Atac de tip frecvență maximă a cifrurilor de substituție

Fie  $t_1, \dots, t_M$  text clar de lungime  $M$  cu  $t_i \in \{0, \dots, 2^L - 1\}$  (textul este codificat pe  $L$ -biți). Notăm prin  $\sigma \in S_{2^L}$  cheia (permutarea) și prin  $c_1, \dots, c_M$  textul permutat (cifrat). Energia informațională a textului este:

$$\sum_{i=0}^{2^L-1} \left( \frac{f_i}{M} \right)^2$$

unde  $f_i$  este frecvența de apariție a caracterului  $i$  ( $i = 0, \dots, 2^L - 1$ ). Utilizând inegalitatea Cauchy-Buniakovsky-Schwartz obținem:

$$\sum_{i=0}^{2^L-1} \frac{f_i}{M} p_{\sigma_i} \leq \sqrt{\sum_{i=0}^{2^L-1} \left(\frac{f_i}{M}\right)^2 \sum_{i=0}^{2^L-1} p_{\sigma_i}^2},$$

unde  $p_{\sigma_i}$  este probabilitatea de apariție a caracterului  $i$  în textul  $T$ . Deoarece

$$\sum_{i=0}^{2^L-1} \left(\frac{f_i}{M}\right)^2 = \sum_{i=0}^{2^L-1} p_{\sigma_i}^2,$$

vom obține

$$\sum_{i=0}^{2^L-1} \frac{f_i}{M} p_{\sigma_i} \leq \sum_{i=0}^{2^L-1} \left(\frac{f_i}{M}\right)^2.$$

Avem egalitate dacă și numai dacă:

$$p_{\sigma_i} = \frac{f_i}{M} \text{ pentru orice } i.$$

Deci pentru a obține permutarea optimă  $\sigma$  (cheia sistemului de cifrare) trebuie rezolvată problema următoare:

$$\left\{ \begin{array}{l} \max_{\sigma} \sum_{i=0}^{2^L-1} \frac{f_i}{M} \hat{p}_{\sigma_i} \\ \sum_{i=0}^{2^L-1} \frac{f_i}{M} = 1, \end{array} \right.$$

unde  $\hat{p}_{\sigma_i}$  este un estimator al lui  $p_{\sigma_i}$ .

**Observația 6.9.1.** Această procedură se poate folosi și ca un test de confirmare a rezultatului obținut prin atacul de tip stereotip.

## 6.10. Concluzii

Funcțiile de test definite în cadrul acestui capitol sunt invariante în cazul folosirii anumitor sisteme de cifrare cum ar fi sistemele de cifrare de tip transpoziție și de substituție mono sau polialfabetică. Aceste funcții de test pot fi folosite în identificarea limbii folosite în textul clar, a sistemului de cifrare folosit, a cheii și în cele din urmă a textului clar. Funcția *Kappa* este folosită în identificarea sistemului de cifrare și în identificarea limbii folosite în textul clar. Procedura de identificare a

limbii se bazează pe compararea valorilor lui *Psi* și *Phi* ale textului cifrat cu valorile lui *Psi* și *Phi* ale fiecărei limbi (aceste funcții de test devin în acest moment teste de confirmare). Funcțiile *Chi* sunt folosite în atacuri de *tip isolog* (texte diferite cifrate cu aceeași cheie). Testele *Kappa*, *Chi*, *Psi* și *Phi* se pot efectua pe digrame sau trigrame, etc.

Alte aplicații criptografice sunt cele de tip căutare inteligentă a parolelor sau de tip dicționar modificat (conform unor reguli lexicale).

Cele mai bune rezultate sunt obținute dacă se analizează textul format numai din litere mari sau mici. Textele pot fi cu delimitator de cuvânt (spațiu) sau fără delimitator de cuvânt.

Descrierea completă a acestor tipurilor de cifruri prezentate în cadrul acestui capitol precum și principalele moduri de atac se poate găsi de exemplu în *Bauer* [6].

## 6.11. Aplicații

**Exercițiul 6.11.1.** Care este diferența dintre proprietatea de confuzie și proprietatea de difuzie a unui algoritm de cifrare?

**Exercițiul 6.11.2.** Să se construiască alfabetul cifrat cu ajutorul parolei de cifrare *TESTARE SISTEM* iar apoi să se cifreze mesajul "IN CRIPTOGRAFIE NICI O REGULA NU ESTE ABSOLUTA". Permutarea ce realizează corespondența este:

0	1	2	3	4	5	6	7	8	9	10	11	12
25	24	23	22	21	20	19	18	17	16	15	14	13

13	14	15	16	17	18	19	20	21	22	23	24	25
12	11	10	9	8	7	6	5	4	3	2	1	0

*Răspuns.* Corepondența dintre alfabetul clar și alfabetul cifrat (înainte de realizarea permutării) este:

A	B	C	D	E	F	G	H	I	J	K	L	M
T	E	S	A	R	I	M	B	C	D	F	G	H

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
J	K	L	N	O	P	Q	U	V	W	X	Y	Z

Corepondența dintre alfabetul clar și alfabetul cifrat după realizarea permutării este:

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

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

Mesajul clar se procesează astfel încât spațiul este înlocuit cu cea mai puțin frecventă literă:

*INQCRIPTOGRAFIEQNICIQREGULAQNUQESTEQAABSOLUTA.*

Mesajul cifrat va fi:

*OHDXCOFMGQCZUOV DHOXODCVQIKZDHIDVBMVDZYBGKIMZ.*

**Exercițiul 6.11.3.** Să se cifreze mesajul "SI IN CRIPTOGRAFIE TACEREA ESTE AUR" cu ajutorul metodei celor 4 tabele inițializate cu parolele de cifrare *CRIPTOGRAFIE* și *TEST*.

**Exercițiul 6.11.4.** Să se construiască matricea de cifrare Playfair cu ajutorul parolei *CRIPTOGRAFIE* iar apoi să se cifreze mesajul "SI IN CRIPTOGRAFIE TACEREA ESTE AUR".

*Răspuns.* Matricea Playfair se obține trecând literele din parolă o singură dată în careul de  $5 \times 5$ , iar apoi celelalte litere ale alfabetului în ordinea lexicografică:

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

Mesajul este preprocesat prin introducerea literei *Q* ca delimitator de cuvânt, adjuncționându-se la finalul mesajului (pentru ca acesta să aibă lungime pară) litera *Q*:

*SIQINQCRIPTOGRAFIEQTACEREAQESTEQAURQ.*

Respectând regulile de cifrare Playfair mesajul cifrat devine:

*QPXAQSRIPTCEDGFETAUIOIGTOFU AUPAUEQIN.*

**Exercițiul 6.11.5.** Să se cifreze prin metoda transpoziției ( $N = 12$ ), pornind de la parola *CRIPTOGRAFIE* mesajul "SI IN CRIPTOGRAFIE TACEREA ESTE AUR".

*Răspuns.* Vom construi secvența numerică de cifrare asociind fiecărei litere din parolă indicele din ordinea lexicografică: astfel literele din parolă, scrise în ordine lexicografică sunt:

1	2	3	4	5	6	7	8	9	10	11	12
A	C	E	F	G	I	I	O	P	R	R	T

deci parola *CRIPTOGRAFIE* produce permutarea: 2 10 6 9 12 8 5 11 1 4 7 3.

Textul clar este scris într-o tabelă cu 12 coloane:

2	10	6	9	12	8	5	11	1	4	7	3
S	I	Q	I	N	Q	C	R	I	P	T	O
G	R	A	F	I	E	Q	T	A	C	E	R
E	A	Q	E	S	T	E	Q	A	U	R	Q

Deoarece lungimea textului nu este divizibilă cu 12 vom completa ultimul rând cu o secvență cunoscută (în acest caz caracterul *Q*). Textul cifrat se obține citind coloanele tabelului de cifrare în ordinea indicată de parola numerică:

*IAASGEORRQPCUCQEQAQTERQETIFEIRARTQNIS.*

Descifrarea se va realiza în mod similar folosind permutarea inversă  $\sigma^{-1}$ .

Dacă dimensiunea transpoziției  $N$  este mai mică decât lungimea parolei atunci se vor reține  $N$  caractere din parolă.

**Exercițiul 6.11.6.** Studiați comutativitatea operatorilor de cifrare substituție mono/polialfabetică și a transpoziției.

**Exercițiul 6.11.7.** Implementați algoritmul de decriptare a unei transpoziții.

**Exercițiul 6.11.8.** Implementați algoritmul de decriptare a unei substituții simple.

**Exercițiul 6.11.9.** Implementați algoritmul de decriptare a unui cifru obținut prin compunerea unei transpoziții și a unei substituții simple.

**Exercițiul 6.11.10.** Implementați algoritmul de decriptare a unui cifru polialfabetic.