Curs 3 + 4

Obs Cel mai ieftin algoritm Euclidian

$g = 1$

while ($a \bmod 2 = 0$ **and** $b \bmod 2 = 0$) **do**

$a = a/2$
$b = b/2$
$g = 2g$

end

while $a \neq 0$ **do**

while $a \bmod 2 = 0$ **do** $a = a/2$
while $b \bmod 2 = 0$ **do** $b = b/2$
(acum ambii sunt impari!)
if $a \geq b$ **then** $a = (a-b)/2$
else $b = (b-a)/2$

end

return $g \cdot b$;

Obs Lema chineză a resturilor, varianta efectivă

$m_1, \dots m_r$: $i \neq j \rightarrow gcd(m_i, m_j) = 1$

$x = ?$ a.î. $\forall i$ $x = a_i \bmod m_i$

Soluție
$$\left.\begin{aligned} M &= m_1 m_2 \dots m_r; \\ M_i &= M / m_i \\ y_i &= M_i^{-1} \bmod m_i \end{aligned}\right\} \quad x = \sum_{i=1}^{r} a_i M_i y_i \bmod M$$

Exemplu Caut x a.î.

$$\begin{cases} x = 5 \bmod 7 \\ x = 3 \bmod 11 \\ x = 10 \bmod 13 \end{cases}$$

$M = 7 \cdot 11 \cdot 13 = 1001$

$M_1 = 11 \cdot 13 = 143$, $y_1 = 143^{-1} \bmod 7 = 3^{-1} \bmod 7 = 5$

$M_2 = 7 \cdot 13 = 91$, $y_2 = 91^{-1} \bmod 11 = 3^{-1} \bmod 11 = 4$

$M_3 = 77$, $y_3 = 77^{-1} \bmod 13 = 12^{-1} \bmod 13 =$
$= (-1)^{-1} \bmod 13 = -1 \bmod 13 = 12$

$x = \sum_i a_i M_i y_i \bmod M = 5 \cdot 143 \cdot 5 + 3 \cdot 91 \cdot 4 + 10 \cdot 77 \cdot 12$

$\bmod 1001 = 894.$

$$\boxed{\begin{array}{c} \mathbb{Z}_{1001} = \mathbb{Z}_7 \times \mathbb{Z}_{11} \times \mathbb{Z}_{13} \\ 894 \longleftarrow (5, 3, 10) \end{array}}$$

Aplicație directă a aritmeticii modulare: Codurile liniare

$\mathcal{A}$ alfabet, $|\mathcal{A}| = n$, se identifică $\mathcal{A}$ cu $\mathbb{Z}_n$

$c : \mathcal{A}^k \to \mathcal{A}^k$ dată de $c(a_1 \dots a_k) = M \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$

unde $M \in \mathcal{M}_{k \times k}(\mathbb{Z}_n)$ inversabilă.

R inel, $M \in \mathcal{M}_{K\times K}(R)$

$$M^{-1} = \det(M)^{-1} \left(\cdots (-1)^{i+j} \det M_{i,j} \cdots \right)$$

Teoremă M inversabilă $\iff \det(M) \in R^{\times}$

Exemplu $\mathcal{A} = \{A, B, C \ldots\}$ $|\mathcal{A}| = 26$ $x_1 x_2 \rightsquigarrow y_1 y_2$

$$\begin{cases} y_1 = 6x_1 + 2x_2 \bmod 26 \\ y_2 = 5x_1 + 2x_2 \bmod 26 \end{cases}$$

Să se găsească cuvintele $x_1 x_2$ și $x_1' x_2'$ care au aceeași codificare $y_1 y_2$. Concluzie nu se poate folosi în criptografie.

$$\begin{cases} y_1 = 6x_1 + x_2 \bmod 26 \\ y_2 = 5x_1 + x_2 \bmod 26 \end{cases}$$

Să se afle regula de decriptare.

Securitate bazată pe teoria informației

P = mulțimea textelor clare (plaintext)

K = mulțimea cheilor

C = mulțimea textelor codificate.

$Enc : P \times K \to C$

$Dec : C \times K \to P$

$Dec_K(Enc_K(m)) = m$

(criptare simetrică)

Definiție (Enc, Dec) are securitate perfectă $\iff$

$\forall m \in P \; \forall c \in C \quad p(m \mid c) = p(m)$ $\quad p$ = probabilitate

Cu alte cuvinte, a îl vedea pe c nu oferă nici o informație despre m!

<u>Lemă</u> Dacă securitatea este perfectă

$$\# K \geq \# C \geq \# P$$

Dem. Enc_k injectivă $\forall k \in K$ fixat $\Rightarrow \# C \geq \# P$.

$\forall c \in C \quad p(c) > 0$ fiindcă altfel excludem c.

Deci $\forall m \in P \ \forall c \in C$

$$p(c|m) = p(c) > 0$$

Deci $\forall m \in P \ \forall c \in C \ \exists k \in K$

$$Enc_k(m) = c$$

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

probabilitatea condiționată.

Deci $\# K \geq \# C$.

<u>Teoremă</u> (Shannon) Dacă $\# P = \# C = \# K$.

securitate perfectă $\Leftrightarrow$
- orice cheie este folosită cu probabilitate egală $1/\# K$
- $\forall m \in P \ \forall c \in C \ \exists! \ k \in K$ $Enc_k(m) = c$

"$\Rightarrow$" $\forall m \in P \ \forall c \in C \ \exists k \ Enc_k(m) = c$

Dar $\# C = \# K \Rightarrow \# \{ Enc_k(m) \mid k \in K \} = \# K$

$\Rightarrow \forall m \in P \ \forall c \in C \ \exists! \ k \in K \ Enc_k(m) = c$

Vrem să arătam că fiecare cheie e folosită cu probabilitate egală, adică $p(K) = \frac{1}{\# K}$ pt orice $K \in K$.

Fie $\# K = n$, $P = \{ m_i \mid 1 \le i \le n \}$, fixăm $c \in C$.

Indexăm cheile $K_1, \dots K_n$ a.î. $Enc_{K_i}(m_i) = c$. $\forall i$.

Securitate perfectă $\Leftrightarrow p(m_i \mid c) = p(m_i)$, deci

$$p(m_i) = p(m_i \mid c) = \frac{p(c \mid m_i)\, p(m_i)}{p(c)} = \frac{p(K_i)\, p(m_i)}{p(c)}$$

Deci pt $\forall i$: $p(K_i) = p(c)$ deci toate cheile au probabilitate egală! și ea este $\frac{1}{n}$.

"$\Leftarrow$" Să arătăm că
$$\left.\begin{array}{l} \# K = \# P = \# C \\ p(K) = \frac{1}{\# K} \quad \forall K \in K \\ \forall m, c\ \exists!\, K\ Enc_K(m) = c \end{array}\right\} \Rightarrow$$

$$\underbrace{p(m \mid c) = p(m)}_{\shortparallel}$$

chei cu prob. egală

$$p(c) = \sum_{K \in K} p(K)\, p(m = Dec_K(c)) = \frac{1}{\# K} \sum_{K} p(m = Dec_K(c))$$

$\forall m \forall c\ \exists!\, K\ Enc_K(m) = c \Rightarrow$

$$\sum_{K \in K} p(m = Dec_K(c)) = \sum_{m} p(m) = 1$$

Deci $p(c) = \frac{1}{\# K}$. Dacă $c = Enc_K(m)$
$$p(c \mid m) = p(K) = \frac{1}{\# K}$$

Bayes:

$$p(m \mid c) = \frac{p(m)\, p(c \mid m)}{p(c)} = \frac{p(m) \cdot \frac{1}{\# K}}{\frac{1}{\# K}} = p(m)$$

qed.

Exemple

(1) $\mathcal{A} = \{A, B, \dots Z\}$

$$\# K = \# P = \# C = 26^n, \quad p(K) = \frac{1}{26^n}$$

$$c = m + K \mod 26 \text{ pe componente.}$$

(2) Vernam's Code (OTP)

$$\# K = \# P = \# C = 2^n, \quad p(K) = 2^{-n}$$

$c = m \oplus K$ unde $\oplus$ este $+$ în $\mathbb{F}_2 = \mathbb{Z}_2$.

Obs A nu se folosi aceeași cheie, fiindcă

$c_1 \oplus c_2 = m_1 \oplus K \oplus m_2 \oplus K = m_1 \oplus m_2$ pe care se poate face analiză de frecvență!

Definiție Entropie X variabilă random, ia valori $(x_i)_{1 \le i \le n}$ cu probabilități $p_i = p(x_i)$.

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i \quad \text{cu convenția } p_i = 0 \Rightarrow p_i \log p_i = 0$$

<u>Exemple</u>

Dacă eu răspund întotdeauna "da"

$$p_1 = 1, \quad p_2 = 0$$

$$H(X) = -1 \log_2 1 - 0 \log_2 0 = 0$$

adică nu îți ofer nici o informație

Dacă eu răspund la întâmplare "da" sau "nu"

$$p_1 = p_2 = \frac{1}{2}$$

$$H(X) = \left(-\log_2 \frac{1}{2} - \log_2 \frac{1}{2}\right) \frac{1}{2} = 1$$

adică îți ofer 1 bit de informație. la fiecare răspuns.

- $H(X) \geq 0$
- $H(X) = 0 \iff \exists!\, i \ p_i = 1 \wedge \forall j \neq i \ p_j = 0$
- $p_i = \frac{1}{n} \ \forall i \Rightarrow H(X) = \log_2 n$

Inegalitatea lui Jensen

$$\sum_{i=1}^{n} a_i = 1 \Rightarrow \sum_{i=1}^{n} a_i \log_2 x_i \leq \log_2 \left(\sum_{i=1}^{n} a_i x_i \right)$$

cu egalitate $\iff x_1 = x_2 = \dots = x_n$

<u>Teoremă</u> X ia n valori posibile

$$0 \leq H(X) \leq \log_2 n$$

deoarece $H(X) = -\sum_i p_i \log_2 p_i = \sum_i p_i \log_2 \frac{1}{p_i} \leq \log_2 \underbrace{\sum_i p_i \frac{1}{p_i}}_{= \log_2 n}$!

<u>Def</u> $H(X|y) = -\sum_x p(X=x \mid Y=y) \log_2 p(X=x \mid Y=y)$

entropie condițională $H(X|Y) = \sum_y p(Y=y) H(X|y)$

X, Y variabile random

$$r_{ij} = p(X = x_i \wedge Y = y_j)$$

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2 r_{ij}$$

entropia comună

$$H(X,Y) \leq H(X) + H(Y)$$

cu egalitate $\Leftrightarrow X, Y$ independente.

$$H(X,Y) = H(Y) + H(X|Y)$$

$$H(X|Y) \leq H(X)$$

cu egalitate $\Leftrightarrow X, Y$ independente.

<u>În criptografie</u>

$H(P|K,C) = 0$ dacă știm ~~cifrul~~ mes criptat și cheia, știm mesajul

$H(C|P,K) = 0$ dacă știm mesajul și cheia, știm mes criptat

$$H(K,P,C) = H(P,K) + \underbrace{H(C|P,K)}_{0}$$ deoarece $H(X,Y) = H(Y) + H(X|Y)$

$= H(P,K)$

$= H(K) + H(P)$ deoarece K și P sunt independente!

$$\Rightarrow \boxed{H(K,C) = H(K) + H(P)}$$

$H(K|C)$ = key equivocation
= amount of uncertainty about the key left after one cyphertext is revealed.

$$H(K|C) = H(K,C) - H(C) = H(K) + H(P) - H(C)$$

Exemplu

$P = \{a, b, c, d\}$, $K = \{K_1, K_2, K_3\}$, $C = \{1, 2, 3, 4\}$

$p(a) = 0,25$; $p(b) = p(d) = 0,3$; $p(c) = 0,15$

$p(K_1) = p(K_3) = 0,25$ $\quad p(K_2) = 0,5$

$p(1) = p(2) = p(3) = 0,2625$; $p(4) = 0,2125$

$H(P) = 1,9527$
$H(K) = 1,5$
$H(C) = 1,9944$

$$H(K|C) = 1,9527 + 1,5 - 1,9944 = 1,4583$$

Deci dacă vedem un mesaj cifrat, mai trebuie să găsim cam 1,5 bits de informație despre cheie. Asta este foarte puțin! => foarte nesigur! Cifrarea există într-adevăr:

	a	b	c	d
K_1	3	4	2	1
K_2	3	1	4	2
K_3	4	3	1	2

$p(1) = p(K_1)p(d) + p(K_2)p(b) + p(3)p(c) = 0,2625$ etc...

L = limbaj natural

H_L = entropia pe literă (informația pe literă...)

Random string are $H = \log_2 26 = 4,70$

Deci $H_L \leq 4,70$

$p(A) = 0,082$; ... ; $p(E) = 0,127$; ... $p(Z) = 0,001$

$H_L \leq H(p) \approx 4,14$ bits de informație pe literă în engleză.

În realitate Q întotdeauna urmat de U, TH foarte frecvent, etc... Mai bine considerăm grupuri de două litere.

"P^2" = variabila aleatoare a bigramelor.

$$H(P^2) = -\sum_{i,j} p(P=i, P'=j) \log(P=i, P'=j)$$

$H(P^2) \approx 7,12$

$H_L \leq H(P^2)/2 \approx 3,56$ etc...

<u>Definiție</u> Entropia limbajului natural L

$$H_L = \lim_{n\to\infty} \frac{H(P^n)}{n}$$

$$1.0 \leq H_L \leq 1.5$$

Deci o literă în engleză : folosește 5 bits dar conține 1,5 bits de informație pagină suplimentară

$p(a) = p(b) = p(c) = p(d) = \frac{1}{4}$

$p(K_1) = p(K_2) = p(K_3) = \frac{1}{3}$

$p(1) = 3 \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{4}$

la fel $p(2), p(3), p(4)$.

$H(P) = -4 \frac{1}{4} \log_2 (\frac{1}{4}) = 2$

$H(K) = -3 \frac{1}{3} \log_2 (\frac{1}{3}) = \log_2 3$

$H(C) = -4 \frac{1}{4} \log_2 (\frac{1}{4}) = 2$

$$H(K|C) = 2 + \log_2 3 - 2 = \log_2 3 =$$

$$= 1,5849 > 1,4583$$

Deci se știe mai puțin despre cheie ...

On** up** a t**e t**re **s a

rl **ll Sn** Wh**e.

Def Redundanța limbajului

$$R_L = 1 - \frac{H_L}{\log_2 \# P}$$

Dacă $H_L = 1,25$, redundanța în engleză

$$R_L = 1 - \frac{1.25}{\log_2 26} = 0,75.$$

Deci putem comprima texte în această limbă de la 10 MB la 2,5 MB.

Despre cheile false: $c \in C$, $|c| = n$ peste false în decriptare.

$K(c) = \{ k \in K \mid Ent_K(c)$ "are sens"$\}$

$\# K(c) - 1$ = numărul cheilor false.

Numărul "mediu" de chei false (piste false)

$$s_n = \sum_{c \in C}' p(c) (\# K(c) - 1) = \left(\sum_{c \in C}' p(c) \cdot \# K(c) \right) - 1$$

n mare

$\# P = \# K \Rightarrow \log_2 (s_n + 1) = \log_2 \sum_{c \in C}' p(c) \cdot \# K(c)$

$\geq \sum_{c}' p(c) \log_2 (\# K(c))$ Jensen $\geq$

$$\geq \sum_{c \in C} p(c)\, H(K|c) = H(K|C)$$

$$= H(K) + H(P) - H(C) \approx H(K) + n H_L - H(C)$$

n foarte mare

$$= H(K) - H(C) + n(1 - R_L) \log_2 \# P \quad \text{(din def. redundanței)}$$

$$\geq H(K) - n \log_2 \# C + n(1-R_L) \log_2 \# P$$

(deoarece $H(C) \leq n \log_2 \# C$)

$$= H(K) - n R_L \log_2 \# P \quad \text{deoarece } \# P = \# C.$$

$$s_n \geq \frac{\# K}{(\# P)^{n R_L}} - 1$$

<u>Def</u> Unicity distance n_0 = the value of n such that the expected number of "spurious keys" becomes 0

$$n_0 \approx \frac{\log_2 \# K}{R_L \log_2 \# P}$$

Substitutie

$\# P = 26$

$\# K = 26! \approx 4 \cdot 10^{26}$

$R_L = 0{,}75 \quad \Rightarrow n_0 \approx \dfrac{88{,}4}{0{,}75 \cdot 4{,}7} \approx 25$

Deci pentru $|C| \geq 25$ se presupune că există o unică descifrare cu sens!

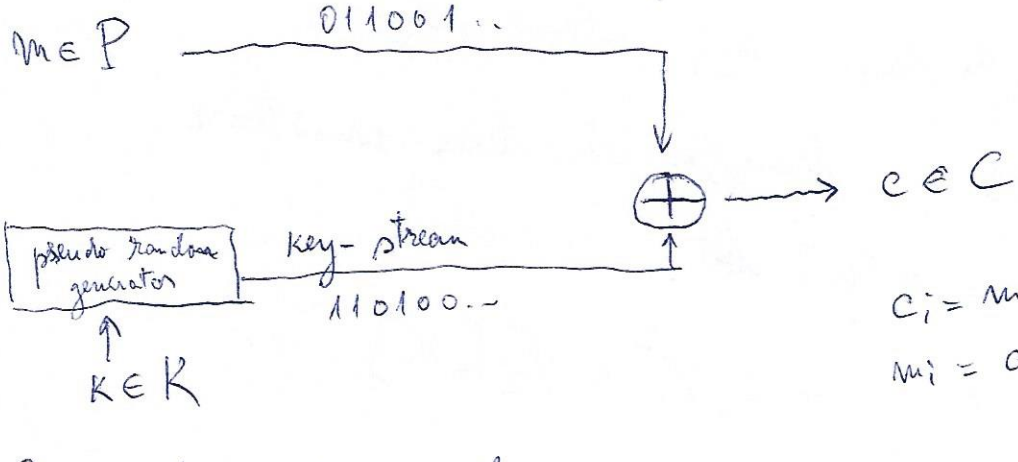
Bit strings + keys of length l

$\# P = 2$

$\# K = 2^{l}$ $\qquad n_0 \approx \frac{l}{0,75} = \frac{4l}{3}$

$R_L = 0,75$

Dacă comprimăm datele înainte de a le transmite $\qquad n_0 = \frac{l}{0} = \infty$ deci atacatorul va avea o muncă mult mai dificilă!



$c_i = m_i + K_i$

$m_i = c_i + K_i$

Condiții pt pseudo-random

- Perioadă lungă $\quad K_i = K_{i+N}$, N foarte mare $\quad$ (N există fiindcă este determinist)
- proprietăți pseudo-random
- mare complexitate liniară!

Linear feedback stream registers

L = lungimea registrului

$c_1 \dots c_L$ = bits

Starea inițială $[s_{L-1}, \dots s_1, s_0]$

Output sequence: $s_0, s_1, \dots s_{L-1}, s_L, s_{L+1}, \dots$

unde $s_j = c_1 \cdot s_{j-1} \oplus c_2 \cdot s_{j-2} \oplus \dots c_L \cdot s_{j-L} \quad \forall j \geq L$

$s_{i+N} = s_i$; $N \leq 2^L - 1$

N = perioada

$$M = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ & & & & 1 \\ c_L & c_{L-1} & c_{L-2} & \dots & c_1 \end{pmatrix}$$

$v = (1, 0, \dots 0)$

$s = (s_1, s_2, \dots s_L)$ internal state

$s = M \cdot s$ tranziția la starea următoare

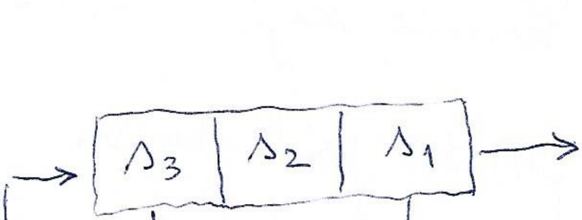
$v \cdot s$ = output bit.

$C(X) = 1 + c_1 X + \dots + c_L X^L \in \mathbb{F}_2[X]$

polinomul de conexiune.

$C(X) = \det(XM - I_{L \times L})$

Examples $\quad X^3 + X + 1$



Definiție $C(X)$ primitiv $\iff$ $C(X)$ ireductibil,

$$\mathbb{F}_2[X]/(C(X)) = \mathbb{F}_{2^L}$$

și $\mathbb{F}_{2^L}^{\times} = \langle \theta \rangle$ unde $C(\theta) = 0$.

- $c_L = 0$ (singular) $\Rightarrow$ șirul devine periodic mai târziu
- $c_L = 1$ și C ireductibil $\Rightarrow$ șir periodic, perioada $N \mid 2^L - 1$ (cea mai mică valoare a.î. $C(X) \mid 1 + X^N$)
- $c_L = 1$ și C primitiv $\Rightarrow$ $N = 2^L - 1$

Exemplu $L = 4$, $C(X) = X^3 + X + 1$ singular

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\begin{tikzcd}
s_{12} \arrow[d] & s_2 \arrow[d] & s_5 \arrow[d] & s_6 \arrow[d] & & \\
s_9 \arrow[dr] & s_4 \arrow[l] & s_{10} \arrow[l] & s_{13} \arrow[l] & & \\
& s_3 \arrow[r] & s_7 \arrow[r] & s_{14} \arrow[u] & s_8 \arrow[r] & s_0 \\
& s_1 \arrow[u] & s_{11} \arrow[u] & s_{15} \arrow[u] & &
\end{tikzcd}$$

$$C(X) = X^4 + X^3 + X^2 + 1 = (X+1)(X^3 + X + 1)$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\begin{tikzcd}
s_8 \arrow[dr] & s_{12} \arrow[l] & s_6 \arrow[l] & s_{11} \arrow[l] \\
& s_1 \arrow[r] & s_2 \arrow[r] & s_5 \arrow[u]
\end{tikzcd}$$

$$\begin{tikzcd}
s_4 \arrow[dr] & s_{10} \arrow[l] & s_{13} \arrow[l] & s_{14} \arrow[l] \\
& s_9 \arrow[r] & s_3 \arrow[r] & s_7 \arrow[u]
\end{tikzcd}$$

$$s_{15} \qquad s_0$$

$C(X) = X^4 + X + 1$ irreducible and primitive

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Cycle of length 15!