Securitatea sistemelor informatice

Sisteme istorice de criptare. Securitate perfecta.

Curs 2

Anul III, Informatica
2022-2023

Adela Georgescu
Facultatea de Matematica – Informatica
Universitatea Bucuresti

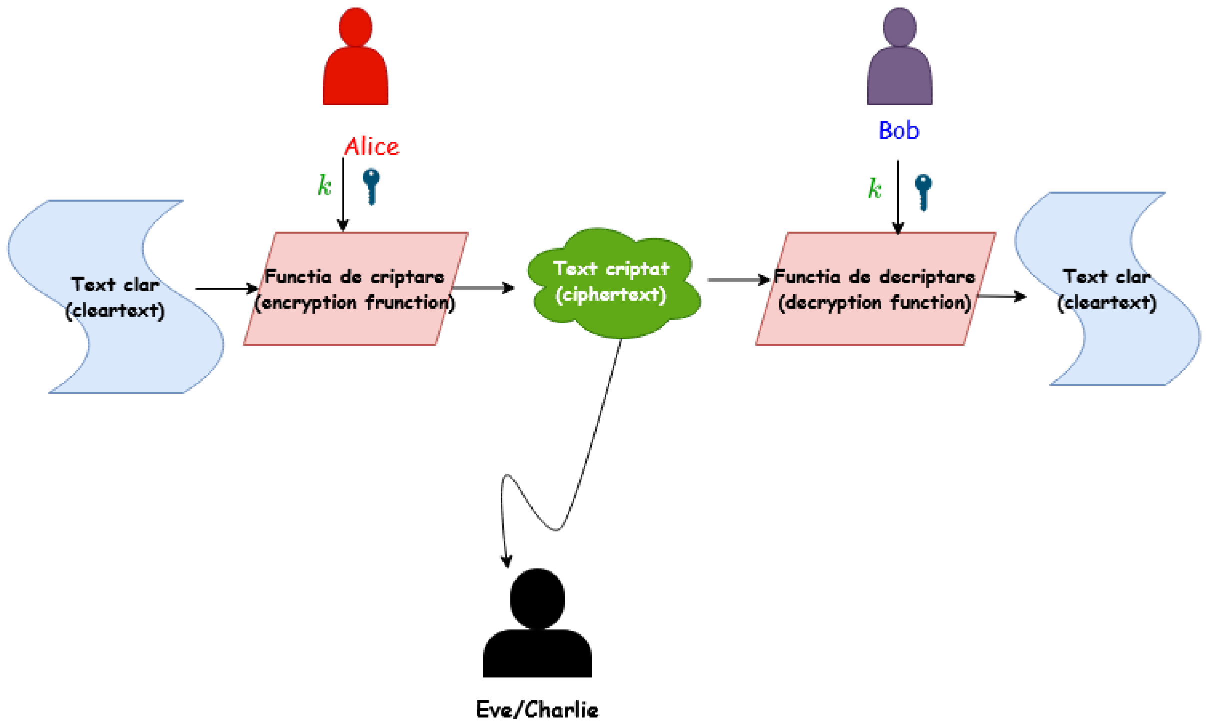
Bibliografie

- Introduction to Modern Cryptography, 3rd edition. Johnathan Katz, Yehuda Lindell

- Encyclopedia of Cryptography and Security
  https://www.researchgate.net/publication/230674947_Springer_Encyclopedia_of_Cryptography_and_Security

- https://www.coursera.org/learn/crypto

Criptarea simetrica (cu cheie secreta)



Criptarea simetrica (cu cheie secreta)

Definitie

Un sistem de criptare simetric definit peste $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, *cu:*

- $\mathcal{K}$ = *spațiul cheilor*
- $\mathcal{M}$ = *spațiul textelor clare (mesaje)*
- $\mathcal{C}$ = *spațiul textelor criptate*

este un triplet $(\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$, *unde:*

1. Gen: *este algoritmul probabilistic de generare a cheilor care întoarce o cheie* k *conform unei distribuții*
2. $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$
3. $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

a.î. $\forall m \in \mathcal{M}, k \in \mathcal{K} : Dec_k(Enc_k(m)) = m$.

Criptarea simetrica (cu cheie secreta)

- $K =$ variabilă aleatoare ce reprezintă valoarea cheii returnată de Gen
- $Pr[K = k] =$ probabilitatea cheii generate de Gen de a lua valoarea k, $\forall k \in \mathcal{K}$
- $M =$ variabilă aleatoare ce reprezintă mesajul care se criptează
- $Pr[M = m] =$ probabilitatea ca mesajul de criptat să ia valoarea $m \in \mathcal{M}$

Criptarea simetrica (cu cheie secreta)

- K = variabilă aleatoare ce reprezintă valoarea cheii returnată de Gen
- $Pr[K = k]$ = probabilitatea cheii generate de Gen de a lua valoarea k, $\forall k \in \mathcal{K}$
- M = variabilă aleatoare ce reprezintă mesajul care se criptează
- $Pr[M = m]$ = probabilitatea ca mesajul de criptat să ia valoarea $m \in \mathcal{M}$
- expl. de probabilitate de distributie peste $\mathcal{M}$: $\mathcal{M} = \{\mathrm{atacati\ azi}, \mathrm{nu\ atacati}\}$ cu $Pr[M = \mathrm{atacati\ azi}] = 0.2$ si $Pr[M = \mathrm{nu\ atacati}] = 0.8$

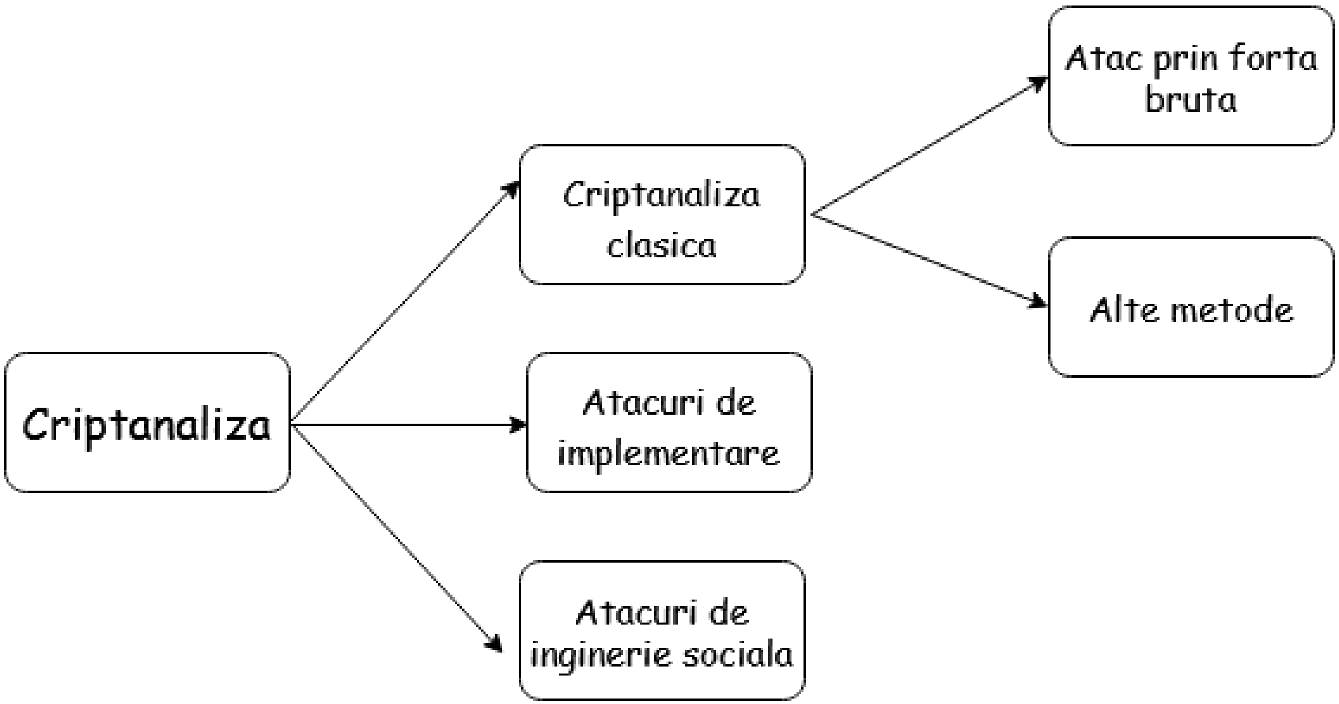
Terminologie

- Mesajul în forma originară se numește text clar;
- Expeditorul rescrie mesajul folosind un sistem de criptare, adică îl criptează și obține un text criptat;
- Destinatarul îl decriptează cunoscând metoda folosită pentru criptare;
- Procesul de determinare a cheii aferente unui sistem de criptare, cunoscând doar textul criptat (eventual și alte informații auxiliare) se numește criptanaliză;
- Decriptarea și criptanaliza au același scop: găsirea textului clar; diferența constă în faptul că la criptanaliză nu se cunoaște cheia de decriptare.

Scenarii de atac

- Atac cu text criptat (ciphertext-only attack): Atacatorul știe doar *textul criptat* - poate încerca un atac prin forță brută prin care se parcurg toate cheile până se găsește cea corectă;
- Atac cu text clar (known-plaintext attack): Atacatorul cunoaște una sau mai multe perechi *(text clar, text criptat)*;
- Atac cu text clar ales (chosen-plaintext attack): Atacatorul poate obține criptarea unor texte clare alese de el;
- Atac cu text criptat ales (chosen-ciphertext attack): Atacatorul are posibilitatea să obțină decriptarea unor texte criptate alese de el.

Criptanaliza



Sisteme de criptare istorice

Cifruri de permutare / transpoziție

Definitie

Un cifru de permutare presupune rearanjarea literelor în textul clar pentru a obține textul criptat.

Cifruri de permutare / transpoziție

- sistemul Rail Fence >>> curs
- cifruri generale de transpoziție >>> laborator

Rail Fence

M	A	R	T
E	J	I	A
S	C	P	T

Text clar: mesaj criptat
Cheia: $k = 3$
Text criptat: MARTEJIASCPT

Cifruri de substitiuţie monoalfabetice

- cifrul lui Cezar
- substituţie simplă
- sistemul Cavalerilor de Malta

Cifrul lui Cezar

a	b	c	d	e	f	g	h	i	j	k	l	m
D	E	F	G	H	I	J	K	L	M	N	O	P

n	o	p	q	r	s	t	u	v	w	x	y	z
Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Text clar: mesaj criptat
Text criptat: PHVDM FULSWDW

Cifrul lui Cezar

- $\mathcal{K} = \{0, 1, \ldots, 25\}$
- $\mathcal{M} = \{a, b, \ldots, z\}^*$
- $\mathcal{C} = \{A, B, \ldots, Z\}^*$
- $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$

- $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

Cifrul lui Cezar

- $\mathcal{K} = \{0, 1, \ldots, 25\}$
- $\mathcal{M} = \{a, b, \ldots, z\}^*$
- $\mathcal{C} = \{A, B, \ldots, Z\}^*$
- $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$

$$Enc_k(m) = m + k \pmod{26}$$

- $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

Generalizare - Shift cipher

- $\mathcal{K} = \{0, 1, \ldots, 25\}$
- $\mathcal{M} = \{a, b, \ldots, z\}^*$
- $\mathcal{C} = \{A, B, \ldots, Z\}^*$
- $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$

$$Enc_k(m) = m + k \pmod{26}$$

- $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

$$Dec_k(c) = c - k \pmod{26}$$

Criptanaliză - Atac prin forță brută

- $|\mathcal{K}| = 26$
- atac prin forță brută (căutare exhaustivă): încercarea, pe rând, a tuturor cheilor posibile până când se obține un text clar cu sens

Principiul cheilor suficiente: O schemă sigură de criptare trebuie să aibă un spațiu al cheilor suficient de mare a.î. să nu fie vulnerabilă la căutarea exhaustivă.

Substituția simplă

a	b	c	d	e	f	g	h	i	j	k	l	m
F	I	L	O	R	U	X	A	D	G	J	M	P

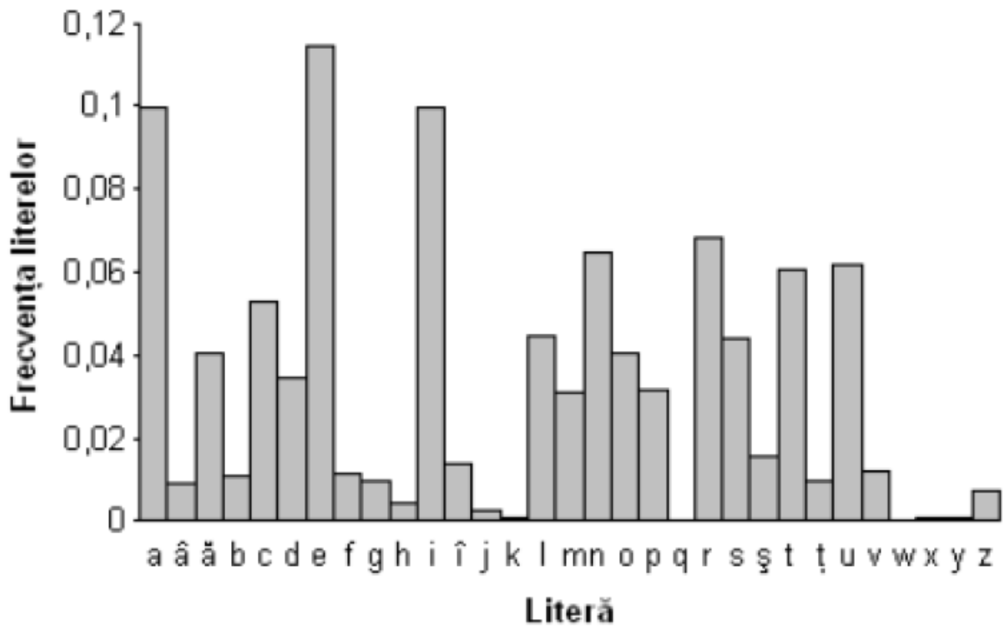
n	o	p	q	r	s	t	u	v	w	x	y	z
S	V	Y	B	E	H	K	N	Q	T	W	Z	C

Text clar: mesaj criptat
Text criptat: PRHFG LEDYKFK

Criptanaliză - Analiza de frecvență

- $|\mathcal{K}| = 26!$
- atacul prin forță brută devine mai dificil
- analiza de frecvență: determinare corespondenței între alfabetul clar și alfabetul criptat pe baza frecvenței de apariție a literelor în text, cunoscând distribuția literelor în limba textului clar
  - se cunoaște limba textului clar
  - lungimea textului permite analiza de frecvență

Criptanaliză - Analiza de frecvență



[Wikipedia]

Cifruri de substitiuţie polialfabetice / poligrafice

- sistemul Playfair
- sistemul Hill
- sistemul Vigenére

Cifrul Vigenére

Text clar:	c	u	r	s	c	r	i	p	t	o	g	r	a	f
Cheie:	c	h	e	i	e	c	h	e	i	e	c	h	e	i
Text criptat:	E	B	V	A	G	T	P	T	B	S	I	Y	E	N

Cifrul Vigenére

- $\mathcal{K}$
- $\mathcal{M}$
- $\mathcal{C}$
- $\text{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$
- $\text{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

Cifrul Vigenére

- $\mathcal{K} = \{0, 1, \ldots, 25\}$
- $\mathcal{M} = \{0, 1, \ldots, 25\}^*$
- $\mathcal{C} = \{0, 1, \ldots, 25\}^*$
- $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$
  Textul clar: $m_0 m_1 \ldots m_{n-1}$
  Cheia: $k_0 k_1 \ldots k_x$
- $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

Cifrul Vigenére

- $\mathcal{K} = \{0, 1, \ldots, 25\}$
- $\mathcal{M} = \{0, 1, \ldots, 25\}^*$
- $\mathcal{C} = \{0, 1, \ldots, 25\}^*$
- $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$
  Textul clar: $m_0 m_1 \ldots m_{n-1}$
  Cheia: $k_0 k_1 \ldots k_x$

  $$Enc_{k_j}(m_i) = m_i + k_j \pmod{26}$$

  unde $j = i \bmod x$
- $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

  $$Dec_{k_j}(c_i) = c_i - k_j \pmod{26}$$

Vigenére – criptanaliza

- Pentru o cheie de **n** caractere, spatiul cheilor $|\mathcal{K}| = 26^n$
- Un text criptat c = $c_1 c_2 \ldots$ poate fi impartit in **s** parti in care fiecare parte a fost criptata cu aceeasi litera din alfabet.

  pentru j ∈ $\{1,2,..,s\}$

$$c_j = m_j + k_j$$
$$c_{j+s} = m_{j+s} + k_j$$
$$c_{j+2s} = m_{j+2s} + k_j$$

- Daca se cunoaste lungimea *p* a cheii, problema se reduce la criptanaliza a *p* texte criptate cu shift cipher
- Putem face analiza de frecventa pe fiecare sir separat

Sisteme mecanice de criptare

Masina Enigma

Sisteme mecanice

Criptografia moderna

- Se bazeaza pe 3 principii moderne

Principiul 1 – orice problema criptografica necesita o definitie clasica si riguroasa – *discutat in Curs 1*

(scheme construite dupa acest principiu sunt folosite azi in TLS, SSH, IPSec)

Principiul 2 - securitatea primitivelor criptografice se bazeaza pe prezumptii clare de securitate

Principiul 3 – orice constructie criptografica trebuie sa fie insotita de o demonstratie de securitate conform principiilor anterioare

Principiul 2 – prezumptii (ipoteze) de securitate

- majoritatea constructiilor criptografice moderne **nu pot fi** demonstrate ca fiind **sigure neconditionat**

- ipotezele de securitate trebuie sa fie explicite
  - permit cercetatorilor sa valideze aceste ipoteze
  - permit comparatia intre doua scheme bazate pe ipoteze diferite de securitate
  - implicatii practice in cazul unor erori aparute in cadrul ipotezei de securitate
  - necesare pentru demonstratiile de securitate

Exemplu de ipoteza de securitate (problema dificila)

- Factorizarea numerelor mari
  - Se da un numar compus N si se cere descompunerea lui in factori primi.
  - Expl: 85 = 17 * 5
  - Astazi nu se cunoaste nici un algoritm care sa factorizeze un numar de 400 cifre intr-un timp practic
- Totusi:
  1. Un algoritm mai rapid ar putea exista
  2. Un calculator cuantic factorizeaza rapid (inca nu a fost construit dar se fac eforturi in acest sens)
  3. Criptografia "post-cuantica" este in plina dezvoltare - competitia de standardizare post-cuantica NIST, criptografia bazata pe latici etc.

Principiul 3 - demonstratii de securitate

- ofera o demonstratie riguroasa a faptului ca o constructie satisface definitia data in ipoteze de securitate clare

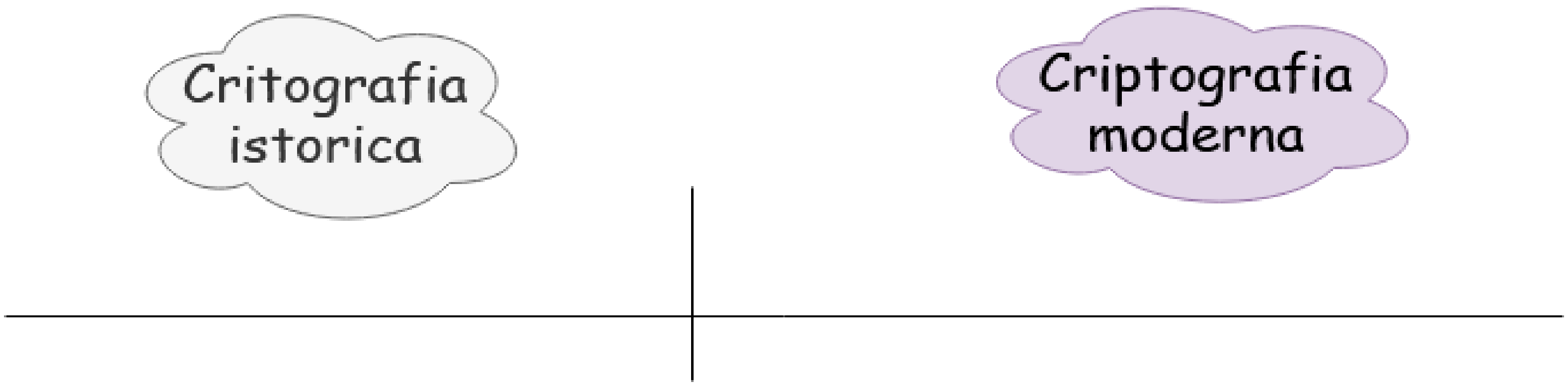
- fara o demonstratie riguroasa, intuitia ca o schema este corecta poate avea consecinte dezastruoase

- majoritatea demonstratiilor folosesc o abordare reductionista

Teorema *Constructia Y este sigura conform definitiei daca prezumptia X este adevarata.*

- demonstratia va arata cum un adversar care sparge schema Y poate incalca prezumptia X.

Retineti



- Abordare ad-hoc

1980

- Abordare riguroasa

Securitate perfecta (neconditionata)

- Sistemele de criptare istorice (substitutie, permutare, Vigenere, Playfair etc.) pot fi sparte cu un efort computational foarte mic

- In cursul de azi - scheme perfect sigure care rezista in fata unui adversar cu putere computationala nelimitata

- Insa ... limitarile sunt inevitabile

Generare aleatorism

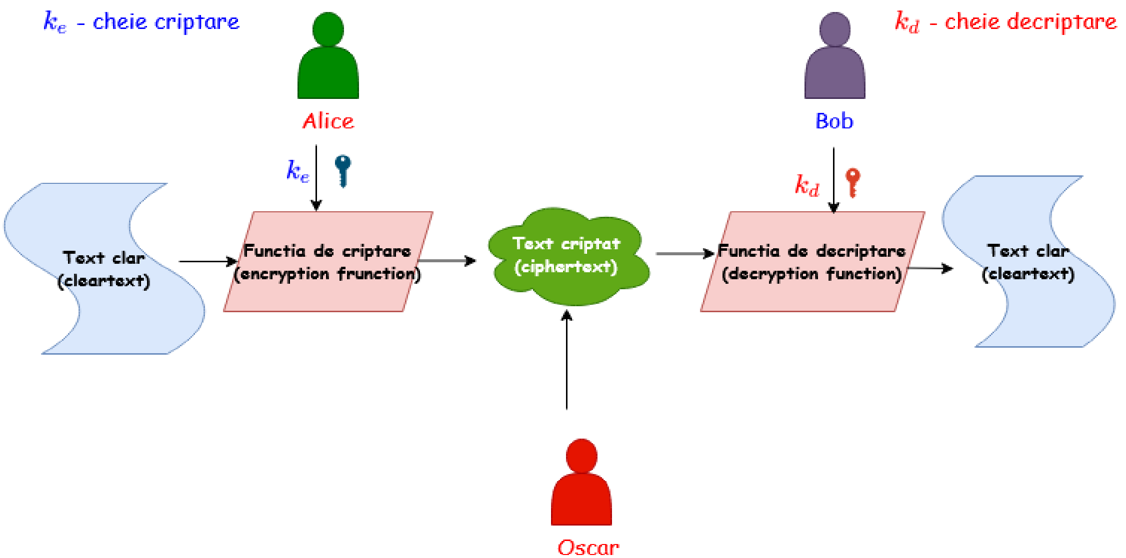
- In exemplul urmator cheile sunt generate aleator, deci avem nevoie de o sursa buna de aleatorism

- Trebuiesc folosite generatoare de numere aleatoare create in scop criptografic iar nu unele generale care nu sunt destinate aplicatiilor criptografice

- Expl.: functia **rand()** din libraria stdlib.h din C nu este sigura din punct de vedere criptografic

- Sunt necesare doua proprietati pentru criptografie:
  1. **Generala** - poseda proprietati statistice bune (nu poate fi reprodus)
  2. **Specifica** – este impredictibil: fiind dati n biti de iesire, este imposibil de calculat (infezabil computational) care sunt urmatorii biti

Securitate perfecta (Shannon 1949)



Ipoteza: Oscar cunoaste distributia peste M

Securitate perfecta:

1. daca Oscar afla textul criptat nu are nici un fel de informatie in plus decat daca nu l-ar fi aflat (nu schimba ceea ce stie el despre distributia peste M).

(textul criptat nu ofera nici un fel de informatie despre textul clar)

2. Oscar nu poate ghici care din doua posibile mesaje clare a fost criptat, doar vazand textul criptat

Securitate perfecta (Shannon 1949)

Definiție

O schemă de criptare peste un spațiu al mesajelor $\mathcal{M}$ este perfect sigură dacă pentru orice probabilitate de distribuție peste $\mathcal{M}$, pentru orice mesaj $m \in \mathcal{M}$ și orice text criptat c pentru care $Pr[C = c] > 0$, următoarea egalitate este îndeplinită:

$$Pr[M = m|C = c] = Pr[M = m]$$

- $Pr[M = m]$ - probabilitatea *a priori* ca Alice să aleagă mesajul m;
- $Pr[M = m|C = c]$ - probabilitatea *a posteriori* ca Alice să aleagă mesajul m, chiar dacă textul criptat c a fost văzut ;
- securitate perfectă - dacă Oscar afla textul criptat nu are nici un fel de informație în plus decât dacă nu l-ar fi aflat.

Securitate perfecta (Shannon 1949)

Definitie

O schemă de criptare (Enc, Dec) *este perfect sigură dacă pentru orice mesaje* $m_0, m_1 \in \mathcal{M}$ *cu* $|m_0| = |m_1|$ *și* $\forall c \in \mathcal{C}$ *următoarea egalitate este îndeplinită:*

$$Pr[Enc_k(m_0) = c] = Pr[Enc_k(m_1) = c]$$

unde $k \in \mathcal{K}$ *este o cheie aleasă uniform.*

- fiind dat un text criptat, este imposibil de ghicit dacă textul clar este m_0 sau m_1
- cel mai puternic adversar nu poate deduce nimic despre textul clar dat fiind textul criptat

One Time Pad (OTP)

- Patentat in 1917 de Vernam (mai poartă denumirea de Cifrul Vernam)
- Algoritmul:
  1. Fie $l > 0$ iar $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^l$
  2. Cheia k se alege cu distribuţie uniformă din spaţiul cheilor $\mathcal{K}$
  3. **Enc**: dată o cheie $k \in \{0,1\}^l$ şi un mesaj $m \in \{0,1\}^l$, întoarce $c = k \oplus m$.
  4. **Dec**: dată o cheie $k \in \{0,1\}^l$ şi un mesaj criptat $c \in \{0,1\}^l$, întoarce $m = k \oplus c$.

One Time Pad (OTP)

mesaj clar:	0	1	1	0	0	1	1	1	1	$\oplus$
cheie:	1	0	1	1	0	0	1	1	0	
text criptat:	1	1	0	1	0	1	0	0	1	

- avantaj - criptare și decriptare rapide
- dezavantaj - cheia foarte lungă (la fel de lungă precum textul clar)
- Este OTP sigur?

One Time Pad (OTP)

mesaj clar:	0	1	1	0	0	1	1	1	1	$\oplus$
cheie:	1	0	1	1	0	0	1	1	0	
text criptat:	1	1	0	1	0	1	0	0	1	

mesaj clar:	1	1	0	0	0	0	1	1	0	$\oplus$
cheie:	0	0	0	1	0	1	1	1	1	
text criptat:	1	1	0	1	0	1	0	0	1	

- Același text criptat poate să provină din orice text clar cu o cheie potrivită
- Dacă adversarul nu știe decât textul criptat, atunci nu știe nimic despre textul clar!

Securitatea perfecta

- Securitatea perfecta nu este imposibila dar..
  - cheia trebuie sa fie la fel de lunga precum mesajul
  - inconveniente practice (stocare, transmisie)
  - cheia trebuie sa fie folosita o singura data – one time pad
- Exercitiu: *Ce* se intampla daca folosim o aceeasi cheie de doua ori cu sistemul OTP ?

OTP

- Daca un adversar obtine

$$C = M \oplus K \text{ si } C' = M' \oplus K$$

- atunci el poate calcula

$$C \oplus C' = (M \oplus K) \oplus (M' \oplus K) = (M \oplus M')$$

ceea ce invalideaza proprietatea de securitate perfecta

Limitarile securitatii perfecte

Teoremă

Fie o schemă (Enc, Dec) de criptare perfect sigură peste un spatiu al mesajelor $\mathcal{M}$ și un spațiu al cheilor $\mathcal{K}$. Atunci $|\mathcal{K}| \geq |\mathcal{M}|$.

Sau altfel spus

Limitarile securitatii perfecte

Teoremă

Fie o schemă (Enc, Dec) *de criptare perfect sigură peste un spatiu al mesajelor* $\mathcal{M}$ *și un spațiu al cheilor* $\mathcal{K}$. *Atunci* $|\mathcal{K}| \geq |\mathcal{M}|$.

Sau altfel spus

Teoremă

Nu există nici o schemă de criptare (Enc, Dec) *perfect sigură în care mesajele au lungimea* n *biți iar cheile au lungimea (cel mult)* $n-1$ *biți.*