Securitatea Sistemelor Inform

- Curs 8.5 -

Prezumpții criptografice dificile

Adela Georgescu

Facultatea de Matematică și Informatică Universitatea din București Anul universitar 2022-2023, semestrul I

Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;

- Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;
- Până acum am văzut că schemele de criptare și de autentificare se bazează pe prezumpția existenței permutărilor pseudoaleatoare;

- Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;
- Până acum am văzut că schemele de criptare și de autentificare se bazează pe prezumpția existenței permutărilor pseudoaleatoare;
- Dar această prezumpție e nenaturală și foarte puternică;

- Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;
- Până acum am văzut că schemele de criptare şi de autentificare se bazează pe prezumpţia existenţei permutărilor pseudoaleatoare;
- Dar această prezumpție e nenaturală și foarte puternică;
- În practică, PRF pot fi instanțiate cu cifruri bloc;

- Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;
- Până acum am văzut că schemele de criptare și de autentificare se bazează pe prezumpția existenței permutărilor pseudoaleatoare;
- Dar această prezumpție e nenaturală și foarte puternică;
- În practică, PRF pot fi instanțiate cu cifruri bloc;
- Însă metode pentru a demonstra pseudoaleatorismul construcțiilor practice relativ la alte prezumpții "mai rezonabile" nu se cunosc;

- Criptografia modernă se bazează pe prezumpţia că anumite probleme nu pot fi rezolvate în timp polinomial;
- Până acum am văzut că schemele de criptare şi de autentificare se bazează pe prezumpţia existenţei permutărilor pseudoaleatoare;
- Dar această prezumpție e nenaturală și foarte puternică;
- În practică, PRF pot fi instanțiate cu cifruri bloc;
- Însă metode pentru a demonstra pseudoaleatorismul construcțiilor practice relativ la alte prezumpții "mai rezonabile" nu se cunosc;
- Dar e posibil a demonstra existenţa permutărilor pseudoaleatoare pe baza unei prezumpţii mult mai slabe, cea a existenţei funcţiilor one-way;

▶ În continuare vom introduce câteva două considerate "dificile": problema factorizării şi problema logaritmului discret şi vom prezenta funcţii conjecturate ca fiind one-way bazate pe aceste probleme;

- În continuare vom introduce câteva două considerate "dificile": problema factorizării şi problema logaritmului discret şi vom prezenta funcții conjecturate ca fiind one-way bazate pe aceste probleme;
- ► Tot materialul ce urmează se bazează pe noțiuni de teoria numerelor;

- ▶ În continuare vom introduce câteva două considerate "dificile": problema factorizării şi problema logaritmului discret şi vom prezenta funcţii conjecturate ca fiind one-way bazate pe aceste probleme;
- Tot materialul ce urmează se bazează pe noțiuni de teoria numerelor;
- ▶ La criptografia simetrică (cu cheie secretă) am văzut primitve criptografice (i.e. funcții hash, PRG, PRF) care pot fi construite eficient fără a implica teoria numerelor;

- În continuare vom introduce câteva două considerate "dificile": problema factorizării şi problema logaritmului discret şi vom prezenta funcții conjecturate ca fiind one-way bazate pe aceste probleme;
- Tot materialul ce urmează se bazează pe noțiuni de teoria numerelor;
- ► La criptografia simetrică (cu cheie secretă) am văzut primitve criptografice (i.e. funcții hash, PRG, PRF) care pot fi construite eficient fără a implica teoria numerelor;
- ► La criptografia asimetrică (cu cheie publică) construcțiile cunoscute se bazează pe probleme matematice dificile din teoria numerelor;

1. Problema factorizării

 O primă problemă conjecturată ca fiind dificilă este problema factorizării numerelor întregi sau mai simplu problema factorizării;

1. Problema factorizării

- O primă problemă conjecturată ca fiind dificilă este problema factorizării numerelor întregi sau mai simplu problema factorizării;
- Fiind dat un număr compus N, problema cere să se găsească două numere prime x_1 și x_2 pe n biți așa încât $N = x_1 \cdot x_2$;

1. Problema factorizării

- O primă problemă conjecturată ca fiind dificilă este problema factorizării numerelor întregi sau mai simplu problema factorizării;
- Fiind dat un număr compus N, problema cere să se găsească două numere prime x_1 și x_2 pe n biți așa încât $N = x_1 \cdot x_2$;
- Cele mai greu de factorizat sunt numerele care au factori primi foarte mari.

Primalitate și factorizare

"The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length... The dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated."

(C.F.Gauss 1777 – 1855)

► Pentru a putea folosi problema în criptografie, trebuie să generăm numere prime aleatoare *în mod eficient*;

- ► Pentru a putea folosi problema în criptografie, trebuie să generăm numere prime aleatoare *în mod eficient*;
- Putem genera un număr prim aleator pe n biți prin alegerea repetată de numere aleatoare pe n biți până când găsim unul prim;

- ► Pentru a putea folosi problema în criptografie, trebuie să generăm numere prime aleatoare *în mod eficient*;
- Putem genera un număr prim aleator pe n biţi prin alegerea repetată de numere aleatoare pe n biţi până când găsim unul prim;
- Pentru eficiență, ne interesează două aspecte:

- ► Pentru a putea folosi problema în criptografie, trebuie să generăm numere prime aleatoare *în mod eficient*;
- Putem genera un număr prim aleator pe n biţi prin alegerea repetată de numere aleatoare pe n biţi până când găsim unul prim;
- Pentru eficiență, ne interesează două aspecte:
 - 1. probabilitatea ca un număr aleator de n biți să fie prim;

- ► Pentru a putea folosi problema în criptografie, trebuie să generăm numere prime aleatoare *în mod eficient*;
- Putem genera un număr prim aleator pe n biţi prin alegerea repetată de numere aleatoare pe n biţi până când găsim unul prim;
- Pentru eficiență, ne interesează două aspecte:
 - 1. probabilitatea ca un număr aleator de n biți să fie prim;
 - 2. cum testăm eficient că un număr dat *p* este prim.

Pentru distribuţia numerelor prime, se cunoaște următorul rezultat matematic:

Teoremă

Pentru orice n > 1, proportia de numere prime pe n biți este de cel puțin 1/3n.

Pentru distribuţia numerelor prime, se cunoaște următorul rezultat matematic:

Teoremă

Pentru orice n > 1, proportia de numere prime pe n biţi este de cel puţin 1/3n.

▶ Rezultă imediat că dacă testăm $t = 3n^2$ numere, probabilitatea ca un număr prim să nu fie ales este cel mult e^{-n} , deci neglijabilă.

Pentru distribuția numerelor prime, se cunoaște următorul rezultat matematic:

Teoremă

Pentru orice n > 1, proportia de numere prime pe n biți este de cel puțin 1/3n.

- ▶ Rezultă imediat că dacă testăm $t = 3n^2$ numere, probabilitatea ca un număr prim să nu fie ales este cel mult e^{-n} , deci neglijabilă.
- ► Deci avem un algoritm în timp polinomial pentru gasirea numerelor prime

► Cei mai eficienți algoritmi sunt probabiliști:

- Cei mai eficienți algoritmi sunt probabiliști:
 - ▶ Dacă numărul p dat este prim atunci algoritmul întotdeauna întoarce rezultatul prim;

- Cei mai eficienți algoritmi sunt probabiliști:
 - ▶ Dacă numărul p dat este prim atunci algoritmul întotdeauna întoarce rezultatul prim;
 - ▶ Dacă p este compus, atunci cu probabilitate mare algoritmul va întoarce compus;

- Cei mai eficienți algoritmi sunt probabiliști:
 - ▶ Dacă numărul p dat este prim atunci algoritmul întotdeauna întoarce rezultatul prim;
 - ▶ Dacă p este compus, atunci cu probabilitate mare algoritmul va întoarce compus;
 - ► Concluzie: dacă outputul este *compus*, atunci *p* sigur este compus, dacă outputul este *prim*, atunci cu probabilitate mare *p* este prim dar este posibil și să se fi produs o eroare;

- Cei mai eficienți algoritmi sunt probabiliști:
 - ▶ Dacă numărul p dat este prim atunci algoritmul întotdeauna întoarce rezultatul prim;
 - ▶ Dacă p este compus, atunci cu probabilitate mare algoritmul va întoarce compus;
 - ► Concluzie: dacă outputul este *compus*, atunci *p* sigur este compus, dacă outputul este *prim*, atunci cu probabilitate mare *p* este prim dar este posibil și să se fi produs o eroare;
- Un algoritm determinist polinomial a fost propus în 2002, dar este mai lent decât algoritmii probabilişti;

- Cei mai eficienți algoritmi sunt probabiliști:
 - ▶ Dacă numărul p dat este prim atunci algoritmul întotdeauna întoarce rezultatul prim;
 - ▶ Dacă p este compus, atunci cu probabilitate mare algoritmul va întoarce compus;
 - Concluzie: dacă outputul este compus, atunci p sigur este compus, dacă outputul este prim, atunci cu probabilitate mare p este prim dar este posibil şi să se fi produs o eroare;
- Un algoritm determinist polinomial a fost propus în 2002, dar este mai lent decât algoritmii probabilişti;
- Un algoritm probabilist foarte răspândit este Miller-Rabin care acceptă la intrare un număr N și rulează în timp polinomial în |N|

▶ Deocamdată nu se cunosc *algoritmi polinomiali* pentru problema factorizării;

- Deocamdată nu se cunosc algoritmi polinomiali pentru problema factorizării;
- Dar există algoritmi mult mai buni decât forța brută;

- Deocamdată nu se cunosc algoritmi polinomiali pentru problema factorizării;
- Dar există algoritmi mult mai buni decât forța brută;
- Prezentăm în continuare câțiva algoritmi de factorizare.

Reamintim: Fiind dat un număr compus N, problema factorizării cere să se găsească 2 numere prime p şi q a.î. N = pq;

- Reamintim: Fiind dat un număr compus N, problema factorizării cere să se găsească 2 numere prime p şi q a.î. N = pq;
- ► Considerăm |p| = |q| = n și deci $n = O(\log N)$;

- Reamintim: Fiind dat un număr compus N, problema factorizării cere să se găsească 2 numere prime p şi q a.î. N = pq;
- ► Considerăm |p| = |q| = n și deci $n = O(\log N)$;
- Metoda cea mai simplă este împărțirea numărului N prin toate numerele p impare din intervalul $p=3,...,\left|\sqrt{N}\right|$.

- Reamintim: Fiind dat un număr compus N, problema factorizării cere să se găsească 2 numere prime p şi q a.î. N = pq;
- ► Considerăm |p| = |q| = n și deci $n = O(\log N)$;
- Metoda cea mai simplă este împărțirea numărului N prin toate numerele p impare din intervalul $p=3,...,\left|\sqrt{N}\right|$.
- ► Complexitatea timp este $O(\sqrt{N} \cdot (\log N)^c)$ unde c este o constantă, adica exponentiala in numarul b de biti al lui N; $b = \log_2 N$. Complexitatea este $O(N^{1/2}) = O((2^{\log_2 N})^{1/2})$

- Reamintim: Fiind dat un număr compus N, problema factorizării cere să se găsească 2 numere prime p şi q a.î. N = pq;
- ► Considerăm |p| = |q| = n și deci $n = O(\log N)$;
- Metoda cea mai simplă este împărțirea numărului N prin toate numerele p impare din intervalul $p=3,...,\left|\sqrt{N}\right|$.
- ▶ Complexitatea timp este $O(\sqrt{N} \cdot (\log N)^c)$ unde c este o constantă, adica exponentiala in numarul b de biti al lui N; $b = \log_2 N$. Complexitatea este $O(N^{1/2}) = O((2^{\log_2 N})^{1/2})$
- Pentru $N < 10^{12}$ metoda este destul de eficientă.

Există însă algoritmi mai sofisticați, cu timp de execuție mai bun, între care:

- Există însă algoritmi mai sofisticați, cu timp de execuție mai bun, între care:
 - Metoda **Pollard p** -1: funcționează atunci când p-1 are factori primi "mici";

- Există însă algoritmi mai sofisticați, cu timp de execuție mai bun, între care:
 - Metoda **Pollard p** -1: funcționează atunci când p-1 are factori primi "mici";
 - Metoda **Pollard rho**: timpul de execuție este $O(N^{1/4} \cdot (\log N)^c)$, deci tot exponențial în lungimea lui N;

- Există însă algoritmi mai sofisticați, cu timp de execuție mai bun, între care:
 - Metoda **Pollard p** -1: funcționează atunci când p-1 are factori primi "mici";
 - Metoda **Pollard rho**: timpul de execuție este $O(N^{1/4} \cdot (\log N)^c)$, deci tot exponențial în lungimea lui N;
 - ► Algoritmul sitei pătratice rulează în timp sub-exponențial în lungimea lui N.

- Există însă algoritmi mai sofisticați, cu timp de execuție mai bun, între care:
 - Metoda **Pollard p** -1: funcționează atunci când p-1 are factori primi "mici";
 - Metoda **Pollard rho**: timpul de execuție este $O(N^{1/4} \cdot (\log N)^c)$, deci tot exponențial în lungimea lui N;
 - Algoritmul sitei pătratice rulează în timp sub-exponențial în lungimea lui N.
- ▶ Deocamdată, cel mai rapid algoritm de factorizare este o îmbunătățire a sitei pătratice care factorizează un număr N de lungime O(n) în timp $2^{O(n^{1/3} \cdot (\log n)^{2/3})}$.

▶ În 1991, Laboratoarele RSA lansează RSA Challenge;

- ▶ În 1991, Laboratoarele RSA lansează RSA Challenge;
- Aceasta presupune factorizarea unor numere N, unde $N = p \cdot q$, cu p, q 2 numere prime mari;

- ▶ În 1991, Laboratoarele RSA lansează RSA Challenge;
- Aceasta presupune factorizarea unor numere N, unde $N = p \cdot q$, cu p, q 2 numere prime mari;
- ► Au fost lansate mai multe provocări, câte 1 pentru fiecare dimensiune (în biţi) a lui *N*:

- ▶ În 1991, Laboratoarele RSA lansează RSA Challenge;
- Aceasta presupune factorizarea unor numere N, unde $N = p \cdot q$, cu p, q 2 numere prime mari;
- Au fost lansate mai multe provocări, câte 1 pentru fiecare dimensiune (în biţi) a lui N:
- Exemple includ: RSA-576, RSA-640, RSA-768, ..., RSA-1024, RSA-1536, RSA-2048;

- ▶ În 1991, Laboratoarele RSA lansează RSA Challenge;
- Aceasta presupune factorizarea unor numere N, unde $N = p \cdot q$, cu p, q 2 numere prime mari;
- Au fost lansate mai multe provocări, câte 1 pentru fiecare dimensiune (în biţi) a lui N:
- Exemple includ: RSA-576, RSA-640, RSA-768, ..., RSA-1024, RSA-1536, RSA-2048;
- Provocarea s-a încheiat oficial în 2007;

- ▶ În 1991, Laboratoarele RSA lansează RSA Challenge;
- Aceasta presupune factorizarea unor numere N, unde $N = p \cdot q$, cu p, q 2 numere prime mari;
- Au fost lansate mai multe provocări, câte 1 pentru fiecare dimensiune (în biţi) a lui *N*:
- Exemple includ: RSA-576, RSA-640, RSA-768, ..., RSA-1024, RSA-1536, RSA-2048;
- Provocarea s-a încheiat oficial în 2007;
- Multe provocări au fost sparte în cursul anilor (chiar şi ulterior închiderii oficiale), însă există numere încă nefactorizate:

► RSA-1024

 $13506641086599522334960321627880596993888147560566\\ 70275244851438515265106048595338339402871505719094\\ 41798207282164471551373680419703964191743046496589\\ 27425623934102086438320211037295872576235850964311\\ 05640735015081875106765946292055636855294752135008\\ 52879416377328533906109750544334999811150056977236\\ 890927563$

[http://www.emc.com/emc-plus/rsa-labs/historical/the-rsa-challenge-numbers.htm]

► RSA-2048

[http://www.emc.com/emc-plus/rsa-labs/historical/the-rsa-challenge-numbers.htm]

 O altă prezumție dificilă este DLP (Discrete Logarithm Problem) (sau PLD (Problema Logaritmului Discret));

- O altă prezumție dificilă este DLP (Discrete Logarithm Problem) (sau PLD (Problema Logaritmului Discret));
- ► Considerăm \mathbb{G} un grup ciclic de ordin q;

- O altă prezumție dificilă este DLP (Discrete Logarithm Problem) (sau PLD (Problema Logaritmului Discret));
- ► Considerăm \mathbb{G} un grup ciclic de ordin q;
- Există un generator $g \in \mathbb{G}$ a.î. $\mathbb{G} = \{g^0, g^1, ..., g^{q-1}\};$

- O altă prezumție dificilă este DLP (Discrete Logarithm Problem) (sau PLD (Problema Logaritmului Discret));
- ► Considerăm \mathbb{G} un grup ciclic de ordin q;
- ightharpoonup Există un generator $g \in \mathbb{G}$ a.î. $\mathbb{G} = \{g^0, g^1, ..., g^{q-1}\};$
- Echivalent, pentru fiecare $h \in \mathbb{G}$ există un *unic* $x \in \mathbb{Z}_q$ a.î. $g^x = h$;

- O altă prezumție dificilă este DLP (Discrete Logarithm Problem) (sau PLD (Problema Logaritmului Discret));
- ► Considerăm \mathbb{G} un grup ciclic de ordin q;
- ightharpoonup Există un generator $g \in \mathbb{G}$ a.î. $\mathbb{G} = \{g^0, g^1, ..., g^{q-1}\};$
- Echivalent, pentru fiecare $h \in \mathbb{G}$ există un *unic* $x \in \mathbb{Z}_q$ a.î. $g^x = h$;
- x se numește logaritmul discret al lui h în raport cu g și se notează

$$x = \log_g h$$

1. Generează (\mathbb{G}, q, g) unde \mathbb{G} este un grup ciclic de ordin q (cu |q| = n) iar g este un generator al lui \mathbb{G} .

- 1. Generează (\mathbb{G} , q, g) unde \mathbb{G} este un grup ciclic de ordin q (cu |q|=n) iar g este un generator al lui \mathbb{G} .
- 2. Alege $h \leftarrow^R \mathbb{G}$. (se poate alege $x' \leftarrow^R \mathbb{Z}_q$ si apoi $h := g^{x'}$.)

- 1. Generează (\mathbb{G}, q, g) unde \mathbb{G} este un grup ciclic de ordin q (cu |q| = n) iar g este un generator al lui \mathbb{G} .
- 2. Alege $h \leftarrow^R \mathbb{G}$. (se poate alege $x' \leftarrow^R \mathbb{Z}_q$ și apoi $h := g^{x'}$.)
- 3. \mathcal{A} primește \mathbb{G}, q, g, h și întoarce $x \in \mathbb{Z}_q$;

- 1. Generează (\mathbb{G}, q, g) unde \mathbb{G} este un grup ciclic de ordin q (cu |q| = n) iar g este un generator al lui \mathbb{G} .
- 2. Alege $h \leftarrow^R \mathbb{G}$. (se poate alege $x' \leftarrow^R \mathbb{Z}_q$ și apoi $h := g^{x'}$.)
- 3. \mathcal{A} primește \mathbb{G}, q, g, h și întoarce $x \in \mathbb{Z}_q$;
- 4. Output-ul experimentului este 1 dacă $g^x = h$ și 0 altfel.

- 1. Generează (\mathbb{G}, q, g) unde \mathbb{G} este un grup ciclic de ordin q (cu |q| = n) iar g este un generator al lui \mathbb{G} .
- 2. Alege $h \leftarrow^R \mathbb{G}$. (se poate alege $x' \leftarrow^R \mathbb{Z}_q$ și apoi $h := g^{x'}$.)
- 3. \mathcal{A} primește \mathbb{G}, q, g, h și întoarce $x \in \mathbb{Z}_q$;
- 4. Output-ul experimentului este 1 dacă $g^x = h$ și 0 altfel.

Definiție

Spunem că problema logaritmului discret (DLP) este dificilă dacă pentru orice algoritm PPT $\mathcal A$ există o funcție neglijabilă negl așa $\widehat{\operatorname{nncat}}$

$$Pr[DLog_{\mathcal{A}}(n) = 1] \leq negl(n)$$

► Există câteva clase de grupuri ciclice pentru care DLP este considerată dificilă;

- Există câteva clase de grupuri ciclice pentru care DLP este considerată dificilă;
- ► Una dintre ele este clasa grupurilor ciclice de *ordin prim* (în aceste grupuri, problema este "cea mai dificilă");

- Există câteva clase de grupuri ciclice pentru care DLP este considerată dificilă;
- ► Una dintre ele este clasa grupurilor ciclice de ordin prim (în aceste grupuri, problema este "cea mai dificilă");
- ▶ DLP nu poate fi rezolvată în timp polinomial în grupurile care nu sunt de ordin prim, ci doar este mai ușoară;

- Există câteva clase de grupuri ciclice pentru care DLP este considerată dificilă;
- ► Una dintre ele este clasa grupurilor ciclice de ordin prim (în aceste grupuri, problema este "cea mai dificilă");
- ▶ DLP nu poate fi rezolvată în timp polinomial în grupurile care nu sunt de ordin prim, ci doar este mai ușoară;
- ▶ În aceste grupuri căutarea unui generator și verificarea că un număr dat este generator sunt triviale.

Lucrul în \mathbb{Z}_p^*

▶ DLP este considerată dificilă în grupuri ciclice de forma \mathbb{Z}_p^* cu p prim;

Lucrul în \mathbb{Z}_p^*

- ▶ DLP este considerată dificilă în grupuri ciclice de forma \mathbb{Z}_p^* cu p prim;
- ▶ Insă pentru p > 3 grupul \mathbb{Z}_p^* NU are ordin prim;

Lucrul în \mathbb{Z}_p^*

- ▶ DLP este considerată dificilă în grupuri ciclice de forma \mathbb{Z}_p^* cu p prim;
- ▶ Insă pentru p > 3 grupul \mathbb{Z}_p^* NU are ordin prim;
- Aceasta problemă se rezolvă folosind un *subgrup* potrivit al lui \mathbb{Z}_p^* ;

Important de reținut!

- Cel mai rapid algoritm de factorizare necesită timp sub-exponențial;
- ► Problema logaritmului discret este dificilă.