DAUGAVPILS UNIVERSITĀTE
Dabaszinātņu un matemātikas fakultāte
Matemātikas katedra
Bakalaura studiju programma "Matemātika"

Studiju kurss

SKAITĻU TEORIJA

11.lekcija

Docētājs: Dr. P. Daugulis

2012./2013.studiju gads

Saturs

Lekcijas mērķis:

- apgūt kriptoloģijas un publisko šifrēšanas sistēmu pamatjēdzienus.

Lekcijas kopsavilkums:

- kriptogrāfija ir saistīta ar matemātiku,
- veselo skaitļu teoriju var izmantot kriptogrāfijas vajadzībām.

Svarīgākie jēdzieni: kriptogrāfija, šifrs, šifrēšana, dešifrēšana, Cēzara šifrs, Vigenère šifrs, Hilla šifrs, simetriskās atslēgas šifrēšana, publiskās atslēgas šifrēšana, RSA kriptosistēma.

Svarīgākie fakti un metodes: Cēzara, Vigenère un Hilla šifra dešifrēšana, iespējamās operācijas publiskās atslēgas kriptosistēmās, RSA šifrēšana un dešifrēšana.

1. Skaitļu teorijas lietojumi informācijas aizsardzībā

1.1. Pamatfakti par šifrēšanu un informācijas aizsardzību

1.1.1. Kriptogrāfijas pamatjēdzieni

Kā aizsargāt datus no nesankcionētas pieejas?

- *fiziskā/loģiskā barjera*- izolēt informāciju no ienaidniekiem,
- *šifrēšana*- uzglabāt un pārsūtīt datus tā, lai ienaidnieki nevarētu tos izmantot, pat ja viņi tiem piekļūst.

Kriptogrāfija (kriptoloģija) - mācība par informācijas slēpšanu, tiek pielietota saistībā ar informācijas drošību (interneta sakari, interneta komercija, militārie sakari u.c.).

Kriptogrāfijas pamatjēdzieni:

- *šifrēšana* - atklātas informācijas (atklātā teksta) pārvēršana nesaprotamā formā (šifrētajā tekstā);
- *dešifrēšana* - operācija, kas ir inversa attiecībā uz šifrēšanu: šifrētā teksta pārvēršana atklātajā tekstā;
- *šifri* - šifrēšanas un dešifrēšanas algoritmi;
- *šifra atslēga* - šifra parametri, kas parasti ir slepeni un tiek bieži mainīti (svarīgos gadījumos tiek izmantoti tikai vienu reizi);

Šifrēšana ir funkcija

$$E : A \times K_E \to S, \text{ kur}$$

- A ir atklāto tekstu kopa,
- K_E ir šifrēšanas atslēgu kopa,
- S ir šifrēto tekstu kopa.

Dešifrēšana ir funkcija

$$D : S \times K_D \to A, \text{ kur}$$

K_D ir dešifrēšanas atslēgu kopa.

Dešifrēšanai bez informācijas par atslēgu (šifra uzlaušanai) ir jābūt ļoti grūtai problēmai, ko nav iespējams atrisināt reālā laikā.

Kā noteikt šifra vērtību? Ja šifra uzlaušanas izmaksa ir lielāka kā šifrētās informācijas vērtība, tad šifru var uzskatīt par labu.

Šifrēšanas metode var tikt uzskatīta par drošu, ja izpildās šādi nosacījumi:

- ir pierādīts vai tiek uzskatīts, ka dešifrēšana nav iespējama bez noteiktas matemātiskas problēmas $\mathcal{P}$ atrisināšanas,
- ir pierādīts vai tiek uzskatīts, ka problēmas $\mathcal{P}$ atrisināšana nav iespējama īsā laikā.

Pirms datoru parādīšanās kriptogrāfija nodarbojās galvenokārt ar informācijas slepenības nodrošināšanu militāriem, diplomātiskiem vai ekonomiskiem mērķiem.

Līdz ar datoru ēras sākumu un ar to saistīto informācijas plūsmas palielināšanos kriptogrāfija nodrošina šādas papildus funkcijas:

- sūtījuma integritātes pārbaude;
- sūtītāja un saņēmēja autentificēšana;
- digitālā paraksta nodrošināšana;
- slepenas informācijas dalīta glabāšana (kādā personu grupā katra persona zina daļu no informācijas, katra k personu grupa var rekonstruēt informāciju, bet grupa, kurā ir mazāk nekā k personu - nevar).

1.1.2. Kriptogrāfijas vēsture

Cēzara šifrs

Senos laikos lielākā daļa cilvēku neprata lasīt, tāpēc tika izmantoti šādi vienkāršākie šifrēšanas veidi:

- burtu kārtības maiņa vārdos,
- burtu aizvietošana ar citiem burtiem (piemēram, *Cēzara šifrs* (ap 50BC), kurā katrs burts tika aizvietots ar to burtu, kas atrodas n pozīcijas tālāk alfabētā, šajā gadījumā atslēga ir skaitlis n).

Cēzara šifru var interpretēt veselo skaitļu terminos šādā veidā:

- katru burtu aizstājam ar atlikumu mod 26: $A = 0$, $B = 1$,..., $Z = 25$;
- šifrēšana tiek veikta neatkarīgi katram burtam ar formulu

$$E(x) \equiv x + n \pmod{26};$$

- dešifrēšana tiek veikta ar formulu

$$D(x) = E^{-1}(x) \equiv x - n \pmod{26}.$$

Aizvietošanas šifri ir viegli uzlaužami, ja izmanto statistisko informāciju par burtu biežumu dotajā valodā, katrā valodā ir specisks burtu sadalījums:

- lai atrastu nobīdi, ir jāatrod simbols x, kas šifrētajā tekstā atkārtojas visbiežāk,
- ja dotajā valodā visbiežāk atkārtojas burts y, tad nobīde ir vienāda ar soļu skaitu no y līdz x.

Vigenère šifrs

Ap 1467.gadu tika izgudrota *polialfabētiskā aizvietošana*, kurā burti tika aizvietoti ar dažādiem burtiem atkarībā no to atrašanās vietas tekstā. Populārākais polialfabētiskās aizvietošanas šifrs - *Vigenère šifrs*, kas darbojas saskaņā ar šādu algoritmu:

- tiek fiksēts atslēgas vārds $K = k_0k_1...k_{m-1}$ - parasti vārds vai teikums tajā pašā valodā, kurā tiek rakstīts teksts;

- K tiek savienots pats ar sevi tik ilgi, kamēr savienojums pārsniedz šifrējamā teksta garumu, tiek iegūts vārds $\mathcal{K} = KK...K$;
- $\mathcal{K}$ tiek rakstīts tieši zem šifrējamā teksta;
- šifrējamā teksta burts τ, zem kura atrodas burts κ tiek aizvietots ar to burtu, kas ir nobīdīts no τ par tādu pašu attālumu, par kuru κ ir nobīdīts no burta a.

Vigenère šifrēšana var tikt interpretēta kā vektoru jeb lineāras telpas elementu pārveidojums šādā veidā:

- sākotnējais šifrējamais teksts $X = x_0x_1...$ tiek sadalīts virknēs

$$X_1 = [x_0|...|x_{m-1}],$$
$$X_2 = [x_m|...|x_{2m-1}],$$
$$...$$

- katra virkne X_i tiek interpretēta kā m-dimensionāls vektors $\mathbf{x}_i$,
- atslēgas vārds K tiek interpretēts kā m-dimensionāls vektors $\mathbf{k}$,

- ar katru vektoru $\mathbf{x}_i$ tiek veikts šifrēšanas pārveidojums

$$E(\mathbf{x}_i) = \mathbf{x}_i + \mathbf{k}.$$

19.gs vidū tika atklātas metodes Vigenère šifra atšifrēšanai, kas balstījās uz šifrētā teksta aritmētisko progresiju apakšvirkņu (formā $(x_i, x_{i+d}, x_{i+2d}, ...)$) statistisko analīzi - lai uzlauztu šifru, ir jāatrod tādas aritmētiskās progresijas apakšvirknes, kurās simbolu biežuma sadalījums ir tuvs burtu biežuma sadalījumam dotajā valodā.

Kerkhofa princips

19.gs beigās kriptogrāfijā tika pieņemts *Kerkhofa princips* - lai šifrēšanas metode būtu droša, tai ir jābalstās tikai uz atslēgas slepenību (ir jāpieņem, ka ienaidnieks zina metodi).

Hilla šifrs

Ap 1930.gadu tika izgudrota grūtāk uzlaužama polialfabētiskās šifrēšanas metode - *Hilla šifrs*, kurā šifrēšana atšķirībā no Vigenère

šifrēšanas tiek veikta ar *afīno pārveidojumu*

$$E(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}, \text{ kur } \mathbf{A} \text{ invertējama matrica.}$$

Dešifrēšana tiek veikta ar pārveidojumu

$$D(\mathbf{x}) = E^{-1}(\mathbf{x}) = \mathbf{A}^{-1}(\mathbf{x} - \mathbf{b}).$$

Uz Otrā pasaules kara beigu laiku polialfabētiskā šifrēšana ar elektromehāniskām ierīcēm sasniedza augstu līmeni (piemēram, vācu *Enigma* šifrēšanas mašīna). Šifrēšana tika veikta ar Vigenère vai Hilla tipa metodēm un garām (šifrējamā teksta garumā) *vienu reizi izmantojamām atslēgām*.

Dešifrēšana bija motivējošs faktors elektronisko skaitļošanas ierīču attīstībai. Alans Tjūrings (teorētiskās datorzinātnes pamatlicējs) kara laikā bija viens no Lielbritānijas dešifrēšanas darba aktīviem dalībniekiem.

Pēc Otrā pasaules kara līdz ar datoru attīstību tika izgudrotas sarežģītākas šifrēšanas metodes:

- tiek šifrēti dažāda formāta teksti (piemēram, binārās virknes);
- šifrēšanas tiek veikta simbolu grupām - blokiem;
- tiek intensīvi pielietoti matemātikas (skaitļu teorijas, varbūtību teorijas) sasniegumi.

Datori tika izmantoti arī dešifrēšanā.

Ja abas atslēgas ir slepenas, tādu šifrēšanas metodi sauc par *simetriskās atslēgas* kriptosistēmu.

70.gadu vidū Lielbritānijā (Ellis-Cocks-Williamson) un ASV (Diffie-Hellman) tika izgudrotas vairākas *publiskās atslēgas* kriptosistēmas, kurās viena no divām atslēgām ir zināma visiem, bet otra ir slepena.

80.-90.gados tika izgudrotas vairākas kriptosistēmas (piemēram, Imai-Matsumoto kriptosistēmas), kuru dešifrēšana balstās vienlaicīgi uz vairākām matemātikas sadaļām (veselo skaitļu teoriju, polinomu teoriju u.c.) Mūsdienās kriptogrāfijas attīstība turpinās, jo tai ir liels

pieprasījums (militārie un drošības pielietojumi, interneta drošība un komercija u.c.)

1.1.3. Simetriskās atslēgas kriptosistēmas

Simetriskās atslēgas kriptosistēma (SAK) ir kriptosistēma, kurā sūtītājam un saņēmējam ir kopīga informācija par atslēgu. Līdz 1976. gadam bija tikai tādas šifrēšanas metodes.

Ir divu veidu simetriskās atslēgas kriptosistēmas:

- plūsmas šifrēšana;
- bloku šifrēšana.

Plūsmas šifrēšana ir līdzīga Vigenere šifrēšanai, kurā atslēga tiek ģenerēta ar slepenu algoritmu palīdzību izmantojot nejaušos skaitļus.

Bloku šifrēšana ir šifrēšanas metode, kurā teksts tiek dalīts apakšvirknēs - blokos (parasti blokā ir 64 vai 128 biti) un katrs bloks tiek šifrēts ar slepenas atslēgas palīdzību.

1.1.4. Publiskās atslēgas kriptosistēmas

Publiskās atslēgas jeb *asimetriskā* kriptosistēma (PAK) ir salīdzinoši jauna (70.gadi) metode, kurā viena no šifrēšanas/dešifrēšanas operācijām ir atklāta, bet otra - slepena.

Šī metode tika izstrādāta, lai mazinātu grūtības, kas ir saistītas ar slepeno atslēgu nodošanu visām pusēm, kas piedalās sakaros. Ar šī tipa šifrēšanas metožu palīdzību var vieglāk pārraidīt informāciju pa neaizsargātiem sakaru kanāliem.

Precīzāk, katram lietotājam ir divas atslēgas:

- *publiskā atslēga* (publiski pieejama visiem) un
- *privātā atslēga* (slepena, zināma tikai noteiktam lietotāju lokam).

Abas atslēgas ir saistītas ar matemātiskiem algoritmiem, bet zinot publisko atslēgu, ir praktiski neiespējami atrast privāto atslēgu.

Tādējādi, katram lietotājam X ir definētas divas savstarpēji inversas funkcijas E_X un $D_X = E_X^{-1}$ tādas, ka

$$E_X \circ D_X = \mathrm{id},$$

$$D_X \circ E_X = \mathrm{id}\,.$$

Funkcija E_X katram X ir publiski zināma.

Šādā kriptosistēmā ir iespējamas divas operācijas:

- (*šifrēšana ar publisko atslēgu*) lai nosūtītu lietotājam X slepenu sūtījumu M, šis sūtījums ir jāaizšifrē ar funkciju E_X, šādā gadījumā tikai X varēs izlasīt sūtījumu $E_X(M)$, pielietojot tam savu slepeno funkciju D_X:

$$D_X\Big((E_X(M)\Big) = \Big(D_X \circ E_X\Big)(M) = \mathrm{id}(M) = M.$$

- (*digitālā paraksta nodrošināšana*) lai lietotājs X varētu pierādīt savu identitāti, X aizšifrē kādu noteiktu tekstu N (piemēram, savu vārdu) ar savu slepeno funkciju D_X, jebkurš sūtījuma

$D_X(N)$ saņēmējs var pielietot šim sūtījumam publisko funkciju E_X, izlasīt rezultātu

$$E_X\Big(D_X(N)\Big) = \Big(E_X \circ D_X\Big)(N) = \mathrm{id}(N) = N$$

un pārliecināties, ka sūtītājs ir bijis X (vai vismaz kāds, kas zina X slepeno atslēgu).

1.1. piezīme. PAK var tikt lietota kombinācijā ar SAK.

PAK mēģina uzlauzt, provocējot X aizšifrēt vienu vai vairākus tekstus ar noteiktām īpašībām, kas atvieglo dešifrēšanu.

1.2. Publisko atslēgu kriptosistēmu konkrēta realizācija

1.2.1. Ievads

PAK balstās uz matemātiķiem zināmu novērojumu, ka ir funkcijas/algoritmi, kuriem ir grūti atrast inversās funkcijas/algoritmus.

Piemēri:

1. reizināt ir vieglāk nekā dalīt;
2. kāpināt ir vieglāk nekā atrast sakni;
3. reizināt pirmskaitļus ir vieglāk nekā atrast skaitļa sadalījumu pirmskaitļu reizinājumā;
4. izjaukt puzli ir vieglāk nekā salikt.

- PAKā vieglā funkcija - publiski pieejamā funkcija,
- grūtā (inversā) funkcija - slepenā privātā funkcija.

1.2.2. Rivest-Shamir-Adleman (RSA) kriptosistēma

Pamatideja:

- meklēsim šifrējošo funkciju pakāpes funkcijas formā

$$a \to a^e (\text{mod } n),$$

- $\Longrightarrow$ dešifrējošā funkcija arī var būt pakāpes funkcija ar kāpinātāju d:

$$(a^e)^d \equiv a^{ed} \equiv a (\text{mod } n),$$

- $\Longrightarrow$ $ed \equiv 1 (\text{mod } \varphi(n))$,
- lai būtu grūti dešifrēt - atrast d mod $\varphi(n)$, nepieciešams izvēlēties n tā, lai $\varphi(n)$ ir grūti atrast,
- $\Longrightarrow$ n jābūt tādam, lai n ir grūti faktorizēt.

RSA kriptosistēmas atslēgu ģenerēšanas algoritms:

1. Izvēlēties divus lielus pirmskaitļus p un q.
2. Atrast $n = pq$, n tiek lietots kā palīginformācija (modulis) abās atslēgās.
3. Atrast $\varphi(n) = (p-1)(q-1)$.
4. Atrast netriviālu elementu $e \in \mathcal{U}_{\varphi(n)}$:

$$\begin{cases} LKD(e, \varphi(n)) = 1 \\ 1 < e < \varphi(n) - 1 \end{cases}$$

   e ir publiskā atslēga.
5. Atrast $d \equiv e^{-1} \pmod{\varphi(n)}$, piemēram, izmantojot Eiklīda algoritmu, d ir privātā (slepenā) atslēga.

Publiski pieejamā informācija - (n, e).
Slepenā informācija - (p, q, d).

RSA šifrēšanas algoritms (funkcija E_X):

1. Sadalīt sūtāmo tekstu M daļās $m_1 m_2 ... m_k$ tā, lai $\forall\ m_i$ pēc iekodēšanas decimālajā vai binārajā pierakstā ar publiski zināmu algoritmu būtu mazāka nekā n.
2. $\forall\ m_i$ pārveidot par $c_i = m_i^e \pmod{\text{n}}$. Šifrēšanas rezultāts ir virkne $c_1 c_2 ... c_k$.

RSA dešifrēšanas algoritms (funkcija D_X):

1. Katram i atrast $m_i = c_i^d (\text{mod } n)$.
2. Savienot atšifrētās daļas m_i vienā virknē.

1.1. teorēma. *(RSA dešifrēšanas algoritma pamatojums)*

$$\Big(c \equiv m^e (\text{mod } n)\Big) \implies \Big(m \equiv c^d (\text{mod } n)\Big).$$

PIERĀDĪJUMS

$$ed \equiv 1\ (\text{mod } \varphi(n)) \implies ed = 1 + l\varphi(n) \implies$$

$$c^d \equiv (m^e)^d \equiv m^{ed} \equiv m^{1+l\varphi(n)} \equiv m \cdot (m^{\varphi(n)})^l (\text{mod } n).$$

Jāapskata divi gadījumi: $m \in \mathcal{U}_n$ vai $m \notin \mathcal{U}_n$.

$$\underline{m \in \mathcal{U}_n} \Longrightarrow m^{\varphi(n)} \equiv 1(\text{mod } n)(\text{Eilera teorēma}) \Longrightarrow$$

$$c^d \equiv m(\text{mod } n).$$

$$\underline{m \notin \mathcal{U}_n} \Longrightarrow LKD(m, \underbrace{n}_{=pq}) > 1 \Longrightarrow$$

$$m \equiv 0(\text{mod } p) \vee m \equiv 0(\text{mod } q).$$

Apskatīsim gadījumu $m \equiv 0(\text{mod } p)$.

$$m \equiv 0(\text{mod } p) \Longrightarrow m^{ed} \equiv 0 \equiv m(\text{mod } p) \Longrightarrow$$

$$m^{ed} \equiv m^{1+l(p-1)(q-1)} \equiv m \cdot (m^{q-1})^{p-1} \equiv m(\text{mod } q).$$

Attiecībā uz m^{ed} esam ieguvuši sistēmu

$$\begin{cases} m^{ed} \equiv m(\text{mod } p) \\ m^{ed} \equiv m(\text{mod } q). \end{cases}$$

Saskaņā ar ķīniešu atlikumu teorēmu $\Longrightarrow$ $m^{ed} \equiv c^d \equiv m (\mathrm{mod}\ pq)$.

Līdzīgi tiek pierādīts gadījums, kad $m \equiv 0 (\mathrm{mod}\ q)$. ■

1.1. piemērs. Lai uzlauztu RSA kriptosistēmu, ir jāzina $d (\mathrm{mod}\ \varphi(n))$. Lai atrastu d un $\varphi(n)$, ir jāzina p un q, kas ir grūti.

1.2. piemērs. Atradīsim RSA kriptosistēmas atslēgas un algoritmus, ja $p = 17$ un $q = 19$:

1. Ir izvēlēti divi pirmskaitļi $p = 17$ un $q = 19$.
2. Atrast $n = pq = 323$, n tiek lietots kā palīginformācija (modulis) abās atslēgās.
3. Atrast $\varphi(n) = (p-1)(q-1) = 288$.
4. Atrast invertējamu elementu $e \in \mathcal{U}_{288}$ tādu, ka $1 < e < 288$ $(LKD(e, 288) = 1)$: izvēlēsimies $e = 5$, e ir publiskā atslēga.
5. Atrast $d \equiv e^{-1} (\mathrm{mod}\ \varphi(n))$: $d \equiv 173 (\mathrm{mod}\ 288)$, d ir privātā (slepenā) atslēga.

Šifrējošā funkcija ir $E(m) \equiv m^5 \pmod{323}$, dešifrējošā funkcija $D(c) \equiv c^{173} \pmod{323}$. Ja $m = 300$, tad

$$E(300) \equiv 300^5 \equiv 78 \pmod{323}$$

$$D(E(300)) = D(78) \equiv 78^{173} \equiv 300 \pmod{323}.$$

2. 11.mājasdarbs

11.1 Izmantojot Cēzara šifru ar nobīdi 11 un latīņu alfabēta standarta kārtību, aizšifrējiet tekstu "Gallia est omnis divisa in partes tres".

11.2 Izmantojot Vigenère šifru ar atslēgas vārdu "livonija" atšifrējiet tekstu "iģtohi" izmantojot latviešu alfabēta standarta kārtību.

11.3 Latīņu alfabēts un vēl 3 simboli [.],[ ],[?] ir iekodēti kā atlikumi mod 29: $A = 0$, $B = 1$, ..., $Z = 25$, [.] = 26, [ ] = 27, [?] = 28. Teksts tiek sadalīts kolonnās pa 4 simboli, katra kolonna tiek šifrēta ar Hilla šifru $E(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kur $\mathbf{A} = \left[\begin{array}{c|c|c|c} 1 & 1 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \end{array}\right]$,

$\mathbf{b} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$. Tika pārtvertas vairākas šifrētas kolonnas -

$$\begin{bmatrix} 0 \\ 5 \\ 19 \\ 8 \end{bmatrix}, \begin{bmatrix} 9 \\ 28 \\ 0 \\ 22 \end{bmatrix}, \begin{bmatrix} 8 \\ 27 \\ 17 \\ 14 \end{bmatrix}, \begin{bmatrix} 7 \\ 6 \\ 20 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 14 \\ 8 \end{bmatrix}, \begin{bmatrix} 2 \\ 22 \\ 26 \\ 23 \end{bmatrix}$$

Atšifrējiet pārtverto sūtījumu (teksts ir angļu valodā, kolonnu kārtība var būt sajaukta).

11.4 Atrodiet atslēgas un aprakstiet šifrēšanas/dešifrēšanas algoritmus RSA sistēmā ar pirmskaitļiem 83 un 89.