Luentorunko ja harjoitustehtävät

SALAUSMENETELMÄT (801346A) 4 op, 2 ov

Keijo Väänänen

I JOHDANTO

Salakirjoitukset kurssilla tarkastelemme menetelmiä, jotka mahdollistavat tiedon siirtämisen tai tallentamisen niin, että ainoastaan tarkoitettu vastaanottaja saa viestin selville. Lähettäjän tehtävänä on salakirjoittaa (encrypt) selväkielinen teksti (plaintext) salakirjoitukseksi (cryptotext) ja vastaanottajan tehtävänä puolestaan avata (decrypt) salakirjoitus selväkieliseksi tekstiksi. Menettelyn tulee olla sellainen, että mahdollinen salakirjoituksen sieppaaja ei kykene murtamaan sitä (ts. selvittämään selväkielistä tekstiä) ainakaan nopeasti.

Aikaisemmin salakirjoituksia tarvittiin lähinnä sotilaallisiin tai diplomaattisiin tarkoituksiin. Viimeisten runsaan 20 vuoden aikana tietokoneisiin perustuvan tiedonvälityksen yleistyminen on merkinnyt sitä, että salausmenetelmiä tarvitaan päivittäin hyvin monilla muillakin yhteiskunnan alueilla (pankit, yritykset ym.).

Esimerkki (Caesar). Salakirjoitus tehdään siirtämällä kirjaimia k askelta eteenpäin, esim $k = 3$ antaa seuraavaa:

$$A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ X\ Y\ Z\ \ddot{A}\ \ddot{O}$$
$$\downarrow$$
$$D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ X\ Y\ Z\ \ddot{A}\ \ddot{O}\ A\ B\ C$$

selväkielinen teksti	: $ALOITAMME$
	$\downarrow$
salakirjoitus	: $DORLXDPPH$
salakirjoitus	: $KHOS$
	$\downarrow$
avattuna selväkieliseksi	: $HELP$

Tämä salakirjoitus on helppoa murtaa kokeilemalla eri siirtovaihtoehdot (26 kpl).

Selväkielinen teksti ja salakirjoitettu teksti kirjoitetaan käyttämällä jotakin aakkostoa (kirjaimet, numerot, muut merkit). Jotta matemaattisia menetelmiä voitaisiin käyttää, korvataan kirjaimet usein numeroilla, esimerkiksi $A = 0,\ B = 1, ..., \ddot{O} = 26$. Selväkielinen teksti ja salakirjoitus jaetaan viestiyksikköihin ja salaaminen tehdään yksikkö kerrallaan. Viestiyksikkö voi olla kirjain (kuten äskeisessä esimerkissä), kirjainpari tai tietyn pituinen kirjainjono. Salakirjoittamiseen käytetään yleensä bijektiivistä funktiota $E : P \to C$, missä

$$P = \{\text{selväkieliset viestiyksiköt}\} = \{m\},$$
$$C = \{\text{salakirjoitetut viestiyksiköt}\} = \{c\}.$$

Avaaminen tapahtuu tällöin käänteisfunktion $D = E^{-1}$ avulla.

Salakirjoitusjärjestelmään kuuluvat siis: (i) P, (ii) C, (iii) avainjoukko $K = \{k\}$, missä kukin avain k määrää salausfunktion E_k ja avausfunktion D_k, jolle $D_k(E_k(m)) = m$. Lähettäjä tuntee ennakolta E_k:n ja vastaanottaja D_k:n, jolloin järjestelmä toimii seuraavan kaavion mukaisesti.

Hyvältä salakirjoitusjärjestelmältä edellytetään:

(i) $E_k(m)$ ja $D_k(c)$ voidaan laskea helposti.

(ii) Jollei tunneta D_k:ta, niin m ei selviä c:stä (ts. sieppaajalla on vaikea tehtävä).

Kaikissa perinteisissä salakirjoitusjärjestelmissä D_k saadaan välittömästi E_k:sta (ei tosin aina niin helposti kuin äskeisessä esimerkissä). Näin ollen lähettäjän ja vastaanottajan tulee sopia jollakin tavalla avaimesta ja pitää tämä sopimuksensa salassa muilta. Tästä syystä näistä järjestelmistä käytetään nimitystä yksityisen avaimen salakirjoitus. Runsaat 20 vuotta sitten kehitettiin ensimmäiset julkisen avaimen järjestelmät, joille on ominaista se, että E_k ei paljasta D_k:ta (ainakaan helposti). Nämä perustuvat ns. yksisuuntaisiin funktioihin (one-way function) f, joiden käänteisfunktiota on käytännössä mahdotonta (tai ainakin hyvin vaikeaa) määrittää.

Useimmat käytössä olevat salakirjoitusmenetelmät perustuvat lukuteorian tuloksiin. Tästä syystä aloitamme kurssin kertaamalla eräitä lukuteorian alkeiden tuloksia.

II LUKUTEORIAA

1. Jakoalgoritmi ja eri kantaiset esitykset

Luvuista puhuttaessa tarkoitamme seuraavassa luonnollisia lukuja $\mathbb{N} = \{0, 1, 2, ...\}$ tai kokonaislukuja $\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$.

Lause 1 (jakoalgoritmi). Jos $a, b \in \mathbb{Z}, b \neq 0$, niin on olemassa sellaiset yksikäsitteiset $q, r \in \mathbb{Z}$, että

$$a = qb + r, \quad 0 \leq r < |b|.$$

Olemme tottuneet käsittelemään lukuja 10-järjestelmässä, esimerkiksi $2367 = 2 \cdot 10^3 + 3 \cdot 10^2 + 6 \cdot 10 + 7$. Luku 10 on luonnollinen biologisista syistä, mutta mikään ei estä valitsemasta kantalukua $b \geq 2$ muutenkin. Annettu $a \in \mathbb{N}$ voidaan esittää helposti jakoalgoritmia käyttäen b-kantaisessa järjestelmässä. Esimerkiksi luvun 319 8-kantainen esitys saadaan jakamalla toistuvasti 8:lla:

$$319 = 39 \cdot 8 + 7, \quad 39 = 4 \cdot 8 + 7,$$

joten

$$310 = (4 \cdot 8 + 7) \cdot 8 + 7 = 4 \cdot 8^2 + 7 \cdot 8 + 7.$$

Merkitsemme $319 = 477_8$. Vastaavasti esimerkiksi $251_6 = 2 \cdot 6^2 + 5 \cdot 6 + 1 = 103$. Jos siis kantaluku $b \neq 10$, merkitsemme sen näkyviin alaindeksinä.

10-järjestelmän ohella erityisen tärkeä on 2-kantainen eli binäärijärjestelmä, jota tietokoneet ymmärtävät.

2. Jaollisuus

Määritelmä. Jos $a, b \in \mathbb{Z}$ ja on olemassa sellainen $c \in \mathbb{Z}$, että $a = bc$, niin b on a:n tekijä (tai a on jaollinen b:llä). Tällöin merkitään $b|a$.

Jaollisuudella on seuraava ominaisuus.

Lause 2. Jos $n|a$ ja $n|b$, niin $n|(ra + sb)$ $\forall$ $r, s \in \mathbb{Z}$.

Seurauksia: 1) $n|a$ ja $n|b \Rightarrow n|(a \pm b)$,
2) $n|a \Rightarrow n|ra$ $\forall$ $r \in \mathbb{Z}$,
3) $n|a$ ja $a|b \Rightarrow n|b$.

Jokaisella $a \in \mathbb{Z}$ on aina ns. triviaalit tekijät ± 1 ja $\pm a$.

Määritelmä. Luku $p \geq 2$ on alkuluku, jos sillä on vain triviaalit tekijät (± 1 ja $\pm p$).

Alkulukuja on äärettömän monta ja kaikki positiiviset kokonaisluvut voidaan esittää oleellisesti yksikäsitteisesti niiden avulla, sillä on voimassa

Lause 3 (aritmetiikan peruslause). Jokainen positiivinen kokonaisluku ≥ 2 voidaan esittää järjestystä vaille yksikäsitteisesti alkulukujen tulona.

Seurauksia: 1) Jos p on alkuluku ja $p|ab$, niin $p|a$ tai $p|b$.
2) Jos $r|a$ ja $s|a$ ja luvuilla r ja s ei ole yhteisiä alkulukutekijöitä, niin $rs|a$.
3) Jos tunnetaan luvun esitys alkulukujen tulona ($\exists$ $\mathcal{L}3$ perusteella), niin positiivisten tekijöiden lukumäärä on helppo laskea. Esimerkiksi $4200 = 2^3 \cdot 3 \cdot 5^2 \cdot 7$, joten tekijöiden lukumäärä on (3+1)(1+1)(2+1)(1+1)=48.

Määritelmä. Olkoot $a, b \in \mathbb{Z}$ ja ainakin toinen $\neq 0$. Lukujen a ja b suurin yhteinen tekijä $syt(a, b)$ on suurin luonnollinen luku, joka on sekä a:n että b:n tekijä.

Nähdään helposti, että $d > 0$ on $syt(a, b)$, jos
1) $d|a$ ja $d|b$,
2) $n|a$ ja $n|b \Rightarrow n|d$.
Jos a:n ja b:n alkutekijäesitykset tunnetaan, niin on helppoa antaa $syt(a, b)$. Esimerkiksi $4200 = 2^3 \cdot 3 \cdot 5^2 \cdot 7$ ja $10780 = 2^2 \cdot 5 \cdot 7^2 \cdot 11$, joten

$$syt(4200, 10780) = 2^2 \cdot 5 \cdot 7 = 140.$$

Suuria lukuja tarkasteltaessa niiden alkutekijäesitysten löytäminen on vaikeaa. Kuitenkin on olemassa menettely, ns. Eukleideen algoritmi, jonka avulla syt löytyy helposti. Kyse on jakoalgoritmin toistuvasta soveltamisesta. Voimme olettaa, että $a \geq b > 0$.

Eukleideen algoritmi:

$$a = q_1 b + r_1,\ 0 < r_1 < b,$$
$$b = q_2 r_1 + r_2,\ 0 < r_2 < r_1,$$
$$r_1 = q_3 r^2 + r^3,\ 0 < r_3 < r_2,$$
$$\dots$$
$$r_{n-2} = q_n r_{n-1} + r_n,\ 0 < r_n < r_{n-1},$$
$$r_{n-1} = q_{n+1} r_n.$$

Lause 4. Viimeinen nollasta eroava jakojäännös $r_n = syt(a, b)$. Edelleen on olemassa sellaiset $u, v \in \mathbb{Z}$ (löytyvät helposti Eukleideen algoritmilla), että

$$syt(a, b) = ua + vb.$$

Lukuja a ja b sanotaan keskenään jaottomiksi tai suhteellisiksi alkuluvuiksi, jos $syt(a, b) = 1$.

Määritelmä. Ns. Eulerin funktio φ on määritelty $\forall n \geq 1$ ja $\varphi(n)$ on niiden kokonaislukujen b lukumäärä, joille $0 \leq b < n$ ja $syt(b, n) = 1$, ts.

$$\varphi(n) = |\{0 \leq b < n |\ syt(b, n) = 1\}|.$$

($|A|$ tarkoittaa joukon A alkioiden lukumäärää.)

Lause 5. Eulerin funktiolla φ on ominaisuudet:

1) $\varphi(1) = 1$;

2) Jos p on alkuluku, niin $\varphi(p^k) = p^{k-1}(p-1)$, erityisesti $\varphi(p) = p - 1$;

3) Jos p ja $q \neq p$ ovat alkulukuja, niin $\varphi(pq) = (p-1)(q-1)$.

3. Kongruensseista

Esittelemme nyt jaollisuuteen perustuvan kongruenssikäsitteen, jonka otti käyttöön Gauss.

Määritelmä. Jos $a, b, n \in \mathbb{Z},\ n \geq 2$, niin a on kongruentti b modulo n, merk. $a \equiv b \pmod{n}$, jos $m|(a - b)$.

Ominaisuuksia:

1) $a \equiv a \pmod{n}$; $a \equiv b \pmod{n} \Leftrightarrow b \equiv a \pmod{n}$; $a \equiv b \pmod{n}$ ja $b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$.

2) Jos $a \equiv b \pmod{n}$ ja $c \equiv d \pmod{n}$, niin

$$a \pm c \equiv b \pm d \pmod{n} \text{ ja } ac \equiv bd \pmod{n}.$$

Ominaisuuden 1) perusteella kongruenssi (mod n) on ekvivalenssirelaatio, joten se jakaa $\mathbb{Z}$:n alkiot ekvivalenssiluokkiin, ns. jäännösluokkiin (mod n). Luvun a määräämä jäännösluokka on

$$\bar{a} = \{x \in \mathbb{Z} | x \equiv a \pmod{n}\} = \{a + kn | k \in \mathbb{Z}\}.$$

Selvästi $\bar{a} = \bar{b} \Leftrightarrow a \equiv b \pmod{n}$. Jos $a \in \mathbb{Z}$, niin jakoalgoritmin mukaan on olemassa yksikäsitteiset $q, r \in \mathbb{Z}$, joille

$$a = qn + r,\ 0 \leq r < n.$$

Tällöin $\bar{a} = \bar{r}$, joten jokainen kokonaisluku kuuluu täsmälleen yhteen luokista $\bar{0}, \bar{1}, \bar{2}, ..., \overline{n-1}$. Nämä jäännösluokat ovat erillisä, joten jäännösluokkia (mod n) on n kpl. Käytämme niiden joukolle merkintää $\mathbb{Z}_n$, joten edellisen mukaan

$$\mathbb{Z}_n = \{\bar{0}, \bar{1}, \bar{2}, ...\overline{n-1}\}.$$

Ominaisuuteen 2) nojautuen voimme määritellä jäännosluokkien (mod n) summan ja tulon asettamalla

$$\bar{a} + \bar{b} = \overline{a+b}, \ \bar{a}\bar{b} = \overline{ab}.$$

Asettamalla lisäksi $-\bar{a} = \overline{-a}$ ja $\bar{a} - \bar{b} = \bar{a} + (-\bar{b})$, voimme suorittaa jäännösluokilla (mod n) yhteen-, vähennys- ja kertolaskuja. Lisäksi joukosta $\mathbb{Z}$ tutut laskusäännöt ovat voimassa, ts. $\mathbb{Z}_n$ on vaihdannainen ykkösellinen rengas ykkösalkionaan $\bar{1}$.

Tarkastelemme seuraavaksi käänteisalkion olemassaoloa. Olkoon $\bar{a} \neq \bar{0}$. Milloin on olemassa sellainen $\bar{x}$, että $\bar{a}\bar{x} = \bar{1}$? Voimme yhtäpitävästi kysyä, milloin kongruenssilla $ax \equiv 1 \pmod{n}$ on ratkaisu x?

Lause 6. Kongruenssilla $ax \equiv 1 \pmod{n}$ on ratkaisu jos ja vain jos $syt(a, n) = 1$. Ratkaisut muodostavat täsmälleen yhden jäännösluokan (mod n).

Lauseen 6 todistuksesta käy ilmi, miten yhtälön $\bar{a}\bar{x} = \bar{1}$ ratkaisu $\bar{a}^{-1} = \bar{x}$ löytyy Eukleideen algoritmin avulla. Luokkaa $\bar{a}$, jolle $\bar{a}^{-1}$ on olemassa, sanotaan alkuluokaksi (mod n). Alkuluokkien lukumäärä on $\varphi(n)$ ja niiden muodostamalle joukolle käytetään merkintää $\mathbb{Z}_n^*$. $\mathbb{Z}_n^*$ on vaihdannainen ryhmä kertolaskun suhteen. Jos $n = p$ on alkuluku, niin $\mathbb{Z}_p^* = \mathbb{Z}_p - \{\bar{0}\}$, joten $\mathbb{Z}_p$:n jokaisella nollasta eroavalla alkiolla on käänteisalkio. Joukossa $\mathbb{Z}_p^*$ voidaan siis suorittaa rajoituksetta kaikkia neljää laskutoimitusta, vain nollalla jakaminen ei ole mahdollista, ja kaikki rationaaliluvuilta tutut laskusäännöt toimivat. $\mathbb{Z}_p$ on näin esimerkki äärellisestä kunnasta. Nämä ovat tärkeitä joissakin salakirjoitusmenetelmissä.

Kun otetaan yksi alkio jokaisesta jäännösluokasta (mod n), saadaan ns täydellinen jäännössysteemi (mod n). Vastaavasti ottamalla yksi alkio kustakin alkuluokasta (mod n), saadaan supistettu jäännössysteemi (mod n). Helposti todetaan, että jos $a_1, a_2, ..., a_{\varphi(n)}$ on supistettu jäännössysteemi ja $syt(a, n) = 1$, niin myös $aa_1, aa_2, ..., aa_{\varphi(n)}$ on supistettu jäännössysteemi. Tähän tietoon nojautuen ei ole vaikea todistaa

Lause 7 (Eulerin lause). Jos $syt(a, n) = 1$, niin $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Seuraus (Fermat'n pieni lause). Jos p on alkuluku ja $p \nmid a$, niin

$$a^{p-1} \equiv 1 \pmod{p}.$$

Lause 7 voidaan myös lausua muodossa: Jos $\bar{a} \in \mathbb{Z}_n^*$, niin $\bar{a}^{\varphi(n)} = \bar{1}$. Tästä käy ilmi, että $\bar{a}^{-1} = \bar{a}^{\varphi(n)-1}$, joten käänteisalkio $\bar{a}^{-1}$ saadaan potenssiinkorotuksella laskemalla $\bar{a}^{\varphi(n)-1}$.

Olkoon seuraavassa $a \pmod{n}$ jakojäännös, kun a jaetaan n:llä. Tarkastelemme potenssin $a^\ell \pmod{n}$ nopeaa laskemista, kun $syt(a, n) = 1$. Jakoalgoritmin mukaan

$$\ell = q\varphi(n) + r, \ 0 \leq r < \varphi(n).$$

Eulerin lauseen mukaan $a^{\varphi(n)} \equiv 1 \pmod{n}$, joten

$$a^\ell = (a^{\varphi(n)})^q a^r \equiv a^r \pmod{n}$$

eli $a^\ell \pmod{n} = a^r \pmod{n}$. Olkoon luvun r binääriesitys

$$r = e_k \cdot 2^k + e_{k-1} \cdot 2^{k-1} + ... + e_1 \cdot 2 + e_0,\ e_i \in \{0,1\},\ e_k = 1.$$

Lasketaan peräkkäisillä neliöönkorotuksilla:

$$a_1 = a^2 \pmod{n},$$
$$a_2 = a_1^2 \pmod{n} (= a^{2^2} \pmod{n}),$$
$$a_3 = a_2^2 \pmod{n} (= a^{2^3} \pmod{n}),$$
$$\vdots$$
$$a_k = a_{k-1}^2 \pmod{n} (= a^{2^k} \pmod{n}).$$

Tämän jälkeen saadaan

$$a^\ell \equiv a^r \equiv a_k^{e_k} a_{k-1}^{e_{k-1}} ... a_2^{e_2} a_1^{e_1} a^{e_0} \pmod{n},$$

josta $a^\ell \pmod{n}$ on nopeasti laskettavissa.

III PERINTEISIÄ SALAKIRJOITUSJÄRJESTELMIÄ

1. Caesar ja sen yleistyksiä

Oletamme, että käytämme aakkostoa, jossa on N symbolia.

- suomenkielinen aakkosto, $N = 27$,
- suomenkielinen aakkosto ja väli, $N = 28$,
- suomenkielinen aakkosto, väli ja numerot, $N = 38$,
- englanninkielinen aakkosto, $N = 26$

  ($\ddot{a}$ ja $\ddot{o}$ puuttuvat, w lisää).

Tällöin on yksinkertaista asettaa joukon $\mathbb{Z}_N$ alkiot vastaamaan aakkoston symboleja. Esimerkiksi suomenkielistä aakkostoa vastaa $\mathbb{Z}_{27}$,

$$A = 0, B = 1, ..., \ddot{O} = 26,$$

missä jätämme jäännösluokkien (mod 27) yläviivat pois.

Aikaisemmin tarkastelemamme Caesarin menetelmän salausfunktio E_k ja avausfunktio D_k ovat yo. merkinnöillä yksinkertaisesti

$$E_k(x) = x + k;\ D_k(x) = x - k,$$

missä laskutoimitukset tehdään joukossa $\mathbb{Z}_N$.

Kuten aikaisemmin totesimme, tämä järjestelmä on helppo murtaa kokeilemalla kaikki avaimet k, joita on $N - 1$ kpl.

Tarkastelemme nyt yleisempää järjestelmää, missä yhteenlasku korvataan affiinilla kuvauksella. Tätä varten valitsemme $a \in \mathbb{Z}_N^*,\ b \in \mathbb{Z}_N$. Avaimemme on nyt $k = (a, b)$ ja salausfunktio on

$$E(x) = E_{a,b}(x) = ax + b.$$

Avainten lukumäärä on tällöin $\varphi(N) \cdot N$.

Erikoistapaukset:

$a = 1 \Rightarrow$Caesar: $E(x) = x + b$.

$b = 0 \Rightarrow$ ns. kertolasku Caesar:$E(x) = ax$.

Selvästi $E : \mathbb{Z}_N \to \mathbb{Z}_N$ on bijektio ja sen käänteisfunktio on avausfunktio D:

$$y = ax + b \Leftrightarrow ax = y - b \Leftrightarrow x = a^{-1}y - a^{-1}b.$$

Edellä $a \in \mathbb{Z}_N^*$, joten $a^{-1}\exists$. Siis avausfunktio

$$D(y) = D_{a,b}(y) = a^{-1}y - a^{-1}b.$$

Yhteenveto affiinista järjestelmästä:

$P = C = \mathbb{Z}_N$.

$K = \{(a, b)\},$ missä $a \in \mathbb{Z}_N^*$ ja $b \in \mathbb{Z}_N$.

Salausfunktio $E(x) = ax + b$.

Avausfunktio $D(y) = a^{-1}y - a^{-1}b$.

Affiinin järjestelmän murtaminen.

1) Kokeillaan kaikki avaimet.

2) Käytetään hyväksi kirjainten esiintymistiheyksiä
(tarvitaan aika pitkä salakirjoitusteksti).

Em. järjestelmät ovat erikoistapauksia yksinkertaisesta sijoitusjärjestelmästä. Siinä on avainjoukkona $\mathbb{Z}_N$:n permutaatioiden muodostama joukko S_N, jonka alkiot σ kirjoitetaan usein muotoon

$$\sigma = \begin{pmatrix} 0 & 1 & \dots & N-1 \\ \sigma(0) & \sigma(1) & \dots & \sigma(N-1) \end{pmatrix}.$$

Nämä ovat bijektioita, joten niillä on käänteiskuvaus σ^{-1}.

Yksinkertainen sijoitusjärjestelmä:

$P = C = \mathbb{Z}_N$.
$K = S_N$.
Salausfunktio $E(x) = \sigma(x)$.
Avausfunktio $D(y) = \sigma^{-1}(y)$.

Joukon S_N alkioiden lukumäärä on $N!$, joten avaimia on runsaasti. Eräs tapa välittää avain on käyttää avainsanaa (sana tai lyhyt lause, josta jätetään toistuvat kirjaimet pois).

Avainsana Caesar: avain $k = (a,\text{avainsana})$, $0 \le a < N$. Olkoon $a = 4$ ja avainsana "kesä meni"

$$\begin{array}{lcccccccccccccccccc}
 & & & & \overset{4}{D} & & & & & & & & & & & & & & \\
\sigma\downarrow & A & B & C & D & E & F & G & H & I & J & K & L & M & N & O & P & Q & R & S \\
 & Y & Z & O & K & E & S & Ä & M & N & I & A & B & C & D & F & G & H & J & L \\
 & T & U & V & X & Y & Z & Ä & Ö & \uparrow \sigma^{-1} & & & & & & & & & & \\
 & O & P & Q & R & T & U & V & X & & & & & & & & & & &
\end{array}$$

$$\begin{array}{lcccccccccccccccccc}
 & & & & & & & & & & & & & & & & & & & \\
\overset{selvä}{\downarrow \sigma} & V & I & D & E & O & Y & H & T & E & Y & S & & L & U & E & N & T & O & \overset{selvä}{\uparrow \sigma^{-1}} \\
\overset{sala}{} & Q & N & K & E & F & T & M & O & E & T & L & & B & P & E & D & O & F & \overset{sala}{}
\end{array}$$

Yksinkertainen sijoitusjärjestelmä on murrettavissa kirjainten esiintymistiheyksiä tutkimalla, koska kirjaimen x kuva $E(x)$ on sama koko ajan. Tästä johtuen on kehitetty myös ns. moniaakkosjärjestelmiä, joista esimerkkinä tarkastelemme Vigenéren järjestelmää.

Vigenére järjestelmän avain koostuu useammasta Caesar järjestelmän avaimesta, joita sovelletaan jaksollisesti.
Oletetaan, että $k_1, k_2, ..., k_r \in \mathbb{Z}_N$. Jaetaan selväkielinen teksti r:n pituisiin osiin $x_1, x_2, ..., x_r$. Kunkin osan i:s kirjain x_i salakirjoitetaan Caesar salausfunktion E_{k_i} avulla, ts,

selvä sala
$$x_i \to E_{k_i}(x_i) = x_i + k_i = y_i.$$

Vastaava avausfunktio on $D_{k_i}(y_i) = y_i - k_i$. Avaimet k_i annetaan usein avainsanan avulla.

Esimerkki. Suomenkielinen aakkosto $\mathbb{Z}_{27}$.

Salasana ÖLJY $\Rightarrow r = 4$ ja $k_1 = 26, k_2 = 11, k_3 = 9, k_4 = 23$.

Tässä järjestelmässä sama kirjain kuvautuu eri kirjaimiksi paikasta riippuen, joten yksinkertainen kirjainten esiintymistiheyteen perustuva analyysi ei toimi. Kuitenkin, jos avaimen pituus r selviää, voidaan kunkin k_i murtaa erikseen (kuten Caesarissa). Avaimen pituuden määrittämiseksi on olemassa menetelmiä.

2. Salakirjoitus matriiseilla

Jos viestiyksikön pituus on $r > 1$ kirjainta, on luontevaa tarkastella viestiyksikköjä $\mathbb{Z}_N^r$:n vektoreina ja käyttää salauksessa $r \times r$ matriiseja. Rajoitumme seuraavassa tapaukseen $r = 2$.

Kertaamme lyhyesti matriisien laskusääntöjä:

$\mathbb{Z}_N^2$:n alkiot ovat $\begin{pmatrix} x \\ y \end{pmatrix}$; 2×2 matriisit $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

- yhtäsuuruus $\Leftrightarrow$ samoilla paikolla olevat alkiot yhtäsuuria,

- yhteenlasku $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x+u \\ y+v \end{pmatrix}$; $\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} x & u \\ y & v \end{pmatrix} = \begin{pmatrix} a+x & b+u \\ c+y & d+v \end{pmatrix}$,

- skalaarilla kertominen $a \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax \\ ay \end{pmatrix}, u \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} au & bu \\ cu & du \end{pmatrix}$,

- kertolasku $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$,

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & u \\ y & v \end{pmatrix} = \begin{pmatrix} ax+by & au+bv \\ cx+dy & cu+dv \end{pmatrix}$, ei vaihdannainen,

- yksikkömatriisi $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ on kertolaskun neutraalialkio,

- käänteismatriisi: Matriisin $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ käänteismatriisi on matriisi A^{-1}, jolle

$$AA^{-1} = A^{-1}A = I,$$

mikäli tällainen matriisi on olemassa. Voidaan osoittaa, että A^{-1} on olemassa jos ja vain jos $D = ad - bc \in \mathbb{Z}_N^*$, ja tällöin

$$A^{-1} = D^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Esimerkki $\mathbb{Z}_{26}$, $A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}$. Määritä A^{-1}.

Lause 1. Jos $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $a, b, c, d \in \mathbb{Z}_N$, $D = ad - bc \in \mathbb{Z}_N^*$,

ja $B = \begin{pmatrix} e \\ f \end{pmatrix} \in \mathbb{Z}_N^2$, niin affiini kuvaus

$$E : \mathbb{Z}_N^2 \to \mathbb{Z}_N^2 \ ; \ E(X) = AX + B,$$

on bijektio, jonka käänteiskuvaus on

$$D : \mathbb{Z}_N^2 \to \mathbb{Z}_N^2 \ ; \ D(Y) = A^{-1}(Y - B).$$

Matriisisalakirjoitus:

$P = C = \mathbb{Z}_N^2 = \{\begin{pmatrix} x \\ y \end{pmatrix},\ x, y \in \mathbb{Z}_N\}$.

Avain $k = \{A, B\},\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ D = ad - bc \in \mathbb{Z}_N^*,\ B = \begin{pmatrix} e \\ f \end{pmatrix} \in \mathbb{Z}_N^2$.

Salausfunktio $E(X) = AX + B$.

Avausfunktio $D(Y) = A^{-1}Y - A^{-1}B$.

Huom. Jos valitsemme $A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, niin $E(X) = X + B$, joten kyseessä on Vigenére järjestelmä, missä $r = 2$. Vastaavasti saamme matriisisalakirjoituksen erikoistapauksena yleisen Vigenére järjestelmän, jos tarkastelemme $r \times r$ matriiseja.

Tarkastelemme nyt matriisisalakirjoituksen murtamista, kun $B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ja tunnemme kaksi paria selvätekstiä ja vastaavat salakirjoitukset. Tämä on mahdollista esimerkiksi niin, että sieppaaja onnistuu jotenkin lähettämään tekstiä tarkasteltavan kanavan kautta.

Olkoot tuntemamme selvätekstit P_1 ja P_2 ja vastaavat salakirjoitukset C_1 ja C_2. Tällöin

$C_1 = AP_1$ ja $C_2 = AP_2$

eli

$C = AP$, missä $P = (P_1, P_2)$ ja $C = (C_1, C_2)$.

Jos nyt P_1 ja P_2 on valittu niin, että P^{-1} on olemassa, niin saamme

$A = CP^{-1}$

ja järjestelmä on murrettu.

Tapauksessa $B \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ tarvitsemme murtamiseen vielä kolmannen parin P_3 ja C_3:

$C_1 = AP_1 + B,\ C_2 = AP_2 + B,\ C_3 = AP_3 + B$.

Tällöin pääsemme yo:n kaltaiseen tilanteeseen vähentämällä kolmannen yhtälön kahdesta ensimmäisestä:

$C_1 - C_3 = A(P_1 - P_3),\ C_2 - C_3 = A(P_2 - P_3)$.

Jos tästä saadaan A, niin sen jälkeen $B = C_1 - AP_1$.

IV JULKISEN AVAIMEN SALAKIRJOITUS
(PUBLIC KEY CRYPTOGRAPHY)

1. Yleinen periaate

Nykyaikaisissa tiedonvälitysjärjestelmissä perinteisillä salakirjoitusmenetelmillä on esimerkiksi seuraavat ongelmat:

- Avaimista sopiminen ja niiden välittäminen. Jos verkossa on n käyttäjää, tarvitaan $\binom{n}{2} = n(n-1)/2$ avainta. Jos joku käyttäjä haluaa vaihtaa avaimensa (tai uusi käyttäjä liittyy verkkoon), tarvitaan $n-1$ (tai n) sopimusta uusista avaimista.

-Allekirjoitusongelma, koska normaali allekirjoitus on korvattava jotenkin.

Aikaisemmin, kun salausta käytettiin lähinnä sotilaallisissa tai diplomaattisissa tarkoituksissa, nämä seikat eivät tuottaneet suurta ongelmaa.

Vuonna 1976 Diffie ja Hellman esittivät ajatuksen julkisen avaimen järjestelmistä, joissa yo. ongelmat on selvitetty. Tällaisessa järjestelmässä kukin käyttäjä $\mathcal{U}$ (user) muodostaa oman salausmenettelynsä $E_{\mathcal{U}}$ ja avausmenettelynsä $D_{\mathcal{U}}$, jotka toteuttavat ehdon

$(JA\ 1)\quad D_{\mathcal{U}}(E_{\mathcal{U}}(m)) = m \ \ \forall m \in P$ (selväkieliset viestiyksiköt).

Kukin käyttäjä $\mathcal{U}$ julkaisee salausmenettelynsä $E_{\mathcal{U}}$ avainkirjassa, joka on kaikkien käytettävissä. Avausmenettely $D_{\mathcal{U}}$ pidetään vain $\mathcal{U}$:n omana tietona.

Jos käyttäjä A haluaa lähettää selväkielisen viestiyksikön m käyttäjälle B, hän etsii avainkirjasta B:n salakirjoitusmenettelyn E_B ja salakirjoittaa sanomansa sen avulla, ts. $c = E_B(m)$. Kun B saa salakirjoitetun viestin c, hän saa sanoman selville käyttämällä (salaista) avausmenettelyään D_B:

$$D_B(c) = D_B(E_B(m)) \underset{(JA\ 1)}{=} m.$$

Jotta menettely olisi toimiva ja salaisuus säilyisi, tarvitsemme seuraavat ehdot, joiden tulee olla voimassa kaikille käyttäjille $\mathcal{U}$:

$(JA\ 2)$ Menettelyt $E_{\mathcal{U}}$ ja $D_{\mathcal{U}}$ ovat nopeita eivätkä tarvitse liian paljon muistia.

$(JA\ 3)$ On käytännössä mahdotonta määrittää $E_{\mathcal{U}}$:n avulla menettelyä $D^*_{\mathcal{U}}$, jolle

$$D^*_{\mathcal{U}}(E_{\mathcal{U}}(m)) = m \ \ \forall m \in P.$$

Ominaisuus (JA 3) säilyttää järjestelmän salaisena ja mahdollistaa salausmenettelyn julkaisemisen avainkirjassa. Jos joku käyttäjistä vaihtaa avaimensa, riittää vaihtaa uusi $E_{\mathcal{U}}$ avainkirjaan. Uuden käyttäjän mukaantulo on yhtä yksinkertaista.

Em. menettely muodostetaan usein käyttämällä yksisuuntaista funktiota tai salaovifunktiota.

Määritelmä. Funktio f on yksisuuntainen, jos sen arvot ovat helposti laskettavissa ja käänteisfunktion f^{-1} (joka $\exists$) määrääminen on hyvin vaikeaa. Salaovifunktio on sellainen yksisuuntainen funktio, jonka käänteisfunktio on helppoa määrätä jonkin lisätiedon (salaovi) avulla.

Funktiota ei yleensä onnistuta todistamaan yksisuuntaiseksi, joten joudutaan käyttämään funktioita, joiden uskotaan olevan yksisuuntaisia. Tällaisten funktioiden muodostaminen perustuu usein vaikeisiin ja paljon tutkittuihin lukuteorian ongelmiin, esimerkiksi RSA-järjestelmässä suurten lukujen tekijöiden jakamisen vaikeuteen. Jos meillä on käytössämme salaovifunktioita $f_{\mathcal{U}}$, niin voimme muodostaa ehdot (JA 1), (JA 2) ja (JA 3) toteuttavan järjestelmän valitsemalla

$$E_{\mathcal{U}} = f_{\mathcal{U}} \text{ ja } D_{\mathcal{U}} = f_{\mathcal{U}}^{-1}.$$

Tarkastelemme nyt allekirjoitusongelmaa, jonka ratkaisemiseksi asetamme kaksi uutta vaatimusta, joiden tulee olla voimassa kaikille käyttäjille $\mathcal{U}$.

(JA 4) $E_{\mathcal{U}}(D_{\mathcal{U}}(c)) = c \; \forall c \in C$.

(JA 5) On käytännössä mahdotonta määrittää $E_{\mathcal{U}}$:n avulla menettelyä $D^*_{\mathcal{U}}$, jolle

$$E_{\mathcal{U}}(D^*_{\mathcal{U}}(c)) = c \; \forall c \in C.$$

Julkisen avaimen järjestelmässä on usein $P = C$ sama kaikilla käyttäjillä. Jos A lähettää B:lle selväkielisen viestin, hän lähettää allekirjoituksen s muodossa $m = D_A(s)$, jolloin B etsii avainkirjasta E_A:n ja laskee

$$E_A(m) = E_A(D_A(s) \underset{(JA\ 4)}{=} s.$$

Allekirjoituksen muodostaa pari $(s, D_A(s) = m)$. Menettely edellyttää ehtoja (JA 2), (JA 4) ja (JA 5), erityisesti ehto (JA 5) varmistaa, että vain A voi toimia lähettäjänä. A ei voi myöskään jälkikäteen kieltää viestiä.

$$\boxed{s} \overset{A}{\longrightarrow} \boxed{D_A(s) = m} \overset{m}{\longrightarrow} \boxed{m} \overset{B}{\longrightarrow} \boxed{E_A(m) = s}$$

$$\uparrow E_A$$

Allekirjoitus on pari $(s, D_A(s) = m)$ $\qquad \boxed{Avainkirja}$

Eo. menettely voidaan toteuttaa salaovifunktiolla f valitsemalla $E_A = f,\ D_A = f^{-1}$.

Jos halutaan suorittaa salakirjoitus ja allekirjoitus, tarvitaan ominaisuudet (JA 1) – (JA 5). Tällöin A muodostaa allekirjoitusosasta s tekstin $c = E_B(D_A(s))$. B soveltaa tähän funktiota $E_A \circ D_B$:

$$E_A(D_B(c)) = E_A(D_B(E_B(D_A(s)))) = E_A(D_A(s)) = s.$$

E_A löytyy avainkirjasta, mutta vain B tuntee D_B:n ja voi määrittää s:n. B käyttää A:n allekirjoituksena paria $(s, D_B(c) = D_A(s))$ kuten edellä.

$$\overset{A}{\boxed{s}} \longrightarrow \boxed{E_B(D_A(s)) = c} \longrightarrow \boxed{c} \xrightarrow{B} \boxed{E_A(D_B(c)) = s}$$
$$\uparrow E_B \qquad\qquad\qquad \uparrow E_A$$
$$\longrightarrow \boxed{Avainkirja} \longrightarrow$$

Koska julkisen avaimen salakirjoitus vaatii yleensä paljon enemmän aikaa kuin perinteiset menetelmät, sitä käytetään usein perinteisen menetelmän ohella avainten vaihdossa, ts. käytettävän perinteisen menetelmän avaimet vaihdetaan julkisen avaimen menetelmällä.

2. RSA-menetelmä

(1978, Rivest, Shamir ja Adleman)

RSA-menetelmässä salaovifunktion muodostaminen perustuu vuosisatoja tutkittuun lukuteorian ongelmaan: Kuinka annettu (suuri) luku n jaetaan alkutekijöihin? Yksisuuntaisen funktion perusta on nyt se, että annettujen lukujen tulo on nopeasti laskettavissa, mutta tekijöiden löytäminen tulosta vaatii nopeiltakin koneilta liikaa aikaa.

Olkoot nyt p ja q kaksi alkulukua ja $n = pq$. Tällöin $\varphi(n) = (p-1)(q-1)$. Valitsemme nyt sellaisen luvun e, $1 < e < \varphi(n)$, että $\text{syt}(e, \varphi(n)) = 1$ (mikä tahansa alkuluku e väliltä $max\{p, q\} < e < \varphi(n)$ käy). Eukleideen algoritmin avulla löydämme ehdon $1 < d < \varphi(n)$ toteuttavan luvun d, jolle $ed \equiv 1 \pmod{\varphi(n)}$. Tällöin $ed = 1 + \ell\varphi(n)$ eräällä $\ell \in \mathbb{N}$.

Eulerin lauseeseen nojautuen saamme seuraavan lauseen.

Lause 1. Kaikilla $a \in \mathbb{Z}_n$ on $a^{ed} = a$.

Olkoon nyt $f : \mathbb{Z}_n \to \mathbb{Z}_n$, $f(a) = a^e$, ja $g : \mathbb{Z}_n \to \mathbb{Z}_n$, $g(a) = a^d$. Tällöin $g = f^{-1}$ Lauseen 1 perusteella ja molempien funktioiden arvot ovat nopeasti laskettavissa potenssiin korotuksella (mod n), kun e ja d tunnetaan. Jos nyt n ja e tunnetaan, niin $f(a)$ on laskettavissa, mutta d:n laskeminen ei onnistu ilman lukua $\varphi(n)$, jonka löytäminen edellyttää p:n ja q:n tuntemista eli n:n tekijöihinjakoa. Luku d on siis salaovi, jota ilman f^{-1} ei löydy, ts. f on salaovifunktio.

RSA-järjestelmässä kukin käyttäjä $\mathcal{U}$ valitsee yo:n tapaan alkuluvut $p_\mathcal{U}$ ja $q_\mathcal{U}$, laskee luvun $n_\mathcal{U} = p_\mathcal{U} q_\mathcal{U}$, valitsee sitten luvun $e_\mathcal{U}$, $1 < e_\mathcal{U} < \varphi(n_\mathcal{U})$, $e_\mathcal{U} d_\mathcal{U} \equiv 1 \pmod{\varphi(n_\mathcal{U})}$.

Avainkirjassa kukin käyttäjä julkaisee avaimensa $K_\mathcal{U} = (n_\mathcal{U}, e_\mathcal{U})$, mutta pitää omana tietonaan luvun $d_\mathcal{U}$ (ja luvut $p_\mathcal{U}, q_\mathcal{U}$), joka muodostaa salaoven.

Kun käyttäjä A haluaa lähettää viestin käyttäjälle B, hän muuntaa viestinsä $\mathbb{Z}_{n_B}$:n alkioiksi (seuraavassa esitettävällä tavalla) ja toimii eo. yleisen periaatteen mukaisesti.

Alkion $m \in \mathbb{Z}_{n_B}$ salakirjoitus on

$$E_B(m) = m^{e_B} = c \in \mathbb{Z}_{n_B}.$$

Vain B tietää luvun d_B, joten hänen avausfunktionsa on

$$D_B(c) = c^{d_B} = m^{e_B d_B} = m.$$

Allekirjoitus onnistuu myös yleisen periaatteen mukaisesti, A lähettää allekirjoituksensa s muodossa $D_A(s) = s^{d_A} = m \in \mathbb{Z}_{n_A}$. Pari (s, m) muodostaa allekirjoituksen.

Jos halutaan salaus ja allekirjoitus, niin voimme toimia yleisen periaatteen mukaisesti, jos $n_A < n_B$.

Tällöin A laskee ensin $s^{d_A} (\text{mod } n_A) = m$ ($a(\text{mod } n) = x,\ 0 \leq x < n,\ x \equiv a(\text{mod } n)$) ja sitten $m^{e_B} (\text{mod } n_B) = c$, ts. salattu allekirjoitus on

$$c = (s^{d_A} (\text{mod } n_A))^{e_B} (\text{mod } n_B).$$

Saatuaan c:n B laskee ensin

$$c^{d_B} (\text{mod } n_B) = m^{e_B d_B} (\text{mod } n_B) = m$$

ja sitten

$$m^{e_A} (\text{mod } n_A) = s^{d_A e_A} (\text{mod } n_A) = s.$$

Tämä menettely ei toimi, jos $n_B < n_A$, koska tällöin voisi $m(\text{mod } n_B)$ olla sama useammalla arvolla m. Jos $n_B < n_A$, vaihdetaan järjestystä: A laskee ensin $s^{e_B} (\text{mod } n_B) = m_1$ ja sitten $m_1^{d_A} (\text{mod } n_A) = c$ eli

$$c = (s^{e_B} (\text{mod } n_B))^{d_A} (\text{mod } n_A).$$

Vastaavasti B avaa c:n laskemalla ensin $c^{e_A} (\text{mod } n_A) = m_1$, ja sitten $m_1^{d_B} (\text{mod } n_B) = s$.

3. Tekstin esittäminen $\mathbb{Z}_n$:n alkioina

Edellä olemme muuntaneet N:n symbolin aakkoston yleensä luvuiksi $0, 1, ..., N-1$, jotka voidaan tulkita myös $\mathbb{Z}_N$:n alkioiksi. Esimerkiksi RSA järjestelmässä toimitaan joukossa $\mathbb{Z}_n$, missä n on suuri. Jos jaamme N-aakkostolla kirjoitetun tekstin k-pituisiin osiin $a = a_1, a_2, ..., a_k,\ 0 \leq a_i < N$, voidaan jokainen tällainen osa esittää kääntäen yksikäsitteisesti N-kantaisessa järjestelmässä muodossa

$$a = a_1 N^{k-1} + a_2 N^{k-1} + ... + a_k.$$

Nyt $a \leq N^k - 1$, joten jos valitsemme k:n niin, että $N^k \leq n$, niin N-aakkoston k pituiset sanat voidaan esittää kääntäen yksikäsitteisesti $\mathbb{Z}_n$:n osajoukkona luonnollisella tavalla. Jos taas $n \leq N^\ell$, niin jokaista $\mathbb{Z}_n$:n alkiota vastaa eräs N-aakkoston ℓ-pituinen sana.

Jos siis käytämme RSA-järjestelmää ja jokainen käyttäjä $\mathcal{U}$ valitsee lukunsa $n_\mathcal{U}$ niin, että

$$N^k \leq n_\mathcal{U} \leq N^\ell,\ k, \ell \in \mathbb{N},$$

niin kaikki käyttäjät voivat valita

$$P = \{k-\text{kirjaimiset sanat}\} \subset \mathbb{Z}_{n_{\mathcal{U}}}$$

$$C = \mathbb{Z}_{n_{\mathcal{U}}} \subset \{\ell\text{-kirjaimiset sanat}\}.$$

Esimerkki.

4. Diskreetti logaritmi

Reaalilukujen joukossa lukujen b^x ja $\log_b x$ laskeminen tietyllä tarkkuudella on yhtä helppoa. Tarkastelemme nyt vastaavaa tilannetta joukossa $\mathbb{Z}_n^*$. Jos $b \in \mathbb{Z}_n^*$, niin b^x on laskettavissa nopeasti suurillakin $x \in \mathbb{N}$. Oletamme nyt, että tiedämme alkion $y \in \mathbb{Z}_n^*$ olevan muotoa $y = b^x$. Kuinka $x = \log_b y$ saadaan selville (tässä logaritmi on ns. diskreetti logaritmi, ei $\mathbb{R}$:n logaritmifunktio)? Tälle ns. "diskreetin logaritmin ongelmalle" ei tunneta nopeaa ratkaisua yleisessä tapauksessa, joten voimme jälleen rakentaa yksisuuntaisen funktio.

Määritelmä. Luku g, $1 \leq g \leq n-1$, on primitiivijuuri (mod n), jos

$$\mathbb{Z}_n^* = \{\bar{g}^k | k = 0, 1, ..., \varphi(n)-1\} = \{\bar{1}, \bar{g}, \bar{g}^2, ..., \bar{g}^{\varphi(n)-1}\}.$$

Esimerkki.

Lause 2. Primitiivijuuri (mod n) on olemassa jos ja vain jos n on

$$2, 4, p^\ell, 2p^\ell, \ \ell = 1, 2, ..., \text{ missä } p > 2 \text{ on alkuluku.}$$

Jos g on primitiivijuuri (mod n) (n on $\mathcal{L}2$:n muotoa), niin sen määräämää jäännösluokkaa sanotaan $\mathbb{Z}_n^*$:n primitiiviseksi alkioksi.

Määritelmä. Olkoon $g \in \mathbb{Z}_n^*$ primitiivinen. Alkion $y \in \mathbb{Z}_n^*$ diskreetti logaritmi kannan g suhteen on sellainen luku $k \in \{0, 1, ..., \varphi(n)-1\}$, jolle $y = g^k$. Tällöin merkitsemme $k = \log_g y$.

Usein valitsemme edellä $n = p$ (alkuluku), jolloin $\mathbb{Z}_p$ on äärellinen kunta.

5. Diffie-Hellman avaimenvaihto

Oletamme tässä, että joukossa $\mathbb{Z}_n^*$ on primitiivinen alkio g. Tiedonsiirtojärjestelmän avaimen vaihtomenettely voidaan hoitaa seuraavasti:

Kaikki käyttäjät tuntevat luvun n ja primitiivialkion g. Kukin käyttäjä $\mathcal{U}$ valitsee luvun $m_{\mathcal{U}} \in \{1, ..., \varphi(n)-1\}$ ja laskee $g^{m_{\mathcal{U}}}$, jonka hän julkaisee. Käyttäjien A ja B keskinäisen yhteydenpidon avaimen määrää $g^{m_A m_B}$.

Käyttäjä A saa avaimen potenssiinkorotuksella $(g^{m_B})^{m_A} = g^{m_A m_B}$ ja käyttäjä B vastaavasti $(g^{m_A})^{m_B} = g^{m_A m_B}$.

Sieppaaja tuntee vain alkiot g^{m_A} ja g^{m_B}. Jos hän ratkaisisi diskreetin logaritmin ongelman, ts. m_A:n tai m_B:n, avain löytyisi. Muuten sen löytäminen ei ilmeisesti onnistu ($g^{m_A} g^{m_B} = g^{m_A+m_B}$).

Esimerkki.

6. El Gamal salakirjoitusjärjestelmä

Olkoon $b \in \mathbb{Z}_p^*$ annettu (yleensä valitaan b primitiivialkioksi, mutta se ei ole välttämätöntä). Kaikki käyttäjät tuntevat p:n ja b:n. Lisäksi oletamme, että $P = C = \mathbb{Z}_p^*$ (tai jotkin osajoukot).

Kukin käyttäjä $\mathcal{U}$ valitsee luvun $m_{\mathcal{U}} \in \{1, 2, ..., p-2\}$ ja pitää sen omana tietonaan. Julkiseen avainkirjaan annetaan $b^{m_{\mathcal{U}}} \in \mathbb{Z}_p^*$.

Jos joku käyttäjä haluaa lähettää viestin $m \in \mathbb{Z}_p^*$ käyttäjälle A, hän valitsee $k \in \mathbb{N}$ ja lähettää parin

$$(b^k, mb^{m_A k}) \in \mathbb{Z}_p^{*^2}.$$

Koska $b^{m_A k} = (b^{m_A})^k$ ja b^{m_A} on avainkirjassa, on pari laskettavissa nopeasti potenssiinkorotuksella. Käyttäjä A (ja vain hän) tuntee luvun m_A, joten hän laskee parin ensimmäisen komponentin avulla $(b^k)^{m_A} = b^{m_A k}$. Jakamalla näin saadulla alkiolla toisen komponentin A saa

$$\frac{mb^{m_A k}}{b^{m_A k}} = m.$$

Jos sieppaaja ratkaisee diskreetin logaritmin ongelman, on murtaminen helppoa. Muuten on ilmeisesti vaikea saada $b^{m_A k}$ alkioista b^k ja b^{m_A}.

Edellä viesti m on naamioitu alkion $b^{m_A k}$ avulla ja ensimmäinen komponentti on johtolanka b^k, jonka avulla A (ja vain hän) voi riisua naamion.

7. Selkäreppuongelma

Oletetaan, että k positiivista kokonaislukua sisältävä joukko $\{v_1, v_2, ..., v_k\}$ ja positiivinen kokonaisluku V on annettu. Kysytään, onko olemassa sellainen osajoukko $\mathcal{I} \subset \{1, 2, ..., k\}$, että

$$\sum_{i \in \mathcal{I}} v_i = V.$$

(Jos sekkärepun tilavuus on V ja on k esinettä, joiden tilavuudet ovat $v_1, v_2, ..., v_k$, niin voidaanko reppu pakata näillä esineillä tasan täyteen?) Tämä kysymys on ns. selkäreppuongelma. Se voidaan muotoilla myös seuraavasti:

Selkäreppuongelma. Olkoot $v_1, v_2, ..., v_k$ ja V annettuja positiivisia kokonaislukuja. Onko olemassa sellaiset luvut $\varepsilon_i \in \{0, 1\}, i = 1, ..., k$, että

$$\varepsilon_1 v_1 + \varepsilon_2 v_2 + ... + \varepsilon_k v_k = V?$$

Esimerkki.

Jos k on suuri, niin on osoitettu, että selkäreppuongelma on yleisessä tapauksessa hyvin vaikea ratkaista. Eräissä erikoistapauksissa ongelma on kuitenkin helppo ratkaista.

Määritelmä. k-jono $v_1, v_2, ..., v_k$ on superkasvava, jos $\forall\ j = 2, ..., k$

$v_j > v_1 + ... + v_{j-1}$.

Vastaava selkäreppua sanotaan myös superkasvavaksi.

Esimerkki.

Superkasvava selkäreppu on helppo ratkaista seuraavasti: $\{v_1, v_2, ..., v_k\}$ superkasvava, V.

Jos $v_1 + v_2 + ... + v_k < V$, ratkaisua ei ole. Oletamme, että $V \leq v_1 + v_2 + ... + v_k$. Mahdollisessa ratkaisussa $\varepsilon_1, \varepsilon_2, ..., \varepsilon_k$ on

$\varepsilon_k = 1$, jos ja vain jos $V \geq v_k$;

tämän jälkeen saamme rekursiivisesti kaikilla $j = k-1, k-2, ..., 1$

$\varepsilon_j = 1$ jos ja vain jos $V - (\varepsilon_k v_k + ... + \varepsilon_{j+1} v_{j+1}) \geq v_j$.

Näin saamme luvut $\varepsilon_1, \varepsilon_2, ..., \varepsilon_k \in \{0, 1\}$. Jos $V = \varepsilon_1 v_1 + \varepsilon_2 v_2 + ... + \varepsilon_k v_k$, niin selkäreppu ratkeaa ja $\varepsilon_1, \varepsilon_2, ..., \varepsilon_k$ on ratkaisu, muuton selkäreppu ei ratkea.

Esimerkki.

8. Selkäreppujärjestelmä

Sovitaan, että selväkieliset viestiyksiköt ovat binäärilukuja, joiden pituus on $\leq k$ (tähän päästään käyttämällä aakkostolle lukuvastineita ja muuntamalla tiettyä kirjainmäärää vastaava luku binääriluvuksi). Tyypillinen selväkielinen viestiyksikkö on

$m = (m_1 m_2 ... m_k)_2 = m_1 2^{k-1} + m_2 2^{k-1} + ... + m_k,\ m_i \in \{0, 1\}$.

Esimerkki.

Kukin käyttäjä $\mathcal{U}$ valitsee superkasvavan k-jonon $u_1, u_2, ...u_k$ ja sellaiset kokonaisluvut $n_\mathcal{U}$ ja $a_\mathcal{U}$, että

(1) $n_\mathcal{U} > u_1 + u_2 + ... + u_k,\ 1 < a_\mathcal{U} < n_\mathcal{U}$, syt$(a_\mathcal{U}, n_\mathcal{U}) = 1$.

Tämän jälkeen $\mathcal{U}$ määrittää (Eukleideen algoritmilla) luvun $b_\mathcal{U}$, jolle

(2) $a_\mathcal{U} b_\mathcal{U} \equiv 1 \pmod{n_\mathcal{U}}$ (ts. $a_\mathcal{U}^{-1} = b_\mathcal{U}$ $\mathbb{Z}_{n_\mathcal{U}}^*$:ssä).

Seuraavaksi käyttäjä $\mathcal{U}$ laskee k-jonon

$u_i^* = a_\mathcal{U} u_i (\text{mod } n_\mathcal{U}),\ i = 1, 2, ..., k.$

$\mathcal{U}$-pitää luvut $u_i,\ n_\mathcal{U},\ a_\mathcal{U}$ ja $b_\mathcal{U}$ vain omana tietonaan, mutta julkaisee k-jonon $u_1^*, u_2^*, ..., u_k^*$ avaimenaan, ts. $K_\mathcal{U} = \{u_1^*, u_2^*, ..., u_k^*\}$. Jos joku haluaa lähettää $\mathcal{U}$:lle viestin $m = (m_1 m_2 ... m_k)_2$, hän käyttää salausfunktiota

$E_\mathcal{U}(m) = m_1 u_1^* + m_2 u_2^* + ... + m_k u_k^* = c$,

joka on helppoa laskea avaimesta $K_{\mathcal{U}}$. Avausmenettely $D_{\mathcal{U}}$ koostuu kahdesta osasta. Salaoven muodostaa luku $b_{\mathcal{U}}$, jonka $\mathcal{U}$ tuntee. Ensimmäisessä vaiheessa $\mathcal{U}$ laskee (2):n avulla

$$b_{\mathcal{U}}c \equiv b_{\mathcal{U}}(m_1 a_u u_1 + m_2 a_{\mathcal{U}} u_2 + ... + m_k a_{\mathcal{U}} u_k)$$
$$\equiv m_1 u_1 + m_2 u_2 + ... + m_k u_k (\text{mod } n_{\mathcal{U}}).$$

Ehdon (1) mukaisesti yo. kongruenssista seuraa yhtälö

$$m_1 u_1 + m_2 u_2 + ... + m_k u_k = b_{\mathcal{U}} c \ (\text{mod } n_{\mathcal{U}}).$$

Toisessa vaiheessa ratkaistaan tämä superkasvava selkäreppu ja saadaan $m = (m_1 m_2 ... m_k)_2$.

Kun A haluaa lähettää sanoman $m = (m_1 m_2 ... m_k)_2$ B:lle, toimitaan seuraavasti:

Menettelyn idea on luonnollisesti siinä, että B:n superkasvava jono $b_1, b_2, ..., b_k$ sotketaan kertomalla a_B:llä B:n avainjonoksi $b_1^*, b_2^*, ..., b_k^*$, joka ei ole superkasvava. Sieppaajan tulisi salakirjoituksen murtamiseksi ratkaista tämän jonon avulla muodostettu yleiseltä näyttävä selkäreppuongelma $m_1 b_1^* + m_2 b_2^* + ... + m_k b_k^* = c$.

Yo:n järjestelmän esittivät vuonna 1978 Merkle ja Hellman ja se saavutti yksinkertaisuutensa ja nopeutensa takia suuren suosion. Kuitenkin Shamir onnistui murtamaan menetelmän jo vuonna 1982. Myös menetelmän parannuksia on murrettu, joten selkäreppujärjestelmää ei pidetä kovin luotettavana.

9. Huomautuksia

Edellä emme ole käsitelleet esiteltyjen julkisen avaimen järjestelmien heikkouksia, esim. murtamismahdollisuuksia. RSA:ta pidetään varsin turvallisena, kun p ja q ovat riittävän suuria (≈ 150 numeroa) sekä niiden valinnassa otetaan huomioon eräitä rajoituksia. Samoin diskreetti logaritmi lienee luotettava, kun kunta $\mathbb{Z}_p$ (tai yleisempi äärellinen kunta) on riittävän suuri. Voimme myös kysyä täsmällisempää tietoa eri menetelmien vaatimien laskutoimitusten määristä (tarvittavasta ajasta). Luonnollisesti tulee esille myös kysymys siitä, miten esim. RSA:n tarvitsemia suuria alkulukuja on löydettävissä. Syventävien opintojen kurssilla "Kryptografia" tarkastellaan lähemmin näitä kysymyksiä sekä esitellään menetelmiä, jotka vaativat syvällisempiä matematiikan tietoja.

HARJOITUSTEHTÄVÄT:

1. Ystäväsi K lähettää sinulle Caesarin yhteenlaskumenetelmällä kirjoitetun viestin

   "$\ddot{O}\ H\ X\ H\ H\ T\ T\ L\ O\ H\ U\ P\ S\ S\ H\ S\ S\ H\quad R$".

   Avaa viesti.

2. Avaa seuraava Caesarin yhteenlaskumenetelmällä laadittu englanninkielinen salakirjoitus

   $A\ L\ Y\ U\ N\ Y\ M\ N\ W\ I\ G\ G\ I\ H\ X\ C\ P\ C\ M\ I\ L.$

   Mikä on ollut avain?

3. Jaa luvut

   211, 212, 213, 721

   alkutekijöihin. Määritä myös lukujen binääriesitykset.

4. Esitä luku 1995

   a) 5-kantaisessa,

   b) 8-kantaisessa,

   c) 32 kantaisessa lukujärjestelmässä.

5. Laske $212_3 \cdot 122_3$.

6. Tarkastellaan 26-kantaista järjestelmää, missä englanninkieliset aakkoset $A - Z$ esittävät numeroita 0-25. Laske $(YES) \cdot (NO)$.

7. Määrää lukujen

   a) 101, 3040

   b) 1690, 650

   suurin yhteinen tekijä.

8. Etsi sellaiset kokonaisluvut x ja y, että

   $213x - 89y = 1$,

   $1 \leq x, y \leq 212$.

9. Todista jaollisuussäännöt luvuille 2, 3, 5, 9 ja 11.

10. Ratkaise kongruenssit
    a)$3x \equiv 4 \pmod 7$,
    b) $14x \equiv 1 \pmod{27}$,
    c) $2x \equiv 1 \pmod p$,
    missä $p \geq 3$ on alkuluku.

11. Ratkaise kongruenssit
    a) $5x \equiv 1 \pmod 7$ ja $\pmod 5$,
    b) $3z \equiv 3 \pmod 3$ ja $\pmod 9$,
    c) $4x^2 \equiv 2 \pmod 7$.

12. Kirjoita yhteen- ja kertolaskutaulut joukoille $\mathbb{Z}_7$ ja $\mathbb{Z}_8$.

13. Laske a^{-1}, kun
    a) $a = 15 \in \mathbb{Z}_{17}$,
    b) $a = 16 \in \mathbb{Z}_{19}$,
    c) $a = 3 \in \mathbb{Z}_9$.

14. Laske
    $4^{10} \pmod 5$, $5^{101} \pmod{25}$, $101^{101} \pmod 3$, $101^{101} \pmod 9$,
    $15^{55} \pmod{91}$ ja $2^{1000000} \pmod{77}$.

15. Käytetään suomenkielistä aakkostoa täydennettynä tyhjällä välillä (=27) ja affiinia järjestelmää, jonka avain on (5,24). Salakirjoita teksti "$TULKAA\ APUUN$". Mikä on avausfunktio?

16. Seuraava teksti on muodostettu affiinilla järjestelmällä:
    $$VCLMPCAVESVFYDVOVIZHPCYXAXGBDATYZ.$$
    Mikä on selväkielinen teksti?

17. Tehtäväsi on selvittää affiinilla järjestelmällä tehty salakirjoitus, jossa käytetään 37-kirjaimista aakkostoa, jonka kirjaimet ovat luvut 0-9 (vastaavat lukuja 0-9), englan-

ninkieliset aakkoset $A-Z$ (vastaavat lukuja 10-35) ja tyhjä (=36). Salakirjoitusteksti on

$$OH7F86BB46R3627O266BB9.$$

Tiedät, että sanoman on lähettänyt ”007”. Mikä on sanoma?

18. Ovatko funktiot

    a) $\sigma(i) = i + 17$,

    b) $\sigma(i) = 17i$

    permutaatioita joukossa $\mathbb{Z}_{26}$?

19. Laske Caesarin yhteenlasku- ja kertolaskumenetelmien sekä affiinin järjestelmän suhteelliset osuudet kaikista yksinkertaisista sijoitusjärjestelmistä, kun N=26, 27, 28 ja yleisessä tapauksessa.

20. Käytetään avainsana Caesaria, jonka avain on (5,$SYYSKUU$). Salakirjoita teksti $SYKSY\ ON\ TULLUT$.

21. Salakirjoita teksti

    $$DOES\ SURJECTION\ BECOME\ INJECTION$$

    Vigenéren järjestelmällä, kun avainsana on $INJECTION$.

22. Valitse jokin 2. pituinen avain Vigenéren järjestelmässä. Salakirjoita jokin 25-30 merkkiä pitkä teksti. Anna se toiselle harjoitusryhmän jäsenelle murrettavaksi.

23. Seuraavien matriisien alkiot ovat joukon $\mathbb{Z}_N$ alkioita. Määritä A^{-1}, kun

    a) $A = \begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix}, N = 5$; b) $A = \begin{pmatrix} 15 & 17 \\ 4 & 9 \end{pmatrix}, N = 26$; c) $A = \begin{pmatrix} 3 & 6 \\ 2 & 5 \end{pmatrix}, N = 28$.

24. Käytämme suomenkielistä aakkostoa, missä $A - \ddot{O}$ on 0-26 ja tyhjä=27, joten $N = 28$. Salakirjoita matriisisalakirjoituksella käyttämällä edellisen tehtävän c)-kohdan matriisia A teksti $YHTEYS\ TOIMII$.

25. Käytämme englanninkielistä aakkostoa, missä $A - Z$ on 0-25, tyhjä=26, ?=27 ja !=28, joten N=29. Saat matriisisalakirjoituksella (missä $B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$) tehdyn sanoman

$GFPYJP\ X?UYXSTLADPLW$

ja tiedät, että viisi viimeistä kirjainta on lähettäjän nimi $KARLA$. Avaa sanoma.

26. Oletetaan, että $n = pq$, missä p ja q ovat erisuuria alkulukuja. Osoita, että luvun $\varphi(n)$ tunteminen on yhtäpitävää tekijöiden p ja q tuntemisen kanssa.

27. Salakirjoita teksti $SELKÄREPPU$ RSA-järjestelmän avaimella $(n, e) = (91, 11)$. Murra tämä salausmenettely määräämällä salaovi d.

28. Laske tulot

    $117 \cdot 103,\ 7008 \cdot 6992$

    käyttäen kaavaa $(a + b)(a - b) = a^2 - b^2$. Esitä seuraavat luvut

    250997, 1689999

    ainakin kahden tekijän tulona.

29. a) Salakirjoita teksti TAKE IT AWAY RSA-järjestelmän avaimella (n, e)= (2773, 17) muuntamalla se ensin 2 kirjaimen osissa $\mathbb{Z}_n$:n alkioiksi.

    b) Murra a)-kohdan salausmenettely määräämällä salaovi d.

    c) Avaa a)-kohdan salakirjoitus ja totea, että saat alkuperäisen tekstin.

30. Tietoverkon käyttäjien A ja B RSA-avaimet ovat $(n_A, e_A) = (2773, 17)$ ja $(n_B, e_B) = (2047, 179)$. Missä muodossa A lähettää allekirjoituksensa AK B:lle? Entä salattu allekirjoitus? Selvitä samat kysymykset B:n lähettäessä allekirjoituksensa BG A:lle.

31. a) Määritä primitiivijuuri (mod n), merkitään g, kun n=3, 4, 5, 6, 7, 9, 10, 11.

    b) Laske tapauksissa n=5, 9 ja 11 $\mathbb{Z}_n^*$:ssä $\log_g(-k)$, kun k=1,2.

32. Määritä $\mathbb{Z}_{25}^*$:n primitiivinen alkio g ja laske $\log_g(-1)$, $\log_g 2$ ja $\log_g 3$.

33. Osoita, että 2 on kunnan $\mathbb{Z}_{37}$ primitiivinen alkio ja laske

    $\log_2$ 28, $\log_2$ 8, $\log_2$ (-10).

34. Olkoon n=37 ja g=2. Diffie-Hellman avaimenvaihdossa käyttäjän A salainen eksponentti m_A=18 ja käyttäjän B m_B=23. Mitkä ovat A:n ja B:n julkaisemat alkiot ja mikä on yhteinen avain?

35. Käytetään El Gamal salakirjoitusjärjestelmää, missä p=37 ja b=2. Salakirjoita viesti NYT käyttäjälle A, jonka salainen eksponentti on $m_A = 23$.

36. Olkoon alkuluku p=65537 ja b=5. Saat viestin (29095, 23846), jonka ystäväsi on lähettänyt käyttäen El Gamal salakirjoitusjärjestelmää $\mathbb{Z}_p$:ssä ja antamaasi julkista avainta 5^m, missä salainen eksponenttisi m=13908. Olette sopineet, että $\mathbb{Z}_p$:n alkiot muunnetaan 31-kirjaimisen aakkoston 3-pituisiksi jonoiksi esittämällä ne 31-kantaisessa muodossa (kirjaimet $A - Z$ ovat 0-25, tyhjä=26, .=27, ?=28, !=29 ja '=30). Avaa saamasi viesti.

37. Tutki, onko seuraava jono superkasvava ja ratkaise annettu selkäreppuongelma:
    a) $\{2, 3, 7, 20, 35, 69\}$, V=45;
    b) $\{4, 5, 10, 30, 50, 101\}$, V=186;
    c) $\{1, 2, 2^2, ..., 2^{k-1}\}$, $V < 2^k$, $k \in \mathbf{N} - \{0\}$.

38. a) Muodosta superkasvavasta 5-jonosta 2, 3, 7, 15, 31 käyttäjälle U julkinen avainjono, kun n_U=61 ja a_U=7.
    b) Salakirjoita teksti $YLLK$ (A-Ö vastaavat lukuja 0-26, jotka muunnetaan binääriluvuiksi).
    c) Avaa salakirjoitus 65, 141, 99, 99, 97.

39. Muodosta superkasvavasta 10-jonosta 1, 3, 5, 11, 21, 44, 87, 175, 349, 701 käyttäjälle U julkinen avainjono, kun n_U=1590 ja a_U=43. Salakirjoita teksti $KNAPSACK$.

40. Oletamme, että selvätekstin yksiköt ovat 3 kirjaimen jonoja 32 kirjaimisessa aakkostossa, missä $A - Z$ on 0-25, tyhjä=26, ?=27, !=28, .=29, '=30 ja \$=31. Olet antanut selkäreppujärjestelmän julkisen avainjonon 24038, 29756, 34172, 34286, 38334, 1824, 18255, 19723, 143, 17146, 35366, 11204, 32395, 12958, 6479, jonka muodostamisessa olet käyttänyt lukuja n=47107 ja a, missä $a^{-1} = b = 30966$. Avaa saamasi viesti
    152472, 116116, 68546, 165420, 168261.

41. Avaa salakirjoitus
    $KOKOOKOKOONKOKOKOKKOKOKOKOKKOKOKOKOKOKKO$.