Lukuteoriaa Diffie-Hellmania varten

- Eulerin lauseen mukaan kaikille n ja a, joille $\gcd(n, a) = 1$, pätee $a^{\phi(n)} \equiv 1 \pmod n$.

  Onko olemassa pienempiä i, joille $a^i \equiv 1 \pmod n$?

  Esimerkki: lukujen 2 ja 5 potenssit modulo 7.

i	1	2	3	4	5	6	7
$2^i \bmod 7$	2	4	1	2	4	1	2
$5^i \bmod 7$	5	4	6	2	3	1	5

  Huomataan, että 5 generoi kaikki luvut $1 \ldots 6$, mutta 2 vain osajoukon.

- **Määritelmä**: Olkoon p alkuluku. p:n redusoitu jäännösjoukko $\mathbb{Z}_n^* = \{1, \ldots, p-1\}$ on kertolaskuryhmä. Luku a on **generaattori** $g \bmod p$ (**primitiivinen alkio**, generoi ryhmän), joss

$$\{a^1 \bmod p,\ a^2 \bmod p, \ldots, a^p \bmod p\} = \{1, 2, \ldots, p-1\}$$

  eli kaikki $\mathbb{Z}_n^+$:n luvut voidaan esittää a:n potensseina.

  Esimerkki: 3 ja 5 ovat generaattoreja $\bmod 7$.

- Generaattoreiden tunnistaminen on yleisessä tapauksessa vaikeaa. Se on helppoa, jos tunnetaan $(p-1)$:n alkutekijät.

  Lause: Olkoon p alkuluku. a on generaattori $\bmod p$, joss kaikille $p-1$:n alkutekijöille q pätee

$$a^{(p-1)/q} \not\equiv 1 \pmod p.$$

Esimerkki: $7 - 1 = 6$. Alkutekijät ovat 2 ja 3.

$$5^{6/2} = 6 \pmod 7$$
$$5^{6/3} = 4 \pmod 7$$

Siis 5 on generaattori.

$$2^{6/2} = 1 \pmod 7$$
$$2^{6/3} = 4 \pmod 7$$

Siis 2 ei ole generaattori.

- Generaattorit löydetään valitsemalla satunnainen luku ja testaamalla. (Niitä on $\phi(p-1)$ kappaletta.)

- **Diskreetti eksponenttifunktio**

$$y = a^x \bmod n.$$

  Käänteisfunktio (x, kun a, n ja y tunnetaan) on vaikea laskea, kun n on alkuluku tai tekijöitä ei tunneta. Vaikeinta se on, jos a on generaattori modulo n.

  Tämä **diskreetin logaritmin** ongelma on suunnilleen yhtä vaikea kuin tekijöihin jako samansuuruisille luvuille.

- Jos tiedetään n:n tekijät p ja q, helpottuu logaritmin laskeminen:

$$a^x \equiv y \pmod{p\,q}$$

$$a^x + k\,p\,q = y$$

$$a^x \equiv y \pmod p \text{ ja } a^x \equiv y \pmod q$$

Diffie–Hellman-avaintenvaihto

- Sovitaan yhteisestä salaisesta avaimesta ilman, että ulkopuolinen saa sen selville.

  Ei edellytä aiempaa yhteistä salaisuutta.

- Turvallisuus riippuu diskreetin logaritmin ongelmasta.

- Tuloksena on ns. avainmateriaalia, josta avaimet yleensä tuotetaan yksisuuntaisella hajautusfunktiolla.

- Ensimmäinen julkisen avaimen protokolla 1976.

  Patentti rauennut vuonna 1997.

Diffie–Hellman-avaintenvaihto

- Alice ja Bob haluavat yhteisen salaisen avaimen. He sopivat suuresta alkuluvusta n ja generaattorista g modulo n.

1. Alice valitsee satunnaisen suuren luvun x ja lähettää Bobille

$$X = g^x \bmod n.$$

2. Bob valitsee satunnaisen suuren luvun y ja lähettää Alicelle

$$Y = g^y \bmod n.$$

3. Alice laskee

$$k = Y^x \bmod n.$$

4. Bob laskee

$$k = X^y \bmod n.$$

Molemmat laskevat saman avaimen arvon $k = g^{xy}$.

Varsinainen istuntoavain tuotetaan tästä yksisuuntaisella hajautusfunktiolla, esim. $h(k)$.

Diffie–Hellman-esimerkki

On sovittu moduliksi alkuluku $n = 43261$ ja generaattoriksi $g = 2$.

Alice valitsee satunnaisen suuren luvun $x = 37339$ ja lähettää Bobille luvun

$$X = 2^{37339} \bmod 43261 = 11255.$$

Bob valitsee satunnaisen suuren luvun $y = 5664$ ja lähettää Alicelle luvun

$$Y = 2^{5664} \bmod 43261 = 28841.$$

Alice laskee

$$k = 28841^{37339} \bmod 43261 = 9977.$$

Bob laskee

$$k = 11255^{5664} \bmod 43261 = 9977.$$

Diffie-Hellmannin turvallisuus

- Turvallisuus perustuu diskreetin logaritmin laskemisen vaikeuteen.

  Vastustaja saa tietoonsa vain luvut n, g, g^x ja g^y. Hän ei (oletettavasti) pysty selvittämään avainta g^{xy} laskematta diskreettiä logaritmia.

  Jos osaa laskea diskreettejä logaritmeja tehokkaasti, osaa myös murtaa Diffie-Hellmannin. Entä toisinpäin?

- g:ksi valitaan generaattori $\bmod n$, jotta eri arvoja X, Y ja k olisi mahdollisimman paljon.

  Oikeastaan g:n ei tarvitse olla primitiivinen, mutta sen pitää generoida mahdollisimman suuri $\mathbb{Z}_n^*$:n aliryhmä. ($\mathbb{Z}_n^*$ on kertolaskuryhmä $\{1, \ldots, n-1\}$.)

- Diffie-Hellmannin tuottamat avaimet $g^{xy} \bmod n$ eivät jakaudu tasaisesti joukkoon $\mathbb{Z}_n^*$, vaan todennäköisyydet voivat vaihdella enintään suhteessa 1:6.

  Avainmateriaalin entropia on siis 2–3 bittiä pienempi kuin mitä näyttää. Tällä ei ole merkistystä.

- Jos g^x generoi pienen aliryhmän $\mathbb{Z}_n^*$:ssä (sen **aste** on pieni), kuuluu $(g^x)^y$ tuohon pieneen aliryhmään. Tällöin avain on helppo löytää kokeilemalla.

  Ratkaisu:

  – luotetaan siihen, että x ja y valitaan satunnaisesti

  – jos n:llä on suuri alkutekijä, todennäköisyys, että g^x:n aste on alhainen, on pieni.

  Paras valinta on **turvallinen alkuluku**

  $$n = 2p + 1,$$

  missä myös p on alkuluku. Tällöin kaikki $\{2, \ldots, n-2\}$ ovat joko primitiivisiä tai niiden aste on p.

- Turvallisen alkuluvun n ja generaattorin g valinta:

  1. valitaan satunnaisia alkulukuja p, kunnes $n = 2p + 1$ on alkuluku.
  2. valitaan satunnainen g, kunnes se on generaattori $\bmod n$. Todennäköisyys onnistua kullakin yrityksellä on noin $1/2$.

- Hyökkääjä voi korottaa molemmat julkiset avaimet samaan potenssiin: $g^x \longrightarrow g^{xz}$ ja $g^y \longrightarrow g^{yz}$. Avaimeksi tulee g^{xyz}.

  Jos g^z:n aste on pieni, löytyy avain g^{xyz} kokeilemalla.

  Ratkaisu:

  – käytetään turvallista alkulukua g.
  Silti voi olla $g^z = 1$ (aste = 1) tai $g^z = -1$ (aste = 2).

  – **allekirjoitetaan julkiset avaimet**.

- Itse Diffie-Hellman-protokollassa vastapuolen henkilöllisyyttä ei tarkisteta. Aktiivinen hyökkääjä saattaa esiintyä väärällä nimellä.

  Syntyy yhteinen salaisuus, mutta kenen kanssa?
  Vastapuoli on aina tunnistettava erikseen.

  Julkiset avaimet pitää sitoa lähettäjäänsä esim. allekirjoituksin. Muuten on mahdollinen ns. *man-in-the-middle*-hyökkäys.

1.	$A \longrightarrow B$	$g^x,\ S_A(g^x)$
2.	$B \longrightarrow A$	$g^y,\ S_B(g^y)$
		Istuntoavain $K = h(g^{xy})$
3.	$A \longrightarrow B$	$E_K(M)$

Man-in-the-middle

n ja g ovat julkisesti tunnettuja.
Alice generoi avaimensa x ja $X = g^x \bmod n = 8389$.
Bob generoi avaimensa y ja $Y = g^y \bmod n = 1771$.
Charlie generoi avaimensa z ja $g^z \bmod n = 5876$.

1. Alice lähettää Bobille tarkoitetun sanoman: "$X = 8389$. T: Alice." Charlie nappaa sanoman.
2. Charlie vastaa Bobin puolesta: "$Y = 5876$. T: Bob."
3. Charlie lähettää Bobille sanoman: "$X = 5876$. T: Alice."
4. Bob vastaa: "$Y = 1771$. T: Bob." Charlie nappaa sanoman.

Nyt Alice ja Bob uskovat, että heillä on yhteinen avain.
Oikeasti Alicella ja Charliella on avain $K_1 = g^{xz} \bmod n$
ja Bobilla ja Charliella on avain $K_2 = g^{yz} \bmod n$.

1. Alice lähettää Bobille sanoman, jonka Charlie nappaa: E_{K_1}("Tavataan klo 13 sillalla.").
2. Charlie välittää sanoman Alicelle: E_{K_2}("Tavataan klo 14 puistossa.")

Charlie voi

– välittää sanomia ja siten salakuunnella keskustelua
– muuttaa välittämiään sanomia
– väärentää kokonaan uusia sanomia.

Diffie-Hellman vakioavaimilla

- Vakioavainten ja lähettäjän määräämien avainten etu:
  - avainneuvottelua ei tarvita, vaan voidaan aloittaa salatulla sanomalla
  - lähettäjä voi myös suorittaa salauksen etukäteen ennen yhteyden avaamista.
- Koska potenssiinkorotus on raskas operaatio, voidaan samoja eksponentteja käyttää useamman kerran samojen tahojen välillä.

  Ongelma: avain sama joka kerta.

  Ratkaisu: vaihdetaan julkisten avainten lisäksi nonsseja.

1.	$A \longrightarrow B$	$g^x,\ R_A,\ S_A(g^x, R_A)$
2.	$B \longrightarrow A$	$g^y,\ R_B,\ S_B(g^y, R_B)$
		Istuntoavain $K = h(g^{xy}, R_A, R_B)$

  DH takaa avaimen salaisuuden, nonssit tuoreuden.

- **Nonssi** on parametri, joka saa joka protokolla-ajossa uuden arvon. Yleensä nonssin pitää myös olla mahdoton arvata.

  Nonssi voi olla satunnaisluku tai satunnaisluvun ja aikaleiman yhdistelmä.

- Julkiset Diffie-Hellman-avaimet voidaan tallettaa julkiseen hakemistoon. Avaimesta ei tarvitse erikseen sopia. Riittää yksisuuntainen lähetys.
- Ongelma: Joku voi kopioida toisen henkilön julkisen avaimen itselleen. Esim.
  1. Alice kommunikoi Bobin kanssa. Avain K on generoitu Diffie-Hellmanilla. Alice lähettää Bobille sanoman $E_K(M)$. Charlie kuulee sen.
  2. Charlie lukee hakemistosta Alicen julkisen avaimen X ja lähettää sen hakemistoon omana avaimenaan. Istuntoavain Charlien ja Bobin välillä on nyt sama kuin Alicen ja Bobin välillä.
  3. Charlie lähettää Bobille sanoman $E_K(M)$. Bob luulee Charlien lähettäneen sanoman M, vaikka Charlie ei edes tunne sen sisältöä.

  Ratkaisu 1: Vaihdetaan nonsseja ja luodaan nonsseista riippuva avain $h(K, R_A, R_b)$. Charlie ei voi kopioida avainta, mutta nonssien vaihto vaatii kaksisuuntaista kommunikaatiota.

  Ratkaisu 2: avaimen pitää riippua osapuolten nimistä tms. istuntoja erottavista tekijöistä. Esim.

$$K = HASH(A, B, \text{IP-addr}_A, \text{IP-addr}_B, g^{xy}).$$

Monenkeskinen Diffie-Hellman

- Kolme tai useampi henkilöä voi laskea yhteisen avaimen samalla tavalla.

1. Alice valitsee satunaisen suuren luvun x ja lähettää Bobille $X = g^x \bmod n$.
2. Bob valitsee satunaisen suuren luvun y ja lähettää Carolille $Y = g^y \bmod n$.
3. Carol valitsee satunaisen suuren luvun z ja lähettää Alicelle $Z = g^z \bmod n$.
4. Alice lähettää Bobille $\bar{Z} = Z^x \bmod n$.
5. Bob lähettää Carolille $\bar{X} = X^y \bmod n$.
6. Carol lähettää Alicelle $\bar{Y} = Y^z \bmod n$.
7. Alice laskee $k = \bar{Y}^x \bmod n$.
8. Bob laskee $k = \bar{Z}^y \bmod n$.
9. Carol laskee $k = \bar{X}^z \bmod n$.

ElGamal

- Allekirjoitus (ja salaus)
- Perustuu diskreetin logaritmin laskemisen vaikeuteen.
- Ei kovin paljon käytetty sellaisenaan, mutta pohjana suurelle määrälle erilaisia algoritmeja (esim. DSA).

- **Avaimet**: Valitaan alkuluku p, **generaattori** $\bmod p$ ja satunnaisluku x.

  $$y = g^x \bmod p.$$

  Julkinen avain on (y, g, p), salainen x.

  Salaisen avaimen ratkaiseminen vaatii diskreetin logaritmin laskemisen.

- **Allekirjoitus:** Halutaan allekirjoittaa sanoma M. Valitaan satunnainen luku k siten, että $\gcd(k, p-1) = 1$. Lasketaan

  $$a = g^k \bmod p.$$

  Ratkaistaan seuraavasta yhtälöstä b Eukleideen algoritmilla.

  $$h(M) = (xa + kb) \bmod (p-1).$$

  Allekirjoitus on lukupari (a, b).

  Huom! **Sanomasta pitää laskea yksisuuntainen funktio.** Muuten satunnaisia allekirjoituksia on helppo väärentää.

  Huom! Satunnaisluvun k täytyy pysyä salaisena. Jokaiselle sanomalle pitää olla uusi k.

- **Allekirjoituksen tarkistaminen:**

  $$1 \overset{?}{\leq} a \overset{?}{\leq} p-1 \qquad \text{ja} \qquad y^a a^b \bmod p \overset{?}{=} g^{h(M)} \bmod p.$$

  Huom! a**:n koko pitää tarkistaa.** Muuten yhdestä allekirjoituksesta voi johtaa uusia.

ElGamal-esimerkki

- Alkuluku $p = 65537$, generaattori $g = 19251$ ja satunnaisluku $x = 12674$.

  $$y = g^n \bmod p = 19251^{12674} \bmod 65537 = 38347.$$

  Julkinen avain on $(38347, 19251, 65537)$ ja yksityinen 12674.

- Allekirjoitetaan sanoma, jonka sormanjälki $h(M) = 1234$. Generoidaan uusi satunnaisluku $k = 60871$.

  $$a = 19251^{60871} \bmod 65537 = 28902.$$

  Ratkaistaan b yhtälöstä

  $$1234 = (12674 \cdot 28902 + 60871 \cdot b) \bmod (65537 - 1).$$

  Saadaan
  $b = 60871^{-1} \cdot (1234 - 12674 \cdot 28902) \bmod (65536) = 40394.$

  Allekirjoitus on $(28902, 40394)$.

- Verifioidaan allekirjoitus. $1 \leq 28902 < 65537$ ja

  $$\begin{aligned} 38347^{28902} \cdot 28902^{40394} \bmod 65537 &= 5129 \\ 19251^{1234} \bmod 65537 &= 5129. \end{aligned}$$

  Allekirjoitus ok.

- Oikeasti lukujen pitäisi olla noin 1000 bittiä pitkiä.