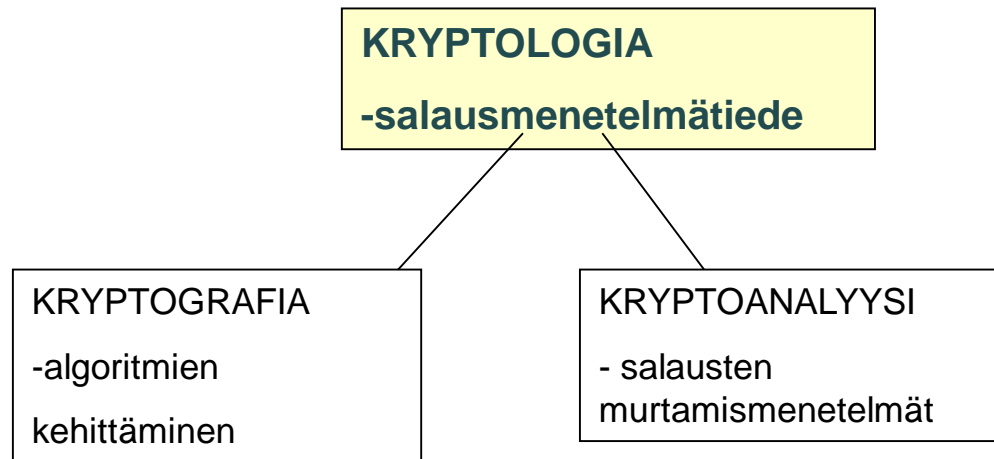


Aiheet:

Salauskäsitteistöä ja periaatteita

Klassiset salaukset

Cryptology, Cryptography, Cryptanalysis



Suomessa on merkittävää salausmenetelmäalan teollisuutta. Esimerkkinä mainittakoon SSH, joka on suomalaisen Tatu Ylösen kehittämä salausohjelmisto. Tuotteita myy SSH Communications Security – yhtiö. Asiakkaina on mm. EU ja NASA.

Internet Engineering Task Force (IETF) on määritellyt tietoturvan palvelut alla olevan listan mukaisesti. Kryptografia antaa keinoja useimpien tavoitteiden saavuttamiseksi

- | | |
|---|---------------------------------------|
| 1. luottamuksellisuus (confidentiality), | - tietoliikenteen ja tietojen salaust |
| 2. eheys (integrity), | - tiivistefunktiot |
| 3. todennus (authentication), | - esim. RSA - autentikointi |
| 4. kiistämättömyys (non-repudiation), | - digitaalinen allekirjoitus |
| 5. pääsynvalvonta (access control) ja | - salasana todennus, kryptografiset |
| 6. käytettävyys (availability). | (vahvat) autentikointimenetelmät |

Selitykset kohtiin 1 - 4

1. Vain tietoihin oikeutetut pääsevät lukemaan tietoja.
2. Tieto ei muutu siirrettäessä ja tiedon lähettäjä on varmennettu
3. Tiedon lähettäjä on kiistattomasti tunnistettu
4. Jonkin tapahtuman osapuoli ei voi jälkeinpäin kiistää osuuttaan siihen.

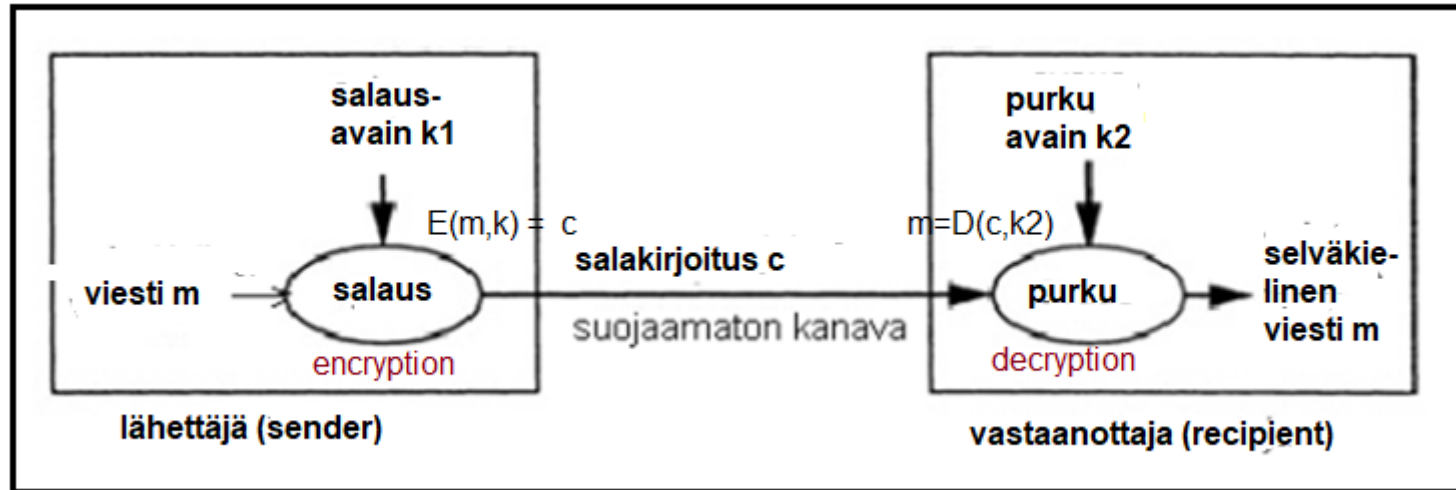
Salauksen peruskäsitteitä

Yleisiä periaatteita

Aiheet:

1. Tietoliikenteen salaus kaaviona
2. Symmetrinen ja asymmetrinen salaus
3. Kerckhoffin periaate
4. Avainavaruus ja efektiivinen avainavaruus
5. Salaukseen liittyviä matemaattisia käsitteitä

1. SALAUKSEN TOIMINTA KAAVIONA



Salauksen "kaava"

$$c = e(m, k_1)$$

m = selväkielinen viesti

k_1 = salausavain

e = salausfunktio

c = salakirjoitettu viesti

(plaintext, message)

(encryption key)

(encryption function)

(cipher, ciphertext)

Purkaminen

$$m = d(c, k_2)$$

d = purkufunktio

k_2 = purkuavain

decryption function

decryption key

2. SYMMETRINEN JA ASYMMETRINEN SALAUS

SYMMETRINEN SALAUS: Salausavain k_1 ja Purkuavain k_2 ovat samat

* Yhteyden osapuolet sopivat etukäteen salausavaimesta (yleensä käyttäen "key exchange"- protokollaa: DH, RSA tai ECDHE)

ASYMMETRINEN SALAUS: Salausavain k_1 ja Purkuavain k_2 ovat erisuuret

- Yleisin asymmetrisen salauksen muoto on ns. **julkisen avaimen salaus** (public key encryption), jossa jokaisella käyttäjällä on kaksi avainta: **julkinen avain** , jolla ko. käyttäjälle lähetettävät viesti salataan , sekä sen parina **yksityinen avain**, joka on vain käyttäjän tiedossa ja jolla hän purkaa saamansa viestit.

3. Kerckhoffin periaate

Kerckhoffin periaate:

"Salausjärjestelmän tulisi olla turvallinen siinäkin tapauksessa, että kaikki järjestelmässä avainta lukuunottamatta on julkista".

(Auguste Kerckhoff 1835 - 1903 oli hollantilainen kielitieteilijä ja kryptologi, joka toimi Pariisin kauppakorkeakoulussa kielten professorina)

Mm. DES – ja GSM – salauksien algoritmeja yritettiin pitää salassa vastoin Kerckhoffin periaatetta, mutta ne vuotivat julkisuuteen. Nykyään algoritmit pyritään alusta saakka pitämään julkisina, jotta jokaisella tutkijalla ym. taholla olisi mahdollisuus testata ja arvioida niiden turvallisuutta ja varmistua siitä, että menetelmät eivät sisällä heikkouksia tai takaportteja.

4. Avainavaruus (key space)

Hyvässä salausmenetelmässä ei ole rakenteellisia heikkouksia eikä takaportteja. Hyökkääjän paras mahdollisuus on murtaa salaus ns. **Brute Force** hyökkäyksellä: käymällä systemaattisesti läpi kaikki mahdolliset avainvaihtoehdot.

Tällöin turvallisuuden kannalta keskeisin tekijä on kaikkien mahdollisten avainten lukumäärä, eli ns. avainavaruuden koko. Avainavaruutta mitataan bitteinä.

Ehdottomana alarajana symmetrisen salausavaimen turvalliselle koolle pidetään 80 bittiä. Tämä tarkoittaa, että mahdollisia avaimia on $2^{80} = 1.2 \cdot 10^{24}$ kpl

Käytännön miniminä symmetrisen salausavaimen pituudelle pidetään 128 bittiä.

Efekiivinen avainavaruus (effective key space)

Salausmenetelmissä on usein joitain heikkouksia: ts. löydetään menetelmiä, joilla salausavain voidaan murtaa lyhyemmässä ajassa kuin kuin Brute Force hyökkäyksessä.

Efekiivisellä avainavaruudella tarkoitetaan keskimääräistä avainvaihtoehtojen määrää, joka läpi käymällä salausavain kyetään murtamaan hyödyntäen parhaita, tunnettuja menetelmiä. Efekiivinen avainavaruus on siten pienempi kuin nimellinen avainavaruus. Hyvissä salauksissa ero on erittäin pieni.

Esim. AES128	key space 128 bits, effective 126.1
AES256	key space 256 bits, effective 254.4
DES	key space 64 bits, effective 54
3DES	key space 168 bits, effective 112

Laskelmia salasanojen turvallisista pituuksista

Avainavaruuden minimikoko $2^{80} = 1.2 \cdot 10^{24}$ soveltuu myös salasanojen turvallisuuteen Brute Force – hyökkäyksiä vastaan. Tällöin voidaan puhua salasana-avaruuden koosta. Seuraavassa pari laskuesimerkkiä.

1. Salasanan pituus on 8 merkkiä, salasanassa saa olla vain englanninkielisiä aakkosia, joita on 26 kpl. Isoja ja pieniä kirjaimia ei eroteta.

Salasana-avaruus on $26^8 = 2 \cdot 10^{11}$ turvaton

2. Salasanan pituus on 11 merkkiä, salasanan merkistönä englanninkielisiä isot ja pienet aakkoset + numerot 0...9: yhteensä 62 merkkiä.

Salasana-avaruus on $62^{11} = 5 \cdot 10^{19}$ turvaton

3. Salasanan pituus on 13 merkkiä, salasanan merkistönä englanninkielisiä isot ja pienet aakkoset + numerot 0...9 + 10 erikoismerkkiä: yhteensä 72 merkkiä.

Salasana-avaruus on $72^{13} = 1.4 \cdot 10^{24}$ riittävä

5. Matemaattisia käsitteitä

HARD PROBLEM

= matemaattinen ongelma, joka on yleisesti tunnustettu erittäin vaikeaksi ja kompleksiseksi ratkaista. Salausmenetelmien turvallisuus perustuu usein tunnettuun ”kovaan” ongelmaan. Esim. RSA perustuu suurten kokonaislukujen tekijöihinjaon vaikeuteen, Diffie-Hellman protokolla ns. Diskreetin logaritmin ongelmaan

YKSISUUNTAISET FUNKTIOT (one way functions)

= **funktio $y = f(x)$, jonka arvo on nopea ja helppo laskea, kun x tunnetaan, mutta jonka käänteisfunktio eli x :n laskeminen kun y tunnetaan, on äärimmäisen hankalaa tai mahdotonta.**

TAKAPORTTI (backdoor)

= **Lisätieto, jonka avulla muutoin yksisuuntaisen funktion käänteisfunktio on helppo laskea.**

Esim. RSA:ssa yksityisen purkuavaimen laskeminen julkisesta avaimesta on lähes mahdotonta. Mutta jos joku tuntee julkisen avaimen (erittäin suuri kokonaisluku) alkulukutekijät, tehtävä on helppo. Julkisen avaimen tekijöiden tunteminen on siis takaportti RSA:han.

Satunnaislukugeneraattorissa Dual_EC_DRBG, joka oli hyväksytty USA:n standardiviranomaisen toimesta standardiksi, ja oli levinnyt yleiseen käyttöön, oli takaportti, joka antoi NSA:lle mahdollisuuden suojattujen tietoliikenneyhteyksien salakuunteluun.

Klassisia salausmenetelmiä

1. Klassisten salausten jaottelua
2. Caesar salaus
3. Frekvenssianalyysi
4. Affiini salaus
5. Yksinkertainen substituutio
6. Vigneren salaus
7. One Time Pad: "the only unbreakable cipher"
8. Enigma ja Lorentz 40/42 salauslaitteet

1. Klassisten salausten jaottelua

JAKO MERKKIEN KUVA-ALKIODEN LUKUMÄÄRÄN PERUSTEELLA

A) **Monoalphabetic cipher** on englanninkielinen termi salausmenetelmälle, jossa viesti salataan merkki kerrallaan ja kullakin merkillä on aina sama kuva-alkio (merkki) salauksessa. Tällaisia ovat esim. Caesar -salaus ja yksinkertainen substituutio.

B) **Polyalphabetic cipher** tarkoittaa salausjärjestelmää, jossa sama merkki voi kuvautua eri kohdissa viestiä eri kuvamerkiksi. Esimerkkinä tästä on Vigenèren salaus.

Tyyppiä B olevia salausmenelmiä pidetään tyyppiä A turvallisempina, koska tyyppiin A on olemassa tehokas kryptoanalyysimenetelmä: frekvenssianalyysi.

2. Caesar sala

- Salaus perustuu aakkoston rotaation. Salausavaimena on rotaation määrä.



Kuvan kiekkojen asennolla viestin
“AAMU” salakirjoitus olisi “NNZH”

Avainavaruuden koko (mahdollisten rotaatioiden lukumäärä) on 25 , joten murtaminen on helppoa jopa Brute Force -menetelmällä.

3. Frekvenssianalyysi

Monia klassisia salakirjotuksia voidaan tehokkaasti murtaa frekvenssianalyysillä.

Alla on yleisimpien englanninkielen aakkosten suhteellisia frekvenssejä (prosenttiosuutena englanninkielisessä teksteissä).

<i>e</i>	<i>t</i>	<i>a</i>	<i>o</i>	<i>n</i>	<i>i</i>	<i>s</i>	<i>r</i>	<i>h</i>
12.3	9.6	8.1	7.9	7.2	7.2	6.6	6.0	5.1

E on englannin kielen selvästi yleisin kirjain. Salakirjoitetun viestin merkki, jolla on suurin frekvenssi, on useimmin e:n kuvamerkki.

Vastaavasti suomen kielen kuusi yleisintä merkkiä ovat a, i, t, n, e, s

Esimerkki Caesar salauksen frekvenssianalyysistä

Seuraava salakirjoitus on tehty Caesar salauksella.

```
cqnujcnbcvxernj kxdclahycxpajyqhrbwxfrcqnjcnabrbcnuubj kxdcbnlxwmfxaumfja jwm  
cqnjccnvycbvjmnrwnwpujwmcxkajtnpnavjwnwrpvjlryqna
```

Lasketaan salakirjoituksen merkkien frekvenssit

n : 15 , c : 14, j : 13 , w : 9 , a : 8 , x : 8 ,

SUURIN FREKVENSSI ON KIRJAIMELLA n. Tehdään olettamus, että merkin n alkukuva on e, joka on englannin kielen yleisin kirjain. Aakkoston rotaation määrä, eli salausavain olisi siten 9.

Testataan analyysin paikkansapitävyys kohdistamalla salakirjoituksen merkkeihin aakkoston rotaatio 9 merkkiä vastakkaiseen suuntaan. Tulos on järkevä englannin kielinen viesti.

```
thelatestmovieaboutcryptographyisnowintheatersistellsaboutsecondworldwarand:  
theattempts made in england to brake german enigma cipher
```


4. Affiini salaus

1. Viesti koodataan luvuiksi välillä 0 – 25 ao. taulukon mukaan

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

2. Viestin luvut m salataan kaavalla

$$c = a m + b \pmod{26}$$

3. Salakirjoituksen luvut c voidaan dekodata takaisin merkkimuotoon

Salakirjoituksen purku

1. Salakirjoituksen lukuihin c sovelletaan käänteiskuvausta

$$m = a' c - a' b \pmod{26}$$

missä a' on a :n käänteisluku mod 26

(kurssin matemaattisessa osuudessa on esitetty Laajennettu Eukleideen algoritmi käänteisluvun laskemiselle)

2. Lukujono m dekodataan takaisin merkkimuotoon

- **Avaimena on lukupari (a, b)**
- Merkit koodataan kokonaisluvuiksi 0-25 yo. taulukon mukaisesti. Salaukseen ja purkuun käytetään jakojäännösaritmetiikkaa
- Purkukaavassa käytetään avaimen a käänteislukua mod 26. Menetelmä edellyttää siten käänteisluvun laskemisen hallitsemista.
- Caesar salausta voidaan pitää affiinin salauksen erikoistapauksena, kun $a = 1$.
Tällöin $c = m + b \pmod{26}$, joka vastaa aakkoston rotaatioita
- Avainavaruuden koko on $25 \cdot \phi(26) = 312$, koska b voidaan valita 26 eri tavalla, mutta a :n voidaan valita vain 12 eri tavalla siitä syystä että a :lla on käänteisluku vain jos sillä ei ole yhteisiä tekijöitä luvun 26 kanssa. Näitä lukuja on $\phi(26) = 12$ kpl

Laskuesimerkki affiinista salauksesta

Salaa viesti "kemi" avainparilla $a = 11$ ja $b = 3$

* Numeroiksi koodattuna viesti $m = (10, 4, 12, 8)$

* Salakirjoitus $c = (11 \cdot 10 + 3, 11 \cdot 4 + 3, 11 \cdot 12 + 3, 11 \cdot 8 + 3) \bmod 26$
 $= (9, 21, 5, 13) = \text{"jvfn"}$

Purku tapahtuu kaavalla $m = a' c - a' b \pmod{26}$

• $a' = a$:n käänteisluku $\bmod 26 = 11^{-1} \bmod 26 = 19$ *)

* Viesti $m = (19 \cdot 9 - 19 \cdot 3, 19 \cdot 21 - 19 \cdot 3, 19 \cdot 5 - 19 \cdot 3, 19 \cdot 13 - 19 \cdot 3) \bmod 26$
 $= (10, 4, 12, 8) = \text{"kemi"}$

*) Käänteislukuja $\bmod n$ voi helposti laskea online- laskimella [wolframalpha.com](https://www.wolframalpha.com), kirjoittamalla syöttökenttään: $11^{-1} \bmod 26$. (kokeile). Laskin käyttää Laajennettua Eukleideen algoritmia

*) Oikealla puolen oleva keskeneräinen analyysi perustuu olettamukseen, että salakirjoituksen suurimman frekvenssin omaava merkki s on alkuperäisessä viestissä e, ja toiseksi suurimman frekvenssin u on alunperin t. Lisäksi yhdistelmien t_e välikirjain, salakirjoituksen o on alun perin h, koska "the" on yleinen artikkeli ja sama yhdistelmä esiintyy myös sanoissa these, then, j.n.e.

6. Vigenèren salaus

Blaise de Vigenère 1523 -1596 oli ranskalainen kielitieteilijä ja kryptologi

- Salausavaimena käytetään salasanaa
- Salasanaa laajennetaan monistamalla salasanaa, kunnes se on viestin mittainen
- Salakirjoitus saadaan lisäämällä viestin merkkeihin salasanan merkit seuraavan kalvon yhteenlaskutaulukon avulla
- Vigenèren salaus oli yleinen sotilassalaus uuden ajan alussa
- Vigenèren salaus liittyy Caesar salaukseen siten, että se koostuu Caesar – salauksia avaimilla, jotka saadaan avainsanasta.
- Murtomenetelmän kehitti 1800- luvulla preussilainen upseeri Kasiski. Se perustuu frekvenssianalyysiin, jossa tutkitaan 2- 3 merkin mittaisten ”tavujen” frekvenssejä salakirjoituksissa

Vigeneren yhteenlaskutaulukko

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Esim.

Salataan viesti
“helsinki”
 salasanalla **“oulu”**

HEL S I N K I
 O U L U O U L U

=====

V Y W M W H V C

Avainavaruus on 25^n , missä n = salasanan pituus. Esim. jos satunnaisen salasanan pituus $n = 20$, avainavaruuden koko = $9 \cdot 10^{27}$ (kestää siis Brute Forcen)

Vigeneren salaus on murtamaton tietyin edellytyksin

- Jos Vigenèren salauksessa käytetään kertakäyttöistä, satunnaista, viestin mittaista salausavainta, on salaus mahdotonta murtaa.
- Tämä on ilmeistä, koska jokaista salakirjoitusta ja jokaista mahdollista viestiä kohden on olemassa jokin avain, jolla kyseisestä viestistä saadaan ko. salakirjoitus. Mahdollisten viestien joukosta on mahdotonta erottaa oikeaa viestiä.

7. One Time Pad (Vernam 1919)

Täydellinen, murtamaton salaus, jossa viesti muutetaan binäärimuotoon, salaus-avaimena on kertakäyttöinen, satunnaisesti generoitu, viestin pituinen bittijono.

Salaus suoritetaan laskemalla viesti ja avain yhteen XOR yhteenlaskua käyttäen.

Purku suoritetaan laskemalla salakirjoitus ja avain yhteen.

XOR yhteenlasku : $0 + 0 = 0$, $1 + 1 = 0$, $1 + 0 = 1$, $0 + 1 = 1$

Salaus:

Viesti	1	0	1	1	0	0	1	0
Avain	0	1	1	0	1	0	1	1
Salakirjoitus	1	1	0	1	1	0	0	1

Purku:

Salakirjoitus	1	1	0	1	1	0	0	1
Avain	0	1	1	0	1	0	1	1
Purettu viesti	1	0	1	1	0	0	1	0

Moskova- Washington kuuma linja käytti tätä salausta telex -yhdeyden salaamiseen kylmän sodan aikana.

8. Enigma ja Lorentz SZ 40/42



Viimeisiä klassisia salauslaitteita oli saksalaisten II maailmansodassa käyttämä Enigma, jota käytettiin viestintään sukellusveneiden kanssa.

Englantiin perustettiin kryptoanalyysiä varten keskus, "Bechley Park" johon kutsuttiin parhaat insinöörit ja matemaatikot, mm. Alan Turing.

Saksan salattujen viestien murtamisyritykset johtivat mm. ensimmäisen tietokoneen keksimiseen ("Tummy – machine")

Enigman salaus murrettiin ilmeisesti jo 1941, mutta saksalaiset kehittivät siitä parannetun version Lorenz SZ 40/42 – koneen.

Videot , yht. n. 15 min

<https://www.youtube.com/watch?v=GBsfWSQVtYA>

Lorentz machine

<https://www.youtube.com/watch?v=b4WBINgRMTY>

Tummy machine