Fejlkorrigerende koder, secret sharing (og kryptografi)

Olav Geil

Afdeling for Matematiske Fag
Aalborg Universitet

Møde for Matematiklærere i Viborg og Ringkøbing amter
7. november, 2006

Oversigt

Fejlkorrigerende koder

- Lineære koder
- Dekodning
- Endelige legemer
- Reed-Solomon koder
- Dekodning af Reed-Solomon koder
- Reed-Muller koder, Hyperbolske koder og Hermitekoder

Secret sharing

- Alle tilstede
- Lagrange-interpolation
- Shamirs secret sharing scheme

Offentlig nøgle kryptering

- McElieces kryptosystem

Fejlkorrigerende koder

Sikrer troværdig kommunikation i støjfyldte miljøer.

- Kommunikation fra Mars
- Stregkoder (forhåbentlig snart chip)
- Digitalt TV og radio
- CD
- DVD
- FTP (filoverførsel)
- osv...

Repetitionskoden

Hvis vi skal transmitere $(0, 1, 1, 0)$ så kunne vi jo bare sende $(0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0)$. Hvis der så sker en enkelt fejl, da kan den rettes:
$(0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0)$ skulle med stor sandsynlighed være $(0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0)$.

Er miljøet mere støjfyldt kunne vi jo indkode $(0, 1, 1, 0)$ til $(0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0)$ eller til noget endnu længere.

Ulempen er, at dyrt.

Indtil 1948 troede man, at dette ikke kunne gøres bedre. Kodningsteori handler om at kunne rette fejl med fornuftlig transmissionshastighed.

Tom som animation...gå til popularization->index.html

En binær lineær kode

Eksempel:
$C = \{(0,0,0,0,0), (1,0,1,1,0), (0,0,1,1,1), (1,0,0,0,1)\}$
Kan opfattes som linearkombinationer af basisvektorerne $g_1 = (1,0,1,1,0)$ og $g_2 = (0,0,1,1,1)$.

Vi har $(0,0,0,0,0) = 0g_1 + 0g_2$, $(1,0,1,1,0) = 1g_1 + 0g_2$, $(0,0,1,1,1) = 0g_1 + 1g_2$ og $(1,0,0,0,1) = 1g_1 + 1g_2$.

Dvs. koden svarer til rækkerummet af

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Indkodning: Besked (m_1, m_2) indkodes til $(m_1, m_2) = m_1 g_1 + m_2 g_2 = \vec{m}G$, hvor $\vec{m} = (m_1, m_2)$.

Enhver matrix G definerer en kode. En (lineær) kode er et vektorrum (indtil videre over det binære alfabet).

Model

$$\vec{m} \rightarrow \text{Indkoder} \xrightarrow{\vec{c}} \text{Kanal} \xrightarrow{\vec{r} = \vec{c} + \vec{e}} \text{Dekoder} \rightarrow \vec{m}'$$

$$\uparrow \vec{e}$$

$\vec{m} = (m_1, \ldots, m_k) \in \mathbb{F}_2^k$, $k < n$, $\vec{c} = (c_1, \ldots, c_n) \in \mathbb{F}_2^n$.
Indkodning $\vec{m} \mapsto \vec{c}$ vha. matrix G.
$\vec{e} = (e_1, \ldots, e_n) \in \mathbb{F}_2^n$, $\vec{m}' \in \mathbb{F}_2^k$

$Pr(e_i = 0) = 1 - p > \frac{1}{2}$, $Pr(e_i = 1) = p$. Kanalen antages hukommelsesfri, dvs. $Pr(e_i), Pr(e_j)$ er uafhængige for $i \neq j$.

Minimumafstand

$$w_H((w_1, \ldots, w_n)) = \#\{i \mid w_i \neq 0\}$$

$$w_H((1,1,1,1,0)) = 4$$

$$\text{dist}_H(\vec{w}_1, \vec{w}_2) = w_H(\vec{w}_1 - \vec{w}_2)$$

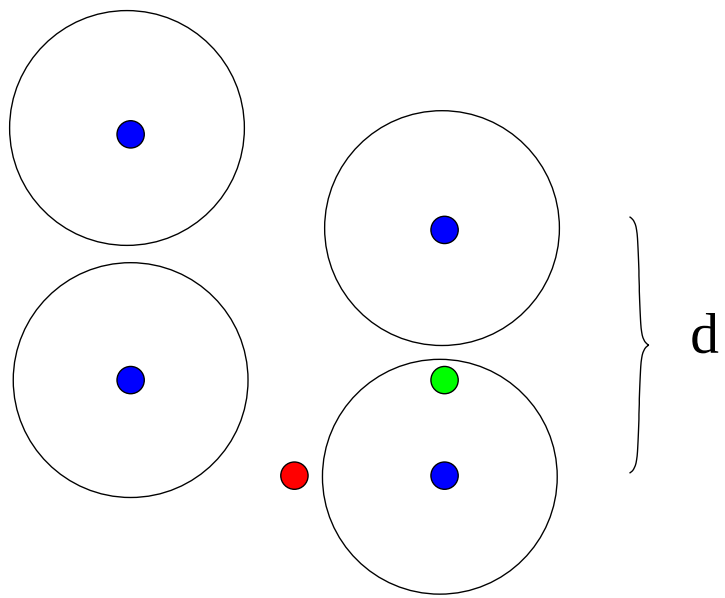
$$\text{dist}_H((0,1,0,1,1),(1,1,1,1,0)) = 3$$

$$\begin{aligned} d &= \min\{\text{dist}_H(\vec{c}_1, \vec{c}_2) \mid \vec{c}_1, \vec{c}_2 \in C, \vec{c}_1 \neq \vec{c}_2\} \\ &= \min\{w_H(\vec{c}) \mid \vec{c} \in C, \vec{c} \neq \vec{0}\} \end{aligned}$$

Minimumafstandsdekodning

Indenfor en afstand af $\left\lfloor \frac{d-1}{2} \right\rfloor$ fra et ord $\vec{w}$ kan man højst finde 1 kodeord.

Hvis der højst er sket $\left\lfloor \frac{d-1}{2} \right\rfloor$ fejl, da kan vi rette dem ved simpelthen at vælge det kodeord, som er tættest på den modtagne ord.



De tre parametre

Længden n, dimensionen k og minimumafstanden d.
$[n, k, d]$

Hvis $\frac{k}{n}$ er stor, da er vi sikret hurtig transmission.
Hvis $\frac{d}{n}$ er stor, da kan vi rette mange fejl.

Udfordringen er at vælge koder, så både $\frac{k}{n}$ og $\frac{d}{n}$ er store samtidig.

Dette er ikke en triviel opgave.

Assymptotisk gode koder

Man kender uendelige følger af koder $(C_1, C_2, \ldots)$ med tilhørende parametre $([n_1, k_1, d_1], [n_2, k_2, d_2], \ldots)$, som opfylder:

- $n_i \to \infty$ for $i \to \infty$
- $\frac{k_i}{n_i} \geq R >> 0$ for alle i
- $\frac{d_i}{n_i} \geq \delta >> 0$ for alle i

Betragt fejlmønstre $\vec{e}$ af voksende længde n. Lad kanalantagelserne gælde. De store tals lov siger, at $w_H(\vec{e})/n \to p$ i sandsynlighed.

Så hvis $\delta/2 > p$ kan vi kommunikere vilkårligt sikkert med hastighed mindst R, blot meget lange koder benyttes.

Shanons kanalkodningssætning garanterer endnu bedre fejlretning, hvis koderne vælges tilfældigt og af voksende længde. Et meget teoretisk resultat.

Endelige legemer

Eksempel:
$\mathbb{F}_5 = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Regneregel: $5 = 0$, så $3 + 4 = 2$ og $4 \cdot 4 = 1$.

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

$-3 = 2$ fordi $3 + 2 = 0$.

$3^{-1} = 2$ fordi $3 \cdot 2 = 1$.

$$2/3 = 2 \cdot 3^{-1} = 2 \cdot 2 = 4.$$

$\mathbb{F}_p = \mathbb{Z}_p = \{0, 1, \ldots, p-1\}$. Regneregel: $p = 0$.

Givet a, p da entydigt r, så $a = qp + r$ med $0 \leq r < p$.
$r = a (\text{mod } p)$.

Regneoperationer i $\mathbb{F}_p$ svarer til
$(a + b)(\text{mod } p)$ og $(a \cdot b)(\text{mod } p)$.

For $a = 0$ haves $-a = 0$
For $a \neq 0$ haves $-a = p - a$ og a^{-1} findes vha. den udvidede Euklidsalgoritme (EEA) idet $sfd(a, p) = 1$. $EEA(a, p)$ returnerer s, t, så $sa + tp = 1$ og dermed $a^{-1} = s (\text{mod } p)$.

Eksempel: $\mathbb{F}_{2^2} = \mathbb{F}_4 = \{0, 1, \alpha, \alpha + 1\}$. Regneregler: $2 = 0$ og $\alpha^2 + \alpha + 1 = 0$.

$+$	0	1	α	$\alpha + 1$
0	0	1	α	$\alpha + 1$
1	1	0	$\alpha + 1$	α
α	α	$\alpha + 1$	0	1
$\alpha + 1$	$\alpha + 1$	α	1	0

$\cdot$	0	1	α	$\alpha + 1$
0	0	0	0	0
1	0	1	α	$\alpha + 1$
α	0	α	$\alpha + 1$	1
$\alpha + 1$	0	$\alpha + 1$	1	α

$\alpha \cdot \alpha = \alpha^2 = -\alpha - 1 = \alpha + 1$

$\alpha(\alpha + 1) = \alpha^2 + \alpha = -1 = 1 \quad (\alpha + 1)(\alpha + 1) = \alpha^2 + 1 = \alpha.$

Eksempel:
$\mathbb{F}_{2^3} = \mathbb{F}_8 = \{0, 1, \alpha, \alpha + 1, \alpha^2, \alpha^2 + \alpha, \alpha^2 + \alpha + 1, \alpha^2 + 1\}$.
Regneregler: $2 = 0$ og $\alpha^3 + \alpha + 1 = 0$.

Det generelle set-up:
Legemet $\mathbb{F}_{p^m} = \{a(\alpha) \mid a(X) \in \mathbb{F}_p[X], \deg(a) < m\}$.
Regneregler $p = 0$ og $f(\alpha) = 0$. Polynomiet $f(X)$ er irreducibelt over $\mathbb{F}_p$, dvs. kan ikke faktoriseres. Den udvidede Euklids algoritme give $(a(\alpha))^{-1}$ idet, $sfd(a(X), f(X)) = 1$.

Lineære koder over $\mathbb{F}_q$

$\mathbb{F}_q^n = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{F}_q\}$ er et vektorrum over $\mathbb{F}_q$.

Et underrum $C \subseteq \mathbb{F}_q^n$ kaldes en lineær kode.

Enhver lineær kode kan realiseres som rækkerummet af en matrix G, svarende til indkodning $\vec{c} = \vec{m}G$.

Al gennemgået teori genereliserer til lineære koder over $\mathbb{F}_q$.

Reed-Solomon koderne

$q = p^m$ og $\mathbb{F}_q = \{P_1, P_2, \ldots, P_q\}$.

Givet $F(X) = F_0 + F_1 X + \cdots F_{k-1} X^{k-1} \in \mathbb{F}_q[X]$, da er $F(P_1), F(P_2), \ldots, F(P_q))$ en vektor af længde $n = q$ over $\mathbb{F}_q$.

$$\mathrm{RS}_q(k) \quad = \quad \{(F(P_1), F(P_2), \ldots, F(P_q)) \mid F(X) \in \mathbb{F}_q[X], \deg(F) < k\}$$

$$G = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ P_1 & P_2 & \cdots & P_q \\ P_1^2 & P_2^2 & \cdots & P_q^2 \\ \vdots & \vdots & \ddots & \cdots \\ P_1^{k-1} & P_2^{k-1} & \cdots & P_q^{k-1} \end{bmatrix}$$

Besked $\vec{m} = (F_0, F_1, \ldots, F_{k-1})$ indkodes til

$$(F(P_1), F(P_2), \ldots, F(P_q)) = \vec{m}G.$$

Reed-Solomon kodens parametre

Et ikke-nul polynomium af grad højst $k-1$ kan højst have $k-1$ nulpunkter.

Lad $k \leq q$ og betragt

$$\mathrm{RS}_q(k) \quad = \quad \{(F(P_1), F(P_2), \ldots, F(P_q)) \mid F(X) \in \mathbb{F}_q[X], \deg(F) < k\}$$

Minimumafstand lig minimumvægt. Højst $k-1$ nuller og dermed mindst $q-(k-1) = q-k+1 = n-k+1$ ikke-nuller. Altså $d = n-k+1$.

Hvis rækkerne i G var lineært afhængige, så ville vi have ikke-nul polynomium af grad højst $k-1 \leq q-1$ med q nulpunkter. Så G må have fuld rang.

$[n = q, k, d = n-k+1]$

Minimumafstands-dekodning af Reed-Solomon koder

Betragt en Reed-Solomon kode
$\mathrm{RS}_q(k) = \{(F(P_1), \ldots, F(P_q)) \mid \deg(F) < k\}$.
Definer $t = \lfloor (d-1)/2 \rfloor = \lfloor (q-k)/2 \rfloor$.

Hvis vi modtager $\vec{r} = (r_1, \ldots, r_q)$ da bestemmer vi et ikke-nul polynomium

$$Q(X, Y) = Q_0(X) + YQ_1(X)$$

som opfylder følgende

- $Q(P_1, r_1) = 0, Q(P_2, r_2) = 0, \ldots, Q(P_q, r_q) = 0$
- $\deg(Q_0) \leq q - 1 - t = l_0$
- $\deg(Q_1) \leq t = l_1$

Hvordan kan vi nu være sikre på, at et sådant polynomium $Q(X, Y)$ eksisterer?

Lad $Q_0(X) = Q_{0,0} + Q_{0,1}X + Q_{0,2}X^2 + \cdots + Q_{0,l_0}X^{l_0}$ og $Q_1(X) = Q_{1,0} + Q_{1,1}X + \cdots + Q_{1,l_1}X^{l_1}$. Vi får

$$
\begin{aligned}
& Q(P_1, r_1) = 0 \\
\Updownarrow & \\
& Q_{0,0} + Q_{0,1}P_1 + Q_{0,2}P_1^2 + \cdots + Q_{0,l_0}P_1^{l_0} \\
& \quad + Q_{1,0}r_1 + Q_{1,1}r_1P_1 + \cdots + Q_{1,l_1}r_1P_1^{l_1} = 0
\end{aligned}
$$

Dette er en homogen ligning med $(l_0 + 1) + (l_1 + 1) = q + 1$ ubekendte (nemlig $Q_{i,j}$'erne).

Der er q sådanne ligninger. Da altså homogent ligningssystem med flere ubekendte end ligninger haves ikke-nul løsning.

På matrixform ser ligningssystemet ud som følger:

$$\begin{bmatrix} 1 & P_1 & P_1^2 & \cdots & P_1^{l_0} & r_1 & r_1P_1 & \cdots & r_1P_1^{l_1} \\ 1 & P_2 & P_2^2 & \cdots & P_2^{l_0} & r_2 & r_2P_2 & \cdots & r_2P_2^{l_1} \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & P_q & P_q^2 & \cdots & P_q^{l_0} & r_q & r_qP_q & \cdots & r_qP_q^{l_1} \end{bmatrix} \begin{bmatrix} Q_{0,0} \\ Q_{0,1} \\ Q_{0,2} \\ \vdots \\ Q_{0,l_0} \\ Q_{1,0} \\ Q_{1,1} \\ \vdots \\ Q_{1,l_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Antag, at $\vec{c} = (F(P_1), F(P_2), \ldots, F(P_q))$ var afsendt (men er ukendt) og at der højst skete t fejl under transmissionen.

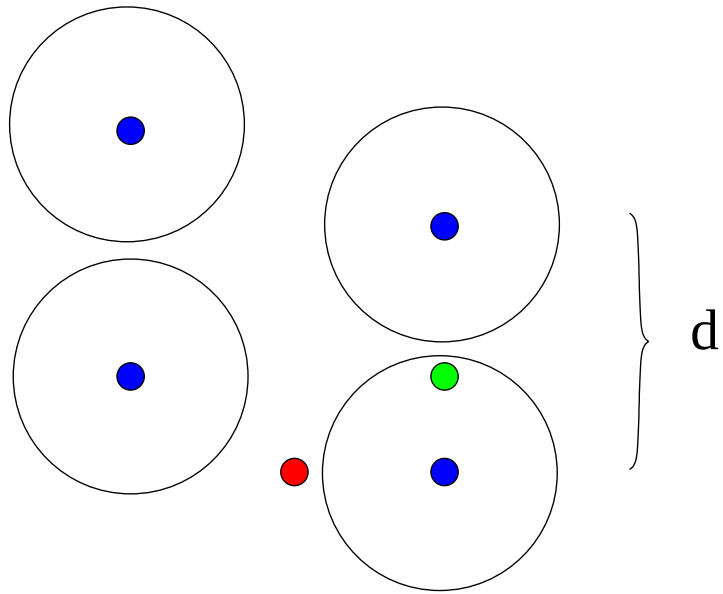
Vi har $Q(P_1, r_1) = Q(P_2, r_2) = \cdots = Q(P_q, r_q) = 0$ og da der højst er sket t fejl, må der altså være mindst $q - t$ nuller blandt

$$Q(P_1, F(P_1)), Q(P_2, F(P_2)), \ldots, Q(P_q, F(P_q)).$$

Men opfattes $Q(X, F(X)) = Q_0 + F(X)Q_1(X)$ som polynomium i X, da har det grad højst $\max\{q - 1 - t, (k - 1) + t\} = q - 1 - t$. Et polynomium af grad højst $q - 1 - t$, som har mindst $q - t$ nulpunkter må være nulpolynomiet. Vi får

$$\begin{aligned} & Q(X, F(X)) = 0 \\ \Updownarrow & \\ & Q_0(X) + F(X)Q_1(X) = 0 \\ \Updownarrow & \\ & F(X) = -\frac{Q_0(X)}{Q_1(X)} \end{aligned}$$

Listedekodning



Det er langt fra altid, at der findes et kodeord i afstanden $t = \lfloor (d-1)/2 \rfloor$ fra et modtaget ord $\vec{r}$. Det giver derfor god mening, at betragte kugler med større radius end t. Gør vi det, må vi så blot acceptere, at vi en sjælden gang imellem finder flere kandiderende kodeord.

Generelisering af minimumafstandsdekodningsmetoden til listedekodning som følger.

Søger $Q(X,Y) = Q_0(X) + Q_1(X)Y + \cdots + Q_m(X)Y^m$ således at

- $Q(P_i, r_i) = 0$ for $i = 1, \ldots, q$
- Visse gradbetingelser på Q_i'erne skal være opfyldt

Bestemmer herefter samtlige faktorer $Y - F(X)$ i $Q(X,Y)$. Der kan højst være m sådanne faktorer.

Metoden kan yderligere forbedres, hvis nulpunkter regnes med multiplicitet. Multiplicitet af polynomier i flere variabler er ikke trivielt. Der er mange forskellige definitioner.

Reed-Muller koder og hyperbolske koder

$F(X,Y) \in \mathbb{F}_4[X,Y]$.
$\mathbb{F}_4 \times \mathbb{F}_4 = \{(a,b) \mid a,b \in \mathbb{F}_4\} = \{P_1, P_2, \ldots, P_{16}\}$.

Hvor mange nulpunkter kan F have?

Definerende ligninger for $\mathbb{F}_4$ er $X^4 - X = 0$ og $Y^4 - Y = 0$, altså

$$
\begin{array}{c}
(a,b) \in \mathbb{F}_4 \times \mathbb{F}_4 \\
\Updownarrow \\
a^4 - a = 0 \ \text{ og } \ b^4 - b = 0
\end{array}
$$

Hvor mange fælles nulpunkter mellem $X^4 - X$, $Y^4 - Y$ og $F(X,Y)$?

Eksempel: $F(X,Y) = X^2Y + \alpha XY^2 + XY + 1$. Hvad er ledende monomium? Mange monomielle ordninger, når flere variable. Nogle ville give $\mathrm{lm}(F) = X^2Y$ andre ville give $\mathrm{lm}(F) = XY^2$.

Vælg monomiel ordning så $\mathrm{lm}(F) = X^2Y$. Vi har $\mathrm{lm}(X^4 - X) = X^4$ og $\mathrm{lm}(Y^4 - Y) = Y^4$.

Y^4	$\square$	$*$	$*$	$*$	$*$
Y^3	$\cdot$	$\cdot$	$*$	$*$	$*$
Y^2	$\cdot$	$\cdot$	$*$	$*$	$*$
Y	$\cdot$	$\cdot$	$\square$	$*$	$*$
1	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\square$
	1	X	X^2	X^3	X^4

Antal fælles nulpunkter for $F(X, Y)$, $X^4 - X$ og $Y^4 - Y$ er højst 10.

Y^3	12	13	14	15
Y^2	8	10	12	14
Y	4	7	10	13
1	0	4	8	12
	1	X	X^2	X^3

Antal nulpunkter for $F(X, Y)$ med $\mathrm{lm}(F) = XY^3$ er højst 13 osv.

$$\mathrm{RM}_4(s, 2) = \{(F(P_1), \ldots, F(P_{16})) \mid \text{total grad af } F \text{ er højst } s - 1\}$$

Y^3	12	13	14	15
Y^2	8	10	12	14
Y	4	7	10	13
1	0	4	8	12
	1	X	X^2	X^3

Koder	n	k	d
$\mathrm{RM}_4(1,2)$	16	1	16
$\mathrm{RM}_4(2,2)$	16	3	12
$\mathrm{RM}_4(3,2)$	16	6	8
$\mathrm{RM}_4(4,2)$	16	10	4
$\mathrm{RM}_4(5,2)$	16	13	3
$\mathrm{RM}_4(6,2)$	16	15	2
$\mathrm{RM}_4(7,2)$	16	16	1
Hyp	16	11	4

Hermitekoder

$X^3 + Y^2 + Y$ har 8 nulpunkter $\{P_1, P_2, \ldots P_8\}$ i $\mathbb{F}_4 \times \mathbb{F}_4$.

Givet $F(X, Y)$, hvor mange nuller kan der risikere at være i ordet $(F(P_1), F(P_2), \ldots, F(P_8))$?
Svar: lad $w(X^i Y^j) = i2 + j3$. Definer wdeg(F) til at være højeste $w(X^i Y^j)$-værdi for monomier $X^i Y^j$ i $F(X, Y)$.
Eksempel: wdeg$(X^2 + XY) = \max\{4, 5\} = 5$.

Resultat: Antal nulpunkter i $(F(P_1), F(P_2), \ldots, F(P_8))$ er højst lig wdeg(F).

Algebraisk-geometri koder

Reed-Solomon koder, Reed-Muller koder og Hermite koder de allersimpleste eksempler på AG-koder

- Algebraisk geometri er en meget dyb matematisk teori
- Meget har vist sig at kunne laves vha. simplere Gröbner basis teori.
- Vægte (eller valuationer) spiller stor rolle. Ikke nødvendigvis numeriske.

Secret sharing - problemet

Bank med 5 bankdirektører.

Simpleste opgave:
Elektronisk lås laves således, at

- alle 5 skal være tilstede for at pengeskab kan åbnes

Lidt mere avanceret opgave:
Elektronisk lås laves således, at

- vilkårlige 3 bankdirektører kan sammen åbne pengeskabet
- intet sæt af blot 2 bankdirektører kan alene åbne pengeskabet

Alle tilstede - naivt set-up

Bankdirektør b_1 får koden s_1, bankdirektør b_2 får koden s_2, $\cdots$ bankdirektør b_5 får koden s_5.

$s_1, s_2, \ldots, s_5 \in \{0000, 0001, \ldots, 1999\}$.

Den hemmelige nøgle er

$$s = s_1 + s_2 + s_3 + s_4 + s_5 \in \{0000, 0001, \ldots, 9995\}$$

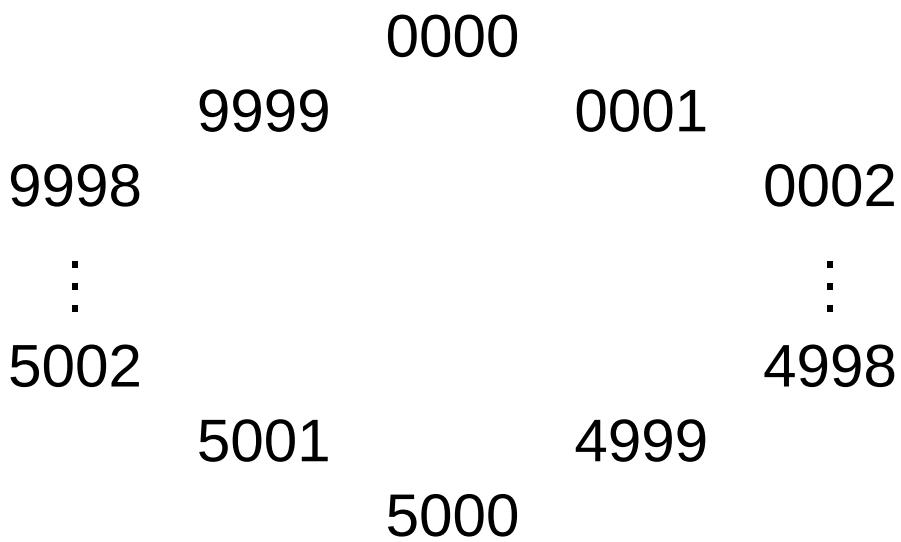
Imidlertid har $\{b_1, b_2, b_3, b_4\}$ information om s, idet de ved

$$s \in \{s_1 + s_2 + s_3 + s_4 + 0000, \ldots, s_1 + s_2 + s_3 + s_4 + 1999\}$$

Alle tilstede - smart set-up

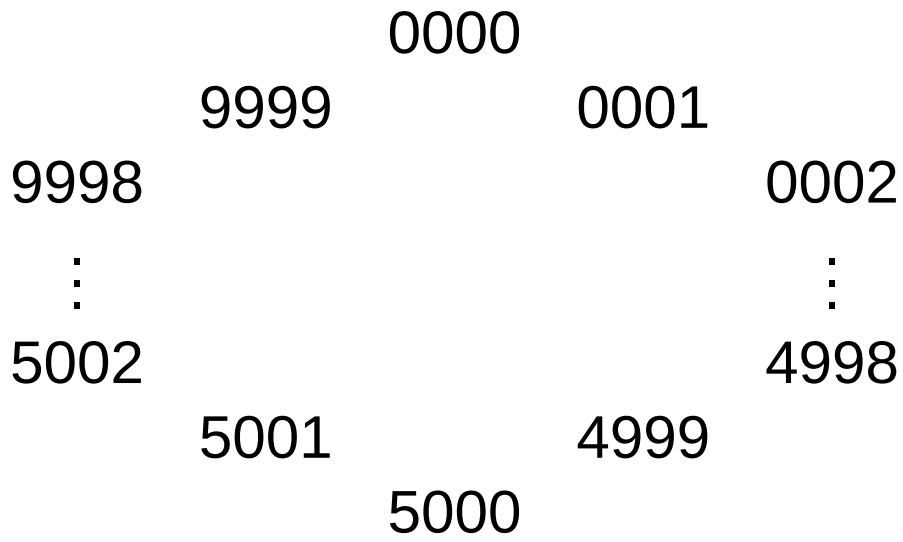
Nu istedet: $s_1, s_2, \ldots, s_5 \in \{0000, \ldots, 9999\}$.

Regner, modulo 10000.
Dvs. $5012 + 6013 \pmod{10000} = 1025$.



$s = (s_1 + s_2 + s_3 + s_4 + s_5) \pmod{10000}$.

Alle tilstede - smart set-up



$\{b_1, b_2, b_3, b_4\}$ har ingen information om s. De ved blot, at

$$\begin{aligned} s &\in \{(s_1 + s_2 + s_3 + s_4 + 0000) \pmod{10000}, \ldots, \\ &\qquad (s_1 + s_2 + s_3 + s_4 + 9999) \pmod{10000}\} \\ &= \{0000, \ldots, 9999\} \end{aligned}$$

Legemet $\mathbb{Z}_p$

$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ er et legeme.
$a + b \pmod{p}$ og $a \cdot b \pmod{p}$.

Eksempel:
$\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$, $3 + 4 = 2$ og $4 \cdot 4 = 1$.

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$-3 = 2$ fordi $3 + 2 = 0$.

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

$3^{-1} = 2$ fordi $3 \cdot 2 = 1$.

Lagrange-interpolation

$u_1 \neq u_2$. Der findes entydigt polynomium $F(X)$ af grad højst 1 så $F(u_1) = v_1$ og $F(u_2) = v_2$.

Entydighed:
Antag både $F_1(X)$ og $F_2(X)$. Men så har $F(X) = F_1(X) - F_2(X)$ nulpunkter u_1 og u_2. Dermed $F(X) = 0$.

Eksistens:

$$F(X) = \frac{X - u_2}{u_1 - u_2} v_1 + \frac{X - u_1}{u_2 - u_1} v_2$$

Lagrange-interpolation

u_1, u_2 og u_3 forskellige. Der findes entydigt polynomium $F(X)$ af grad højst 2, så $F(u_1) = v_1$, $F(u_2) = v_2$ og $F(u_3) = v_3$.

Eksistens:

$$\begin{aligned} F(X) &= \frac{(X-u_2)(X-u_3)}{(u_1-u_2)(u_1-u_3)}v_1 + \frac{(X-u_1)(X-u_3)}{(u_2-u_1)(u_2-u_3)}v_2 \\ &\quad + \frac{(X-u_1)(X-u_2)}{(u_3-u_1)(u_3-u_2)}v_3 \end{aligned}$$

Det generelle set-up: $u_1, u_2, \ldots, u_n$ forskellige $F(X)$ af grad højst $n-1$ osv.

Lagrange-interpolation

Eksempel: Legeme er $\mathbb{R}$
$F(X)$ af grad højst 1, så $F(3) = 2$ og $F(4) = 6$ er givet ved

$$\begin{aligned} F(X) &= \frac{X-4}{3-4}2 + \frac{X-3}{4-3}6 \\ &= -10 + 4X \end{aligned}$$

Eksempel: Legeme er $\mathbb{Z}_5$
$F(X)$ af grad højst 1, så $F(3) = 2$ og $F(4) = 1$ er givet ved

$$F(X) = 4X$$

Shamirs secret sharing scheme

3 bankdirektører. Vilkårlige 2 må åbne pengeskab sammen, men ingen må alene.

Hemmelig nøgle $f_0 \in \mathbb{Z}_p$. Dealer vælger tilfældigt $f_1 \in \mathbb{Z}_p$ og danner $F(X) = f_0 + f_1 X$.

Bankdirektør b_1 får: $u_1 = 1$ og $v_1 = F(u_1) = F(1)$
Bankdirektør b_2 får: $u_2 = 2$ og $v_2 = F(u_2) = F(2)$
Bankdirektør b_3 får: $u_3 = 3$ og $v_3 = F(u_3) = F(3)$

Vha. Lagrangeinterpolation kan vilkårlige 2 direktører genskabe $F(X)$ og dermed specielt den hemmelige nøgle f_0.

Hvad ved b_1 alene? Kald $v_2 = T$.

$$F(X) = \frac{X-2}{1-2} v_1 + \frac{X-1}{2-1} T = (2v_1 - T) + (T - v_1)X$$

Så $f_0 = 2v_1 - T$ og b_1 ved intet.

Shamirs secret sharing scheme

5 bankdirektører $b_1, b_2, \ldots, b_5$. Vilkårlige 3 må genskabe nøgle, men ej 2.

Hemmelig nøgle $f_0 \in \mathbb{Z}_p$, $p > 5$. Dealer vælger tilfældigt $f_1, f_2 \in \mathbb{Z}_p$ og danner $F(X) = f_0 + f_1 X + f_2 X^2$.

$$
\begin{aligned}
b_1: &\quad u_1 = 1 \text{ og } v_1 = F(1) \\
b_2: &\quad u_2 = 2 \text{ og } v_2 = F(2) \\
&\quad \vdots \\
b_5: &\quad u_5 = 5 \text{ og } v_5 = F(5)
\end{aligned}
$$

b_1, b_2 og b_3 danner Lagrangepolynomiet

$$\begin{aligned} F(X) &= \frac{(X-u_2)(X-u_3)}{(u_1-u_2)(u_1-u_3)}v_1 + \frac{(X-u_1)(X-u_3)}{(u_2-u_1)(u_2-u_3)}v_2 \\ &\quad + \frac{(X-u_1)(X-u_2)}{(u_3-u_1)(u_3-u_2)}v_3 \end{aligned}$$

og finder heraf

$$f_0 = \frac{v_1u_2u_3}{(u_1-u_2)(u_1-u_3)} + \frac{u_1v_2u_3}{(u_2-u_1)(u_2-u_3)} + \frac{u_1u_2v_3}{(u_3-u_1)(u_3-u_2)}$$

Uden kendskab til v_1 vides intet om første led. Uden kendskab til v_2 vides intet om andet led. Uden kendskab til v_3 vides intet til sidste led.

(åbn Maple 10 og vælg “recent documents”) Maple-worksheet

Secret sharing generelt

Givet $\{b_1, b_2 \ldots, b_n\}$ specificer en familie af delmængder af $\{b_1, b_2, \ldots, b_n\}$ som skal have adgang til at generere hemmelig nøgle. Lad eventuelt nogle personers share vægte højere end andres.

Enhver access-struktur kan realiseres matematisk.

Effektivitet: Forhold mellem størrelsen på hemmelig nøgle og størrelsen på share. Ikke altid lige effektivt, derfor forskningsområde stadig.

Hvad, hvis en af personerne er en pirat?

Er det smart, hvis share er stor?

Offentlig nøgle kryptosystem

- Funktion f med invers f^{-1}
- Offentliggør f
- Hemmeligholder f^{-1}
- Kan bruges til kryptering, hvis det er meget svært at beregne f^{-1} alene ud fra viden om f.

$$\begin{array}{ccccc} \text{Alice} & \rightarrow & \text{Kanal} & \rightarrow & \text{Bob} \\ & & \downarrow & & \\ & & \textit{Eve} & & \end{array}$$

Alice (indkodning): $f(\vec{m}) = \vec{c}$

Bob (afkodning): $\vec{m} = f^{-1}(\vec{c})$

McElieces kryptosystem

G' : $k \times n$ generator matrix for binær irreducibel $[n, k]$ Goppa-kode $\mathcal{G}$, som kan rette t fejl

S : $k \times k$ randomiseret binær matrix af fuld rang

P : $n \times n$ permutationsmatrix

G : $SG'P$

Offentlig nøgle: (G, t)

Hemmelig nøgle: $(S, \mathcal{D}_{\mathcal{G}}, P)$, hvor $\mathcal{D}_{\mathcal{G}}$ er en effektiv dekodningsalgoritme for koden $\mathcal{G}$.

Alice (indkodning): Vælg $\vec{e}$ med $w_H(e) = t$ tilfældigt. Indkod $\vec{c} = f(\vec{m}) = \vec{m}G + \vec{e}$.

Bob (afkodning): Beregn
$\vec{c}P^{-1} = (\vec{m}SG'P + \vec{e})P^{-1} = mSG' + \vec{e}P^{-1}$.
Bemærk, at $w_H(\vec{e}P^{-1}) = t$ og at $\vec{m}SG'$ er et kodeord.
Dekod vha. $\mathcal{D}_{\mathcal{G}}$ til mSG'.
Find herefter $\vec{m}S$ og til sidst $\vec{m} = (\vec{m}S)S^{-1}$.

Hvorfor virker McElieces kryptosystem?
Svar: Fordi det er uhyre svært at dekode, når man ikke har en god beskrivelse af ens kode. Dermed er det svært at finde f^{-1}.