Samoopravné kódy.

$x \longrightarrow y \longrightarrow x$

Def: Abeceda – konečná množina $\mathcal{E}$

Def: Prvek abecedy – symbol

Def: Slovo délky n – uspořádaná n-tice symbolů $\in \mathcal{E}$

$\mathcal{E}^n$ → množina všech slov délky n

Def: Hammingova vzdálenost mezi slovy x, y,

kde $x = x_1 \dots x_n$, $y_1 \dots y_n$, je

$d(x,y) = |\{i \in \{1 \dots n\} : x_i \neq y_i\}|$ ← počet pozic, kde se liší

Def: Blokový kód $C \subseteq \mathcal{E}^n$

– slova z C → kódová slova

Pomocí kódu C jsme schopni opravit nanejvýš t chyb, pokud $\forall y$ slovo délky n existuje nanejvýš jedno kódové slovo, které je od y vzdáleno nejvíce t.

Parametry kódu:

– délka slov $= n$

– velikost abecedy $= q$

– dimenze $= k = \log_q(C)$

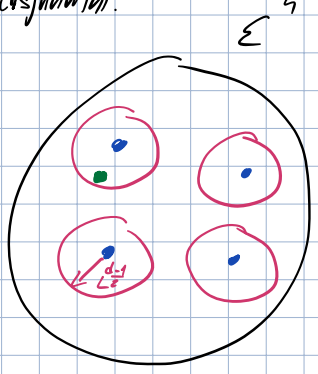
– vzdálenost $= d$ = min. kam. vzdálenost 2 různých kódových slov

Kód s parametry n, k, d, q značíme $(n,k,d)_q$.

Kód $(n,d,k)_q$ zvládne opravit $\leq \lfloor \frac{d-1}{2} \rfloor$ chyb. ← délky min. chybě d

– množina slov ve vzd. $\leq \lfloor \frac{d-1}{2} \rfloor$ od kódových slov jsou disjunktní.

– tedy každé slovo spadne do nanejvýš jedné kuličky.



Každý kód délky n opraví $\leq \lfloor \frac{n-1}{2} \rfloor$ chyb.

1) Opakovací kód:

$|\mathcal{E}| = q$, kódová slova vzniknou opakováním nějakého symbolu n-krát

$C = \{\underbrace{11\dots1}_{n}, \underbrace{22\dots2}_{n}, \underbrace{33\dots3}_{n}, \dots, \underbrace{qq\dots q}_{n}\}$, parametry: $(n,1,n)_q$

$k = \log_q |C| = \log_q q = 1$

2) Charakteristické vektory přímek UPR:

Mějme UPR $(X, \mathcal{P})$ řádu n, $\mathcal{E} = \{0,1\}$, $C = \{v \in \{0,1\}^{|X|} : v = \text{char. vekt. přímky}\}$

$\Leftrightarrow v_i = 1$, pokud vrchol i je přímce

Parametry $(n^2+n+1,\ \log_2(n^2+n+1),\ 2n)_2$

$\hookrightarrow |X|$; $\hookrightarrow$ # přímek

$\forall$ dvě přímky se protínají v jedném bodě. Obsahuje kvěčí $n+1$ bodů. Tedy chybí mi ve dvojici $2n$ pozic odlišných.

3) Hadamardovy kódy:

Def: Hadamardova matice: $H \in \{-1,1\}^{n \times n}$, kde $H \cdot H^T = n \cdot I_n$

- Dva různé řádky se liší na právě $\frac{n}{2}$ pozicích.

Kódové slovo $C = \{\text{řádky } H\} \cup \{\text{řádky } -H\} \subseteq \{-1,1\}^n$

Když se řádky liší, násobím -1 a 1, počet se mi změní. Na diagonále ale násobím řádek sám se sebou, dostanu $n \cdot 1$.

parametry $(n, \log_2(2n), \frac{n}{2})_2$

$\hookrightarrow$ délka odlišnosti řádků

Sylvestrova konstrukce: Jak si vyrobit H velikosti 2^i?

$$H_1 = (1), \quad H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad n \geq 2: \ H_{2n} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}$$

Hadamardova domněnka: $\forall n$ dělitelné 4 $\exists$ had. matice $n \times n$. Platí pro $n < 668$

Def: Ekvivalentní kódy:

Kódy C, C' jsou ekvivalentní (strukturálně shodné), pokud

$\exists \Pi$ permutace na $\{1 \dots n\}$ t.ž. $\forall \underbrace{x \in \mathcal{E}^n}_{x_1 \dots x_n}: x \in C \Leftrightarrow (x_{\Pi(1)}, \dots, x_{\Pi(n)}) \in C'$.

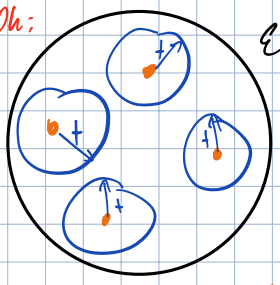
Pro jaké parametry existují kódy?:

Věta: Hammingův odhad:

$\forall$ kód C s parametry $(n, k, 2t+1)_q$ platí:

$$|C| \leq \frac{q^n}{V(t)}$$

Dk:



podle pozor. víme, že $B(x,t)$ $\forall x$ jsou disj.

$$\bigcup_{x \in C} B(x,t) \subseteq \mathcal{E}^n \Rightarrow |C| \cdot V(t) \leq |\mathcal{E}^n| = q^n \quad (\subseteq \Rightarrow \leq)$$

$$\Rightarrow |C| \leq \frac{q^n}{V(t)}. \quad \boxed{X}$$

Df: Kombinatorická koule $B(x,t)$ = (slovo, poloměr)

$$= \{ y \in \mathcal{E}^n : d(x,y) \leq t \}$$

Pozorování: $\forall C$ (kód) se vzdáleností $2t+1$:

$$\forall \underset{x \neq y}{x,y \in C} : B(x,t) \cap B(y,t) = \emptyset$$

Sporem: Nechť $z \in B(x,t) \cap B(y,t)$. (díky kouli)

Díky Δnerov.: $d(x,y) \leq d(x,z) + d(z,y) \leq t + t = 2t$

$2t+1 \not\leq 2t$ ↯

Objem koule $B(x,t)$: (tedy počet slov v kouli)

$$V(t) = |B(x,t)| = \sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i$$

(i je vzdálenost; $\binom{n}{i}$ výběr i různých se pozic; $(q-1)^i$ do vybraných symbolů)

Df: Perfektní kód:

Je kód, který s parametry $(n, k, 2t+1)_q$ s $\frac{q^n}{V(t)}$ kódovými slovy.

Věta: Gilbertův - Varshamův odhad:

$$\forall n, d, q \in \mathbb{N} \quad \exists \text{ kód } C : |C| \geq \frac{q^n}{V(d-1)}$$

Dk:

C vyrobíme hladově, z $\mathcal{E}^n$ odebíráme kódová slova x spolu s $B(x, d-1)$, to lze iterovat po $\frac{q^n}{V(d-1)} = \frac{\#\text{ slov}}{\text{objem koule}}$ krocích.

Lineární kódy:

– vekt. podprostor prostoru $\mathbb{K}^n$, kde $\mathbb{K}$ = konečné těleso, značíme $[n, k, d]_q$.

($\mathbb{K}^n$: $\mathbb{K}$ = $\mathcal{E}$, n = k; $\mathbb{K}$ = $\mathcal{E}$)

$$|C| = |\mathbb{K}^{\dim(C)}| \quad (k \in \mathbb{N})$$

Př.: a) Opakovací kód $\mathcal{E} = \mathbb{Z}_q$ ✓

b) Charakter. vekt. kódů x nejsou lin. kódy nad $\mathbb{Z}_2$

c) Hadamardovy kódy: obecně x, ze Sylvestera ✓

Min. vzdálenost $d = MIN\left(d(x,y)\ \forall x,y : x \neq y\right)$

V lin. kódu: $\forall x,y,z \in C : d(x,y) = d(x+z, y+z) = d(x-y, 0)$,

($z = -y$; $x-y \in C$, $0 \in C$)

pak $d = MIN\left(d(x,0) : x \in C \setminus \{0\}\right)$. → je to daleko rychlejší.

→ pro kódování

Def: Buď C – lineární kód. Generující matice kódu C je $M \in \mathbb{F}_q^{k \times n}$, kde řádky jsou vektory báze C.

Na $\mathbb{F}_q^n$ definujeme sk. součin: pro slova $x = (x_1 \dots x_n)$, $y = (y_1 \dots y_n) \in \mathbb{F}_q^n$

jako $\langle x, y \rangle = \sum_{i=1}^{n} x_i \cdot y_i$ (neplatí $\langle x, x \rangle = 0 \iff x = 0$)

např.: $x = (1,1,0)$ nad $\mathbb{Z}_2$

Def: Duální kód $C^\perp$ k C je ortogonální doplněk kódu C.

$$C^\perp \{ x \in \mathbb{F}_q^n : \langle x, y \rangle = 0 : \forall y \in C \}$$

→ může být $C \cap C^\perp \neq \{0\}$

Platí: $\dim(C) + \dim(C^\perp) = n$

$(C^\perp)^\perp = C$

Def: Kontrolní matice $M^\perp$, generující matice kódu $C^\perp$

→ pro dekódování

Máme $C = \{ y \in \mathbb{F}_p^n : M^\perp y = 0 \}$

→ je to jádrem maticového zobrazení $M^\perp$

Mějme $C =$ kód $[n,k,d]_q$

Kódování:

vstup: $z \in \mathbb{F}_q^k$, chceme vytvořit kódové slovo $x \in C \subseteq \mathbb{F}_q^n$

BÚNO: Generující matice M je ve standardním tvaru

$$M = \underbrace{\left[\, \overbrace{I_k}^{k} \;\middle|\; B \,\right]}_{n} \Big\} k$$

→ díky Gaussově eliminaci + permutaci sloupců
<→ to furt drží ekv. kód

Kódujeme: $x = M^T z$

$$M^T z = \begin{matrix} k \{ \\ n-k \{ \end{matrix} \begin{bmatrix} I_k \\ B^T \end{bmatrix} \overbrace{\begin{bmatrix} z \end{bmatrix}}^{k} \Big\}k = \begin{bmatrix} z \\ \end{bmatrix} \Big\} n$$

→ informační symboly
→ kontrolní symboly

Dekódování:

Odešle kódované $x \in C$, příjemce obdrží $y \in \mathbb{F}_q^n$, chce from y dostat z.

Buď $M^\perp$ kontrolní matice C. Pokud generující M byla ve standardním tvaru, tak

$$M^\perp = \overbrace{\left[-B^T \,\middle|\, I_{n-k} \right]}^{n} \Big\} n-k$$

– proč? $M^\perp \cdot M^T = -B^T \cdot I_k + I_{n-k} \cdot B^T = -B^T + B^T = 0$

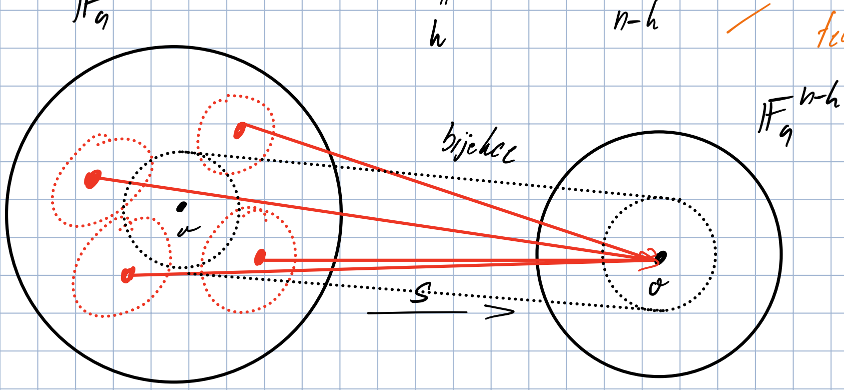
Df: Syndrom slova y je $s(y)$, kde $s: \mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$, zobrazení $s(y) = M^\perp y$

$C = \mathrm{Ker}(s)$

Vlastnosti syndromu s:

- je „na": $\underbrace{\dim(\mathrm{Ker}(s))}_{\dim(C) = k} + \underbrace{\dim(\mathrm{Im}(s))}_{= n-k} = \dim(\mathbb{F}_q^n) = n$

→ ta dimenze je ale stejná jako $\mathbb{F}_q^{n-k}$, tedy musí být na.



tedy existuje na okolí počátku s^{-1} pro $s \upharpoonright B(0,t)$

Lemma: Zobrazení s je prosté na $B(0,t)$, kde $t = \left\lfloor \frac{d-1}{2} \right\rfloor$

Dk: $y, y' \in B(0,t)$, $y \neq y'$. $d(y,y') \leq \underbrace{d(0,y)}_{\leq t} + \underbrace{d(0,y')}_{\leq t} \leq 2t$ → platí Δ nerovnost

Spor: Nechť s není prosté a $s(y) = s(y')$. Pak $0 = s(y) - s(y') = s(y-y')$ (lineární → ten, že patří do jádra zob. → to je ale kód)

$\Rightarrow y - y' \in C$

$2t+1 \leq d \leq d(0, y-y') \leq d(y,y') \leq 2t$ (z volby t; pokud $y-y' \neq 0$) $\Rightarrow y - y' = 0 \Rightarrow y = y'$ ↯

Víme:

1) $s(y) = s(y-x)$ (→ chyba) $\quad s(y-x) = s(y) - s(x) = s(y)$ $\quad s(x) = 0$, $C = \ker(s)$, $x \in C$

2) Chybu $y-x$ lze vyjádřit pomocí $s(y-x)$ (→ nemuselo moc chyb → tudíž je s prosté)

Mějme $y \in B(x,t) \Rightarrow y - x \in B(0,t)$, $y - x = \underbrace{s^{-1}}_{id}(s(y-x))$

Dekódování dopředu: (→ obdrženo → bez vzniklé chyby)

$x = y - (y-x) \overset{2)}{=} y - s^{-1}(s(y-x)) \overset{1)}{=} y - s^{-1}(s(y))$

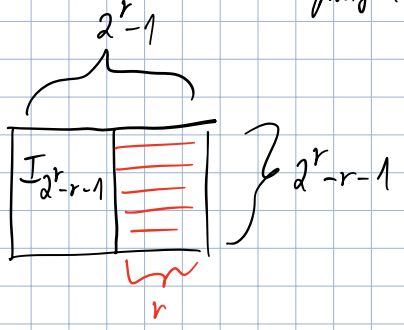
$x = y - s^{-1}(s(y))$

Za předpokladu, že nemuselo příliš chyb.

Pro přijaté $y \in \mathbb{F}_q^n$ spočítej syndrom $s(y) = M^{\perp} y$, zjisti $s^{-1}(s(y))$ pomocí tabulky s q^{n-k} prvky zachycující s^{-1}

Hammingovy kódy:

- perfektní lineární kód
- $q = 2$, parametr $r \geq 2$

generující matice $M =$ 

Parametry:

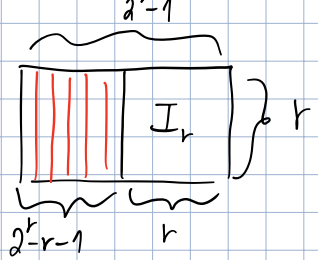
$n = 2^r - 1$

$k = 2^r - r - 1$

$d = 3$

$q = 2$

kontrolní matice $M^{\perp}$. všechny nenulové vektory $\in \mathbb{F}_2^r \setminus \{$vekt. kan. bází$\}$



Tvrzení: V $[n,k,d]_q$ kódu je d = min. počet LZ sloupců kontrolní matice.

Dk: $d = \min_{\substack{x \in C \\ x \neq 0}} d(x,0)$

$x \in C \Leftrightarrow M x^T = 0$

↪ lin. kombinace sloupců M^T určené nenulovými prvky x dá nulový vektor, tedy je to LZ. ☒