Gütekriterien für Code

- erkennt und korrigiert möglichst viele Fehler (e Stück)
  (großer Minimalabstand, $|B_e(\bar{c})|$ wird groß mit wachsendem e und n)
- möglichst viele Codewörter (im Verhältnis zur Länge n)
  $|C|$ bzw $\dim C$, falls C linear
  ($\rightsquigarrow$ Packungsproblem: $B_e(\bar{c})$ möglichst dicht für $\bar{c} \in C$)
- Ver- und Entschlüsselung möglichst effektiv (zeitsparend)
  (Tabellalgorithmen sind tendenziell langsam;
  schnellere Algorithmen setzen innere "Struktur des Codes voraus)
  z.B. algebraisch

Schranken

Binomialkoeffizient $\binom{n}{i} = \frac{n!}{i!\,(n-i)!}$

q-ärer Code C der Länge n

$$|B_r(\bar{c})| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i$$

($(q-1)^i$: Anzahl möglicher Einträge)

Abstand zu $\bar{c}$

$\binom{n}{i}$: Anzahl der Möglichkeiten für die i Stellen, an denen Fehler auftreten

$q=2$ $\quad |B_r(\bar{c})| = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{r}$

Gilbert-Schranke: Gegeben n, q, d. Dann gibt es einen (linearen) q-ären Code der Länge n und vom Minimalabstand $\geq d$ mit

mindestens $\quad q^n / \underbrace{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}_{= |B_{d-1}(\bar{c})|} =: k$

$= |H(n,q)|$

Beweis: C Code vom Minimalabstand $\geq d$, $|C| < k$

Dann $|C| \cdot |B_{d-1}(\bar{c})| < q^n$, also existiert $\bar{x} \in \mathbb{F}_q^n$ mit

$d(\bar{x}, \bar{c}) \geq d$ für alle $\bar{c} \in C$

Dann ist $C \cup \{\bar{x}\}$ größerer Code vom Minimalabstand $\geq d$.

linearer Fall: C linearer Code wie oben, $|C| < k$

Folge: $C' = \langle C, \bar{x} \rangle$ hat Minimalabstand $\geq d$
(der von C und $\bar{x}$ erzeugte Unterraum)

$\Downarrow$

$\alpha\bar{x} + \beta\bar{c}$ $\quad \alpha, \beta \in \mathbb{F}_q$, $\bar{c} \in C$

$d(\alpha\bar{x} + \beta\bar{c}, \bar{0}) = ?$

1. Fall $\alpha = 0$ $\quad d(\alpha\bar{x} + \beta\bar{c}, 0) = d(\underbrace{\beta\bar{c}}_{\in C}, 0) \geq d$ nach Annahme

2. Fall $\alpha \neq 0$

$$d(\alpha\bar{x} + \beta\bar{c}, 0) = d\left(\bar{x}, \underbrace{-\frac{\beta}{\alpha} \cdot \bar{c}}_{\in C}\right) \geq d \quad \text{nach Wahl von } \bar{x}$$

$\Rightarrow$ Minimalgewicht von $C' \geq d$ ☑

Bsp: $q = 2$, $n = 7$, $d = 3$ Gilbert-Schranke: $2^7 / \sum_{i=0}^{2} \binom{7}{i} = 128/(1+7+21) \geq 4{,}4$

linearer Fall: $|C|$ linear, $|C| \geq 4{,}4$
C Untervektorraum von $\mathbb{F}_2^7$, d.h. $|C| =$ 2er Potenz $\Rightarrow |C| \geq 8$

Hamming-Schranke Jeder q-äre Code der Länge n und von Minimalabstand $\geq d$ hat höchstens $q^n \Big/ \sum_{i=0}^{\left\lfloor \frac{d-1}{2} \right\rfloor} \binom{n}{i} (q-1)^i$ Codewörter

(Beim: perfekter Code $\Leftrightarrow$ " $=$ ")

Bsp: a) $q=2, n=7, d=3$ $\quad 2^7 / (1+7) = 16$

b) $q=2, n=6$ $\quad 2^6 = 64$

$\binom{6}{0} + \binom{6}{1} + \binom{6}{2} + \dots$

$= 1 + 6 + 15 + 20 + 15 + 6 + 1$

7

22

42

57

63

64

einzige Teiler von 64:

1 und 64

entspricht trivialen Codes

↳ alle Elemente von $\mathbb{H}(n,q)$ sind Codewörter Minimalabstand 1

↳ nur 1 Codewort Minimalabstand $\geq n$

Linear Codes

C linearer Code der Dimension k, Wortlänge n,
wird beschrieben durch eine Erzeugermatrix G
$(k \times n)$-Matrix, deren Zeilen Basis von C (nicht eindeutig)

Im Beispiel des $[7,4,3]$-Hamming-Codes

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

Kodierung: eigentliche Information ist Zahl zwischen 0 und 15, z.B. 11
Schreibe Binärdarstellung als Spaltenvektor,
z.B. $11_{10} = 1011_2 \rightsquigarrow v_{11} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$

Code von 11 ist $G^T \cdot v_{11} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$

Durch Basiswechsel / Skalierung kann man stets erreichen, dass

C die Form $(Id_k \mid A)$ hat (evtl. Übergang zu äquivalentem Code)

C kann auch beschrieben werden durch eine Prüfmatrix H, d.h.

$H \cdot \bar{c} = 0$ für alle $\bar{c} \in C = \mathrm{Kern}(H)$

und Zeilen von H linear unabhängig

$\dim \mathrm{Bild}(H)$
$= n - \dim \mathrm{Kern}(H)$
$= L - \dim C = n-k$,
d.h. H $((n-k) \times n)$-Matrix

im Bsp: $H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$

beachte: H ist nicht eindeutig!

G $(k \times n)$-Matrix jeweils: linear unabhängige Zeilen

H $((n-k) \times n)$-Matrix

$$G \cdot H^T = 0 \qquad (H \cdot G^T = 0)$$

Bem: Umkehrung gilt: 2 Matrizen mit diesen Eigenschaften sind Erzeuger- und Prüfmatrix eines k-dimensionalen Codes!

H kann in die Form gebracht werden: $(-A^T \mid Id_{n-k})$

falls $G = (Id_k \mid A)$

Decodierung Empfange $\bar{w} \in \mathbb{F}_q^n$

$\bar{w}$ wird zunächst decodiert als das $\bar{c} \in C$ mit $d(\bar{c}, \bar{w}) \leq \frac{d-1}{2}$

„nächstgelegenes Codewort"

d.h. $\bar{w} = \bar{c} + \bar{f}$ $d(\bar{f}, 0) \leq \frac{d-1}{2}$

Berechne $H \cdot \bar{w} = H \cdot (\bar{c} + \bar{f}) = \underbrace{H \cdot \bar{c}}_{0} + H \cdot \bar{f} = H \cdot \bar{f}$ → $H \cdot \bar{f}$ ist Teilmenge der Syndrome der möglichen Fehler nachgeschalten

„Syndrom von $\bar{w}$"

Falls $G = (Id_k \mid A)$, dann ist die eigentliche Information = die ersten k Stellen von $\bar{c}$

Hamming-Codes

Satz. C linearer Code

C hat Minimalabstand $\geq d$ $\Leftrightarrow$ je $d-1$ Spalten der Prüfmatrix sind linear unabhängig

Spezialfall $d=3$: keine Spalte ist 0
keine Spalte ist ein Vielfaches einer anderen

Def: C Hamming-Code: Mindestabstand 3, maximale Prüfmatrix
d.h. alle Vektoren $\neq 0$ kommen, bis auf ein skalares Vielfaches, als Spalte vor
(d.h. alle eindimensionalen Unterräume haben einen Vertreter als Spaltenvektor)

Spezialfall $q=2$: H besteht aus allen Vektoren $\neq 0$ als Spaltenvektoren

gegeben $\ell = n-k$ (Anzahl der Zeilen von H), dann ist $n = 2^\ell - 1$, k = dim Hamming-Code $= 2^\ell - 1 - \ell$

$\ell = 3$: $n = 2^3 - 1 = 7$, $k = 7 - 3 = 4$

$\ell = 4$: $n = 2^4 - 1 = 15$, $k = 15 - 4 = 11$

Hamming-Codes sind perfekte Codes vom Minimalabstand 3
d.h. korrigieren einen Fehler.

Bem für Decodierung im Fall $q = 2$
Fehler $f = e_i$ Standardbasisvektor
Syndrom von e_i ist i-te Spalte von H

Bem: Nach Definition von H sind die Spaltenvektoren alle verschieden,
d.h. das Syndrom gibt eindeutig an, an welcher Stelle der zu
korrigierende Fehler aufgetreten ist

oder umgekehrt: $i \neq j$, $d(e_i, e_j) = 2$ (bzw $\leq 2e$ im allgemeinen Fall)

also $d(e_i - e_j, 0) = ?$

d.h. $e_i - e_j \notin C$, somit $H \cdot (e_i - e_j) \neq 0$

$\Leftrightarrow He_i \neq He_j$ wollgemeiner

Liste aller perfekten q-ären Codes, q = Primzahlpotenz:

korrigiert e Fehler

- trivialer Code $[n, 0, n]$ (nur 1 Wort) $e = n$

  $[n, n, 1]$ (Wiederholungscode, alle Wörter sind Codewörter) $e = 0$

- Hamming-Codes $\left[\frac{q^\ell - 1}{q-1}, \frac{q^\ell - 1}{q-1} - \ell, 3\right]$ $e = 1$

(einige nicht-lineare Codes mit gleichen Parametern)

- binärer Wiederholungscode ungerader Länge $[2e+1, 1, e]$ $e \in \mathbb{N}$

  (nur 2 Wörter: $(\underbrace{0, 0, \ldots, 0}_{2e+1})$ $(\underbrace{1, 1, \ldots, 1}_{2e+1})$)

- binärer Golay-Code ($q=2$) $[23, 12, 7]$ $e = 3$

  ternärer — " — ($q=3$) $[11, 6, 5]$ $e = 2$

q kein Primfaktorpotenz: man weiß sehr wenig

allgemeines Problem: sehr gute (nicht perfekte) Codes finden
viele offene Fragen

Lüke bohmert „Codierungstheorie" (Vieweg)

Conway, Sloane „Sphere Packings, Lattices and Groups" (Springer)