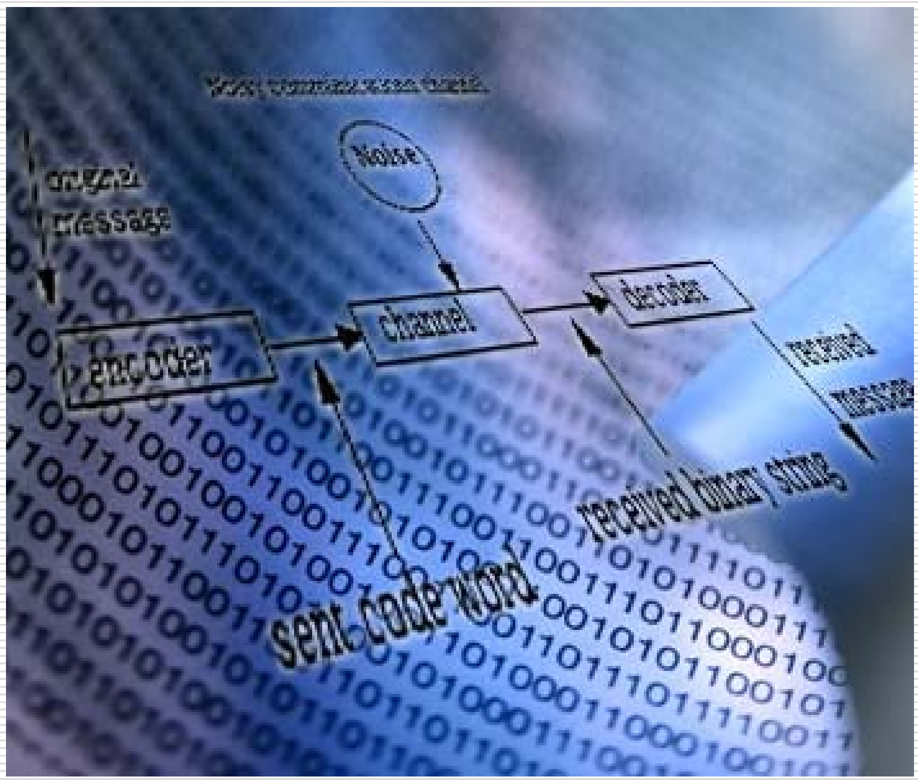
Kapitel 13: Syndromcodierung / Hamming Codes



Ziele des Kapitels

- Lineare Codes
- Zyklische Codes

Parity-Check-Matrix

- Theorem: Die Minimaldistanz eines linearen Codes mit Parity-Check-Matrix H ist gleich der minimalen Anzahl von Spalten in H, die linear abhängig sind

Parity-Check-Matrix

- Beweis: Wenn H insgesamt t linear abhängige Spalten h_i enthält, wobei

$$\sum_{i=1}^{N} c_i h_i = 0$$

- Mit höchstens t verschiedenen $c_i \neq 0$, dann ist $[c_1,\dots,c_N]$ ein Codewort mit minimalem Gewicht $w = d_{\min}$

Parity-Check-Matrix

- Umgekehrt, wenn $[c_1,\ldots,c_N]$ ein Codewort mit minimalem Gewicht $w=d_{\min}$
- Dann sind die den Symbolen $c_i\ !=0$ entsprechenden w Spalten von H linear abhängig

Syndromcodierung

- Wir betrachten das Problem der Fehlerkorrektur für lineare Codes
- Fehlerkorrektur ist immer komplexer als Fehlerdetektion
- Der allgemeine Aufwand zum Vergleich eines empfangenen Codewortes mit allen möglichen Einträgen ist $O(q^K)$
- Für lineare Codes gibt es ein effizienteres Verfahren in $O(q^{N-K})$
- Dazu verwenden wir folgende Annahme

Syndromcodierung

- Das Codewort *c* werde im Kanal durch einen Fehlervektor *e* in ein Wort (*c*+*e*) verfälscht
- Durch Multiplikation mit *H* erhalten wir das so-genannte Syndrom *s*, so dass

$$s = (c+e)\cdot H^{\mathrm{T}} = c\cdot H^{\mathrm{T}} + e\cdot H^{\mathrm{T}} = 0 + e\cdot H^{\mathrm{T}} = e\cdot H^{\mathrm{T}}$$

- Bei Annahme von Gleichverteilung der Codewörter kann Fehlerkorrektur wie folgt erreicht werden:
  - Wir schätzen *e* aufgrund von *s*
  - Wir korrigieren das empfangene Codewort durch Subtraktion
- Implementierung mittels einer Tabelle, die jedem der $O(q^{N-K})$ Syndrome ein Fehlermuster minimalen Gewichts zuweist

Syndromcodierung

- Gegeben: Binärer [5,2] Code mit

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

I_K $\quad$ A

$$H = \begin{bmatrix} 1 & 1 & \vdots & 1 & 0 & 0 \\ 1 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix} \quad n\text{-}k$$

- Auflisten aller möglichen $2^{N-K}=8$ Syndrome

Syndrom	Muster (e)	
000	00000	$w=0$
001	00001	$w=1$
010	00010	$w=1$
011	00011	$w=2$

Jeder Spalte der Matrix H wird ein Fehlervektor zugeordnet bei dem 1 Bit gesetzt ist

Syndromcodierung

Syndrom	Muster (e)	
100	00100	$w=1$
101	01000	$w=1$
110	00110	$w=2$
111	10000	$w=1$

- Prinzip: Minimales Gewicht w_{min} bezüglich aller mit Syndrom konsistenter Fehlermuster

$$r = c + e = 11110 \text{ empfangen}$$

$$\rightarrow s = r \cdot H^T = 100 \rightarrow e = 00100 \text{ (aus der Syndromtabelle)}$$

$$\rightarrow r - e = 11\ 010 = c \rightarrow a = 11$$

Parity-Check Prüfbits

Anmerkungen

- Bei einem Linearcode entsteht dann und nur dann eine nichterkennbare Verfälschung, wenn $w(e) \geq d_{\min}$ ist
- Dann könnte e ein Kanalcodewort sein
- Gilt die Bedingungen zur Mindestdistanz, so ist r ein Kanalcodewort, wenn

$$r \cdot H^{\mathrm{T}} = 0$$

- Sollen alle Fehlermuster korrigierbar sein, deren Distanz $w(e) \leq s$ ist, so kann man die Anzahl der $(N\text{-}K)$ erforderlichen Kontrollbits berechnen

Anzahl von Kontrollbits

- Wir berechnen die Mindestanzahl der erforderlichen Kontrollbits
- Gegeben sei ein Kanalcode mit Codewörtern $[c_1,\ldots,c_N]$ der Länge N, sowie $d_{\min}$
- Aufgrund von Verfälschungen gibt es

$$\binom{N}{i} = \frac{N!}{i!(N-i)!}$$

- Binärfolgen der Distanz i
- Davon sind nur 2^K Folgen Kanalwörter mit Distanz $\geq d_{min}$

Anzahl von Kontrollbits

- Damit alle Fehler ≤ s korrigierbar sind, muss gelten

$$2^N \geq 2^K \left(1 + \binom{N}{1} + \binom{N}{2} \cdots + \binom{N}{r} \right)$$

- Damit ist K berechenbar
- Es gilt:

$$2^K \geq \sum_{i=1}^{\left\lfloor \frac{d_{\min}-1}{2} \right\rfloor} \binom{N}{i}$$

Anzahl von Kontrollbits

- Kanalcode mit $N=7$ und $d_{\min}=3$. Berechnen K
- Ein Fehler korrigierbar, zwei erkennbar

$$2^7 \geq 2^K \left(1 + \binom{7}{1}\right)$$

$$2^7 \geq 2^3 \cdot 2^4$$

$$K \leq 4$$

- Das heisst bei 4 Nutzbits erhält man 16 Quellcodewörter sowie 3 Kontrollbits pro Wort

Hamming Codes

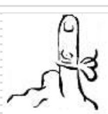
- Eine bedeutende Klasse linearer Codes sind die Hamming Codes
- Für jedes $r > 1$ gibt es einen Code der Länge $N=2^r-1$ mit $K=N-r$ Informationsbits und $d_{min}=3$
- Je grösser K, desto näher an *1* ist die Rate
- Jeder Hamming-Code kann einen Fehler pro Block korrigieren
- Je länger der Block, desto kleiner die tolerierbare Fehlerwahrscheinlichkeit des Kanals
- Beliebt bei kleinen Fehlerraten

Hamming Codes

- Die (r x (2^r-1)) Parity-Check-Matrix kann einfach konstruiert werden
- Sie besteht aus allen 2^r-1 vom Nullvektor verschiedenen Spalten
- Die Reihenfolge der Spalten ist damit noch nicht festgelegt
- Die Minimaldistanz folgt aus der Tatsache, dass sich keine zwei Spalten von H zu Null addieren, da sie verschieden sind
- Es gibt jedoch viele Spaltentripel, die linear abhängig sind

Praktische Konstruktion

- Damit kann man einen Hamming-Code wie folgt konstruieren:
  1. Lege r und damit N fest, z. B. $r=3$, $N=7$
  2. Schreibe alle 2^r-1 vom 0-Vektor verschiedenen Spalten in H
  3. Ordne sie strukturell gemäss $H=[-A^T|I_{N-K}]$
  4. Extrahiere A^T aus $H=[-A^T|I_{N-K}]$
  5. Konstruiere die Generatormatrix $G=[I_K|A]$



Dies ist offensichtlich eine sehr einfache Methode, einen fehlerkorrigierenden Code zu konstruieren

Hamming - Code

- [7,4] das heisst $r=3$ und $N=7$
- $\mathrm{Dim}(H)=r\mathrm{x}(2^r-1)=3\mathrm{x}7$ Matrix

$$H=\left[-A^T \mid I_{N-K}\right]=\left[\begin{array}{cccc:ccc} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$-A^T$

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

Hamming - Code

$$G = [I_K \mid A] = \left[\begin{array}{cccc:ccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]$$

Duale Hamming Codes

- Interessante Codes erhält man, wenn man die Parity-Check-Matrix eines Codes zur Generatormatrix eines weiteren Codes macht
- Definition: Ein Code C' ist dual zu einem Code C, wenn die Generatormatrix von C' eine Parity-Check-Matrix von C ist
- Im dualen Hamming-Code ist die Anzahl der Informationsbits kleiner und die Minimaldistanz sehr gross
- Es gilt für den dualen Hamming-Code

$$d_{\min} = 2^{r-1}$$