

Kapitel 12: Codierungstheorem und Fehlerkorrektur



Ziele des Kapitels

ETH

- Shannon'sches Kanalcodierungstheorem
- Fehlerkorrektur
- Parity-Check-Matrix

Fehlerwahrscheinlichkeit

ETH

- Annahme (o.B.d.A.): $[U_1, \dots, U_K]$ sind zufällig und gleichverteilt.
- Damit gilt: $H(U^K) = K$
- Wir betrachten die mittlere Bitfehlerwahrscheinlichkeit, d.h. den Bruchteil der falschen Bits am Ausgang

$$\bar{P}_e = \frac{1}{K} \sum_{i=1}^K P[\tilde{U}_i \neq U_i]$$



Eine beliebige Bitfolge einer Informationsquelle kann zunächst ideal komprimiert werden. Nach dieser Quellcodierung approximiert die Bitfolge der Codewörter eine zufällige, gleichverteilte Bitfolge beliebig gut. Warum?

Fehlerwahrscheinlichkeit

ETH

- Es gilt

$$H(U^K | \tilde{U}^K) \geq K - NC = K \left(1 - \frac{C}{R}\right)$$
- Hieraus erhalten wir mit der Fano-Ungleichung den ersten Teil des Shannon'schen Kanalcodierungstheorems

Theorem Shannon, Teil 1

ETH

- **Theorem (Shannon I):** Wenn ein DGK mit Kapazität C zur Übertragung echt zufälliger Informationsbits mit Rate $R > C$ benutzt wird, so gilt für die mittlere Bitfehlerwahrscheinlichkeit beim Empfänger

$$h(\bar{P}_e) \geq 1 - \frac{C}{R}$$



Dieses Theorem erlaubt es uns, eine einfache Unterschranke für den Fehler anzugeben, wenn wir mit Raten oberhalb der Kapazität übertragen

Theorem Shannon, Teil 1

ETH

- Der zweite Teil des Theorems von Shannon enthält ein verblüffendes Resultat
- Die Lehrmeinung war, dass die Kommunikation mit zunehmender Rate schlechter wird
- Dies ist unterhalb der Kapazität nicht der Fall!
- $R < C$ ist also nicht nur eine **notwendige** Bedingung, sondern auch eine **hinreichende** für zuverlässige Kommunikation

Theorem Shannon, Teil 2 ETH

- **Theorem (Shannon II):** Gegeben sei ein DGK mit Inputalphabet χ , Outputalphabet γ und Kapazität C .
- Für jede Rate $R < C$ und für jedes $\varepsilon > 0$ existiert für genügend grosse Blocklänge N ein Code C mit

$$M = \lfloor 2^{R \cdot N} \rfloor$$

Codewörtern, für den die maximale Decodierfehlerwahrscheinlichkeit über alle Codewörter kleiner als ε ist:

$$\max_{1 \leq j \leq M} P(F | X^N = c_j) < \varepsilon$$

Interpretation ETH

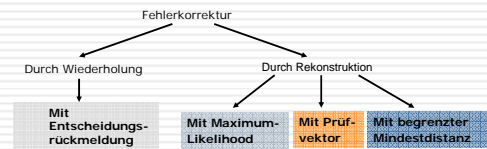
- Beweis für BSK im Skript
- Entwurf eines guten Codes ist leichter, als angenommen
- Auch eine zufällige Wahl eines Codes liefert mit hoher Wahrscheinlichkeit einen Code, der das Theorem erfüllt
- Zufällig heisst, dass alle MN Symbole unabhängig nach einer Verteilung P_X gewählt werden, welche $I(X; Y)$ maximiert und die Kapazität des Kanals erreicht
- Beim BSK sind alle Codewörter unabhängige, zufällige Bitfolgen der Länge N

Interpretation ETH

- In der Praxis werden Codes dennoch gezielt konstruiert
- Decodierung unstrukturierter Codes ist in der Regel sehr ineffizient
- Es müssen alle Codewörter durchsucht werden
- Effiziente Decodierungsverfahren sind nur für Codes mit starker algebraischer Struktur bekannt
- Unzählige Folgearbeiten zur Konstruktion praktischer Codes
- Seit 1993 Aufbruch in Richtung der Shannon-Grenze mit „Turbo-Codes“

Prinzipien der Fehlerkorrektur ETH

- Redundanz ist Voraussetzung für ein Korrekturverfahren
- Wir unterscheiden folgende Prinzipien



Fehlerkorrektur ETH

- Bei **Wiederholung** fügt der Kanalcodierer Kontrollinformation hinzu, welche eine Fehlerdetektion ermöglicht
- Falls Fehler, wird nochmals übertragen
- Es kann sein, dass Fehler nicht erkannt werden
- Bei Rekonstruktion wird der Fehler sowohl erkannt, als auch lokalisiert
- Durch die Lokalisierung kann der Fehler dann korrigiert werden

Fehlerkorrektur ETH

- Die Redundanz entscheidet über **Restfehlerwahrscheinlichkeit**
- Bei **Maximum Likelihood** wird jeweils das wahrscheinlichste Zeichen im Codewortalphabet gewählt
- Bei **begrenzter Mindestdistanz** wird nur korrigiert, wenn ein Zeichen innerhalb einer Korrekturkugel liegt
- Bei **Prüfvektoren** enthält der Decoder Strings, mit denen er testen kann, ob ein Zeichen zum Codealphabet gehört

Kriterien zum Codeentwurf **ETH**

- ❑ **Restfehlerwahrscheinlichkeit:** Sie bestimmt die Güte unter der Bedingung des verwendeten Kanals
- ❑ **Zeit:** Diese umfasst die Zeit zur Codierung, Decodierung und Fehlerkorrektur, incl. dem Rückkanal
- ❑ **Aufwand:** Dieser beschreibt den (schaltungstechnischen oder algorithmischen) Aufwand zur Realisierung eines Decodierers

Blockcodes **ETH**

- ❑ **Definition:** Ein Blockcode C mit Blocklänge N für einen Kanal mit Inputalphabet χ ist eine Teilmenge $C = \{c_1, \dots, c_M\}$ von χ^N der N -Tupel über χ
- ❑ Typischerweise betrachten wir Codes mit $M = q^K$ Codewörtern für eine ganze Zahl $K < N$ für $|\chi| = q$
- ❑ Wir betrachten nur binäre Codes $q=2$
- ❑ Für CDs beispielsweise verwendet man $q=256$
- ❑ Mit einem Code werden K q -äre Zeichen in Codewörtern der Länge N codiert
- ❑ K/N beschreibt den Anteil der Informationsbits, hängt also mit der Redundanz zusammen

Blockcodes **ETH**

- ❑ K/N heisst auch die **dimensionslose Rate**
- ❑ $C_2 = \{00000, 11100, 00111, 11011\}$ ist ein binärer Code mit $N=5$ und $K=2$
- ❑ $C_3 = \{0000, 0112, 1100, 0221, 2200, 1212, 2012, 1021, 2121\}$ ist ein ternärer Code mit $N=4$ und $K=2$
- ❑ **Definition:** Die **Hammingdistanz** $d(a, b)$ zweier Wörter a und b ist die Anzahl von Positionen, in denen sich a und b unterscheiden.
- ❑ Die **Minimaldistanz** $d_{\min}(C)$ eines Codes C ist die kleinste Hammingdistanz zwischen zwei Codewörtern

Blockcodes **ETH**

- ❑ $d_{\min}(C_2) = 3, d_{\min}(C_3) = 2$
- ❑ Ein Code kann r Fehler detektieren, wenn für jedes Codewort und jedes Fehlermuster mit höchstens r Fehlern ein Fehler festgestellt werden kann
- ❑ Dies ist genau dann, wenn $d_{\min}(C) \geq r + 1$
- ❑ Ein Code kann s Fehler korrigieren, wenn für jedes Codewort und jedes Fehlermuster mit höchstens s Fehlern das Codewort wieder eindeutig gefunden werden kann
- ❑ Dies ist genau dann, wenn $d_{\min}(C) \geq 2s + 1$

Blockcodes **ETH**

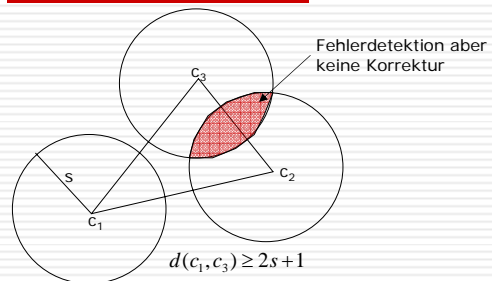
- ❑ Die Minimaldistanz ist also ein wichtiges Mass zur Worst-Case Betrachtung eines Codes
- ❑ **Theorem:** Ein Code mit Minimaldistanz d erlaubt, $d-1$ Fehler zu detektieren oder

$$\text{floor}\left(\frac{d-1}{2}\right)$$

Fehler zu korrigieren

- ❑ Man spricht auch oft von Korrekturkugeln

Korrekturkugeln **ETH**



$$d(c_1, c_3) \geq 2s + 1$$

$$d(c_2, c_3) < 2s + 1$$

aber $d(c_2, c_3) \geq s + 1 \rightarrow$ Detektion möglich

Lineare Codes

ETH

- Neben grosser Minimaldistanz ist auch effiziente Codierbarkeit und Decodierbarkeit wichtig
- Lineare Codes erlauben eine sehr effiziente Codierung, allerdings nicht unbedingt eine effiziente Decodierung
- **Definition:** Ein **linearer Blockcode** mit q^K Codewörtern der Blocklänge N über einem endlichen Körper $GF(q)$ ist ein Unterraum der Dimension K des Vektorraums der N -Tupel über $GF(q)$
- Ein solcher Code wird als $[N, K]$ -Code bezeichnet

Lineare Codes

ETH

- Eine alternative Definition kann ebenfalls in der Literatur gefunden werden
- Dieser trägt der algebraischen Struktur des Codes direkt Rechnung
- **Definition:** Ein **linearer Blockcode** ist ein Code, bei dem der Codierer zur Transformation von Quellencodewörtern a der Länge K in Kanalcodewörter c der Länge N ausschliesslich Operationen verwendet, die in der algebraischen Struktur einer Gruppe definiert sind

Linearcodes als Gruppen

ETH

- Das Codealphabet bestehe aus den Wörtern
 $a_0 = (00000)$ $a_1 = (10010)$ $a_2 = (01011)$ $a_3 = (00101)$
 $a_4 = (11001)$ $a_5 = (10111)$ $a_6 = (01110)$ $a_7 = (11100)$
zeige, dass dieser Code die Gruppenaxiome erfüllt
- Axiom G1: Abgeschlossenheit
 $a_0 \oplus a_1 = a_1$, $a_1 \oplus a_3 = a_5$
 $a_2 \oplus a_3 = a_6$, $a_2 \oplus a_5 = a_7$
 $a_5 \oplus a_6 = a_4$, usw.

Linearcodes als Gruppen

ETH

- Axiom G2: Assoziatives Gesetz
 $(a_1 \oplus a_2) \oplus a_3 = a_1 \oplus (a_2 \oplus a_3)$
 $(a_4 \oplus a_5) \oplus a_3 = a_4 \oplus (a_5 \oplus a_3)$
usw.
- Axiom G3: Neutrales Element
 $a_1 \oplus a_0 = a_1$
 $a_2 \oplus a_0 = a_2$
usw.

Linearcodes als Gruppen

ETH

- Axiom G4: Inverses Element
 $(a_1 \oplus a_1) = a_0$
 $(a_2 \oplus a_2) = a_0$
usw.
- Ferner gilt das Kommutativgesetz
 $a_1 \oplus a_2 = a_2 \oplus a_1$
 $a_1 \oplus a_3 = a_3 \oplus a_1$
usw.

Lineare Codes

ETH

- Jeder lineare Code enthält das Nullwort 0
- Die Distanz eines Codewortes c zu 0 wird **Hamminggewicht** $w(c)$ genannt
- **Theorem:** Die Minimaldistanz eines linearen Codes ist gleich dem minimalen Hamminggewicht eines von 0 verschiedenen Codewortes
- Beweis: Seien c_1 und c_2 zwei Codewörter mit minimaler Distanz
- Aufgrund der Vektorraum-Eigenschaft ist die Differenz $c_1 - c_2$ ebenfalls ein Codewort mit Gewicht d_{\min}
- Umgekehrt ist d_{\min} nicht grösser als $w(c_1 - c_2)$, da die 0 im Code ist

Generatormatrix

ETH

- Die Abbildung der Informationsvektoren $a = [a_1, \dots, a_K]$, dargestellt als K -Tupel, auf die q^K Codewörter $c = [c_1, \dots, c_N]$ kann mittels der $K \times N$ **Generatormatrix** G dargestellt werden (modulo-Multiplikation)

$$c = a \cdot G$$

Generatormatrix (1)

ETH

- Die Generatormatrix G von C_2 ist

$$G = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$c([2,1]) = [2 \ 1] \cdot \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{bmatrix} = [1 \ 0 \ 2 \ 1]$$

Generatormatrix (2)

ETH

- Kanalcodealphabet mit $N=7$, $q=2$ und $K=3$

$$c_0 = (0000000)$$

$$c_1 = (1001101)$$

$$c_2 = (0101010)$$

$$c_3 = (0010010)$$

$$c_4 = c_1 \oplus c_2 = (1100111)$$

$$c_5 = c_1 \oplus c_3 = (1011111)$$

$$c_6 = c_2 \oplus c_3 = (0111000)$$

$$c_7 = c_1 \oplus c_2 \oplus c_3 = (1110101)$$

Generatormatrix (2)

ETH

- C ist ein Unterraum der Grösse 2^3
- V ist 2^7 gross
- Mögliche G sind:

$$G_1 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Generatormatrix

ETH

- Die Zeilen der Matrix bilden eine Basis des Codes
- Da Vektorräume verschiedene Basen besitzen, gibt es auch verschiedene Generatormatrizen für gleiche Codes
- Eine speziell geeignete Form enthält die ersten K Spalten als Einheitsmatrix I_K

$$G = [I_K \ A]$$

- A ist eine $K \times (N-K)$ Matrix.
- Eine solche Generatormatrix heisst **systematisch**



Dies bedeutet, dass das Codewort aus K Informationssymbolen besteht und $N-K$ angefügten "Parity-Checks"

Parity-Check-Matrix

ETH

- Eine weitere bedeutende Matrix ist die **Parity-Check- oder Kontrollmatrix**
- Definition:** Eine $(N-K) \times N$ -Matrix H heisst Parity-Check-Matrix eines linearen $[N, K]$ -Codes C , falls

$$c \in C \Leftrightarrow cH^T = 0$$

- Die Zeilen von H spannen den $(N-K)$ -Unterraum der Vektoren v auf, für die $cv^T = 0$



Dies bedeutet, dass wir damit schnell prüfen können, ob ein Codewort am Kanalausgang auch Element von C ist

Parity-Check-Matrix

ETH

- **Theorem:** Sei $G=[I_K|A]$ eine systematische Generatormatrix eines linearen Codes. Die $(N-K) \times N$ -Matrix $H=[-A^T|I_{N-K}]$ ist eine Parity-Check Matrix des Codes

- Teil 1: Zeige, dass für jedes Codewort $c=aG$, gilt

$$c \cdot H^T = 0$$

- Teil 2: Wenn, $cH^T=0$, zeige, dass c ein Codewort ist
- Beweis: Triviale Matrizenalgebra sowie Aufspalten von c in $[c_1|c_2]$