Institut für Nachrichtentechnik

Grundlagen der Informationstechnik - Nachrichtentechnik

Vorlesung: Eduard A. Jorswieck
Übung: Dr. Bile Peng

Wintersemester 2023-2024, 30. November 2023

Entscheidungstheorie I

Kapitel 7 in M. Bossert 'Einführung in die Nachrichtentechnik'

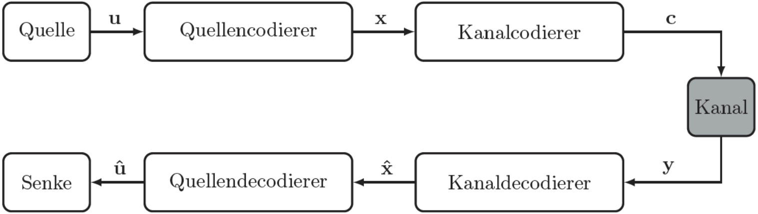

Abbildung 1: Der Kanal im Modell der Informationstheorie

- Nun betrachten wir den Fall ohne Kanalcodierung $\boldsymbol{x} = \boldsymbol{c}$ und nehmen an, dass der Kanal $f_{Y|X}(y|x)$ gegeben ist.

Entscheidungstheorie II

Satz: Wahrscheinlichkeiten von Hypothesen

Seien X und Y zwei Zufallsvariablen mit Wahrscheinlichkeitsdichten $f_X(x)$ und $f_Y(y)$, sowie der bedingten Wahrscheinlichekitsdichte $f_{Y|X}(y|x)$. Wenn der Wert y_b beobachtet wurde, ist die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert x_h hat, mit der Regel von *Bayes*:

$$f_{X|Y}(x_h|y_b) = \frac{f_{Y|X}(y_b|x_h) \cdot f_X(x_h)}{f_Y(y_b)}.$$

Entscheidungstheorie III

- Die Maximierung über alle möglichen Hypothesen ergibt die Entscheidung x_d mit

$$x_d = \arg \max_{x_h \in \mathcal{A}_X} f_{X|Y}(x_h|y_b).$$

Maximum a-posteriori Entscheider (MAP)

Die MAP Entscheidung mit perfekter Kenntnis der bedingten Übergangswahrscheinlichkeit des Kanals, ist

$$x_d = \arg \max_{x_h \in \mathcal{A}_X} f_{X|Y}(x_h|y_b) = \max_{x_h \in \mathcal{A}_X} f_{X|Y}(y_b|x_h) \cdot f_X(x_h).$$

Entscheidungstheorie IV

Beispiel MAP-Entscheider

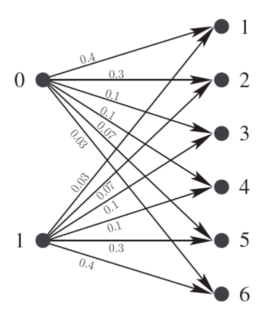

Abbildung 2: Multi-Ausgangs Kanalmodell

Entscheidungstheorie V

Maximum-Likelihood Entscheider (ML)

Die ML-Entscheidung unter der Annahme, dass X gleichverteilt ist, ergibt sich durch

$$x_d = \max_{x_h \in \mathcal{A}_X} f_{X|Y}(y_b|x_h).$$

- Unter der Annahme der Gleichverteilung ist die ML-Entscheidung equivalent zu der MAP-Entscheidung.

Entscheidungstheorie VI

- Das grundlegende Theorem zur Fehlerwahrscheinlichkeit bei einer Entscheidung stammt von **Neyman und Pearson** von 1933.
- Man geht von einer binären Hypothese aus $x_h = 0$ oder $x_h = 1$.
- Man kennt die bedingten Wahrscheinlichkeiten für $y \in \mathcal{A}_Y$ und die beiden Hypothesen als $f_{Y|X}(y|0)$ und $f_{Y|X}(y|1)$.
- Der Beobachtungsraum $\mathcal{A}_Y$ wird in Entscheidungsgebiete $\mathcal{A}_0$ und $\mathcal{A}_1$ aufgeteilt, so dass $\mathcal{A} = \mathcal{A}_0 \cup \mathcal{A}_1$ und $\mathcal{A}_0 \cap \mathcal{A}_1 = \emptyset$.
- Fehlerwahrscheinlichkeiten vom Typ I und Typ II sind

$$\alpha = \sum_{y \in \mathcal{A}_1} f_{Y|X}(y|0) \qquad \text{und} \qquad \beta = \sum_{y \in \mathcal{A}_0} f_{Y|X}(y|1). \tag{1}$$

Entscheidungstheorie VII

Neyman-Pearson Theorem

Sei $\theta \in \mathbb{R}$ eine Entscheidungsschwelle und die Entscheidungsregionen seien

$$\mathcal{A}_0(\theta) = \{y : f_{Y|X}(y|1) \leqslant f_{Y|X}(y|0)\exp(-\theta)\} \tag{2}$$

$$\mathcal{A}_1(\theta) = \{y : f_{Y|X}(y|1) > f_{Y|X}(y|0)\exp(-\theta)\}. \tag{3}$$

Für die damit definierten Fehlerwahrscheinlichkeiten von Typ I und Typ II (α und β) gilt, dass für jedes $\theta' \neq \theta$ entweder

$$\alpha' < \alpha \quad \text{und} \quad \beta' > \beta$$

oder

$$\alpha' > \alpha \quad \text{und} \quad \beta' < \beta.$$

Entscheidungstheorie VIII

- Es gibt also einen Abtausch zwischen Fehlern vom Typ I und Typ II.
- Oft wird die Summe $\alpha + \beta$ minimiert.
- Schreiben wir das Entscheidungsgebiet durch das *Log-Likelihood Verhältnis*

$$\mathcal{A}_0(\theta) = \left\{ y : \log \frac{f_{Y|X}(y|0)}{f_{Y|X}(y|1)} \geqslant \theta \right\},$$

so ergibt sich der Erwartungswert bezüglich der bedingten Wahrscheinlichkeitsdichte $f_{Y|X}(y|0)$ des Log-Likelihood Verhältnisses zur relativen Entropie

$$D(f_{Y|X}(y|0) || f_{Y|X}(y|1)) = \sum_y f_{Y|X}(y|0) \log \frac{f_{Y|X}(y|0)}{f_{Y|X}(y|1)}.$$

Kanalcodierung I

Kapitel 8 in M. Bossert 'Einführung in die Nachrichtentechnik'

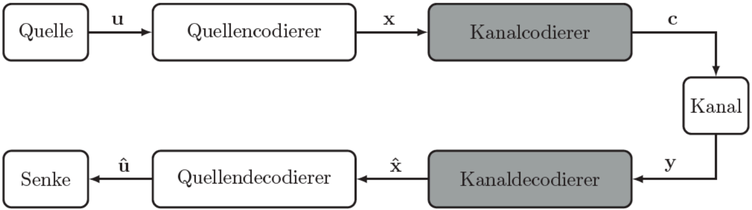

Abbildung 3: Der Kanal im Modell der Informationstheorie

- Jetzt konzentrieren wir uns auf den Kanalcodierer und -decodierer.

Kanalcodierung II

- Hinzufügen von *Redundanz* erlaubt es, Fehler bei der Übertragung zu erkennen und womöglich zu korrigieren.
- Eine einfache Möglichkeit zur Fehlererkennung besteht im Hinzufügen einer Prüfsumme.
- Einer Folge von k Bits sollen $n-k$ Bits Redundanz hinzugefügt werden. Dazu werden die k Bits als Koeffizienten eines binären Polynoms $\hat{c}(x)$ geschrieben. Die Daten in die oberen k Koeffizienten und die unteren Koeffizienten werden zu Null gesetzt:

$$\hat{c}(x) = 0 + 0x + 0x^2 + \ldots + 0x^{n-k+1} + \hat{c}_{n-k}x^{n-k} + \ldots + \hat{c}_{n-1}x^{n-1}.$$

- Die Redundanz wird mit Hilfe der Polynomdivision berechnet. (Cyclic Redundancy Check - CRC)

Kanalcodierung III

1. Wir wählen eine Generatorpolynom $g(x)$ vom Grad $n-k$, d.h., $g(x) = g_0 + g_1 x + g_2 x^2 + ... g_{n-k} x^{n-k}$.
2. Wir teilen nun $\hat{c}(x)$ durch das Generatorpolynom $g(x)$ und erhalten einen Quotienten $q(x)$ und den Rest $r(x)$ mit Grad $< n-k$: $\hat{c}(x) : g(x) = q(x)$ Rest $r(x)$.
3. Das Codewort $c(x)$ wird berechnet durch
$$c(x) = \hat{c}(x) - r(x).$$
4. Die Daten bleiben unverändert, da der Grad $r(x)$ kleiner als $n-k$ ist.
5. Das Codewort $c(x)$ wird durch $g(x)$ ohne Rest geteilt: $c(x) = g(x)q(x)$.
6. Bei vorliegender Folge von Bits kann sofort einfach überprüft werden, ob es ein gültiges Codewort (ohne Fehler) ist durch Teilen mit $g(x)$.

Kanalcodierung IV

Satz: Nicht erkennbare Fehler bei CRC

Fehler $e(x) \neq 0$, die durch das Generatorpolynom $g(x)$ ohne Rest teilbar und damit gültige Codeworte sind, können nicht erkannt werden.

Definition: binärer Vektorraum $\mathbb{F}_2^n$

Alle Vektoren $\boldsymbol{a}$ der Länge n und Komponenten aus $\mathbb{F}_2$ stellen den binären Vektorraum $\mathbb{F}_2^n$ über $\mathbb{F}_2$ dar. Wir schreiben

$$\boldsymbol{a} = [a_0, a_1, \ldots, a_{n-1}] \in \mathbb{F}_2^n, a_i \in \mathbb{F}_2, i = 0, 1, \ldots, n-1.$$

Kanalcodierung V

- Die Menge $\mathbb{F}_2$ bildet einen *Körper* bezüglich der Addition und Multiplikation.

Definition: Hamming-Distanz

Seien $\boldsymbol{a} = [a_0, a_1, ..., a_{n-1}]$ und $\boldsymbol{b} = [b_0, b_1, ..., b_{n-1}]$ zwei binäre Vektoren der Länge n. Die Hamming-Distanz zwischen $\boldsymbol{a}$ und $\boldsymbol{b}$ ist definiert durch

$$\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=0}^{n-1} \mathrm{dist}(a_i, b_i) \quad \text{mit} \quad \mathrm{dist}(a_i, b_i) = \begin{cases} 1 & \text{für } a_i \neq b_i \\ 0 & \text{sonst} \end{cases}.$$

- Die Distanz $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{0}) = \mathrm{wt}(\boldsymbol{a})$ ist das Hamming-Gewicht des Vektors $\boldsymbol{a}$. Das ist die Anzahl der Komponenten, die nicht Null sind.

Kanalcodierung VI

- Die Hamming-Distanz ist eine *Metrik* und erfüllt die folgenden Eigenschaften:
  1. Positive Definitheit: $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) \geqslant 0$ und $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) = 0$ genau dann, wenn $\boldsymbol{a} = \boldsymbol{b}$.
  2. Symmetrie: $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) = \mathrm{dist}(\boldsymbol{b}, \boldsymbol{a})$.
  3. Dreiecksungleichung: $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) \leqslant \mathrm{dist}(\boldsymbol{a}, \boldsymbol{c}) + \mathrm{dist}(\boldsymbol{c}, \boldsymbol{b})$.

Definition: Linearität eines Codes

Ein Code heißt linear, wenn die Linearkombination von zwei beliebigen Codeworten $\boldsymbol{a}, \boldsymbol{c} \in \mathcal{C} \subseteq \mathbb{F}_2^n$ ebenfalls ein Codewort ist. Im binären Fall:

$$\forall \boldsymbol{a}, \boldsymbol{c} \in \mathcal{C} : u\boldsymbol{a} + v\boldsymbol{c} \in \mathcal{C}, \ u, v \in \mathbb{F}_2.$$

Kanalcodierung VII

Definition: Mindestdistanz d

$\mathcal{C} \subseteq \mathbb{F}_2^n$ sei ein Code. Dann heißt

$$d = \min \operatorname{dist}(\boldsymbol{a}, \boldsymbol{c}), \quad \boldsymbol{a}, \boldsymbol{c} \in \mathcal{C}, \quad \boldsymbol{a} \neq \boldsymbol{c}$$

die Mindestdistanz des Codes.

Definition: Parameter binärer linearer Blockcode $\mathcal{C}(n, k, d)$

Ein binärer linearer Blockcode $\mathcal{C}(n, k, d) \subseteq \mathbb{F}_2^n$ über dem Alphabet $\mathbb{F}_2$ hat die Länge n und die Dimension k, d.h., $|\mathcal{C}| = 2^k$. Die Mindestdistanz ist $d = \min \operatorname{dist}(\boldsymbol{a}, \boldsymbol{c}))$, $\boldsymbol{a}, \boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{a} \neq \boldsymbol{c}$.

Institut für Nachrichtentechnik

Kanalcodierung VIII

Definition: Generatormatrix $\boldsymbol{G}$

Mit der $k \times n$ Generatormatrix $\boldsymbol{G}$ vom Rang k werden die 2^k Codewörter $\boldsymbol{c}$ durch Multiplikation der 2^k möglichen Informationswörter $\boldsymbol{i} \in \mathbb{F}_2^k$ mit der Generatormatrix $\boldsymbol{G}$ erzeugt:

$$\mathcal{C} = \{\boldsymbol{c} \in \mathbb{F}_2^n | \boldsymbol{c} = \boldsymbol{i} \cdot \boldsymbol{G}\}.$$

Definition: Prüfmatrix $\boldsymbol{H}$

Eine $(n-k) \times n$ Matrix $\boldsymbol{H}$ definiert einen Code durch

$$\mathcal{C} = \{\boldsymbol{c} \in \mathbb{F}_2^n | \boldsymbol{H}\boldsymbol{c}^T = \boldsymbol{0}^T\},$$

wobei $\boldsymbol{c}^T$ der transponierte, d.h., Spaltenvektor, ist.

Kanalcodierung IX

Satz: Prüfmatrix und Mindestdistanz

Eine Code $\mathcal{C}(n, k, d) \subseteq \mathbb{F}_2^n$ besitzt die Mindestdistanz d genau dann, wenn beliebige $d-1$ Spalten der Prüfmatri $\boldsymbol{H}$ linear unabhängig sind und d Spalten existieren, die linear abhängig sind.

Definition: Parity-Check (PC)-Code $\mathcal{C}(n, n-1, 2)$

Ein PC-Code hängt an die $k = n-1$ Bits des Informationswortes ein Bite an, so dass das Hamming-Gewicht des Codeworts gerade ist. Die Decodierung erfolgt durch

$$c_{n-1} = \sum_{j=0}^{n-2} i_j \mod 2, \; c_j = i_j, j = 0, \dots, n-2.$$

Kanalcodierung X

Definition: Wiederholungs-Code oder Repetition-Code $\mathcal{C}(n, 1, n)$

Der Wiederholungscode besteht aus zwei Codeworten, dem Allnullwort und dem Alleinswort. Das Informationsbit 0 oder 1 wird n-mal wiederholt. Die Dimension ist damit $k = 1$ und die Mindestdistanz ist $d = n$.

Definition: Hamming-Code

Ein Code, dessen Prüfmatrix aus allen binären Vektoren außer dem Nullvektor $\mathbf{0}^T$ besteht, heißt Hamming-Code

$$\mathcal{C}_H(n = 2^h - 1, k = n - h, d = 3).$$

Kanalcodierung XI

Definition: Systematische Codierung

Eine Codierung heißt systematisch, wenn in jedem Codewort $\boldsymbol{c}$ das zugehörige Informationswort $\boldsymbol{i}$ unverändert vorkommt.

Definition: Zyklische Codes

Ein Code $\mathcal{C}$ heißt zyklisch, wenn für alle Codeworte $\boldsymbol{c} \in \mathcal{C}$ gilt:

$$\left(c_0, c_1, \ldots, c_{n-1}\right) \in \mathcal{C} \iff \left(c_{n-1}, c_0, c_1, \ldots, c_{n-2}\right) \in \mathcal{C}.$$

- Zyklische Codes können acuh durch Polynome statt Matrizen beschrieben werden und bieten daher praktisch einige Vorteile.