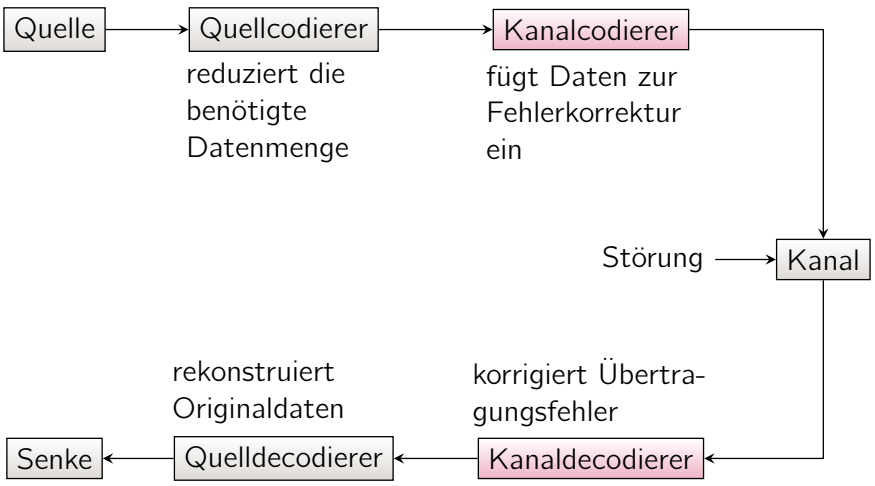
7 Grundlagen der Kanalcodierung

7 Grundlagen der Kanalcodierung



7.1 Einfache Kanalcodes

Beispiel 7.1 (Wiederholungscode)

Der Kanalcodierer führt die Codierung

$$0 \mapsto 000, \qquad 1 \mapsto 111$$

durch; die Decodierung erfolgt gemäß

$$\begin{aligned} 000 &\mapsto 0 & 111 &\mapsto 1 \\ 001 &\mapsto 0 & 110 &\mapsto 1 \\ 010 &\mapsto 0 & 101 &\mapsto 1 \\ 100 &\mapsto 0 & 011 &\mapsto 1 \end{aligned}$$

(Mehrheitsentscheidung). Man erkennt:

- Die Datenmenge wird verdreifacht.
- Einzelne Bitfehler können korrigiert werden.
- Bei zwei Bitfehlern kann erkannt werden, dass mindestens ein Fehler aufgetreten ist.

Beispiel 7.2 (Parity Check)

Der Kanalcodierer führt die Codierung

$$00 \mapsto 000, \qquad 01 \mapsto 011, \qquad 10 \mapsto 101, \qquad 11 \mapsto 110$$

durch; es wird also ein Bit ergänzt, so dass die Gesamtzahl der Einsen gerade wird. Die Decodierung erfolgt gemäß

$$\begin{aligned} 000 &\mapsto 00 & 001 &\mapsto ? \\ 011 &\mapsto 01 & 010 &\mapsto ? \\ 101 &\mapsto 10 & 100 &\mapsto ? \\ 110 &\mapsto 11 & 111 &\mapsto ?. \end{aligned}$$

Man erkennt:

- Die Datenmenge erhöht sich um 50%.
- Einzelne Bitfehler können erkannt, aber nicht korrigiert werden.

7.2 Decodierung

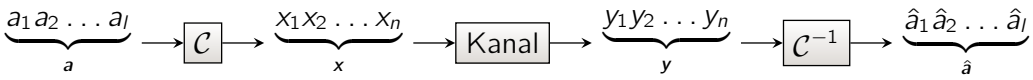
Definition 7.1

Eine (binäre) Kanalcodierung $\mathcal{C} : \{0,1\}^l \to \{0,1\}^n$ bildet je l Bits einer (quellcodierten) Bitsequenz auf $n > l$ Bits zur Übertragung über einen gestörten Kanal ab. Die Bildmenge $C \subset \{0,1\}^n$ wird als Code bezeichnet.

Definition 7.2

Die entsprechende Decodierung $\mathcal{C}^{-1} : \{0,1\}^n \to \{0,1\}^l$ rekonstruiert die ursprünglichen l Datenbits aus den n gestörten Bits am Kanalausgang.

Im Folgenden verwendete Nomenklatur:



Die Decodierung soll die Wahrscheinlichkeit maximieren, dass $\hat{\boldsymbol{a}} = \boldsymbol{a}$, also

$$\mathcal{C}^{-1}(\boldsymbol{y}) = \operatorname*{argmax}_{\hat{\boldsymbol{a}}} p_{X^n|Y^n}\big(\mathcal{C}(\hat{\boldsymbol{a}})|\boldsymbol{y}\big) \tag{7.1}$$

$$= \operatorname*{argmax}_{\hat{\boldsymbol{a}}} \frac{p_{X^n}\big(\mathcal{C}(\hat{\boldsymbol{a}})\big)}{p_{Y^n}(\boldsymbol{y})} p_{Y^n|X^n}\big(\boldsymbol{y}|\mathcal{C}(\hat{\boldsymbol{a}})\big) \qquad \text{(Bayes)} \tag{7.2}$$

$$= \operatorname*{argmax}_{\hat{\boldsymbol{a}}} p_{X^n}\big(\mathcal{C}(\hat{\boldsymbol{a}})\big) p_{Y^n|X^n}\big(\boldsymbol{y}|\mathcal{C}(\hat{\boldsymbol{a}})\big) \qquad (p_{Y^n}(\boldsymbol{y}) \text{ von } \hat{\boldsymbol{a}} \text{ unabh.}) \tag{7.3}$$

und für gleichverteilte Nutzdaten $\boldsymbol{a}$, also konstantes $p_{X^n}\big(\mathcal{C}(\hat{\boldsymbol{a}})\big) = 2^{-l}$, folgt

$$\mathcal{C}^{-1}(\boldsymbol{y}) = \operatorname*{argmax}_{\hat{\boldsymbol{a}}} p_{Y^n|X^n}\big(\boldsymbol{y}|\mathcal{C}(\hat{\boldsymbol{a}})\big). \tag{7.4}$$

Für den symmetrisch gestörten Binärkanal mit Fehlerwahrscheinlichkeit p gilt

$$p_{Y^n|X^n}(\boldsymbol{y}|\boldsymbol{x}) = p^{\mathrm{d}(\boldsymbol{x},\boldsymbol{y})} \cdot (1-p)^{n-\mathrm{d}(\boldsymbol{x},\boldsymbol{y})}, \tag{7.5}$$

wobei $\mathrm{d}(\boldsymbol{x},\boldsymbol{y})$ die Hamming-Distanz zwischen $\boldsymbol{x}$ und $\boldsymbol{y}$, also die Anzahl der gestörten Bits angibt.

Definition 7.3

Die Hamming-Distanz $\mathrm{d}(\boldsymbol{x},\boldsymbol{y})$ bezeichnet die Anzahl der Stellen, in denen sich $\boldsymbol{x}$ und $\boldsymbol{y}$ unterscheiden.

Für $p < 0{,}5$ ist $p_{Y^n|X^n}(\boldsymbol{y}|\boldsymbol{x})$ streng monoton fallend in $\mathrm{d}(\boldsymbol{x},\boldsymbol{y})$ und damit

$$\mathcal{C}^{-1}(\boldsymbol{y}) = \operatorname*{argmax}_{\hat{\boldsymbol{a}}} p_{Y^n|X^n}\left(\boldsymbol{y}|\mathcal{C}(\hat{\boldsymbol{a}})\right) = \operatorname*{argmin}_{\hat{\boldsymbol{a}}} \mathrm{d}\left(\mathcal{C}(\hat{\boldsymbol{a}}),\boldsymbol{y}\right). \tag{7.6}$$

Der Decoder muss also das Codewort $\hat{\boldsymbol{x}} = \mathcal{C}(\hat{\boldsymbol{a}})$ suchen, das von dem empfangenen $\boldsymbol{y}$ in den wenigsten Stellen abweicht.

Beispiel 7.3

Wie lautet der Decodierer zu der Codierung gemäß

$$00 \mapsto 000, \qquad 01 \mapsto 001, \qquad 10 \mapsto 011, \qquad 11 \mapsto 111,$$

wenn gleichverteilte Nutzdaten und ein symmetrisch gestörter Kanal mit $p = \frac{1}{10}$ angenommen werden?
Die Codewörter selbst werden trivial decodiert:

$$000 \mapsto 00, \qquad 001 \mapsto 01, \qquad 011 \mapsto 10, \qquad 111 \mapsto 11.$$

Aus den Distanzen $\mathrm{d}\left(\mathcal{C}(\boldsymbol{a}),\boldsymbol{y}\right)$

$\boldsymbol{a}$ / $\mathcal{C}(\boldsymbol{a})$ / $\boldsymbol{y}$	00	01	10	11
	000	001	011	111
010	1	2	1	2
100	1	2	3	2
101	2	1	2	1
110	2	3	2	1

ergibt sich die Decodiervorschrift

$$010 \mapsto 00 \text{ oder } 10, \qquad 100 \mapsto 00, \qquad 101 \mapsto 01 \text{ oder } 11, \qquad 110 \mapsto 11.$$

7.3 Kenngrößen von Kanalcodes

Definition 7.4 (Coderate)

Das Verhältnis von Nutzdatenbits l zu Codewortlänge n wird als Coderate $R = \frac{l}{n}$ bezeichnet.

- Der Wiederholungscode aus Beispiel 7.1 hat die Coderate $R = \frac{1}{3}$.
- Der Parity-Check-Code aus Beispiel 7.2 hat die Coderate $R = \frac{2}{3}$.

Definition 7.5 (Minimaler Abstand)

Der minimale Abstand $\mathrm{d}(C)$ eines Codes C ist die minimale Hamming-Distanz zweier Codewörter

$$\mathrm{d}(C) = \min_{\substack{x,x' \in C \\ x \neq x'}} \mathrm{d}(\boldsymbol{x}, \boldsymbol{x'}). \tag{7.7}$$

- Der Wiederholungscode aus Beispiel 7.1 hat den minimalen Abstand $\mathrm{d}(C) = 3$.
- Der Parity-Check-Code aus Beispiel 7.2 hat den minimalen Abstand $\mathrm{d}(C) = 2$.

Satz 7.1

Bei der Decodierung eines Codes C können genau dann bis zu t Fehler sicher detektiert werden, wenn $t < \mathrm{d}(C)$.

Beweis.

Seien $\boldsymbol{x}$ das gesendete und $\boldsymbol{y}$ das empfangene Codewort. Falls $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = t < \mathrm{d}(C)$, so folgt unmittelbar $\boldsymbol{y} \notin C$, der Übertragungsfehler wird also erkannt. Gilt andererseits $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = t \geq \mathrm{d}(C)$, so könnte $\boldsymbol{y} = \boldsymbol{x'} \in C$ sein, also fälschlicherweise eine korrekte Übertragung von $\boldsymbol{x'}$ erkannt werden. □

Satz 7.2

Bei der Decodierung eines Codes C können genau dann bis zu t Fehler sicher korrigiert werden, wenn $2t < \mathrm{d}(C)$.

Beweis.

Seien $\boldsymbol{x}$ das gesendete und $\boldsymbol{y}$ das empfangene Codewort. Korrekte Decodierung erfordert $t = \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) < \mathrm{d}(\boldsymbol{x}', \boldsymbol{y})$ für alle $\boldsymbol{x}' \in C, \boldsymbol{x}' \neq \boldsymbol{x}$.
Die Dreiecksungleichung liefert

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{d}(\boldsymbol{y}, \boldsymbol{x}') \geq \mathrm{d}(\boldsymbol{x}, \boldsymbol{x}') \qquad \text{bzw.} \qquad \mathrm{d}(\boldsymbol{y}, \boldsymbol{x}') \geq \mathrm{d}(\boldsymbol{x}, \boldsymbol{x}') - \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}). \tag{7.8}$$

Somit erfordert sicher korrekte Decodierung

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{x}') - \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) > \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{bzw.} \qquad \mathrm{d}(\boldsymbol{x}, \boldsymbol{x}') > 2\,\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}), \tag{7.9}$$

also $\mathrm{d}(\boldsymbol{x}, \boldsymbol{x}') > 2t$ für alle $\boldsymbol{x}' \in C, \boldsymbol{x}' \neq \boldsymbol{x}$, mithin $\mathrm{d}(C) > 2t$. □

7.4 Perfekte Codes

Lemma 7.1

Bei einem Code der Länge n gibt es

$$N = \sum_{i=0}^{t} \binom{n}{i} \tag{7.10}$$

Wörter $\boldsymbol{y}$ zu einem Codewort $\boldsymbol{x}$ mit $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) \leq t$.

Beweis.

Es gibt

$$N_i = \frac{n!}{i! \cdot (n-i)!} = \binom{n}{i} \tag{7.11}$$

Wörter $\boldsymbol{y}$ mit $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = i$. Summation über alle Distanzen $i = 0, \ldots, t$ liefert das gesuchte Ergbenis. □

Satz 7.3 (Hamming-Schranke)

Für einen Code, der t Fehler korrigieren kann, müssen die Anzahl der Nutzbits l und die Codewortlänge n die Bedingung

$$2^{n-l} \geq \sum_{i=0}^{t} \binom{n}{i} \tag{7.12}$$

erfüllen.

Beweis.

Da der Code t Fehler korrigieren kann, wird jedes empfangene $\boldsymbol{y}$, für das $\mathrm{d}(\boldsymbol{y}, \boldsymbol{x}) \leq t$, zu $\boldsymbol{x}$ decodiert. Nach Lemma 7.1 gibt es $\sum_{i=0}^{t} \binom{n}{i}$ solche $\boldsymbol{y}$.
Da die Mengen der $\boldsymbol{y}$, die zu den insgesamt 2^l verschiedenen $\boldsymbol{x}$ decodiert werden, disjunkt sind, es aber 2^n verschiedene $\boldsymbol{y}$ gibt, folgt

$$2^l \sum_{i=0}^{t} \binom{n}{i} \leq 2^n. \tag{7.13}$$

□

Definition 7.6

Codes, die die Hamming-Schranke mit Gleichheit erfüllen, heißen perfekte Codes.

Beispiel 7.4

Der Wiederholungscode aus Beispiel 7.1 hat die Länge $n = 3$, trägt $l = 1$ Nutzdatenbit und kann $t = 1$ Fehler korrigieren. Es gilt

$$2^{n-l} = 2^2 = 4$$

sowie

$$\sum_{i=0}^{t} \binom{n}{i} = \binom{3}{0} + \binom{3}{1} = 1 + 3 = 4,$$

der Code ist also perfekt.

Beispiel 7.5

Für den Parity-Check-Code aus Beispiel 7.2 gilt $n = 3$, $l = 2$, $t = 0$, und damit $2^{n-l} = 2^1 = 2$ und

$$\sum_{i=0}^{t} \binom{n}{i} = \binom{3}{0} = 1,$$

der Code ist also *nicht* perfekt.

7.5 Zweites Shannonsches Codierungstheorem

Lemma 7.2

Die Wahrscheinlichkeit, dass sich zwei zufällig gewählte Wörter $\boldsymbol{x}$ *und* $\boldsymbol{y}$ *der Länge* n *in maximal* t *Stellen unterscheiden, beträgt*

$$P\big(\mathrm{d}(\boldsymbol{x},\boldsymbol{y}) \leq t\big) = \frac{1}{2^n} \sum_{i=0}^{t} \binom{n}{i}. \tag{7.14}$$

Beweis.

Nach Lemma 7.1 gilt für

$$\sum_{i=0}^{t} \binom{n}{i}$$

von 2^n möglichen $\boldsymbol{y}$, dass $\mathrm{d}(\boldsymbol{x},\boldsymbol{y}) \leq t$. □

Lemma 7.3

Die Wahrscheinlichkeit, dass bei der Übertragung von n *Bits über einen symmetrisch gestörten Binärkanal mit der Fehlerwahrscheinlichkeit* p *für die Anzahl der gestörten Stellen* $\frac{1}{n}|t - np| < \epsilon$ *gilt, geht mit beliebig kleinem* $\epsilon > 0$ *für* $n \to \infty$ *gegen 1.*

Beweis.

Der Erwartungswert der Anzahl t der gestörten Bits beträgt offenbar np und nach dem Gesetz der großen Zahlen folgt die Behauptung. □

Lemma 7.4

$$\sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} \leq 2^{-n\cdot(p\,\mathrm{ld}\,p+(1-p)\,\mathrm{ld}(1-p))} \qquad \text{für } 0 \leq p \leq 0{,}5 \tag{7.15}$$

Beweis.

Mit $\bar{p} = 1 - p$ gilt

$$1 = (p+\bar{p})^n = \sum_{i=0}^{n} \binom{n}{i} p^i \bar{p}^{n-i} \geq \sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} p^i \bar{p}^{n-i}, \tag{7.16}$$

und weil

$$p^i \bar{p}^{n-i} = \underbrace{\left(\frac{p}{\bar{p}}\right)^i}_{\leq 1} \bar{p}^n \geq \underbrace{\left(\frac{p}{\bar{p}}\right)^{np}}_{\leq 1} \bar{p}^n = p^{np}\bar{p}^{n-np} = p^{np}\bar{p}^{n\bar{p}} = \left(p^p \bar{p}^{\bar{p}}\right)^n \tag{7.17}$$

folgt $1 \geq \sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} \left(p^p \bar{p}^{\bar{p}}\right)^n$, also

$$\sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} \leq \left(p^p \bar{p}^{\bar{p}}\right)^{-n} = 2^{\mathrm{ld}\left(\left(p^p \bar{p}^{\bar{p}}\right)^{-n}\right)} = 2^{-n\left(\mathrm{ld}(p^p)+\mathrm{ld}(\bar{p}^{\bar{p}})\right)} = 2^{-n(p\,\mathrm{ld}\,p+\bar{p}\,\mathrm{ld}\,\bar{p})}. \tag{7.18}$$

□

Satz 7.4 (Zweites Shannonsches Codierungstheorem)

Seien $\delta, \epsilon > 0$, so lässt sich über einen symmetrisch gestörten Binärkanal mit Fehlerwahrscheinlichkeit p und Kapazität $K = 1 + p\,\mathrm{ld}\,p + (1-p)\,\mathrm{ld}(1-p)$ zu jeder hinreichend großen Codewortlänge n ein Code der Rate R, $K > R \geq K - \epsilon$ angeben, sodass die Falschdecodierwahrscheinlichkeit $P_f < \delta$.

Man kann also bei verschwindend geringer Fehlerwahrscheinlichkeit beliebig nah an der Kanalkapazität codieren, wenn man n groß genug wählt.

Beweisskizze.

Der Code C bestehe aus 2^{Rn} *zufällig* gewählten Codewörtern.
Nach Lemma 7.3 gilt $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = np$, es werden bei der Übertragung eines Codeworts $\boldsymbol{x}$ genau np Bits gestört.
Zu einer fehlerhaften Decodierung kommt es, wenn ein $\boldsymbol{x'} \in C$ existiert, sodass $\mathrm{d}(\boldsymbol{x'}, \boldsymbol{y}) < \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = np$.
Daraus ergibt sich

$$P_f \leq \sum_{\substack{\boldsymbol{x'} \in C \\ \boldsymbol{x'} \neq \boldsymbol{x}}} P\left(\mathrm{d}(\boldsymbol{x'}, \boldsymbol{y}) < np\right) \tag{7.19}$$

und da es $2^{Rn} - 1$ Codewörter $\boldsymbol{x'} \neq \boldsymbol{x}$ gibt und diese zufällig gewählt sind, folgt

$$P_f \leq (2^{Rn} - 1) \cdot P\left(\mathrm{d}(\boldsymbol{x'}, \boldsymbol{y}) < np\right) < 2^{Rn} \cdot P\left(\mathrm{d}(\boldsymbol{x'}, \boldsymbol{y}) < np\right). \tag{7.20}$$

Einsetzen von Lemma 7.2 liefert

$$P_f < 2^{Rn} \cdot P\left(\mathrm{d}(\boldsymbol{x'}, \boldsymbol{y}) < np\right) = 2^{Rn} \cdot \frac{1}{2^n} \sum_{i=0}^{np} \binom{n}{i} = 2^{(R-1)n} \sum_{i=0}^{np} \binom{n}{i} \tag{7.21}$$

und mit Lemma 7.4 folgt

$$P_f < 2^{(R-1)n} \sum_{i=0}^{np} \binom{n}{i} \leq 2^{(R-1)n} \cdot 2^{-n \cdot (p \operatorname{ld} p + (1-p) \operatorname{ld}(1-p))} = 2^{(R-K)n}, \tag{7.22}$$

was für große n gegen Null geht, da $R - K < 0$. □

Anmerkungen zum zweiten Shannonschen Codierungstheorem:

- Obwohl nur für den symmetrisch gestörten Binärkanal gezeigt, gilt das Theorem allgemein für alle Kanäle mit Kapazität K.
- Es lässt sich auch zeigen, dass für große n die Fehlerwahrscheinlichkeit eines Codes mit $R > K$ gegen Eins geht.
- Nur für hinreichend große n liefert ein zufällig gewählter Code gute Ergebnisse. Für die Konstruktion eines Codes mit beschränktem n liefert das Theorem keine Hinweise.

7.6 Zusammenfassung

- Falls gleichverteilte Nutzdaten codiert über einen symmetrisch gestörten Kanal übertragen werden, minimiert die Decodierung anhand der Hamming-Distanz die Wahrscheinlichkeit falscher Decodierung.
- Wichtige Kenngrößen eines Codes C sind neben der Codewortlänge n die Coderate R und der minimale Abstand $\mathrm{d}(C)$, der bestimmt, wie viele Fehler korrigiert (oder detektiert) werden können.
- Die Hamming-Schranke gibt eine Grenze für die Beziehung von Codewortlänge n, Anzahl der Nutzdatenbits l und Anzahl der korrigerbaren Fehler t an. Codes, die diese Grenze erreichen, heißen perfekte Codes.
- Über einen Kanal der Kapazität K gibt es für hinreichend große n einen Code, dessen Rate R bei verschwindend geringer Fehlerwahrscheinlichkeit beliebig nah an die K herankommt. (Zweites Shannonsches Codierungstheorem).