Fortsetzung: Kanalcodierung / Blockcodes

Was passiert, wenn in der Prüfsumme, d.h. in den Paritätsbits, Fehler sind?

Daten	1	0	1	1	P_1	P_2	P_3
	_	_		_	0		
	_		_	_		1	
		_	_	_			0
Senden	1	0	1	1	0	1	0
Empfangen	1	0	1	1	1	1	0
Prüfung	1	0	1	1			
	_	_		_	0		
	_		_	_		1	
		_	_	_			0
Syndromwort					1	0	0

In der Matrix zur Erstellung der Paritäten ergibt sich mit diesem Syndromwort folgender Fall:

Syndromwort	Matrix der Paritäten-Erstellung				Paritätenmatrix		
1	_	_		_	_		
0	_		_	_		_	
0		_	_	_			_

Nur beim Paritätsbit befindet sich genau der Fall, dass nur die obere Zeile markiert ist, deswegen muss der Fehler an dieser Stelle vorliegen.

In jener Formel, welche den Zusammenhang zwischen der Codelänge $n = m + k$ und der Prüflänge m für eine Anzahl zu korrigierender Fehler t darstellt, ergibt sich entsprechend:

$2^m \geq n + 1 \Rightarrow 2^3 \geq 7 + 1$ (m = Länge der Prüfsumme, n = Wortlänge im Code)

Für den Fall $t = 2$, $n = 7$ ergibt sich entsprechend

$$2^m \geq 7 + 1 + \binom{7}{2}$$

$$2^m \geq 29 \Rightarrow m = 5$$

Bei $t = 3$ würde ein weiterer Binomialkoeffizient $\binom{7}{3}$ addiert, usw.

- ➔ „Perfekte“ Codes nutzen den Fall der Gleichheit der Formelseiten, weil dadurch die verfügbaren Codezustände perfekt ausgenutzt werden
- ➔ Wichtig ist, zu längeren Rahmen (größeren Werten für n)zu kommen, um die Redundanz im Code gering zu halten (→ Tabelle der letzten Unterrichtseinheit)

Typische Werte für n in der Praxis: Einige Hundert … einige Tausend Bit

Vergleich von kleinem und großen Rahmen:

a. $n = 7, m = 3, k = 4$

   Wir suchen die aus $2^k = 2^4 = 16$ aus $2^n = 2^7 = 128$ Wörtern mit dem Abstand $d_{\min} = 3$.

b. $n = 1000, m = 10, k = 990$

   Wir suchen die aus $2^k = 2^{990} \approx 10^{298}$ aus $2^n = 2^{1000} \approx 10^{301}$ Wörtern mit möglichst gleichem Abstand.

➔ Auf Grund der größeren Wortlänge kann effizienter Codiert werden, es können mehr Wörter als richtige Wörter deklariert werden.

➔ Zahlenwerte-Vergleich:

- Masse des Universuns → $10^{53} ... 10^{54}\ kg$
- Anzahl Atome im Universum → $10^{78} ... 10^{81}$

<u>Zyklischer Hamming Code</u>

- Verwendung von nicht zerlegbaren Polynomen („Primzahlen“ der Polynome), z.B. $x^3 + x + 1$
- Ablauf für einen Code der Länge 7:
  - 1. Zeile ist eine Nullfolge, 2. Zeile enthält das Generatorpolynom
  - 3.-8. Zeile → es wird jeweils im Register um 1 nach links geschoben
  - 9. Zeile enthält Modulo Addition von 2. und 3. Zeile
  - 10.-15. Zeile → es wird jeweils im Register um 1 nach links geschoben
  - 16. Zeile enthält Modulo Addition von 2. und 11. Zeile → Einsfolge
- Eine weitere Eigenschaft des Codes besteht darin, dass er ein Polynom von $x^n + 1$ ist. Dies kann durch eine entsprechende Polynomdivision festgestellt werden. In diesem Fall ist $x^n + 1 = x^7 + 1 = 10000001$:

$$
\begin{array}{l}
10000001 : (1011) \\
\underline{1011} \\
\quad 1100 \\
\quad \underline{1011} \\
\quad\ 1110 \\
\quad\ \underline{1011} \\
\quad\quad 1011 \\
\quad\quad \underline{1011} \\
\quad\quad\ \ 0
\end{array}
$$

- Neben dieser restlosen Division müssen auch alle anderen Zeilen restlos durch das Generatorpolynom (Zeile 2) teilbar sein
- Das Gewicht eines Codes zeigt grafisch dargestellt auch die Distanzverteilung, in der folgenden Tabelle ist diese dargestellt:

Anzahl der Einsen im Codewort	0	3	4	7
Häufigkeit der Worte	1	7	7	1

- Der geringste Abstand zwischen der Nullfolge und einem korrekten Wort ist $d_{\min} = 3$, es existieren aber auch Wörter, deren Abstand größer ist
- Die Codewörter, in denen 4 Einsen stehen, sind auch alle ein Vielfaches des Generatorpolynoms → alle ohne Rest teilbar

➔ Durch dieses Verfahren können korrekte Wörter direkt hergeleitet werden

➔ Daraus lässt sich ein Verfahren zum Finden von Prüfsummen in langen Codes entwickeln

Eine zu übertragende Codesequenz $C(x)$ enthält die um x^{n-k} verschiedenen Informationsstellen und die Paritätsbits (Prüfsumme) $p(x)$.

$$C(x) = I(x) \cdot x^{n-k} \underset{\substack{\uparrow \\ \text{Modulo 2 Addition}}}{+} p(x)$$

Auf beiden Seiten durch $\mod g(x)$ ($g(x)$ = Generatorpolynom) teilen

$$C(x) \bmod g(x) = 0$$

$$\underbrace{p(x)}_{\text{Grad 2}} \bmod \underbrace{g(x)}_{\text{Grad 3}} = p(x)$$

$$0 = I(x) \cdot x^{n-k} \bmod g(x) + p(x)$$

$$p(x) = I(x) \cdot x^{n-k} \bmod g(x)$$

Erklärung:

- $p(x)$ kann nicht durch $g(x)$ geteilt werden, weil der Grad der Prüfsumme (Zähler) kleiner dem Grad des Generatorpolynoms (Nenner) ist
- Die Umformung, dass $p(x)$ auf die andere Seite der Gleichung gebracht wird, funktioniert auf Grund der Modulo 2 Rechnung, welche dazu führt, dass man das Vorzeichen entfernen kann.

Rechenbeispiel: Die ersten 4 Stellen der Zeile 11 werden um drei Stellen nach links verschoben und anschließend per Modulo Rechnung durch das Generatorpolynom geteilt.

$$1110000 : (1011)$$

$$\begin{array}{r} \underline{1011} \\ 1010 \\ \underline{1011} \\ 0100 \end{array}$$

Der Wert 100 entspricht den weiteren 3 Stellen der Zeile 11, also der entsprechenden Prüfsumme für diese Zeile.

Blockfehlerwahrscheinlichkeit (→ Codierungsgewinn)

Um zu klären, wie effizient die Kanalcodierung bei verschiedenen Blocklängen und Fehlerzahlen ist, nutzt man die Betrachtung durch binomialverteilte Zufallsgrößen:

- Die Wahrscheinlichkeit, dass ein Bit falsch ist, wird als p_{Bit} festgelegt
- Die Wahrscheinlichkeit, dass in einem Block n bestimmte Bits falsch sind, beträgt folglich $\left(p_{Bit}\right)^n$
- Die Wahrscheinlichkeit, dass n bestimmte Bits in einem Block korrekt gesendet werden, ist die Umkehrung von $\left(p_{Bit}\right)^n$, also $\left(1-p_{Bit}\right)^n$

Betrachtet man nun den gesamten Block und möchte die Wahrscheinlichkeit dafür angeben, dass es bei einer bestimmten Blocklänge n genau i Fehler passieren, muss man noch einen Binomialkoeffizienten hinzu multiplizieren, welcher dafür sorgt, dass alle möglichen Fälle abgedeckt werden, bei denen die angegebene Situation zutrifft:

$$p_{Block}\left(n,i,p_{Bit}\right)=\binom{n}{i}\cdot\left(p_{Bit}\right)^i\cdot\left(1-p_{Bit}\right)^{n-i}$$

Betrachtet man die Graphen von $p_{Block}\left(i\right)$ bei fester Blocklänge $n=10$ für verschiedenene Wahrscheinlichkeiten p_{Bit}, so wird deutlich, dass bei $p_{Bit}=0{,}1$ der Fehlerfall $i=1$ am wahrscheinlichsten ist, allerdings auch der fehlerfreie Fall und mehrere Fehler durchaus wahrscheinlich sind.
Bei $p_{Bit}=0{,}01$ liegt die höchste Wahrscheinlichkeit beim fehlerfreien Fall, allerdings sind immer noch kleine Werte (also relativ unwahrscheinliche Ereignisse) für die Fehlerfälle zu sehen.

➔ Die Übertragung funktioniert in den meisten Fällen korrekt, es ist aber nie auszuschließen, dass Fehler passieren, auch wenn sie noch so unwahrscheinlich sind.

Numerisches Beispiel anhand eines typischen Rahmens

Gegeben:

$$n=1000,\binom{n}{0}=1,\binom{n}{1}=1000,\binom{n}{2}=499500,\binom{n}{3}=166167000,\ p_{Bit}=10^{-8}$$

Daraus ergibt sich folgende Verteilung der Wahrscheinlichkeiten

$$p_{Block}(0) = \binom{1000}{0} \cdot \left(10^{-8}\right)^0 \cdot \left(1-10^{-8}\right)^{1000} = 1 \cdot 1 \cdot 0,9999 = 0,99 = 99,9\,\%$$

$$p_{Block}(1) = \binom{1000}{1} \cdot \left(10^{-8}\right)^1 \cdot \left(1-10^{-8}\right)^{999} = 1000 \cdot 10^{-8} \cdot 0,9999 = 9,99 \cdot 10^{-6} \approx 10^{-5}$$

$$p_{Block}(2) = \binom{1000}{2} \cdot \left(10^{-8}\right)^2 \cdot \left(1-10^{-8}\right)^{998} = 499500 \cdot 10^{-16} \cdot 0,9999 = 4,99 \cdot 10^{-11}$$

$$p_{Block}(3) = \binom{1000}{3} \cdot \left(10^{-8}\right)^3 \cdot \left(1-10^{-8}\right)^{997} = 166167000 \cdot 10^{-24} \cdot 0,9999 = 1,662 \cdot 10^{-16}$$

- ➔ Die Doppel- oder Dreifachfehler sind sehr unwahrscheinlich, aber trotzdem nicht unmöglich. In den meisten Fällen geht die Übertragung gut, jeder 100000ste Rahmen hat allerdings einen Bitfehler.
- ➔ Qualität macht sich an der Blocklänge fest. Ist diese größer, werden Fehler in Blöcken unwahrscheinlicher.

<u>Effekt der Kanalcodierung</u>

In der folgenden Übersicht wird deutlich, wie eine Fehlerkorrektur durch eine Kanalcodierung sich bei der Übertragung bemerkbar macht und wie die Wahrscheinlichkeit für Fehler und somit Probleme bei der Übertragung verringert werden kann.

Fall:	ohne Kanalcodierung $t=0$	mit Kanalcodierung $t=1$	mit Kanalcodierung $t=2$
$p_{Block}(0)$	gut	gut	gut
$p_{Block}(1)$	Problem	gut	gut
$p_{Block}(2)$	Problem	Problem	gut
$p_{Block}(3)$	Problem	Problem	Problem
Wahrscheinlichkeit für Fehler	10^{-5}	10^{-11}	10^{-16}

- ➔ Die Physik einer Übertragung hat sich in den letzten Jahrzehnten kaum verändert, BER und Rahmenlängen (Blocklängen) auch nicht
- ➔ Das Datenvolumen einer Übertragung ist in den letzten 20 Jahren um den Faktor 10000 gestiegen
- ➔ Fehler tauchen bei größerem Datenvolumen häufiger auf, weil die Blocklängen sich nicht verändert haben, ein 10000ster, fehlerbehafteter Block wird häufiger erreicht

Verändert man die Werte für p_{Bit} z.B. auf 10^{-6}, so spiegelt sich der Unterschied der Größenordnungen in den Wahrscheinlichkeiten für p_{Block} wieder.
Dies wird an den zwei folgenden Beispielen deutlicher:

Beispiel 1 (GSM-Daten):

$n = 10000 = 10^4$, $p_{bit} = 10^{-4}$

Die Größenordnung für einen Fehler ($p_{Block}(1)$) ergibt sich zu:

$$\underset{\text{Exponent von } n}{\underline{4}} \quad \underset{\text{Exponent von } p_{Bit}}{\underline{-4}} = 0 \rightarrow p_{Block}(1) \approx 10^0 = 1$$

➔ Das bedeutet, dass beim Datentransfer mit GSM der Regelfall ist, dass ein Fehler in einem Block vorliegt.

Beispiel 2 (Glasfaser):

$n = 10^5$, $p_{bit} = 10^{-12}$

Die Größenordnung für einen Fehler ($p_{Block}(1)$) ergibt sich zu:

$$\underset{\text{Exponent von } n}{\underline{5}} \quad \underset{\text{Exponent von } p_{Bit}}{\underline{-12}} = -7 \rightarrow p_{Block}(1) \approx 10^{-7},\ p_{Block}(2) \approx 10^{-14},\ p_{Block}(3) \approx 10^{-21}$$