Prof. Dr. Matthias Kriesell
Fakultät für Mathematik und Naturwissenschaften

TECHNISCHE UNIVERSITÄT
ILMENAU

5. Studienbrief zur Informations- und Codierungstheorie

In dieser Woche beginnen wir das Kapitel über Kanalcodierung. Anders als im Abschnitt über die Quellencodierung, wo es um die weitgehende Auslöschung von Redundanz in der Quelle geht („Komprimation"), wird hier Redundanz zum Chiffrat *hinzugefügt*.

Warum könnte man das tun wollen? Der Grund besteht darin, daß die Übertragungskanäle im allgemeinen nicht ganz fehlerfrei arbeiten. Der einfachste Weg, diese Fehler zu kompensieren, besteht darin, jeden Buchstaben mehrmals hintereinander zu übetragen, zum Beispiel fünfmal. Der Empfänger sieht dann die Fünferpäckchen an und pickt sich den darin häufigsten Buchstaben heraus. (Dabei wird natürlich davon ausgegangenen, daß Buchstaben bei der Übertragung niemals verlorengehen, sondern schlimmstenfalls „verrauscht" sind.) Immerhin kann man dadurch das vom Empfänger intendierte Wort auch dann noch rekonstruieren, wenn bis zu zwei Zeichen verfälscht sind, was bei einer (praktisch schon hohen) buchstabenweisen Fehlerrate von $1/1000$ recht unwahrscheinlich ist. Außerdem wird natürlich erkannt, daß bei der Übetragung ein Fehler geschehen ist, wenn bis zu vier Zeichen verrauscht sind; der Empfänger kann dann (bei entsprechendem Protokoll) um die erneute Übertragung „bitten". Entsprechend erhält man einen 2-*fehlerkorrigierenden* bzw. 4-*fehlererkennenden Code*. — Allerdings sendet man für diese Vorteile auch die fünffache Datenmenge, und man kann sich fragen, ob sich das nicht ganz wesentlich verbessern läßt.

Der erste Teil des zweiten Kapitels führt bis zum Satz von Shannon und wird Sie die kommenden zwei Wochen beschäftigen. Die ersten vier Seiten beschreiben vor allem die Konzepte. Die anderen vier (für die kommende Woche) bestehen aus dem Beweis des Satzes von Shannon und sind ein wenig dichter. Lesen Sie gegebenenfalls noch einmal in Ihren Stochastikunterlagen zur bedingten Wahrscheinlichkeiten nach.

*

Lesepensum bis zum 22. Mai 2020:

Kapitel 2 Teil 1 Seiten 1–4.

*

Übungspensum bis zum 29. Mai 2020:

1. Für zwei zwei — nicht notwendig gleichlange — Wörter $u = a_1 \dots a_n$ und $w = b_1 \dots b_m$ über dem Alphabet A sei $d(u,w) := |\{i \in \{1,\dots,\min\{n,m\}\} : a_i \neq b_i\}| + |n-m|$. Man zeige, daß hierdurch eine Metrik $d : A^* \times A^* \to \mathbb{N}$ auf A^* definiert wird. Man spricht hier auch vom *erweiterten Hamming-Abstand*.

2. Man zeige, daß die im Beweis des Satzes von Shannon zu festem d definierte modifizierte Decodierung $m'_w : \mathbb{Z}_2^n \to \{w_1,\dots,w_m,?\}$ mit $m'_w(x) = w_i$ falls $w_i \in B_d(x)$ und $w_j \notin B_d(x)$ für alle $j \in \mathbb{N}_m \setminus \{i\}$ und $m'_w(x) = ?$ sonst wohldefiniert ist und aus $m_w(w_i + x) \neq w_i$ stets $m'_w(w_i + x) \neq w_i$ folgt.

3. Man zeige: Für einen perfekt r-fehlerkorrigienden Code C ist $d(C) = 2r + 1$.

4. Sei A ein Alphabet, $|A| \geq 2$ und
$$C := \{\underbrace{xx\dots x}_{n\text{-mal}} : x \in A\}$$
der n-*Wiederholungscode* über A. Man zeige: Genau dann ist C perfekt r-fehlerkorrigierend, wenn $|A| = 2$ und $n = 2r + 1$ gilt.

5. Man betrachte den durch die Zeilen der Matrix
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$
gegebenen 4-Code C über $A = \mathbb{Z}^2$ und bestimme Minimalabstand und Informationsrate. Der Code D entstehe aus C, indem man für jedes Wort auch sein Gegenwort, das durch Tausch von 0 gegen 1 und 1 gegen 0 entsteht, hinzufügt. Welche Eigenschaften hat D?

*

Fragen und Anregungen gerne per email an mich — IC am Anfang der Betreffzeile nicht vergessen!

Ilmenau, den 18. Mai 2020 · Matthias Kriesell