Zusammenfassung Information Theory and Coding

Markus Velm

Inhaltsverzeichnis

1. Einleitung

1.1. Informationstheorie

Bit: binary unit → Einheit für Information

bit: binary digit → bit als binäres Symbol

Informationgehalt

je unwahrscheinlicher ein Symbol x auftritt, desto mehr Information enhält es:

$$I(x) = ld\left(\frac{1}{P(x)}\right) - ld(P(x))$$

P: Wahrscheinlichkeit eines Symbols
I: Informationsgehalt $[I] = Bit$

Entropie

gemittelter Informationsgehalt einer Quelle X:

$$H(X) = \sum_i P(x_i) \cdot I(x_i) = -\sum_i P(x_i) \cdot ld(P(x_i))$$

H: Entropie $[H] = Bit/Symbol$

Entscheidungsgehalt

Entropie wird maximal, wenn alle Symbole gleichwahrscheinlich sind → Entscheidungsgehalt

$$H_0 = ld(N)$$

H_0: Entscheidungsgehalt $[H_0] = Bit/Symbol$
N: Anzahl der Symbole eines Alphabets

Redundanz

$$R = H_0 - H$$

$$r = \frac{R}{H_0}$$

R: Redundanz $[R] = Bit/Symbol$
r: relative Redundanz

1.2. Quellcodierung

1.2.1. Huffman-Code

1.2.2. Arithmetische Codierung

1.3. Kanalmodell

1.4. Kanalkapazität

1.5. Shannon-Theoreme

2. Blockcodes

3. Galois-Felder

4. Reed-Solomon-Code

n: Länge Codewort
k: Länge Informationswort
d: Mindestabstand

4.1. Wunsch und Idee

Wunsch

Konstruktion eines Codes mit vorgegebener Korrekturfähigkeit
→ Vorgabe des Mindestabstandes d

$$e = \left\lfloor \frac{d-1}{2} \right\rfloor$$
$$d = 2e - 1$$

bei linearem Code ist Mindestabstand = Mindestgewicht

→ Codeworte haben mind. d von 0 verschiedene Koeffizienten

d'Alembert: Polynom vom Grad n hat n komplexe (oder höchstens n reelle) Nullstellen; auch im Galois-Feld

Idee

Konstruktion des Informationswortes als Polynom $A(x)$ mit Grad $k-1$ (damit höchstens $k-1$ Nullstellen)

Im $GF(p)$ mit Ordnung $n = p - 1$ kann man $A(x)$ an n Stellen auswerten, danach wiederholen sich die Werte

→ Auswertung des Polynoms für verschiedene x (bzw. α^i) ergeben die Koeffizienten a_i des Polynoms $a(x)$

$a_i = A(\alpha^i)$ IDFT

von diesen sind höchstens $k-1$ Null (weil $grad(A(x)) = k-1$)
von diesen sind also mind. $n-(k-1)$ von Null verschieden
→ Mindestgewicht d

$$d = n - (k-1) = n - k + 1$$

4.2. Codierung

Verschiedene Möglichkeiten aus einem Informationswort ein Codewort zu generieren

4.2.1. IDFT (nicht systematisch)

$$a_i = A(\alpha^i)$$

$A(x)$: Informationswort
a_i: Koeff. des Codewortes

4.2.2. Generatorpolynom (nicht systematisch)

$$a_i = g(x) \cdot i(x)$$

mit Generatorpolynom

$$g(x) = \prod_{i=k}^{n-1} (x - \alpha^{-i})$$

$g(x)$: Generatorpolynom
$i(x)$: Informationspolynom

4.2.3. Polynomdivision (systematisch)

Informationswort ist Teil des Codewortes (an den hohen Potenzen)

$a^*(x) = i_{k-1}x^{n-1} + i_{k-2}x^{n-2} + ... + i_1x^{n-k+1} + i_0x^{n-k}$

jedes Codewort muss durch Generatorpolynom teilbar sein
$\rightarrow$ ist für $a^*(x)$ i.A. nicht der Fall

$\frac{a^*(x)}{g(x)} = b(x) + \frac{rest(a^*(x))}{g(x)}$
$\rightarrow \frac{a^*(x) - rest(a^*(x))}{g(x)} = b(x)$
$a(x) = a^*(x) - rest(a^*(x))$

$rest(a^*(x))$: Divisionsrest

4.2.4. Über Prüfpolynom (systematisch)

Prüfpolynom:

$$h(x) = \prod_{i=0}^{k-1}(x - \alpha^{-i})$$

Produkt aus Generator- und Prüfpolynom ist 0

$g(x) \cdot h(x) = 0$

4.3. Decodierung

Idee:

Addition des Fehlerpolynoms $f(x)$ mit t Koeffizienten (d.h. t Fehler sind auf dem Kanal aufgetreten) zum gesendeten Codewort $a(x)$

im Zeitbereich:

$r(x) = a(x) + f(x)$

im Frequenzbereich:

$R(x) = A(x) + F(x)$

gedanklich wird ein Polynom $c(x)$ aufgestellt, welches t Nullen an den Fehlerstellen hat

Da die Koeffizienten von $c(x)$ die Auswertung ihrer Fouriertransformierten $C(x)$ ist, ist der Grad von $C(x)$ t

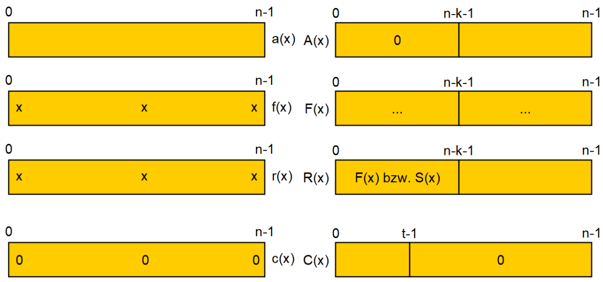
Da $c(x)$ gerade dort 0 ist, wo $f(x)$ ungleich 0, ist das Produkt $f_i \cdot c_i$ immer 0 (Achtung, keine Polynommultiplikation gemeint, sondern punktweise Multiplikation)

$f_i \cdot c_i = 0$

wenn Zeitbereich $= 0 \rightarrow$ Frequenzbereich $= 0$

$F(x) \cdot C(x) = 0$

Achtung: hier Polynommultiplikation/ Faltung/ Filterung gemeint
$\hookrightarrow$ Aufstellen der Schlüsselgleichungen



4.3.1. Schlüsselgleichungen

beschreiben, dass Faltung von $C(x)$ und $F(x)$ Null sind (Achtung: zyklische Faltung, siehe Abschnitt 3)

F_0 bis F_{n-k-1} (bzw. F_{d-2}) sind bekannt, da diese direkt an den Syndromstellen von $R(x)$ stehen

Alle C-Koeff. sind unbekannt, außer C_{t-1}, dieser wird zu 1 gesetzt

$C_{t-1} = 1$

da Anzahl der Fehler (t) unbekannt sind, muss ausprobiert werden, welche minimale Anzahl an Fehlern die Schlüsselgleichungen erfüllt

Berlekamp-Massey-Algorithmus

effizientes Verfahren zur Lösung der Schlüsselgleichungen

Euklidscher Algorithmus

Suche des ggT zweier Zahlen

Kann zur Lösung der Schlüsselgleichungen verwendet werden

4.4. Horner-Schema

5. Erweiterungskörper

Erweitern des Grundkörpers (z.B. 2) mit Exponent (z.B. 4)
$\rightarrow GF(2^4)$

primitives Element wird zu primitivem Polynom, z.B. $p(x) = x^4 + x + 1$

6. BCH-Codes

A. Polynome

A.1. Polynommultiplikation

A.2. Polynomdivision