Codierung und Verschlüsselung

Prof. P. Hauck[1]

WS 2009

[1]Mitschrift: Christian Stegmann und Sebastian Brandt

Inhaltsverzeichnis

Kapitel 1

Kryptologie

1.1 Grundbegriffe und einfache Verfahren

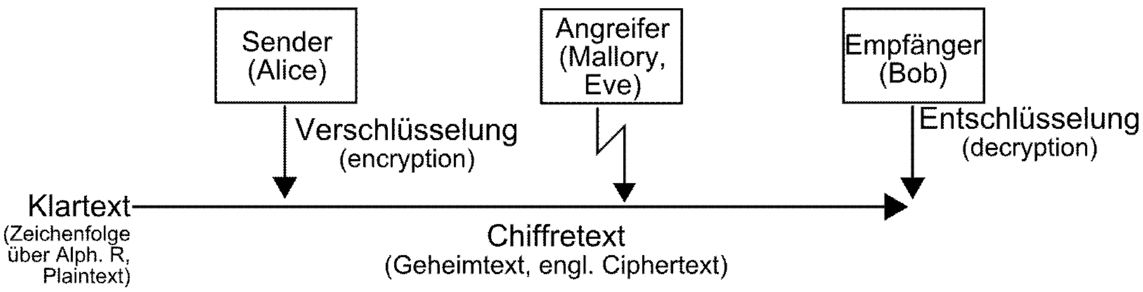

Abbildung 1.1: Schaubild der Kryptologie

1.1.1 Verschlüsselung erfordert

- Verschlüsselungsverfahren, Algorithmus (Funktion)
- Schlüssel k_e (encryption key)

$$E(m, k_e) = c$$

E=Verschlüsselungs Funktion, m=Klartext, c=Chiffretext

$$E(m_1, k_e) \neq E(m, k_e) \text{ für } m_1 \neq m_2$$

$$D(c, k_d) = m$$

(k_d zu k_e gehöriger Dechiffrierschlüssel!)
$k_d = k_e$ (oder k_d leicht aus k_e zu berechnen):
symmetrisches Verschlüsselungsverfahren, ansonsten **asymmetrische Verschlüsselungsverfahren**. Ist k_d nur sehr schwer (oder gar nicht) zu k_e berechenbar, so kann k_e veröffentlicht werden:
Public-Key-Verfahren.

1.1.2 Beispiel für (nicht sicheres) symmetrische Verfahren

a) $R = S = \{0, 1, \ldots, 25\}$
Verfahren: Verschiebechiffre
Schlüssel: $i \in \{0, 1, \ldots, 25\}$
Verfahren $x \in \mathbb{R} \longrightarrow x + i \mod 26 = y$
$y \longmapsto y - i \mod 26 = y$
$m = x_1...x_2 \longrightarrow c = (x_1 + i \mod 26) \ldots (x_n + i \mod 26)$, $E(m, i)$
Unsicher, weil Schlüsselmenge klein ist (Brute Force Angriff).

b) R,S, Schlüsselmenge=Menge aller Permutationen von $\{1, \ldots, 25\} = S_{26}$
Verschl.: Wähle Permuation π
$x \in \mathbb{R} \longrightarrow \pi(x) = y$
Entschl.: $y \longrightarrow \pi^{-1}(y) = x$
$m = x_1 \ldots x_r \rightarrow c = \pi(x_1) \ldots \pi(x_r)$
$\begin{pmatrix} 0 & 1 & 2 & \ldots & 25 \\ 3 & 17 & 4 & \ldots & 13 \end{pmatrix} \longrightarrow \pi(0) = 3$, u.s.w.
Anzahl der Permutationen: $|S_{26}| = 26! \approx 4 \cdot 10^{26} \longrightarrow$ Brute-Force Angriff nicht mehr möglich!
Warum? Man muss im Schnitt 50% der Permutationen testen. Angenommen man könnte 10^12 Perm. pro Sekunde testen.
Aufwand: $2 \cdot 10^{14}$ Sekunden $\approx$ 6.000.000 Jahre
Trotzdem unsicher!
Grund: Charakteristisches Häufigkeitsverteilung von Buchstaben in natürlichsprachigen Texten.

Verfahren beinhalten viele Verschlüsselungsmöglichkeiten, abhängig von der Auswahl des Schlüssels.
Verfahren bekannt, aber Schlüssel k_d geheim!

1.1.3 Prinzip von Kerkhoffs (1835-1903)

Sicherheit eines Verschlüsselungsverfahren darf nicht von der Geheimhaltung des Verfahrens, sondern nur von der Geheimhaltung des verwendeten Schlüssels abhängen!

Kryptologie besteht aus Kryptographie (Entwurf) und der Kryptoanalyse (Angriff).
Angriffserfolge:

- Schlüssel k_d wird gefunden
- Eine zu der Dechiffrierfunktion $D(\cdot, k_d)$ äquivalente Funktion finden ohne Kenntnis von k_d
- gewisse Chiffretexte werden entschlüsselt

1.1.4 Arten von Angriffen

- Ciphertext-Only Angriff
- Known-Plaintext Angriff
- Chosen-Plaintext Angriff
- Chosen-Ciphertext Angriff

1.2 One-Time-Pad und perfekte Sicherheit

Lauftextverschlüsselung

Alphabet $\mathbb{Z}_k = \{0, 1, \ldots, k-1\}$
In $\mathbb{Z}_k$ kann man addieren und multiplizieren mit mod k.

Klartext $x_1, x_2, \ldots, x_n$
Schlüsselwort $k_1, k_2, \ldots, k_n$
$x_1 + k_1 \mod k$, $x_n + k_n \mod k \leftarrow$ Chiffretext

Mit natürlichsprachlichen Texten ist das Verfahren unsicher.
$\mathbb{Z}_2 = \{0, 1\}$, $1 \oplus 1 = 0 = 0 \oplus 0$, $0 \oplus 1 = 1 = 1 \oplus 0 \Rightarrow XOR$
Klartext in $\mathbb{Z}_2^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{Z}_2\}$ Schlüssel: Zufallsfolge über $\mathbb{Z}_2$ der Länge n. m Klartext, k Zufallsfolge (beide Länge n)

$$c = m \oplus k, (x_1, \ldots, x_n) \oplus (k_1, \ldots, k_n) := (x_1 \oplus k_1, \ldots, x_n \oplus k_n)$$

1.2.1 One-Time-Pad

Schlüssel k darf nur einmal verwendet werden!

$$m_1 \oplus k = c_1, m_2 \oplus k = c_2, c_1 \oplus c_2 = m_1 \oplus k \oplus m_2 \oplus k = m_1 \oplus m_2$$

Wieder nur Lauftext $\rightarrow$ unsicher!
m_1 und m_2 lässt sich ermitteln.
Zufallsfolge der Länge n: eigentlich unsinniger Begriff. Da jedes Bit unabhängig von anderen mit Wahrscheinlichkeit $\frac{1}{2}$ erzeugt wird (Output einer binär symmetrischen Quelle)
Jede Folge der Länge n ist gleich wahrscheinlich (Wahrscheinlichkeit $\frac{1}{2}n$
One-Time-Pad ist perfekt sicher.

1.2.2 Perfekte Sicherheit

Ein Verschlüsselungsverfahren ist perfekt sicher, falls gilt: Für jeden Klartext m und jedem Chiffretext c (der festen Länge n)

$$pr(m|c) = pr(m)$$

$pr(m|c) \rightarrow$ A-posteriori-Wahrscheinlichkeit (Wahrscheinlichkeit, dass m Klartext, wenn c empfangen wurde)
$pr(m) \rightarrow$ A-priori-Wahrscheinlichkeit

Beispiel: Substitutionschiffre aus Kapitel 2.
$n = 5, m = HALLO, pr(m) > 0$
Ang:$c = QITUA$ wird empfangen, $LL \neq TU \rightarrow pr(m|c) = 0$
nicht perfekt sicher.

One-Time-Pad ist perfekt sicher.
(Bayes'sche Formel) $m \oplus k$
Jede Folge c lässt sich mit geeignetem k in der Form $c = m \oplus k$ erhalten.
Wähle $k = m \oplus c$, $m \oplus k = m \oplus m \oplus c = c$
Bei gegebenem m und zufällige gewählten Schlüssel k ist jeder Chiffretext gleichwertig.

1.3 Symmetrische Blockchiffre

1.3.1 Blockchiffre

Zerlege Klartext in Blöcke (Strings) der Länge n. Jeder Block wird einzeln verschlüsselt (in der Regel wieder in einem Block der Länge n). Gleiche Blöcke werden gleich verschlüsselt.

Wie viele Blockchiffren der Länge n gibt es?
Alphabet $\mathbb{Z}_2 = \{0, 1\}$
$|\{\underbrace{(0, \ldots, 0)}_{Block}, (0, \ldots, 1), \ldots, (1, \ldots, 1)\}| = 2^n$
Blockchiffre = Permutation der 2^n Blöcke.
$(2^n)!$ Blockchiffre

Wenn alle verwendet werden:
Schlüssel = Permutation der 2^n Blöcke
$(x_{1,1}, \ldots, x_{1,n}, x_{2,1}, \ldots, x_{2,n}, \ldots)$ $\boxed{n \cdot 2^n \text{ Bit}}$
Zur Speicherung eines Schlüssels werden $n \cdot 2^n$ Bit benötigt.
Zum Beispiel:
$n = 64,\ 64 \cdot 2^{64} = 2^{70} = 1$ ZiB ≈ 1 Milliarde Festplatten à 1 TB
Illusional!

Konsequenz: Verwende Verfahren, wo nur ein kleiner Teil der Permutation als Schlüssel verwendet wird und so sich die Schlüssel dann in kürzerer Form darstellt.

1.3.2 Vorbemerkung

$n \times m$-Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

$1 \times n$ = Zeilenvektor = $(a_1, \dots, a_m)$

$n \times 1$ = Spaltenvektor = $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

z.B. $a_{ij} \in \mathbb{R},\ a_{ij} \in \mathbb{Z}$ oder $a_{ij} \in R,\ R$ Ring

$n \times m$-Matrix A,B

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1m} + b_{1m} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nm} + b_{nm} \end{pmatrix}$$

$$A = n \times m,\ B = m \times k,$$

$$A \cdot B \begin{pmatrix} c_{1l} & \dots & c_{1k} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mk} \end{pmatrix} = n \times k$$

$$c_{1l} = (a_{i1} \cdot b_{ij}) + (a_{i2} \cdot b_{2j}) + \dots + (a_{im} \cdot b_{mj})$$

$$(A + B) \cdot C = A \cdot B + B \cdot C$$

Im Allgemeinem: $A \cdot B \neq B \cdot A$

Quadritsche Matrix ($n \times n$)

$$E_n = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

$$A = n \times n,\ A \cdot E_n = E_n \cdot A = A$$

A $n \times n$-Matrix über kommutativen Ring R mit Eins.
Wann existiert Matrix A^{-1}(Inverse Matrix) mit $A^{-1} \cdot A = A \cdot A^{-1} = E_n$?
$det(A) \in R$ Determinante von A

2×2-Matrix: $det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{21} \end{pmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$

A besitzt inverse Matrix $\Leftrightarrow$ $det(A)$ in R ein inverses besitzt (z.B. R Körper, $\mathbb{Z}, \mathbb{Q}, \mathbb{Z}_p,\ det(A) \neq 0$

$$A^{-1} = \begin{pmatrix} \frac{1}{det(A)} \cdot b_{11} & \dots & \frac{1}{det(A)} \cdot b_{1m} \\ \vdots & & \vdots \\ \frac{1}{det(A)} \cdot b_{n1} & \dots & \frac{1}{det(A)} \cdot b_{nm} \end{pmatrix}$$

$b_{ij} = (-1)^{i+j}\ det(A_{ji})$

$A_{ji} = (n-1) \times (n-1)$-Matrix, die aus A durchstreichen der j-ten Zeile und i-ten Spalte entsteht.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\ A^{-1} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$R = \mathbb{Z}_k\ \{0, 1, \dots, k\}$
Addition und Multiplikation in $\mathbb{Z}_k(\oplus, \odot)$
normale Add. und Mult. mit mod k

1.3.3 Affin-lineare Chiffren

Klartextalphabet = Chiffretextalphabet = $\mathbb{Z}_k$ $(k = 2,\ k = 26)$
Wähle $n \times n$-Matrix A über $\mathbb{Z}_k$ und Zeilenvektor b der Länge n über $\mathbb{Z}_k$. Dies wird der Schlüssel sein für die Chiffrierung.
Blockchiffre der Länge n.
Block = Zeilenvektor der Länge n über $\mathbb{Z}_k$.
Klartextblock v
Chiffretextblock

$$\begin{aligned} v \cdot A + b &=: w \\ v \to v \cdot A + b &=: w \\ w - b &= v \cdot A \end{aligned}$$

benötigen: A^{-1} existiert (d.h. $ggT(det(A), k) = 1$)
Dechiffrierung:

$$(w - b) \cdot A^{-1} = v \cdot A \cdot A^{-1} = v \cdot E_n = v$$

(wenn immer b=0 gewählt wird, dann lineare Chiffren, Hill-Chiffren)
Beispiel:

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \quad \mathbb{Z}_6$$

Blockchiffre der Länge n $det(A) = 1 \cdot 2 - 3 \cdot 3 = -7 = 5$ inverse in $\mathbb{Z}_6$

$$\frac{1}{det(A)} = det(A)^{-1} = 5$$

$$A^{-1} = 5 \cdot \begin{pmatrix} 2 & -3 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 10 & -15 \\ -15 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 5 \end{pmatrix}$$

Test:

$$A \cdot A^{-1} = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 4+9 & 3+15 \\ 12+6 & 9+10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Verschlüsselung: Schlüssel:

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \quad b = (3, 5)$$

Klartextblock: $(1, 2)$
Chiffretextblock:

$$w = (1, 2) \cdot \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} + (3, 5) = (1, 1) + (3, 5) = (4, 0)$$

Entschlüsselung:

$$(w - b) \cdot A^{-1} = (1, 1) \cdot \begin{pmatrix} 4 & 3 \\ 3 & 5 \end{pmatrix} = (1, 2)$$

$\mathbb{Z}_2$: $n^2 + n$ Bit zur Speicherung eines Schlüssels.
Wie viele inverse Matrizen über $\mathbb{Z}_2$ mit $n = 64$?

$$(2^{64} - 1) \cdot (2^{64} - 2) \cdot \ldots \cdot (2^{64} - 2^{63}) \approx 0.29 \cdot 2^{4096}$$

Verfahren ist unsicher gegenüber Known-Plaintext-Angriffe.
(A, b) Schlüssel, A inverse $n \times n$-Matrix über $\mathbb{Z}_k, b \in \mathbb{Z}_k^n$
Angenommen Angreifer kennt $n + 1$ Klartext/Chiffretextpaare verschlüsselt mit (A, b), $v_0, v_1, \ldots, v_n$ $w_0, \ldots, w_n$
Dann kann er häufig (A, b) bestimmen.

$$V = \begin{pmatrix} v_1 - v_0 \\ v_2 - v_0 \\ \vdots \\ v_n - v_0 \end{pmatrix} n \times n - \text{Matrix}$$

Angenommen: V ist invertierbar. Setze $W = \begin{pmatrix} w_1 - w_0 \\ \vdots \\ w_n - w_0 \end{pmatrix}$

$$V \cdot A = \begin{pmatrix} (v_1 - v_0) \cdot A \\ \vdots \\ (v_n - v_0) \cdot A \end{pmatrix} = \begin{pmatrix} v_1 \cdot A + b - v_0 \cdot A + b \\ \vdots \\ v_n \cdot A + b - v_0 \cdot A + b \end{pmatrix} = \begin{pmatrix} w_1 - w_0 \\ \vdots \\ w_n - w_0 \end{pmatrix} = W$$

$V \cdot A$ bekannt, also auch V^{-1}:
$A = V^{-1} \cdot w$
$b = w_0 - v_0 \cdot A$
Beispiel: $n = 2,\ k = 25\ \{A, \ldots, Z\} = \{0, \ldots, 25\}$

$$\texttt{HERBST} \longrightarrow \texttt{NEBLIG}$$

$$\begin{array}{c|c|c c c|c|c}
H & 7 & \multirow{2}{*}{v_0} & & N & 13 & \multirow{2}{*}{w_0} \\
E & 4 & & & E & 4 & \\
\hline
R & 17 & \multirow{2}{*}{v_1} & \longrightarrow & B & 1 & \multirow{2}{*}{w_1} \\
B & 1 & & & L & 11 & \\
\hline
S & 18 & \multirow{2}{*}{v_2} & & I & 8 & \multirow{2}{*}{w_2} \\
T & 19 & & & G & 6 &
\end{array}$$

$$V = \begin{pmatrix} 10 & -3 \\ 11 & 15 \end{pmatrix} = \begin{pmatrix} 10 & 23 \\ 11 & 15 \end{pmatrix}, \quad W = \begin{pmatrix} 14 & 7 \\ 21 & 2 \end{pmatrix}$$

$$det(V) = 10 \cdot 15 + 33 = 183 \equiv 1 (\text{mod } 26)$$

$$V^{-1} = \begin{pmatrix} 15 & 3 \\ -11 & 10 \end{pmatrix} = \begin{pmatrix} 15 & 3 \\ 15 & 10 \end{pmatrix}$$

$$A = V^{-1} \cdot W = \begin{pmatrix} 15 & 3 \\ 15 & 10 \end{pmatrix} \cdot \begin{pmatrix} 14 & 7 \\ 21 & 2 \end{pmatrix} = \begin{pmatrix} 210 + 63 & 105 + 6 \\ 210 + 210 & 105 + 20 \end{pmatrix} = \begin{pmatrix} 13 & 7 \\ 4 & 21 \end{pmatrix}$$

$$b = w_0 - v_0 \cdot A = (13, 4) - (7, 4) \cdot \begin{pmatrix} 13 & 7 \\ 4 & 21 \end{pmatrix} = (10, 1)$$

Test:

$$v_1 \cdot A + b = w_1, \quad v_2 \cdot A + b = w_2$$

1.4 Der Advanced Encryption Standard (AES)

1.4.1 Mathematische Methoden gebraucht für AES

Seit 70er Jahren gab es DES (Blocklänge 64 Bit, Schlüssellänge 56 Bit)

Nachfolger des DES: Daemen, Rijmen (Belgier)
Rijndael-Verfahren → AES (2002 FIPS 197)

Iterierte Blockchiffre
Version mit 128 Bit Block und Schlüsselänge.
Vorbemerkung: 128-Bit Blöcke werden dargestellt als:

$$\begin{pmatrix} a_{00} & a_{01} & \ldots & a_{03} \\ a_{10} & a_{11} & \ldots & a_{13} \\ \vdots & \vdots & \vdots & \vdots \\ a_{30} & \ldots & \ldots & a_{33} \end{pmatrix}$$

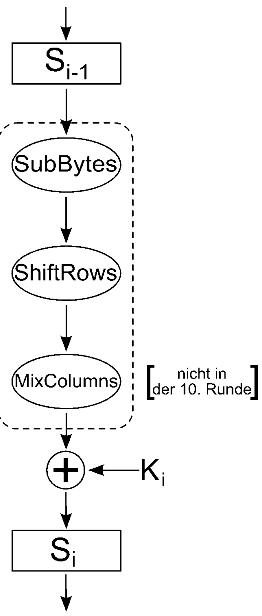

Abbildung 1.2: Eine Runde vom AES

Jedes a_{ij} = Byte
128er Block $\hat{=}$ $a_{00}a_{10}a_{20}\dots a_{01}a_{11}\dots a_{33}$ (spaltenweise gelesen)

endlicher Körper: einfachste Möglichkeit $\mathbb{Z}_p$ (p Primzahl)
$\mathbb{F}_{2^8}$ Körper mit $2^8 = 256$ Elementen

Menge: Polynome vom Grad < 8 über $\mathbb{Z}_2$
$b_7x^7 + \dots + b_1x + b_0, b_i \in \mathbb{Z}_2$
$(b_7, b_6, \dots, b_0)$ Byte

Addition = normale Addition von Polynomen
Multiplikation = normale Multiplikation von Polynomen + Reduktion modulo irreduzibler Polynom vom Grad 8. $(x^8 + x^4 + x^3 + x + 1)$

Bsp.
$(x^7 + x + 1) \odot (x^3 + x) = x^{10} + x^8 + x^4 + x^3 + x^2 + x$

$x^{10} + x^8 + x^4 + x^3 + x^2 + x \mod x^8 + x^4 + x^3 + x + 1$

$$
\begin{array}{l}
x^{10} + x^8 + x^4 + x^3 + x^2 + x \div x^8 + x^4 + x^3 + x + 1 = x^2 + 1 \\
\underline{x^{10} + x^6 + x^5 + x^3 + x^2} \\
\qquad x^8 + x^6 + x^5 + x^4 + x \\
\qquad \underline{x^8 + x^4 + x^3 + x + 1}
\end{array}
$$

$x^6 + x^5 + x^3 + 1 \leftarrow$

$(x^7 + x + 1) \odot (x^3 + x) = x^6 + x^5 + x^3 + 1$

In $\mathbb{F}_{2^8}$ hat jedes Element $\neq 0$ ein Inverses bzgl. $\odot$:
$g \neq 0. Ex. g^{-1} \in \mathbb{F}_{2^8} : g \odot g^{-1} = 1$

Erweiterte Euklid. Algo. für Polynome:
$g \neq 0$ (Grad ≤ 7) $\quad h = x^8 + x^4 + x^3 + x + 1$ irred. $\quad \mathrm{ggT}(g, h) = 1$

EEA: $u, v \in \mathbb{Z}_2[x] : u \cdot g + v \cdot h = 1$
$u \mod h =: g^{-1}$
$g^{-1} \odot g = ((u \mod h) \cdot g) \mod h = u \cdot g \mod h = (1 - vh) \mod h = 1 \mod h = 1$

1.4.2 SubBytes-Transfer

$$S_{i-1} = \begin{pmatrix} a_{00} & a_{01} & \dots & a_{03} \\ a_{10} & a_{11} & \dots & a_{13} \\ \vdots & \vdots & \vdots & \vdots \\ a_{30} & \dots & \dots & a_{33} \end{pmatrix}, a_{ij} \text{ Bytes}$$

Sei g eines dieser Bytes, $g = (b_7 b_6 \dots b_0), b_i \in \mathbb{Z}_2$

1. Schritt: Fasse g als Element in $\mathbb{F}_{2^8}$ auf.
   Ist $g = (0, \dots, 0)$, so lasse g unverändert.
   Ist $g \neq (0, \dots, 0)$, so ersetzte g durch g^{-1}.

2. Schritt: Ergebnis nach Schritt 1: $\tilde{g}$ wird folgenderm. Transformiert
   $\tilde{g} \cdot A + b = \tilde{\tilde{g}}$ (affin-lin. Transformation) ($\tilde{g}$: g-schlange, $\tilde{\tilde{g}}$:g-doppel-schlange)

A wird durch zyklischer Shift der vorherigen Zeile um 1 Stelle nach rechts erzeugt.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Schritt 1 und 2 werden kombiniert, nicht jedes mal berechnet. Alle möglichen Sub-Bytes (2^8 viele) sind in einer 16x16 Matrix und wird per Table-Lookup nachgeschlagen.
$g = (b_7 b_6 b_5 b_4 b_3 b_2 b_1 b_0)$ $\;b_7 b_6 b_5 b_4 = 0$ bis 15 (Zeile) $\;b_3 b_2 b_1 b_0$ (Spalte)

1.4.3 Shift Rows Transformation

4x4-Matrix von Bytes:

$$\begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} \begin{array}{l} \leftarrow \text{Erste Zeile unverändert} \\ \leftarrow \text{1 Stelle nach links zykl.} \\ \leftarrow \text{2 Stellen nach links zykl.} \\ \leftarrow \text{3 Stellen nach links zykl.} \end{array}$$

1.4.4 Mix Columns Transformation

4x4-Matrix, Einträge als Elemente in $\mathbb{F}_{2^8}$ auffassen.
Multiplikation von links mit Matrix (Mult. der Eintr. in $\mathbb{F}_{2^8}$):

$$\begin{pmatrix} x & x+1 & 1 & 1 \\ 1 & x & x+1 & 1 \\ 1 & 1 & x & x+1 \\ x+1 & 1 & 1 & x \end{pmatrix}$$

$$x \mathrel{\hat{=}} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

1.4.5 Schlüsselerzeugung

Ausgangsschlüssel hat 128 Bit. (16er String in Hexcode)

Schreibe als 4x4-Matrix von Bytes. 4 Spalten $w(0), w(1), w(2), w(4)$. Definiere weitere 40 Spalten à 4 Bytes.

$w(i-1)$ sei schon definiert.

$$4 \nmid i : w(i) := w(i-4) \oplus w(i-1) \quad \text{(byteweise } XOR\text{)}$$
$$4 \mid i : w(i) := w(i-4) \oplus T(w(i-1)) \quad (T \text{ Transformation})$$

T?

$$w(i-1) = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \quad a, \ldots, d\text{Bytes}$$

Wende auf b, c, d, a SubBytes-Transformation an $\rightarrow e, f, g, h$

$$r(i) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}^{\frac{(i-4)}{4}} \quad \text{Potenz. in } \mathbb{F}_{2^8}$$

$$T(w(i-1)) = \begin{pmatrix} e \oplus r(i) \\ f \\ g \\ h \end{pmatrix}$$

Rundenschlüssel K_i: 4x4-Matrix mit Spalten $w(4i), ; w(4i+1), ; w(4i+2), ; w(4i+3)$

(Nebenbemerkung: Linear heißt $f(x + y) = f(x) + f(y)$)

1.4.6 Public-Key-Systeme

1.4.7 Grundidee

Diffie, Hellman, 1976 Jeder Teilenehmer hat ein Paar von Schlüsseln:

- öffentlichen Schlüssel P_A
- geheimen Schlüssel G_A

Zu P_A gehört öffentlich bekannte Verschlüsselungsfunktion E_{P_A} (=$E(\cdot, P_A)$)
$B \xrightarrow{m} A : E_{P_A}(m) = c$

1. m darf mit „realistischen Aufwand“ nicht aus $E_{P_A}(m)$ berechenbar sein. E_{P_A} ist **Einwegfunktion**
   (E_{P_A} muss effizient berechenbar sein, aber $E_{P_A}^{-1}$ nicht!)

2. A muss mit Hilfe einer Zusatzinformation (=G_A) in der Lage sein, $E_{P_A}^{-1}$ effizient zu berechnen.
   $$D_{G_A}(c) = m = E_{P_A}^{-1}(c)$$
   Injektive Einwegfunktionen, die mit Zusatzinformation effizient invertierbar sind: **Geheimtürfunktion** (trapdoor function)

Aus 1) und 2) folgt:

3 G_A darf aus P_A nicht schnell berechenbar sein!

Es ist unbekannt ob Einwegfunktion existieren! Notwendig für die Existenz von Einwegfunktionen:
$$P \neq NP$$
Es gibt Kandidaten für Einwegfunktionen.

1.5 RSA-Verfahren

(Rivest, Shamir, Adleman, 1977)
Beruht auf Schwierigkeit große Zahlen zu faktorisieren!

1.5.1 Schlüsseslerzeugung

Wähle zwei große Primzahlen $p, q (p \neq q)$ (mindestens 500 Bit Länge)
Bilde $n = p \cdot q$

$$\varphi(n) = ||\{a \in \mathbb{N} : 1 \leq a < n, ggt(a,n) = 1\}||$$

$$n = p \cdot q : \varphi(n) = (p-1) \cdot (q-1)$$

[nicht teilerfremd zu n:$1 \cdot p, 2 \cdot p, \ldots, (q-1) \cdot p = n = 1 \cdot q, 2 \cdot q, \ldots, (p-1) \cdot q$
$(p-1) + (q-1) + 1$
$\varphi(n) = n-(p-1)-(q-1)-1 = n-p-q+1 = p \cdot q - p - q + 1 = (p-1) \cdot (q-1)$]
Wähle e, $1 < e < \varphi(n)$ mit $ggT(e, \varphi(n)) = 1$
Zufallswahl, bestimme $ggT(e, \varphi(n))$ mit Euklidischer Algorithmus, so lange, bis e mit $ggT(e, \varphi(n)) = 1$ gefunden ist.)

öffentlicher Schlüssel

$(n, e) = P_A$
Wähle $d < \varphi(n)$ mit $e \cdot d \equiv 1 (\mathrm{mod}\ n)$ (d.h. $\varphi(n) \mid e \cdot d - 1, e \cdot d = 1 + k \cdot \varphi(n)$ für $k \in \mathbb{N}$)
(Wende erweiterten Euklidischen Algorithmus auf e, $\varphi(n)$ an:
Liefert $u, v \in \mathbb{Z}$ mit $u \cdot e + v \cdot \varphi(n) = ggT(e, \varphi(n)) = 1$
$d = u \mod \varphi(n)$
$u \cdot e + v \cdot \varphi(n) \mod \varphi(n) = 1$
$(\underbrace{u \mod \varphi(n)}_{d} \cdot e) \mod \varphi(n) = 1)$

Geheimerschlüssel

$G_A = d$

1.5.2 Verschlüsselung

B Nachrichtan A. Codiere Nachricht als Zahl. Zerlege in Blöcke deren Zahlwert $< n$. Sei m so ein Block. $(m < n)$
$m^e \mod n = c$

1.5.3 Entschlüsselung

$c^d \mod n = m$
Gültigkeit basiert auf kleinem Satz von Fermat:
r Primzahl, $ggT(a, r) = 1$ (d.h. $r \nmid a$)
$a^{r-1} \equiv 1 (\mathrm{mod}\ r)$

Sei $m < n = p \cdot q$

$$\begin{aligned} c = m^e \mod n,\ c^d \mod n &= m^{e\cdot d} \mod n \\ e \cdot d &= 1 + k \cdot \varphi(n) \\ &= 1 + k \cdot (p-1) \cdot (q-1) \end{aligned}$$

Ist $p \nmid m$, so

$$\begin{aligned} m^{e\cdot d} &= m^{1+k\cdot(p-1)\cdot(q-1)} \\ &= m \cdot \underbrace{(m^{p-1})}_{\equiv 1 \mod p}{}^{k\cdot(q-1)} \\ &\overset{\mod p}{\Rightarrow} m \cdot 1^{k\cdot(q-1)} (\mathrm{mod}\ p) \\ &\equiv m (\mathrm{mod}\ p) \end{aligned}$$

Ist $p \mid m$:

$$m \equiv 0 \equiv m^{e\cdot d} (\mathrm{mod}\ p)$$

In jedem Fall:

$$m^{e\cdot d} \equiv m (\mathrm{mod}\ p)$$

Genauso:

$$\begin{aligned} m^{e\cdot d} &\equiv m (\mathrm{mod}\ q) \\ p \mid m^{e\cdot d} - m,\ q \mid m^{e\cdot d} - m, p \neq q \Rightarrow n = p \cdot q \mid m^{e\cdot d} - m \\ m^{e\cdot d} &\equiv m (\mathrm{mod}\ n),\ m^{e\cdot d} \mod n = m \end{aligned}$$

Schnelle Berechnung von modularen Potenzen ($m^e \mod n$)

$$\begin{aligned} e &= \sum_{i=0}^{k} e_i \cdot 2^k, \quad e_i \in \{0, 1\}, \quad e_k = 1 \\ m^e &= m^{2\cdot k + e_{k-1}\cdot 2^{k-1} + \ldots + e_1 \cdot 2 + e_0} \end{aligned}$$

$$((\ldots((m^2 \cdot m^{e_{k-1}})^2 \cdot m^{e_{k-2}})^2 \ldots)^2 \cdot m^{e_1})^2 \cdot m^{e_0}$$

gelöst im worst case mit $2 \cdot k$ Multiplikationen.

$$k = \lfloor log_2(e) \rfloor$$

$$\boxed{\text{Nach jedem Rechenschritt mod } n \text{ reduzieren!}}$$

1.5.4 Sicherheit vom RSA-Verfahren

Falls $p,\ q$ bekannt $\Rightarrow \varphi(n),\ d$ bekannt.
$\varphi(n)$ bekannt $\Rightarrow\ p,\ q$ bekannt.
$\varphi(n) = n - q - p + 1$ bekannt $\Rightarrow\ p + q = s$ bekannt, $p \cdot q = n$ bekannt.
$p \cdot (s - p) = n\ p^2 - s \cdot p + n = 0$ quadratische Gleichung für p

Es gilt auch: Bestimmung von d ist „genauso schwierig" wie die Faktorisierung von n.
Komplexität der besten Faktorisierungsalgorithmen:

$$O(e^{c \cdot (\log n)^{\frac{1}{3}} \cdot ((\log \log n)^{\frac{2}{3}})})$$

Um eine 640 Bit Zahl zu faktorisieren braucht man 30-CPU-Jahre auf einer 2.2 GHz CPU.

Häufig wird $e = 3$ gewählt.

$$A \xrightarrow{m} (n_1, 3)\ B_1$$
$$A \xrightarrow{m} (n_2, 3)\ B_2$$
$$A \xrightarrow{m} (n_3, 3)\ B_2$$

$ggT(n_i, n_j) = 1$

$$\begin{aligned} c_1 &= m^3 \mod n_1 \\ c_2 &= m^3 \mod n_2 \\ c_3 &= m^3 \mod n_3 \end{aligned}$$

Eve fängt $c_1,\ c_2,\ c_3$ ab: Chinesisches Restsatz:

$$\begin{aligned} 0 \leq x \leq n_1 \cdot n_2 \cdot n_3 \\ \text{mit } x = c_i \mod n_i, \quad i = 1, 2, 3 \\ x \text{ ist eindeutig bestimmbar} \\ m^3 \equiv c_i \mod n_i,\ m^3 < n_1 \cdot n_2 \cdot n_3 \\ \Rightarrow x = m^3 \Rightarrow m = \sqrt[3]{x} \end{aligned}$$

Wenn $e = 5$, dann braucht man 5 Nachrichten.

1.5.5 Wie bestimmt man große Primzahlen?

$$p \text{ Primzahl}, \quad a \in \mathbb{Z}, \quad ggT(a, p) = 1$$
$$a^{p-1} \equiv 1 (\text{mod } p) \text{ [kl. Satz von Fermat]}$$

gegeben: n, $\quad ggT(a, n) = 1 \quad a^{n-1} \equiv 1 \text{ (mod } n)$?

1.5.6 Fermat-Test

Wenn nicht, so ist n keine Primzahl. Wenn ja, so keine Aussage möglich. Wähle neues a!
Es gibt zusammengesetze Zahlen n (Carmichael-Zahlen) mit:

$$a^{n-1} \equiv 1 \pmod{n} \quad \forall\, a \text{ mit } ggT(a,n) = 1$$

1.5.7 Miller-Rabin-Test

$$\begin{aligned} ggT(a,p) &= 1 \\ p \text{ Primzahl } p-1 &= 2^s \cdot t, \quad 2 \nmid t \\ a^{2^s \cdot t} &\equiv 1 \pmod{p} \\ (a^{2^{s-1} \cdot t})^2 &= b \\ a^{2^{s-1} \cdot t} &= \begin{cases} 1 & \mod p \\ -1 & \mod p \end{cases} \\ b^2 &\equiv 1 \pmod{p} \\ (b \mod p)^2 &= 1 \in \mathbb{Z}_p \\ x^2 - 1 &\in \mathbb{Z}_p[x] \end{aligned}$$

$\boxed{\text{Entweder } a^t \equiv 1 \pmod{p} \text{ oder } a^{s^i \cdot t} \equiv -1 \pmod{p} \text{ für ein } 0 \le i \le s}$
Teste dies mit n statt p.
Wenn n keine Primzahl ist, dann gibt es mindestens $\frac{3}{4}\varphi(n)$ viele a, so dass der Test fehlschlägt.
→ probabilistischer Primzahltest
p Primzahl $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ Gruppe bezüglich Multiplikation (zyklisch)

$$\begin{aligned} \exists g \in \mathbb{Z}_p^* : \{g^0, g^1, g^2, \dots, g^{p-2}\} = \mathbb{Z}_p \\ g^{p-1} \equiv 1 (\mathrm{mod}\ p) \end{aligned}$$

Primitivwurzel $\mod p$
$0 \le a \le p-2 : a \mapsto g^a \mod p$ Kandidat für Einwegfunktion.
$g^a \mod p \to a$ (diskreter Logarithmus) ist nach heutigem Stand schwer!

1.5.8 Diffie-Hellman-Verfahren zur Schlüsselvereinbarung

A, B wollen gemeinsamen Schlüssel K für ein symmetrische Verfahren vereinbaren; es steht nur unsichere Kommuniaktionskanal zur Verfügung.

Lösung: p, g (Bitlänge von p > Bitlänge von K) (können öffentlich bekannt sein).

1. A wählt zufällig $a \in \{2, \ldots, p-2\}$
   A berechnet $x = g^a \mod p$
   (a geheim halten)
2. B wählt zufällig $b \in \{2, \ldots, p-2\}$
   B berechnet $y = g^b \mod p$
   (b geheim halten)
3. $A \xrightarrow{x=g^a} B$
   $B \xrightarrow{y=g^b} A$
   $A : y^a \mod p = g^{b \cdot a} \mod p = K$
   $B : x^b \mod p = g^{a \cdot b} \mod p = K$

1.5.9 Sicherheit

Angreifer: $p, g,\ g^a \mod p,\ g^b \mod p$
gesucht: $g^{a \cdot b} \mod p$
Einzig bekannte Möglichkeit ist das Berechnen a aus $g^a : (g^b)^a \mod p = K$ müsste diskretes Logarithmus-Problem lösen.

1.5.10 Man-in-the-Middle

M fängt g^a und g^b ab und wählt $c \in \{2, \ldots, p\}$ und schickt $g^c \mod p$ an A und B.
$A :\ g^{c \cdot a} \mod p,\ B :\ g^{c \cdot b} \mod p$. Beide Schlüssel kennt auch M

1.6 ElGamal-Public Key Verfahren (1984)

1.6.1 Schlüsselerzeugung

A wählt p, g wie bei Diffie-Hellman. Wählt $a \in \{2, \ldots, p-2\},\ x = g^a \mod p$.
Öffentlicher Schlüssel: (p, g, x)
Geheimer Schlüssel: a

1.6.2 Verschlüsselung

Klartext $m :\ 1 \leq m \leq p-1$
$B \xrightarrow{m} A$
B wählt zufällig $b \in \{2, \ldots, p-2\}$
$y = g^b \mod p$
Er berechnet $x^b \mod p$ und $f = m \cdot x^b \mod p$, sendet (y, f) an A,

1.6.3 Entschlüsselung

$y^a \mod p \ (= x^b \mod p)$

Berechnet $(y^a)^{-1} \mod p \ [(y^a)^{-1} = (y^{\overbrace{p-1-a}^{\geq 0}})]$
$f \cdot (y^a)^{-1} \mod p = m$

Nachteil zu RSA

Doppelte Länge wird gebraucht, da Nachricht (Chiffre) und Teilschlüssel versendet werden.

1.7 Signaturen, Hashfunktionen, Authentifizierung

1.7.1 Anforderung an digitale Signaturen

Identitätseigenschaft: ID des Unterzeichners des Dokuments wird sichergestellt

Echtheitseigenschaft: des signiertem Dokument

Verifikationseigenschaft: Jeder Empfänger muss digitale Signatur verifizieren können.

1.7.2 RSA-Signatur (vereinfachte Version)

A will Dokument m signieren.
A besitzt öffentlichen RSA-Schlüssel (n, e), geheimen Schlüssel d. Signatur: $m^d \mod n$ sendet $(m,\ m^d \mod n)$ an B.
$(m^d \mod n)^e = m^{e \cdot d} \mod n = m \ (\mathrm{mod}\ n)$
$m < n$
Wenn $m^{e \cdot d} \mod n = m$, dann akzeptiert B die Signatur.
$m > n \ \ m^d \mod n \ B \mod n$. Ist $m' \mod n = m \mod n$, dann $(m', m^d \mod n)$ gültige Signatur.

1.7.3 Wie lassen sich lange RSA-Signaturen vermeiden?

Definition Sei R ein endliches Alphabet.
Hashfunktion $H : \mathbb{R}^* \to R^k (k \in \mathbb{N}$ fest $)$ soll effizient berechenbar sein.

1.7.4 RSA-Signatur mit HASH-Funktion

H öffentlich bekannte Hashfunktion.
A will Nachricht m signieren.
Bildet $H(m)$ und signiert $H(m) : H(m)^d \mod n$ sendet $(m, H(m)^d \mod n)$

Verifikation durch B: $m \to H(m)$
$(H(m)^d \mod n)^e \mod n = H(m)$

1.7.5 Angriffsmöglichkeiten

- Angreifer kann $H(m)$ bestimmen wenn es ihm gelingt, $m' \neq m$ zu finden, so $(m', H(m)^d \mod n)$ gültige Signatur von m durch A.
- Angreife wählt zufällig y und berechnet $y^e \mod n = z$
  Gelingt es ihm, m zu finden mit $H(m) = z$, dann ist (m, y) gütlige Signatur von m durch A
  $H(m)\ \ y^e = H(m)$

Definition: Eine **kryptographische Hashfunktion** ist eine Hashfunktion, die folgende Bedinungen erfüllt.

1. H ist Einwegfunktion (um Angriffe des zweiten Typs zu vermeiden)
2. H ist **schwach kollisionsresistent**, d.h. zu gegebenem $m \in R^*$, soll es effizient nicht möglich sein ein $m' \neq m$, mit $H(m) = H(m')$, zu finden. (um Angriffe des ersten Typs zu vermeiden)

Verschärfung von 2.

2′ H ist **stark kollisionsresistent**, wenn es effizient nicht möglich ist $m \neq m'$ zu finden, mit $H(m) = H(m')$.

Da R^* unendlich und $\left|R^k\right| = |R|^k$ endlich ist, existiert unendlich viele Paare $(m, m'),\ m \neq m'$ mit $H(m) = H(m')$.
(Bilde $|R|^k + 1$ viele Hashwerte: Kollision)
Kollisionen lassen sich nicht vermeiden, sie sollten aber nicht schnell herstellbar sein.

1.7.6 Satz: Geburtstagsparadoxon

Ein Merkmal komme in m verschiedenen Ausprägungen vor. Jede Person besitze genau eine dieser Merkmalsausprägungen. Ist $c \geq \frac{1+\sqrt{1+8\cdot m\cdot \ln 2}}{2} \approx 1.18\sqrt{m}$, so ist die Wahrscheinlichkeit, dass unter l Personen zwei die gleiche Merkmalsausprägung haben, mindestens $\frac{1}{2}$ (Geburtstage: $m = 366, l = 23$).

Beweis l Personen
Alle Möglichkeiten $(g_1, g_2, \ldots, g_l), g_i \in \{1, \ldots, m\}$ m^l Möglichkeiten.
Alle Merkmalausprägungen verschieden: $m \cdot (m-1) \cdot (m-2) \cdot \ldots \cdot (m-(l-1))$
Wahrscheinlichkeit, dass l Personen lauter verschiedene Geburtstage haben.

$$q = \frac{m \cdot (m-1) \cdot (m-2) \cdot \ldots \cdot (m-(l-1)}{m^l} = \prod_{0}^{l-1} 1 - \frac{i}{m}$$

Wann ist $q \leq \frac{1}{2}$?
$e^x \geq 1 + x$

$$\prod_{0}^{l-1} 1 - \frac{i}{m} \leq \prod_{0}^{l-1} e^{-\frac{i}{m}} = e^{\prod_0^{l-1} -\frac{i}{m}}$$
$$= e^{-\frac{1}{m}\sum_0^{l-1} i}$$
$$= e^{-\frac{1}{m}\cdot\frac{l\cdot(l-1)}{2}}$$

$$\ln a \leq -\frac{1}{m} \cdot \frac{l \cdot (l-1)}{2} = -\frac{l^2 - l}{2 \cdot m}$$

1.7.7 Hashfunktion

$H(m) = H(m'), m \neq m'$
$H : \mathbb{Z}_2^* \to \mathbb{Z}_2^n$ (2^n Hashwerte)
Bei Erzeugung von circa $2^{\frac{n}{2}}$ Hashwerten ist die Wahrscheinlichkeit, dass zwei gleich sind ungefähr $\frac{1}{2}$.
$n = 64 : 2^{32}$ Hashwerte ($4 \cdot 10^9$) unsicher.
Weit verbreitet waren und sind:

MD5 (message digest / Ron Rivest, 1991, 128 Bit)

SHA-1 (Secure Hash Algorithm, NSA, 1992/1993, 160 Bit)

1.7.8 Authentifizierung

Nachweise bzw. Überprüfung, dass jemand derjenige ist für den er sich ausgibt.
Möglichkeiten der Authentifizierung durch:

Wissen

Besitz

biometrische Merkmale

gängiste Methode: Passwort
Im Allgemeinem: Passwort w abgespeichert als $f(w)$ f Einwegfunktion.
w $f^n(w) = w_0 \xrightarrow{sicher}$ Id. überprüfer f Einweg.

1. Auth. $w_1 = f^{n-1}(w) \to f(f^{n-1}(w)) = w_0$ ersetzt w_0 durch w_1

2. Auth. $w_2 = f^{n-2}(w) \to \ldots$

Passwortsicherheit: http://www.schneier.com/crypto-gram-0701.html

1.7.9 Challenge-Response-Authentifizierung

RSA-Verfahren $A \xrightarrow{auth.} B$
Öffentlicher Schlüssel: (n, e)
geheimer Schlüssel: d
$A \xleftarrow{\text{Zufallszahl } r} B,\ r < n\ \leftarrow$ **Challenge**
$A \xrightarrow{r^d \mod n} B$ überprüft, ob $r^{d^e} \mod n = r\ \leftarrow$ **Response**
Damit B sich sicher seien kann, dass es wirklich A ist, kann B so oft wie es für nötig hält neue r schicken und dadurch die Chance verringern, dass A nicht A ist.

1.8 Secret Sharing Scheme

Geheimnis wird auf mehrere Teilnehmer verteilt (Teilgeheimnisse), so dass gewisse Teilmengen der Teilnehmer das Geheimnis mit ihren Teilgeheimnissen rekonstruieren können, die anderen nicht.

$T = \{t_1, \ldots, t_n\}, \quad k < n \quad$ (T Menge der Teilnehmer)

Jede Teilmenge von T mit mindestens k Teilnehmer sollen Geheimnis rekonstruieren können, Teilmengen von T mit weniger als k Teilnehmer nicht.

1.8.1 (k, n) - Schwellenwertsysteme

1979 Shamir (How to share a secret)

Konstruktion

Vereinbarung von großer Primzahl p, mindestens $p \geq n + 1$

$$g \in \mathbb{Z}_p = \{0, \ldots, p - 1\}$$

Verteilung der Teilgeheimnisse

Dealer wählt zufällig $a_1, \ldots, a_{k-1} \in \mathbb{Z}_p, a_{k-1} \neq 0, k =$ Schwelle

$$f(x) = g + a_1 x + \ldots + a_{k-1} x^{k-1} \in \mathbb{Z}_p[x]$$

($a_1, \ldots, a_{k-1}$ hält er geheim, natürlich auch g)

Dealer wählt zufällig $x_1, \ldots, x_n \in \mathbb{Z}_p$ (paarweise verschieden).
Teilnehmer t_i erhält als Teilgeheimnis $(x_i, f(x_i))$ (Punkt auf Polynom)
Bei $x = 0$ hast du g.

Rekonstruktion(sversuch) des Geheimnisses

k Teilnehmer $(x_{i_1}, f(x_{i_1})), \ldots, (x_{i_k}, f(x_{i_k}))$
Durch diese Punkte ist f eindeutig bestimmt, z.B. durch Lagrange-Interpol.:

$$\begin{aligned}
f(x_{i_j}) &= g_{i_j} \\
f(x) &= \sum_{j=1}^{k} g_{i_j} \cdot \frac{(x - x_{i_1}), \ldots, (x - x_{i_{j-1}})(x - x_{i_{j+1}}), \ldots, (x - x_{i_k})}{(x_{i_j} - x_{i_1}), \ldots, (x_{i_j} - x_{i_{j-1}})(x_{i_j} - x_{i_{j+1}}), \ldots, ((x_{i_j} - x_{i_k})} \\
f(0) &= g \\
g &= \sum_{j=1}^{k} g_{i_j} \prod_{l \neq j} \frac{x_{i_l}}{(x_{i_l} - x_{i_j})}
\end{aligned}$$

Bei mehr als k Teilnehmer selbe Ergebnis.
Weniger als k Teilnehmer (k'): Anderes Polynom wegen weniger Punkte, also wahrscheinlich anderer g.
Erzeugen Polynom vom Grad $\leq k' - 1$
Für alle $k \in \mathbb{Z}_p$ existiert gleich viele Polynome vom Grad $\leq k' - 1$ durch die vorgegebene k' Punkte, die bei h durch y-Achse gehen.

Kapitel 2

Codierungstheorie

2.1 Grundbegriffe und einfache Beispiele

2.1.1 Codierung

(Kanalcodierung)
Sicherung von Daten/Nachrichten gegen zufällig auftretenden Fehler bei Speicherung/Übertragung.

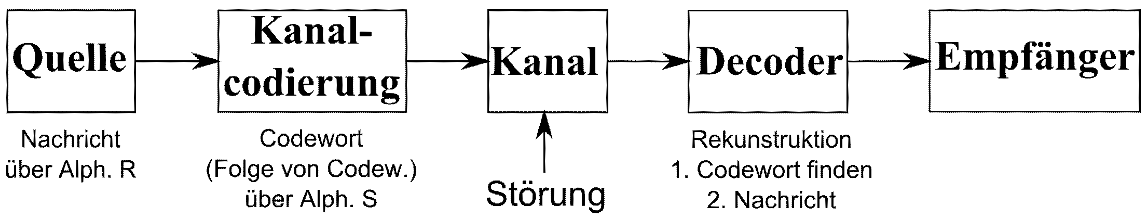

Abbildung 2.1: Schaubild der Codierung

2.1.2 Ziele

- Möglichst viele Fehler erkennen und gegebenenfalls korrigieren.
- Aufwand für Codierung und Decodierung möglichst gering.

2.2 Grundprinzip

Hinzufügen von Redundanz
Es gibt zwei Typen um Redundanz zu erzeugen.

2.2.1 FEC-Verfahren (Forward Error Correction)

Aufgetretene Fehler sollen erkannt und korrigiert werden.
Vorteil: keine Verzögerung der Übertragung aber ggf. große Redundanz notwendig.

2.2.2 ARQ-Verfahren (Automatic Repeat Request)

Aufgetretene Fehler sollen erkannt werden, werden nicht korrigiert. Stattdessen wiederholt die Übertragung beim Sender anfordern.
Vorteil: geringe Redundanz, aber Verzögerung.

Beispiele

1. Parity-Check-Codes
   z.B. Nachrichten: 00, 01, 10, 11
   Codierung:

$$\begin{aligned} 00 &\to 000 \\ 01 &\to 011 \\ 10 &\to 101 \\ 11 &\to 110 \end{aligned}$$

   (gerade Anzahl von Einsen in den Codewörtern)
   1 Fehler wird erkannt, nicht korrigiert.
   2 Fehler werden nicht erkannt.

2. Wiederholungscode
   Nachrichten wie in 1.

   Codierung:

$$\begin{aligned} 00 &\to 000000 \\ 01 &\to 010101 \\ 10 &\to 101010 \\ 11 &\to 111111 \end{aligned}$$

   (3-Fache Wiederholung)

   1 Fehler wird erkannt und korrigiert.
   $\underline{01}11\underline{01} \overset{\text{Decod}}{\curvearrowright} 010101 \curvearrowright 01$

3. Nachrichten wie in 1. und 2.

Codierung:

$$\begin{aligned} 00 &\to 00000 \\ 01 &\to 01101 \\ 10 &\to 10110 \\ 11 &\to 11011 \end{aligned}$$

Je zwei Codewörter unterscheiden sich an mindestens 3 Positionen.
Angenommen 1 Fehler tritt bei Übertragung auf. Dann gibt es genau ein Codewort, dass sich vom empfangenen Wort an genau einer Stelle unterscheidet; in das wird decodiert.

4. (ehmaliger) ISBN-Code
International Standard Book Number

10-Stelliger Code
Erste 9 Ziffern haben inhaltliche Bedingung ($\hat{=}$ Nachricht)
10. Ziffer: Prüfziffer

Beispiel: 3-540-26121-? (Land - Verlag - Buchnummer - Prüfziffer)

Uncodierte Wörter sind gebildet über $R = \{0, \dots, 9\}$
Codierte Wörter sind gebildet über $S = \{0, \dots, 9, X\}$
ISBN-Wort $C_{10}C_9 \dots C_2C_1$
$C_{10} \dots C_2$ inhaltliche Bedeutung, C_1 wird so gewählt, dass

$$\sum_{k=1}^{10} k \cdot C_k \equiv 0 (\text{mod } 11)$$
$$10 \cdot C_{10} + \dots + 2 \cdot C_2 + C_1 \equiv 0 (\text{mod } 11)$$

falls $C_1 = 10$ so setzte $C_1 = X$
C_1 vom Beispiel ausrechnen.

$$10 \cdot 3 + 9 \cdot 5 + 8 \cdot 4 + 7 \cdot 0 + 6 \cdot 2 + 5 \cdot 6 + 4 \cdot 1 + 3 \cdot 2 + 2 \cdot 1 + C_1 \equiv 0 (\text{mod } 11)$$
$$161 + C_1 \equiv 0 (\text{mod } 11) \Rightarrow C_1 = 4$$

Ändern einer Ziffer wird erkannt:
$C_{10}C_9 \dots C_2C_1 \curvearrowright C_i$ wird $X_i \neq C_i$ ersetzt
$C_{10} \dots C_{i+1}X_iC_{i-1} \dots C_1$

$$\sum_{k=1, k\neq i}^{10} k \cdot C_k + i \cdot x_i = \underbrace{\sum_{k=1, k\neq i}^{10} k \cdot C_k}_{\equiv 0 (\text{mod } 11)} + \overbrace{\underset{\substack{\uparrow \\ \not\equiv 0 (\text{mod } 11)}}{i} \cdot (\underbrace{x_i - c_i}_{\not\equiv 0 (\text{mod } 11)})}^{\not\equiv 0 (\text{mod } 11) \to 11 = Prim} \not\equiv 0 (\text{mod } 11)$$

Fehler wird erkannt, Korrektur nicht möglich.

$3-540-26121-4 \equiv 0 (\text{mod } 11)$

$\left.\begin{array}{l} 3-540-26121-\mathbf{6} \\ 3-540-2612\mathbf{2}-4 \end{array}\right\}$ Prüfsumme 2.
Vertauschung von Zwei Ziffern wird erkannt.
C_i und C_j vertauscht.
O.B.d.A $C_i \neq C_j$
$C_{10} \dots \underset{\substack{\uparrow \\ i}}{C_j} \dots \underset{\substack{\uparrow \\ j}}{C_i} \dots C_1$

$$\begin{aligned}\sum_{k=1,k\neq i,j}^{10} k \cdot C_k + i \cdot C_j + j \cdot C_i &= \sum_{k=1}^{10} k \cdot C_k + i(C_j - C_i) + j(C_i - C_j) \\ &= \underbrace{\sum_{k=1}^{10} k \cdot C_k}_{\equiv 0 (\text{mod } 11)} + \underbrace{(C_j - C_i)}_{\not\equiv 0 (\text{mod } 11)} \underbrace{(i-j)}_{\not\equiv 0 (\text{mod } 11)} \not\equiv 0 (\text{mod } 11)\end{aligned}$$

Vertauschung wird durch gewichtete Quersummen erkannt.

5. EAN-13-Code
European Article Number
13-Stelliger Code, erste 12 Ziffer sind inhaltlich festgelegt. 13. Ziffer ist Prüfziffer.
$R = S = \{0, \dots, 9\}$

$$C_1 \dots C_{12} C_{13}$$

$C_1 \dots C_{12}$ inhaltliche Angabe (in der Regel):
$C_1 C_2$ Herstellerland (40-43 Deutschland)
$C_6 \dots C_7$ Hersteller $C_8 \dots C_{12}$ interne Produktions Nummer
C_{13} so gewählt, dass

$$C_1 + 3 \cdot C_2 + C_3 + 3 \cdot C_4 + \dots + 3 \cdot C_{12} + C_{13} \equiv 0 (\text{mod } 10)$$

1 Fehler wird erkannt. $x \to 3 \cdot x (\text{mod } 10)$ Permutation auf $\mathbb{Z}_{10}$, da $ggT(3, 10) = 1$, Vertauschung werden nicht immer erkannt.
Übersetzung in Barcode:

$$C_1 C_2 \dots C_7 C_8 \dots C_{13}$$

Jede der Ziffern $C_2, \dots, C_{13}$ wird durch einen 0-1-String der Länge 7 binär codiert.
$0 \mathrel{\hat{=}}$ weißer Balken, $1 \mathrel{\hat{=}}$ schwarzer Balken.

7 501031 311309

Abbildung 2.2: EAN-13 Barcode

Codierung sorgt dafür, dass nie mehr als 4 weiße oder schwarze Balken nebeneinander stehen.

Schmalen Balken in Mitte und am Rand, sind nur Abtrennzeichen, die nichts mit EAN zu tun haben und nur beim einscannen helfen.

$$5 \text{ zu } 0110001_2$$

$C_2, \ldots, C_7$ werden nach Code A oder Code B codiert. C_1 bestimmt welcher dieser beiden Codes verwendet wird.
$C_8, \ldots, C_{13}$ werden nach Code C codiert.
C_1 ergibt sich aus der Art der Codierung von $C_2, \ldots, C_7$

	Ziffern $\mathbf{C_2 - C_7}$		**Ziffern $\mathbf{C_8 - C_{13}}$**	**bestimmt durch $\mathbf{C_1}$**
Zeichen	**Code A**	**Code B**	**Code C**	**Code D**
0	0001101	0100111	1110010	AAAAAA
1	0011001	0110011	1100110	AABABB
2	0010011	0011011	1101100	AABBAB
3	0111101	0100001	1000010	AABBBA
4	0100011	0011101	1011100	ABAABB
5	0110001	0111001	1001110	ABBAAB
6	0101111	0000101	1010000	ABBBAA
7	0111011	0010001	1000100	ABABAB
8	0110111	0001001	1001000	ABABBA
9	0001011	0010111	1110100	ABBABA

Codewörter von Code A,B oder C kommen nur einmal vor. Daher treten nie mehr als 4 gleiche Balken nebeneinander auf.

2.3 Blockcodes

$$
\begin{aligned}
00 &\rightarrow 00000\\
01 &\rightarrow 01101\\
10 &\rightarrow 10110\\
11 &\rightarrow 11011
\end{aligned}
$$

2.3.1 Definition

S endl. Menge (=Alphabet), $n \in \mathbb{N}$.
Ein Blockcode C der (Block-)Länge n über S ist Teilmenge von $S^n = \underset{\longleftarrow\ n\ \longrightarrow}{S \times \ldots \times S}$

Elemente von C heißen **Codewörter**.
Ist $|S| = 2$ (i.d.R. $S = \{0, 1\}$, so **binär** Code.
$|C| = m$, so ist $m \leq |S|^n$.

Dann lassen sich n Informationssymbole (oder Strings von Informationssymbolen) codieren (Codierungsfunktion). Folge von Informationssymbolen (oder Strings) werden dann in Folge von Codewörtern codiert.

2.3.2 Definition: Hamming-Abstand

S endl. Alphabet, $n \in \mathbb{N}$.
$a, b \in S^n$ $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$
$d(a, b) = \sharp\{i : a_i \neq b_i\}$
Hamming-Abstand von a und b (Anzahl der unterschiedlichen Stellen).
(Richard W. Hamming, 1915-1998, Begründer der Codierungstheorie)

Eigenschaften

a) $d(a, b) = 0 \Leftrightarrow a = b$

b) $d(a, b) = d(b, a)$

c) $d(a, b) \leq d(a, c) + d(c, b)$ (Dreiecksungleichung)
$(a_i \neq b_i \Rightarrow a_i \neq c_i$ oder $b_i \neq c_i)$

d) Wenn $(S, +)$ komm. Gruppe, dann auch S^n
$[(a_1, \ldots a_n) + (b_1, \ldots b_n) = (a_1 + b_1, \ldots, a_n + b_n)]$
$d(a, b) = d(a + c, b + c)$ (Translationsinvarianz)

Also: Wird $x \in C$ gesendet und $y \in S^n$ wird empfangen und $d(x, y) = k$, so sind k Fehler aufgetreten.

2.3.3 Definition

a) Hamming-Decodierung

für Blockcode $C \subseteq S^n$
Wird $y \in S^n$ empfangen, so wird y zu einem Codewort $x' \in C$ decodiert, das unter allen Codewörtern minimalen Hamming-Abstand zu y hat.

$$d(x', y) = min\ d(x, y), x \in C$$

(x' muss nicht eindeutig bestimmt sein)
z.B. $C = \{(0000), (1111)\}$
Empfangen: 0011 x' nicht eindeutig in diesem Fall.

($|S| = 2$: Hamming-Decodierung ist bestmöglich, falls jedes Symbol in einem Codewort mit der gleichen Wahrscheinlichkeit $p < \frac{1}{2}$ verändert wird und wenn jedes Codewort gleich wahrscheinlich ist.)

b) Minimalabstand

C Blockcode in S^n, Minimalabstand von C:

$$d(C) = min\ d(x, x'),\ \ x, x' \in C, x \neq x'$$

(Ist $|C| = 1$, so $d(C) = n$)
$[Bsp : C = \{(00000), (01101), (10110), (11011)\}, d(C) = 3]$

c)

Ein Blockcode C ist **t-Felder-korrigierend**, falls $d(C) \geq 2t + 1$, und er heißt **t-Fehler-erkennend**, falls $d(C) \geq t + 1$.

Begründung für die Bezeichnung in c)
„Kugel“ vom Radius t um $x \in C : K_t(x) = \{y \in S^n : d(x, y) \leq t\}$

Ist $d(C) \geq 2t + 1$, so sind Kugelm vom Radius t um Codewörter disjunkt.

Angenommen es existiert $y \in S^n$ mit $y \in K_t(x) \cap K_t(x')$, $x, x' \in C, x \neq x'$. Dann $d(x, x') \leq d(x, y) + d(y, x') \leq t + t = 2t$. Widerspruch
$x \in C$ gesendet, y wird empfangen, und angenommen maximal t-Fehler sind aufgetreten, dann $y \in K_t(x)$ und Abstand zu jedem anderem Codewort ist $> t$
$\Rightarrow$ Hamming-Decodierung ist korrekt.
$d(C) \geq t + 1$ und es treten maximal t minimal 1 Fehler auf, so ist y kein Codewort.

Bsp:

a) n-fach Wiederholungscode

$$\begin{array}{lcl} S_n & \to & \underbrace{S_1 S_1 \dots S_1}_{\longleftarrow\ n\ \longrightarrow} \\ \vdots & & \\ S_k & \to & \underbrace{S_k S_k \dots S_k}_{\longleftarrow\ n\ \longrightarrow} \end{array}$$

$C = \{(s, s, \dots, s) : s \in S\} \subseteq S^n$
$d(C) = n$

$\left\lfloor \frac{n-1}{2} \right\rfloor$-Fehler-korr.

b) ISBN, EAN-Codes, $d(C) = 2$, 1-Fehler-erkennend.

2.3.4 Definition: Perfekter Code

Code $C \subseteq R^n$ heißt perfekt, falls es ein $t \in \mathbb{N}_0$ gibt, mit der Eigenschaft:

$$R^n = \bigcup_{x \in C} K_t(x) \quad \text{und} \quad K_t(x) \cap K_t(x') = \emptyset \quad \text{für} \quad x, x' \in C,\ x \neq x'$$

Dann ist $d(C) = 2 \cdot t + 1$, falls $|C| > 1$:
Ang. $d(C) \leq 2 \cdot t$. Wähle $x, x' \in C,\ x \neq x'$, mit $d(x, x') = d(C) \leq 2 \cdot t$.
Wähle $y \in R^n$ mit $d(x, y) = t,\ d(y, x') \leq t$
$y \in K_t(x) \cap K_t(x')$ Widerspruch
$d(C) \leq 2 \cdot t + 1$

Wähle $x \in C$, wähle $y \in R^n$ mit $d(x, y) = t + 1$. Nach Vorraussetzung existiert $x' \in C$ mit $y \in K_t(x')$.

$$d(x, x') \leq d(x, y) + d(y, x') \leq t + 1 + t = 2 \cdot t + 1$$
$$d(C) \leq 2 \cdot t + 1$$

2.3.5 Gibt es perfekte Codes?

Trivial Beispiele:

- einelementige Codes (t=n)
- $C = R^n$ (t=0)
  (Jedes Element ist ein Codewort)
- n-fache Wiederholungscode über Z_2
  $n = 2 \cdot t + 1$
  $C = \{(\underbrace{0, \dots, 0}_{\longleftarrow n \longrightarrow}), (\underbrace{1, \dots, 1}_{\longleftarrow n \longrightarrow})\}$

2.3.6 Lemma

$|R| = q,\ x \in R^n,\ t \in \mathbb{N}$
Dann ist $|K_t(x)| = \sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i$
$\left(\binom{n}{i} = \frac{n!}{i! \cdot (n-i)!}\right)$

Beweis

Abstand 0 zu x: 1 Wort (nämlich x): $\binom{n}{0} \cdot (q-1)^0 = 1$
Abstand $i > 0$ zu x:
Anzahl der Auswahl von i Positionen aus n Positionen: $\binom{n}{i}$
An jeder Position $q-1$ Änderungsmöglichkeiten.
$\rightarrow$ insgesamt $(q-1)^i$ Möglichkeiten,
Anzahl der Wörter vom Abstand i von x: $\binom{n}{i} \cdot (q-1)^i$

Satz

Sei C ein Code der Länge n über R, $|C| > 1, |R| = q$. Sei $t \in \mathbb{N}_0$ maximal mit $d(C) \geq 2 \cdot t + 1,\ t = \left\lfloor \frac{d(C)-1}{2} \right\rfloor$.

a) (Kugelpackungsschranke)
$|C| \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i}$

b) C ist perfekt $\Leftrightarrow$ in a) gilt Gleichheit, d.h.
$|C| = \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i}$

Beweis

a)

$$d(C) \geq 2 \cdot t + 1,\ \text{daher } K_t(x) \cap K_t(x') = \emptyset,\ x \neq x',\ x, x' \in C$$

$$R^n \geq \dot{\bigcup_{x \in C}} K_t(x)$$

$$q^n = \left|R^n\right| \geq \left|\dot{\bigcup_{x \in C}} K_t(x)\right| = \sum_{x \in C} |K_t(x)| \underset{Lemma}{=} |C| \cdot \sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i$$

b)

$$\Rightarrow: d(C) = 2 \cdot t + 1$$

$$R^n = \dot{\bigcup_{x \in C}} K_t(x) \Rightarrow \text{Gleichheit in a)}$$

$$\Leftarrow: \text{Gleichheit} \Rightarrow R^n = \dot{\bigcup_{x \in C}} K_t(x) \Rightarrow C \text{ perfekt.}$$

2.3.7 Binärer Hamming-Code der Länge 7

$R = \mathbb{Z}_2 = \{0, 1\}$, C perfekt, $d(C) = 3$, $|C| = 16$, (korrigiert 1 Fehler)

$$\begin{aligned} C = \{(C_1, \ldots, C_7) : C_i \in \mathbb{Z}_2,\ & C_1 + C_4 + C_6 + C_7 = 0, \\ & C_2 + C_4 + C_5 + C_7 = 0, \\ & C_3 + C_5 + C_6 + C_7 = 0 \\ & \} \subseteq \mathbb{Z}_2^7 \end{aligned}$$

C ist Unterraum von $\mathbb{Z}_2^7$
$(C_1, \ldots, C_7) \in C,\ (C'_1, \ldots, C'_7) \in C$
$(C_1 + C'_1, \ldots, C_7 + C'_7) \quad (C_1 + C'_1) + (C_4 + C'_4) + (C_6 + C'_6) + (C_7 + C'_7) = 0$
$dim(C) = 4,\ C_4, C_5, C_6, C_7$ frei wählbar $\curvearrowright$ C_1, C_2, C_3 festgelegt
Basis:

$$\begin{aligned} &(\ldots 1000) \to (1101000) \\ &(\ldots 0100) \to (0110100) \\ &(\ldots 0010) \to (1010010) \\ &(\ldots 0001) \to (1110001) \\ &|C| = 2^4 = 16 \end{aligned}$$

$d(C) = 3$:
Ang. $d(C) = d$. Wähle $x, x' \in C$ mit $d(x, x') = d$
Translationsinvarianz der Metrik:

$$\begin{aligned} d &= d(x, x') = d(x + x, x + x') = d(0, x + x') \\ wt(x) &= \text{ Anzahl der Einsen in } x \\ &= d(0, x) \\ d(C) &= \min wt(x), \quad x \in C, \quad x \neq \vec{0} \end{aligned}$$

Zeige: Jeder Vektor $\neq \vec{0}$ in C enthält mind. 3 Einsen.
$d(C) = 3$ weist man nach durch überprüfen aller 15 von $\vec{0}$ verschiedenen Co-

dewörtern oder durch Analyse der Gleichung.

$(C_1, \ldots, C_7) \in C$ Ang. $C_7 = 1$
$\Rightarrow C_1 + C_4 + C_6 = 1$. Wenn alle Eins $\checkmark$, sonst 3 Fälle:
$C_1 = 1, C_4 = C_6 = 0$
$C_4 = 1, C_1 = C_6 = 0$
$C_6 = 1, C_1 = C_4 = 0$
Addition alle Gleichungen $\Rightarrow C_1 + C_2 + C_3 + C_7 = 0$
C_1, C_2 oder $C_3 = 1$
2. Fall: $C_7 = 1, C_4 = 1, C_1 = 0, C_2$ oder $C_3 = 1$
3. Fall: analog zu Fall 2.
1. Fall: $C_1 = 1, C_4 = C_6 = 0, C_7 = 1,$ o.B.d.A. $C_2 = C_3 = 0 \Rightarrow C_5 = 1$
$d(C) \leq 3,\ d(C) = 3 = 2 \cdot 1 + 1$

Prüfe nach, ob bei Kugelpackungsschranke Gleichheit gilt:

$$|C| = 16$$

$$|C| \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i} \qquad (q = 2, t = 1, n = 7)$$

$$= \frac{2^7}{1 + \binom{7}{1}} = \frac{2^7}{2^3} = 2^4 = 16$$

C perfekt!

2.4 Lineare Codes

2.4.1 Definition: linearer Code

Sei K ein endlicher Körper, $n \in \mathbb{N}$. Ein linearer Code C der Länge n ist ein Unterraum von K^n. (Zeilenvektoren) [Alphabet = K]
Ist $dim(C) = k$, so heißt C $[n, k]$-Code.
Ist $d(C) = d$, so $[n, k, d]$-Code.
Beachte: $|K| = q \Rightarrow |C| = q^k$.

2.4.2 Definition: Informationsrate

Informationsrate (Rate) von C : $\frac{k}{n}$.

2.4.3 Bemerkung über endliche Körper

a) p Primzahl, $\mathbb{Z}_p$ ist Körper der Ordnung p

b) K endlicher Körper $\Rightarrow |K| = p^m$, p Primzahl, $m \in \mathbb{N}$.

c) Zu jeder Primzahlpotenz p^m existiert (bis auf Isomorphie) genau ein Körper der Ordnung p^m.

d) f sei irreduzibles Polynom vom Grad m über $\mathbb{Z}_p$.
$K = \{g \in \mathbb{Z}_p[x] : Grad(g) \leq m - 1\}, \quad |K| = p^m$
K wird Körper:
Addition = übliche Addition von Polynomen
Multiplikation = normale Multiplikation + Reduktion mod f
(AES : $|K| = 2^8$)

2.4.4 Bsp

a) n-facher Wiederholungscode über $\mathbb{Z}_p$
$C = \{(\underbrace{0, \ldots, 0}_{\longleftarrow n \longrightarrow}), (1, \ldots, 1), \ldots, (p-1, \ldots, p-1)\}$
C ist linearer Code, $C =< (1, \ldots, 1) >$
$[n, 1, n]$-Code

b) Hamming-Code ist linearer [7,4,3]-Code über $\mathbb{Z}_2$

c) $C = \{(C_1, \ldots, C_n) : C_i \in \mathbb{Z}_p, \sum_{i=1}^n C_i = 0\}$
($p = 2$: Parity Check Code)
linear $[n, n-1, 2]$-Code über $\mathbb{Z}_p$
Basis von C : $(1, 0, \ldots, 0, p-1), (0, 1, 0, \ldots, 0, p-1), \ldots, (0, \ldots, 0, 1, p-1)$

2.4.5 Definition: Gewicht und Minimalgewicht

K endl. Körper

a) $x \in K^n$, so Gewicht von x, $wt(x)$, definiert durch

$$wt(x) = \sharp\{i : x_i \neq 0\}$$

b) $\{\vec{0}\} \neq C \subseteq K^n$, so ist das Minimalgewicht von C definiert durch

$$wt(C) = \min_{x \in C, x \neq \vec{0}} wt(x)$$

2.4.6 Satz

Ist $C \neq \{\vec{0}\}$ ein linearer Code, so ist $d(C) = wt(C)$. (Beweis wie beim [7,4,3]-Hamming Code)

2.4.7 Definition: Erzeugermatrix

Sei C ein $[n,k]$-Code über K, sei $g_1 = (g_{11}, \dots, g_{1n}), \dots, g_k = (g_{k1}, \dots, g_{kn})$ eine Basis von C.

Dann heißt die $k \times n$ -Matix $G = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} = \begin{pmatrix} g_{11} & \dots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \dots & g_{kn} \end{pmatrix}$ Erzeugermatrix von C

2.4.8 Satz

Sei G ein Erzeugermatrix von C. Dann ist $C = \{ \underbrace{\overset{u \cdot G}{1 \times k \cdot k \times n}}_{1 \times n} : u \in K^k \}$

Beweis:

$$u = (u_1, \dots, u_k), u_i \in K$$
$$uG = (u_1, \dots, u_k) \cdot (g_1, \dots, g_k)^t = u_1 g_1 + \dots + u_k g_k \in C$$

2.4.9 Bemerkung

a) Die Abb $\begin{cases} K^k & \longrightarrow C \\ u & \longmapsto uG \end{cases}$ ist bijektiv.
   $u \in K^k$ Informationswörter
   Codiert in Codewörter durch uG.

b) Elementare Zeilenumformungen an Erzeugermatrix liefern Erzeugermatrix.

2.4.10 Beispiel: Hamming-$[7,4]$-Code über $\mathbb{Z}_2$

$$\begin{aligned} C = \{(C_1, \dots, C_7) : C_i \in \mathbb{Z}_2, & \; C_1 + C_4 + C_6 + C_7 = 0, \\ & C_2 + C_4 + C_5 + C_7 = 0, \\ & C_3 + C_5 + C_6 + C_7 = 0 \\ & \} \subseteq \mathbb{Z}_2^7 \end{aligned}$$

Erzeugermatrix:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Codierung eines Informationswort (u_1, u_2, u_3, u_4) mit G

$(u_1, u_2, u_3, u_4) \to (u_1, u_2, u_3, u_4) \cdot G = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, u_1 + u_3 + u_4, u_2 + u_3 + u_4, u_1 + u_2 + u_3)$

2.4.11 Definition: Standardform

C $[n, k]$-Code Erzeugermatrix C ist in Standardform, falls sie folgende Gestalt hat.

$$G = \begin{pmatrix} 1 & & 0 & & \\ & \ddots & & & * \\ 0 & & 1 & & \end{pmatrix}$$

Codierung $(u_1, \ldots, u_k) \cdot G = (u_1, \ldots, u_k, *, \ldots, *)$

2.4.12 Satz

Sei C ein $[n, k]$-Code über K. Dann existiert $(n-k) \times n$-Matrix H über K mit folgenden Eigenschaften:
Sei $y \in K^n$. Dann: $y \in C \Leftrightarrow H \cdot y^t = \vec{0}$
H heißt Kontrollmatrix von C ($\Leftrightarrow y \cdot H^t = \vec{0}$)
Es ist $rg(H) = n - k$ (Dann ist $H \cdot G^t = 0$)

2.4.13 Beweis

Sei $g_1, \ldots, g_k$ Basis von C, $G = \begin{pmatrix} g_1 \\ \vdots \\ g_2 \end{pmatrix}$
$g_i = (g_{i1}, \ldots, g_{in})$
Betrachte LGS:

$$\begin{aligned} g_{11}x_1 + \ldots + g_{1n}x_n &= 0 \\ &\vdots \\ g_{k1}x_1 + \ldots + g_{kn}x_n &= 0 \end{aligned}$$

d.h. $G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$. Koeffizientmatrix G hat Rang k.
Dimension des Löungsraums dieses LGS $= n - k$
Sei $h_1, \ldots, h_{n-k} \in K^n$ Basis des Lösungsraums dieses LGS.

$$H = \begin{pmatrix} m \\ \vdots \\ h_{n-k} \end{pmatrix}, \quad H \cdot g_i^t = \begin{pmatrix} h_1 g_i^t \\ \vdots \\ h_{n-k} g_i^t \end{pmatrix} = 0, i = 1, \ldots, k$$

$Hy^t = 0$ für alle $y \in C$.
$rg(H) = n - k \Rightarrow dim\ Kern(H) = k = dim(C)$
$C = Kern(H)$

2.4.14 Bermerkung

- Kontrollmatrix kann zur Fehlererkennung verwendet werden.
- Beweis liefert Verfahren: Erzeugermatrix $\to$ Kontrollmatrix
- Umgekehrt: Kontrollmatrix $\to$ Erzeugermatrix (Bilde Basis des Lösungsraums von $Hy^t = 0$)

2.4.15 Beispiel

a) Parity-Check-Code über $\mathbb{Z}_p$
$C = \{(c_1, \ldots, c_n) : \sum_{i=1}^n c_i = 0\}$
$H = (1, 1, \ldots, 1)$
$H \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0 \Leftrightarrow c_1 + \ldots c_n = 0 \Leftrightarrow (c_1, \ldots, c_n) \in C$

b) $[7,4]$-Hamming-Code

$$\begin{aligned} C = \{(C_1, \ldots, C_7) : C_i \in \mathbb{Z}_2, \; & C_1 + C_4 + C_6 + C_7 = 0, \\ & C_2 + C_4 + C_5 + C_7 = 0, \\ & C_3 + C_5 + C_6 + C_7 = 0 \\ & \} \subseteq \mathbb{Z}_2^7 \end{aligned}$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

c) C Code mit Erzeugermatrix

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad [4,2]\text{-Code über } \mathbb{Z}_2 \quad G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix} = 0$$

$$\begin{aligned} x_1 + x_2 + x_4 &= 0 \\ x_2 + x_4 &= 0 \end{aligned}$$

x_5, x_4 frei wählen, x_1, x_2 fesgelegt.

$$\text{Basis } (0010), (0101)$$

$$\text{Kontrollmatrix } H = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$$C\{(c_1, \ldots, c_4) : c_3 = 0, c_2 + c_4 = 0\}$$

2.4.16 Satz

C $[n,k]$-Code, $C \neq \{\vec{0}\}$, K^n, Kontrollmatrix H.

$$\begin{aligned} d(C) = wt(C) &= \min\{r : \text{ in H gibt es } r \text{ linear abhänige Spalten}\} \\ &= \max\{r : \text{ je } r-1 \text{ Spalten linear unabhängig}\} \end{aligned}$$

Beweis

$s_1, \ldots, s_n$ Spalten von H, Länge $n-k$.
$C \neq \{\vec{0}\}, k > 1, n-k < n \Rightarrow s_1, \ldots, s_n$ lin. abhängig.
Sei $\min\{r : \ldots\} = w$. $s_{i_1}, \ldots, s_{i_w}$ lin. abhängig.
Existiert $c_{i_1}, \ldots, c_{i_w} \in K$, nicht alle $= 0$, $c_{i_1}s_{i_1} + \ldots + c_{i_w}s_{i_w} = 0$
w min $\Rightarrow$ alle $c_{i_1}, \ldots, c_{i_w} \neq 0$.
Def. $c = (c_1, \ldots, c_n)$ mit den c_{i_j} an den Stellen i_j, übrige $c_i = 0$

$$\sum_{i=1}^{n} c_i s_i = c_{i_1}s_{i_1} + \ldots + c_{i_w}s_{i_w} = 0$$

$$\sum_{i=1}^{n} c_i s_i^t = 0$$

$$Hc^t = 0. \quad c \in C.$$

$wt(c) = w$, Min. Gewicht von $C \leq wt(c) = w$
Ang. es ex. $0 \neq c' \in C$, $wt(c') = w' < w$. $\quad Hc'^t = 0$
$c' = (c'_1, \ldots, c'_n) \quad \sum c'_i s_{i\,i=1}^{\;n} = 0 \Rightarrow w'$ der Spalten $c_1, \ldots, c_n$ sind linear abhänging.
Widerspruch!

$$wt(C) = w$$

2.4.17 Beispiel: $[7,4]$-Hamming-Code über $\mathbb{Z}_2$

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \text{ Kontrollmatrix}$$

Keine Spalte ist Nullspalte, keine zwei Spalten sind gleich, 1.,2.,4. Spalte sind linear abhänging.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad d(C) = 3$$

2.4.18 Korollar: (Singleton-Schranke)

Ist C ein linearer $[n,k]$-Code, $d(C) = d$, so gilt:

$$d \leq n - k + 1$$

Beweis

2. Gleichheit: $d \leq rg(H) + 1 = n - k + 1$ (Zeilen von H sind lin. unabhängig)

2.4.19 Bemerkung: (Nebenklassen von Unterräumen in Vektorräumen)

C ein Umterraum von Vektorraum V. Für jedes $v \in V$:

$$v + C = \{v + x : x \in C\}$$

Nebenklasse von C zu v.

a) $v_1, v_2 \in V$. Dann:
$v_1 + C = v_2 + C$ oder $(v_1 + C) \cap (v_2 + C) = \emptyset$

b) $v_1 + C = v_2 + C \Leftrightarrow v_1 - v_2 \in C$
$(v + C = C(= \vec{0} + C) \Leftrightarrow v \in C)$

c) Wähle aus jeder Nebenklasse einen Vektor v_i:

$$V = \dot{\bigcup}(v_i + C)$$

d) V Vektorraum über endl. Körper: $|v + C| = |C|$

e) C $[n, k]$-Code ($V = K^n, dim(C) = k, |C| = q^k,$ falls $|K| = q$)
Anzahl der Nebenklassen ist q^{n-k}

2.5 Syndrom-Decodierung linearer Code

C $[n, k]$-Code über K, $|K| = q$, Kontrollmatrix H, $(n - k) \times n$-Matrix.
Ist $y \in K^n$, so heißt $Hy^t \in K^{n-k}$ **Syndrom** von y.

a) $x \in C \Leftrightarrow Hx^t = 0$ (x hat Syndrom 0)

b) $y_1, y_2 \in K^n$. y_1, y_2 liegen in der gleichen Nebenklasse zu C (d.h. $y_1 + C = y_2 + C$)
$\Leftrightarrow y_1, y_2$ haben gleiches Syndrom
(d.h. $Hy_1^t = Hy_2^t$)

$[y_1 + C = y_2 + C \Leftrightarrow y_1 - y_2 \in C \Leftrightarrow 0 = H(y_1 - y_2)^t = Hy_1^t - Hy_2^t \Leftrightarrow Hy_1^t = Hy_2^t]$

c) Jedes $z \in K^{n-k}$ tritt als Syndrom auf.

Ang. $x \in C$ wird gesendet, $y = x + f$, wird empfangen. f „Fehlervektor“.
$y + C = f + C$, y und f haben das gleiche Syndrom, nämlich Hy^t.
Bestimmt in der Nebenklasse von y ein e mit kleinstmögliche Gewicht (**Nebenklassenführer**)
Decodierung: $y \to y - e \in C$ (Hamming-Decodierung)

Ordne die Nebenklassenführer nach der lexikogr. ihrer Syndome.
Speicherbedarf: q^{n-k} Nebenklassenführer, jeder hat Länge n
(Besser als Durchforsten der Liste aller Codewörter (q^k), falls $k \geq \frac{n}{2}$)
C $[70, 50]$-Code über $\mathbb{Z}_2$. 2^{20} Nebenklassenführer, je 70 BitLänge.
Speicher: $70 \cdot 2^{20} Bit \approx 8,75$ MegaByte
Speicher für Codewörter: $70 \cdot 2^{50}$ Bit = 9 PetaByte

2.5.1 Beispiel

C $[5, 2]$-Code über $\mathbb{Z}_2$, Kontrollmatrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

$$d(C) = 3$$

$$\begin{aligned} (x_1, \ldots, x_5) \in C \Leftrightarrow & x_1 + x_5 = 0 \\ & x_2 + x_3 = 0 \\ & x_2 + x_4 + x_5 = 0 \end{aligned}$$

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Nebenklassen von C.

$$\begin{aligned} C =& (00000) + C = \{(00000), (11101), (01110), (10011)\} \\ & (10000) + C = \{(10000), (01101), (11110), (00011)\} \\ & \quad \text{Nebenklassenführer:}(10000) \\ & (01000) + C = \{(01000), (10101), (00110), (11011)\} \\ & \quad \text{Nebenklassenführer:}(01000) \\ & (00100) + C = \{(00100), (11001), (01010), (10111)\} \\ & \quad \text{Nebenklassenführer:}(00100) \\ & (00010) + C = \{(00010), (11111), (01100), (10001)\} \\ & \quad \text{Nebenklassenführer:}(00010) \\ & (00001) + C = \{(00001), (11100), (01111), (10010)\} \\ & \quad \text{Nebenklassenführer:}(00001) \\ & (00111) + C = \{(00111), (11010), (01001), (10100)\} \\ & \quad \text{Mögliche Nebenklassenführer:}(01001), (10100) \\ & (00101) + C = \{(00101), (11000), (01011), (10110)\} \\ & \quad \text{Mögliche Nebenklassenführer:}(00101), (11000) \end{aligned}$$

Angenommen als Nebenklassenführer werden gewählt:
$f_0 = (00000), f_1 = (10000), \dots, f_5 = (00001), f_6 = (01001), f_7 = (00101)$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Syndrome:

$$Hf_0^t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, Hf_1^t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, Hf_2^t = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, Hf_3^t = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, Hf_4^t = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$Hf_5^t = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, Hf_6^t = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, Hf_7^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(Ordnung: $f_0, f_4, f_3, f_2, f_1, f_5, f_6, f_7$)

Empfangen: $y = (10110)$

$$Hy^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Decodierung: $y \to y + f_7 = (10011) \in C$
(Hätte man für die Nebenklasse $f_7 + C$ als Nebenklassenführer (11000) gewählt, so wäre decodiert worden in $y + (11000) = (01110) \in C$)

2.6 Beispiele guter linear Codes

2.6.1 Hamming-Codes

Sei q ein Primzahlpotenz, K Körper mit $|K| = q$
Sei $l \in \mathbb{N}$, $l \geq 2$. $n = \frac{q^l-1}{q-1}$, $k = n - l$
Denn ex. perfekter $[n, k]$-Code C über K, $d(C) = 3$. Hamming-Code.

Konstruktion

$\left|K^l \setminus \left\{\vec{0}\right\}\right| = q^l - 1$, je $q - 1$ von 0 versch. Vektoren erzeugen den gleichen 1-dim. Unterräume in K^l, d.h.

$$n = \frac{q^l - 1}{q - 1} \quad \text{1-dim Unterraum}$$

Bilde $l \times n$-Matrix H: Wähle aus jedem der 1-dim. Unterräume von K^l einen Vektor $\neq 0$ aus und schreibe ihn als Spalte in H.

$$C = \{x \in K^n : Hx^t = 0\}$$

$rg(H) = l$, denn H enthält l lin. unabhängige Spalten.
$dim(C) = n - l = k$, $|C| = q^k$
$d(C) = 3$
Nach Konstruktion von H sind je zwei Spalten linear unabhänging. Es gibt drei linear abhängige Spalten:

$$\begin{pmatrix} a \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ b \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} c \\ c \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad a, b, c \neq 0$$

$$\frac{c}{a} \begin{pmatrix} a \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \frac{c}{b} \begin{pmatrix} 0 \\ b \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c \\ c \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Kugelpackungsbed.:

$$\sum_{j=0}^{1} \binom{n}{j} (q-1)^i = 1 + n \cdot (q-1) = 1 + \frac{q^l - 1}{q - 1} = q^l$$

$$\frac{q^n}{q^l} = q^{n-l} = q^k = |C|$$

C perfekt.

2.6.2 Beispiel

a) $q = 2(K = \mathbb{Z}_2), l = 3, n = 7$

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

[7, 4]-Hamming-Code über $\mathbb{Z}_2$

b) $q = 3(K = \mathbb{Z}_3), l = 3, n = \frac{3^3-1}{3-1} = 13$

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 1 & 2 & 0 & 0 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 2 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \end{pmatrix}$$

$dim(C) = 10$ [13, 10]-H.C. über $\mathbb{Z}_3$ $|C| = 3^{10} = 59049$

c) $q = 2(K = \mathbb{Z}_2), l = 4, n = \frac{2^4.1}{2-1} = 15$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$dim(C) = 11$, $[15, 11]$-Code über Z_2, $|C| = 2^{11} = 2048$
Ang. $y = (110101110000110)$ empfangen.
Ist $y \in C$?

$$H \cdot y^t = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow y \notin C$$

Decodierung : $y \rightarrow x \in C \quad (d(x, y) = 1)$
y unterscheidet sich von x an genau einer Stelle i, statt y_i steht in $x : y_i + 1$.

$$\begin{aligned} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} &= H \cdot y^t \\ &= H \cdot (x^t + (0, \ldots, \underset{\substack{\uparrow \\ i}}{1}, 0, \ldots, 0)^t) \\ &= \underbrace{H \cdot x^t}_{\vec{0}} + H \cdot (0, \ldots, 0, \underset{\substack{\uparrow \\ i}}{1}, 0, \ldots, 0)^t \\ &= \begin{pmatrix} h_{1i} \\ h_{2i} \\ h_{3i} \\ h_{4i} \end{pmatrix} \Rightarrow i = 2 \end{aligned}$$

$$x = (100101110000110) \in C$$

2.6.3 Decodierung binärer Hamming-Codes

y empfangen. Ist $Hy^t = 0$ so ist $y \in C$
Ist $Hy^t \neq 0$, so wähle in H diejenige Spalte, die mit Hy^t übereinstimmt. Ist dies die $i - te$ Spalte von H, so wird y an Stelle i geändert $\curvearrowright x \in C$.

2.7 Reed-Solomon-Codes

K Körper, $|K| = q,\ 1 \leq k \leq n \leq q$
$K[x]_{k-1} = \{f \in K[x] : grad(f) \leq k - 1\}$
K-Vektorraum der Dimension k
Wähle $M = \{a_1, \ldots, a_n\} \subseteq K, \quad a_i \neq a_j$, für $i \neq j$ (geht, da $n \leq q$)
$C = C_M = \{(f(a_i), \ldots, f(a_n) : f \in K[x]_{k-1}\}$

2.7.1 Reed-Solomon-Code (Auswertungscode)

(Besonderes wichtiger Fall: $n = q - 1$
$a_i = a^i$, wobei a ein Erzeuger von K^*, d.h. $K^* =< a >= \{a^1, a^2, \ldots, a^{q-1}\}$)

(1) $dim(C) = k$:

$$\alpha \begin{cases} K[x]_{k-1} & \longrightarrow C \\ f & \longmapsto (f(a_1), \ldots, f(a_n)) \end{cases}$$

α ist K-lineare Abbildung, surjektiv.
$f \in Kern(\alpha)$, d.h. $(f(a_1), \ldots, f(a_n)) = (0, \ldots, 0)$

$$f = 0 \Leftarrow \begin{cases} f \text{ hat mind. } n \text{ Nullstellen} \\ Grad(f) = k - 1 < n \end{cases}$$

$Kern(\alpha) = \{\vec{0}\}$, α bijektiv. $dim(C) = dim(K[x]_{k-1}) = k$

(2) $d(C) = n - k + 1$:
Jedes $f \neq 0$ in $K[x]_{k-1}$ hat höchstens $k - 1$ Nullstellen, d.h., mind. $n - (k - 1)$ Einträge jedes Codeworts $\neq 0$ sind $\neq 0$, $d(C) = wt(C) \geq n - k + 1$
$f \prod_{i=1}^{k-1} x - a_i \in K[x]_{k-1},\ f(a_1) = \ldots = f(a_{k-1}) = 0$
$f \longmapsto c = (\underset{\leftarrow k-1 \rightarrow}{0, \ldots, 0}, \underbrace{*, \ldots, *}_{\neq 0}) \quad wt(c) = n - k + 1$
$d(C) = n - k + 1$

(3) Reed-Solomon-Codes sind MDS-Codes (für $q = 2$ nur triviale Codes)

(4)

$$G = \begin{pmatrix} 1 & \ldots & 1 \\ a_1 & \ldots & a_n \\ a_1^2 & \ldots & a_n^2 \\ & \vdots & \\ a_1^{k-1} & \ldots & a_n^{k-1} \end{pmatrix} \text{ ist Erzeugermatrix von } C$$

$K[x]_{k-1} \cong C$
Basis: $1, x, x^2, \ldots, x^{k-1} \underset{\alpha}{\longrightarrow} \underbrace{(1, \ldots, 1), (a_1, \ldots, a_n), \ldots, (a_1^{k-1}, \ldots, a_n^{k-1})}_{\text{Basis von } C}$

2.8 MDS-Codes

(1)

$$\tilde{G} = \begin{pmatrix} 1 & \ldots & 1 & 0 \\ a_1 & \ldots & a_n & 0 \\ a_1^2 & \ldots & a_n^2 & 0 \\ & \vdots & & \vdots \\ a_1^{k-1} & \ldots & a_n^{k-1} & 1 \end{pmatrix}$$

$|K| = q$
$\tilde{C}$ der Code mit Erzeugermatrix $\tilde{G}$.
Länge: $n+1$, Dim. k
$d(\tilde{G}) = n - k + 2$
$d(\tilde{G}) \leq (n+1) - k + 1 = n - k + 2$
$0 \neq \tilde{x} \in \tilde{C}$
Falls $\tilde{x}$ Linearkombination des ersten $k-1$ Zeilen von $\tilde{G}$, so

$$\tilde{x} = (x, 0),\ \ x \in C$$

$x = (f(a_1), \ldots, f(a_n))$, wobei f Polynom von Grad $\leq k-2$

$$wt(\tilde{x}) = wt(x) \geq n - (k-2) = n - k + 2$$

Falls die letzte Zeile von $\tilde{G}$ in der Linearkombination von $\tilde{x}$ vorkommt, so

$$\tilde{x} = (x, a),\ \ x \in K,\ \ a \neq 0,\ \ x \in C.$$
$$wt(\tilde{x}) = wt(x) + 1 \geq n - k + 1 + 1 = n - k + 2$$

$$d(\tilde{C}) = n - k + 2 = (n+1) - k + 1 \text{ MDS-Code.}$$

Wähle $n = q$. $\tilde{C}$ MDS-Code der Länge $q+1$

(2)
$$\overline{G} = \begin{pmatrix} 1 & \ldots & 1 & 0 & 0 \\ a_1 & \ldots & a_n & 1 & 0 \\ a_1^2 & \ldots & a_n^2 & 0 & 1 \end{pmatrix}$$

Code $\overline{C}$ der Länge $q+2$ Dimension 3.
$d(\overline{C}) = q = (q+2) - 3 + 1$ MDS-Code, falls $q = 2^l$

(3) MDS-Vermutung
Falls C $[n, k]$-Code über K, $|K| = q$, mit $d(C) = n - k + 1$ (MDS-Code), so

$$n \leq q + 2 \text{ (und falls } q = 2^l \text{ sogar } n \leq q + 1)$$

(4) Ist C $[n, k]$-MDS-Code, $n \geq 2$, $d = d(C) = n - k + 1$
$\check{C} = \{(C_1, \ldots, C_{n-1} : (C_1, \ldots, C_{n-1}, 0) \in C\}$ (Verkürzung von C)
ist $[n-1, k-1, d]$-Code

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

$\check{C}$ MDS-Code.

2.9 Codierung von Audio-CDs

Auslöschung: Bit kann nicht mehr gelesen werden. (Man weiß wo der Fehler liegt)
Bündelfehlern:

2.9.1 Lemma: $[n, k, d]$-Code

kann Bündelfehler bis zur Länge $d - 1$ korrigieren (pro Codewort). (Ebenso: $d - 1$ Auslöschungen)
Bew: c gesendet, y empfangen $(c_1 \dots c_i \underbrace{|* \dots *|}_{a \le d-1} c_j \dots c_n)$ Sei $c' \in C$, das um Stelle $1, \dots, i, j, \dots, n$ mit y übereinstimmen. $d(c, c') \le a \le d - 1$

2.9.2 Interleaving (Spreizung)

Def: Sei C ein $[n, k] - Code$.
Interleaving von C zur Tiefe s liefert lin. Code $C(s) \subseteq K^{n \cdot s}$ auf folgende Art:
Wähle s Codewörter von C, schreibe diese als Spalten einer Matrix:

$$c_1 = (c_{11}, \dots, c_{n1}, \dots, c_s = (c_{1s}, \dots, c_{ns}) \begin{pmatrix} c_{11} & \dots & c_{1s} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{ns} \end{pmatrix}$$

s Codewörter aus $C \rightarrow \boxed{\text{Interleaver}} \rightarrow$ Codewörter der Länge ns.
$C(s)$ $[ns, ks]$-Code, $d(C(s)) = d(C)$
In $C(s)$ können Bündelfehler der Länge $b \le (d - 1) \cdot s$ korrigieren.

$$\begin{pmatrix} c_{11} & \dots & c_{1s} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{ns} \end{pmatrix}$$

Bündelfehler der Länge $b \le (d - 1) \cdot s$ bewirkt Bündelfehler der Länge $\le d - 1$ zu jedem der ursprünglichen Codewörter aus C. Korrigierbar nach Lemma.

2.9.3 Cross-Interleaving

Def: C_1, C_2 $[n_i, k_i]$-Code $(i = 1, 2)$ über K, C_1 besitzt Erzeugermatrix in Standardform $(E_{k_1}|A)$. Das Cross-Interleaving des inneren Codes C_1 mit äußern Code C_2 ist ein Code $C \subseteq K^{n_1 n_2}$, der so entsteht:
Wähle k_1 Codewörter $c_1, \dots, c_{k_1}$ aus $C_2 \subseteq K^{n_2}$ $c_i = (c_{1i}, \dots, c_{n_2 i}), i = 1, \dots, k_1$, und schreibe sie spaltenweise in eine Matrix:

$$\begin{pmatrix} c_{11} & \dots & c_{1k_1} \\ \vdots & & \vdots \\ c_{n_2 1} & \dots & c_{n_2 k_1} \end{pmatrix}$$

Codiere jede Zeile mit $(E_{k_1}|A)$ zu Codewörtern zu C_1. $C \cdot (E_{k_1}|A)$: Letzte $n_1 - k_1$ Spalten dieser Matrix sind L.K., der ersten k_1 Spalten. Lese zeilenweise aus!
Liefert $[n_1 \cdot n_2, k_1 \cdot k_2, d_1 \cdot d_2]$-Code.
k_1 Codewörter aus $C_1 \rightarrow \boxed{\text{Cross-Interleaver}} \rightarrow$ Codewörter aus C der Länge $n_1 \cdot n_2$

(n_2 hintereinander folgende Codewörter aus C_1). C kann Bündelfehler der Länge $(d2 - 1) \cdot n_1$ korrigieren.
Das lässt sich durch sogenanntes Cross-Interleaving mit t-facher Verzögerung mit Burst der Länge $(d_2 - 1) \cdot n_1 \cdot t$ erhöhen.

2.9.4 Audio-CD

Ausgangscode Reed-Solomon-Code $R = [255, 251, 5]$ über $\mathbb{F}_{2^8}$
C_1 $[32, 28, 5]$-Code 223-fache Verkürzung von R.
C_2 $[28, 24, 5]$-Code 227-fache Verkürzung von R.
Cross-Interleaving mit C_1, C_2 mit 4-facher Verzögerung. (CIRC = Cross Interleave Reed-Solomon Code)
Kann Bursts bis zur Länge $4 \cdot 32 \cdot 4 = 512$