Decodierung von Reed–Solomon–Codes

Reinhold Hübl

1 Verallgemeinerte Reed–Solomon–Codes

Wir betrachten einen endlichen Körper $\mathbb{F}_q$ mit $q = p^e$ Elementen und n paarweise verschiedene Punkte $u_1, \ldots, u_n \in \mathbb{F}_q$ sowie Koeffizienten $b_1, \ldots, b_n \in E(\mathbb{F}_q)$ $(= \mathbb{F}_q \setminus \{0\})$. Wir setzen

$$\mathbf{u} = (u_1, \ldots, u_n), \quad \mathbf{b} = (b_1, \ldots, b_n)$$

Definition 1.1. Der verallgemeinerte $[n, k]$–Reed–Solomon–Code $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ zu $\mathbf{u}$ und $\mathbf{b}$ ist der lineare $[n, k]$–Code mit Erzeugermatrix

$$G = \begin{pmatrix} b_1 & b_2 & \ldots & b_n \\ b_1 \cdot u_1 & b_2 \cdot u_2 & \ldots & b_n \cdot u_n \\ b_1 \cdot u_1^2 & b_2 \cdot u_2^2 & \ldots & b_n \cdot u_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ b_1 \cdot u_1^{k-1} & b_2 \cdot u_2^{k-1} & \ldots & b_n \cdot u_n^{k-1} \end{pmatrix}$$

Nach Matzat, (3.11) gibt es $d_1, \ldots, d_n \in E(\mathbb{F}_q)$, sodass

$$\mathrm{RS}_k(\mathbf{u}, \mathbf{b}) = \mathrm{RS}_{n-k}(\mathbf{u}, \mathbf{d})^{\perp}$$

Damit hat $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ eine Paritätsprüfmatrix H der Form

$$H = \begin{pmatrix} d_1 & d_2 & \ldots & d_n \\ d_1 \cdot u_1 & d_2 \cdot u_2 & \ldots & d_n \cdot u_n \\ d_1 \cdot u_1^2 & d_2 \cdot u_2^2 & \ldots & d_n \cdot u_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ d_1 \cdot u_1^{n-k-1} & d_2 \cdot u_2^{n-k-1} & \ldots & d_n \cdot u_n^{n-k-1} \end{pmatrix}$$

(denn das ist die Erzeugermatrix von $\mathrm{RS}_{n-k}(\mathbf{u}, \mathbf{d})$).

2 Syndrome

Wir betrachten einen verallgemeinerten Reed–Solomon–Code $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ und dazu ein $\mathbf{d}$ mit

$$\mathrm{RS}_k(\mathbf{u}, \mathbf{b}) = \mathrm{RS}_{n-k}(\mathbf{u}, \mathbf{d})^{\perp}$$

sodass also $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ die Paritätsprüfmatrix

$$H = \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ d_1 \cdot u_1 & d_2 \cdot u_2 & \dots & d_n \cdot u_n \\ d_1 \cdot u_1^2 & d_2 \cdot u_2^2 & \dots & d_n \cdot u_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ d_1 \cdot u_1^{n-k-1} & d_2 \cdot u_2^{n-k-1} & \dots & d_n \cdot u_n^{n-k-1} \end{pmatrix}$$

hat. Ferner betrachten wir ein $a = (a_1, \dots, a_n) \in \mathbb{F}_q^n$.

Definition 2.1. Für $0 \leq l \leq n-k-1$ heißt

$$[a, X^l] = \sum_{i=1}^{n} a_i \cdot d_i \cdot b_i^l$$

das l–**te Syndrom** von a.

Bemerkung 2.1. Es ist $[a, X^l]$ die $(l+1)$–te Komponente von $H \cdot \overrightarrow{a}$. Speziell gilt also

$$a \in C \iff [a, X^l] = 0 \quad \text{für alle } l \in \{0, \dots, n-k-1\}$$

Ist nun a ein n–Tupel, dass aus einem Codewort c durch möglicherweise fehlerhafte Übertragung entstanden ist, so schreiben wir

$$a = c + e = (c_1, \dots, c_n) + (e_1, \dots, e_n)$$

mit einem Fehlerterm $e = (e_1, \dots, e_n)$.

Bemerkung 2.2. Für alle $0 \leq l \leq n-k-1$ gilt

$$[a, X^l] = [c, X^l] + [e, X^l] = [e, X^l]$$

(denn $[c, X^l] = 0$ nach Bemerkung 2.1).

Der Begriff des Syndroms wird nun wie folgt verallgemeinert:

Für ein beliebiges Polynom $f(X) \in \mathbb{F}_q[X]$ vom Grad höchstens $n-k-1$ heißt

$$[a, f] = \sum_{i=1}^{n} a_i \cdot d_i \cdot f(b_i)$$

das Syndrom von a bezüglich $f(X)$.

Ist speziell $f(X) = X^l$ so gilt

$$[a, f] = \sum_{i=1}^{n} a_i \cdot d_i \cdot f(b_i) = \sum_{i=1}^{n} a_i \cdot d_i \cdot b_i^l = [a, X^l]$$

es handelt sich also tatsächlich um eine Verallgemeinerung des Syndrombegriffs.

Für zwei Polynome $f(X), g(X) \in \mathbb{F}_q[X]$ vom Grad höchstens $n-k-1$ gilt dann

$$\begin{aligned} [a, f+g] &= \sum_{i=1}^{n} a_i \cdot d_i \cdot (f+g)(b_i) \\ &= \sum_{i=1}^{n} a_i \cdot d_i \cdot (f(b_i) + g(b_i)) \\ &= \sum_{i=1}^{n} a_i \cdot d_i \cdot f(b_i) + \sum_{i=1}^{n} a_i \cdot d_i \cdot g(b_i) \\ &= [a, f] + [a, g] \end{aligned}$$

die Syndrombildung ist also linear in der zweiten Komponente (und genauso in der ersten). Wenden wir das wiederholt auf $f(X) = \sum_{l=0}^{n-k-1} r_l \cdot X^l$ an, so erhalten wir

$$[a, f] = \sum_{l=0}^{n-k-1} r_l \cdot [a, X^l]$$

und daraus folgt

Bemerkung 2.3. Es gilt

$a \in C \iff [a, f] = 0$ für jedes $f(X) \in \mathbb{F}_q[X]$ vom Grad höchstens $n-k-1$

sowie

Bemerkung 2.4. Ist $a = c + e$ mit einem Codewort c und einem Fehlerterm e, und ist $f(X)$ ein Polynom vom Grad höchstens $n-k-1$, so gilt

$$[a, f] = [c, f] + [e, f] = [e, f]$$

3 Decodierung von Reed–Solomon–Codes

Wir betrachten einen verallgemeinerten $[n, k]$–Reed–Solomon–Code $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ und wählen ein $\mathbf{d}$ so, dass

$$\mathrm{RS}_k(\mathbf{u}, \mathbf{b}) = \mathrm{RS}_{n-k}(\mathbf{u}, \mathbf{d})^{\perp}$$

sodass also

$$H = \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ d_1 \cdot u_1 & d_2 \cdot u_2 & \dots & d_n \cdot u_n \\ d_1 \cdot u_1^2 & d_2 \cdot u_2^2 & \dots & d_n \cdot u_n^2 \\ \vdots & & \ddots & \vdots \\ d_1 \cdot u_1^{n-k-1} & d_2 \cdot u_2^{n-k-1} & \dots & d_n \cdot u_n^{n-k-1} \end{pmatrix}$$

eine Paritätsprüfmatrix von $\mathrm{RS}_k(\mathbf{u}, \mathbf{b})$ ist.

Ferner betrachten wir ein empfangenes Wort $a = (a_1, \dots, a_n)$ und schreiben

$$a = c + e = (c_1, \dots, c_n) + (e_1, \dots, e_n)$$

mit dem tatsächlich gesendeten Codewort c und einem durch die Übertragung verursachten Fehlerterm $e = (e_1, \dots, e_n)$. Wir nehmen jetzt an, dass die Anzahl der Fehler bei der Übertragung die Fehlerkorrekturschranke

$$t = \lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{n-k}{2} \rfloor$$

nicht übersteigt (dass also höchstens t Komponenten von e von Null verschieden sind). Wie in Abschnitt 2 bilden wir die Syndrome $[a, X^l]$ für alle $l = 0, \dots, n-k-1$ und damit das Gleichungssystem

$$\begin{array}{rrrl} [a, X^0] \cdot Y_0 + & [a, X^1] \cdot Y_1 + \dots + & [a, X^t] \cdot Y_t & = 0 \\ [a, X^1] \cdot Y_0 + & [a, X^2] \cdot Y_1 + \dots + & [a, X^{t+1}] \cdot Y_t & = 0 \\ \vdots & & & \vdots \\ [a, X^{n-k-t-1}] \cdot Y_0 + & [a, X^{n-k-t}] \cdot Y_1 + \dots + & [a, X^{n-k-1}] \cdot Y_t & = 0 \end{array} \tag{1}$$

mit $n-k-t$ Gleichungen und $t+1$ Unbekannten $Y_0, Y_1, \dots, Y_t$.

Eine Lösung $(\eta_0, \dots, \eta_t)$ dieses Gleichungssystems erfüllt also die Gleichungen

$$\sum_{l=0}^{t} \eta_l \cdot [a, X^{j+l}] = 0 \tag{2}$$

für alle $j = 0, \dots, n-k-t-1$.

Lemma 3.1. *Das Gleichungssystem* (1) *hat immer eine nicht–triviale Lösung* $\eta = (\eta_0, \ldots, \eta_t)$. *Setzen wir*

$$L(X) = \eta_0 + \eta_1 \cdot X + \eta_2 \cdot X^2 + \cdots + \eta_t \cdot X^t$$

und ist $r \in \{1, \ldots, n\}$ *mit* $e_r \neq 0$ *(ist also* r *ein Fehlerstelle von* a*), so gilt*

$$L(b_r) = 0$$

Beweis: Mit τ bezeichnen wir die Anzahl der Fehler, die bei der Übertragung von a aufgetreten sind. Wir wissen, dass $\tau \leq t$. Zur Vereinfachung der Notation nehmen wir an, dass die Fehler an den Stellen $1, \ldots, \tau$ aufgetreten sind.

Da $\tau \leq t$, gibt es ein (nicht–triviales) Polynom $L(X)$ vom Grad höchstens t mit $L(b_i) = 0$ für alle $i \in \{1, \ldots, \tau\}$ (dh. ein fehlerlokalisierendes Polynom vom Grad höchstens t existiert auf jeden Fall, etwa $L(X) = \prod_{i=1}^{\tau}(X - b_i)$; falls kein Fehler aufgetreten ist, so können wir $L(X) = 1$ nehmen). Schreiben wir $f(X) = \sum_{l=0}^{t} r_l \cdot X^l$, so gilt für $j = 0, \ldots, n-k-t-1$ aufgrund der Bemerkungen 2.3 und 2.4:

$$\begin{aligned}
\sum_{l=0}^{t} r_l \cdot [a, X^{l+j}] &= [a, \sum_{l=0}^{t} r_l \cdot X^{l+j}] \\
&= [a, f(X) \cdot X^j] \\
&= [e, f(X) \cdot X^j] \\
&= \sum_{i=1}^{\tau} e_i \cdot d_i \cdot f(b_i) \cdot b_i^j \\
&= 0
\end{aligned}$$

(denn $f(b_i) = 0$ für jede Fehlerstelle b_i, also für alle i mit $e_i \neq 0$) , wobei wir auch ausgenutzt haben, dass $\deg(f(X) \cdot X^j) \leq n-k-1$ für jedes $j \in \{0, \ldots, n-k-t-1\}$. Das bedeutet aber gerade, dass $(r_0, r_1, \ldots, r_n)$ Gleichung 2 erfüllt, dass also $(r_0, r_1, \ldots, r_t)$ eine nichttriviale Lösung des Gleichungssystems (1) ist.

Ist nun umgekehrt eine nicht–triviale Lösung $(\eta_0, \ldots, \eta_t)$ des Gleichungssystems (1) gegeben und ist

$$L(X) = \eta_0 + \eta_1 \cdot X + \eta_2 \cdot X^2 + \cdots + \eta_t \cdot X^t$$

so ist noch zu zeigen, dass $L(b_i) = 0$ für $i = 1, \ldots, \tau$. Dazu nehmen wir an, dass das nicht der Fall ist und mindestens eine der Fehlerstellen b_i keine Nullstelle von $L(X)$ ist. Zur Vereinfachung der Notation können wir annehmen, dass b_1 eine solche Nichtnullstelle ist, dass also $L(b_1) \neq 0$.

Da

$$n - k - 1 - t = n - k - 1 - \lfloor \frac{n-k}{2} \rfloor \geq \lfloor \frac{n-k}{2} \rfloor - 1 = t - 1 \geq \tau - 1$$

gibt es ein Polynom $f(X)$ vom Grad höchstens $n - k - t - 1$ mit

$$f(b_1) \neq 0,\ f(b_2) = \cdots = f(b_\tau) = 0$$

(dh. $f(X)$ verschwindet an allen Fehlerstellen außer an b_1, hierfür können wir etwa $f(X) = \prod_{i=2}^{\tau}(X - b_i)$ nehmen; falls $\tau = 1$, nehmen wir $f(X) = 1$). Dann gilt

$$\deg(f(X) \cdot L(X)) \leq n - k - 1$$

und (wegen Bemerkung 2.4)

$$\begin{aligned} [a, f(X) \cdot L(X)] &= [e, f(X) \cdot L(X)] \\ &= \sum_{i=1}^{\tau} e_i \cdot f(b_i) \cdot L(b_i) \\ &= e_1 \cdot f(b_1) \cdot L(b_1) \\ &\neq 0 \end{aligned} \tag{3}$$

Beachten Sie dabei, dass wir ausgenutzt haben, dass $e_{\tau+1} = \cdots = e_n = 0$ und $f(b_2) = \cdots = f(b_\tau) = 0$.

Schreiben wir $f(X) = \sum_{j=0}^{n-k-t-1} r_j \cdot X^j$, so gilt

$$f(X) \cdot L(X) = \sum_{j=0}^{n-k-t-1} r_j \cdot \sum_{l=0}^{t} \eta_l \cdot X^{j+l}$$

und damit aufgrund der Linearität das Syndroms in der zweiten Komponente

$$[a, f(X) \cdot L(X)] = \sum_{j=0}^{n-k-t-1} r_j \cdot \sum_{l=0}^{t} \eta_l \cdot [a, X^{j+l}] = 0 \tag{4}$$

denn $\sum_{l=0}^{t} \eta_l \cdot [a, X^{j+l}] = 0$ für alle $j = 0, \ldots, n - k - t - 1$, denn $(\eta_0, \ldots, \eta_t$ ist eine Lösung des Gleichungssystems (1).

Gleichung (4) steht aber im Widerspruch zu Gleichung (3), und damit war unsere Annahme falsch. Also ist $L(X)$ ein fehlerlokalisierendes Polynom.

Zur Fehlerkorrektur von a reicht es, einen Fehlertupel $e = (e_1, \ldots, e_n)$ zu finden, das nur an maximal t Stellen von 0 verschieden ist und für das $c = a - e$ ein Codewort ist (wir gehen immer noch davon aus, dass bei der Übertragung von a höchstens t Fehler aufgetreten sind). Dazu reicht es, ein Tupel $e = (e_1, \ldots, e_n)$ zu finden, das nur an maximal t Stellen von 0 verschieden ist und für das

$$[e, X^l] = [a, X^l] \quad \text{für alle } l \in \{0, \ldots, n-k-1\}$$

Dann gilt nämlich für $c = a - e$:

$$[c, X^l] = [a, X^l] - [e, X^l] = 0 \quad \text{für alle } l \in \{0, \ldots, n-k-1\}$$

und daher ist c ein Codewort nach Bemerkung 2.1.

Mithilfe des fehlerlokalisierenden Polynoms $L(X)$ aus Lemma 3.1 können zunächst alle Fehlerstellen gefunden werden. Dazu setzen wir

$$N(L) = \{i = \{1, \ldots, n\} | \, L(u_i) = 0\}$$

Dann enthält $N(L)$ alle Fehlerstellen von a. Wir schreiben

$$N(L) = \{i_1, \ldots, i_\tau\}$$

(für ein $\tau \leq t$) und betrachten das lineare Gleichungssystem

$$\begin{array}{rcl}
d_{i_1} \cdot E_{i_1} + & d_{i_2} \cdot E_{i_2} \; + \ldots + & d_{i_\tau} \cdot E_{i_\tau} \; = [a, X^0] \\
d_{i_1} \cdot u_{i_1} \cdot E_{i_1} + & d_{i_2} \cdot u_{i_2} \cdot E_{i_2} \; + \ldots + & d_{i_\tau} \cdot u_{i_\tau} \cdot E_{i_\tau} \; = [a, X^1] \\
\vdots & & \vdots \\
d_{i_1} \cdot u_{i_1}^{n-k-1} \cdot E_{i_1} + & d_{i_2} \cdot u_{i_2}^{n-k-1} \cdot E_{i_2} \; + \ldots + & d_{i_\tau} \cdot u_{i_\tau}^{n-k-1} \cdot E_{i_\tau} \; = [a, X^{n-k-1}]
\end{array} \tag{5}$$

mit $n - k$ Gleichungen in den Unbekannten $E_{i_1}, \ldots, E_{i_\tau}$.

Lemma 3.2. *Das Gleichungssystem* (5) *hat eine eindeutige Lösung* $(e_{i_1}, \ldots, e_{i_\tau})$. *Setzen wir* $e_i = 0$ *für* $i \in \{1, \ldots, n\} \setminus N(L)$, *so ist*

$$e = (e_1, \ldots, e_n)$$

der Fehlerterm von a *und*

$$c = a - e$$

ist das Codewort c, *das zu* a *gehört.*

Beweis: Da wir annehmen, dass bei der Übertragung von a höchstens t Fehler aufgetreten sind, und da wir aus Lemma 3.1 wissen, dass diese Fehlerstellen in den Nullstellen von $L(X)$ enthalten sind, können wir schreiben

$$a = c + e = (c_1, \ldots, c_n) + (e_1, \ldots, e_n)$$

wobei e höchstens t von Null verschiedene Einträge hat und dabei auch noch gilt

$$e_i \neq 0 \Longrightarrow i \in N(L)$$

Da nach Bemerkung 2.2 $[a, X^l] = [e, X^l]$ für $l = 0, \ldots, n-k-1$ gilt, erhalten wir

$$[a, X^l] = [e, X^l] = \sum_{i=1}^{n} d_i \cdot e_i \cdot u_i^l = \sum_{j=1}^{\tau} d_{i_j} \cdot u_{i_j}^l \cdot e_{i_j}$$

für alle $l = 0, \ldots, n-k-1$, und das bedeutet gerade, dass e (bzw. der Anteil $(e_{i_1}, \ldots, e_{i_\tau})$) eine Lösung von Gleichungssystem (5) ist.

Damit hat das Gleichungssystem auch schon mindestens eine Lösung, und es bleibt zu zeigen, dass diese Lösung eindeutig ist. Dazu beachten wir, dass wir in diesem Gleichungssystem $\tau \leq t < n - k$ Unbekannte haben, also weniger Unbekannte als Gleichungen. Wenn wir nur die ersten τ Gleichungen betrachten, so haben diese die Koeffizientenmatrix

$$A = \begin{pmatrix} d_{i_1} & d_{i_2} & \ldots & d_{i_\tau} \\ d_{i_1} \cdot u_{i_1} & d_{i_2} \cdot u_{i_2} & \ldots & d_{i_\tau} \cdot u_{i_\tau} \\ \vdots & & \ddots & \vdots \\ d_{i_1} \cdot u_{i_1}^{\tau-1} & d_{i_2} \cdot u_{i_2}^{\tau-1} & \ldots & d_{i_\tau} \cdot u_{i_\tau}^{\tau-1} \end{pmatrix}$$

Dann gilt nach den Regeln zum Rechnen mit Determinanten (und der Formel zur Bestimmung der Vandermonde–Determinante)

$$\begin{aligned} \det(A) &= d_{i_1} \cdots d_{i_\tau} \cdot \det \begin{pmatrix} 1 & \ldots & 1 \\ u_{i_1} & \ldots & u_{i_\tau} \\ \vdots & \ddots & \vdots \\ u_{i_1}^{\tau-1} & \ldots & u_{i_\tau}^{\tau-1} \end{pmatrix} \\ &= d_{i_1} \cdots d_{i_\tau} \cdot \prod_{1 \leq r < s \leq \tau} (u_{i_s} - u_{i_r}) \\ &\neq 0 \end{aligned}$$

(da alle $d_i \neq 0$ und die u_j paarweise verschieden sind).

Damit gibt es also für die ersten τ Gleichungen eine eindeutige Lösung und daher für das ganze Gleichungssystem höchstens eine Lösung.

Da schon gezeigt wurde, dass der Fehlerterm $e = (e_1, \ldots, e_n)$ eine Lösung von Gleichungssystem (5) ist, berechnet sich der Fehlerterm als eindeutige Lösung dieses Gleichungssystems.

Bemerkung 3.1. Bei Lemma 3.1 und Lemma 3.2 wurde jeweils vorausgesetzt, dass bei der Übertragung von a höchstens $t = \lfloor \frac{n-k}{2} \rfloor$ Fehler aufgetreten sind. Daraus ergibt sich

a) War die Übertragung von a fehlerhaft, hat aber das durch Gleichungssystem (1) bestimmte fehlerlokalisiernde Polynom keine Nullstellen unter den Punkten $u_1, \ldots, u_n$, so sind bei der Übertragung von a mehr als t Fehler aufgetreten.

b) War die Übertragung von a fehlerhaft und hat das durch Gleichungssystem (1) bestimmte fehlerlokalisiernde Polynom Nullstellen in $u_1, \ldots, u_n$ (wurden also potentielle Fehlerstellen gefunden), hat aber das daraus abgeleitete Gleichungssystem (5) keine Lösung, so sind bei der Übertragung von a mehr als t Fehler aufgetreten.