Kódování

Cílem kódování je přenášet informace tak, abychom postřehli, pokud by došlo k jejich zkreslení náhodnou chybou, v ideálním případě dokonce takto vzniklou chybu nejen odhalit, ale i opravit.

Přenášená informace bude zapsána pomocí abecedy, která má p symbolů (tj. jakýchsi písmen), kde p je pevně zvolené prvočíslo. Tuto abecedu tedy můžeme ztotožnit s množinou $\mathbb{Z}_p$ všech zbytkových tříd modulo p. Přenášet budeme slova délky n, každé takové kódové slovo $a_1 a_2 a_3 \dots a_{n-1} a_n$ lze tedy chápat jako polynom

$$a(x) = a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n \in \mathbb{Z}_p[x]$$

stupně $\mathrm{st}(a) < n$. Číslo n nazýváme délka kódu.

Kdyby každý polynom stupně menšího než n bylo některé z kódových slov, tak bychom nemohli postřehnout, že při přenosu došlo k nějaké náhodné chybě.

Polynomiální kód délky n daný polynomem $g(x) \in \mathbb{Z}_p[x]$

Máme tedy pevně zvolené prvočíslo p a přirozené číslo n.

Zafixujme také polynom $g(x) \in \mathbb{Z}_p[x]$ stupně $\mathrm{st}(g) = k < n$.

Kódová slova budou ty polynomy $a(x) \in \mathbb{Z}_p[x]$ stupně $\mathrm{st}(a) < n$, které jsou dělitelné polynomem $g(x)$ v okruhu $\mathbb{Z}_p[x]$, jsou to tedy polynomy $a(x) = g(x) \cdot h(x)$, kde $h(x) \in \mathbb{Z}_p[x]$ je libovolný polynom stupně $\mathrm{st}(h) < n - k$.

Pokud chceme odeslat zprávu zapsanou pomocí $n - k$ symbolů, tj. polynom $b(x) = b_1 x^{n-k-1} + \cdots + b_{n-k-1}x + b_{n-k} \in \mathbb{Z}_p[x]$ stupně $\mathrm{st}(b) < n - k$, vydělíme polynom $x^k \cdot b(x)$ polynomem $g(x)$ se zbytkem a dostaneme polynomy $q(x)$, $r(x) \in \mathbb{Z}_p[x]$ tak, že $x^k \cdot b(x) = g(x) \cdot q(x) + r(x)$, kde $\mathrm{st}(r) < k$.

Odešleme pak polynom $g(x) \cdot q(x) = x^k \cdot b(x) - r(x)$.

Každé kódové slovo se tedy skládá z $n - k$ významových symbolů (daných polynomem $b(x)$) následovaných k kontrolními symboly (daných polynomem $-r(x)$). Je však nutné vhodně zvolit polynom $g(x)$. Určitě by nebyla vhodná volba $g(x) = x^k$, protože pak bychom každou zprávu $b(x)$ doplnili nulovým polynomem.

Hammingova vzdálenost kódových slov

Pro každé dva polynomy $a(x)$, $b(x) \in \mathbb{Z}_p[x]$ stupnů $\mathrm{st}(a) < n$, $\mathrm{st}(b) < n$, definujeme jejich vzdálenost jako počet nenulových koeficientů rozdílu $a(x) - b(x)$, tj. počet koeficientů, v nichž se oba polynomy liší. Uvědomte si, že jde o metrický prostor.

Máme-li být schopni postřehnout, že došlo k chybě, pokud při přenosu byl právě na jedné pozici přenášený symbol náhodné změněn, je třeba, aby vzdálenost libovolných dvou různých kódových slov byla alespoň 2. Pokud bude alespoň 3, budeme dokonce schopni takovou chybu i opravit.

Obecněji, je-li pro nějaké $t \in \mathbb{N}$ vzdálenost libovolných dvou různých kódových slov alespoň $t + 1$, pak lze chybu detekovat, když došlo při přenosu ke změně na nejvýše t pozicích. Je-li tato vzdálenost alespoň $2t + 1$, pak takovou chybu lze dokonce i opravit.

Protože u polynomiálního kódu je rozdíl libovolných dvou kódových slov opět kódové slovo, lze místo o nejmenší vzdálenosti dvou různých kódových slov hovořit o nejmenší vzdálenosti nenulového kódového slova od nuly.

Příklad

Zvolme $p = 2$, tedy symboly jsou 0 a 1. Dále položme $n = 5$, $g(x) = x^2 + x + 1$. Pak kódová slova jsou polynomy

$$0, \quad x^2 + x + 1, \quad x^3 + x^2 + x, \quad x^3 + 1,$$
$$x^4 + x^3 + x^2, \quad x^4 + x^3 + x + 1, \quad x^4 + x, \quad x^4 + x^2 + 1.$$

Budeme-li polynomy psát jako posloupnosti koeficientů, budeme kódovat takto:

$$\begin{aligned} 000 &\mapsto 00000, & 100 &\mapsto 10010, \\ 001 &\mapsto 00111, & 101 &\mapsto 10101, \\ 010 &\mapsto 01001, & 110 &\mapsto 11011, \\ 011 &\mapsto 01110, & 111 &\mapsto 11100. \end{aligned}$$

Je ihned vidět, že nejmenší vzdálenost nenulového kódového slova od nuly je 2, jsme tedy schopni detekovat chybu v jednom symbolu. Opravit tuto chybu nejsme obecně schopni, například posloupnost 01000 by mohla vzniknout jednou chybou na druhé pozici anebo jednou chybou na páté pozici.

Kód opravující chybu na jedné pozici

Věta 1. *Zvolme prvočíslo p a přirozené číslo $m > 1$. Nechť K je těleso mající právě p^m prvků, nechť $\alpha \in K$ je libovolný generátor multiplikativní grupy $K^\times$ tělesa K a $g(x)$ je minimální polynom prvku α nad tělesem $\mathbb{Z}_p$. Pak polynomiální kód délky $n = \frac{p^m-1}{p-1}$ daný polynomem $g(x)$ je schopen opravit chybu na jedné pozici.*

Důkaz. Platí $\mathrm{st}(g) = m < 1 + p + \cdots + p^{m-1} = n$. Stačí tedy ukázat, že žádné kódové slovo nemá vzdálenost 1 nebo 2 od nuly. Předpokládejme naopak, že takové kódové slovo existuje, tj. pro nějaké $i \in \{0, 1, \ldots, n-1\}$, $a \in \mathbb{Z}_p^\times$, je polynom ax^i kódové slovo, anebo pro nějaká $0 \leq i < j < n$, $a, b \in \mathbb{Z}_p^\times$, je polynom $ax^j + bx^i$ kódové slovo. Protože polynom g je nesoudělný s polynomem x, první případ $g(x) \mid ax^i$ vede ke sporu. Druhý případ $g(x) \mid ax^j + bx^i = a(x^{j-i} + ba^{-1})x^i$ dává $g(x) \mid x^{j-i} + ba^{-1}$, tedy $\alpha^{j-i} = -ba^{-1} \in \mathbb{Z}_p^\times$, odkud $\alpha^{(j-i)(p-1)} = 1$. Protože řád prvku α v grupě $K^\times$ je $p^m - 1$, dostáváme $p^m - 1 \mid (j-i)(p-1)$, tj. $n \mid j - i$, což je ve sporu s tím, že $0 < j - i < n$.

Ještě lepší (pro $p > 2$) kód opravující chybu na jedné pozici

Věta 2. *Zvolme prvočíslo p a přirozené číslo m. Nechť K je těleso mající právě p^m prvků, nechť $\alpha \in K$ je libovolný generátor multiplikativní grupy $K^\times$ tělesa K a $f(x)$ je minimální polynom prvku α nad tělesem $\mathbb{Z}_p$. Je-li $p^m - 1 > m + 1$, pak polynomiální kód délky $n = p^m - 1$ daný polynomem $g(x) = (x-1) \cdot f(x)$ je schopen opravit chybu na jedné pozici.*

Důkaz. Platí $\mathrm{st}(g) = m + 1 < n$. Stačí tedy ukázat, že žádné kódové slovo nemá vzdálenost 1 nebo 2 od nuly. Předpokládejme naopak, že pro nějaké $i \in \{0, 1, \ldots, n-1\}$, $a \in \mathbb{Z}_p^\times$, je polynom ax^i kódové slovo, anebo pro nějaká $0 \le i < j < n$, $a, b \in \mathbb{Z}_p^\times$, je polynom $ax^j + bx^i$ kódové slovo. Protože polynom g je nesoudělný s polynomem x, první případ $g(x) \mid ax^i$ vede ke sporu. Druhý případ $g(x) \mid ax^j + bx^i = a(x^{j-i} + ba^{-1})x^i$ dává $g(x) \mid x^{j-i} + ba^{-1}$, odkud $x - 1 \mid x^{j-i} + ba^{-1}$, a proto $ba^{-1} = -1$. Dále odtud plyne $f(x) \mid x^{j-i} + ba^{-1} = x^{j-i} - 1$, tedy $\alpha^{j-i} = 1$. Protože řád prvku α v grupě $K^\times$ je $n = p^m - 1$, dostáváme $n \mid j - i$, což je ve sporu s tím, že $0 < j - i < n$.

Kódy opravující chyby na více pozicích

Věta 3. *Zvolme prvočíslo p a přirozené číslo m. Nechť K je těleso mající právě p^m prvků, nechť $\alpha \in K$ je libovolný generátor multiplikativní grupy $K^\times$ tělesa K. Nechť $r \geq -1$, $t > 0$ jsou celá čísla. Předpokládejme, že polynom $g(x)$ je dělitelný minimálním polynomem prvku α^{r+j} pro každé $j = 1, 2, \ldots, 2t$ a platí $\mathrm{st}(g) < p^m - 1$. Pak polynomiální kód délky $n = p^m - 1$ daný polynomem $g(x)$ je schopen opravit chybu na t pozicích.*

Důkaz. Protože $\mathrm{st}(g) < p^m - 1$, existuje v $K^\times$ prvek, který není kořen $g(x)$, a tedy $2t < n$. Předpokládejme, že existuje nenulové kódové slovo, jehož vzdálenost od nuly nepřevyšuje $2t$. Existují tedy $b_1, \ldots, b_{2t} \in \mathbb{Z}_p$, ne všechny nuly, a $0 \leq k_1 < k_2 < \cdots < k_{2t} < n$ tak, že polynom $f(x) = \sum_{i=1}^{2t} b_i x^{k_i}$ je kódové slovo, tj. $g(x) \mid f(x)$ a $f(\alpha^{r+j}) = 0$ pro každé $j = 1, 2, \ldots, 2t$. Pak $(\alpha^{(r+j)k_i})_{j,i=1,2,\ldots,2t}$ je matice s lineárně závislými sloupci. Ovšem pro $k = \sum_{i=1}^{2t} k_i$ platí $0 = \det(\alpha^{(r+j)k_i})_{j,i=1,2,\ldots,2t} = \alpha^{(r+1)k} \cdot \prod_{1 \leq i < j \leq 2t} (\alpha^{k_j} - \alpha^{k_i})$ užitím vzorce pro Vandermondův determinant. Ale to je součin mající pouze nenulové činitele, neboť α má řád n, spor.

Příklad

Věta 3 je zobecněním věty 2, stačí zvolit $r = -1$ a $t = 1$. Je-li $p = 2$, je i věta 1 důsledkem věty 3, zvolte $r = 0$, $t = 1$ a uvědomte si, že $\alpha^p = \alpha^2$ je také kořen polynomu $g(x)$.

Užijme větu 3 pro konstrukci kódu, který by měl mít skutečně reálné využití (užívá se prý při vyhodnocování QR kódů):

Nechť $K = \mathbb{Z}_2[x]/(x^4 + x + 1)$, označme $\alpha = x + (x^4 + x + 1)$.

Minimální polynom prvků $\alpha, \alpha^2, \alpha^4, \alpha^8$ je $x^4 + x + 1$.

Minimální polynom prvků $\alpha^3, \alpha^6, \alpha^{12}, \alpha^9$ je $x^4 + x^3 + x^2 + x + 1$.

Minimální polynom prvků α^5, α^{10} je $x^2 + x + 1$.

Proto předpoklady předchozí věty pro $p = 2$, $m = 4$, $n = 15$, $r = 0$, $t = 3$ splňuje polynom

$$g(x) = (x^4 + x + 1) \cdot (x^4 + x^3 + x^2 + x + 1) \cdot (x^2 + x + 1) =$$
$$= x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1.$$

Odpovídající kód délky 15 má 5 významových symbolů, 10 kontrolních symbolů a je schopen opravit chyby až na třech pozicích.