Algorithms of Information Security: Samoopravné kódy

Olha Jurečková, Martin Jureček
{jurecolh,jurecmar}@fit.cvut.cz

Faculty of Information Technology
Czech Technical University in Prague

October 7, 2020

Základní definice

Definition

Hammingova vzdálenost $d(x, y)$ dvou vektorů x a y je rovna počtu souřadnic, ve kterých se liší.

Příklad. $d(1000111, 1010110) = 2.$

Definition

Generující matici lineárního (n, k) kódu C v F^n je $k \times n$ matice G, s prvky v F, taková, že její řádky tvoří báze C.

Matice G je ve standardním tvaru, platí-li $G = (I \mid A)$, kde I je jednotková matice $k \times k$ a A je libovolná matice $k \times (n-k)$. Generující matice má rozměr $k \times n$ a musí splňovat 3 základní pravidla:

1. každý řádek matice je kódovým slovem
2. řádky matice jsou lineárně nezávisle, takže hodnost matice G je rovna k
3. každé kódové slovo je lineární kombinaci řádků matice.

Má-li kód C generující matici $G = (I \mid A)$, pak jeho kontrolní matice odpovídá $H = (-A^T \mid I)$, kde I je jednotková matice $(n-k) \times (n-k)$.

Příklad. Uvažujme těleso F_3 a nechť kontrolní matice kódu (5,2) je následující:

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

Převeďte matici H do standardní formy a najděte generující matici G daného kódu.

Řešení: Máme matici

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

vynásobíme druhý řádek matici 2 a dostaneme následující matici:

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

Lineární kódy

Dostaneme $H' = (I_3 A)$, takže

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

Poznámka. Kontrolní matice lineárního kódu C je generující maticí jeho duálního kódu.
Protože H je generující matice kódu $C^{\perp}$, potom najdeme kontrolní matici kódu $C^{\perp}$, což je generující matice kódu C. Dostaneme

$$G = (-A^T \mid I_2) = \begin{pmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Lineární kódy

Příklad. Je dána generující matice G nad tělesem F_3. Určete kontrolní matici H lineárního kódu generovaného následující maticí

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

Řešení: Máme matici

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

odečteme první řádek od 3. řádku a dostaneme následující matici:

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

Lineární kódy

Dále 3. řádek výnásobíme 2 a druhý řádek odečteme od prvního řádku. Potom 3. řádek odečteme od 2. řádku a dostaneme následující matici:

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

Použijeme následující vztah $H = (-A^T \mid I)$ a dostaneme

$$H = (-A^T \mid I_2) = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 \end{pmatrix}$$

Lineární kódy

Příklad. Uvažujme následující binární kód
$C = \{(0,0,0),(1,0,1),(0,1,1),(1,1,0)\}$.

- Dokažte, že C je lineární kód.
- Nájděte minimální vzdálenost kódu C.
- Nájděte generující matici kódu C.

Řešení:

- Vektor $(0,0,0) \in C$, operace sčítání vektorů z F_2^3 je uzavřená a každý prvek(vektor) z C má opačný prvek.
- Postupně spočteme Hammingovou váhu všech nenulových slov a zjistíme, že minimální váha je rovna 2. Podle vety (Nechť C je lineární kód nad F_q^n. Potom $d(C) = wt(C)$) plyne, že minimální vzdálenost kódu C je rovna 2.

- Velikost kódu je 4, takže $k = 2$ a generující matice G musí mít dva řádky. Můžeme vzít např. první dva nenulové vektory a dostaneme:

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

Cyklické kódy

Definition

Cyklický kód je lineární kód, jehož generující matice je tvořena kódovými slovy (vektory). Tato kódová slova vzniknou cyklickým posunem. Lineární kód C délky n nad tělesem F_q je tedy invariantní vzhledem k cyklickému posunu jeho souřadnic.

Pro každé slovo $a = (a_0, \ldots, a_{n-1}) \in F_q^n$ platí: $(a_0, \ldots, a_{n-1}) \in C \Rightarrow (a_1, \ldots, a_{n-1}, a_0) \in C$. Každé slovo (vektor) a můžeme ztotožnit s polynomem nad tělesem $F_q : a = (a_0, \ldots, a_{n-1})$ je reprezentován $a(x) = a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ nebo také jinak:

$$\sum_{i=0}^{n-1} a_i x^i \in F_q^n[x].$$

Cyklické kódy

Kódové polynomy jsou pak násobky generujícího polynomu, neboť cyklickému posunu odpovídá násobení polynomem x. Pro cyklický kód s polynomem $a(x) = a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ platí, že je jeho generující matice je:

$$G = \begin{pmatrix} a_0 & a_1 & \ldots & a_{n-1} & 0 & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & a_{n-1} & \ldots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & & \\ 0 & 0 & \ldots & a_0 & a_1 & \ldots & a_{n-1} \end{pmatrix}$$

Příklad.
Najděte generující matici pro cyklický kód (6,3), jehož generující polynom je následující: $1 \cdot x^3 + 0 \cdot x^2 + 1 \cdot x + 1$.
Řešení. Okamžitě ze znalosti koeficientů polynomu $x^3 + x + 1$ dostaneme posunem generující matici

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

Cyklické kódy

Příklad. Najděte generující a kontrolní matici pro binární cyklický kód délky 6 s generujícím polynomem: $g(x) = x^3 + 1$.

Řešení. Máme $n = 6$. Všimněte si, že jsme definovali k tak, že $deg(g(x)) = n - k$, potom $n - k = 3$. Odkud dostaneme $k = 3$. Okamžitě ze znalosti koeficientů polynomu $g(x) = x^3 + 1$ dostaneme posunem generující matici

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

Dále spočteme $h(x) = (x^n - 1) : g(x)$, tj. $h(x) = (x^6 - 1) : (x^3 + 1) = (x^6 + 1) : (x^3 + 1) = x^3 + 1$. Potom kontrolní matici je následující:

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

Cyklické kódy

Příklad. Uvažujme binární cyklický kód C délky 7 s generujícím polynom: $g(x) = x^3 + x + 1$. Potom

- Ověřte, že kód C je cyklický (tj. g dělí $x^7 - 1$).
- Najděte generující a kontrolní matici pro daný binární cyklický kód C.

Řešení.

- Snadno ověříme, že $x^7 - 1 = 1 + x^7 = (1 + x + x^3)(1 + x + x^2 + x^4)$ nad F_2, takže $g(x)$ dělí $x^7 + 1$ (nebo $x^7 - 1$) a tedy kód C je cyklický $[7, 4]$ kód.
- Máme $n = 7$. Všimněte si, že jsme definovali k tak, že $deg(g(x)) = n - k$, potom $n - k = 3$. Odkud dostaneme $k = 4$. Okamžitě ze znalosti koeficientů polynomu $g(x) = x^3 + x + 1$ dostaneme posunem generující matici

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

Dále spočteme $h(x) = (x^n - 1) : g(x)$, tj.
$h(x) = (x^7 - 1) : (x^3 + x + 1) = 1 + x + x^2 + x^4$. Potom kontrolní matice je následující:

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

Cyklické kódy

Jak získat všechny cyklické kódy dané délky n?
Vše, co musíme udělat, je najít všechny faktory $x^n - 1$.
Příklad. Najděte všechny binární cyklické kódy délky 3.
Řešení. Pokud chceme určit $g(x)$, potom potřebujeme najít rozklad mnohočlenu $x^3 - 1$ nad tělesem F_2, který má člen stupně $n-k$.
Poznamenejme, že $x^3 - 1 = (x+1)(x^2+x+1)$. Dostaneme následující výsledky:

generující polynom	kód v R_3
1	R_3
$x+1$	$\{0, 1+x, x+x^2, 1+x^2\}$
x^2+x+1	$\{0, 1+x+x^2\}$
x^3-1	$\{0\}$

Cyklické kódy

Příklad. Najděte všechny cyklické kódy délky 4 nad F_3.

Řešení. Pokud chceme určit $g(x)$, potom potřebujeme najít rozklad mnohočlenu $x^4 - 1$ nad tělesem F_3, který má člen stupně $n - k$. Nechť $k = 1$, potom budeme postupně dělit $x^4 - 1$ děliteli $x, x - 1, x - 2 = x + 1$ (všechny polynomy stupně 1.) Podobně budeme postupovat pro $k = 2$ a pro $k = 3$ a pro $k = 4$. Poznamenejme, že $x^4 - 1 = (x - 1)(x + 1)(x^2 + 1)$. Dostaneme následující výsledky, kromě triviálních případů (R_4 a $\{0\}$.)

- Kód (4,3) generovaný pomocí $x - 1 = x + 2$.
- Kód (4,3) generovaný pomocí $x + 1$.
- Kód (4,2) generovaný pomocí $x^2 + 1$.
- Kód (4,2) generovaný pomocí $x^2 - 1 = x^2 + 2$.

- Kód (4,1) generovaný pomocí $(x-1)(x^2+1) = x^3+2x^2+x+2$.
- Kód (4,1) generovaný pomocí $(x+1)(x^2+1) = x^3+x^2+x+1$.

Konečná tělesa

Definition

Mějme konečné těleso F_q a v něm libovolný nenulový prvek a. Nejmenší přirozené číslo n takové, že $a^n = 1$, se nazývá řád prvku.

Uvažujme těleso F_{2^3}. Toto těleso je tvořeno polynomy nad Z_2 modulo ireducibilní polynom $x^3 + x + 1$. Obsahuje prvky $\{0, 1, x, x+1, x^2, x^2+x, x^2+1, x^2+x+1\}$. Charakteristika tohoto tělesa je $p = 2$. Všechny prvky kromě 0 a 1 mají řád $n = q - 1 = 8 - 1 = 7$, a tudíž jsou všechny primitivní.

Dále uvedeme tabulku sčítání:

$+$	0	1	x	x^2
0	0	1	x	x^2
1	1	0	$x+1$	x^2+1
x	x	$x+1$	0	x^2+x
x^2	x^2	x^2+1	x^2+x	0
$x+1$	$x+1$	x	1	x^2+x+1
x^2+x	x^2+x	x^2+x+1	x^2	x
x^2+x+1	x^2+x+1	x^2+x	x^2+1	$x+1$
x^2+1	x^2+1	x^2	x^2+x+1	1

Pokračování tabulky sčítání:

$+$	$x+1$	x^2+x	x^2+x+1	x^2+1
0	$x+1$	x^2+x	x^2+x+1	x^2+1
1	x	x^2+x+1	x^2+x	x^2
x	1	x^2	x^2+1	x^2+x+1
x^2	x^2+x+1	x	$x+1$	1
$x+1$	0	x^2+1	x^2	x^2+x
x^2+x	x^2+1	0	1	$x+1$
x^2+x+1	x^2	1	0	x
x^2+1	x^2+x	$x+1$	x	0

Dále uvedeme tabulku násobení:

$\cdot$	0	1	x	x^2
0	0	0	0	0
1	0	1	x	x^2
x	0	x	x^2	$x+1$
x^2	0	x^2	$x+1$	x^2+x
$x+1$	0	$x+1$	x^2+x	x^2+x+1
x^2+x	0	x^2+x	x^2+x+1	x^2+1
x^2+x+1	0	x^2+x+1	x^2+1	1
x^2+1	0	x^2+1	1	x

Pokračování tabulky násobení:

$\cdot$	$x+1$	x^2+x	x^2+x+1	x^2+1
0	0	0	0	0
1	$x+1$	x^2+x	x^2+x+1	x^2+1
x	x^2+x	x^2+x+1	x^2+1	1
x^2	x^2+x+1	x^2+1	1	x
$x+1$	x^2+1	1	x	x^2
x^2+x	1	x	x^2	$x+1$
x^2+x+1	x	x^2	$x+1$	x^2+x
x^2+1	x^2	$x+1$	x^2+x	x^2+x+1

Při práci s Reed Solomonovými kódy pro nás bude výhodné reprezentovat nenulové prvky konečného tělesa jako mocniny primitivního prvku. Vyberme si jeden z primitivních prvků v tělese F_{2^3} (například x) a označme jej α. Prvkem α^2 rozumíme součin $\alpha \cdot \alpha = x \cdot x = x^2$. Pokračujeme dále s $\alpha^3 = \alpha^2 \cdot \alpha = x^2 \cdot x = x + 1$. Zbylé přiřazení uvedeme v následující tabulce:

α	x
α^2	x^2
α^3	$x+1$
α^4	x^2+x
α^5	x^2+x+1
α^6	x^2+1
α^7	1

Reed Solomonové kódy

Příklad. Rozhodněte, zda existuje RS kód s parametry $[7,5,3]_q$. Existuje-li takový kód, najděte nejmenší q a jeho kontrolní matici.
Řešení.
Hledáme nejmenší q, pro které 7 dělí $q-1$, zřejmě je to právě $q=2^3$. Reprezentujme si prvky tělesa F_8 pomocí kořenu α polynomu x^3+x+1 ireducibilního nad F_2, tedy $F_8=\left\{a_0+a_1\alpha+a_2\alpha^2 \mid a_i\in F_2\right\}$. Protože je grupa ${F_8}^*$ cyklická, je každý nejednotkový prvek řádu 7, proto dopočteme matici

$$H=\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1\\ 1 & \alpha & \alpha^2 & \alpha+1 & \alpha^2+\alpha & \alpha^2+\alpha+1 & \alpha^2+1\end{pmatrix}$$

Reed Solomonové kódy

Příklad. Uvažujme konečné těleso F_5 a ať $\alpha = 2$. Najděte:

- generující polynom pro $RS(4, 2)$
- generující matici pro $RS(4, 2)$
- kontrolní matici pro $RS(4, 2)$.

Řešení.

- Uvažujme konečné těleso F_5 a $\alpha = 2$. Snadno se zkontroluje, že $ord(\alpha) = 4$, a $\beta = \alpha$ je tedy primitivní 4tý kořen jednotky. Poznámka. Vytvoříme generující polynom $g(x)$ RS kódu pomocí následujícího vzorečku:

$$g(x) = (x - \alpha)(x - \alpha^2) \ldots (x - \alpha^{n-k}),$$

  kde α je primitivní prvek.
  Potom generující polynom je:

$$g(x) = (x - 2)(x - 4) = 3 + 4x + x^2.$$

- Dále mužeme napsat generující matici pro $RS(4, 2)$:

$$G = \begin{pmatrix} 3 & 4 & 1 & 0 \\ 0 & 3 & 4 & 1 \end{pmatrix}$$

Reed Solomonové kódy

- Víme generující matici a potřebujeme najít kontrolní matici pro $RS(4,2)$. Nejprve upravíme generující matici do standardní formy a získáme následující matice:

$$\begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 3 & 2 \end{pmatrix}$$

  Teď máme generující matici ve tvaru $G = (I \mid A)$, pak jeho kontrolní matice je $H = (-A^T \mid I)$, kde I je jednotková matice. V našem případě

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 2 \end{pmatrix}$$

Potom dostaneme následující matici

$$(-A^T \mid I) = \begin{pmatrix} 2 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix}$$

Různými úpravami získáme následující matice:

$$H = \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

Cyklické kódy

Příklad.
Najděte generující matici pro cyklický kód (7,3), jehož generující polynom je následující: $x^4 + x^2 + x + 1$.

Řešení. Okamžitě ze znalosti koeficientů polynomu $x^4 + x^2 + x + 1$ dostaneme posunem generující matici

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

Příklad. Najděte generující matici pro binární cyklický kód délky 9 s generujícím polynomem: $g(x) = x^6 + x^3 + 1$.

Řešení. Máme $n = 9$. Všimněte si, že jsme definovali k tak, že $deg(g(x)) = n - k$, potom $n - k = 6$. Odkud dostaneme $k = 3$. Okamžitě ze znalosti koeficientů polynomu $g(x) = x^6 + x^3 + 1$ dostaneme posunem generující matici

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