Kódolástechnika

ZH felkészítő gyakorlat, 2013. 11. 21.

Igaz-hamis

I	Egy $C(n, k)$ paraméterű ciklikus kód paritásellenőrző polinomja osztja az x^n-1 polinomot.
H	Az LZ77 futtatásához ismerni kell a forráseloszlást.
I	A GF(8)-ban 2 konjugált gyökcsoport van.
H	A BCH kód mindig MDS tulajdonságú.
H	A Shannon–Fano kód rövidebb átlagos kódszóhosszat ér el a Huffmann-nál.
I	A forráskódolásnál az egyértelmű dekódoláshoz nem lehetnek a kódszavak tetszőlegesen rövidek.
H	Véges forrás ABC esetén van olyan eloszlás, hogy az entrópia negatív.
H	Egy t hibát javítani képes lineáris ciklikus kódnál a hibacsapda algoritmus a hibavektorban tetszőleges helyen előforduló t vagy annál kisebb számú hibát képes javítani.

Rövid kérdések

Mennyi két független bináris egyenletes eloszlású valószínűségi változó közös entrópiája?

$H(X) = \log_2 N$, most $N = 2$.

$H(X) = \log_2 2 = 1$, $H(X, Y) = H(X) + H(Y) = 1 + 1 = 2$

Mekkora egy emlékezet nélküli, 4 állapotú, egyenletes eloszlást követő forrás 4 hosszúságú blokkjainak entrópiája?

$$H(x_k, x_{k-1}, x_{k-2}, x_{k-3}) = \sum_{i=1}^{4} H(x_{k-i+1}) = 4H(x_k) = 4\log_2 4 = 4 \cdot 2 = 8$$

$4 \cdot \log_2 4$ magyarázata: az első 4-es a blokkhossz, a második 4-es az állapotok száma.

Adott egy 5 szimbólumot kibocsátó forrás, $p_1 > p_2 > p_3 > p_4 > p_5$, és adottak az l_1, l_2, l_3, l_4 Huffmann-kódolással kapott kódszóhosszak. Mennyi l_5?

$l_5 = l_4$, a Huffmann kódolás azon tulajdonsága miatt, hogy a két legkisebb valószínűségű, azaz a két leghosszabb kódszó hossza biztosan megegyezik, mert ha nem így lenne, nem lenne optimális a kód.

Van egy $C(15, 11)$ kódunk. Lehet-e ez egy Hamming-kód?

Mivel a Hamming-kódok perfekt kódok, teljesülnie kell a $2^{n-k} = n + 1$ egyenlőségnek, azaz $2^4 = 16$-nak, ami igaz, tehát a kód lehet Hamming-kód.

Ha egy Hamming-kódból kódátfűzéssel $\lambda = 12$ paraméterű burst hiba javítására alkalmas kódot csinálunk, milyen hosszúságú burst hiba javítható?

Burst hiba javításánál, ha az eredeti kód t hibát volt képes javítani, akkor az új kód $t \cdot \lambda$ hosszúságú burst hiba javítására képes. Mivel az eredeti kód Hamming-kód volt, így 1 hibát javít, tehát $1 \cdot 12 = 12$ hosszúságú burst hiba javítására képes a kód.

Lehet-e a generátorpolinom egy ciklikus RS kódnál a következő:
$g(x) = y^6x^{12} + y^8x^{11} + \ldots + y^3$?

Nem, mert a generátorpolinom főegyütthatójának a 1-nek kell lennie.

1. feladat

Adott egy hibajavító lineáris bináris kód a következő generátormátrixszal:

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

A bináris szimmetrikus csatornán a leadott kódszóhoz az **e** = (01100) hibavektor adódik.

Milyen típusú a kód?

A típust $C(n, k)$-val adjuk meg, ahol n a kódszóhossz, k az üzenethossz. A generátormátrix $k \times n$-es mátrix, így k=5, n=2, tehát a kód típusa: $C(5,2)$.

Ilyen hibavektor (lásd fent) hatására milyen szindrómavektor keletkezik?

Az **s** szindrómavektort a következőképpen kapjuk:

$\mathbf{Hv}^T = \mathbf{s}^T$, ahol **H** a paritásellenőrző mátrix, **v** a vett vektor. Mivel **v** = **c** + **e**, azaz a vett vektor a kódszó és a hibavektor összege, ezért:

$\mathbf{Hv}^T = \mathbf{H}(\mathbf{c}+\mathbf{e})^T = \mathbf{Hc}^T + \mathbf{He}^T = \mathbf{s}^T$, de $\mathbf{Hc}^T = \mathbf{0}^T$, hiszen **c** kódszó, tehát: $\mathbf{He}^T = \mathbf{s}^T$.

Ehhez meg kell határoznunk a **H** mátrixot. Mivel a kód szisztematikus – azaz a **G** generátormátrix baloldalán az $\mathbf{E}_{k \times k}$, most $\mathbf{E}_{2 \times 2}$ egységmátrix található – a **H**-t a következőképpen kapjuk:

Ha a **G** $k \times n$-es, akkor a **H** $(n-k) \times n$-es, esetünkben $\mathbf{G}_{2\times5} \Rightarrow \mathbf{H}_{3\times5}$. Mivel szisztematikus a kód, a **H** jobboldalán az $\mathbf{E}_{(n-k)\times(n-k)}$, azaz $\mathbf{E}_{3\times3}$ egységmátrix van. A maradék rész – tehát most az első két oszlop – a **G** jobb oldali, nem egységmátrix részének a transzponáltja. Tehát:

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

Most már csak a mátrix-vektor szorzást kell végrehajtanunk:

$$\mathbf{s}^T = \mathbf{He}^T = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Ehhez a szindrómavektorhoz milyen hibavektor identifikálódik?

Először meg kell határoznunk az **s** = (010) szindrómavektorunkhoz tartozó $E_{\mathbf{s}}$, azaz $E_{(010)}$ hibacsoportot. Ez azon **e** hibavektorok csoportja, melyekre fennáll, hogy $\mathbf{He}^T = \mathbf{s}^T$, azaz most $\mathbf{He}^T=(010)^T$.

Egy hibavektort már tudunk a feladatból, ez a (01100). Mivel az azonos csoportba tartozó hibavektorok csak egy kódszóban különböznek, ezért ha ehhez hozzáadjuk az összes nem **0** kódszót, akkor készen vagyunk.

A kódszavakat úgy kapjuk, hogy a **G**-t balról szorozzuk az összes nem **0 u** üzenetvektorral. A szóba jöhető **u** vektorok: (01), (10), (11). A (01)-gyel való szorzás a 2. sor, az (10)-val való szorzás az 1. sor kiválasztását jelenti, az (11)-gyel való szorzás pedig a két sor (modulo 2) összeadását. A kapott kódszavak így:

$\mathbf{c}_1=(01)\cdot\mathbf{G}$ = (01110), $\mathbf{c}_2=(10)\cdot\mathbf{G}$ = (10111), $\mathbf{c}_3=(11)\cdot\mathbf{G}$ = (11001).

A (01100)-hoz ezeket hozzáadva kapjuk a hibacsoport tagjait (természetesen a (01100)-val együtt):

$E_{(010)}$ = {(01100), (00010), (11011), (10101)}.

Az identifikált $\mathbf{e}_{\mathbf{s}}$ hibavektor ezek közül a legnagyobb valószínűséggel előforduló. Egy **e** hibavektor valószínűsége a következőképpen adható meg:

$$P(\mathbf{e}) = \left(\frac{P_b}{1-P_b}\right)^{w(\mathbf{e})}(1-P_b)^n$$

P_b a csatorna bithiba-valószínűségét jelenti, $w(\mathbf{e})$, pedig az **e** vektor súlyát (a benne lévő 1-esek számát), n a kódszóhossz. Mivel P_b és n adott, így látható, hogy egy hibacsoportban mindig a legkisebb súlyú hibavektor lesz a legvalószínűbb. Ezt nevezzük csoportvezetőnek (group leadernek). A csoportvezető itt a (00010), így ő lesz az identifikált $\mathbf{e}_{\mathbf{s}}$ hibavektor.

2. feladat

Hatványtábla a GF(4) felett:

$y^{-\infty}$	0	0
y^0	1	1
y^1	y	2
y^2	$y+1$	3
y^3	1	4
y^4	y	5
…	…	…

Az irreducibilis polinom: $p(y) = y^2 + y + 1$ (fokszáma 2, mert GF(4)-ben vagyunk, és $4=2^2$).

Adja meg az 1. konjugált gyökcsoportot a GF(4) felett!

Tudjuk, hogyha β gyök, akkor β^{2^i} is gyök lesz. Az 1. konjugált gyökcsoport felírásához válasszuk ki a primitív elemet, y-t! Ez lesz az első tagja a gyökcsoportnak. Ekkor y^2 is gyök, de mivel $y^4 = y$, ezért nincs több gyök a csoportban, hiszen újra y-t kaptunk. Tehát az 1. konjugált gyökcsoport:

$$C_g^{(1)} = \{y, y^2\}$$

Adja meg az ehhez tartozó minimálpolinomot!

A hatványtábla alapján dolgozunk:

$$\Phi_1(x) = (x + y)(x + y^2) = x^2 + yx + y^2x + y^3 = x^2 + (y + y^2)x + 1 = x^2 + x + 1$$

Ha ez egy BCH kód generátorpolinomja, hány hibát képes javítani a kód?

Egy t hibát javító BCH kód $g(x)$ generátorpolinomja:

$$g(x) = \Phi_1(x)\Phi_3(x) \dots \Phi_{2t-1}(x)$$

Most $g(x) = \Phi_1(x)$, tehát 2t – 1 = 1, azaz t = 1, tehát 1 hibát javít a kód.

3. feladat

Adott egy lineáris bináris kód, melynek kódszavai: $\mathbf{c}_0$ = (00000), $\mathbf{c}_1$= (01111), $\mathbf{c}_2$ = (10100) és $\mathbf{c}_3$ = (11011).

Adja meg a kód típusát!

A kód típusa $C(n, k)$ alakú, n a kódszóhossz, k az üzenethossz. Itt most n = 5, k = 2, mivel a négyféle kódszóhoz négy üzenet tartozik, ezt pedig 2 biten lehet binárisan kódolni, az üzenetek (00), (01), (10), (11) lehetnek. Így a kód típusa $C(5, 2)$.

Adja meg a generátormátrixot!

Az n×k-s azaz most 2×5-ös **G** mátrix első sora az (10)-hoz tartozó kódszó, azaz $\mathbf{c}_2$, a második sora a (01)-hez tartozó kódszó, azaz $\mathbf{c}_1$. A **G** mátrix:

$$\mathbf{G}_{2\times5} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Adja meg a paritásellenőrző mátrixot!

Mivel szisztematikus a kód, a **H** (n-k)×n-es, azaz 3×5-ös mátrixot az 1. feladatban már ismertetett módon kapjuk:

$$\mathbf{H}_{3\times5} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Hány hibát javít a kód?

$t = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor$, ahol t a javított hibák száma, d_{min} a minimális Hamming-távolság. Mivel bináris kódról van szó, $d_{min} = w_{min}$, azaz a minimális súly (a **0**-t leszámítva), ami itt a $\mathbf{c}_2$ súlya, azaz 2. Tehát itt $\left\lfloor \frac{2-1}{2} \right\rfloor = 0$, azaz a kód csak hibajelzésre alkalmas, hibát nem javít.

4. feladat

Hatványtábla a GF(8) felett:

$y^{-\infty}$	0	0
y^0	1	1
y^1	y	2
y^2	y^2	3
y^3	$y+1$	4
y^4	y^2+y	5
y^5	y^2+y+1	6
y^6	y^2+1	7
y^7	1	8
y^8	y	9
…	…	…

Az irreducibilis polinom: $p(y) = y^3 + y + 1$

Adja meg egy $C(7, 5)$ RS kódnak a generátorpolinomját a standard alakban!

Ciklikus RS kód esetén:

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

Azaz esetünkben:

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

Az egyes átalakításoknál a hatványtáblát használtuk.

Hány hibát tud javítani a kód?

Mivel az RS kód MDS kód, így $d_{min} = n - k + 1$, tehát:

$$t = \left\lfloor \frac{d_{min} - 1}{2} \right\rfloor = \left\lfloor \frac{n - k + 1 - 1}{2} \right\rfloor = \left\lfloor \frac{n-k}{2} \right\rfloor = \left\lfloor \frac{7-5}{2} \right\rfloor = 1$$

Tehát a kód 1 hibát képes javítani.