Hibajavító kódok

Témavázlat

- Hibajavító kódolás
  - Blokk-kódok
    - Hamming-távolság, Hamming-súly
    - csoportkód
    - S^n-beli u középpontú t sugarú gömb
    - hibajelző képesség
    - minimális távolságú dekódolás
    - hibajavító képesség
    - Hamming-korlát, tökéletes kód
  - Lineáris kódok
    - Generátor- és paritásellenörző mátrix
    - Kódolás lineáris kóddal
    - Singleton-korlát, maximális távolságú kód
    - Hamming-kód
    - BCH-kód
    - Polinomkódok
  - Irodalomjegyzék

Hibajavító kódolás

A csatornán áthaladó jelsorozat szükségszerűen időnként meghibásodik.

Hibajavító kódolás esetén a szükségesnél több jelet rendelünk hozzá az üzenetekhez, és ez a redundancia teszi lehetővé, hogy felismerhessük, ha hiba történik, illetve a keletkező hibát kijavíthassuk.

A hibafelismerés, illetve a hibajavítás mindig csak bizonyos számú hibáig működik helyesen.

Definíció:

K *blokk-kód*, ha A, $S \neq \emptyset$ véges halmazok, $f: A \rightarrow S^*$ injektív leképezés, $K=f(A) \subseteq S^n$. K elemei a *kódszavak.*

1. példa

Az üzenetek bináris jelsorozatok, az S={0,1} halmaz elemeiből épülnek fel.

Rendeljünk hozzá a kettő hosszú üzenetekhez öt hosszú kódokat a következő szabály szerint.

(0 0) → (0 0 0 0 0)
(0 1) → (0 1 1 1 0)
(1 0) → (1 0 1 0 1)
(1 1) → (1 1 0 1 1)

Négy elemből álló blokk-kód.

Ha a csatornán a (0 1 0 0 0) jelsorozat érkezik, tudjuk, hogy hiba történt, hiszen ez a szó nem szerepel a kódszavak között.

Az n hosszú jelsorozatok halmazán távolságot definiálunk.

Definíció.

$S \neq \emptyset$ véges halmaz, $u, v \in S^n$, $K \subseteq S^n$.

u és v *Hamming távolsága* u és v különböző komponenseinek a száma.

$$d(u,v) = \left| \{ i \mid 1 \leq i \leq n,\ u_i \neq v_i \} \right| .$$

Legyen az előbbiek mellett még $|K| \geq 2$.

K (*minimális*) *távolsága* az egymáshoz legközelebbi két kódszónak a távolsága.

$$d(K) = \min \{ d(u,v) \mid u,v \in K,\ u \neq v \}.$$

Blokk-kód jelölése: (n, M), vagy $(n, M)_s$ illetve $(n, M, d)_s$, ahol $M = |K|$ és $d = d(K)$.

A továbbiakban kizárólag blokk-kódokkal foglalkozunk.

Tétel.

A Hamming-távolság metrika az S^n halmazon:

bármely két S^n -beli vektornak létezik Hamming-távolsága, egy valós szám, amely kielégíti az alábbi három tulajdonságot:

- 1. $d(u,v) \geq 0$, és $d(u,v)=0 \leftrightarrow u=v$.
- 2. $d(u,v)=d(v,u)$.
- 3. $d(u,v) \leq d(u,w)+d(w,v)$. (Háromszög-egyenlőtlenség.)

Definíció.

Legyen S Abel csoport és $\underline{0}=\{0,...,0\}$, ahol 0 az S egységeleme.

$u \in S^n$ *Hamming súlya* az u nem nulla komponenseinek a száma.

$$w(u)=\left|\{i \mid 1 \leq i \leq n,\ u_i \neq 0\}\right|.$$

Ha K-nak van eleme a $\underline{0}$ -n kívül, akkor K (*minimális*) *súlya* a nem nulla elemei közül a minimális súlyúnak a súlya.

$$w(K)=\min\{w(u) \mid u \in K,\ u \neq \underline{0}\}.$$

2. példa.

(0 0) → (0 0 0 0 0)
(0 1) → (0 1 1 1 0)
(1 0) → (1 0 1 0 1)
(1 1) → (1 1 0 1 1)

Az u=(0 1 1 1 0) és v=(1 0 1 0 1) kódszavak távolsága
d(u,v)=4

A kód távolsága: 3

A z=(1 1 0 1 1) sorozat súlya: w(z)=4

A kód súlya: 3

Definíció.

Legyen K blokk-kód. K *csoportkód*, ha S Abel-csoport és $K \leq S^n$ (K részcsoport S^n-ben, a művelet a komponensenként végzett S-beli művelet.)

3. példa.

Az 1. példa S halmaza legyen $\mathbf{F}_2$ a kételemű test.

(0 0) → (0 0 0 0 0)
(0 1) → (0 1 1 1 0)
(1 0) → (1 0 1 0 1)
(1 1) → (1 1 0 1 1)

$\mathbf{F}_2$ az összeadásra csoportot alkot.

Az $\mathbf{F}_2$ elemeiből készített n hosszú vektorok is csoportot alkotnak az elemenkénti összeadásra nézve.

A K kódhalmaz részcsoport S^n -ben, így K csoportkód: bármely két kódszó különbsége is kódszó.

$\mathbf{F}_2$ -ben 1= -1, ezért elegendő azt megvizsgálnunk, hogy bármely két kódszó összege is kódszó-e.

Definíció.

Legyen $S \neq \emptyset$ véges halmaz, $n \in \mathbf{N}^+$, $t \in \mathbf{N}$, $u \in S^n$.

Az *S^n-beli u középpontú t sugarú gömb* azokat a sorozatokat tartalmazza S^n-ből, amelyek u-tól legfeljebb t távolságra vannak.

$$C_t(u)=\{v \mid v \in S^n,\ d(u,v) \leq t\}.$$

4. példa.

Legyen K az 1. példabeli kód, u=(0 0 0 0 0), t=1.
Az u körüli 1 sugarú gömbben a következő sorozatok vannak.

(0 0 0 0 0)
(1 0 0 0 0)
(0 1 0 0 0)
(0 0 1 0 0)
(0 0 0 1 0)
(0 0 0 0 1)

Figyeljük meg, hogy nincs a gömb elemei között a (0 0 0 0 0)-t leszámítva kódszó, ami összhangban van azzal, hogy a kód távolsága nagyobb 1-nél (3).

Definíció.

Egy kód *t-hiba jelző*, ha tetszőleges üzenetben előforduló legfeljebb t számú hibát képes jelezni, *pontosan t-hiba jelző*, ha t-hiba jelző, de van olyan t+1 hiba, amit nem jelez.

Egy kód *t-hiba javító*, ha tetszőleges üzenetben előforduló legfeljebb t számú hibát képes javítani, *pontosan t-hiba javító*, ha t-hiba javító, de van olyan t+1 hiba, amit hibásan javít.

Minimális távolságú dekódolásról beszélünk, ha minden S^n-beli szóhoz a hozzá legközelebb eső (egyik) kódszót választjuk.

A minimális távolságú dekódolás nem mindig a leghelyesebb stratégia. Ez a helyzet például, ha a minimális távolság nagyon nagy.

5. példa.

Az 1. példa kódját alkalmazva tegyük fel, hogy a
(0 1 0 0 0) hibás jelsorozat érkezik.

Minimális távolságú dekódolás esetén a hozzá legközelebbi kódszót, a
(0 0 0 0 0) szót választjuk helyette.

4. tétel.

Legyen K blokk-kód, tehát $K \subseteq S^n$, és d(K)=d. A kód akkor és csak akkor t-hiba jelző, ha t<d,

pontosan t-hiba jelző, ha t=d-1.

Minimális távolságú dekódolás esetén a d távolságú kód t-hiba javító, ha $t \leq \frac{d-1}{2}$

pontosan t-hiba javító, ha $t = \left\lfloor \frac{d-1}{2} \right\rfloor$

Bizonyítás.

- Ha t<d, és t helyen romlik el valamelyik kódszó, nem kapunk másik kódszót, s így fel tudjuk ismerni, hogy hiba történt, tehát a kód minden t-hibát tud jelezni.

  Ha azonban d helyen történt hiba, előfordulhat, hogy egy kódszó másik kódszóba megy át, s így nem érzékeljük, hogy hiba történt.

- Ha $t \leq \frac{d-1}{2}$

  és legfeljebb t helyen romlott el valamelyik kódszó, akkor a hibás szó az eredetihez van legközelebb a kódszavak közül, tehát minimális távolságú dekódolásnál a hibát helyesen javítjuk.

  Ha azonban a hibák száma $> \frac{d-1}{2}$

  akkor előfordulhat, hogy az elromlott szó az eredetitől különböző más kódszóhoz kerül közelebb, vagy több kódszótól is ugyanolyan távolságra van, és így minimális távolságú dekódolásnál a hibát rosszul javítjuk.

6. példa.

Legyen $S=\mathbf{F}_2$ a közleményszavak pedig k hosszú sorozatok.
Állapítsuk meg az alábbi kódok minimális távolságát, hibajelző- és hibajavító képességét.

- Kétszeri ismétlés kódja. A kódszót megkapjuk, ha a közleményszót kétszer egymás után leírjuk

  közleményszó: $(\alpha_1\ \alpha_2 \ldots \alpha_k)$

  kódszó: $(\alpha_1\ \alpha_2 \ldots \alpha_k\ \alpha_1\ \alpha_2 \ldots \alpha_k)$

  A kódszavak n=2k hosszúak.

  Az n hosszú bináris szavak közül azok a kódszavak, amelyekre az i-edik és k+i-edik elem megegyezik minden $1 \le i \le k$ esetén.

  A kód távolsága d=2, 1-hibajelző, hiba javítására nem alkalmas.

  (Bizonyos 2-hibák nem derülnek ki, pl. ha az első és a k+1-edik jegy hibás)

- Háromszori ismétlés kódja.

A kódszót megkapjuk, ha a közleményszót háromszor írjuk egymás után

közleményszó: $(\alpha_1\ \alpha_2 \ldots \alpha_k)$

kódszó: $(\alpha_1\ \alpha_2 \ldots \alpha_k\ \alpha_1\ \alpha_2 \ldots \alpha_k\ \alpha_1\ \alpha_2 \ldots \alpha_k)$

A kódszavak n=3k hosszúak.

Az n hosszú bináris szavak közül azok a kódszavak, amelyekre az i-edik, k+i-edik és 2k+i-edik elem megegyezik minden $1 \le i \le k$ esetén.

A kód távolsága d=3, 2-hibajelző, 1-hiba javító.

- Paritásvizsgálat kódja.

A kódszót megkapjuk, ha a közleményszó végére az elemek összegét írjuk.

közleményszó: $(\alpha_1 \alpha_2 \ldots \alpha_k)$

kódszó: $\alpha_1 \alpha_2 \ldots \alpha_k \beta$ $\qquad \beta = \sum_{i=1}^{k} \alpha_i$

A közleményszó végére a kódba 1 vagy 0 kerül aszerint, hogy a közleményszó jegyeinek az összege páratlan vagy páros volt.

A kódszavak n=k+1 hosszúak.

Az n hosszú bináris szavak közül azok a kódszavak, amelyek páros sok 1-est tartalmaznak.

A kód távolsága d=2, 1-hibajelző, hiba javítására nem alkalmas.

Az 1. és 3. kód hibajelző és hibajavító képessége megegyezik, ugyanakkor az 1. kód hossza k értékével sokkal gyorsabban nő, mint a 3. kód hossza, tehát a 3. alkalmazása gazdaságosabb.

5. tétel.(Hamming-korlát)

t-hiba javító $(n, M, d)_S$ kód esetén (s=|S|):

$$M \cdot \sum_{i=0}^{t} \binom{n}{i} (s-1)^i \leq s^n$$

Bizonyítás.

S^n-ben egy adott kódszótól i távolságra azok a szavak vannak, amelyek pontosan i helyen különböznek tőle. Az i különböző helyet $\binom{n}{i}$-féleképpen választhatjuk ki, mindegyik helyre $(s-1)^i$ jel kerülhet.

Az adott kódszó körüli t sugarú gömbben $\sum_{i=0}^{t} \binom{n}{i} (s-1)^i$ S^n-beli szó található, beleértve magát a kódszót is. M kódszavunk van. A kód t-hiba javító, ezért az M különböző kódszó körüli t-sugarú gömbök diszjunktak kell legyenek.

Összesen nem tartalmazhatnak több elemet, mint az S^n halmaz, vagyis s^n-et.

Definíció.

Az $(n, M, d)_S$ kód *tökéletes (perfekt),* ha a Hamming-korlát egyenlőséggel teljesül.

7. példa.

Állapítsuk meg, hogy van-e 5 minimális távolságú 13 hosszú perfekt bináris kód?

n=13, d=6, és $\left\lfloor \frac{d-1}{2} \right\rfloor = 2$ így a kód t=2-hibajavító.

Ha a közleményszavak k hosszúak, akkor a közleményszavak száma 2^k, ami megegyezik a kódszavak számával, $M= 2^k$.

Ha a kód perfekt, akkor a Hamming-korlát egyenlőséggel teljesül:

$$2^k\left(\binom{13}{0}+\binom{13}{1}+\binom{13}{2}\right)=2^{13} \qquad 2^k\left(1+13+\frac{12\cdot 13}{2}\right)=2^{13}$$

$$2^k \cdot 92 = 2^{13}$$

A jobb oldalon 2 hatványa áll, a bal oldalon szereplő 92 azonban nem 2 hatványa, ez az egyenlőség nem teljesülhet semmilyen k esetén. Ezért ilyen kód nem létezik