Linearni kodovi

Petrušić, Doris

Sveučilište u Rijeci

Fakultet za matematiku

Preddiplomski studij Matematika

Doris Petrušić

LINEARNI KODOVI

Završni rad

Rijeka, rujan 2022.

Sveučilište u Rijeci

Fakultet za matematiku

Preddiplomski studij Matematika

Doris Petrušić

LINEARNI KODOVI

Mentorica: dr. sc. Andrea Švob

Završni rad

Rijeka, rujan 2022.

Sadržaj

Sažetak

U radu je opisan način na koji se prenose poruke, odnosno kodovi kroz komunikacijski kanal i pogreške do kojih dolazi tijekom njihovog slanja. Budući da živimo u svijetu prepunom raznih informacija, uveliko nam je potrebna učinkovita i točna metoda za njihov što bolji prijenos. Cilj je poslati poruku sa što manje pogrešaka te, ako do njih dođe, pronaći najučinkovitiju metodu za njihovo ispravljanje. Želimo konstruirati što bolje kodove, čije nam postojanje garantira Shannonov teorem. Shannon je otkrio broj imenom kapacitet kanala i dokazao da je proizvoljno pouzdana komunikacija moguća pri bilo kojoj vrijednosti ispod kapaciteta kanala. Od svih vrsta kodova najviše se proučavaju linearni kodovi, koji su glavna tema ovog rada. Zbog njihove algebarske strukture, lakše ih je opisati, kodirati i dekodirati od nelinearnih kodova. Opisujemo ih pomoću generirajuće matrice i matrice provjere pariteta o kojima će se također govoriti u radu. Tri vrlo česta polja u istraživanjima linearnih kodova su binarno polje s dva elementa, ternarno polje s tri elementa i polje s četiri elementa. Posebna važnost posvećena je binarnim kodovima te je naveden binarni Hammingov kod i definirani pojmovi dualnih kodova, Hammingove udaljenosti i Hammingove težine. Na kraju je opisan postupak kodiranja i dekodiranja Hammingovim kodom.

Ključne riječi

Komunikacijski kanal, binarno polje, ternarno polje i polje s četiri elementa, linearni kodovi, generirajuće matrice, matrice provjere pariteta, dualni kodovi, Hammingova udaljenost, Hammingova težina, Hammingov kod.

1 Uvod

Teorija kodiranja bavi se dizajniranjem kodova za ispravljanje grešaka koji omogućuju pouzdan prijenos podataka kroz zvučni kanal te proučavanjem njihovih svojstava. U radu će se proučavati linearni kodovi te će se opisati neka njihova svojstva.

U prvom poglavlju će se opisati osnovni koncepti linearnih kodova i važnost kodiranja poruke kako bi se smanjile pogreške prilikom prenošenja. U drugom poglavlju će se navesti definicija polja i dati neki primjeri konačnih polja, bit će govora o osnovnim pojmovima vezanim uz njih te će se dati primjeri binarnog, ternarnog i polja s četiri elementa. U trećem poglavlju navest će se definicija vektorskog prostora i neki načini pomoću kojih se opisuje linearni kod. Govorit će se o linearnim kodovima, generirajućim matricama i matricama provjere pariteta te Hammingovom kodu. U četvrtom poglavlju govorit će se o dualnim, odnosno ortogonalnim kodovima te definirati pojmovi samoortogonalnosti i samodualnosti. U petom poglavlju govorit će se o težini i udaljenosti, definirati pojmovi Hammingove udaljenosti, minimalne udaljenosti, Hammingove težine te distribucija težine, odnosno težinski spektar. U posljednjem poglavlju će se opisati Hammingovi kodovi te, kroz primjere, objasniti postupak kodiranja i dekodiranja pomoću njih.

2 Osnovni pojmovi vezani uz linearne kodove

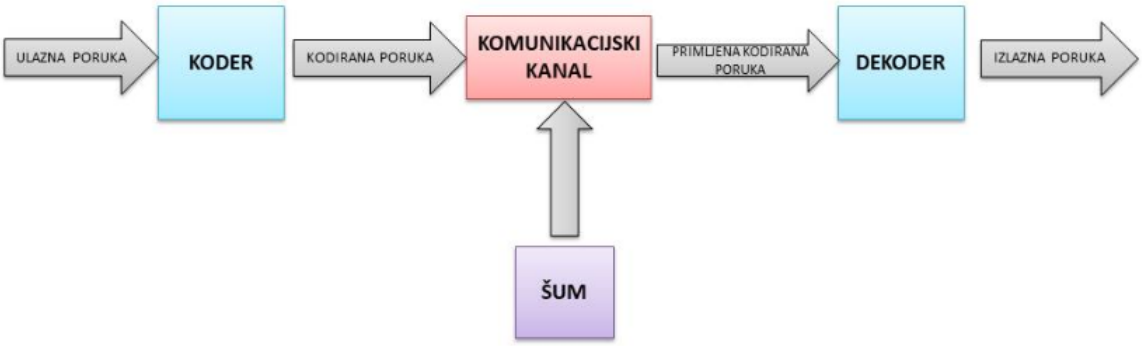
Claude Shannon objavio je 1948. značajan rad pod nazivom "Matematička teorija komunikacije" koji je označio početak teorije informacija i teorije kodiranja. S obzirom na to da komunikacijski kanal može iskvariti informacije poslane putem njega, Shannon je otkrio broj imenom kapacitet kanala i dokazao da je proizvoljno pouzdana komunikacija moguća pri bilo kojoj vrijednosti ispod kapaciteta kanala. Primjerice, prilikom prijenosa slika planeta iz dubokog svemira nepraktično je ponovno prenositi slike. Stoga, ako su dijelovi podataka koji tvore slike izmijenjeni zbog šuma koji nastaje u prijenosu, podaci se mogu pokazati beskorisnima. Shannonovi rezultati jamče da se podaci mogu kodirati prije prijenosa kako bi se izmijenjeni podaci mogli dekodirati do određenog stupnja točnosti. Primjeri ostalih komunikacijskih kanala uključuju magnetske uređaje za pohranu, kompaktne diskove i bilo koje vrste elektroničkih komunikacijskih uređaja poput mobilnih telefona.

Zajedničko obilježje komunikacijskih kanala je da informacije polaze od izvora i šalju se prijemniku preko kanala na drugoj strani. Na primjer, prilikom komunikacije s dubokim svemirom, izvor poruke je satelit, kanal je svemir zajedno s hardverom koji šalje i prima podatke, a prijamnik je zemaljska satelitska stanica (naravno, poruke putuju i sa Zemlje do satelita.) Kod kompaktnog diska poruka je glas, glazba ili podaci koji se prenose na disk, komunikacijski kanal je sam disk, a prijamnik je slušatelj. Komunikacijski kanal je "bučan" u smislu da ono što je primljeno nije uvijek jednako onome što je poslano. Stoga, ako se binarni podaci prenose preko komunikacijskog kanala, kad se pošalje 0, nadamo se primiti ju kao 0, ali ponekad ćemo ju primiti kao 1 (ili kao neprepoznatljivo). Buka u komunikaciji s dubokim svemirom može, primjerice, biti uzrokovana toplinskim smetnjama. Buku na kompaktnom disku mogu uzrokovati otisci prstiju ili ogrebotine diska. Osnovni problem u teoriji kodiranja je na temelju primljene poruke utvrditi što je poslano.

Teorija kodiranja je bazirana na sljedećem komunikacijskom modelu. Pošiljatelj želi poslati poruku primatelju. Poruke se šalju kroz komunikacijski kanal, koji nije savršen, pa može dodati grešku na originalnu poruku. Na primjer, dok se gleda televizijski program,

često se javljaju šumovi i slab prijem slike upravo zbog atmosferskih pojava.

Komunikacijski model izgleda ovako:



Slika 1: Komunikacijski sustav

Na slici je prikazano sljedeće. U koder ulazi ulazna poruka. U njemu se kodira i tako dobivamo kodiranu poruku. Kodirana poruka se šalje putem komunikacijskog kanala u kojemu se dodaje greška (šum). Dekoder prima izmijenjenu kodiranu poruku (primljena kodirana poruka) i dekodira ju kako bi se dobila izlazna poruka.

3 Primjeri konačnih polja

Najprije navedimo definiciju polja:

Definicija 3.1. *Neka je F neprazan skup na kojem su zadane dvije binarne operacije, zbrajanja ($+ : F \times F \to F$) i množenja ($\cdot : F \times F \to F$). Uređena trojka $(F, +, \cdot)$ je polje ako vrijede sljedeća svojstva:*

1. *$(F, +)$ je Abelova grupa*
2. *$(F\backslash\{0\}, \cdot)$ je Abelova grupa*
3. *Vrijedi distributivnost množenja prema zbrajanju:*
   $x \cdot (y + z) = x \cdot y + x \cdot z$, za sve $x, y, z \in F$.

Od svih vrsta kodova najviše se proučavaju linearni kodovi. Zbog njihove algebarske strukture lakše ih je opisati, kodirati i dekodirati od nelinearnih kodova. Kodna abeceda za linearne kodove je konačno polje iako se ponekad i druge algebarske strukture (poput cijelih brojeva modula 4) mogu koristiti za definiranje kodova koji se također nazivaju "linearnima".

Proučavat ćemo linearne kodove čija je abeceda polje F_q, tj. konačno polje s q elemenata koje ćemo označivati s $GF(q)$.

Definicija 3.2. *Neka je F konačan skup koji se sastoji od q elemenata. Skup F nazivamo **abecedom**, a njegove elemente **simbolima**. q-naran **kod** C duljine n nad poljem F je podskup $C \subseteq F^n$.*

Iako ćemo opće rezultate predstaviti preko proizvoljnih polja, najčešće ćemo se baviti poljima s dva, tri ili četiri elementa.

Tri vrlo česta polja u istraživanjima linearnih kodova su binarno polje s dva elementa, ternarno polje s tri elementa i polje s četiri elementa. S tim se poljima može raditi poznavanjem zbrajanja i tablice množenja, što ćemo pokazati u sljedeća tri primjera:

Primjer 3.1. *Binarno polje F_2 s dva elementa {0, 1} ima sljedeće tablice zbrajanja i množenja:*

+	0	1
0	0	1
1	1	1

.	0	1
0	0	0
1	0	1

Primjer 3.2. *Ternarno polje F_3 s tri elementa {0, 1, 2} ima sljedeće tablice zbrajanja i množenja:*

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$.$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Primjer 3.3. *Polje F_4 s četiri elementa {0, 1, x, y} složenije je. Ima sljedeće tablice zbrajanja i množenja:*

$+$	0	1	x	y
0	0	1	x	y
1	1	0	y	x
x	x	y	0	1
y	y	x	1	0

$.$	0	1	x	y
0	0	0	0	0
1	0	1	x	y
x	0	x	y	1
y	0	y	1	x

Možemo uočiti da su u ovim tablicama korištene neke jednadžbe. Na primjer, vidimo da je $a + a = 0$ za sve elemente a iz polja F_4. Također, vrijedi, $y = x^2 = y = 1 + x$.

4 Linearni kodovi, generirajuće matrice i matrice provjere pariteta

Da bismo definirali linearni kod, najprije moramo poznavati definicije polja i vektorskog prostora. U prethodnom poglavlju smo se prisjetili definicije polja pa nam samo preostaje navesti definiciju vekotorskog prostora.

Definicija 4.1. *Neka je V neprazan skup i F polje. Neka su zadane operacije $+ : V \times V \to V$ i $\cdot : F \times V \to V$ $\cdot$. Uredena trojka $(V, +, \cdot)$ se naziva vektorski prostor nad poljem F ako vrijedi:*

1. *$(V, +)$ Abelova grupa,*
2. *$(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$, za sve $\alpha, \beta \in F$, $a \in V$ (distributivnost operacije $\cdot$ u odnosu na zbrajanje u F),*
3. *$\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$, za sve $\alpha \in F$, $a, b \in V$, (distributivnost operacije $\cdot$ u odnosu na zbrajanje u V),*
4. *$\alpha \cdot (\beta \cdot a) = (\alpha\beta) \cdot a$, za sve $\alpha, \beta \in F$, $a \in V$, (kvaziasocijativnost),*
5. *$1 \cdot a = a$, za sve $a \in V$.*

Definicija 4.2. *Neka je F_q polje reda q. Tada je skup $F_q^n = \{(x_1, \ldots, x_n) | x_i \in F_q, 1 \leqslant i \leqslant n\}$ svih uredjenih n-torki s elementima iz F_q n-dimenzionalan vektorski prostor uz operacije zbrajanja $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ i množenja skalarom $\alpha(x_1, ..., x_n) = (\alpha x_1, \ldots, \alpha x_n)$, za sve $\alpha \in F_q$.*

Sada možemo navesti definiciju linearnog koda:

Definicija 4.3. *Neka je C k-dimenzionalan potprostor vektorskog prostora F_q^n. Tada se kod C zove $[n, k]$* ***linearni kod*** *nad F_q.*

Linearni kod C ima q^k kodnih riječi jer svaka baza linearnog $[n, k]$ koda C nad poljem F sadrži k kodnih riječi čije linearne kombinacije generiraju cijeli skup C, stoga vrijedi $|C| = q^k$.

Neka je F_q^n vektorski prostor svih n-torki nad konačnim poljem F_q. Zapis (n, M) označava kod C nad poljem F_q kao podskup od F_q^n veličine M. Vektore $(a_1, a_2, \ldots, a_n)$ u F_q^n obično zapisujemo $a_1 a_2 \ldots a_n$ i te vektore u C nazivamo kodnim riječima. Klasični primjer je polinomni prikaz koji se koristi za kodne riječi u cikličkim kodovima. Polje F_2 iz Primjera 3.1. imalo je vrlo posebno mjesto u povijesti teorije kodiranja te se kodovi nad F_2 nazivaju binarnim kodovima. Slično tome, kodovi nad F_3 nazivaju se ternarnim kodovima, dok se kodovi nad F_4 nazivaju kodovima s četiri elementa.

Bez dodavanja daljnjih struktura kodu, njegova upotrebljivost donekle je ograničena. Linearnost je najkorisnija dodatna struktura koja se može nametnuti.

Dva najčešća načina predstavljanja linearnog koda su pomoću generirajuće matrice i matrice provjere pariteta.

Definicija 4.4. ***Generirajuća matrica*** *za $[n, k]$ kod C je bilo koja matrica G reda $k \times n$ čiji redovi čine bazu (bazu vektorskog prostora) za kod C.*

Primjer 4.1. *Matrica*

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

jedna je od generirajućih matrica za $[3, 2]$ kod nad poljem F_2.

Općenito, postoji puno generirajućih matrica za kod. Za bilo koji skup k nezavisnih stupaca generirajuće matrice G, odgovarajući skup koordinata tvori informacijski skup (engl. information set) za C. Preostale $r = n - k$ koordinate nazvane su redundantnim skupom (engl. redundancy set) te se r naziva redundancijom od C. Ako prvih k koordinata čini informacijski skup, kod ima jedinstvenu generirajuću matricu oblika

$$\left[\; I_k \;\middle|\; A \;\right]$$

pri čemu je I_k jedinična matrica reda $k \times k$. Kažemo da je takva generirajuća matrica zapisana u **standardnom obliku**. Budući da je linearni kod potprostor vektorskog prostora, on je jezgra neke linearne transformacije.

Generirajuće matrice uvodimo zbog pojednostavljivanja kodiranja i dekodiranja, te zbog skraćivanja zapisa. Kod C je skup svih linearnih kombinacija redova matrice G.

Kod C dobivamo tako da množimo G sa svim $1 \times k$ vektorima redaka s lijeve strane i tako dobivamo sve linearne kombinacije, $C = \{\alpha G | \alpha \in F_q^n\}$.

Primjer 4.2. *Neka je C $[5,3]$ kod vektorskog prostora F_2^5 i G generirajuća matrica u standardnom obliku*

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

Ako pomnožimo G s osam različitih binarnih vektora redaka duljine tri, dobijemo osam kodnih riječi. Primjer jedne kodne riječi:
$(1,1,1)G = (1,1,1,0,1)$

Sve kodne riječi: $(0,0,0,0,0)$, $(1,0,0,1,1)$, $(0,1,0,0,1)$, $(0,0,1,1,1)$, $(1,1,0,1,0)$, $(1,0,1,0,0)$, $(0,0,0,1,1)$, $(1,1,1,0,1)$

Definicija 4.5. ***Matrica provjere pariteta*** *za $[n,k]$ kod C je $(n-k) \times n$ dimenzionalna matrica H, definirana izrazom*

$$C = \{x \in F_q^n | Hx^T = 0\}.$$

Redci od H će također biti nezavisni. Općenito, postoji također nekoliko mogućih matrica provjere pariteta za C. Sljedeći teorem daje jednu od njih kada C ima generirajuću matricu u standardnom obliku. U ovom je teoremu A^T transponirana matrica matrice A.

Teorem 4.1. *Ako je*

$$G = \left[\; I_k \;\middle|\; A \; \right]$$

$[n,k]$ koda C zapisana u standardnom obliku, onda je

$$H = \left[\; -A^T \;\middle|\; I_{n-k} \; \right]$$

matrica provjere pariteta za C.

Dokaz: Jasno je da imamo $HG^T = -A^T + A^T = O$. Dakle, C je sadržan u jezgri linearne transformacije $x \mapsto Hx^T$. Kako je matrica H ranga $n-k$, ova linearna transformacija ima jezgru dimenzije k koja je ujedno i dimenzija od C. □

Primjer 4.3. *Najjednostavniji način kodiranja informacija sa svrhom njihova oporavka u prisutnosti buke je ponavljanje svakog simbola poruke određeni broj puta. Pretpostavimo da je naša informacija binarna sa simbolima iz polja F_2, a svaki simbol ponavljamo n puta. Ako je, na primjer, $n = 7$, onda kad god želimo poslati 0, šaljemo 0000000 i kad god želimo poslati 1, šaljemo 1111111. Ako se u prijenosu naprave najviše tri pogreške i ako dekodiramo prema "većini glasova", možemo točno utvrditi informacijski simbol, 0 ili 1. Općenito, naš kod C je $[n, 1]$ binarni linearni kod koji se sastoji od dvije kodne riječi $0 = 00\cdots0$ i $1 = 11\cdots1$ te se naziva binarnim kodom ponavljanja duljine n. Ovaj kod može ispraviti do $e = \lfloor (n-1)/2 \rfloor$ pogrešaka: ako se u primljenom vektoru napravi najviše e pogrešaka, većina će koordinata biti točna, a time će se moći oporaviti i izvorno poslanu kodnu riječ. Ako je napravljeno više od e pogrešaka, te se pogreške ne mogu ispraviti. Međutim, ovaj kod može otkriti $n-1$ pogrešaka pošto primljeni vektori s pogreškama između 1 i $n-1$ definitivno neće biti kodne riječi. Generirajuća matrica za kod ponavljanja je:*

$$G = \left[\begin{array}{c|c} 1 & 1\cdots1 \end{array}\right]$$

naravno, zapisana u standardnom obliku. Odgovarajuća matrica provjere pariteta iz Teorema 4.1 je:

$$H = \left[\begin{array}{c|c} 1 & \\ 1 & \\ \vdots & I_{n-1} \\ 1 & \end{array}\right]$$

Prva koordinata je informacijski skup, a posljednje $n-1$ koordinate čine redundantni skup.

Primjer 4.4. *Matrica*

$$G = \left[\begin{array}{c|c} I_4 & A \end{array} \right]$$

gdje je

$$G = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]$$

je generirajuća matrica zapisana u standardnom obliku za binarni kod $[7, 4]$ *koji označavamo kao* H_3*. Prema Teoremu 4.1, matrica provjere pariteta za* H_3 *je*

$$H = \left[\begin{array}{c|c} A^T & I_3 \end{array} \right] = \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

Ovaj kod se zove Hammingov kod $[7, 4]$

U ovome tekstu ćemo se često pozivati na potkod koda C. Ako C nije linearan kod (ili nije poznato je li linearan), potkod C je bilo koji podskup C. Ako je C linearan, potkod će biti podskup C koji također mora biti linearan; u ovom slučaju potkod C je podskup C.

5 Dualni kodovi

Generirajuća matrica G $[n, k]$ koda C je, jednostavno rečeno, matrica čiji su redci nezavisni te obuhvaćaju kod. Redci matrice provjere pariteta H su nezavisni: stoga je H generirajuća matrica nekog koda, koji se naziva **dualnim ili ortogonalnim kodom** koda C te se označava simbolom $C^{\perp}$.

Primijetimo da je $C^{\perp}$ $[n, n-k]$ kod.

Drugi način definiranja dualnog koda je korištenjem skalarnih umnožaka.

Prisjetimo se da je običan skalarni umnožak vektora $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$ u F_q^n jednak

$$\langle x \cdot y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Ovdje se zapravo radi o simetričnoj bilinearnoj formi te prema tome vrijedi da je $\langle x \cdot y \rangle = \langle y \cdot x \rangle$, $\langle \alpha x + \beta x_0, y \rangle = \alpha \langle x, y \rangle + \beta \langle x_0, y \rangle$ i $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$.

Definicija 5.1. *Bilinearna forma na F_q^n je funkcija $B : F_q^n \times F_q^n \to F_q$ koja je linearna u obje koordinate, to jest takva da za sve $x, y, z \in F_q^n$ i $\alpha, \beta \in F_q$ vrijedi*

1. $B(\alpha x + \beta y, z) = \alpha B(x, z) + \beta B(y, z)$
2. $B(x, \alpha y + \beta z) = \alpha B(x, y) + \beta B(x, z)$

Bilinearna forma B je simetrična ako za sve $x, y \in F_q^n$ vrijedi $B(x, y) = B(y, x)$.

Ortogonalni komplement od C (skup svih vektora koji su okomiti na sve vektore iz C) kojeg označavamo s $C^\perp$ je vektorski potprostor.

Prema tome vidimo da se $C^\perp$ može definirati i izrazom

$$C^\perp = \{x \in F_n^q | \langle x, c \rangle = 0 \text{ za sve } c \in C\}.$$

Primjer 5.1. *Generirajuća matrica i matrica provjere pariteta za $[n, 1]$ kod ponavljanja C dane su u Primjeru 4.3.. Dualni kod $C^\perp$ je $[n, n-1]$ kod s generirajućom matricom H te se stoga sastoji od svih binarnih n- torki $a_1 a_2 \ldots a_{n-1} b$, gdje je $b = a_1 + a_2 + \cdots + a_{n-1}$ (zbrajanje u polju F_2). N-ta koordinata b je ukupna provjera pariteta za prvih $n-1$ odabranih koordinata, dakle, zbroj svih koordinata iznosi 0. Zbog toga je lako uočiti da je G doista matrica provjere pariteta za kod $C^\perp$. Kod $C^\perp$ ima svojstvo da se može otkriti jedna pogreška u prijenosu (s obzirom na to da zbroj svih koordinata neće biti 0), ali se ona ne može ispraviti (jer bi promjena bilo koje od primljenih koordinata dala vektor čiji bi zbroj koordinata bio 0).*

Definicija 5.2. *Kažemo da je kod C **samoortogonalan** ako je $C \subseteq C^\perp$, a **samodualan** ako je $C = C^\perp$. Duljina n samodualnog koda je uvijek paran broj, a dimenzija je $\frac{n}{2}$.*

6 Težine i udaljenosti

Važna konstanta koda je **minimalna udaljenost** između kodnih riječi.

Definicija 6.1. ***(Hammingova) udaljenost*** $d(x, y)$ *između dvaju vektora* $x, y \in F_q^n$ *je broj koordinata u kojima se* x *i* y *razlikuju. Odnosno,* $d(x, y) = |\{i | 1 \leqslant i \leqslant n, x_i \neq yi\}|$,

Teorem 6.1. *Funkcija udaljenosti* $d(x, y)$ *zadovoljava sljedeća četiri svojstva:*

1. $d(x, y) \geqslant 0$ *za svaki* $x, y \in F_q^n$
2. $d(x, y) = 0$ *ako i samo ako je* $x = y$
3. $d(x, y) = d(y, x)$ *za svaki* $x, y \in F_q^n$
4. $d(x, z) \leqslant d(x, y) + d(y, z)$ *za svaki* $x, y, z \in F_q^n$

Dokaz:

1. Neka su $x, y \in F_q^n$ proizvoljni. Budući da je $d(x, y) = |\{i | 1 \leqslant i \leqslant n, x_i \neq yi\}|$ slijedi da je $d(x, y) \geqslant 0$.
2. Neka su $x, y \in F_q^n$ proizvoljni. Pretpostavimo da je $d(x, y) = 0$, to znači da se x i y ne razlikuju ni u jednoj koordinati, prema tome je $x = y$. Pretpostavimo da je $x = y$ prema tome su x i y jednaki, tj. ne razlikuju se ni u jednoj koordinati pa je $d(x, y) = 0$.
3. Neka su $x, y \in F_q^n$ proizvoljni. Ako je $x = y$, onda je i $y = x$ pa je prema tome i $d(x, y) = d(y, x)$
4. Neka su $x, y, z \in F_q^n$ proizvoljni. Ako su x i z jednaki slijedi $0 \leqslant d(x, y) + d(y, z)$. Ako su x i z različiti i $x = y$ ili $z = y$ slijedi $d(x, z) = d(x, y) + d(y, z)$. Te ako su svi različiti, vrijedi $d(x, z) \leqslant d(x, y) + d(y, z)$.

□

Prema ovom teoremu (F_q^n, d) je metrički prostor.

Definicija 6.2. ***(Minimalna) udaljenost koda*** C *je najmanja udaljenost između različitih kodnih riječi iz* C. *Odnosno,* $min\{d(x, y) | x \in C, y \in C, x \neq y\}$.

Udaljenost koda važna za određivanje sposobnosti ispravljanja pogrešaka koda C. Što je veća minimalna udaljenost, kod može ispraviti više pogrešaka.

Definicija 6.3. ***(Hammingova) težina*** *$wt(x)$ vektora $x \in F_q^n$ udaljenost od x do 0, gdje je $0 = (0, 0, \ldots, 0)$. Odnosno, $w(x) := d(x, 0)$.*

Definicija 6.4. ***(Minimalna) težina koda*** *C je najmanja udaljenost između različitih kodnih riječi iz C i 0. Odnosno, $min\{w(x)|x \in C, x \neq 0\}$.*

(Hammingova) težina $wt(x)$ vektora $x \in F_q^n$ može se definiratu i kao broj koordinata koje nisu nula u x.

Sljedeći teorem govori o odnosu između udaljenosti i težine koda:

Teorem 6.2. *Ako je $x, y \in F_q^n$, onda je $d(x, y) = wt(x - y)$. Ako je C linearni kod, minimalna udaljenost jednaka je minimalnoj težini kodnih riječi od C koje nisu nula.*

Dokaz: Neka su $x, y \in C$. Budući da je C vektorski postor, slijedi da je $x - y \in C$. Sada imamo $d(x, y) = d(x + y, 0)$ što je težina od $x - y$. Zaključujemo da su udaljenosti kodnih riječi težina neke kodne riječi, prema tome je najmanja udaljenost jednaka najmanjoj težini. □

Kao rezultat ovog teorema minimalna udaljenost kod linearnih kodova također se naziva i minimalnom težinom koda. Ako je poznata minimalna težina d koda $[n, k]$ onda taj kod navodimo $[n, k, d]$ kod.

Primjer 6.1. *Prisjetimo se Primjera 4.2. gdje je zadan C $[5, 3]$ kod vektorskog prostora F_2^5 i G generirajuća matrica u standardnom obliku*

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

Sve kodne riječi su bile: $(0, 0, 0, 0, 0)$, $(1, 0, 0, 1, 1)$, $(0, 1, 0, 0, 1)$, $(0, 0, 1, 1, 1)$, $(1, 1, 0, 1, 0)$, $(1, 0, 1, 0, 0)$, $(0, 0, 0, 1, 1)$, $(1, 1, 1, 0, 1)$.

Vidimo da su tri kodne riječi težine dva, tri kodne riječi težine tri, jedna kodna riječ težine četiri i jedna težine nula.. Budući da je ovo linearni kod, najmanja udaljenost ovog koda je dva. Prema tome ovo je $[5, 3, 2]$ *kod.*

Kada se bavimo kodovima nad poljima F_2, F_3 ili F_4, postoje određeni korisni osnovni rezultati o težinama kodnih riječi:

Napomena 6.1. *Vrijede sljedeće tvrdnje:*

1. *Ako su* $x, y \in F_2^n$, *onda je* $wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y)$, *gdje je* $x \cap y$ *vektor u* F_2^n *koji ima jedinice (1) točno u onim pozicijama gdje i* x *i* y *imaju jedinice (1).*

2. *Ako su* $x, y \in F_2^n$, *onda je* $wt(x \cap y) \equiv x \cdot y (mod 2)$.

3. *Ako je* $x \in F_2^n$, *onda je* $wt(x) \equiv x \cdot x (mod 2)$

4. *Ako je* $x \in F_3^n$, *onda je* $wt(x) \equiv x \cdot x (mod 3)$

5. *Ako je* $x \in F_4^n$, *onda je* $wt(x) \equiv \langle x, x \rangle (mod 2)$.

Definicija 6.5. *Neka je* A_i, *također označen kao* $Ai(C)$, *broj kodnih riječi težine* i *u* C. A_i *za* $0 \leqslant i \leqslant n$ *naziva se* ***distribucijom težine ili težinskim spektrom*** *koda* C.

Puno istraživanja posvećeno je izračunu distribucije težine pojedinih kodova ili skupina kodova.

Primjer 6.2. *Neka je* C *binarni kod s generatorskom matricom*

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

Uzmimo sve linearne kombinacije redaka kako bismo napisali osam kodnih riječi koda C:

kodna riječ	težina	kodna riječ	težina
(0,0,0,0,0,0)	0	(1,1,1,1,0,0)	4
(0,0,0,0,1,1)	2	(0,0,1,1,1,1)	4
(0,0,1,1,0,0)	2	(1,1,0,0,1,1)	4
(1,1,0,0,0,0)	2	(1,1,1,1,1,1)	6

Distribucija težine koda C je:
$A_0 = A_6 = 1$*, odnosno, ima jedna kodna riječ težine nula i jedna kodna riječ težine šest,*
$A_2 = A_4 = 3$*, odnosno, imaju tri kodne riječi težine dva i tri kodne riječi težine četiri*
Primijetimo da su navedeni samo A_i koji nisu nula.

7 Hammingov kod

Definicija 7.1. *Hammingov kod je linearni kod čija matrica provjere pariteta H ima r redaka te u stupcima ima sve mogućnosti vektora dimenzija $r > 1$, osim nul-vektora. Za $r \geqslant 2$ Hammingov kod je linearni $[2^r - 1, 2^r - 1 - r]$ kod.*

Hammingov kod omogućuje nam otkrivanje jednostruke pogreške, mjesta na kojem se pogreška dogodila te njeno ispravljanje.

Sastoji se od podatkovnih bitova (oznaka P) i zaštitnih bitova (oznaka Z).

Zaštitni bitovi postavljaju se na mjesta u kodnoj riječi koja su potencija broja dva (prvo, drugo, četvrto …mjesto). Broj zaštitnih bitova ovisi o broju podatkovnih.

Prije nego što počnemo kodirati i dekodirati pomoću Hammingovog koda spomenut ćemo Metodu pariteta. Metoda pariteta koristi se za otklanjanje pogreške jedne neispravne znamenke. Razlikujemo metodu parnog pariteta (svaka kodna riječ ima paran broj logičkih jedinica) i metodu neparnog pariteta (svaka kodna riječ ima neparan broj logičkih jedinica).

Način kodiranja i dekodiranja Hammingovim kodom bit će prikazan u sljedećim primjerima:

Primjer 7.1. *Kodiraj podatak* $1010_{(2)}$ *Hammingovim kodom.*

Na određene položaje stavimo podatkovne i zaštitne bitove. Na mjesto podatkovnih bitova upišemo podatak koji želimo kodirati:

	2^0	2^1		2^2			
Položaj:	1	2	3	4	5	6	7
Uloga (podatkovni/zaštitni):	Z_1	Z_2	P_1	Z_3	P_2	P_3	P_4
Kod:	?	?	1	?	0	1	0

Preostale dijelove koda, odnosno zaštitne podatke Z_1, Z_2, Z_3 *dobijemo tako da:*

- Z_1 *dobijemo tako da promatramo sljedeće dijelove koda:*

Kod:	?	?	1	?	0	1	0

iz čega dobivamo ?100 pa zaključujemo da nam na mjesto ? dolazi 1 prema metodi parnog pariteta (metoda koja nam kaže da svaka kodna riječ mora imati paran boj jedinica). Dakle $Z_1 = 1$.

- Z_2 *dobijemo tako da promatramo sljedeće dijelove koda:*

Kod:	?	?	1	?	0	1	0

iz čega dobivamo ?110 pa zaključujemo da nam na mjesto ? dolazi 0.

Dakle $Z_2 = 0$.

- *Z_3 dobijemo tako da promatramo sljedeće dijelove koda:*

Kod:	?	?	1	?	0	1	0

iz čega dobivamo ?010 pa zaključujemo da nam na mjesto ? dolazi 1.

Dakle $Z_3 = 1$.

Na kraju tablica izgleda ovako:

	2^0	2^1		2^2			
Položaj:	1	2	3	4	5	6	7
Uloga (podatkovni/zaštitni):	Z_1	Z_2	P_1	Z_3	P_2	P_3	P_4
Kod:	?	?	1	?	0	1	0
	1	0	1	1	0	1	0

Hammingov kod za zadani podatak $1010_{(2)}$ *je* $1011010_{(Hamming)}$.

Primjer 7.2. *Provjeri je li u poruci 10111100 kodiranoj Hammingovim kodom došlo do pogreške te, ako jeste, ispravi ju. Koji podatak je poslan prije kodiranja?*

Da bismo otkrili je li došlo do pogreške, izračunat ćemo zaštitne bitove na način kao u prethodnom primjeru. Najprije konstruirajmo tablicu:

	2^0	2^1		2^2			
Položaj:	1	2	3	4	5	6	7
Uloga (podatkovni/zaštitni):	Z_1	Z_2	P_1	Z_3	P_2	P_3	P_4
Kod:	1	0	1	1	1	1	0

Nakon provođenja postupka iz prethodnog primjera dobivamo da je $Z_1 = Z_2 = Z_3 = 0$ *iz čega zaključujemo da se zaštitni podatci ne podudaraju na prvom i četvrtom položaju. Mjesto pogreške u kodu dobijemo tako da zbrojimo položajna mjesta na kojima nam se zaštitni podatci ne podudaraju (*$1 + 4 = 5$*) iz čega dobivamo da se pogreška dogodila na petom položaju. Prema tome ispravljena kodna riječ glasi:* $1011010_{(Hamming)}$.

Ono što nam još preostaje otkriti je koji podatak je bio poslan prije kodiranja. To lako iščitamo iz tablice (nakon ispravljanja) tako da pročitamo podatkovne bitove. Prema tome poslani podatak bio je $1010_{(2)}$.

8 Zaključak

U radu je dan kratak uvod u teoriju kodiranja i linearne kodove. Ukratko je naveden problem kodiranja i opisan komunikacijski sustav. Kako bi se informacije što bolje prenosile, potrebno je konstruirati što bolje kodove. Linearne kodove o kojima se govorilo u radu zanimljivi su zbog svoje algebarske strukture zbog koje ih je lakše opisati, kodirati i dekodirati od nelinearnih kodova. Opisani su pomoću generirajuće matrice i matrice provjere pariteta.

Popis slika

Literatura

[1] W. C. Huffman, V. Pless: *Fundamentals of error correcting codes*, Cambridge University Press, 2003.

[2] R. Hill: *A First Course in Coding Theory*, Oxford University Press Inc., New York, 2004.

[3] J. H. van Lint: *Introduction to Coding Theory*, Springer-Verlag Berlin Heidelberg, 1992.

[4] I. S. Pandžić, A. Bažant, Z. Ilič, Z. Vrdoljak, M. Kos, V. Sinković: *Uvod u teoriju informacije i kodiranje*, Element, Zagreb, 2007.