Kodiranje koje ispravlja greške

Gaurina, Šime

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Šime Gaurina

KODIRANJE KOJE ISPRAVLJA GREŠKE

Diplomski rad

Voditelj rada:
prof. dr. sc. Filip Najman

Zagreb, veljača, 2023.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

Uvod

U razmjeni poruka lako može doći do pogrešaka. Neke pogreške mogu biti jako problematične. Primjera radi, pilot u avionu dobija dva tipa poruka, poruku 0 i poruku 1, pri čemu poruka 1 označava da može sletjeti, a 0 da ne može. Ukoliko se dogodi pogreška, npr. 0 se u komunikacijskom kanalu pretvori u 1, posljedice mogu biti katastrofalne. Zbog toga se uvode razne metode kako neku poruku enkodirati tako da se originalna poruka može iščitati čak i ako se dogodi pogreška.

Sam čovjek je u stanju ispraviti neku poruku ukoliko broj pogrešaka nije velik. Primjera radi, neka primatelj dobije poruku (preko nekog sredstva komunikaciji, u pismu ili u usmenom obliku) "DAPLOMSKA RUD". Postoji velika vjerojatnost da će primatelj uspjeti otkriti pravo značenje, tj. primatelj će poruku uspjeti dekodirati. U ovom slučaju ispravljena poruka je "DIPLOMSKI RAD". Kodiranje koje ispravlja greške funkcionira tako da se poruka koju treba poslati proširi sa tzv. redundantnim podatcima, odnosno podatcima koji omogućuju ispravljanje poruke.
Ovaj diplomski rad dobrim dijelom prati knjigu [2]. U prvom poglavlju daje se formalna definicija koda, odnosno kodiranja koja ispravljaju greške.

U drugom poglavlju prvo je objašnjeno jedno kodiranje, ponajviše zato da se da uvid kako kodiranja funkcioniraju i da se na primjeru pokažu pojmovi koji se definiraju u ovom radu. Nakon toga je opisan jedan jako važan pojam kod kodova koji ispravljaju greške: Hammingova udaljenost.

U trećem poglavlju se opisuju blok kodovi, koji predstavlju jednu od dvije klase kodova koji ispravljaju greške. U istom poglavlju je ponešto rečeno i o broju riječi koje neki kod sadrži. Naglasilo se da je poželjno da kod sadrži što više riječi. Opisali su se i savršeni kodovi. Oni su, pojednostavljeno, kodovi koji imaju zadovoljavajući broj riječi. O savršenim kodovima može se više pročitati u članku [6]. Rečeno je ponešto i o ekvivalentnim kodovima.

Tema četvrtog poglavlja, koji dobrim dijelom prati knjige [2] i [4], su linearni kodovi. Linearni kodovi su, pojednostavljeno, skupovi koji se mogu poistovjetiti sa nekim vektorskim prostorom. Dane su definicije dviju važnih matrica kod linearnih kodova: generirajuće matrice i matrice provjere pariteta. U istom poglavlju se opisuju i dvije metode dekodiranja, odnosno ispravljanja poruke: dekodiranje pomoću standardnog niza i sin-

dromsko dekodiranje.

I na kraju, u petom poglavlju je prvo formalizirana Hammingova familija kodova. Zatim su spomenuti simpleks kodovi, sa zanimljivim svojstvom ekvidistantnosti. Potom je, sa svrhom odmaka od Hammingovih kodova, opisano je još jedno kodiranje, koje se definira koristeći Hadamard-Sylvesterovu matricu, odnosno Hadamardovo kodiranje. Dijelovi petog poglavlja, točnije dio posvećen Hadamardovom kodu, prati članak [3].

Diplomski rad napravljen je u sklopu aktivnosti Projekta KK.01.1.1.01.0004 - Znanstveni centar izvrsnosti za kvantne i kompleksne sustave te reprezentacije Liejevih algebri.

Poglavlje 1

Osnovno o kodovima koji ispravljaju greške

Tijekom komunikacije izmedu pošiljatelja i primatelja lako može doći do grešaka. Do pogreške najčešće dolazi u komunikacijskom kanalu, koji originalnu poruku pogrešno prenese primatelju. U tu svrhu javlja se potreba enkodiranja poruke, tj. zapisivanje poruke na način da ukoliko i dođe do pogreške primatelj može redundacijom otkriti polaznu poruku. Pretpostavlja se sljedeće:

1. Greške su neovisne jedna o drugoj, tj. ukoliko dođe do greške na jednom mjestu, o tome ne ovisi hoće li se dogoditi greška na drugom mjestu.

2. Jednaka je vjerojatnost da se znak a pretvori u znak b i da se znak b pretvori u znak a, pri čemu a i b označavaju neki znak poruke.

Definicija 1.0.1. *Kod koji ispravlja greške je kod koji se koristi u telekomunikacijama i računarstvu, sa svrhom da se originalna poruka može razaznati iako dođe do pogreške u prijenosu.*

U idućem primjeru opisuje se jedna jednostavna metoda enkodiranja poruke koja je vjerodostojna uz pretpostavku da je najviše jedna znamenka krivo poslana.

Primjer 1.0.2. *Potrebno je poslati poruku koja je jednaka ili* 0 *ili* 1*. Evidentno je da su greške u ovakvim porukama nepoželjnje jer ukoliko primatelj umjesto* 0 *dobije* 1 *nemoguće je razaznati je li došlo do pogreške ili je to bila originalna poruka. Stoga se uvodi jednostavno enkodiranje koje iako duljinu poruke koju je potrebno poslati udvostruči, omogućava mehanizam za detekciju pogreške. Kodiranje se provodi na sljedeći način:*

- *Ukoliko je potrebno poslati* 0 *poruka se enkodira u* 00*.*

- *Ukoliko je potrebno poslati* 1 *poruka se enkodira u* 11.

Na ovaj način lako se zaključi ukoliko je primljena poruka oblika 10 *ili* 01 *da je došlo do pogreške u prijenosu poruke. Ukoliko su obje znamenke primljene poruke jednake i uz uvjet da može doći do najviše jedne pogreške tijekom prijenosa poruke, primatelj je siguran da nije došlo do pogreške tijekom prijenosa.*

Veliki nedostatak kodiranja u primjeru iznad je ta da se duljina poruke udvostručuje. S druge strane pretpostavka da može doći do najviše jedne pogreške nije pretjerano zahtjevna. Naime, od bilo kakvog ozbiljnog kanala komunikacije se očekuje da vjerojatnost slanja jednog krivog podatka ne bude velika, pa samim tim bi vjerojatnost slanja dva kriva podatka trebala biti još znatno manja.

Napomena 1.0.3. *U ostatku rada će se poruka i riječ smatrati sinonimima. Obje predstavljaju neki niz znakova.*

Napomena 1.0.4. *Poruka može biti sastavljena od npr. slova, brojeva ili bitova. Da se ne izgubi na općenitosti u ovom radu će se dijelovi od kojih je sastavljena poruka nazivati znakovima.*

Definicija 1.0.5. *Duljina riječi je broj znakova od kojih je riječ sastavljena.*

Definicija 1.0.6. *Kodna riječ je riječ nastala enkodiranjem.*

Definicija 1.0.7. *Kod je skup svih kodnih riječi.*

Definicija 1.0.8. *Kod C je δ-detektirajući ako može detektirati δ krivo poslanih znakova.*

Definicija 1.0.9. *Kod C je δ-ispravljajući ako može detektirati i ispraviti δ krivo poslanih znakova.*

Poglavlje 2

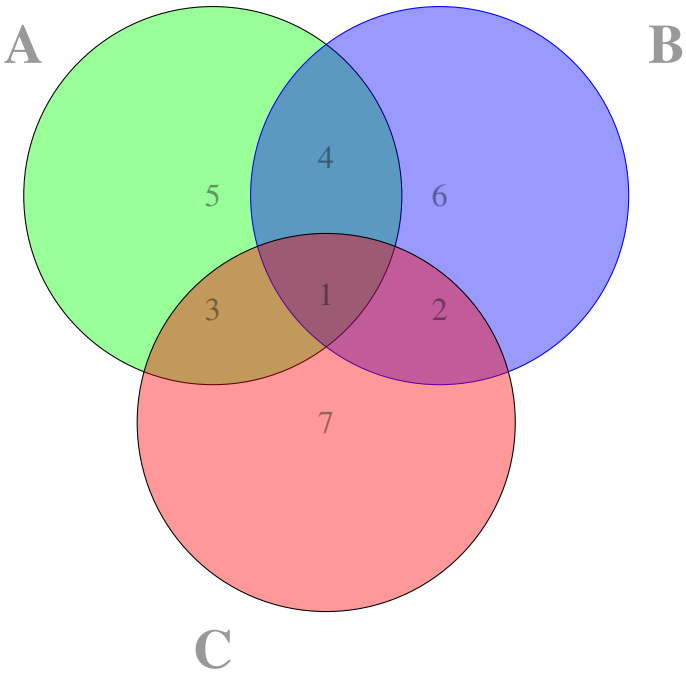
Hammingov (7,4) kod i Hammingova udaljenost

2.1 Hammingov (7,4) kod

Definicija 2.1.1. *Hamming (7,4) kod je kod za ispravljanje grešaka u prijenosu poruka koji riječ od* 4 *znaka enkodira u riječ od* 7 *znakova.*

Iz definicije se vidi da je potrebno dodati 3 redundantna znaka na poruku duljine 4, pa je broj znakova uvećan za 75%, dok je u kodiranju u primjeru (1.0.2) broj znakova uvećan za 100%. Također, kod osim detekcije omogućuje i ispravljanje podataka uz uvjet da je najviše jedan znak pogrešno poslan. Postupak enkodiranja riječi provodi se pomoću Vennova dijagrama sa slike (2.1) na sljedeći način:

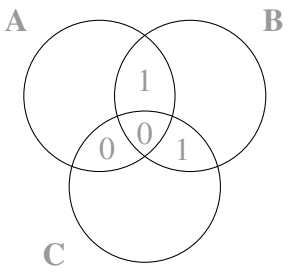
1. Četiri znaka riječi stavljaju se po redu u polja od 1 do 4.
2. U preostala 3 polja se dodaje po jedan znak tako da broj jedinica u svakom od krugova A, B i C bude paran.



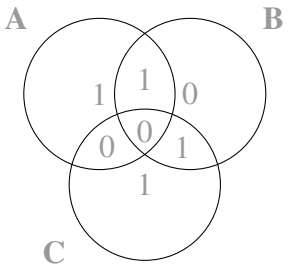
Slika 2.1: Vennov dijagram

Primjer 2.1.2. *Neka je poruka od* 4 *znaka koju je potrebno poslati jednaka* 0101. *Četiri znaka poruke prvo se po redu stavljaju u polja od* 1 *do* 4. *Pripadni Vennov dijagram je prikazan na slici* (2.2).

Nakon toga, za svaki od krugova A, B i C treba provjeriti je li broj znakova jednakih 1 *paran. U krugu A jedan znak ima vrijednost* 1, *dakle u polje* 5 *treba dodati znak vrijednosti* 1. *U krugu B postoje dva znaka vrijednosti* 1, *iz čega slijedi da u polje* 6 *treba dodati znak vrijednosti* 0. *U krugu C jedan znak ima vrijednost* 1, *dakle u polje* 7 *treba dodati znak vrijednosti* 1. *Pripadni Vennov dijagram je prikazan na slici* (2.3).



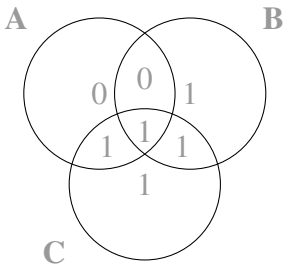
Slika 2.2: Prva četiri bita poruke 0101



Slika 2.3: Endodirana poruka 0101

Sljedeći primjer opisuje kako primatelj može uočiti pogrešku u primljenoj poruci.

Primjer 2.1.3. *Neka je primatelj primio poruku* 1110011. *Primatelj redom u polja od* 1 *do* 7 *upisuje poruku i nakon toga radi provjeru pariteta, tj. za svaki od krugova A, B i C provjerava je li broj znakova vrijednosti* 1 *paran. Pripadni Vennov dijagram je prikazan na slici* (2.4). *Pošto je u krugu B broj znakova vrijednosti* 1 *jednak* 3 *došlo je do pogreške u prijenosu podataka.*



Slika 2.4: Vennov dijagram za poruku 1110011

Dekodiranje poruke označava proces kojim primatelj pokušava dokučiti originalnu poruku, tj. ispraviti poruku koja je krivo poslana.

Postupak dekodiranja poruke uz pretpostavku da se dogodila najviše jedna pogreška također se provodi pomoću Vennova dijagrama sa slike (2.1) na sljedeći način:

1. Znakovi se po redu stavljaju u polja od 1 do 7.

2. Zatim se radi provjera pariteta za krugove A, B i C.

Napomena 2.1.4. *Provjera pariteta za krug unutar Vennova dijagrama se odvija tako da se provjeri je li u krugu paran broj znakova sa vrijednosti* 1.

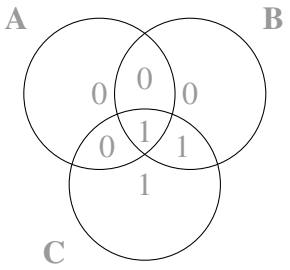
Rezultati provjere pariteta za krugove unutar Vennova dijagrama generiraju sljedeće zaključke:

- Ukoliko nije otkrivena pogreška poruka je ispravna.
- Ukoliko provjera nije točna za samo jedan od krugova A, B ili C, do greške je došlo na pozicijima 5, 6 ili 7 respektivno.
- Ukoliko provjera nije točna za krugove B i C ili A i C ili A i B, do greške je došlo na pozicijama 2, 3 ili 4 respektivno.
- Ukoliko je provjera pogrešna za svaki od krugova do greške je došlo na poziciji 1.

Primjer 2.1.5. *Neka je primatelj primio poruku* 1100001. *Pripadni Vennov dijagram je prikazan na slici* (2.5). *Rezulati provjere pariteta su sljedeći:*

- *Provjera pariteta za krug A nije točna.*
- *Provjera pariteta za krug B je točna.*
- *Provjera pariteta za krug C nije točna.*

Dakle, pogreška je na poziciji 3*, odnosno ispravna poruka je* 1110001.



Slika 2.5: Vennov dijagram za poruku 1100001

Uz pretpostavku o najviše jednoj pogrešci Hammingov $(7, 4)$ kod radi konkretno, no čim se ta pretpostavka otkloni može doći do pogrešnog zaključka na kojem mjestu se dogodila pogreška.

Primjer 2.1.6. *Neka je poruka koju je potrebno poslati* 1111. *Enkodirana poruka bi trebala biti* 1111111. *No tijekom slanja dolazi do pogreške i primatelj primi poruku* 1111010. *Koristeći se koracima opisanima iznad može se zaključiti da je do pogreške došlo na trećem znaku tj. da je originalna poruka trebala biti* 1101010.

2.2 Hammingova udaljenost

Definicija 2.2.1. *Neka su ω i ω' kodne riječi. Hammingova udaljenost, u oznaci $d(\omega, \omega')$, je broj mjesta na kojima se riječi razlikuju.*

Definicija 2.2.2. *Neka je C kod i neka za ω, $\omega' \in C$, $\omega \neq \omega'$, vrijedi da im je Hammingova udaljenost jednaka δ i ne postoje druge dvije različite kodne riječi u C sa manjom udaljenosti, tada se δ, u oznaci $d(C)$, naziva minimalna udaljenost koda C.*

Propozicija 2.2.3 (Nejednakost trokuta kod Hammingove udaljenosti)**.** *Neka su x, y i z neke riječi jednake duljine nad istim alfabetom. Vrijedi sljedeća nejednakost:*

$$d(x, y) \leq d(x, z) + d(z, y) \tag{2.1}$$

Dokaz. Neka se riječi x i y razlikuju na najmanje jednom položaju, inače je tvrdnja trivijalna jer vrijedi $d(x, y) = 0$. Neka je a neki znak riječi x, a b neki znak riječi y na nekom od položaja na kojem se riječi razlikuju. Ukoliko se riječi y i z razlikuju na tom položaju onda je tvrdnja dokazana, a ukoliko ne onda se sigurno riječi x i z razlikuju na tom položaju. Dakle, na svakom položaju na kojem se x i y razlikuju, razlikovat će se ili x i z ili y i z. □

Teorem 2.2.4. *Minimalna udaljenost Hammingova* $(7, 4)$ *koda je* 3.

Dokaz teorema se provodi tako da se uoči da se promjenom znaka na jednoj poziciji automatski moraju promjeniti još barem dva znaka na neke druge dvije pozicije. Dokaz se može pronaći u knjizi [2, str. 16].

Napomena 2.2.5. *U idućim teoremima se pretpostavlja da se primljena riječ ispravlja u njoj najbližu riječ po Hammingovoj udaljenosti. Takva vrsta dekodiranja se često naziva dekodiranje najbližim susjedom.*

Teorem 2.2.6. *Neka je C neki kod i neka je δ neki broj. Kod C je $\delta - 1$-detektirajući ako i samo ako je $d(C) \geq \delta$.*

Dokaz. Neka je C neki $\delta - 1$-detektirajući kod takav da je $d(C) = \gamma < \delta$. Neka su $y, z \in C$ takve da vrijedi $d(y, z) = \gamma$. Ukoliko se y greškom u komunikaciji pretvori u z došlo je do γ pogrešaka u komunikaciji. No, pošto je $z \in C$, primatelj neće uočiti pogrešku. Dakle, kod nije γ-detektirajući, a pošto je $\gamma < \delta$, odnosno $\gamma \leq \delta - 1$ kod nije ni $\delta - 1$-detektirajući, što je kontradikcija sa pretpostavkom teorema.
U suprotom smjeru, neka je $d(C) \geq \delta$. Ukoliko se nekoj poslanoj riječi promjeni najviše $\delta - 1$ znakova, primljena riječ sigurno neće biti element koda C zbog pretpostavke da je minimalna udaljenost veća ili jednaka δ, tj. pogreška će se ispravno detektirati. □

Teorem 2.2.7. *Neka je C neki kod i neka je ε neki broj. Kod C je ε-ispravljajući ako i samo ako je $d(C) \geq 2\varepsilon + 1$.*

Dokaz. Neka je C neki ε-ispravljajući kod takav da vrijedi $\gamma = d(C) < 2\varepsilon + 1$. Neka su $x, y \in C$ takve da vrijedi $d(x, y) = \gamma$. Neka je u neka riječ koja je nastala tako da se na $\lfloor \frac{\gamma}{2} \rfloor$ pozicija riječi x na kojima se riječ razlikuje od y, stave znakovi sa odgovarajuće pozicije riječi y. Neka je sada x poslana, a u primljena riječ. Broj pogrešaka ovog prijenosa je $\lfloor \frac{\gamma}{2} \rfloor \leq \gamma < 2\varepsilon + 1$. Ukoliko je γ paran broj tada vrijedi $d(x, u) = d(y, u) = \frac{\gamma}{2} \leq \frac{2\varepsilon+1}{2}$, pa se ne može znati kako ispraviti u, a ukoliko je γ neparan tada vrijedi $d(x, u) = \frac{\gamma+1}{2} \leq \frac{2\varepsilon+2}{2}$, a $d(y, u) = \frac{\gamma-1}{2} \leq \frac{2\varepsilon}{2}$, pa će se u dekodirati u y. Dakle kod nije ε-ispravljajući jer vrijedi $d(y, u) \leq \varepsilon$.
Neka je sada $d(C) \geq 2\varepsilon + 1$. Neka je x neka poslana riječ, a u neka primljena riječ takva da je došlo do ε pogrešaka u prijenosu. Neka je y bilo koja kodna riječ od C. Vrijede sljedeće nejednakosti:

$$2\varepsilon + 1 \leq d(x, y) \leq d(x, u) + d(u, y) \leq \varepsilon + d(u, y). \tag{2.2}$$

Prva nejednakost slijedi iz definicije minimalne udaljenosti, a druga iz nejednakosti trokuta. Iz nejednakosti sada slijedi $\varepsilon + 1 \leq d(u, y)$. Dakle, udaljenost između u i bilo koje koje druge riječi koda C je veća od ε. Slijedi da će se u ispraviti u x. □

Korolar 2.2.8. *Dekodiranjem Hammingova* $(7, 4)$ *koda može se ispraviti samo jedan znak.*

Dokaz. Iz teorema (2.2.4) slijedi da je minimalna udaljenost Hammingova $(7, 4)$ koda jednaka 3, a iz teorema (2.2.7) slijedi $3 \leq 2\varepsilon + 1$, odnosno $\varepsilon \leq 1$. Dakle, Hammingov kod je najmanje 1-ispravljajući, no iz primjera (2.1.6) slijedi da nije 2-ispravljajući, pa nije ni ispravljajući ni za bilo koji veći broj od 2. □

Poglavlje 3

Blok kodovi

Definicija 3.0.1. *Blok kod je kod za ispravljanje grešaka koji poruku dijeli na blokove. Poruka se enkodira na način da joj se redundantni znakovi dodaju na kraj kao jedan blok. Također vrijedi da svaka poruka ima jednak broj znakova.*

Napomena 3.0.2. *Drugi tip kodova koji ispravaljaju greške su konvolucijski kodovi. Kod njih duljina enkodirane poruke nije fiksna i redundantni znakovi se ne moraju nalaziti na kraju poruke. Na primjer, iza svakog znaka se postavi jedan redundatni znak.*

Napomena 3.0.3. *Hammingov* $(7, 4)$ *kod je blok kod.*

Svaki kod se može okarakterizirati sa duljinom riječi koju sadrži, u oznaci n, sa brojem kodnih riječi koje sadrži, u oznaci M, i sa minimalnom udaljenosti, u oznaci d. Ukoliko kod nije binaran, tada ga karakterizira i broj simbola alfabeta nad kojim je izgrađen q. Broj n zapravo predstavlja cijenu slanja svake poruke, broj M predstavlja bogatstvo jezika, a iz prethodnog odjeljka je jasno da je minimalna udaljenost izravno povezana sa brojem grešaka koje kod može detektirati i ispraviti (uz pretpostavku dekodiranja najbližim susjedom). Iz svega opisanog lako se zaključi da bi za svaki kod bilo poželjno da ima što manji n, a što veće M i d.

Napomena 3.0.4. *U ovom odjeljku će se blok kod sa gore opisanim značenjima pojedinih brojeva nazivati* $q-$*aran* (n, M, d)*, a ukoliko je kod izgrađen nad binarnim alfabetom, označavat će se sa* (n, M, d)*.*

Primjer 3.0.5. *Za Hammingov* $(7, 4)$ *kod vrijedi:*

- $n = 7$*, jer je duljina riječi* 7*.*
- $M = 16$*, jer prva četiri znaka mogu biti ili* 0 *ili* 1*, što daje* $2^4 = 16$ *mogućnosti, a zadnja* 3 *znaka(redundantni znakovi) su unaprijed određeni pomoću prva četiri.*

- $d = 3$, *slijedi iz teorema* (2.2.4).

Dakle, Hammingov $(7, 4)$ *kod je primjer jednog* $(7, 16, 3)$ *koda.*

Napomena 3.0.6. *Često se blok kod izgrađen nad binarnim alfabetom označava i sa* (n, k)*, pri čemu n označava duljinu riječi, a k broj slobodnih pozicija, tj. onih koji na kojima nisu redundatni znakovi. Zbog toga se Hammingov kod iz prethodnog poglavlja označava sa* $(7, 4)$.

3.1 Gornja granica za maksimalni broj elemenata koda

Već je rečeno da je poželjno da kod sadrži što više kodnih riječi. U ovom poglavlju će se pokazati nekoliko rezultata koji određuju gornju granicu na broj kodnih riječi. Pretpostavlja se da su n, d i q fiksne veličine.

Propozicija 3.1.1. *Vrijedi* $M \leq q^n$.

Dokaz. Broj q^n odgovara broju riječi nad alfabetom sa brojem znakova q duljine n. □

Propozicija 3.1.2 (Singleton granica)**.** *Neka je C neki* $q-$*aran* (n, M, d) *kod. Vrijedi sljedeća nejednakost:*

$$M \leq q^{n-d+1}. \tag{3.1}$$

Dokaz. Neka je L skup konstruiran pomoću koda C na način da se ukloni prvih $d-1$ znakova riječi iz koda C. Kada bi u skupu L postojale dvije iste riječi, tada bi za riječi iz koda C od kojih su riječi dobivene vrijedilo da se razlikuju samo na nekim od prvih $d-1$ mjesta. Dakle, udaljenost tih dviju riječi je sigurno manja od d, što je kontradikcija sa pretpostavkom propozicije. Dakle, skup L se se ne sastoji od dvije iste riječi. Maksimalni veličina skupa L je q^{n-d+1}, a pošto kod C ima jednako riječi kao L vrijedi da je M manji ili jednak od q^{n-d+1}. □

Napomena 3.1.3. *Od sada će se često broj znakova koje neka riječ v sadrži formalno označavati sa* $|v|$.

Lema 3.1.4. *Neka je C neki* δ*-ispravljajući kod i neka je:*

$$S(u, r) = \{v : d(u, v) \leq r \; i \; |v| = n\} \tag{3.2}$$

krug sa središtem u u i radijusom r s obzirom na Hammingovu udaljenost. Tada nijedna dva kruga sa radijusom δ *i sa središtima u različitim kodnim riječima nemaju zajednički element.*

Dokaz. Neka su riječi x,$y \in C$ takve da se krugovi $S(x,\delta)$ i $S'(y,\delta)$ sijeku u nekoj točki z. Tada bi vrijedilo:

$$d(x,y) \leq d(x,z) + d(y,z) \leq \delta + \delta = 2\delta. \tag{3.3}$$

Prva nejednakost slijedi iz teorema (2.2.3), a druga iz definicije kruga s obzirom na Hammingovu udaljenost. Pošto je C δ-ispravljajući kod iz teorema (2.2.7) vrijedi $d(C) \geq 2\delta + 1$. Dakle:

$$2\delta + 1 \leq d(C) \leq d(x,y) \leq 2\delta. \tag{3.4}$$

Odnosno, $1 \leq 0$. Dakle nijedna dva kruga sa središtem u različitim kodnim riječima i sa radijusom δ nemaju zajednički element. □

Propozicija 3.1.5 (Hammingova granica)**.** *Neka je C neki $q-$aran (n,M,t) kod. Vrijedi sljedeća nejednakost:*

$$M \leq \frac{q^n}{\sum_{i=o}^{t} \binom{n}{i}(q-1)^i}. \tag{3.5}$$

Dokaz. Neka je u neka kodna rječ. Broj riječi koje su od riječi u udaljene za i je jednak $\binom{n}{i}(q-1)^i$. Naime, mjesta na kojima se riječi razlikuju mogu se odabrati na $\binom{n}{i}$ načina i na tim mjestima može biti koji od $(q-1)$ znakova. Iz toga slijedi da je broj riječi unutar kruga $S(u,t)$ jednak:

$$\sum_{i=0}^{t} \binom{n}{i}(q-1)^i. \tag{3.6}$$

Broj kodnih riječi, odnosno broj krugova, je jednak M. Dakle, broj kodnih riječi može biti:

$$M \sum_{i=0}^{t} \binom{n}{i}(q-1)^i. \tag{3.7}$$

Otprije je pokazano da broj kodnih riječi ne može bit veci od q^n, iz čega slijedi:

$$M \sum_{i=0}^{t} \binom{n}{i}(q-1)^i \leq q^n. \tag{3.8}$$

Dijeljenjem lijeve i desne strane nejednadžbe sa sumom dobije se formula (3.6). □

Korolar 3.1.6. *Za neki (n,M,d) kod koji je izgrađen nad alfabetom s brojem znakova jednakim 2 ($q = 2$) vrijedi:*

$$M \leq \frac{2^n}{\sum_{i=0}^{d} \binom{n}{i}}. \tag{3.9}$$

Odnosno:

$$M \leq 2^{n-d+2}. \tag{3.10}$$

Prva nejednakost predstavlja Hammingovu, a druga Singleton granicu.

Sljedeća propozicija daje odmak u odnosu na naziv poglavlja. Naime, u njoj se daje jedna donja granica na M.

Propozicija 3.1.7 (Gilbert-Varshamova donja granica)**.** *Za svaka tri broja n, q i $d \leq n$ postoji neki $q-aran$ (n, M, d) kod C takav da za broj riječi M koje sadrži vrijedi:*

$$M \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i}. \tag{3.11}$$

Dokaz. Konstrukcija koda C počinje tako da se prvo uzme neka riječ duljine n nad alfabetom duljine q. Zatim druga za koju mora vrijediti samo da je udaljenost od prve izabrane riječi veća ili jednaka d. Treća riječ se bira tako da joj udaljenost mora biti najmanje d od prvih dviju izabranih riječi. Taj proces se nastavlja dok god postoji riječ čiji krug s radijusom $d-1$ ne sadrži ni jednu već izabranu riječ. Ovaj proces je sigurno konačan jer riječi duljine n nad alfabetom duljine q ima konačno mnogo. Očito je konstruirani skup C jedan $q-aran$ (n, M, d) kod. Potrebno je sada odrediti donju granicu za broj M. Može se uočiti iz same konstrukcije koda C da skup svih krugova radijusa $d-1$ sadrži sve riječi duljine n. Neka je U skup definiran na sljedeći način:

$$U = \bigcup_{c \in C} S(c, d-1). \tag{3.12}$$

Veličina skupa U je jednaka q^n. Svaka riječ se nalazi u više krugova a svaki krug sadrži $\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i$ riječi (uz isto zaključivanje kao u propoziciji (3.6)). Neka b označava tu sumu. Vrijedi sljedeća nejednakost:

$$Mb \geq |U|. \tag{3.13}$$

Odnosno:

$$M \sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i \geq q^n. \tag{3.14}$$

Dijeljenjem sa $(\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i)^{-1}$ slijedi nejednakost (3.11). □

3.2 Savršeni kodovi

Definicija 3.2.1. *Za neki $q-aran$ (n, M, d) t-ispravljajući kod se kaže da je savršen ako vrijedi:*

$$M = \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i}(q-1)^i}. \tag{3.15}$$

Propozicija 3.2.2. *Hammingov* $(7, 4)$ *kod je savršen.*

Dokaz. Broj kodnih riječi je $M = 2^4 = 16$. Naime, prva 4 znaka mogu biti ili 0 ili 1, a zadnja 3 su redundantni znakovi koji su određeni pomoću prva 4 znaka. S druge strane, uvrštavanjem vrijednosti $n = 7$, $q = 2$ i $t = 1$ u jednakost (3.15) dobije se:

$$M = \frac{2^7}{\sum_{i=0}^{1} \binom{7}{i}(2-1)^i} = 16. \tag{3.16}$$

Dakle, Hammingov $(7, 4)$ kod je savršen. □

Teorem 3.2.3. *Za t-ispravljajući* $q-$*aran* (n, M, d) *kod C vrijedi da je savršen ako i samo vrijedi da je skup svih krugova sa središtima u kodnim riječima i sa radijusom t, u parovima disjunktan i u uniji daje skup svih riječi duljine n nad alfabetom s brojem simbola q.*

Dokaz. Neka je t-ispravljajući kod C savršen. Tvrdnja da je skup

$$S = \{S(c, t); c \in C\} \tag{3.17}$$

u parovima disjunktan slijedi iz leme (3.1.4). Za broj riječi skupa S vrijedi:

$$|S| = M \sum_{i=0}^{t} \binom{n}{i}(q-1)^i. \tag{3.18}$$

Pošto je kod C savršen vrijedi $M \sum_{i=0}^{t} \binom{n}{i}(q-1)^i = q^n$, odnosno $|S| = q^n$, što odgovara broju riječi duljine n nad alfabetom od q elemenata.
Obratno, ako za skup S vrijedi da u uniji daje sve riječi duljine n nad alfabetom od q elemenata, onda je $|S| = q^n$, pa dijeljenjem jednakosti (3.18) sa sumom slijedi (3.15). □

Definicija 3.2.4. *Neka je C neki kod duljine n nad alfabetom s brojem znakova q. Za najmanji broj r takav da je skup S definiran sa:*

$$S = \{S(c, r) : c \in C\} \tag{3.19}$$

jednak skupu svih riječi duljine n nad alfabetom od q elemenata, se kaže da je prekrivajući radijus koda C. Broj r se često označava s ρ.

Propozicija 3.2.5. *Prekrivajući radijus Hammingova* $(7, 4)$ *koda je* 1.

Dokaz. Slijedi direktno iz činjenice da je kod 1-ispravljajući, propozicije (3.2.2) i teorema (3.2.3). □

Definicija 3.2.6. *Za neki* $q-$*aran* (n, M, d) *kod se kaže da je optimalan ako se u njega ne može dodati riječ bez da se smanji vrijednost broja d.*

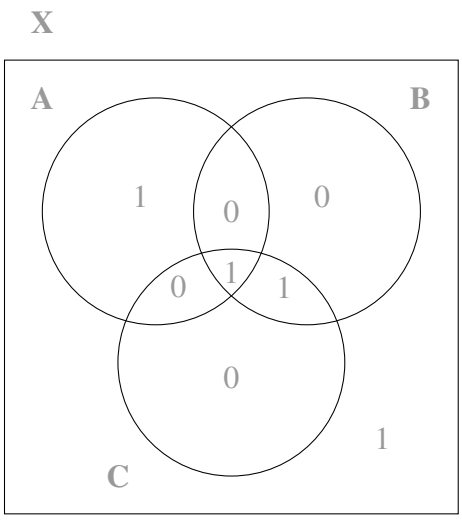
Propozicija 3.2.7. *Hammingov* $(7,4)$ *kod je optimalan.*

Dokaz. Neka kod nije optimalan, to jest postoji riječ x koja je udaljena od bilo koje riječi koda za više od 3, a nije element koda. Iz same konstrukcije koda slijedi da postoji neka riječ c koja je element koda, takva da se c na prve četiri pozicije poklapa sa x. Dakle, riječ x nije dobro enkodirana, tj. nije kodna riječ Hammingova koda. □

3.3 Hammingov (8,4) kod

Hammingov $(8,4)$ predstavlja poboljšanje u odnosu na $(7,4)$ kod. U odnosu na $(7,4)$ kod pokazat će se da $(8,4)$ kod prepoznaje da su se dogodile dvije pogreške u prijenosu, no i dalje je samo 1-ispravljajući. Za razliku od $(7,4)$ koda Hammingov $(8,4)$ kod ima još jedan redundantni znak na kraju riječi. Postupak enkodiranja poruke za prvih 7 znakova je isti, a zadnji znak se dodaje tako da ukupni broj jedinica u poruci bude paran.

Primjer 3.3.1. *Potrebno je poslati poruku* 1100*. Prva* 3 *redundantna znaka se dodaju na isti način kao i prije. Dakle enkodirana poruka je sada* 1100100*. No, za razliku od prije dodaje se još jedan znak koji mora biti takav da ukupan broj jedinica bude paran. Dakle, enkodirana poruka je* 11001001*. Pripadni Vennov dijagram je prikazan na slici* (3.1)*. Na dijagramu se vidi da se posljednji znak stavlja izvan krugova.*



Slika 3.1: Vennov dijagram za poruku 11001001

U primjeru (2.1.6) je pokazano da za $(7,4)$ kod ukoliko dođe do dvije pogreške u prijenosu poruke, nije moguće ispraviti poruku tj. naći mjesta na kojima se pogreška dogodila.

S druge strane poruka iz primjera (2.1.6) kodirana sa $(8,4)$ kodom bi bila 11111111. I ukoliko bi se opet pogreške dogodile na istom mjestu, tj. primljena poruka bi bila 11110101 primatelj bi mogao detektirati da je do pogreške došlo na dva mjesta jer je zadnji znak ispravan. Iz ovoga se naslućuje da je Hammingov $(8,4)$ kod 2-detektirajući.

Postupak dekodiranja poruke uz uvjet da su se dogodile najviše dvije pogreške je sljedeći:

1. Znakovi se stavljaju redom u polja od 1 do 8, pri čemu polje 8 predstavlja prostor van krugova.

2. Radi se provjera pariteta za zadnji znak poruke (provjerava se je li ukupni broj jedinica paran) i za svaki od krugova.

- Ukoliko su provjere pariteta za svaki od krugova i za zadnji znak ispravne, poruka je ispravna.

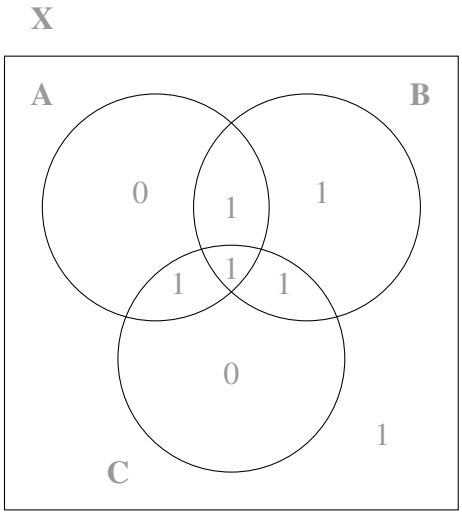
- Ukoliko nije ispravna provjera pariteta za zadnji znak i nije ispravna neka provjera pariteta za krugove A, B i C došlo je do jedne pogreške koja se ispravlja na isti način kao i za $(7,4)$ kod.

- Ukoliko provjera nije ispravna samo za zadnji znak, on je pogrešan.

- Ukoliko je provjera pariteta točna za zadnji znak, a neka provjera pariteta nije ispravna za krugove A, B i C, dva podatka su krivo poslana.

Sljedeći primjer pokazuje da Hammingov $(8,4)$ kod nije 2-ispravljajući.

Primjer 3.3.2. *Neka je pošiljatelj poslao poruku* 11111111*, a primatelj primio poruku* 11110101*, tj. došlo je do dvije pogreške u prijenosu. Primatelj sada radi provjeru pariteta pomoću Vennova dijagrama prikazanog na slici* (3.2)*. Primatelj uočava dvije pogreške. Naime, provjera pariteta za zadnji znak je ispravna, no provjera nije ispravna ni za krug A ni C. No primatelj ne može doznati na koja dva mjesta se dogodila pogreška. Npr. pogreška se mogla dogoditi na pozicijama* 5 *i* 7*, ali i na pozicijama* 1 *i* 6*.*



Slika 3.2: Vennov dijagram za poruku 11110101

Iz postupka dekodiranja Hammingova $(8,4)$ koda, prethodnog primjera i činjenice da je Hammingov $(7,4)$ kod 1-ispravljajući slijedi sljedeći teorem.

Teorem 3.3.3. *Hammingov* $(8,4)$ *kod je* 1*-ispravljajući.*

Napomena 3.3.4. *Može se zaključiti da je Hammingov* $(8,4)$ *kod* 2*-detektirajući, kao i Hammingov* $(7,4)$*. Razlika je u tomu što* $(7,4)$ *kod ne prepoznaje je li se dogodila jedna ili dvije pogreške dok* $(8,4)$ *kod prepoznaje. Zadnji znak dakle služi kao provjera jesu li se dogodile jedna ili pak dvije pogreške. Ukoliko je došlo do dvije pogreške primatelj npr. može zatražiti ponovno slanje poruke.*

Propozicija 3.3.5. *Hammingov* $(8,4)$ *kod nije savršen.*

Dokaz. Istim zaključivanjem kao u propoziciji (3.2.2) slijedi da je ukupan broj riječi koda jednak 16. S druge strane evaluiranjem izraza (3.15) za $q = 2$, $t = 1$ i $n = 8$, se dobiva:

$$\frac{2^8}{\sum_{i=0}^{1}\binom{8}{i}(2-1)^i} = \frac{256}{9} \neq 16. \tag{3.20}$$

Sada se može zaključiti da kod nije savršen. □

Teorem 3.3.6. *Prekrivajući radijus Hammingova* $(7, 4)$ *koda je* 2*.*

Dokaz. Neka je $x = x_1x_2 \dots x_8$ neka riječ nad binarnim alfabetom i neka je $x' = x_1 \dots x_7$ riječ duljine 7. Pošto je prekrivajući radijus Hammingova $(7, 4)$ koda jednak 1, x' je udaljene od neke kodne riječi c' Hammingova $(7, 4)$ koda za najviše 1. Neka je c riječ duljine 8 koja se poklapa sa riječi c' na prvih 7 pozicija. Udaljenost riječi c do riječi x je najvise 2. Dakle, prekrivajući radijus je manji ili jednak 2. Neka je sada 00000011 neka riječ duljine 8. U ovoj riječi je zadnji znak točan ali provjera pariteta ne prolazi za krug C. Dakle, dva znaka su krivo poslana, odnosno riječ je udaljena od neke riječi Hammingova $(8, 4)$ koda za najmanje 2. □

3.4 Ekvivalentni kodovi

Za potrebe ovog odjeljka treba definirati dva tipa permutacija.

Definicija 3.4.1. *Pozicijska permutacija nad kodom C je zamjena znakova na i-tom i j-tom položaju nad svakom riječi iz C.*

Primjer 3.4.2. *Neka je* $C = \{abba, bbba, abab, abbb\}$ *kod. Jedna pozicijska permutacija koja zamjenjuje znakove na pozicijama* 1 *i* 4 *je* $C' = \{abba, abbb, bbaa, bbba\}$.

Definicija 3.4.3. *Neka je C neki kod i neka je i, takav da vrijedi* $1 \leq i \leq n$*, neki položaj unutar kodne riječi. Permutacija simbola nad kodom C na poziciji i je zamjena na svim kodnim riječima znaka na poziciji i nekim drugim istim znakom, pri čemu dva različita znaka ne mogu bit zamjenjena istim znakom.*

Primjer 3.4.4. *Neka je* $C = \{abbcbc, bcbcca, bcbcba, ababab\}$ *kod. Jedna permutacija simbola koja zamjenjuje znak c sa znakom b, znak b sa znakom a te znak a sa znakom c na poziciji* 5 *je* $C' = \{abbcac, bcbcba, bcbcaa, ababcb\}$.

Teorem 3.4.5. *Izvođenjem pozicijskih permutacija nad kodom C ne mijenja se minimalna udaljenost koda.*

Dokaz. Neka su $x, y \in C$. Neka je nad C izvedena neka pozicijska permutacija. Neka su i i j položaji na kojima je došlo do zamjene znakova. Ukoliko su se riječi x i y razlikovale na poziciji i one će se nakon izvođenja permutacije razlikovati na poziciji j, a ukoliko se nisu razlikovale na poziciji i nakon izvođenja permutacije neće se razlikovati na poziciji j. Potpuno istim zaključivanjem zaključuje se i za poziciju j. Iz svega opisanog slijedi da se udaljenost bilo koje dvije riječi nakon izvođenja permutacije ne mijenja, pa samim time ne dolazi ni do promjene minimalne udaljenosti.

□

Teorem 3.4.6. *Izvođenjem permutacija simbola nad kodom C ne mijenja se minimalna udaljenost koda.*

Dokaz. Neka su $x, y \in C$. Neka je nad C izvedena neka permutacija simbola. Neka je i položaj na kojem je došlo do zamjene. Ukoliko su riječi prije zamjene imale isti znak na poziciji i one će i nakon djelovanja permutacije imati isti znak na toj poziciji. Ukoliko nisu imale isti znak tada ni nakon djelovanja permutacije neće imati isti znak jer permutacija ne može dva različita znaka zamjeniti istim znakom. Dakle, udaljenost između dvije riječi se ne mijenja djelovanja permutacije simbola pa samim tim ni minimalna udaljenost. □

Definicija 3.4.7. *Dva koda su ekvivalentna ako se jedan može dobiti iz drugog nizom pozicijskih permutacija i permutacija simbola.*

Definicija 3.4.8. *Neka su C i C′ neka dva koda nad istim alfabetom, sa istom duljinom riječi i istim brojem elemenata, za C i C′ se kaže da su izomorfni po udaljenosti ako se njihove riječi mogu poredati na način da za svaki c_i, c_j i za svaki c'_i, c'_j vrijedi $d(c_i, c_j) = d(c'_i, c'_j)$.*

Može se lako vidjeti da su svaka dva ekvivalentna koda izomorfna po udaljenosti. Za riječi c_i i c_j iz koda C uzmu se riječi c'_i i c'_j koje su dobivene nizom permutacija nad tim riječima.

Poglavlje 4

Linearni kodovi

U ovom odjeljku smatra se da je alfabet nad kojim se grade riječi, odnosno kodovi jednak polju $\mathbf{F}_p$ pri čemu p predstavlja neki prosti broj.

Napomena 4.0.1. *Neka je $x = x_1x_2 \dots x_n$ neka riječ nad alfabetom $\mathbf{F}_p$. Vektorski prikaz riječi x je vektor $(x_1, x_2, \dots, x_n) \in \mathbf{F}_p^n$.*

Napomena 4.0.2. *Često će se od sada pa na dalje riječ i njen vektorski prikaz smatrati sinonimima.*

Definicija 4.0.3. *Neka su $x = (x_1, x_2, \dots, x_n)$ i $y = (y_1, y_2, \dots, y_n)$ neke riječi.*

- *Zbroj riječi x i y, u oznaci $x + y$, se definira kao:*
  $x + y = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$.
- *Umnožak riječi x sa skalarom $\gamma \in \mathbf{F}_p$, u oznaci γx, se definira kao:*
  $\gamma x = \gamma(x_1, x_2, \dots, x_n) = (\gamma x_1, \gamma x_2, \dots, \gamma x_n)$.

Definicija 4.0.4. *Kod C je linearan ako vrijedi sljedeće:*

- *za $c, c' \in C$ vrijedi $c + c' \in C$.*
- *za $c \in C$ i za $\gamma \in \mathbf{F}_p$ vrijedi $\gamma c \in C$.*

Iz definicije lagano slijedi da svaki linearni kod sadrži riječ koja na svakom položaju ima znak 0, tzv. nul riječ.

Definicija 4.0.5. *Težina riječi x u oznaci $\omega(x)$ je broj pozicija na kojima se ne nalazi znak 0.*

Lema 4.0.6. *Neka su c_1 i c_2 dvije riječi jednake duljine, tada je $\omega(c_1 - c_2) = d(c_1, c_2)$.*

Dokaz. Riječ $c_1 - c_2$ na nekom položaju ima znamenku različitu od 0 ako i samo ako su pripadne znamenke riječi c_1 i c_2 različite. Dakle, broj znamenki različitih od 0 riječi $c_1 - c_2$ je jednak broju položaja na kojima se riječi c_1 i c_2 razlikuju. □

Teorem 4.0.7. *Za svaki linearni kod C, minimalna udaljenost koda C je jednaka najmanjoj težini medu svim ne nul riječima.*

Dokaz. Neka je d minimalna udaljenost koda C i neka ω označava težinu najmanje ne nul riječi. Neka su c_1 i $c_2 \in C$ takve da $d(c_1, c_2) = d$. Vrijedi

$$\omega \leq \omega(c_1 - c_2) = d(c_1, c_2) = d. \tag{4.1}$$

Ovdje $\omega(c_1 - c_2)$ označava težinu riječi $c_1 - c_2$. Iz linearnosti koda C slijedi $c_1 - c_2 \in C$. Neka je c riječ iz koda C s najmanjom težinom. Vrijedi

$$\omega = \omega(c) = \omega(c - 0) = d(c, 0) \geq d. \tag{4.2}$$

Pri čemu 0 označava nul riječ. Dakle $\omega = d$. □

Definicija 4.0.8. *Neka je C kod, neka je c poslana kodna riječ, a neka je r primljena riječ. Riječ $e = r - c$ se naziva uzorak pogreške ovog prijenosa.*

Teorem 4.0.9. *Za svaki linearni kod C, koristeći metodu dekodiranja najbližim susjedom, je li primljena riječ r ispravno dekodirana ovisi samo o e, a ne o r.*

Dokaz. Neka su c, c' dvije riječi linearnog koda C. Neka je e uzorak pogreške takav da se $r = c + e$ dekodira ispravno, a $r' = c' + e$ ne. Pošto se r' ne dekodira ispravno znači da postoji kodna riječ c'' takva da vrijedi $d(r', c'') \leq d(r', c')$ u koju se r dekodira. Neka je e' uzorak pogreške prijenosa poruke c'' u r' tj. $e' = r' - c''$. Iz $d(r', c'') \leq d(r', c')$ slijedi i sljedeća nejednakost: $\omega(e') \leq \omega(e)$.
Neka je $u = r - e'$. Pošto je kod C linearan, riječ u je sadržana u njemu. Vrijedi

$$d(u, r) = \omega(e') \leq \omega(e) = d(c, r). \tag{4.3}$$

Iz gornje nejednakosti slijedi da je udaljenost izmedu u i r manja od udaljenosti od c do r, što je kontradikcija s pretpostavkom da je kodna riječ r ispravno dekodirana metodom najbližeg susjeda. □

4.1 Generirajuća matrica i Matrica provjere pariteta

U Hammingovu (7,4) kodu koristio se Vennov dijagram da bi se riječ duljine 4 enkodirala u riječ duljine 7. Potrebno je izgraditi jedan općenitiji sustav enkodiranja poruke za linearne

kodove. U tu svrhu definiraju se generirajuće matrice koda. Blok enkodiranje se provodi tako da se uzmu sve poruke duljine k i na njih se nadoda još $n - k$ redundantnih znakova. Primjera radi, u Hammingovu $(7, 4)$ kodu duljina enkodirane riječi je 7, a broj redundantnih znakova je 3.

Napomena 4.1.1. *Lagano se može uočiti, iz definicije linearnog koda i činjenice da je nul riječ sadržana u svakom linearnom kodu, da je linearni kod, s duljinom riječi n nad alfabetom $\mathbf{F}_p$, jedan vektorski potprostor od $\mathbf{F}_p^n$. Zbog toga će se od od sada nadalje pojmovi lineanog koda i vektorskog potprostora poistovjećivati.*

Da bi se definirala generirajuća matrica potrebno je definirati određene pojmove iz Linearne algebre.

Definicija 4.1.2. *Neka su $x_1, x_2, \ldots, x_t \in \mathbf{F}_p^n$. Riječ $\sum_{i=1}^{t} \lambda_i x_i \in \mathbf{F}_p^n$ se naziva linearna kombinacija riječi $x_1, x_2, \ldots, x_t$, pri čemu vrijedi $\lambda_i \in \mathbf{F}_p$ za svaki $i \in \{1, \ldots, t\}$.*

Definicija 4.1.3. *Skup riječi $S \subseteq \mathbf{F}_p^n$ je linearno nezavisan ako ne postoji riječ $x \in S$ koja se može izraziti kao linearna kombinacija ostalih riječi iz skupa S. Za skup koji nije linearno nezavisan se kaže da je linearno zavisan.*

Definicija 4.1.4. *Neka je C neki linearni kod (vektorski potprostor) i neka je $S \in C$. Skup S je baza koda C ako vrijede sljedeće dvije tvrdnje:*

- *Skup S je linearno nezavisan.*
- *Svaka riječ iz C se može dobiti kao linearna kombinacija riječi iz S.*

Definicija 4.1.5. *Dimenzija vektorskog potprostora (linearnog koda) je broj elemenata bilo koje njegove baze.*

Definicija je korektna jer se može dokazati da svaka baza nekog vektorskog prostora ima isti broj elemenata.

Napomena 4.1.6. *Slobodne pozicije neke kodne riječi su pozicije na kojima se ne nalaze redundantni znakovi. Broj slobodnih pozicija se označava sa k.*

Definicija 4.1.7. *Neka je C neki linearni kod sa duljinom riječi n i brojem slobodnih pozicija k. Generirajuća matrica koda C je $k \times n$ matrica čiji retci odgovaraju jednom elementu neke baze koda C.*

Za svaku riječ m sa k elemenata odgovarajuća enkodirana poruka sadržana u linearnom kodu C je jednaka mG, pri čemu je G neka $k \times n$ generirajuća matrica koda C.

Primjer 4.1.8. *Neka je $C \subseteq \mathbf{F}_3^5$ neki linearni kod kojem je broj slobodnih pozicija jednak 3. Neka je $S = \{12020, 21000, 11111\}$ neka baza koda C. Generirajuća matrica koda C dobivena pomoću baze S je:*

$$\begin{pmatrix} 1 & 2 & 0 & 2 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{4.4}$$

Neka je $m = (1, 2, 0)$ riječ koju je potrebno enkodirati. Enkodirana riječ m je jednaka:

$$mG = (1, 2, 0)\begin{pmatrix} 1 & 2 & 0 & 2 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = (2, 1, 0, 2, 0). \tag{4.5}$$

Pomoću slike (2.1) uvjeti na poruku enkodiranu Hammingovim $(7, 4)$ kodom mogu se zapisati kao sustav linearnih jednadžbi u aritmetici modulo 2.

$$\begin{cases} x_1 + x_3 + x_4 + x_5 = 0 \\ x_1 + x_2 + x_4 + x_6 = 0 \\ x_1 + x_2 + x_3 + x_7 = 0 \end{cases} \tag{4.6}$$

Svaka jednadžba odgovara jednom krugu Vennova dijagrama, a nepoznanice $x_1, x_2, \ldots, x_7$ odgovaraju vrijednostima unutar svakog od 7 polja Vennova dijagrama. Ukupan broj jedinica mora biti paran, što u aritmetici modulo 2 odgovara broju 0.

Matrica koja odgovara sustavu jednadžbi (4.6) je

$$H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.7}$$

Neka je c neka riječ duljine 7. Riječ c je kodna riječ Hammingova $(7, 4)$ koda ako vrijedi

$$cH^T = 0. \tag{4.8}$$

U jednadžbi c označava riječ c zapisanu kao vektor, 0 označava nul riječ, a H^T označava transponiranu matricu H.

Definicija 4.1.9. *Neka su $x = x_1x_2 \ldots x_n$ i $y = y_1y_2 \ldots y_n$ riječi. Umnožak riječi x i y, u oznaci $x{\cdot}y$, je jednak broju:*

$$x_1y_1 + x_2y_2 + \ldots + x_ny_n \pmod{p} \tag{4.9}$$

Definicija 4.1.10. *Dvije riječi $x, y \in \mathbf{F}_p^n$ su međusobno ortogonalne ako vrijedi:*

$$x \cdot y = 0. \tag{4.10}$$

Definicija 4.1.11. *Neka je $C \subseteq \mathbf{F}_p^n$ neki kod. Skup svih riječi koje su ortogonalne na svaku riječ iz C, u oznaci $C^\perp$, se naziva dualni kod.*

Iz prethodnih definicija sada se može zaključiti da je riječ c iz jednakosti (4.8), ortogonalna na skup $S = \{(1,0,1,1,1,0,0),(1,1,0,1,0,1,0),(1,1,1,0,0,0,1\}$. Dakle, Hammingov $(7,4)$ kod je jednak skupu $S^\perp$.

Definicija 4.1.12. *Neka je H neka $k \times n$ matrica sa koeficijentima u $\mathbf{F}_p$. Matrica H je matrica provjere pariteta linearnog koda C, sa duljinom riječi n i brojem slobodnih koeficijenata $n-k$, ako vrijedi da su retci matrice linearno nezavisni i C je jednak skupu svih riječi c koje zadovoljavaju jednadžbu $cH^T = 0$.*

Napomena 4.1.13. *U poglavlju 3.1 je rečeno da će se kod s duljinom riječi n, s brojem kodnih riječi M, minimalnom udaljenosti d i alfabetom nad kojim je izgrađen q označavati sa $q-aran$ (n,M,d). Od sada pa na dalje kod će označavati sa $[n,k]$ pri čemu će k označavati broj slobodnih pozicija. Naime, brojevi M i d nisu od značaja za sljedeći dio, a za alfabet je vec rečeno da je jednak polju $\mathbf{F}_p$.*

Sada slijedi iskaz teorema koji povezuje kodove C i $C^\perp$. Formalan dokaz teorema može se pronaći u knjizi [2, Teorem 5.8, str. 83].

Teorem 4.1.14. *Ako je C neki $[n,k]$ linearni kod nad $\mathbf{F}_p$, tada je $C^\perp$ neki $[n,n-k]$ linearni kod.*

Da bi se dokazao teorem (4.1.17) koji slijedi potrebno je iskazati jedan važan teorem iz Linearne algebre. Dokaz teorema može se pronaći u knjizi [1, Teorem 5.1.11., str. 125]

Teorem 4.1.15 (Teorem o rangu i defektu)**.** *Neka je A neka $k \times n$ matrica sa koeficijentima u $\mathbf{F}_p$. Slika matrice A, u oznaci $Im(A)$, je linearni kod koji odgovara svim riječima koje se mogu dobiti linearnom kombinacijom redaka matrice A, dok je jezgra matrice A, u oznaci $Null(A)$, linearni kod koji odgovara svim riječima c, takvima da vrijedi $cA^T = 0$. Rang matrice A je broj koji označava dimenziju slike, a defekt broj koji označava dimenziju jezgre. Vrijedi sljedeća jednakost:*

$$n = dim(Null(A)) + dim(Im(A)). \tag{4.11}$$

Korolar 4.1.16. *Hammingov $(7,4)$ kod je linearan kod.*

Dokaz. Tvrdnja slijedi iz Teorema o rangu i defektu (točnije iz toga da je jezgra matrice linearni kod), matrice (4.7) i jednakosti (4.8). □

Teorem 4.1.17. *Neka je H neka $(n-k) \times n$ matrica sa koeficijentima u $\mathbf{F}_p$. Matrica H je matrica provjere pariteta za neki $[n,k]$ kod C ako i samo ako je H generirajuća matrica za $C^\perp$.*

Dokaz. Neka je H matrica provjere pariteta koda C. Iz definicije matrice pariteta slijedi $c \in C$ ako vrijedi $cH^T = 0$. Dakle, vrijedi

$$C = Null(H). \tag{4.12}$$

Pošto je dimenzija koda jednaka k, slijedi $k = dim(Null(H))$. Nadalje, iz teorema o rangu i defektu slijedi da je $k = Dim(Null(H)) = n - dim(Im(H))$, odnosno $dim(Im(H)) = n - k$. Iz teorema (4.1.14) slijedi $dim(C^\perp) = n - k$, odnosno:

$$dim(C^\perp) = dim(Im(H)). \tag{4.13}$$

Također vrijedi i $Im(H) \subseteq C^\perp$. Sada se može zaključiti da je $Im(H) = C^\perp$. Iz čega slijedi da je H generirajuća matrica koda C^T.
Neka je H generirajuća matrica koda $C^\perp$. Iz definicije generirajuće matrice slijedi $C^\perp = Im(H)$. Iz definicije ortogonalnog koda i jezgre matrice slijedi $C \subseteq Null(H)$. Za dimenziju koda C vrijedi iz teorema (4.1.14):

$$dim(C) = k = n - (n - k) = n - dim(C^\perp) = n - dim(Im(H)) = dim(Null(H)). \tag{4.14}$$

Odnosno $C = Null(H)$. Dakle, vrijedi da je H matrica provjere pariteta koda C. □

Iz teorema (4.1.17) i činjenice da je dualni kod linearan ako je kod linearan, slijedi zaključak da svaki linearni kod ima matricu provjere pariteta. Jednostavno se uzme bilo koja generirajuća matrica koda $C^\perp$ koja odgovara matrici provjere pariteta koda C.

U nastavku se daje jedna metoda kako iz generirajuće matrice dobiti matricu provjere pariteta linearnog koda.

Definicija 4.1.18. *Za generirajuću matricu G nekog linearnog $[n, k]$ koda C se kaže da je u standardnom obliku ako se može zapisati kao*

$$G = [I_k|A]. \tag{4.15}$$

Ovdje I_k označava jediničnu matricu reda k, a A neku $k \times (n - k)$ matricu.

Napomena 4.1.19. *Generirajuće matrice u standardnom obliku su pogodne jer će riječ duljine k koju je potrebno enkodirati pomoću generirajuće matrice, nakon enkodiranja na prvih k mjesta imati nepromijenjene znakove.*

Teorem 4.1.20. *Neka je C neki $[n, k]$ linearni kod sa generirajućom matricom $G = [I_k|A]$ u standardnom obliku, tada za matricu provjere pariteta H koda C vrijedi*

$$H = [-A^T|I_{n-k}]. \tag{4.16}$$

Napomena 4.1.21. *Teorem se lako može dokazati tako da se pokaže da je svaki redak matrice H ortogonalan na svaki redak matrice G. Dakle, retci matrice H čine jednu bazu dualnog koda u odnosu na C, pa iz teorema* (4.1.17) *slijedi da je H matrica provjere pariteta koda C. O dokazu teorema može se više pročitati u knjizi [2, str. 88].*

Sada je pomoću prethodnog teorema moguće odrediti jednu generirajuću matricu Hammingova $(7, 4)$ koda.

Primjer 4.1.22. *Već je pokazano da je matrica* (4.7) *matrica provjere pariteta Hammingova* $(7, 4)$ *koda. Očito se matrica H može zapisati kao*

$$H = [-A^T | I_3]. \tag{4.17}$$

Ovdje je matrica A jednaka:

$$A = \begin{pmatrix} -1 & -1 & -1 \\ 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}. \tag{4.18}$$

Dakle iz prethodnog teorema slijedi da je jedna generirajuća matrica G koda jednaka:

$$G = [I_4 | A]. \tag{4.19}$$

Odnosno:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 \end{pmatrix}. \tag{4.20}$$

Pošto je kod izgrađen na poljem $\mathbf{F}_2$ *vrijedi*

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{4.21}$$

4.2 Metode dekodiranja linearnih kodova

Dekodiranje konstruiranjem standardnog niza

Definicija 4.2.1. *U teoriji kodiranja zajedničko ime za sve skupove koji čine particiju nekog skupa, takvu da svaki skup ima jednak broj elemenata, je koskupovi, a vođa koskupa predstavlja jedan element koskupa koji ima najmanju težinu.*

Definicija 4.2.2. *Standardni niz za neki linearni* $[n, k]$ *nad* $\mathbf{F}_p$ *kod je dvodimenzionalni niz kojem je broj redaka jednak* p^{n-k}*, a broj stupaca jednak* p^k*, takav da vrijedi sljedeće:*

- *Prvi red niza čine sve riječi koda C s nul riječi na prvom mjestu.*
- *Svaki drugi red predstavlja jedan koskup.*

U idućem primjeru opisat će se kako se konstruira standardni niz i kako se pomoću njega dekodira riječ.

Primjer 4.2.3. *Neka je C neki* $[4, 2]$ *linearni kod nad* $\mathbf{F}_2$*, sa generiririajućom matricom G koja je definirana na sljedeći način:*

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}. \tag{4.22}$$

Iz definicije generirajuće matrice slijedi da je kod C jednak:

$$C = \{\alpha(1, 0, 1, 1) + \beta(1, 1, 0, 1) : \alpha \in \{0, 1\}, \beta \in \{0, 1\}\} \tag{4.23}$$

Uvrštavanjem vrijednosti 0 *i* 1 *za* α *i* β *dobiju se sve kodne riječi koda C:*

$$C = \{(0, 0, 0, 0), (1, 1, 0, 1), (1, 0, 1, 1), (0, 1, 1, 0)\}. \tag{4.24}$$

Riječi koda C čine prvi red standarnog niza kojeg je potrebno konstruirati. Niz će se zapisivati u obliku tablice, pri čemu će jednom retku odgovarati jedan korak postupka. Također, potrebno je nul riječ, koja je sadražana u svakom linearnom kodu, postaviti na prvo mjesto u tablici.
Sada je potrebno uzeti bilo koju riječ iz $\mathbf{F}_2^4$ *koja već nije u tablici, uz uvjet da riječ ima najmanju težinu među svim dosad neodabranim riječima. Odabrana riječ se naziva vođa koskupa. Neka je* 1000 *odabrana riječ, odnosno vođa koskupa. Sada je potrebno vođu koskupa zbrojiti sa svakim elementom prvog reda. Tablica je sada sljedećeg oblika:*

$$\begin{array}{rcccc} C & = & 0000 & 1101 & 1011 & 0110 \\ C + 1000 & = & 1000 & 0101 & 0011 & 1110 \end{array}$$

Neka je sljedeći vođa koskupa riječ 0100*. Tablica je sada sljedećeg oblika:*

$$\begin{array}{rcrrrr} C & = & 0000 & 1101 & 1011 & 0110 \\ C+1000 & = & 1000 & 0101 & 0011 & 1110 \\ C+0100 & = & 0100 & 1001 & 1111 & 0010 \end{array}$$

Sljedeći vođa koskupa je riječ 0001*. Riječ* 0010 *također ima minimalnu težinu no ona je već sadržana u tablici. Tablica je sada sljedećeg oblika:*

$$\begin{array}{rcrrrr} C & = & 0000 & 1101 & 1011 & 0110 \\ C+1000 & = & 1000 & 0101 & 0011 & 1110 \\ C+0100 & = & 0100 & 1001 & 1111 & 0010 \\ C+0001 & = & 0001 & 1100 & 1010 & 0111 \end{array}$$

Postupak je gotov (standardni niz je kreiran) jer tablica sadrži sve riječi iz $\mathbf{F}_2^4$.
Neka je sada primljena neka riječ iz $\mathbf{F}_2^4$*. Postupak dekodiranja poruke je poprilično jednostavan, primljena riječ se pronađe u tablici i dekodira kao riječ koja se nalazi na prvom mjestu u stupcu u kojoj se primljena riječ nalazi. Na primjer, ukoliko je primljena riječ* 1010*, potrebno je pogledati u tablicu i vidjeti da se na prvom mjestu u stupcu nalazi riječ* 1011*. Dakle, riječ* 1010 *se dekodira kao* 1011.

Teorem 4.2.4. *Dekodiranje konstruiranjem standardnog niza je dekodiranje najbližim susjedom.*

Dokaz. Neka je x neka riječ koju je potrebno dekodirati. Neka se x dekodira u c pomoću standardnog niza. Iz definicije standardnog niza slijedi da je $d(c, x) = k$, pri čemu je k neki vođa koskupa, odnosno

$$x = c + k. \tag{4.25}$$

Neka postoji neka riječ c' takva da

$$x = c' + e \tag{4.26}$$

i vrijedi $d(x, c') < d(x, c)$. Iz jednadžbi (4.25) i (4.26) slijedi $e - k = c - c'$, odnosno $e - k$ je element koda. Dakle, riječi e i k se nalaze u istom koskupu. S druge strane iz $d(x, c') < d(x, c)$, slijedi $\omega(x - c') < \omega(x - c)$, odnosno $\omega(e) < \omega(k)$, što je konstradikcija s činjenicom da je k vođa koskupa.

□

Teorem 4.2.5. *Uzorci pogreške koji će biti ispravno dekodirani pomoću standardnog niza su koskup vođe.*

Dokaz. Neka je c poslana riječ, r primljena riječ, a e uzorak pogreške tog prijenosa, odnosno $e = r - c$. Riječ r će biti ispravno dekodirana ako se r i c nalaze u istom stupcu standardnog niza, tj. ukoliko postoji koskup k takav da vrijedi $r = c + k$. Uvrštavanjem $r = e + c$ u jednadžbu slijedi $e = k$. □

Sindromsko dekodiranje

Definicija 4.2.6. *Neka je H matrica provjere pariteta nekog linearnog koda i neka je u neka riječ. Sindrom riječi u, u oznaci $syn(u)$, je jednak uH^T.*

Teorem 4.2.7. *Neka je H matrica provjera pariteta linearnog koda C i neka su u i v neke riječi. Vrijedi $syn(u) = syn(v)$ ako i samo ako se riječi u i v nalaze u istom koskupu s obzirom na standardni niz.*

Dokaz. Riječi u i v pripadaju istom koskupu ako i samo ako vrijedi $u - v \in C$. Nadalje, vrijedi

$$u - v \in C \Leftrightarrow (u - v)H^T = 0 \Leftrightarrow uH^T = vH^T \Leftrightarrow syn(u) = syn(v).$$

□

Pošto je uzorak pogreške svake riječi vođa koskupa u koskupu u kojem se riječ nalazi, iz prethodnog teorema vrijedi $syn(e) = syn(v)$, pri čemu je u e uzorak pogreške za v.

Teorem 4.2.8. *Neka je C neki $[n, k]$ kod sa matricom provjere pariteta H i neka je $e = e_1e_2 \dots e_n$ uzorak pogreške riječi v. Vrijedi*

$$syn(v) = (\sum_{i=1}^{n} e_i h_i)^T \tag{4.27}$$

pri čemu h_i označava i-ti stupac matrice H.

Dokaz. Iz definicije sindroma vrijedi

$$syn(v) = syn(e) = (e_1, e_2, \dots, e_n) \begin{pmatrix} h_{11} & . & . & h_{n-k1} \\ . & . & . & . \\ . & . & . & . \\ h_{1n} & . & . & h_{n-kn} \end{pmatrix}. \tag{4.28}$$

Ovdje prva jednakost slijedi iz teorema (4.2.7), a druga iz definicije sindroma. Množenjem matrice H s riječi e dobije se sljedeća riječ zapisana kao vektor:

$$\begin{aligned} (e_1h_{11} + \dots + e_nh_{1n}, \dots, e_1h_{n-k1} + \dots + e_nh_{n-kn}) = \\ e_1(h_{11} \dots h_{n-k1}) + \dots + e_n(h_{1n} \dots h_{n-kn}) = \\ e_1(h_1)^T + \dots e_n(h_n)^T = (\sum_{i=1}^{n} e_i h_i)^T. \end{aligned}$$

□

Rezultati prethodnog teorema daju novu mogućnost za dekodiranje, tzv. sindromsko dekodiranje. Sindromsko dekodiranje se provodi tako da se odrede koskup vode i standardni niz. Standardni niz se nakon toga odbacuje, a koskup vode (odnosno uzorci pogreške (4.2.5)) se spremaju u niz. Prednost sindromskog dekodiranja je ta što nije potrebno pamtiti cijeli standardni niz nego samo koskup vode, pa je prostorna složenost algoritma znatno manja, iako je vremenska složenost ista u odnosu na dekodiranje pomoću tablice. Koraci sindromskog dekodiranja su sljedeći:

- Izračunati sindrom primljene riječi r.
- U listi vođa koskupa pronaći vođu sa istim sindromom.
- Dekodirati riječ r kao $c = r - e$.

Poglavlje 5

Neki kodovi koji ispravljaju greške

5.1 Hammingovi kodovi

Dosad su već opisana dva primjera Hammingovih kodova, pa će se zato u ovom poglavlju poopćeniti pojam Hammmingovih kodova i dati neki rezultati koji za njih vrijede.

Definicija 5.1.1. *$Ham(r, q)$ je skup svih linearnih $[n, k]$ kodova nad alfabetom $\mathbf{F}_q$ čije su matrice provjere pariteta dimenzija $r \times n$, pri čemu n predstavlja najveći mogući broj stupaca matrice provjere pariteta takvih da nijedna dva stupca nisu linearno zavisna.*

Postavlja se pitanje zašto bi svaki kod u $Ham(r, q)$ trebao imati istu duljinu riječi n i broj slobodnih pozicija k. Za broj slobodnih pozicija vrijedi $k = n - r$, pa sljedeći teorem daje odgovor na oba pitanja.

Teorem 5.1.2. *Za duljinu riječi svakog koda iz skupa $Ham(r, q)$ vrijedi*

$$n = \frac{q^r - 1}{q - 1}. \tag{5.1}$$

Dokaz. Neka je u neka riječ iz $\mathbf{F}_q^r$. Neka $m(u)$ označava skup svih ne nul riječi iz $\mathbf{F}_q^r$ koje se mogu dobiti množenjem skalarom riječi u. Pošto skalari moraju biti iz $\mathbf{F}_q$, slijedi da $m(u)$ sadrži $q - 1$ riječ. Ne nul riječi duljine r nad $\mathbf{F}_q$ ima $q^r - 1$. Neka je v neka riječ duljine r nad alfabetom $\mathbf{F}_q$, takva da postoji riječ a sadržana i u $m(u)$ i u $m(v)$. Dakle postoje skalari α i β takvi da vrijedi $a = \alpha u$ i $a = \beta v$. Neka je $x \in m(u)$. Slijedi da postoji skalar γ takav da vrijedi $x = \gamma u$. Iz prethodnih jednakosti slijedi:

$$x = \gamma u = \gamma\alpha^{-1}a = \gamma\alpha^{-1}\beta v. \tag{5.2}$$

Dakle, $x \in m(v)$, odnosno $m(u) \subseteq m(v)$. Istim postupkom se dokaže da za neki $y \in m(v)$, vrijedi $y \in m(u)$, odnosno $m(v) \subseteq m(u)$. Sada slijedi $m(u) = m(v)$, odnosno različiti skupovi oblika $m(x)$, za neki $x \in \mathbf{F}_q^p$, čine jednu particiju od $\mathbf{F}_q^p$. Ukupan broj skupova je

$$\frac{q^r - 1}{q - 1}. \tag{5.3}$$

Iz konstrukcije skupova slijedi da ukoliko točno jedna riječ iz svakog skupa postane jedan stupac matrice provjere pariteta, nijedna dva stupca neće biti linearno zavisna. Kada bi matrica provjere pariteta imala više stupaca od (5.3) tada bi matrica imala više stupaca iz istog skupa, koji bi bili zavisni. □

Definicija 5.1.3. *Dva linearna koda C i C' sa generirajućim matricama G_1 i G_2 su linearno ekvivalentna ako se matrica G_2 može dobiti iz matrice G_1 izvođenjem konačno mnogo sljedećih operacija:*

- *Zamjenom neka dva stupca matrice.*
- *Množenjem stupca matrice sa skalarom različitim od* 0.

U članku [5] je opisano kako pronaći permutaciju koja povezuje linearno ekvivalentne kodove. Slijedi iskaz teorema čiji se dokaz može pronaći u knjizi [2, Teorem 6.2, str. 104].

Teorem 5.1.4. *Za svaki r i q vrijedi da su svi $Ham(r, q)$ kodovi linearno ekvivalentni.*

Teorem 5.1.5. *Minimalna udaljenost d linearnog koda C je jednaka veličini najmanjeg mogućeg skupa zavisnih stupaca matrice provjere pariteta H.*

Dokaz. Neka je $c = c_1c_2 \dots c_n \in C$ neka riječ težine $d > 0$. Pošto je c kodna riječ koda C vrijedi $cH^T = 0$. Odnosno:

$$c_1h_1 + c_2h_2 + \ldots + c_nh_n = 0. \tag{5.4}$$

Ovdje h_i označava i-ti stupac matrice H. Iz činjenice da je riječ c težine d, slijedi da je skup $\{h_i : 1 \leq i \leq n\}$ linearno zavisan. Kada bi vrijedilo da je skup $\{h_{\gamma_1}, \dots, h_{\gamma_t}\}$, za neki $t < d$, linearno zavisan, tj. $k_{\gamma_1}h_{\gamma_1} + \ldots + k_{\gamma_t}h_{\gamma_t} = 0$ za neke netrivijalne k-ove, tada bi riječ $c' = k_{\gamma_1} \dots k_{\gamma_t}0 \dots 0$ čijih je zadnjih $n - t$ znakova jednako 0, imala težinu jednaku t, što je kontradikcija s teoremom (4.0.7). □

Teorem 5.1.6. *Za svaki r i q vrijedi da svi $Ham(r, q)$ kodovi imaju minimalnu udaljenost jednaku* 3.

Dokaz teorema se može pronaći u knjizi [2, Teorem 6.3, str. 104].

Teorem 5.1.7. *Svi Hammingovi kodovi su savršeni.*

Dokaz. Neka je $C \in Ham(r, q)$. Vrijedi da je C jedan $q - aran(n, M, d)$ kod pri čemu je $d = 3$ i vrijedi

$$n = \frac{q^r - 1}{q - 1} \text{ i } k = n - r. \tag{5.5}$$

Uvrštavnjem vrijednosti za d, n i k u jednakost (3.15) dobije se:

$$\frac{q^n}{\sum_{i=0}^{1} \binom{n}{i}(q-1)^i} = \frac{q^n}{1 + n(q-1)} = q^{n-r} = q^k = M. \tag{5.6}$$

Dakle, kod C je savršen. □

5.2 Simpleks kodovi

Definicija 5.2.1. *Dualni kod svakog $Ham(r, q)$ koda se naziva simpleks kod.*

Definicija 5.2.2. *Za neki kod se kaže da je ekvidistantan ako svake dvije različite riječi, sadržane u kodu, imaju jednaku udaljenost.*

Teorem 5.2.3. *Za svaki linearni kod broj kodnih riječi težine ω je jednak umnošku broja parova kodnih riječi sa udaljenosti jednakom ω i broja kodnih riječi M, pri čemu je ω neki broj.*

Dokaz. Neka je $C = \{c_1, \ldots, c_M\}$ linearni kod sa brojem kodnih riječi M i neka je A_ω broj kodnih riječi sa težinom ω. Neka je c neka riječ težine ω. Tada je udaljenost svaka dva para riječi oblika $(c_1, c_1 - c), (c_2, c_2 - c), \ldots (c_M, c_M - c)$ jednaka ω. Iz toga slijedi zaključak da kod C ima A_ω kodnih riječi težine ω ako i samo ako ima MA_ω parova kodnih riječi sa udaljenosti jednakom ω. □

Teorem 5.2.4. *Svi simpleks kodovi su ekvidistantni.*

Dokaz. Neka je $C \in Ham(r, q)$ i neka je H pripadna matrica provjere pariteta. Iz teorema (4.1.17) slijedi da je H generirajuća matrica koda $C^\perp$ koji ima q^r kodnih riječi. Skup svih parova kodnih riječi, u oznaci M^2, sadrži M parova sa udaljenosti 0 (oblika (c, c), za $c \in C^\perp$). Potrebno je dokazati da ostalih $M^2 - M$ parova kodnih riječi ima jednaku udaljenost. Zbog prethodnog teorema slijedi da je dovoljno dokazati da $M - 1$ riječi koda $C^\perp$ imaju jednaku težinu, što je jednako q^{r-1}. Neka sada kod $C^\perp$ ima riječ s težinom većom od q^{r-1} i neka je G generirajuća matrica koda $C^\perp$ sa tom riječi u prvom redu. Sada je potrebno svaki stupac kojemu je prvi znak različit od 0 pomnožiti sa inverzom tog znaka. Pošto je takvih znakova više od q^{r-1}, u novoj matrici G' će više od q^{r-1} stupaca imati prvi znak jednak 1. Svaki takav stupac se može popuniti znakovima na q^{r-1} način (generirajuća matrica ima r redaka). Dakle, postoje barem dva jednaka stupca, što je kontradikcija s

teoremima (5.1.5) i (5.1.6). Na sličan način se pokaže da $C^{\perp}$ ne sadrži riječ s duljinom manjom od q^{r-1}. □

Napomena 5.2.5. *U prethodnom teoremu se matrica G' dobila iz matrice G nekom transformacijom stupaca, dakle ona ne mora generirati isti kod, ali sigurno generira neki linearno ekvivalentan kod.*

5.3 Hadamardovi kodovi

Hadamardov kod je kod za detekciju i ispravljanje grešaka, nazvan po francuskom matematičaru Jacquesu Hadamardu. Ponekad se naziva i Walsh-Hadamardov kod po američkom matematičaru Josephu Leonardu Walshu.

Da bi se pokazala struktura Hadamardova koda potrebno je definirati Hadamardovu matricu.

Definicija 5.3.1. *Hadamardova matrica reda n, u oznaci H_n, je matrica s koeficijentima* 1 *ili* -1 *takva da vrijedi $H_n H_n^T = I_n$, pri čemu I_n predstavlja jediničnu matricu reda n.*

Slijedi nekoliko primjera Hadamardovih matrica:

- Hadamardova matrica reda 1:
$$H_1 = (1). \tag{5.7}$$
- Hadamardova matrica reda 2:
$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.8}$$
- Hadamardova matrica reda 4:
$$H_4 = \begin{pmatrix} -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix}. \tag{5.9}$$

Napomena 5.3.2. *Iz definicije Hadamardove matrice slijedi da svaka dva različita retka moraju biti međusobno ortogonalna, tj. njihov umnožak mora biti* 0.

Napomena 5.3.3. *Uz već definirane Hadamardove matrice H_1, H_2 i H_4, matematičari pretpostavljaju da Hadamardova matrica reda n postoji ukoliko je n višekratnik broja* 4.

Definicija 5.3.4. *Neka su A i B neke matrice. Kroneckerov produkt matrica A i B, u oznaci $A \otimes B$, se definira kao:*

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \dots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}. \tag{5.10}$$

Definicija 5.3.5. *Hadamardove matrice definirane sa*

$$H_{2^n} = H_2 \otimes H_{2^{n-1}}, \tag{5.11}$$

odnosno

$$H_{2n} = \begin{pmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{pmatrix}, \tag{5.12}$$

se nazivaju Hadamard-Sylvesterove matrice.

Napomena 5.3.6. *Može se dokazati da su matrice definirane u* (5.11) *i* (5.12)*, Hadamardove matrice za svaki n.*

Hadamardov kod sa duljinom riječi 2^k dobije se tako da se uzmu svi retci iz matrica H'_{2^k} i H''_{2^k}, pri čemu se matrica H'_{2^k} dobije iz Hadamard-Sylvesterove matrice H_{2^k} na način da joj svi koeficijenti jednaki -1 zamjene sa 0, a matrica H''_{2^k} se na isti način dobije iz matrice $-H_{2^k}$.

Napomena 5.3.7. *U načelu se za bilo koji kod konstruiran pomoću Hadamardove matrice kaže da je Hadamardov, ali u praksi se najviše radi s kodovima konstruiranima pomoću Hadamard-Sylvesterove matrice, ponajviše zbog toga što kod koji nije generiran pomoću Hadamard-Sylvesterove matrice ne mora biti linearan.*

Napomena 5.3.8. *Nastavno na prethodnu napomenu, od sada pa na dalje će se za Hadamardov kod smatrati da je konstruiran pomoću Hadamard-Sylvesterove matrice.*

Primjer 5.3.9. *Neka je*

$$H_{2^2} = H_4 = \begin{pmatrix} -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix} \tag{5.13}$$

Hadamardova matrica reda 4*. Zamjenom koeficijenata jednakih* -1 *sa* 0 *dobiju se sljedeće matrice:*

$$H'_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \tag{5.14}$$

i

$$H_4'' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{5.15}$$

Slijedi da je Hadamardov kod generiran matricom H_4 sljedećeg oblika:

$$C_4 = \{0000, 0101, 0011, 0110, 1111, 1010, 1100, 1001\} \tag{5.16}$$

Sada je potrebno opisati kako enkodirati riječ duljine $k + 1$. Postoji nekoliko različitih načina kako se riječ može enkodirati. Riječi od kojih se sastoji Hadamardov kod čine jedan vektorski potprostor. Iz toga slijedi da svaki Hadamardov kod ima pripadajuću generirajuću matricu dimenzija $(k + 1) \times 2^k$. Broj redaka generirajuće matrice mora odgovariti duljini riječi koju je potrebno enkodirati, a broj stupaca mora biti jednak duljini riječi koda. Generirajuća matrica se lako može dobiti iz matrica H'_{2^k} i H''_{2^k}, koje odgovaraju Hadamardovim matricama H_{2^k} i $-H_{2^k}$ sa zamjenjim vrijednostima -1 i 0, tako da se uzme $k + 1$ linearno nezavisnih redaka matrica H'_{2^k} i H''_{2^k}.

Primjer 5.3.10. *Iz matrica H'_4 i H''_4 definiranih u primjeru* (5.3.9)*, generirajuća matrica se može dobiti tako da se uzme drugi stupac matrice H'_4 i drugi i treći stupac matrice H''_4. Generirajuća matrica, u oznaci G, je sljedećeg oblika:*

$$G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \tag{5.17}$$

Lako se provjeri da su retci matrice linearno nezavisni i da generiraju Hadamardov kod C_4.

Napomena 5.3.11. *Hadamarov kod koji je konstruiran pomoću Hadamard-Sylvesterove matrice H_{2^k} će se označavati sa $[2^k, k + 1]$, što u skladu s napomenom* (4.1.13) *označava da kod ima duljinu riječi jednaku 2^k i broj slobodnih pozicija jednak $k + 1$.*

Bibliografija

[1] D. Bakić, *Linearna algebra*, (2020).

[2] J. Baylis, *Error-Correcting Codes*, A Mathematical Introduction (1998).

[3] H Evangelaras, Christos Koukouvinos i J Seberry, *Applications of Hadamard matrices*, Journal of telecommunications and information Technology (2003), br. 2, 3–10.

[4] Florence Jessie MacWilliams i Neil James Alexander Sloane, *The theory of error-correcting codes*, sv. 16, Elsevier, 1977.

[5] N. Sendrier, *Finding the permutation between equivalent linear codes: the support splitting algorithm*, IEEE Transactions on Information Theory **46** (2000), br. 4, 1193–1203.

[6] Jacobus Hendricus Van Lint, *A survey of perfect codes*, The Rocky Mountain Journal of Mathematics **5** (1975), br. 2, 199–224.

Sažetak

Cilj ovoga rada je opisati kodiranja koja ispravljaju greške. U prvom poglavlju su opisani osnovni pojmovi kod kodiranja koja ispravljaju greške i dan je jedan jednostavni primjer kodiranja. U drugom poglavlju se opisuje Hammingov $(7, 4)$ kod i definira se Hammingova udaljenost. U trećem poglavlju opisuju se blok kodovi, tj. kodovi u kojima se podatci dijele na blokove. Također je dana gornja granica na broj kodnih riječi. Opisani su savršeni kodovi i definiran pojam ekvivalentnosti kodova. Opisan je i Hammingov $(8, 4)$ kod koji detektira dvije pogreške u prijenosu poruuke. U četvrtom poglavlju su opisani linearni kodovi i dana je definicija dviju važnih matrica: generirajuće matrice i matrice provjere pariteta. Opisano je kako pomoću njih dobiti enkodiranu riječ, odnosno kako provjeriti pripada li riječ kodu. U istom poglavlju su opisane i neke metode dekodiranja poruke. U petom poglavlju se formaliziraju Hammingovi kodovi. Opisani su i simpleks kodovi te Hadamardovi kodovi.

Summary

The goal of this paper is to describe error-correcting codes. In the first chapter, the basic concepts of error-correcting codes are described and a simple example is given. In the second chapter, the Hamming $(7, 4)$ code is described and the Hamming distance is defined. In the third chapter, block codes are described, i.e. codes in which data is divided into blocks. In the same chapter, the upper limit for the number of codewords is given. Perfect codes are described and the concept of code equivalence is defined. The Hamming $(8, 4)$ code, which detects two errors in message transmission is described. In the fourth chapter, linear codes are described and two important matrices are defined: generator matrix and parity check matrix. It is described how to use them to encode a word and how to check whether the word belongs to the code. In the same chapter, some methods of message decoding are described. In the fifth chapter, Hamming codes are formalized. Simplex codes and Hadamard codes are also described.

Životopis

Rođen sam 1997. godine u Šibeniku. U istom gradu 2016. godine završavam Gimnaziju Antuna Vrančića. Iste godine na Prirodoslovno-matemetičkom fakultetu u Zagrebu upisujem studij Matematike. Preddiplomski studij završavam 2020. godine i upisujem diplomski studij Računarstvo i matematika na istom fakultetu. Na kraju druge godine diplomskog studija dobivam posao u tvrtki Mireo.