

SVEUČILIŠTE U ZAGREBU
Fakultet elektrotehnike i računarstva

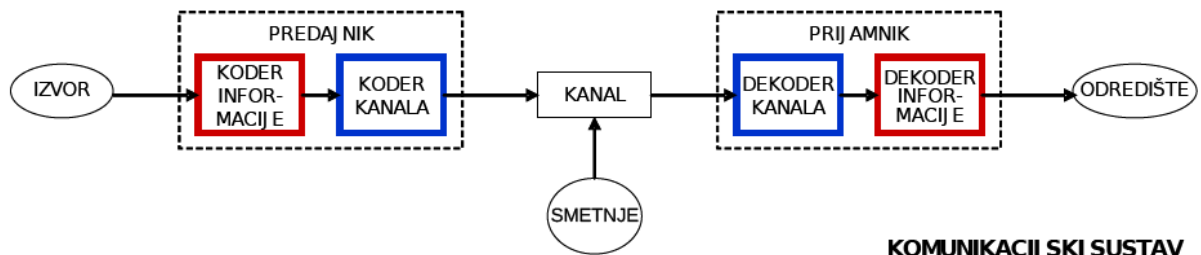
Predmet: Teorija informacije
Ak. godina: 2010./2011.

Laboratorijske vježbe:
I Z V J E Š T A J

Grupa MK17:

1. Mario Jagar, 0036437847
2. Mario Jambrišak, 0036426189
3. Domagoj Jerčić, 0036447294

Zadatak



Sl 1. Model komunikacijskog sustava i prijenosni kanal

Na sl. 1 zadan je komunikacijski sustav koji se sastoji od izvora informacije, koda informacije, koda kanala, kanala na koji utječu smetnje, dekoda kanala, dekoda informacije i odredišta. Izvorište generira simbole sa zadanim vjerojatnostima pojavljivanja.

Te simbole potrebno je kodirati LZW metodom entropijskog kodiranja, tako da je polazni rječnik definiran zadanom izvorišnom abecedom. Nakon toga, novonastale slijedove simbola (tzv. kodirana poruka) potrebno je zaštitno kodirati metodom cikličnog kôda zaštitnog čime dobivamo zaštitno kodiranu poruku. Zaštitno kodirana poruka prenosi se komunikacijskim kanalom u kojem:

- i) u jednom slučaju nema smetnji, dok
- ii) u drugom slučaju smetnje djeluju na poslane podatke, gdje je vjerojatnost pogreške bita $p = 1/5$.

Podatke koji su prošli kroz komunikacijski kanal potrebno je dekodirati na odredištu.

Simboli se na izvorištu pojavljuju sa vjerojatnostima:

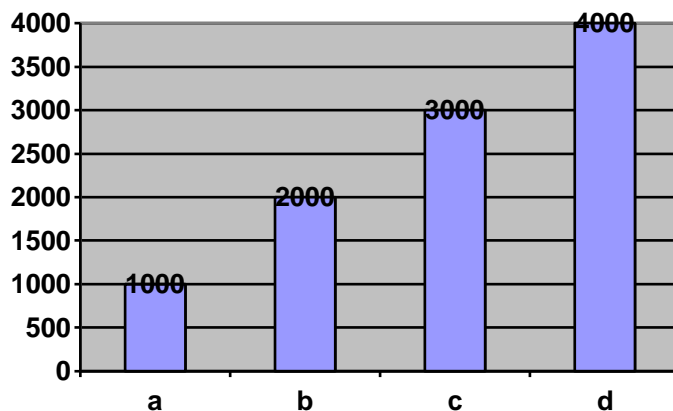
$$\begin{aligned}p(a) &= 0.1 \\p(b) &= 0.2 \\p(c) &= 0.3 \\p(d) &= 0.4\end{aligned}$$

Zadatak je generirati slučajni slijed od 10000 simbola u kojem će se simboli pojavljivati s zadanim vjerojatnostima i zadovoljavati sve prethodno navedene zahtjeve.

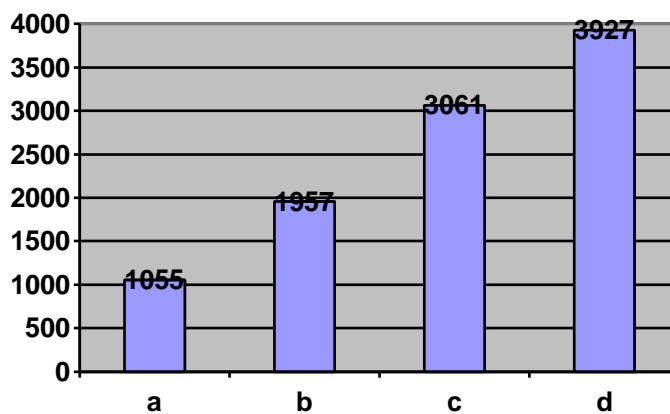
Rješenje

Informacijska svojstva kanala

1. Histogram generiranog slijeda.



Slika 2. Idealno izvoriste



Slika 3. Realno izvoriste

Vjerojatnosti pojavljivanja simbola u generiranom slijedu:

$$p(a) = \frac{1055}{10000} = 0.1055$$

$$p(b) = \frac{1957}{10000} = 0.1957$$

$$p(c) = \frac{3061}{10000} = 0.3061$$

$$p(d) = \frac{3927}{10000} = 0.3927$$

Kod realnog izvorišta dolazi do malog odstupanja od zadanih vrijednosti, u konkretnom primjeru odstupanje iznosi 0.0073, odnosno 0.73%.

2. Izračunajte entropiju izvorišnog skupa simbola.
 $H(\text{izv.}) = \underline{1.8464}$ [bita/simbolu]
3. Izračunajte entropiju odredišnog skupa simbola u slučaju da se koristi zaštitno kodiranje.
 $H(\text{odr.}) = \underline{1.2808}$ [bita/simbolu]
4. Izračunajte transinformaciju u slučaju da se koristi zaštitno kodiranje:
 $I(X;Y) = \underline{0.037295}$ [bita/simbolu]
5. Izračunajte entropiju odredišnog skupa simbola u slučaju da se NE koristi zaštitno kodiranje.
 $H(\text{odr.}) = \underline{1.281443}$ [bita/simbolu]
6. Izračunajte transinformaciju u slučaju da se NE koristi zaštitno kodiranje:
 $I(X;Y) = \underline{0.002141}$ [bita/simbolu]

Zaštitno kodiranje

Napomena: Sve niže definirane vrijednosti je generiranjem velikog broja ulaznih simbola potrebno odrediti statistički. Ako možete odrediti tu vjerojatnost računski, tada provjerite da li rezultat računa odgovara rezultatu dobivenom statističkom analizom.

1. Statistički dobivena udaljenost za zadani zaštitni kôd iznosi: $d(K) = 3$
2. Računski dobivena udaljenost za zadani zaštitni kôd iznosi $d(K) = 3$
3. Kôd može ispraviti 1 pogrešku.
4. Kôd može detektirati 2 pogreške.

Navedite primjer u kojem zaštitni kôd ne bi mogao ispraviti niti detektirati pogrešku:

Zaštitni kôd ne može detektirati niti ispraviti pogrešku u slučaju da se greška dogodi na točno 3 bita koliko iznosi udaljenost za zaštitni kôd, odnosno kada jedna kodna riječ prijeđe u drugu. Npr., kada se na kodnoj riječi:

000000001011111

dogodi greška na 12., 14. i 15. bitu te ona prijeđe u kodnu riječ:

000000001001100

Vjerojatnost ispravnog dekodiranja poruke iznosi (statistički dobiven podatak):
0.044

5. Kodna brzina zaštitnog kôda iznosi: $R = \frac{11}{15}$
6. Navedite generirajuću matricu **G**:

$$G =$$

8. Navedite matricu provjere pariteta \mathbf{H} za zadani kôd:

$$H =$$

Entropijsko kodiranje

1. Prosječna duljina kodne riječi: $L = 15 \text{ bita/simbolu}$
2. Za zadani kôd navesti postupak kodiranja zadanog izvorišta. Ako se kodiranje temelji na rječniku, tada je potrebno kodirati proizvoljni niz simbola. U oba slučaja potrebno je postupak potkrijepiti detaljnim komentarima.

Kodiranje se provodi algoritmom LZW koji se temelji na rječniku. Pseudo kod po kojem se provodi LZW kodiranje:

1. RadnaRiječ = sljedeći simbol poruke s ulaza
2. WHILE (ima još simbola na ulazu) DO
3. NoviSimbol = sljedeći simbol s ulaza
4. IF RadnaRiječ + NoviSimbol postoji u rječniku THEN
5. RadnaRiječ = RadnaRiječ + NoviSimbol
6. ELSE
7. IZLAZ: kod za RadnaRiječ
8. dodaj (RadnaRiječ + NoviSimbol) u rječnik i dodijeli tom izrazu novi kod
9. RadnaRiječ = NoviSimbol
10. END IF
11. END WHILE
12. IZLAZ: kod za RadnaRiječ

Za primjer ćemo kodirati niz:

MJESTO	1	2	3	4	5	6	7
SIMBOL	a	b	b	a	d	b	c

Potrebno je definirati početni rječnik. Definiramo ga kao:

Indeks	Riječ
(1)	a
(2)	b
(3)	c
(4)	d

Korištenje pseudokoda korak po korak možemo ukratko prikazati u sljedećoj tablici:

KORAK	MJESTO	SADRŽAJ RIJEČNIKA	IZLAZ
1	1	(5) ab	(1)
2	2	(6) bb	(2)
3	3	(7) ba	(2)
4	4	(8) aba	(4)
5	5	(9)ad	(1)
6	6	10 (db)	(4)
7	7	11 (bc)	(2)
8	8		(3)

Dodatak

Prilikom izrade laboratorijske vježbe javljali su se i neki problemi. Slijedi dio komunikacijskog kanala gdje se problem javio, te kratak opis i uspješnost rješenja:

LZW koder: problem pri implementaciji – problem sa zadnjim znakom u nizu. Npr: pojavi se *ab* koji nije u riječniku, u riječnik se dodaje *ab*, te *b* postaje radna riječ, a pošto nema više znakova, znak *b* se dodaje u riječnik. Problem je uspješno riješen.

LZW dekode: problem pri implementaciji – kada naiđe znak koji se ne nalazi u riječniku. U tom slučaju na radnu riječ se dodaje prvi znak te riječi te se radna riječ stavlja u riječnik. Nova radna riječ je trenutna radna riječ (znači radna riječ se ne mijenja). Problem je uspješno riješen.