Dejan E. LAZIĆ
Vojin ŠENK
Radomir ZAMUROVIĆ
Fakultet tehničkih nauka
Institut za računare, automatiku i merenja
Novi Sad, V.Vlahovića 3.

O SOPSTVENOM KAPACITETU ZAŠTITNIH BLOK KODOVA

ON THE INDIVIDUAL CAPACITY OF BLOCK ERROR CORRECTING CODES

SADRŽAJ - U ovom radu sistematski je analiziran problem odredjivanja asimptotskih performansi zaštitnih blok kodova pri porastu dimenzije kodnih reči. Pokazano je da za vremenski diskretan Gausov kanal bez memorije postoje blok kodovi kod kojih kodne reči na minimalnom rastojanju ne utiču na prosečnu verovatnoću greške pri vrlo velikim dimenzijama. Ovo je pokazano na primeru BCH kodova. Na osnovu te osobine postavljena je hipoteza o egzistanciji sopstvenog kapaciteta pojedinih klasa blok kodova, koji je različit od nule, a manji ili jednak kapacitetu kanala.

ABSTRACT - The problem of determining the asymptotical performances of block error correcting codes of large dimensions is systematically analysed in this paper. It is demonstrated that for a time discrete memoryless Gaussian channel there exists block codes whose probability of error is unaffected by the codewords on the minimal distance in high dimensions. This is shown on the BCH code example. According to this property, a hypothesis is made that there exists an individual capacity of certain classes of block codes, different from zero, and less than or equal to the channel capacity.

1. U V O D

Svaki zaštitni blok kod C predstavlja skup od M različitih nizova *simbola*

$$C = C(M,N,A) = \{\vec{X}_i = (X_{i1},\ldots,X_{iN}) \mid i=1,\ldots,M\},$$

koji čine *kodne reči* jednakih *dimenzija* N, pri čemu svi simboli pripadaju nekom *alfabetu* A. Alfabet može imati konačan broj simbola q, kada je $A=A_q = \{S_1,\ldots,S_q\}$, ili neograničen broj simbola, kada je $A=A_E \subset \mathbb{R}$, gde je $\mathbb{R}$ skup realnih brojeva. Blok kodovi nad konačnim alfabetom A_q nazivaju se *diskretni blok kodovi*, a nad alfabetom A_E *euklidski blok kodovi* (jer se uvek mogu predstaviti u N-dimenzionalnom euklidskom prostoru $\mathbb{R}^N$).

Euklidski blok kod je *energetski ograničen* po srednjoj energiji E_s ako je ispunjen sledeći uslov

$$\sum_{i=1}^{M} P(\vec{X}_i)|\vec{X}_i|^2 = E_s, \tag{1}$$

a ako je

$$|\vec{X}_i|^2 = E \quad \text{za} \quad i=1,\ldots,M, \tag{2}$$

onda je reč o *sferičnom blok kodu* čije sve kodne reči leže na površini N-dimenzionalne hipersfere Ω_N poluprečnika $r = \sqrt{E}$

Svaki blok kod se može snabdeti nekom metrikom d koja odredjuje *rastojanja* $d(\vec{X}_i,\vec{X}_\ell)$ izmedju kodnih reči. Najčešće se diskretni blok kodovi snabdevaju Hemingovom metrikom d_H, a euklidski blok kodovi Euklidovom metrikom d_E. Skup svih rastojanja $D = \{d_{i\ell} = d(\vec{X}_i,\vec{X}_\ell) \mid i < \ell\}$ bitno karakteriše svaki blok kod.

Svaki diskretan blok kod uvek se može preslikati u euklidski blok kod pomoću neke injektivne funkcije

$$f_E : C(M,N,A_q) \to C(M,N,A_E). \tag{3}$$

Ovo preslikavanje omogućava da se svi zaštitni blok kodovi analiziraju u $\mathbb{R}^N$. Medjutim, ovakva analiza ima smisla ako preslikavanje f_E čuva usvojenu metriku za diskretne kodove u domenu (iskazuje je kroz euklidsku metriku u svom kodomenu).

Kada se bilo koja kodna reč nekog blok koda C(N,M,A) propusti kroz *kodni kanal* K na njegovom izlazu se dobija *odziv kodnog kanala* $\vec{y}$, na osnovu koga *zaštitni dekoder* donosi odluku koja je kodna reč bila poslata. Ako je odluka zaštitnog dekodera $\hat{\vec{X}}$, a poslata kodna reč $\vec{X}_i$ verovatnoća

$$P(\hat{\vec{X}} \neq \vec{X}_i|\vec{X}_i,\vec{y}) = P_{ei} \tag{4}$$

predstavlja *verovatnoću greške (dekodovanja)* i-te *kodne reči. Prosečna verovatnoća greške (dekodovanja)* blok koda C(N,M,A) je

$$P_e = \sum_{i=1}^{M} P(\vec{X}_i)\, P_{ei}. \tag{5}$$

Ova verovatnoća, u opštem slučaju, zavisi od kodnog kanala, blok koda (zaštitnog kodera), zaštitnog dekodera i apriornih verovatnoća kodnih reči.

Uobičajeno je da se u analizi performansi zaštitnog kodovanja uzimaju uniformne apriorne verovatnoće kodnih reči tako da je $P(\vec{X}_i) = 1/M$; $i=1,\ldots,M$. Za dati kodni kanal K uvek postoji zaštitni dekoder koji odlučuje sa najmanjom verovatnoćom greške P_e, bez obzira na zaštitni blok kod koji se kroz kodni kanal propušta. Takav dekoder se naziva *optimalni zaštitni dekoder.*

2. ASIMPTOTSKE PERFORMANSE ZAŠTITNIH BLOK KODOVA

Skup svih blok kodova $\{C_A\}$ nad nekim alfabetom A može se potpuno razložiti na klase $\{C_A(N,M)\}$ po različitim vrednostima parametara N i M, ili na klase $\{C_A(R)\}$ po različitim vrednostima *bitske kodne brzine* $R=(\ell dM)/N$. Jedna klasa $\{C_A(R)\}$ predstavlja uniju svih klasa $\{C_A(N,M)\}$ za koje je odnos $R=(\ell dM)/N$; $(R \in \mathcal{B} \subset \mathbb{R})$ konstantan. Svaka klasa $\{C_A(R)\}$ može se urediti po rastućim dimenzijama, tako da je

$$\{C_A(R)\} = \{C_A(N_1,M_1)\} \cup \ldots \cup \{C_A(N_\ell,M_\ell)\} \cup \ldots , \qquad (6)$$

pri čemu je $N_\ell < N_{\ell+1} \Rightarrow M_\ell < M_{\ell+1}$, $(\ell dM_\ell)/N_\ell = R$ za svako $\ell=1,2,\ldots$. Ako se za svako ℓ iz odgovarajuće klase $\{C_A(N_\ell,M_\ell)\}$ na neki način izabere samo po jedan blok kod $C_A(N_\ell,M_\ell)=C_A(N_\ell,R)$, dobija se niz blok kodova konstantne bitske kodne brzine nad alfabetom A uredjen po rastućim dimenzijama

$$\mathcal{R}(R,A) = (C(N_1,R),\ldots,C(N_\ell,R),\ldots), \qquad (7)$$

koji se može nazvati *R-nizom* zaštitnih blok kodova.

Članovi *diskretnog* R-niza $\mathcal{R}(R,q)$ su diskretni blok kodovi nad nekim alfabetom A_q sa q simbola, tako da razlikujemo *binarne* (q=2), *ternarne* (q=3), ... i q-*arne* R-nizove. Članovi *energetski ograničenog* R-niza $\mathcal{R}(R,E_s)$ su euklidski blok kodovi sa jednakim srednjim energijama E_s. Članovi *sferičnog* R-niza $\mathcal{R}(R,E)$ su sferični blok kodovi sa jednakim energijama E. Ako je E=1 $(r = \sqrt{E} = 1)$, reč je o *normalizovanom sferičnom* R-nizu $\mathcal{R}(R,1)$. Članovi *optimalnog* R-niza $\mathcal{R}_{opt}(R,\mathcal{K})$ za dati kodni kanal $\mathcal{K}$ su *optimalni blok kodovi,* koji u odgovarajućoj klasi $\{C_A(N_\ell,N_\ell) = C_A(N_\ell,R)\}$ imaju najmanju verovatnoću P_e kada se dekoduju pomoću optimalnog dekodera za dati kodni kanal $\mathcal{K}$. Optimalni R-niz može biti: (1) diskretan $\mathcal{R}_{opt}(R,\mathcal{K},q)$, (2) energetski ograničen $\mathcal{R}_{opt}(R,\mathcal{K},E_s)$ i (3) sferičan $\mathcal{R}_{opt}(R,\mathcal{K},E)$. Članovi *asimptotski dobrog* R-niza $\mathcal{R}(Rd^*)$ su blok kodovi za koje je normalizovano

minimalno rastojanje $d_m^* = d_m/E_s$; ($d_m = \min_{i<\ell} \{d_{i\ell}\}$) uvek veće ili jednako od neke pozitivne konstante d* (pri čemu je srednja energija za diskretne kodove E_s=N). Pokazano je [2] da postoje (1) diskretni $\mathcal{R}(R,q,d_H^*)$, (2) energetski ograničeni $\mathcal{R}(R,E_s,d_E^*)$ i (3) sferični $\mathcal{R}(R,E,d_E^*)$ asimptotski dobri R-nizovi. Članovi *slučajnog* R-niza $\mathcal{R}_s(R)$ su blok kodovi koji se biraju na slučaj (po uniformnom zakonu) iz odgovarajućih klasa $\{C_A(N_\ell,R)\}$. Slučajan R-niz može biti (1) diskretan $\mathcal{R}_s(R,q)$, (2) energetski ograničen $\mathcal{R}_s(R,E_s)$ i (3) sferičan $\mathcal{R}_s(R,E)$.

O osobinama R-nizova zaštitnih blok kodova se veoma malo zna. Osnovni i najznačajniji dosadašnji rezultat je Fundamentalna teorema teorije informacija, koju je postavio i dokazao C.Shannon [1]. Ova teorema je do danas dokazana za široku klasu raznih tipova kodnih kanala i sadrži dva stava:

- *pozitivan stav:* za članove slučajnog R-niza prosečna verovatnoća greške P_e teži nuli, ako i samo ako je njihova bitska kodna brzina manja od odredjene vrednosti R_c, koja se naziva *kapacitet kodnog kanala;*

- *negativan stav:* za bitske kodne brzine iznad kapaciteta kanala prosečna verovatnoća greške teži jedinici u svakom R-nizu.

Kapacitet kanala je jedan od najznačajnijih parametara teorije informacija koji zavisi od osobina kodnog kanala i podrazumeva optimalan zaštitni dekoder. Do danas nije poznat ni jedan neslučajan R-niz koji zadovoljava pozitivan stav Fund. teor. teor. inform. Jedino je A.Wyner [2] uspeo da dokaže da (pored $\mathcal{R}_{opt}(R,\mathcal{K})$, što je po definiciji jasno) još kod asimptotski dobrih R-nizova P_e teži nuli za vrednosti R koje su znatno manje od kapaciteta kanala, dok za ostale brzine (do kapaciteta) ponašanje ovih nizova nije poznato. Medjutim, postupak za konstrukciju članova $\mathcal{R}_{opt}(R,\mathcal{K})$ i članova $\mathcal{R}(R,d^*)$ za sada je praktično izvodljiv samo za relativno male vrednosti dimenzije N; [5].

3. PRIKAZ DOBIJENIH REZULTATA

Svi do danas poznati R-nizovi (diskretni i euklidski), kod kojih se za sve članove zna postupak konstrukcije, nisu asimptotski dobri, već njihovo normalizovano minimalno rastojanje uvek teži ka nuli. Ovakvi R-nizovi (zaštitni kodovi) se tretiraju kao neperspektivni u vrlo visokim dimenzijama, jer se veruje (mada to u opštem slučaju nije dokazano) da sa opadanjem normalizovanog minimalnog rastojanja ka nuli verovatnoća greške P_e mora da

raste ka jedinici. Ovo je slučaj (dokazan)i za binarne BCH kodove u binarnom simetričnom kodnom kanalu (BSC). Medjutim, autori ovog rada *pretpostavljaju* da će, kad se binarni BCH kodovi preslikaju pomoću funkcije (3) u obliku f_E : „0" → $-\sqrt{E/N}$, „1" → $+\sqrt{E/N}$; u odgovarajuće sferične (euklidske) BCH kodove, verovatnoća greške težiti nuli u takvom R-nizu, ako je u pitanju vremenski diskretan Gausov kodni kanal bez memorije sa optimalnim dekoderom po minimalnom rastojanju, bez obzira što minimalno normalizovano euklidsko rastojanje takodje teži nuli. Verovatnoća greške bi trebala da opada ka nuli u ovakvom R-nizu, koji ćemo označiti sa $\mathcal{R}$(R,BCH,E), za sve brzine $R \in B \subset \mathbb{R}$ manje od neke granične R_{BCH}, koja bi se nazvala *sopstveni kapacitet* BCH kodova za vremenski diskretan Gausov kanal bez memorije.

Osnovu za gore navedenu pretpostavku autori vide u činjenici da deo verovatnoće greške P_e koji potiče od svih kodnih reči na minimalnom (euklidskom) normalizovanom rastojanju d_m^* teži nuli u $\mathcal{R}$(R,BCH,E) za sve R<1 i nezavisno od odnosa *signal-šum* $SNR = E/N\cdot\sigma^2$, (gde je σ^2 varijansa jedne komponente Gausovog šuma), iako d_m^* takodje teži nuli. Ova činjenica će biti nadalje dokazana.

Prosečna verovatnoća greške u blok kodu sa dve kodne reči na euklidskom rastojanju d_E je data relacijom:

$$P_e = P_{e1} = P_{e2} = P_e\,(M=2,\ d_E) = Q\left(\frac{d_E}{\sigma}\right), \tag{8}$$

gde je

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-z^2/2)dz. \tag{9}$$

Deo verovatnoće greške dekodovanja i-te kodne reči u nekom euklidskom blok kodu $C(M,N,A_E)$, koji potiče od uticaja M_m kodnih reči na minimalnom rastojanju d_{Em} od i-te kodne reči, očigledno će biti

$$P_{ei}\,(M_m,d_{Em}) \leq M_m\cdot P_e(M=2,d_{Em}). \tag{10}$$

Za BCH kodove, koji su linearni, uvek će biti $P_e(M_m,d_{Em}) = P_{ei}(M_m,d_{Em})$, kada se prevedu iz diskretnog u euklidski oblik pomoću preslikavanja (3) datog sa f_E : „0" → $-\sqrt{E/N}$; „1" → $+\sqrt{E/N}$. Berlekamp [3] je pokazao da je minimalno Hemingovo rastojanje d_{Hm} za BCH kodove u diskretnom obliku, pri velikim dimenzijama N dato sa

$$d_{Hm} \sim 2N\cdot[\mathfrak{L}n(1/R)/\mathfrak{L}d(N)], \tag{11}$$

koje se pomoću gore navedenog preslikavanja prevodi u ekvivalentno euklidsko rastojanje

$$d_{Em} \sim \sqrt{8E} \cdot \sqrt{\ell n(1/R)/\ell d(N)}. \qquad (12)$$

Procenu broja kodnih reči u BCH kodu na minimalnom rastojanju M_m dao je Sideljnikov [4]:

$$M_m \sim \frac{1}{2^{N(1-R)}} \binom{N}{d_{Hm}} \quad (1+C_1 N^{-0,1}), \qquad (13)$$

gde je C_1 neka konstanta.

Korišćenjem Stirlingove formule, binomni koeficijent u prethodnom izrazu može se aproksimirati sa

$$\binom{N}{d_{Hm}} \sim \frac{[\mu^{\mu} \cdot (1-\mu)^{(1-\mu)}]^N}{[2\pi N\mu\,(1-\mu)]^{\frac{1}{2}}} \qquad (14)$$

gde je

$$\mu = 2\,\frac{\ell n(1/R)}{\ell d(N)}\,.$$

Zamenom (12) u (8), a zatim (8), (13) i (14) u izraz (10) i posmatranjem graničnog slučaja kad $N \to \infty$ dobija se:

$$P_e\,(M_m, d_{Em}) \sim \lim_{N\to\infty} 2^{N(R-1)+O[1/\ell d(N)]} \qquad (15)$$

Kako je za BCH kodove uvek $R<1$, to $P_e(M_m, d_{Em})$ sigurno teži nuli kada $N \to \infty$, nezavisno od odnosa signal-šum.

4. ZAKLJUČAK

Na osnovu prethodnih razmatranja može se zaključiti da kod BCH kodova u vremenski diskretnom Gausovom kanalu bez memorije kodne reči na minimalnim rastojanjima uopšte ne utiču na povećanje prosečne verovatnoće greške pri vrlo velikim dimenzijama. Razlog tome je što tada takvih kodnih reči ima relativno malo u odnosu na udaljenije kodne reči, kojih ima značajno više. Iz tih razloga te udaljenije reči imaju veći ukupni uticaj na verovatnoću greške, bez obzira što su na većem rastojanju. Ukoliko bi se pokazalo da za kodne reči na svim rastojanjima postoje vrednosti parametra

SNR $\neq$ 0 pri kojima P_e teži nuli, BCH kodovi bi imali sopstveni kapacitet različit od nule a manji ili jednak kapacitetu kanala.

5. LITERATURA

[1] Shannon, C.E., "A Mathematical Theory Communication", BSTJ, Vol. 27, 379-423 i 623-656, 1948.

[2] Wyner, A.D., "Capabilities of Bounded Discrepancy Decoding, BSTJ, Vol. 44, 1061-1122, 1965.

[3] Berlekamp, E.R., "Long Primitive BCH codes have distance $d \sim 2N \ell n R^{-1}$ / /$\ell d(N)$ ...." IEEE Trans. Inf. Theory, Vol. 18, 415-426, 1972.

[4] Sideljnikov, V.M., "O spektre vesov dvoičnih kodov BCH", Problemi peredači inf., Vol. VII, 14-22, 1971.

[5] Lazić, D.E., Bece, T.B., "Generisanje najboljih sferičnih blok kodova metodom sukcesivnih aproksimacija", XXVI konf. ETAN-a, Subotica, II.375--II.382, 1982.