

TEHNIKE PRENOSA V DOSTOPOVNIH OMREŽJIH

Sašo Tomažič, Fakulteta za elektrotehniko, Univerza v Ljubljani

Povzetek—V prispevku je narejen kratek pregled različnih tehnik kodiranja in modulacije ter tehnik sodostopa do skupne prenosne poti, ki se uporabljajo v dostopovnih omrežjih

Ključne besede—kodiranje, modulacija, sodostop

I. UVOD

Prav gotovo nismo prav daleč od resnice, če zatrdimo, da čas ob koncu dvajsetega in na začetku enaindvajsetega stoletja označuje izreden razvoj predvsem pa globalizacija telekomunikacij. Marsikomu se bo morda zdelo, da ta trditev ni upravičena, saj predstavlja razvoj na področju genetike in razkritje celotnega genoma bistveno večji korak, ki ima lahko še mnogo večje in nepredvidljive posledice za razvoj človeške vrste. Mogoče ima celo prav. Zelo verjetno bo razvoj na tem področju omogočil ozdraviti marsikatero neozdravljivo bolezen, morda bodo uspeli bistveno podaljšati življenjsko dobo, nadomestiti poškodovane ali izrabljene organe in kdo ve kaj vse. Morda, vendar je razvoj na področju genetike približno tam, kjer je bil razvoj telekomunikacij, ko je Samuel F. B. Morse razmišljal, kako bi prenesel sporočilo na daljavo in prav gotovo tudi o tem, kakšne posledice bi to lahko imelo. Po drugi strani pa je razvoj telekomunikacij na točki, ko je resno posegel v naše življenje, spremenil naše navade naš način dela in zabave ter začel spreminjati družbo kot celoto.

Razvoj tehnike je omogočil izgradnjo vedno večjih in vedno bolj razvejanih omrežij, prenosne hitrosti po teh omrežjih pa so narasle za nekaj velikostnih razredov, kar omogoča ponudnikom storitev, da zadovoljujejo potrebe vse zahtevnejših uporabnikov in ponudijo storitve, ki zahtevajo visoke pasovne širine, prenos v realnem času, veliko zanesljivost in tudi veliko dostopnost: kjerkoli in kadarkoli. In ravno to zadnje je tema 11. delavnice o telekomunikacijah Vitel.

Da bi omogočili uporabnikom dostop storitev, ki zahtevajo veliko pasovno širino, predvsem sproten prenos video signala, ne zadošča samo velika zmogljivost hrbteničnega omrežja, temveč je potrebno zagotoviti veliko zmogljivost povezave do končnega uporabnika, to je do njegove pisarne, njegovega doma in tudi kjerkoli na terenu. Visoke zmogljivosti optičnih povezav omogočajo relativno enostavno povečevanje zmogljivosti hrbteničnega omrežja, izgradnja novih optičnih povezav do pisarne predvsem pa do doma, pa

bi bila zaenkrat za večino uporabnikov predraga, za mobilne uporabnike pa tako ali tako nedostopna.

Zmogljivost, ki je na voljo uporabnikom, je tako še vedno odvisna od zmogljivosti prenosne poti, ki je tem uporabnikom na voljo, to pa je lahko bakrena parica v telefonskem kablu, koaksialni kabel kabelskega omrežja ali pa omejen frekvenčni pas pri mobilnih uporabnikih. Vprašanje je, kaj omejuje zmogljivost teh prenosnih poti in kako te prenosne poti čim bolje izkoristiti.

V tem prispevku želim predstaviti osnovne mehanizme, ki omejujejo zmogljivosti prenosne poti in tehnike prenosa, ki omogočajo bolj ali manj optimalno izkoriščanje teh zmogljivosti. Pri tem ne želim zaiti pregloboko v teorijo komunikacij in v utemeljevanje teoretičnih rezultatov, kot tudi ne v tehnične podrobnosti, ki so potrebne pri načrtovanju prenosnih sistemov. Rezultati teorije telekomunikacij nam bodo služili zgolj za oceno uspešnosti različnih tehnik prenosa in dali odgovor na vprašanje kaj je moč od njih največ pričakovati.

II. ZMOGLJIVOST PRENOSNE POTI

Da bi lahko govorili o zmogljivosti prenosne poti, moramo najprej razjasniti, kaj pravzaprav prenašamo in kaj razumemo pod pojmom zmogljivost.

Kadar v telekomunikacijah prenašamo signalne oblike, ki neposredno ustrezajo obliki signala (govornega, avdio, video, ...), ki ga v osnovi želimo prenašati, govorimo o analognem prenosu, kadar pa prenašamo sporočilo, ki ga ne predstavlja signalna oblika temveč je zapisano z nizom znakov (simbolov), govorimo digitalnem prenosu.

Pri analognem prenosu zelo težko enoumno definiramo zmogljivost, saj na kvaliteto prenosa vpliva pasovna širina, linearno in nelinearno popačenje, šum in slabljenje.

Pri digitalnem prenosu je zmogljivost nekoliko lažje definirati. Vsako sporočilo lahko na namreč zapišemo kot niz binarnih simbolov – bitov, to je simbolov, ki lahko zavzamejo le dve vrednosti: nič ali ena $\{0,1\}$. Zmogljivost prenosne poti je zato določena z maksimalnim bitnim pretokom, to je z največjim številom bitov, ki jih lahko po prenosni poti prenesemo v eni sekundi.

Samo podatek o največjem bitnem pretoku pa še ne pove dovolj o zmogljivosti. Pri prenosu lahko namreč pride do napake, to je, pri sprejemu vrednost 0 zamenjamo z vrednostjo 1 ali obratno. Teoretično zato

govorimo o največjem možnem bitnem pretoku brez napake, praktično pa govorimo o maksimalnem pretoku ob sprejemljivi pogostosti napak - BER (Bit Error Rate). Danes se zahtevane vrednosti BER gibljejo med 10^{-7} pa vse do 10^{-12} , kar pomeni, da je v povprečju le na vsakih 10 milijonov oziroma na vsakih 1000 milijard bitov eden sprejet napačno.

Vendar tudi podatek o maksimalnem pretoku brez napake ali maksimalnem pretoku pri omejeni pogostosti napaki ne pove dovolj. Za zapis sporočila, ki ga želimo prenašati smo lahko namreč porabili več bitov, kot bi bilo to nujno potrebno. Tedaj pravimo da so v zapisu odvečni biti ali redundanca. Večino sporočil, ki jih neposredno pretvorimo v binarno obliko, vsebuje redundanco, ki jo je mogoče delno izločiti z uporabo ustreznega kodiranja (npr. pkzip) in na ta način zapis skrajšati. Če sporočila pred prenosom skrajšamo, lahko pri istem enakem bitnem pretoku v istem času prenesemo več sporočil, kar je enako, kot da bi povečali bitni pretok.

Da bi lahko resnično ocenili prenosno zmogljivost moramo zato dopustiti tudi stiskanje sporočila pred oddajo in njegovo razširjanje pri sprejemu. V definicijo zmogljivosti prenosne poti zato dodamo še pogoj, da gre za prenos sporočil iz katerih je izločena vsa redundanca. Največji možen bitni pretok pri prenosu sporočila brez redundance, ne da bi naredili napako pri sprejemu, imenujemo tudi kapaciteta prenosne poti oziroma *kapaciteta komunikacijskega kanala*. Kapaciteta kanala predstavlja zgornjo teoretično mejo, ki se ji v praksi lahko le približamo. Koliko smo se tej meji približali, pa je lahko zelo dober kriterij za oceno uspešnosti nekega postopka prenosa.

Taka definicija zmogljivosti lahko dokaj dobro služi tudi za oceno zmogljivosti prenosne poti pri analognem prenosu, saj lahko vsak pasovno omejen analogen signal enoumno zapišemo z njegovimi vzorci. Za to potrebujemo najmanj dva vzorca na sekundo za vsak Hz pasovne širine signala. Vsak vzorec lahko nato zapišemo s končnim številom bitov, pri čemer velja, da z dodajanjem vsakega bita zmanjšamo kvantizacijski šum za 6 dB. Tako lahko določimo tudi bitni pretok, ki bi bil potreben za tak signal in ocenimo ali je prenosni kanal dovolj zmogljiv.

III. KAJ OMEJUJE ZMOGLJIVOST

V prejšnjem razdelku smo definirali kapaciteto kanala, vprašati pa se moramo, kaj to kapaciteto omejuje. Zakaj je po nekaterih prenosnih poteh moč prenašati mnogo večje bitne pretoke kot po drugih? Zakaj ni mogoče prenašati poljubno velikega bitnega pretoka?

Za prenos v sodobnih telekomunikacijskih sistemih uporabljamo skoraj izključno elektromagnetno valovanje, pa naj gre za prenos po žičnih vodih, optičnih vlaknih ali atmosferi. Niz simbolov moramo zato pred prenosom pretvoriti v elektromagnetni signal, ki se širi po prenosni poti. Na prenosni poti se ta signal pokvari, kar povzroči napačno prepoznavo pri

sprejemu in privede do napake. Poglavitni vzroki za to so:

- slabljenje signala,
- omejitev pasovne širine,
- popačenje oblike signala in
- dodajanje motenj oziroma šuma signalu.

Različne prenosne poti različno vplivajo na signal. Glede na vse zgornje kriterije imajo optična vlakna prednost pred vsemi drugimi prenosnimi potmi, zato imajo tudi največjo zmogljivost.

Slabljenje samo po sebi ne predstavlja večjega problema, ker lahko signal ponovno ojačimo in tako pravzaprav ničesar ne izgubimo.

Pasovna širina je lahko omejena s samimi lastnostmi prenosne poti lahko pa je omejena tudi zato, ker več uporabnikov uporablja isto prenosno pot in je posameznemu uporabniku le določen frekvenčni pas. Pasovna širina signala omejuje maksimalno število simbolov, ki jih je mogoče prenašati po neki prenosni poti, na da bi se ti medsebojno prekrivali. Prenašati je mogoče namreč največ dva simbola vsako sekundo na vsak Hz pasovno širine. Če imamo torej na voljo frekvenčni pas 1000 Hz, je mogoče v eni sekundi prenesti največ 2000 simbolov.

Popačenje signala je posledica neravne frekvenčne karakteristike prenosne poti. Frekvenčno karakteristiko lahko izravnamo z uporabo ustreznega sprejemnega sira (izravnalnika), ki ima prenosni poti inverzno frekvenčno karakteristiko. Da bi to lahko naredili, bi morali vnaprej poznati karakteristiko prenosne poti. Ker te karakteristike običajno ne poznamo vnaprej, moramo uporabiti sito, ki se je sposobno samo prilagajati (adaptivni izravnalnik).

Običajno pa izravnavo frekvenčne karakteristike povzroči poslabšanje razmerja med šumom in koristnim signalom pri sprejemu. Večji problem predstavlja tudi izravnavo frekvenčne karakteristike prenosne poti v mobilnih komunikacijah, kjer se ta zaradi gibanja mobilnega uporabnika stalno in zelo hitro spreminja.

Šum je v največji meri posledica zunanjih motilnih virov. V atmosferi je teh izredno veliko, pri prenosu po kovinskem vodniku v kablu pa so to predvsem sosednji vodniki, po katerih se odvija prenos. To vrsto šuma imenujemo presluh. Tudi, če ne bi bilo zunanjih virov šuma, se šumu ne bi mogli popolnoma izogniti, ker nekaj šuma nastaja zaradi samih fizikalnih lastnosti prenosne poti in oddajnika in sprejemnika. Te vrste šum je poglaviten pri prenosu po optičnih vlaknih.

Če v sprejetem signalu ne bi bilo šuma, bi lahko z enim samim simbolom prenašali poljubno število bitov. En bit namreč predstavlja 2 različni vrednosti, dva bita 4, trije biti 8 in tako naprej. Tako je na primer prenos simbol as 1024 različnimi vrednostmi enakovreden prenosu 10 bitov. Ravno šum, ki je pri

soten v sprejetem signalu pa onemogoča prenos prevelikega števila različnih vrednosti simbola, ker onemogoča ločevanje med posameznimi vrednostmi.

Če poznamo frekvenčno karakteristiko kanala $S(f)$, frekvenčno karakteristiko šuma pri sprejemu $N(f)$ in pasovno širino, ki jo omejujeta spodnja frekvenčna meja f_s in zgornja frekvenčna meja f_z , lahko izračunamo z upoštevanjem vseh zgoraj naštetih omejitev kapaciteto komunikacijskega kanala po enačbi:

$$C = \int_{f_s}^{f_z} \log_2 \left(1 + \frac{S(f)}{N(f)} \right) df$$

(1)

Kot smo že omenili predstavlja zgornja enačba zgolj teoretično mejo, ki bi jo bilo mogoče doseči le, če bi v celoti izkoristili prenosno pot z uporabo optimalnih postopkov kodiranja in dekodiranja, ki pa so v celoti praktično neizvedljivi.

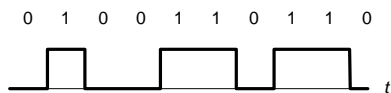
IV. PRENOS V OSNOVNEM FREKVENČNEM PASU

O prenosu v osnovnem frekvenčnem pasu govorimo, kadar uporabljamo za prenos frekvenčno področje od frekvence nič naprej. Tak prenos se uporablja predvsem pri prenosu po kovinskih vodnikih (paricah), ki jih ima uporabnik na voljo popolnoma za svoje potrebe in jih ni potrebno deliti z drugimi uporabniki.

Čeprav bi bili za prenos v osnovnem pasu možni tudi drugačni načini, se za tak prenos najpogosteje uporablja prenos simbolov z impulzi omejenega trajanja (dolžine), pri čemer je vrednost simbola zapisana v amplitudi impulza. Uporabljamo lahko impulze različnih oblik in simbole z različno velikimi nabori vrednosti.

A. Binarni prenos

Najbolj preprost je vsekakor neposreden prenos binarnih simbolov s pravokotnimi impulzi. Za tak primer je oblika signala prikazana na sliki 1.



Slika 1: Binarni prenos

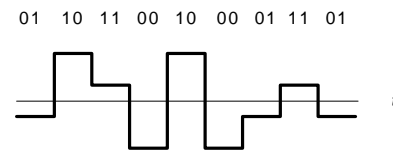
Tak način prenosa bomo uporabili, predvsem zaradi njegove preprostosti, takrat kadar imamo na voljo precej večjo kapaciteto prenosne poti, kot jo potrebujemo. Njegove osnovne slabosti so:

- Ker smo se omejili na binarne simbole prenašamo le en bit z vsakim simbolom, čeprav bi nam morda šumne razmere dopuščale prenos bistveno večjega števila bitov na simbol.

- V signalu je prisotna enosmerna komponenta, kar zahteva galvanski spoj oddajnika in sprejemnika na kovinski vodnik.
- Pri dolgih zaporedjih samih enic oziroma samih ničel lahko izgubimo podatek za sinhronizacijo simbolnega takta.
- Zaradi pravokotne oblike impulzov je frekvenčna širina signala izredno velika, za prenos pa nujno potrebujemo pasovno širino, ki je vsaj dvakrat večje od minimalne. Po frekvenčnem pasu 1000 Hz bi na ta način lahko prenašali največ 1000 bit/s, ne glede na to, kakšne so šumne razmere.

B. Večnivojski prenos

Precej slabosti, ki jih ima binarni prenos lahko odpravimo z uporabo večnivojskega prenosa. Primer signala, ki ga prenašamo pri štirinivojskem prenosu, je prikazan na sliki 2.



Slika 2: Štiri nivojski prenos

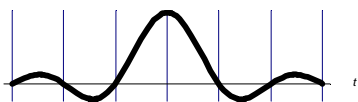
Oglejmo si nekaj osnovnih lastnosti večnivojskega prenosa.

- Ker pasovna širina signala ni odvisna od števila nivojev, je pri večnivojskem prenosu enaka kot pri binarnem prenosu, čeprav smo povečali bitni pretok.
- S povečanjem števila nivojev se pri isti razdalji med nivoji poveča moč signala. Če ohranimo enako moč signala, se torej zmanjša razdalja med nivoji.
- Ker je od razdalje med sosednjimi nivoji odvisna občutljivost na šum, se z večanjem števila nivojev povečuje verjetnost napake. Šum torej omejuje število nivojev, ki jih lahko uporabimo.
- Napaka pri sprejemu enega simbola lahko povzroči napako več bitov. Ker je najbolj verjetna napaka med sosednjima nivojema, dodelimo bite posameznim nivojem tako, da se sosednji nivoji med seboj razlikujejo le za en bit (Grejeva koda).
- Dolga zaporedja enakih simbolov lahko še vedno povzročijo izpad sinhronizacije.

C. Oblikovanje signala

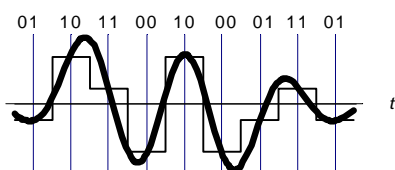
Namesto, da bi uporabili pravokotne impulze, lahko impulze oblikujemo in na ta način dosežemo zoženje potrebnega frekvenčnega pasu za prenos. Ker se s širjenjem impulza oži frekvenčni pas, je smiselna uporaba impulzov, ki se raztezajo čez več simbolov.

Pri tem je potrebno paziti, da se sosednji impulzi med seboj ne motijo, da ne povzročajo intersimbolne interference. Pomembno je, da imajo vrednost nič v trenutkih, ko pri sprejemu vzorčimo vrednosti sosednjih simbolov (Nyquistov kriterij). Primer takega impulza je prikazan na sliki 3.



Slika 3: Oblikovan impulz

Signal, ki bi ga dobili z zaporedjem simbolov s slike 2, in impulzom s slike 3, je prikazan na sliki 4.



Slika 4: Štiri nivojski prenos z oblikovanim signalom

Vidimo, da gre tako dobljen signal ob trenutkih vzorčenja skozi iste točke kot signal sestavljen iz pravokotnih impulzov. Iz same oblike signala pa je razvidno, da bo zavzemal ožji frekvenčni pas, kot signal oblikovan s pravokotnimi impulzi.

D. Dodajanje redundance v signal

Pri prenosu signala po prenosni poti lahko se pa popači, zato prihaja pri prenosu do napak. Njihovo število začne hitro naraščati že pri bitnih pretokih precej pod kapaciteto kanala.

Pri prenosu podatkov zadošča, če lahko napako odkrijemo. Ko pri sprejemu odkrijemo, da je prišlo do napake, zahtevamo ponovni prenos dela podatkov, v katerem je prišlo do napake.

Pri sprotnem prenosu govornega ali video signala zgolj odkrivanje napak ne zadošča več. Del signala, pri katerem je bila odkrita napaka, bi ob zahtevi po ponovnem prenosu, na cilj prispel prepozno, da bi ga še lahko vključili v sprotno tek govora ali slike, zato je tak signal izgubljen. Pri sprotnem prenosu moramo zato napake sproti odpravljati ali pa preprečiti njihov nastanek.

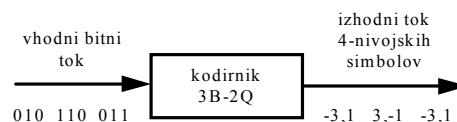
Tako odkrivanje, kot tudi odpravljanje napak lahko dosežemo z dodajanjem redundance v niz simbolov, ki ga prenašamo. Za odkrivanje napak je potrebno načeloma dodati manj redundance kot za njihovo odpravljanje.

Redundanco dodajamo s predhodnim kodiranjem bitnega niza. Osnovni niz razdelimo na bloke dolžine N , ki jih pripeljemo na vhod kodirnika.

Kodirnik te bloke pretvori v bloke dolžine M , simboli na izhodu pa so lahko binarni ali pa več nivojski. Če označimo dolžino bloka na izhodu z M in število nivojev simbolov na izhodu z L , govorimo o NB- ML kodiranju. Redundanco, ki smo jo dodali, izrazimo v procentih in jo podaja izraz:

$$R = \left(1 - \frac{N}{M \log_2 L}\right) 100\% \quad (2)$$

Shematsko je tak kodirnik prikazan na sliki 5.



Slika 5: Shematski prikaz kodirnika 3B-2Q

Postopek odkrivanja napak z dodajanjem redundance lahko najenostavneje razložimo s postopkom dodajanja paritetnega bita, ki se uporablja pri prenosu po serijskem vodilu RS 232. Pri tem postopku vsakemu bloku 7 bitov, s katerimi je zapisan en alfa-numerični znak, dodamo osmi paritetni bit. Vrednost tega bita je odvisna od predhodnih sedmih bitov. Pri sodi pariteti je paritetni bit enak ena če je število enic v predhodnih bitih sodo, drugače pa je paritetni bit enak nič. Redundanca ki smo jo dodali je v skladu z izrazom (2) enaka 12,5%. Dodajanje paritetnega bita nam omogoča odkrivanje ene napake v bloku sedmih bitov. Če pride v bloku sedmih do sodega števila napak, tega ne moremo več odkriti. Dodajanje paritetnega bita bi lahko označili kot kodiranje 7B-8B.

Zgoraj opisani postopek je le najpreprostejši način dodajanja redundance za odkrivanje napak. Veliko pogosteje se uporablja dodajanje krožne redundančne kode (CRC), ki jo izračunamo iz vseh bitov vhodnega bloka in jo dodamo bloku. Tak način omogoča odkrivanje več napak v bloku.

Dodajanje redundance, ki omogoča odkrivanje napak je nekoliko zahtevnejše. Poleg tega, da odkrijemo, da je do napake prišlo, moramo iz redundance razbrati tudi točne, kateri bit je bil napačno sprejet, da lahko napako odpravimo. Tudi to bomo najlažje razumeli, če si ogledamo preprost primer. Vhodni bitni tok razdelimo na okvirje, ki vsebujejo 7 blokov s sedmimi biti. In jih razvrstimo v tabelo dimenzije 7 x 7. Vsaki vrstici in stolpcu izračunamo paritetni bit in ga dodamo na konec vrstice oziroma stolpca, kot je to prikazano na sliki 6.

0	1	0	1	1	0	1	1
1	0	0	0	0	0	1	1
0	1	0	1	1	1	0	1
0	0	1	1	1	0	1	0
1	1	1	0	0	0	1	1
0	1	0	1	0	0	1	0
1	0	0	0	1	0	1	0
0	1	1	0	1	0	1	

Slika 6: Popravljanje napak

V zgornji tabeli vidimo, da sta paritetna bita v četrthi vrstici in v četrtem stolpcu napačna, iz tega lahko sklepamo, da je prišlo do napake v bitu, ki je v četrti vrstici in v četrtem stolpcu in ga lahko popravimo.

Na ta način lahko odpravimo celo večje število napak v bloku, če le ne nastopajo v isti vrstici ali stolpcu. Če nastopiti več napak v eni vrstici ali stolpcu, jih lahko odkrijemo, ne moremo pa jih več odpraviti. Dodana redundanca je pri opisanem postopku 23%.

Redundanco lahko poleg tega, da jo uporabljamo za odkrivanje napak, izkoriščamo tudi v druge namene, najpogosteje za izločanje enosmerne komponente ali za oblikovanje spektra signala.

E. Dekodiranje na osnovi največje verjetnosti

Pri vseh kodirnik postopkih, ki smo jih obravnavali doslej, smo se pri sprejemu omejili na detekcijo posameznega simbola. Verjetnost napake je bila pri tem enaka verjetnosti, da vrednost šuma v trenutku odločitve preseže polovico razdalje med posameznimi nivoji in zato detektiramo napačen simbol.

Če smo pri kodiranju dodali redundanco, zaporedja na izhodu kodirnika ne morejo biti poljubna. Določene zaporedja simbolov na izhodu kodirnika namreč niso mogoča – so prepovedana. Pri detekciji napak smo na osnovi prepovedane kombinacije simbolov v sprejetem bloku detektirali napako. Pri dekodiranju na osnovi največje verjetnosti pa se namesto tega raje odločimo za najbližje možno zaporedje, to je zaporedje, ki je bilo z največjo verjetnostjo oddano. Ker v tem primeru napake ne odkrivamo, temveč jo skušamo že vnaprej preprečiti, imenujemo takšen način tudi dekodiranje z vnaprejšnjim odpravljanjem napak.

Najbolj verjetno je, da je bilo oddano zaporedje, ki od ima od sprejetega zaporedja najmanjšo srednjo kvadratično vrednost odstopanja (najmanjšo evklidsko razdaljo). Kadar dekodiramo na osnovi največje verjetnosti moramo zato paziti, da imajo zaporedja, ki so dovoljena na izhodu kodirnika, čim večjo medsebojno evklidsko razdaljo. Na ta način zmanjšujemo možnost, da bi eno zaporedje zamenjali z drugim in s tem tudi verjetnost napake.

Opisani postopek je popolnoma enakovreden korelacijskemu sprejemniku, ki se na osnovi korelacije s signalom znane oblike odloča o tem, kateri signal je bil oddan. Korelacija se v tem primeru izvaja s tako imenovanim signalu prilagojenim sitom, ki ima sistemsko funkcijo enako časovno obrnjeni funkciji znanega signala.

Z opisanimi postopkoma se lahko zelo približamo kapaciteti kanala v primerih, ko ima kanal bolj ali manj ravno frekvenčno karakteristiko in šum bel spekter.

Verjetnost napake se pri dekodiranju na osnovi največje verjetnosti znižuje z večanjem števila simbolov, ki jih dekodirnik hkrati obravnava, zato bi želeli, da bi dekodirniki obravnavali čimveč simbolov hkrati. To pa je tudi največja težava dekodiranja na osnovi največje verjetnosti. Sprejeti niz moramo namreč primerjati z vsemi možnimi oddanimi nizi, število teh pa eksponentno narašča z njihovo dolžino in ravno tako hitro narašča tudi zahtevnost dekodirnika.

F. Mrežno kodiranje in Viterbijev dekodirnik

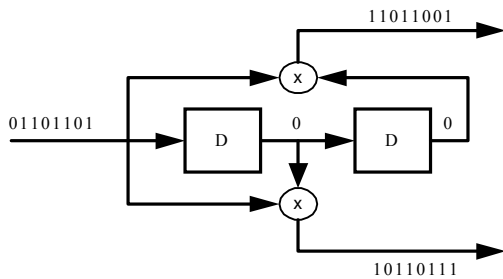
Za dodajanje redundance lahko namesto blokovnega kodiranja uporabimo mrežno kodiranje. Mrežni kodirnik je naprava z več stanji. Stanje v katerem je, je odvisno od preteklih vhodnih bitov, izhodni simboli pa so odvisni od stanja kodirnika in vhodnih bitov. Mrežni kodirnik na vsakih N vhodnih bitov tvori M izhodnih bitov, ki tvorijo enega ali več izhodnih simbolov. Za razliko od blokovnega kodiranja izhodnih M bitov ni odvisno samo od zadnjih N vhodnih bitov temveč tudi od stanja kodirnika. Mrežne kodirnike običajno označujemo z razmerjem med številom vhodnih in izhodnih bitov:

$$r = \frac{M}{N} \quad (3)$$

Običajno imamo opravka z 1/2, 2/3, 3/4, ... kodirniki. Od tega razmerja je neposredno odvisna tudi dodana redundanca:

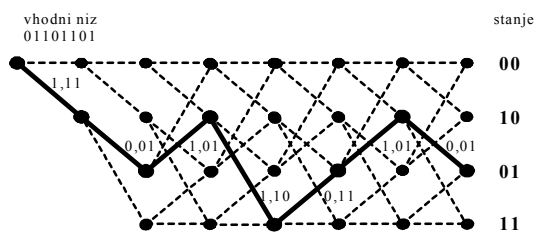
$$R = (1 - r)100\% \quad (4)$$

Primer mrežnih kodirnikov so tako imenovani konvolucijski kodirniki, pri katerih dobimo izhodne bite s pomočjo XOR operacij med trenutnim in K preteklimi vhodnimi biti. K pri tem predstavlja dolžino kodirnika in določa število možnih stanj kodirnika. Primer konvolucijskega kodirnika ($r = 1/2$, $K=2$) je prikazan na sliki 7.



Slika 7: Konvolucijski kodirnik $R = 1/2, K=2$

Mrežne kodirnike lahko predstavimo v mrežnem diagramu. Vozlišča v diagram predstavljajo stanja kodirnika, ki jih določajo vsebine D celic. Kodirnik na sliki 5 ima tako 4 različna stanja. Prehajanje med stanji določajo vhodni biti, izhodni biti pa so odvisni od teh prehodov. Vhodne in izhodne bite označujemo nad povezavami, ki označujejo prehajanje stanj. Za primer kodirnika na sliki 7 je mrežni diagram podan na sliki 8.



Slika 8: Mrežni diagram konvolucijski kodirnika $R = 1/2, K=2$

Vsak vhodni niz predstavlja določeno pot skozi mrežni diagram. Vse možne poti so v mrežnem diagramu na sliki 8 označene s črtkanimi črtami, z izbranim vhodnim nizom določena pot pa s polno črto.

Pri sprejemu uporabimo Viterbijev dekodirnik. Viterbijev dekodirnik se ne odloča o vsakem simbolu posebej, temveč se odloči za tisto od vseh možnih poti v mrežnem diagramu, ki ima od sprejetega signala najmanjšo evklidsko razdaljo. Prednost Viterbijevega dekodirnika je v Viterbijevem algoritmu, ki omogoča, da kompleksnost dekodirnika ne narašča več eksponentialno z njegovo dolžino, temveč samo še linearno, drugače pa ima vse lastnosti dekodiranja na osnovi največje verjetnosti.

V. PRENOS V VIŠJIH FREKVENČNIH LEGAH

Kadar več uporabnikov uporablja skupni prenosni medij ali kadar osnovni frekvenčni pas ni na voljo, je potrebno signal prestaviti v višjo frekvenčno lego. V višjo frekvenčno lego prestavimo signal z modulacijskimi postopki. Glede kapacitete kanala je prenos v višji frekvenčni pri dani pasovni širini prenos

enakovreden prenosu v osnovnem pasu. Z modulacijskimi postopki lahko ohranimo pasovno širino, ki jo je imel signal v osnovnem pasu ali pa jo povečamo. Pasovne širine z nobenim modulacijskim postopkom ne moremo zmanjšati. Za prestavitev v višjo frekvenčno lego lahko uporabljamo naslednje modulacijske postopke:

- amplitudna modulacija,
- frekvenčna modulacija,
- fazna modulacija in
- amplitudno fazna modulacija.

A. Amplitudna modulacija

Pri amplitudni modulaciji s signalom v osnovnem pasu moduliramo amplitudo nosilca nosilne frekvence:

$$y(t) = A x(t) \cos(2\pi f_0 t) \quad (5)$$

kjer je $x(t)$ signal v osnovnem pasu, A ojačevalni faktor modulatorja, f_0 nosilna frekvenca in $y(t)$ moduliran signal na izhodu modulatorja. Kadar je signal v osnovnem nastal s pretvorbo simbolov v impulze, govorimo o impulzno amplitudni modulaciji (PAM – Pulse Amplitude Modulation). Za amplitudno modulacijo velja:

- Spekter signala v osnovnem pasu se preslika levo in desno od nosilne frekvence. Pri tem se podvoji spektralna širina signala.
- Če je v modulacijskem signalu prisotna enosmerna komponenta, je v moduliranem signalu prisoten nosilec. Prisotnost nosilca v splošnem ni zaželena, ker je zaradi tega večja moč signala, ne da bi bila večja prenosna hitrost.
- Informacija je kodirana v amplitudi signala, zato moramo pri prenosu ohranjati amplitudna razmerja in potrebujemo linearne ojačevalnike.

B. Frekvenčna modulacija

Pri frekvenčni modulaciji s signalom v osnovnem pasu spreminjamo trenutno frekvenco signala okrog nosilne frekvence:

$$f(t) = f_0 + \Delta f x(t) \quad (5)$$

$$y(t) = A \cos\left(\int f(t) dt\right)$$

kjer je $f(t)$ označena trenutna frekvenca signala in Δf frekvenčni odmik. Pri modulaciji z nizom impulzov imenujemo frekvenčno modulacijo tudi frekvenčno kodiranje (FSK – Frequency Shift Keying). V tem primeru lahko določimo modulacijski indeks z razmerjem:

$$m = \frac{\Delta f}{f_s} \quad (6)$$

kjer je s f_s označen simbolni pretok. Za frekvenčno modulacijo velja:

- Spektralna širina je odvisna od modulatorskega indeksa. Pri majhnih modulatorskih indeksih je enaka kot pri amplitudni modulaciji, pri velikih pa je približno enaka dvojnemu frekvenčnemu odmiku ($2\Delta f$), kar je precej večja pasovna širina kot pri amplitudni modulaciji.
- Informacija se prenaša s trenutno frekvenco, amplituda pa je konstantna, zato modulacija ni občutljiva na nelinearna popačenja.
- Pri velikem modulatorskem indeksu je na račun večje pasovne širine, ki jo porabimo za prenos, signal veliko manj občutljiv na šum kot pri amplitudni modulaciji.
- Kadar pri spremembi frekvence ohranjamo trenutno fazo, govorimo o modulaciji z zvezno fazo.
- Dvonivojska frekvenčna modulacija z modulatorskim indeksom $m = 0,5$ je znana pod imenom MSK (Minimum Shift Frequency Keying). Izpeljava iz MSK je GMSK, ki se uporablja v sistemu GSM.

C. Fazna modulacija

Pri fazni modulaciji z modulatorskim signalom spreminjamo trenutno fazo signala:

$$\begin{aligned} \Phi(t) &= f_0 t + \Delta\Phi x(t) \\ y(t) &= A \cos(\Phi(t)) \end{aligned} \quad (7)$$

Kadar fazo moduliramo z impulzi govorimo faznem kodiranju (PSK – Phase Shift Keying). Lastnosti fazne modulacije so zelo podobne lastnostim frekvenčne modulacije, saj sta postopka zelo podobna. Razlika je le v tem, da v nastopa argumentu kosinusne funkcije pri frekvenčni modulaciji integral modulatorskega signala, pri fazni modulaciji pa nastopa modulatorski signal brez integrala. Za fazno modulacijo velja:

- Spektralna učinkovitost in odpornost proti šumu sta enaki kot pri frekvenčni modulaciji.
- Kadar moduliramo fazo s signalom, ki nima nezveznosti (oblikovani impulzi), govorimo o modulaciji z zvezno fazo (CPM – Continuous Phase Modulation).
- Ker je informacija kodirana v fazi signala, modulacija ni občutljiva na nelinearna popačenja.
- Pri postopku demodulacije moramo poznati absolutno fazo signala, da se lahko odločimo o oddanem simbolu. Temu se izognemo tako, da informacijo kodiramo v spremembi faze namesto v

absolutni fazi. Govorimo o diferenčnem faznem kodiranju (DPSK). DPSK je, če spreminjamo fazo zvezno, popolnoma enak FSK.

D. Amplitudno fazna modulacija

Pri amplitudno fazni modulaciji z modulatorskim signalom spreminjamo amplitudo in fazo signala. To običajno naredimo s kvadraturnim modulatorjem (QAM – Quadrature Amplitude Modulator), tako da iz osnovnega niza simbolov tvorimo dva niza s polovično simbolno hitrostjo in dobimo dva signala v osnovnem pasu: signal v fazi $x_i(t)$ in signal v kvadraturi $x_q(t)$. Zaradi polovične simbolne hitrosti imata ta dva signala tudi polovično pasovno širino. Z njima amplitudno moduliramo ortogonalna nosilca kosinus in sinus:

$$\begin{aligned} y(t) &= A (x_i(t) \cos(2\pi f_0 t) + x_q(t) \sin(2\pi f_0 t)) = \\ &= A(t) \cos(\Phi(t)) \end{aligned} \quad (8)$$

kjer je z $A(t)$ trenutna amplituda in s $\Phi(t)$ trenutna faza signala. Amplitudno fazno modulacijo običajno označujemo kratico mQAM, kjer m pove koliko nivojev prenašamo. Za amplitudno fazno modulacijo velja:

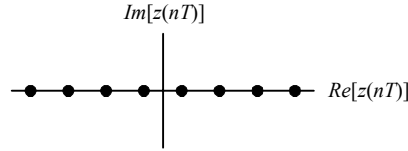
- Spektralno je postopek enako učinkovit kot prenos v osnovnem pasu. To pomeni, da potrebuje glede na amplitudno modulacijo polovično pasovno širino. To je spektralno najbolj učinkovit postopek. Njegova spektralna učinkovitost je enaka kot pri enobočni modulaciji (SSB – Single Side Band), ki se uporablja pri analognem prenosu.
- Ker je informacija kodirana tudi v amplitudi moduliranega signala, je signal občutljiv na nelinearna popačenja)
- Pri demodulaciji moramo poznati frekvenco in fazo nosilca. Potrebi po poznavanju absolutne faze se lahko izognemo, če uporabimo diferenčno kodiranje.

E. Konstalcijski diagram

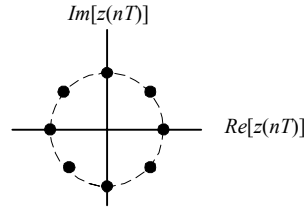
PAM, PSK in QAM postopke lahko predstavimo tudi v konstalcijskem diagramu. V kompleksni ravnini označimo s točkami vse možne vrednosti ekvivalentnega signala v osnovnem pasu:

$$z(t) = y(t)e^{2\pi f_0 t} \quad (9)$$

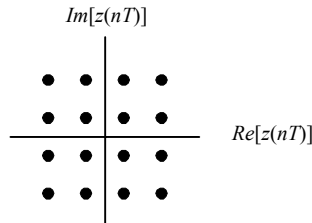
ob odločitvenih trenutkih nT . Oddaljenost točke od izhodišča predstavlja amplitudo, kot ki ga tvori daljica med točko in izhodiščem z realno osjo pa fazo signala. Konstalcijski diagrami za 8PAM, 8PSK in 16 QAM so prikazani na slikah 9, 10 in 11.



Slika 9: Konstalacijski diagram 8PAM



Slika 10: Konstalacijski diagram 8PSK



Slika 11: Konstalacijski diagram 16QAM

Iz primerov na slikah 9, 10 in 11 lahko vidimo, da sta PAM in PSK le posebna primera QAM. PAM dobimo kadar je kvadratni modulacijski signal enak 0, PSK pa kadar je izpolnjen pogoj:

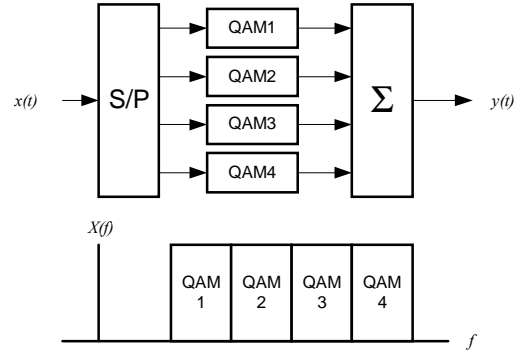
$$x_i^2(t) + x_q^2(t) = A^2 \quad (10)$$

PSK modulacijo lahko zato izvedemo s pomočjo QAM modulatorja. V tem primeru govorimo o QPSK modulacijskem postopku.

F. Prenos z več nosilci

Pri prenosu v širokem frekvenčnem pasu se pogosto soočamo z dejstvom, da karakteristika kanala v tem frekvenčnem pasu ni ravna, zato jo je potrebno izravnati s pomočjo izravnalnika. Z izravnavo frekvenčne karakteristike poslabšamo razmerje med signalom in šumom in na ta način znižujemo kapaciteto kanala.

V zelo ozkem frekvenčnem pasu spremembe frekvenčne karakteristike niso velike, zato pri prenosu ne potrebujemo izravnave, ki bi poslabšala razmerje med signalom in šumom. Ker pa v ozkem frekvenčnem pasu ne moremo doseči velikega bitnega pretoka, si pomagamo na ta način, da osnovni bitni tok razdelimo na več tokov in vsakega moduliramo s svojim nosilcem, kot je to prikazano na sliki 12.



Slika 12: Prenos z več nosilci

Če uporabimo pri tem postopku dovolj ozke frekvenčne pasove, da ni potrebna izravnava, in v vsakem pasu uporabimo dekodiranje na osnovi največje verjetnosti, se lahko približamo kapaciteti kanala, kot jo podaja enačba (1).

Na prvi pogled izgleda ta postopek izjemoma zahteven, saj potrebujemo veliko število QAM modulatorjev in tudi veliko število dekodirnikov. V praksi lahko stvar močno poenostavimo, tako da s pomočjo hitrega Fourierovega transformata (FFT) naredimo vse modulacije naenkrat, pri dekodiranju pa tokove najprej združimo in šele nato izvedemo dekodiranje na osnovi največje verjetnosti. Na tak način so izvedljivi tudi modulatorji 256, 1024 ali več nosilci.

VI. TEHNIKE SODOSTOPA DO SKUPNEGA MEDIJA

Kot smo že omenjali, si mora pogosto več uporabnikov deliti skupno prenosno pot. Prenosno pot si lahko delijo na osnovi:

- delitve zmogljivosti med posamezne uporabnike,
- začasnega zasedanja celotne zmogljivosti ali
- kodno deljen sodostop.

A. Delitev zmogljivosti

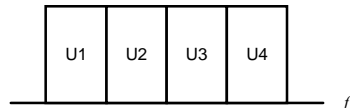
Več uporabnikov lahko uporablja skupno prenosno pot tako, da si med seboj razdelijo njeno zmogljivost. V času ko je del zmogljivosti (kanal) dodeljen nekemu uporabniku, ta ni na voljo drugim uporabnikom, zato ima lahko uporabnik zagotovljen konstanten bitni pretok, oziroma je mogoče uporabniku zagotoviti določeno kvaliteto storitve (QoS – Quality of Service). Tak način je zato najbolj primeren tam, kjer je kvaliteta storitve zahtevana, to je pri sprotnem prenosu govornih, avdio in video signalov.

Poznamo predvsem dva načina delitve zmogljivosti:

- Frekvenčno deljen sodostop (FDMA – Frequency Division Multiple Access) in
- Časovno deljen sodostop (TDMA – Time Division Multiple Access).

Frekvenčno deljen sodostop

Pri frekvenčno deljenem sodostopu vsakemu uporabniku dodelimo svoj frekvenčni kanal, kot je to prikazano na sliki 13.

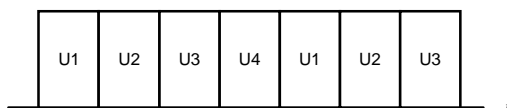


Slika 13: Frekvenčno deljen sodostop

Vsak uporabnik zaseda svoj kanal ves čas komunikacije, omejen pa je njegov frekvenčni pas. Število hkratnih uporabnikov je omejeno s pasovno širino, ki je na voljo vsem uporabnikom in pasovno širino, ki je dodeljena posameznemu uporabniku.

Časovno deljen sodostop

Pri časovno deljenem sodostopu je vsakemu uporabniku dodeljena svoja časovna rezina (časovni okvir) kot je to prikazano na sliki 14.



Slika 13: Časovno deljen sodostop

Pri časovnem sodostopu vsak uporabnik zaseda celoten frekvenčni pas, ki je dodeljen skupini uporabnikov, omejen pa je na svojo časovno rezino. Število hkratnih uporabnikov v določenem frekvenčnem pasu je lahko enako kot pri frekvenčno deljenem sodostopu, vendar je tu potrebna časovna sinhronizacija celotnega sistema, tako da lahko vsak uporabnik najde svojo časovno rezino.

B. Zaseganje zmogljivosti

Zaseganje zmogljivosti prenosne poti je na nek način podobno časovni delitvi zmogljivosti, saj tudi tu uporabnik za določen čas zasede celoten frekvenčni pas, ki je na voljo. Razlika med časovno deljenim sodostopom in zaseganjem zmogljivosti je v tem, da pri slednjem čas ni razdeljen na rezine, ki so dodeljene posameznim uporabnikom, temveč lahko vsak uporabnik zaseže celotno zmogljivost, ko je prenosna pot prosta. Kadar prenosna pot ni na voljo, ker jo že zaseda nek uporabnik, morajo ostali uporabniki počakati.

Pri zaseganju zmogljivosti morajo veljati določena pravila lepega obnašanja. Uporabnik, ki je prenosno pot zasegel, jo mora po določenem času tudi sprostiti

in jo dati na voljo drugim uporabniku, pa četudi bi jo tudi sam še potreboval.

Ker dobi uporabnik na voljo prenosno pot le, ko je ta prosta, tak način ne zagotavlja kvalitete storitve, zato ni najbolj primeren za sprotni prenos, primeren pa je za prenos podatkov.

Ločimo predvsem dva načina zaseganja zmogljivosti prenosne poti:

- zaseganje na osnovi poslušanja in detekcije trkov (collision detection) in
- zaseganje na osnovi žetona (token).

Zaseganje na osnovi detekcije trkov

Pri zaseganju na osnovi poslušanja uporabnik posluša, kaj se dogaja na prenosni poti. Ko ugotovi, da je prenosna pot prosta, jo zaseže in prične z oddajo svojega signala.

Ker so uporabniki med seboj oddaljeni in ker ima elektromagnetno valovanje končno hitrost, traja nekaj časa, preden pride signal od enega uporabnika do drugega. Čeprav je nek uporabnik že oddal svoj signal, se lahko zgodi, da začne drug uporabnik oddajati, preden je ta signal prispel do njega. Pride do trka.

Uporabniki morajo zato poslušati tudi med tem ko oddajajo in, ko zaznajo trk, morajo z oddajo prenehati. Da ob poskusu ponovne oddaje ne bi prihajalo do ponovnih trkov, po prenehanju oddaje vsak uporabnik počaka naključen čas, tako da je verjetnost, da bi po zaznavi trka ista dva uporabnika ponovno trčila, minimalna.

Verjetnost, da pride do trkov se veča z medsebojno razdaljo med uporabniki, zato je tak način primeren le za manjše razdalje.

Pri veliki zasedenosti prenosne zmogljivosti (nad 50%) postanejo trki zelo pogosti, zato pri ponovnem poskušanju pride do ponovnih trkov, kar pripelje do zasičenosti prenosne poti, ko nihče ne more več komunicirati, temveč vsi smo še poskušajo.

Tak postopek se uporablja v lokalnem omrežju Ethernet.

Zaseganje na osnovi žetona

Pri zaseganju na osnovi žetona uporabniki ne poskušajo oddajati takoj, ko je prenosna pot prosta temveč šele, ko jim je to dovoljeno. Pravico do oddajanja daje uporabniku žeton. Uporabnik, ki ima žeton lahko odda svoj paket podatkov, nato pa mora žeton predati naslednjemu uporabniku v vrsti. Z naslednjo oddajo mora počakati, dokler ponovno ne dobi žetona.

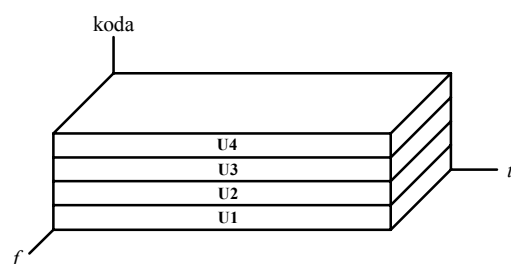
Pri veliki zasedenosti prenosne zmogljivosti tu ne pride do zasičenosti, le čakalni čas za žeton se podaljša. Ravno tako tu ni problema z medsebojno oddaljenostjo uporabnikov, saj do trkov v nobenem primeru ne prihaja. Z ustrezno dodelitvijo prioriteta pri dodeljevanju žetona in z omejevanjem števila hkratnih uporabnikov, je pri tem načinu možno zagotavljati kvaliteto storitev.

Edina pomanjkljivost načina zaseganja zmogljivosti na osnovi žetona nastopi pri relativno majhni zasedenosti omrežja, ko morajo uporabnik čakati, da pride žeton do njega, klub temu, da nihče drug ne želi oddajati.

Način z zaseganjem na osnovi žetona se uporablja v lokalnem omrežju Token Ring.

C. Kodno deljen sodostop

Kodno deljen sodostop (CDMA – Code Division Multiple Access) predstavlja nekakšno mešanico med obema prejšnjima načinoma ali tudi mešanico med časovno in frekvenčno deljenim sodostopm. Pri kodnem sodostopu uporabljajo vsi uporabniki ves čas komunikacije celoten frekvenčni pas. Med seboj so razdeljeni s kodami, kot je to prikazano na sliki 14.



Slika 14: Kodno deljen sodostop

Namesto, da bi z modulatorskim signalom modulirali harmonični nosilec, kot pri frekvenčnem sodostopu, tu moduliramo psevdonaključen nosilec, ki ga imenujemo koda. Pri sprejemu lahko posamezne uporabnike med seboj ločimo le, če so kode med seboj ortogonalne. Kadar so kode med seboj ortogonalne je tak sodostopm enakovreden frekvenčnemu ali časovnemu sodostopu, tako glede števila hkratnih uporabnikov kot glede kvalitete storitve.

Običajno pa namesto ortogonalnih kod uporabljamo semiortogonalne kode. Teh kod je lahko bistveno večje število kot ortogonalnih kod, zato pa se lahko posamezni uporabniki med seboj rahlo motijo. Zaradi medsebojnega motenja je število uporabnikov, ki lahko hkrati uporabljajo prenosno pot manjše kot pri časovnem ali frekvenčnem multipleksu. Kljub temu pa ima uporaba semiortogonalnih kod bistvene prednosti pred uporabo ortogonalnih kod, predvsem v celičnih omrežjih, ker omogočajo:

- ponovno uporabo iste frekvence v sosednjih celicah in s tem boljši izkoristek frekvenčnega pasu in lažje frekvenčno planiranje
- dinamično zaseganje kanala, ki je na nek način podobno prej opisanim tehnikam z zaseganjem kanala, in povečuje celotno kapaciteto sistema,
- sektorizacijo celic, kar tudi povečuje kapaciteto celotnega sistema,
- mehke prehode med celicami in
- odpornost proti frekvenčno selektivnemu presihu.

Zadnja lastnost ni odvisna od ortogonalnosti kod temveč je posledica uporabe celotnega frekvenčnega pasu.

Zaradi vseh omenjenih prednosti, je uporaba kodno deljenega sodostopa predvidena za uporabo v celičnih omrežjih tretje generacije UMTS. Več o tem pa bo govora v naslednjih prispevkih.

VII. ZAKLJUČEK

V prispevku je bil narejen kratek pregled različnih postopkov kodiranja in modulatorskih postopkov, ki so potrebni za izkoriščanje kapacitet prenosne poti, ki je na voljo uporabniku za dostop do telekomunikacijskih storitev. Ker si mora pogosto več uporabnikov deliti skupno prenosno pot, so opisane tudi različne tehnike sodostopa do prenosne poti.

Ker je narava prispevka pregledna in informativna je izpuščena tako matematična analiza kot tudi tehnične podrobnosti.

Če bo dobil bralec osnovno informacijo o tem, kakšne možnosti obstajajo, kakšne so njihove prednosti in slabosti in kaj je največ mogoče pričakovati od posameznih postopkov, potem je bil namen tega prispevka dosežen.

BIOGRAFIJA

Sašo Tomažič je izredni profesor na Fakulteti za elektrotehniko Univerze v Ljubljani, kjer predava predmete iz področja digitalnih komunikacij in digitalne obdelave signalov. Od leta 1999 je predstojnik katedre za telekomunikacije in od leta 2000 nacionalni koordinator za področje telekomunikacij na MŠZŠ.