

UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET



Marija D. Antić

**OPTIMIZACIJA NEBLOKIRAJUĆIH PAKETSKIH
MREŽA UPOTREBOM PRAKTIČNOG PROTOKOLA ZA
RUTIRANJE SA BALANSIRANJEM SAOBRAĆAJA**

– Doktorska disertacija –

Beograd, 2014.

UNIVERSITY OF BELGRADE
SCHOOL OF ELECTRICAL ENGINEERING



Marija D. Antić

**OPTIMIZATION OF NON-BLOCKING PACKET
NETWORKS USING THE PRACTICAL ROUTING
PROTOCOL WITH LOAD BALANCING**

– PhD Thesis –

Belgrade, 2014.

MENTOR:

dr Aleksandra Smiljanić, redovni profesor
Univerzitet u Beogradu, Elektrotehnički fakultet

ČLANOVI KOMISIJE:

dr Irini Reljin, redovni profesor
Univerzitet u Beogradu, Elektrotehnički fakultet

dr Dragana Perić, naučni saradnik
Vlatacom, Beograd

dr Zoran Jovanović, redovni profesor
Univerzitet u Beogradu, Elektrotehnički fakultet

Datum odbrane: _____

REZIME

Kvalitetan odabir strategije rutiranja u paketskoj mreži u velikoj meri utiče na pouzdanost mreže i kvalitet prenosa podataka u njoj. Odabrana strategija rutiranja utiče i na stepen iskorišćenja različitih mrežnih resursa. Loše odabrano rutiranje može prouzrokovati stvaranje uskih grla u mreži i dovesti do blokade određenih linkova. Takođe, od primenjene strategije rutiranja zavisi kašnjenje pri prenosu paketa, kao i cena mreže. U cilju odabiranja ispravne strategije rutiranja za neku mrežnu topologiju i date korisničke zahteve, vrši se *optimizacija rutiranja* u mreži. Proces optimizacije rutiranja za cilj ima određivanje putanja paketa tako da se postigne što bolja vrednost neke od mrežnih performansi. U zavisnosti od problema koji se rešava, optimizacija rutiranja može imati za cilj maksimizaciju propusne moći mreže, minimizaciju cene mreže, minimizaciju kašnjenja u prenosu i slično.

U ovome radu bavićemo se problemima optimizacije rutiranja u cilju *maksimizacije propusne moći* i *minimizacije cene* mreže. Pri tome, želimo da komunikaciona mreža bude neblokirajuća. To znači da mreža mora garantovati servis za sve korisnike koji generišu ili primaju saobraćaj u okviru dozvoljenih granica. Nezavisno od trenutnog stanja saobraćaja u mreži, korisnički zahtevi za slanje ili prijem paketa moraju biti usluženi, ukoliko ne izlaze iz okvira dozvoljenih vrednosti. Okosnicu ovog rada predstavlja predlog nove šema rutiranja zasnovane na balansiranju saobraćaja i rutiranju po najkraćoj putanji, koju nazivamo *rutiranje sa balansiranjem saobraćaja* (RBS). Predložena šema rutiranja može se optimizovati tako da se ostvari maksimalna propusna moć mreže ili minimalna cena mreže, a za optimizaciju se koristi linearno programiranje.

Glavnu prednost predložene strategije rutiranja predstavlja realna mogućnost njene implementacije. Naime, rutiranje po najkraćoj putanji je u širokoj primeni, a predložena šema rutiranja je modifikacija koja se relativno lako realizuje. Dimenzija linearnog programa koji koristimo je prihvatljiva za praktičnu primenu, kao i vreme

neophodno za njegovo izvršavanje. Primenom predložene šeme može se ostvariti servis za veći broj korisnika nego u slučaju kada se primenjuje regularno rutiranje po najkraćoj putanji, a da pri tom nema potrebe za promenama postojeće mrežne topologije. Takođe, u već postojećoj mreži, predložena šema omogućava servisiranje određenog broja korisnika po nižoj ceni u odnosu na cenu servisiranja istog skupa korisnika u mreži koja koristi rutiranje po najkraćoj putanji.

U ovom radu biće prvo definisana strategija rutiranja sa balansiranjem saobraćaja. Zatim će biti data analiza performansi predložene strategije rutiranja i izvršeno njeno poređenje sa običnim rutiranjem po najkraćoj putanji. Konkretno, biće analizirana dva slučaja. U prvom slučaju vršićemo optimizaciju rutiranja sa balansiranjem saobraćaja, tako da se ostvari maksimalna propusna moć mreže. Propusna moć u kojoj se primenjuje rutiranje sa balansiranjem saobraćaja zatim će biti poređena sa propusnom moći mreže u kojoj je primenjeno rutiranje po najkraćoj putanji. Biće pokazano da se primenom balansiranja saobraćaja može ostvariti značajno povećanje propusne moći mreže, uz korišćenje već postojećih resursa. U drugom slučaju, biće vršena optimizacija rutiranja sa balansiranjem saobraćaja, tako da se ostvari minimalna cena mreže koja može da podrži neki zadati saobraćaj. Cena mreže koja koristi rutiranje sa balansiranjem saobraćaja zatim će biti poređena sa cenom mreže koja koristi obično rutiranje po najkraćoj putanji. Biće pokazano je da je cena mreže koja može da podrži neki zadati saobraćaj manja ukoliko se primenjuje predloženo rutiranje sa balansiranjem saobraćaja.

Analiza će biti vršena kako za regularne, tako i za realne mrežne topologije. Takođe, optimizacija rutiranja biće vršena i za slučajeve u kojima se zahteva da mreža bude pouzdana. Pouzdana mreža mora biti u stanju da i u slučaju otkaza nekog resursa (linka ili rutera) omogući nesmetanu komunikaciju korisnika. U slučaju otkaza nekog resursa, veće opterećenje se postavlja na one resurse koji su ostali u funkciji. Zato se u ovakvoj mreži garantovani saobraćaj svakog korisnika mora dodatno ograničiti (čime se smanjuje propusna moć mreže), odnosno kapacitet svakog linka dodatno povećati (čime se povećava i cena mreže). Biće pokazano da su performanse predloženog rutiranja sa balansiranjem saobraćaja u svakom od analiziranih slučajeva iste ili bolje od performansi rutiranja po najkraćoj putanji, koje je danas u praksi široko rasprostranjeno.

KLJUČNE REČI: neblokirajuće paketske mreže, rutiranje, rutiranje po najkraćoj putanji, balansiranje saobraćaja, optimizacija rutiranja, pouzdanost

NAUČNA OBLAST: tehničke nauke, elektrotehnika

UŽA NAUČNA OBLAST: telekomunikacije

UDK BROJ: 621.3

ABSTRACT

The choice of the routing strategy in a packet network has the great impact on the reliability and the overall quality of data transmission. The chosen routing strategy also has an impact on the usage of different network resources, and bad routing decisions can cause bottlenecks in the network, and limit the throughput. Additionally, network latency as well as network cost also depend on the chosen routing scheme. In order to determine the best routing strategy for the given network and expected traffic, routing optimization is performed. The goal of routing optimization is achieving the best possible value for the chosen network performance metric. Depending on the problem that is being solved, optimization can be performed to maximize the network throughput, minimize the network cost, minimize the latency, etc.

In this paper, we will address the problems of network throughput maximization, and network cost minimization. We will focus on the non-blocking networks, which guarantee the service to all users that are generating and receiving the traffic within the defined boundaries, regardless of the current traffic distribution in the network. In this paper, we will propose the novel routing scheme, based on load balancing and shortest path routing. We will refer to the proposed scheme as *load balanced routing* (LBR). The proposed routing protocol can be optimized using linear programming, in order to minimize the cost, or maximize the network throughput.

The proposed routing scheme can easily be deployed in practice. Namely, shortest path routing is widely used, and the proposed routing is a modification that can relatively easy be implemented. The dimension of the linear program is acceptable, as well as the time needed to perform the optimization. With the proposed routing, it is possible to guarantee the service to the greater number of users, then in case of the regular shortest path routing. Also, the cost of servicing a given number of users is smaller if the proposed routing is applied instead of the shortest path routing.

First, we will introduce the proposed load balanced routing, and the linear program for its optimization. Then, we will analyze the performance of the proposed routing scheme, and compare it with the performance of the shortest path routing. In particular,

we will focus on two cases. In the first one, we will optimize the load balanced routing to achieve the maximum network throughput. It will be shown that the proposed load balanced routing can significantly increase the guaranteed network throughput, when compared to the regular shortest path routing. In the second case, we will optimize the load balanced routing in order to achieve minimum cost of the network that can support given user demands. It will be shown that the cost of the network with load balanced routing is lower than the cost of the network using shortest path routing.

Analysis will be performed for both regular and realistic network topologies. Also, the analysis will be performed for the networks in which service reliability is demanded, in the case of single resource (link or router) failure. It will be shown that the performance of the proposed load balanced routing is better than the performance of shortest path routing in all of these cases.

KEYWORDS:	non-blocking packet networks, routing, shortest path routing, load balancing, routing optimization, reliability
SCIENTIFIC AREA:	technical science, electrical engineering
SPECIFIC SCIENTIFIC AREA:	telecommunications
UDK NUMBER:	621.3

SADRŽAJ

SADRŽAJ	3
1. UVOD.....	5
1.1. RAZLOZI ZA OPTIMIZACIJU RUTIRANJA	5
1.1.1. Problem minimizacije cene mreže	6
1.1.2. Problem maksimizacije propusne moći mreže	7
1.2. RUTIRANJE SA BALANSIRANJEM SAOBRAĆAJA	7
1.3. SADRŽAJ I ORGANIZACIJA TEZE	9
2. RUTIRANJE	12
2.1. OSNOVNI KONCEPTI: PROCES RUTIRANJA, TABELE RUTIRANJA, ALGORITMI ZA RUTIRANJE	12
2.1.1. Određivanje putanje	13
2.1.2. Prosleđivanje paketa	14
2.2. PODELA ALGORITAMA ZA RUTIRANJE	14
2.3. MATEMATIČKO MODELOVANJE RUTIRANJA	16
2.3.1. Načini opisivanja stanja saobraćaja u mreži	16
2.3.2. Matematička definicija rutiranja	18
2.4. RUTIRANJE PO NAJKRAĆOJ PUTANJI	19
2.4.1. Definicija najkraće putanje	19
2.4.2. Dijkstra algoritam	20
3. KRITERIJUMI OPTIMIZACIJE RUTIRANJA	23
3.1. MAKSIMIZACIJA PROPUSNE MOĆI MREŽE	23
3.2. MINIMIZACIJA CENE MREŽE	26
4. LINEARNA OPTIMIZACIJA	28
4.1. MATEMATIČKA OPTIMIZACIJA U OPŠTEM SLUČAJU	28
4.2. LINEARNA OPTIMIZACIJA	29
4.3. SIMPLEKS ALGORITAM	31
4.3.1. Faza I	32
4.3.2. Faza II	33
4.3.3. Faza III	37
4.4. DUALNOST LINEARNIH PROGRAMA	37
4.5. PRIMENJENI OPTIMIZACIONI ALAT <i>LP SOLVE</i>	38
5. NEBLOKIRAJUĆE MREŽE	39
5.1. NEBLOKIRAJUĆE MREŽE SA KOMUTACIJOM VEZE – KLOSOVE MREŽE	40
5.1.1. Trostepena Klosova mreža	42
5.1.2. Petostepena Klosova mreža	43
5.1.3. Opšti slučaj Klosove mreže sa neparnim brojem stepeni	45
5.2. NEBLOKIRAJUĆE PAKETSKE MREŽE	45
5.2.1. Uslov neblokiranja u slučaju mreža sa vremenskim multipleksiranjem	46
5.2.2. Uslov neblokiranja u asinhronim paketskim mrežama	47
6. BALANSIRANJE SAOBRAĆAJA U REGULARNIM TOPOLOGIJAMA	51
6.1. BALANSIRANJE SAOBRAĆAJA U VIŠESTEPENIM SVIČEVIMA SA KLOSOVIM MREŽAMA	52
6.2. OPTIMALNO BALANSIRANJE SAOBRAĆAJA U SVIČEVIMA	54
6.2.1. Garantovana propusnost uniformne meš mreže	57
6.2.2. Garantovana propusnost polarisane meš mreže	59
7. RBS ALGORITAM – RUTIRANJE SA BALANSIRANJEM SAOBRAĆAJA	60

7.1.	OPIS ŠEME RUTIRANJA SA BALANSIRANJEM SAOBRAĆAJA	61
7.2.	OPTEREĆENJE LINKA KAO FUNKCIJA SAOBRAĆAJA ČVOROVA	63
7.3.	OPTIMIZACIJA RUTIRANJA SA BALANSIRANJEM SAOBRAĆAJA	65
8.	MAKSIMIZACIJA PROPUSNE MOĆI MREŽE.....	69
8.1.	PROPUSNA MOĆ MREŽE SA RUTIRANJEM SA BALANSIRANJEM SAOBRAĆAJA (RBS).....	69
8.2.	PROPUSNA MOĆ MREŽE SA RUTIRANJEM PO NAJKRAĆOJ PUTANJI (RNP).....	71
8.2.1.	<i>Pridruživanje grafa linku u mreži</i>	<i>72</i>
8.2.2.	<i>Algoritam za određivanje maksimalnog toka (maximum flow)</i>	<i>73</i>
8.2.3.	<i>Određivanje propusne moći mreže sa rutiranjem po najkraćoj putanji</i>	<i>76</i>
8.3.	POREĐENJE PROPUSNE MOĆI MREŽE ZA RBS I RNP	77
8.3.1.	<i>Regularne mrežne topologije</i>	<i>78</i>
8.3.2.	<i>Realne mrežne topologije</i>	<i>84</i>
9.	MINIMIZACIJA CENE MREŽE	87
9.1.	CENA MREŽE SA RUTIRANJEM SA BALANSIRANJEM SAOBRAĆAJA (RBS).....	87
9.2.	CENA MREŽE SA RUTIRANJEM PO NAJKRAĆOJ PUTANJI (RNP).....	89
9.3.	POREĐENJE CENE MREŽE ZA RBS I RNP	90
10.	OPTIMIZACIJA U POUZDANIM MREŽAMA	92
10.1.	POUZDAN KOMUNIKACIONI SISTEM.....	92
10.2.	MAKSIMIZACIJA PROPUSNE MOĆI MREŽE	94
10.2.1.	<i>Propusna moć pouzdane mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)</i>	<i>94</i>
10.2.2.	<i>Propusna moć mreže sa rutiranjem po najkraćoj putanji (RNP).....</i>	<i>95</i>
10.2.3.	<i>Poređenje propusne moći za RBS i RNP</i>	<i>96</i>
10.3.	MINIMIZACIJA CENE MREŽE	101
10.3.1.	<i>Cena mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)</i>	<i>102</i>
10.3.2.	<i>Cena mreže za rutiranjem po najkraćoj putanji (RNP)</i>	<i>103</i>
10.3.3.	<i>Poređenje cene mreže za RBS i RNP</i>	<i>103</i>
11.	ZAKLJUČAK	106
	LITERATURA	107
A.	PRIMENJEN OPTIMIZACIONI ALAT LP_SOLVE.....	112
	BIOGRAFIJA KANDIDATA	117

1. UVOD

Stalni razvoj Interneta i rast broja njegovih korisnika praćeni su velikim promenama u strukturi saobraćaja koji se mrežom prenosi. Naime, prenos elektronske pošte i tekstualnih podataka se odavno već podrazumeva, i danas je interesovanje korisnika pre svega usmereno ka raznim multimedijalnim sadržajima koji se mogu naći na mreži. Samim tim, raste udeo video i audio sadržaja u ukupnom saobraćaju. Ovi sadržaji su osetljivi na kašnjenje i gubitak paketa, i za njihov prenos potrebno je obezbediti veće kapacitete. Takođe, usled razvoja i popularizacije različitih *peer-to-peer* aplikacija, svi korisnici u mreži danas mogu generisati i primati velike količine saobraćaja, pri čemu raspodela tog saobraćaja među parovima korisnika može biti proizvoljna. Stoga se klasičan klijentsko-serverski model saobraćaja ne može više primenjivati, a predviđanje stanja saobraćaja u mreži je postalo znatno komplikovanije nego ranije.

Nemogućnost da se korisnički zahtevi i raspodela intenziteta među parovima korisnika precizno definišu otežava projektovanje mreže, koja mora biti u stanju da i u slučajevima maksimalnog opterećenja obezbedi neometan servis svim korisnicima. Ekonomski uslovi ne ostavljaju mnogo slobode prilikom projektovanja, već diktiraju potrebu da se što veći broj korisnika servisira korišćenjem raspoloživih resursa, ili uz minimalna ulaganja u nove kapacitete mreže. Dakle, neophodno je raspoložive resurse iskoristiti na optimalan način, kako bi se obezbedio kvalitetan i pouzdan prenos podataka između što većeg broja korisnika.

1.1. Razlozi za optimizaciju rutiranja

Kvalitetan odabir strategije rutiranja u paketskoj mreži u velikoj meri utiče na pouzdanost mreže i kvalitet prenosa podataka u njoj. Odabrana strategija rutiranja utiče i na stepen iskorišćenja različitih mrežnih resursa. Loše odabrano rutiranje može prouzrokovati stvaranje uskih grla u mreži i dovesti do blokade određenih linkova. Takođe, od primenjene strategije rutiranja zavisi kašnjenje pri prenosu paketa, kao i cena mreže. Većina protokola za rutiranje koji su danas u upotrebi ne uzima u obzir trenutna opterećenja linkova u mreži, već putanje određuje isključivo na osnovu mrežne topologije

i cene linkova u mreži. Cena dodeljena linkovima može zavisiti od različitih parametara, kao što su cena prenosa, kašnjenje, dužina putanje i slično. Od pravilnog izbora parametara koji utiču na cenu linkova, a samim tim i odabir putanja paketa u mreži, zavisi propusna moć mreže i njena ukupna cena.

U cilju odabiranja ispravne strategije rutiranja za neku mrežnu topologiju i date korisničke zahteve, vrši se *optimizacija rutiranja* u mreži. Proces optimizacije rutiranja za cilj ima određivanje putanja paketa tako da se postigne što bolja vrednost neke od mrežnih performansi. U zavisnosti od problema koji se rešava, optimizacija rutiranja može imati za cilj maksimizaciju propusne moći mreže, minimizaciju cene mreže, minimizaciju kašnjenja u prenosu i slično.

U ovome radu bavićemo se problemima optimizacije rutiranja u cilju *maksimizacije propusne moći* i *minimizacije cene* mreže. Pri tome, želimo da komunikaciona mreža bude neblokirajuća. To znači da mreža mora garantovati servis za sve korisnike koji generišu ili primaju saobraćaj u okviru dozvoljenih granica. Nezavisno od trenutnog stanja saobraćaja u mreži, korisnički zahtevi za slanje ili prijem paketa moraju biti usluženi, ukoliko ne izlaze iz okvira dozvoljenih vrednosti.

1.1.1. Problem minimizacije cene mreže

Kod problema minimizacije cene mreže, cilj je odrediti raspodelu kapaciteta linkova i strategiju rutiranja, tako da se omogući servisiranje korisničkih zahteva po što manjoj ceni. Cena mreže zavisi kako od cene za uspostavljanje novih linkova, tako i od cene rutera u mreži. Pri tome, cena rutera direktno zavisi od cene njegovih portova, dok cena portova zavisi pre svega od njihove brzine (odnosno, od kapaciteta linkova koji su na te portove prikačeni). Iz ovoga možemo zaključiti da je cena rutera u mreži proporcionalna sumi kapaciteta linkova koji su na njega povezani, te se minimizacija ukupne cene svih rutera u mreži svodi na određivanje optimalne raspodele kapaciteta linkova u mreži, tako da suma kapaciteta svih linkova bude minimalna.

Kod problema maksimizacije propusne moći mreže, cilj je odabrati rutiranje tako da se korisnicima u mreži garantuje maksimalni mogući saobraćaj. Garantovani saobraćaj za nekog korisnika je onaj koji se uvek može rutirati kroz mrežu, bez obzira na to koliko je drugih korisnika u mreži u tom trenutku aktivno i koja su trenutna opterećenja linkova, a pod uslovom da nijedan od aktivnih korisnika ne generiše više saobraćaja nego što mu je dozvoljeno.

1.1.2. Problem maksimizacije propusne moći mreže

Da bi se izvršila optimizacija rutiranja u cilju maksimizacije propusne moći mreže, pored poznavanja topologije mreže potrebna je i informacija o saobraćaju koji korisnici u mreži generišu i primaju. Međutim, kao što je već spomenuto, izuzetno dinamična priroda saobraćaja na Internetu danas u velikoj meri otežava predviđanje korisničkih zahteva, a samim tim otežava i proces optimizacije rutiranja. Naročito je nezahvalno i neprecizno predviđanje intenziteta saobraćaja među parovima čvorova u mreži. Naime, danas svaki korisnik može generisati i primati velike količine saobraćaja i usmeravati ih ka proizvoljnim čvorovima u mreži, te su moguće velike fluktuacije u intenzitetu saobraćaja između neka dva čvora.

Problemom optimizacije rutiranja u cilju maksimizacije propusne moći mreže bavili su se mnogi autori. Nepolinomijalno rešenje za opšti slučaj, zasnovano na dekompoziciji grafa mreže u strukturu na osnovu koje se donosi odluka o rutiranju, predloženo je u [1]. Poboljšanje ovog rešenja, koje radi u polinomijalnom vremenu, dato je u [2], [3]. Linearno programiranje u cilju pronalaženja optimalne strategije rutiranja primenjeno je u [4], [5]. Svi pomenuti radovi bave se pronalaženjem optimalnog rutiranja za one slučajeve kada je informacija o saobraćaju u mreži data u formi matrice, čiji su elementi intenziteti saobraćaja među parovima čvorova. Kao što je već pomenuto, procena saobraćajnih zahteva korisnika u ovakvoj formi je izuzetno komplikovana i neprecizna. Sa druge strane, znatno je lakše proceniti ukupan ulazni/izlazni saobraćaj u nekom čvoru mreže, nego tačnu raspodelu saobraćaja među parovima čvorova.

Autori u [6] i [7] uvode model koji omogućava da se optimizacija rutiranja izvrši samo uz poznavanje ukupnog ulaznog i izlaznog saobraćaja čvorova u mreži, i bez potrebe za poznavanjem tačne raspodele intenziteta tokova među parovima čvorova. Oni uvode takozvano rutiranje u dve faze, koje je zasnovano na balansiranju saobraćaja. Balansiranje saobraćaja im omogućava da saobraćajne zahteve između parova čvorova u mreži izraze u funkciji ulaznih/izlaznih saobraćaja u čvorovima mreže. Zatim uz pomoć linearnog programiranja pronalaze optimalno rutiranje za tako određene saobraćajne zahteve, tako da se minimizira cena mreže [6] ili maksimizira njena propusna moć [7].

1.2. Rutiranje sa balansiranjem saobraćaja

Okosnicu ovog rada predstavlja predlog nove šema rutiranja zasnovane na balansiranju saobraćaja i rutiranju po najkraćoj putanji, koju nazivamo *rutiranje sa*

balansiranjem saobraćaja (RBS). Rutiranje sa balansiranjem saobraćaja uvedeno je u [8], a zatim je u [9]-[17] vršena analiza njegovih performansi. Predložena šema rutiranja može se optimizovati tako da se ostvari maksimalna propusna moć mreže ili minimalna cena mreže, a za optimizaciju se koristi linearno programiranje.

Glavnu prednost predložene strategije rutiranja predstavlja realna mogućnost njene implementacije. Naime, rutiranje po najkraćoj putanji je u širokoj primeni, a predložena šema rutiranja je modifikacija koja se relativno lako realizuje. Realizacija predložene šeme rutiranja, zasnovana na implementaciji OSPF rutiranja koju je dao J. T. Moy [18], predstavljena je i analizirana u [13]-[16]. Dimenzija linearnog programa koji koristimo je prihvatljiva za praktičnu primenu, kao i vreme neophodno za njegovo izvršavanje. Primenom predložene šeme može se ostvariti servis za veći broj korisnika nego u slučaju kada se primenjuje regularno rutiranje po najkraćoj putanji, a da pri tom nema potrebe za promenama postojeće mrežne topologije. Takođe, u već postojećoj mreži, predložena šema omogućava servisiranje određenog broja korisnika po nižoj ceni u odnosu na cenu servisiranja istog skupa korisnika u mreži koja koristi rutiranje po najkraćoj putanji.

U predloženoj šemi se, kao i u [6], [7], svaki tok rutira u dva koraka. U prvom koraku, saobraćaj se od izvora usmerava ka čvorovima posrednicima u mreži, dok se u drugom koraku šalje od čvora posrednika ka svom odredištu. Razliku u odnosu na [6], [7] predstavlja to što se saobraćaj u svakom od koraka rutira po najkraćoj putanji. Fiksirana je, dakle, strategija rutiranja u dvama fazama predložene šeme, tako da ne možemo uticati na odabir putanje između izvorišnog čvora i čvora posrednika, odnosno između čvora posrednika i konačnog odredišta saobraćaja. Ono što možemo menjati je udeo saobraćaja koji se balansira preko određenog čvora posrednika. Fiksiranjem strategije rutiranja u obema fazama predložene šeme uspevamo da u velikoj meri uprostim linearni program – umesto $O(N^2M)$ promenljivih koliko ima model za optimizaciju rutiranja u [6], [7], predloženi model ih ima svega $O(N)$ za slučaj maksimizacije propusne moći [8]-[10], odnosno $O(M + N)$ za slučaj minimizacije cene mreže [11], [12], gde je N broj čvorova, a M broj linkova u mreži. Pokazuje se da ovakvim postupkom možemo ostvariti solidno povećanje garantovanog saobraćaja u čvoru mreže, odnosno smanjenje cene mreže, u odnosu na slučaj kada je u mreži primenjeno klasično rutiranje po najkraćoj putanji [8]-[17].

Pored svega navedenog, predložena šema značajno pojednostavljuje i proces rezervacije kapaciteta u mreži. Rezervacije kapaciteta su veoma značajne za multimedijalni saobraćaj, koji je osetljiv na kašnjenje i zahteva velike brzine prenosa. Obično se rezervacije kapaciteta vrše centralizovano, na osnovu informacija o trenutnom zauzeću svih čvorova i linkova na putanji od izvora od odredišta saobraćaja. Međutim, u slučaju primene predložene šeme rutiranja, dovoljno je izvršiti upit samo u izvorišnom i odredišnom čvoru za sesiju za koju se želi izvršiti rezervacija. Svaki od ovih čvorova raspolaže informacijom o trenutno rezervisanim kapacitetima, kao i o tome koliki je saobraćaj koji mu je u mreži garantovan (garantovani saobraćaj može se u svakom trenutku rutirati, bez obzira na aktivnost ostalih čvorova i trenutno zauzeće pojedinih linkova). Dovoljno je, stoga, prilikom uspostave rezervacija za neku sesiju, proveriti samo da li je razlika garantovanog i rezervisanog saobraćaja za odredišni i izvorišni čvor dovoljna da podrži zahteve sesije. Ukoliko je to slučaj, može se izvršiti rezervacija kapaciteta i ažurirati informacije o rezervisanim kapacitetima u izvorišnom i odredišnom čvoru sesije, bez potrebe da se o tome obaveštavaju drugi čvorovi u mreži.

1.3. Sadržaj i organizacija teze

U ovom radu biće prvo definisana strategija rutiranja sa balansiranjem saobraćaja. Zatim će biti data analiza performansi predložene strategije rutiranja i izvršeno njeno poređenje sa običnim rutiranjem po najkraćoj putanji. Konkretno, biće analizirana dva slučaja. U prvom slučaju vršićemo optimizaciju rutiranja sa balansiranjem saobraćaja, tako da se ostvari maksimalna propusna moć mreže. Propusna moć u kojoj se primenjuje rutiranje sa balansiranjem saobraćaja zatim će biti poređena sa propusnom moći mreže u kojoj je primenjeno rutiranje po najkraćoj putanji. Biće pokazano da se primenom balansiranja saobraćaja može ostvariti značajno povećanje propusne moći mreže, uz korišćenje već postojećih resursa. U drugom slučaju, biće vršena optimizacija rutiranja sa balansiranjem saobraćaja, tako da se ostvari minimalna cena mreže koja može da podrži neki zadati saobraćaj. Cena mreže koja koristi rutiranje sa balansiranjem saobraćaja zatim će biti poređena sa cenom mreže koja koristi obično rutiranje po najkraćoj putanji. Biće pokazano je da je cena mreže koja može da podrži neki zadati saobraćaj manja ukoliko se primenjuje predloženo rutiranje sa balansiranjem saobraćaja.

Analiza će biti vršena kako za regularne, tako i za realne mrežne topologije. Takođe, optimizacija rutiranja biće vršena i za slučajeve u kojima se zahteva da mreža

bude pouzdana. Pouzdana mreža mora biti u stanju da i u slučaju otkaza nekog resursa (linka ili rutera) omogući nesmetanu komunikaciju korisnika. U slučaju otkaza nekog resursa, veće opterećenje se postavlja na one resurse koji su ostali u funkciji. Zato se u ovakvoj mreži garantovani saobraćaj svakog korisnika mora dodatno ograničiti (čime se smanjuje propusna moć mreže), odnosno kapacitet svakog linka dodatno povećati (čime se povećava i cena mreže). Biće pokazano da su performanse predloženog rutiranja sa balansiranjem saobraćaja u svakom od analiziranih slučajeva iste ili bolje od performansi rutiranja po najkraćoj putanji, koje je danas u praksi široko rasprostranjeno.

U početnim poglavljima teze iznet je pregled osnovnih teorijskih koncepata. U Poglavlju 2 formalno se definiše rutiranje, kao i matematički modeli koji se koriste za opisivanje stanja saobraćaja u mreži. U ovom poglavlju dat je i detaljniji prikaz algoritma za rutiranje po najkraćoj putanji.

U Poglavlju 3 definisani su kriterijumi na osnovu kojih se vrši optimizacija rutiranja – minimalna cena, odnosno maksimalna propusna moć mreže. Detaljno su predstavljeni uslovi koje rutiranje treba da zadovolji da bi bilo smatrano optimalnim po nekom od zadatih kriterijuma. Zatim je u Poglavlju 4 dat kratak osvrt na linearno programiranje, simpleks metodu za rešavanje linearnih programa, a ukratko je opisan i softverski alat koji je korišćen za optimizaciju rutiranja.

U Poglavlju 5 definisan je pojam neblokirajućih mreža sa komutacijom veze i komutacijom paketa, dok je u Poglavlju 6 ukratko opisana primena balansiranja saobraćaja u paketskim svičevima velikog kapaciteta.

Opis predložene šeme rutiranja sa balansiranjem saobraćaja dat je u Poglavlju 7. U okviru ovog poglavlja dat je i opšti model linearnog programa koji se koristi za optimizaciju predložene šeme rutiranja, za saobraćaj proizvoljne distribucije. Pokazano je kako primena predložene šeme omogućava formulaciju optimizacionog problema u funkciji ukupnog saobraćaj generisanog ili primljenog u čvorovima mreže, bez potrebe za poznavanjem matrice saobraćaja.

U Poglavlju 8 izvršeno je poređenje predložene šeme rutiranja sa običnim rutiranjem po najkraćoj putanji, pri čemu je kao kriterijum na osnovu kog se vrši poređenje uzeta propusna moć mreže. Dat je model linearnog programa za optimizaciju rutiranja sa balansiranjem saobraćaja u cilju maksimizacije propusne moći mreže. Takođe, dat je i model koji se koristi za određivanje propusne moći mreže u kojoj je

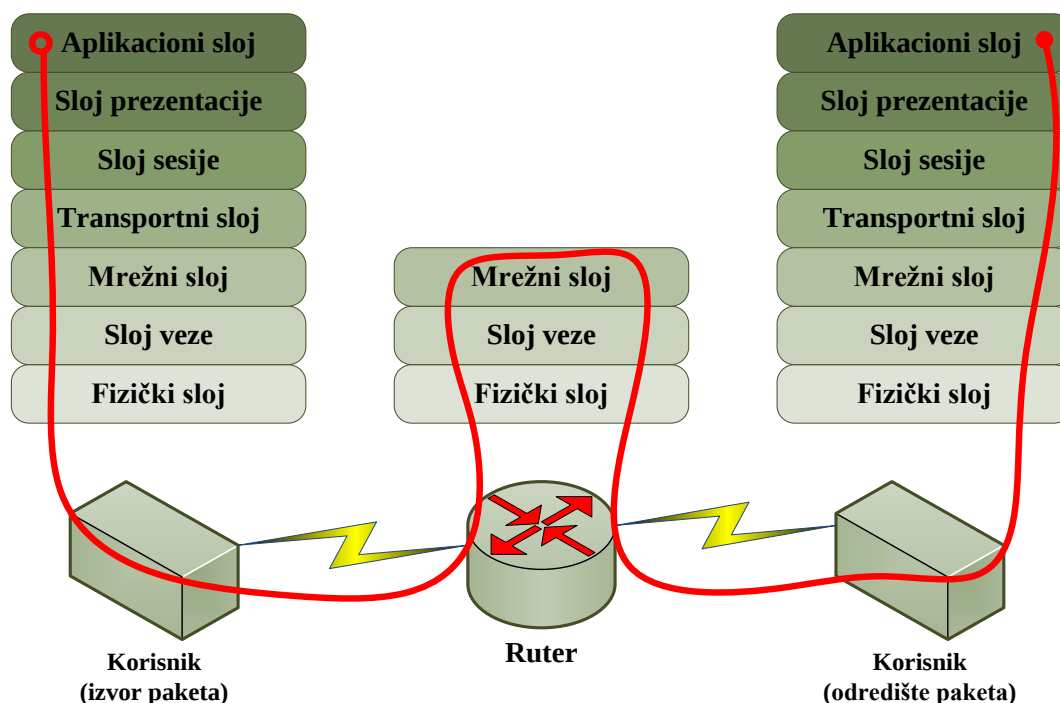
primenjeno rutiranje po najkraćoj putanji. Pokazano je da su performanse rutiranja sa balansiranjem saobraćaja značajno bolje od performansi običnog rutiranja po najkraćoj putanji, za saobraćaj proizvoljne distribucije. Poređenje je vršeno za regularne i realne mrežne topologije, i pokazano je da povećanje propusne moći raste sa stepenom čvorova u mreži.

U Poglavlju 9, poređena je cena mreže koja koristi rutiranje sa balansiranjem saobraćaja, sa cenom mreže koja koristi klasično rutiranje po najkraćoj putanji. Kao i u prethodnom slučaju, dat je model linearnog programa za optimizaciju rutiranja sa balansiranjem, tako da se odredi topologija mreže koja ima minimalnu cenu, a može da podrži rutiranje saobraćaja proizvoljne distribucije. Opisan je i način određivanja cene mreže sa rutiranjem po najkraćoj putanji. Zatim je izvršeno poređenje cena za realne mrežne topologije i pokazano da se upotrebom rutiranja sa balansiranjem može ostvariti ušteda.

Poglavlje 10 bavi se problemima maksimizacije propusne moći i minimizacije cene pouzdanih mreža. Pouzdana mreža treba da omogući neometan servis korisnicima i u slučaju otkaza nekog od resursa. Vršena je analiza za slučajeve jednostrukih otkaza resursa (čvorova ili linkova) u mreži. Za slučaj otkazivanja linkova analiza je vršena za realne mreže, dok je slučaj otkazivanja čvorova analiziran i za realne i za regularne mrežne topologije. Kao i u prethodnim poglavljima, ukazano je na prednost primene predložene šeme rutiranja u odnosu na klasično rutiranje po najkraćoj putanji.

2. RUTIRANJE

Pod pojmom rutiranja podrazumeva se prenos informacije u mreži, od njenog izvora do konačnog odredišta. Ovaj proces odvija se na trećem, mrežnom sloju OSI referentnog modela, kao na slici 2.0.1. Mrežni uređaji koji rade na trećem sloju su *ruteri*, a informacije se prenose u jedinicama koje se nazivaju *paketi*. Paketi na svom putu od izvora do odredišta obično prolaze kroz veći broj rutera. Za određivanje putanje kojom se paketi kreću kroz mrežu zaduženi su *algoritmi za rutiranje*.



Slika 2.0.1. Prenos informacije između dve aplikacije

U ovom poglavlju biće prvo reči o fazama koje obuhvata rutiranje i klasifikaciji algoritama za rutiranje. Zatim će biti definisani načini opisivanja saobraćaja u mreži, i data odgovarajuće matematičke definicije rutiranja za različite načine opisivanja saobraćaja. Na kraju će biti detaljnije opisano rutiranje po najkraćoj putanji.

2.1. Osnovni koncepti: proces rutiranja, tabele rutiranja, algoritmi za rutiranje

Kao što je već rečeno, proces rutiranja predstavlja prenos informacije (paketa) od izvora do odredišta u mreži. Ovaj proces odvija se na trećem (mrežnom) sloju OSI

referentnog sistema. Pri tome, paketi na svojoj putanji prolaze kroz veći broj uređaja koji rade na mrežnom sloju i koji se nazivaju ruterima. Kako bi pravilno odabrao putanju između izvora i odredišta paketa, mrežni sloj mora poznavati topologiju mreže. Putanja se bira na osnovu neke zadate metrike, obično tako što se linkovima u mreži dodele cene u skladu sa tom metrikom, a zatim se među raspoloživim putanjama između dva čvora mreže odabira ona sa minimalnom ukupnom cenom linkova.

U procesu rutiranja razlikujemo dve faze. U okviru prve faze vrši se *određivanje putanje* paketa između parova čvorova u mreži, dok drugu fazu predstavlja fizičko *prosleđivanje paketa* kroz mrežu. Za određivanje putanje paketa koriste se algoritmi za rutiranje. Prosleđivanje paketa zatim vrši za to specijalizovani deo rutera, na osnovu kontrolnih informacija koje dobija kao rezultat izvršavanja algoritma za rutiranje. Najčešće se radi o specijalizovanom hardveru, dizajniranom tako da omogući što brži prolazak paketa kroz ruter. U ovom odeljku objasnićemo ukratko postojeće faze u procesu rutiranja.

2.1.1. *Određivanje putanje*

Za određivanje putanja kojima se šalju paketi između čvorova u mreži koriste se algoritmi za rutiranje. Oni imaju zadatak da na osnovu informacija o mrežnoj topologiji odrede putanje koje će se koristiti za komunikaciju između pojedinih parova čvorova u mreži. Koji algoritam za rutiranje će se koristiti definisano je protokolom za rutiranje. Pored toga, protokolom su definisani i formati kontrolnih i korisničkih paketa, način na koji se kontrolne informacije i informacije o odabranim putanjama čuvaju, vremenski intervali značajni za dobijanje informacija o mrežnoj topologiji i sl.

Odluka o tome koja od više postojećih putanja između dva čvora u mreži će biti upotrebljena za rutiranje saobraćaja donosi se na osnovu unapred odabrane metrike. Metrika može biti propusni opseg, kašnjenje, broj skokova između posredničkih rutera na putanji, cena prenosa, opterećenje, pouzdanost putanje, itd. Najčešće se linkovima u mreži dodeljuju cene, koje na određen način zavise od nekih ili svih ovih pomenutih parametara, a zatim se saobraćaj između dva čvora u mreži rutira po putanji sa najmanjom ukupnom cenom. Informacije o odabranim putanjama čuvaju se u tabeli rutiranja.

Način na koji se formira tabela rutiranja, kao i njen sadržaj, zavise kako od primenjenog protokola za rutiranje, tako i od konkretne implementacije tog protokola. Prvenstveno, u tabeli rutiranja se mora naći informacija o sledećem ruteru na putanji

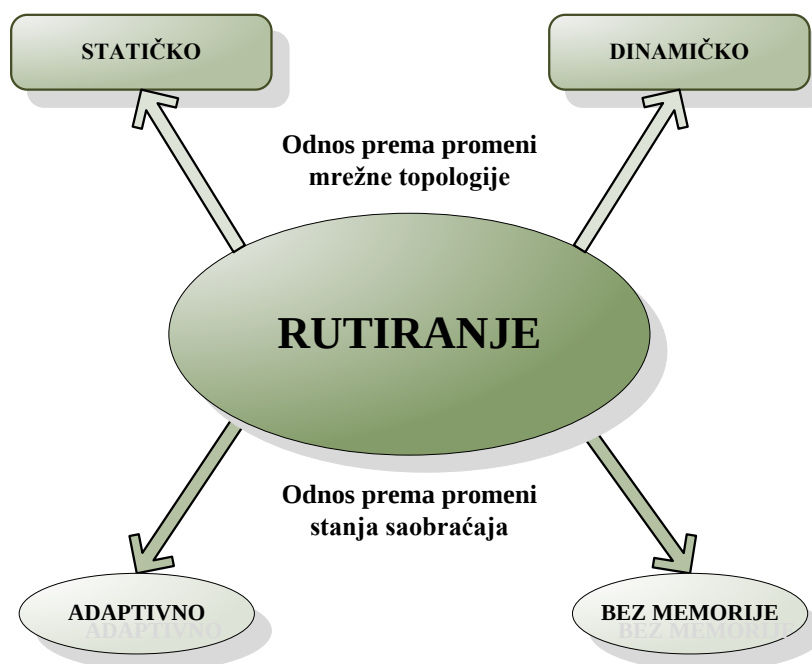
paketa ka konačnom odredištu, odnosno o interfejsu na koji treba proslediti primljeni paket. Pored toga, u tabeli rutiranja se mogu naći i alternativne putanje paketa ka odredištu, kao i parametri kojima se definiše poželjnost određene putanje, ukoliko postoji više različitih putanja iste ili slične cene.

2.1.2. *Prosleđivanje paketa*

Prosleđivanje paketa je proces za koji je obično zadužen specijalizovani hardver rutera. Prebacivanjem ove faze rutiranja na hardver, dobija se na brzini obrade paketa. Naime, komunikacija između dva rutera u mreži odvija se na drugom sloju, na osnovu MAC adresa. Po prijemu paketa i uklanjanju zaglavlja protokola drugog sloja, ruter analizira zaglavlje protokola trećeg sloja. Ukoliko na osnovu adrese odredišta ruter zaključi da primljeni paket nije namenjen korisniku na nekoj od njegovih lokalnih mreža, ruter će izvršiti pretraživanje tabele rutiranja, kako bi paket prosledio dalje. Uvidom u odgovarajuću tabelu rutiranja određuje se koji je sledeći ruter na putanji paketa ka odredištu. Ruter prosleđuje paket na interfejs na kome se nalazi sledeći ruter, dodajući zaglavlje drugog sloja u skladu sa medijumom i protokolom koji se na datom interfejsu koriste. Kao odredišna adresa na drugom sloju postavlja se MAC adresa sledećeg rutera na putanji paketa. Dakle, paket kroz mrežu prolazi nepromenjen, a u svakom skoku između posredničkih rutera na putanji menja se zaglavlje drugog sloja.

2.2. Podela algoritama za rutiranje

Kao što je rečeno u prethodnom odeljku, algoritmi za rutiranje imaju zadatak da odrede putanju kojom se kreću paketi između čvorova mreže. Podela algoritama za rutiranje može se izvršiti na osnovu različitih karakteristika, poput načina na koji se prikupljaju kontrolne informacije, distribuiranosti izvršavanja, mesta na kome se donosi odluka o putanji paketa (izvor ili trenutni ruter), i sl. U ovom odeljku bavićemo se podelom algoritama za rutiranje na osnovu dva osnovna kriterijuma: odnosa prema promenama mrežne topologije i odnosa prema promenama stanja saobraćaja u mreži. Naime, rutiranje se može (ali i ne mora) prilagođavati ovim promenama. U zavisnosti od konkretnog problema i mreže koja se posmatra, moguće je primeniti različite pristupe. Klasifikacija rutiranja predstavljena je na slici 2.2.1.



Slika 2.2.1. Podela algoritama za rutiranje

Sa obzirom na jednostavnost manjih mreža, u njima je moguće primeniti *statičko rutiranje*, odnosno administrativno zadati putanje od izvora do odredišta, tako da one budu optimalne u skladu sa unapred odabranim kriterijumom. Složene topologije većih komunikacionih mreža zahtevaju automatizaciju procesa rutiranja. Automatizacija treba da omogući brže prilagođavanje rutiranja promenama mrežne topologije, koje nastaju usled priključivanja novih i odlazaka starih korisnika iz mreže. Ovakvo rutiranje naziva se *dinamičko rutiranje*. Naravno, veća prilagodljivost promenama u mrežnoj topologiji koju ovakav pristup omogućava plaćena je osetljivošću na probleme poput pojave petlji rutiranja i oscilacija na rutama.

Kada je u pitanju odnos rutiranja prema promenama stanja saobraćaja u mreži, ono može biti *adaptivno* ili *bez memorije (oblivious)*. Adaptivno rutiranje prilagođava se postojećem saobraćaju i prilikom odabiranja ruta u obzir uzima trenutna opterećenja linkova, kao i aktivnost svih čvorova u mreži. Pri odabiru ruta izbegavaju se veoma opterećeni linkovi. Sa druge strane, rutiranje bez memorije je ono u kome se odluke o rutiranju donose samo na osnovu konačnog odredišta saobraćaja. Kod ovakvog rutiranja, putanja ili skup mogućih putanja od izvora ka nekom odredištu su unapred određeni, tako da se zadovolje neka ograničenja po pitanju odabrane metrike. U praksi se koncept adaptivnog rutiranja ređe koristi, jer isuviše česte promene saobraćaja u mreži zahtevaju

podjednako česte promene u načinu rutiranja, a ovo se može negativno odraziti na kvalitet servisa.

Dakle, u praksi se uglavnom susrećemo sa dinamičkim rutiranjem bez memorije. Drugim rečima, odabir putanje zavisi od topologije mreže, ali obično ne i od trenutog stanja saobraćaja u njoj.

2.3. Matematičko modelovanje rutiranja

U prethodnom odeljku razmatrali smo neke tehničke aspekte vezane za proces rutiranja, način određivanja putanje paketa i prosleđivanje paketa u mreži. Međutim, kako bi se mogla izvršiti formalna analiza rutiranja, neophodno je usvojiti matematički model kojim će ono biti predstavljeno. U ovom odeljku biće uvedena notacija koja se koristi za matematičko opisivanje mreže i definisani načini na koje se predstavlja saobraćaj u mreži, a zatim će biti data matematička definicija rutiranja.

U cilju matematičke analize, mreža se predstavlja kao graf $\Gamma = (V, E)$, gde je V skup čvorova grafa – kojima odgovaraju ruteri unutar mreže, a E skup grana grafa – kojima odgovaraju linkovi između rutera. Ukupan broj čvorova u mreži označen je sa $|V| = N$, a broj linkova sa $|E| = M$. Granama grafa dodeljuju se težine, koje predstavljaju kapacitete odgovarajućih linkova.

2.3.1. Načini opisivanja stanja saobraćaja u mreži

Od toga na koji način će se opisivati saobraćaj u mreži u velikoj meri zavisi kompleksnost matematičkog modela rutiranja i manipulacija tim modelom. Najčešće se susrećemo sa dva načina opisivanja stanja saobraćaja: na bazi intenziteta saobraćaja između parova čvorova u mreži (*matrica saobraćaja*) ili na bazi ukupnog saobraćaja generisanog, tj. primljenog u čvoru mreže (*vektor saobraćaja*). U ovom pododeljku biće više reči o ova dva načina opisivanja saobraćaja u mreži, tj. o njihovim prednostima i manama.

i) *Matrica saobraćaja*

Matrica saobraćaja je najčešće korišćeni način opisivanja stanja saobraćaja u mreži. Elementi matrice saobraćaja su intenziteti tokova između parova korisnika u mreži.

Definicija 2.3.1. Za mrežu sa N čvorova saobraćajna matrica saobraćaja data je kvadratnom matricom $\mathbf{TM}=[d_{ij}]_{N \times N}$, gde element d_{ij} odgovara intenzitetu saobraćaja između čvorova i i j date mreže:

$$\mathbf{TM} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & \dots & d_{NN} \end{bmatrix} \quad (2.3.1)$$

Predviđanje mogućih oblika matrice saobraćaja je izuzetno teško. I najkvalitetnije tehnike koje se koriste za predviđanje vrednosti elemenata ove matrice imaju grešku od oko 20% u najboljem slučaju, dok u ekstremnim slučajevima greška može biti i veća od 100% [19]. Odnosno, greška pri proceni intenziteta saobraćaja između neka dva čvora u mreži može biti i veća od stvarne vrednosti tog saobraćaja. U skorije doba, proces predviđanja vrednosti elemenata matrice saobraćaja dodatno je otežan porastom udela saobraćaja generisanog od strane aplikacija između ravnopravnih korisnika (*peer-to-peer*). Za ovakav saobraćaj izuzetno je komplikovano proceniti ko će kada i sa kim želeći da ostvari komunikaciju, kao i koliki će biti intenzitet saobraćaja između dva korisnika. Stanje saobraćaja u mreži se menja znatno brže nego ranije, a pri tome ne postoje jasno utvrđena pravila ponašanja korisnika na način na koji su postojala kod serversko – klijentskih aplikacija, što otežava proces predviđanja.

ii) *Vektor ulaznog i izlaznog saobraćaja čvorova*

Alternativni način opisivanja saobraćaja u mreži predstavljaju vektori ulaznog i izlaznog saobraćaja u čvorovima mreže.

Definicija 2.3.2. Za mrežu sa N čvorova, vektor izlaznog saobraćaja $\mathbf{S}$ dat je izrazom

$$\mathbf{S}=[s_1, s_2, \dots, s_N], \quad (2.3.2)$$

gde je sa s_i označen ukupan saobraćaj generisan u čvoru i date mreže.

Uočimo da između elemenata vektora izlaznog saobraćaja i saobraćajne matrice važi sledeća relacija:

$$s_i = \sum_{j \in V} d_{ij} . \quad (2.3.3)$$

Definicija 2.3.3. Za mrežu sa N čvorova, vektor izlaznog saobraćaja $\mathbf{R}$ dat je izrazom

$$\mathbf{R}=[r_1, r_2, \dots, r_N], \quad (2.3.4)$$

gde je sa r_j označen ukupan saobraćaj primljen u čvoru j date mreže.

Veza između elemenata vektora izlaznog saobraćaja i elemenata matrice saobraćaja data je izrazom

$$r_j = \sum_{i \in V} d_{ij}. \quad (2.3.5)$$

Za razliku od matrice saobraćaja, vektori ulaznog i izlaznog saobraćaja su znatno jednostavniji za predviđanje. Nije neophodno poznavati distribuciju saobraćaja među parovima čvorova u mreži, već samo ukupan saobraćaj generisan i primljen u nekom čvoru. U svakom slučaju znamo barem gornju granicu za elemente ovih vektora – ukupan saobraćaj generisan u čvoru ne može premašiti kapacitet izlaznih linkova, a ukupan primljeni saobraćaj ne može biti veći od kapaciteta ulaznih linkova za dati čvor.

2.3.2. Matematička definicija rutiranja

Kao što smo već rekli, stanje saobraćaja u mreži može biti opisano matricom saobraćaja ili vektorima ulaznog i izlaznog saobraćaja. U ovom odeljku biće date matematičke definicije rutiranja za oba načina predstavljanja saobraćaja u mreži.

Definicija 2.3.4. *Neka je mreža predstavljena grafom, čiji čvorovi $i \in V$ odgovaraju ruterima, a grane $l \in E$ odgovaraju linkovima u mreži. Za slučaj kada je saobraćaj opisan matricom saobraćaja $\mathbf{TM}=[d_{ij}]_{N \times N}$, rutiranje je definisano kao skup koeficijenata f_{ij}^l , takvih da zadovoljavaju uslove*

$$\begin{aligned} \forall l \in E \quad \forall i, j \in V: 0 \leq f_{ij}^l \leq 1 \\ \forall n, i, j \in V: \sum_{l \in OUT(n)} f_{ij}^l - \sum_{l \in IN(n)} f_{ij}^l = \begin{cases} 1, & n = i \\ 0, & \text{inače} \\ -1, & n = j. \end{cases} \end{aligned} \quad (2.3.6)$$

Ovde je sa $OUT(n)$ označen skup svih linkova koji izlaze iz čvora n , dok je $IN(n)$ skup svih ulaznih linkova za isti čvor.

Ukupan broj koeficijenata kojima je rutiranje definisano iznosi MN^2 , gde je N broj čvorova, a M broj linkova u mreži. Praktično, koeficijent rutiranja f_{ij}^l određuje koji udeo toka d_{ij} se rutira po linku l . Ukupan saobraćaj koji po linku l razmenjuje par čvorova (i, j) jednak je $f_{ij}^l d_{ij}$.

Posmatrajmo sada slučaj kada je saobraćaj u mreži opisan vektorima ulaznog i izlaznog saobraćaja.

Definicija 2.3.4. *Neka je mreža predstavljena grafom, čiji čvorovi $i \in V$ odgovaraju ruterima, a grane $l \in E$ odgovaraju linkovima u mreži. Pretpostavimo da na putanji saobraćaja od nekog izvora ka konačnom odredištu ne može biti petlji. Za slučaj kada je saobraćaj opisan vektorom ulaznog saobraćaja $\mathbf{S}=[s_1, s_2, \dots, s_N]$ i vektorom izlaznog saobraćaja $\mathbf{R}=[r_1, r_2, \dots, r_N]$, rutiranje je definisano koeficijentima g_i^l , takvima da važi*

$$\begin{aligned} \forall i \in V, \forall l \in E: 0 \leq g_i^l &\leq 1 \\ \forall i \in V: \sum_{l \in OUT(i)} g_i^l &= 1 \\ \forall n \in V: \sum_{l \in OUT(n)} \sum_{i \in V} g_i^l s_i - \sum_{l \in IN(n)} \sum_{i \in V} g_i^l s_i &= s_n - r_n. \end{aligned} \quad (2.3.7)$$

Koeficijenti g_i^l određuju udeo saobraćaja generisanog u nekom čvoru i koji se rutira preko linka l . Ovih koeficijenata ima ukupno MN , odnosno N puta manje nego u slučaju kada je stanje saobraćaja u mreži opisano saobraćajnom matricom.

Već je rečeno da su elementi vektora saobraćaja jednostavniji za predviđanje od elemenata matrice saobraćaja. Takođe, znatno manji broj koeficijenata kojima se opisuje rutiranje ukazuje na prednost upotrebe vektora saobraćaja u odnosu na matricu saobraćaja.

2.4. Rutiranje po najkraćoj putanji

Većina algoritama za rutiranje koji se danas susreću u praksi za prenos paketa između neka dva čvora mreže koristi jedinstvenu putanju, sa minimalnom cenom. Ova putanja određuje se Dijkstra algoritmom, o kome će biti više reči u tekstu koji sledi. Od načina na koji se dodeljuju cene linkovima, može se raditi o najkraćoj, najbržoj, putanji sa najmanje koraka u prenosu i sl. Radi jednostavnosti, u daljem tekstu ovakvo rutiranje nazivaćemo rutiranjem po najkraćoj putanji, iako u opštem slučaju cena linkova ne mora biti proporcionalna njihovoj dužini.

2.4.1. Definicija najkraće putanje

U ovom odeljku biće data matematička definicija rutiranja po najkraćoj putanji. Prvo ćemo definisati pojam najkraće putanje.

Posmatrajmo mrežu datu grafom $\Gamma = (V, E)$, gde je V skup čvorova, a E skup linkova u mreži. Neka su sve grane grafa Γ usmerene, i neka je cena grane između čvorova a i b data sa w_{ab} .

Definicija 2.4.1. Neka je sa P_{ij} označen skup svih mogućih putanja između čvorova i i j mreže $\Gamma = (V, E)$. Najkraća putanja između čvorova i i j je ona putanja $p \in P_{ij}$, za koju je ostvarena minimalna vrednost ukupne cene $\sum_{l \in p} w_l$, gde je w_l cena dodeljena linku l koji se nalazi na putanji p .

U slučaju kada postoji više putanja sa istom (minimalnom) cenom, algoritam za određivanje najkraće putanje odabira jednu od njih i proglašava je za najkraću putanju između dva čvora u mreži. Sav saobraćaj između čvorova i i j usmerava se ovom putanjom, tako da koeficijenti rutiranja mogu imati samo jednu od dve vrednosti $f_{ij}^l \in \{0, 1\}$. Poznavanje matrice saobraćaja je neophodno da bi se odredilo trenutno opterećenje linkova u mreži u kojoj je primenjeno rutiranje po najkraćoj putanji.

Definisanjem cene linka postiže se da se kao najkraća odabira putanja željenih performansi. Naime, ukoliko cena linka odgovara njegovoj dužini, to je najjeftinija putanja istovremeno i najkraća putanja između dva čvora. Ukoliko se, pak, svim linkovima dodele iste cene, onda je odabrana putanja ona sa najmanjim brojem rutera (hopova) koje paket mora da prođe od izvora do odredišta. Linkovima se mogu dodeliti i cene obrnuto proporcionalne kapacitetima linkova. Ovakvim odabirom metrike, forsira se rutiranje saobraćaja po linkovima koji su većeg kapaciteta. Upravo na ovaj način se dodeljuju cene linkovima u protokolu za rutiranje OSPF (*Open Shortest Path First*), koji se često susreće u praksi.

Nezavisno od toga kako se u konkretnoj implementaciji algoritma modeluje cena linka, u daljem tekstu govorićemo uvek o rutiranju po najkraćoj putanji. U odeljku koji sledi, biće detaljnije opisan algoritam koji se koristi za određivanje najkraće putanje.

2.4.2. Dijkstra algoritam

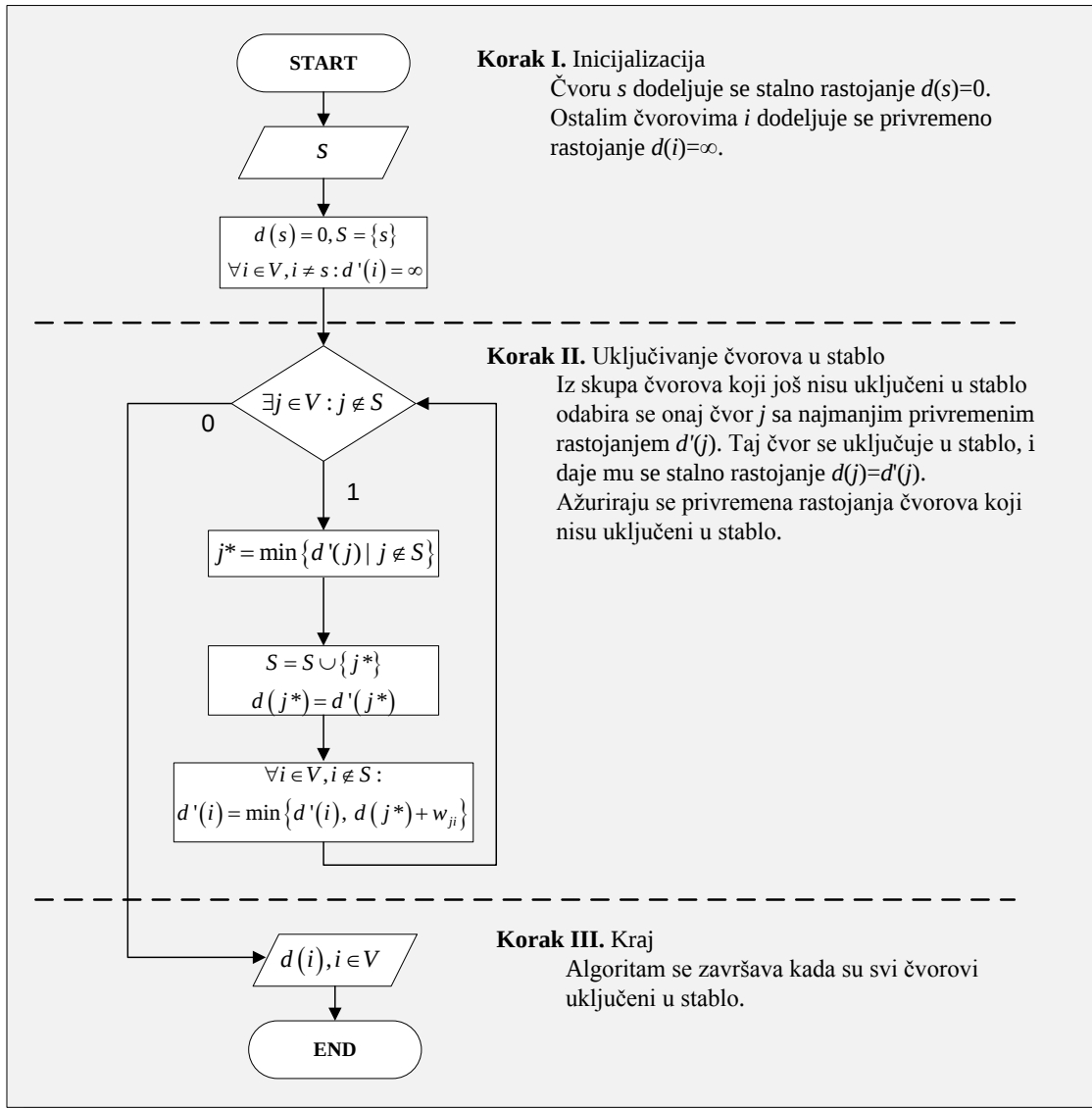
Za određivanje najkraće putanje od izvorišnog čvora s do ostalih čvorova u mreži koristi se Dijkstra algoritam, koji će biti opisan u ovom pododeljku.

Dijkstra algoritmom formira se stablo najkraćih putanja, sa izvorišnim čvorom s kao korenom stabla. Svakom čvoru iz mreže $G = (V, E)$ pridružena je funkcija rastojanja,

$d: V \rightarrow \mathbb{R}^+ \cup \{0\}$. U svakom trenutku izvršavanja algoritma, postoje dve grupe čvorova: oni za koje je funkcija rastojanja privremena i oni za koje je ona stalna. Privremena vrednost funkcije rastojanja $d(i)$ za neki čvor i predstavlja gornju granicu za rastojanje čvora i od izvora s . Ukoliko je vrednost funkcije rastojanja $d(i)$ stalna, onda je ona jednaka najkraćem rastojanju čvora i od izvora s . Čvorovi sa stalnom vrednošću funkcije rastojanja su uključeni u stablo najkraćih putanja, dok oni sa privremenom vrednošću nisu. U toku inicijalizacije, čvoru s dodeljuje se stalno rastojanje $d(s) = 0$, dok se svim ostalim čvorovima u mreži dodeljuje beskonačno privremeno rastojanje od čvora s , $d(i) = \infty, i \neq s$. Čvor s postavlja se u koren stabla najkraćih putanja. Ostali čvorovi mreže u ovom trenutku nisu uključeni u stablo.

Algoritam zatim uključuje ostale čvorove u stablo, na način koji ćemo sada opisati. U svakom koraku odabira se među čvorovima koji nisu uključeni u stablo onaj sa minimalnim privremenim rastojanjem od izvora s . Neka je u pitanju čvor j . Čvor j se uključuje u stablo. Rastojanje $d(j)$ postaje stalno rastojanje čvora j od čvora s . Zatim se vrši ažuriranje privremenih rastojanja čvorova koji još nisu uključeni u stablo. Privremeno rastojanje za neki čvor i postaje $\min(d(i), d(j) + w_{ji})$, gde je w_{ji} težina (cena) linka od čvora j do čvora i . Postupak se ponavlja dok svi čvorovi mreže ne budu uključeni u stablo. Tako dobijeno stablo predstavlja *stablo najkraćih putanja* od čvora s koji se nalazi u korenu, do svih ostalih čvorova mreže.

Na slici 2.5.1 dat je šematski prikaz Dijkstra algoritma.



Slika 2.5.1. Dijkstra algoritam

3. KRITERIJUMI OPTIMIZACIJE RUTIRANJA

Optimizacija rutiranja vrši se u cilju poboljšanja njegovih performansi. Kao kriterijum na osnovu kog se vrši optimizacija uzima se neki od parametara, na primer cena rutiranja, brzina prenosa, količina ostvarenog saobraćaja, zagušenje i sl. Optimalno rutiranje je ono rutiranje koje za koje se postiže najbolja vrednost (maksimalna ili minimalna) nekog posmatranog parametra.

U tekstu koji sledi, koncentrisaćemo se na dva kriterijuma koji nas interesuju: *propusna moć mreže* i *cena mreže*. Biće definisana ova dva kriterijuma i objašnjeno šta se podrazumeva pod optimalnim rutiranjem u odnosu na svaki od njih.

3.1. Maksimizacija propusne moći mreže

Posmatrajmo slučaj kada je mrežna topologija u potpunosti poznata – dakle, poznati su i raspored čvorova i linkova u mreži, kao i raspoloživi kapacitet na svim linkovima. Želimo da na što kvalitetniji način iskoristimo resurse koji su nam na raspolaganju i da omogućimo servis za što veći broj korisnika. Definisaćemo prvo osnovne pojmove, kao što su propusna moć mreže i garantovani saobraćaj za čvor mreže.

Definicija 3.1.1. *Garantovani saobraćaj čvora predstavlja onu (konačnu) vrednost saobraćaja koja uvek može biti rutirana kroz postojeću mrežu, pod uslovom da i ostali čvorovi ne generišu više saobraćaja od vrednosti garantovanog saobraćaja koja im je dodeljena.*

Definicija 3.1.2. *Propusna moć mreže predstavlja sumu garantovanih saobraćaja za sve čvorove u mreži.*

Iz gore navedenog sledi da je, da bi se ostvarila maksimalna propusna moć mreže i servisirao maksimalan broj korisnika, neophodno maksimizovati garantovani saobraćaj svakog čvora u mreži. Pokazuje se da je problem maksimizacije garantovanog saobraćaja čvora ekvivalentan problemu minimizacije zagušenja u mreži, što ćemo sad objasniti. Definišimo prvo pojam zagušenja.

Definicija 3.1.3. *Zagušenje Q predstavlja maksimalno iskorišćenja linka u mreži i dato je izrazom*

$$Q = \max_{l \in E} U^l = \max_{l \in E} \frac{L^l}{C^l}, \quad (3.1.1)$$

gde je sa L^l označeno opterećenje linka l , a sa C^l njegov kapacitet.

Link sa maksimalnim iskorišćenjem prvi je u opasnosti da u slučaju povećanja intenziteta tokova postane preopterećen. S obzirom na to da zagušenje direktno zavisi od opterećenja linkova, to je ono u opštem slučaju funkcija kako odabranog rutiranja, tako i stanja saobraćaja u mreži. Ukoliko je za neko stanje saobraćaja i primenjeno rutiranje postignuto zagušenje Q , tada je maksimalan faktor kojim se intenziteti svih tokova u mreži mogu skalirati (tako da matrica saobraćaja ili vektor ulaznog saobraćaja ne izlazi iz dozvoljenih okvira, koji mogu biti rutirani) jednak $1/Q$. Dakle, što je zagušenje manje, faktor kojim se svi tokovi u mreži mogu skalirati je veći. Stoga je problem maksimizacije saobraćaja ekvivalentan problemu minimizacije zagušenja. Praktično, rutiranje za koje se ostvaruje minimalno zagušenje je ono koje na uravnotežen način raspoređuje opterećenje među svim linkovima mreže. Na ovaj način izbegava se stvaranje uskih grla, i povećava propusna moć.

Pojedini autori prilikom rešavanja problema maksimizacije propusne moći mreže umesto zagušenja posmatraju propusnost mreže.

Definicija 3.1.4. *Propusnost mreže definisana je kao faktor θ kojim se mogu skalirati svi elementi u matrici ili vektoru saobraćaja, tako da saobraćaj bude rutiran kroz mrežu.*

Na osnovu prethodne diskusije, očigledno je da važi da je $\theta = 1/Q$, te je problem minimizacije zagušenja ekvivalentan problemu maksimizacije propusnosti. Mi ćemo u našim analizama rešavati problem minimizacije zagušenja, ali ćemo u Poglavlju 6 u pojedinim dokazima preuzetim iz literature koristiti i pojam propusnosti mreže.

Označimo stanje saobraćaja u mreži sa $\mathbf{TS}$, nezavisno od toga da li je ono opisana matricom saobraćaja ili vektorima ulaznog i izlaznog saobraćaja. Rekli smo već da je zagušenje funkcija stanja saobraćaja i primenjenog rutiranja. Dakle,

$$Q = Q(\mathbf{RS}, \mathbf{TS}), \quad (3.1.2)$$

gde je sa $\mathbf{RS}$ označena primenjena šema rutiranja, koja je data ili kao skup koeficijenta $\{f_{ij}\}$ ili kao skup koeficijenata $\{g_i\}$, već u zavisnosti od toga kako je definisano stanje

saobraćaja u mreži. Sada za dato stanje saobraćaja u mreži možemo formalno definisati optimalno rutiranje.

Definicija 3.1.5. *Za dato stanje saobraćaja u mreži $\mathbf{TS}$, optimalno rutiranje $\mathbf{RS}_{(\mathbf{TS})}^{opt}$ je ono za koje je ostvareno minimalno zagušenje*

$$Q(\mathbf{RS}_{(\mathbf{TS})}^{opt}, \mathbf{TS}) = \min_{\mathbf{RS}} Q(\mathbf{RS}, \mathbf{TS}). \quad (3.1.3)$$

Naravno, saobraćaj u mreži se menja sa vremenom. Rutiranje koje je optimalno za jedno stanje saobraćaja u mreži, za neko drugo može biti daleko od zadovoljavajućeg. Sa druge strane, kod uobičajenog pristupa u kome se definiše rutiranje bez memorije nema adaptacije na promene u saobraćaju. Rute se biraju unapred i ne zavise od stanja linkova. Stoga je prilikom odabiranja šeme rutiranja neophodno proceniti moguća stanja saobraćaja u mreži i rutiranje odabrati tako da bude optimalno za skup mogućih stanja saobraćaja. Pogledajmo sad na koji način se definiše optimalno rutiranje na takvom jednom skupu.

Posmatrajmo neko stanje saobraćaja $\mathbf{TS}$ i neko rutiranje $\mathbf{RS} \neq \mathbf{RS}_{(\mathbf{TS})}^{opt}$, koje nije optimalno rutiranje za dati saobraćaj. Pokazatelj kvaliteta rutiranja $\mathbf{RS}$ je odnos zagušenja $Q(\mathbf{RS}, \mathbf{TS})$ prema minimalnom mogućem zagušenju za dati saobraćaj, $Q(\mathbf{RS}_{(\mathbf{TS})}^{opt}, \mathbf{TS})$ [4]. Dakle, performanse rutiranja $\mathbf{RS}$ opisane su sa

$$\text{perf}(\mathbf{RS}, \mathbf{TS}) = \frac{Q(\mathbf{RS}, \mathbf{TS})}{Q(\mathbf{RS}_{(\mathbf{TS})}^{opt}, \mathbf{TS})}. \quad (3.1.4)$$

Ukoliko posmatramo skup mogućih stanja saobraćaja $T = \{\mathbf{TS}\}$, tada je kvalitet rutiranja $\mathbf{RS}$ na ovom skupu definisan kao

$$\text{perf}(\mathbf{RS}, T) = \max_{\mathbf{TS} \in T} \text{perf}(\mathbf{RS}, \mathbf{TS}) = \max_{\mathbf{TS} \in T} \frac{Q(\mathbf{RS}, \mathbf{TS})}{Q(\mathbf{RS}_{(\mathbf{TS})}^{opt}, \mathbf{TS})}. \quad (3.1.5)$$

Drugim rečima, u obzir se uzima najgori slučaj, odnosno ono stanje saobraćaja za koje posmatrano rutiranje najviše odstupa od optimalnog. Što je niža vrednost za $\text{perf}(\mathbf{RS}, T)$, to je rutiranje bolje. Pri tome važi da je uvek $\text{perf}(\mathbf{RS}, T) \geq 1$.

Definicija 3.1.6. *Optimalno rutiranje za skup T je ono za koje se postiže minimalna vrednost performansi $\text{perf}(\mathbf{RS}, T)$:*

$$\text{perf}(\mathbf{RS}_{(T)}^{\text{opt}}, T) = \min_{\mathbf{RS}} \text{perf}(\mathbf{RS}, T) \quad (3.1.6)$$

Problem je, naravno, odrediti skup mogućih stanja saobraćaja, s obzirom na to da intenziteti tokova mogu uzimati kontinualne vrednosti iz nekog skupa. Problem je utoliko veći ukoliko je saobraćaj opisan matricom saobraćaja. Pokazaće se da nam primena balansiranja saobraćaja omogućava da problem postavimo isključivo u funkciji vektora ulaznog i izlaznog saobraćaja, koji su jednostavniji za previđanje, o čemu će biti više reči u Poglavlju 5, prilikom opisa predložene šeme rutiranja.

3.2. Minimizacija cene mreže

U slučaju kada nije poznata mrežna topologija, ali su poznati saobraćajni zahtevi koje mreža treba da podrži, kriterijum koji se može zadati kao pokazatelj kvaliteta rutiranja jeste cena takve mreže.

Definicija 3.2.1. *U opštem slučaju, cena mreže modeluje se kao zbir cena njenih linkova i čvorova.*

$$\mathbf{C} = \sum_{l \in E} \mathbf{C}_L^l + \sum_{n \in V} \mathbf{C}_N^n \quad (3.2.1)$$

Cena linka l označena je sa $\mathbf{C}_L^l$, zavisi od njegove dužine, i predstavlja troškove postavljanja novog linka. Cena čvora n označena je sa $\mathbf{C}_N^n$ i zavisi prvenstveno od cene njegovih portova. Cena porta zavisi od brzine porta, koja mora odgovarati kapacitetu priključenog linka. Stoga možemo smatrati da je cena čvora mreže proporcionalna ukupnom kapacitetu linkova koji su vezani za taj čvor. Kao u slučaju minimizacije zagušenja, cena mreže koja može da podrži zadati saobraćaj zavisi od odabranog rutiranja i stanja saobraćaja u mreži:

$$\mathbf{C}(\mathbf{RS}_{(\mathbf{TS})}^{\text{opt}}, \mathbf{TS}) = \min_{\mathbf{RS}} \mathbf{C}(\mathbf{RS}, \mathbf{TS}). \quad (3.2.2)$$

Analogno kao u odeljku 3.1, za neko stanje saobraćaja $\mathbf{TS}$ i rutiranje koje nije optimalno za to stanje saobraćaja, $\mathbf{RS} \neq \mathbf{RS}_{(\mathbf{TS})}^{\text{opt}}$, može se kao pokazatelj kvaliteta rutiranja definisati odnos cene mreže ukoliko je primenjeno rutiranje $\mathbf{RS}$ i cene mreže kada je primenjeno optimalno rutiranje $\mathbf{RS}_{(\mathbf{TS})}^{\text{opt}}$:

$$\text{perf}(\mathbf{RS}, \mathbf{TS}) = \frac{\mathbf{C}(\mathbf{RS}, \mathbf{TS})}{\mathbf{C}(\mathbf{RS}_{(\mathbf{TS})}^{\text{opt}}, \mathbf{TS})}. \quad (3.2.3)$$

Kvalitet rutiranja $\mathbf{RS}$ na skupu mogućih stanja saobraćaja $T = \{\mathbf{TS}\}$ definisan je onda izrazom

$$\text{perf}(\mathbf{RS}, T) = \max_{\mathbf{TS} \in T} \text{perf}(\mathbf{RS}, \mathbf{TS}) = \max_{\mathbf{TS} \in T} \frac{C(\mathbf{RS}, \mathbf{TS})}{C(\mathbf{RS}_{(\mathbf{TS})}^{\text{opt}}, \mathbf{TS})}. \quad (3.2.4)$$

Kao i u slučaju maksimizacije propusne moći mreže, važi da je $\text{perf}(\mathbf{RS}, T) \geq 1$. Što je niža vrednost $\text{perf}(\mathbf{RS}, T)$, to je odabrano rutiranje bolje za dati skup stanja saobraćaja u mreži.

4. LINEARNA OPTIMIZACIJA

Linearna optimizacija i linearno programiranje predstavljaju jedan od često primenjivanih alata za rešavanje raznih problema vezanih za saobraćaj u mrežama. Problem određivanja optimalnog rutiranja može se formulisati kao problem linearne optimizacije.

Problemi linearne optimizacije predstavljaju poseban slučaj optimizacionih problema. U ovom poglavlju biće prvo dat osvrt na probleme matematičke optimizacije u opštem slučaju, a zatim će biti detaljnije obrađena klasa problema linearne optimizacije, kao i metode za rešavanje ovih problema.

4.1. Matematička optimizacija u opštem slučaju

Postupak matematičke optimizacije obično se koristi za rešavanje složenih problema odlučivanja i alokacije resursa, sa kojima se možemo susresti u čitavom nizu oblasti – od analize fizičkih procesa, preko organizacije industrijske proizvodnje, do procesa upravljanja preduzećima. To su problemi u kojima je neophodno odrediti vrednosti većeg broja promenljivih, koje su međusobno na neki način povezane. Pažnja se prilikom analize usmerava na ciljnu funkciju, koja predstavlja vezu između promenljivih. Vrednosti promenljivih moraju se odabrati tako da se ostvari maksimum ili minimum ciljne funkcije. Pri tome vrednosti promenljivih moraju zadovoljavati neka data ograničenja, odnosno promenljive moraju pripadati unapred zadatom skupu. U opštem slučaju ciljna funkcija je realna funkcija promenljivih čije vrednosti treba odrediti.

Formalno se optimizacioni problemi mogu definisati na sledeći način.

Definicija 4.1.1. *Neka su dati skup $D \subseteq \mathbb{R}^n$ i funkcija $f : D \rightarrow \mathbb{R}$. Optimizacioni problem minimizacije ciljne funkcije je onaj u kome je cilj odrediti $\mathbf{x}_0 \in D$, tako da je $f(\mathbf{x}_0) \leq f(\mathbf{x})$ za svako $\mathbf{x} \in D$. Analogno, problem maksimizacije ciljne funkcije f je onaj kod koga je cilj odrediti $\mathbf{x}_0 \in D$ tako da je $f(\mathbf{x}_0) \geq f(\mathbf{x})$ za svako $\mathbf{x} \in D$. Funkcija f naziva se ciljna funkcija, skup D je domen ili oblast pretraživanja, a njegovi elementi predstavljaju dopustiva rešenja (feasible solutions).*

U opštem slučaju, može se dogoditi da ne postoji ekstremna vrednost funkcije f na skupu D . Međutim, ukoliko je skup D konačan, tada postoje i minimum i maksimum ciljane funkcije na tom skupu.

Skup D je obično implicitno zadan skupom ograničenja (datih u vidu jednakosti ili nejednakosti) koja njegovi elementi moraju da zadovolje. Dakle, neka je dat skup od m funkcija ograničenja, $\{g_i : \mathbb{R}^n \rightarrow \mathbb{R}\}_{i=1}^m$ i vektor $\mathbf{b} = (b_1, b_2, \dots, b_m) \in \mathbb{R}^m$. Tada se optimizacioni problem može zapisati na sledeći način:

$$\begin{aligned} \min / \max f(\mathbf{x}) \\ g_i(\mathbf{x}) \leq b_i, i = 1, \dots, m \end{aligned} \quad (4.1.1)$$

Ovako zapisan problem naziva se još i *matematički program*.

Rešavanje optimizacionih problema u opštem slučaju nije uvek moguće. Problem leži u činjenici da ne postoji u svakom slučaju garancija da je pronađena lokalna ekstremna vrednost funkcije f u nekoj tački skupa D istovremeno i globalni ekstremum te funkcije na celom skupu D . Iz tog razloga su naročito interesantni problemi *konveksne optimizacije*, kod kojih možemo sa sigurnošću tvrditi da postoji ekvivalencija između lokalnog i globalnog ekstremuma funkcije na posmatranom domenu.

Definicija 4.1.2. *Optimizacioni problem spada u grupu konveksnih problema onda kada su i ciljna funkcija i ograničenja konveksne funkcije realnih promenljivih, odnosno kada za svako $\mathbf{x}, \mathbf{y} \in D$ i za svako $t \in [0, 1]$ važi*

$$\begin{aligned} f(t\mathbf{x} + (1-t)\mathbf{y}) &\leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \\ g_i(t\mathbf{x} + (1-t)\mathbf{y}) &\leq tg_i(\mathbf{x}) + (1-t)g_i(\mathbf{y}), i = 1, \dots, m. \end{aligned} \quad (4.1.2)$$

4.2. Linearna optimizacija

Problemi linearne optimizacije čine posebnu klasu optimizacionih problema i spadaju u grupu konveksnih problema. Kod ovakvog tipa problema i ciljna funkcija i ograničenja su linearne funkcije ulaznih promenljivih.

Definicija 4.2.1. *Kanonički oblik linearnog programa ima sledeći oblik:*

$$\begin{aligned} \min / \max \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{aligned} \quad (4.2.1)$$

gde je $\mathbf{x}$ vektor promenljivih, $\mathbf{b}$ i $\mathbf{c}$ vektori, a $\mathbf{A}$ matrica realnih koeficijenata.

Ograničenja zadata u modelu linearnog programa definišu neku konveksnu oblast u n -dimenzionom prostoru (politop) u okviru koje se nalaze potencijalna rešenja. Naime, sve tačke n -dimenzionog prostora za koje je ispunjen uslov $\mathbf{A}_k \mathbf{x} \leq \mathbf{b}_k$, gde su $\mathbf{A}_k$ i $\mathbf{b}_k$ k -te vrste matrica $\mathbf{A}$ i $\mathbf{b}$, nalaze se sa jedne strane $(n-1)$ -dimenzione ravni koja je zadata jednakošću $\mathbf{A}_k \mathbf{x} = \mathbf{b}_k$.

Veliki broj problema sa kojima se susrećemo u analizi telekomunikacionih mreža spada u klasu problema linearne optimizacije. Takvi su prvenstveno problemi maksimizacije toka i minimizacije cene puta u mreži, ali i čitav niz drugih problema koji se odgovarajućom formulacijom mogu svesti na neki od ova dva. Najčešće korišćeni metod za rešavanje linearnih programa je simpleks metod.

Za rešavanje problema simpleks metodom neophodno je da on bude postavljen u takozvanoj *standardnoj (proširenoj) formi*. U tom slučaju, ograničenja moraju biti zadata u vidu jednakosti, a ne kao nejednakosti. Zbog toga je potrebno uvesti dodatne, *izravnavajuće promenljive* $\mathbf{x}_s$, tako da važi $\mathbf{A}\mathbf{x} + \mathbf{x}_s = \mathbf{b}$, $\mathbf{x}_s \geq \mathbf{0}$. Rekli smo već da jednakost u ograničenju važi za tačke $(n-1)$ -dimenzionih ravni koje ograničavaju politop sa dopustivim rešenjima. Dakle, uvođenjem izravnavajućih promenljivih, skup potencijalnih rešenja se svodi samo na omotač politopa koji je definisan ograničenjima. Takođe, uobičajeno je da se kod simpleks metoda polazi od minimizacionog problema. Pokazuje se da postoji dualnost između minimizacionih i maksimizacionih problema i da se odgovarajućim transformacijama početni maksimizacioni problem može svesti na minimizacioni, o čemu će biti reči kasnije u ovom poglavlju. Razmatranjem isključivo minimizacionog problema ne gubi se na opštosti. Linearni program sada ima oblik

$$\begin{aligned} \min \mathbf{c}^T \mathbf{x} \\ [\mathbf{A} \quad \mathbf{I}] \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_s \end{bmatrix} &= \mathbf{b} \\ \mathbf{x}, \mathbf{x}_s &\geq \mathbf{0} \end{aligned} \quad (4.2.2)$$

U daljem tekstu smatraćemo, takođe bez gubitka opštosti, da su ograničenja u linearnom programu (4.2.2) data u vidu jednakosti $\mathbf{A}\mathbf{x} = \mathbf{b}$ (što je ekvivalentno tome da pod matricom $\mathbf{A}$ smatramo matricu $[\mathbf{A} \quad \mathbf{I}]$ iz linearnog programa (3.2.2), a pod vektorom $\mathbf{x}$ vektor $[\mathbf{x} \quad \mathbf{x}_s]$ iz istog programa).

Neka posmatrana matrica koeficijenata $\mathbf{A}$ ima m vrsta i n kolona ($n \geq m$). Možemo smatrati da je rang ove matrice jednak m . Naime, ukoliko je rang matrice $\mathbf{A}$ manji od m , to znači da se neka od vrsta matrice $\mathbf{A}$ može izraziti kao linearna kombinacija ostalih vrsta. Neka je to i -ta vrsta; označimo je sa A^i . Možemo pisati da je $A^i = \sum_{j \neq i} \lambda_j A^j$. Tada mora važiti i da je $b_i = \sum_{j \neq i} \lambda_j b_j$, inače sistem ograničenja nema rešenja, domen je prazan skup i optimizacioni problem je nerešiv. Ukoliko je, pak, ovaj uslov ispunjen, to znači da je posmatrano ograničenje suvišno u linearnom programu, tako da se može izostaviti.

Definicija 4.2.2. Za dati skup $B \subseteq \{1, 2, \dots, n\}$, označimo sa $\mathbf{A}_B$ matricu koja sadrži samo one kolone iz $\mathbf{A}$ čiji indeksi pripadaju skupu B . Bazično rešenje linearnog programa (basic feasible solution) je takvo rešenje $\mathbf{x}$ iz skupa dopustivih rešenja D za koje postoji skup $B \subseteq \{1, 2, \dots, n\}$ sa m elemenata, takav da je kvadratna matrica $\mathbf{A}_B$ regularna (odnosno da su kolone čiji su indeksi u B linearno nezavisne) i da je $x_j = 0$ za svako $j \notin B$. Skup B nazivamo bazom, koja odgovara datom bazičnom rešenju.

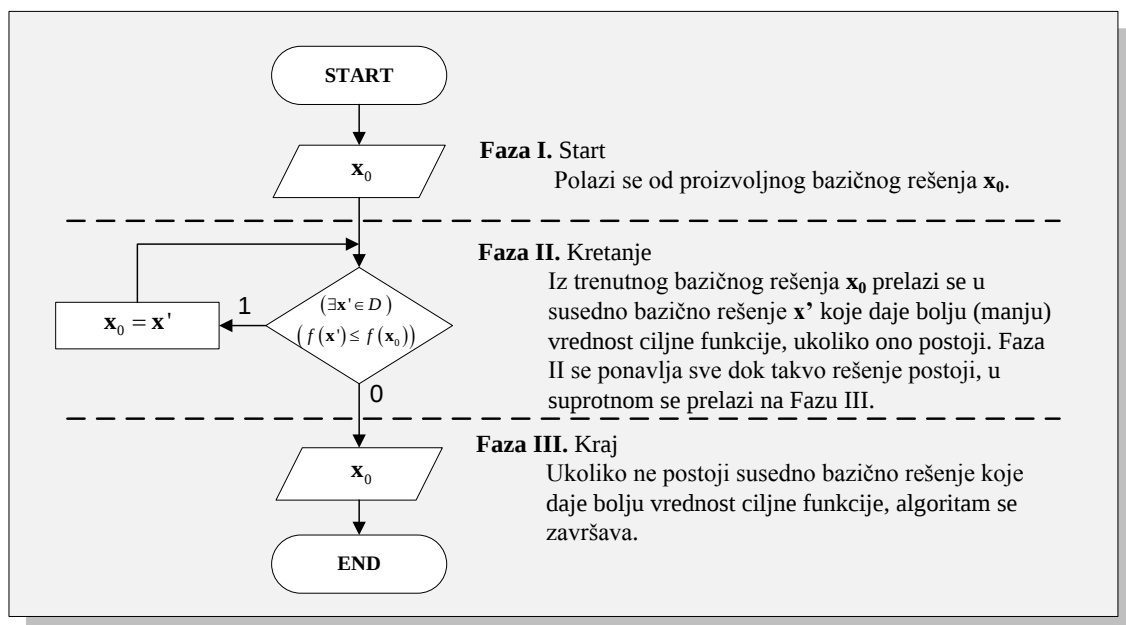
Može se pokazati da su bazična rešenja linearnog programa ekvivalentna temenima politopa koji predstavlja domen, odnosno skup mogućih rešenja [20], [21]. Takođe, ako postoji optimalno rešenje linearnog programa, onda postoji i bazično rešenje koje je optimalno. Jedan od načina na koji se, stoga, može doći do rešenja linearnog programa jeste pretraživanje po svim temenima poliedra koji čini domen funkcije i pronalaženje onog za koje se ostvaruje najbolja vrednost ciljne funkcije.

4.3. Simpleks algoritam

Rekli smo već da se u praksi za rešavanje problema linearne optimizacije najčešće koristi simpleks algoritam. Ovu metodu predložio je 1947. godine Dantzig [22]. Simpleks algoritam polazi od nekog bazičnog rešenja linearnog programa i kreće se kroz temena politopa koji ograničava domen (od bazičnog rešenja do bazičnog rešenja), težeći da u svakom koraku ostvaruje poboljšanje (smanjenje) vrednosti ciljne funkcije. Algoritam se sastoji od tri faze, koje su ilustrovane na Slici 4.3.1.

Simpleks algoritam u praksi se pokazuje zadovoljavajućim za veliki broj problema linearne optimizacije, čak i za veoma velike linearne programe. Eksperimenti pokazuju da se optimalno rešenje za linearni program sa m ograničenja obično pronalazi u nekih

$2m$ do $3m$ koraka u okviru druge faze algoritma (faze kretanja). Međutim, u opštem slučaju simpleks algoritam radi u eksponencijalnom vremenu. Do sada nikome nije pošlo za rukom da osmisli pravilo kretanja po temenima poliedra, tako da se može dokazati da je broj koraka u drugoj fazi algoritma uvek polinomijalna funkcija broja promenljivih i ograničenja u linearnom programu. Za većinu predloženih pravila nađeni su primeri linearnih programa za čije je rešavanje algoritmu neophodan eksponencijalan broj koraka. Bez obzira na ove specijalne slučajeve, u praksi se simpleks metod ipak veoma uspešno koristi.



Slika 4.3.1. Simpleks algoritam

4.3.1. Faza I

U početnoj fazi simpleks algoritma treba odrediti neko bazično rešenje iz koga će započeti iterativni proces. Pronalaženje bazičnog rešenja može u opštem slučaju biti podjednako kompleksno kao i pronalaženje optimalnog rešenja linearnog programa. Jedan od načina da se relativno lako dođe do bazičnog rešenja ćemo sada opisati.

Želimo, dakle, da nađemo bazično rešenje linearnog programa (4.3.1). Pretpostavićemo, bez gubitka opštosti, da su svi elementi vektora $\mathbf{b}$ nenegativni – ukoliko to nisu, možemo pomnožiti odgovarajuće vrste u matrici $\mathbf{A}$ i vektoru $\mathbf{b}$ sa -1 i ova pretpostavka će biti ispunjena, dok se domen linearnog programa neće promeniti (pošto posmatramo samo omotač konveksne oblasti u n – dimenzionom prostoru).

$$\begin{aligned}
& \min \mathbf{c}^T \mathbf{x} \\
& \mathbf{Ax} = \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}.
\end{aligned} \tag{4.3.1}$$

Posmatrajmo sada problem

$$\begin{aligned}
& \min \sum_i y_i \\
& \mathbf{Ax} + \mathbf{y} = \mathbf{b} \\
& \mathbf{x}, \mathbf{y} \geq \mathbf{0}
\end{aligned} \tag{4.3.2}$$

Ukoliko bismo našli rešenje ovog programa tako da je $\sum_i y_i = 0$, imali bismo i rešenje za $\mathbf{Ax} = \mathbf{b}$ koje se nalazi u temenu politopa, odnosno, predstavlja bazično rešenje za početni program. Ovaj problem može se lako rešiti primenom simpleks algoritma na problem (4.3.2), pošto je u njemu početna faza jednostavna: rešenje $\mathbf{x} = \mathbf{0}, \mathbf{y} = \mathbf{b}$ predstavlja bazično rešenje ovog linearnog programa. Kao optimalno rešenje linearnog programa dobićemo vektor $[\mathbf{x}^{opt} \quad \mathbf{y}^{opt}]$. Razlikujemo tri slučaja:

- 1) $\mathbf{y}^{opt} \neq \mathbf{0}$. U ovom slučaju ne postoji bazično rešenje za početni program (3.3.1), odnosno optimizacioni problem nije rešiv.
- 2) $\mathbf{y}^{opt} = \mathbf{0}$. Neka je sa I označen skup indeksa pozitivnih elemenata vektora $\mathbf{x}^{opt}$, $I = \{i \mid x_i^{opt} > 0\}$. Ukoliko je $|I| = m$, $\mathbf{x}^{opt}$ je bazno rešenje početnog linearnog programa, a I je skup indeksa linearno nezavisnih kolona u tom rešenju. Tada je baza koja odgovara rešenju jednaka $B = I$.
- 3) $\mathbf{y}^{opt} = \mathbf{0}$ i $|I| < m$. U ovome slučaju, polazimo od nekog skupa indeksa $J = I$. Proveravamo da li je kolona $k, k \notin J$ matrice $\mathbf{A}$ linearno nezavisna od kolona $j \in J$ i ukoliko jeste, dodajemo k u skup J . Na ovaj način proširujemo skup J sve dokle je to moguće, i dobijamo odgovarajuće bazično rešenje. Baza koja odgovara rešenju je $B = J$.

4.3.2. Faza II

U okviru druge faze, simpleks algoritam se kreće od temena do temena politopa koji ograničava domen linearnog programa. Polazi se iz nekog temena $\mathbf{x}_0 = [x_{10} \quad x_{20} \quad \dots \quad x_{n0}]^T$ kome odgovara baza $B = \{k_1, \dots, k_m\}$. S obzirom na to da su

svi elementi vektora $\mathbf{x}_0$ čiji se indeksi ne nalaze u bazi jednaki nuli, to možemo pisati da je

$$\sum_{i=1}^m x_{k_i 0} A_{k_i} = \mathbf{b} \quad (4.3.3)$$

gde je A_{k_i} kolona matrice $\mathbf{A}$ sa rednim brojem $k_i \in B$. Svaka kolona A_j koja nije u bazi može se izraziti kao linearna kombinacija kolona iz baze

$$A_j = \sum_{i=1}^m p_{k_i j} A_{k_i}, \quad (4.3.4)$$

gde su $p_{k_i j}$ realni koeficijenti.

Baza se proširuje dodavanjem neke kolone A_j . Posmatrajmo linearne kombinacije kolona iz ovako proširene baze, takve da je

$$\sum_{i=1}^m x_{k_i} A_{k_i} + \theta A_j = \mathbf{b}. \quad (4.3.5)$$

Promena parametra θ odgovara kretanju po jednoj od ivica poliedra koja polazi iz temena (baznog rešenja) u kome se trenutno nalazimo. S obzirom na (4.3.3) i (4.3.4), možemo zaključiti da u (4.3.5) mora biti

$$x_{k_i} = x_{k_i 0} - \theta p_{k_i j}. \quad (4.3.6)$$

Promena parametra θ utiče na promenu onih koordinata koje odgovaraju kolonama iz nove baze. Koordinate koje odgovaraju kolonama koje su se već nalazile u bazi, $k_i \in B$, menjaju se prema jednačini (4.3.6). Na osnovu (4.3.5) možemo zaključiti da koordinata koja odgovara koloni koju dodajemo, A_j , takođe zavisi od θ , i da važi da je $x_j = \theta$. Ostale koordinate koje se nisu nalazile u bazi se ne menjaju. One ostaju jednake nuli, kao što su i bile u baznom rešenju $\mathbf{x}_0$. U nekom trenutku bi trebalo da sa povećanjem θ neka od koordinata x_{k_i} postane jednaka nuli, odnosno da stignemo do susednog temena poliedra. Temenu poliedra koji ograničava skup dopustivih rešenja odgovara novo bazično rešenje [20]. U trenutku kada je $x_{k_i} = 0$ važi da je $\theta = x_{k_i 0} / p_{k_i j}$.

Neka se parametar θ kreće u granicama $0 \leq \theta \leq \theta_0^{(j)}$. Pošto uvek mora biti ispunjen uslov nenegativnosti koordinata, $\mathbf{x} \geq 0$, maksimalna vrednost parametra $\theta_0^{(j)}$ jednaka je

$$\theta_0^{(j)} = \min_{\{i | p_{k_i j} > 0\}} \frac{x_{k_i 0}}{p_{k_i j}} = \frac{x_{l 0}}{p_{lj}}. \quad (4.3.7)$$

Gornja granica za parametar θ zavisi od ivice politopa kojom se krećemo. Drugim rečima, zavisi od indeksa kolone A_j koju uključujemo u novu bazu, tj. od j . Postoje tri tipična slučaja za gornju granicu u okviru koje se može menjati parametar θ .

- 1) $\theta_0^{(j)} > 0$. Ovo je uobičajen slučaj, u kome stižemo do novog bazičnog rešenja.
- 2) $\theta_0^{(j)} = 0$. Ovo je takozvani degenerativni korak, u kome je $x_{k_i 0} = 0$, a $p_{k_i j} > 0$.

Ne možemo povećavati vrednost parametra θ u jednačini (4.3.6), a da ne narušimo uslov nenegativnosti koordinate x_{k_i} , ali možemo formirati novu bazu, tako što ćemo kolonu A_{k_i} zameniti kolonom A_j .

- 3) $\theta_0^{(j)} \rightarrow \infty$. Do ovog slučaja dolazi kada su svi koeficijenti $p_{k_i j} \leq 0$. Tada ne možemo koristiti (4.3.7) da odredimo $\theta_0^{(j)}$, jer je skup $\{i | p_{k_i j} > 0\}$ prazan. Iz (4.3.6) vidimo da proizvoljno povećanje parametra θ ne narušava uslov nenegativnosti koordinata, ali nas ni ne dovodi do novog bazičnog rešenja. Problem linearne optimizacije nije ograničen.

Posmatrajmo slučaj kad je $\theta_0^{(j)} > 0$. Definišimo vektor pomeraja $\mathbf{v}$, takav da je

$$v_q = \begin{cases} -p_{k_i j}, & q = k_i, k_i \in B \\ 1, & q = j \\ 0, & \text{inače.} \end{cases} \quad (4.3.8)$$

Sada je novo bazično rešenje jednako

$$\mathbf{x}' = \mathbf{x}_0 + \theta_0^{(j)} \mathbf{v}. \quad (4.3.9)$$

Vrednost ciljne funkcije u ovoj tački jednaka je

$$\mathbf{c}^T \mathbf{x}' = \mathbf{c}^T (\mathbf{x}_0 + \theta_0^{(j)} \mathbf{v}) = \mathbf{c}^T \mathbf{x}_0 + \theta_0^{(j)} \mathbf{c}^T \mathbf{v}. \quad (4.3.10)$$

Pošto je cilj pronaći minimum ciljne funkcije, onda ima smisla usvojiti novo bazično rešenje samo ukoliko je

$$\bar{c}_j = \mathbf{c}^T \mathbf{v} = c_j - \sum_{i=1}^m p_{k_i j} c_{k_i} < 0. \quad (4.3.11)$$

Postoji više pravila na osnovu kojih možemo odabrati kolonu A_j , koju ćemo uvrstiti u novu bazu. Jedan od načina je da se uvek odabira ono $j \notin B$, za koje je ostvarena minimalna vrednost $\bar{c}_j$ odnosno

$$j = \arg \min_{j | \bar{c}_j < 0} \bar{c}_j. \quad (4.3.12)$$

Ovakav pristup naziva se *pravilom najvećeg koeficijenta* i može dovesti do pojave petlji pri obilasku temena poliedra koji ograničava domen linearnog programa.

Drugi način je da se odabere ona kolona za koju se ostvaruje najveće poboljšanje ciljne funkcije

$$j = \arg \max_{j | \bar{c}_j < 0} (\mathbf{c}^T \mathbf{x}_0 - \mathbf{c}^T \mathbf{x}') = \arg \max_{j | \bar{c}_j < 0} (-\theta_0^{(j)} \bar{c}_j). \quad (4.3.13)$$

Ovo pravilo naziva se *pravilom najvećeg poboljšanja* i takođe može dovesti do pojave petlji.

Još jedan način je takozvano *Blandovo pravilo*, po kome se odabira kolona sa najmanjim indeksom za koju se postiže poboljšanje ciljne funkcije,

$$j = \min \{ j | \bar{c}_j < 0 \}. \quad (4.3.14)$$

Ovo pravilo garantuje da neće doći do pojave petlji.

Postoji i mogućnost *slučajnog odabiranja* kolone koja će se uvrstiti u bazu.

U praksi se ipak najčešće koristi *pravilo najstrmije ivice*, koje obično najbrže nalazi optimalno rešenje. U ovom slučaju, u bazu se uvrštava ona kolona za koju je

$$j = \arg \max_{j | \bar{c}_j < 0} \frac{(\mathbf{c}^T \mathbf{x}_0 - \mathbf{c}^T \mathbf{x}')}{\|\mathbf{x}' - \mathbf{x}_0\|}, \quad (4.3.15)$$

gde je $\|\mathbf{x}' - \mathbf{x}_0\| = \theta_0^{(j)} \|\mathbf{v}\|$ dužina ivice poliedra po kojoj se krećemo, tj. euklidsko rastojanje početnog i krajnjeg baznog rešenja.

4.3.3. Faza III

Može se pokazati da ukoliko iz nekog bazičnog rešenja nema mogućnosti za prelazak u novo bazično rešenje u kome se ostvaruje bolja vrednost ciljne funkcije, tada trenutno posmatrano bazično rešenje predstavlja istovremeno i optimalno rešenje linearnog programa. Dokaz nećemo navoditi.

4.4. Dualnost linearnih programa

Prilikom analize simpleks metoda posmatrali smo problem minimizacije linearne funkcije. Međutim, simpleks metod može se primeniti i na rešavanje maksimizacionih problema. Ključno svojstvo linearnih programa, na koje se u tom slučaju oslanjamo, predstavlja svojstvo dualnosti. Posmatrajmo početni linearni program (*primal*), dat sa (3.4.1), kao i dualni linearni program (*dual*), dat sa (4.4.2).

$$\begin{aligned} \max \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{4.4.1}$$

$$\begin{aligned} \min \mathbf{b}^T \mathbf{y} \\ \mathbf{A}^T \mathbf{y} \geq \mathbf{c} \\ \mathbf{y} \geq \mathbf{0}. \end{aligned} \tag{4.4.2}$$

Ovde je $\mathbf{y}$ vektor realnih promenljivih, $\mathbf{b}$ i $\mathbf{c}$ su, kao i do sada, vektori a $\mathbf{A}$ matrica realnih koeficijenata. Prema teoremi o dualnosti, čiji dokaz nećemo ovde navoditi ali se može naći u literaturi [20], [21], moguća su četiri slučaja:

- 1) Ni jedan od ova dva linearna programa nema mogućih rešenja.
- 2) Početni program je neograničen, a dualni nema mogućih rešenja.
- 3) Početni program nema mogućih rešenja, a dualni je neograničen.
- 4) Oba programa imaju bar jedno moguće rešenje. U tom slučaju, oba programa imaju i optimalno rešenje. Ako je $\mathbf{x}^*$ optimalno rešenje početnog programa, a $\mathbf{y}^*$ optimalno rešenje dualnog programa, tada je

$$\mathbf{c}^T \mathbf{x}^* = \mathbf{b}^T \mathbf{y}^*. \tag{4.4.3}$$

Drugim rečima, maksimum početnog ekvivalentan je minimumu dualnog programa.

Znači, problem maksimizacije ciljne funkcije može se rešiti primenom simpleks metode, tako što se program na odgovarajući način preformuliše u dualni minimizacioni problem.

4.5. Primenjeni optimizacioni alat *LP_Solve*

Za optimizaciju je korišćen nekomercijalni (*open source*) alat, *LP_Solve*. Ovaj softver namenjen je prvenstveno rešavanju problema linearne optimizacije, mada može raditi i sa modelima u kojima promenljive uzimaju samo celobrojne vrednosti. Softver je moguće preuzeti kao biblioteku funkcija koje se mogu pozivati iz programskog jezika C++, Visual Basic, Delphi, Java, a može se integrisati i u Excel, Matlab, itd. Za rešavanje problema linearne optimizacije softver koristi simpleks algoritam, dok kod kombinovanog ili celobrojnog programiranja koristi metod grananja i ograničavanja (*branch and bound*).

Model linearnog programa moguće je zadati u okviru fajla određenog formata, koji *LP_Solve* može preuzeti, a postoji i mogućnost formiranja modela korišćenjem funkcija koje se pozivaju iz odabranog programskog jezika. U našem slučaju korišćena je mogućnost pozivanja funkcija iz *LP_Solve* biblioteke, model je formiran iz programskog jezika C++, izdavana naredba za njegovo rešavanje, a zatim su rezultati preuzimani odgovarajućim komandama.

Više detalja o ovom alatu, kao i pregled osnovnih funkcija koje se koriste za formiranje modela linearnog programa i njegovo rešavanje dati su u Prilogu A.

5. NEBLOKIRAJUĆE MREŽE

Problem kojim se bavimo u ovom radu je linearna optimizacija neblokirajućih mreža. U ovom poglavlju definisaćemo neblokirajuće svojstvo mreža sa komutacijom veze i komutacijom paketa. Zatim ćemo navesti uslove koje različite mreže treba da zadovolje, da bi se mogle smatrati neblokirajućim.

Pojam neblokirajućih mreža potiče još iz doba razvoja fiksne telefonije, i uveden je i detaljno analiziran 1953. godine, u radu Čarlsa Klosa [23]. Prema Klosovoj definiciji, neblokirajućom se smatra ona mreža u kojoj je moguće u svakom trenutku uspostaviti vezu između neka dva do tada neaktivna korisnika, bez obzira na broj veza koje su već uspostavljene u sistemu. Kasnije je definicija neblokirajućih mreža prilagođena i za mreže sa komutacijom paketa [24]- [26], tako da se pod neblokirajućim paketskim mrežama smatraju se one mreže u kojima je uvek moguće uspostaviti sesiju između dva korisnika, bez obzira na opterećenje koje u mreži prouzrokuju ostali korisnici, a pod uslovom da korisnici koji stupaju u vezu imaju dovoljno raspoloživog kapaciteta na svom ulazu, tj. izlazu.

Porast broja telefonskih pretplatnika tokom prve polovine dvadesetog veka doveo je do potrebe za efikasnijim i jeftinijim načinom povezivanja korisnika od onog koji je do tada korišćen. Povezivanje svakog ulaza centrale direktno sa svim izlazima centrale garantovalo je neblokirajuće svojstvo mreže, ali su komutaciona polja takvih centrala postala previše skupa. Sa druge strane, ograničavanje kapaciteta telefonske centrale omogućavalo je ostvarivanje ušteda pri postavljanju i održavanju sistema, ali je prouzrokovalo probleme vezane za kvalitet pružene usluge. Naime, centrale manjeg kapaciteta imale su izvesnu verovatnoću blokade. U slučaju opterećenja većeg od projektovanog, dolazilo bi odbijanja zahteva za uspostavom veze između novog para pretplatnika, sve dok se ne prekine neka od već uspostavljenih veza u centrali.

Klov rad pokazao je da je moguće ostvariti neblokirajuće svojstvo komutacionog polja centrale i bez povezivanja svakog ulaza u centralu direktno sa svim izlazima iz centrale. Rezultati njegovog rada imali su veliki uticaj na razvoj neblokirajućih mreža sa komutacijom veze, a kasnije su primenjeni i na razvoj neblokirajućih paketskih mreža. U ovom poglavlju prvo ćemo se osvrnuti na originalne

rezultate Klosovog rada vezanog za dizajn neblokirajućeg komutacionog polja za telefonske centrale, tj. mreže sa komutacijom veze. Zatim će u sledećem odeljku biti reči o modifikacijama i primenama Klosovih mreža, za dizajn komutacionog polja rutera u paketskim mrežama.

5.1. Neblokirajuće mreže sa komutacijom veze – Klosove mreže

Kao što smo već pomenuli, sa porastom broja telefonskih pretplatnika četrdesetih godina prošlog veka, pojavila se potreba za razvojem ekonomičnog komutacionog polja u telefonskim centralama. Ovakvo polje trebalo je da omogući što jeftinije povezivanje većeg broja korisnika, a da zadrži neblokirajuća svojstva kakva poseduje mreža u kojoj su svi korisnici međusobno povezani jedni sa drugima.

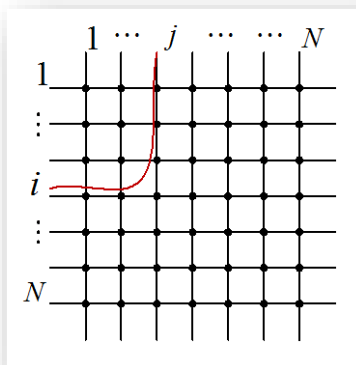
U ovom odeljku najpre ćemo formalno definisati pojam neblokirajuće mreže sa komutacijom veze. Zatim ćemo se osvrnuti na rezultate originalnog Klosovog rada [23], kojima se definiše arhitektura neblokirajuće mreže, i pokazuje ušteda koja se može ostvariti ovakvom arhitekturom.

Definicija 5.1.1 *Pod neblokirajućom mrežom sa komutacijom veze podrazumeva se ona u kojoj je uvek moguće uspostaviti vezu između dva neaktivna korisnika, bez obzira na to koliko je u datom trenutku veza u sistemu već uspostavljeno.*

Definicija 5.1.2 *Striktno neblokirajuća mreža je neblokirajuća mreža u kojoj nije potrebno vršiti rekonfigurisanje postojećih veza u sistemu, kako bi se nova veza ostvarila.*

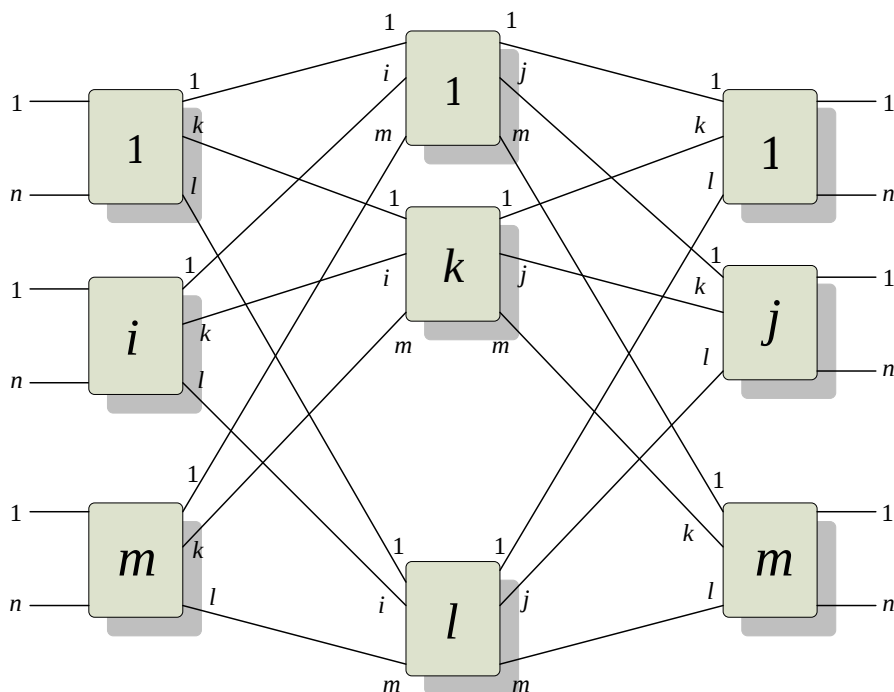
Najjednostavniji način za međusobno povezivanje N korisnika u neblokirajuću mrežu je formiranje kvadratnog komutacionog polja (krosbara), kao na slici 5.1.1. Komutaciono polje centrale sa N ulaza i N izlaza ima N^2 čvorova, u kojima se nalaze prekidači. Veza između korisnika na ulazu i i korisnika na izlazu j uspostavlja se zatvaranjem prekidača na preseku vrste i i kolone j u komutacionom polju centrale.

U ovakvom polju, moguće je uspostaviti vezu između svakog ulaza i svakog izlaza. Korisnik sa ulaza i uvek može uspostaviti vezu sa korisnikom sa izlaza j , ukoliko je korisnik j slobodan, bez obzira na to koliko je ostalih korisnika aktivno i koliko je veza trenutno uspostavljeno. Međutim, ovakvo komutaciono polje u Klosovo doba nije bilo ekonomično, s obzirom na veliki broj čvorova – cena komutacionog polja u to vreme bila je proporcionalna broju čvorova u njemu.



Slika 5.1.1. Kvadratno komutaciono polje (krosbar)

Klosova analiza u [23] pokazala je da se neblokirajuće svojstvo može postići i sa manje od N^2 čvorova u komutacionom polju. Arhitektura koju Klos predlaže sastoji se od nekoliko nizova (stepeni) komutacionih polja, kojima je tako odabran broj ulaza i izlaza, da mreža bude neblokirajuća. U tekstu koji sledi, najpre ćemo analizirati Klosovu mrežu sa tri stepena, i izvesti uslov neblokiranja za nju. Zatim ćemo analizu proširiti na mrežu sa pet stepeni, kao i opšti slučaj mreže sa neparnim brojem stepeni.

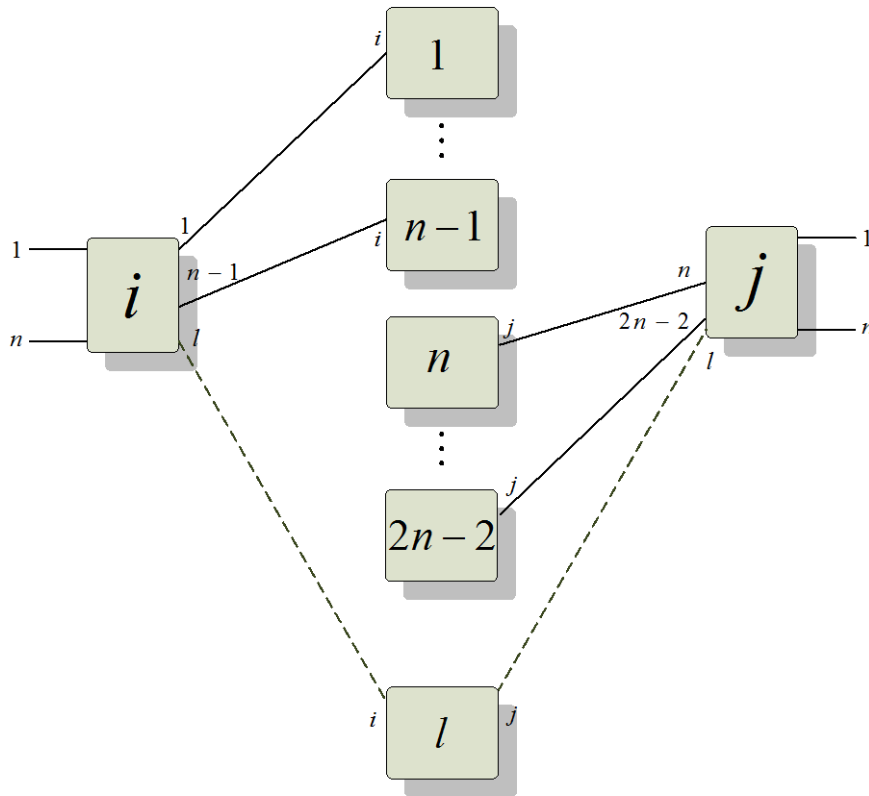


Slika 5.1.2. Klosova mreža sa tri stepena

5.1.1. Trostepena Klosova mreža

Najjednostavniji oblik Klosove mreže ima tri stepena, i prikazan je na slici 5.1.2. Koristićemo oznaku $K(d)$ za Klosovu mrežu sa d stepeni, te u skladu sa tom notacijom označavamo mrežu sa tri stepena sa $K(3)$.

Signal svakog od $N = m \cdot n$ korisnika vodi se na neku od raspoloživih linija ulaznog niza komutacionih polja (levo na slici 5.1.2). Komutaciona polja u ovom nizu su dimenzije $n \times l$, i mora ih biti m u ulaznom nizu, kako bi svaki korisnik imao liniju na raspolaganju. Takođe, u izlaznom nizu komutacionih polja (desno na slici 5.1.2) imamo m komutacionih polja dimenzije $n \times l$. Središnji niz komutacionih polja čine kvadratna polja dimenzije $m \times m$. Broj ovih polja određuje se takozvanom Klosovom teoremom.



Slika 5.1.3. Ilustracija Klosove teoreme

Teorema 5.1.1. (Klosova teorema) Klosova mreža sa tri stepena je neblokirajuća ukoliko za broj komutacionih polja u središnjem nizu l važi $l \geq 2n-1$, gde je n dimenzija ulaznog komutacionog polja.

Dokaz: Posmatrajmo polje i iz ulaznog niza komutacionih polja, i polje j iz izlaznog niza komutacionih polja. Neka je u posmatranom trenutku slobodan samo jedan

korisnik vezan na polje i i samo jedan korisnik vezan na polje j , dok su sve ostale linije ova dva polja zauzete. Pretpostavimo sada da upravo ova dva slobodna korisnika žele da uspostave vezu.

U najgorem slučaju, ni jedan od $n-1$ aktivnih korisnika vezanih na polje i nije uspostavio vezu sa nekim od $n-1$ korisnika vezanih na polje j . Pri tome, u najgorem slučaju svaka od uspostavljenih veza iz polja i i j koristi različito komutaciono polje iz središnjeg niza komutacionih polja. Stoga su već uspostavljene veze u sistemu angažovale ukupno $2 \cdot (n-1)$ komutacionih polja iz središnjeg niza – slika 5.1.3. Da bi se zahtevana veza između dva preostala korisnika uspostavila, neophodno je postojanje još barem jednog komutacionog polja u središnjem nizu. Dakle:

$$l \geq 2(n-1)+1 = 2n-1. \quad (5.1.1)$$

Ovim je dokaz završen. ■

Pogledajmo sada kolika je cena Klosove mreže sa tri stepena. Pomenuli smo već da je cena mreže proporcionalna broju čvorova, tj. prekidača u njoj. Naime, u vreme u kome je Klosov rad nastao, prekidači u telefonskim centralama bili su elektromehanički i njihova cena predstavljala je veliki udeo u ukupnoj ceni sistema.

Kao što smo već rekli, u kvadratnom komutacionom polju koje treba da servisira N korisnika, broj čvorova iznosi

$$C(1, N) = N^2 \quad (5.1.2)$$

Ukupan broj čvorova u Klosovoj mreži sa tri stepena, koja servisira $N = n^2$ korisnika, iznosi

$$C(3, N) = 3N(2N^{1/2} - 1) = 6N^{3/2} - 3N. \quad (5.1.3)$$

Za $N \geq 36$ važi da je $C(3, N) < C(1, N)$. Dakle, sve centrale sa više od 36 korisnika mogu imati koristi od primene trostepene Klosove arhitekture komutacionog polja.

5.1.2. Petostepena Klosova mreža

Analiza petostepene Klosove mreže može se svesti na analizu trostepene Klosove mreže, date na slici 5.1.4. Pri tome, ulazni i izlazni niz modula čine komutaciona polja dimenzije $n \times l$, označena sa $K(1)$ na slici 5.1.4. Središnji niz čine trostepene Klosove mreže sa m ulaza i izlaza, označene sa $K(3)$. Prema (5.1.1), da bi petostepena mreža bila neblokirajuća, u središnjem nizu mora biti $l \geq 2n-1$ trostepenih Klosovih mreža.

Posmatrajmo petostepenu mrežu u kojoj je $m = N^{2/3}$ i $n = N^{1/3}$. Da bi mreža bila neblokirajuća, mora biti $l = 2N^{1/3} - 1$. Znači, ovakva mreža ima $m = N^{2/3}$ komutacionih polja dimenzije $N^{1/3} \times (2N^{1/3} - 1)$ u ulaznom i izlaznom nizu, dok središnji niz čini $l = 2N^{1/3} - 1$ trostepenih Klosovih mreža sa po $N^{2/3}$ ulaza. Ukupan broj čvorova u ulaznom i izlaznom nizu iznosi

$$C'(5, N) = 2N^{2/3}N^{1/3}(2N^{1/3} - 1) = 2N(2N^{1/3} - 1). \quad (5.1.4)$$

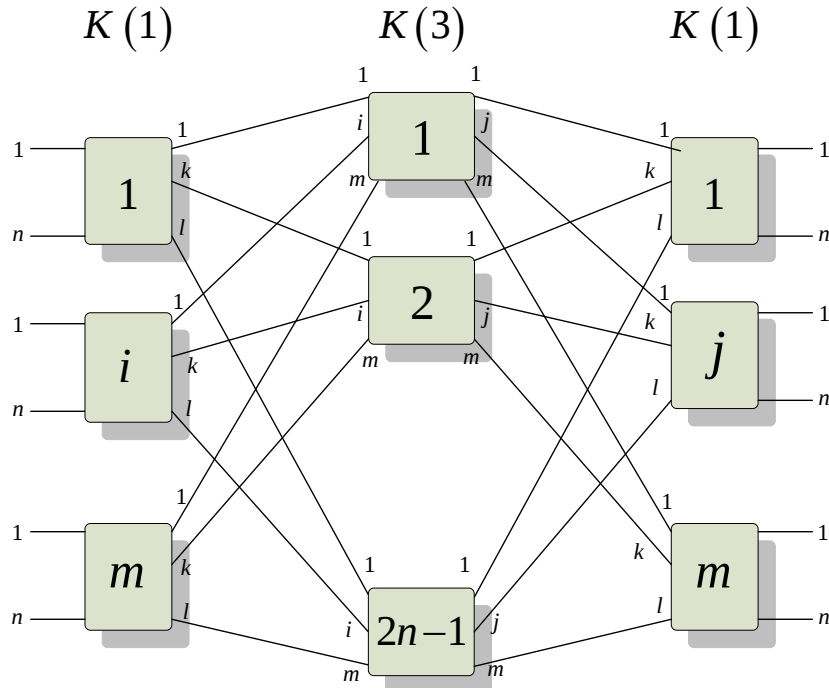
U skladu sa jednačinom (5.1.3), broj čvorova u svakoj od trostepenih Klosovih mreža u središnjem nizu dat je sa $C(3, N^{2/3}) = 3N^{2/3}(2N^{1/3} - 1)$. Ovakvih polja ima $2N^{1/3} - 1$, tako da je broj čvorova u središnjem nizu trostepenih Klosovih mreža jednak

$$C''(5, N) = (2N^{1/3} - 1)3N^{2/3}(2N^{1/3} - 1) = (2N^{1/3} - 1)^2 3N^{2/3}. \quad (5.1.5)$$

Stoga je ukupan broj čvorova u petostepenoj Klosovoj mreži jednak

$$C(5, N) = (2N^{1/3} - 1)^2 3N^{2/3} + (2N^{1/3} - 1)2N = 16N^{4/3} - 14N + 3N^{2/3}. \quad (5.1.6)$$

Za $N \geq 729$ važi da je $C(5, N) < C(3, N)$, pa je za komutaciona polja sa više od 729 ulaza i izlaza bolje primeniti petostepenu Klosovu mrežu u odnosu na Klosovu mrežu sa tri stepena.



Slika 5.1.4. Petostepena Klosova mreža

5.1.3. Opšti slučaj Klosove mreže sa neparnim brojem stepeni

Posmatrajmo opšti slučaj Klosove mreže $K(2t+1)$. Ova mreža ima $2t+1$ stepeni, i opslužuje N korisnika. Analogno načinu posmatranja petostepene Klosove mreže, možemo smatrati da je ova mreža ima niz običnih komutacionih polja $K(1)$ na ulazu i izlazu, dok je u sredini niz Klosovih mreža nižeg stepena $K(2t-1)$. Ako su u ulaznom i izlaznom nizu komutaciona polja dimenzije $n \times l$, tada središnji niz mora sadržati $l \geq 2n-1$ Klosovih mreža $K(2t-1)$.

Svako od komutacionih polja ulaznog niza ima $N^{1/(t+1)}$ ulaza, i ovakvih polja ima ukupno $N^{t/(t+1)}$. Usled simetrije, izlazni niz čini $N^{t/(t+1)}$ polja sa po $N^{1/(t+1)}$ izlaza. Klosovih mreža u središnjem stepenu ima $2N^{1/t} - 1$.

Broj čvorova u svakom od komutacionih polja ulaznog i izlaznog niza je $N^{1/(t+1)}(2N^{1/t} - 1)$, tako da je ukupan broj čvorova u ova dva niza dat sa

$$C'(2t+1, N) = 2N^{t/(t+1)}N^{1/(t+1)}(2N^{1/(t+1)} - 1) = 2N(2N^{1/(t+1)} - 1). \quad (5.1.7)$$

Središnji stepen čini $2N^{1/t} - 1$ Klosovih mreža stepena $d = 2t-1$, sa $N^{t/(t+1)}$ ulaza i izlaza. Svaka od ovih mreža ima $C(2t-1, N^{t/(t+1)})$ čvorova, pa je ukupan broj čvorova u središnjem nizu dat sa

$$C''(2t+1, N) = (2N^{1/t} - 1)C(2t-1, N^{t/(t+1)}) \quad (5.1.8)$$

Rešavanjem rekurzije (5.1.8) može se pokazati [23] da je

$$C(2t+1, N) = \frac{N^{2/(t+1)}(2N^{1/(t+1)} - 1)}{N^{1/(t+1)} - 1} \left[(5N^{1/(t+1)} - 3)(2N^{1/(t+1)} - 1)^{t-1} - 2N^{t/(t+1)} \right] \quad (5.1.9)$$

5.2. Neblokirajuće paketske mreže

Originalni Klosov rad bavio se dizajnom neblokirajućeg komutacionog polja u mrežama sa komutacijom veze. U to vreme, u telefonskim centralama upotrebljavani su elektromehanički prekidači, čija je cena bila visoka, te je cena centrale zavisila prvenstveno od broja prekidača u njoj. Zato je bilo neophodno voditi računa o minimizaciji broja čvorova u komutacionom polju, tako da se po što manjoj ceni usluga garantuje što većem broju korisnika. Kasniji razvoj tehnologije omogućio je prelazak sa elektromehaničkih na digitalne prekidače, što je dovelo do smanjenja cene centrala. Stoga su Klosovi rezultati vezani za minimizaciju broja čvorova u komutacionoj mreži izgubili

na značaju, ali je predložena arhitektura našla primenu i u paketskim mrežama, zahvaljujući svom neblokirajućem svojstvu [24]- [26].

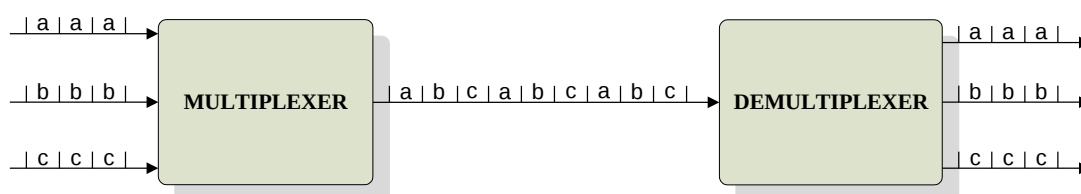
Suštinsku razliku pri definisanju problema neblokirajućih mreža u paketskim mrežama predstavlja činjenica da različiti servisi u paketskim mrežama zahtevaju različite brzine (propusne opsege), kao i da intenzitet saobraćaja među čvorovima varira tokom vremena. Stoga se prilikom analize sviča moraju u obzir uzeti i težine dodeljene vezama između pojedinih portova. Ovim težinama se opisuje kapacitet neophodan za ostvarivanje veze između različitih parova portova. U skladu sa tim, menja se i definicija neblokirajućeg svojstva mreža.

Definicija 5.2.1. *Neblokirajuće paketske mreže su one mreže u kojima je uvek moguće uspostaviti sesiju između dva korisnika koji imaju dovoljno raspoloživog kapaciteta, bez obzira na opterećenje koje u mreži prouzrokuju ostali korisnici (ukoliko se oni pridržavaju ograničenja saobraćaja, i ne ostvaruju veći saobraćaj od onog koji im je dozvoljen).*

U tekstu koji sledi, biće reči o uslovima neblokiranja za Klosove mreže u mrežama sa komutacijom paketa. Rezultati o kojima je reč odnose se pre svega na sisteme u kojima je dužina paketa fiksna, poput ATM svičeva. Teoretski, ovi rezultati važe i u opštem slučaju komutacije paketa, za pakete proizvoljne dužine, pod uslovom da se paketi pri prolasku kroz svič podele na ćelije fiksne dužine

5.2.1. Uslov neblokiranja u slučaju mreža sa vremenskim multipleksiranjem

Kod mreža sa vremenskim multipleksiranjem (TDM), prenos preko linka vrši se sinhrono, u vremenskim intervalima tačno određene dužine. Pri tome, osnovni vremenski interval podeljen je na k podintervala, a svaki od podintervala dodeljen je jednoj vezi. Na ovaj način istovremeno se može ostvariti veza između k korisnika, bez potrebe da se svakom od korisnika dodeli odvojeni link, čime se postiže ušteda kapaciteta. Na slici 5.2.1. prikazano je vremensko multipleksiranje tri kanala.



Slika 5.2.1. Vremensko multipleksiranje

Ukoliko određena veza zahteva veći propusni opseg od opsega osnovnog kanala u vremenskom multipleksu, moguće joj je dodeliti više vremenskih slotova.

Teorema 5.2.1. *Posmatrajmo paketski svič sa Klosovom strukturom, prikazanom na slici 5.1.2. Neka se ovaj svič koristi u mreži sa vremenskim multipleksiranjem, u kojoj je svakoj vezi potrebno maksimalno B osnovnih kanala. Prema [24], uslov neblokiranja sviča dat je sa*

$$l \geq 2 \left\lfloor \frac{nk - B}{k - B + 1} \right\rfloor + 1, \quad (5.2.1)$$

gde je n broj ulaznih linkova na ulaznim modulima Klosove strukture, l broj modula u središnjem nizu Klosove strukture, a k kapacitet pojedinačnih linkova u sviču.

Dokaz: Da bi se uspostavila nova veza koja zahteva B vremenskih slotova, modul u ulaznom nizu preko koga se ostvaruje ova veza mora imati barem jedan link na kome je opterećenje manje od $k - B$. Dakle, ukupno opterećenje linkova modula u ulaznom nizu mora biti manje od $nk - B$. Pri tome, nova veza se ne može uspostaviti preko linka koji povezuje modul iz ulaznog niza i modul iz središnjeg niza, ukoliko je opterećenje tog linka veće od $k - B + 1$. Stoga modula u središnjem nizu, preko kojih se veza ne može uspostaviti, ima maksimalno $\lfloor (nk - B) / (k - B + 1) \rfloor$. Analogna analiza može se izvršiti i za vezu središnjeg modula sa izlaznim modulom, tako da važi (5.2.1). ■

5.2.2. Uslov neblokiranja u asinhronim paketskim mrežama

U prethodnom odeljku razmatrali smo uslov neblokiranja u mrežama sa vremenskim multipleksiranjem. Međutim, u ATM mrežama propusni opseg dodeljen nekoj vezi ne mora biti umnožak kapaciteta osnovnog kanala. Pored toga, ATM svičevi obično imaju ubrzanje na putanjama paketa unutar sviča, tako da se prenos paketa kroz komutaciono polje vrši većom brzinom nego što je brzina spoljašnjeg linka. U ovom odeljku osvrnućemo se na rezultate analize neblokirajućih svojstava ATM svičeva prezentovane u [26]. Prvo ćemo definisati uslov neblokiranja u opštem slučaju, a zatim analizirati slučajeve sa ubrzanjem unutrašnjih linkova sviča, odnosno slučaj sa ograničenjem propusnog opsega veze koja se servisira.

i) Uslov neblokiranja u opštem slučaju

Posmatrajmo vezu između korisnika kao uređenu trojku (x, y, w) , gde je x izvor, y odredište paketa, a w težina koja odgovara propusnom opsegu koji je neophodan za

uspostavljanje veze. Smatramo da je taj propusni opseg normalizovan na raspoloživi propusni opseg linka, tj, da je $0 \leq w \leq 1$, kao i da su svi linkovi u mreži istog kapaciteta. Putanja r kojom se ostvaruje veza (x, y, w) naziva se rutom težine w .

Teorema 5.2.2. *Klosova mreža sa l modula u središnjem nizu je neblokirajuća u odnosu na vezu (x, y, w) ako važi*

$$l \geq 2 \left\lceil \frac{n-1}{1-w} \right\rceil + 1 \quad (5.2.2)$$

Dokaz: Posmatrajmo modul i u ulaznom nizu za koji je vezan korisnik x . U najgorem slučaju, svih ostalih $n-1$ ulaza ovog modula su potpuno zauzeti, a postojeće opterećenje na jedinom linku koji nije potpuno zauzet iznosi $1-w$. Dakle, ukupno opterećenje linkova koji povezuju ulazni modul i i module iz središnjeg stepena Klosove mreže iznosi

$$n-1+1-w = n-w. \quad (5.2.3)$$

Nova veza se ne može uspostaviti preko linkova ka središnjem modulu koji već imaju opterećenje veće od $1-w$. Ovakvih linkova ima maksimalno

$$\left\lceil \frac{n-w}{1-w} \right\rceil - 1, \quad (5.2.4)$$

i to je maksimalan broj modula iz središnjeg niza koji nisu dostupni. Analogno razmatranje može se izvesti i za vezu između izlaznog modula j na koji je vezan korisnik y , tako da je maksimalan broj modula preko kojih se veza ne može uspostaviti u najgorem slučaju jednak

$$2 \left\lceil \frac{n-w}{1-w} \right\rceil - 1 = 2 \left\lceil \frac{n-1}{1-w} \right\rceil + 1. \quad (5.2.5)$$

Stoga je neophodan broj modula u središnjem stepenu dat u (5.2.2). ■

Može se uočiti da u (5.2.2) $l \rightarrow \infty$ onda kada $w \rightarrow 1$. Dakle, postoje situacije (zahtevi za uspostavom veze) pri kojima je praktično nemoguće ostvariti neblokirajuće svojstvo sviča. Jedan od načina da se smanji broj neophodnih modula u središnjem nizu Klosove mreže predstavlja povećanje kapaciteta unutrašnjih linkova sviča (takozvano ubrzanje). Drugi način je ograničavanje zahteva za uspostavom veze na interval $w \in [b, B]$. Oba ova slučaja ćemo analizirati u tekstu koji sledi.

ii) *Uslov neblokiranja u slučaju sviča sa ubrzanjem unutrašnjih linkova*

Teorema 5.2.3. *Klosova mreža sa faktorom ubrzanja unutrašnjih linkova S je neblokirajuća u odnosu na vezu (x, y, w) ako važi*

$$l \geq 2 \left\lceil \frac{n-w}{S-w} \right\rceil - 1. \quad (5.2.6)$$

Dokaz: U odnosu na unutrašnje linkove, veza (x, y, w) može se posmatrati kao veza $(x, y, w/S)$ u sviču bez ubrzanja. Kao i pri dokazu Teoreme 5.2.1, možemo zaključiti da je u najgorem slučaju suma opterećenje svih linkova koji povezuju ulazni modul sa središnjim modulima jednako

$$\frac{n-1}{S} + \frac{1-w}{S} = \frac{n-w}{S}. \quad (5.2.7)$$

Takođe, broj linkova koji povezuju modul iz ulaznog niza sa modulima iz središnjeg niza, a koji imaju opterećenje veće od $1-w/S$ je manji od

$$\left\lceil \frac{n-w}{S} \right\rceil / \left(1 - \frac{w}{S}\right) = \left\lceil \frac{n-w}{S-w} \right\rceil, \quad (5.2.8)$$

tako da je maksimalan broj modula iz središnjeg niza preko kojih se ne može ostvariti veza jednak

$$\left\lceil \frac{n-w}{S-w} \right\rceil - 1. \quad (5.2.9)$$

Analogno se analiza može sprovesti za vezu modula iz središnjeg niza i modula iz izlaznog niza, te dolazimo do zaključka da je za ostvarivanje neblokirajućeg svojstva neophodno da u središnjem modulu bude barem

$$2 \left(\left\lceil \frac{n-w}{S-w} \right\rceil - 1 \right) + 1 = 2 \left\lceil \frac{n-w}{S-w} \right\rceil - 1. \quad (5.2.10)$$

■

iii) *Uslov neblokiranja u slučaju ograničenog kapaciteta veze*

Kao što je već spomenuto, jedan od načina da se smanji broj modula u središnjem nizu Klosove mreže jeste i ograničenje propusnog opsega veze koja se u sistemu može uspostaviti. Pretpostavimo da je veza ograničena na opseg $w \in [b, B]$. Teorema koja sledi preuzeta je iz [26].

Teorema 5.2.4. *Posmatrajmo vezu zadatu uređenom trojkom (x, y, w) , takvom da je propusni opseg neophodan za uspostavljanje veze ograničen sa $w \in [b, B]$, gde je*

$0 \leq b, B < 1$. Tada je dovoljan uslov da Klosova mreža bude neblokirajuća u odnosu na vezu (x, y, w) sledeća nejednakost

$$l \geq 2 \left\lceil \frac{n-1}{1-B} \right\rceil + 1. \quad (5.2.11)$$

Ako je $b > 1-B$, potreban i dovoljan uslov da Klosova mreža bude neblokirajuća u odnosu na vezu (x, y, w) je dat nejednakošću

$$l \geq 2(n-1) \lfloor 1/b \rfloor + 1. \quad (5.2.12)$$

Dokaz: Postupak dokazivanja tvrdnje (5.2.11) je analogan dokazu teoreme 5.2.2. U najgorem slučaju, $w = B$ i (5.2.5) postaje (5.2.11). Pošto je $B < 1$, to je vrednost l uvek konačna.

Sada ćemo dokazati tvrdnju (5.2.12). U najgorem slučaju je neophodno uspostaviti novu vezu težine $w = B$. Pošto je $b > 1-B$, svi ulazni linkovi preko kojih su uspostavljanje čak i veze sa najmanjom težinom b ne mogu se koristiti za uspostavu nove veze $w = B$. Stoga, jedan link modula iz ulaznog niza mora biti slobodan u potpunosti, dok su u najgorem slučaju ostali potpuno zauzeti. Maksimalan broj veza koje se mogu uspostaviti preko jednog od zauzetih ulaznih linkova jednak je $\lfloor 1/b \rfloor$. Svaka od njih može opteretiti različit link iz skupa linkova koji povezuju ulazni sa središnjim modulom. Takav link onda ne može koristiti novoprispela veza težine $w = B$. Ukupno, ovakvih linkova ima maksimalno $(n-1) \lfloor 1/b \rfloor$. Analognu analizu onda možemo sprovesti i za vezu središnjeg sa izlaznim nizom modula. Ukupan broj modula u središnjem nizu koji se ne mogu koristiti za uspostavu nove paketske veze u najgorem slučaju iznosi $2 \cdot (n-1) \lfloor 1/b \rfloor$. Da bi Klosova mreža bila neblokirajuća, mora postojati bar još jedan modul u središnjem nizu, tj. mora važiti (5.2.12).

6. BALANSIRANJE SAOBRAĆAJA U REGULARNIM TOPOLOGIJAMA

Balansiranje saobraćaja u telekomunikacionoj mreži jedan je od slučajeva u kojima se primenjuje tehnika balansiranja opterećenja. U ovom poglavlju prvo ćemo definisati šta se podrazumeva pod balansiranjem opterećenja i balansiranjem saobraćaja, a zatim ćemo se osvrnuti na rezultate primene balansiranja saobraćaja u dizajnu paketskih svičeva velikog kapaciteta.

Definicija 6.0.1 *Balansiranje opterećenja je metodologija kojom se ukupno opterećenje nekog sistema raspoređuje između više različitih resursa u sistemu. Zadatak tih resursa je procesiranje datog opterećenja.*

U zavisnosti od slučaja o kome je reč, opterećenje sistema može predstavljati saobraćaj koji treba preneti, podaci koje treba obraditi, podaci koje treba skladištiti, i sl. U skladu sa opterećenjem koje se procesira, resursi mogu biti ruteri i linkovi u paketskoj mreži, kanali u sistemima za bežični prenos podataka, različiti računari i procesorska jezgra koji vrše obradu podataka, diskovi na kojima se skladište podaci, itd.

Balansiranje opterećenja može se primeniti u cilju boljeg iskorišćenja postojećih resursa, maksimizacije ostvarenog saobraćaja, minimizacije vremena obrade podataka ili radi izbegavanja preopterećenja pojedinih delova sistema. Takođe, sistemi u kojima se primenjuje balansiranje opterećenja su pouzdaniji, pošto postoji redundantnost resursa.

Predmet analize u ovome radu jesu paketske mreže sa balansiranjem saobraćaja. U ovakvim mrežama, opterećenje sistema predstavlja saobraćaj koji treba preneti. Resursi koji služe za prenos podataka su linkovi i čvorovi u mreži.

Balansiranje saobraćaja može se primeniti kako pri određivanju putanje paketa od izvora do odredišta u mreži, tako i pri određivanju putanje paketa unutar komutacionog paketskog sviča. U ovom poglavlju osvrnućemo se na rezultate primene balansiranja saobraćaja pri dizajnu komutacionog polja paketskih svičeva. Naime, pokazuje se da se primenom balansiranja saobraćaja može ostvariti neblokirajuće svojstvo svičeva sa Klosovom arhitekturom, bez potrebe za uvođenjem unutrašnjeg ubrzanja linkova ili dodavanjem modula u središnji niz [27]. Takođe, primena balansiranja saobraćaja

omogućava postizanje garantovanih vrednosti kašnjenja paketa prilikom obrade u sviču. O ovome će biti reči u prvom odeljku ovog poglavlja. Zatim ćemo se u drugom odeljku osvrnuti na rezultate analize iz [28], koja pokazuje da se u opštem slučaju, nezavisno od unutrašnje arhitekture sviča, neblokirajuće svojstvo može garantovati ukoliko se koristi balansiranje saobraćaja i unutrašnje ubrzanje linkova $S = 2$.

Primena balansiranja saobraćaja pri rutiranju paketa u višelokacijskim mrežama neregularnih topologija u kojima se koristi rutiranje po najkraćoj putanji (kakve su realne računarske mreže) biće predmet analize u Poglavlju 7.

6.1. Balansiranje saobraćaja u višestepenim svičevima sa Klosovim mrežama

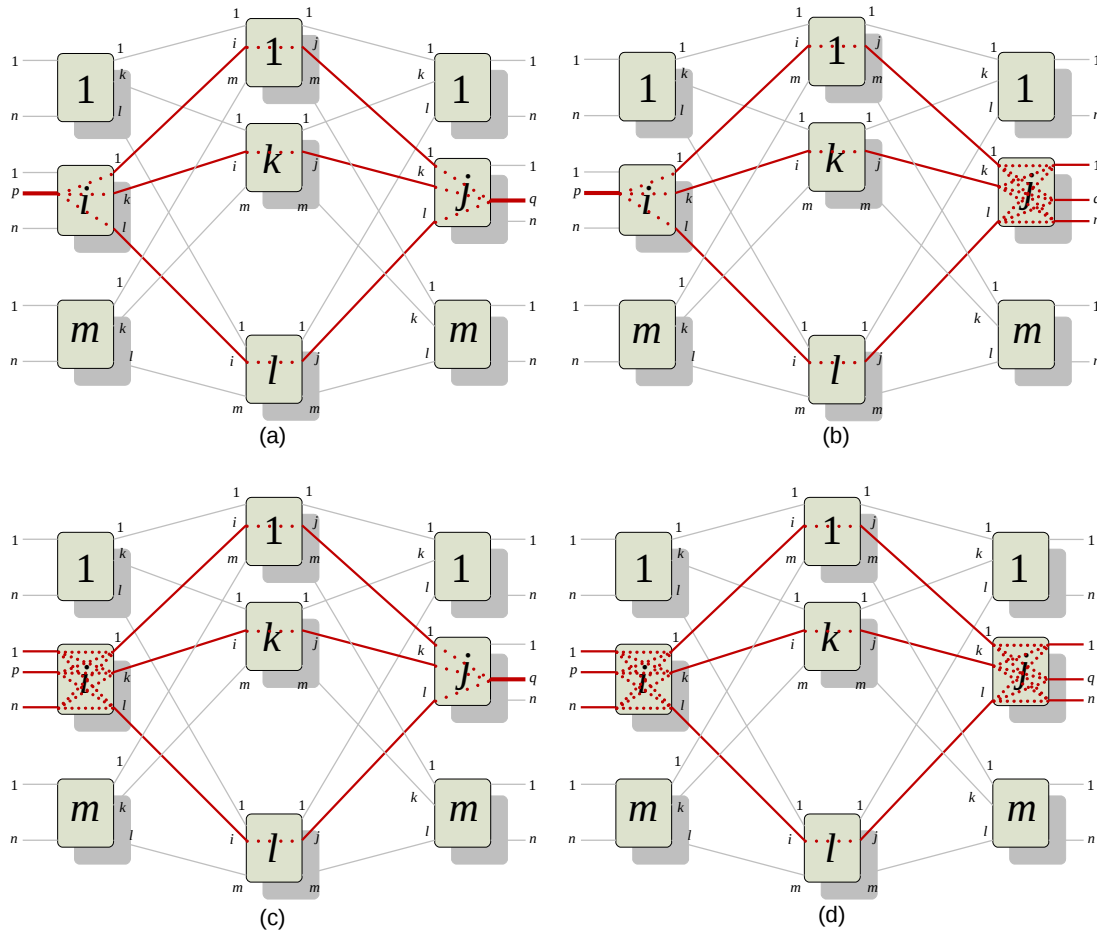
Neblokirajuće svojstvo paketskih svičeva sa Klosovim mrežama i balansiranjem saobraćaja pokazano je u [27]. Pri tome, zahvaljujući primeni balansiranja saobraćaja, neblokirajuće svojstvo sviča može se ostvariti i bez unutrašnjeg ubrzanja linkova u sviču. U ovom odeljku objasnićemo u kratkim crtama spomenute zaključke.

Posmatrajmo svič sa Klosovom strukturom, poput one na slici 5.1.2. Moduli u Klosovoj mreži mogu imati arhitekturu krosbara, ili neku drugu njoj ekvivalentnu, koja omogućava povezivanje svakog ulaza sa svakim izlazom.

Prilikom prenosa paketa kroz svič, vrši se podela njegovog sadržaja na manje jedinice približno iste dužine, koje se nazivaju ćelijama. Balansiranje saobraćaja pri prenosu paketa postiže se tako što se, u prvoj fazi rutiranja kroz svič, ćelije jednog toka usmeravaju preko različitih modula iz središnjeg niza Klosove mreže. Zatim se u drugoj fazi rutiranja kroz svič ćelije šalju na odgovarajući izlaz.

Prema [27], tok kome pripadaju ćelije možemo definisati na više načina. U prvom slučaju, pod tokom se podrazumevaju ćelije koje se sa jednog tačno određenog ulaza $i_p, 0 < i \leq m, 0 < p \leq n$ šalju na jedan tačno određen izlaz $j_q, 0 < j \leq m, 0 < q \leq n$. Ovaj slučaj predstavljen je na slici 6.1.1.(a). Na slici 6.1.1. crvenom bojom označene su moguće putanje za ćelije toka. U drugom slučaju, pod tokom se podrazumevaju sve ćelije koje se sa nekog tačno određenog ulaza $i_p, 0 < p \leq n$ usmeravaju na neki modul $j, 0 < j \leq m$ izlaznog niza – slika 6.1.1.(b). U trećem slučaju, tok čine sve ćelije koje sa jednog ulaznog modula $i, 0 < i \leq m$ šaljemo na tačno određen izlaz $j_q, 0 < j \leq m, 0 < q \leq n$ – slika 6.1.1.(c). U četvrtom slučaju, tok čine sve ćelije koje se

usmeravaju sa jednog ulaznog modula $i, 0 < i \leq m$ na jedan izlazni modul $j, 0 < j \leq m$ – slika slika 6.1.1.(d).



Slika 6.1.1. Tokovi ćelija u paketskom sviču sa balansiranjem saobraćaja

Prvo ćemo odrediti neophodan kapacitet linkova unutar sviča, koji garantuje neblokirajuće svojstvo [27].

Teorema 6.1.1. *Paketski svič sa Klosovom strukturom i balansiranjem saobraćaja ima neblokirajuće svojstvo ukoliko važi*

$$R_c \geq \frac{nR}{l}. \quad (6.1.1)$$

gde je R brzina ulaznih linija sviča, R_c kapacitet unutrašnjih linkova u sviču, l broj modula u središnjem nizu sviča, a n broj ulaza svakog od ulaznih modula sviča.

Dokaz: Nezavisno od toga šta se pod tokom podrazumeva, sve ćelije jednog toka ravnomerno se raspoređuju preko l modula u središnjem nizu Klosove mreže. Pretpostavimo da su sve ulazne linije sviča brzine R . Ukupno opterećenje koje jedan ulazni modul predaje sviču iznosi maksimalno nR , u slučaju kada su svi ulazi tog modula

u potpunosti iskorišćeni. Pošto se ovo opterećenje ravnomerno raspoređuje na l modula iz središnjeg niza, to je opterećenje svakog od unutrašnjih linkova u sviču najviše nR/l . Da bi svič bio neblokirajući, za kapacitet R_c unutrašnjeg linka mora važiti $R_c \geq nR/l$. ■

Teorema 6.1.2. *Za paketski svič sa Klosovom strukturom i balansiranjem saobraćaja nije neophodno uvoditi dodatno ubrzanje linkova unutar sviča da bi se garantovalo svojstvo neblokiranja.*

Dokaz: Ubrzanje sviča S definisano je kao odnos ukupnog kapaciteta unutrašnjih linkova prema ukupnom kapacitetu spoljašnjih linkova sviča. Ukupan kapacitet spoljašnjih linkova jednak je nmR , dok je ukupan kapacitet unutrašnjih linkova dat sa lmR_c . Dakle, ubrzanje je dato formulom

$$S = \frac{lmR_c}{nmR}. \quad (6.1.2)$$

Iz jednačina (6.1.1) i (6.1.2) možemo zaključiti da je za

$$S \geq 1. \quad (6.1.3)$$

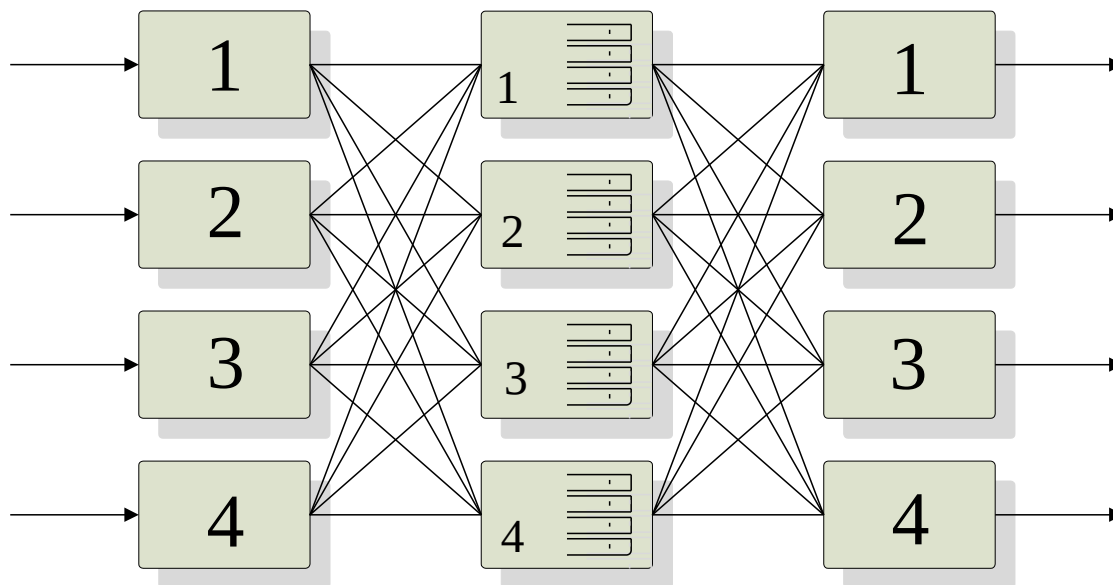
Drugim rečima, ukoliko se primeni balansiranje saobraćaja, moguće je ostvariti neblokirajuće svojstvo paketskog sviča sa klasičnom Klosovom arhitekturom i bez uvođenja ubrzanja unutrašnjih linkova sviča (tj. sa $S = 1$), ili povećanja kompleksnosti hardvera.

6.2. Optimalno balansiranje saobraćaja u svičevima

Balansiranje saobraćaja preko višestrukih putanja unutar sviča ili u komunikacionoj mreži, analizirano je u [28]. U ovom slučaju, ne pretpostavlja se postojanje klasične Klosove strukture sa tri niza modula, u kojoj svi ulazi priključeni na jedan ulazni modul dele kapacitet linkova koji povezuju taj ulazni modul sa modulima iz središnjeg niza. Pretpostavljena arhitektura sviča bazirana je na postojanju dve unifomne meš mreže koje omogućavaju da svaka ulazna linija može poslati svoj saobraćaj ka proizvoljnom modulu iz središnjeg niza, nezavisno od toga da li su ostali ulazi aktivni. Topologija sviča sa četiri ulazne i izlazne linije data na slici 6.2.1.

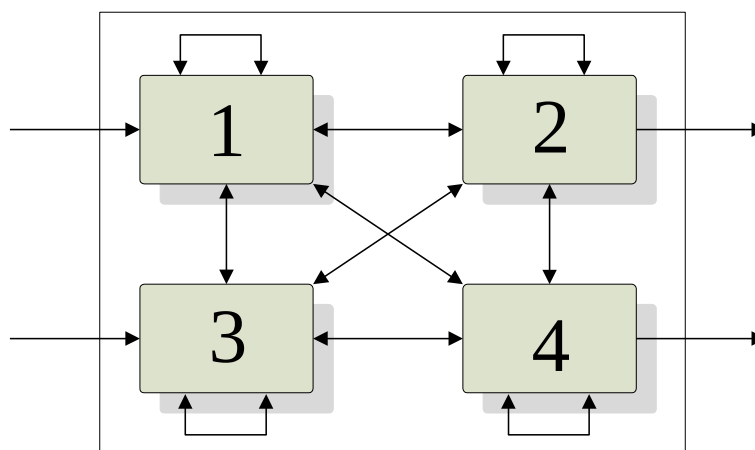
Paketi se sa nekog ulaza (linijske kartice) ravnomerno balansiraju preko središnjih modula. Središnji moduli sadrže bafere, koji su podeljeni u onoliko redova čekanja koliko ima izlaznih modula (tj. linijskih kartica). Paketi se iz redova čekanja u središnjem

modulu konstantnom brzinom šalju na odgovarajući izlaz. Dakle, u prvoj fazi vrši se balansiranje opterećenja, dok se u drugoj fazi paketi prosleđuju ka odredištu.



Slika 6.2.1. Arhitektura sviča sa balansiranjem saobraćaja

Topologija ekvivalentna topologiji sa slike 6.2.1 može se ostvariti i direktnim povezivanjem linijskih kartica u potpuni meš, kao na slici 6.2.2. U tom slučaju, svaki paket kroz svič prolazi dva puta, te je neophodno uvesti ubrzanje unutrašnjih linkova sviča, $S = 2$. Analiza u [28] pokazuje da je ovo ubrzanje dovoljno da se ostvari neblokirajuće svojstvo sviča. U ovom odeljku objasnićemo ukratko zašto.



Slika 6.2.2. Ekvivalentna topologija unutrašnjosti sviča sa balansiranjem saobraćaja

Usvojimo prvo notaciju. Neka su u mreži sa N identičnih čvorova ($N \geq 2$) kapaciteti linkova dati matricom $\mathbf{C} = [C_{ij}]_{1 \leq i, j \leq N}$. Takođe, neka je saobraćaj u mreži opisan matricom saobraćaja $\mathbf{T} = [T_{ij}]_{1 \leq i, j \leq N}$. Pretpostavlja se da je maksimalna brzina kojom neki čvor šalje ili prima saobraćaj jednaka jedinici. Drugim rečima, za neki čvor $n \in V$ važi

$$\begin{aligned} \sum_{j \in V} C_{nj} &\leq 1 \\ \sum_{i \in V} C_{in} &\leq 1. \end{aligned} \quad (6.2.1)$$

Posmatrajmo par čvorova i i j , i označimo skup svih mogućih putanja p koje između njih postoje sa $P(i, j)$. Saobraćaj između ova dva čvora balansira se preko svih raspoloživih putanja $p \in P(i, j)$, i to tako što se po putanji p šalje saobraćaj T_{ij}^p . Naravno, za svaki par (i, j) mora važiti

$$T_{ij} = \sum_{p \in P(i, j)} T_{ij}^p. \quad (6.2.2)$$

Prilikom prenosa saobraćaja od i ka j , svaki link l koji se nalazi na nekoj od putanja $p \in P(i, j)$ trpi opterećenje T_{ij}^p . Ukupno opterećenje proizvoljnog linka $l \in E$ dato je kao suma opterećenja po svim parovima čvorova i putanjama između njih:

$$L_l = \sum_{(i, j) \in V \times V} \sum_{p \in P(i, j), l \in p} T_{ij}^p. \quad (6.2.3)$$

Autori u [28] rešavaju problem maksimizacije propusnosti mreže. Prema definiciji 3.1.4, propusnost mreže definisana je kao faktor θ kojim se mogu skalirati svi elementi u matrici saobraćaja $\mathbf{T}$, tako da saobraćaj bude rutiran kroz mrežu. Propusnost mreže zavisi kako od topologije mreže (koja je opisana matricom kapaciteta linkova), tako i od saobraćaja koji mreža treba da podrži (a koji je dat matricom saobraćaja). Za mrežu sa kapacitetima linkova $\mathbf{C}$, u slučaju saobraćaja koji je dat matricom $\mathbf{T}$, optimalno balansiranje saobraćaja je ono za koje se postiže maksimum propusnosti mreže $\theta(\mathbf{C}, \mathbf{T})$. Raspodela saobraćaja i maksimum propusnosti mreže za neki dati saobraćaj određuju se sledećim linearnim programom:

$$\theta(\mathbf{C}, \mathbf{T}) = \max_{\theta} \theta$$

$$(1) \forall i, j \in V : \sum_{p \in P(i, j)} T_{ij}^p = \theta T_{ij}$$

$$(2) \forall i, j \in V : L_{l=(i, j)} \leq C_{ij}$$

$$(3) \forall i, j \in V, \forall p \in P(i, j) : T_{ij}^p \geq 0$$
(6.2.4)

Garantovana propusnost mreže $\theta(\mathbf{C})$ je ona vrednost propusnosti koja se može ostvariti u najgorem slučaju. Dakle, svaka matrica saobraćaja iz skupa mogućih matrica saobraćaja $\mathbf{T} = \{\mathbf{T}\}$ može se skalirati faktorom $\theta(\mathbf{C})$, a da pri tom saobraćaj može biti rutiran kroz mrežu. Drugim rečima, garantovana propusnost mreže određena je matricom saobraćaja $\mathbf{T}$ za koju se ostvaruje minimalna vrednost faktora $\theta(\mathbf{C}, \mathbf{T})$:

$$\theta(\mathbf{C}) = \min_{\mathbf{T}} \theta(\mathbf{C}, \mathbf{T}).$$
(6.2.5)

Optimalna topologija mreže je ona za koju se postiže maksimum $\theta(\mathbf{C})$, tj. maksimalna moguća garantovana propusnost mreže:

$$\theta^* = \max_{\mathbf{C}} \theta(\mathbf{C})$$
(6.2.6)

6.2.1. Garantovana propusnost uniformne meš mreže

Kod uniformne meš mreže, svi čvorovi su međusobno povezani i svi linkovi imaju jednake kapacitete $C_{ij} = 1/N$.

Definicija 6.2.1. Kapaciteti linkova uniformne meš mreže dati su matricom $\mathbf{C}_{UM}$, gde je $C_{ij} = 1/N$ za svaki par (i, j) .

Teorema 6.2.1. Garantovana propusnost $\theta(\mathbf{C}_{UM})$ uniformne meš mreže iznosi

$$\theta(\mathbf{C}_{UM}) = 50\%.$$
(6.2.7)

Dokaz: Posmatrajmo saobraćaj po linku između dva čvora i i j . On se sastoji iz dve komponente. Prvu čini saobraćaj koji se generiše u i i balansira ka proizvoljnom čvoru k u mreži, koristeći putanju preko j . Drugu komponentu predstavlja saobraćaj koji j prima od ostalih čvorova u mreži, a koji se balansira preko putanje koja sadrži i . Pošto smo pretpostavili da je ukupan saobraćaj koji čvor može da primi ili pošalje jednak jedinici, a u najgorem slučaju sav svoj saobraćaj i usmerava, a j prima preko posmatranog linka, to je opterećenje linka

$$L_{ij} = \sum_{k \in V} T_{ik} + \sum_{k \in V} T_{kj} \leq 2.$$
(6.2.8)

Dakle, skaliranje svakog od tokova faktorom $\theta(\mathbf{C}_{UM}) \geq 50\%$ obezbeđuje da nijedan link ne bude blokiran. Sa druge strane, može se pokazati da postoji slučaj matrice saobraćaja za koji je $\theta(\mathbf{C}_{UM}) \leq 50\%$ [28]. Naime, posmatrajmo matricu saobraćaja

$$\mathbf{T} = \begin{bmatrix} 0 & x & 0 & \cdots & 0 \\ 0 & 0 & x & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & x \\ x & 0 & \cdots & 0 & 0 \end{bmatrix} \quad (6.2.9)$$

Čvor i šalje čvoru $j = (i+1) \bmod N$ saobraćaj intenziteta x . Pri tome, maksimalno $1/N$ se može poslati po linku C_{ij} , dok ostatak saobraćaja $x - 1/N$ mora ići alternativnim putanjama, koje imaju barem dva linka. Dakle, suma opterećenja svih linkova u mreži koji prenose saobraćaj od i ka j jednaka je barem $1/N + 2(x - 1/N)$.

Suma opterećenja svih linkova u mreži sa N čvorova i matricom $\mathbf{T}$ iz (6.2.9) jednaka je $N(1/N + 2(x - 1/N)) = 2Nx - 1$.

S obzirom da kod ovakve matrice saobraćaja nema smisla koristiti linkove C_{ii} , to je maksimalan kapacitet koji je na raspolaganju za prenos saobraćaja jednak $N(1 - 1/N) = N - 1$.

Suma opterećenja linkova koji prenose saobraćaj od i ka j mora biti manja od ukupnog raspoloživog kapaciteta, te mora važiti

$$2Nx - 1 \leq N - 1, \quad (6.2.10)$$

odnosno

$$x \leq 1/2. \quad (6.2.11)$$

Matrica $\mathbf{T}$ u kojoj je $x = 1$ zadovoljava uslov da svaki čvor generiše i prima maksimalno jedinični saobraćaj, tako da je propusnost mreže u tom slučaju jednaka maksimalno 50%. Ovim je dokaz završen. ■

Ukoliko se uvede ubrzanje unutrašnjih linkova u sviču, $S = 2$, onda mreža postaje neblokirajuća i sav saobraćaj koji dođe na neku od ulaznih linijskih kartica biće rutiran do odredišnog izlaza.

6.2.2. *Garantovana propusnost polarisane meš mreže*

U prethodnom pododeljku pokazano je da je propusnost uniformne meš mreže jednaka 50%. Međutim, uniformna meš mreža nije optimalna. Pokazuje se da se bolja propusnost može ostvariti za takozvanu polarisanu meš mrežu [28].

Definicija 6.2.2. *Polarisana meš mreže date matricom kapaciteta linkova $\mathbf{C}_{BM}$, takvom da za svako $i \in V$ i svako $j \neq i$ važi*

$$C_{ij} = 2C_{ii}. \quad (6.2.12)$$

Drugim rečima, kapacitet linkova koji povezuju čvorove u mreži sa samima sobom je dvostruko manji od kapaciteta linkova koji povezuju dva različita čvora.

Može se pokazati [28] da je garantovana propusnost polarisane meš mreže jednaka

$$\theta(\mathbf{C}_{BM}) = \frac{N}{2N-1} \quad (6.2.13)$$

Takođe, u [28] je pokazano da za svaku moguću matricu kapaciteta $\mathbf{C} \neq \mathbf{C}_{BM}$ važi da je

$$\theta(\mathbf{C}) \leq \theta(\mathbf{C}_{BM}). \quad (6.2.14)$$

Znači, polarisana meš mreža je ona topologija za koju se postiže maksimalna vrednost garantovane propusnosti, te predstavlja optimalnu topologiju za primenu balansiranja saobraćaja.

7. RBS ALGORITAM – RUTIRANJE SA BALANSIRANJEM SAOBRAĆAJA

U prethodnom poglavlju bilo je reči o primeni balansiranja saobraćaja pri dizajnu svičeva velikog kapaciteta. Pokazano je da upotreba balansiranja saobraćaja omogućava garantovanje neblokirajućeg svojstva sviča, bez uvođenja unutrašnjeg ubrzanja za linkove unutar sviča [27]. Takođe, primena balansiranja omogućava maksimizaciju propusne moći sviča [28].

Šema rutiranja uvedena u [8], [9], koja će biti opisana u ovom poglavlju, zasnovana je na primeni principa balansiranja saobraćaja u realnim paketskim mrežama. U ovom slučaju, umesto direktnog rutiranja saobraćaja između neka dva čvora mreže, rutiranje se sprovodi u dve faze. Pri tome se saobraćaj u prvoj fazi usmerava od izvora ka čvorovima posrednicima, a zatim u drugoj fazi od čvorova posrednika ka konačnom odredištu. Za razliku od situacije sa kojom se susrećemo u komutacionim poljima telefonskih centrala i paketskih svičeva, u slučaju realnih mreža ne postoji regularnost i simetrija mrežne topologije. Ipak, pokazuje se da i u takvim uslovima postoje prednosti primene balansiranja saobraćaja na predloženi način, i upravo te prednosti biće izložene u ovom radu.

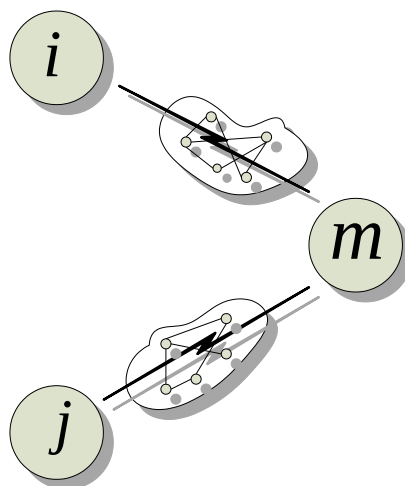
Osnovna prednost primene balansiranja saobraćaja leži u činjenici da se time omogućava formulisanje linearnog programa za pronalaženje optimalnog rutiranja isključivo u funkciji vektora ulaznog i izlaznog saobraćaja u čvoru. Već je pomenuto da je mnogo jednostavnije predvideti stanje saobraćaja u mreži ukoliko je ono dato u obliku ova dva vektora, a ne u obliku matrice saobraćaja. Takođe, predloženi način rutiranja omogućava i brz, jednostavan i distribuiran proces rezervacije kapaciteta za multimedijalni saobraćaj. Naime, svaki ruter u mreži raspolaže informacijom o tome koji je maksimalan saobraćaj koji mu je dozvoljeno da generiše. U slučaju pojave zahteva za uspostavljanjem nove sesije, neophodno je da ruteri kroz koje saobraćaj date sesije ulazi, odnosno napušta mrežu, provere da li na raspolaganju imaju dovoljno slobodnog kapaciteta. Ukoliko je to slučaj, sesija se može uspostaviti. Pri tome, nije neophodno

poznavanje aktivnosti ostalih rutera u mreži niti stanja na linkovima, a u procesu pregovaranja učestvuju samo ova dva rutera, granična za saobraćaj posmatrane sesije.

U ovom odeljku biće najpre opisana predložena šema rutiranja. Zatim će biti reči o načinu na koji se vrši optimizacija predložene šeme rutiranja sa balansiranjem saobraćaja.

7.1. Opis šeme rutiranja sa balansiranjem saobraćaja

U ovom odeljku opisaćemo način na koji se vrši rutiranje sa balansiranjem saobraćaja. Kao što je to uobičajeno, posmatraćemo mrežu kao graf $\Gamma = (V, E)$. Ovde je sa V označen skup čvorova (rutera), a E skup grana (linkova) u grafu, tj. posmatranoj mreži. Ukupan broj čvorova u mreži označavamo sa $|V| = N$, a broj linkova sa $|E| = M$. Takođe, pretpostavljamo da je saobraćaj u mreži opisan vektorima ulaznog i izlaznog saobraćaja, $\mathbf{S}$ i $\mathbf{R}$, kako je to definisano u Poglavlju 2.

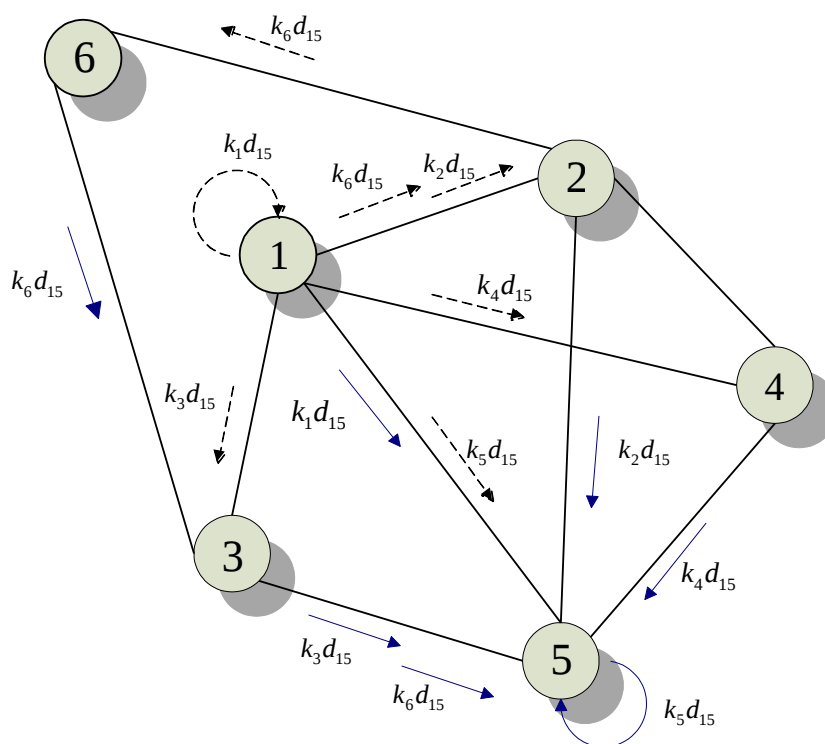


Slika 7.1.1. Dve faze rutiranja u predloženoj šemi

Usvajamo koncept balansiranja saobraćaja, te umesto da saobraćaj između dva čvora rutiramo direktnom putanjom, rutiranje sprovodimo u dve faze. U okviru prve faze rutiranja, određeni delovi toka generisanog u čvoru i a namenjeni čvoru j usmeravaju se preko čvorova posrednika $m \in V$. Zatim se, u okviru druge faze rutiranja, saobraćaj primljen od strane čvora posrednika usmerava ka krajnjem odredištu. Čvorovi posrednici nazivaju se balansirajućim čvorovima (ruterima). Dakle, logički posmatrano, putanja između neka dva čvora i i j sastoji se iz dva dela: od izvora i ka balansirajućem čvoru m i od balansirajućeg čvora m ka konačnom odredištu j , kao na Slici 7.1.1. Fizički

Svaki čvor m u mreži može imati ulogu balansirajućeg čvora. Udeo toka između i i j koji se usmerava preko nekog balansirajućeg čvora $m \in V$ određen je koeficijentom k_m koji je dodeljen čvoru m . Nezavisno od toga koji je čvor izvor, a koji krajnje odredište toka, uvek isti udeo k_m balansira se preko čvora m . Za koeficijente k_m , naravno, moraju važiti sledeća ograničenja:

Naravno, balansiranje saobraćaja vršiće se preko onih čvorova za koje je $k_m \neq 0$, ali u opštem slučaju svi čvorovi mogu biti oni preko kojih se balansira.



Na Slici 7.1.2. je prikazana ilustracija predložene šeme. U konkretnom primeru, posmatramo tok između čvorova 1 i 5 u datoj mreži. Prva faza rutiranja označena je

isprekidanom, a druga faza punom linijom na slici. Kada bi se koristilo klasično rutiranje po najkraćoj putanji, sav saobraćaj usmeravao bi se preko linka između ova dva čvora, koji predstavlja najkraću putanju između njih. Umesto da se saobraćaj usmerava samo po jednom linku, upotrebom balansiranja saobraćaja u kombinaciji sa rutiranjem po najkraćoj putanji, praktično skoro svi linkovi u mreži sa slike koristiće se za prenos paketa. Na taj način postiže se ravnomernija raspodela opterećenja među svim linkovima. Takođe, ukoliko ostali čvorovi nisu aktivni, čvor 1 može poslati čvoru 5 znatno veći saobraćaj nego kada koristi samo direktan link.

7.2. Opterećenje linka kao funkcija saobraćaja čvorova

Već je pomenuto da upotreba balansiranja saobraćaja omogućava da se problem pronalaženja optimalnog rutiranja formuliše isključivo u funkciji ukupnog saobraćaja generisanog, tj. primljenog u čvoru mreže. Naime, pokazuje se da je opterećenje linka u mreži moguće izraziti preko vrednosti elemenata vektora saobraćaja i da nije potrebno poznavati konkretan oblik matrice saobraćaja. Pokazaćemo to sada i na primeru.

Posmatrajmo saobraćaj između neka dva čvora i i m u mreži. U okviru prve faze rutiranja, čvor i usmerava udeo k_m svakog toka d_{ij} preko čvora m , nezavisno od toga koje je konačno odredište saobraćaja, j . Ukupan saobraćaj koji čvor i balansira preko čvora m jednak je

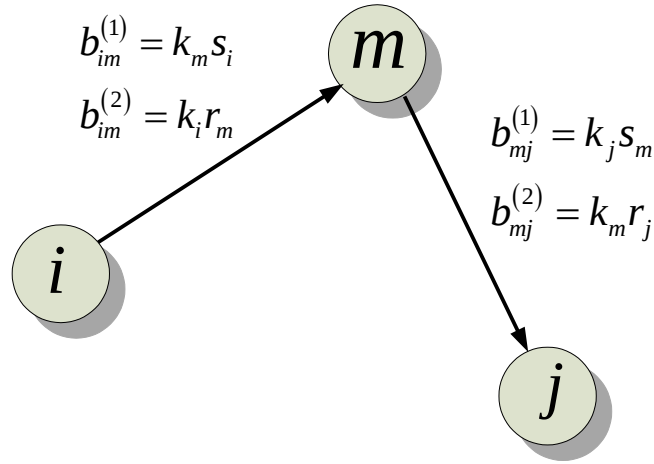
$$b_{im}^{(1)} = \sum_{j \in V} k_m d_{ij} = k_m s_i. \quad (7.2.1)$$

Sa druge strane, čvor i igra ulogu balansirajućeg čvora za saobraćaj generisan u proizvoljnom čvoru mreže, a namenjen čvoru m kao konačnom odredištu. Stoga se, u okviru druge faze rutiranja, od čvora i do čvora m usmerava saobraćaj

$$b_{im}^{(2)} = \sum_{p \in V} k_i d_{pm} = k_i r_m. \quad (7.2.2)$$

Ukupan saobraćaj između i i m jednak je, dakle:

$$b_{im} = b_{im}^{(1)} + b_{im}^{(2)} = \sum_{j \in V} k_m d_{ij} + \sum_{p \in V} k_i d_{pm} = k_m s_i + k_i r_m \quad (7.2.3)$$



Slika 7.2.1. Intenziteti tokova u različitim fazama rutiranja

Uvešćemo promenljivu F_{im}^l , koja je jednaka jedinici ukoliko se saobraćaj od i ka m usmerava preko linka $l \in E$, odnosno jednaka je nuli ukoliko to nije slučaj. Ova promenljiva izračunavaće se na osnovu tabela rutiranja po najkraćoj putanji, odnosno na osnovu rezultata Dijkstra algoritma. Uvođenjem ove promenljive, za opterećenje linka l možemo pisati da je jednako:

$$L^l = \sum_{(i,m)} F_{im}^l (b_{im}^{(1)} + b_{im}^{(2)}) \quad (7.2.4)$$

Vidimo da ovaj saobraćaj zavisi samo od koeficijenata pridruženih čvorovima i i m , kao i ukupnog saobraćaja generisanog, odnosno primljenog u tim čvorovima.

U slučaju koji analiziramo, koristi se rutiranje po najkraćoj putanji u obe faze rutiranja (od izvora do balansirajućeg čvora, kao i od balansirajućeg čvora do odredišta). Stoga je putanja između neka dva čvora u mreži jedinstvena i važi da su koeficijenti F_{im}^l binarne promenljive, $F_{im}^l \in \{0,1\}$. Međutim, u opštem slučaju, rutiranje u svakoj fazi može biti proizvoljno, definisano koeficijentima rutiranja $f_{im}^l \in [0,1]$, kao u (2.3.4). Autori u [6] i [7] koriste ovakav pristup da na osnovu vektora ulaznog i izlaznog saobraćaja proračunaju elemente matrice saobraćaja, a zatim rešavaju linearni problem pronalaženja optimalnih vrednosti koeficijenata rutiranja i koeficijenata balansiranja, kako bi se dobilo rutiranje sa zadatim performansama – maksimalnim saobraćajem u čvoru ili minimalnom cenom mreže. Program koji je u ovom slučaju neophodno rešiti

ima $MN^2 + N + 1$ promenljivih i u opštem slučaju sadrži ograničenja vezana kako za kapacitete linkova, tako i za konzervaciju protoka u čvorovima. Strogo matematički posmatrano, ovakav pristup je uopšteniji i rešenje koje se dobija je bliže optimalnom. Međutim, dimenzija problema koji se rešava nije pogodna za upotrebu u praksi, zbog dugog vremena neophodnog za optimizaciju. Upotreba rutiranja po najkraćoj putanji u obe faze predložene šeme sa balansiranjem saobraćaja značajno uprošćava problem. S obzirom na to da su pri ovakvoj postavci koeficijenti rutiranja unapred određeni, jedine vrednosti koje se mogu menjati su koeficijenti balansiranja, tako da linearni program ima $N + 1$ promenljivih. Linearni program je zato znatno manji i brže se pronalazi optimalno rešenje. Zahvaljujući brzini izračunavanja optimalnih vrednosti koeficijenata balansiranja, kao i činjenici da nije neophodno menjati način rutiranja u celoj mreži, predloženo rutiranje se može relativno lako implementirati u praksi, pri čemu se ostvaruje kompatibilnost sa mrežama u kojima modifikacija nije primenjena [13], [15].

7.3. Optimizacija rutiranja sa balansiranjem saobraćaja

Kao što je već spomenuto, rutiranje sa balansiranjem saobraćaja može se optimizovati tako da se ostvari maksimalan saobraćaj u čvoru postojeće mreže ili da se omogući projektovanje mreže sa minimalnom cenom.

U opštem slučaju, linearni program za optimizaciju rutiranja u mreži $\Gamma = (V, E)$ ima sledeći oblik:

$$LP(\Gamma) = \begin{cases} \text{ciljna funkcija} \\ (C1) \text{ ograničenje za koeficijente balansiranja} \\ (C2) \text{ ograničenja vezana za opterećenje linkova} \\ (C3) \text{ ograničenja vezana za konzervaciju toka u čvorovima} \\ (C4) \text{ dodatna ograničenja} \end{cases} \quad (7.3.1)$$

Ciljna funkcija, kao i broj promenljivih u linearnom programu, zavisi od tipa optimizacionog problema koji rešavamo. Ograničenja tipa (C1) za koeficijente balansiranja data su jednačinama (7.1.1). Ograničenja tipa (C2) služe da se uspostavi veza između opterećenja i kapaciteta linkova u mreži – ni za jedan link opterećenje ne sme biti veće od kapaciteta tog linka:

$$\forall l \in E: L^l \leq C^l. \quad (7.3.2)$$

Ograničenja tipa (C3) tiču se konzervacije protoka u čvorovima mreže – razlika opterećenja ulaznih i izlaznih linkova za neki čvor u mreži mora biti jednaka razlici primljenog i generisanog saobraćaja u tom čvoru:

$$\forall n \in V: \sum_{l \in IN(n)} L^l - \sum_{l \in OUT(n)} L^l = r_n - s_n. \quad (7.3.3)$$

Ovde je sa $IN(n)$ označen skup ulaznih linkova, sa $OUT(n)$ skup izlaznih linkova za posmatrani čvor n , a r_n i s_n su primljeni, odnosno generisani saobraćaj u tom čvoru.

Ograničenja tipa (C4) predstavljaju eventualna dodatna ograničenja, koja zavise od prirode optimizacionog problema.

Obratićemo sada pažnju na jednu od prednosti koja proizlazi iz načina na koji je rutiranje sa balansiranjem saobraćaja definisano, odnosno iz činjenice da se rutiranje po najkraćoj putanji koristi u obe faze predložene šeme rutiranja. Naime, može se pokazati da su, uz pretpostavku jednakosti ulaznog i izlaznog saobraćaja u čvoru $r_n = s_n$, ograničenja tipa (C3) redundantna i mogu se izostaviti.

Lema 7.3.1: *Ukoliko su ulazni i izlazni saobraćaj u svim čvorovima mreže međusobno jednaki, $r_n = s_n$ za svako $n \in V$, tada su ograničenja tipa (C3) uvek zadovoljena i mogu se izostaviti iz modela linearnog programa.*

Dokaz: Ukoliko zamenimo odgovarajuće izraze za L^l , uz činjenicu da je $r_n = s_n$, dobijamo sledeći oblik ograničenja (C3) za neki čvor n :

$$\sum_{\substack{(i,j) \\ l \in IN(n)}} F_{ij}^l (k_i s_j + k_j s_i) = \sum_{\substack{(i,j) \\ l \in OUT(n)}} F_{ij}^l (k_i s_j + k_j s_i) \quad (7.3.4)$$

Posmatrajmo prvo slučaj kada je $i \neq n, j \neq n$. Tada se čvor n može naći na najkraćoj putanji između i i j , a može biti i izvan nje. Ako se n ne nalazi na toj putanji, onda ne postoje linkovi koji vode ka n ili ga napuštaju, takvi da je $F_{ij}^l = 1$. Dakle, članovi oblika $(k_i s_j + k_j s_i)$ ne pojavljuju se ni na jednoj strani jednakosti. Ukoliko se n , pak, nalazi na putanji između čvorova i i j , tada postoji tačno jedan ulazni link i tačno jedan izlazni link za n , takav da je $F_{ij}^l = 1$. Drugim rečima, članovi oblika $(k_i s_j + k_j s_i)$ javljaju se sa obe strane jednakosti i mogu se iz nje eliminisati. Na ovaj način ostaju nam samo članovi vezani za slučajeve kada je $i = n$ ili $j = n$.

Analizirajmo sada slučaj kada je $i = n$. Kako je n izvorišni čvor za saobraćaj ka j u ovom slučaju, to ne može postojati link $l \in IN(n)$ takav da je $F_{ij}^l = F_{nj}^l = 1$, jer bi u tom slučaju putanja od n ka j sadržala petlju, i ne bi bila najkraća. Slično, ukoliko je $j = n$, ne može postojati link $l \in OUT(n)$ takav da je $F_{ij}^l = F_{nj}^l = 1$.

Dakle, možemo jednačinu prepisati na sledeći način:

$$\sum_{l \in IN(n)} F_{in}^l (k_i s_n + k_n s_i) = \sum_{l \in OUT(n)} F_{nj}^l (k_n s_j + k_j s_n) \quad (7.3.5)$$

U povezanoj mreži mora postojati putanja između svaka dva čvora i ukoliko se koristi rutiranje po najkraćoj putanji, ona je jedinstvena. To znači da za svako i mora postojati tačno jedan link $l \in IN(n)$ takav da je $F_{in}^l = 1$. I obrnuto, za svako j mora postojati link $l \in OUT(n)$ takav da je $F_{nj}^l = 1$. Pošto je

$$\sum_i k_i s_n = \sum_j k_j s_n = s_n \quad (7.3.6)$$

kao i

$$\sum_i k_i s_i = \sum_j k_j s_j \quad (7.3.7)$$

to su leva i desna strana jednačine praktično identične. Jednakost je, dakle, uvek zadovoljena. To znači da je i početni uslov konzervacije protoka uvek ispunjen i da nije neophodno uvoditi ga kao ograničenje u linearni program. ■

Kao što je već rečeno, optimizacija rutiranja u cilju maksimizacije propusne moći ili minimizacije cene mreže biće vršena za realne i regularne mrežne topologije. Biće analizirana dva slučaja: slučaj kada pretpostavljamo da nikada ne dolazi do kvarova resursa u mreži, kao i slučaj kada pretpostavljamo da su mogući jednostruki kvarovi resursa (rutera ili linkova).

U skladu sa kriterijumom za optimizaciju rutiranja, razlikovaće se i interpretacija vrednosti elemenata vektora saobraćaja $\mathbf{S}$ i $\mathbf{R}$. Naime, u slučaju minimizacije cene mreže, elementi vektora saobraćaja predstavljaju vrednosti saobraćaja koji treba rutirati. U slučaju maksimizacije propusne moći mreže, elementi vektora saobraćaja biće proporcionalni težinama čvorova. Težinama čvorova određuje se međusobni odnos garantovanih saobraćaja za različite čvorove u mreži, tako što se čvorovima veće težine garantuje se veći saobraćaj. Težine čvorova mogu biti međusobno jednake, u slučaju kada su svi čvorovi ravnopravni i želimo da im garantujemo jednake saobraćaje. U praksi,

čvorovima sa većim brojem korisnika treba dodeliti više garantovanog kapaciteta. Odabirom težina čvorova tako da one budu proporcionalne broju korisnika koje čvor treba da servisira postiže se bolja raspodela garantovanog kapaciteta među čvorovima.

8. MAKSIMIZACIJA PROPUSNE MOĆI MREŽE

Optimizacijom rutiranja u cilju maksimizacije propusne moći mreže omogućava se servisiranje maksimalnog broja korisnika, uz korišćenje već postojećih mrežnih resursa. Kao što je već rečeno, broj korisnika koji se u mreži može servisirati u direktnoj je vezi sa garantovanim saobraćajem u čvorovima te mreže. Naime, garantovani saobraćaj čvora se u svakoj situaciji može propustiti kroz mrežu. Sve dok su generisane vrednosti saobraćaja čvora ispod garantovane, možemo biti sigurni da će servis biti dostupan svim korisnicima vezanim za taj čvor. Što je garantovani saobraćaj za čvor mreže veći, to se više korisnika može povezati na čvor, bez bojazni da će doći do prekoračenja dozvoljene vrednosti i blokade servisa. Stoga je, u slučaju već izgrađene mreže, u kojoj su definisani i topologija i kapaciteti svih linkova, neophodno izvršiti optimizaciju rutiranja kako bi se maksimizovao garantovani saobraćaj u čvoru mreže i omogućilo servisiranje maksimalnog mogućeg broja korisnika bez potrebe za dodatnim ulaganjem u povećanje raspoloživih kapaciteta linkova. Sa maksimizacijom garantovanog saobraćaja u svakom čvoru, maksimizira se i ukupna propusna moć mreže.

U ovom poglavlju biće opisan način na koji se vrši optimizacija rutiranja sa balansiranjem saobraćaja, u cilju maksimizacije propusne moći mreže. Zatim će biti objašnjeno kako se procenjuje propusna moć mreže u kojoj je primenjeno obično rutiranje po najkraćoj putanji. Na kraju će biti upoređene propusne moći mreže u ova dva slučaja. Biće pokazano da upotreba rutiranja sa balansiranjem saobraćaja omogućava rutiranje većeg saobraćaja u odnosu na slučaj kada je primenjeno obično rutiranje po najkraćoj putanji, kako za neke regularne mrežne topologije, tako i za realne mreže.

8.1. Propusna moć mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)

Kao što je već pokazano u odeljku 5.2, opterećenje linka u mreži u kojoj je primenjeno balansiranje saobraćaja dato je izrazom

$$L^l = \sum_{(i,j)} F_{ij}^l (b_{ij}^{(1)} + b_{ij}^{(2)}). \quad (8.1.1)$$

Iskorišćenje linka je onda dato formulom

$$U^l = L^l / C^l = \sum_{(i,j)} F_{ij}^l (b_{ij}^{(1)} + b_{ij}^{(2)}) / C^l \quad (8.1.2)$$

gde je C^l kapacitet linka.

Kao što je objašnjeno u odeljku 3.1.1, problem maksimizacije garantovanog saobraćaja u čvoru mreže ekvivalentan je problemu minimizacije zagušenja. Što je niža vrednost zagušenja, to se više mogu povećavati intenziteti svih tokova u mreži, a da pri tome ne dođe do preopterećivanja nekog od linkova u mreži. Dakle, cilj je minimizacija zagušenja Q , koje je definisano kao maksimalno iskorišćenje linka u mreži, $Q = \max_{l \in V} U^l$.

U opštem slučaju linearni program koji optimizuje koeficijente k_m u cilju minimizacije zagušenja ima oblik:

$$\begin{aligned} & \underline{\min Q} \\ & (C1) \sum_{i=1}^N k_i = 1 \\ & (C2) \forall l \in E : \sum_{(i,m)} F_{im}^l (k_i r_m + k_m s_i) / C^l \leq Q \\ & (C3) \forall n \in V : \sum_{l \in IN(n)} L^l - \sum_{l \in OUT(n)} L^l = r_n - s_n \end{aligned} \quad (8.1.3)$$

Ovde je sa $IN(n)$ označen skup linkova koji ulaze u čvor n , a sa $OUT(n)$ skup linkova koji iz njega izlaze. Uz pretpostavku jednakosti ulaznog i izlaznog saobraćaja za sve čvorove u mreži ($\forall n \in V : r_n = s_n$), ograničenja tipa (C3) se mogu izostaviti, kao što je pokazano u odeljku 7.3 (lema 7.3.1). Konačni oblik linearnog programa ima, dakle, $N + 1$ promenljivu i $M + 2$ ograničenja (ne računajući podrazumevana ograničenja o nenegativnosti promenljivih) i sledeći oblik:

$$\begin{aligned} & \underline{\min Q} \\ & (C1) \sum_{i=1}^N k_i = 1 \\ & (C2) \forall l \in E : \frac{\sum_{(i,j)} F_{ij}^l (k_i s_j + k_j s_i)}{C^l} \leq Q \end{aligned} \quad (8.1.4)$$

Garantovani saobraćaj u čvoru mreže obrnuto je proporcionalan zagušenju.

$$\forall i \in V : g_i^{rbs} = \frac{s_i}{Q} \quad (8.1.5)$$

Elementi vektora saobraćaja $\mathbf{S} = [s_i]$ proporcionalni su težinama čvorova u mreži, koje ćemo označiti sa w_i . Zadavanjem težina čvorovima, možemo kontrolisati saobraćaj koji se pojedinim čvorovima garantuje – što je veća težina nekog čvora, veći će biti i njegov garantovani saobraćaj. Analizirani su slučajevi u kojima su svi čvorovi iste težine, kao i slučajevi u kojima su težine čvorova proporcionalne ukupnom kapacitetu izlaznih linkova čvora, odnosno broju korisnika koje čvor servisira. Odgovarajuća definicija težina čvorova biće data za svaki analizirani slučaj u odeljku 8.3.

Propusna moć mreže jednaka je sumi garantovanih saobraćaja svih čvorova:

$$T^{rbs} = \sum_{i \in V} g_i^{rbs} = \frac{1}{Q} \sum_{i \in V} s_i. \quad (8.1.6)$$

8.2. Propusna moć mreže sa rutiranjem po najkraćoj putanji (RNP)

Da bismo uporedili performanse rutiranja po najkraćoj putanji sa performansama rutiranja sa balansiranjem saobraćaja, neophodno je proceniti garantovani saobraćaj u čvoru za slučaj klasičnog rutiranja po najkraćoj putanji. Problem sa kojim se u ovom slučaju susrećemo predstavlja činjenica da je opterećenje linka kod rutiranja po najkraćoj putanji funkcija intenziteta tokova između parova čvorova izvor-odredište, a ne ukupnog saobraćaja u čvoru. Zato je neophodno usvojiti model koji nam omogućava da ove dve veličine povežemo. Usvojeni model biće opisan u ovom odeljku.

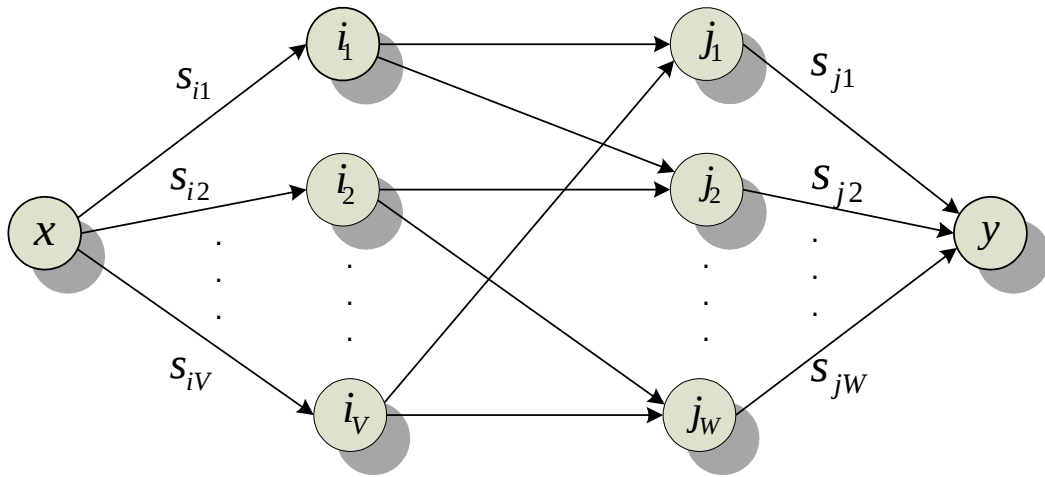
Posmatraćemo najgori slučaj za neki link, kada su svi čvorovi koji preko njega šalju ili primaju saobraćaj aktivni. Opterećenje linka u ovakvoj situaciji je maksimalno, i određuje ograničenja za saobraćaje čvorova koji se preko tog linka mogu rutirati. Na osnovu ograničenja za saobraćaje čvorova u mreži, možemo proceniti propusnu moć mreže sa rutiranjem po najkraćoj putanji.

Koristićemo sledeći postupak. Prvo ćemo svakom linku u mreži pridružiti graf, čiji su čvorovi oni čvorovi u mreži koji koriste posmatrani link za međusobnu komunikaciju. Naći ćemo zatim vrednost maksimalnog toka (*maximum flow*) u tako pridruženom grafu. Ova vrednost predstavlja maksimalno opterećenje linka u slučaju rutiranja po najkraćoj putanji u mreži sa čvorovima zadate težine. Na osnovu ove vrednosti odredićemo ograničenje za garantovani saobraćaj u čvoru koje važi na posmatranom linku, a najstrožije ograničenje po svim linkovima u mreži predstavljaće konačnu vrednost za garantovani saobraćaj na nivou mreže.

U daljem tekstu, prvo ćemo opisati način na koji se linku u mreži pridružuje graf. Zatim ćemo opisati algoritam za određivanje maksimalnog toka u tom grafu. Na kraju, objasnićemo kako se dobijeni rezultat koristi za procenu propusne moći mreže sa rutiranjem po najkraćoj putanji.

8.2.1. Pridruživanje grafa linku u mreži

Da bismo odredili maksimalno opterećenje linka u mreži, neophodno je prvo tom linku pridružiti graf, u kome ćemo odrediti intenzitet maksimalnog toka. U ovom pododeljku biće objašnjeno na koji način se formira pridruženi graf za proizvoljan link u mreži.



Slika 8.2.1. Graf pridružen linku u mreži

Posmatrajmo link $l \in E$. Označimo skup parova čvorova koji koriste link l sa

$$P^l = \{(i, j) \mid F_{ij}^l = 1\} \quad (8.2.1)$$

Definišimo skup izvora koji šalju saobraćaj po linku l ,

$$I^l = \{i \mid \exists j, (i, j) \in P^l\} \quad (8.2.2)$$

kao i skup odredišta koja primaju saobraćaj preko linka l ,

$$J^l = \{j \mid \exists i, (i, j) \in P^l\} \quad (8.2.3)$$

Formiraćemo bipartitni graf u kome su skupovi čvorova I^l i J^l . Za neki graf kažemo da je bipartitan ukoliko se skup njegovih čvorova V može predstaviti kao disjunktna unija dva podskupa, L i R , tako da ne postoji grana između dva proizvoljna čvora iz istog podskupa. Neka u grafu postoji usmerena grana beskonačnog kapaciteta

između svih parova čvorova i i j , takvih da je $(i, j) \in P^l$ - Slika 8.2.1. Zatim ćemo tako formiranom grafu dodati još dva čvora: izvor, označen na Slici 8.2.1. sa x i odredište, označeno sa y . Iz izvora x ka svakom čvoru $i \in I^l$ biće usmeren link kapaciteta s_i , a od svakog čvora $j \in J^l$ ka odredištu y biće usmeren link kapaciteta s_j . Praktično, preko čvora $i \in I^l$ neće se moći poslati više nego što on generiše, niti će se preko čvora $j \in J^l$ moći primiti više nego što je to moguće u analiziranoj mreži $G=(V, E)$, pri pretpostavljenom vektoru saobraćaja $\mathbf{S}$. Maksimalno opterećenje linka l jednako je vrednosti maksimalnog toka od x ka y u pridruženom grafu. Označićemo tu vrednost sa $f_{\max}^l$, i u narednom pododeljku biće opisan algoritam za određivanje maksimalnog toka.

8.2.2. Algoritam za određivanje maksimalnog toka (*maximum flow*)

Kod problema maksimalnog toka cilj je odrediti maksimalnu vrednost saobraćaja koja u mreži može biti poslata između neka dva čvora, koje označavamo kao izvor a i odredište b . Postoji veliki broj algoritama koji se koriste za određivanje maksimalnog toka u mreži. U ovom odeljku biće opisan implementirani algoritam sa potiskivanjem pratoka (*preflow-push algorithm*).

Neka je dat graf $Q(V_Q, E_Q)$ i neka su njegovim granama e pridružene težine $u(e)$ koje odgovaraju kapacitetu grana. Neka je saobraćaj koji se šalje od a ka b granom e jednak $t(e)$, $0 \leq t(e) \leq u(e)$. Rezidualna mreža definisana je kao graf $R(V_Q, F)$ sa istim čvorovima, ali promenjenim skupom grana i težina grana. U ovome grafu, svaka grana $e=(v, w)$ sa tokom $t(e)$ zamenjena je sa dve grane: granom $e'=(v, w)$, sa težinom (kapacitetom) $r(e')=u(e)-t(e)$, i granom $e''=(w, v)$, sa težinom (kapacitetom) $r(e'')=t(e)$. Ukoliko u grafu Q već postoji i grana (w, v) , tada će se dobiti po dve paralelne grane u svakom smeru između dva čvora; njih zamenjujemo jednom granom sa kapacitetom jednakim zbiru kapaciteta dvaju paralelnih grana.

Sada definišemo i funkciju rastojanja proizvoljnog čvora mreže u odnosu na odredišni čvor b . Svakom čvoru dodeljujemo neku celobrojnu vrednost rastojanja

$d: V_Q \rightarrow \mathbb{N}^+ \cup \{0\}$. Kaže se da je funkcija rastojanja *validna* za neku rezidualnu mrežu $R(V_Q, F)$ ukoliko je ispunjeno

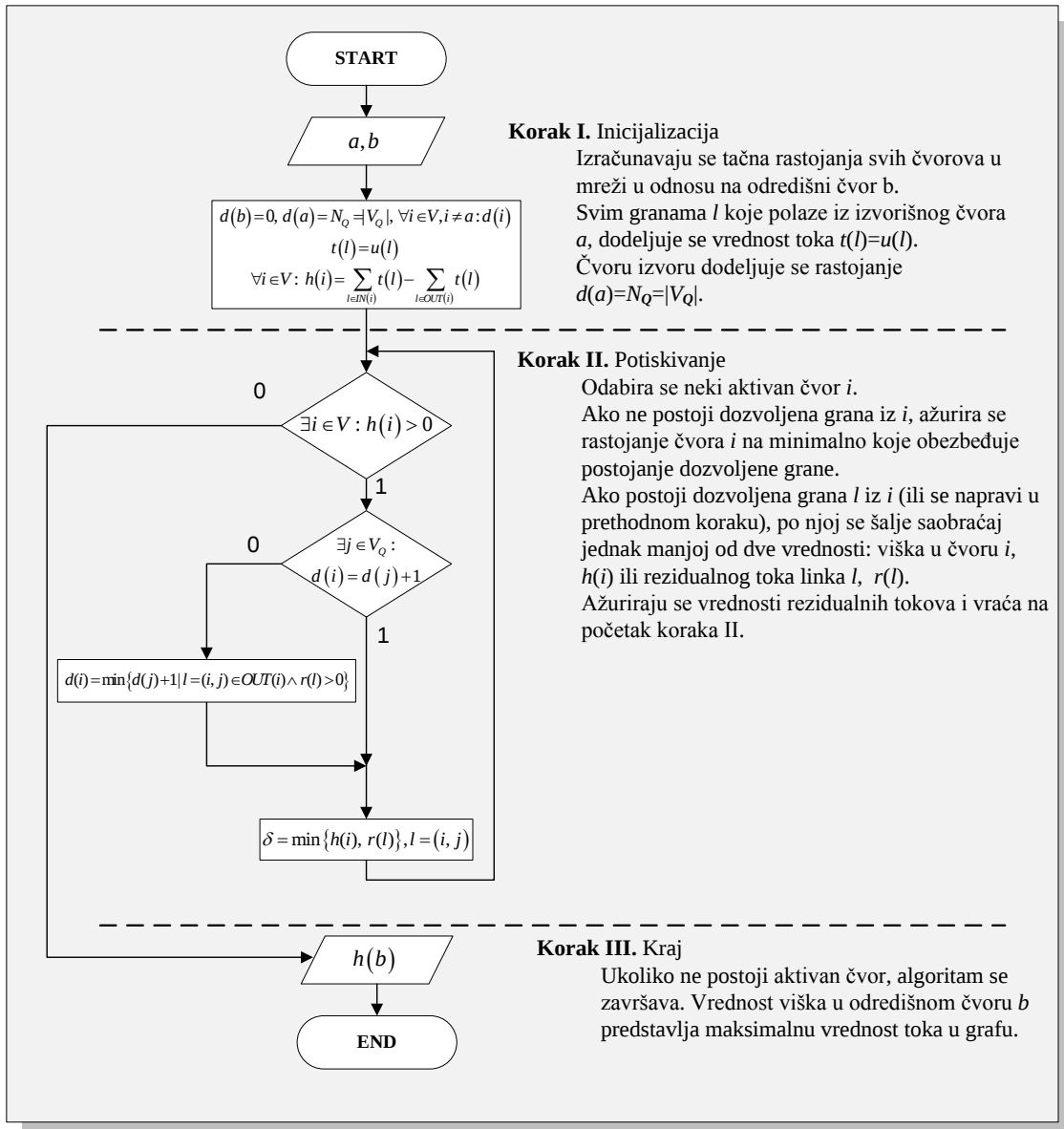
$$\begin{aligned} d(b) &= 0 \\ \forall (i, j) \in F \quad d(i) &\leq d(j) + 1. \end{aligned} \tag{8.2.4}$$

Praktično, funkcija rastojanja za odredišni čvor b mora biti jednaka nuli, dok se vrednosti funkcija rastojanja za susedne čvorove u mreži mogu razlikovati za maksimalno 1. Pri tome, funkcija rastojanja za čvor iz koga neka usmerena grana polazi ne sme biti manja od funkcije rastojanja za čvor u kom se ta grana završava.

U opštem slučaju, vrednost funkcije rastojanja za bilo koji čvor u mreži je manja ili jednaka dužini najkraće usmerene putanje od tog čvora do odredišta b . Rastojanja su *tačna* ukoliko je za svaki čvor rastojanje jednako dužini najkraće putanje od njega ka odredišnom čvoru b . Grana (i, j) u rezidualnoj mreži je *dozvoljena* (*admissible*) ako je ispunjen uslov da je $d(i) = d(j) + 1$. Putanja između izvora a i odredišta b je dozvoljena ako sadrži samo dozvoljene grane. U uobičajenoj analizi tokovi $t(e)$ u mreži moraju biti takvi da se zadovolje kako ograničenja vezana za kapacitete linkova, $0 \leq t(e) \leq u(e)$, tako i konzervacija protoka u svakom čvoru mreže, izuzev izvora i odredišta. Međutim, za potrebe pronalaženja maksimalnog toka u mreži upotrebićemo takozvani *pratok* (*preflow*), za koji konzervacija protoka u čvorovima ne mora da važi. U tom slučaju definišaćemo višak saobraćaja za čvor u mreži

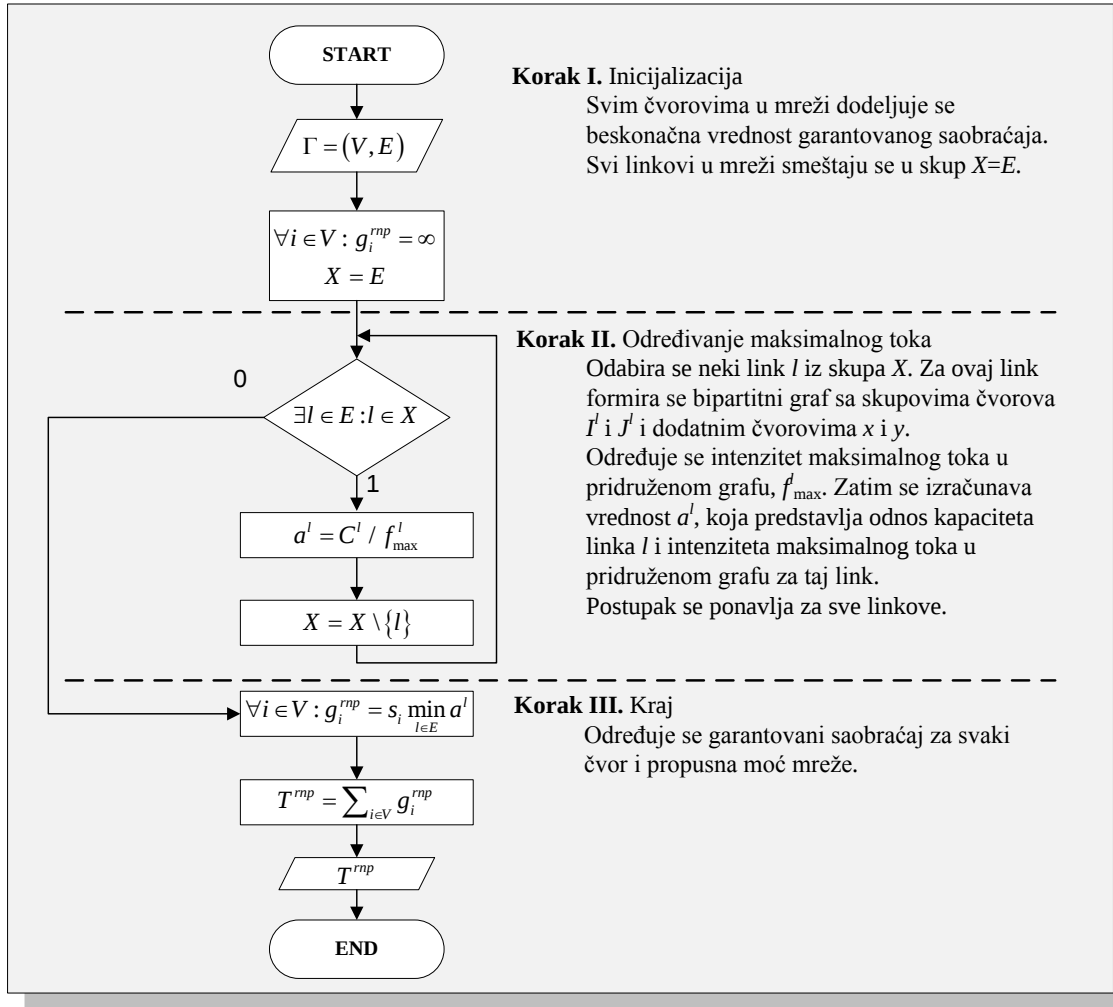
$$h(n) = \sum_{e \in IN(n)} t(e) - \sum_{e \in OUT(n)} t(e) \tag{8.2.5}$$

Jedini čvor u mreži koji može da ima negativan višak je izvorišni čvor, dok je za sve ostale čvorove $h(n) \geq 0$.



Slika 8.2.2. Algoritam za određivanje maksimalnog toka

Čvor za koji je višak pozitivan, $h(n) > 0$, naziva se aktivnim čvorom. Za potrebe algoritma smatraćemo da izvorišni i odredišni čvor ne mogu biti aktivni. Algoritam sa potiskivanjem prtoka (*preflow-push*) teži da odmah pošalje što veću ukupnu vrednost saobraćaja od izvora ka odredištu. Zato u početnoj fazi on po svim linkovima koji napuštaju izvor a šalje maksimalnu moguću vrednost saobraćaja, jednaku kapacitetu linkova. Na taj način čvorovi susedni izvoru postaju aktivni. Zatim algoritam traži aktivne čvorove i pokušava da eliminiše njihov višak saobraćaja i da ga pogura ka čvorovima koji su bliže konačnom odredištu. Algoritam je prikazan na slici 8.2.2.



Slika 8.2.3. Određivanje propusne moći mreže sa rutiranjem po najkraćoj putanji

8.2.3. Određivanje propusne moći mreže sa rutiranjem po najkraćoj putanji

Maksimalno opterećenje linka l jednako je vrednosti maksimalnog toka od x ka y u pridruženom grafu, koju smo označili sa $f_{\max}^l$. Saobraćaj svakog čvora se može povećati maksimalno

$$a^l = \frac{C^l}{f_{\max}^l} \quad (8.2.6)$$

puta, a da se ne preoptereći posmatrani link. Ograničenje za maksimalno dozvoljeno povećanje će važiti i za čvorove koji ne šalju svoj saobraćaj preko linka l , s obzirom na to da su svi tokovi u mreži proporcionalni u skladu sa zadatim težinama u vektoru ulaznog saobraćaja. Dakle, za čvor kome je dodeljena težina $s_i = w_i$, maksimalna dozvoljena

vrednost saobraćaja po linku l jednaka je $s_l a$. Najstrožije ograničenje po svim linkovima u mreži predstavlja garantovani saobraćaj u čvoru mreže,

$$g_i^{rnp} = s_i \min_{l \in E} a^l. \quad (8.2.7)$$

Propusna moć mreže sa rutiranjem po najkraćoj putanji jednaka je sumi garantovanih saobraćaja za sve čvorove:

$$T^{rnp} = \sum_{i \in V} g_i^{rnp} = \min_{l \in E} a^l \sum_{i \in V} s_i. \quad (8.2.8)$$

Algoritam za određivanje propusne moći mreže sa rutiranjem po najkraćoj putanji predstavljen je na Slici 8.2.3. Za svaki link $l \in E$ formira se bipartitni graf kao na slici 8.2.1. i određuje vrednost maksimalnog toka za taj link $f_{\max}^l$. Prema formuli (8.2.6) određuje se faktor a^l kojim se može skalirati svaki tok koji koristi link l , a da ne dođe do blokade linka. Kada su izračunate vrednosti a^l za sve linkove $l \in E$, garantovani saobraćaj svakog čvora određuje se prema formuli (8.2.7), i na osnovu njega izračunava se propusna moć mreže, kao u (8.2.8).

8.3. Poređenje propusne moći mreže za RBS i RNP

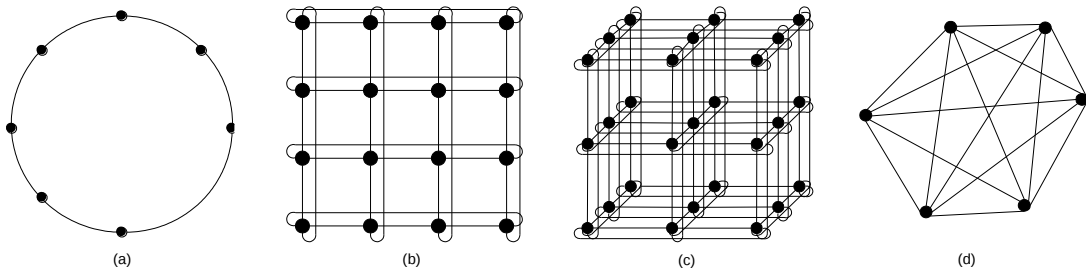
Analiziraćemo povećanje propusne moći mreže u slučaju kada je primenjeno rutiranje sa balansiranjem saobraćaja, u odnosu na slučaj kada je primenjeno regularno rutiranje po najkraćoj putanji. Ovo povećanje dato je izrazom

$$G = \frac{T^{rbs}}{T^{rnp}}, \quad (8.3.1)$$

gde su vrednosti T^{rbs} i T^{rnp} definisane izrazima (8.1.6) i (8.2.8), respektivno.

Vrednosti elemenata vektora saobraćaja $\mathbf{S}$ jednake su težinama čvorova u mreži. Čvorovima kojima je dodeljena veća težina biće garantovan i veći saobraćaj, u skladu sa jednačinama (8.1.5) za rutiranje sa balansiranjem saobraćaja, odnosno (8.2.7) za rutiranje po najkraćoj putanji.

Poređenje propusne moći vršeno je za regularne i realne mrežne topologije. Posmatrani su slučajevi jednakih težina svih čvorova, ili su čvorovima dodeljivane težine proporcionalne kapacitetima ulaznih/izlaznih linkova, odnosno broju korisnika koje čvor servisira.



Slika 8.3.1. Regularne mrežne topologije. (a) Prsten. (b) Menhetn. (c) Torus. (d) Kompletan meš.

8.3.1. Regularne mrežne topologije

Analiza propusne moći regularnih topologija vršena je u [29] i [9]. Posmatrane su sledeće topologije: prsten, Menhetn, torus i meš, predstavljene na Slici 8.3.1. Pretpostavljena je jednakost saobraćaja u svim čvorovima mreže, tj. svim čvorovima dodeljene su težine w_i jednake jedinici, pa su i elementi vektora saobraćaja jednaki jedinici, $s_i = w_i = 1$.

Usled simetrije, u posmatranim topologijama postoje višestruke putanje jednake (najkraće) dužine među parovima čvorova. Algoritam za rutiranje po najkraćoj putanji odabira jednu od njih i proglašava je jedinstvenom najkraćom putanjom. Putanja se može odabrati tako da se paketi rutiraju nekim redosledom po koordinatnim osama mreže, kao u [29]. Na primer, ako se odabere redosled osa $x \rightarrow y \rightarrow z$, prenos paketa od tačke sa koordinatama (x_0, y_0, z_0) u tačku sa koordinatama (x_1, y_1, z_1) vrši se po putanji $(x_0, y_0, z_0) \rightarrow (x_1, y_0, z_0) \rightarrow (x_1, y_1, z_0) \rightarrow (x_1, y_1, z_1)$. Takav način odabira putanje ravnomernije raspoređuje opterećenje po linkovima mreže, u odnosu na realniji slučaj u kome se putanja bira na osnovu adresa (identifikacionih brojeva) čvorova u mreži. U ovom drugom slučaju, koji je analiziran u [9], ako između čvorova i i j postoji više putanja iste dužine, prilikom odabira putanje uzimaju se u obzir identifikacioni brojevi suseda čvora i preko kojih vode moguće putanje ka j . Odabira se ona putanja koja vodi preko čvora sa najmanjim identifikacionim brojem.

Poređenje propusne moći regularnih mreža sa rutiranjem sa balansiranjem saobraćaja i rutiranjem po najkraćoj putanji vršeno je za oba tipa rutiranja po najkraćoj putanji.

i) *Rutiranje po koordinatnim osama mreže*

Kao što je već rečeno, pretpostavljamo da svi čvorovi imaju iste težine $s = s_i = w_i = 1$. Usled simetrije mreže i stanja saobraćaja u njoj, možemo smatrati da su svi koeficijenti balansiranja k_m međusobno jednaki, i iznose $k_m = 1/N$, $\forall m \in V$. Tada je, na osnovu jednačine (8.1.1), saobraćaj između svaka dva čvora u mreži jednak

$$b = b_{ij}^{(1)} + b_{ij}^{(2)} = 2s / N = 2 / N . \quad (8.3.2)$$

a) **Prsten:**

Topologija prstena predstavljena je na Slici 8.3.1.(a). Svaki čvor u mreži sa topologijom prstena ima stepen čvora $d = 2$.

Prosečna dužina putanje koju prelazi tok generisan u nekom čvoru mreže sa topologijom prstena iznosi

$$l_{av} = \sum_{i=1}^{\lfloor N/2 \rfloor} \frac{2i}{N} \approx \frac{N}{4} \quad (8.3.3)$$

Najduža putanja između dva čvora u mreži sa topologijom prstena ima dužinu $\lfloor N/2 \rfloor$. Pošto je ukupan broj usmerenih linkova u prstenu jednak $2N$, a postoji N^2 parova čvorova, očekivano opterećenje linka u mreži u slučaju kada je primenjeno balansiranje saobraćaja iznosi

$$L^l = \frac{l_{av}}{2N} b N^2 \approx \frac{N}{4} . \quad (8.3.4)$$

Pošto su svi linkovi isto opterećeni, to je $Q = L / C = N / 4C$, a garantovani saobraćaj u čvoru mreže jednak je

$$g_i^{rbs} = s_i / Q = 1 / Q = 4C / N , \quad (8.3.5)$$

dok je propusna moć mreže jednaka sumi garantovanih saobraćaja svih čvorova

$$T^{rbs} = N (4C / N) = 4C . \quad (8.3.6)$$

U slučaju običnog rutiranja po najkraćoj putanji, najgori slučaj je kad međusobno komuniciraju najudaljeniji čvorovi, jer tad saobraćaj iz jednog čvora opterećuje najveći broj linkova. Pošto se neki link može naći na putanji od i ka j ili obrnuto, ali ne može u isto vreme biti na obe putanje, to će u ovom najgorem slučaju po linku biti rutiran saobraćaj iz $\lfloor N/2 \rfloor$ čvorova. Samim tim, biće $Q = L / C \approx N / 2C$, pa garantovani saobraćaj čvora u ovom slučaju iznosi

$$g_i^{rnp} = 1/Q = 2C/N, \quad (8.3.7)$$

a propusna moć mreže

$$T^{rbs} = N(2C/N) = 2C \quad (8.3.8)$$

Dakle, dobitak pri primeni balansiranja saobraćaja iznosi

$$G = 2. \quad (8.3.9)$$

b) Menhetn:

Mreža sa Menhetn topologijom predstavljena je na slici 8.3.1.(b). Ona ima formu dvodimenzionalne cirkularno povezane rešetke, tako da svaki čvor ima stepen $d = 4$. U slučaju simetričnog Menhetna, svaka stranica mreže ima dimenziju D , pa važi da je $N = D^2$.

Prosečna dužina putanje toka generisanog u nekom čvoru rešetke u svakom od pravaca iznosi $l_{av} \approx D/4$. Naime, i u horizontalnom i u vertikalnom pravcu posmatrano, čvor pripada prstenima dužine D , pa ovaj rezultat dobijamo zamenom N sa D u jednačini (8.3.3). Ukupan broj horizontalnih, odnosno vertikalnih linkova iznosi $2N$, tako da je očekivano opterećenje linka kada je primenjeno balansiranje saobraćaja jednako

$$L^l = \frac{l_{av}}{2N} bN^2 \approx \frac{D}{4}. \quad (8.3.10)$$

Odavde dobijamo da je $Q = D/4C$, pa je garantovani saobraćaj u čvoru mreže

$$g_i^{rbs} = s_i / Q = 1/Q = 4C/D. \quad (8.3.11)$$

Propusna moć mreže data je izrazom

$$T^{rbs} = N(4C/D) = 4CD. \quad (8.3.12)$$

U slučaju običnog rutiranja po najkraćoj putanji, najgori slučaj je kada komuniciraju čvorovi na međusobnoj udaljenosti D , što je i maksimalna udaljenost dva čvora u mreži. Tada po jednom linku prolazi $\lfloor D/2 \rfloor$ tokova, zagušenje je $Q = L/C \approx D/2C$ pa je garantovani saobraćaj čvora

$$g_i^{rnp} = 2C/D, \quad (8.3.13)$$

dok je propusna moć mreže

$$T^{rbs} = N(2C/D) = 2CD. \quad (8.3.14)$$

Drugim rečima, i u ovom slučaju ostvaruje se poboljšanje garantovanog saobraćaja

$$G = 2. \quad (8.3.15)$$

c) Torus:

Mreža sa topologijom torusa predstavljena je na slici 8.3.1.(c), i ima oblik trodimenzionalne cirkularno povezane rešetke. Svaki čvor ove mreže ima stepen $d = 6$. Analiziramo slučaj simetričnog torusa u kome je dimenzija sve tri stranice rešetke jednaka D a ukupan broj čvorova $N = D^3$.

Neka je redosled rutiranja po osama $x \rightarrow y \rightarrow z$. Slično kao i kod Menhetn mreže, možemo zaključiti da je prosečna dužina putanje u svakom pravcu jednaka $l_{av} \approx D/4$. Očekivano opterećenje linka u slučaju kada je primenjeno balansiranje saobraćaja jednako je:

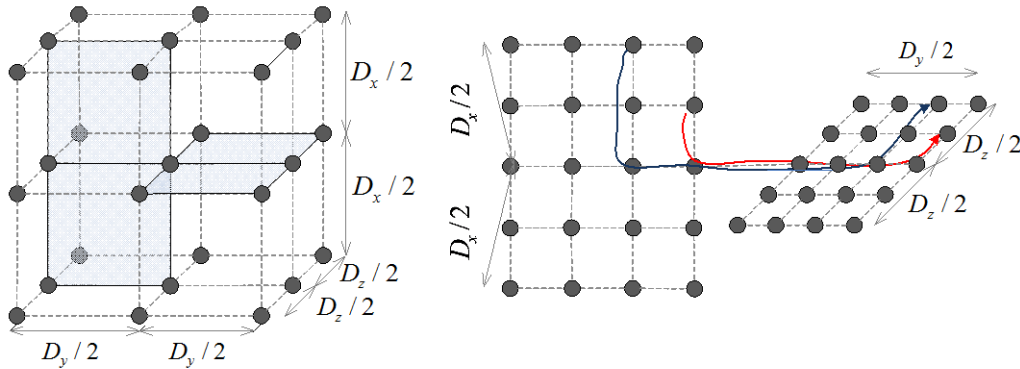
$$L^l = \frac{l_{av}}{2N} bN^2 \approx \frac{D}{4}. \quad (8.3.16)$$

Odavde možemo zaključiti da je $Q = L/C \approx D/4C$ garantovani saobraćaj čvora kada je primenjeno balansiranje saobraćaja jednak

$$g_i^{rbs} = 4C/D, \quad (8.3.17)$$

dok je propusna moć mreže

$$T^{rbs} = N(4C/D) = 4ND^2. \quad (8.3.18)$$



Slika 8.3.2. Najgori slučaj za rutiranje po najkraćoj putanji u torusnoj mreži

U najgorem slučaju za regularno rutiranje po najkraćoj putanji, kada komuniciraju parovi međusobno najudaljenijih čvorova, najveće opterećenje je na linkovima na y osi. U ovom slučaju preko svakog od linkova na y osi komunicira po $D^2/2$ parova., kao na

slici 8.3.2. Odatle imamo da je $Q = L / C \approx D^2 / 2C$, pa je garantovani saobraćaj u čvoru mreže

$$g_i^{mp} = 2C / D^2, \quad (8.3.19)$$

a propusna moć mreže sa rutiranjem po najkraćoj putanji iznosi

$$T^{mp} = N(2C / D^2) = 2CD \quad (8.3.20)$$

Dobitak u ovom slučaju iznosi

$$G = 2D. \quad (8.3.21)$$

d) Kompletan meš:

U kompletnom mešu svaki čvor mreže povezan je sa svim ostalim čvorovima u mreži, kao na slici 8.3.1.(d), i ima stepen $D = N - 1$. Po svakom linku komunicira samo jedan par čvorova.

Kada je primenjeno rutiranje sa balansiranjem saobraćaja, opterećenje linka sastoji se iz dve komponente, kao u (8.3.2), i iznosi

$$L^l = \frac{2}{N}. \quad (8.3.22)$$

Stoga je zagušenje jednako $Q = NC / 2$, pa je garantovani saobraćaj čvora mreže jednak

$$g_i^{rbs} = CN / 2, \quad (8.3.23)$$

a propusna moć mreže

$$T^{rbs} = CN^2 / 2. \quad (8.3.24)$$

U slučaju regularnog rutiranja po najkraćoj putanji, s obzirom na to da je samo jedan tok rutiran po jednom linku, imamo da je garantovana vrednost saobraćaja

$$g_i^{rnp} = C, \quad (8.3.25)$$

a propusna moć mreže

$$T^{rbs} = CN \quad (8.3.26)$$

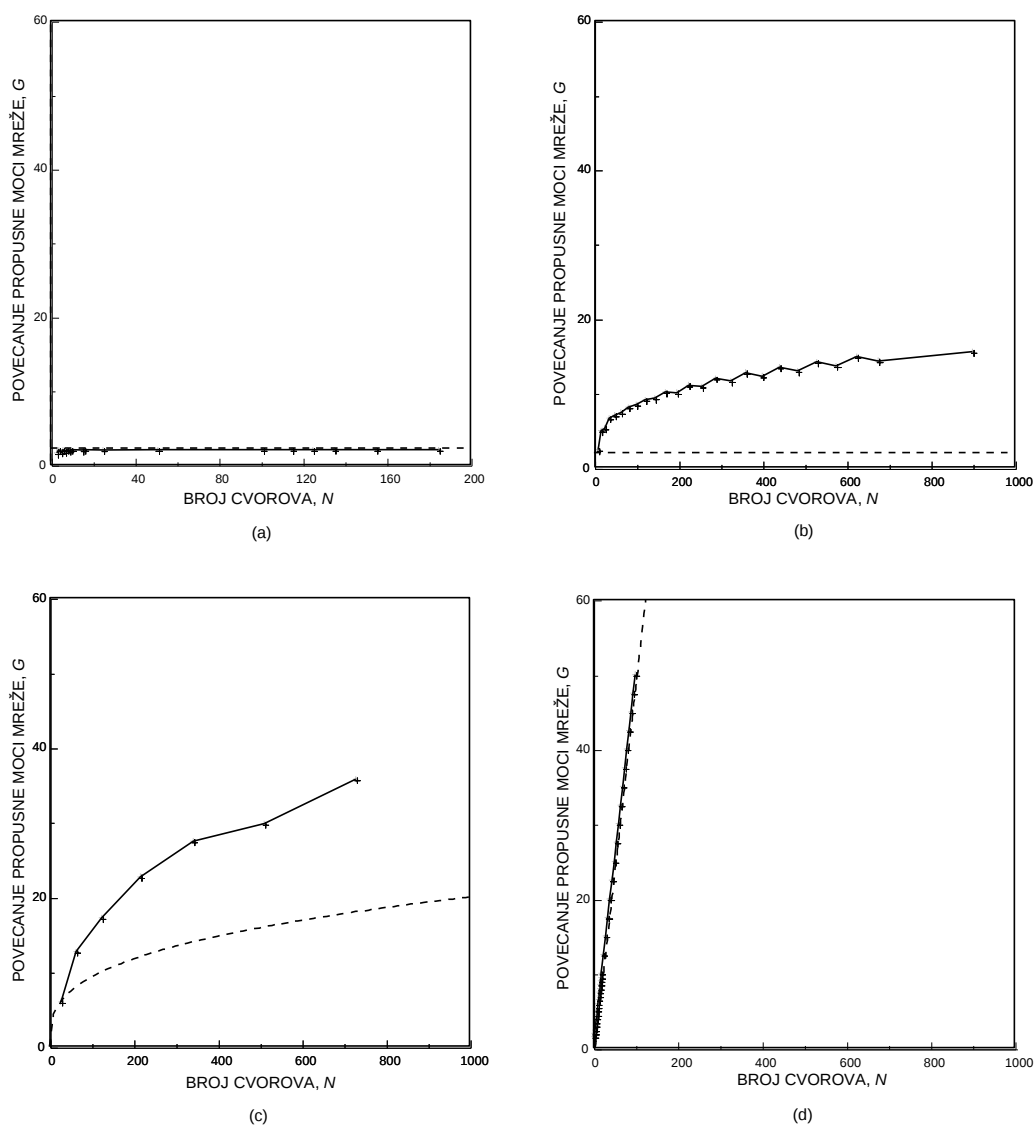
Dakle, povećanje propusne moći mreže iznosi

$$G = N / 2. \quad (8.3.27)$$

ii) Rutiranje na osnovu numeracije čvorova

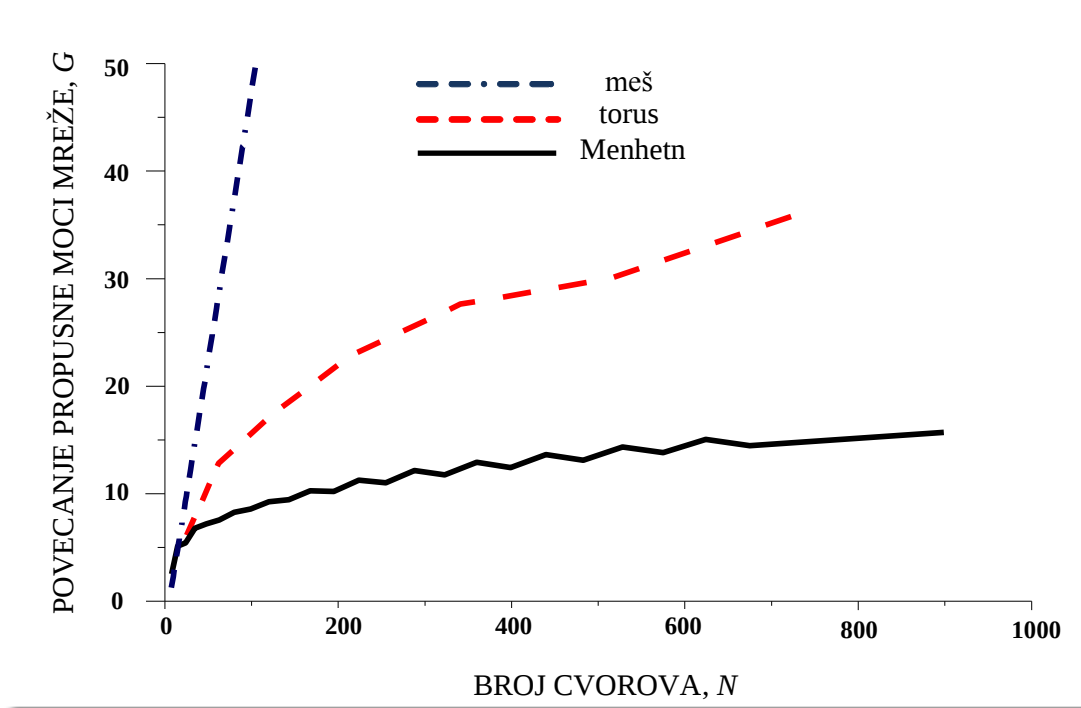
U praksi se ne vodi računa o uravnoteženom rasporedu opterećenja među linkovima u mreži, već se rutiranje vrši na osnovu identifikacionog broja čvorova u mreži – među putanjama iste dužine iz nekog čvora, prednost se daje onoj putanji koja ide preko suseda sa manjim identifikacionim brojem. Ovaj slučaj analiziran je u [9]. Za slučaj kada se koristi balansiranje saobraćaja određivano je optimalno rutiranje prema postupku

izloženom u odeljku 8.1 i propusna moć mreže prema jednačini (8.1.6). Propusna moć mreže bez balansiranja određivana je prema postupku izloženom u odeljku 8.2. i prema jednačini (8.2.8). Povećanje propusne moći koje se u ovom slučaju ostvaruje primenom balansiranja saobraćaja veće je nego u prethodno analiziranom slučaju, u kome se najkraća putanja odabira tako da se ide redom po osama mreže, što se može videti na slici 8.3.3. Na njoj su isprekidanim linijama prikazana povećanja propusne moći mreže u slučajevima kada se rutiranje po najkraćoj putanji vrši ravnomerno po osama regularne topologije, dok je punom linijom prikazano povećanje propusne moći ukoliko primenjeno rutiranje po najkraćoj putanji zavisi od numeracije čvorova u mreži.



Slika 8.3.3. Povećanje propusne moći za različite topologije. (a) Prsten. (b) Menhetn. (c) Torus. (d) Potpuni meš.

Zavisnost povećanja propusne moći mreže za topologije Menhetn, torus i meš od broja čvorova mreže prikazana je na slici 8.3.3. Može se uočiti da povećanje propusne moći raste sa porastom stepena čvora u mreži. Drugim rečima, mreže sa više alternativnih putanja među čvorovima imaju veću korist od primene balansiranja saobraćaja prilikom rutiranja.



Slika 8.3.4. Povećanje garantovanog saobraćaja za regularne mrežne topologije i slučaj rutiranja po najkraćoj putanji na osnovu numeracije čvorova.

8.3.2. Realne mrežne topologije

Analiza povećanja propusne moći mreže vršena je i za realne topologije, publikovane u okviru *Rocketfuel* projekta [30]. Ovaj projekat bavio se procenom realnih topologija najvećih Internet provajdera u SAD, Evropi i Australiji, na osnovu informacija do kojih se moglo doći praćenjem ICMP poruka. Kao rezultat dati su detaljni podaci za šest velikih provajdera. Originalni podaci sadrže listu linkova u formi uređenih trojki: (*ruter_1*, *ruter_2*, *težina_linka*). Pošto podaci ne sadrže informaciju o kapacitetu linkova, pretpostavljeno je da je kapacitet inverzno proporcionalan težini linka.

Primenjena je linearna optimizacija u cilju određivanja optimalnih koeficijenata k_m i maksimizacije propusne moći mreže. Dimenzije programa za optimizaciju rutiranja u mrežama sa balansiranjem saobraćaja date su u Tabeli 8.3.1. Zatim je izvršena procena

propusne moći mreže u slučaju rutiranja po najkraćoj putanji. Ove vrednosti su upoređene i izračunat dobitak, kao u jednačini (8.3.1). Posmatrana su tri slučaja težina čvorova u mreži. U daljem tekstu biće objašnjena sva tri slučaja dodeljivanih težina.

i) *Jednaki saobraćaji svih čvorova*

U prvom slučaju, smatrano je da su svi čvorovi međusobno jednaki, odnosno da je $s_i = w_i = 1$ za sve čvorove mreže. U ovom slučaju za povećanje garantovanog saobraćaja dobijeni su rezultati označeni sa G_1 u Tabeli 8.3.2. Povećanje se kreće od 2 do 7.5 u zavisnosti od topologije posmatrane mreže, i veće je za topologije sa većim prosečnim stepenom čvora. Vreme neophodno za optimizaciju rutiranja sa balansiranjem saobraćaja označeno je u Tabeli 8.3.2. sa t_1 .

Tabela 8.3.1. Dimenzija optimizacionog problema za realne mrežne topologije

MREŽA	Broj čvorova	Broj linkova	DIMENZIJA LP MODELA	
			broj promenljivih	broj ograničenja
Exodus (US) 3967	79	294	80	282
Ebone (Europe) 1755	87	322	88	316
Telstra (Australia) 1221	104	302	105	304
Abovenet (US) 6461	138	744	139	720
Tiscali (Europe) 3257	161	656	162	629
Sprintlink (US) 1239	315	1944	316	1923

Tabela 8.3.2. Rezultati analize za realne mrežne topologije

MREŽA	G_1	$t_1[s]$	G_2	$t_2[s]$	G_3	$t_3[s]$
Exodus (US) 3967	6.057	2.812	6.018	2.766	5.744	2.875
Ebone (Europe) 1755	2.012	3.438	2.023	3.375	3.358	3.563
Telstra (Australia) 1221	2.141	4.078	2.243	3.531	2.367	4.125
Abovenet (US) 6461	4.576	10.953	4.932	9.360	3.040	9.219
Tiscali (Europe) 3257	7.217	13.625	5.953	15.922	7.998	13.375
Sprintlink (US) 1239	7.548	56.047	7.688	60.359	5.567	61.422

ii) *Saobraćaj proporcionalan kapacitetu izlaznih linkova*

U drugom slučaju, smatrano je da je težina nekog čvora proporcionalna ukupnom kapacitetu izlaznih linkova za taj čvor. Označimo ukupan kapacitet izlaznih linkova za čvor i sa

$$K_i^{out} = \sum_{l \in OUT(i)} C^l \quad (8.3.28)$$

Sada ćemo smatrati da je

$$s_i = w_i = aK_i^{out} \quad (8.3.29)$$

gde je a neka konstanta. Povećanje garantovanog saobraćaja u slučaju ovako pretpostavljenih odnosa za saobraćaje čvorova označeno je sa G_2 u Tabeli 8.3.1, i kreće se od 2 do 7.7. Kao i u slučaju kad smo smatrali da svi čvorovi generišu iste saobraćaje, i ovde je povećanje veće za razgranatije topologije. Vreme neophodno za optimizaciju označeno je sa t_2 u istoj tabeli.

iii) *Saobraćaj proporcionalan broju stanovnika*

U trećem slučaju smatrano je da je saobraćaj u čvoru proporcionalan broju korisnika koje neki čvor treba da usluži. Neka je u nekom mestu ukupan broj stanovnika P , a broj čvorova posmatrane mreže u istom mestu x . Smatra se da se saobraćaj ravnomerno raspoređuje među tim čvorovima, tj. da na svaki čvor i koji se nalazi u posmatranom mestu dolazi $w'_i = P/x$ stanovnika, odnosno proporcionalan broj korisnika. Sada ćemo smatrati da su vrednosti elemenata u inicijalnom vektoru izlaznog saobraćaja date sa (8.3.4), gde $\lfloor w'_i/100000 \rfloor$ označava ceo deo od $w'_i/100000$.

$$s_i = w_i = \begin{cases} 1, & w'_i \leq 100000 \\ \lfloor w'_i/100000 \rfloor, & u \leq w'_i/100000 \leq u + 0.5, u \in N \\ \lfloor w'_i/100000 \rfloor + 1, & u + 0.5 \leq w'_i/100000 \leq u + 1, u \in N \end{cases} \quad (8.3.30)$$

Rezultati za ovaj slučaj označeni su sa G_3 u Tabeli 8.4.1, a vreme za optimizaciju sa t_3 . Povećanje garantovanog saobraćaja kreće se od 2.4 za mrežu Telstra, koja ima topologiju sličnu prstenu, do skoro 8 u mrežama sa više putanja između čvorova.

Možemo zaključiti da je za sve analizirane slučajeve ostvareno značajno povećanje propusne moći mreže, uz prihvatljivo vreme za optimizaciju rutiranja. Povećanje propusne moći mreže raste sa stepenom čvora u mreži, tj. veće je u mrežama koje imaju razgranatiju topologiju.

9. MINIMIZACIJA CENE MREŽE

U slučaju koji je razmatran u prethodnom poglavlju, poznata je bila mrežna topologija. Dakle, bio je poznat raspored čvorova i linkova, kao i kapaciteti linkova koji su na raspolaganju. Cilj je bila optimizacija rutiranja kako bi se ostvarila maksimalna propusna moć tako zadate mreže.

Posmatrajmo sada nešto drugačiji problem. Pretpostavimo da su nam poznati raspored čvorova i linkova u mreži, ali da ne postoji ograničenje po pitanju kapaciteta pojedinih linkova. Ovakva pretpostavka je opravdana usled činjenice da se danas na važnijim rutama uglavnom polaže veliki broj optičkih vlakana, od kojih svako obezbeđuje terabitske kapacitete, tako da možemo smatrati da je kapacitet linkova koji imamo na raspolaganju praktično neograničen. Želimo sada da projektujemo mrežu koja je u stanju da podrži zadate saobraćajne zahteve, a da cena te mreže bude što niža.

U ovom poglavlju, biće opisano način na koji se rutiranje sa balansiranjem saobraćaja optimizuje tako da se ostvari minimalna cena mreže koja može da podrži zadati saobraćaj. Zatim će biti izvršena procena cene mreže sa rutiranjem po najkraćoj putanji, i upoređene cene mreže za dva slučaja rutiranja. Analiza će biti vršena za realne mrežne topologije publikovane u okviru *Rocketfuel* projekta.

9.1. Cena mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)

U opštem slučaju, kao što je to već objašnjeno u Odeljku 3.2, cena mreže zavisi od cene čvorova i cene linkova:

$$C = \sum_{l \in E} C_L^l + \sum_{n \in V} C_N^n. \quad (9.1.1)$$

U optimizaciji koju vršimo polazimo, kao što je rečeno, od pretpostavke da već imamo poznat raspored linkova, ali ne i raspodelu kapaciteta. Pretpostavljamo da tamo gde postoji link, imamo neograničene kapacitete na raspolaganju, te je cena postavljanja dodatnih optičkih kablova na toj trasi jednaka nuli. Samim tim, cena mreže jednaka je sumi cena čvorova u njoj.

$$C = \sum_{n \in V} C_N^n \quad (9.1.2)$$

Sa druge strane, može se smatrati da je cena čvora proporcionalna ceni njegovih portova. Cena porta rutera direktno zavisi od brzine porta, koja mora biti takva da podrži očekivani saobraćaj na linku vezanom za taj port. Dakle, portovi moraju imati brzine koje su veće ili jednake kapacitetu linkova koji su na njih povezani. Stoga možemo smatrati da je cena čvora jednaka sumi kapaciteta njegovih izlaznih linkova:

$$C_N^n = \sum_{l \in OUT(n)} C^l, \quad (9.1.3)$$

pa je cena mreže jednaka sumi kapaciteta svih linkova u njoj.

$$C = \sum_{l \in E} C^l \quad (9.1.4)$$

Pogledajmo kako se vrši optimizacija rutiranja sa balansiranjem saobraćaja u cilju minimizacije cene mreže. Za razliku od slučaja u kome je vršena maksimizacija propusne moći mreže, u ovom slučaju nepoznate promenljive u linearnom programu su i koeficijenti balansiranja k_m , kao i kapaciteti linkova C^l . One nisu međusobno nezavisne, jer od koeficijenata balansiranja zavisi opterećenje linkova u mreži, a od opterećenja linkova zavisi i dodeljeni kapacitet. Koeficijente balansiranja k_m treba odrediti tako da se minimizira ukupna suma dodeljenih kapaciteta linkova C^l , a da se istovremeno propusti zadati saobraćaj. U opštem slučaju, linearni program za optimizaciju rutiranja ima sledeći oblik:

$$\begin{aligned} & \min \sum_{l \in E} C^l \\ & (C1) \quad \sum_{i=1}^N k_i = 1 \\ & (C2) \quad \forall l \in E: \sum_{(i,m)} F_{im}^l (k_i r_m + k_m s_i) \leq C^l \\ & (C3) \quad \forall n \in V: \sum_{l \in IN(n)} L^l - \sum_{l \in OUT(n)} L^l = r_n - s_n. \end{aligned} \quad (9.1.5)$$

Pod uslovom da su ukupan saobraćaj generisan i primljen u čvoru mreže jednaki, $r_i = s_i$ za svaki čvor i , ograničenja (C3) o konzervaciji protoka mogu se ukloniti iz linearnog programa (lema 7.3.1), tako da on u konačnom obliku ima $M + 1$ ograničenja, gde je M broj linkova u mreži. Finalni oblik linearnog programa koji ćemo rešavati dat je jednačinom (9.2.1). Za razliku od linearnog programa za određivanje zagušenja, u kome su vektorima ulaznog i izlaznog saobraćaja zadavani međusobni odnosi saobraćaja generisanih/primljenih u čvorovima mreže, a ne i konkretne vrednosti saobraćaja, ovde

vrednosti s_i i r_i odgovaraju konkretnim vrednostima saobraćaja koje treba propustiti kroz mrežu. Takođe, pošto su kapaciteti linkova promenljive u ovom linearnom programu, ukupan broj promenljivih jednak je $M + N$.

$$\begin{aligned} & \min \sum_{l \in E} C^l \\ & (C1) \sum_{i=1}^N k_i = 1 \\ & (C2) \forall l \in E : \sum_{(i,m)} F_{im}^l (k_i r_m + k_m s_i) \leq C^l. \end{aligned} \quad (9.1.6)$$

Rezultat izvršavanja linearnog programa (9.1.2) biće optimalni kapaciteti linkova i koeficijenti rutiranja. Suma optimalnih kapaciteta biće minimalna, i možemo je predstaviti jednačinom

$$C^{rbs} = \min_{\mathbf{k} | \sum_{i=1}^N k_i = 1} \sum_{l \in E} C^l. \quad (9.1.7)$$

9.2. Cena mreže sa rutiranjem po najkraćoj putanji (RNP)

Za određivanje cene mreže sa rutiranjem po najkraćoj putanji koristimo postupak sličan onom koji je opisan u odeljku 8.2. I u slučaju procene cene mreže, kao i u slučaju procene propusne moći mreže, interesuje nas najgori slučaj za svaki link l u mreži. Naime, u najgorem slučaju, kada su aktivni svi čvorovi koji preko posmatranog linka šalju ili primaju saobraćaj, opterećenje linka je maksimalno. Linku mora biti dodeljen kapacitet koji odgovara ovom opterećenju, ukoliko se želi omogućiti rutiranje zadatog saobraćaja u svim okolnostima. Kao i u slučaju procene propusne moći mreže, najveće opterećenje linka jednako je intenzitetu maksimalnog toka u pridruženom bipartitnom grafu.

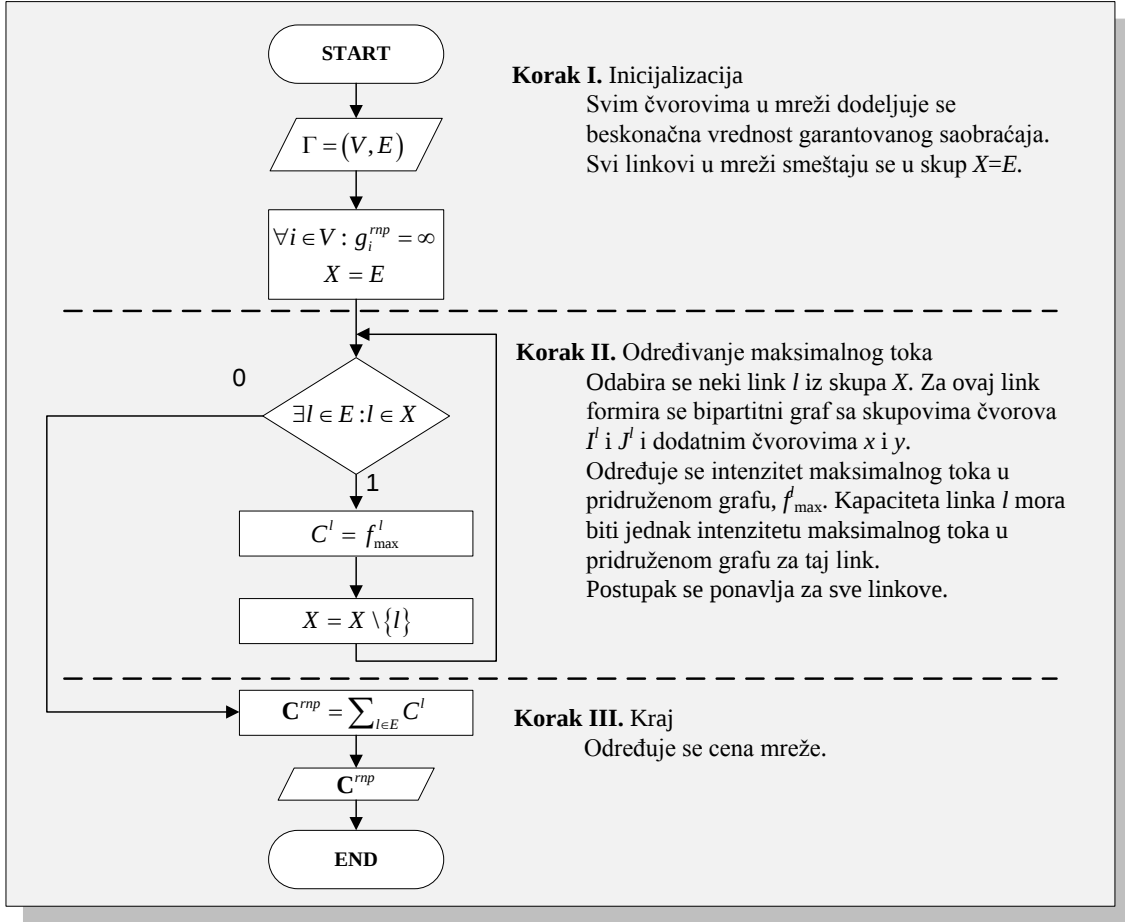
Minimalan kapacitet linka l koji može da podrži rutiranje saobraćaja datog vektorom $\mathbf{S}$ jednak je vrednosti maksimalnog toka u pridruženom bipartitnom grafu:

$$C^l = f_{\max}^l. \quad (9.2.1)$$

Cena mreže sa rutiranjem po najkraćoj putanji data je kao suma minimalnih kapaciteta linkova:

$$C^{rnp} = \sum_{l \in E} C^l = \sum_{l \in E} f_{\max}^l. \quad (9.2.2)$$

Algoritam za određivanje cene mreže sa rutiranjem po najkraćoj putanji prikazan je na Slici 9.2.1.



Slika 9.2.1. Određivanje cene mreže sa rutiranjem po najkraćoj putanji

9.3. Poređenje cene mreže za RBS i RNP

Ušteda koja se ostvaruje upotrebom rutiranja sa balansiranjem saobraćaja data je odnosom cena mreže sa rutiranjem po najkraćoj putanji i mreže sa rutiranjem sa balansiranjem saobraćaja.

$$G = \frac{C^{mp}}{C^{rbs}}. \quad (9.3.3)$$

Cena mreže sa u kojoj je primenjeno rutiranje sa balansiranjem saobraćaja je G puta niža od cene mreže u kojoj je primenjeno obično rutiranje po najkraćoj putanji.

Analiza je vršena za realne mrežne topologije iz *Rocketfuel* projekta. Iz raspoloživih podataka, preuzet je raspored čvorova i postojećih linkova među njima. Prilikom proračuna najkraćih putanja Dijkstra algoritmom, pretpostavljeno je da su svi linkovi jednake težine. Na taj način se već i prilikom određivanja najkraće putanje za saobraćaj vrši minimizacija broja linkova preko kojih se saobraćaj prenosi, a samim tim se minimizira i cena mreže u kojoj je primenjeno rutiranje po najkraćoj putanji.

Pretpostavljeno je da su saobraćajni zahtevi u čvorovima mreže proporcionalni broju korisnika koje treba servisirati. Vrednosti za poboljšanje koje se ostvaruje primenom rutiranja sa balansiranjem saobraćaja date su u Tabeli 7.3.1. Poboljšanje se kreće od 1.82 do 3.39. Kao i u slučaju optimizacije zagušenja, najlošiji rezultati postižu se za australijsku mrežu Telstra.

Poboljšanje je nešto niže nego u slučaju garantovanog saobraćaja u čvoru. Razlika potiče od načina na koji se računa garantovani saobraćaj za rutiranje po najkraćoj putanji – pri toj analizi posmatra se najopterećeniji link, koji najstrožije ograničava garantovani saobraćaj. U ovome slučaju, pak, računamo ukupan neophodan kapacitet svih linkova, a u proseku ostali linkovi i ne moraju biti toliko opterećeni.

Tabela 9.3.1. Rezultati analize za realne mrežne topologije u slučaju saobraćaja proporcionalnih broju stanovnika

MREŽA	Broj čvorova	Broj linkova	G	DIMENZIJA LP MODELA	
				broj promenljivih	broj ograničenja
Exodus (US) 3967	79	294	2.11	373	295
Ebone (Europe) 1755	87	322	2.07	409	323
Telstra (Australia) 1221	104	302	1.82	406	303
Abovenet (US) 6461	138	744	3.19	882	745
Tiscali (Europe) 3257	161	656	2.90	817	657
Sprintlink (US) 1239	315	1944	3.39	2259	1945

U Tabeli 9.3.2. data su vremena neophodna za izvršavanje linearnog programa za optimizaciju rutiranja.

Tabela 9.3.2. Vreme za optimizaciju u slučaju saobraćaja proporcionalnih broju stanovnika

Mreža	VREME [s]			
	load data	presolve	solver	total
Exodus (US) 3967	0.656	0.016	0.281	0.953
Ebone (Europe) 1755	0.937	0.016	0.297	1.250
Telstra (Australia) 1221	1.546	0.032	0.203	1.781
Abovenet (US) 6461	3.297	0.047	2.250	5.594
Tiscali (Europe) 3257	5.265	0.063	1.672	7.000
Sprintlink (US) 1239	37.313	0.250	20.218	57.781

10. OPTIMIZACIJA U POUZDANIM MREŽAMA

Do sada smo razmatrali mreže u kojima ne dolazi do kvarova resursa. Međutim, u praksi se može dogoditi da dođe do otkaza nekog od linkova ili rutera u mreži, čemu se dinamičko rutiranje prilagođava. Mogućnost da do kvara resursa dođe utiče kako na problem određivanja propusne moći mreže, tako i na problem minimizacije cene mreže.

Posmatrajmo prvo propusnu moć mreže u kojoj su mogući kvarovi resursa. Očigledno, mogućnost da u nekom trenutku ne budu svi mrežni resursi na raspolaganju ograničava saobraćaj koji se kroz postojeću mrežu garantovano može rutirati, te smanjuje propusnu moć mreže. Sa druge strane, kod određivanja kapaciteta linkova tako da cena mreže bude što manja, neophodno je dodeliti više kapaciteta nekim linkovima u mreži, kako bi se ostavila rezerva za slučajeve u kojima dođe do kvara. Samim tim, povećava se cena mreže.

U ovom pododeljku prvo ćemo definisati pojam pouzdanog komunikacionog sistema, a zatim predstaviti matematički model koji se koristi za optimizaciju rutiranja. Zatim ćemo izvršiti poređenje performansi rutiranja sa balansiranjem saobraćaja sa performansama rutiranja po najkraćoj putanji.

10.1. Pouzdan komunikacioni sistem

Pouzdanim komunikacionim sistemom smatra se onaj koji može da obezbedi povezanost i komunikaciju krajnjih korisnika i u slučajevima kada dođe do kvara pojedinačnih resursa. U slučaju IP mreža koje razmatramo, može doći do kvara na linkovima ili na ruterima u mreži. Smatraćemo da je situacija u kojoj postoje višestruki kvarovi malo verovatna, odnosno da u nekom trenutku u kvaru može biti samo jedan link ili jedan ruter.

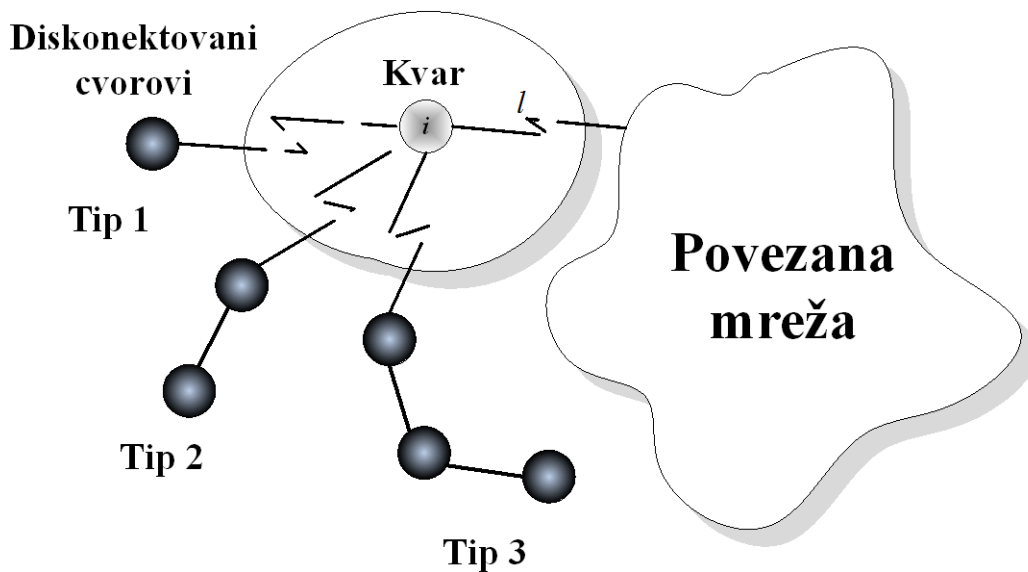
Da bi se omogućila komunikacija korisnika i u slučaju kvara nekog od resursa, potrebno je ili dodatno ograničiti saobraćaj koji oni generišu (u slučaju kada nam je poznata mrežna topologija i raspoloživi kapaciteti linkova) ili obezbediti dodatni kapacitet na linkovima, kako bi oni mogli podržati očekivani saobraćaj i u slučaju kvara (u slučaju projektovanja mreže koja treba da obezbedi rutiranje zadatog saobraćaja).

Prvo uvodimo notaciju za označavanje različitih stanja mreže. Stanje mreže definisano je funkcionalnošću linkova i rutera u njoj. Razlikujemo stanje potpune funkcionalnosti svih resursa – bez kvarova, kao i različita stanja u kojima je neki od resursa u kvaru. Označimo sa $t \in \{0,1\}$ tip kvara, tako da vrednost $t=0$ označava kvar rutera, a $t=1$ kvar linka. Takođe, obeležimo sa r resurs (ruter ili link) koji je u kvaru. Neka vrednost $r=0$ označava slučaj u kome nema kvara u mreži, odnosno kada je kompletna mreža funkcionalna. Dakle, u slučaju kvara na ruterima (čvorovima), vrednost indikatora kvara je $t=0$, dok indikator r može uzeti neku od vrednosti $r \in R_0 = \{0,1,\dots,N_0\}$, gde je N_0 broj čvorova u mreži bez kvara. Analogno, u slučaju kvara linka, imamo da je $t=1$, a $r \in R_1 = \{0,1,\dots,M_0\}$, gde je M_0 broj linkova u početnoj mreži bez kvarova.

Za svako od mogućih stanja mreže (bilo stanje potpune funkcionalnosti svih resursa, bilo stanje u kom je neki resurs u kvaru) mora se rešiti odgovarajući optimizacioni problem. Najstrožije ograničenje po svim mogućim stanjima predstavljaće konačnu vrednost promenljive koja nas interesuje. U zavisnosti od prirode optimizacionog problema, to može biti garantovani saobraćaj čvora, ili neophodan kapacitet linka u mreži.

U radu ćemo analizu vršiti na regularnim topologijama, kao i na realnim mrežnim topologijama publikovanim u okviru *Rocketfuel* projekta [30]. Iako pretpostavljamo da u nekom trenutku može doći samo do pojedinačnog kvara u mreži, u analiziranim realnim mrežama takav kvar može prouzrokovati diskonekciju jednog ili više čvorova, koji su preko resursa koji je u kvaru (rutera ili linka) povezani sa ostatkom mreže, kao što je to ilustrovano na Slici 10.1.1. U tom slučaju, moguće je ostvariti komunikaciju samo među korisnicima iz preostalog povezanog dela mreže. Ovaj deo mreže, koji ostaje povezan i nakon nekog od kvarova, označićemo sa $\Gamma(r,t) = (V(r,t), E(r,t))$. Ovakva mreža sada ima $N(r,t)$ čvorova i $M(r,t)$ linkova. Mreža bez kvarova biće označena sa $\Gamma_0 = \Gamma(0,0) = \Gamma(0,1) = (V_0, E_0)$, gde je $|V_0| = N_0$, a $|E_0| = M_0$.

Nezavisno od toga koji je optimizacioni problem koji posmatramo, označimo linearni program kojim se vrši optimizacija rutiranja u mreži $\Gamma(r,t)$ sa $LP(\Gamma(r,t))$. Kao i kod mreža u kojima nema kvarova resursa, vršićemo optimizaciju u cilju maksimizacije propusne moći, odnosno minimizacije cene mreže.



Slika 10.1.1. Mogući tipovi diskonekcije u slučaju kvara čvora i , odnosno linka l .

10.2. Maksimizacija propusne moći mreže

Propusna moć mreže definisana je kao suma garantovanih saobraćaja čvorova. Garantovani saobraćaj čvora predstavlja saobraćaj koji uvek može biti rutiran kroz mrežu. Dakle, garantovani saobraćaj mora biti rutiran i u mreži u kojoj je došlo do kvara nekog od resursa. Očigledno, potreba da se saobraćaj rutira i u situacijama u kojima nisu svi resursi na raspolaganju nameće dodatna ograničenja za garantovani saobraćaj.

Određivanje propusne moći pouzdane mreže vršeno je za slučaj padova čvorova u regularnim topologijama, kao i za slučajeve padova čvorova i linkova u realnim topologijama. U ovom odeljku biće prvo opisan linearni program za optimizaciju rutiranja sa balansiranjem saobraćaja u slučaju kvara nekog resursa, kao i način određivanja propusne moći pouzdane mreže. Zatim će biti objašnjeno kako se određuje propusna moć mreže sa rutiranjem po najkraćoj putanji, ukoliko su mogući kvarovi resursa. Na kraju, biće upoređene vrednosti propusne moći mreže za ova dva slučaja rutiranja.

10.2.1. Propusna moć pouzdane mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)

Da bismo odredili garantovani saobraćaj čvora u mreži u kojoj dolazi do kvara resursa, neophodno je prvo odrediti ograničenje za saobraćaj čvora u svakom od mogućih stanja mreže $\Gamma(r, t)$. Najstrožije ograničenje za saobraćaj čvora po svim mogućim

stanjima mreže predstavljaće garantovani saobraćaj tog čvora, a propusna moć mreže biće jednaka sumi garantovanih saobraćaja svih čvorova.

Za svako od mogućih stanja mreže $\Gamma(r, t)$ vršimo optimizaciju rutiranja u cilju postizanja minimalnog zagušenja, odnosno maksimizacije saobraćaja u čvoru mreže. U tu svrhu primenjujemo linearni program $LP(\Gamma(r, t))$:

$$\begin{aligned} & \min Q(r, t) \\ (C1) \quad & \sum_{i=1}^N k_i(r, t) = 1 \\ (C2) \quad & \forall l \in E: \sum_{(i,m)} F_{im}^l(r, t) (k_i(r, t)r_m + k_m(r, t)s_i) / C^l \leq Q(r, t) \end{aligned} \quad (10.2.1)$$

Kao rezultat optimizacije dobićemo vrednost garantovanog saobraćaja za čvor i u mreži $\Gamma(r, t)$, koju označavamo sa $g_i^{rbs}(r, t)$. Za čvor i , garantovana vrednost saobraćaja u mreži rezistentnoj na kvarove resursa tipa t u slučaju kada je primenjeno rutiranje sa balansiranjem saobraćaja data je izrazom:

$$g_i^{rbs}(t) = \min_{r \in R_t} g_i^{rbs}(r, t), \quad (10.2.2)$$

dok je propusna moć mreže jednaka sumi garantovanih saobraćaja svih čvorova

$$T^{rbs}(t) = \sum_{i \in V} g_i^{rbs}(t). \quad (10.2.3)$$

10.2.2. Propusna moć mreže sa rutiranjem po najkraćoj putanji (RNP)

Garantovani saobraćaj za neki čvor i u mreži rezistentnoj na kvarove resursa tipa t određujemo na osnovu najstrožijeg ograničenja za saobraćaj čvora u svakom od mogućih stanja mreže (r, t) .

U svakom stanju mreže (r, t) određujemo garantovani saobraćaj $g_i^{rnp}(r, t)$, po postupku opisanom u Odeljku 8.2. Dakle, prvo svakom linku u mreži $\Gamma(r, t)$ pridružujemo bipartitni graf i računamo intenzitet maksimalnog toka u tom grafu, $f_{\max}^l(r, t)$. Zatim na osnovu intenziteta maksimalnog toka procenjujemo ograničenje za saobraćaj čvora na posmatranom linku l . Najstrožije ograničenje po linkovima iz mreže $\Gamma(r, t)$ predstavlja vrednost $g_i^{rnp}(r, t)$. Na nivou mreže, garantovana vrednost saobraćaja u ovom slučaju onda iznosi:

$$g_i^{rnp}(t) = \min_{r \in R_t} g_i^{rnp}(r, t). \quad (10.2.4)$$

Propusna moć mreže definisana je izrazom

$$T^{rnp}(t) = \sum_{i \in V} g_i^{rnp}(t). \quad (10.2.5)$$

10.2.3. Poređenje propusne moći za RBS i RNP

U odnosu na propusnu moć mreže sa rutiranjem po najkraćoj putanji, povećanje propusne moći u mreži rezistentnoj na kvarove resursa tipa t , u kojoj je primenjeno rutiranje sa balansiranjem saobraćaja, jednako je:

$$G(t) = \frac{T^{rbs}(t)}{T^{rnp}(t)}, \quad (10.2.6)$$

gde su vrednosti $T^{rbs}(t)$ i $T^{rnp}(t)$ definisane u prethodnom odeljku.

Analiza povećanja propusne moći vršena je za regularne topologije u slučaju otkaza čvorova, kao i za realne topologije u slučaju otkaza čvorova i linkova.

i) Regularne mrežne topologije

Za regularne mrežne topologije vršena je analiza za slučaj pada čvorova u mreži. Smatrano je da svi čvorovi imaju iste težine $s_i = w_i = 1$. Kao i u slučaju mreža u kojima nema otkaza resursa, postoje dva načina na koje se vrši rutiranje po najkraćoj putanji: redom po osama mreže ili na osnovu numeracije čvorova u mreži.

a) Rutiranje po koordinatnim osama mreže

Posmatraćemo prvo slučaj u kome se rutiranje po najkraćoj putanji vrši po osama mreže, i to redosledom $x \rightarrow y \rightarrow z$. Ukoliko dođe do pada nekog od čvorova koji se nalazi na putanji od izvorišnog čvora do odredišta, ciklus po osama se ponavlja onoliko puta koliko je potrebno, da bi paket stigao do konačnog odredišta. Na primer, pretpostavimo da je neophodno izvršiti rutiranje od čvora sa koordinatama (x_0, y_0, z_0) , do čvora sa koordinatama (x_1, y_1, z_1) . Neka je do kvara došlo u čvoru koji se nalazi na x -osi, između čvorova (x_0, y_0, z_0) i (x_1, y_0, z_0) . Neka je poslednji čvor do koga se može stići krećući se x -osom iz (x_0, y_0, z_0) čvor sa koordinatama (x^*, y_0, z_0) . Tada se rutiranje vrši po putanji $(x_0, y_0, z_0) \rightarrow (x^*, y_0, z_0) \rightarrow (x^*, y_1, z_0) \rightarrow (x^*, y_1, z_1) \rightarrow (x_1, y_1, z_1)$. Ukoliko se čvor koji je u kvaru nalazi na osi po kojoj se poslednjoj vrši rutiranje, onda se on zaobilazi tako što se napravi korak po osi koja je prva sledeća na redu, obiđe čvor u kvaru, a zatim ponovo vrati na osu po kojoj treba nastaviti kretanje.

Kao i u slučaju analiziranom u Odeljku 8.3.1, smatramo da su usled simetrije mreže i jednakih saobraćaja svih čvorova koeficijenti balansiranja isti za sve čvorove i iznose $k_m = 1/N$. Tada je saobraćaj između svaka dva čvora jednak

$$b = b_{ij}^{(1)} + b_{ij}^{(2)} = 2s/N = 2/N. \quad (10.2.7)$$

1) Prsten:

Posmatrajmo prvo slučaj pada čvora u mreži sa topologijom prstena, kada je primenjeno balansiranje saobraćaja. Očigledno, koeficijenti balansiranja prilagodiće se tako da se sav saobraćaj balansira preko čvora koji se nalazi dijametralno suprotno od čvora koji je u kvaru, i predstavlja centar simetrije novonastale mreže.

Obratimo sada pažnju na sve čvorove koji se nalaze na jednoj strani čvora preko koga se sav saobraćaj balansira, recimo da je to desna strana. Tih čvorova ima $\lfloor (N-1)/2 \rfloor$. Sav saobraćaj koji ti čvorovi primaju kreće se po linku desno od balansirajućeg čvora. Pošto svi čvorovi generišu i primaju iste saobraćaje, $s_i = 1$, opterećenje linka desno od balansirajućeg čvora iznosi $\lfloor (N-1)/2 \rfloor$. Stoga je ograničenje za saobraćaj čvora u svakom od mogućih stanja mreže (r, t) jednako $2C/(N-1)$, pa je i garantovani saobraćaj čvora u mreži rezistentnoj na kvarove resursa tipa $t = 0$ (kvarovi čvorova) jednako:

$$g_i^{rbs}(0) = \frac{2C}{N-1}. \quad (10.2.8)$$

Propusna moć mreže iznosi onda

$$T^{rbs}(0) = 2C. \quad (10.2.9)$$

Sa druge strane, za rutiranje po najkraćoj putanji najveće opterećenje nekog linka ostvaruje se kada međusobno najudaljeniji parovi čvorova međusobno komuniciraju. U tom slučaju, $\lfloor (N-1)/2 \rfloor$ tokova se rutira po najopterećenijem linku u mreži. Stoga je garantovani saobraćaj za čvor mreže dati izrazom

$$g_i^{rnp}(0) = \frac{2C}{N-1}, \quad (10.2.8)$$

a propusna moć mreže iznosi

$$T^{rnp}(0) = 2C. \quad (10.2.9)$$

Povećanja propusne moći nema u ovom slučaju, odnosno

$$G(0)=1. \quad (10.2.10)$$

Linija koja odgovara pojačanju datom sa (10.2.10) predstavljena je isprekidanom (crnom) linijom sa trouglićima na Slici 10.2.1.

2) Menhetn:

Smatraćemo da u slučaju kada je primenjeno balansiranje saobraćaja a N je dovoljno veliko, nakon prilagođavanja koeficijenata balansiranja padu čvora u mreži ne dolazi do bitnije promene opterećenja linkova u mreži usled preusmeravanja saobraćaja na čijoj putanji se nalazi čvor koji je u kvaru. Stoga je garantovani saobraćaj u čvoru mreže jednak garantovanom saobraćaju za čvor mreže u kojoj ne dolazi do kvarova resursa, i koji je izveden u Odeljku 8.3.1:

$$g_i^{rbs}(0) = \frac{4C}{D}, \quad (10.2.11)$$

Propusna moć mreže iznosi

$$T^{rbs}(0) = (N-1)(4C/D) \approx 4CD. \quad (10.2.12)$$

Kada je u pitanju rutiranje po najkraćoj putanji, u najgorem slučaju pre otkaza čvora, $D/2$ parova čvorova komuniciraju koristeći jedan link. Po padu čvora (x_0, y_0) , sav saobraćaj koji bi se inače usmeravao po dvema paralelnim osama mreže ($x = x_0$ i $x = x_0 - 1$), sada se rutira po osi $x = x_0 - 1$. Stoga se opterećenja linkova na ovoj osi dupliraju, pa u najgorem slučaju imamo D parova čvorova koji komuniciraju po jednom linku. Zato je garantovani saobraćaj za rutiranje po najkraćoj putanji dat sa

$$g_i^{rnp}(0) = C/D, \quad (10.2.13)$$

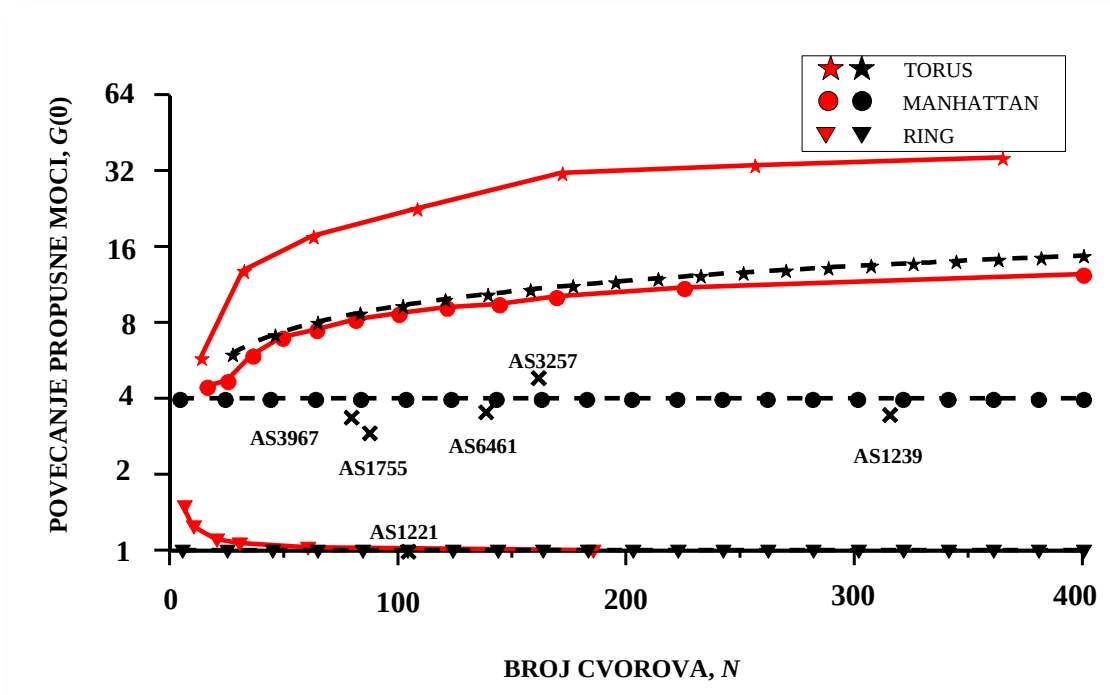
dok je propusna moć mreže

$$T^{rnp}(0) = (N-1)(C/D) \approx CD. \quad (10.2.14)$$

Povećanje propusne moći mreže iznosi

$$G(0) = 4, \quad (10.2.15)$$

i predstavljeno je isprekidanom (crnom) linijom sa tačkama na Slici 10.2.1.



Slika 10.2.1. Povećanje propusne moći pouzdane mreže za regularne i realne mrežne topologije, u slučaju pada čvorova. Isprekidanom crnom linijom označeni su rezultati za regularne mrežne topologije i RNP po osama, a punom crvenom za RNP na osnovu numeracije čvorova. Krstići predstavljaju rezultate za realne mreže.

3) Torus:

Kao i kod Menhetn mreže, smatramo da nema bitnijih promena opterećenja linkova usled pada čvora mreže, te da garantovani saobraćaj čvora za slučaj kada je primenjeno rutiranje sa balansiranjem saobraćaja ostaje približno jednak garantovanom saobraćaju određenom jednačinom (8.3.16):

$$g_i^{rbs}(0) = 4C / D. \quad (10.2.16)$$

Propusna moć mreže iznosi

$$T^{rbs}(0) = (N-1)(4C / D) \approx 4CD. \quad (10.2.17)$$

Ako sada posmatramo rutiranje po najkraćoj putanji, možemo zaključiti da u najkritičnijem slučaju $D^2 / 2$ parova čvorova komunicira preko linka na y -osi mreže. Naime, ako do čvora koji je u kvaru dođemo krećući se po x -osi, sav saobraćaj koji bi se rutirao po x -osi se usmerava po y -osi mreže. Međutim, u najgorem slučaju može se raditi o saobraćaju koji generišu $D^2 / 2$ parova čvorova iz osenčenih oblasti na slici 8.2.3. Ukoliko se na čvor koji je u kvaru naiđe tokom kretanja po y -osi, sav saobraćaj se usmerava po z -osi. Saobraćaj koji se inače rutira po z -osi je zanemarljiv u odnosu

na maksimalnih $D^2 / 2$ koji se mogu preusmeriti. Dakle, u najgorem slučaju opterećenje linka iznosi $D^2 / 2$, pa je garantovani saobraćaj u čvoru mreže jednak

$$g_i^{mp}(0) = 2C / D^2, \quad (10.2.18)$$

dok je propusna moć mreže

$$T^{mp}(0) = (N-1)(2C / D^2) \approx 2C. \quad (10.2.17)$$

Povećanje propusne moći mreže predstavljeno je isprekidanom (crnom) linijom sa zvezdicama na Slici 10.2.1. i iznosi

$$G(0) = 2D = 2\sqrt[3]{N}. \quad (10.2.15)$$

b) Rutiranje na osnovu numeracije čvorova

Rezultati za rutiranje na osnovu numeracije čvorova predstavljeni su na Slici 8.2.1. punom crvenom linijom. Može se uočiti da su povećanja propusne moći mreže veća za ovaj slučaj nego za slučaj rutiranja po koordinatnim osama mreže. Kao i u slučajevima analiziranim u Odeljku 6.3.1, ovakav rezultat može se objasniti činjenicom da rutiranje po koordinatnim osama ravnomernije raspoređuje opterećenje po linkovima u mreži.

ii) Realne mrežne topologije

Za realne mrežne topologije razmatrani su kvarovi oba tipa resursa, kako čvorova tako i linkova. Pored vrednosti iz izraza (8.2.6), za realne mrežne topologije računata je i vrednost povećanja propusne moći za svako od stanja mreže:

$$G(r, t) = \frac{T^{rbs}(r, t)}{T^{mp}(r, t)}, \quad (10.2.16)$$

gde je $T^{rbs}(r, t) = \sum_{i \in V} g_i^{rbs}(r, t)$, a $T^{mp}(r, t) = \sum_{i \in V} g_i^{mp}(r, t)$.

Prosečna vrednost povećanja garantovanog saobraćaja označena je sa

$$\overline{G(t)} = \frac{\sum_{r \in R_t} G(r, t)}{|R_t|}, \quad (10.2.17)$$

a standardna devijacija sa $\sigma_{G(t)}$. Maksimalna vrednost povećanja garantovanog saobraćaja za slučaj kvara resursa tipa t označena je sa $G_{\max}(t) = \max_{r \in R_t} G(r, t)$, a minimalna sa $G_{\min}(t) = \min_{r \in R_t} G(r, t)$.

Rezultati za slučaj kvarova čvorova u mreži ($t = 0$) dati su u Tabeli 10.2.1, a za slučaj kvarova linkova ($t = 1$) u Tabeli 10.2.2. Najlošiji rezultati postižu se za australijsku mrežu Telstra, koja ima topologiju prstena sa izdvojenim kracima.

Tabela 10.2.1. Rezultati za slučaj otkazivanja čvorova

MREŽA	Broj čvorova	Broj slučajeva	$\overline{G(0)}$	$\sigma_{G(0)}$	$G_{\max}(0)$	$G_{\min}(0)$	$G(0)$
Exodus (US) 3967	79	80	5.59	0.51	6.34	3.40	3.40
Ebone (Europe) 1755	87	88	3.22	0.32	4.62	2.01	2.95
Telstra (Australia) 1221	104	105	2.29	0.26	2.49	1	1
Abovenet (US) 6461	138	139	3.03	0.15	3.97	2.36	3.57
Tiscali (Europe) 3257	161	162	7.69	0.78	8.95	2.98	4.88
Sprintlink (US) 1239	315	316	5.51	0.27	6.17	2.63	3.48

Rezultati za slučaj kvarova čvorova ($t = 0$) predstavljeni su i na Slici 10.2.1, te se mogu uporediti sa rezultatima za regularne mrežne topologije. Dok za mrežu Telstra rezultati očekivano odgovaraju rezultatima za mreže sa topologijom prstena, povećanje propusne moći ostalih mreža je u rangu mreža sa Menhethn topologijom. Kao i u slučaju mreža u kojima nema kvarova resursa, najbolji rezultati ostvaruju se za topologiju Tiscali.

Tabela 10.2.2. Rezultati za slučaj otkazivanja linkova

MREŽA	Broj linkova	Broj slučajeva	$\overline{G(1)}$	$\sigma_{G(1)}$	$G_{\max}(1)$	$G_{\min}(1)$	$G(1)$
Exodus (US) 3967	294	148	5.72	0.26	6.59	3.40	3.40
Ebone (Europe) 1755	322	162	3.32	0.25	5.10	2.23	3.33
Telstra (Australia) 1221	302	152	2.34	0.19	3.30	1	1
Abovenet (US) 6461	744	373	3.04	0.10	4.22	2.52	3.65
Tiscali (Europe) 3257	656	329	7.86	0.48	8.42	3.76	5.35
Sprintlink (US) 1239	1944	973	5.56	0.12	6.16	3.57	3.96

10.3.Minimizacija cene mreže

U prethodnom odeljku analizirana je propusna moć mreže u kojoj dolazi do kvarova resursa. Sada ćemo analizirati cenu pouzdane mreže koja treba da podrži rutiranje zadanog saobraćaja.

Da bi se omogućila komunikacija korisnika u mreži i u slučaju otkazivanja linkova, neophodno je dimenzionisati kapacitet svakog od linkova u odnosu na najgoru

moću situaciju, tj. onu u kojoj je taj link najopterećeniji. U isto vreme, cilj je minimizacija cene mreže, u skladu sa definicijom cene kao sume kapaciteta linkova u mreži, koja je korišćena u dosadašnjoj analizi.

U ovom odeljku biće opisan postupak određivanja cene pouzdane mreže u slučajevima kada se koriste rutiranje sa balansiranjem saobraćaja i obično rutiranje po najkraćoj putanji, a zatim će biti upoređene cene mreže u ova dva slučaja.

10.3.1. Cena mreže sa rutiranjem sa balansiranjem saobraćaja (RBS)

Svakome linku $l \in E_0$ iz početne mreže Γ_0 neophodno je dodeliti dovoljno kapaciteta, tako da može da propusti saobraćaj u onim slučajevima kada je najviše opterećen. Istovremeno, cena mreže treba da bude što niža. Za svako stanje (r, t) , u kome je došlo do kvara resursa r tipa t , optimalna raspodela kapaciteta, za koju se postiže minimalna cena mreže, biće određena sledećim linearnim programom:

$$\begin{aligned}
 & \min \sum_{l \in E(r, t)} C^{l_{rbs}}(r, t) \\
 & (C1) \sum_{i \in V(r, t)} k_i(r, t) = 1 \\
 & (C2) \forall l \in E(r, t): \\
 & \quad \sum_{(i, m) \in (V(r, t) \times V(r, t))} F_{im}^l(r, t) (k_i(r, t) s_m + k_m(r, t) s_i) \leq C^{l_{rbs}}(r, t).
 \end{aligned} \tag{10.3.1}$$

Kao i do sad, $F_{im}^l(r, t)$ je binarna promenljiva koja ima vrednost jednaku jedinici ukoliko se link l u mreži $\Gamma(r, t)$ nalazi na najkraćoj putanji od čvora i do čvora m . U suprotnom, vrednost ove promenljive jednaka je nuli. Koeficijent balansiranja dodeljen čvoru i u mreži $\Gamma(r, t)$ označen je sa $k_i(r, t)$, dok je sa $C^{l_{rbs}}(r, t)$ označen kapacitet koji je neophodno dodeliti linku l u mreži $\Gamma(r, t)$. Praktično, kao rezultat optimizacije rutiranja u cilju minimizacije cene mreže $\Gamma(r, t)$, dobiće se raspodela kapaciteta linkova $C^{l_{rbs}}(r, t), l \in E$. Ukoliko link l ne pripada povezanom delu mreže u slučaju kvara, smatramo da je $C^{l_{rbs}}(r, t) = 0$.

Kapacitet svakog od linkova mora biti takav da se omogući rutiranje saobraćaja u najgorem mogućem slučaju, odnosno u onom stanju mreže $\Gamma(r, t)$ za koje je opterećenje

linka maksimalno. Stoga je neophodan kapacitet linkova u mreži rezistentnoj na kvarove resursa tipa t dat izrazom:

$$\forall l \in E : C^{l_{rbs}}(t) = \max_{r \in R_t} C^{l_{rbs}}(r, t). \quad (10.3.2)$$

Cena mreže jednaka je sumi neophodnih kapaciteta linkova:

$$C^{rbs}(t) = \sum_{l \in E_0} C^{l_{rbs}}(t) \quad (10.3.3)$$

Sumiranje se vrši po svim linkovima originalne mreže, uzimajući u obzir pretpostavku da je $C^l(r, t) = 0$ ako se link l ne nalazi u povezanom delu mreže $\Gamma(r, t)$.

10.3.2. Cena mreže za rutiranjem po najkraćoj putanji (RNP)

Kada je u pitanju mreža u kojoj je upotrebljeno rutiranje po najkraćoj putanji, svakom linku mora biti dodeljen kapacitet tako da se omogući rutiranje u slučaju kada je link najopterećeniji. U mreži $\Gamma(r, t)$, maksimalno opterećenje linka određujemo koristeći algoritam maksimalnog toka, kao u odeljku 7.2:

$$\forall l \in E(r, t) : C^{l_{rnp}}(r, t) = f_{\max}^l(r, t)$$

Ovde je sa $f_{\max}^l(r, t)$ označen intenzitet maksimalnog toka u bipartitnom grafu pridruženom linku l u mreži $\Gamma(r, t)$. Ukoliko se link $l \in E_0$ ne nalazi u povezanoj mreži $\Gamma(r, t)$ posle kvara resursa r tipa t , smatramo da je $C^l(r, t) = 0$.

Naravno, kapacitet linka $l \in E_0$ je neophodno dimenzionisati tako da može da podrži saobraćaj u onom stanju mreže $\Gamma(r, t)$ za koje je on najopterećeniji. Dakle:

$$\forall l \in E : C^{l_{rnp}}(t) = \max_{r \in R_t} C^{l_{rnp}}(r, t) \quad (10.3.4)$$

Ukupna cena mreže u kojoj je primenjeno obično rutiranje po najkraćoj putanji jednaka je sumi kapaciteta koje je neophodno dodeliti linkovima u mreži.

$$C^{rnp}(t) = \sum_{l \in E_0} C^{l_{rnp}}(t). \quad (10.3.5)$$

10.3.3. Poređenje cene mreže za RBS i RNP

Ušteda koja se ostvaruje primenom rutiranja sa balansiranjem saobraćaja umesto običnog rutiranja po najkraćoj putanji u mreži rezistentnoj na kvarove resursa tipa t data je izrazom

$$G(t) = \frac{C^{np}(t)}{C^{rs}(t)}, \quad (10.3.6)$$

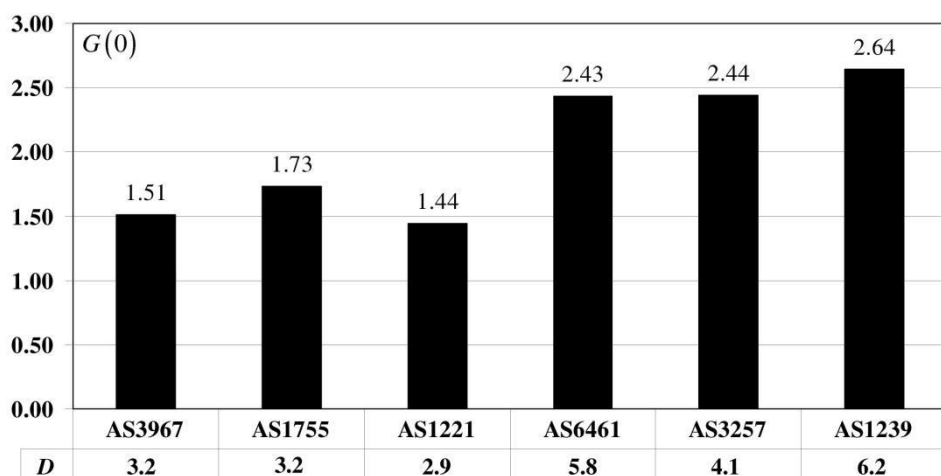
gde su cene mreže sa rutiranjem po najkraćoj putanji i rutiranjem sa balansiranjem saobraćaja definisane jednačinama (10.3.5) i (10.3.3) respektivno.

U ovom odeljku biće izloženi rezultati analize koja je sprovedena za realne mrežne topologije publikovane u okviru *Rocketfuel* projekta. Analizirani su, kao što je već spomenuto, slučajevi kvarova na čvorovima ($t = 0$), odnosno na linkovima u mreži ($t = 1$). U slučaju kvara na ruterima (čvorovima) mreže, imamo ukupno N_0 mogućih stanja u kojima je neki od čvorova u kvaru. U slučaju kvara na linkovima, postoji M_0 mogućih stanja u kojima postoji kvar. Kvar resursa r tipa t može prouzrokovati diskonekciju izvesnog broja čvorova, u slučajevima koji su ilustrovani na slici 10.1.1. Ukupan broj stanja (r, t) u kojima dolazi do ovakvih diskonekcija jednak je P . Broj stanja u kojima je d čvorova diskonektovano označićemo sa p_d . Naravno, mora važiti da je $\sum_d p_d = P$. Raspodela broja diskonektovanih čvorova (d, p_d) , kao i ukupan broj stanja u kojima dolazi do diskonekcija za analizirane mrežne topologije, dati su u Tabeli 10.3.1.

Tabela 10.3.1. Raspodela otkaćenih čvorova mreže usled kvara nekog od resursa za analizirane realne topologije

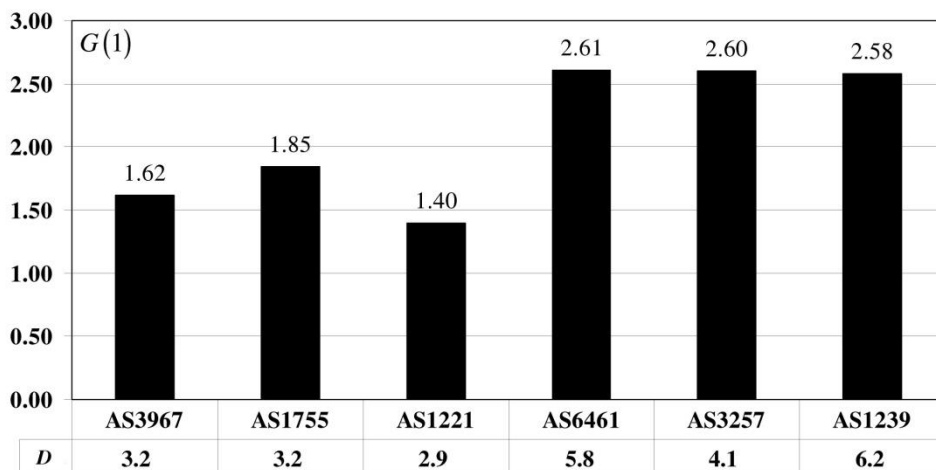
MREŽA	Exodus AS 3967		Ebony AS 1755		Telstra AS 1221						Abovenet AS 6461			Tiscali AS 3257						Sprintlink AS 1239	
N_0	79		87		104						138			161						315	
M_0	294		322		302						744			656						1944	
P	6		11		19						4			24						28	
d	1	2	1	2	1	2	3	6	10	17	1	3	4	1	2	3	4	6	1	2	
p_d	5	1	6	5	9	6	1	1	1	1	2	1	1	8	8	5	2	1	25	3	

Na Slici 10.3.1. prikazani su odnosi cena mreža sa rutiranjem po najkraćoj putanji i cena mreža sa rutiranjem sa balansiranjem saobraćaja, za slučaj mogućih kvarova rutera. Cena mreža je od 1.44 do 2.64 puta manja ukoliko je primenjeno rutiranje sa balansiranjem saobraćaja.



Slika 10.3.1. Odnos cena pouzdanih mreža sa balansiranjem saobraćaja i običnim rutiranjem po najkraćoj putanji, rezistentnih na kvarove rutera

Na Slici 10.3.2. prikazani su odnosi cena mreža sa rutiranjem po najkraćoj putanji i cena mreža sa rutiranjem sa balansiranjem saobraćaja, za slučaj mogućih kvarova linkova. Cene mreža su u ovom slučaju od 1.40 do 2.61 puta manje ukoliko je primenjeno rutiranje sa balansiranjem saobraćaja., u odnosu na slučaj kada je primenjeno obično rutiranje po najkraćoj putanji.



Slika 10.3.2. Odnos cena pouzdanih mreža sa balansiranjem saobraćaja i običnim rutiranjem po najkraćoj putanji, rezistentnih na kvarove linkova.

11. ZAKLJUČAK

U radu je predložena šema rutiranja zasnovana na balansiranju saobraćaja i rutiranju po najkraćoj putanji. Predložena šema omogućava optimizaciju rutiranja u cilju minimizacije cene mreže, ili maksimizacije njene propusne moći. Optimizacija rutiranja vrši se linearnim programom, za čiju je formulaciju dovoljno raspolagati informacijom o ukupnom saobraćaju generisanom i primljenom u čvorovima mreže. Nije neophodno poznavati raspodelu saobraćaja među parovima čvorova. Stoga je predložena šema pogodna za primenu u modernim mrežama, u kojima je udeo *peer-to-peer* saobraćaja veliki i otežava predviđanje intenziteta tokova među parovima čvorova.

Izvršeno je poređenje performansi predloženog rutiranja sa rutiranjem po najkraćoj putanji. Pokazano je da se primenom rutiranja sa balansiranjem saobraćaja umesto rutiranja po najkraćoj putanji može povećati propusna moć postojeće mreže. Takođe, pokazano je da se mogu ostvariti i uštede u ceni mreže koja treba da podrži zadati saobraćaj. Poređenje je vršeno za realne i regularne mrežne topologije. Analizirani su slučajevi u kojima nema kvarova u mreži, kao i slučajevi jednostrukih otkaza resursa. Pokazano je da se u svim slučajevima primenom rutiranja sa balansiranjem saobraćaja ostvaruju značajna poboljšanja (povećanje propusne moći, tj. smanjenje cene mreže). Takođe, pokazano je da poboljšanje koje se može ostvariti zavisi od stepena čvorova u mreži, i veće je za mreže sa razgranatijom topologijom.

Predložena šema može se jednostavno implementirati. Njena primena olakšava i proces rezervacije kapaciteta u mreži. Prilikom rezervacije kapaciteta za neku sesiju, potrebno je uputiti zahtev samo izvorišnom i odredišnom čvoru sesije. Nema potrebe za proverom stanja linkova u mreži, niti je neophodno uzimati u obzir aktivnost ostalih čvorova.

LITERATURA

- [1] H. Raecke, "Minimizing congestion in general networks," in *Proceedings of FOCS 43*, 2002.
- [2] M. Bienkowski, M. Korzeniowski and H. Räcke, "A Practical Algorithm for Constructing Oblivious Routing Schemes," in *Proceedings of SPAA '03*, 2003.
- [3] C. Harrelson, K. Hildrum and S. Rao, "A Polynomial-time Tree Decomposition to Minimize Congestion," in *Proceedings of SPAA '03*, 2003.
- [4] D. Applegate and E. Cohen, "Making intra-domain routing robust to changing and uncertain traffic demands: understanding fundamental tradeoffs," in *Proceedings of ACM SIGCOMM '03*, 2003.
- [5] Y. Azar, E. Cohen, A. Fiat, H. Kaplan and H. Räcke, "Optimal oblivious routing in polynomial time," in *Proceedings of 35th ACM Symposium on the Theory of Computing*, 2003.
- [6] M. Kodialam, T. V. Lakshman and S. Sengupta, "Efficient and Robust Routing of Highly Variable Traffic," in *Proceedings of 3rd Workshop on Hot Topics in Networks*, 2004.
- [7] M. Kodialam, T. V. Lakshman and S. Sengupta, "Traffic-Oblivious Routing for Guaranteed Bandwidth Performance," *IEEE Communications Magazine*, vol. 45, no. 4, pp. 46-51, 2007.
- [8] M. Antić and A. Smiljanić, "Šema rutiranja zasnovana na balansiranju saobraćaja i rutiranju po najkraćoj putanji," in *Proceedings of TELFOR 2007*, 2007.
- [9] M. Antić and A. Smiljanić, "Oblivious Routing Scheme Using Load Balancing Over Shortest Paths," in *Proceedings of IEEE International Conference on Communications (ICC)*, 2008.

- [10] M. Antić and A. Smiljanić, "Routing with Load Balancing: Increasing the Guaranteed Node Traffics," *IEEE Communication Letters*, vol. 13, no. 6, pp. 450-452, 2009.
- [11] M. Antić and A. Smiljanić, "Optimal Capacity Allocation for Load Balanced Shortest Path Routing," in *Proceedings of IEEE Workshop on High Performance Switching and Routing (HPSR)*, 2009.
- [12] M. Antić and A. Smiljanić, "Cost Reduction of Reliable Networks Using Load Balanced Routing," *IEEE Communication Letters*, vol. 14, no. 3, pp. 263-265, 2010.
- [13] M. Antić, N. Maksić, P. Knežević and A. Smiljanić, "Two Phase Load Balanced Routing using OSPF," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 1, pp. 51-59, 2010.
- [14] N. Maksić, P. Knežević, M. Antić and A. Smiljanić, "Influence of Load Balancing on Quality of Real Time Data Transmission," vol. 6, no. 3, pp. 515-524, *Serbian Journal of Electrical Engineering*, 2009.
- [15] N. Maksić, P. Knežević, M. Antić and A. Smiljanić, "On the Performance of the Load Balanced Shortest Path Routing," in *Proceedings of PacRim 2009*, 2009.
- [16] N. Maksić, P. Knežević, M. Antić and A. Smiljanić, "Uticaj rutiranja sa balansiranjem na kvalitet prenosa podataka u realnom vremenu," in *Proceedings of ETRAN*, 2009.
- [17] M. Antić and A. Smiljanić, "Throughput of Reliable Networks with Load Balanced Shortest Path Routing," in *Proceedings of IEEE Globecom*, 2010.
- [18] J. T. Moy, *OSPF Complete Implementation*, Addison-Wesley Professional, 2000.
- [19] A. Medina, N. Taft, K. Salamatian, S. Bhattacharyya and C. Diot, "Traffic Matrix Estimation: Existing Techniques and New Directions," in *Proceedings of the ACM SIGCOMM'02*, 2002.
- [20] J. Matoušek and B. Gärtner, *Understanding and Using Linear Programming*, Springer Verlag, 2007.

- [21] D. Cvetković, M. Čangalović, Đ. Dugošija, V. Kovačević-Vujčić, S. Simić and J. Vuleta, *Kombinatorna optimizacija – Matematička teorija i algoritmi*, Društvo operacionih istraživača Jugoslavije – DOPIS, 1996.
- [22] G. Dantzig and M. Thapa, *Linear Programming 1: Introduction*, Springer Verlag, 1997.
- [23] C. Clos, "A Study of Non-Blocking Switching Networks," *Bell System Technical Journal*, vol. 32, no. 5, p. 406–424, 1953.
- [24] R. Melen and J. S. Turner, "Nonblocking multirate networks," *SIAM Journal on Computing*, vol. 18, no. 2, pp. 301-313, 1989.
- [25] J. S. Turner and R. Melen, "Multirate Clos Networks," *IEEE Communications Magazine*, vol. 41, no. 10, pp. 38 - 44, 2003.
- [26] S. C. Liew, M. H. Ng and C. W. Chan, "Blocking and Nonblocking Multirate Clos Switching Networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 3, pp. 307-318 , 1998.
- [27] A. Smiljanić, "Rate and delay guarantees provided by Clos packet switches with load balancing," *IEEE/ACM Transactions on Networking*, 2008.
- [28] I. Keslassy, C. S. Chang, N. McKeown and D. S. Lee, "Optimal Load Balancing," in *Proceedings of IEEE INFOCOM*, 2005.
- [29] A. Smiljanić, "Neblokirajuće paketske mreže zasnovane na balansiranju saobraćaja," in *Proceedings of PosTel*, 2006.
- [30] N. Spring, R. Mahajan and D. Wetherall, "Measuring ISP topologies with Rocketfuel," in *Proceedings of the ACM SIGCOMM'02*, 2002.
- [31] B. Towels and W. J. Dally, "Worst-case Traffic for Oblivious Routing Functions," *IEEE Computer Architecture Letters*, vol. 1, no. 1, 2002.
- [32] S. R. Thorncraft, H. R. Outhred and D. J. Clements, "Evaluation of Open-Source LP Optimization Codes in Solving Electricity Spot Market Optimization Problems," in *Proceedings of 19th Mini-Euro Conference on Operation Research Models and Methods in the Energy Sector*, 2006.
- [33] T. Chaney, J. A. Fingerhut, M. Flucke and J. S. Turner, "Design of a Gigabit ATM Switch," in *Proceedings of IEEE INFOCOM*, 1997.

- [34] I. Keslassy, S. T. Chuang, K. Yu, D. Miller, M. Horowitz, O. Solgard and N. McKeown, "Scaling Internet Routers Using Optics," in *Proceedings of the ACM SIGCOMM'03*, 2003.
- [35] C. S. Chang, D. S. Lee and Y. S. Jou, "Load Balanced Birkhoff-von Neuman switches, part I: one-stage buffering," in *Proceedings of IEEE Workshop on High Performance Switching and Routing (HPSR)*, 2001.
- [36] C. S. Chang, D. S. Lee and Y. S. Jou, "Load Balanced Birkhoff-von Neuman switches, part II: multi-stage buffering," *Computer Communications*, vol. 25, pp. 623-634, 2002.
- [37] J. F. Kurose and K. W. Ross, *Umrežavanje računara: od vrha ka dnu sa Internetom u fokusu*, CET Beograd, 2005.
- [38] A. S. Tanenbaum, *Računarske mreže*, Mikro knjiga, 2005.
- [39] M. Živković, *Algoritmi*, Beograd: Matematički fakultet, 2000.
- [40] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to Algorithms*, The MIT Press, 2001.
- [41] R. K. Ahuja, T. L. Magnanti and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993.
- [42] J. Gross and J. Yellen, *Graph Theory and its Applications*, CRC Press, 1999.
- [43] D. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, 1984..
- [44] S. Bradley, A. Hax and T. Magnanti, *Applied Mathematical Programming*, Addison-Wesley, 1977.
- [45] G. Dantzig, *Linear Programming and Extensions*, Princeton: Princeton University Press, 1963.
- [46] *LpSolve Reference Guide*, 2005. [Online]. Available: <http://lpsolve.sourceforge.net>.
- [47] *Internetworking Technology Handbook*, Cisco Press, [Online]. Available: http://www.cisco.com/en/US/docs/internetworking/technology/handbook/ito_doc.html.

- [48] V. Glazer, "Introduction to Optimization Problems and Mathematical Programming," University of Toronto, [Online]. Available:
<http://www.cs.toronto.edu/~avner/teaching/S5-2411/ln/lecture1.pdf>.
- [49] G. Taylor, "Different forms of LP. The algebraic objects behind LP. Basic Feasible Solutions," University of Toronto, [Online]. Available:
<http://www.cs.toronto.edu/~avner/teaching/S5-2411/ln/lecture2.pdf>.
- [50] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

A. PRIMENJEN OPTIMIZACIONI ALAT LP_SOLVE

Korišćen optimizacioni alat *LP_Solve* je besplatan softver, razvijen na Univerzitetu u Ajndhovenu. Na adresi projekta <http://sourceforge.net/projects/lpsolve> mogu se preuzeti kompletan kod, primeri i uputstva. Sva pitanja u vezi sa softverom mogu se postaviti u okviru grupe http://groups.yahoo.com/group/lp_solve

Ovaj softver namenjen je prvenstveno rešavanju problema linearne optimizacije, mada može raditi i sa modelima u kojima promenljive uzimaju samo celobrojne vrednosti. Softver je moguće preuzeti kao biblioteku funkcija koje se mogu pozivati iz programskih jezika C++, Visual Basic, Delphi, Java, a može se integrisati i u Excel, Matlab, itd. Za rešavanje problema linearne optimizacije softver koristi simpleks algoritam, dok kod kombinovanog ili celobrojnog programiranja koristi metod grananja i ograničavanja (*branch and bound*).

Zvanično ograničenje veličine modela koji se može zadati alatu *LP_Solve* ne postoji. Naravno, veći modeli linearnih programa mogu zahtevati veliko procesorsko vreme i veliku količinu RAM memorije za svoje izvršavanje. Model za optimizaciju može se zadati alatu *LP_Solve* na jedan od tri načina:

1) API

API čini set rutina koje se mogu pozvati iz nekog programskog jezika. Na ovaj način se model linearnog programa pravi u radnoj memoriji, zatim se taj model rešava i pozivom odgovarajućih rutina rezultati se prosleđuju glavnom programu.

2) Ulazni fajlovi

Model linearnog programa može se zadati i korišćenjem ulaznog fajla u nekom od dva podržana formata. Prvi je MPS format, koji koristi većina postojećih programa za linearnu optimizaciju, i u kome se zadaju vrednosti koeficijenata kolona u linearnom programu. Ovakav oblik fajla je veoma nečitljiv za korisnika. Drugi format je LP format, kojim se zadaju jednačine

linearnog programa, i znatno je čitljiviji. Ovi fajlovi daju se kao ulazni parametar izvršnoj verziji alata *LP_Solve*, koja zatim poziva API, rešava program i vraća rezultate

3) IDE

U ovom slučaju, model se zadaje direktno u okviru Windows aplikacije, koja poziva API i rešava ga, vraćajući rezultate.

U okviru *LP_Solve* biblioteke definisan je tip podataka **lpvec**, koji predstavlja model linearnog programa. Model se čuva kao matrica oblika

$$\begin{bmatrix} \mathbf{c} & \mathbf{0} \\ \mathbf{A} & \mathbf{b} \end{bmatrix}, \quad (\text{A.0.1})$$

gde su **b, c** već pominjani vektori, a **A** matrica realnih koeficijenata iz linearnog programa. U našem slučaju korišćena je mogućnost pozivanja funkcija iz *LP_Solve* biblioteke putem API, model je formiran iz programskog jezika C++, izdavana naredba za njegovo rešavanje, a zatim su rezultati preuzimani odgovarajućim komandama.

Pregled osnovnih komandi alata *LP_Solve* koje su korišćene za formiranje modela za linearnu optimizaciju dat je u Tabeli A.0.1

Tabela A.0.1. Osnovne funkcije LP_Solve

KOMANDA	make_lp(int rows, int columns)
OPIS FUNKCIJE	Rezervise prostor za novu lprec strukturu i vraća pokazivač na nju. Ukoliko je vraćena vrednost NULL, znači da je došlo do greške (najčešće znači da nema dovoljno memorije).
PARAMETRI FUNKCIJE	<i>rows</i> početni broj redova u modelu; može biti jednak nuli pošto se redovi mogu dodavati komandama <i>add_constraint</i> , <i>add_constraintex</i> <i>columns</i> početni broj kolona u modelu; može biti jednak nuli pošto se nove kolone mogu dodavati komandama <i>add_column</i> , <i>add_columnex</i>
KOMANDA	copy_lp(lprec *lp)
OPIS FUNKCIJE	Kopira postojeću lprec strukturu u novu lprec strukturu i vraća pokazivač na novu lprec strukturu. Vraćena vrednost NULL je indikator da je došlo do greške.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već postojeću lprec strukturu
KOMANDA	delete_lp(lprec *lp)
OPIS FUNKCIJE	Oslobađa svu memoriju dodeljenu lprec strukturi.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na lprec strukturu koju brišemo
KOMANDA	write_lp(lprec *lp, char *filename)
OPIS FUNKCIJE	Zapisuje model u fajl sa imenom <i>filename</i> .
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>filename</i> ime fajla u kome će model biti zapisan
KOMANDA	set_add_rowmode(lprec *lp, unsigned char turnon)
OPIS FUNKCIJE	Definiše da li će se model praviti dodavanjem kolona vezanih za promenljive, ili dodavanjem vrsta (ograničenja). Ukoliko se ne zada drugačije, smatra se da se model pravi dodavanjem vrsta.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>turnon</i> uzima vrednost TRUE – model se pravi dodavanjem ograničenja ili FALSE – model se pravi dodavanjem vrsta
KOMANDA	set_obj_fn(lprec *lp, REAL *row) set_obj_fnex(lprec *lp, int count, REAL *row, int *colno)
OPIS FUNKCIJE	Postavlja ciljnu funkciju. Vraća vrednost 1 ako je uspešno obavljen zadatak, a 0 ako je došlo do greške. Funkcija <i>set_obj_fnex</i> omogućava zadavanje samo nenulih elemenata vektora c . Poželjno je da se ciljna funkcija definiše pre dodavanja ograničenja u model, zbog efikasnosti.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>row</i> za <i>set_obj_fn</i> ovo je niz sa <i>columns</i> +1 elemenata koji sadrži koeficijente uz promenljive u ciljnoj funkciji, tj. elemente vektora c (multi element niza se ignoriše, tako da prvoj koloni modela odgovara koeficijent koji u nizu ima indeks 1 i tako redom); za <i>set_obj_fnex</i> u pitanju je niz sa <i>count</i> elemenata koji sadrži nenulte koeficijente uz promenljive u datom ograničenju <i>count</i> broj elemenata u nizu <i>row</i> i nizu <i>colno</i> <i>colno</i>

	niz sa <i>count</i> elemenata koji sadrži brojeve kolona sa nenultim koeficijentima; može imati i vrednost NULL, u tom slučaju <i>i</i> -ti element niza <i>row</i> odgovara <i>i</i> -toj koloni, pri čemu se nulti element ne računa (nulta kolona modela se koristi za smeštanje vrednosti sa desne strane ograničenja)
KOMANDA	add_constraint (<i>lprec</i> * <i>lp</i> , REAL * <i>row</i> , int <i>constr_type</i> , REAL <i>rh</i>) add_constraintex (<i>lprec</i> * <i>lp</i> , int <i>count</i> , REAL * <i>row</i> , int * <i>colno</i> , int <i>constr_type</i> , REAL <i>rh</i>)
OPIS FUNKCIJE	Dodaje red (ograničenje) u model linearnog programa. Ograničenje se dodaje na kraj modela. Funkcija add_constraintex omogućava da se zadaju samo oni koeficijenti odgovarajuće vrste matrice A koji su različiti od nule.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>row</i> za add_constraint ovo je niz sa <i>columns</i> +1 elemenata koji sadrži koeficijente uz promenljive u datom ograničenju (nulti element niza se ignoriše, tako da prvoj koloni modela odgovara koeficijent koji u nizu ima indeks 1 i tako redom); za add_constraintex ovo je niz sa <i>count</i> elemenata koji sadrži nenulte koeficijente uz promenljive u datom ograničenju <i>constr_type</i> tip ograničenja, može imati vrednosti LE(1) – manje ili jednako, GE(2) – veće ili jednako i EQ(3) – jednako <i>rh</i> vrednost sa desne strane jednakosti ili nejednakosti, tj. element vektora b koji odgovara ograničenju koje se dodaje <i>count</i> broj elemenata u nizu <i>row</i> i nizu <i>colno</i> <i>colno</i> niz sa <i>count</i> elemenata koji sadrži brojeve kolona sa nenultim koeficijentima; može imati i vrednost NULL, u tom slučaju <i>i</i> -ti element niza <i>row</i> odgovara <i>i</i> -toj koloni, pri čemu se nulti element ne računa
KOMANDA	add_column (<i>lprec</i> * <i>lp</i> , REAL * <i>column</i>) add_columnnex (<i>lprec</i> * <i>lp</i> , int <i>count</i> , REAL * <i>column</i> , int * <i>rowno</i>)
OPIS FUNKCIJE	Dodaje kolonu u model linearnog programa. Kolona se dodaje na kraj modela. Funkcija add_columnnex omogućava zadavanje samo nenulatih elemenata kolone.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>column</i> za add_column ovo je niz sa <i>rows</i> +1 elemenata koji sadrži koeficijente uz promenljivu u ograničenjima, odgovara koloni u matrici A (nulti element niza odgovara koeficijentu uz promenljivu u ciljnoj funkciji, odnosno odgovarajućem elementu vektora c); za add_columnnex niz sadrži <i>count</i> elemenata koji su različiti od nule <i>count</i> broj elemenata u nizu <i>column</i> i nizu <i>rowno</i> <i>rowno</i> niz sa <i>count</i> elemenata koji sadrži brojeve vrsta sa nenultim koeficijentima; može imati i vrednost NULL, u tom slučaju <i>i</i> -ti element niza <i>column</i> odgovara <i>i</i> -toj vrsti modela

KOMANDA	<code>solve(lprec *lp)</code>
OPIS FUNKCIJE	Rešava model linearnog programa. Vraća vrednost 0 ukoliko je nađeno optimalno rešenje.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model
KOMANDA	<code>get_variables(lprec *lp, REAL *var)</code>
OPIS FUNKCIJE	Preuzima vrednosti promenljivih iz modela nakon što je izvršena komanda solve.
PARAMETRI FUNKCIJE	<i>lp</i> pokazivač na već formiran model <i>var</i> niz u koji treba upisati vrednosti promenljivih; neophodno je da ovaj niz bude dimenzioniran na <i>columns</i> elemenata pre izdavanja naredbe <code>get_variables</code>

BIOGRAFIJA KANDIDATA

Marija Antić rođena je 1980. godine u Beogradu, gde je završila osnovnu školu i IX beogradsku gimnaziju. Na Elektrotehničkom fakultetu Univerziteta u Beogradu diplomirala je 2005. godine, sa prosečnom ocenom 9,36. Na istom fakultetu magistrirala je 2008. godine, sa prosečnom ocenom 10,00 i temom magistarskog rada „Šema rutiranja zasnovana na balansiranju saobraćaja i rutiranju po najkraćoj putanji“, kod mentora dr Aleksandre Smiljanić.

U periodu od februara 2007. do jula 2008. godine bila je angažovana na Elektrotehničkom fakultetu kao stipendista Ministarstva nauke na projektu „Implementacija kontrolera Internet rutera“. Od jula 2008. do jula 2011. godine, kao saradnik Inovacionog centra Elektrotehničkog fakulteta, učestvovala je na projektu „Sistemska integracija Internet rutera“. Od jula 2011. do novembra 2014. zaposlena je kao softverski inženjer u Majkrosoftovom razvojnom centru u Beogradu. Od januara 2015. radi u kompaniji PSTech kao softverski inženjer. Autor je većeg broja radova vezanih za optimizaciju rutiranja u paketskim mrežama, publikovanih u međunarodnim časopisima i na međunarodnim i domaćim konferencijama. Kao deo tima koji je u Majkrosoftu radio na implementaciji importera PDF-a za Word, autor je i nekoliko patenata iz oblasti analize dokumenata sa fiksnim rasporedom elemenata i konverzije dokumenata iz fiksnog formata u format koji se može dinamički prilagođavati uređaju na kome se prikazuje.

Marija Antić dobitnik je velikog broja nagrada i priznanja na srednjoškolskim takmičenjima iz matematike, fizike i srpskog jezika. U periodu od 1997. do 2007. bila je stipendista Republičke fondacije za razvoj naučnog i umetničkog podmlatka. Dobitnik je i stipendije Kraljevine Norveške, nemačke DAAD stipendije, kao i nagrade alumni udruženja ETF Bafa. Od oktobra do decembra 2003. godine boravila je na Univerzitetu San Martin u Bogoti, Kolumbija, u okviru razmene studenata preko IAESTE. Govori engleski, nemački i španski jezik.

Прилог 1.

Изјава о ауторству

Потписани-а: Марија Антић

Изјављујем

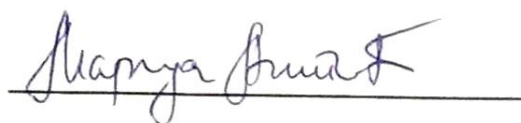
да је докторска дисертација под насловом

„Оптимизација неблокирајућих пакетских мрежа употребом практичног протокола за рутирање са балансирањем саобраћаја”

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 4.5.2015.

Handwritten signature of Marija Antić in blue ink, written over a horizontal line.

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Марија Антић

Број уписа

Студијски програм

Наслов рада Оптимизација неблокирајућих пакетских мрежа употребом практичног протокола за рутирање са балансирањем саобраћаја

Ментор проф. др Александра Смиљанић, Електротехнички факултет Универзитета у Београду

Потписани Марија Антић

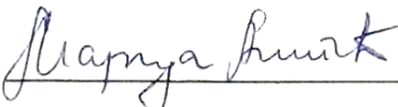
изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, 4.5.2015.



Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Оптимизација неблокирајућих пакетских мрежа употребом практичног протокола за рутирање са балансирањем саобраћаја

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, 4.5.2015.

