Stručni rad

Škola biznisa
Broj 4/2011
UDC 658.818:519.8

Miloš Japundžić*
Ivan Pavkov**

IZRAČUNAVANJE METRIČKIH KARAKTERISTIKA REDOVA ČEKANJA

Sažetak: *Pod metričkim karakteristikama redova čekanja podrazumevamo prosečan broj mušterija prisutnih u sistemu (redu za pružanje usluge, procesu pružanja usluge), kao i prosečno vreme koje mušterija provede u sistemu (redu za pružanje usluge, procesu pružanja usluge). U radu su, za konkretne modele redova čekanja, prikazane metričke karakteristike u zavisnosti od parametara koji predstavljaju prosečnu brzinu pristizanja mušterija u sistem, kao i prosečnu brzinu usluživanja mušterija.*

Ključne reči: *redovi čekanja, metričke karakteristike redova čekanja, prosečan broj mušterija, prosečno vreme čekanja*

CALCULATION OF THE QUEUING METRIC CHARACTERISTICS

Abstract: *By queuing metric characteristics we mean the average number of customers in system (queuing, the service process), as well as the average waiting time of customer in system (queuing, the service process). In this paper, for concrete queuing models, the metric characteristics depending on parametars that represent the average customer arrival rate in system, as well as the average customer service rate are presented.*

Key words: *queuing, queuing metric characteristics, average number of customers, average waiting time*

1. UVOD

Jedna od najvažnijih primena teorije slučajnih procesa jeste predviđanje opterećenosti sistema, mereno kašnjenjima u procesu pružanja usluge, pri čemu su kašnjenja prouzrokovana čekanjem mušterija u redu. Klijenti pristigli u banku, gledaoci na bioskopsku blagajnu karata, kupci na naplatnu kasu u supermarketu, vozači na parking automobila itd., mogu čekanje u redu za pružanje usluge smatrati gubljenjem vremena. Ponavljana i suvišna zadržavanja u redu mogla bi, u krajnjem slučaju, za posledicu imati gubitak mušterija. U tom smislu zanimljiv je i primer koji se može videti u *First*

* Mr Miloš Japundžić, asistent, Visoka poslovna škola strukovnih studija, Novi Sad.

** MSc Ivan Pavkov, saradnik u nastavi, Visoka poslovna škola strukovnih studija, Novi Sad.

Hawaiian Bank, inače jednoj od najvećih banaka američke države Havaji. Naime, u svim ekspoziturama te banke na posterima se može videti slika nasmejane šalterske službenice i upadljiv naslov: „Ako u redu čekate duže od 5 minuta, biće vam isplaćeno 20$". Kratko i jasno. Zaista, ukoliko ste čekali duže od pet minuta, nasmejana službenica pružiće vam novu i sjajnu novčanicu od dvadeset dolara. Ona to čini bez pogovora, a da vi ni reč niste progovorili.

Inače, mušterije ne moraju biti fizički prisutne na mestu pružanja usluge da bi došlo do pojave čekanja u redu. Naime, redovi čekanja se mogu formirati putem telefonskih poziva, npr. prilikom narudžbi kataloških proizvoda, poziva taksi službe, ili poziva korisničkih servisa za softversku podršku, recimo korisničkih servisa mobilne telefonije. Takođe, mnogi ljudi svoje poslove obavljaju preko Interneta, pri čemu se u ovom slučaju dugi redovi javljaju kao posledica usporenog pristupa određenom sajtu usled preopterećenosti servera.

Ispitujući matematičke modele redova čekanja, u mogućnosti smo da damo odgovore na pitanja kao što su:

- Koji deo vremena je svaki server neiskorišćen?
- Koji je očekivani broj mušterija prisutnih u redu?
- Koje je očekivano vreme koje će mušterija provesti čekajući u redu?
- Koja je raspodela verovatnoće broja mušterija prisutnih u redu?
- Koja je raspodela verovatnoće mušterijinog vremena provedenog čekajući u redu?
- Ako menadžer banke želi da obezbedi da samo 1% od svih mušterija mora da čeka na pružanje usluge duže od 5 minuta, koliko blagajnika bi trebalo zaposliti?

Organizacije dizajniraju sopstvene redove čekanja razmatrajući posledice postojanja mušterija koje čekaju u redu, nasuprot troškovima povećanja servisnih kapaciteta. Teorija redova čekanja, čiji začeci datiraju iz prvih godina prošlog veka, pruža mnoštvo analitičkih modela, koji u kombinaciji sa metodama optimizacije, mogu olakšati donošenje odluka u tom pravcu.

2. MATEMATIČKI MODELI REDOVA ČEKANJA

Da bi opisali redove čekanja u određenom sistemu, moramo najpre precizirati ulazne i izlazne procese. Neki od primera ulaznih i izlaznih procesa su dati u Tabeli 1.

Tabela 1. Primeri ulaznih i izlaznih procesa

Sistem	Ulazni proces	Izlazni proces
Banka	Klijenti pristižu u banku	Blagajnici uslužuju klijente
Aerodrom	Avioni pristižu u vazdušni prostor oko aerodroma	Avioni sleću na aerodromsku pistu
Bolnica	Pacijenti pristižu u bolnicu	Doktori pregledaju pacijente
Dispečerska taksi služba	Pristižu telefonski pozivi za vozilo	Dispečeri šalju slobodno vozilo na traženu adresu

Ulazni proces se obično naziva **proces pristizanja (mušterija)**. U svim modelima koje ispitujemo, pretpostavljamo da se u svakom datom vremenskom trenutku može pojaviti najviše jedna mušterija. Takođe, pretpostavljamo da na proces pristizanja ne utiče broj mušterija trenutno prisutnih u sistemu. U kontekstu banke to bi značilo da, bez obzira da li je u banci prisutno 500 ili 5 ljudi, proces pristizanja ostaje nepromenjen. U tom slučaju ovaj proces opisujemo navodeći raspodelu verovatnoće vremena proteklog između sukcesivnih pristizanja mušterija.

Za opisivanje **izlaznog procesa**, koji se često naziva **proces pružanja usluge**, navodimo raspodelu verovatnoće vremena potrebnog za usluživanje mušterije. U većini slučajeva pretpostavljamo da ta raspodela ne zavisi od broja mušterija trenutno prisutnih u sistemu. Ovo implicira, na primer, da server neće raditi brže ukoliko se povećava broj mušterija u sistemu. Inače, razlikujemo dva uređenja servera: serveri povezani u paralelnu vezu i oni povezani u rednu vezu. U prvom slučaju, svi serveri obezbeđuju isti tip usluge, a da bi mušterija kompletirala uslugu dovoljno je da bude uslužen samo od jednog (bilo kojeg) servera. Na primer, blagajnici u bankama su povezani u paralelnu vezu i svaki blagajnik može da pruži željenu uslugu. U drugom slučaju, mušterija mora da bude uslužena od strane svih servera da bi kompletirala uslugu. Fabrika za proizvodnju automobila je primer jednog takvog sistema, gde se automobil montira na svakoj od fabričkih traka.

Da bi u potpunosti opisali sisteme koje ćemo proučavati, moramo takođe da opišemo redosled usluživanja mušterija u odnosu na momenat pristizanja u sistem. Najčešće korišćeni poredak usluživanja mušterija je **FCFS poredak** (*first come, first served*), kod kojeg se mušterije uslužuju po redosledu pristizanja. Kod **LCFS poretka** (*last come, first served*) mušterija koja je poslednja pristigla se prvo uslužuje. Ako izlaženje iz lifta posmatramo kao pružanje usluge, onda lift pun ljudi ilustruje LCFS poredak. Međutim, ponekad redosled u kojem mušterije pristižu u sistem ne utiče na redosled usluživanja. To bi bio slučaj ako bi naredni kandidat za usluživanje bio slučajno izabran od strane onih mušterija koje čekaju u redu na pružanje usluge. Takav poredak usluživanja je poznat kao **SIRO poredak** (*service in random order*). U situacijama kada nije bitno precizirati redosled usluživanja mušterija, radi se o **GD poretku** (*general queue discipline*), to je poredak koji razmatramo u svim modelima koje ispitujemo.

Konačno, spomenimo i **poretke prioriteta**. Kod ovih poredaka svaka pristigla mušterija se klasifikuje u jednu od nekoliko kategorija. Pri tome je svakoj kategoriji dodeljen nivo prioriteta, i unutar svakog nivoa prioriteta mušterije se uslužuju po FCFS poretku. Poreci prioriteta se često koriste u hitnoj pomoći, za određivanje redosleda u kojem će pristigli pacijenti primiti lekarsku pomoć.

Neka broj mušterija prisutnih u sistemu u vremenskom trenutku t bude stanje sistema u tom trenutku. Za $t = 0$, stanje sistema biće jednako broju mušterija inicijalno prisutnih u sistemu. Od velikog interesa za nas je veličina $P_{ij}(t)$, koja predstavlja verovatnoću da će j mušterija biti prisutno u sistemu u vremenskom trenutku t, ako je u trenutku $t = 0$ bilo prisutno i mušterija. Inače, proizilazi da će za većinu sistema u kojima figurišu redovi čekanja, veličina $P_{ij}(t)$, za veliko t, težiti ka π_j, koje je nezavisno od inicijalnog stanja i. U našoj terminologiji veličinu π_j spominjemo kao ravnomerno stanje ili ravnotežnu verovatnoću stanja j.

Za sisteme o kojima ćemo diskutovati, π_j može biti posmatrano kao verovatnoća da će u daljoj budućnosti u sistemu biti prisutno j mušterija. Alternativno, π_j može biti posmatrano kao deo vremena (u daljoj budućnosti) u kome je u sistemu prisutno j mušterija. Takođe, za većinu sistema vrednost $P_{ij}(t)$ će zavisiti od i, tj. broja mušterija inicijalno prisutnih u sistemu. Na pitanje – *Koliko veliko t mora biti pre nego što je ravnomerno stanje aproksimativno dostignuto?* – teško je odgovoriti. Iz tog razloga, u svim analizama redova čekanja pretpostavićemo da je ravnomerno stanje prethodno dostignuto. To nam dopušta da baratamo sa veličinama π_j umesto sa $P_{ij}(t)$.

Ukoliko definišemo da ravnomerno stanje π_j bude jednako sa $\lim_{t\to\infty} P_{ij}(t)$, tada će za ravnotežna stanja važiti relacija:

$$\pi_j = \pi_0 c_j \,, j = 1,2,... \tag{1}$$

pri čemu je:

$$c_j = \frac{\lambda_0 \lambda_1 ... \lambda_{j-1}}{\mu_1 \mu_2 ... \mu_j},$$

dok λ_j i μ_j predstavljaju brzinu pristizanja mušterija u sistem, odnosno brzinu usluživanja mušterija redom, u situaciji kada je u sistemu već prisutno j mušterija.

Koristeći činjenicu da u bilo kom vremenskom trenutku sistem mora biti u nekom stanju, proizilazi:

$$\sum_{j=0}^{\infty} \pi_j = 1,$$

pa dobijamo izraz za π_0:

$$\pi_0 = \frac{1}{1 + \sum_{j=1}^{\infty} c_j}.$$

Na kraju, koristeći relaciju (1) i prethodnu jednakost možemo izračunati i ostala ravnomerna stanja π_1, π_2 ,...

2.1. KENDALOVA NOTACIJA ZA REDOVE ČEKANJA

Da bi opisao sisteme u kojima se javljaju redovi čekanja, Kendal je predložio sledeću notaciju. Naime, svaki takav sistem je opisan sa šest karakteristika:

$$1/2/3/4/5/6.$$

Prva karakteristika se odnosi na prirodu procesa pristizanja i pri tome su uvedene sledeće skraćenice:

M – Slučajne promenljive koje opisuju vreme proteklo između sukcesivnih pristizanja mušterija su nezavisne, identički distribuirane (iid) slučajne promenljive sa eksponencijalnom raspodelom.

Druga karakteristika se odnosi na prirodu procesa pružanja usluge:

M – Slučajne promenljive koje opisuju vreme potrebno za usluživanje mušterija su iid slučajne promenljive sa eksponencijalnom raspodelom.

Treća karakteristika je broj paralelno povezanih servera.

Četvrta karakteristika opisuje redosled usluživanja mušterija:

FCFS = *first come, first served*;

LCFS = *last come, first served*;

SIRO = *service in random order*;

GD = *general queue discipline*.

Peta karakteristika precizira maksimalno dopušten broj mušterija u sistemu (uključujući mušterije koje čekaju u redu kao i one koje su u procesu pružanja usluge).

Na kraju, šesta karakteristika nam daje veličinu populacije iz koje mušterije pristižu. U mnogim važnim modelima 4/5/6 je GD/∞/∞, pa je u tim slučajevima 4/5/6 izostavljeno.

2.2. M / M / 1 / GD / ∞ / ∞ REDOVI ČEKANJA

Kod M / M / 1 / GD / ∞ / ∞ redova čekanja, vreme proteklo između sukcesivnih pristizanja mušterija kao i vreme jedinog servera potrebno za pružanje usluge, modeliramo slučajnim promenljivama sa eksponencijalnom raspodelom (pretpostavljamo da je brzina pristizanja mušterija λ, a brzina pružanja usluge servera μ). Sistemi kod kojih su prisutni ovakvi redovi čekanja, mogu biti modelirani kao procesi rađanja/umiranja, pri čemu su odgovarajući parametri:

$$\lambda_j = \lambda \ (j = 0,1,2,...) \qquad \mu_0 = 0$$

$$\mu_j = \mu \ (j = 1,2,3,...).$$

U ovom slučaju jednakost za ravnotežne verovatnoće π_j , $j = 1,2,...$ ima oblik:

$$\pi_j = \frac{\lambda^j \pi_0}{\mu^j}.$$

Ukoliko definišemo $\rho = \frac{\lambda}{\mu}$, onda ρ predstavlja deo vremena u kome je server zauzet. Iz tog razloga veličinu ρ zovemo **iskorišćenost servera**. Uz pretpostavku $0 \le \rho < 1$, jednostavnim izračunavanjem dobijamo:

$$\pi_0 = 1 - \rho,$$

kao i:

$$\pi_j = \rho^j (1-\rho).$$

Inače, ukoliko je $\rho \geq 1$, tj. ukoliko je brzina pristizanja mušterija λ najmanje jednaka brzini pružanja usluge μ, onda verovatnoće ravnotežnih stanja neće postojati. Zato u daljoj analizi pretpostavljamo da je zadovoljen uslov $\rho < 1$, koji nam obezbeđuje postojanje verovatnoća ravnotežnih stanja.

U nastavku rada se bavimo izračunavanjem metričkih karakteristika ovog sistema kao što su:

L = prosečan broj mušterija prisutnih u sistemu,

L_q = prosečan broj mušterija koje čekaju u redu,

L_s = prosečan broj mušterija u procesu pružanja usluge,

W = prosečno vreme koje mušterija provede boraveći u sistemu,

W_q = prosečno vreme koje mušterija provede čekajući u redu,

W_s = prosečno vreme koje mušterija provede u procesu pružanja usluge.

Pretpostavljajući da je ravnotežno stanje dostignuto, prosečan broj mušterija prisutnih u sistemu (L) je dat sa:

$$L = \sum_{j=0}^{\infty} j\pi_j = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}.$$

Što se tiče prosečnog broja mušterija koje čekaju u redu L_q, možemo primetiti da, ako je sistem prazan ili je samo jedna mušterija prisutna, niko ne čeka u redu. Sa druge strane, ukoliko je u sistemu prisutno j mušterija ($j > 1$), tada tačno $j - 1$ mušterija čeka u redu. Prema tome, ako je ravnotežno stanje dostignuto, imamo:

$$L_q = \sum_{j=1}^{\infty}(j-1)\pi_j = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)}.$$

Takođe, na sličan način dobijamo i izraz za prosečan broj mušterija u procesu pružanja usluge L_s :

$$L_s = 0\pi_0 + 1(\pi_1 + \pi_2 + \ldots) = \rho.$$

Budući da svaka mušterija koja je prisutna u sistemu ili čeka u redu, ili je u procesu pružanja usluge, za sve redove čekanja (ne samo M / M / 1 / GD / ∞ / ∞) važi $L = L_q + L_s$.

Ostale metričke karakteristike sistema W, W_q i W_s se dobijaju korišćenjem sledeće teoreme:

Teorema: Za sve redove čekanja u kojima postoje raspodele ravnotežnih stanja, važe sledeće relacije:

$$L = \lambda W$$

$$L_q = \lambda W_q$$

$$L_s = \lambda W_s.$$

Na taj način dobijamo:

$$W = \frac{1}{\mu-\lambda}$$

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$$

$$W_s = \frac{1}{\mu}$$

Pošto je vreme koje mušterija provede u sistemu jednako vremenu provedenom u redu uvećanim za vreme usluživanja, za sve redove čekanja važi relacija: $W = W_q + W_s$.

2.3. M / M / 1 / GD / C / ∞ REDOVI ČEKANJA

Najpre primetimo da su ovi redovi čekanja u stvari M / M / 1 / GD / ∞ / ∞ redovi sa kapacitetom od c mušterija, pri čemu postoji izuzetak – kada je u tom sistemu prisutno tačno c mušterija, sva pristizanja mušterija u sistem bivaju zaustavljena, a te mušterije izgubljene za sistem. Prema tome, i ovaj sistem može biti modeliran kao proces rađanja/umiranja, pri čemu su u ovom slučaju odgovarajući parametri:

$$\lambda_j = \lambda$$

$$\lambda_c = 0$$

$$\mu_0 = 0$$

$$\mu_j = \mu$$

$$(j = 0,1,\ldots,c\text{-}1)$$

$$(j = 1,2,\ldots,c)$$

Budući da je $\lambda_c = 0$, stanje sistema nikada neće biti $c + 1$. Kao u prethodnom slučaju, sada pretpostavljajući da je $\rho \neq 1$, ravnotežne verovatnoće su date sa:

$$\pi_0 = \frac{1-\rho}{1-\rho^{c+1}}$$

$$\pi_j = \rho^j \pi_0 \ (j = 1,2,\ldots,c)$$

$$\pi_j = 0 \ (j = c+1, c+2, \ldots).$$

Kombinujući ove jednakosti sa činjenicom da je sada $L = \sum_{j=0}^{c} j\pi_j$, dobijamo:

$$L = \frac{\rho\left[1-(c+1)\rho^c + c\rho^{c+1}\right]}{\left(1-\rho^{c+1}\right)\left(1-\rho\right)}.$$

Ukoliko je $\rho = 1$, tada za sve c_j u jednakosti (1) važi: $c_j = 1$, pa sve ravnotežne verovatnoće π_j moraju biti jednake. Sa druge strane, s obzirom da sistem može biti u jednom od $c + 1$ stanja $(0,1,2,\ldots,c)$, proizilazi:

$$\pi_j = \frac{1}{c+1} \ (j = 0,1,2,\ldots,c),$$

kao i:

$$L = \frac{c}{2}.$$

Slično kao kod M / M / 1 / GD / ∞ / ∞ redova čekanja imamo:

$$L_s = 0\pi_0 + 1(\pi_1 + \pi_2 + \ldots + \pi_c) = \rho,$$

a L_q možemo odrediti iz relacije $L_q = L - L_s$.

Određivanje veličina W i W_q se u ovom slučaju donekle razlikuje nego kod M / M / 1 / GD / ∞ / ∞ redova. Naime, kod sistema sa takvim redovima λ, koja se pojavljuje u teoremi, predstavlja prosečan broj mušterija po jedinici vremena koje pristignu u sistem. Međutim, kod sistema sa konačnim kapacitetom mušterija, λ takođe predstavlja prosečan broj mušterija po jedinici vremena, ali $\lambda\pi_c$ od tih mušterija zatiče sistem popunjen i odlazi. Prema tome, prosečno $\lambda - \lambda\pi_c = \lambda(1-\pi_c)$ mušterija po jedinici vremena će u stvari pristići u sistem. Kombinujući ovu činjenicu sa teoremom dobijamo:

$$W = \frac{L}{\lambda(1-\pi_c)}$$

$$W_q = \frac{L_q}{\lambda(1-\pi_c)}$$

$$W_s = \frac{L_s}{\lambda(1-\pi_c)}.$$

Na kraju, primetimo da će za M / M / 1 / GD / c / ∞ redove čekanja ravnotežna stanja postojati čak i ukoliko je $\rho \geq 1$. To sledi iz činjenice da konačni kapacitet sistema kontroliše nagomilavanje mušterija prisutnih u sistemu.

LITERATURA

[1] Chiang, M., Sutivong, A., Boyd, S., (2002) *Efficient Nonlinear Optimizations of Queuing Systems*, Stanford University, Electrical Engineering Department.

[2] Cooper, R. B., (1981) *Introduction to Queueing Theory*, Second Edition. New York, North Holland.

[3] Japundžić, M., (2008) *Redovi čekanja*, XXXV simpozijum o operacionim istraživanjima, SYM-OP-IS, Sokobanja, 14–17. septembar 2008, str. 35–38, Univerzitet u Beogradu, Saobraćajni fakultet.

[4] Sheldon, R. M., (2002) *Introduction to Probability Models*, San Diego, Academic Press.

[5] Winston, L. W., (1994) *Operations Research: Applications and Algorithms*, Belmont, California, Duxbury Press.

Primljeno: 23.11.2011.
Odobreno: 05.12.2011.