

Teória informačných sietí

Prvé vydanie

Martin Klimo
Žilinská univerzita v Žiline

Jana Uramová
Žilinská univerzita v Žiline

Juraj Smieško
Žilinská univerzita v Žiline

Obsah

KAPITOLA 1	Modelovanie tokov v sieťach	1
1.1	Deterministický popis IP tokov	4
1.1.1	Použitie lineárneho (σ, ρ) ohraničenia	4
1.1.2	Lineárne ohraničenie pre periodický tok	8
1.1.3	Subaditívne vstupné krivky	10
1.2	Bernoulliho model IP toku	13
1.2.1	Pravdepodobnostné charakteristiky	16
1.2.2	Rozdelenie pravdepodobnosti medzier	20
1.2.3	Identifikácia Bernoulliho procesu	22
1.2.4	Rozdelenie pravdepodobnosti paketových zhlukov	23
1.3	Zhlukové On/Off modely	25
1.3.1	2-stavový Markovov reťazec	26
1.3.2	Markovov modulovaný regulárny proces (MMRP)	30
1.3.3	Rozdelenie pravdepodobnosti medzier a zhlukov	34
1.3.4	Porovnanie Bernoulliho procesu a MMRP	36
1.4	Poissonov model	39
1.4.1	Momenty a charakteristiky	40
1.4.2	Rozdelenie medzier	42
1.4.3	Multiplex a routing	44
1.5	Efektívna šírka pásma (EB)	44
1.5.1	Základné vlastnosti Eektívnej šírky pásma	44
1.5.2	EB pre Regulárny proces	47
1.5.3	EB pre Bernoulliho proces	47
1.5.4	EB pre Poissonov proces	48
1.5.5	EB pre Markovov modulovaný proces (MMP)	49
1.5.6	EB pre MMP s diskretným časom	50
1.5.7	EB pre MMP so spojitým časom	52
KAPITOLA 2	Modelovanie spoločného komunikačného prostredia	55
2.1	Charakteristiky spoločného komunikačného prostredia	55
2.1.1	Model prostredia so spätnou väzbou	56
2.1.2	Rozdelenie sietí podľa typu riadenia a prenosu	57
2.2	Asynchrónne prostredie s úplnou informáciou	58
2.2.1	Model s diskretným časom $Geo/D/1/\infty$	59
2.2.2	Model s diskretným časom $Geo/Geo/1/\infty$	63
2.2.3	Markovove reťazce so spojitým časom	67
2.2.4	Zovšeobecnený proces vzniku a zániku	72
2.2.5	Analýza Markovovho modelu	74
2.2.6	Markovov model	78
2.2.7	Nemarkovove modely	78

2.3	Použitie princípu veľkých odchýlok	93
2.3.1	Systém s konštantnou kapacitou	93
2.3.2	Dimenzovanie $Geo/D/1/\infty$	96
2.3.3	$M/D/1/\infty$	96
2.3.4	Systém s variabilnou kapacitou	98
2.3.5	$M/M/1/\infty$	100
2.3.6	Markovove systémy s MMPP vstupom	103
2.3.7	$M-M/M/1/\infty$	104
2.4	Sieťový kalkul	106
2.4.1	Dimenzovanie pomocou (σ, ϱ) - Algebry	106
	Literatúra	109

1

Modelovanie tokov v sieťach

OBSAH KAPITOLY

1.1	Deterministický popis IP tokov	4
1.1.1	Použitie lineárneho (σ, ρ) ohraničenia	4
1.1.2	Lineárne ohraničenie pre periodický tok	8
1.1.3	Subaditívne vstupné krivky	10
1.2	Bernoulliho model IP toku	13
1.2.1	Pravdepodobnostné charakteristiky	16
1.2.2	Rozdelenie pravdepodobnosti medzier	20
1.2.3	Identifikácia Bernoulliho procesu	22
1.2.4	Rozdelenie pravdepodobnosti paketových zhlukov	23
1.3	Zhlukové On/Off modely	25
1.3.1	2-stavový Markovov reťazec	26
1.3.2	Markovov modulovaný regulárny proces (MMRP)	30
1.3.3	Rozdelenie pravdepodobnosti medzier a zhlukov	34
1.3.4	Porovnanie Bernoulliho procesu a MMRP	36
1.4	Poissonov model	39
1.4.1	Momenty a charakteristiky	40
1.4.2	Rozdelenie medzier	42
1.4.3	Multiplex a routing	44
1.5	Efektívna šírka pásma (EB)	44
1.5.1	Základné vlastnosti Efektívnej šírky pásma	44
1.5.2	EB pre Regulárny proces	47
1.5.3	EB pre Bernoulliho proces	47
1.5.4	EB pre Poissonov proces	48
1.5.5	EB pre Markovov modulovaný proces (MMP)	49
1.5.6	EB pre MMP s diskretným časom	50
1.5.7	EB pre MMP so spojitým časom	52

VZDELÁVACIE CIELE

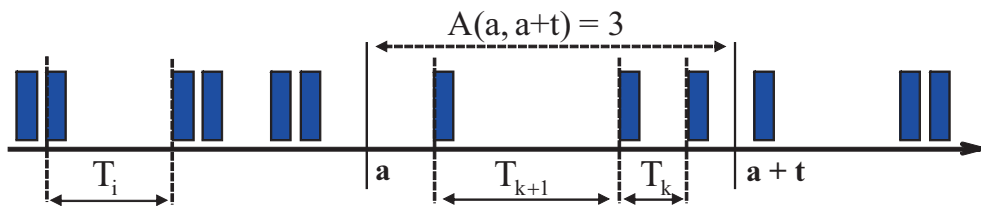
Po dôkladnom preštudovaní tejto kapitoly budete vedieť:

1. Naučiť sa niečo o modelovaní tokov v IP sieťach

Úvod

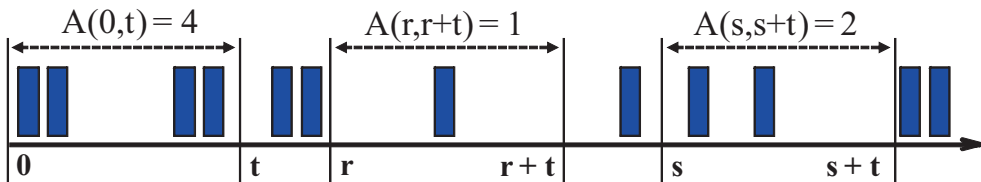
Dôležitou oblasťou výskumu v kontexte sieťových technológií je vývoj modelov IP prevádzky, ktoré môžu byť aplikované na internet, alebo všeobecnejšie na akúkoľvek komunikačnú sieť. Význam týchto modelov môžeme rozdeliť do dvoch skupín. V prvom rade, modely IP prevádzky sú potrebné ako vstup do sieťových simulácií. Tieto simulácie musia byť vykonávané pre účel štúdia a validácie sieťových algoritmov a protokolov, ktoré budú aplikované do reálnej prevádzky a pre účel analýzy ako IP prevádzka reaguje na rôzne podmienky a situácie v sieti, napr. na preťaženie siete a pod. Preto je dôležité, aby navrhnuté modely čo najvernejšie odrážali relevantné charakteristiky prevádzky, ktorú by mali reprezentovať. V druhom rade, dobrý model môže viesť k lepšiemu pochopeniu charakteristík prevádzky, čo môže pomôcť pri navrhovaní smerovačov a iných zariadení, ktoré riadia sieťovú prevádzku.

Paketový tok generovaný nejakým sieťovým zdrojom je jedným zo základných elementov IP siete. Existuje viacero koncepcií ako popísať tok udalostí. Jedným zo spôsobov je popis pomocou tzv. vstupného procesu $A(r, s)$, ktorý predstavuje kumulatívny počet výskytu udalostí (paketov) v intervale $\langle r, s \rangle$. Iný spôsob je použitie premenných T_i , ktoré popisujú dĺžku trvania časovej medzery medzi výskytom jednotlivých udalostí (paketov):



Obr. 1.1: Vstupný proces popísaný pomocou procesu $A(t)$ a premenných T_i .

Vo všeobecnosti budeme predpokladať, že tok predstavujúci IP prevádzku, je náhodný, pričom $A(r, s)$ bude predstavovať nejaký stochastický proces a T_i odpovedajúce náhodné premenné. Ďalším zjednodušením bude uvažovať iba o tzv. stacionárnych náhodných procesoch. Zjednodušene povedané, všetky pravdepodobnostné charakteristiky stacionárneho procesu sú invariantné vzhľadom na čas. Nezáleží, v ktorom časovom intervale sme začali proces pozorovať, záleží iba na dĺžke pozorovania.



Obr. 1.2: Stacionárny vstupný tok

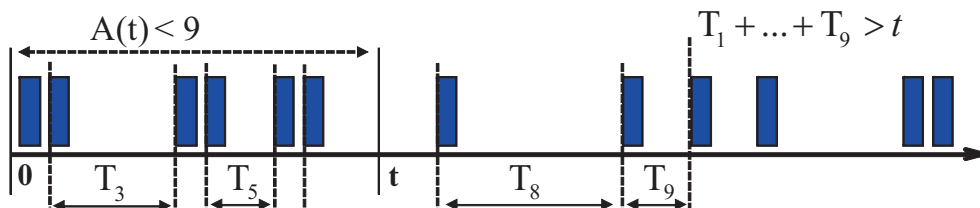
Pre stacionárny proces zavedieme zjednodušené označenie $A(t) = A(0, t)$.

Medzi rozdeleniami pravdepodobnosti stacionárneho procesu $A(t)$ a náhodných premenných T_i , ktoré popisujú medzery medzi udalosťami, platí nasledujúci vzťah. Ak za čas t sa v toku vyskytne menej než k udalostí, potom súčet prvých k

medzier je väčší ako doba t . Preto pravdepodobnosti nastatia uvedených udalostí sa rovnajú.

Pre stacionárny proces $A(t)$ a náhodné premenné T_i popisujúce medzery platí

$$\forall k; \quad \Pr(A(t) < k) = \Pr(T_1 + \dots + T_k > t)$$

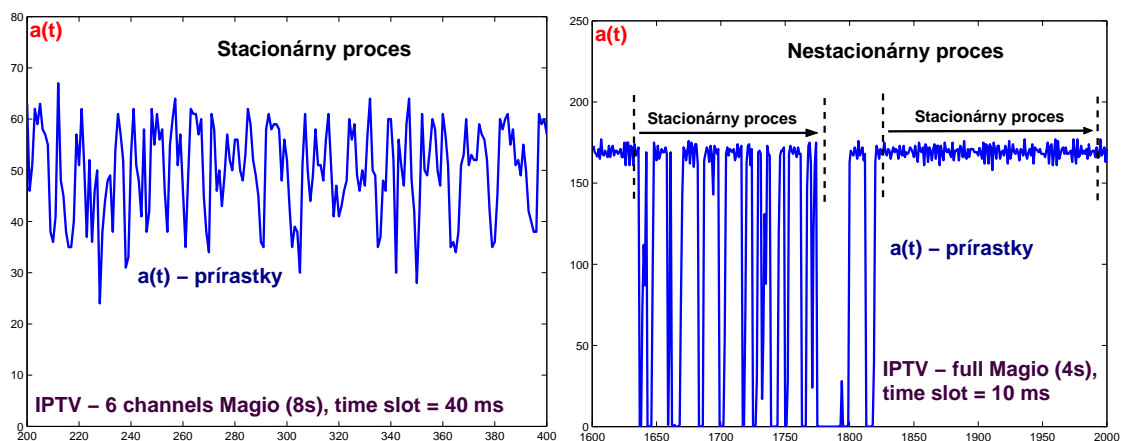


Obr. 1.3: Vzťah medzi procesom $A(t)$ a medzerami T_i

Ak predpokladáme diskretný čas, vstupný proces $A(t)$ sa skladá z tzv. prírastkov $a(i)$, ktoré zaznamenávajú výskyt paketov v čase i . Predpokladáme, že v čase 0 sa nevyskytol žiaden paket:

$$A(t) = \sum_{i=0}^t a(i), \quad a(t) = A(t) - A(t-1), \quad A(0) = a(0) = 0$$

Na nasledujúcich obrázkoch je ukážka merania prírastkov IPTV tokov, ktoré vychádzajú zo zdroja Mágio (ST-com). Na prvom obrázku je meranie 6 kanálov (4s). Podľa priebehu môžeme tok považovať za stacionárny. Na druhom obrázku je nameraná plná prevádzka Mágia (4s). Vidíme, že sa priebeh toku výrazne zmenil. Aj keď ide o nestacionárny tok, dá sa rozdeliť na stacionárne časti.



Obr. 1.4: Príklad stacionárneho a nestacionárneho procesu

Na otestovanie stacionarity procesu by sme potrebovali štatistické nástroje, čo je však mimo rámec tejto publikácie. Čisto laicky však môžeme povedať, že proces je stacionárny ak sa istým spôsobom pravidelne "vlní".

Ak je proces stacionárny, náhodné premenné T_i , ktoré popisujú medzery medzi udalosťami majú rovnaké rozdelenie pravdepodobnosti. Rovnako aj prírastky

toku $a(i)$ majú rovnaké rozdelenie. Ak navyše sú prírastky navzájom nezávislé, potom hovoríme, že proces $A(t)$ je stacionárny s nezávislými prírastkami, resp. (i.i.d) proces (*identical and independent distribution increments*).

Stacionárny proces s nezávislými prírastkami

Ak náhodný proces $A(t)$ má rovnako rozdelené a navzájom nezávislé prírastky $a(i)$, hovoríme, že proces je stacionárny a nezávislý, resp. (i.i.d) proces. V takomto prípade sú aj medzery medzi udalosťami T_i rovnako rozdelené a navzájom nezávislé.

1.1 DETERMINISTICKÝ POPIS IP TOKOV

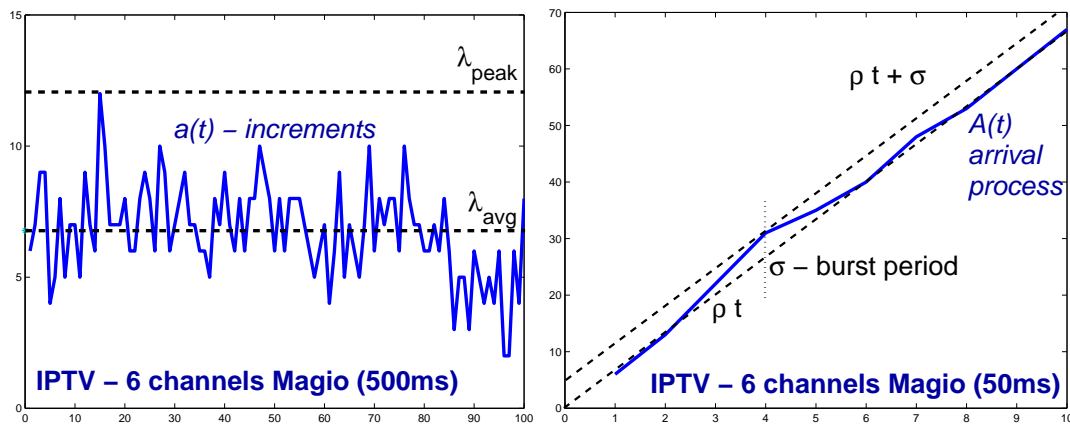
Zostaviť stochastický model pre nameraný IP tok je často komplikované, preto sa budeme najprv zaoberať deterministickými modelmi, ktoré využívajú subaditívne horné ohraňovania pre daný tok.

1.1.1 Použitie lineárneho (σ, ρ) ohraňovania

Na úplne základný popis IP prevádzky sa používajú dva parametre, intenzita toku, resp. stredný počet výskytu udalostí za časovú jednotku (*time slot*), $\lambda_{avg} = \rho$ (*average rate*) a tzv. špičková rýchlosť $\lambda_{peak} = p$ (*peak rate*). Tretí parameter σ (*burst period*), zaviedol Chang v [2]. Predstavuje maximálny rozdiel medzi vstupným procesom $A(t)$ a funkciou intenzity ρt . Burst periódu môžeme chápať ako najväčšiu "vlnu" v dátach, ktorá môže vstúpiť do daného zariadenia.

Pre dané meranie dĺžky n tieto parametre ľahko vypočítame:

$$\lambda_{avg} = \rho = \frac{A(n)}{n}, \quad \lambda_{peak} = \max_{0 \leq t \leq n} [a(t)], \quad \sigma = \max_{0 \leq t \leq n} [A(t) - \rho t]$$



Obr. 1.5: Základné parametre pre prírastky $a(t)$ a kumulatívny proces $A(t)$

Pomocou intenzity toku a burst periódy vieme určiť priamku $\sigma + \rho t$, ktorá daný tok, s ohľadom na dĺžku merania, zhora ohraňuje.

(σ, ρ)
ohraničenie

Hovoríme, že proces $A(t)$ je (σ, ρ) zhora ohraničený, ak platí

$$\forall s, t; \quad 0 \leq s \leq t; \quad A(t) - A(s) \leq \rho(t - s) + \sigma$$

PRÍKLAD 1.1

Majme nameraný nejaký paketový tok o dĺžke 15. Prírastky toku v jednotlivých časových slotoch sú zaznamenané v tabuľke:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$a(t)$	2	3	3	4	6	6	5	4	1	2	3	2	2	1	1

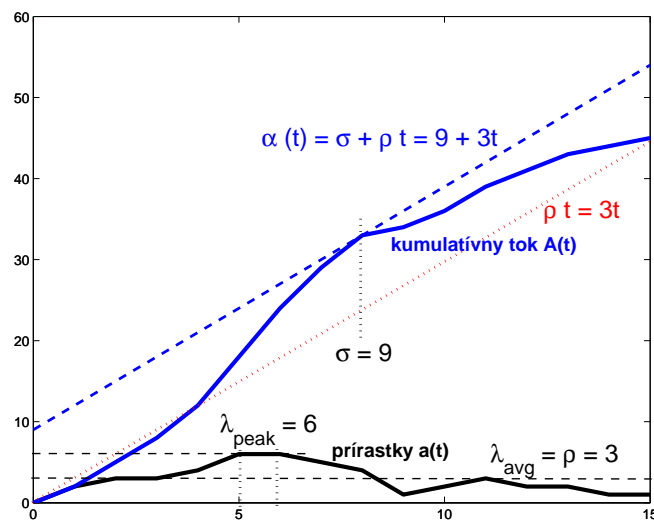
Jednoducho zistíme špičkovú a priemernú intenzitu:

$$\lambda_{peak} = \rho = 6, \quad \lambda_{avg} = \frac{1}{15} \sum_{t=1}^{15} a(t) = \frac{45}{15} = 3$$

Pre výpočet burst parametra musíme vytvoriť kumulatívny tok $A(t)$ a porovnať jeho hodnoty s funkciou ρt , ktorá predstavuje priemernú intenzitu toku za časový interval $\langle 0, t \rangle$.

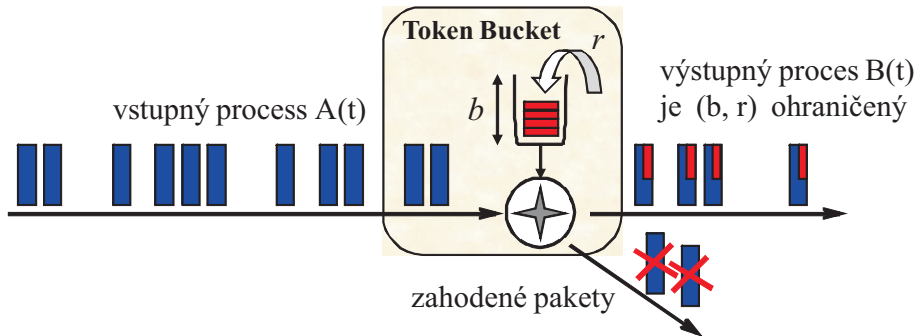
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$A(t)$	2	5	8	12	18	24	29	33	34	36	39	41	43	44	45
ρt	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45
$A(t) - \rho t$	-1	-1	-1	0	3	6	8	9	7	6	6	5	4	2	0

Pre daný tok je burst perióda $\sigma = \max_{0 \leq t \leq n} [A(t) - \rho t] = 9$ pre $t = 8$, a na ohraničenie toku môžeme použiť priamku $\alpha(t) = 9 + 3t$.



Obr. 1.6: Ohraničenie vstupného toku funkciou $\alpha(t) = 9 + 3t$

Na riadenie vstupu IP tokov do chrbticovej siete sa používajú tzv. Token Buckety (TB) alebo Peak limitéry (PR), ktoré sú priamo implementované na vstupných portoch prístupového uzla. Z hľadiska modelovania Token Bucket predstavuje zásobník na povolenia (tokeny) o veľkosti b , ktoré sú generované intenzitou r . Vstupujúci paket je ďalej preposlaný do siete, ak sa v token buckete nachádza aspoň jedno povolenie. V prípade, že zásobník je prázdny, prichádzajúci paket je zahodený. Token Bucket za jednotku času môže vyslať maximálne $b + r$ paketov, preto výstupný tok je ohraničený krivkou $\gamma_{b,r}(t) = b + rt$.



Obr. 1.7: Schéma zariadenia Token Bucket

Zariadenie Peak Limiter je verziou Token Bucketu s parametrom $b = 0$, to znamená, že neexistuje zásobník na povolenia a zásobník prepustí do siete len taký objem pravádzky, koľko dokáže vygenerovať povolení za časovú jednotku. Peak Limiter je vlastne obmedzovač špičkovej kapacity pre danú prevádzku. Na rozdiel od Token Bucketu neumožňuje "našetriť" si kapacitu počas prázdnej periódy v toku. V oboch zariadeniach dochádza k stratám paketov v prípade, keď nie je k dispozícii žiaden token, preto ohraničenie pre výstupný tok musí byť menšie, nanajvýš rovné ohraničeniu vstupného toku.

PRÍKLAD 1.2

Nech daný paketový tok o dĺžke $n = 15$ prechádza cez Peak Limitér, ktorý umožňuje vyslať do siete len prevádzku so špičkovou intenzitou $r = 4$ pakety za časovú jednotku. Určíme prírastky výstupného toku $b(t)$ a lineárne ohraničenie pre výstupný proces $B(t)$:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$a(t)$	2	3	3	4	6	6	5	4	1	2	3	2	2	1	1
$b(t)$	2	3	3	4	4	4	4	4	1	2	3	2	2	1	1
$B(t)$	2	5	8	12	16	20	24	28	29	31	34	36	38	39	40

Parametre výstupného toku sú

$$\lambda_{peak} = 4, \quad \lambda_{avg} = \varrho = \frac{40}{15} = 2.667, \quad \sigma = \max_{0 \leq t \leq n} [B(t) - \varrho t] = 6.667$$

Burst perióda vo výstupnom toku $B(t)$ nastáva v rovnakom časovom okamihu $t = 8$ ako pri vstupnom toku $A(t)$, pričom dosahuje nižšiu hodnotu vďaka obmedzeniu špičkovej intenzity peak limitérom. Výstupný tok je ohraničený priamkou $\beta(t) = 6.667 + 2.667t$, pričom platí $\beta(t) < \alpha(t), \forall t$.

PRÍKLAD 1.3

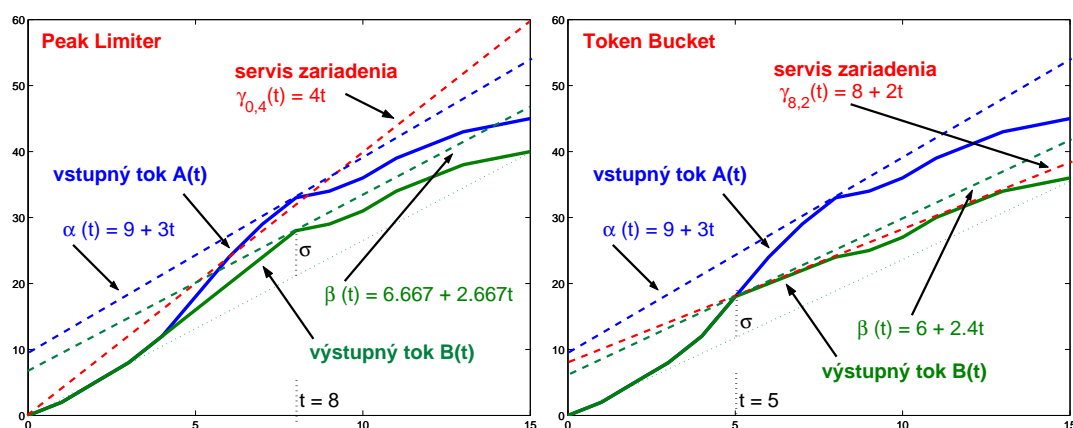
Nech rovnaký paketový tok ako v predchádzajúcom príklade, prechádza cez Token Bucket s parametrami $r = 2$ a $b = 8$. Premenná $tb(t)$ bude popisovať počet tokenov v zásobníku v danom časovom slot. Určíme parametre a lineárne ohraňenie pre výstupný proces $B(t)$:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$a(t)$	2	3	3	4	6	6	5	4	1	2	3	2	2	1	1
$tb(t)$	8	7	6	4	0	0	0	0	1	1	0	0	0	1	2
$b(t)$	2	3	3	4	6	2	2	2	1	2	3	2	2	1	1
$B(t)$	2	5	8	12	18	20	22	24	25	27	30	32	34	35	36

Parametre výstupného toku sú

$$\lambda_{peak} = 6, \quad \lambda_{avg} = \varrho = \frac{36}{15} = 2.4, \quad \sigma = \max_{0 \leq t \leq n} [B(t) - \varrho t] = 6$$

Výstupný tok $B(t)$ je ohraňený $\beta(t) = 6 + 2.4t$, pričom platí $\beta(t) < \alpha(t)$.



Obr. 1.8: Ohraňenie výstupného toku funkciou $\beta(t)$ pri prechode cez Peak Limiter a Token Bucket

Na predchádzajúcich obrázkoch sú zobrazené vstupné a výstupné toky pre zariadenia Peak Limiter a Token Bucket. Výstupné toky sú samozrejme navzájom odlišné. V prvom prípade sme obmedzili iba špičkovú intenzitu, preto ak hodnota prírastkov je menšia alebo rovná $r = 4$, výstupný tok kopíruje v čase vstupný tok. Burst perióda nastane v rovnakom časovom okamihu, ale s menšou hodnotou. V druhom prípade sme okrem povolenej intenzity $r = 2$ mohli využívať aj povolenia z token bucketu, $b = 8$. To znamená, že špičková intenzita toku môže byť až $\lambda_{avg} = 10$ za podmienky, že predtým bola dostatočne dlhý čas slabá prevádzka, menšia alebo rovná ako $r = 2$. Úlohu bucketu na povolenia môžeme chápať ako nejakú rezervu, ktorá sa v prípade príchodu relatívne veľkej záťaže (burst periódy) oproti intenzite r do zariadenia.

V obidvoch zariadeniach dochádza k stratám paketov. V prípade presného merania môžeme počet stratených paketov vyrátať presne rozdielom $A(t) - B(t)$. Pri praktických situáciách môžeme určiť iba horný odhad strát.

Na obrázkoch ďalej vidíme, že pre výstupné toky máme niekoľko lineárnych ohraňení. Funkciu $\beta(t)$ sme získali priamo pomocou výpočtu parametrov toku.

Samozrejme výstupné toky ohraňuje funkcia $\alpha(t)$, ktorá pôvodne popisovala vstupné toky $A(t)$, pričom platí $\beta(t) < \alpha(t), \forall t$. Pribudlo nám ešte ohraňovanie pomocou funkcie $\gamma_{b,r}(t)$, ktoré popisuje servis daného zariadenia, ktoré v čase t_0 môže vyslať do siete maximálne $b + rt_0$ paketov. Preto výstupný tok je ohraňovaný aj priamkou $\gamma_{b,r}(t)$, ktorá ale už nie je v jednoznačnom vzťahu s funkciami $\alpha(t)$ a $\beta(t)$, preto nám zostáva na výber, ktorú z uvedených priamok je rozumnejšie ponechať ako ohraňovanie výstupného toku.

1.1.2 Lineárne ohraňovanie pre periodický tok

V praxi sa môžeme stretnúť s tokmi, ktoré vykazujú istý periodický charakter prenášaných paketov. V ideálnom prípade môžeme uvažovať o deterministickom periodickom toku, ktorého prírastky v čase sa po nejakej perióde p opakujú.

Periodický tok

Vstupný proces $A(t)$ nazveme periodický s periódou $p > 0$, ak pre prírastky toku $a(i)$ platí

$$t > 0; \quad a(t + p) = a(t)$$

Určíme horné ohraňovanie pre rozdiel $A(t) - A(s)$. Číslo $A(p)$ predstavuje počet paketov, ktoré sa vyskytli počas periódy p . Funkcia $\lceil x \rceil$ zaokrúhli reálne číslo x na najmenšie celé číslo, ktoré je väčšie alebo rovné číslu x .

$$A(t) - A(s) \leq \left\lceil \frac{t-s}{p} \right\rceil A(p) \leq \left(\frac{t-s}{p} + 1 \right) A(p) = A(p) + \frac{A(p)}{p}(t-s)$$

Využili sme horné ohraňovanie pre počet periód p , ktoré sa nachádzajú v intervale o dĺžke $t - s$. Odvodenie si vysvetlíme na nasledujúcom príklade.

PRÍKLAD 1.4 Majme daný periodický tok $A(t)$ s periódou p .

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	...
$a(t)$	16	14	4	3	2	1	1	0	16	14	4	3	2	1	1	...
$A(t)$	16	30	34	37	39	40	41	41	57	71	75	78	80	81	82	...

Dĺžka periódy je $p = 8$ a počas periódy sa vyskytne $A(8) = 41$ paketov. Budeme uvažovať o rozdele hodnôt napríklad $A(35) - A(14)$:

$$A(35) - A(14) = [4 \cdot A(8) + A(3)] - [A(8) + A(6)] = 117$$

Rozdiel ohraňujeme pomocou zaokrúhlenia počtu periód v intervale o dĺžke $t - s$ smerom nahor:

$$A(35) - A(14) = 117 \leq \left\lceil \frac{35-14}{8} \right\rceil A(8) = \lceil 2.625 \rceil A(8) = 3 \cdot A(8) = 123$$

V poslednom kroku zväčšíme presný počet periód o jednotku:

$$A(35) - A(14) = 117 \leq 123 \leq (2.625 + 1) \cdot A(8) = 148.625$$

Číslo 148.625 je hodnota horného ohraničenia toku $A(t)$ v bode $t - s = 21$:

$$\alpha_p(t-s) = A(p) + \frac{A(p)}{p}(t-s) = 41 + 5.125 \cdot (t-s) \Rightarrow \alpha_p(21) = 148.625$$

ohraničenie periodického toku

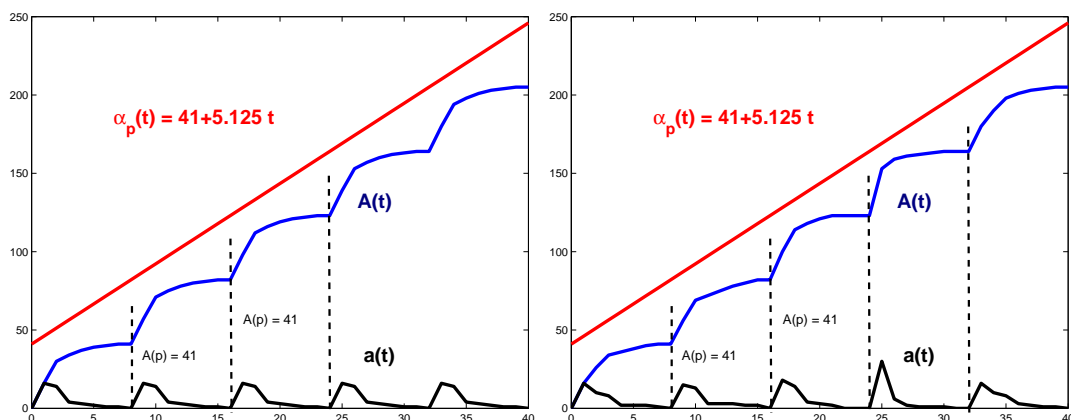
Nech je vstupný proces $A(t)$ periodický s periódou $p > 0$. Proces $A(t)$ je zhora ohraničený priamkou $\alpha_p(t)$:

$$A(t) - A(s) \leq \alpha_p(t-s) = A(p) + \frac{A(p)}{p}(t-s) = \sigma + \varrho(t-s)$$

kde $\sigma = A(p)$ je burst parameter a $\varrho = A(p)/p$ priemerná intenzita toku.

Na určenie burst parametra sme nepoužili namerané hodnoty, ale sme využili periodicitu toku. Pri priamom výpočte musí mať meranie dĺžku celočíselného násobku periódy p , v opačnom prípade môžeme získať zavádzajúce ohraničenie v závislosti na tom, v akej časti periódy bolo meranie ukončené.

Uvažovaný periodický tok je veľmi idealistický. V praxi dochádza k narušeniu periodicity vďaka rôznym náhodným faktorom. Napríklad výskyt paketov počas jednej periódy je náhodný, ale celkový počet paketov v periódach ostáva rovnaký. Aj v takomto prípade je použitie funkcie $\alpha_p(t)$ vhodným ohraničením:

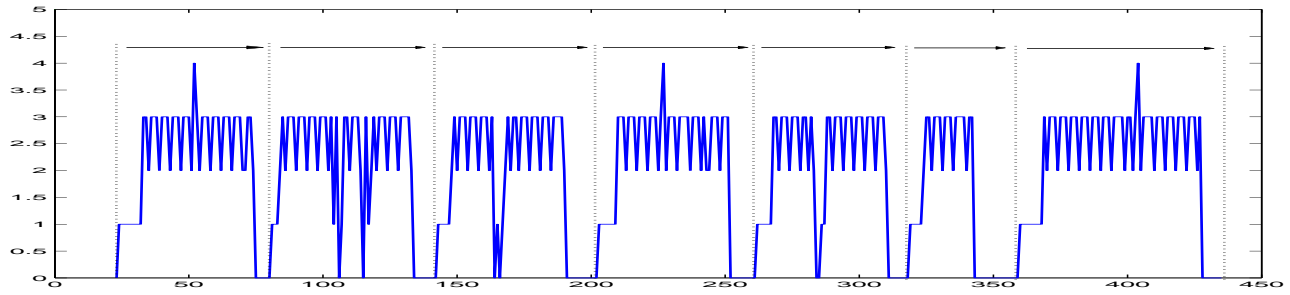


Obr. 1.9: Lineárne ohraničenie dvoch typov periodického toku s rovnakou hodnotou $A(p)$

Na druhom grafe vidíme, že rovnaké ohraničenie $\alpha_p(t)$ ostalo odolné voči náhodnému výskytu paketov počas jednotlivých periód.

V reálnej prevádzke sa môže tok, v ktorom výskyt paketov je náhodný bez obmedzenia na nejakú danú dĺžku periód, javiť v istom zmysle ako periodický. Na nasledujúcom obrázku je ukážka IPTV toku, ktorý predstavuje program Markíza zo zdroja ST-COM Mágio.

Časový slot má dĺžku 10 ms, celková dĺžka ukážky je 4500ms=4.5s, prírastky toku v slote majú hodnoty $a(i) = 0, 1, \dots, 4$. Vidíme, že tok je v istom zmysle periodický. Tok sme rozdelili do 7 periód, pričom za ukončenie periódy sme považovali sekvenciu viacej než 4 núl. Nová perióda potom začína prvým nenulým



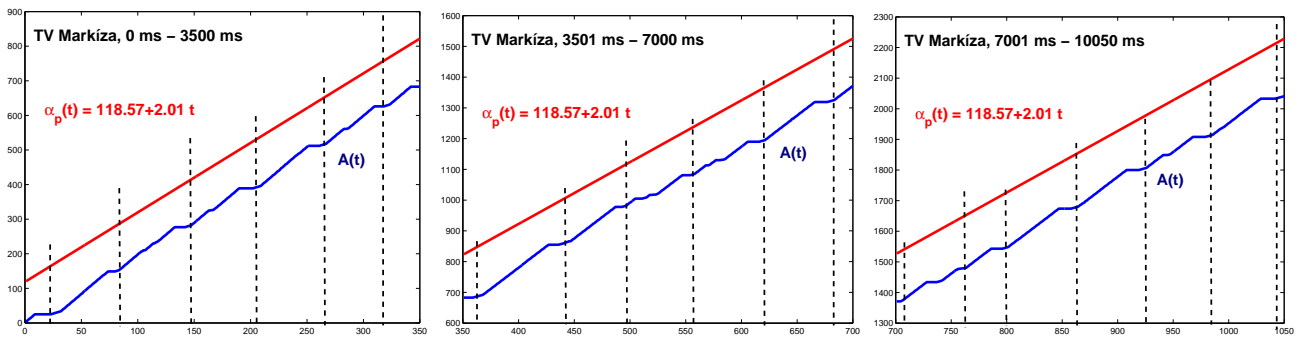
Obr. 1.10: Záznam IPTV Markíza zo zdroja ST=COM Mágioa o dĺžky 4.5s

prírastkom. V nasledujúcej tabuľke sú zaznamenané veľkosti jednotlivých periód p_i , počet paketov v perióde $A(p_i)$ a intenzita výskytu paketov ϱ_i v perióde:

i - perióda	1	2	3	4	5	6	7
p_i [ts=10ms]	57	62	60	59	57	41	77
$A(p_i)$ [paket]	124	128	112	123	114	57	172
ϱ_i [paket/ts]	2.17	2.06	1.87	2.08	2.00	1.39	2.23

Pre lineárne ohraňenie $\alpha_p(t)$ periodického toku použijeme priemery z name-
raných hodnôt, $\hat{p} = 59$ [ts] a $\hat{A}(p) = 118.574$ [paket]:

$$\alpha_p(t) = \hat{A}(p) + \frac{\hat{A}(p)}{\hat{p}} t = 118.574 + 2.0097t$$



Obr. 1.11: Lineárne ohraňenie pre periodický tok TV Markíza o dĺžke 10,05s

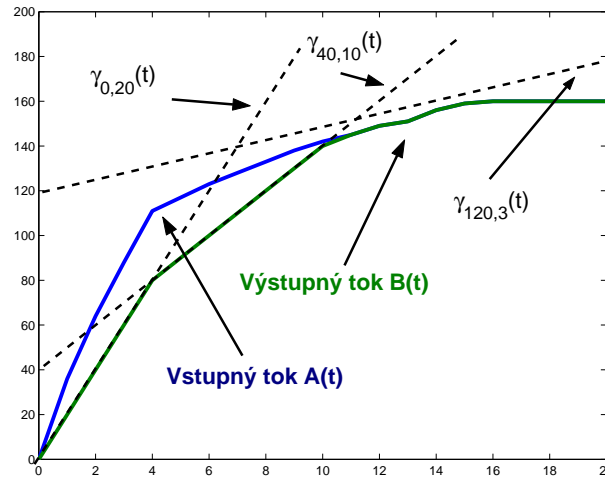
Aj keď sme na určenie funkcie $\alpha_p(t)$ použili iba prvých 7 periód toku (5.4 s), vďaka nemennosti toku (stacionarite), ohraňenie je vhodné pre celú dĺžku merania. Vďaka hodnote burst parametra funkcia $\alpha_p(t)$ garantuje horné ohraňenie s rezervou o veľkosti počtu výskytov paketov počas jednej periódy.

1.1.3 Subaditívne vstupné krivky

V predchádzajúcich kapitolách sme ukázali, že pre tok vychádzajúci z daného zariadenia môžeme určiť niekoľko lineárnych ohraňení $\gamma_{i,j}(t)$. Pre zachovanie linearitý musíme z nich vybrať jedno najvhodnejšie. Avšak pre získanie lepšieho modelu pre tok je optimálne použiť minimum z ohraňení $\gamma_{i,j}(t)$.

Na obrázku je znázornený prechod vstupného toku $A(t)$ cez tri Token Buckety, ktorých práca je popísaná pomocou funkcií $\gamma_{i,j}(t)$. Výstupný tok $B(t)$ je ohraňovaný všetkými troma funkciami $\gamma_{i,j}(t)$. Pre získanie presnejšieho modelu pre výstupný tok použijeme funkciu $\alpha(t)$, ktorá je funkcionálnym minimom všetkých ohraňujúcich funkcií,

$$\alpha(t) = \min [\gamma_{0,20}(t), \gamma_{40,10}(t), \gamma_{120,3}(t)],$$



Obr. 1.12: Prechod toku $A(t)$ cez tri Token Buckety

subaditívna funkcia

Funkcia $f(x)$ je subaditívna, ak platí: $\forall s, t \geq 0; \quad f(t+s) \leq f(t) + f(s)$

PRÍKLAD 1.5

Funkcia $f(x) = x$ je subaditívna:

$$f(t+s) = t+s \leq f(t) + f(s)$$

Funkcia $f(x) = x^2$ nie je subaditívna:

$$f(t+s) = (t+s)^2 = t^2 + 2ts + s^2 \geq t^2 + s^2 = f(t) + f(s)$$

Funkcia $f(x) = \sqrt{x}$ je subaditívna, overíme nerovnosť

$$\sqrt{t+s} \leq \sqrt{t} + \sqrt{s} \quad \Rightarrow \quad t+s \leq t + 2\sqrt{ts} + s \quad \Rightarrow \quad 0 \leq 2\sqrt{ts}$$

subaditívna vstupná krivka

Subaditívnu funkciu $\alpha(t)$ nazveme subaditívna vstupná krivka toku $A(t)$, ak

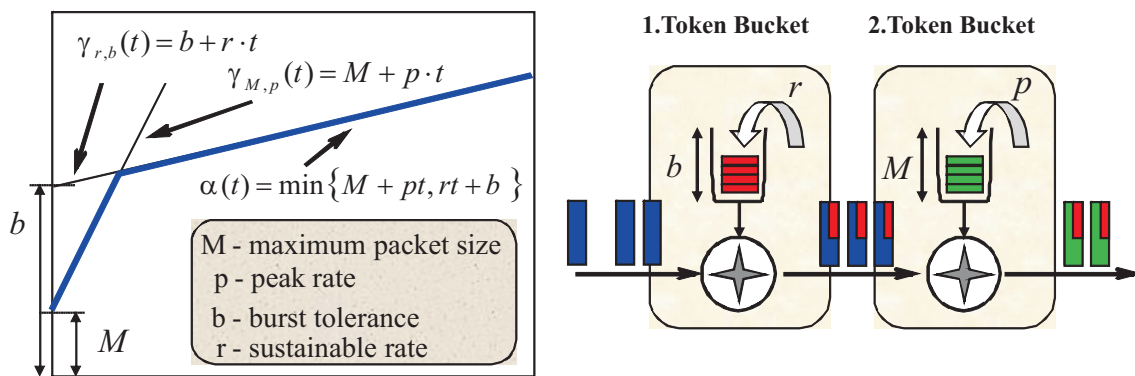
$$\forall s, t \geq 0; \quad 0 \leq s \leq t; \quad A(t) - A(s) \leq \alpha(t-s)$$

Treba si uvedomiť, že všetky funkcie $\alpha(t)$, ktoré sú funkcionálnym minimom lineárnych ohraňčení $\gamma_{b_i, r_i}(t)$, $\alpha(t) = \min_{\forall i} [\gamma_{b_i, r_i}(t)] = \min_{\forall i} [b_i + r_i t]$ sú konkávne funkcie, a dá sa dokázať, že konkávna funkcia je aj subaditívna.

Podľa odporúčaní ITU-T [3] sa pre služby IntServ používa tzv. T-SPEC (špecifikácia) pomocou parametrov: M - maximálna veľkosť paketu, p - špičková intenzita, b - burst parameter a r - priemerná intenzita.

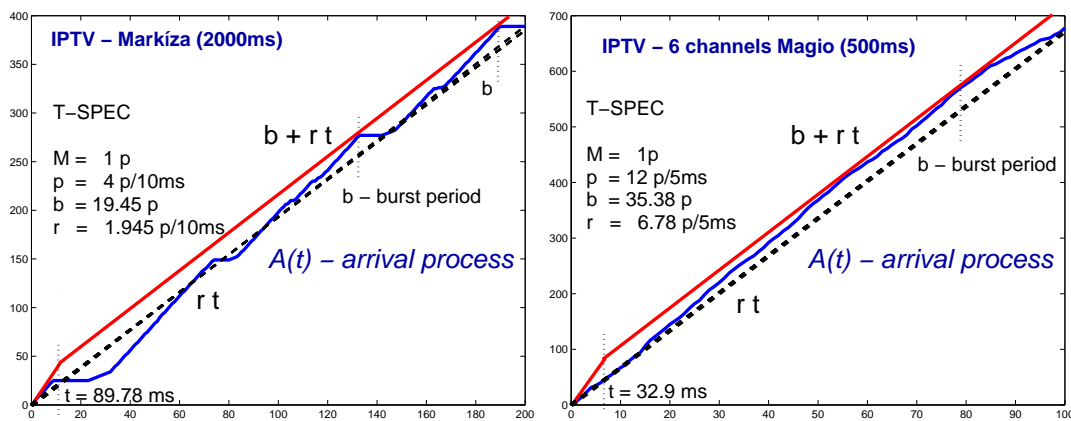
Vstupný tok je vlastne riadený dvoma Token Bucketmi, ktoré ohraňčujú tok krivkami $\gamma_{M,p}(t) = M + p \cdot t$ a $\gamma_{b,r}(t) = b + r \cdot t$. Výsledná krivka je subaditívna:

$$\alpha(t) = \min [M + p \cdot t, b + r \cdot t] \quad \text{resp.} \quad \alpha = \gamma_{M,p} \oplus \gamma_{b,r} \quad (\oplus - \text{minimum})$$



Obr. 1.13: T-špecifikácia pre služby IntServ, obmedzenie toku dvoma Token Bucketmi

Na obrázkoch je ukážka určenia T-špecifikácie pre dve krátke vzorky TV Markíza a 6 kanálov TV. Vďaka použitému kódovaniu v zdroji boli veľkosti paketov rovnaké, preto sme nastavili parameter $M = 1$.



Obr. 1.14: T-špecifikácia pre dva rôzne typy IPTV tokov

V rámci projektu [6] sme pracovali s rôznymi vzorkami záznamov IPTV zo zdroja ST-Magio o veľkosti 2 minúty. V prípade použitia celého záznamu napr. pre 6 kanálov sa zmenia len druhé dve hodnoty $b = 194p$, $r = 6p/5ms$. Priemerná intenzita r ostala v podstate zachovaná, zväčšil sa burst parameter b

oproti polsekundovému záznamu, počas 2 minútového záznamu sa pravdepodobne vyskytla v toku relatívne veľká dátová vlna.

V tejto kapitole sme sa zaoberali deterministickým popisom IP tokov. Najjednoduchší spôsob je použiť lineárne ohraňenie, ktorého parametre, stredná hodnota a burst perióda, vyplývajú buď z charakteru toku, alebo ho ovplyvňujú obmedzovače, ktoré riadia vstup toku do siete. Pre presnejší popis toku sme zaviedli subaditívne vstupné krivky, ktoré budú neskôr tvoriť základ pre použitie Network Kalkulu. Tento nástroj umožňuje z deterministického hľadiska analyzovať a dimenzovať rôzne situácie v sieti.

Vo všeobecnosti však použitie deterministických horných ohraňení môže byť veľmi štedré pri pridelení nutných prostriedkov na prevádzku. Napríklad ojedinelý výskyt burst periódy môže spôsobiť neadekvátne zvýšenie horného ohraňenia, pričom však potlačenie (zahodenie paketov) tejto periódy by možno spôsobilo len minimálne straty, ktoré by boli ďaleko pod hranicou požadovanej kvality obsluhy QoS.

Použitie deterministických ohraňení je jednoduché, efektívne a často plne postačujúce. V snahe však získať podrobnejší popis paketových tokov, napríklad kvôli presnejšiemu dimenzovaniu zariadení, siahneme po stochastických modeloch. Na ich určenie potrebujeme analýzu pravdepodobnostných charakteristík meraných tokov. Základné typy stochastických modelov uvedieme v nasledujúcej kapitole.

1.2 BERNOULLIHO MODEL IP TOKU

Najjednoduchší stochastický model paketového toku, resp. IP prevádzky, predstavuje Bernoulliho proces. Je definovaný dvoma vlastnosťami - výskyt udalostí (paketov, rámcov, buniek, ATM cells, bitov ...) v toku je navzájom nezávislý, a pravdepodobnosť výskytu udalosti v časom slotu sa v čase nemení, čiže je pre všetky sloty rovnaká. Predpokladáme, že v časovom slotu sa môže vyskytnúť iba jedna udalosť. Tento predpoklad je odôvodnený samotnou technológiou IP siete, v kde na najspodnejšej fyzickej vrstve sa nemôžu naraz vyskytnúť 2 bity, bez ohľadu na to, aký rýchly prenos je. Časový slot, v ktorom môže nastať výskyt maximálne jedného paketu (rámca, bunky), budeme nazývať *elementárny slot*.

Bernoulliho proces

Bernoulliho proces nazveme stochastický proces, v ktorom sa je výskyt paketov nezávislý s rovnakou pravdepodobnosťou p pre každý elementárny slot. Nech $q = 1 - p$. Prírastky toku $a(i)$ v elementárnom časovom slotu sú náhodné premenné s alternatívnym rozdelením:

$$\Pr(a(i) = k) = p^k(1 - p)^{1-k} = p^k q^{1-k}, \quad k = 0, 1$$

Náhodnú premennú $a(i)$ nazývame tiež Bernoulliho premenná resp. Bernoulliho indikátor (nadobúda hodnoty 0 a 1, čiže indikuje výskyt paketu v slotu), alebo hovoríme, že prírastky toku $a(i)$ majú alternatívne rozdelenie, $a(i) \sim \text{Alt}(p)$.

Počet udalostí (napr. paketov resp. bitov) v časovom slotе dĺžky n (jeden slot predstavuje súčet n elementárnych slotov) popíšeme pomocou náhodnej premennej $A(n)$. Z elementárnej pravdepodobnosti vieme, že jej rozdelenie je binomické:

Binomické rozdelenie

Pravdepodobnosť, že náhodná premenná $A(n)$, ktorá popisuje počet výskytu paketov n elementárnych slotov, nadobudne hodnotu k , sa riadi binomickým rozdelením:

$$A(n) \sim Bi(n, p); \quad \Pr(A(n) = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

PRÍKLAD 1.6 Nech nasledujúca postupnosť núl a jednotiek predstavuje výskyt bitov, resp. paketov v nejakom IP toku na sieti

100101001000011111001100000101111010001110101101110000010100111010

Z týchto dát vieme odhadnúť pravdepodobnosť výskytu paketu p v jednom časovom slotе. Počet všetkých vzoriek je 66 (čo je samozrejme málo, ale postačuje pre ukážku), počet výskytu paketu je 31, preto $p = 32/66 = 0.4849$. Ak by sme ten istý tok merali v dvoch časových slotoch, z predchádzajúceho merania dostaneme nasledujúci záznam:

1 1 1 0 1 0 1 2 2 0 2 0 0 1 1 2 1 1 0 2 1 1 2 1 2 0 0 1 1 0 2 1 1

Odhadneme rozdelenie pravdepodobnosti výskytu paketov v dvoch slotoch:

$$p_0 = 9/33 = 0.2727 \quad p_1 = 16/33 = 0.4849 \quad p_2 = 8/32 = 0.2424$$

Ak použijeme ako model pre meraný tok Bernoulliho proces s pravdepodobnosťou $p = 0.4849$ (a s predpokladom že výskyt paketov je nezávislý) dostávame binomické rozdelenie pravdepodobnosti:

$$p_0 = p^2 = 0.2351 \quad p_1 = 2pq = 0.4995 \quad p_2 = q^2 = 0.2653$$

Aj bez použitia štatistických metód (napr. testu dobrej zhody o rovnosti rozdelení) vidíme, že pravdepodobnosti odhadnuté z pozorovania a pravdepodobnosti získané z Bernoulliho procesu sa na seba "podobajú".

Ak by sme realizovali meranie tej istej realizácie toku v troch časových slotoch, záznam výskytu paketov by bol

1 2 1 0 2 3 1 1 0 2 3 1 1 2 2 2 2 0 1 1 3 1

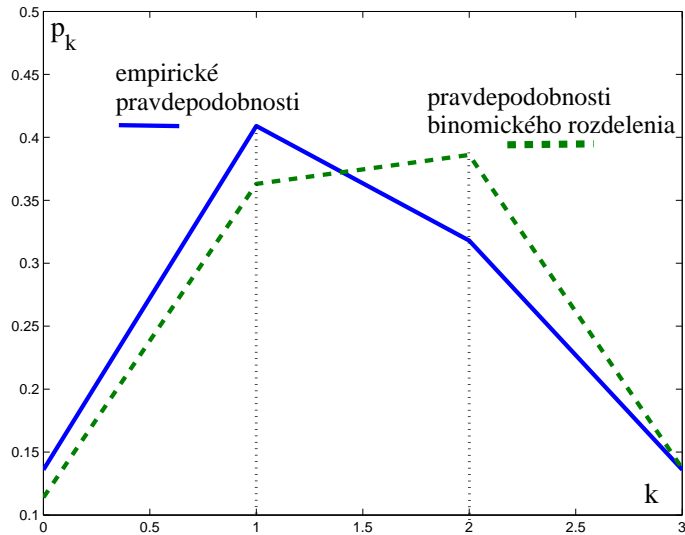
Odhadneme rozdelenie pravdepodobnosti výskytu paketov v troch slotoch:

$$p_0 = \frac{3}{22} = 0.136, \quad p_1 = \frac{9}{22} = 0.409, \quad p_2 = \frac{8}{22} = 0.318, \quad p_3 = \frac{3}{22} = 0.136$$

Použijeme Bernoulliho proces s $p = 0.4849$ ako model pre meraný tok.

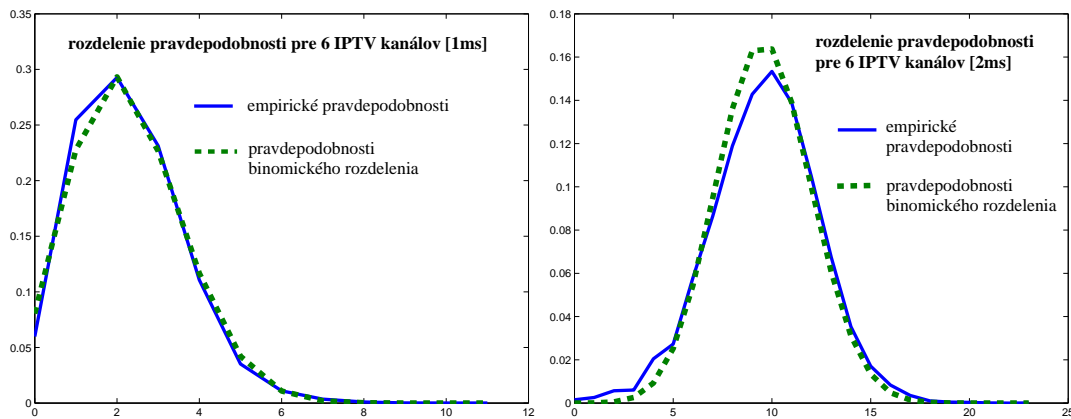
$$p_0 = p^3 = 0.114, \quad p_1 = 3p^2q = 0.363, \quad p_2 = 3pq^2 = 0.386, \quad p_3 = q^3 = 0.137$$

Aj pri malom počte dát (22) vidíme výraznú podobnosť medzi odhadnutými pravdepodobnosťami a hodnotami, ktoré sme získali z Bernoulliho procesu.



Obr. 1.15: Porovnanie odhadnutých a binomických pravdepodobností

Na nasledujúcich obrázkoch sa vrátíme k meraniam získaných zo zdroja ST-Magio, konkrétne k záznamu 6 kanálov IPTV kanálov a porovnáme tento záznam s Bernoulliho modelom IP toku.



Obr. 1.16: modelovanie toku 6 kanálov IPTV Bernoulliho procesom

Na prvom obrázku je znázornené rozdelenie pravdepodobností výskytu paketov, pričom časový slot je 1ms. Maximálny počet paketov v 1ms sa rovnal 11 paketov, a pravdepodobnosť výskytu 1 paketa je 0.2048. Namerané hodnoty sme aproximovali binomickým rozdelením $Bi(11, 0.2048)$. Na druhom obrázku sa časový slot rovnal 8ms a na aproximáciu sme použili rozdelenie $Bi(23, 0.4179)$. To znamená, že maximálne sa vyskytlo 23 paketov za 8ms, a pravdepodobnosť výskytu

paketa je 0.4179. Aj bez prevedenia štatistických testov môžeme aproximáciu v obidvoch prípadoch považovať za dostatočnú.

V predchádzajúcej kapitole sme sa zaoberali IPTV tokom Markíza, teda iba jedným kanálom. Tento tok vykazoval nezameniteľné periodické vlastnosti, aj keď dĺžka periódy a intenzita toku v periódach boli náhodné. Periodicita je typická pre IPTV toky. Na predchádzajúcom obrázku sa zaoberáme 6 kanálmi Mágia, pričom pri ujme na všeobecnosti môžeme predpokladať, že každý z kanálov môže v istej miere vykazovať periodické vlastnosti. Avšak pri súčte šiestich kanálov periodičnosť jednotlivých kanálov zaniká, a prejavuje sa Bernoulliho vlastnosť, výskyt jednotlivých paketov je navzájom nezávislý (jednotlivé pakety patria k rôznym kanálom).

1.2.1 Pravdepodobnostné charakteristiky

Základné číselné charakteristiky náhodnej premennej X sú **stredná hodnota** EX a **disperzia** DX . Pre náhodnú premennú $a(i)$, ktorá popisuje prírastky v elementárnom slotе, ich určíme veľmi jednoducho:

$$Ea(i) = 0q + 1p = p, \quad Da(i) = (0 - p)^2q + (1 - p)^2p = pq(q + p) = pq$$

Momenty prírastkov $a(i)$

Nech náhodná premenná $a(i)$ má alternatívne rozdelenie pravdepodobnosti $\Pr(a(i) = k) = p^k(1 - p)^{1-k}$, $k = 0, 1$. Potom stredná hodnota a disperzia premennej $a(i)$ sú: $Ea(i) = p$, $Da(i) = p(1 - p) = pq$

Strednú hodnotu náhodnej premennej $A(n)$, ktorá popisuje výskyt paketov v slotе o dĺžke n , môžeme získať rôznymi spôsobmi. Najbežnejší spôsob je priamo použitie definície strednej hodnoty pre diskretnú náhodnú premennú:

$$\begin{aligned} EA(n) &= \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^n k \frac{n!}{(n-k)!k!} p^k q^{n-k} = \\ &= npq \sum_{k=1}^n \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} q^{n-k+1} = npq \sum_{k=0}^{n-1} \frac{(n-1)!}{(n-1-k)!k!} p^k q^{n-k-1} = \\ &= \sum_{k=0}^n k \binom{n-1}{k} p^k q^{n-1-k} = npq(p + q)^{n-1} = npq \end{aligned}$$

Podobne, s použitím rozkladu $k^2 = k(k-1) + k$ by sme z rozdelenia vypočítali druhý moment $EA(n)^2$ a z neho disperziu $DA(n) = EA(n)^2 - (EA(n))^2$. Postup je zdĺhavý, a čitateľ ho môže nájsť v ľubovoľnej učebnici o elementárnej pravdepodobnosti

Pri znalosti elementárnej pravdepodobnosti môžeme k získanému výsledku dospieť jednoduchou elementárnou úvahou, ktorá vychádza priamo z definície Bernoulliho procesu. Pravdepodobnosť výskytu paketu v jednom slotе je p . Túto pravdepodobnosť môžeme tiež pokladať za strednú hodnotu výskytu paketu v slotе. Preto stredná hodnota výskytu paketov v n -slotoch je np .

Pre určenie disperzie $DA(n)$ použijeme tvrdenie: *disperzia súčtu nezávislých náhodných premenných je súčet disperzií jednotlivých premenných*. Keďže výskyt paketov v Bernoulliho procese je nezávislý, disperzia $A(n)$ je súčtom disperzií výskytu paketov v jednotlivých slotoch:

**Momenty
náhodnej
premennej
 $A(n)$**

Nech náhodná premenná $A(n)$ má binomické rozdelenie pravdepodobnosti $\Pr(A(n) = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \dots, n$. Potom stredná hodnota a disperzia premennej $A(n)$ sú: $EA(n) = np$, $DA(n) = npq$

Náhodnú premennú jednoznačne popisuje jej rozdelenie pravdepodobnosti, resp. distribučná funkcia. Ďalšími spôsobmi, ako popísať náhodnú premennú je použitie rôznych vytvárajúcich funkcií. Najčastejšie sa uvádza Momentová vytvárajúca funkcia $\varphi(\theta)$ (*MGF - moment generation function*), ktorá je strednou hodnotou parametrickej exponenciálnej transformácie náhodnej premennej. My sa budeme zaoberať kumulatívnu vytvárajúcou funkciou, ktorá je prirodzeným logaritmom MGF, pretože priamo súvisí s dimenzovaním kapacity linky pomocou efektívnej šírky pásma, s ktorým sa stretneme v neskorších kapitolách.

**Kumulatívna
vytvárajúca
funkcia CGF**

Kumulatívna vytvárajúca funkcia (*CGF - cumulant generation function*), pre nezápornú náhodnú premennú X , má tvar:

$$\lambda(\theta) = \ln \varphi(\theta) = \ln E[e^{\theta X}] = \ln E \left[\sum_{\forall k} e^{\theta k} p_k \right]$$

Kumulatívna vytvárajúca funkcia pomocou n -tých derivácií kumuluje, resp. vytvára lineárne kombinácie n -tých počiatočných momentov premennej X . Nám bude stačiť informácia, že prvá derivácia v nule je stredná hodnota EX a druhá derivácia v nule je disperzia $DX = EX^2 - (EX)^2$.

**Výpočet
momentov
pomocou
CGF**

$$\lambda'(0) = \ln E[e^{\theta X}]|_{\theta=0} = \frac{E[Xe^{\theta X}]}{E[e^{\theta X}]} \Big|_{\theta=0} = EX$$

$$\lambda''(0) = \frac{E[X^2 e^{\theta X}] E[e^{\theta X}] - E[Xe^{\theta X}]^2}{E[e^{\theta X}]^2} \Big|_{\theta=0} = EX^2 - (EX)^2 = DX$$

PRÍKLAD 1.7

Určíme CGF $\lambda(\theta)$ pre náhodnú premennú $a(i)$:

$$\lambda(\theta) = \ln \left[e^{\theta \cdot 0} P(a(i) = 0) + e^{\theta \cdot 1} P(a(i) = 1) \right] = \ln [q + pe^\theta] \quad (1.1)$$

Vypočítame prvú a druhú deriváciu v bode $\theta = 0$:

$$\lambda'(0) = \frac{pe^\theta}{q + pe^\theta} \Big|_{\theta=0} = p, \quad \lambda''(0) = \frac{pe^\theta(q + pe^\theta) - (pe^\theta)^2}{(q + pe^\theta)^2} \Big|_{\theta=0} = pq$$

Dostali sme samozrejme strednú hodnotu $Ea(i)$ a disperziu $Da(i)$.

PRÍKLAD 1.8

Určíme CGF $\lambda(\theta)$ pre náhodnú premennú $A(n)$:

$$\begin{aligned} \lambda(\theta) &= \ln \left[\sum_{k=0}^n e^{\theta k} \binom{n}{k} p^k q^{n-k} \right] = \ln \left[\sum_{k=0}^n \binom{n}{k} (pe^\theta)^k q^{n-k} \right] = \ln [q + pe^\theta]^n \\ &\Rightarrow \lambda(\theta) = n \ln [q + pe^\theta] \end{aligned} \quad (1.2)$$

Vypočítame prvú a druhú deriváciu v bode $\theta = 0$:

$$\lambda'(0) = n \frac{pe^\theta}{q + pe^\theta} \Big|_{\theta=0} = np, \quad \lambda''(0) = n \frac{pe^\theta(q + pe^\theta) - (pe^\theta)^2}{(q + pe^\theta)^2} \Big|_{\theta=0} = npq$$

Dostali sme samozrejme strednú hodnotu $EA(n)$ a disperziu $DA(n)$.

Výsledok potvrdil vetu: *CGF funkcia súčtu nezávislých náhodných premenných je súčtom CGF funkcií jednotlivých premenných.* Nech $X = \sum X_k$,

$$\lambda_X(\theta) = \ln E \left[e^{\theta \sum_k X_k} \right] = \ln E \left[\prod_k e^{\theta X_k} \right] = \ln \prod_k E \left[e^{\theta X_k} \right] = \sum_k \ln E \left[e^{\theta X_k} \right]$$

**CGF pre
Binomické
rozdelenie**

Kumulatívna vytvárajúca funkcia náhodnej premennej, ktorá má Binomické rozdelenie $A(n) \sim Bi(n, p)$ má tvar

$$\lambda(\theta) = n \ln [q + pe^\theta]$$

Ďalšia funkcionálna charakteristika náhodnej premennej, ktorú budeme používať, je **distribučná funkcia** $F(x)$:

$$F(x) = P(X \leq x) = \sum_{k; k \leq x} p_k$$

Pre náhodnú premennú K^n má tvar:

$$F(x) = P(A(n) \leq x) = \sum_{k=0}^x \binom{n}{k} p^{n-k} q^k$$

Pri úlohách o dimenzovaní má význam tzv. doplnková distribučná funkcia

$$\bar{F}(x) = 1 - F(x) = P(X > x)$$

ktorá vyjadruje pravdepodobnosť prekročenie istej "kritickej" hodnoty x pre daný proces, ktorý je popísaný náhodnou premennou X .

PRÍKLAD 1.9 Cez informačný kanál prenášame 8 bitové slová. Pravdepodobnosť chybného prenosu je 0.1, pričom výskyt chyby v jednotlivých bitoch sa navzájom neovplyvňuje. Kodek na prímači dokáže opraviť 2 chybné bity. Aká je pravdepodobnosť úspešného prenosu?

V tomto prípade, vďaka nezávislosti, Bernoulliho proces tvorí výskyt jednotlivých chýb v bitoch. Ich počet budeme modelovať náhodnou premennou E , ktorá sa riadi binomickým rozdelením $Bi(8, 0.1)$:

$$p_k = \Pr(E = k) = \binom{8}{k} 0.1^k 0.9^{8-k}$$

Nech udalosť A predstavuje nastatie úspešného prenosu, to znamená, že všetky bity sú bezchybné. Kodek dokáže opraviť maximálne 2 chybné bity

$$\Pr(A) = \Pr(E \leq 2) = p_0 + p_1 + p_2 = 0.9^8 + 8 \cdot 0.9^7 \cdot 0.1 + 28 \cdot 0.9^6 \cdot 0.1^2 = 0.9619$$

Pravdepodobnosť úspešného prenosu, resp. výskytu maximálne dvoch chybných bitov v 8-bitovom slove, je 0.9619.

PRÍKLAD 1.10 Majme nameraný nejaký reálny IP tok, ktorý má formu záznamu počtu paketov, resp. prírastkov toku $a(i) = 0, 1, \dots, 4$, v časovom slotu o dĺžke 20 ms:

4100200000 0012444444 0142300104 0332110002 4103401124 4330034400

Jednoduchým spôsobom získame dva základné parametre na popis toku:

$$\lambda_{peak} = 4 \text{ [paket]}, \quad \lambda_{avg} = \frac{102}{60} = 1.7000 \text{ [paket/20ms]}$$

Chceme overiť, či Bernoulliho model je vhodný na popis nameraného toku. Bernoulliho model závisí od dvoch parametrov, n a p . Vypočítame ich z parametrov nameraného toku:

$$\lambda_{peak} = n = 4 \text{ [paket]}, \quad \lambda_{avg} = np \Rightarrow p = \frac{\lambda_{avg}}{n} = 0.4250$$

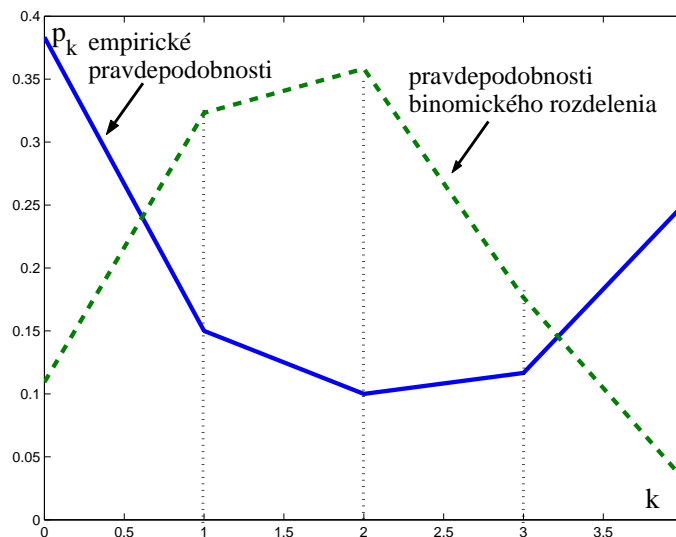
Hodnota 0.4250 je pravdepodobnosť výskytu paketu elementárnom časovom slotu, ktorý predstavuje jednu štvrtinu časového slotu meraného toku, našom

príklade teda 5 ms. Maximálny počet paketov je 4 za 20 ms, teda paket sa vyskytol v každom elementárnom 5 ms slot. Tak to predpokladá model, v realite nemusí byť 20 ms rovnomerne rozdelený na elementárne sloty. Meranie je ale obmedzené na 20 ms, neumožňuje analyzovať proces v rámci 20 ms, preto predpoklady modelu o rovnosti elementárnych tokov môžeme akceptovať.

V tabuľke uvedieme empirické pravdepodobnosti \bar{p}_k vypočítané z dát a pravdepodobnosti binomického rozdelenia p_k s odhadnutými hodnotami n a p :

k	0	1	2	3	4
\bar{p}_k	0.3833	0.1500	0.1000	0.1167	0.2500
p_k	0.1093	0.3232	0.3583	0.1766	0.0326

Medzi pravdepodobnosťami panuje minimálna zhoda, preto s istotou môžeme tvrdiť, že nameraný tok neodpovedá Bernoulliho modelu.



Obr. 1.17: Porovnanie odhadnutých a binomických pravdepodobností

V praxi, napríklad pri použití programu Wireshark, môžeme dostať záznam reálneho toku vo forme hodnôt, ktoré predstavujú doby výskytov paketov. Pri vhodnej voľbe časového slotu môžeme dostať záznam prírastkov toku s vopred zvoleným píkom. Ak by sme za veľkosť časového slotu zobrali veľkosť minimálnej medzery, dostali záznam iba 0 a 1. V takomto prípade môžeme analyzovať medzery medzi paketmi v meranom toku.

1.2.2 Rozdelenie pravdepodobnosti medzier

Medzeru v Bernoulliho procese budeme definovať ako postupnosť núl (v elementárnom slotu sa nevyskytol paket), ktorú ukončí výskyt jednotky (paketu). Napríklad v postupnosti 110101000111 sa vyskytujú dve medzery o dĺžke 1, a jedna medzera o dĺžke 3. Pravdepodobnosť výskytu medzery o dĺžke 3 je vďaka nezávislosti rovná $\Pr(T = 3) = q^3p$. Pre úplnosť pravdepodobnostného modelu musíme uvažovať aj o medzere dĺžky nula, ktorej pravdepodobnosť je $\Pr(T = 0) = p$.

**Rozdelenie
pravdepodob-
nosti
medzier**

Nech náhodná premenná T popisuje veľkosť medzery v Bernoulliho toku. Rozdelenie pravdepodobnosti n.pr. T je geometrické:

$$T \sim Geo(q); \quad P(T = t) = (1 - p)^t p = q^t p, \quad t = 0, 1, 2, \dots$$

Vypočítame strednú hodnotu dĺžky medzier:

$$ET = \sum_{t=0}^{\infty} t q^t p = p q \sum_{t=1}^{\infty} t q^{t-1} = p q \sum_{t=1}^{\infty} \frac{d}{dq} q^t = p q \frac{d}{dq} \frac{q}{1 - q} = \frac{p q}{(1 - q)^2} = \frac{q}{p}$$

Pri výpočte sme zamenili poradie sumácie a derivovania ($q < 1$) a použili sme vzťah $1 - q = p$. Pri výpočte disperzie by sme postupovali podobným spôsobom, ale technicky náročnejším, v priebehu sumácie je potrebné urobiť deriváciu druhého rádu. My si ukážeme podstatne jednoduchší spôsob výpočtu použitím kumulatívnej vytvárajúcej funkcie.

PRÍKLAD 1.11

Pomocou kumulatívnej vytvárajúcej funkcie vypočítajte strednú hodnotu a disperziu n.pr. T .

Najprv určíme CGF funkciu pre geometrické rozdelenie:

$$\begin{aligned} \lambda(\theta) &= \ln \left[\sum_{t=0}^{\infty} e^{\theta t} q^t p \right] = \ln \left[\sum_{t=0}^{\infty} (e^{\theta} q)^t p \right] = \ln \left[\frac{p}{1 - e^{\theta} q} \right] = \\ &= \ln p - \ln [1 - e^{\theta} q], \quad \theta < -\ln q \end{aligned}$$

Vypočítame prvú a druhú deriváciu v bode $\theta = 0$:

$$\lambda'(0) = \frac{e^{\theta} q}{1 - e^{\theta} q} \Big|_{\theta=0} = \frac{q}{p}, \quad \lambda''(0) = \frac{e^{\theta} q (1 - e^{\theta} q) + (e^{\theta} q)^2}{(1 - e^{\theta} q)^2} \Big|_{\theta=0} = \frac{q}{p^2}$$

Dostali sme $ET = q/p$ a $DT = q/p^2$.

**Momenty
náhodnej
premennej T**

Nech náhodná premenná T má geometrické rozdelenie pravdepodobnosti $\Pr(T = t) = q^t p$, $t = 0, 1, 2, \dots$. Stredná hodnota a disperzia sú:

$$ET = \frac{q}{p}, \quad DT = \frac{q}{p^2}$$

Určíme distribučnú funkciu geometrického rozdelenia:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - \sum_{k=t+1}^{\infty} q^k p = 1 - q^{t+1}$$

Doplnková distribučná funkcia má tvar:

$$\bar{F}(t) = \Pr(T > t) = 1 - F(t) = q^{t+1}$$

Geometrické rozdelenie má vlastnosť "zabúdanie na minulosť". Pravdepodobnosť, že medzera medzi paketmi bude ešte trvať viac než čas t , ak vieme že už trvala τ , nezáleží od τ :

$$P(T > t + \tau | T > \tau) = \frac{P(T > t + \tau)}{P(T > \tau)} = \frac{q^{t+\tau+1}}{q^{\tau+1}} = q^t = P(T \geq t)$$

Výsledná pravdepodobnosť je závislá iba od času t , rozdelenie "zabudlo", že medzera už nejaký čas trvala. Táto vlastnosť sa používa pri tvorbe Markovových reťazcov s diskretným časom.

Koeficient variabilnosti

Koeficientom variabilnosti náhodnej premennej T označíme $\nu_T = \frac{\sigma_T}{ET}$

Budeme predpokladať, že náhodná premenná T je nezáporná. Koeficient ν_T vyjadruje mieru variabilnosti procesu popísaného premennou T . Ak $\nu_T \rightarrow 0$, proces je "skoro" deterministický, rozptyl σ_T sa blíži k nule a medzery môžeme považovať za rovnako veľké. Ak $\nu_T = 1$, je splnená nutná podmienka pre Poissonov proces, ak $\nu_T \rightarrow \infty$, dostávame sa do oblasti ťažkých chvostov (Heavy-tailed distribution) a samopodobnosti.

Pre Bernoulliho proces platí

$$\nu_T = \frac{\sigma_T}{ET} = \frac{\sqrt{q/p^2}}{q/p} = \frac{1}{\sqrt{1-p}}$$

Keďže parameter p je pravdepodobnosť, $p \in (0, 1)$, koeficient variabilnosti pre geometrické rozdelenie je $\nu_T \in (1, \infty)$.

1.2.3 Identifikácia Bernoulliho procesu

Predchádzajúca kapitola poskytuje návod, ako určiť parameter p z 0/1 záznamu:

$$\frac{1}{\sqrt{1-p}} = \frac{\sigma_T}{ET} \Rightarrow p = 1 - \frac{(ET)^2}{DT}$$

Nutnou podmienkou pre modelovanie pomocou Bernoulliho procesu je, aby pre odhadnutú hodnotu p platilo $p \in (0, 1)$. Parameter p môžeme vypočítať aj ako strednú hodnotu prírastkov $a(i)$, $p = Ea(i)$. Ak odhadnuté hodnoty parametra podstatne odlišné, skúmaný tok netvorí Bernoulliho proces.

1.2.4 Rozdelenie pravdepodobnosti paketových zhlukov

Paketový zhluk v Bernoulliho procese budeme definovať ako postupnosť jednotiek (v elementárnom slotе sa vyskytol paket), ktorú ukončí výskyt nuly. Veľkosť zhluku budeme popisovať náhodnou premennou Z . Ide vlastne o rovnakú koncepciu ako pri úvahách od medzerách v predchádzajúcej kapitole, s tým, že nahradíme 0 za 1 a naopak. Distribúcia dĺžky paketových zhlukov má preto geometrické rozdelenie:

$$Z \sim Geo(p); \quad P(Z = z) = p^z q, \quad z = 0, 1, 2, \dots,$$

Geometrickým rozdelením sme sa zaoberali v predchádzajúcej kapitole. Popis paketových zhlukov sa od popisu medzier medzi paketmi líši iba parametrom rozdelenia $p = 1 - q$. Momenty a funkcionálne charakteristiky sme vypočítali v predchádzajúcej kapitole.

Charakteristiky náhodnej premennej Z

Nech náhodná premenná Z popisuje veľkosť paketových zhlukov v Bernoulliho toku. Náhodná premenná má potom geometrické rozdelenie pravdepodobnosti $\Pr(Z = z) = p^z q$, $z = 0, 1, 2, \dots$. Stredná hodnota, disperzia sú, kumulatívna vytvárajúca funkcia a distribučná funkcia majú tvar:

$$EZ = \frac{p}{q}, \quad DZ = \frac{p}{q^2}, \quad \lambda(\theta) = \ln q - \ln [1 - e^\theta p], \quad F(z) = 1 - p^{z+1}$$

Poznámka: V Bernoulliho procese distribúcia počtu paketov v časovom slotе o veľkosti n má rozdelenie binomické, medzery medzi paketmi a veľkosti paketových zhlukov sa riadia podľa geometrického rozdelenia, pričom medzi strednou hodnotou medzier a strednou hodnotou veľkosti paketových zhlukov je vzájomná väzba (viď nižšie). Tento fakt sa môže použiť ako nutná podmienka pri testovaní, či sa reálna prevádzka môže modelovať Bernoulliho procesom.

$$A(n) \sim Bi(n, p), \quad T \sim Geo(q), \quad Z \sim Geo(p), \quad EZ = \frac{1}{ET}$$

PRÍKLAD 1.12 V Bernoulliho toku je pravdepodobnosť výskytu paketu v 1 ms je rovná 0.7. Aká je pravdepodobnosť, že veľkosť paketového zhluku nie je väčšia ako 2?

$$\Pr(Z \leq 2) = 1 - \Pr(Z > 2) = 1 - 0.7^3 = 0.657$$

Aká je pravdepodobnosť, že nasledujú za sebou práve 3 pakety?

$$\Pr(Z = 3) = 0.7^3 \cdot 0.3 = 0.1029.$$

Aká je pravdepodobnosť, že medzera medzi paketmi je minimálne 4 ms?

$$\Pr(T \geq 4) = 0.3^4 = 0.0081$$

Aká je pravdepodobnosť, že sa v priebehu 3 ms vyskytne aspoň 1 paket?

$$\Pr(A(3) \geq 1) = 1 - \Pr(A(3) = 0) = 1 - 0.3^3 = 0.973$$

Aká je pravdepodobnosť, že sa v priebehu 6 ms vyskytnú práve 3 pakety?

$$\Pr(A(6) = 3) = \binom{6}{3} 0.7^3 \cdot 0.3^3 = 0.18522$$

Aký je stredný počet paketov v 100 ms?

Výsledok:

$$EA(100) = 100 \cdot 0.7 = 70p/100ms$$

Aká je pravdepodobnosť, že veľkosť paketového zhluku je väčšia než 3?

$$\Pr(Z > 3) = 0.7^4 = 0.2401$$

Aká je pravdepodobnosť, že veľkosť paketového zhluku je maximálne 5?

$$\Pr(Z \leq 5) = 1 - \Pr(Z > 5) = 1 - 0.7^6 = 0.88235.$$

Aká je pravdepodobnosť, že v 3 ms budú maximálne 2 pakety?

$$\Pr(A(3) \leq 2) = p_0 + p_1 + p_2 = 0.7^3 + 3 \cdot 0.7 \cdot 0.3^2 + 3 \cdot 0.7^2 \cdot 0.3 = 0.973.$$

Aká je pravdepodobnosť, že 5 ms slot nie je prázdny?

$$\Pr(A(5) \geq 0) = 1 - p_0 = 1 - 0.3^5 = 0.83193.$$

Aká je pravdepodobnosť, že medzera medzi paketmi je minimálne 6 ms a maximálne 8 ms?

$$\Pr(6 \leq T \leq 8) = 0.3^6 \cdot 0.7 + 0.3^7 \cdot 0.7 + 0.3^8 \cdot 0.7 = 0.00071.$$

Bernoulliho proces predstavuje základný model pre popis IP tokov s diskretným časom. Má dve základné vlastnosti, pravdepodobnosť výskytu paketu v časovom slotu sa v čase nemení, a výskyt paketov je navzájom nezávislý. Preto je proces stacionárny, jeho pravdepodobnostné charakteristiky nezávisia na čase, a má nezávislé prírastky. Rozdelenie pravdepodobnosti počtu paketov v n slotoch sa riadi Binomickým rozdelením, ktoré je dvojparametrické, preto pri analýze nameraného toku môžeme použiť dva základné parametre, strednú intenzitu a peak. V prípade, že máme k dispozícii podrobný záznam meraného toku so znalosťou o výskyte paketov v čase, môžeme pri analyzovaní použiť fakt, že veľkosť paketových zhlukov a medzery medzi paketami sa riadia Geometrickými rozdeleniami s navzájom doplnkovými parametrami p a q .

Opodstatnenosť použitia Bernoulliho toku je hlavne na modelovanie súčtu viacerých IP tokov. Toto tvrdenie sme demonštrovali na ukážke aproximácie rozde-

lenia pravdepodobnosti výskytu paketov v IPTV toku, ktorý predstavoval súčet 6 kanálov zo zdroja Mágio.

V súčasnosti je však Bernoulliho model často nepostačujúci, napríklad keď IP tok vykazuje istú zhlukovitosť výskytu paketov. Vhodným modelom pre takúto situáciu sa budeme zaoberať v nasledujúcej kapitole.

1.3 ZHLUKOVÉ ON/OFF MODELÝ

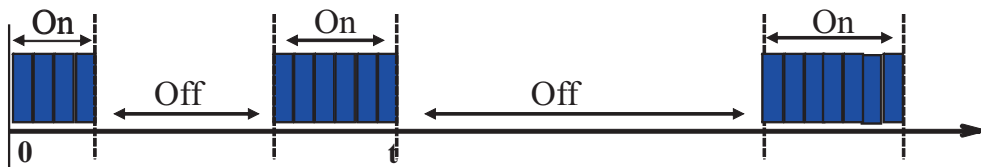
V tejto kapitole sa budeme zaoberať modelmi, ktorých intenzita toku paketov sa v mení v čase. Model predstavuje zdroj paketov, ktorý sa "prepína" medzi koneným počtom stavov, a v každom zo stavov môže "zapnúť" nejaký tok paketov s odlišnou intenzitou. Takéto modely budeme označovať ako modulované procesy (MP).

Bernoulliho model v predchádzajúcej kapitole mal intenzitu toku nemennú, $\lambda_{avg} = np$. Avšak aj tento model môžeme považovať za modulovaný proces s jedným stavom, v ktorom sa "zapne" Bernoulliho tok paketov.

V modulovanom procese pripúšťame zmenu intenzity toku a vo všeobecnosti aj zmenu pravdepodobnostného charakteru toku, preto je proces vo všeobecnosti nestacionárny. Aby sme dosiahli stacionárnosť, musíme proces modulovať správnym spôsobom, prepínanie medzi stavmi preto riadi Markovov reťazec, ktorým sa budeme zaoberať v nasledujúcej kapitole. Takéto modely sa nazývajú Markovove modulované procesy MMP.

V praxi sa najčastejšie používajú 2-stavové modulované procesy. Zdroj paketov sa prepína len medzi dvoma odlišnými tokmi. Dôvod použitia 2-stavových modelov je ich jednoduchá analýza a možnosť získania explicitných výsledkov pri dimenzovaní daných zariadení.

Populárne modely pre tzv. "bursty" zdroje je sú On/Off modely, v ktorých sa zdroj prepína medzi aktívnym stavom produkujúcim dáta (perióda On), a neaktívnym stavom (perióda Off). Najjednoduchším modelom zdroj, ktorý počas periódy On generuje deterministický tok paketov s danou konštantnou intenzitou. Model sa označuje ako Markovov modulovaný regulárny proces (MMRP). Tento model sa zvykne používať na popis prevádzky s konštantnou intenzitou toku dát (CBR - *constant bit rate*), ktorá sa náhodne vypína:



Obr. 1.18: MMRP - Markovov modulovaný regulárny proces

1.3.1 2-stavový Markovov reťazec

Náhodný proces s diskretným časom $X(t)$ je postupnosť náhodných premenných $\{X_i\}_i$, ktoré v čase $i = t_i$ nadobúdajú hodnotu s_i . Hovoríme tiež, že reťazec sa v čase t_i nachádza v stave s_i , $X(t_i) = X_i = s_i$. Náhodný proces, ktorý nadobúda hodnoty v diskretných časoch, nazveme náhodný reťazec. Množinu $S = \{s_0, s_1, s_2, \dots\}$ nazveme množinou stavov náhodného reťazca $X(t)$. Pravdepodobnosť $p_k(t)$ bude označovať pravdepodobnosť, že reťazec sa v čase t nachádza v stave s_k :

$$p_k(t) = \Pr(X(t) = s_k)$$

Vektor $\mathbf{p}(t) = (p_0(t), p_1(t), p_2(t), \dots)$ nazveme rozdelenie pravdepodobnosti reťazca v čase t a vektor $\mathbf{p}(0) = (p_0(0), p_1(0), p_2(0), \dots)$ počiatočným rozdelením v čase $t = 0$.

Markovov reťazec

Náhodný proces $X(t)$ nazveme Markovov reťazec, ak má spočítateľnú množinu stavov $S = \{s_0, s_1, \dots, s_n\}$ a pre podmienené pravdepodobnosti platí Markovova vlastnosť: $\forall m_n, m_{n-1}, \dots, m_0$;

$$\begin{aligned} \Pr(X(t_n) = s_{m_n} / X(t_{n-1}) = s_{m_{n-1}}, \cap X(t_{n-2}) = s_{m_{n-2}} \cap \dots \cap X(t_0) = s_{m_0}) = \\ = \Pr(X(t_n) = s_{m_n} / X(t_{n-1}) = s_{m_{n-1}}) \end{aligned}$$

Markovova vlastnosť znamená, že podmienená pravdepodobnosť javu, že reťazec sa bude v čase t_n nachádzať v stave s_n , ak vieme, že v predchádzajúcom okamihu sa nachádzal v stave s_{n-1} nezávisí od toho, ako sa do stavu s_{n-1} dostal. Markovova vlastnosť iným spôsobom vyjadruje tzv. zabúdanie na minulosť, s ktorým sme sa zaoberali v kapitole o Bernoulliho procese.

Homogénny Markovov reťazec

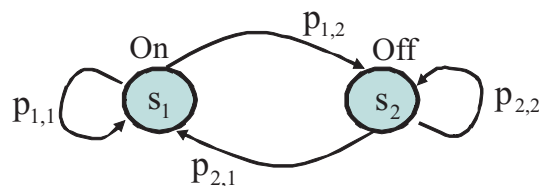
Markovov reťazec nazveme homogénny, ak pravdepodobnosť prechodu medzi stavmi nezáleží od času h :

$$\Pr(X(h_1) = s_j / X(h_1 - 1) = s_i) = \Pr(X(h_2) = s_j / X(h_2 - 1) = s_i) = p_{i,j}$$

V homogénnom Markovovom reťazci pravdepodobnosť $p_{i,j}$ bude označovať pravdepodobnosť prechodu zo stavu s_i do stavu s_j za jeden časový krok bez ohľadu na to, kde v čase sa reťazec nachádza:

$$p_{i,j} = \Pr(X(h) = s_j / X(h - 1) = s_i)$$

Markovov reťazec je jednoznačne určený množinou svojich stavov a pravdepodobnosťami prechodov medzi stavmi. V našom prípade je množina stavov dvojprvková, $S = \{s_1, s_2\} = \{\text{On}, \text{Off}\}$. Pravdepodobnosti prechodov zapíšeme do dvojrozmernej matice $\mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}$. Markovov reťazec môžeme úplne



Obr. 1.19: Prechodový graf 2-stavového Markovovho reťazca

popísať tzv. prechodovým grafom, kde vrcholy grafu predstavujú stavy reťazca a orientované hrany pravdepodobnosti prechodov medzi stavmi:

PRÍKLAD 1.13

Budeme sa zaoberať On/Off zdrojom prevádzky, ktorý je modulovaný pomocou 2-stavového Markovovho reťazca s diskrétnym časom. Nech pravdepodobnosť, že sa zdroj za časový krok vypne je 0.1, a pravdepodobnosť že sa zapne je 0.7. Nech časový krok je 1ms. Našou úlohou bude vypočítať koľko percent z celkovej prevádzky zdroj vysiela.

Máme dané pravdepodobnosti prechodov vypnutia a : $p_{1,2} = 0.1$, $p_{2,1} = 0.7$. Matica pravdepodobností prechodov $\mathbf{P} = \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix}$. Predpokladáme, že na začiatku zdroj vysiela, preto počiatočné rozdelenie reťazca je $\mathbf{p}(0) = (1, 0)$.

Vypočítame pravdepodobnosť, že zdroj po 1ms stále vysiela, pritom musíme brať do úvahy všetky možnosti:

$$p_1(1) = p_1(0)p_{1,1} + p_2(0)p_{2,1} = 1 \cdot 0.9 + 0 \cdot 0.7 = 0.9$$

Podobne vypočítame pravdepodobnosť, že zdroj po 1ms prestal vysielať:

$$p_2(1) = p_1(0)p_{1,2} + p_2(0)p_{2,2} = 1 \cdot 0.1 + 0 \cdot 0.3 = 0.1$$

Obidva predchádzajúce výpočty môžeme vyjadriť maticovo:

$$\mathbf{p}(1) = \mathbf{p}(0) \cdot \mathbf{P} = (1, 0) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.9, 0.1)$$

Rovnako vypočítame v akom stave sa bude zdroj nachádzať po 2 ms, resp. rozdelenie reťazca v čase $t = 2$:

$$\mathbf{p}(2) = \mathbf{p}(1) \cdot \mathbf{P} = (0.9, 0.1) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.88, 0.12)$$

Rozdelenie reťazca v čase n získame rekurentne:

$$\mathbf{p}(n) = \mathbf{p}(n-1) \cdot \mathbf{P} = \mathbf{p}(n-2) \cdot \mathbf{P}^2 = \mathbf{p}(n-3) \cdot \mathbf{P}^3 = \dots = \mathbf{p}(0) \cdot \mathbf{P}^n$$

Rozdelenie reťazca v čase n môžeme získať dvoma spôsobmi, rekurentne pomocou rozdelenia v predchádzajúcom čase, alebo pomocou n -tej mocniny matice.

Vždy musíme poznať počiatočné rozdelenie reťazca v čase $t = 0$.

Budeme ďalej pokračovať vo výpočtoch rozdelenia pravdepodobnosti reťazca čase, pričom použijeme rekurentné vzťahy $\mathbf{p}(n) = \mathbf{p}(n-1) \cdot \mathbf{P}$, ktoré môžeme interpretovať, tak, že na to, aby sme zistili s akou pravdepodobnosťou sa reťazec bude nachádzať v danom stave v čase n , musíme vedieť pravdepodobnosti výskytu v stavoch v čase $n-1$ a tie vynásobiť pravdepodobnosťami prechodov zo stavov do daného stavu. Pre $t = 3$ dostávame:

$$\mathbf{p}(3) = \mathbf{p}(2) \cdot \mathbf{P} = (0.88, 0.12) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.876, 0.124)$$

Rozdelenie reťazca v čase $t = 4$:

$$\mathbf{p}(4) = \mathbf{p}(3) \cdot \mathbf{P} = (0.876, 0.124) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.8752, 0.1248)$$

Rozdelenie reťazca v čase 5:

$$\mathbf{p}(5) = \mathbf{p}(4) \cdot \mathbf{P} = (0.8752, 0.1248) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.87504, 0.12496)$$

Rozdelenie reťazca v čase 6:

$$\mathbf{p}(6) = \mathbf{p}(5) \cdot \mathbf{P} = (0.87504, 0.12496) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.875008, 0.124992)$$

Vidíme, že postupne sa začali pravdepodobnosti rozdelenia reťazca stabilizovať v čase, t.j. medzi $\mathbf{p}(5)$ a $\mathbf{p}(6)$ je minimálny rozdiel. K stabilizácii na 4 desatinné miesta došlo už pri 6 milisekunda prevádzky zdroja. Preto môžeme tvrdiť, že On/Off zdroj bude vysielat 87.5% z celkovej doby prevádzky.

Iný spôsob ako získať rozdelenie reťazca v čase n bol pomocou n -tej mocniny matice prechodov, $\mathbf{p}(n) = \mathbf{p}(0) \cdot \mathbf{P}^n$. Pozrieme sa najprv, ako vyzerá druhá mocnina matice \mathbf{P} vo všeobecnosti:

$$\mathbf{P}^2 = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \cdot \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} = \begin{bmatrix} p_{1,1}p_{1,1}+p_{1,2}p_{2,1} & p_{1,1}p_{1,2}+p_{1,2}p_{2,2} \\ p_{2,1}p_{1,1}+p_{2,2}p_{2,1} & p_{2,1}p_{1,2}+p_{2,2}p_{2,2} \end{bmatrix}$$

Všimnime si napríklad výraz $p_{2,1}p_{1,1}+p_{2,2}p_{2,1}$, ide výpočet pravdepodobnosti, ako sa dostať zo stavu s_2 do stavu s_1 za dva časové kroky, pričom vo výpočte sú zahrnuté všetky možnosti.

Postupne budeme počítat mocniny matice \mathbf{P} :

$$\mathbf{P}^1 = \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix}, \quad \mathbf{P}^2 = \begin{pmatrix} 0.88 & 0.12 \\ 0.84 & 0.16 \end{pmatrix}, \quad \mathbf{P}^3 = \begin{pmatrix} 0.876 & 0.124 \\ 0.868 & 0.132 \end{pmatrix},$$

$$\mathbf{P}^4 = \begin{pmatrix} 0.8752 & 0.1248 \\ 0.8736 & 0.1264 \end{pmatrix}, \quad \mathbf{P}^5 = \begin{pmatrix} 0.87504 & 0.12496 \\ 0.87472 & 0.12528 \end{pmatrix}$$

$$\mathbf{P}^6 = \begin{pmatrix} 0.875008 & 0.124992 \\ 0.874944 & 0.125056 \end{pmatrix}, \quad \mathbf{P}^7 = \begin{pmatrix} 0.8750016 & 0.1249984 \\ 0.8749888 & 0.1250112 \end{pmatrix}$$

Pri zaokrúhlení na 6 desatinných miest môžeme tvrdiť, že matica \mathbf{P}^7 má rovnaké riadky. To znamená, že po vynásobení ľubovoľným počiatočným rozdelením, napr. $(1, 0)$, $(0, 1)$, $(0.5, 0.5)$, $(0.7, 0.3)$, ..., dostávame vždy rovnaký výsledok, a to ľubovoľný riadok matice, ktorý predstavuje rozdelenie reťazca v čase $t = 7$. Vidíme opäť, že rozdelenie reťazca sa postupne stabilizuje v čase. O tomto efekte hovorí nasledujúca Markovova veta.

Markovov reťazec nazveme tranzitívny, ak všetky stavy sú navzájom dosiahnuteľné, tzn. pre všetky stavy s_i, s_j existuje počet krokov t_n, t_m , takých, že prechod zo stavu s_i do stavu s_j za čas t_n , a opačne z s_j do stavu s_i za čas t_m má nenulovú pravdepodobnosť:

$$\forall s_i, s_j, \quad \exists t_n, t_m; \quad p_{i,j}(t_n) > 0, \quad p_{j,i}(t_m) > 0$$

Doteraz uvažovaný 2-stavový reťazec je tranzitívny, s $t_n = 1$ a $t_m = 1$.

Markovova veta

Nech náhodný proces $X(t)$ s diskretným časom je tranzitívny a obsahuje aspoň v jednom stave slučku, $p_{i,i}(1) > 0$. Potom pravdepodobnosti stavov a prevdepodobnosti prechodov prestanú závisieť od času:

$$\forall j; \quad \lim_{n \rightarrow \infty} p_j(n) = \pi_j, \quad \text{resp.} \quad \forall i, j; \quad \lim_{n \rightarrow \infty} p_{i,j}(n) = \pi_j$$

$$\text{V maticovom tvare dostávame:} \quad \lim_{n \rightarrow \infty} \mathbf{p}(n) = \boldsymbol{\pi}, \quad \lim_{n \rightarrow \infty} \mathbf{P}_{j,\cdot}^n = \boldsymbol{\pi}$$

Markovova veta hovorí, že Markovov reťazec s istými vlastnosťami sa stabilizuje v čase, rozdelenie pravdepodobností reťazca prestane byť závislé od času. Toto rozdelenie označíme $\boldsymbol{\pi}$ a nazveme ho invariantné rozdelenie reťazca (steady-states probabilities). Ako sme ukázali v predchádzajúcom príklade, jeho hodnota je tiež invariantná na počiatočnom rozdelení.

Invariantné rozdelenie Markovovho reťazca

Vektor pravdepodobnosti $\boldsymbol{\pi}$ nazveme invariantné rozdelenie Markovovho reťazca (steady-states probabilities). Získame ho limitným prechodom v rekurentných vzťahoch:

$$\mathbf{p}(n) = \mathbf{p}(n-1) \cdot \mathbf{P} \Rightarrow \boldsymbol{\pi} = \boldsymbol{\pi} \cdot \mathbf{P}$$

Systém rovníc je lineárne závislý, pri riešení ľubovoľnú rovnicu nahradíme normovaciou podmienkou $\sum_{\forall j} \pi_j = 1$

PRÍKLAD 1.14 Pre Markovov reťazec z predchádzajúceho príkladu vypočítame invariantné rozdelenie pravdepodobnosti:

$$\begin{aligned}
 (\pi_1, \pi_2) &= (\pi_1, \pi_2) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} \Rightarrow \begin{aligned} \pi_1 &= 0.9\pi_1 + 0.7\pi_2 \\ \pi_2 &= 0.1\pi_1 + 0.3\pi_2 \end{aligned} \Rightarrow \\
 \Rightarrow \begin{aligned} 0.1\pi_1 - 0.7\pi_2 &= 0 \\ -0.1\pi_1 + 0.7\pi_2 &= 0 \end{aligned} \Rightarrow \begin{aligned} 0.1\pi_1 - 0.7\pi_2 &= 0 \\ 0 &= 0 \end{aligned} \Rightarrow \pi_2 = \frac{1}{7} \pi_1
 \end{aligned}$$

Z normovacej podmienky pre pravdepodobnosti dostávame

$$\pi_1 + \pi_2 = 1 \Rightarrow \pi_1 + \frac{1}{7} \pi_1 = 1 \Rightarrow \pi_1 = \frac{7}{8} = 0.875, \quad \pi_2 = \frac{1}{8} = 0.125$$

On/Off zdroj sa stabilizoval v čase a z celkového času prevádzky 87.5% vysielal. Ak konštantná doba vysielania bola napr. $R = 1$ [paket/ms], tak stredná intenzita toku, ktorú generuje takto nastavený On/Off zdroj je

$$\lambda_{avg} = R \cdot \pi_1 = 1 \cdot 0.875 = 0.875 \text{ [p/ms]} = 875 \text{ [p/s]}$$

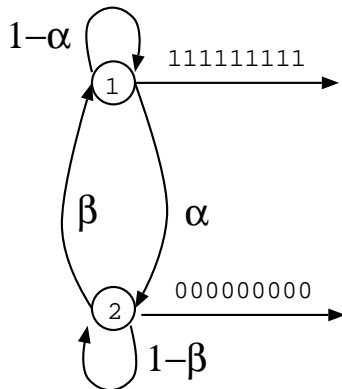
1.3.2 Markovov modulovaný regulárny proces (MMRP)

Markovov modulovaný regulárny proces predstavuje najjednoduchší stacionárny model On/Off zdroja IP prevádzky. Zdroj sa môže nachádzať v dvoch stavoch, v stave On (s_1) - zdroj vysielá pakety s konštantnou intenzitou, a stave Off (s_2) - zdroj nevysielá, resp. vysielá nuly. Označíme α a β pravdepodobnosti prechodov z On do Off, resp. z Off do On za jeden časový krok. O modulovanie procesu sa stará 2-stavový Markovov reťazec. Bez ujmy na všeobecnosti budeme predpokladať, že v stave On vysielá zdroj jednotky (bity, pakety, rámce, ATM bunky) s konštantnou rýchlosťou 1 [paket/ts].

Pre výpočet invariantného rozdelenia môžeme použiť nasledujúce úpravy:

$$\pi = \pi \cdot \mathbf{P} \Rightarrow \pi [\mathbf{E} - \mathbf{P}] = \mathbf{0} \Rightarrow [\mathbf{E} - \mathbf{P}]^T \pi^T = \mathbf{0} \quad (1.3)$$

Vypočítame stacionárne rozdelenie:



$$\mathbf{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad [\mathbf{E} - \mathbf{P}]^T = \begin{pmatrix} \alpha & -\beta \\ -\alpha & \beta \end{pmatrix}$$

$$\Rightarrow \alpha\pi_1 - \beta\pi_2 = 0 \Rightarrow \pi_2 = \frac{\alpha}{\beta} \pi_1 \Rightarrow$$

$$\Rightarrow \pi_1 + \frac{\alpha}{\beta} \pi_1 = 1 \Rightarrow \pi_1 \left[1 + \frac{\alpha}{\beta} \right] = 1$$

$$\Rightarrow \pi_1 = \frac{\beta}{\alpha + \beta}, \quad \pi_2 = \frac{\alpha}{\alpha + \beta}$$

Obr. 1.20: Prechodový graf pre MMRP

Vstupný proces modelovaný pomocou MMRP sa veľmi jednoducho simuluje. Ak sa proces nachádza v stave On, vygeneruje sa 1, ak v stave Off, vygeneruje sa 0.

Prepínanie medzi stavmi zdroja sa riadi pomocou pravdepodobností prechodov $p_{1,2} = \alpha$ a $p_{2,1} = \beta$.

V nasledujúcej ukážke sú vygenerované prírastky toku $a(i)$ v prvom riadku, v druhom riadku sú odpovedajúce pravdepodobnosti prepínania medzi stavmi. Proces začal v stave On, $a(1) = 1$, preto platí $\Pr(a(1) = 1) = 1$, v ďalšom kroku proces zostal v stave On, čomu odpovedá podmienená pravdepodobnosť $\Pr(a(2) = 1/a(1) = 1) = 1 - \alpha$, atď.....

$a(i)$	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0
$\Pr(.)$	1	$1-\alpha$	α	$1-\beta$	$1-\beta$	β	α	β	$1-\alpha$	$1-\alpha$	β	$1-\beta$	$1-\beta$	$1-\beta$	$1-\beta$

Označíme $p_k^m = P(a(i) = k)$, rozdelenie pravdepodobnosti prírastkov v časovom slotu, ktorý má dĺžku m elementárnych slotov. Ak sa budeme zaoberať prírastkom priamo v elementárnom slotu, $m = 1$, symbol m nebudeme používať.

Vypočítame rozdelenie pravdepodobnosti prírastkov toku v elementárnom časovom slotu, pričom použijeme vetu o úplnej pravdepodobnosti:

$$p_1 = \pi_1(1 - \alpha) + \pi_2\beta = \frac{\beta}{\alpha + \beta} = \pi_1, \quad p_0 = \pi_1\alpha + \pi_2(1 - \beta) = \frac{\alpha}{\alpha + \beta} = \pi_2$$

Výsledok sme mohli očakávať, hodnota prírastku $a(i)$ v elementárnom slotu závisí od toho, či sa zdroj nachádza v stave On alebo Off. Vypočítame momenty prírastkov $a(i)$:

$$Ea(i) = 0 \cdot \pi_2 + 1 \cdot \pi_1 = \pi_1, \quad Da(i) = (0 - \pi_1)^2 \pi_2 + (1 - \pi_1)^2 \pi_1 = \pi_1 \pi_2$$

Podľa získaných výsledkov by sme mohli povedať, že MMRP sa ničím nelýsi od Bernoulliho procesu, stačí položiť $p = \pi_1$ a $q = \pi_2$. Potom aj momenty prírastkov vyšli rovnako, ako pri Bernoulliho procese. Musíme si však uvedomiť, že MMRP proces riadi dvojstavový Markovov reťazec, v ktorom je závislosť medzi dvoma po sebe nasledujúcimi stavmi. Preto analyzovať MMRP proces len na základe situácie v elementárnych slotoch je zavádzajúce.

Odvodili sme rozdelenie pravdepodobnosti náhodných premenných $A(2)$, resp. prírastkov v dvoch elementárnych slotoch. Symboli v zátvorkách $\{.,.\}$ predstavujú situáciu v toku, ktorú pravdepodobnosti popisujú:

$$p_0^2 = \pi_2(1 - \beta) \quad \{00\}, \quad p_1^2 = \pi_1\alpha + \pi_2\beta \quad \{10, 01\}, \quad p_2^2 = \pi_1(1 - \alpha) \quad \{11\}$$

Odvodíme charakteristiky premennej $A(2)$:

$$EA(2) = 1 \cdot p_1^2 + 2 \cdot p_2^2 = \pi_1\alpha + \pi_2\beta + 2\pi_1 - 2\pi_1\alpha = \pi_2\beta + 2\pi_1 - \pi_1\alpha = 2\pi_1$$

$$\text{Využili sme vzťah } \pi_2\beta - \pi_1\alpha = \frac{\alpha}{\alpha + \beta} \beta - \frac{\beta}{\alpha + \beta} \alpha = 0.$$

Výslednú strednú hodnotu $EA(2) = 2\pi_1$ sme mohli očakávať, ide o strednú hodnotu v dvoch elementárnych slotoch.

$$EA^2(2) = 1 \cdot p_1^2 + 4 \cdot p_2^2 = \pi_1\alpha + \pi_2\beta + 4\pi_1 - 4\pi_1\alpha = \pi_2\beta + 4\pi_1 - 3\pi_1\alpha =$$

$$\begin{aligned}
&= 4\pi_1 - 2\pi_1\alpha \\
DA(2) &= EA^2(2) - E^2A(2) = 4\pi_1 - 2\pi_1 - 4\pi_1^2 = 4\pi_1\pi_2 - 2\pi_1\alpha = \\
&= 2\pi_1\pi_2 [2 - (\alpha + \beta)]
\end{aligned}$$

Pre Bernoulliho proces ale platí $DA(2) = 2pq$. V prípade MMRP hodnoty disperzie pre dva elementárne sloty ovplyvňujú istým spôsobom aj hodnoty pravdepodobnosti prechodov α a β v Markovovom reťazci.

Charakteristiky premennej $A(2)$

Nech náhodná premenná $A(2)$ popisuje veľkosť prírastkov MMRP toku počas dvoch elementárnych slotov. Potom rozdelenie pravdepodobnosti a momenty n.pr. $A(2)$ majú tvar:

$$\begin{aligned}
p_0^2 &= \pi_2(1 - \beta) & p_1^2 &= \pi_1\alpha + \pi_2\beta & p_2^2 &= \pi_1(1 - \alpha) \\
\pi_1 &= \frac{\beta}{\alpha + \beta}, & \pi_2 &= \frac{\alpha}{\alpha + \beta}, & EA(2) &= 2\pi_1, & DA(2) &= 2\pi_1\pi_2 [2 - (\alpha + \beta)]
\end{aligned}$$

Nepodarilo sa nám odvodiť všeobecné rozdelenie pravdepodobnosti náhodnej premennej $A(n)$. Uvádzame len rozdelenie pre $A(3)$ a $A(4)$:

$$\begin{aligned}
p_0^3 &= \pi_2(1 - \beta)^2 & \{000\} \\
p_1^3 &= (\pi_1\alpha + \pi_2\beta)(1 - \beta) + \alpha\beta\pi_2 & \{100, 010, 001\} \\
p_2^3 &= (\pi_1\alpha + \pi_2\beta)(1 - \alpha) + \alpha\beta\pi_1 & \{110, 101, 011\} \\
p_3^3 &= \pi_1(1 - \alpha)^2 & \{111\}
\end{aligned}$$

Rozdelenie pravdepodobnosti náhodnej premennej $A(4)$

$$\begin{aligned}
p_0^4 &= \pi_2(1 - \beta)^3 \\
p_1^4 &= (\pi_1\alpha + \pi_2\beta)(1 - \beta)^2 + 2\alpha\beta(1 - \beta)\pi_2 \\
p_2^4 &= (\pi_1\alpha + \pi_2\beta)(1 - \alpha)(1 - \beta) + \alpha\beta(\pi_1\alpha + \pi_2\beta + \pi_1(1 - \beta) + \pi_2(1 - \alpha)) \\
p_3^4 &= (\pi_1\alpha + \pi_2\beta)(1 - \alpha)^2 + 2\alpha\beta(1 - \alpha)\pi_1 \\
p_4^4 &= \pi_1(1 - \alpha)^3
\end{aligned}$$

Vo všeobecnosti môžeme zhrnúť poznatky toku v n -časových slotoch:

Všeobecné charakteris- tiky MMRP

Nech náhodné premenné $A(n)$ predstavujú prírastky MMRP procesu v n elementárnych slotoch. Potom platí:

$$\begin{aligned}
\lambda_{peak} &= \max_{\forall i} A(n) = n, & \lambda_{avg} &= EA(n) = n\pi_1 \\
p_0^n &= \pi_2(1 - \beta)^{n-1}, & p_{peak} &= p_n^n = \pi_1(1 - \alpha)^{n-1}
\end{aligned}$$

Budeme sa zaoberať otázkou, ako určiť parametre procesu α a β z nejakého nameraného toku. Ideálna je situácia, keď prírastky toku nadobúdajú hodnoty $A(2) = 0, 1, 2$, resp. sa zaoberáme meraním v dvoch časových slotoch. V takomto prípade vieme presný tvar strednej hodnoty a disperzie prírastkov, a pomocou ich odhadu z daného merania vieme zostaviť dve rovnice na výpočet pravdepodobností α a β .

V prípade bitového záznamu, resp. ak hodnoty prírastkov sú len $A(1) = 0, 1$, je situácia problematická. Nemôžeme pre odhad neznámych parametrov použiť iba informácie z jedného elementárneho slotu, pretože v procese vďaka modulovaniu Markovovým reťazcom je závislosť medzi dvoma po sebe idúcimi stavmi. V takomto prípade musíme buď meranie "naslotovať" na dva elementárne sloty, čiže sčítat vždy dva po sebe idúce prírastky. Získame tak ideálny tok pre analýzu s novými prírastkami $A(n) = 0, 1, 2$, pritom sa nám ale zmenší rozsah nameraných dát na polovicu. Iným riešením je analyzovať medzery medzi paketmi, o ktorých vďaka 0/1 toku máme plnú informáciu. Touto úlohou sa budeme zaoberať v nasledujúcej kapitole.

Vo všeobecnosti môžu mať prírastky meraného toku hodnoty $A(n) = 0, 1, \dots, n$. V takom prípade vieme zo základných parametrov toku λ_{peak} a λ_{avg} určiť vzťah medzi pravdepodobnosťami α a β :

$$\lambda_{peak} = n, \quad \lambda_{avg} = n\pi_1 \quad \Rightarrow \quad \pi_1 = \frac{\lambda_{avg}}{\lambda_{peak}} \quad \Rightarrow \quad \beta = \frac{\lambda_{avg} \cdot \alpha}{\lambda_{peak} - \lambda_{avg}}$$

Pre určenie parametra α môžeme použiť znalosť pravdepodobnosti prázdneho slotu alebo pravdepodobnosti peaku, a dať ju do rovnosti s odpovedajúcou pravdepodobnosťou procesu. Pri zostavovaní a následnej optimalizácii modelu IP siete je snaha pokryť okamihy špičkovej prevádzky. Preto pri odhadovaní parametrov procesu si vyberiem zhodu medzi pravdepodobnosťami peaku.

$$p_{peak} = \pi_1(1 - \alpha)^{n-1} \quad \Rightarrow \quad \alpha = 1 - \left(p_{peak} \cdot \frac{\lambda_{peak}}{\lambda_{avg}} \right)^{\frac{1}{n-1}}$$

V prípade, že $\lambda_{peak} > 4$ pre porovnanie rozdelenia pravdepodobností prírastkov meraného toku a prírastkov procesu MMRP musíme použiť dostatočne dlhú simuláciu.

Určenie parametrov MMRP

Nech nejaký nameraný tok má hodnoty prírastkov $A(n) = 0, 1, \dots, n$. Parametre α a β procesu MMRP určíme podľa vzťahov:

$$\beta = \frac{\lambda_{avg} \cdot \alpha}{\lambda_{peak} - \lambda_{avg}}, \quad \alpha = 1 - \left(p_{peak} \cdot \frac{\lambda_{peak}}{\lambda_{avg}} \right)^{\frac{1}{n-1}}$$

V takomto prípade sme dosiahli zhodu pravdepodobností výskytu preaku medzi modelom MMRP a nameranými dátami.

PRÍKLAD 1.15

Budeme sa zaoberať nameraným IP tokom, ktorým sme sa zaoberali v kapitole o Bernoulliho procese. Modelovanie pomocou Bernoulliho procesu sa ukázalo ako absolútne nevhodné. Tok mal formu záznamu prírastkov $A(4) = 0, 1, \dots, 4$, v časovom slotu 20 ms:

4100200000 0012444444 0142300104 0332110002 4103401124 4330034400

Z nameraných dát vieme určiť parametre

$$\lambda_{peak} = 4 \text{ [paket]}, \quad \lambda_{avg} = 1.7000 \text{ [paket/20ms]}, \quad p_{peak} = 0.2500$$

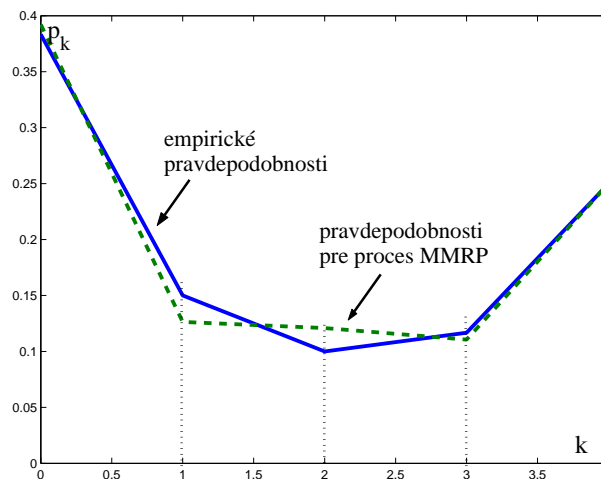
Učíme parametre procesu α a β vzhľadom na rovnosť pravdepodobnosti peaku.

$$\alpha = 1 - \left(0.25 \cdot \frac{4}{1.70}\right)^{\frac{1}{3}} = 0.1621 \quad \Rightarrow \quad \beta = \frac{1.70 \cdot 0.1621}{4 - 1.70} = 0.1198$$

V tabuľke uvedieme empirické pravdepodobnosti \bar{p}_k vypočítané z dát a pravdepodobnosti p_k pre MMRP proces s odhadnutými hodnotami α a β :

k	0	1	2	3	4
\bar{p}_k	0.3833	0.1500	0.1000	0.1167	0.2500
p_k	0.3921	0.1264	0.1209	0.1106	0.2500

Medzi pravdepodobnosťami panuje zhoda, preto s istotou môžeme tvrdiť, že nameraný tok odpovedá MMRP modelu.



Obr. 1.21: Porovnanie odhadnutých pravdepodobností s MMRP procesom

1.3.3 Rozdelenie pravdepodobnosti medzier a zhlučov

Medzeru v MMRP procese budeme definovať ako postupnosť núl (v elementárnom slotu sa nevyskytol paket), ktorú ukončí výskyt jednotky (paketu). Medzera začína prepnutím zdroja do stavu Off, preto pravdepodobnosť medzery o dĺžke 5, 1000001, má pravdepodobnosť $\alpha(1 - \beta)^4\beta$. Aby od j -teho časového slotu začínala medzera o dĺžke t , musí byť prírastok toku v čase j rovný nule, $a(j) = 0$.

Nech náhodná premenná T popisuje veľkosť medzery medzi paketmi v toku. Výskyt paketov je navzájom nezávislý.

$$p_0 = \Pr(T = 0) = 1 - \alpha, \quad p_t = \Pr(T = t) = \alpha(1 - \beta)^{t-1}\beta, \quad t = 1, 2, \dots$$

Vypočítame MGF pre náhodnú premennú T :

$$\begin{aligned} \varphi(\theta) &= \sum_{t=0}^{\infty} e^{\theta t} p_t = 1 - \alpha + \sum_{t=1}^{\infty} e^{\theta t} \alpha(1 - \beta)^{t-1} \beta = \\ &= 1 - \alpha + \alpha\beta e^{\theta} \sum_{t=1}^{\infty} (e^{\theta}(1 - \beta))^{t-1} = 1 - \alpha + \frac{\alpha\beta e^{\theta}}{1 - (1 - \beta)e^{\theta}} \end{aligned}$$

Pomocou derivácii MGF určíme prvý a druhý počiatkový moment:

$$\begin{aligned} ET &= \varphi'(\theta)|_0 = \frac{\alpha\beta e^{\theta}}{(1 - (1 - \beta)e^{\theta})^2} \Big|_0 = \frac{\alpha}{\beta} \\ ET^2 &= \varphi'(\theta)|_0 = \frac{\alpha\beta e^{\theta}(1 + (1 - \beta)e^{\theta})}{(1 - (1 - \beta)e^{\theta})^3} \Big|_0 = \frac{\alpha}{\beta^2}[2 - \beta] \end{aligned}$$

Určíme disperziu medzier:

$$DT = ET^2 - (ET)^2 = \frac{\alpha}{\beta^2}[2 - \beta] - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}[2 - (\alpha + \beta)]$$

Koeficient variabilnosti medzier v MMRP procese má tvar:

$$\nu_T = \frac{\sigma_T}{ET} = \frac{\sqrt{\alpha(2 - (\alpha + \beta))/\beta}}{\alpha/\beta} = \sqrt{\frac{2 - (\alpha + \beta)}{\alpha}}$$

Pre určenie parametrov α a β môžeme použiť nasledujúce vzťahy:

$$\begin{aligned} DT &= \frac{\alpha}{\beta^2}[2 - (\alpha + \beta)] = \frac{\alpha}{\beta} \left[\frac{2}{\beta} - \left(\frac{\alpha}{\beta} + 1 \right) \right] = ET \left[\frac{2}{\beta} - (ET + 1) \right] \Rightarrow \\ \Rightarrow \quad \beta &= \frac{2ET}{ET^2 + ET}, \quad \alpha = \beta \cdot ET \end{aligned}$$

Rozdelenie pravdepodobnosti medzier

Nech náhodná premenná T popisuje veľkosť medzery v MMRP. Rozdelenie pravdepodobnosti n.pr. T má tvar:

$$\Pr(T = 0) = 1 - \alpha, \quad \Pr(T = t) = \alpha(1 - \beta)^{t-1}\beta, \quad t = 1, 2, \dots,$$

Momenty náhodnej premennej sú:

$$ET = \frac{\alpha}{\beta}, \quad DT = \frac{\alpha}{\beta^2}(2 - (\alpha + \beta)), \quad \nu_T = \sqrt{\frac{2 - (\alpha + \beta)}{\alpha}}$$

Situácia s rozdelením pravdepodobnosti paketových zhlučkov je analogická.

Rozdelenie pravdepodobnosti zhlukov

Nech náhodná premenná Z popisuje veľkosť paketového zhluku v MMRP. Rozdelenie pravdepodobnosti n.pr. Z má tvar:

$$\Pr(Z = 0) = 1 - \beta, \quad \Pr(Z = z) = \beta(1 - \alpha)^{z-1}\alpha, \quad z = 1, 2, \dots,$$

Momenty náhodnej premennej sú:

$$EZ = \frac{\beta}{\alpha}, \quad DZ = \frac{\beta}{\alpha^2}(2 - (\beta + \alpha)), \quad \nu_Z = \sqrt{\frac{2 - (\beta + \alpha)}{\beta}}$$

1.3.4 Porovnanie Bernoulliho procesu a MMRP

Budeme sa zaoberať vzájomným porovnaním Bernoulliho a MMRP procesov, pričom najprv si zhrnieme ich základné charakteristiky. Obidva procesy môžeme považovať za zdroj bitového toku, reps. toku núl a jednotiek, pričom jednotka predstavuje výskyt paketu (rámca, ATM bunky, bitu). V Bernoulliho procese je výskyt paketov navzájom nezávislý, s rovnakou pravdepodobnosťou v elementárnom slot. Proces MMRP je riadený 2-stavovým Markovovým reťazcom, ktorý zdroj prepína medzi stavmi On a Off. V stave On zdroj vygeneruje jednotku, v stave Off nulu. Pravdepodobnosť výskytu napr. 1 v danom slot závisí od toho, v ako stave sa reťazec nachádzal v predchádzajúcom slot:

$$\Pr(a(i) = 1/a(i-1) = 0) = \beta, \quad \Pr(a(i) = 1/a(i-1) = 1) = 1 - \alpha$$

Vo všeobecnosti ide o dva navzájom odlišné stochastické procesy. Avšak pri správnom nastavení pravdepodobností prechodov α a β Markovov modulovaný proces sa zhoduje s Bernoulliho procesom. Na určenie parametrov MMRP procesu použijeme disperziu výskytu paketov pre dva elementárne sloty, $A(2)$. Pre Bernoulliho proces má tvar $2pq$, pre MMRP proces disperzia okrem súčinnu pravdepodobností stavov zahŕňa aj väzbu s α a β . Na určenie parametrov použijeme zhodu medzi disperziami:

$$2pq \equiv 2\pi_1\pi_2[2 - (\alpha + \beta)] \Rightarrow [2 - (\alpha + \beta)] = 1 \Rightarrow \alpha + \beta = 1$$

Ak sú pravdepodobnosti prechodov doplnkové, $\beta = 1 - \alpha$, potom MMRP proces je zhodný s Bernoulliho procesom. Pre pravdepodobnosti stavov bude platiť:

$$\pi_1 = \frac{\beta}{\alpha + \beta} = \beta, \quad \pi_2 = \frac{\alpha}{\alpha + \beta} = \alpha$$

Rozdelenie pravdepodobnosti napr. premennej $A(2)$ s zmení na binomické.

$$p_0^2 = \pi_2(1 - \beta) = \alpha^2, \quad p_1^2 = \pi_1\alpha + \pi_2\beta = 2\alpha\beta, \quad p_2^2 = \pi_1(1 - \alpha) = \beta^2$$

PRÍKLAD 1.16

V nasledujúcich ukážkach simulácií MMRP procesu s prírastkami $a(i) = 0, 1, 2$, budeme meniť pravdepodobnosti α a β , ale strednú intenzitu prírastkov necháme rovnakú, $Ea(i) = EA(2) = 2\pi = 1$. To znamená, že musí platiť $\alpha = \beta$.

V prvej ukážke necháme $\alpha = \beta = 0.5$, pôjde vlastne o Bernoulliho tok so smerodajnou odchýlkou $\sigma = \sqrt{DA(2)} = \sqrt{0.5} = 0.7071$:

12101101102121001112101201120211121111111200111112111122100

V druhej ukážke sa budeme snažiť, aby disperzia $DA(2)$ bola čo najmenšia:

$$2\pi_1\pi_2[2 - (\alpha + \beta)] \rightarrow 0 \Rightarrow (\alpha + \beta) \rightarrow 2 \Rightarrow \alpha, \beta \rightarrow 1$$

Parametre MMRP nastavíme napr. na $\alpha = \beta = 0.9$, potom $\sigma = 0.2357$:

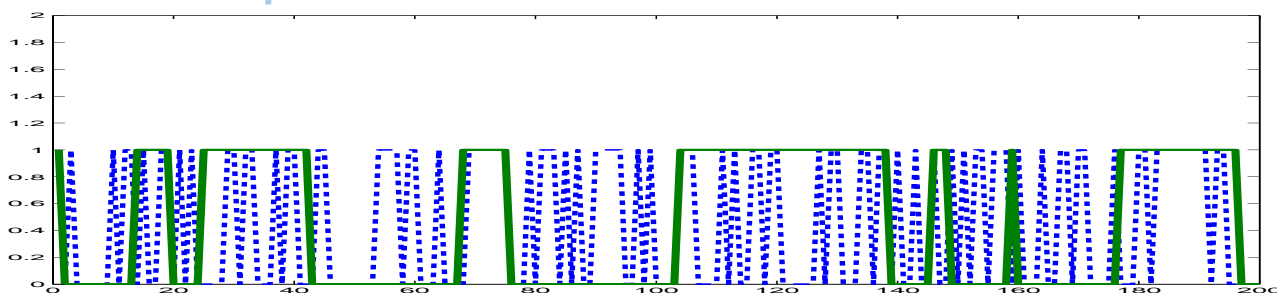
11111111111111111111110110111111111121111111111111111111111101112

Dostali sme tok, o ktorom by sme mohli považovať za skoro deterministický.

V tretej ukážke nastavíme $\alpha = \beta = 0.1$, potom $\sigma = 2.1213$:

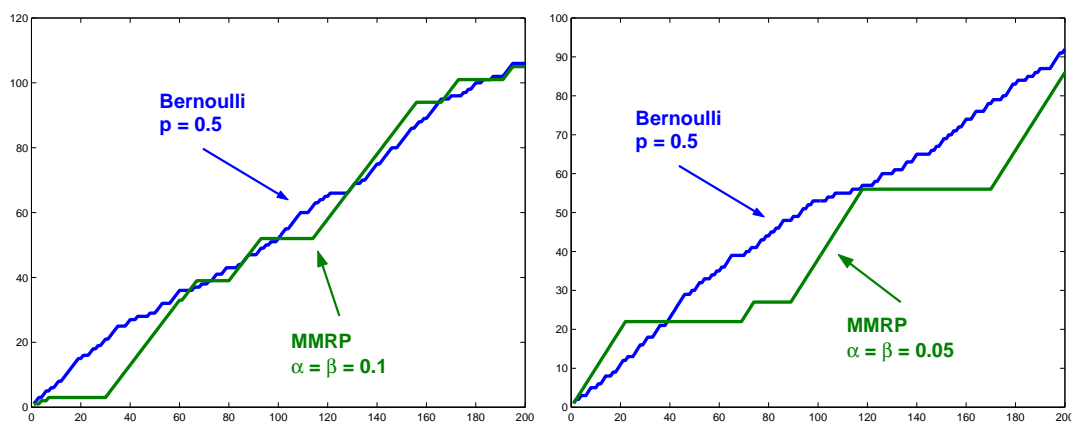
22222222222222220000002222222222012222211000011000000000000000

Dostali sme zaujímavý tok, v ktorom sa striedajú hlavne zhlučky dvojiek a núl. Porovnáme ho s Bernoulliho procesom s rovnakou strednou hodnotou $Ea(i)$, pričom prírastky majú hodnotu $a(i) = 0, 1$:



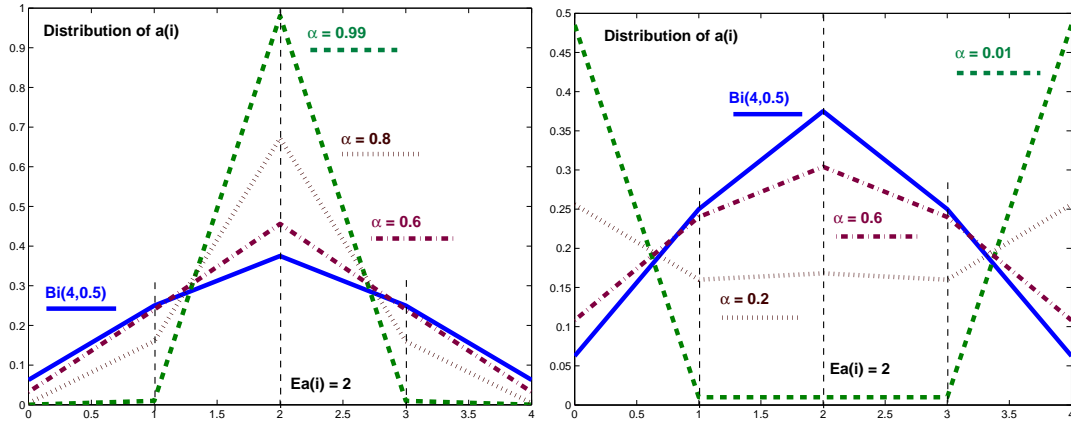
Obr. 1.22: Porovnanie MMRP (plná čiara) a Bernoulliho procesu (prerušovaná čiara)

Pri vhodnom nastavení pravdepodobností MMRP modeluje zhlučovitý proces. Vzťah medzi veľkosťou zhlučkov a pravdepodobnosťami α a β ukážeme na kumulatívnych procesoch $A(t)$:



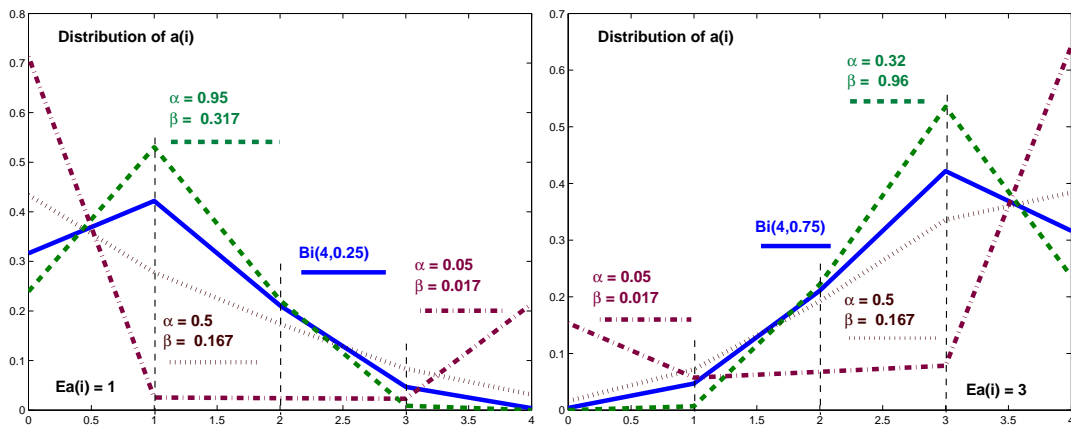
Obr. 1.23: Porovnanie kumulatívnych procesoch s rovnakou strednou hodnotou $Ea(i) = 0.5$

Variabilnosť procesu MMRP budeme demonštrovať na rozdelení pravdepodobnosti prírastkov v 4 elementárnych slotov. Modulovaný proces porovnáme s Bernoulliho procesom rovnakou strednou hodnotou $Ea(i) = 4p = 4\pi_1$. V prvom prípade nastavíme strednú hodnotu na $Ea(i) = 2$. Bernoulliho proces popisuje binomické rozdelenie $A(4) \sim Bi(4, 0.5)$. Pre MMRP proces máme mnoho možností, pričom musí platiť $\alpha = \beta$. Všetky rozdelenia sú symetrické:



Obr. 1.24: Bernoulliho proces $Bi(4, 0.5)$ a MMRP $\alpha, \beta \rightarrow 1$ a $\alpha, \beta \rightarrow 0$

V prípade, že sa pravdepodobnosti prechodov blížia k 1, dostávame takmer deterministický tok, v ktorom pre väčšinu prírastkov platí $a(i) = 2$. Ak sa pravdepodobnosti prechodov blížia k 0, dostávame zhlukovitý tok, s prírastkami $a(i) = 0, a(i) = \lambda_{peak}$. V prípade, že $\alpha = 0.5$, MMRP sa zmenil na Bernoulliho proces. Na ďalších obrázkoch si ukážeme variabilnosť rozdelenia MMRP pre prípady $Ea(i) = 1$ a $Ea(i) = 3$.



Obr. 1.25: Bernoulli process and particular MMRP distribution

Oproti Bernoulliho procesu poskytuje Markovov modulovaný regulárny proces podstatne viac možností pre modelovanie rôznych typov tokov, od skoro deterministického, cez samotný Bernoulli až po tok s vysokou zhlukovitou dát. Jeho nevýhodou je, že neexistuje všeobecné rozdelenie pravdepodobnosti prírastku, čo významne komplikuje analýzu nameraných dát.

1.4 POISSONOV MODEL

Doteraz sme sa zaoberali Bernoulliho procesom a Markovovým modulovaným regulárnym procesom. Obidva sú procesy s diskretným časom, to znamená, že medzeru medzi výskytom paketov modelovala nejaká diskretná náhodná premenná (v obidvoch prípadoch geometrická). Inými slovami, medzera medzi paketmi bola tvorená konečným počtom elementárnych slotov. Pre obidva procesy tiež platí, že hodnoty prírastkov v elementárnom slotu sú len $a(i) = 0, 1$. V tejto kapitole sa budeme zaoberať procesom, ktorého medzery sú modelované spojitou náhodnou premennou a prírastky v elementárnom slotu sú neohraničené.

Poissonov proces

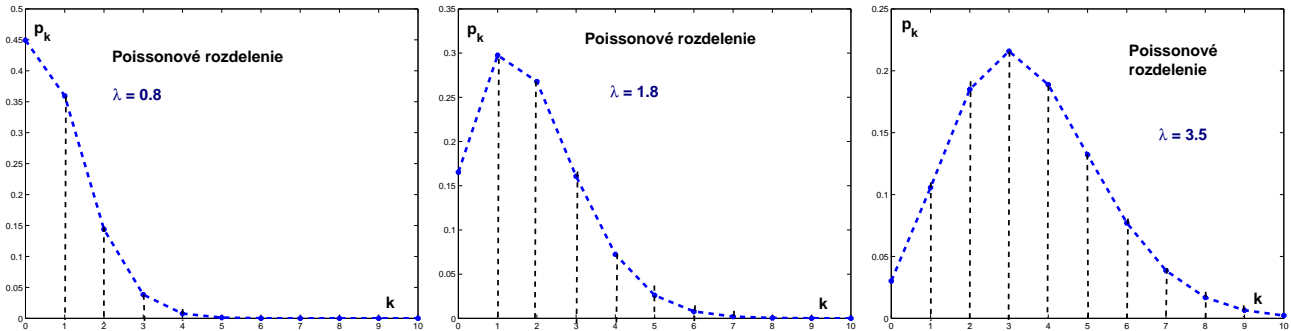
Poissonov proces nazveme stochastický proces, ktorý má nezávislé, rovnako rozdelené prírastky s Poissonovým rozdelením pravdepodobnosti $a(i) \sim Po(\lambda)$:

$$p_k = \Pr(a(i) = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Poissonov proces označíme $A(t) \sim Po(\lambda t)$.

Poissonové rozdelenie pravdepodobnosti môžeme vyjadriť rekurentne:

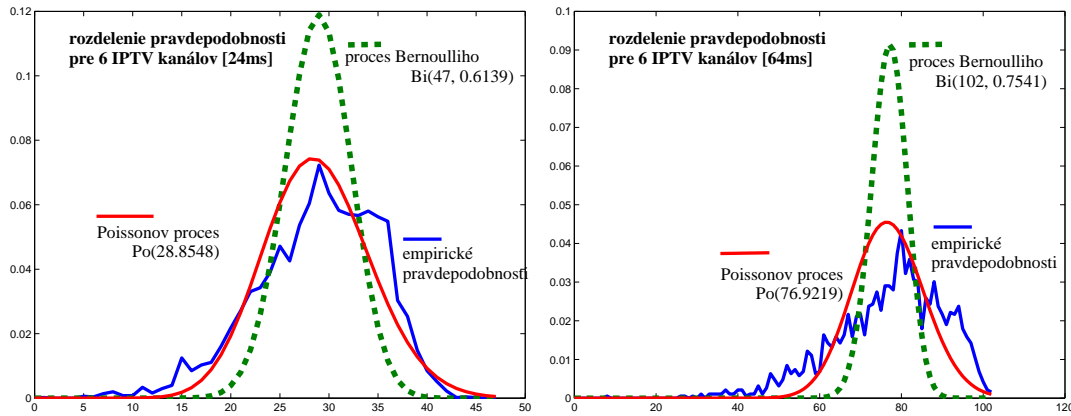
$$\frac{p_k}{p_{k-1}} = \frac{\frac{\lambda^k}{k!} e^{-\lambda}}{\frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}} = \frac{\lambda}{k} \Rightarrow p_k = \frac{\lambda}{k} p_{k-1}, \quad k = 1, 2, \dots \quad p_0 = e^{-\lambda}$$



Obr. 1.26: Ukážky Poissonovho rozdelenia prírastkov $a(i)$

Prírastky v Poissonovom procese môžu nadobúdať hodnoty $a(i) = 0, 1, 2, \dots$, resp. inými slovami, neexistuje peak procesu, $\nexists \max_{\forall i} a(i)$, čo môže predstavovať nedostatok pri modelovaní reálneho toku. Na druhú stranu, od nejakého čísla n_0 je pravdepodobnosť výskytu javu $\Pr(a(i) = n_0)$ veľmi malá až zanedbateľná. Historicky sa Poissonove rozdelenie používalo na modelovanie tzv. riedkych javov, resp. javov, ktorých pravdepodobnostne významný počet výskytov bol len po nejaké prirodzené číslo n_0 . Poissonovo rozdelenie má však niekoľko zaujímavých vlastností, vďaka ktorým sa stalo zaujímavým pri vytváraní modelov v Telekomunikáciách.

Nasledujúci obrázok ukazuje prípad, kedy modelovanie toku pomocou Poissonovho procesu dáva lepšiu aproximáciu, než použitie Bernoulliho procesu.



Obr. 1.27: porovnanie EB pre IPTV a Bernoulliho proces $Bi(11t, 0.2048)$

1.4.1 Momenty a charakteristiky

Vypočítame strednú hodnotu n.pr. s Poissonovým rozdelením $X \sim Po(\lambda)$:

$$EX = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{\lambda} e^{-\lambda} = \lambda$$

Ak majú prírastky toku Poissonove rozdelenie, $a(i) \sim Po(\lambda)$, potom pre ich strednú hodnotu platí $\lambda_{avg} = Ea(i) = \lambda$. Náhodný proces $A(t)$ má potom strednú hodnotu $EA(t) = \lambda t$. Parameter λ nazveme intenzita procesu (resp. toku) a predstavuje stredný počet výskytu paketov za časovú slotu.

Poissonov proces môžeme použiť ako "dobrú" aproximáciu Bernoulliho procesu v prípade, keď pravdepodobnosť výskytu paketu je veľmi malá, $p \rightarrow 0$, a zároveň uvažujeme o veľkom časovom slotu, $n \rightarrow \infty$. Parameter pre Poissonove rozdelenie získame s rovnosťou stredných hodnôt, $\lambda = np$.

$$\begin{aligned} p_k &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} = \left(1 - \frac{\lambda}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{\lambda^k}{k!} \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} \end{aligned}$$

V prípade, že $n \rightarrow \infty$, platí $\forall c \in R$, $\lim_{n \rightarrow \infty} \frac{c}{n} = 0$ a $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$. Z binomického rozdelenia sme dostali rozdelenie Poissonovo, $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$.

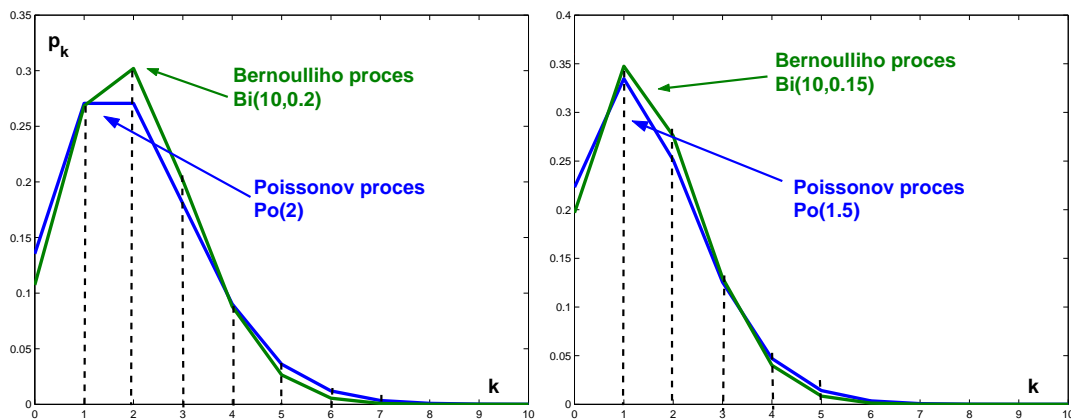
PRÍKLAD 1.17 | Nech stredná intenzita a peak nejakého nameraného toku sú $\lambda_{avg} = 2$ [p/ms] a $\lambda_{peak} = 10$ [p/ms]. Vypočítame parametre pre Bernoulliho a Poissonov proces a navzájom porovnáme pravdepodobnosť rozdelenia počtu paketov v 1ms.

Pre Bernoulliho proces dostávame: $n = 10, p = \lambda_{avg}/\lambda_{peak} = 0.2$, pre Poissonov proces: $\lambda = \lambda_{avg} = 2$. Pre rozdelenie prírastkov platí:

$$\text{Bernoulli: } \Pr(a(i) = k) = \binom{n}{k} p^k q^{n-k}, \quad \text{Poisson: } \Pr(a(i) = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

k	0	1	2	3	4	5	6	7	8	9	10
$Bi(n, p)$	0.107	0.268	0.302	0.201	0.088	0.026	0.005	0.001	0	0	0
$Po(\lambda)$	0.135	0.271	0.271	0.180	0.090	0.036	0.012	0.003	0	0	0

V tabuľke sú uvedené rozdelenia pravdepodobnosti prírastkov pre Bernoulliho a Poissonov proces. Aj keď zďaleka neplatí $p \rightarrow 0$ a $n \rightarrow \infty$, na tvare rozdelení môžeme pozorovať istú zhodu.



Obr. 1.28: porovnanie rozdelení prírastkov pre Bernoulliho a Poissonov proces

PRÍKLAD 1.18

Určíme CGF $\lambda(\theta)$ pre Poissonovo rozdelenie:

$$\begin{aligned} \lambda(\theta) &= \ln \left[\sum_{k=0}^{\infty} e^{\theta k} \frac{\lambda^k}{k!} e^{-\lambda} \right] = \ln \left[e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{\theta})^k}{k!} \right] = \ln \left[e^{-\lambda} e^{\lambda e^{\theta}} \right] \\ &\Rightarrow \lambda(\theta) = \lambda [e^{\theta} - 1] \end{aligned} \quad (1.4)$$

Vypočítame prvú a druhú deriváciu v bode $\theta = 0$:

$$\lambda'(0) = \lambda e^{\theta} \Big|_{\theta=0} = \lambda, \quad \lambda''(0) = \lambda e^{\theta} \Big|_{\theta=0} = \lambda$$

Dostali sme samozrejme strednú hodnotu $Ea(i) = \lambda$ a disperziu $Da(i) = \lambda$.

Charakteristiky Poissonovho procesu

Nech náhodný proces $A(t)$ je Poissonov s prírastkami $a(i)$. Potom rozdelenie pravdepodobnosti, momenty a kumulatívna generujúca funkcia majú tvar:

$$\Pr(a(i) = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad Ea(i) = \lambda, \quad Da(i) = \lambda, \quad \lambda(\theta) = \lambda [e^\theta - 1]$$

1.4.2 Rozdelenie medzier

Rozdelenie medzier medzi výskytom paketov v Poissonovom procese popíšeme náhodnými premennými T . Určíme ich distribučnú funkciu $F(t)$:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - P(A(t) = 0) = 1 - e^{-\lambda t} \quad (1.5)$$

Pomocou derivácie podľa t vypočítame hustotu rozdelenia pravdepodobnosti doby medzier v toku:

$$f(t) = F'(t) = (1 - e^{-\lambda t})' = \lambda e^{-\lambda t} \quad (1.6)$$

Odvodili sme distribučnú funkciu, resp. hustotu exponenciálneho rozdelenia, $T \sim \exp(\lambda)$. Ak výskyt paketov tvorí Poissonov proces, potom medzery medzi paketmi sú exponenciálne. Platí aj opačné tvrdenie: Ak medzery medzi udalosťami sú nezávislé s exponenciálnym rozdelením, potom výskyt udalostí tvorí Poissonov proces.

Určíme distribučnú funkciu náhodných premenných $A(t)$ keď vieme, že medzery sú nezávislé, s exponenciálnym rozdelením. Udalosť $A(t) < k$ znamená, že v čase t sa ešte nevyskytla k -ta udalosť (výskyt k -tej udalosti nastane až po dobe dlhšej ako t). To ale znamená, že súčet nasledujúcich k medzier bude väčší než t , $T_1 + \dots + T_k > t$. Keďže predpokladáme, že náhodné premenné T_i sú nezávislé s rozdelením $\exp(\lambda)$, ich súčet má rozdelenie Erlangovo $E(k, \lambda)$:

$$F_{A(t)}(k) = P[A(t) < k] = P[T_1 + \dots + T_k > t] = \int_t^\infty \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} dx$$

Integrál vyriešime postupným integrovaním metódou per partes:

$$\begin{aligned} \int_t^\infty \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} dx &= \frac{(\lambda t)^{k-1} e^{-\lambda t}}{(k-1)!} + \int_t^\infty \frac{\lambda^{k-1} x^{k-2} e^{-\lambda x}}{(k-2)!} dx = \dots = \sum_{i=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^i}{i!} \\ \Rightarrow F_{A(t)}(k) &= \sum_{i=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^i}{i!} \end{aligned}$$

Pomocou distribučnej funkcie vypočítame rozdelenie pravdepodobnosti $A(t)$:

$$\begin{aligned} P[A(t) = k] &= P[A(t) < k+1] - P[A(t) < k] = \\ &= \sum_{i=0}^k e^{-\lambda t} \frac{(\lambda t)^i}{i!} - \sum_{i=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^i}{i!} = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad \Rightarrow \quad A(t) \sim Po(\lambda t) \end{aligned}$$

**Poissonov
proces**

Ak sa výskyt udalostí vo vstupnom procese $A(t)$ riadi Poissonovým rozdelením s intenzitou λ , potom medzery medzi udalosťami T majú exponenciálne rozdelenie so strednou hodnotou $ET = 1/\lambda$. Resp. naopak, ak sú medzery medzi udalosťami nezávislé, s rovnakým exponenciálnym rozdelením, potom výskyt udalostí sa riadi podľa Poissonovho procesu:

$$A(t) \sim Po(\lambda), \quad EA(t) = \lambda \quad \Leftrightarrow \quad T \sim exp(\lambda), \quad ET = \frac{1}{\lambda}$$

V predchádzajúcom tvrdení sme použili strednú hodnotu exponenciálneho rozdelenia. Ukážeme si, ako ju môžeme vypočítať, spolu s disperziou, použitím kumulatívne vytvárajúcej funkcie.

PRÍKLAD 1.19

Určíme CGF $\lambda(\theta)$ pre exponenciálne rozdelenie:

$$\begin{aligned} \lambda(\theta) &= \ln \left[\int_{t=0}^{\infty} e^{\theta t} \lambda e^{-\lambda t} dt \right] = \ln \left[\lambda \frac{e^{-(\lambda-\theta)t}}{-(\lambda-\theta)} \right]_0^{\infty} = \ln \left[\frac{\lambda}{\lambda-\theta} \right] \\ \Rightarrow \quad \lambda(\theta) &= \ln \lambda - \ln(\lambda - \theta) \quad \theta < \lambda \end{aligned} \quad (1.7)$$

Vypočítame prvú a druhú deriváciu v bode $\theta = 0$:

$$\lambda'(0) = \frac{1}{\lambda - \theta} \Big|_{\theta=0} = \frac{1}{\lambda}, \quad \lambda''(0) = \frac{1}{(\lambda - \theta)^2} \Big|_{\theta=0} = \frac{1}{\lambda^2}$$

Dostali sme strednú hodnotu $ET = 1/\lambda$ a disperziu $DT = 1/\lambda^2$.

Aj exponenciálne rozdelenie má vlastnosť "zabúdanie na minulosť". Pravdepodobnosť, že medzera medzi paketmi bude ešte trvať viac než čas t , ak vieme že už trvala τ , nezáleží od τ :

$$\begin{aligned} P(T > t + \tau / T > \tau) &= \frac{P(T > t + \tau \text{ \& } T > \tau)}{P(T > \tau)} = \frac{P(T > t + \tau)}{P(T > \tau)} = \\ &= \frac{e^{-(t+\tau)\lambda}}{e^{-\tau\lambda}} = e^{-t\lambda} = P(T > t) \end{aligned}$$

Výsledná pravdepodobnosť je závislá iba od času t a doba trvania obsluhy τ nemá vplyv na výslednú pravdepodobnosť. Exponenciálne rozdelenie je spojitým ekvivalentom geometrického rozdelenia a používa sa pri tvorbe Markovových reťazcov so spojitým časom.

1.4.3 Multiplex a routing

Odvodíme distribučnú funkciu medzier medzi paketmi v súčte dvoch nezávislých Poissonových procesov, $\tau_1 \sim Po(\lambda_1 t)$ a $\tau_2 \sim Po(\lambda_2 t)$:

$$\begin{aligned} F(t) &= P(\tau \leq t) = 1 - P(\tau > t) = 1 - P(\tau_1 > t) \cdot P(\tau_2 > t) = \\ &= 1 - e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = 1 - e^{-(\lambda_1 + \lambda_2)t} \end{aligned} \quad (1.8)$$

Súčet dvoch nezávislých Poissonových procesov je opäť proces Poissonov so súčtom intenzít, $A(t) \sim Po((\lambda_1 + \lambda_2)t)$.

Chceme zistiť proces, ktorý vznikne, keď z toku modelovaného Poissonovým proces náhodne presmerujeme $100p$ % prevádzky, resp. paket bol do novej vetvy presmerovaný z pôvodného toku s pravdepodobnosťou p . Odvodíme distribučnú funkciu $F_{new}(t)$ medzier τ_{new} medzi paketmi v novovzniknutom toku (medzery v pôvodnom toku majú rozdelenie $\tau \sim exp(\lambda)$):

$$\begin{aligned} F_{new}(t) &= P(\tau_{new} \leq t) = 1 - P(\tau_{new} > t) = 1 - \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t} \cdot (1-p)^k = \\ &= 1 - \sum_{k=0}^{\infty} \frac{[\lambda t(1-p)]^k}{k!} \cdot e^{-\lambda t} = 1 - e^{-\lambda t} \cdot e^{\lambda t(1-p)} = 1 - e^{-\lambda tp} \end{aligned} \quad (1.9)$$

Proces na novej vetve je Poissonov s intenzitou λp , $A_{new}(t) \sim Po(\lambda p t)$.

Uvedené dve vlastnosti umožňujú modelovať IP sieť pomocou *Jacksonovej siete*, v ktorej uzol predstaviť jednolinkový Markovov server s nekonečnou frontou. Táto problematika je dobre rozpracovaná v Teórii hromadnej obsluhy a my sa s ňou nebudeme zaoberať.

1.5 EFEKTÍVNA ŠÍRKA PÁSMO (EB)

1.5.1 Základné vlastnosti Efektívnej šírky pásma

Jeden zo spôsobov, ako popísať vstupný tok paketov $A(t)$ je tzv. Efektívna šírka pásma (Effective Bandwidth - EB). V tejto kapitole si ukážeme vlastnosti EB a výpočet pre jednotlivé toky. Priamym využitím Efektívnej šírky pásma pri dimenzovaní prenosovej rýchlosti sa budeme zaoberať nasledujúcich kapitolách.

Pojem Efektívna šírka pásma definoval Kelly v [5] pre ľubovoľný kumulatívny náhodný proces $A(t)$ nasledovne:

$$\alpha(\theta, t) = \frac{1}{\theta t} \sup_{s \geq 0} \ln E \left[e^{\theta(A(s+t) - A(s))} \right] \quad 0 < \theta, t < \infty \quad (1.10)$$

Parametre t a θ budeme nazývať časový parameter (*time parameter*) a rozmerový parameter (*scale parameter*).

Ak má proces $A(t)$ **stacionárne** prírastky $a(i)$, rozdelenie pravdepodobnosti rozdielu $A(s+t) - A(s) = \sum_{i=s+1}^{s+t} a(i)$ je rovnaké ako rozdelenie pravdepodobnosti $A(t) = \sum_{i=1}^t a(i)$. V takomto prípade môžeme položiť $s = 0$. Tvar Efektívnej šírky pásma sa zjednoduší:

$$\alpha(\theta, t) = \frac{1}{\theta t} \ln E \left[e^{\theta A(t)} \right] \quad (1.11)$$

EB má niekoľko zaujímavých vlastností, ako prvú si ukážeme, čomu sa rovná hodnota EB v bode $\theta = 0$:

$$\alpha(0, t) = \lim_{\theta \rightarrow 0} \frac{1}{\theta t} \ln E \left[e^{\theta A(t)} \right] = \lim_{\theta \rightarrow 0} \frac{1}{\theta t} \frac{E \left[A(t) e^{\theta A(t)} \right]}{E \left[e^{\theta A(t)} \right]} = \frac{E \left[A(t) \right]}{t}$$

Pri určení limity sme použili L'Hospitalove pravidlo. Hodnota EB v $\theta = 0$ sa dá chápať ako intenzita toku.

Ďalej určíme hodnotu derivácie EB v nule:

$$\alpha'(0, t) = \left. \frac{\frac{E \left[A(t) e^{\theta A(t)} \right]}{E \left[e^{\theta A(t)} \right]} \theta - \ln E \left[e^{\theta A(t)} \right]}{t \theta^2} \right|_{\theta=0} = \frac{E \left[A(t)^2 \right] - E^2 \left[A(t) \right]}{2t}$$

Dostali sme polovicu variancie procesu $D[A(t)]$ za čas t .

Na určenie derivácie sme viac krát použili L'Hospitalove pravidlo pre výraz:

$$\dots = \lim_{\theta \rightarrow 0} \frac{1}{2t\theta} \frac{E \left[A(t)^2 e^{\theta A(t)} \right] E \left[e^{\theta A(t)} \right] - E^2 \left[A(t) e^{\theta A(t)} \right]}{E \left[e^{\theta A(t)} \right]^2} \cdot \theta = \dots$$

Presné odvodenie prenecháme ako jednoduché cvičenie na derivácie pre čitateľa.

Vlastnosti EB pre stacionárny proces

Nech $\alpha(\theta, t)$ je Efektívna šírka pásma pre stacionárny proces $A(t)$. Potom platí

$$\alpha(0, t) = \frac{E[A(t)]}{t}, \quad \alpha'(0, t) = \frac{E[A(t)^2] - E^2[A(t)]}{2t} = \frac{D[A(t)]}{2}$$

Ak má stacionárny proces $A(t)$ zhora ohraničené prírastky, $\exists c \in R, \forall i, a(i) \leq c$ potom jeho EB sa nachádza medzi strednou hodnotou $m = E a(i)$ a horným ohraničením c , ktoré nazveme pík (*peak*) procesu. Pre stacionárny proces $A(t)$ s ohraničenými prírastkami sa dajú ľahko ukázať nasledujúce vzťahy:

$$EA(t) = \sum_{i=1}^t Ea(i) = t \cdot m, \quad A(t) \leq \sum_{i=1}^t a(i) \leq t \cdot c$$

Použijeme Jensenovú nerovnosť, ktorá tvrdí, že pre ľubovoľnú konvexnú funkciu $f(\cdot)$ a diskretnú náhodnú premennú X platí: $f(EX) \leq E[f(X)]$. Ak položíme $f(x) = e^{\theta x}$, dostávame vzťah $e^{\theta EA(t)} \leq E[e^{\theta A(t)}]$. Postupne môžeme písať nerovnosti:

$$\begin{aligned} e^{\theta EA(t)} &\leq E[e^{\theta A(t)}] \leq e^{\theta tc} \Rightarrow \theta t Ea(i) \leq \ln E[e^{\theta A(t)}] \leq \theta tc \Rightarrow \\ &\Rightarrow m \leq \frac{1}{\theta t} \ln E[e^{\theta A(t)}] \leq c \Rightarrow m \leq \alpha(\theta, t) \leq c \end{aligned}$$

Efektívna šírka pásma v závislosti od θ a t leží niekde v "páse" medzi strednou hodnotou prírastkov a maximálnou hodnotou (píkom) procesu.

Ohraničenie

EB

Nech $\alpha(\theta, t)$ je Efektívna šírka pásma nejakého stacionárneho procesu $A(t)$ s ohraničenými prírastkami $\exists c \in R, \forall i, a(i) \leq c$. Potom platí

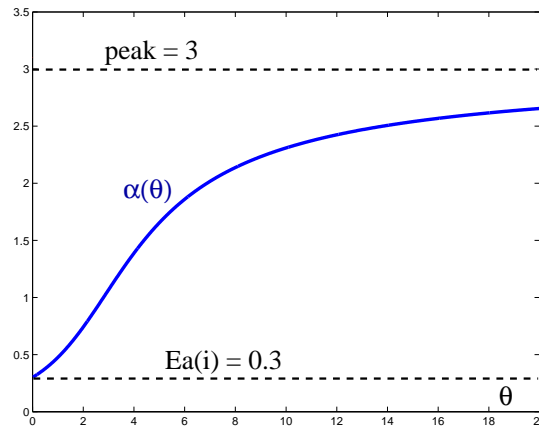
$$\forall \theta > 0, \forall t > 0, \quad m \leq \alpha(\theta, t) \leq c$$

PRÍKLAD 1.20

Nech v časovom slotu (napr. 1ms) sa môžu vyskytnúť najviac 3 pakety s danými pravdepodobnosťami:

$$p_0 = 0.729, \quad p_1 = 0.243, \quad p_2 = 0.027, \quad p_3 = 0.001$$

Stredný počet paketov v slotu: $Ea(i) = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 = 0.3$



Efektívna šírka pásma procesu je zdola ohraničená strednou hodnotou počtu paketov (0.3) a zhora maximálnym počtom paketov v časovom slotu (3).

Ak je proces stacionárny s nezávislými prírastkami (i.i.d.), tvar EB sa podstatne zjednoduší. Nech $\lambda_a(\theta)$ a $\varphi_a(\theta)$ sú CGF a MGF náhodných premenných $a(i)$.

Pre zjednodušenie označíme $a(i)$ premennými X . Pre EB platí:

$$\begin{aligned}\alpha(\theta, t) &= \frac{1}{\theta t} \ln E \left[e^{\theta \sum_{i=1}^t X} \right] = \frac{1}{\theta t} \ln E \left[\prod_{i=1}^t e^{\theta X} \right] = \frac{1}{\theta t} \ln \varphi_a(\theta)^t \Rightarrow \\ \Rightarrow \quad \alpha(\theta, t) &= \alpha(\theta) = \frac{1}{\theta} \ln \varphi_a(\theta) = \frac{\lambda_a(\theta)}{\theta}\end{aligned}\quad (1.12)$$

Pre proces s i.i.d. prírastkami je EB rovná vyškálovanej CGF. Budeme ju označovať $\alpha(\theta)$. Výpočet derivácii v $\theta = 0$ sa podstatne zjednoduší. Použijeme už odvodené vzťahy pre CGF: $\lambda'_a(0) = EX$, $\lambda''_a(0) = DX$. Pri výpočte použijeme L'Hospitalove pravidlo:

$$\begin{aligned}\alpha(0) &= \lim_{\theta \rightarrow 0} \frac{\lambda_a(\theta)}{\theta} = \lim_{\theta \rightarrow 0} \frac{\lambda'_a(\theta)}{1} = \lambda'_a(0) = EX \\ \alpha'(0) &= \left. \frac{\lambda'_a(\theta)\theta - \lambda_a(\theta)}{\theta^2} \right|_{\theta=0} = \lim_{\theta \rightarrow 0} \frac{\lambda''_a(\theta) \cdot \theta - \lambda'_a(\theta) + \lambda'_a(\theta)}{2\theta} = \frac{\lambda''_a(0)}{2} = \frac{DX}{2}\end{aligned}$$

1.5.2 EB pre Regulárny proces

Budeme sa zaoberať konkrétnymi procesmi, ktoré sa najčastejšie používajú na modelovanie datových prenosov na sieti. Samozrejme najjednoduchším modelom je **Regulárny proces** s konštantnou intenzitou toku dát r . Efektívna šírka pásma takéhoto toku je Samozrejme najjednoduchším modelom je **Regulárny proces** s konštantnou intenzitou toku dát r . Efektívna šírka pásma takéhoto toku je

$$\alpha(\theta, t) = \frac{1}{\theta t} \sup_{s \geq 0} \ln E \left[e^{\theta(r \cdot (s+t) - r \cdot (s))} \right] = \frac{1}{\theta t} \ln \left[e^{\theta r t} \right] = \frac{\theta r t}{\theta t} = r \quad (1.13)$$

Tento proces je síce jednoduchý, ale ani zďaleka neodpovedá reálnej situácii v IP sieti. Preto sa ďalej budeme zaoberať stochastickými procesmi, ktoré poskytujú široké možnosti modelovania reálnych tokov. Obmedzíme sa na stacionárne toky $A(t)$ s nezávislými prírastkami $a(i)$, $A(t) = \sum_{i=1}^t a(i)$, ktoré považujeme pre reálne použitie za postačujúce. Efektívnu šírku pásma pre takéto procesy budeme označovať zjednodušene $\alpha(\theta)$.

1.5.3 EB pre Bernoulliho proces

Efektívna šírka pásma pre Bernoulliho proces $Bi(nt, p)$ má tvar:

$$\alpha(\theta) = \frac{\lambda_a(\theta)}{\theta} = \frac{n}{\theta} \ln \left[1 - p + pe^\theta \right] \quad (1.14)$$

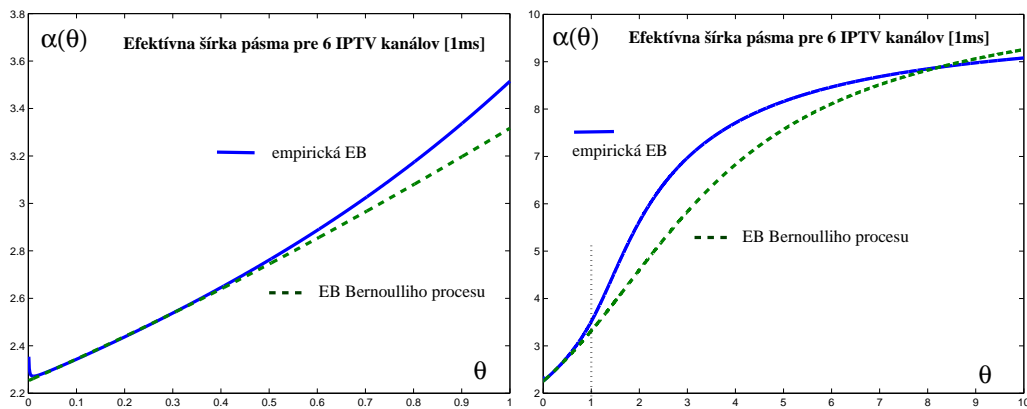
Určíme hodnotu šírky pásma v nule (L'Hospital):

$$\alpha(0) = \lim_{\theta \rightarrow 0} \frac{n \cdot \ln(1 - p + p \cdot e^\theta)}{\theta} = \lim_{\theta \rightarrow 0} \frac{np \cdot e^\theta}{1 - p + p \cdot e^\theta} = np$$

Efektívna šírka pásma Bernoulliho procesu $Bi(nt, p)$ je zhora ohraničená. V jednom časovom slot, ktorý predstavuje súčet n elementárnych slotov, sa môže vyskytnúť najviac n paketov (s pravdepodobnosťou p^n). Preto pre prírastky procesu platí, že sú zhora ohraničené hodnotou n :

$$\forall i; a(i) \leq n \Rightarrow \alpha(\theta) = \frac{1}{\theta t} \ln E \left[e^{\theta A(t)} \right] \leq \frac{1}{\theta t} \ln E \left[e^{\theta nt} \right] = \frac{\ln [e^{\theta nt}]}{\theta t} = n$$

Vrátíme sa k príkladu modelovania toku dát 6 IPTV kanálov, pričom sa budeme zaoberať len prípadom, keď je časový slot 1ms. Na nasledujúcich obrázkoch je porovnanie priebehu empiricky získanej efektívnej šírky pásma priamo z nameraných dát, a jej aproximácie pomocou šírky pásma Bernoulliho procesu $A(t) \sim Bi(11t, 0.2048)$.



Obr. 1.29: porovnanie EB pre IPTV a Bernoulliho proces $Bi(11t, 0.2048)$

Na prvom obrázku je porovnanie EB pre tok 6 IPTV kanálov a Bernoulliho proces v rozpätí $\theta \in \langle 0, 1 \rangle$, na druhom $\theta \in \langle 0, 10 \rangle$. Pre potreby dimenzovania kapacity linky (vid. nasledujúca kapitola) nám väčšinou stačí rozsah priestorového parametra $\theta \in \langle 0, 1 \rangle$. Na tomto intervale môžeme považovať obidve šírky pásma za zhodné.

1.5.4 EB pre Poissonov proces

Efektívna šírka pásma pre Poissonov proces $A(t) \sim Po(\lambda t)$ má tvar:

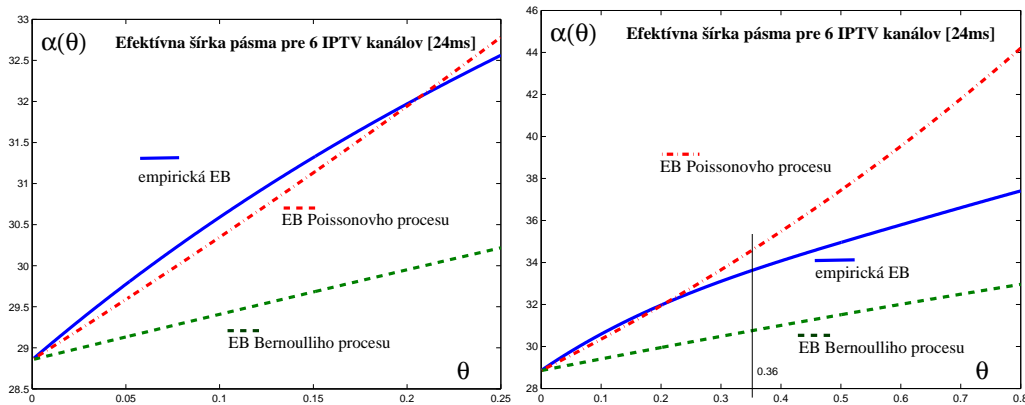
$$\alpha(\theta) = \frac{\lambda_a(\theta)}{\theta} = \lambda \frac{e^\theta - 1}{\theta} \quad (1.15)$$

Určíme hodnotu EB v bode $\theta = 0$ pomocou L'Hospitalového pravidla:

$$\alpha(0) = \lim_{\theta \rightarrow 0} \lambda \frac{e^\theta - 1}{\theta} = \lim_{\theta \rightarrow 0} \lambda \frac{e^\theta}{1} = \lambda = Ea(i)$$

Prírastky procesu nie sú ohraničené, a preto neexistuje peak procesu.

V predchádzajúcej kapitole sme videli, že pri aproximácii rozdelenia pravdepodobnosti výskytu paketov v toku 6 kanálov IPTV dosahuje Poissonov proces podstatne lepšie výsledky než proces Bernoulliho. Táto skutočnosť sa samozrejme prejaví aj na tvare odpovedajúcich efektívnych šíriek pásma:



Obr. 1.30: Efektívna šírka pásma pre tok 6 kanálov IPTV

Podľa priebehu grafov EB vidíme, že aproximáciu pomocou Poissonovho procesu môžeme použiť až po veľkosť rozmerového parametra $\theta = 0.36$ (odchýlka od empirickej EB je menej než jeden paket).

1.5.5 EB pre Markovov modulovaný proces (MMP)

Podstatne komplikovanejší popis IP tokov predstavuje Markovov modulovaný proces (MMP). V predchádzajúcich prípadoch sme uvažovali procesy, pri ktorých sa nemenila intenzita toku. V MMP procese pripúšťame zmenu intenzít, a vo všeobecnosti aj zmenu pravdepodobnostného charakteru toku, podľa istých pravidiel. Tieto pravidla predstavuje Markovov reťazec X_t . Podľa realizácii Markovovho reťazca sa MMP proces $A(t)$ "prepína" medzi rôznymi i.i.d. procesmi. Pre každý stav i Markovovho reťazca je definovaný i.i.d. proces, ktorý je popísaný momentovou vytvárajúcou funkciou $\varphi_i(\theta)$:

$$\varphi_i(\theta) = E \left[e^{\theta a(\tau)} | X_t = i \right] \quad (1.16)$$

Keďže proces $A(t)$ je stacionárny, jeho prírastky sme označili $a(\tau)$, pričom τ predstavuje ľubovoľný čas (resp. index).

V praxi sú najčastejšie používané 2-stavové MMP procesy, pretože pre tieto procesy existuje analyticky vyjadrenie Efektívnej šírky pásma. Proces sa môže prepínať medzi:

- dvoma regulárnymi procesmi - popis pre rôzne prevádzky (MMRP)
- dvoma Poissonovými procesmi s rozdielnymi intenzitami - Markovov modulovaný Poissonov proces (MMPP)
- dvoma Bernoulliho procesmi s rozdielnymi intenzitami - Markovov modulovaný Bernoulliho proces (MMBP)
- procesom vysielaným v stave On a stavom Off, v ktorom je nulová intenzita (On-Off zdroje)

Vo všeobecnosti môžu byť pre stavy Markovovho reťazca X definované ľubovoľné procesy s i.i.d. prírastkami.

1.5.6 EB pre MMP s diskretným časom

Budeme uvažovať o Markovovom reťazci X_l s diskretným časom, s konečnou množinou stavov $S = \{1, 2, \dots, m\}$ a maticou pravdepodobností prechodov medzi stavmi $\mathbf{P} = \{p_{i,j}\}$. Jeho stacionárne rozdelenie označíme $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_m)$. Našou snahou bude určiť MGF $\varphi(\theta, t)$ procesu $A(t)$. Na výpočet použijeme tzv. "backward" rovnice:

$$\varphi(\theta, t) = E \left[e^{\theta A(t)} \right] = \sum_{i=1}^m E \left[e^{\theta A(t)} | X_l = i \right] \cdot P(X_l = i)$$

Ak označíme vektor $\boldsymbol{\psi}(\theta, t) = (E[e^{\theta A(t)} | X_l = 1], \dots, E[e^{\theta A(t)} | X_l = m])$, dostávame maticový zápis:

$$\varphi(\theta, t) = \boldsymbol{\pi} \boldsymbol{\psi}(\theta, t)^T \quad (1.17)$$

Určíme podmienené stredné hodnoty:

$$\begin{aligned} E[e^{\theta A(t)} | X_l = i] &= \\ &= E[e^{\theta a(1)} | X_l = i] \cdot E[e^{\theta(A(t)-A(1))} | X_l = i] = \\ &= \varphi_i(\theta) \sum_{j=1}^m E[e^{\theta(A(t)-A(1))} | X_{l+1} = j, X_l = i] \cdot P(X_{l+1} = j | X_l = i) = \\ &= \varphi_i(\theta) \sum_{j=1}^m E[e^{\theta(A(t)-A(1))} | X_{l+1} = j] \cdot p_{i,j} = \\ &= \varphi_i(\theta) \sum_{j=1}^m E[e^{\theta(A(t-1))} | X_l = j] \cdot p_{i,j} \end{aligned}$$

V prvom kroku sme použili vzťah $a(1) = A(1)$ resp. $a(1) = A(t) - A(1)$, v druhom kroku Vetu o úplnej pravdepodobnosti, v tretom kroku Markovovu vlastnosť a vo štvrtom kroku fakt, že proces $A(t)$ je stacionárny, preto platí $A(t) - A(1) = A(t-1)$, a Markovovu vetu, preto sa zhoduje podmienka $X_{l+1} = j$ s podmienkou $X_l = j$.

Nech $\boldsymbol{\phi}(\theta)$ je diagonálna matica MGF funkcií, $\text{diag}(\varphi_1(\theta), \dots, \varphi_m(\theta))$. Uvažovaný vzťah pre podmienenú strednú hodnotu prepíšeme do maticového tvaru:

$$\boldsymbol{\psi}(\theta, t)^T = \boldsymbol{\phi}(\theta) \mathbf{P} \boldsymbol{\psi}(\theta, t-1)^T = \dots = (\boldsymbol{\phi}(\theta) \mathbf{P})^{t-1} \boldsymbol{\psi}(\theta, 1)^T \quad (1.18)$$

Nech $\mathbf{1}^T$ je jednotkový vektor $(1, \dots, 1)^T$. Dostávame počiatočnú podmienku $\boldsymbol{\psi}(\theta, 1)^T = \boldsymbol{\phi}(\theta) \mathbf{1}^T$ Potom vzťah (1.17) má tvar:

$$\varphi(\theta, t) = E[e^{\theta A(t)}] = \boldsymbol{\pi} (\boldsymbol{\phi}(\theta) \mathbf{P})^{t-1} \boldsymbol{\phi}(\theta) \mathbf{1}^T \quad (1.19)$$

Z Perron-Frobeniusovej vety vyplýva (pozri Chang [2]), že momentová vytvárajúca funkcia $\varphi(\theta, t)$ je rovná spektrálnemu rádiusu $sp(\cdot)$ (maximálne vlastné číslo) matice $\boldsymbol{\phi}(\theta) \mathbf{P}$. EB pre MMP proces $A(t)$ má tvar

$$\alpha(\theta) = \frac{\ln sp(\boldsymbol{\phi}(\theta) \mathbf{P})}{\theta} \quad (1.20)$$

Uvažujme o 2-stavovom MMP. Pre takýto proces vieme odvodiť analytický tvar EB. Nech Markovov reťazec X_l je daný maticou pravdepodobností prechodov

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \text{ a procesy sú popísané maticou } \phi(\theta) = \begin{pmatrix} \varphi_1(\theta) & 0 \\ 0 & \varphi_2(\theta) \end{pmatrix}$$

Na určenie EB musíme vypočítať najväčšie vlastné číslo matice $\phi(\theta)\mathbf{P}$:

$$\det |\phi(\theta)\mathbf{P} - x\mathbf{E}| = 0 \Rightarrow \begin{vmatrix} \varphi_1(\theta)p_{11} - x & \varphi_1(\theta)p_{12} \\ \varphi_2(\theta)p_{21} & \varphi_2(\theta)p_{22} - x \end{vmatrix} = 0 \Rightarrow$$

$$\Rightarrow x^2 - x[\varphi_1(\theta)p_{11} + \varphi_2(\theta)p_{22}] + \varphi_1(\theta)\varphi_2(\theta)[p_{11}p_{22} - p_{21}p_{12}] = 0$$

Najväčší koreň polynómu je spektrálny rádius $sp(\phi(\theta)\mathbf{P})$:

$$\frac{\varphi_1(\theta)p_{11} + \varphi_2(\theta)p_{22} + \sqrt{[\varphi_1(\theta)p_{11} + \varphi_2(\theta)p_{22}]^2 - 4\varphi_1(\theta)\varphi_2(\theta)[p_{11}p_{22} - p_{21}p_{12}]}}{2}$$

Po úpravách dostávame EB pre 2-stavový MMP $\alpha(\theta) =$

$$\frac{1}{\theta} \ln \left[\frac{\varphi_1(\theta)p_{11} + \varphi_2(\theta)p_{22} + \sqrt{[\varphi_1(\theta)p_{11} - \varphi_2(\theta)p_{22}]^2 + 4\varphi_1(\theta)\varphi_2(\theta)p_{21}p_{12}}}{2} \right] \quad (1.21)$$

MMRP On/Off zdroj

Pre Markovov modulovaný Regulárny proces sú prírastky v stave On rovné 1, $a_1(\tau) = 1$, a prírastky v stave Off sú nulové $a_2(\tau) = 0$. Jednotlivým procesom odpovedajú MGF funkcie:

$$\varphi_1(\theta) = E[e^{\theta a_1(\tau)}] = e^\theta \quad \varphi_2(\theta) = E[e^{\theta a_2(\tau)}] = e^{0\theta} = 1$$

Efektívna šírka pásma pre MMRP On/Off zdroj má tvar:

$$\alpha(\theta) = \frac{1}{\theta} \ln \left[\frac{e^\theta p_{11} + p_{22} + \sqrt{[e^\theta p_{11} - p_{22}]^2 + 4e^\theta p_{21}p_{12}}}{2} \right] \quad (1.22)$$

MMBP On/Off zdroj

Markovov modulovaný Bernoulliho proces má v stave On rovné Bernoulliho prírastky $a_1(\tau)$ s parametrom p a v stave Off sú nulové, $a_2(\tau) = 0$. Jednotlivým procesom odpovedajú MGF funkcie:

$$\varphi_1(\theta) = E[e^{\theta a_1(\tau)}] = q + pe^\theta \quad \varphi_2(\theta) = E[e^{\theta a_2(\tau)}] = e^{0\theta} = 1$$

Efektívna šírka pásma pre MMBP On/Off zdroj má tvar:

$$\alpha(\theta) = \frac{1}{\theta} \ln \left[\frac{(q+pe^\theta)p_{11} + p_{22} + \sqrt{[(q+pe^\theta)p_{11} - p_{22}]^2 + 4(q+pe^\theta)p_{21}p_{12}}}{2} \right] \quad (1.23)$$

1.5.7 EB pre MMP so spojitým časom

Odvodenie tvaru EB pre tok so spojitým časom je podstatne komplikovanejší, preto si uvedieme len základné myšlienky a výsledný tvar Efektívnej šírky pásma pre tento typ procesov.

Nech proces X_t je Markovov reťazec so spojitým časom, s konečnou množinou stavov $S = \{1, 2, \dots, m\}$ a maticou intenzít prechodov medzi stavmi $\mathbf{Q} = \{q_{i,j}\}$. Stacionárne rozdelenie reťazca označíme $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_m)$. Momentovú vytvárajúcu funkciu procesu $A(t)$ označíme $\varphi(\theta, t)$:

$$\varphi(\theta, t) = E \left[e^{\theta A(t)} \right] = \sum_{i=1}^m E \left[e^{\theta A(t)} | X_t = i \right] P(X_t = i) = \sum_{i=1}^m \psi_i(\theta, t) \pi_i$$

MMRP

Pre každý stav j Markovovho reťazca máme definovaný regulárny proces s intenzitou λ_j a označíme maticu $\mathbf{\Lambda}(\theta)$ diagonálnu maticu kumulatívnych vytvárajúcich funkcií $\text{diag}(\lambda_1\theta, \dots, \lambda_m\theta)$. Na odvodenie tvaru MGF sa postupne použijú Taylorove rozvoje a "backward" rovnice:

$$\varphi(\theta, t) = E \left[e^{\theta A(t)} \right] = \boldsymbol{\pi} \boldsymbol{\psi}(\theta, t)^T = \boldsymbol{\pi} e^{(\mathbf{Q} + \mathbf{\Lambda}(\theta))t} \mathbf{1}^T \quad (1.24)$$

Z Perron-Frobeniovej vety vyplýva ([2]), že momentová vytvárajúca funkcia $\varphi(\theta, t)$ je rovná spektrálnemu rádiusu $sp(\cdot)$ (maximálne vlastné číslo) matice $e^{\mathbf{Q} + \mathbf{\Lambda}(\theta)}$. Pomocou vlastností spektrálneho rádiusu, $sp(\exp(\mathbf{F})) = \exp(sp(\mathbf{F}))$, dostávame Efektívnu šírku pásma pre MMP proces $A(t)$ so spojitým časom:

$$\alpha(\theta) = \frac{\ln sp(e^{\mathbf{Q} + \mathbf{\Lambda}(\theta)})}{\theta} \Rightarrow \alpha(\theta) = \frac{sp(\mathbf{Q} + \mathbf{\Lambda}(\theta))}{\theta} \quad (1.25)$$

MMRP On/Off zdroj

Nech Markovov reťazec X_t je 2-stavový s intenzitami prechodov medzi stavmi, $q_{12} = \alpha$ a $q_{21} = \beta$. Matica intenzít má tvar $\mathbf{Q} = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$. Pre On/Off zdroj má diagonálna matica tvar $\mathbf{\Lambda}(\theta) = \text{diag}(\theta\lambda, 0)$. Vypočítame najväčšie vlastné číslo matice $\mathbf{Q} + \mathbf{\Lambda}(\theta)$:

$$\begin{aligned} \det |\mathbf{Q} + \mathbf{\Lambda}(\theta) - x\mathbf{E}| = 0 &\Rightarrow \begin{vmatrix} \theta\lambda - \alpha - x & \alpha \\ \beta & -\beta - x \end{vmatrix} = 0 \Rightarrow \\ &\Rightarrow x^2 + x[\alpha + \beta - \theta\lambda] - \theta\lambda\beta = 0 \end{aligned}$$

Vypočítame najväčší koreň polynómu a dostaneme Efektívnu šírku pásma pre 2-stavový MMRP On/Off zdroj:

$$\alpha(\theta) = \frac{\theta\lambda - \alpha - \beta + \sqrt{[\alpha + \beta - \theta\lambda]^2 + 4\theta\lambda\beta}}{2\theta} \quad (1.26)$$

MMPP On/Off zdroj

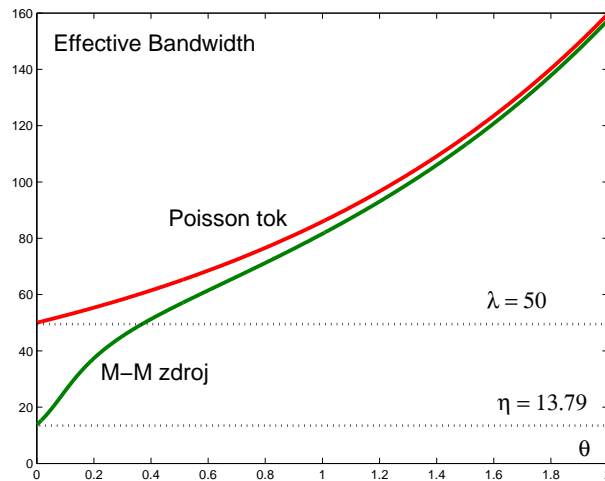
Budeme sa zaoberať 2-stavovým Markovovým modulovaným Poissonovým procesom, ktorý v stave On generuje Poissonov tok s intenzitou λ a v stave Off sú prírastky toku nulové. Diagonálna matica kumulatívnych generujúcich funkcií má tvar $\mathbf{\Lambda} = \text{diag}(\lambda(e^\theta - 1), 0)$. Vypočítame najväčšie vlastné číslo matice $\mathbf{Q} + \mathbf{\Lambda}(\theta)$:

$$\det |\mathbf{Q} + \mathbf{\Lambda}(\theta) - x\mathbf{E}| = 0 \Rightarrow \begin{vmatrix} \lambda(e^\theta - 1) - \alpha - x & \alpha \\ \beta & -\beta - x \end{vmatrix} = 0 \Rightarrow$$

$$\Rightarrow x^2 + x[\alpha + \beta - \lambda(e^\theta - 1)] - \lambda(e^\theta - 1)\beta = 0$$

Efektívna šírka pásma pre MMPP On/Off zdroj má tvar:

$$\alpha_1(\theta, t) = \frac{\lambda(e^\theta - 1) - \alpha - \beta + \sqrt{(\alpha + \beta - \lambda(e^\theta - 1))^2 + 4\lambda(e^\theta - 1)\beta}}{2\theta}$$



Obr. 1.31: EB pre Poissonov tok a MMPP s $\lambda = 50$ p/s

Na obrázku vidíme, že pri zväčšujúcom sa parametre θ hodnota EB pre MMPP zdroj konverguje zdola k EB pre Poissonov zdroj. Je to prirodzené, keby boli intenzity α a β nulové, M-M zdroj by sa zmenil na Poissonov zdroj:

$$\alpha = 0, \beta = 0 \Rightarrow \alpha_1(\theta, t) = \frac{\lambda(e^\theta - 1) + \sqrt{(-\lambda(e^\theta - 1))^2}}{2\theta} = \frac{\lambda(e^\theta - 1)}{\theta}$$

2

Modelovanie spoločného komunikačného prostredia

OBSAH KAPITOLY

2.1	Charakteristiky spoločného komunikačného prostredia	55
2.1.1	Model prostredia so spätnou väzbou	56
2.1.2	Rozdelenie sietí podľa typu riadenia a prenosu	57
2.2	Asynchrónne prostredie s úplnou informáciou	58
2.2.1	Model s diskretným časom $Geo/D/1/\infty$	59
2.2.2	Model s diskretným časom $Geo/Geo/1/\infty$	63
2.2.3	Markovove reťazce so spojitým časom	67
2.2.4	Zovšeobecnený proces vzniku a zániku	72
2.2.5	Analýza Markovovho modelu	74
2.2.6	Markovov model	78
2.2.7	Nemarkovove modely	78
2.3	Použitie princípu veľkých odchýlok	93
2.3.1	Systém s konštantnou kapacitou	93
2.3.2	Dimenzovanie $Geo/D/1/\infty$	96
2.3.3	$M/D/1/\infty$	96
2.3.4	Systém s variabilnou kapacitou	98
2.3.5	$M/M/1/\infty$	100
2.3.6	Markovove systémy s MMPP vstupom	103
2.3.7	$M-M/M/1/\infty$	104
2.4	Sieťový kalkul	106
2.4.1	Dimenzovanie pomocou (σ, ρ) - Algebry	106

2.1 CHARAKTERISTIKY SPOLOČNÉHO KOMUNIKAČNÉHO PROSTREDIA

Dnes už klasickým prístupom (od čias Erlanga) k modelovaniu priepustnosti komunikačných systémov je použitie metód prevádzkového inžinierstva založených na teórii hromadnej obsluhy. Aby sme videli ako množstvo informácie o stave komunikačného prostredia (aj keď nekvanifikované) zvyšuje jeho priepustnosť, predstavíme na základe známych výsledkov z literatúry typické modelové situácie:

- sieť Aloha - komunikačné prostredie bez informácie o stave prostredia
- sieť Ethernet - komunikačné prostredie s čiastočnou informáciou o stave prostredia
- systém M/M/1 - komunikačné prostredie s úplnou informáciou o stave prostredia

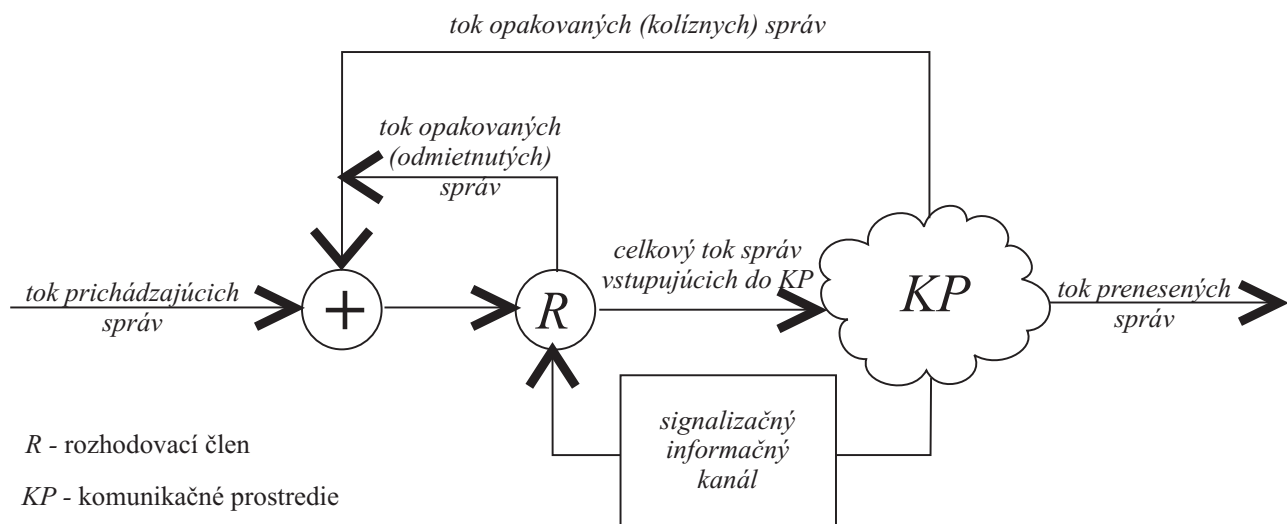
V literatúre sa bežne uvádza model siete Aloha pre pevnú dĺžku rámca [?], [?], tak ako to odpovedá použitej technológii. Pevná dĺžka rámca však predsa dáva akúsi informáciu o stave prostredia, pretože zvyškový čas vysielania rámca má iné rozdelenie ako rozdelenie dĺžky rámca. Spojitým rozdelením, pri ktorom je zvyškový čas vysielania rámca rovnako rozdelený ako dĺžka rámca, je exponenciálne rozdelenie. To teda predstavuje najviac neurčitú situáciu, preto sme dopočítali aj model siete Aloha s exponenciálnym rozdelením dĺžky rámca.

2.1.1 Model prostredia so spätnou väzbou

Komunikačné prostredia (ďalej už len KP) môžeme rozdeliť z pohľadu zdieľania kapacity na prostredia delené a spoločné. Naším záujmom bude štúdium spoločného KP, preto v tejto kapitole zavedieme všeobecný model takéhoto prostredia so spätným kanálom.

Poznámka 2.1.1. *Pojem spätný kanál budeme označovať aj ako signalizačný informačný kanál, keďže ide o kanál, ktorý signalizuje stav prostredia, a budeme používať skratku SIK.*

Každý užívateľ (zdroj správ) využívajúci spoločné KP vytvára tok správ, ktoré postupne prichádzajú do prostredia. Prostredie nemá žiadnu znalosť o tom, kedy ktorý užívateľ potrebuje vstúpiť do prostredia, zároveň užívatelia nevedia o správach od iných užívateľov. V zjednodušenom pohľade, všetky toky z jednotlivých zdrojov tvoria jeden tok správ, ktorý má byť obslužený v komunikačnom prostredí (pozri obrázok 2.1). Spätný kanál pre prichádzajúce požiadavky o

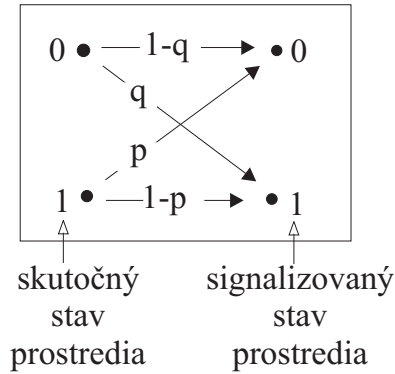


Obr. 2.1: Model spoločného komunikačného prostredia (KP) so spätným kanálom (SIK)

prenos signalizuje stav prostredia: voľné "0" alebo obsadené "1". Rozhodovací člen, podľa signalizovaného stavu prostredia, pustí správu do prostredia alebo ju odmietne. Pokiaľ spätný kanál signalizuje stav voľného prostredia, rozhodovací člen pustí správu do prostredia a zahájí sa jej vysielanie. Pokiaľ kanál signalizuje, že prostredie je v stave obsadené "1", rozhodovací

člen odmietne správu a jej prenos sa zopakuje po náhodnom časovom intervale. Tým vzniká tok odmietnutých správ, ktoré sa opätovne budú pokúšať o vstup do KP.

Obrázok 2.2 ilustruje schému spätného kanála s vyznačenými podmienenými pravdepodobnosťami v kanáli. Podmienená pravdepodobnosť $P(x/y)$ vyjadruje pravdepodobnosť, že kanál



Obr. 2.2: Schéma spätného kanála (SIK)

signalizuje stav prostredia x , za podmienky, že skutočný stav prostredia je y . V prípade, že aspoň jedna z pravdepodobností $p, q \neq 0$, správy nie sú informované o stave prostredia vždy pravdivo. Potom správy, ktoré sú na základe informácie zo spätného kanála pustené do komunikačného prostredia, môžu následne spôsobiť kolíziu. Ďalej budeme predpokladať, že prenos takýchto kolidovaných správ sa po náhodnom čase bude opakovať. Tým vzniká tok opakovaných správ v dôsledku kolízie.

V nasledujúcej kapitole popíšeme z hľadiska priepustnosti spoločné KP bez informácie, t.j. prostredie, v ktorom prichádzajúce správy nemajú žiadnu informáciu o stave KP. V súvislosti so všeobecným modelom na obrázku 2.1 a 2.2, je to prostredie, v ktorom spätný kanál signalizuje pre prichádzajúce požiadavky vždy stav voľného prostredia, $p = 1, q = 0$.

Ďalej preskúmame spoločné KP s čiastočnou informáciou, v ktorom prichádzajúce požiadavky majú len čiastočnú informáciu o stave KP. V súvislosti so všeobecným modelom na obrázku 2.1 a 2.2, pre toto KP platí, že aspoň jedna z pravdepodobností $p, q \in (0, 1)$.

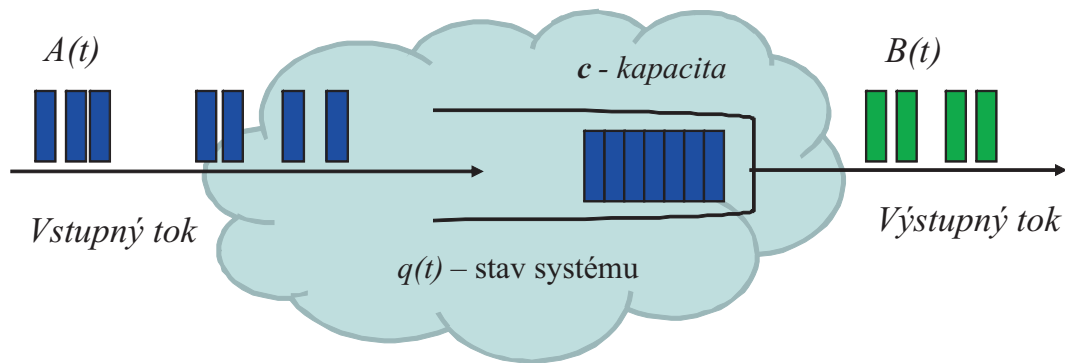
Ak $p = q = 0$, budeme hovoriť o spoločnom KP s úplnou informáciou.

Pojem priepustnosti je používaný v zmysle maximálnej možnej hodnoty zaťaženia prostredia prichádzajúcimi paketmi (čerstvou prevádzkou) tak, že systém zostáva stabilný (pre vyššie hodnoty sa dostáva do nestabilného stavu).

2.1.2 Rozdelenie sietí podľa typu riadenia a prenosu

2.2 ASYNCHRÓNNE PROSTREDIE S ÚPLNOU INFORMÁCIOU

Budeme sa zaoberať spoločným komunikačným prostredím s diskretným časom, do ktorého vstupuje tok požiadaviek $A(t)$ s identickými a nezávislými prírastkami i.i.d. $a(i)$. Prostredie, resp. systém je schopný v priebehu daného časového slotu odoslať práve c požiadaviek. Hovoríme tiež že kapacita systému je c . Predpokladáme, že počet čakajúcich požiadaviek v prostredí je neohraničený. Takéto prostredie nazveme systém hromadnej obsluhy (SHO) a označíme podľa tzv. Kendallovej klasifikácie ako $(GI/D/1/\infty)$. Označenie informuje, že vstupný tok má i.i.d. prírastky s ľubovoľným rozdelením (*General Independent*), spracovanie požiadaviek je konštantné resp. deterministické (D), systém má jeden kanál, resp. ide o spoločne zdieľané prostredie pre všetky požiadavky a front v systéme nie je ohraničený. Predpokladáme, že správa frontu je FIFO, a prvá požiadavka v poradí je práve spracovávaná.



Obr. 2.3: Spoločne zdieľané komunikačné prostredie $GI/D/1/\infty$

Nech náhodný proces $q(t)$ popisuje počet požiadaviek v prostredí. Množina stavov procesiu je $S = \{0, 1, 2, \dots\}$. Obsluha požiadaviek sa v čase nemení a vstup požiadaviek tvorí proces s nezávislými a identickými prírastkami. Preto proces $q(t)$ je Markovov reťazec s diskretným časom. Na jeho popis potrebujeme poznať maticu pravdepodobností prechodov medzi stavmi, $\mathbf{P} = \{p_{ij}\}$.

V prípade, že sa zaoberáme iba simuláciou systému, na zistenie jeho charakteristík nám postačí tzv. Lindleyho rovnica:

Lindleyho rovnica

Nech proces $q(t)$ popisuje počet požiadaviek v systéme v čase t . Nech $A(t)$ predstavuje vstupný porces do systému s prírastkami $a(i)$, $A(t) = \sum_i a(i)$. Nech systém v priebehu jedného časového slotu dokáže spracovať práve c požiadaviek. Potom proces $q(t)$ spĺňa Lindleyho rovnicu:

$$q(t) = \max\{q(t-1) + a(t) - c, 0\} = (q(t-1) + a(t) - c)^+$$

V prípade reálneho použitia Lindleyho rovnice je dôležité, aby časový slot bol

zvolený čo najmenši, resp. maximálne rovný konštantnej dobe spracovania požiadavky. V opačnom prípade dôjde ku zkreseniu charakteristík systému.

2.2.1 Model s diskretným časom $Geo/D/1/\infty$

Nech do prostredia vstupuje Bernoulliho tok požiadaviek $A(2)$ s parametrami p a $q = 1 - p$. Nech systém dokáže odoslať práve jednu požiadavku v priebehu jednej milisekundy, kapacita systému je $c = 1p/ms$. Predpokladáme, že počet čakajúcich požiadaviek v prostredí je neohraničený. Takéto prostredie nazveme systém hromadnej obsluhy (SHO) a označíme podľa tzv. Kendallovej klasifikácie ako $(Geo/D/1/\infty)$. Označenie informuje, že vstupný tok je Bernoulliho, resp. má geometrické medzery medzi vstupmi (Geo).

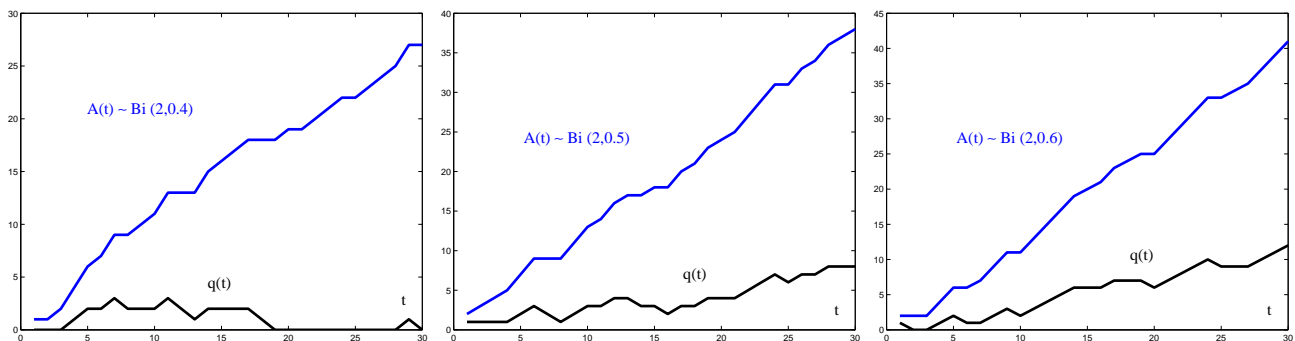
PRÍKLAD 2.1

Pomocou Lindleyho rovnice vytvorte simuláciu systému $(Geo/D/1/\infty)$ so vstupným tokom $A(t) \sim Bi(2, 0.4)$ a s kapacitou $c = 1p/ms$. Určite maximálne oneskorenie paketov v systéme.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$a(t)$	2	2	1	1	0	1	1	0	0	1	2	1	1	0	2	0
$q(t)$	1	2	2	2	1	1	1	0	0	0	1	1	1	0	1	0

Maximálny počet požiadaviek v komunikačnom prostredí v našej simulácii je $q_{\max} = 2$, preto pri správe frontu FIFO by maximálne oneskorenie systému bolo $d_{\max} = q_{\max}/c = 2ms$. Vzhľadom na krátkosť simulácie však túto hodnotu nemôžeme považovať za relevantnú.

Na nasledujúcich obrázkoch je zobrazená dlhšia simuláciu systému, $t = 30ms$. Postupne sme menili hodnotu parametra p . Na prvom obrázku pre $p = 0.4$ vidíme, že $q_{\max} = 3$ a preto aj $d_{\max} = 3ms$. Na ostatných obrázkoch sa dá rozpoznať, že s rastúcou hodnotou parametra p sa fronta v systéme viac plní. Aby sme sa dozvedeli viac o systéme, musíme previesť kompletnú pravdepodobnostnú analýzu.



Obr. 2.4: Systém $Geo/D/1/\infty$ s $c = 1p/ms$ a s $A(t) \sim Bi(2, 0.4), Bi(2, 0.5), Bi(2, 0.6)$

Na úplný popis Markovovho reťazca potrebujeme poznať maticu pravdepodobností prechodov medzi stavmi reťazca, $\mathbf{P} = \{p_{ij}\}$.

Nech časový slot je rovný jednej milisekunde. Nech je systém v čase 0 prázdny, $q(0) = 0$. Aby počas jedného časového slotu systém nezmenil stav, na vstupe sa

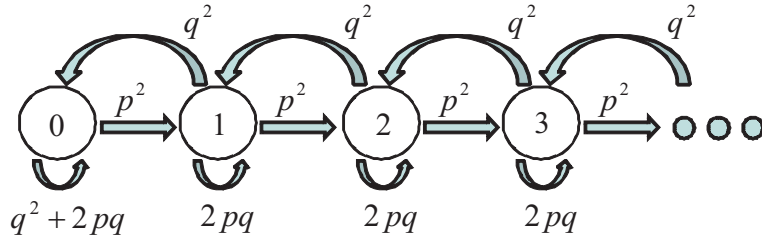
môže vyskytnúť najviac jedna požiadavka, v ktorú za daný slot systém stihne spracovať. Ak sa vyskytnú dve požiadavky, systém sa v nasledujúcom časovom slotе bude nachádzať v stave 1, $q(1) = 1$. Preto pre pravdepodobnosti prechodov zo stavu 0 platí

$$p_{0,0} = q^2 + 2pq, \quad p_{0,1} = p^2$$

Nech sa v systéme nachádza k požiadaviek na spracovanie. Ak sa na vstupe nevyskytne žiadna požiadavka, v ďalšom slotе bude v systéme len $k - 1$ požiadaviek, ak sa vyskytne práve jedna, stav systému ostane nezmenený, ak sa vyskytnú dve, v ďalšom slotе bude čakať na spracovanie $k + 1$ požiadaviek. Pre pravdepodobnosti prechodov platí

$$p_{k,k-1} = q^2, \quad p_{k,k} = 2pq, \quad p_{k,k+1} = p^2$$

Zostrojíme prechodový graf reťazca:



Obr. 2.5: Prechodový graf systému $Geo/D/1/\infty$

Matica pravdepodobností prechodov Markovoho reťazca má tvar:

$$\mathbf{P} = \begin{pmatrix} q^2 + 2pq & p^2 & 0 & 0 & \dots \\ q^2 & 2pq & p^2 & 0 & \dots \\ 0 & q^2 & 2pq & p^2 & \dots \\ 0 & 0 & q^2 & 2pq & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Stacionárne rozdelenie reťazca získame zo vzťahu (1.3), $[\mathbf{E} - \mathbf{P}]^T \boldsymbol{\pi}^T = \mathbf{0}$.

$$[\mathbf{E} - \mathbf{P}]^T = \begin{pmatrix} 1 - q^2 - 2pq & -q^2 & 0 & 0 & \dots \\ -p^2 & 1 - 2pq & -q^2 & 0 & \dots \\ 0 & -p^2 & 1 - 2pq & -q^2 & \dots \\ 0 & 0 & -p^2 & 1 - 2pq & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Ak využijeme vzťah $p^2 + 2pq + q^2 = 1$, maticu koeficientov nekonečného systému lineárnych rovníc môžeme prepísať do tvaru

$$\begin{pmatrix} p^2 & -q^2 & 0 & 0 & \dots \\ -p^2 & p^2 + q^2 & -q^2 & 0 & \dots \\ 0 & -p^2 & p^2 + q^2 & -q^2 & \dots \\ 0 & 0 & -p^2 & p^2 + q^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \sim \begin{pmatrix} p^2 & -q^2 & 0 & 0 & \dots \\ 0 & p^2 & -q^2 & 0 & \dots \\ 0 & 0 & p^2 & -q^2 & \dots \\ 0 & 0 & 0 & p^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Použili sme riadkovo ekvivaltnú úpravu, sčítali sme vždy dva za sebou idúce riadky. Systém rovníc (1.3) sa zmenil na jednoduché rekurentné vzťahy:

$$[\mathbf{E} - \mathbf{P}]^T \boldsymbol{\pi}^T = \mathbf{0} \Rightarrow p^2 \pi_{k-1} - q^2 \pi_k = 0 \Rightarrow \pi_k = \frac{p^2}{q^2} \pi_{k-1}, \quad k = 1, 2, \dots$$

Ak označíme podiel štvorcov pravdepodobností ako $\varrho = p^2/q^2$, môžeme vyjadriť stacionárne rozdelenie reťazca pomocou pravdepodobnosti stavu 0, $\pi_k = \varrho^k \pi_0$. Pravdepodobnosť π_0 získame z normovacej podmienky

$$\sum_{k=0}^{\infty} \pi_k = 1 \Rightarrow \sum_{k=0}^{\infty} \varrho^k \pi_0 = 1 \Rightarrow \frac{1}{1 - \varrho} \pi_0 = 1 \Rightarrow \pi_0 = 1 - \varrho$$

Nekonečný rad konverguje len vďaka podmienke $\varrho < 1$. To ale znamená:

$$\varrho < 1 \Rightarrow \frac{p^2}{q^2} < 1 \Rightarrow p < q \Rightarrow p < 1 - p \Rightarrow p < 0.5$$

Intenzita Bernoulliho toku je np , v našom prípade $2p$. Kapacita zdieľaného prostredia je $c = 1$. Preto aby sa systém nezahltl, intenzita vstupného toku musí byť menšia ako 1, $2p < 1$, preto $p < 0.5$.

Získali sme explicitné vzorce pre stacionárne rozdelenie pravdepodobnosti Markovovho reťazca:

$$\pi_k = \varrho^k (1 - \varrho) = \left[\frac{p^2}{q^2} \right]^k \left[1 - \frac{p^2}{q^2} \right], \quad k = 0, 1, 2, \dots \quad (2.1)$$

Môžeme určiť stredný počet požiadaviek v stabilizovanom systéme:

$$\begin{aligned} Eq &= \sum_{k=0}^{\infty} k \pi_k = \sum_{k=0}^{\infty} k \varrho^k (1 - \varrho) = \varrho(1 - \varrho) \sum_{k=1}^{\infty} k \varrho^{k-1} = \varrho(1 - \varrho) \sum_{k=1}^{\infty} \frac{d}{d\varrho} \varrho^k = \\ &= \varrho(1 - \varrho) \frac{d}{d\varrho} \sum_{k=1}^{\infty} \varrho^k = \varrho(1 - \varrho) \frac{d}{d\varrho} \frac{\varrho}{1 - \varrho} = \varrho(1 - \varrho) \frac{1}{(1 - \varrho)^2} \Rightarrow \\ &\Rightarrow Eq = \frac{\varrho}{1 - \varrho} \end{aligned} \quad (2.2)$$

Ďalšou zaujímavou charakteristikou systému je obsadenosť linky, ktorá vysiela požiadavky zo systému. Popíšeme ju náhodnou premennou S . Linka má len dva stavy, buď je obsadená alebo nevysiela, preto pre jej strednú hodnotu platí

$$ES = 0 \cdot \pi_0 + 1 \cdot (\pi_1 + \pi_2 + \pi_3 + \dots) = 1 - \pi_0 = \varrho \quad (2.3)$$

Poslednú dôležitú charakteristiku, stredný počet požiadaviek čakajúcich na vysielanie EL dostaneme ako rozdiel predchádzajúcich charakteristík:

$$EL = Eq - ES = \frac{\varrho}{1 - \varrho} - \varrho = \frac{\varrho^2}{1 - \varrho} \quad (2.4)$$

Geo/D/1/∞

Nech do systému vstupuje Bernoulliho tok s prírastkami $a(i) \sim Bi(2, p)$, nech doba spracovania požiadavok je konštantná $c = 1p/ts$, a nech počet požiadaviek v systéme je neohraničený. Systém predstavuje Markovov reťazec s diskretným časom, ktorý má geometrické stacionárne rozdelenie $\pi_k = \varrho^k(1 - \varrho)$ a pre stredný počet požiadaviek v systéme, strednú obsadenosť linky a stredný počet požiadaviek čakajúcich na spracovanie platí:

$$Eq = \frac{\varrho}{1 - \varrho}, \quad ES = \varrho, \quad EL = \frac{\varrho^2}{1 - \varrho}, \quad \text{pričom} \quad \varrho = \frac{p^2}{q^2}$$

Na začiatku tejto kapitoly sme za pri simulácii zaoberali maximálnym oneskorením požiadaviek v systéme. Keďže systém je stochastický, vstupný tok je Bernoulliho proces, nemôžeme určiť maximálnu hodnotu oneskorenia s istotou, ale len s nejakou pravdepodobnosťou. Vypočítame napr. pravdepodobnosť, že požiadavka nebude čakať v systéme viac než τ ms:

$$\Pr(q \leq \tau) = 1 - \Pr(q > \tau) = 1 - \sum_{k=\tau+1}^{\infty} \pi_k = 1 - \varrho^{\tau+1} \quad (2.5)$$

PRÍKLAD 2.2 Nech kapacita systému *Geo/D/1/∞* je $c = 1p/ms$ a vstupný tok je Bernoulliho s prírastkami $a(i) \sim Bi(2, 0.4)$. Úlohou bude vypočítať charakteristiky systému a určiť oneskorenie požiadaviek v systéme, ktoré sa na 99.5% neprekročí.

Najprv vypočítame parameter $\varrho = p^2/q^2 = 0.4^2/0.6^2 = 0.444$. Pozorný čitateľ si isto všimol, že stacionárne rozdelenie systému je geometrické, $\pi_k = \varrho^k(1 - \varrho)$, ktoré vďaka podmienke $\varrho < 1$ s rastúcim k limitne zaniká:

$$\pi_0 = 0.556, \quad \pi_1 = 0.247, \quad \pi_2 = 0.110, \quad \pi_3 = 0.049, \quad \pi_4 = 0.022, \quad \pi_5 = 0.010$$

Vypočítame charakteristiky systému:

$$ES = \varrho = 0.444, \quad Eq = \frac{\varrho}{1 - \varrho} = 0.8, \quad EL = \frac{\varrho^2}{1 - \varrho} = 0.356$$

Prenosová linka je využitá na 44.4%, v prostredí je v priemere 0.8 požiadaviek a na spracovanie vo fronte čaká v priemere 0.356 požiadaviek.

Určíme, aké oneskorenie τ s pravdepodobnosťou 99.5% neprekročí.

$$\Pr(q \leq \tau) = 1 - \Pr(q > \tau) = 1 - \sum_{k=\tau+1}^{\infty} \pi_k = 1 - \varrho^{\tau+1} = 0.995 \quad \Rightarrow$$

$$\Rightarrow \varrho^{\tau+1} = 0.005 \quad \Rightarrow \quad \tau + 1 = \frac{\ln(0.005)}{\ln(0.444)} = 6.526 \quad \Rightarrow \quad \tau = 5.526$$

Hodnota τ musí byť celočíselná, preto pre danú pravdepodobnosť 0.995 nemô-

žeme hodnotu τ určiť. Kritériu ale môžeme sprísniť a tvrdiť, že oneskorenie požiadaviek neprekročí 6 ms hranicu s pravdepodobnosťou $1 - \rho^7 = 0.997$.

2.2.2 Model s diskretným časom $Geo/Geo/1/\infty$

Budeme sa zaoberať modelom s diskretným časom, pričom budeme predpokladať, pričom zvolíme veľkosť časového slotu tak, aby sa počas jeho trvania mohla vyskytnúť iba jedna požiadavka a systém mohol vyslať tiež iba jednu požiadavku. Budeme predpokladať, že vstupný tok aj obslužný proces tvoria Bernoulliho toky s prírastkami $a(i) \sim Alt(p_a)$ a $c(i) \sim Alt(p_c)$. Predpokladáme, že počet čakajúcich požiadaviek v prostredí je neohraničený. Takéto prostredie nazveme systém hromadnej obsluhy (SHO) a označíme ako $(Geo/Geo/1/\infty)$. Označenie informuje, že vstupný tok je Bernoulliho, resp. má geometrické medzery medzi vstupmi (Geo) a doba spracovania požiadaviek má geometrické rozdelenie. Rozdielna doba spracovania môže súvisieť s roznyimi veľkosťami požiadaviek.

Aj tento systém môžeme simulovať pomocou Lindleyho rovnice, pričom kapacitu c nahradíme za prírastky obslužného procesu:

$$q(t) = \max\{q(t-1) + a(t) - c(t), 0\} = (q(t-1) + a(t) - c(t))^+$$

PRÍKLAD 2.3 Pomocou Lindleyho rovnice vytvorte simuláciu systému $(Geo/Geo/1/\infty)$ s prírastkami vstupného toku $a(i) \sim Alt(0.7)$ a s prírastkami obslužného procesu $c(i) \sim Alt(0.8)$. Určite maximálnu počet požiadaviek v systéme.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$a(t)$	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0
$c(t)$	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0
$q(t)$	0	0	1	1	1	1	2	2	3	3	3	2	1	0	0	0

Maximálny počet požiadaviek v komunikačnom prostredí v našej simulácii je $q_{\max} = 3$. Keďže je vysielanie náhodné, vo všeobecnosti maximálne oneskorenie nevieme určiť. V našom prípade, ak je správa frontu FIFO, vidíme v tabuľke, že najdlhšie čakali v systéme požiadavky, ktoré vstúpili v časoch $t = 5$ a $t = 6$, ich oneskorenie bolo $d = 5ts$.

Vstupný a servisný proces v tomto modeli tvoria Bernoulliho procesy, preto proces popisujúci počet požiadaviek v systéme, $q(t)$, je Markovov reťazec s diskretným časom. Na jeho úplný popis potrebujeme poznať maticu pravdepodobností prechodov medzi stavmi reťazca, $\mathbf{P} = \{p_{ij}\}$.

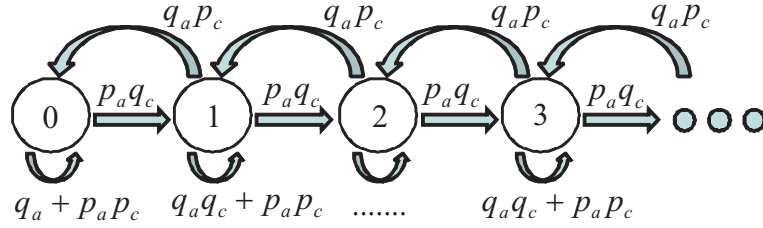
Nech časový slot je rovný jednej milisekunde. Nech je systém v čase 0 prázdny, $q(0) = 0$. Aby počas jedného časového slotu systém nezmenil stav, na vstupe sa nesmie vyskytnúť žiadna požiadavka, resp. môže vyskytnúť jedna požiadavka, ktorú za daný slot systém stihne spracovať. Ak systém požiadavku spracovať nestihne, v nasledujúcom časovom slote sa bude nachádzať v stave 1, $q(1) = 1$:

$$p_{0,0} = q_a + p_a p_c, \quad p_{0,1} = p_a q_c$$

Nech sa v systéme nachádza k požiadaviek na spracovanie. Systém prejde do stavu $k - 1$ ak sa jedna požiadavka obslúži a pritom žiadna nová nevstúpi, systém prejde do stavu $k + 1$ ak sa na vstupe nevyskytne nová požiadavka ale systém žiadnu neobslúži. Stav systému ostane nezmenený, ak požiadavka ani nepríde ani neodíde, alebo zároveň sa vyskytne nová požiadavka a systém jednu požiadavku odovysiela. Pre pravdepodobnosti prechodov platí

$$p_{k,k-1} = q_a p_c, \quad p_{k,k} = p_a p_c + q_a q_c, \quad p_{k,k+1} = p_a q_c$$

Zostrojíme prechodový graf reťazca:



Obr. 2.6: Prechodový graf systému $Geo/Geo/1/\infty$

Matica pravdepodobností prechodov Markovoho reťazca má tvar:

$$\mathbf{P} = \begin{pmatrix} q_a + p_a p_c & p_a q_c & 0 & 0 & \dots \\ q_a p_c & q_a q_c + p_a p_c & p_a q_c & 0 & \dots \\ 0 & q_a p_c & q_a q_c + p_a p_c & p_a q_c & \dots \\ 0 & 0 & q_a p_c & q_a q_c + p_a p_c & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Stacionárne rozdelenie reťazca získame zo vzťahu (1.3), $[\mathbf{E} - \mathbf{P}]^T \boldsymbol{\pi}^T =$

$$= \begin{pmatrix} 1 - q_a - p_a p_c & -q_a p_c & 0 & 0 & \dots \\ -p_a q_c & 1 - q_a q_c - p_a p_c & -q_a p_c & 0 & \dots \\ 0 & -p_a q_c & 1 - q_a q_c - p_a p_c & -q_a p_c & \dots \\ 0 & 0 & -p_a q_c & 1 - q_a q_c - p_a p_c & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Využijeme vzťah $q_a q_c + p_a p_c + p_a q_c + q_a p_c = 1$ a maticu koeficientov nekonečného systému zjednodušíme:

$$\begin{pmatrix} p_a q_c & -q_a p_c & 0 & 0 & \dots \\ -p_a q_c & p_a q_c + q_a p_c & -q_a p_c & 0 & \dots \\ 0 & -p_a q_c & p_a q_c + q_a p_c & -q_a p_c & \dots \\ 0 & 0 & -p_a q_c & p_a q_c + q_a p_c & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Vytvoríme riadkovo ekvivalentnú maticu pomocou sčítania vždy dvoch za sebou idúcich riadkov, a získame rekurentné pre stacionárne rozdelenie reťazca:.

$$\begin{pmatrix} p_a q_c & -q_a p_c & 0 & 0 & \dots \\ 0 & p_a q_c & -q_a p_c & 0 & \dots \\ 0 & 0 & p_a q_c & -q_a p_c & \dots \\ 0 & 0 & 0 & p_a q_c & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \Rightarrow \pi_k = \frac{p_a q_c}{q_a p_c} \pi_{k-1}, \quad k = 1, 2, \dots$$

Označíme podiel pravdepodobností ako $\varrho = p_a q_c / q_a p_c$ a vyjadríme stacionárne rozdelenie reťazca pomocou pravdepodobností π_0 , $\pi_k = \varrho^k \pi_0$. Pravdepodobnosť π_0 získame z normovacej podmienky

$$\sum_{k=0}^{\infty} \pi_k = 1 \Rightarrow \sum_{k=0}^{\infty} \varrho^k \pi_0 = 1 \Rightarrow \frac{1}{1-\varrho} \pi_0 = 1 \Rightarrow \pi_0 = 1 - \varrho$$

Nekonečný rad konverguje len vďaka podmienke $\varrho < 1$. To ale znamená:

$$\varrho < 1 \Rightarrow \frac{p_a q_c}{q_a p_c} < 1 \Rightarrow p_a q_c < q_a p_c \Rightarrow p_a < p_c$$

Pravdepodobnosti p_a a p_c predstavujú intenzity tokov, resp. stredné hodnoty vstupujúcich a obslužených požiadaviek za časovú jednotku. Keďže ide o náhodné procesy, aby sa systém nezahltal, musí v strednej hodnote vstupovať menej požiadaviek, než je systém schopný obslúžiť, preto $p_a < p_c$.

Získali sme explicitné vzorce pre stacionárne rozdelenie pravdepodobností Markovovho reťazca:

$$\pi_k = \varrho^k (1 - \varrho) = \left[\frac{p_a q_c}{q_a p_c} \right]^k \left[1 - \frac{p_a q_c}{q_a p_c} \right], \quad k = 0, 1, 2, \dots \quad (2.6)$$

Charakteristiky systému získame rovnako ako pri modeli $Geo/D/1/\infty$

Geo/Geo/1/∞

Nech do systému vstupuje Bernoulliho tok s prírastkami $a(i) \sim Alt(p_a)$, a nech obsluhu požiadaviek riadi Bernoulliho proces s prírastkami $c(i) \sim Alt(p_c)$. Predpokladáme, že počet požiadaviek v systéme je neohraničenný. Systém predstavuje Markovov reťazec s diskretným časom, ktorý má geometrické stacionárne rozdelenie $\pi_k = \varrho^k (1 - \varrho)$ a pre stredný počet požiadaviek v systéme, strednú obsadenosť linky a stredný počet požiadaviek čakajúcich na spracovanie platí:

$$Eq = \frac{\varrho}{1 - \varrho}, \quad ES = \varrho, \quad EL = \frac{\varrho^2}{1 - \varrho}, \quad \text{pričom} \quad \varrho = \frac{p_a q_c}{q_a p_c}$$

PRÍKLAD 2.4

Nech systému $Geo/Geo/1/\infty$ má parametre $p_a = 0.7$ a $p_c = 0.8$. Úlohou bude vypočítať charakteristiky systému a určiť oneskorenie požiadaviek v systéme, ktoré sa na 99.5% neprekročí.

Najprv vypočítame parameter $\varrho = \frac{p_a q_c}{q_a p_c} = \frac{0.7 \cdot 0.2}{0.3 \cdot 0.8} = 0.583$. Na výpočet stredných charakteristík systému nám stačí poznať hodnotu parametra ϱ :

$$ES = \varrho = 0.583, \quad Eq = \frac{\varrho}{1 - \varrho} = 1.4, \quad EL = \frac{\varrho^2}{1 - \varrho} = 0.817$$

Prenosová linka je využitá na 58.3%, v prostredí je v priemere 1.4 požiadaviek a na spracovanie vo fronte čaká v priemere 0.817 požiadaviek.

Určíme, aké oneskorie τ s pravdepodobnosťou 99.5% neprekročí.

$$\Pr(q \leq \tau) = 1 - \Pr(q > \tau) = 1 - \sum_{k=\tau+1}^{\infty} \pi_k = 1 - \varrho^{\tau+1} = 0.995 \quad \Rightarrow$$

$$\Rightarrow \varrho^{\tau+1} = 0.005 \quad \Rightarrow \quad \tau + 1 = \frac{\ln(0.005)}{\ln(0.444)} = 9.8196 \quad \Rightarrow \quad \tau = 8.8196$$

Hodnota τ musí byť celočíselná, preto pre danú pravdepodobnosť 0.995 nemôžeme hodnotu τ určiť. Kritériu ale môžeme sprísniť a tvrdiť, že oneskorenie požiadaviek neprekročí 9 ms hranicu s pravdepodobnosťou $1 - \varrho^7 = 0.9954$.

PRÍKLAD 2.5

Majme systém *Geo/Geo/1/5*. Posledná hodnota znamená, že v systéme môže čakať na obsluhu iba 5 požiadaviek. Parametre necháme rovnaké ako v predchádzajúcom príklade, $p_a = 0.7$ a $p_c = 0.8$. Parameter ϱ má rovnakú hodnotu ako v systéme s nekonečným frontom, $\varrho = 0.583$. Aby sme mohli určiť charakteristiky systému, musíme vypočítať najprv stacionárne rozdelenie reťazca.

Systém sa môže nachádzať v siedmych stavoch, ktoré predstavujú situácie: prázdny systém, v systéme je 1 požiadavka, ktorá je obsluhovaná, alebo v systéme sa nachádza 1 až 5 požiadaviek čakajúcich na obsluhu. Pre stacionárne rozdelenie reťazca sme odvodili vzťahy závislé od π_0 , pričom pravdepodobnosť prázdneho systému musíme odvodiť z normovacej podmienky:

$$\pi_k = \varrho^k \pi_0, \quad \sum_{k=0}^6 \pi_k = 1 \quad \Rightarrow \quad \pi_0 = \left[\sum_{k=0}^6 \varrho^k \right]^{-1} = \frac{1 - \varrho}{1 - \varrho^7}$$

Pre ručné počítanie a aj ako návod na jednoduchý program si ukážeme nasledujúci postup. Najprv vypočítame všetky potrebné mocniny ϱ^k , z nich určíme pravdepodobnosť π_0 a rekurentným zťahom $\pi_k = \varrho \pi_{k-1}$ dopočítame ostatné pravdepodobnosti. Nakoniec postupne postupne zráťame členy sumy $\sum k \pi_k$:

k	ϱ^k	π_k	$k \cdot \pi_k$
0	1.0000	0.4265	0
1	0.5833	0.2488	0.2488
2	0.3403	0.1451	0.2902
3	0.1985	0.0847	0.2540
4	0.1158	0.0494	0.1975
5	0.0675	0.0288	0.1440
6	0.0394	0.0168	0.1008
\sum	2.3448	1	1.2353

Prvá hodnota sumácie v dolnom riadku tabuľky predstavuje obrátenú hodnotu $\pi_0 = 1/2.3448 = 0.4265$. Druhá hodnota sumácie predstavuje normováciu podmienku $\sum \pi_k = 1$ a posledná suma $\sum k \cdot \pi_k$ je stredný počet požiadaviek v systéme, $Eq = 1.2353$. Obsadenosť linky je $ES = 1 - \pi_0 = 0.5735$, čiže systém je využitý na 57.35%. Vo fronte čaká $EL = Eq - ES = 0.6618$ požiadaviek a systém odmietne požiadavku s pravdepodobnosťou $\pi_6 = 0.0168$.

2.2.3 Markovove reťazce so spojitým časom

Uvažujme o nejakom náhodnom procese $\{X(t)\}_{t \in T}$, ktorý v priebehu času náhodne prechádza určitými stavmi. Predpokladáme spojitý čas. Výraz $X(t) = j$ znamená, že systém sa v čase t nachádza v stave s_j . Pričom v tomto stave sa nachádza s určitou pravdepodobnosťou $\Pr[X(t) = j]$.

Markovov reťazec so spojitým časom

Náhodný proces $\{X(t)\}_{t \in T}$ s množinou stavov S nazveme Markovov reťazec so spojitým časom, ak pre proces platí:

- množina $T = R_0^+$
- množina S je spočítateľná, $S = \{s_1, s_2, \dots\}$
- pre $\forall s_j \in S$ a $\forall t_j \in T$ také, že $t_1 < \dots < t_n$ platí

$$\begin{aligned} \Pr[X(t_n) = s_n / X(t_{n-1}) = s_{n-1}, \dots, X(t_1) = s_1] = \\ = \Pr[X(t_n) = s_n / X(t_{n-1}) = s_{n-1}] \end{aligned}$$

Posledná vlastnosť sa nazýva **Markovova**. Popisuje dôležitú charakteristiku Markovových reťazcov, tzv. "zabúdanie na minulosť". Pravdepodobnosť, že systém sa nachádza v čase t_n v stave s_n , ak predtým bol v čase t_{n-1} v stave s_{n-1} , nezávisí od toho, v akých stavoch sa nachádzal pred časom t_{n-1} .

Homogénny Markovov reťazec

Markovov reťazec $\{X(t)\}_{t \in T}$ so spojitým časom nazveme **homogénny** v čase, ak pre $\forall s_i, s_j \in S$, $\forall t, r \in T$ a $\forall h$ platí:

$$\Pr[X(t+h) = j / X(t) = i] = \Pr[X(r+h) = j / X(r) = i]$$

pričom h je také, že $t+h \in T$ a $r+h \in T$.

Homogenita reťazca znamená, že pravdepodobnosť prechodu zo stavu s_i do stavu s_j za čas h nezávisí od toho, kde v čase je reťazec umiestnený. Ďalej sa budem zaoberať len homogénnymi Markovovými reťazcami so spojitým časom a preto označenie homogénny budeme vynechávať.

Pravdepodobnosť, že sa reťazec v čase t nachádza v stave s_j , $\Pr(X(t) = j)$, nazveme pravdepodobnosť stavu s_j v čase t , resp. pravdepodobnosť výskytu j -teho stavu po čase t a označíme $p_j(t)$:

$$p_j(t) = \Pr[X(t) = j] \quad (2.7)$$

Vektor $\mathbf{p}(t) = (p_0(t), p_1(t), \dots, p_j(t), \dots)$ nazveme **rozdelenie reťazca** v čase t . Vektor $\mathbf{p}(0) = (p_0(0), p_1(0), \dots, p_j(0), \dots)$ nazveme **počiatočné rozdelenie reťazca**. Tento vektor pravdepodobností určuje štartovacie podmienky reťazca.

Pravdepodobnosť prechodu reťazca zo stavu s_i do stavu s_j za čas h vzhľadom na homogenitu označíme: $p_j(t)$:

$$p_{i,j}(h) = \Pr[X(t+h) = j/X(t) = i] \quad (2.8)$$

Definitorky položíme:

$$p_{i,j}(0) = \delta_{i,j}, \text{ kde } \delta_{i,j} = 0 \text{ pre } i \neq j \text{ a } \delta_{i,i} = 1 \quad (2.9)$$

Budeme sa teda zaoberať len takými Markovovými reťazcami, ktoré za nulový čas nemôžu prejsť do iného stavu.

Z vety o úplnej pravdepodobnosti dostávame vzťah medzi pravdepodobnosťami stavov v čase t a pravdepodobnosťami prechodov medzi stavmi za čas t :

$$\forall j \in S, \forall t \in T; \quad p_j(t) = \sum_{i \in S} p_i(0)p_{i,j}(t), \quad \text{resp.} \quad \mathbf{p}(t) = \mathbf{p}(0) \cdot \mathbf{P}(t) \quad (2.10)$$

Chapman-Kolmogorová rovnosť

Nech $\{X(t)\}_{t \in T}$ je Markovov reťazec so spojitým časom, s množinou stavov S . Potom pre $\forall i, j \in S; \forall r, t \in T$ také, že $r+t \in T$ platí:

$$p_{i,j}(r+t) = \sum_{k \in S} p_{i,k}(r)p_{k,j}(t)$$

Dôkaz Chapman-Kolmogorovej rovnosti si uvedieme ako jednoduché cvičenie. Postupne využijeme homogenitu reťazca, vetu úplnej pravdepodobnosti, vzťah

$$\Pr(A \cap B/C) = \Pr(A/B \cap C) \Pr(B/C),$$

a nakoniec Markovovú vlastnosť reťazca:

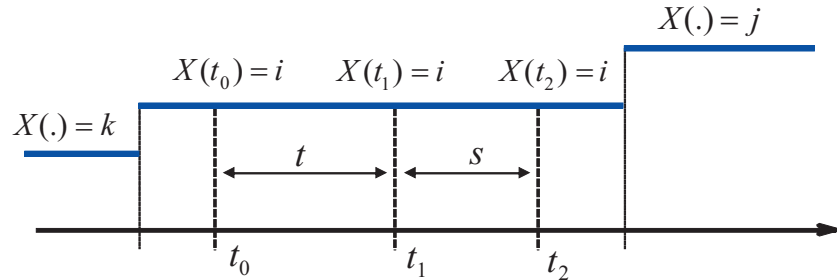
$$\begin{aligned} p_{i,j}(r+t) &= \Pr[X(r+t) = j/X(0) = i] \\ &= \sum_{k \in S} \Pr[X(r+t) = j, X(r) = k/X(0) = i] \\ &= \sum_{k \in S} \Pr[X(r+t) = j/X(r) = k, X(0) = i] \cdot \Pr[X(r) = k/X(0) = i] \\ &= \sum_{k \in S} \Pr[X(r+t) = j/X(r) = k] \cdot \Pr[X(r) = k/X(0) = i] \\ &= \sum_{k \in S} p_{i,k}(r)p_{k,j}(t) \quad \square \end{aligned}$$

O diferencovateľnosti pravdepodobností

Nech $\{X(t)\}_{t \in T}$ je Markovov reťazec so spojitým časom, s množinou stavov S . Pre $\forall i, j \in S$ sú pravdepodobnosti stavov $p_j(t)$ a pravdepodobnosti prechodov $p_{i,j}(t)$ rovnomerne spojitě funkcie premennej t a diferencovateľné sprava v hodnote 0. (Dôkaz čitateľ môže nájsť v [4])

Odvodíme dôležitú vlastnosť Markovových reťazcov so spojitým časom. Pomocou tejto vlastnosti môžeme v praxi určiť ktorý proces môžeme modelovať pomocou Markovovho reťazca:

Uvažujme o časových okamihoch t_0, t_1 a t_2 takých, že $t_0 < t_1 < t_2$. Nech t a s sú časové medzery medzi týmito okamihami, teda $t = t_1 - t_0$ a $s = t_2 - t_1$. Nech sa reťazec nachádzal v čase t_0 v stave i , $X(t_0) = i$, a nech za čas $t + s$ tento stav neopustí:



Obr. 2.7: Vývoj Markovovho reťazca v časoch t_0, t_1 a t_2

Potom z Markovovej vlastnosti vyplýva:

$$\Pr[X(t_2) = i / X(t_1) = i, X(t_0) = i] = \Pr[X(t_2) = i / X(t_1) = i] \quad (2.11)$$

Nech náhodná premenná T_i popisuje dobu pobytu reťazca v stave i resp. s_i . Keďže reťazec zotrúva v stave i počas doby $t + s$, môžeme vzťah (2.11) prepísať:

$$\Pr(T_i > t + s / T_i > t) = \Pr(T_i > s) \quad (2.12)$$

Všimnime si, že tento vzťah vyjadruje "zabúdanie" reťazca na "minulosť", čo predstavuje interpretáciu Markovovej vlastnosti.

žitím elementárnej teórie pravdepodobnosti dostávame:

$$\begin{aligned} \Pr(T_i > t + s / T_i > t) &= \frac{\Pr(T_i > t + s)}{\Pr(T_i > t)} = \Pr(T_i > s) \Rightarrow \\ \Rightarrow \Pr(T_i > t + s) &= \Pr(T_i > t) \cdot \Pr(T_i > s) \end{aligned} \quad (2.13)$$

Vieme, že platí $\Pr(T_i = 0) = p_{i,j}(0) = 0$ pre $i \neq j$. Označíme $f(t) = \Pr(T_i > t)$. Keďže náhodná premenná T_i je spojitá (popisuje spojitý čas), má spojitú a diferencovateľnú distribučnú funkciu $F(t)$. Potom aj uvažovaná pravdepodobnostná funkcia $f(t) = \Pr(T_i > t) = 1 - F(t)$ je spojitá a diferencovateľná. Zo vzťahu (2.13) vyplýva

$$f(t + s) = f(t) \cdot f(s) \quad \text{s podmienkou} \quad f(0) = \Pr(T_i > 0) = 1 - \Pr(T_i = 0) = 1$$

Celý vzťah zlogaritmujeme a zderivujeme podľa s :

$$\begin{aligned} \ln f(t + s) &= \ln f(t) + \ln f(s) \quad \text{s podmienkou} \quad \ln f(0) = 0 \Rightarrow \\ \Rightarrow \frac{f'(t + s)}{f(t + s)} &= \frac{f'(s)}{f(s)} \Rightarrow \text{pre } s = 0 \text{ platí } \frac{f'(t)}{f(t)} = f'(0) \end{aligned}$$

Funkcia $f(t)$ je klesajúca, vyplýva to priamo z definície $f(t) = \Pr(T_i > t)$, v bode $t = 0$ má lokálne maximum $f(0) = 1$ a pre $\forall t \geq 0; f(t) > 0$. Preto môžeme označiť deriváciu sprava v bode nula ako $f'(0) = -\lambda_i$, kde λ_i je kladná konštanta.

$$\frac{f'(t)}{f(t)} = -\lambda_i \Rightarrow \frac{d}{dt} \ln f(t) = -\lambda_i \Rightarrow \ln f(t) = -\lambda_i t + C; \quad C \in \mathbb{R}$$

Vieme, že $f(0) = 1$, preto $C = 0$. Na záver teda dostaneme:

$$\ln f(t) = -\lambda_i t \Rightarrow f(t) = e^{-\lambda_i t} \Rightarrow \Pr(T_i > t) = e^{-\lambda_i t}$$

Pre distribučnú funkciu náhodnej premennej T_i platí

$$F(t) = 1 - \Pr(T_i > t) = 1 - e^{-\lambda_i t}$$

Medzery v Markovovom reťazci

V Markovovom reťazci so spojitým časom sa doby pobytu v jednotlivých stavoch, resp. medzery medzi prechodmi do jednotlivých stavov, riadia podľa exponenciálneho rozdelenia. Nech sa reťazec nachádza v stave i , $X(\cdot) = i$. Potom pravdepodobnosť, že reťazec stav j za čas t neopustí sa rovná

$$\Pr(T_i > t) = e^{-\lambda_i t}$$

V Markovovom reťazci s diskretným časom sa doby pobytu v jednotlivých stavoch, resp. medzery medzi prechodmi do jednotlivých stavov, riadia podľa geometrického rozdelenia s parametrom p_i , ktorý predstavuje pravdepodobnosť, že systém počas jedného časového kroku stav i neopustí. Ľahko sa dá ukázať, že pravdepodobnosť, že reťazec stav i za čas τ neopustí sa rovná

$$\Pr(T_i > \tau) = p_i^{\tau+1}$$

V predchádzajúcich úvahách sme pre zjednodušenie vynechali skutočnosť, že reťazec sa nachádzal v čase t_0 v stave i . Ku každému stavu i môže prislúchať odlišná konštanta λ_i , preto jednotlivé náhodné premenné majú rozdelenie $T \sim \exp(\lambda_i)$.

Poznámka 1.4.1

Hovoríme, že funkcia $f(x)$ je rádu $o(x)$, ak platí:

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = 0$$

To znamená, že funkcia $f(x)$ klesá k nule rýchlejšie, než lineárna funkcia x . Ak sa dá overiť, že súčtom funkcií rádu $o(x)$, násobením konštantou a umocnením opäť dostávame funkciu rádu $o(x)$. Pre zjednodušenie zapisujeme $f(x) = o(x)$.

Ďalej budeme pod označením Δt rozumieť taký "krátky" časový interval, že už jeho *mužovo* *umocnina* je

$$\forall n > 1; \quad \Delta t^n = o(\Delta t)$$

Treba ešte pripomenúť, že symbol $o(\Delta t)$ je len označenie rádu funkcie, nie je to konkrétna funkcia.

Nech náhodná premenná T popisuje časovú medzeru medzi prechodmi do jednotlivých stavov v momente, keď sa reazec nachádza v stave i . Nech za čas Δt reazec v tomto stave zotrúva. Potom z predchádzajúcich úvah vyplýva, že

$$P(T > \Delta t / X(t_0) = i) = e^{-\lambda_i \Delta t}$$

Teraz použijeme rozvoj funkcie e^x do MacLaurinovho radu:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \Rightarrow e^{-\lambda_i \Delta t} = \sum_{k=0}^{\infty} \frac{(-\lambda_i \Delta t)^k}{k!} = 1 - \lambda_i \Delta t + o(\Delta t)$$

Potom pravdepodobnosť, že reazec za čas Δt opustí stav i má tvar:

$$P(T < \Delta t / X(t_0) = i) = 1 - e^{-\lambda_i \Delta t} = \lambda_i \Delta t + o(\Delta t)$$

Pravdepodobnosť $p_{i,j}(\Delta t)$ vyjadruje pravdepodobnosť prechodu zo stavu i do stavu j za čas Δt , pričom stav j nasledoval bezprostredne po stave i , reazec v priebehu času Δt neprechádzal inými stavmi. Je to analogická veličina ako pravdepodobnosť prechodu za jeden časový okamih $p_{i,j}$ pri reazcoch s diskretným časom. Pravdepodobnosti $p_{i,j}(\Delta t)$ vyjadrujú správanie sa reazca v krátkych časových (akoby už ďalej nedeliteľných) okamihoch Δt . Preto môžeme písať:

$$p_{i,i}(\Delta t) = P(T > \Delta t / X(t_0) = i) = e^{-\lambda_i \Delta t} = 1 - \lambda_i \Delta t + o(\Delta t)$$

Vajmame si, že k tomuto výsledku sme dospeli prirodzeným spôsobom len využitím Markovovej vlastnosti reazca. Vzniká otázka, čomu sa rovnajú pravdepodobnosti prechodov zo stavu i do iných stavov j , $p_{i,j}(\Delta t)$ pre $i \neq j$:

$$p_{i,i}(\Delta t) = 1 - \sum_{j \neq i} p_{i,j}(\Delta t) = 1 - \lambda_i \Delta t + o(\Delta t) \Rightarrow \sum_{j \neq i} p_{i,j}(\Delta t) = \lambda_i \Delta t + o(\Delta t)$$

Intuitívne sa dá vycítiť, že ku každej pravdepodobnosti $p_{i,j}(\Delta t)$ by mohla existovať konštanta $\lambda_{i,j}$ tak, že

$$\sum_{j \neq i} \lambda_{i,j} = \lambda_i \quad \text{resp.} \quad p_{i,j}(\Delta t) = \lambda_{i,j} \Delta t + o(\Delta t)$$

Ukáľeme, že takéto konštanty naozaj existujú.

Definícia 1.4.2

Nech $\{X(t)\}_{t \in T}$ je Markovov reazec so spojitým časom, s množinou stavov S . Potom pre všetky $i, j \in S$ nazveme konštantu $\lambda_{i,j} = \left. \frac{d}{dt} p_{i,j}(t) \right|_{t=0+}$ **intenzitou prechodu** zo stavu i do stavu j . Konštantu $\lambda_i = -\lambda_{i,i}$ nazveme **intenzitou výstupu** zo stavu i .

Definícia 1.4.2. je korektná, pretože existenciu konštánt $\lambda_{i,j}$ zaručuje Veta 1.4.2., ktorá hovorí, že pravdepodobnosti $p_{i,j}(t)$ sú diferencovateľné v bode 0.

Dôsledok 1.4.3

Nech $\{X(t)\}_{t \in T}$ je Markovov reazec so spojitým časom, s množinou stavov S .

pre $\forall i \in S$ platí: $p_{i,i}(\Delta t) = 1 - \lambda_i \Delta t + o(\Delta t)$

pre $\forall i, j \in S$ také, že $i \neq j$ platí: $p_{i,j}(\Delta t) = \lambda_{i,j}\Delta t + o(\Delta t)$

Dôkaz

Tvrdenie vyplýva priamo z Definície 1.4.2. Pre každý stav $i \in S$ platí:

$$\begin{aligned} \lambda_{i,i} &= \left. \frac{d}{dt} p_{i,i}(t) \right|_{t=0} = \lim_{\Delta t \rightarrow 0^+} \frac{p_{i,i}(t + \Delta t) - p_{i,i}(t)}{\Delta t} \Big|_{t=0} = \lim_{\Delta t \rightarrow 0^+} \frac{p_{i,i}(\Delta t) - 1}{\Delta t} \Rightarrow \\ &\Rightarrow p_{i,i}(\Delta t) = 1 + \lambda_{i,i}\Delta t + o(\Delta t) \end{aligned}$$

Keďže sme zaviedli $\lambda_{i,i} = -\lambda_i$, dostávame prvé tvrdenie

$$p_{i,i}(\Delta t) = 1 - \lambda_i\Delta t + o(\Delta t)$$

Nech teraz i, j sú také, že $i \neq j$. Potom platí:

$$\begin{aligned} \lambda_{i,j} &= \left. \frac{d}{dt} p_{i,j}(t) \right|_{t=0^+} = \lim_{\Delta t \rightarrow 0^+} \frac{p_{i,j}(t + \Delta t) - p_{i,j}(t)}{\Delta t} \Big|_{t=0} = \lim_{\Delta t \rightarrow 0^+} \frac{p_{i,j}(\Delta t) - 0}{\Delta t} \Rightarrow \\ &\Rightarrow p_{i,j}(\Delta t) = \lambda_{i,j}\Delta t + o(\Delta t) \quad \square \end{aligned}$$

Teraz sa vrátime k úvaham, ktoré predchádzali Definícii 1.4.2., pričom vyulijeme Dôsledok 1.4.3:

$$\forall i \in S; \quad \sum_{j \neq i} p_{i,j}(\Delta t) = 1 - p_{i,i}(\Delta t) \Rightarrow \sum_{j \neq i} [\lambda_{i,j}\Delta t + o(\Delta t)] = \lambda_i\Delta t + o(\Delta t)$$

Poslednú rovnosť predelíme Δt , určíme limitu pre Δt idúce do nuly a vyulijeme pritom to, že pre veličinu $o(\Delta t)$ platí $\lim_{t \rightarrow 0^+} \frac{o(\Delta t)}{\Delta t} = 0$:

$$\sum_{j \neq i} \left[\lambda_{i,j} + \frac{o(\Delta t)}{\Delta t} \right] = \lambda_i + \frac{o(\Delta t)}{\Delta t} \Rightarrow \text{pre } \Delta t \rightarrow 0^+ \quad \sum_{j \neq i} \lambda_{i,j} = \lambda_i$$

Súčet intenzít prechodov zo stavu i do iných stavov j sa asymptoticky rovná intenzite výstupu zo stavu i , teda naše počiatočné úvahy boli správne. Pri označení $\lambda_i = -\lambda_{i,i}$ dostávame, že $\sum_{j \in S} \lambda_{i,j} = 0$.

Základný kameň analýzy Markovových reazcov so spojitým časom tvoria Kolmogorové diferenciálne rovnice. Odvodíme ich len pre reazce s konečným počtom stavov. Dôkaz pre reazce s nekonečným počtom stavov môže čitateľ nájsť v [?].

2.2.4 Zovšeobecnený proces vzniku a zániku

V tejto kapitole sa budeme zaoberať lineárnym procesom, ktorý popisuje kultúru nejakých jedincov. Títo jedinci môžu nejakým spôsobom vznikať alebo zanikať, pričom pojem jedinci chápeme v širších súvislostiach.

Nech sa v kultúre nachádza $i - 1$ jedincov. Vznik alebo zánik jedného jedinca považujeme za elementárne udalosti. Nech pravdepodobnosť, že za časový okamih Δt vznikne nový jedinec, je

$$\boxed{p_{i-1,i}(\Delta t) = \lambda_i\Delta t + o(\Delta t)} \quad \forall i = 1, 2, \dots; \quad \lambda_i \in R^+$$

kde konštanta λ_i je intenzitou prechodu zo stavu $i - 1$ do stavu i .

Nech je teraz v kultúre i jedincov. Pravdepodobnosť, že z i jedincov jeden jedinec zanikne, je

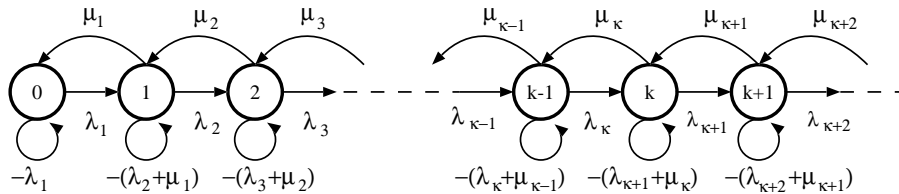
$$p_{i,i-1}(\Delta t) = \mu_i \Delta t + o(\Delta t) \quad \forall i = 1, 2, \dots; \quad \mu_i \in R^+$$

kde konštanta μ_i je intenzitou prechodu zo stavu i do stavu $i - 1$. Ďalej predpokladáme, že pravdepodobnosť, že nastane viac, než jedna elementárna udalosť za časový okamih Δt je veličina rádu $o(\Delta t)$:

$$p_{i,j}(\Delta t) = o(\Delta t) \quad |i - j| > 1$$

Vieme ešte, že jedinci sa chovajú navzájom nezávisle na sebe. Z takéhoto pohľadu sa jedná o homogénny Markovov reťazec so spojitým časom s nekonečnou spočítateľnou množinou stavov, v ktorom významné pravdepodobnosti majú len prechody o jeden stav "dopredu" alebo "dozadu" za časový okamih Δt . Ostatné pravdepodobnosti prechodov za čas Δt sú veličiny rádu $o(\Delta t)$.

Nakreslíme prechodový graf reťazca. Vrcholy v grafe predstavujú počet jedincov v kultúre, orientované šípky predstavujú možné prechody medzi stavmi s pravdepodobnosťou rôznou od $o(\Delta t)$ a ohodnotenie šípok predstavuje nenulové intenzity prechodov.



Pre analýzu tohoto Markovového reťazca použijeme predchádzajúcu teóriu uvedenú v kapitole 3.1. Najprv pomocou prechodového grafu zostrojíme Kolmogorové diferenciálne rovnice pre pravdepodobnosti stav reťazca v čase t , pričom predpokladáme, že v čase $t = 0$ sa reťazec nachádzal v stave nula, teda na počiatku kultúra neobsahovala žiadneho jedinca.

$$\begin{aligned} p'_0(t) &= -\lambda_1 p_0(t) + \mu_1 p_1(t) \\ p'_1(t) &= \lambda_1 p_0(t) - (\lambda_2 + \mu_1) p_1(t) + \mu_2 p_2(t) \\ &\vdots \\ p'_k(t) &= \lambda_k p_{k-1}(t) - (\lambda_{k+1} + \mu_k) p_k(t) + \mu_{k+1} p_{k+1}(t) \\ &\vdots \\ p_0(0) &= 1 \end{aligned}$$

Ďalej nás bude zaujímať len rozdelenie pravdepodobností stabilizovaného reťazca, teda nezávislého na čase t . Z prechodového grafu vidíme, že reťazec je tranzitívny. O stabilizácii tranzitívnych Markovových stavov pojednáva Markovova veta. Markovova veta je ale dokázaná len pre reťazce s konečnou množinou stavov, a náš model vzniku a zániku má nekonečne veľa stavov. Väčšina autorít v obore Teórie hromadnej obsluhy (napr. Chinčín, Feller, ...) predpokladá, že Markovova veta platí aj pre tranzitívne reťazce s nekonečným počtom stavov.

Nech teda Markovova veta platí a náš uvažovaný reťazec sa stabilizuje v čase. Potom ho môžeme popísať nasledujúcim systémom lineárnych rovníc, ktoré doplníme o normováciu pod-

mienku pre súčet pravdepodobností:

$$\begin{aligned}
 0 &= -\lambda_1 p_0 + \mu_1 p_1 \\
 0 &= \lambda_1 p_0 - (\lambda_2 + \mu_1) p_1 + \mu_2 p_2 \\
 &\vdots \\
 0 &= \lambda_k p_{k-1} - (\lambda_{k+1} + \mu_k) p_k + \mu_{k+1} p_{k+1} \\
 &\vdots \\
 \sum k 0 p_k &= 1
 \end{aligned}$$

Tento systém lineárnych rovníc vyriešime nasledujúcimi substitúciami

$$\forall k = 1, 2, \dots \quad \boxed{z_k = \lambda_k p_{k-1} - \mu_k p_k}$$

Po dosadení týchto substitúcií do (3.3.0.) dostávame systém rovníc:

$$\begin{aligned}
 0 &= -z_1 \\
 0 &= z_1 - z_2 \\
 &\vdots \\
 0 &= z_k - z_{k+1} \\
 &\vdots
 \end{aligned}$$

Vidíme, že platí:

$$z_1 = z_2 = \dots = z_k = \dots = 0$$

Pomocou týchto rovností vypočítame pravdepodobnosť, že sa stabilizovaný systém nachádza v stave k :

$$z_k = 0 \quad \Rightarrow \quad \lambda_k p_{k-1} - \mu_k p_k = 0 \quad \Rightarrow \quad p_k = \frac{\lambda_k}{\mu_k} p_{k-1}$$

Postupnými iteráciami dostaneme explicitné vzťahy pre pravdepodobnosti stavov stabilizovaného reťazca:

$$p_k = \frac{\lambda_k}{\mu_k} p_{k-1} = \frac{\lambda_k \lambda_{k-1}}{\mu_k \mu_{k-1}} p_{k-2} = \dots \quad \Rightarrow \quad \boxed{p_k = \prod_{i=1}^k \frac{\lambda_i}{\mu_i} p_0}$$

Pravdepodobnosť p_0 vypočítame z normovacej podmienky:

$$\sum k 0 p_k = 1 \quad \Rightarrow \quad \boxed{p_0 = \left[\sum k 0 \prod_{i=1}^k \frac{\lambda_i}{\mu_i} \right]^{-1}}$$

Zovšeobecnený model vzniku a zániku bude kľúčový pri analýze štandardných Markovových systémov hromadnej obsluhy. Kultúru jedincov budú predstavovať zákazníci nachádzajúci sa v obslužnom systéme. Vznik jedinca potom predstavuje vstup nového zákazníka do systému, zánik jedinca odchod obsluženého zákazníka zo systému. Vzťah (3.3.1.) budeme označovať ako Erlangove alebo zovšeobecnené Erlangove vzorce pre pravdepodobnosti stavov sralilizovaného systému p_k . Pre lepšie pochopenie si celý postup odvodenia týchto vzorcov pri jednotlivých modeloch vždy zopakujeme.

V nasledujúcej kapitole si odvodíme konkrétny tvar intenzít prechodov λ_i a μ_i pre Markovove obslužné systémy.

2.2.5 Analýza Markovovho modelu

Markovovým systémom hromadnej obsluhy nazveme obslužný systém, do ktorého vstupuje elementárny tok zákazníkov a doba obsluhy sa riadi podľa exponenciálneho rozdelenia. Väčšinou budeme intenzitu vstupného toku značiť λ a parameter exponenciálneho rozdelenia, ktoré modeluje obsluhu, značiť μ . Systémy hromadnej obsluhy môžu mať rôzny režim liniek a rôzny systém fronty. Najprv sa budeme zaoberať len takými systémami, v ktorých sú linky zapojené paralelne a pracujú nezávisle na sebe s rovnakým exponenciálnym rozdelením. To znamená, že zákazník si pre obsluhu vyberá ľubovlnú voľnú linku. Každá z liniek má rovnakú strednú dobu obsluhy. Ak systém zahŕňa aj frontu čakajúcich zákazníkov, nezáleží na tom, ktorý zákazník z fronty v okamihu uvoľnenia linky túto linku obsadzuje, ani nezáleží na poradí zákazníkov vo fronte. Jeden zákazník predstavuje jednu požiadavku pre systém.

V priebehu času vstupuje do systému náhodný tok zákazníkov a doba obsluhy liniek sa riadi podľa náhodného rozdelenia. Preto systém hromadnej obsluhy je stochastický dynamický systém závisiaci od času.

Markovov systém hromadnej obsluhy budeme popisovať systémom náhodných premenných $X(t)$, resp. náhodným procesom $\{X(t)\}_{t \in T}$, kde T je množina reálnych nezáporných čísel. Náhodný jav $X(t) = k$ znamená, že v čase t bolo v systéme k zákazníkov. Množinu všetkých hodnôt $S = \{0, 1, \dots\}$ náhodných premenných $X(t)$ nazveme množinou stavov systému. Podľa typu systému môže byť konečná alebo nekonečná spočítateľná. Preto proces $\{X(t)\}_{t \in T}$ je náhodný reťazec so spojitým časom. Dokážeme, že náhodný reťazec $\{X(t)\}_{t \in T}$ je Markovov.

Obsluha zákazníkov je modelovaná náhodnou premennou T s exponenciálnym rozdelením $\exp(\mu)$, ktoré sa v priebehu času nemení. Tak isto vstupný tok zákazníkov do systému je stacionárny (jedna z vlastností elementárneho toku), preto náhodný reťazec $\{X(t)\}_{t \in T}$ je **homogénny**. Jednoducho povedané, nezáleží na tom, v ktorom časovom okamihu systém pozorujeme, záleží iba na dĺžke pozorovania t .

Ak má nejaký reťazec Markovovú vlastnosť, znamená to, že nezávisí od minulosti. Ak v čase t_0 sa reťazec nachádza v stave i , pravdepodobnosť, že za okamih Δt bude v nejakom stave $j \in S$, nezávisí od toho, ktorými stavmi prechádzal pred časom t_0 . Čas t_0 označuje prítomnosť a čas $t_0 + \Delta t$ budúcnosť.

Nech sa v Markovovom systéme hromadnej obsluhy nachádza i zákazníkov v čase t_0 , $X(t_0) = i$. Zmena stavu systému za čas Δt , $X(t_0 + \Delta t)$, závisí od nasledujúcich náhodných javov:

- od okamihov ukončenia obslúh j zákazníkov ($j \leq i$), ktorí sa v systéme nachádzali už v čase t_0
- od vstupu nových zákazníkov do systému po čase t_0
- od okamihov ukončenia obslúh zákazníkov, ktorí do systému vstúpili po čase t_0

Ukážeme, že žiadna z uvedených situácií nezávisí od diania v systéme pred časom t_0 .

(1) Pravdepodobnosť ukončenia obsluhy zákazníka za čas Δt , keď obsluha začala už pred časom t_0 je:

$$P(t_0 < T < t_0 + \Delta t | T > t_0) = \frac{\int_{t_0}^{t_0 + \Delta t} \mu e^{-\mu x} dx}{\int_{t_0}^{\infty} \mu e^{-\mu x} dx} =$$

$$= \frac{e^{-\mu t_0} - e^{-\mu(t_0+\Delta t)}}{e^{-\mu t_0}} = 1 - e^{-\mu \Delta t} = P(T < \Delta t)$$

Pretože je obsluha exponenciálna, pravdepodobnosť skončenia obsluhy zákazníka za čas Δt nezávisí od toho, ako dlho pred časom t_0 bol už zákazník obsluhovaný. Keďže linky pracujú nezávisle, pravdepodobnosť ukončenia obslúh j za čas Δt tiež nezávisí od času t_0 .

(2) Keďže je vstupný tok zákazníkov do systému elementárny, je aj beznásledný. Zákazníci vstupujú do systému nezávisle na sebe. Počet zákazníkov, ktorí vstúpili do systému pred časom t_0 nemá preto vplyv na počet zákazníkov, ktorí vstúpia do systému v časovom intervale $(t_0, t_0 + \Delta t)$ (pozri definíciu beznáslednosti, kapitola 2).

(3) V tomto prípade je nezávislosť od diania pred časom t_0 zrejmalá.

Ukázali sme, že ak sa v systéme v čase t_0 nachádzalo i zákazníkov, pravdepodobnosť, že za časový okamih Δt bude v systéme j zákazníkov, kde $j \in S$, nezávisí od toho, koľko zákazníkov sa v systéme nachádzalo pred časom t_0 . Toto je ale inými slovami vyjadrená Markovová vlastnosť. Preto Markovov systém hromadnej obsluhy $\{N(t)\}_{t \in T}$ je **homogénny Markovov reťazec so spojitým časom**.

Odvodíme tvar pravdepodobností prechodov $p_{i,j}(\Delta t)$, teda pravdepodobnosti zmien stavov systému v priebehu krátkych okamihov Δt . Vstupný tok zákazníkov je elementárny, preto pravdepodobnosť príchodu práve jedného zákazníka za časový okamih Δt je $v_1(\Delta t)$. Túto hodnotu upravíme, využijúc MacLaurinovho rozvoja funkcie e^x a toho, že ľubovoľná mocnina veličiny Δt je rádu $o(\Delta t)$:

$$v_1(\Delta t) = \lambda \Delta t e^{-\lambda \Delta t} = \lambda \Delta t [1 - \lambda \Delta t + o(\Delta t)] = \lambda \Delta t + o(\Delta t)$$

Počet zákazníkov v systéme sa zmenší, ak niektorá z obsadených liniek skončí obsluhu a zákazník opustí systém. Linky pracujú podľa exponenciálneho rozdelenia. Nech náhodná premenná T popisuje dĺžku obsluhy jedného zákazníka. Potom pravdepodobnosť, že linka za čas Δt neskončí obsluhu, je

$$P(T > \Delta t) = \int_{\Delta t}^{\infty} \mu e^{-\mu x} dx = [-e^{-\mu x}]_{\Delta t}^{\infty} = e^{-\mu \Delta t}$$

Ak je obsadených k liniek, pravdepodobnosť, že žiadna z nich neskončí za čas Δt obsluhu (linky pracujú nezávisle na sebe) je

$$\prod_{i=1}^k P(T > \Delta t) = \prod_{i=1}^k e^{-\mu \Delta t} = e^{-k\mu \Delta t}$$

Preto pravdepodobnosť, že z k obsadených liniek sa za čas Δt aspoň jedna uvoľní, je

$$1 - \prod_{i=1}^k P(T > \Delta t) = 1 - e^{-k\mu \Delta t} = 1 - [1 - k\mu \Delta t + o(\Delta t)] = k\mu \Delta t + o(\Delta t)$$

V Markovovom modeli sú elementárne náhodné udalosti príchody zákazníkov do systému a uvoľňovanie obsadených liniek. Ukážeme, že pravdepodobnosť nastatia viac ako jednej elementárnej udalosti za časový okamih Δt je veličina rádu $o(\Delta t)$.

Môžu nastať iba tri situácie:

(1) Príchod viac než jedného zákazníka za čas Δt . Vstupný tok je elementárny a tým pádom aj ordinárny. Preto pre pravdepodobnosť príchodu viac než jedného zákazníka za čas Δt je

$$\psi(\Delta t) = o(\Delta t)$$

(2) Uvoľnenie aspoň dvoch z k obsadených liniek. Pravdepodobnosť uvoľnenia linky za čas Δt je $1 - e^{-\mu\Delta t}$ a pravdepodobnosť, že linka za Δt neskončí obsluhu, je $e^{-\mu\Delta t}$. Pravdepodobnosť, že sa z k obsadených liniek práve i uvoľní je rovná súčinu pravdepodobností, že sa uvoľní i liniek,

$$[1 - e^{-\mu\Delta t}]^i$$

pravdepodobnosti, že sa $k - i$ liniek neuvoľní

$$[e^{-\mu\Delta t}]^{k-i}$$

a počtu všetkých možností $\binom{k}{i}$. Preto pravdepodobnosť, že sa aspoň dve linky z k obsadených uvoľní, je

$$\begin{aligned} & \sum i 2k \binom{k}{i} [1 - e^{-\mu\Delta t}]^i e^{-(k-i)\mu\Delta t} = \\ & = \sum i 2k \binom{k}{i} [\mu\Delta t + o(\Delta t)]^i [1 - \mu\Delta t + o(\Delta t)]^{k-i} = \\ & = o(\Delta t) \end{aligned}$$

(3) Naraz vstúpi do systému aspoň jeden zákazník a aspoň jedna z k obsadených liniek sa uvoľní:

$$[1 - e^{-\lambda\Delta t}] \cdot [1 - e^{-k\mu\Delta t}] = [\lambda\Delta t + o(\Delta t)] \cdot [k\mu\Delta t + o(\Delta t)] = o(\Delta t)$$

Vidíme, že v Markovovom modeli pravdepodobnosť nastatia viac ako jednej elementárnej udalosti v priebehu krátkeho časového okamihu Δt je veličina $o(\Delta t)$. Toto je dôležitá vlastnosť, ktorá umožňuje jednoduchú analýzu Markovových modelov.

Popíšeme pravdepodobnosti prechodov medzi stavmi Markovoveho reťazca so spojitým časom, ktoré ďalej použijeme pri analýze konkrétnych Markovových systémov.

Nech sa reťazec nachádza v stave k , resp. nech sa v obslužnom systéme nachádza k zákazníkov. Pretože v Markovovom systéme sa za čas Δt s významnou pravdepodobnosťou (rôznou od Δt) môže udiať iba jedna udalosť, pravdepodobnosť, že sa počet zákazníkov v systéme za čas Δt zvýši práve o jedného sa rovná pravdepodobnosti príchodu práve jedného zákazníka $v_1(\Delta t)$:

$$p_{k,k+1}(\Delta t) = \lambda\Delta t + o(\Delta t)$$

Pravdepodobnosť, že sa z k obsadených liniek aspoň jedna uvoľní je

$$k\mu\Delta t + o(\Delta t)$$

Ako sme už ukázali, uvoľnenie dvoch a viac liniek z k obsadených liniek je veličina rádu $o(\Delta t)$. Pre veličiny rádu $o(\Delta t)$ platí, že rozdiel dvoch veličín rádu $o(\Delta t)$ je opäť veličina rádu $o(\Delta t)$. Preto pravdepodobnosť, že sa z k obsadených liniek uvoľní práve jedna linka je

$$p_{k,k-1}(\Delta t) = k\mu\Delta t + o(\Delta t)$$

Pravdepodobnosť, že v priebehu okamihu Δt nenastane žiadna udalosť, teda žiadny zákazník do systému nevstúpi, ani sa žiadna z k obsadených liniek neuvolní, je

$$p_{k,k}(\Delta t) = [1 - \lambda \Delta t + o(\Delta t)] \cdot [1 - k\mu \Delta t + o(\Delta t)] \Rightarrow$$

$$\Rightarrow \boxed{p_{k,k}(\Delta t) = 1 - (\lambda + k\mu)\Delta t + o(\Delta t)}$$

Pravdepodobnosť prechodu o viac než jeden stav za čas Δt je:

$$p_{i,k}(\Delta t) = o(\Delta t) \quad \text{pre} \quad |i - k| > 1$$

A nakoniec, ak sa systém nachádzal v stave k , po čase Δt sa musí stále nachádzať v jednom zo svojich stavov, preto platí tzv. normovacia podmienka:

$$\sum j p_{k,j}(\Delta t) = 1$$

Pre konkrétny typ Markovovho modelu budú niektoré pravdepodobnosti prechodov v kritických stavoch (napr. stav naplnenia liniek, stav naplnenia fronty) modifikované.

2.2.6 Markovov model

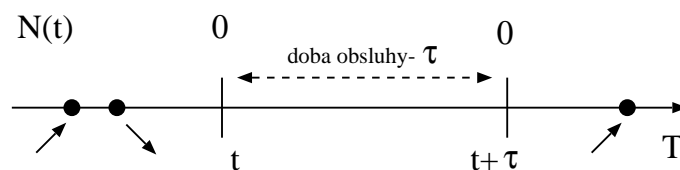
2.2.7 Nemarkovove modely

Nemarkovov model s konštatným vysielaním

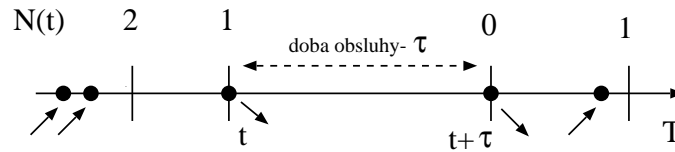
Analýza nemarkovových modelov hromadnej obsluhy, teda modelov, v ktorých buď vstupný tok zákazníkov nie je elementárny, alebo doba obsluhy zákazníkov sa neriadi podľa exponenciálneho rozdelenia, je veľmi náročná a často pomocou analytických metód ani nerealizovateľná, preto sa obmedzíme iba na analýzu niektorých jednolinkových modelov. Komplikovanejšie modely sa pri konkrétnych vstupných parametroch riešia pomocou výpočetnej techniky využitím teórie simulačných metód, to však už presahuje rámec tejto publikácie.

Budeme analyzovať jednolinkový model $M/D/1/\infty$, do ktorého vstupuje elementárny vstupný tok zákazníkov s intenzitou λ a obsluha zákazníkov je nenáhodná, deterministická. To znamená, že linka každého zákazníka obsluží presne za čas τ . Predpokladáme, že systém má neobmedzený počet miest na čakanie vo fronte.

Počet zákazníkov v systéme, resp. stavy systému, popisujeme opäť náhodnými premennými $\{N(t)\}_{t \in T}$. Budeme pozorovať zmeny stavu systému počas doby obsluhy jedného zákazníka. Nech je systém v čase $t + \tau$, kde $t \in R^+$, prázdny. Takýto stav môže nastať len v prípade, keď v čase t bol systém prázdny a počas doby τ (systém nepracoval) nikdo neprišiel,



alebo v čase t bol v systéme jeden zákazník a počas doby obsluhy jedného zákazníka žiaden ďalší zákazník do systému nevstúpil:



Číčky pod grafom znamenajú príchody a odchody zákazníkov do systému, čísla nad grafom označujú počet zákazníkov, ktorí sa momentálne nachádzajú v systéme.

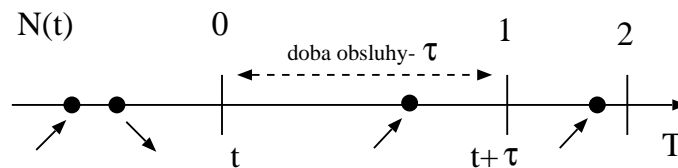
Keby sa systém v čase t nachádzal vo vyššom stave než 1, v čase $t + \tau$, vzhľadom na to, že v systéme je iba jedna linka, by nemohol byť v stave 0, pretože za dobu τ stihne obslúžiť iba jedného zákazníka. Keďže je vstupný tok zákazníkov elementárny, pravdepodobnosť príchodu k zákazníkov za čas τ je

$$v_k(\tau) = \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau}$$

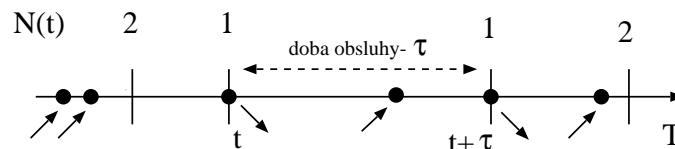
Podľa vety o úplnej pravdepodobnosti dostávame vzťah pre p_0 :

$$p_0(t + \tau) = [p_0(t) + p_1(t)] v_0(\tau)$$

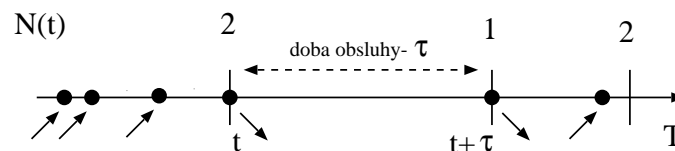
Analogicky, systém sa v čase $t + \tau$ môže nachádzať v stave 1, ak v čase t bol systém prázdny a počas doby τ prišiel do systému zákazník



alebo v čase t už v systéme zákazník bol a v priebehu jeho obsluhy prišiel ďalší zákazník,



alebo sa v systéme v čase t nachádzali dvaja zákazníci a počas doby obsluhy jedného z nich už nikdo neprišiel:



Dostali sme rovnicu

$$p_1(t + \tau) = [p_0(t) + p_1(t)] v_1(\tau) + p_2(t) v_0(\tau)$$

Takýmto istým spôsobom dostávame rovnice aj pre vyššie stavy. Pravdepodobnosti $v_k(\tau)$ nahradíme priamo ich hodnotami, pričom označíme $\lambda\tau = \rho$. Uvedomme si, že systém neodmieta

zákazníkov. Preto ak by do systému vstupovalo viac zákazníkov, než je systém schopný obslúžiť, došlo by k "zahľteniu" systému, fronta by rástla nad všetky medze. Preto, z hľadiska stabilizácie systému, musí byť $\varrho = \lambda\tau < 1$.

$$\begin{aligned} p_0(t + \tau) &= [p_0(t) + p_1(t)] e^{-\varrho} \\ p_1(t + \tau) &= [p_0(t) + p_1(t)] \varrho e^{-\varrho} + p_2(t) e^{-\varrho} \\ &\vdots \\ p_k(t + \tau) &= [p_0(t) + p_1(t)] \frac{\varrho^k}{k!} e^{-\varrho} + \sum_{i=1}^k p_{i+1}(t) \frac{\varrho^{k-i}}{(k-i)!} e^{-\varrho} \\ &\vdots \end{aligned}$$

Predpokladajme, že sa systém po dlhom čase stabilizuje, teda bude stacionárny vzhľadom na čas:

$$\lim tp_k(t) = p_k \quad \text{resp.} \quad \lim tp_k(t + h) = p_k$$

Pravdepodobnosti p_k nazveme pravdepodobnosti stavov stabilizovaného systému. Použitím limitného prechodu dostaneme systém rekurentných homogénnych lineárnych rovníc s normovaciou podmienkou (pre prehľadnosť všetky rovnice vynásobíme konštantou e^ϱ):

$$\begin{aligned} e^\varrho p_0 &= [p_0 + p_1] \\ e^\varrho p_1 &= [p_0 + p_1] \varrho + p_2(t) \\ &\vdots \\ e^\varrho p_k &= [p_0 + p_1] \frac{\varrho^k}{k!} + p_2 \frac{\varrho^{k-1}}{(k-1)!} + \cdots + p_k \varrho + p_{k+1} \\ &\vdots \\ \sum_{k=0}^{\infty} p_k &= 1 \end{aligned}$$

Hoci získanie explicitných tvarov pravdepodobností p_k nie je jednoduché, tieto rekurentné rovnice môžeme použiť pre postupný výpočet pravdepodobností p_k . Stačí si uvedomiť, že pre využitie H jednolinkového systému s nekonečnou frontou platí

$$\boxed{H = ES = \lambda\tau = \alpha = \varrho = 1 - p_0}$$

Systém je nevyužitý iba ak sa nachádza v stave 0, preto $p_0 = 1 - \varrho$. Z predchádzajúceho systému lineárnych rovníc môžeme potom postupne vyjadrovať pravdepodobnosti p_k :

$$\begin{aligned} p_0 &= 1 - \varrho \\ p_1 &= p_0 (e^\varrho - 1) = (1 - \varrho) (e^\varrho - 1) \\ p_2 &= p_1 (e^\varrho - \varrho) - \varrho p_0 = e^\varrho (1 - \varrho) (e^\varrho - \varrho - 1) \\ &\vdots \end{aligned}$$

Pre odvodenie strednej hodnoty EK použijeme vytvárajúcu funkciu $G(x)$ náhodnej premennej K , ktorá popisuje počet zákazníkov v systéme.

$$G(x) = \sum_{k=0}^{\infty} p_k x^k \quad |x| < 1$$

Pomocou tejto vytvárajúcej funkcie môžeme tiež vypočítať pravdepodobnosti stavov p_k ako jej derivácie podľa x v bode nula (pozri kapitolu 1.3.):

$$p_k = \frac{1}{k!} \left[\frac{d}{dx} G(x) \right]_{x=0} \quad k = 0, 1, \dots$$

Pre vyjadrenie vytvárajúcej funkcie $G(x)$ použijeme rekurentné vzťahy pre pravdepodobnosti stavov stabilizovaného systému p_k :

$$\begin{aligned} te^{\varrho} G(x) &= e^{\varrho} p_0 + e^{\varrho} \sum_{k=0}^{\infty} p_k x^k = \\ &= [p_0 + p_1] + \sum_{k=1}^{\infty} \left[(p_0 + p_1) \frac{\varrho^k}{k!} + p_2 \frac{\varrho^{k-1}}{(k-1)!} + \dots \right] x^k = \\ &= \frac{(\varrho x)^k}{k!} + p_1 \sum_{k=0}^{\infty} \frac{(\varrho x)^k}{k!} + \dots + p_{i+1} x^i \sum_{k=i}^{\infty} \frac{(\varrho x)^{k-i}}{(k-i)!} + \dots = \\ &= e^{\varrho x} \left[p_0 + \sum_{k=0}^{\infty} p_{k+1} x^k \right] = \\ &= e^{\varrho x} \left[p_0 + \frac{1}{x} \sum_{k=0}^{\infty} (p_{k+1} x^{k+1} + p_0 - p_0) \right] = \\ &= e^{\varrho x} \left[\frac{1}{x} G(x) - \frac{1-x}{x} p_0 \right] \end{aligned}$$

Dosadíme za $p_0 = 1 - \varrho$ a vyjadríme vzorec pre vytvárajúcu funkciu $G(x)$:

$$G(x) = (1 - \varrho) \frac{1 - x}{1 - x e^{\varrho(1-x)}}$$

Pomocou vytvárajúcej funkcie vypočítame strednú hodnotu EK :

$$\begin{aligned} EK &= \frac{d}{dx} [G(x)]_{x=1} = \frac{d}{dx} (1 - \varrho) \left[\frac{1 - x}{1 - x e^{\varrho(1-x)}} \right]_{x=1} = \\ &= (1 - \varrho) \left[\frac{e^{\varrho(1-x)} (1 - x\varrho + x^2\varrho) - 1}{(1 - x e^{\varrho(1-x)})^2} \right]_{x=1} \end{aligned}$$

Čitateľ ľahko overí, že hodnotu $x = 1$ nemôžeme priamo do výrazu dosadiť. Preto hodnotu EK vypočítame ako limitu pre $x \rightarrow 1$. Keďže máme riešiť limitu typu $\frac{0}{0}$, použijeme L'Hospitalové pravidlo (celkom dva krát):

$$\begin{aligned} EK &= (1 - \varrho) \cdot \lim_{x \rightarrow 1} x \frac{e^{\varrho(1-x)} (1 - x\varrho + x^2\varrho) - 1}{(1 - x e^{\varrho(1-x)})^2} = \\ &= \frac{\varrho(1-\varrho)}{2} \cdot \lim_{x \rightarrow 1} x \frac{x^2\varrho - x(\varrho+2) + 2}{(1-x\varrho) \cdot (1 - x e^{\varrho(1-x)})} = \\ &= \frac{1}{2} \lim_{x \rightarrow 1} x \frac{\varrho(1-\varrho) \cdot (2 + \varrho - 2x\varrho)}{e^{\varrho(1-x)} (1-x\varrho)^2 - \varrho(1 - x e^{\varrho(1-x)})} \Rightarrow EK = \frac{\varrho(2 - \varrho)}{2(1 - \varrho)} \end{aligned}$$

Pomocou EK a $ES = \varrho$ vypočítame strednú dĺžku fronty:

$$EL = EK - ES = \frac{\varrho(2 - \varrho)}{2(1 - \varrho)} - \varrho \Rightarrow EL = \frac{\varrho^2}{2(1 - \varrho)}$$

Toto vyjadrenie strednej dĺžky fronty EL resp. stredného počtu zákazníkov v systéme EK sa nazýva modifikovaná Chinčín-Polaczeková formula. Všimnime si, že ak by bolo využitie systému 100%, teda $H = \varrho = 1$, potom by stredný počet zákazníkov vo fronte, alebo v systéme bol nekonečný a systém by sa preplnil. Pre stredný čas strávený jedným zákazníkom čakaním vo fronte platí rovnako ako pri Markovovom modeli $M/M/n/\infty$:

$$EW = \frac{EL}{\lambda}$$

Určenie kapacity pomocou THO

Nech p_k sú pravdepodobnosti stavov systému. Označíme $\varrho = \lambda/c = \lambda\tau$. Aby sa systém stabilizoval, musí do systému vstupovať v strednom počte menej paketov ako je systém schopný odvyselať, preto $\lambda < c$ resp. $\varrho = \lambda/c < 1$. Pre pravdepodobnosti stavov stabilizovaného systému platia rovnice ([?]):

$$\begin{aligned} p_0 &= 1 - \varrho \\ p_1 &= e^\varrho p_0 - p_0 \\ p_2 &= e^\varrho p_1 - \varrho e^\varrho p_0 \\ &\vdots \\ p_k &= e^\varrho p_{k-1} - \frac{\varrho^{k-1}}{(k-1)!} e^\varrho p_0 - \sum_{i=1}^{k-2} \frac{\varrho^{k-1-i}}{(k-1-i)!} p_{i+1} \quad k = 3, 4, \dots \end{aligned}$$

Pomocou p_k vyjadríme pravdepodobnosť pretečenia buffera

$$p = P(q > n) = \sum_{k=n+1}^{\infty} p_k = 1 - \sum_{k=0}^n p_k$$

Pre hodnoty $n = 1, 2$ dostávame jednoduché vyjadrenie:

$$P(q > 1) = 1 - e^\varrho p_0, \quad P(q > 2) = 1 - [e^{2\varrho} - \varrho e^\varrho] p_0$$

Pre $n = 3, 4, \dots$ je tvar pravdepodobnosti podstatne komplikovanejší

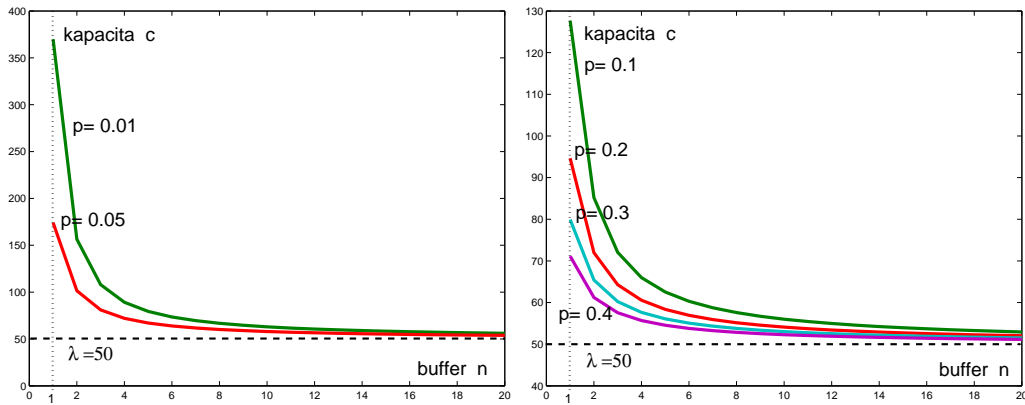
$$P(q > n) = 1 - [e^{2\varrho} - \varrho e^\varrho] p_0 - \sum_{k=3}^n \left[e^\varrho p_{k-1} - \frac{\varrho^{k-1}}{(k-1)!} e^\varrho p_0 - \sum_{i=1}^{k-2} \frac{\varrho^{k-1-i}}{(k-1-i)!} p_{i+1} \right]$$

Vzhľadom na tvar $P(q > n)$ pre "veľké" hodnoty n sme na určenia minimálnej kapacity servera c museli použiť komplikované iteračné algoritmy.

■

Nech v systéme $M/D/1/\infty$ je intenzita vstupného toku $\lambda = 5$ z/hod a konštantná doba obsluhy $\tau = 10$ min/z. Vypočítajte pravdepodobnostné charakteristiky systému.

$n \setminus p$	0.01	0.05	0.1	0.2	0.3	0.4
1	369.94	174.19	127.66	94.64	79.94	71.13
2	156.49	101.53	85.18	71.96	65.41	61.22
3	108.02	81.06	72.05	64.28	60.22	57.55
4	89.16	72.07	65.99	60.55	57.64	55.68
5	79.42	67.09	62.53	58.36	56.09	54.55
6	73.52	63.93	60.30	56.93	55.07	53.80
7	69.58	61.76	58.74	55.91	54.34	53.26
8	66.76	60.17	57.59	55.15	53.79	52.85
9	64.64	58.96	56.71	54.57	53.37	52.54
10	63.00	58.00	56.01	54.10	53.03	52.28

Tabuľka 2.9: Kapacita c pre $M/D/1/\infty$ pomocou THOObr. 2.8: Kapacita c pre $p = 0.01, 0.05$ a $p = 0.1, 0.2, 0.3, 0.4$

Pre využitie systému platí:

$$H = \varrho = \lambda\tau = 5 \cdot \frac{1}{6} = 0.833$$

Vidíme, že $\varrho < 1$, čiže systém sa stabilizuje v čase.

$$ES = \varrho = 1 - p_0 \quad \Rightarrow \quad ES = 0.833 \quad \text{a} \quad p_0 = 0.167$$

Systém bude v priemere $60 \cdot p_0 = 10$ minút z hodiny prázdny a stredný počet zákazníkov, ktorí budú v priebehu hodiny obslužení bez čakania je $\lambda \cdot p_0 = 0.835$ (zákazník musí vstúpiť do systému v okamihu, keď je linka prázdna, v opačnom prípade by totiž čakať vo fronte na obsluhu).

Pre výpočet EL použijeme modifikovanú Chinčin-Polazcekovú formulu:

$$EL = \frac{\varrho^2}{2(1 - \varrho)} = 2.083$$

Vo fronte v priemere čakajú dvaja zákazníci.

$$EK = ES + EL = 2.916$$

Stredný počet zákazníkov v celom systéme je skoro tri.

$$EW = \frac{EL}{\lambda} = \frac{1}{5} \cdot 2.083 = 0.4166$$

Stredná doba, ktorú strávi zákazník čakaním vo fronte je 0.4166 hod. resp. 25 minút.

■

Vieme, že v systéme $M/D/1/\infty$ s deterministickou obsluhou je stredná dĺžka fronty $EL = 8$ zákazníkov čakajúcich vo fronte a konštantná doba obsluhy je $\tau = 5$ min/z. Našou úlohou je vypočítať využitie systému, intenzitu vstupného toku zákazníkov, stredný počet minút z hodiny, kedy je fronta prázdna a pravdepodobnosť, že v počas doby obsluhy jedného zákazníka vstúpia do systému ďalší dvaja zákazníci.

Poznáme stredný počet zákazníkov vo fronte $EL = 8$, preto pre výpočet využitia systému použijeme Chinčín-Polazcekovu formulu:

$$8 = \frac{\varrho^2}{2(1 - \varrho)} \Rightarrow \varrho^2 + 16 \cdot \varrho - 16 = 0 \Rightarrow H = \varrho = 0.9443$$

Systém je využitý na 94,43%. Intenzitu vstupného toku zákazníkov zistíme ľahko:

$$\varrho = \lambda\tau \Rightarrow \lambda = \frac{\varrho}{\tau} = 0.9443 \cdot \frac{1}{5} = 11.33$$

Vstupný tok do nášho systému má intenzitu 11.33 zákazníkov za hodinu. Pomocou rekurentných rovníc pre pravdepodobnosti stavov p_k vypočítame pravdepodobnosť, že je fronta prázdna:

$$p_0 + p_1 = e^{\varrho} p_0 = e^{\varrho}(1 - \varrho) = 0.1432$$

V priebehu hodiny je fronta v priemere prázdna $60 \cdot (p_0 + p_1) = 8.6$ minút.

A dostali sme sa k poslednej otázke, vypočítať pravdepodobnosť vstupu dvoch zákazníkov počas doby obsluhy jedného zákazníka. Doba obsluhy je konštantná, presne $\tau = 5$ minút na každého zákazníka. Vstupný tok zákazníkov do systému je elementárny. V kapitole 2.2. sme označili pravdepodobnosť výskytu k udalostí, ktoré sa riadia podľa elementárne toku, za čas τ ako $v_k(\tau)$:

$$v_k(\tau) = \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} = \frac{\varrho^k}{k!} e^{-\varrho}$$

Preto hľadaná pravdepodobnosť bude $v_2(5 \text{ min})$:

$$v_2(5) = \frac{0.9443^2}{2!} e^{-0.9443} = 0.1734$$

Pravdepodobnosť, že v priebehu obsluhy zákazníka vstúpia do systému ďalší dvaja zákazníci, je 0.1734.

Nemarkovov model s variabilným vysielaním

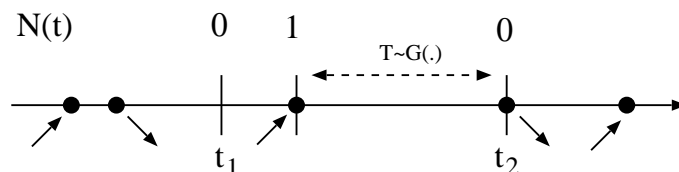
Teraz budeme analyzovať jednolinkový model hromadnej obsluhy s nekonečnou frontou, do ktorého vstupuje elementárny tok zákazníkov s intenzitou λ a dĺžka obsluhy je modelovaná pomocou náhodnej premennej T s ľubovoľným rozdelením pravdepodobnosti, $T \sim G$. Ďalej budeme predpokladať, že T je spojitá náhodná premenná s hustotou rozdelenia pravdepodobnosti $g(\cdot)$. Strednú dobu obsluhy označíme τ a disperziu σ^2 :

$$\tau = \int_0^\infty tg(t)dt \quad \sigma^2 = \int_0^\infty (t - \tau)^2 g(t)dt$$

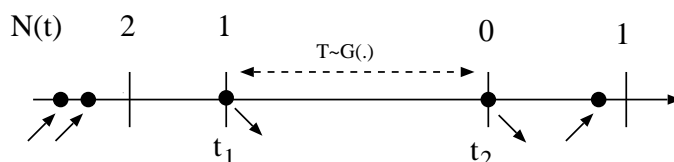
Obslužný systém budeme popisovať náhodným reťazcom $N(t)$ s rozdelením pravdepodobnosti p_k .

Analýzu systému prevedieme obdobne ako pri modeli $M/D/1/\infty$. Označme π_k pravdepodobnosť príchodu k zákazníkov počas doby obsluhy jedného zákazníka. Budeme pozorovať tesne po sebe nasledujúce časové okamihy ukončenia doby obsluhy t_1 a t_2 .

Najprv rozoberieme prípad, keď v čase t_2 bol systém prázdny. Ak v čase t_1 bol systém prázdny, po tomto okamihu prišiel zákazník a počas doby obsluhy tohoto zákazníka už žiadny ďalší do systému nevstúpil.



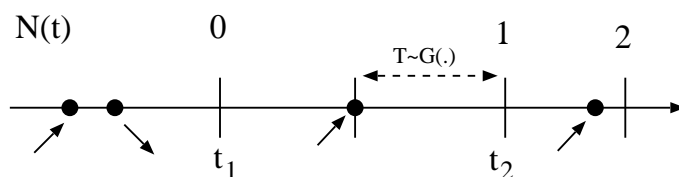
Alebo v čase t_1 bol v systéme jeden zákazník a počas doby obsluhy (do času t_2) už žiaden ďalší zákazník nevstúpil do systému.



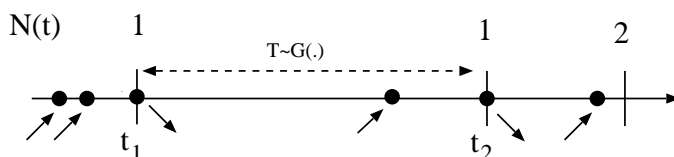
Žiadny iný prípad nastať nemôže, pretože ak by v systéme v čase t_2 boli napr. dvaja zákazníci, do času t_1 systém stihne obslúžiť iba jedného (vyplýva to z definície okamihov t_1 a t_2). Dostali sme teda rovnicu:

$$p_0(t_2) = [p_0(t_1) + p_1(t_1)] \pi_0$$

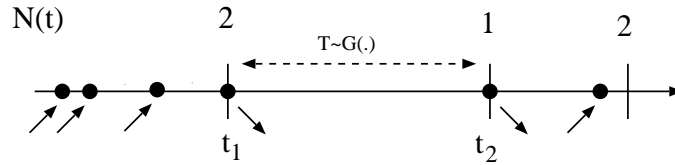
Nech sa teraz nachádza v systéme v čase t_2 jeden zákazník. Ak bol v čase t_1 systém prázdny, prišiel zákazník a počas jeho obsluhovania do systému dorazil práve jeden nový zákazník.



Alebo sa v systéme v čase t_1 už nachádzal zákazník a počas obsluhy prišiel tiež práve jeden zákazník,



alebo boli v čase t_1 v systéme dvaja zákazníci a počas doby obsluhy nikdo neprišiel:



Podľa vety o úplnej pravdepodobnosti potom dostávame vzťah:

$$p_1(t_2) = [p_0(t_1) + p_1(t_1)] \pi_1 + p_2(t_2) \pi_0$$

Tento postup použijeme aj v ostatných stavoch v čase t_2 a dostávame systém rovníc:

$$\begin{aligned} p_0(t_2) &= [p_0(t_1) + p_1(t_1)] \pi_0 \\ p_1(t_2) &= [p_0(t_1) + p_1(t_1)] \pi_1 + p_2(t_2) \pi_0 \\ &\vdots \\ p_k(t_2) &= [p_0(t_1) + p_1(t_1)] \pi_k + \sum_{i=1}^k p_{i+1}(t_1) \pi_{k-i} \\ &\vdots \end{aligned}$$

Teraz odvodíme tvar pravdepodobností π_k . Pravdepodobnosť, že počas doby t vstúpi do systému k zákazníkov, ak výskyt zákazníkov sa riadi podľa elementárneho toku s intenzitou toku λ , je pravdepodobnosť $v_k(t)$. Pravdepodobnosti π_k sú pravdepodobnosti vstupu k zákazníkov do systému počas doby obsluhy jedného zákazníka. Na rozdiel od predchádzajúceho modelu $M/D/1/\infty$ je teraz ale doba obsluhy t realizáciou nejakej spojitej náhodnej premennej T , ktorá má hustotu rozdelenia pravdepodobnosti $g(\cdot)$. Preto pre pravdepodobnosť π_k platí:

$$\pi_k = E[v_k(t)] = \int_0^\infty v_k(t) g(t) dt = \int_0^\infty \frac{(\lambda t)^k}{k!} e^{-\lambda t} g(t) dt$$

Systém je jednolinkový a neodmieta žiadnych zákazníkov. Preto, aby sa systém nepreplnil, nesmie do systému vstupovať viac zákazníkov (v strednej hodnote), než je systém schopný obslužiť. Teda zaťaženie systému je $\alpha < 1$ resp. využitie systému je $H = \rho = \alpha = \lambda \tau < 1$.

Nech teda platí $\rho < 1$ a predpokladáme, že systém sa po dlhom čase prevádzky stabilizuje, stane sa stacionárny vzhľadom na čas. Inými slovami existujú limity:

$$\lim_{t_1} t_1 p_k(t_2) = p_k \quad \lim_{t_2} t_2 p_k(t_1) = p_k$$

Limity p_k nazveme pravdepodobnosti stavov stabilizovaného systému. Sú riešením nasledujúceho systému homogénnych lineárnych rovníc s normovacou podmienkou:

$$\begin{aligned} p_0 &= [p_0 + p_1] \pi_0 \\ p_1 &= [p_0 + p_1] \pi_1 + p_2 \pi_0 \\ &\vdots \\ p_k &= [p_0 + p_1] \pi_k + p_2 \pi_{k-1} + \cdots + p_k \pi_1 + p_{k+1} \pi_0 \\ &\vdots \\ \sum_{k=0}^{\infty} p_k &= 1 \end{aligned}$$

Explicitné vyjadrenie pravdepodobností p_k je náročné a závisí od tvaru hustoty rozdelenia pravdepodobnosti doby obsluhy $g(\cdot)$, resp. od existencie pravdepodobností π_k (uvedomme si, že pre vyjadrenie π_k musíme vypočítať nevlastný integrál).

Rovnako, ako v predchádzajúcich modeloch platí:

$$H = ES = \lambda\tau = \alpha = \varrho = 1 - p_0$$

Vypočítame strednú hodnotu EK . Nech $G(x)$ je vytvárajúca funkcia náhodnej premennej K a $\Pi(x)$ je vytvárajúca funkcia rozdelenia pravdepodobností π_k :

$$G(x) = \sum_{k=0}^{\infty} p_k x^k \quad \Pi(x) = \sum_{k=0}^{\infty} \pi_k x^k \quad |x| < 1$$

Použijeme rekurentné lineárne rovnice pre pravdepodobnosti p_k :

$$\begin{aligned} G(x) &= \sum_{k=0}^{\infty} p_k x^k = \\ &= [p_0 + p_1] \pi_0 + \cdots + [(p_0 + p_1) \pi_k + \cdots + p_{k+1} \pi_0] x^k + \cdots = \\ &= p_0 \cdot \sum_{k=0}^{\infty} \pi_k x^k + p_1 \cdot \sum_{k=0}^{\infty} \pi_k x^k + \cdots + p_k x^{k-1} \cdot \sum_{k=0}^{\infty} \pi_k x^k + \cdots = \\ &= \Pi(x) [p_0 + p_1 + p_2 x + \cdots + p_k x^{k-1} + \cdots] = \\ &= \frac{\Pi(x)}{x} [p_0 x - p_0 + p_0 + p_1 x + \cdots + p_k x^k + \cdots] \end{aligned}$$

Súčet na pravej strane napíšeme ako vytvárajúcu funkciu $G(x)$ a ďalej použijeme vzťah $p_0 = 1 - \varrho$. Dostaneme rovnicu pre vytvárajúcu funkciu, ktorú jednoduchými úpravami vyriešime:

$$G(x) = \frac{\Pi(x)}{x} [(1 - \varrho)(x - 1) + G(x)] \Rightarrow G(x) = (1 - \varrho) \cdot \frac{(x - 1)\Pi(x)}{x - \Pi(x)}$$

Strednú hodnotu EK vypočítame ako prvú deriváciu $G(x)$ v bode $x = 1$:

$$\begin{aligned} EK &= \frac{d}{dx} [G(x)]_{x=1} = \\ &= (1 - \varrho) \cdot \frac{d}{dx} \left[\frac{(x-1)\Pi(x)}{x-\Pi(x)} \right]_{x=1} = \\ &= (1 - \varrho) \cdot \left[\frac{x^2 \Pi'(x) - x \Pi'(x) - \Pi(x)^2 + \Pi(x)}{[x - \Pi(x)]^2} \right]_{x=1} \end{aligned}$$

Upravíme vytvárajúcu funkciu $\Pi(x)$. Pre stručnosť budeme predpokladať, že jednotlivé pravdepodobnosti π_k existujú a zámenny poradia integrovania, derivovania a sčítania môžeme previesť.

$$\begin{aligned} \Pi(x) &= \sum_{k=0}^{\infty} \pi_k x^k = \sum_{k=0}^{\infty} \left[\int_0^{\infty} v_k(t) g(t) dt \right] \cdot x^k = \\ &= \int_0^{\infty} \sum_{k=0}^{\infty} v_k(t) x^k g(t) dt = \int_0^{\infty} e^{\lambda t(x-1)} g(t) dt \end{aligned}$$

Použili sme vytvárajúcu funkciu elementárneho toku (pozri kapitolu 2.2.):

$$F(x, t) = \sum_{k=0}^{\infty} v_k(t) x^k = e^{\lambda t(x-1)}$$

Teraz vypočítame hodnoty funkcie $\Pi(x)$ a jej derivácie $\Pi'(x)$ v bode $x = 1$:

$$\begin{aligned}\Pi(1) &= \lim_{x \rightarrow 1} x \int_0^{\infty} e^{\lambda t(x-1)} g(t) dt = \int_0^{\infty} g(t) dt = 1 \\ \Pi'(1) &= \lim_{x \rightarrow 1} x \frac{d}{dx} \int_0^{\infty} e^{\lambda t(x-1)} g(t) dt = \lim_{x \rightarrow 1} x \int_0^{\infty} \lambda t e^{\lambda t(x-1)} g(t) dt = \\ &= \int_0^{\infty} \lambda t g(t) dt = \lambda \tau = \varrho\end{aligned}$$

Vidíme, že ak by sme do výrazu pre strednú hodnotu EK dosadili $x = 1$, dostali by sme výraz typu $\frac{0}{0}$. Preto hodnotu EK vypočítame ako limitu daného výrazu, pričom použijeme dva krát L'Hospitalove pravidlo (čitateľ tento výpočet ľahko prevedie sám). Potom pre strednú hodnotu EK dostávame:

$$EK = \frac{1 - \varrho}{2} \cdot \frac{\Pi''(1) + 2\Pi'(1)[1 - \Pi'(1)]}{[1 - \Pi'(1)]^2}$$

Vieme, že $\Pi'(1) = \varrho$. Zostáva ešte vypočítať hodnotu $\Pi''(1)$:

$$\begin{aligned}\Pi''(1) &= \lim_{x \rightarrow 1} x \frac{d^2}{dx^2} \int_0^{\infty} e^{\lambda t(x-1)} g(t) dt = \lim_{x \rightarrow 1} x \int_0^{\infty} (\lambda t)^2 e^{\lambda t(x-1)} g(t) dt = \\ &= \lambda^2 \int_0^{\infty} t^2 g(t) dt = \lambda^2 E(T^2)\end{aligned}$$

Z elementárnej teórie pravdepodobnosti vieme, že platí:

$$DT = E(T^2) - (ET)^2 \quad \Rightarrow \quad \Pi''(1) = \lambda^2 E(T^2) = \lambda^2(\tau^2 + \sigma^2)$$

Dosadením do výrazu pre EK dostávame Chinčín-Polaczekovú formulu:

$$EK = \frac{1 - \varrho}{2} \cdot \frac{\lambda^2(\tau^2 + \sigma^2) + 2\varrho(1 - \varrho)}{(1 - \varrho)^2} \quad \Rightarrow \quad \boxed{EK = \frac{\lambda^2(\tau^2 + \sigma^2)}{2(1 - \varrho)} + \varrho}$$

Vieme, že $ES = \varrho$, preto pre strednú dĺžku fronty EL a stredný čas EW dostaneme:

$$\boxed{EL = \frac{\lambda^2(\tau^2 + \sigma^2)}{2(1 - \varrho)}} \quad \Rightarrow \quad EW = \frac{EL}{\lambda} = \frac{\lambda(\tau^2 + \sigma^2)}{2(1 - \varrho)}$$

Odvodili sme strednú dĺžku fronty pre ľubovoľný model $M/G/1/\infty$, kde G predstavuje ľubovoľné spojité rozdelenie pravdepodobnosti trvania doby obsluhy. Ak by sme použili definovali strednú hodnotu ako Liebesque-Stieltesov integrál (pozri kapitolu 1.1.)

$$\pi_k = E[v_k(t)] = \int_R v_k(t) dF(t)$$

kde $F(t)$ je distribučná funkcia rozdelenia doby obsluhy, dokázali by sme obdobným spôsobom platnosť Chinčín-Polaczekovej formuly aj pre modely s diskretným rozdelením doby obsluhy.

Ak je doba obsluhy deterministická, teda ak sa jedná o obslužný systém $M/D/1/\infty$, potom je rozptyl nulový. Ak dosadíme $\sigma^2 = 0$ do Chinčín-Polaczekovej formuly, dostaneme jej modifikovaný tvar, ktorý sme odvodili v kapitole 5.1.

Ak uvažujeme o Markovovom modeli $M/M/1/\infty$, potom $\tau = \frac{1}{\mu}$ a $\sigma^2 = \frac{1}{\mu^2}$. Ak tieto hodnoty dosadíme do Chinčín-Polaczekovej formuly, resp. do vzorca pre EL , opäť dostaneme výsledok, ktorý už poznáme z Teórie Markovových modelov.

■

Vrátime sa k príkladu 6.1.1. s tým rozdielom, že teraz budeme predpokladať, že doba obsluhy je náhodná a správa sa podľa normálneho rozdelenia $T \sim N(\tau, \sigma^2)$. Rovnako, ako v predchádzajúcom príklade, je intenzita vstupného toku $\lambda = 5$ z/hod a stredná doba obsluhy $\tau = 10$ min/z. Nech stredná odchýlka obsluhy na jedného zákazníka je $\sigma = 4$ min/z. Porovnáme, ako sa zmení stredná dĺžka fronty a stredný prestoj jedného zákazníka vo fronte oproti príkladu 6.1.1.

Využitie systému H a stredná hodnota ES zostanú nezmenené, pretože nezávisia od rozptylu obsluhy:

$$H = ES = \rho = \lambda\tau = 0.833$$

Dosadíme do Chinčín-Polaczekovej formuly pre strednú dĺžku fronty hodnoty $\lambda = \frac{1}{12}$ z/min, $\tau = 10$ min/z a $\sigma^2 = 16$ min²/z:

$$EL = \frac{\lambda^2(\tau^2 + \sigma^2)}{2(1 - \rho)} = \frac{10^2 + 16}{2 \cdot 12(1 - 0.833)} = 2.417$$

Zistili sme, že oproti deterministickej obsluhu sa pri modelovaní obsluhy pomocou normálneho rozdelenia $N(\tau, \sigma^2)$ stredná dĺžka fronty zväčší.

$$EW = \frac{EL}{\lambda} = \frac{2.417}{5} = 0.4834$$

Stredný prestoj jedného zákazníka, ktorý vznikne čakaním vo fronte je 0.483 hodiny, resp. 29.0 minút.

■

Našou úlohou bude vypočítať strednú dĺžku fronty pre Erlangov model $M/E_n/1/\infty$, kde $E_n = E(n, b)$ je Erlangovo rozdelenie, a porovnať so strednými dĺžkami fronty v modeloch $M/M/1/\infty$ a $M/D/1/\infty$.

Pre náhodnú premennú T , ktorá má Erlangové rozdelenie $E(n, b)$ platí:

$$\tau = ET = \frac{n}{b} \quad \sigma^2 = DT = \frac{n}{b^2}$$

Tieto hodnoty dosadíme do Chinčín-Polaczekovej formuly pre strednú dĺžku fronty EL :

$$EL = \frac{\lambda^2 \left(\frac{n^2}{b^2} + \frac{n}{b^2} \right)}{2(1 - \rho)} = \frac{\rho^2 + \frac{\rho^2}{n}}{2(1 - \rho)} = \frac{n+1}{n} \cdot \frac{\rho^2}{2(1 - \rho)} \quad \text{pretože} \quad \rho = \lambda \cdot \frac{n}{b}$$

Upravíme vzorec pre strednú dĺžku fronty EL v Markovovom jednolinkovom modeli s nekonečnou frontou $M/M/1/\infty$:

$$EL = \frac{n^n}{n!} \cdot \frac{\varrho^{n+1}}{(1-\varrho)^2} p_0 = \frac{\varrho^2}{(1-\varrho)^2} (1-\varrho) = \frac{\varrho^2}{1-\varrho}$$

Vzorec pre strednú dĺžku fronty EL pre model $M/D/1/\infty$ sme odvodili v predchádzajúcej kapitole.

obslužný model	stredná dĺžka vrnty
$M/M/1/\infty$	$EL = \frac{\varrho^2}{1-\varrho}$
$M/E_n/1/\infty$	$EL = \frac{n+1}{n} \cdot \frac{\varrho^2}{2(1-\varrho)}$
$M/D/1/\infty$	$EL = \frac{\varrho^2}{2(1-\varrho)}$

Ak by sme vo vzorci pre Erlangov model dosadili $n = 1$, dostaneme vzorec pre model Markovov, čo je prirodzené. Stredná dĺžka fronty pri modeli s deterministickou obsluhou je polovičná, než pri modeli Markovovom. Exponenciálne rozdelenie modeluje "najchaotickejšiu obsluhu zákazníkov. So zvyšujúcim sa parametrom n Erlangovho rozdelenia, sa bude obsluha zákazníkov správať "slušnejšie a stredná dĺžka fronty sa začne znižovať, až limitne, pre $n \rightarrow \infty$, dosiahne dĺžku fronty pre model s deterministickou obsluhou.

■

Vrátime sa k príkladom 2.5.1. a 2.5.2. Na základe nameraných resp. pozorovaných dôb trvania obsluhy zákazníkov sme vyberali vhodné Erlangove rozdelenie pre model obsluhy. Zaoberali sme sa len jednou linkou. V prvom prípade bola stredná doba obsluhy $\tau = 1$ min/z a smerodajná odchýlka $\sigma = 20$ sek/z. V druhom prípade sa obsluha zákazníkov spresnila, teda $\tau = 1$ min/z a $\sigma = 5$ sek/z. V prvom prípade sme dobu obsluhy modelovali rozdelením $E(9; 0.15 \text{ z/sek})$, v druhom $E(144; 2.4 \text{ z/sek})$. Budeme predpokladať, že linka je zapojená v modeli $M/E_n/1/\infty$ a využitie systému je 95%. Našou úlohou bude zistiť zmeny pravdepodobnostných charakteristík modelu pred a po spresnení obsluhy.

Pre názornosť si pod linkou môžeme predstaviť pokladničku v predajni potravín. Spresnenie doby obsluhy predstavuje získania skúsenosti pokladničky (samozrejme je príklad silne idealizovaný, v realite by po spresnení práce došlo aj k skráteniu strednej doby obsluhy). Využitie $H = 95\%$ znamená, že pokladnička v priemere $60p_0 = 60 \cdot 0.05 = 3$ minúty z hodiny nemá koho obsluhovať.

Postupne vypočítame strednú dĺžku fronty pre prvý a druhý model:

$$EL_1 = \frac{10}{9} \cdot \frac{0.95^2}{2(1-0.95)} = 10.027$$

$$EL_2 = \frac{145}{144} \cdot \frac{0.95^2}{2(1-0.95)} = 9.087$$

Po spresnení práce sa stredná dĺžka fronty skrátila skoro o jedného zákazníka.

$$EW_1 = \frac{EL_1}{\lambda} = EL_1 \cdot \frac{\tau}{\varrho} = \frac{10.027}{57} = 0.176$$

$$EW_2 = \frac{EL_2}{\lambda} = EL_2 \cdot \frac{\tau}{\varrho} = \frac{9.0876}{57} = 0.159$$

Pokladnička v oboch prípadoch za hodinu obsluži v priemere 57 ľudí:

$$\varrho = \lambda \cdot \tau \Rightarrow \lambda = \frac{\varrho}{\tau} = 0.95 \cdot 60 = 57$$

Stredný čas strávený zákazníkom čakaním vo fronte je v prvom prípade $60 \cdot 0.176 = 10.56$ minút, po spresnení obsluhy sa táto hodnota zníži na $60 \cdot 0.159 = 9.54$ minút. Zníženie rozptylu obsluhy pri zachovaní ostatných parametrov systému nemá vplyv na využitie systému H a na stredný počet obsadených liniek ES .

■

V niektorých praktických úlohách má dôležitý význam použitie zovšeobecneného exponenciálneho rozdelenia (pozri kapitolu 2.5.) pre model doby obsluhy zákazníkov. Je to v príkladoch, kde predpokladáme, že obsluha trvá nejakú minimálnu dobu a po prekročení tejto doby sa začne správať náhodne.

Majme systém $M/G/1/\infty$. Nech v tomto systéme trvá minimálna doba obsluhy zákazníka $A = 10$ min/z. Po uplynutí tejto doby sa predĺženie obsluhy chová exponenciálne so strednou hodnotou $\frac{1}{\mu} = 5$ min/z. Intenzita vstupného toku je $\lambda = 3$ z/hod. Takýto systém označíme $M/M_A/1/\infty$, kde M_A znamená, že doba obsluhy sa riadi podľa zovšeobecneného exponenciálneho rozdelenia $T \sim \exp(A, \mu)$.

Vypočítajte strednú dĺžku fronty EL a pravdepodobnosť, že je fronta v systéme prázdna. Porovnajte výsledky s Markovovým jednolinkovým modelom $M/M/1/\infty$ so strednou dobou obsluhy $\frac{1}{\mu_0} = 15$ min/z.

Pre obidva modely vypočítame strednú dĺžku fronty. Pre zovšeobecnené exponenciálne rozdelenie $\exp(A, \mu)$ platí

$$\tau = A + \frac{1}{\mu} = 15 \text{ min/z} = \frac{1}{4} \text{ hod/z}; \quad \sigma^2 = \frac{1}{\mu^2} = 5^2 \text{ min}^2/\text{z} = \frac{1}{144} \text{ hod}^2/\text{z}$$

Hodnoty dosadíme do Chinčín-Polaczekovej formuly:

$$EL = \frac{\lambda^2(\tau^2 + \sigma^2)}{2(1 - \lambda\tau)} = \frac{9 \left[\frac{1}{16} + \frac{1}{144} \right]}{2 \left[1 - \frac{3}{4} \right]} = 1.25$$

Pre porovnanie vypočítame EL pre model $M/M/1/\infty$:

$$EL = \frac{\frac{\lambda^2}{\mu_0^2}}{1 - \frac{\lambda}{\mu_0}} = \frac{\frac{3^2}{4^2}}{1 - \frac{3}{4}} = 2.25$$

V Markovovom modele je stredná dĺžka fronty podstatne dlhšia. Je to spôsobené tým, že v modele $M/M_A/1/\infty$ pracuje linka menej "chaoticky". Každá obsluha trvá minimálne 10 minút a až potom sa k nej pridajú náhodné doby, ktoré sa správajú podľa exponenciálneho rozdelenia so strednou hodnotou 5 minút. Dlhá obsluha o dobe trvania T sa v exponenciálnom rozdelení so strednou hodnotou 15 minút vyskytne častejšie, než obsluha s dĺžkou trvania $T-A = T-10$ minút v exponenciálnom rozdelení so strednou hodnotou 5 minút, pretože hustota rozdelenia $\exp(\frac{1}{15})$ klesá k nule pomalšie, než hustota rozdelenia $\exp(\frac{1}{5})$.

Teraz vyriešime druhú úlohu. Pravdepodobnosť π_0 vypočítame priamo z jej definície:

$$\begin{aligned} \pi_0 &= \int_0^\infty v_0(t)g(t)dt = \int_0^\infty e^{-\lambda t} \mu e^{-\mu(t-A)} dt = \\ &= \mu e^{\mu A} \cdot \left[\frac{-e^{-t(\lambda+\mu)}}{\lambda+\mu} \right]_A^\infty = \frac{\mu e^{-\lambda A}}{\lambda+\mu} \end{aligned}$$

Pravdepodobnosť, že fronta v systéme je prázdna je:

$$p_0 + p_1 = \frac{p_0}{\pi_0} = \frac{1 - \varrho}{\pi_0} = \frac{\left[1 - \lambda \left(A + \frac{1}{\mu} \right) \right] (\lambda + \mu)}{\mu e^{-\lambda A}} = 0.5152$$

Pre porovnanie, v Markovovom modele je táto pravdepodobnosť menšia:

$$p_0 + p_1 = p_0 + \varrho p_0 = (1 - \varrho) \cdot (1 + \varrho) = 1 - \varrho^2 = 0.4375$$

■

Intenzita vstupného toku zákazníkov do systému hromadnej obsluhy je $\lambda = 6$ z/hod. Z tohoto počtu štyria zákazníci vyžadujú obsluhu presne 6 minút a zvyšní dvaja presne 12 minút. Systém je s nekonečnou frontou a je v ňom zapojená jedna obslužná linka. Určte strednú dĺžku fronty čakajúcich zákazníkov.

Ako sme už spomenuli, Chinčín-Polazcekovú formulu môžeme použiť aj v prípade, keď náhodná premenná, ktorá popisuje dobu obsluhy, je diskretná. V našom prípade náhodná premenná T môže nadobúdať len dve hodnoty s rozdelením pravdepodobnosti:

$$P(T = 6 \text{ min/z}) = \frac{4}{6} \quad \text{a} \quad P(T = 12 \text{ min/z}) = \frac{2}{6}$$

Vypočítame strednú hodnotu a disperziu obsluhy:

$$\begin{aligned} \tau &= ET = 6 \cdot \frac{4}{6} + 12 \cdot \frac{2}{6} = 8 \\ \sigma^2 &= DT = (6 - \tau)^2 \cdot \frac{4}{6} + (12 - \tau)^2 \cdot \frac{2}{6} = 13.333 \end{aligned}$$

Zostáva nám už len dosadiť do vzorca pre EL :

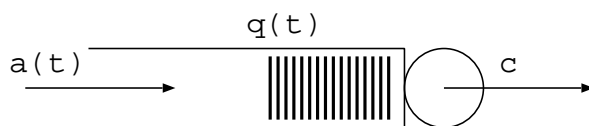
$$EL = \frac{\lambda^2(\tau^2 + \sigma^2)}{2(1 - \lambda\tau)} = \frac{0.1^2(8^2 + 13.333)}{2(1 - 0.1 \cdot 8)} = 1.933$$

V takomto systéme vo fronte čakajú v priemere skoro dvaja zákazníci.

2.3 POUŽITIE PRINCÍPU VEĽKÝCH ODCHÝLOK

2.3.1 Systém s konštantnou kapacitou

Princíp veľkých odchýliek použijeme na určenie distribúcie (rozdelenia pravdepodobnosti) náhodnej premennej q , ktorá popisuje stav fronty v jednolinkovom systéme s konštantnou kapacitou c (*Single Server Queue*, *Work Conserving Link*, *CBR - Constant Bit Rate, Shaper*). V časovom slotu t do systému vstúpi $a(t)$ paketov (rámcov, ATM buniek, kilobitov, ...) a systém je schopný v tom istom slotu vyslať c paketov.



Obr. 2.9: Jednolinkový systém s konštantnou kapacitou

Náhodná premenná $q(t)$ popisuje počet čakajúcich paketov vo fronte v čase t . Predpokladáme, že prevádzka systému spĺňa **Lindleyho rovnicu**:

$$q(t) = \max\{0, q(t-1) + a(t) - c\} = (q(t-1) + a(t) - c)^+ \quad (2.14)$$

Vstupným procesom (*Arrival process*) nazveme náhodný proces $A(t)$, ktorý popisuje kumulatívny počet vstúpivších paketov za časový interval $\langle 0, t \rangle$:

$$A(t) = \sum_{i=1}^t a(i), \quad \text{resp.} \quad a(t) = A(t) - A(t-1), \quad A(0) = 0 \quad (2.15)$$

Ak náhodné premenné $a(t)$ sú nezávislé a rovnako rozdelené (*i.i.d - independent identical distribution*), hovoríme, že vstupný proces má stacionárne a nezávislé prírastky.

Prevedieme niekoľko iterácií Lindleyho rovnice:

$$\begin{aligned} q(t) &= \max\{0, q(t-1) + a(t) - c\} = \\ &= \max\{0, \max\{0, q(t-2) + a(t-1) - c\} + a(t) - c\} = \\ &= \max\{0, a(t) - c, q(t-2) + a(t) + a(t-1) - 2c\} = \\ &= \max\{0, a(t) - c, a(t) + a(t-1) - 2c, \\ &\quad q(t-3) + a(t) + a(t-1) + a(t-2) - 3c\} \end{aligned}$$

Za predpokladu $q(0) = 0$ môžeme napísať:

$$q(t) = \max_{0 \leq s < t} \left\{ 0, \sum_{i=s+1}^t (a(i) - c) \right\} = \max_{0 \leq s \leq t} \{A(t) - A(s) - (t-s)c\} \quad (2.16)$$

Korektný dôkaz môžeme previesť pomocou Matematickej Indukcie:

$$\begin{aligned}
 t = 1 : \quad q(1) &= \max\{0, q(0) + a(1) - c\} = \max\{0, a(1) - c\} = \\
 &= \max_{0 \leq s \leq 1} \{A(1) - A(s) - (1-s)c\} \\
 t + 1 : \quad q(t+1) &= \max\{0, q(t) + a(t+1) - c\} = \\
 &= \max\{0, \max_{0 \leq s \leq t} \{A(t) - A(s) - (t-s)c\} + a(t+1) - c\} = \\
 &= \max\{0, \max_{0 \leq s \leq t} \{A(t+1) - A(s) - (t+1-s)c\}\} = \\
 &= \max_{0 \leq s \leq t+1} \{A(t+1) - A(s) - (t+1-s)c\} \quad \square
 \end{aligned}$$

Dĺžka frontu, resp. počet paketov ktoré zostali v systéme v čase t závisí od rozdielov $a(i) - c$. Hodnota tohto rozdielu predstavuje istú záťaž pre systém. Zavedieme nové náhodné premenné $Z_{t-i+1} = a(i) - c$. Náhodný proces $\{W_t\}_{t \geq 0}$ nazveme záťažový proces (*Workload process*), ak platí

$$W_t = Z_1 + \dots + Z_t \quad \text{pre } t > 0, \quad \text{a} \quad W_0 = 0 \quad (2.17)$$

Pre dĺžku frontu v čase t dostávame vzťah:

$$q(t) = \max\{0, Z_1, Z_1 + Z_2, \dots, Z_1 + \dots + Z_t\} \Rightarrow q(t) = \max_{0 \leq i \leq t} \{W_i\}$$

Predpokladali sme, že náhodné premenné $a(i)$ sú stacionárne a nezávislé (i.i.d.). Potom aj Z_i sú stacionárne a nezávislé. Ak navyše platí $\forall t, E a(t) < c$ ($\forall i, Z_i < 0$), potom dĺžka frontu v čase t konverguje podľa rozdelenia ku konečnej náhodnej premennej q (pozri [?]).

$$q = \lim_{t \rightarrow \infty} q(t) = \lim_{t \rightarrow \infty} \max_{0 \leq i \leq t} \{W_i\} \Rightarrow q = \max_{t \geq 0} \{W_t\} \quad (2.18)$$

Z vety ?? vieme, že pre $\forall n > E W_t$ platí $P(W_t > n) \leq e^{-tI(n/t)}$. Keďže $E Z_i < 0$, potom aj $E W_t < 0$ a tento vzťah môžeme využiť pre všetky $n \in N$. Ďalej sme definovali $W_0 = 0$ a preto $P(W_0 > n) = 0$ pre $n > 0$:

$$P(q > n) = P(\max_{t \geq 0} \{W_t\} > n) \leq \sum_{t \geq 1} P(W_t > n) \leq \sum_{t \geq 1} e^{-tI(n/t)} \quad (2.19)$$

Výraz $-tI(n/t)$ upravíme na $-n \frac{I(n/t)}{n/t}$. Ak je n dostatočne "veľké", člen ktorý "prevládne" nad ostatnými členmi súčtu je ten, pre ktorý je výraz $\frac{I(n/t)}{n/t}$ najmenší (pozri [?]), preto pravdepodobnosť prekročenia danej dĺžky frontu n zaniká exponenciálne v závislosti od n a s tzv. spádovou konštantou δ :

$$P(q > n) \asymp e^{-n\delta}, \quad \delta = \min_x \frac{I(x)}{x} \quad (2.20)$$

Vyjadríme spádovú konštantu δ pomocou CGF. Nech $\theta \leq \delta$, potom

$$\begin{aligned}
 \theta \leq \min_x \frac{I(x)}{x} &\iff \forall x; \theta \leq \frac{I(x)}{x} \iff \forall x; \theta x - I(x) \leq 0 \iff \\
 &\iff \sup_x \{\theta x - I(x)\} \leq 0 \iff \lambda(\theta) \leq 0
 \end{aligned}$$

Dostali sme ekvivalenciu $\theta \leq \delta \Leftrightarrow \lambda(\theta) \leq 0$ a druhý spôsob výpočtu δ :

$$P(q > n) \asymp e^{-n\delta}, \quad \delta = \max\{ \theta ; \lambda(\theta) \leq 0 \} \quad (2.21)$$

Funkciu $\lambda(\theta)$ (CGF premenných Z_j) budeme tiež nazývať **škálová kumulatívna vytvárajúca funkcia** (*SCGF - Scaled Cumulant Generation function*) záťažového procesu W_t , pretože platí

$$\begin{aligned} E[e^{\theta W_t}] &= E[e^{\theta(Z_1 + \dots + Z_t)}] = \prod_{k=1}^t E[e^{\theta Z_k}] = e^{t\lambda(\theta)} \Rightarrow \\ \Rightarrow \lambda(\theta) &= \frac{1}{t} \ln E[e^{\theta W_t}] \end{aligned} \quad (2.22)$$

Získali sme dva spôsoby výpočtu spádovej konštanty, pomocou rýchlostnej funkcie $I(x)$ náhodných premenných Z_j , vzťah (2.20), alebo pomocou SCGF $\lambda(\theta)$ záťažového procesu W_t , vzťah (2.21). Ak má vstupný proces $A(t)$ i.i.d. prírastky, výpočet spádovej konštanty môžeme zjednodušiť. CGF prírastkov $a(i)$ označíme $\lambda_a(\theta)$. Vieme, že $Z_{t-i+1} = a(i) - c$. Bez straty na všeobecnosti preznačíme $j = t - i + 1$ a využijeme stacionárnosť:

$$\begin{aligned} \varphi(\theta) &= E[e^{\theta Z_j}] = E[e^{\theta(a(i)-c)}] = e^{-\theta c} E[e^{\theta a(i)}] = e^{-\theta c} \varphi_a(\theta) \Rightarrow \\ \Rightarrow \lambda(\theta) &= \ln \varphi(\theta) = \ln e^{-\theta c} + \ln \varphi_a(\theta) = \lambda_a(\theta) - \theta c \end{aligned}$$

Preto pre spádovú konštantu δ podľa vzťahu (2.21) platí:

$$\delta = \max\{ \theta ; \lambda(\theta) \leq 0 \} \Rightarrow \delta = \max\{ \theta ; \lambda_a(\theta) \leq \theta c \}$$

Funkcia $\lambda_a(\theta)$ je konvexná a $\lambda_a(0) = 0$. Pre $\theta > 0$ sú obidve funkcie $\lambda_a(\theta)$ a θc rastúce a $\theta = 0$ je nulovým riešením rovnice $\lambda_a(\theta) = \theta c$. To znamená, že existuje nenulové riešenie δ , ktoré je zároveň maximom hodnôt θ , pre ktoré platí že funkcia θc je väčšia nanaajvýš rovná funkcii $\lambda_a(\theta)$ Práve sme dokázali tvrdenie nasledujúcej vety:

Veta 2.3.1. *Nech jednolinkový systém má konštantnú kapacitu c a vstupný proces $A(t)$ s i.i.d. prírastkami $a(t)$. Nech funkcia $\lambda_a(\theta)$ je CGF prírastkov $a(t)$. Spádová konštantu δ je nenulovým riešením rovnice*

$$\lambda_a(\theta) = \theta c \quad (2.23)$$

Poznámka: Tvrdenie vety 2.3.1. hovorí, že ak vstupný proces $A(t)$ do jednolinkového systému má stacionárne a i.i.d. prírastky $a(i)$ s CGF $\lambda_a(\theta)$ a kapacita systému spĺňa podmienku $c = \lambda_a(\delta)/\delta$, potom pravdepodobnosť, že veľkosť fronty prekročí nejakú hodnotu n , $P(q > n)$, zaniká exponenciálne vzhľadom na n so spádovou konštantou δ :

$$P(q > n) \leq e^{-n\delta} \iff c = \frac{\lambda_a(\delta)}{\delta} = \frac{1}{\delta} E[e^{\delta a(i)}] \quad (2.24)$$

Nerovnosť $P(q > n) \leq e^{-n\delta}$ nazveme δ -kritérium QoS. Všetky predchádzajúce tvrdenia vedú k definícii veľmi dôležitého pojmu, **efektívna šírka pásma** (*Effective Bandwidth*), ktorú sformulujeme v nasledujúcej kapitole.

Veta 2.3.2. ($GI/D/1/\infty$)

Nech vstupný proces $A(t) = \sum_{i=1}^t a(i)$ do jednolinkového systému s konštantnou kapacitou c a s dĺžkou frontu q má stacionárne, ergodické a i.i.d. prírastky. Nech jeho efektívna šírka pásma je $\alpha(\theta, t)$. Potom platí

$$\alpha(\delta, t) = c \iff \forall n \in N, \quad P(q > n) \asymp e^{-n\delta} \quad (2.25)$$

Dôkaz. Z vlastnosti (ii) efektívnej šírky pásma dostávame

$$\alpha(\theta, t) = c \iff \frac{\lambda_a(\theta)}{\theta} = c \iff \lambda_a(\theta) = \theta c$$

Z vety 2.3.1 vyplýva:

$$P(q > n) \asymp e^{-n\delta} \quad \text{kde } \delta \text{ je riešením rovnice } \lambda_a(\theta) = \theta c \quad \square$$

□

2.3.2 Dimenzovanie $Geo/D/1/\infty$

Model $Geo/D/1/\infty$ popisuje podľa Kendallovej klasifikácie systém s jednou linkou a nekonečným frontom, pričom doba vysielania τ , resp. kapacita c , je konštantná (D), a do systému vstupuje Bernoulliho proces (Geo) s pravdepodobnosťou výskytu paketu v časovom slotu p . Intenzita toku za časový interval dĺžky t je potom pt .

Na dimenzovanie systému použijeme vzťah (2.24):

$$P(q > n) = e^{-n\delta} \iff c = \alpha(\theta) = \frac{1}{\theta} \ln [1 + p(e^\theta - 1)] \quad (2.26)$$

2.3.3 $M/D/1/\infty$

Nech do LBS vstupuje Poissonov tok rečových paketov s intenzitou $\lambda = 50$ p/s a doba vysielania paketov je konštantná, τ . Takýto systém sa označuje $M/D/1/\infty$. Aby sme mohli porovnať prístup pomocou klasickej teórie hromadnej obsluhy a prístup založený na princípe veľkých odchýlok, budeme za dĺžku frontu považovať celkový počet paketov v systéme. Určíme kapacitu tohto systému $c = 1/\tau$ pre danú pravdepodobnosť prekročenia určitej kritickej hodnoty $p = P(q > n)$.

Vstupný proces $A(t)$ do $M/D/1/\infty$ je Poissonov, $A(t) \sim Po(\lambda t)$, so stacionárnymi a i.i.d. prírastkami, $a(t) \sim Po(\lambda)$. Kapacitu servera sme označili c . Vypočítame SCGF $\lambda(\theta)$ pracovného procesu:

$$\lambda(\theta) = \lambda_a(\theta) - \theta c = \lambda(e^\theta - 1) - \theta c = \lambda e^\theta - (\lambda + \theta c)$$

Určíme rýchlostnú funkciu $I(x)$:

$$I(x) = \sup_{\theta} \{ \theta x - \lambda(e^\theta - 1) + \theta c \}, \quad x = \lambda e^\theta - c \Rightarrow \theta_0 = \ln \frac{x + c}{\lambda} \Rightarrow$$

$$\Rightarrow I(x) = (x + c) \cdot \left[\ln \frac{x + c}{\lambda} - 1 \right] + \lambda \quad \text{pre} \quad x > -c$$

Ak použijeme na výpočet spádovej konštanty δ rýchlostnú funkciu $I(x)$, musíme minimalizovať nasledujúci výraz

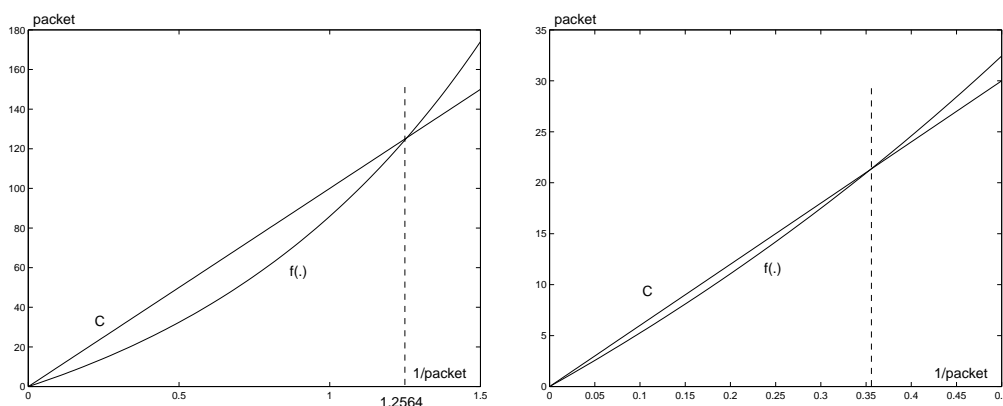
$$\delta = \min_x \frac{I(x)}{x} = \min_x \left[\frac{x + c}{x} \cdot [\ln(x + c) - \ln \lambda - 1] + \frac{\lambda}{x} \right]$$

čo znamená riešiť rovnicu $\frac{d}{dx} \frac{I(x)}{x} = 0 \Rightarrow x + c(\ln \lambda + 1) = c \ln(x + c) + \lambda$

Vidíme, že rovnica nemá analytické riešenie. Preto použijeme vetu 2.3.1:

$$\lambda_a(\theta) = \theta c \Rightarrow \lambda(e^\theta - 1) = \theta c$$

Na zistenie spádovej konštanty δ pomocou danej intenzity λ a kapacity c musíme použiť numerické metódy.



Obr. 2.10: $f(\theta) = \lambda(e^\theta - 1)$ s $\lambda = 50p/s$ a s $c = 100p/s$ a $c = 60p/s$

Pomocou numerických metód sme pre kapacitu $c = 100p/s$ získali hodnotu $\delta = 1.256p^{-1}$ a pre $c = 60p/s$ hodnotu $\delta = 0.354p^{-1}$.

Teraz bude našou úlohou navrhnúť minimálnu kapacitu servera c pre danú pravdepodobnosť prekročenia nejakej dĺžky frontu $P(q > n)$. Vieme, že podmienkou pre stabilizáciu v čase systému $M/D/1/\infty$ je $\lambda < c$. Aby navyše pravdepodobnosť pretečenia klesala exponenciálne, musí podľa LDP platiť:

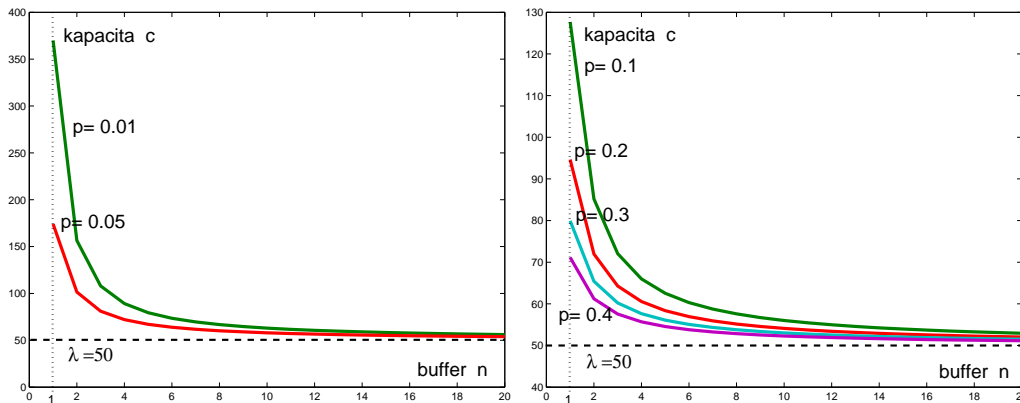
$$\lambda_a(\theta) = \theta c \Rightarrow \lambda(e^\theta - 1) = \theta c \Rightarrow c = \lambda \frac{e^\theta - 1}{\theta} \quad (2.27)$$

V takomto prípade hovoríme, že systém spĺňa θ -kritérium QoS.

Určiť minimálnu kapacitu serveru c pre danú pravdepodobnosť prekročenia nejakej medznej hodnoty n , pričom systém bude spĺňať QoS, je jednoduché:

$$p = P(q > n) \leq e^{-n\delta} = p \Rightarrow \delta = \frac{-\ln p}{n} \Rightarrow c = \lambda \cdot \frac{n(1 - p^{-1/n})}{\ln p}$$

$n \setminus p$	0.01	0.05	0.1	0.2	0.3	0.4
1	1074.88	317.12	195.43	124.27	96.90	81.85
2	195.43	115.90	93.91	76.80	68.58	63.42
3	118.61	85.84	75.20	66.17	61.52	58.48
4	93.91	74.42	67.60	61.56	58.34	56.19
5	82.08	68.48	63.50	58.98	56.53	54.87
6	75.20	64.85	60.95	57.35	55.37	54.02
7	70.73	62.40	59.20	56.22	54.56	53.42
8	67.60	60.65	57.94	55.38	53.96	52.98
9	65.28	59.33	56.99	54.75	53.50	52.63
10	63.50	58.30	56.22	54.25	53.13	52.36

Tabuľka 2.10: Kapacita c pre $M/D/1/\infty$ pomocou LDPObr. 2.11: Kapacita c pre $p = 0.01, 0.05$ a $p = 0.1, 0.2, 0.3, 0.4$

Dostali sme explicitné vyjadrenie kapacity c pre dané n a p .

PRÍKLAD 2.6 | N

ech je tok rečových paketov Poissonov s $\lambda = 50p/s$ a maximálna veľkosť buffera je $n = 10$. Máme určiť kapacitu c tak, aby pravdepodobnosť straty paketu bola menšia než 0.05. Použijeme vzťah (2.27):

$$c = \lambda \frac{e^\delta - 1}{\delta} = \lambda \cdot \frac{n(1 - p^{-1/n})}{\ln p} = 50 \cdot \frac{10(1 - 0.05^{-1/10})}{\ln 0.05} = 58.3 \text{ p/s}$$

Ak by sme mali k dispozícii buffer s polovičnou veľkosťou, potrebná kapacita linky by sa zvýšila:

$$c = 50 \cdot \frac{5(1 - 0.05^{-1/5})}{\ln 0.05} = 68.5 \text{ p/s}$$

2.3.4 Systém s variabilnou kapacitou

Budeme sa zaoberať jednolinkovým systémom hromadnej obsluhy (*single server queue*) s variabilnou kapacitou $c(t)$ a s neohraničenou dĺžkou frontu (*buffer*). Predpokladáme diskretný čas $t = 0, 1, 2, \dots$. **Výstupný proces** (*Output process*) výchádzajúci zo systému popíšeme postupnosťou náhodných premenných $C(.) \equiv \{C(t), t = 0, 1, 2, \dots\}$, pričom $C(t)$ je kumulatívny

počet výstupov za čas t . Ďalej predpokladáme

$$C(0) = 0, \quad \forall s \leq t; \quad C(s) \leq C(t)$$

Označíme $c(t)$ počet paketov vystupujúcich zo systému za čas t . Náhodné premenné $c(t)$ predstavujú variabilnú kapacitu kanála. Budeme predpokladať, že $c(t)$ sú stacionárne, nezávislé a rovnako rozdelené (i.i.d) náhodné premenné s $Ea(t) < Ec(t)$. V takom prípade hovoríme o **výstupnom procese so stacionárnymi a nezávislými prírastkami**. Platí

$$c(t) = C(t) - C(t-1) \quad , \text{ resp. } \quad C(t) = \sum_{i=1}^t c(i)$$

Jednolinkový systém s kapacitou $c(t)$ spĺňa **Lindleyho rovnicu**:

$$q(t) = \max\{0, q(t-1) + a(t) - c(t)\} = (q(t-1) + a(t) - c(t))^+$$

Lema 2.3.1. *Nech náhodný proces s diskretným časom $A(t)$ je vstupný proces vstupujúci do práce zachovávajúceho jednolinkového systému a $C(t)$ je jeho výstupný proces. Nech $q(t)$ je dĺžka frontu v čase t . Ak v čase $t = 0$ je front prázdny, $q(0) = 0$, potom platí*

$$q(t) = \max_{0 \leq s \leq t} \{A(t) - A(s) - (C(t) - C(s))\}$$

Dôkaz. Dôkaz je identický s dôkazom lemy ??

□

Rovnako, ako v predchádzajúcej kapitole zavedieme záťažový proces $\{W_t\}_{t \geq 0}$, $W_t = Z_1 + \dots + Z_t$ pre $t > 0$ a $W_0 = 0$ pričom $Z_{t-i+1} = a(i) - c(i)$. Keďže náhodné premenné $a(i)$ a $c(i)$ sú stacionárne a i.i.d. s $Ea(t) < Ec(t)$, potom aj Z_i sú stacionárne a i.i.d s $EZ_i < 0$. V takom prípade dĺžka frontu v čase t konverguje podľa rozdelenia ku konečnej náhodnej premennej q (pozri [?]).

$$q = \lim_{t \rightarrow \infty} q(t) = \lim_{t \rightarrow \infty} \max_{0 \leq i \leq t} \{W_i\} \quad \Rightarrow \quad q = \max_{t \geq 0} \{W_t\}$$

Veta 2.3.3. *Nech jednolinkový systém má variabilnú kapacitu $c(t)$ a vstupný proces $A(t)$ má stacionárne a i.i.d. prírastky $a(t)$. Označíme CGF náhodných premenných $a(t)$ ako $\lambda_a(\theta)$ a CGF náhodných premenných $c(t)$ ako $\lambda_c(\theta)$. Potom spádová konštanta δ je nenulovým riešením rovnice*

$$\lambda_a(\theta) = -\lambda_c(-\theta) \tag{2.28}$$

Dôkaz. Vieme, že $Z_{t-i+1} = a(i) - c(i)$. Pre zjednodušenie preznačíme $j = t - i + 1$ a využijeme stacionárnosť:

$$\varphi(\theta) = E[e^{\theta Z_j}] = E[e^{\theta(a(i)-c(i))}] = E[e^{\theta a(i)} e^{-\theta c(i)}] = \varphi_a(\theta) \varphi_c(-\theta) \quad \Rightarrow$$

$$\lambda(\theta) = \lambda_a(\theta) + \lambda_c(-\theta)$$

Preto pre spádovú konštantu δ platí (pozri vetu ??):

$$\delta = \max\{\theta ; \lambda(\theta) \leq 0\} \Rightarrow \delta = \max\{\theta ; \lambda_a(\theta) \leq -\lambda_c(-\theta)\}$$

Vieme, že funkcia $f(\theta) = \lambda_a(\theta)$ je konvexná a $g(\theta) = -\lambda_c(-\theta)$ je konkávna. Pre $\theta > 0$ sú obidve funkcie rastúce a pre $\theta = 0$ platí $f(0) = g(0) = 0$. To znamená, že existuje nenulové riešenie rovnice $f(\theta) = g(\theta)$, ktoré je zároveň maximom hodnôt θ a pre ktoré platí, že rýdzo konkávna funkcia $g(\theta)$ je väčšia nanejvýš rovná rýdzo konvexnej $f(\theta)$. Riešením je spádová konštantá δ .

□

Veta 2.3.4. ($GI/G/1/\infty$)

Nech vstupný proces $A(t) = \sum_{i=1}^t a(i)$ a výstupný proces $C(t) = \sum_{i=1}^t c(i)$ jednolinkového systému s dĺžkou frontu q majú stacionárne, ergodické a i.i.d. prírastky. Nech ich efektívne šírky pásma sú $\alpha_1(\theta, t)$ a $\alpha_2(\theta, t)$. Potom platí

$$\alpha_1(\delta, t) = \alpha_2(-\delta, t) \iff P(Q > n) \asymp e^{-n\delta} \quad (2.29)$$

Dôkaz. Z vlastnosti (ii) efektívnej šírky pásma dostávame

$$\alpha_1(\theta, t) = \alpha_2(-\theta, t) \Rightarrow \frac{\lambda_a(\theta)}{\theta} = \frac{\lambda_c(-\theta)}{-\theta} \Rightarrow \lambda_a(\theta) = -\lambda_c(-\theta) = 0$$

Z vety ?? a vety 2.3.3 vyplýva:

$$P(q > n) \asymp e^{-n\delta} \quad \text{kde } \delta \text{ je riešením rovnice } \lambda_a(\theta) = -\lambda_c(-\theta)$$

□

2.3.5 M/M/1/ ∞

Do Markovovho systému M/M/1/ ∞ vstupuje Poissonov tok paketov s intenzitou $\lambda = 50$ p/s a doba vysielania paketov je exponenciálna s intenzitou μ . Vo fronte sa môže hromadiť neobmedzený počet paketov, pričom za dĺžku frontu považujeme celkový počet paketov v systéme. Na určenie kapacity μ najprv využijeme klasickú THO a výsledky porovnáme s použitím LDP.

Nech vstupný a výstupný proces v M/M/1/ ∞ je Poissonov, $A(t) \sim Po(\lambda t)$ a $C(t) \sim Po(\mu t)$, so stacionárnymi a i.i.d. prírastkami $a(t)$ a $c(t)$. Určíme SCGF záťažového procesu:

$$\lambda(\theta) = \lambda(e^\theta - 1) + \mu(e^{-\theta} - 1) = \lambda e^\theta + \mu e^{-\theta} - (\lambda + \mu)$$

Rýchlostná funkcia $I(x)$ má tvar:

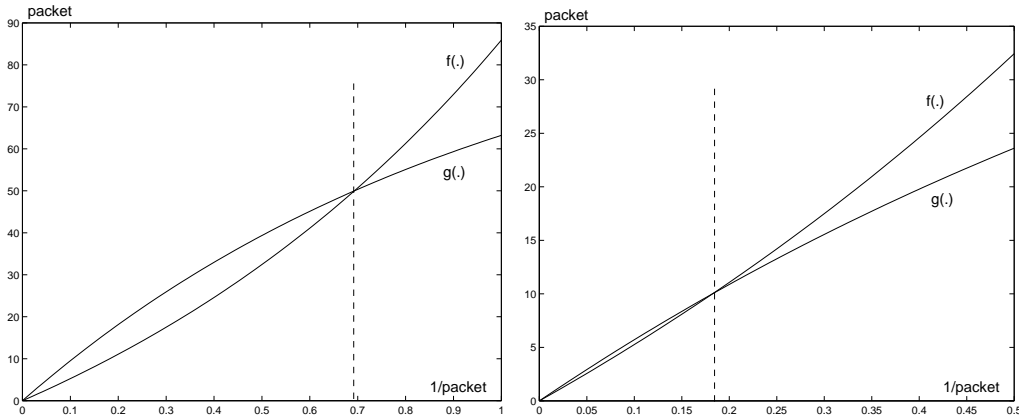
$$\begin{aligned} I(x) &= \sup_{\theta} \{\theta x - \lambda e^\theta - \mu e^{-\theta} + \lambda + \mu\}, \quad x - \lambda e^\theta + \mu e^{-\theta} = 0 \Rightarrow \\ &\Rightarrow I(x) = x \cdot \ln \frac{x + \sqrt{x^2 + 4\lambda\mu}}{2\lambda} - \sqrt{x^2 + 4\lambda\mu} + \lambda + \mu \end{aligned}$$

Ak by sme chceli na výpočet spádovej konštanty δ použiť rýchlostnú funkciu $I(x)$, museli by sme minimalizovať výraz

$$\delta = \min_x \frac{I(x)}{x} = \min_x \left[\ln \frac{x + \sqrt{x^2 + 4\lambda\mu}}{2\lambda} - \frac{\sqrt{x^2 + 4\lambda\mu} + \lambda + \mu}{x} \right]$$

My však zvolíme jednoduchšiu cestu a na výpočet δ použijeme funkciu CGF:

$$\begin{aligned} \lambda_a(\theta) = -\lambda_c(-\theta) &\Rightarrow \lambda(e^\theta - 1) = \mu(1 - e^{-\theta}) \Rightarrow \lambda(e^\theta - 1)e^\theta = \mu(e^\theta - 1) \\ &\Rightarrow \lambda e^\theta = \mu \Rightarrow \delta = \theta = \ln \lambda - \ln \mu \end{aligned} \quad (2.30)$$



Obr. 2.12: $f(\theta) = \lambda(e^\theta - 1)$ s $\lambda = 50$, $g(\theta) = \mu(1 - e^{-\theta})$ s $\mu = 100, 60$

Jednoduchým výpočtom sme pre $\mu = 100$ p/s získali hodnotu spádovej konštanty $\delta = 0.693$ p⁻¹ a pre $\mu = 60$ p/s hodnotu $\delta = 0.182$ p⁻¹.

Pomocou LDP sme ukázali, že podmienka pre intenzity vstupu a výstupu s ohľadom na splnenie θ -kritéria QoS je $\lambda e^\delta = \mu$, resp. $\lambda/\mu = e^{-\delta}$. Overíme či tento poznatok korešponduje s klasickou teóriou hromadnej obsluhy:

$$P(q > n) = (\lambda/\mu)^{n+1} = \varrho^{n+1} = (e^{-\delta})^{n+1} = e^{-(n+1)\delta} \leq e^{-n\delta}$$

Ak je splnená podmienka $\lambda/\mu = e^{-\delta}$, aj podľa klasickej teórie hromadnej obsluhy pravdepodobnosť pretečenia buffera klesá exponenciálne.

Vieme navrhnúť minimálnu kapacitu kanála tak, aby sme zaručili, že pravdepodobnosť pretečenia bucketu bude menšia ako nejaká daná hodnota p :

$$p = P(q > n) \leq e^{-n\delta} \Rightarrow \delta = \frac{-\ln p}{n} \Rightarrow \mu = \lambda e^\delta = \lambda p^{-(1/n)}$$

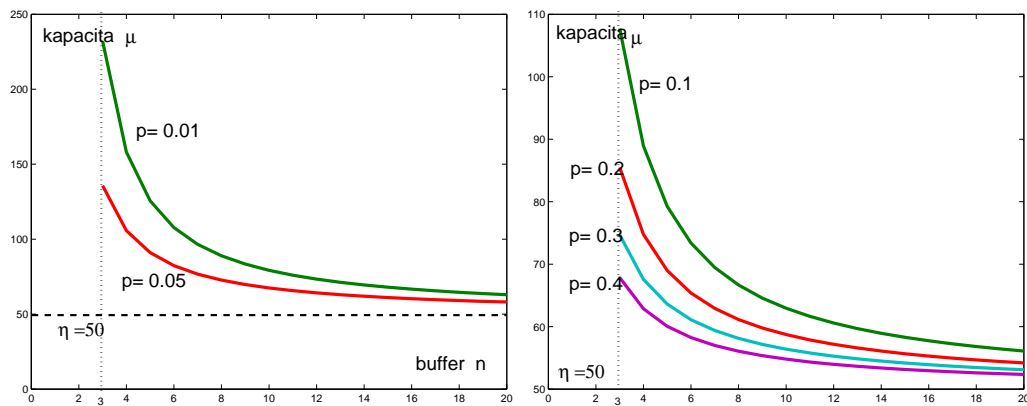
Dostali sme explicitné vyjadrenie kapacity c pre dané n a p .

PRÍKLAD 2.7 | N

ech $\lambda = 50$ p/s a veľkosť buffera je $n = 10$. Treba určiť intenzitu vysielania μ tak, aby pravdepodobnosť straty paketu bola menšia než 0.05.

$$\text{THO : } \mu = 50 \cdot 0.05^{-1/11} = 65.5 \text{ p/s} \quad \text{LDP : } \mu = 50 \cdot 0.05^{-1/10} = 67.5 \text{ p/s}$$

$n \setminus p$	0.01	0.05	0.1	0.2	0.3	0.4
1	5000.00	1000.00	500.00	250.00	166.67	125.00
2	500.00	223.61	158.11	111.80	91.29	79.06
3	232.08	135.72	107.72	85.50	74.69	67.86
4	158.11	105.74	88.91	74.77	67.56	62.87
5	125.59	91.03	79.24	68.99	63.61	60.06
6	107.72	82.38	73.39	65.38	61.11	58.25
7	96.53	76.71	69.47	62.92	59.38	56.99
8	88.91	72.71	66.68	61.14	58.12	56.07
9	83.40	69.75	64.58	59.79	57.16	55.36
10	79.24	67.46	62.95	58.73	56.40	54.80

Tabuľka 2.11: Kapacita μ pre M/M/1/ ∞ pomocou LDPObr. 2.13: Kapacita μ pre $p = 0.01, 0.05$ a $p = 0.1, 0.2, 0.3, 0.4$ **PRÍKLAD 2.8**

Nech intenzita vstupu je $\lambda = 50p/s$ a intenzita vysielania $\mu = 60p/s$. Určte veľkosť buffera tak, aby server odmietal maximálne 5% paketov.

$$\lambda e^{\delta} = \mu \Rightarrow \delta = \ln \frac{60}{50} = 0.182 \Rightarrow n = \frac{-\ln 0.05}{0.182} = 16.4 \Rightarrow n = 17$$

PRÍKLAD 2.9

Ako sa zmení potrebná veľkosť buffera, ak sa intenzita vysielania paketov zväčší na $\mu = 100p/s$?

$$\delta = \ln \frac{100}{50} = 0.693 \Rightarrow n = \frac{-\ln 0.05}{0.693} = 6.64 \Rightarrow n = 7$$

PRÍKLAD 2.10

Intenzita vstupného toku paketov je $\lambda = 50p/s$, intenzita vysielania $\mu = 60p/s$. veľkosť buffera $n = 10$. Určte pravdepodobnosť straty paketu v takomto systéme:

$$\theta = \ln \frac{\mu}{\lambda} = \ln \frac{60}{50} = 0.18232 \quad \Rightarrow \quad p = e^{-0.18232 \cdot 10} = 0.1615$$

2.3.6 Markovove systémy s MMPP vstupom

V tejto kapitole budeme pomocou LDP určovať minimálnu kapacitu servera pre modely MMPP/M/1/ ∞ , čo znamená, že doba vysielania paketu je exponenciálna s intenzitou μ a vstupný tok rečových paketov je Poissonov s intenzitou $\lambda = 50 p/s$, ktorý je generovaný nejakým Markovovým On-Off zdrojom reči. Takýto tok sa v literatúre označuje aj ako Markovov odulovaný Poissonov proces (MMPP). Vstupný a výstupný proces má stacionárne a i.i.d. prírastky, preto podľa vety 2.3.4 pre spádovú konštantu δ platí:

$$\alpha_1(\delta, t) = \alpha_2(-\delta, t) \quad \Longleftrightarrow \quad P(Q > n) \asymp e^{-n\delta}$$

Efektívna šírka pásma (EB) pre výstupný tok je: $\alpha_2(\theta, t) = \mu \frac{e^\theta - 1}{\theta}$

Určíme efektívnu šírku pásma pre Markovov On-Off zdroj. Nech \mathbf{Q}_0 je matica intenzít prechodov medzi stavmi zdroja, a nech $\mathbf{\Lambda}$ je diagonálna matica s intenzitami Poissonových tokov prislúchajúcich k stavom Markovovho zdroja. Podľa [?] platí:

$$\alpha_1(\theta, t) = \frac{1}{\theta t} \ln E [e^{\theta A(t)}] = \frac{1}{\theta} sp [\mathbf{Q}_0 + (\mathbf{e}^\theta - \mathbf{1})\mathbf{\Lambda}] := \frac{h(\theta)}{\theta} \quad (2.31)$$

pričom $sp(\mathbf{F})$ je spektrálny rádius, resp. najväčšie vlastné číslo matice \mathbf{F} .

Pre kapacitu servera μ dostávame vzťah:

$$\alpha_1(\theta, t) = \alpha_2(-\theta, t) \quad \Rightarrow \quad \frac{h(\theta)}{\theta} = \mu \frac{1 - e^{-\theta}}{\theta} \quad \Rightarrow \quad \mu = \frac{e^\theta}{e^\theta - 1} h(\theta)$$

Pre danú pravdepodobnosť pretečenia $P(q > n)$ vieme určiť minimálnu kapacitu μ :

$$\begin{aligned} p = P(q > n) \leq e^{-n\delta} \quad &\Rightarrow \quad \delta = \frac{\ln p}{-n} \quad \Rightarrow \quad \mu = \frac{p^{-1/n}}{p^{-1/n} - 1} h\left(\frac{\ln p}{-n}\right) \quad \Rightarrow \\ &\Rightarrow \quad \mu = \frac{1}{1 - p^{1/n}} sp \left[\mathbf{Q}_0 + \frac{\mathbf{1} - \mathbf{p}^{1/n}}{\mathbf{p}^{1/n}} \mathbf{\Lambda} \right] \end{aligned} \quad (2.32)$$

Dostali sme síce explicitné vyjadrenie pre kapacitu μ v závislosti od pravdepodobnosti pretečenia p a kritickej hodnoty n , ale na určenie spektrálneho rádiusu viacstavových zdrojov (viac než 2) budeme musieť riešiť polynomicke rovnice vyšších rádov.

2.3.7 M-M/M/1/∞

Vypočítame efektívnu šírku pásma (EB) pre exponenciálny M-M zdroj reči s parametrami $\alpha = 4.41 \text{ p}^{-1}$ a $\beta = 1.68 \text{ p}^{-1}$. Zdroj generuje Poissonov tok rečových paketov s intenzitou $\lambda = 50 \text{ p/s}$ a stredný počet paketov vychádzajúcich zo zdroja je $\eta = 13.793 \text{ p/s}$. Matice v (2.31) majú tvar:

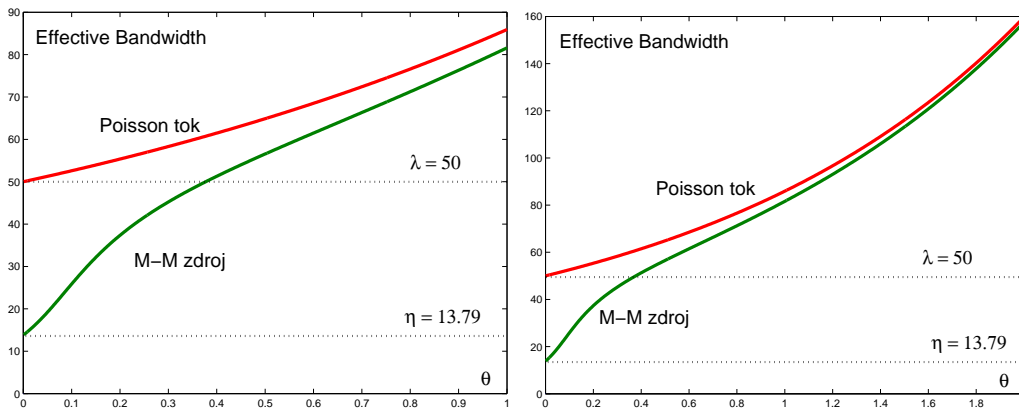
$$\mathbf{Q}_0 + (\mathbf{e}^\theta - \mathbf{1})\mathbf{\Lambda} = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix} + (\mathbf{e}^\theta - \mathbf{1}) \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \lambda(e^\theta - 1) - \alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

Spektrálny rádius $h(\theta) = sp[\mathbf{Q}_0 + \theta\mathbf{\Lambda}]$ označíme x :

$$(\alpha - (e^\theta - 1)\lambda + x)(\beta + x) - \alpha\beta = 0 \Rightarrow x^2 + x(\alpha + \beta - \lambda(e^\theta - 1)) - \lambda(e^\theta - 1)\beta = 0$$

Efektívna šírka pásma pre M-M zdroj má tvar:

$$\alpha_1(\theta, t) = \frac{\lambda(e^\theta - 1) - \alpha - \beta + \sqrt{(\alpha + \beta - \lambda(e^\theta - 1))^2 + 4\lambda(e^\theta - 1)\beta}}{2\theta}$$



Obr. 2.14: EB pre Poissonov tok a M-M s $\lambda = 50 \text{ p/s}$

Na obrázku vidíme, že pri zväčšujúcom sa parametre θ hodnota EB pre M-M zdroj konverguje zdola k EB pre Poissonov zdroj. Je to prirodzené, keby boli intenzity α a β nulové, M-M zdroj by sa zmenil na Poissonov zdroj:

$$\alpha = 0, \beta = 0 \Rightarrow \alpha_1(\theta, t) = \frac{\lambda(e^\theta - 1) + \sqrt{(-\lambda(e^\theta - 1))^2}}{2\theta} = \frac{\lambda(e^\theta - 1)}{\theta}$$

Určíme minimálnu kapacitu μ pre pravdepodobnosť pretečenia buffera p :

$$\mu = \frac{\lambda(p^{-\frac{1}{n}} - 1) - \alpha - \beta + \sqrt{(\alpha + \beta - \lambda(p^{-\frac{1}{n}} - 1))^2 + 4\lambda(p^{-\frac{1}{n}} - 1)\beta}}{2(1 - p^{\frac{1}{n}})}$$

Dostali sme explicitné vyjadrenie pre kapacitu μ v závislosti od p a n .

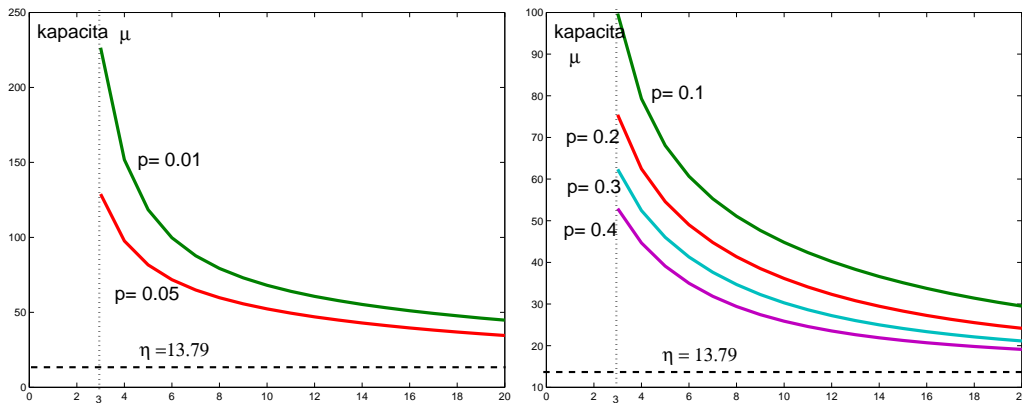
$$\alpha = 4.41 \text{ s}^{-1}$$

$$\beta = 1.68 \text{ s}^{-1}$$

$$\eta = 13.79 \text{ p/s}$$

$n \setminus p$	0.01	0.05	0.1	0.2	0.3	0.4
1	4995.55	995.37	495.12	244.53	160.46	117.82
2	495.12	217.98	151.77	104.05	81.96	67.82
3	226.51	128.88	99.74	75.42	62.35	52.91
4	151.77	97.64	79.30	62.45	52.45	44.67
5	118.44	81.67	68.04	54.57	46.01	39.10
6	99.74	71.79	60.66	49.02	41.29	34.99
7	87.74	64.92	55.28	44.77	37.62	31.85
8	79.30	59.76	51.09	41.35	34.68	29.40
9	72.99	55.67	47.68	38.52	32.28	27.45
10	68.04	52.31	44.81	36.13	30.29	25.88

Tabuľka 2.12: Intenzita μ pre M-M/M/1/ ∞ pomocou LDP



Obr. 2.15: Kapacita μ pre $p = 0.01, 0.05$ a $p = 0.1, 0.2, 0.3, 0.4$

PRÍKLAD 2.11

Našou úlohou bude určiť minimálnu veľkosť buffera tak, aby pri intenzite vysielania $\mu = 60p/s$ server odmietal maximálne 5% paketov. Na rozdiel od príkladu ?? bude zdroj rečových paketov predstavovať exponenciálny M-M model s odporúčanými parametrami.

Získať explicitné vyjadrenie veľkosti frontu n je komplikované, preto na riešenie použiť predchádzajúcu tabuľku. V stĺpci pre pravdepodobnosť $p = 0.05$ zadanej intenzite $\mu = 60p/s$ zhruba zodpovedá hodnota 59.76. V takom prípade je veľkosť frontu $n = 8$. Je to skoro polovica oproti prípadu, keď sme tok rečových paketov modelovali Poissonovým tokom (pozri príklad ??). Je to prirodzené, exponenciálny zdroj generuje v priemere len $13.8p/s$.

2.4 SIETOVÝ KALKUL

2.4.1 Dimenzovanie pomocou (σ, ρ) - Algebry

Network Calculus is a set of recent developments that provide deep insights into flow problems encountered in networking. The foundation of network calculus lies in the mathematical theory of dioids, and in particular, the Min-Plus dioid, also called Min-Plus algebra. With Network Calculus, we are able to understand some fundamental properties of integrated services networks, window flow control, scheduling, shaping and buffer or delay dimensioning.

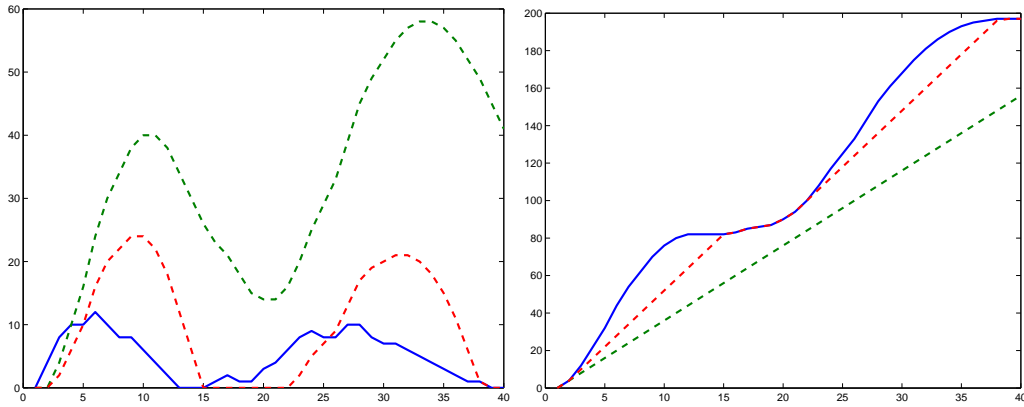
Let $q(t)$ be queue length at time t and A arrival process with increments $a(i)$, $A(t) = \sum_i a(i)$.

Then work conserving link with capacity c satisfies Lindley equation:

$$q(t+1) = (q(t) + a(t+1) - c)^+$$

Let B be **output process** from a work conserving link with capacity c . Then

$$B(t) = q(0) + A(t) - q(t)$$



Obr. 2.16: Queue length and output process of Work conserving link with capacities $c = 4$ and $c = 6$

Lemma 1.3.1.

Suppose $q(0) = 0$. Let B be output from a work conserving link with capacity c . Then

$$(i) \quad q(t) = \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)], \quad t \geq 0$$

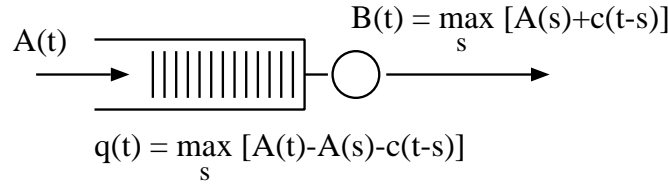
$$(ii) \quad B(t) = \min_{0 \leq s \leq t} [A(s) + c(t-s)], \quad t \geq 0$$

$$(i) \quad 1^o \quad q(1) = \max(0, q(0) + a(1) - c) = \max_{0 \leq s \leq 1} [A(1) - A(s) - c(1-s)]$$

$$\begin{aligned} 2^o \quad q(t+1) &= (q(t) + a(t+1) - c)^+ = \max \left[0, \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)] + a(t+1) - c \right] = \\ &= \max \left[0, \max_{0 \leq s \leq t} [A(t+1) - A(s) - c(t+1-s)] \right] = \max_{0 \leq s \leq t+1} [A(t+1) - A(s) - c(t+1-s)] \end{aligned}$$

$$(ii) \quad B(t) = A(t) - q(t) = A(t) - \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)] = \min_{0 \leq s \leq t} [A(s) + c(t-s)]$$

Lemma 1.3.2.



Obr. 2.17: Queue length and output process of Work Conserving link

Suppose the input A to a work conserving link with capacity c is (σ, ϱ) - upper constrained .

(i) (*Maximum queue length*) Let q be the maximum queue length.

$$\varrho < c \quad \Rightarrow \quad q < \sigma$$

$$q(t) = \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)] \leq \max_{0 \leq s \leq t} [\varrho(t-s) + \sigma - c(t-s)] = \sigma + \max_{0 \leq s \leq t} [(\varrho - c)(t-s)] = \sigma$$

(ii) (*Maximum delay*) Let d be the maximum delay among all the packets.

$$\varrho < c \quad \Rightarrow \quad d \leq \left\lceil \frac{\sigma}{c - \varrho} \right\rceil$$

A busy period begins at time t , if $q(t-1) = 0$ and $a(t) > 0$, a busy period ends at time t , if $q(t) = 0$, if a busy period begins at t_1 and ends at t_2 , then the length of the busy period is $t_2 - t_1 + 1$.

Delay of packet is bounded above by the length of the busy period. Suppose there is a busy period with length $T > \left\lceil \frac{\sigma}{c - \varrho} \right\rceil$. T is an integer in the discrete-time setting:

$$T > \left\lceil \frac{\sigma}{c - \varrho} \right\rceil \quad \Rightarrow \quad T \geq \left\lceil \frac{\sigma}{c - \varrho} \right\rceil + 1 \quad \Rightarrow \quad T \geq \frac{\sigma}{c - \varrho} + 1$$

The number of packets served in that busy period, excluding those served in the ending slot, is $c(T-1)$. The number of packets that arrive during these $T-1$ time slot cannot be less than $c(T-1)$:

$$c(T-1) < A(T) - A(1) \leq \varrho(T-1) + \sigma \quad \Rightarrow \quad cT - c < \varrho T - \varrho + \sigma \quad \Rightarrow \quad T(c - \varrho) < c - \varrho + \sigma \quad \Rightarrow$$

$$T < \frac{\sigma}{c - \varrho} + 1 \quad \textbf{conflict}$$

Lemma 1.3.3.

If the service policy is FCFS, then $\varrho \leq c \quad \Rightarrow \quad d \leq \left\lceil \frac{\sigma}{c} \right\rceil$

Literatúra

- [1] J.L.Boudec, P.Thiran *Network Calculus*, Springer Verlag, 2004
- [2] Cheng Shang Chang, *Performance Guarantees in Communication Networks*, Springer Verlag, 2000
- [3] ITU-T Recommendation Y.1315, *Global information infrastructure, Internet protocol aspect and Next-Generation Networks* 09/2006
- [4] Jozef Kalas, *Markovove reťazce - skriptum*, Polygrafické stredisko UK v Bratislave, 1993
- [5] F. P. Kelly, *Notes on Effective Bandwidth*, "Stochastic Networks: Theory and Application", pages 141-168, Oxford University Press, 1996
- [6] T-COM: *Project: Modeling of IP traffic*, Research report, 2007-2008