به نام خدا

دانشگاه علامه طباطبایی تهران

دانشکده اقتصاد
گروه آمار

پایان نامه جهت اخذ کارشناسی ارشد

عنوان:

سامانه های صف بندی با دو نوع ورودی و دونوع سرویس غیر متجانس با تعطیلی سرویس دهنده یا دورده اولویت در سرویس دهی

استاد راهنما:
دکتر عبدالرحیم بادامچی زاده

استاد مشاور:
دکتر محمدرضا صالحی راد

تهیه و تنظیم:
فاطمه رحیمی

شهریور 88

فهرست مطالب

فصل سوم: سامانه های صف بندی با اولویت در سرویس دهی

فصل چهارم: مثال های عددی

فصل اول

معرفی صف ها

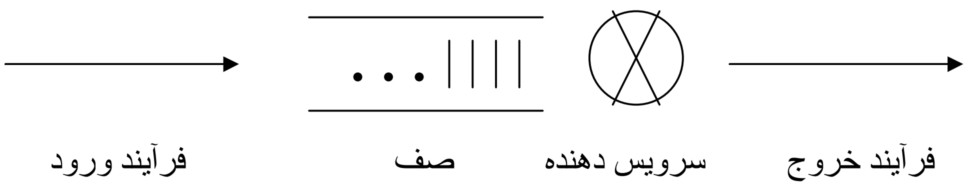
1-1 مقدمه

مدل های تصادفی بخشی از نظریه احتمال را تشکیل می دهند که در توصیف و تحلیل پدیده های طبیعی هم به کار می رود. معمولاً ضابطه سازی بر اساس قوانین تعیینی کار پیچیده ای است و لذا شرایط محیطی وطبیعت مساله منجر به استفاده وسیع از مفاهیم تصادفی می شود. زمینه های عملی مانند ارتباطات راه دور یا بیمه، نتایج و روش هایی از مدل سازی تصادفی را به علوم کاربردی مانند مهندسی یا اقتصاد وارد کرده است. از طرف دیگر اغلب پیشرفت های تکنولوژیکی منجر به مساله های جدید در یک زمینه عملی شده و باعث ظهور مسیرهای تازه ای در تحقیقات و مباحث تصادفی و انگیزه هایی در احتمال کاربردی می شود. مدل سازی تصادفی علمی است که در آن تقابل نزدیکی بین نظریه و کاربرد های عملی وجود دارد. این از ویژگی های جالب آن است که مفاهیم نظری با واقعیت های عملی ترکیب می شوند. از طرف دیگر منطبق کردن مفاهیم وسیع نظری با کاربرد های عملی در طبیعت نیز کار مشکلی است. یکی از مهم ترین زمینه ها در مدل سازی تصادفی نظریه صف می باشد که موضوع اصلی این پایان نامه است.

2-1 مثال ها و مفاهیم مقدماتی

بسیاری از سامانه ها دارای مؤلفه هایی هستند که به کمک نظریه صف، مدل سازی می شوند. بیشتر ایده های این نظریه از تجربه های روزمره در صف باجه های دریافت پول، سوپرمارکت ها، بانک ها و از موارد مشابه گرفته شده اند. یک صف به معنای علمی عبارتست از سامانه ای که درآن جمعیتی از کاربران قبل از ترک سامانه ظرفیتی از آن را اشغال می کنند. این کاربران در سامانه توسط یک سرویس دهنده یا بیشتر سرویس داده می شوند. بنابراین یک سامانه صف را می توان به عنوان توصیفی (تصادفی) از جمعیت ورودی ونوع تقاضای هرکاربرو همچنین خط مشی سرویس تعریف کرد. قبل از تعریف دقیق یک سامانه صف چند مثال می آوریم.

مثال 1-2-1 صف تک سرویس دهنده



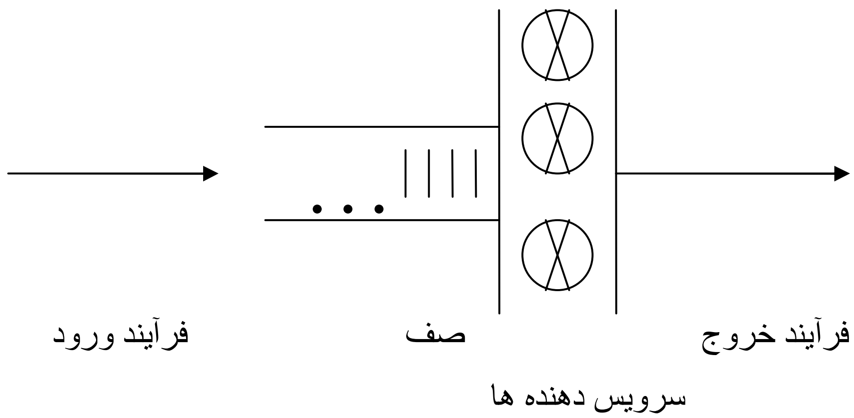
شکل 1.2.1 صف تک سرویس دهنده

صفی که در مقابل باجه شمارش پول یک سوپرمارکت تشکیل می شود، ساده ترین توصیف از یک سامانه است. یک جریان ورودی و یک سرویس دهنده بر اساس ورود متقاضیان به آنها سرویس می دهد.

این خط مشی سرویس که در آن هیچ اولویتی برای کاربران وجود ندارد، صف سرویس به ترتیب ورود (FCFS) نامیده می شود.

■

مثال 1-2-2 صف چند سرویس دهنده:



شکل 1.2.2 صف چند سرویس دهنده

اولین کاربرد واقعی از نظریه صف در طراحی و تحلیل شبکه های تلفن ظاهر شد. پیشتر از این در ابتدای قرن بیستم مکالمات تلفنی قبل از وصل شدن به شخص مورد نظربه یک اپراتور می رسیدند. بنابراین بخش مهمی ازیک شبکه تلفن را می توان به کمک یک سامانه صف مدل سازی کرد به طوری که سرویس دهنده ها ، اپراتورهای مرکز مکالمات بوده و مکالمات رسیده (که به عنوان جریان کاربران ورودی هستند) را به آدرس های مورد نظر می فرستند. در اینجا زمان ارتباط با تقاضای سرویس یک متقاضی تعیین می شود. در این سامانه وقتی یک متقاضی اقدام به مکالمه می کند، احتمال اینکه اپراتورها را مشغول ببیند و در نتیجه نتواند سرویس دریافت کند، مهم است. این مقدار را احتمال زیان[1] سامانه گویند. در مراکزارتباطی پیشرفته به جای کابل های ارتباطی سوالات به طور حضوری پاسخ داده می شوند و زمان سرویس، زمان مکالمه بین متقاضی و اپراتور است.

مثال1-2-3 در حال حاضر شبکه های رایانه ای (که مثال بارزآن اینترنت است) موجب توسعه روز افزون کاربرد های نظریه صف شده اند. به عنوان مثال در شبکه رایانه ای، یک سرویس دهنده تقاضایی از کاربران دریافت کرده و باید به آنها سرویس دهد. تقاضا به عنوان جریان ورودی به سامانه صف، تعبیر می شود که نشانگر بهره دهی سرویس دهنده است. یک خط مشی سرویس که معمولاً در این سامانه به کار می رود، عبارتست از اینکه :

۱) Loss

ظرفیت پردازش دهنده به طور مساوی بین تقاضـاها تقسیم مـی شـود، بـه قسمی کـه هیچ متقاضـی اولـویتی نداشـته ولی زمان سرویس هر متقاضـی بـه تعداد متقاضیان حاضـر وابسـته است. بـه دلیل رواج این مـدل در کـاربرد هـای رایانه ای، این خط مشی را اشتراک پردازشگر[2] نامند.

■

مثال 1-2-4 برای صف ها کاربرد های زیادی در ترافیک فرودگاه ها نیز می توان یافت. در اینجا سرویس دهنده ها باندهای فرود متعدد و قابل دسترس برای هواپیماهای ورودی به عنوان متقاضیان سامانه هستند. واضح است که صفی برای هواپیماهای در حال فرود در آسمان وجود ندارد لذا هواپیمای ورودی برای یافتن باندی قابل دسترس برای فرود باید در آسمان دور بزند و این را تا جایی ادامه می دهد که باندی خالی برای فرود پیدا کند. با هر دور زدن بنزین هواپیما به سرعت مصرف می شود و لازم است برای جلوگیری از حادثه، اولویت ویژه ای برای فرود داده شود. با این اعمال نفوذ در مراحل سرویس، صف را با اولویت[3] گویند.

■

مثال 1-2-5 در سامانه های پیشرفته تولید، تحلیل خط تغذیه[4] مهم است. در این مدل که آن را شبکه صف پی در پی[5] می نامند، یک سری از سامانه های صف طوری چیده شده اند که خروجی یک صف ، ورودی صف بعدی است.

■

مثال 1-2-6 هر خط تولید در یک کارخانه ، مثالی از یک صف است. ورودی شامل قطعاتی است که باید تعمیر و یا پردازش شوند. در این خط امکان دارد قطعه ای در انتهای خط مورد قبول سامانه واقع نشود. لذا لازم است دوباره به ابتدای خط برگشته و مورد سرویس مجدد قرار گیرد. در همین سامانه صف، به دلایلی دستگاه پردازشگر (که همین سرویس دهنده است) خراب شده و از کار می افتد واین اتفاق منجر به تغییراتی در صف ورودی می شود. صف هایی با این ویژگی را صف با تعطیلی[6] نامند. دراین پایان نامه به این صف ها توجه ویژه ای می شود.

با این چند مثال می توان توصیفی از یک صف را به صورت زیر بیان کرد:

یک سامانه صف عبارتست از یک یا چند سرویس دهنده، یک طرح ورودی از متقاضیان، طرح سرویس و نظم صف یعنی روشی که وقتی صف شکل گرفت، متقاضیان را برای سرویس انتخاب می کنیم. تعداد متقاضیان در سامانه در زمان دلخواه برابر تعداد متقاضیان در صف و تعداد متقاضیان در سرویس است. البته این تعداد برحسب زمان ورود و عزیمت متقاضیان تغییرمی کند ودرنتیجه یک متغیر تصادفی است.

در ادامه مفاهیم اساسی در یک سامانه صف آورده می شود:

طرح ورود: طرح ورود متقاضیان مانند ورود بیماران به مطب پزشک، مکالمات تلفنی، ورود اتومبیل ها به پمپ بنزین و غیره را طرح ورود به یک صف گویند. طرح ورود منظم یا نامنظم (یعنی بطور تصادفی در فاصله های زمانی) است.

اگر تعداد متقاضیان خیلی زیاد باشد در این صورت احتمال ورود در بازه زمانی آینده به تعداد متقاضیان حاضر در سامانه بستگی ندارد. از این رو ورود کاملاً تصادفی است و فرآیند پواسون با میانگینی برابر متوسط تعداد ورودها در واحد زمان است. دراین صورت زمان بین ورودها به عنوان متغیرهای تصادفی مستقل دارای توزیع نمایی منفی است. عکس این موضوع نیز برقرار است و با فرض اخیر تعداد متقاضیان ورودی در یک بازه دلخواه، متغیری تصادفی با توزیع پواسون خواهند بود و نرخ ورود λ، برای توصیف فرایند ورود کافی است.

معکوس λ یعنی $(\frac{1}{\lambda})$ متوسط زمان بین دو ورود است.

طرح سرویس: طرح سرویس با اندکی تفاوت، شبیه طرح ورود است. چون در دوره بیکاری که متقاضی وجود ندارد این طرح منقطع می شود، لذا توابع توزیع احتمال مربوط به سرویس، مشروط به غیر تهی بودن سامانه اند.

با افزایش طول صف ممکن است سرویس دهنده ها سریع تر کار کنند. در این حالت ساده ترین توزیع زمان سرویس عبارتست از توزیع نمایی منفی ، که بطور کامل با یک پارامتر یعنی نرخ سرویس μ یا متوسط زمان سرویس $(1/\mu)$ تعیین می شود.

با این حال می توان زمان سرویس را یک توزیع دلخواه و کلی در نظر گرفت.

آرایش های سرویس: برای ارایه سرویس به متقاضیان، یک یا چند محل سرویس (یا سرویس دهنده) چیده می شود که بستگی به تعداد متقاضیان، نرخ ورود، زمان لازم برای ارایه سرویس هر متقاضی وغیره دارد. بر حسب این متغیرها محیط سرویس تکی و یا چند باجه ای است.

تک سرویس دهنده ها: در حالت تک سرویس دهنده یک صف و یک باجه سرویس وجود دارد به قسمی که هر متقاضی به تنهایی سرویس داده می شود.

باجه های سری: در بعضی از فرآیند های صف بندی ممکن است سرویس چند مرحله ای داشته باشیم که درآن هر متقاضی باید مراحل مختلفی را یکی پس از دیگری طی کند. در این حالت باجه های سرویس به صورت سری قرار دارند.

باجه ها موازی: تعدادی باجه با امکانات یکسان برای ارایه سرویس وجود دارند به قسمی که چند متقاضی بطور همزمان می توانند سرویس دریافت کنند.

زمان سرویس: زمان لازم برای اتمام سرویس یک متقاضی را زمان سرویس گوییم. زمان سرویس ممکن است مقداری ثابت بوده و یا برحسب نوع متقاضی تغییر کند. اغلب برای ساده کردن مدل صف فرض می شود که زمان سرویس برای همه متقاضیان مقداری یکسان است. علاوه بر این چون طرح ورودی تصادفی فرض می شود لذا زمان سرویس نیز تصادفی است. از این رو زمان سرویس دارای توزیع نمایی با میانگینی برابر عکس نرخ سرویس می باشد. در حالتی که فرض نمایی بودن زمان سرویس ساقط می شود، توزیع ارلانگ به کار می رود. با این حال می توان زمان سرویس را یک توزیع کلی در نظر گرفت.

خط مشی(نظم) صف: ترتیب انتخاب متقاضیان برای سرویس را خط مشی صف گویند.

یکی از حالتهای زیر رخ می دهد:

۱) اولین ورود، اولین سرویس(**FCFS**): انتخاب متقاضی برای سرویس به ترتیب ورود به صف است.

۲) آخرین ورود، اولین سرویس(**LCFS**): انتخاب متقاضی برای سرویس عکس ورود به صف است.

۳) سرویس به ترتیب تصادفی(**SCRO**): انتخاب متقاضی برای سرویس در زمانهای معین تصادفی است.

٤) سرویس با اولویت(**SIP**): متقاضیان خاصی نسبت به بقیه دارای اولویت برای دریافت سرویس هستند. دو نوع اولویت وجود دارد: اولیت عادی (Non-Emptive) و اولویت مخصوص (Emptive).

اولویت عادی(**Non-Emptive**): سرویس متقاضی معمولی حتی با ورود یک متقاضی با اولویت ادامه و پایان می یابد.

اولویت مخصوص(**Emptive**): به محض ورود یک متقاضی با اولویت، سرویس متقاضی معمولی قطع شده و سرویس متقاضی بااولویت شروع می شود.

رفتار متقاضی: معمولاً رفتار متقاضی بر چهار نوع است:

۱) جا زن: برخی از متقاضیان برای انتظار در صف بی قرارند. آنها به طور صحیح در صف قرار نمی گیرند و سعی در جا زدن و رسیدن به ابتدای صف با تخطی به دیگران دارند.

۲) ناشکیبا: برخی از متقاضیان پس از مدتی انتظار در صف بدون دریافت سرویس سامانه را ترک می کنند.

۳) تبانی: برخی از متقاضیان با همدیگر وارد صف شده و فقط یکی از آنها به جای بقیه در صف باقی می ماند. با این حال در موقع سرویس متقاضیانی که صف را ترک کرده اند (تبانی گرها) برای دریافت سرویس مراجعه می کنند.

٤) تغییر جا: در این حالت بیش از یک صف برای سرویس مشابه وجود دارد، برخی از متقاضیان اقدام به جابجایی از یک صف به صف دیگر می کنند تا موقعیت خود را بهبود بخشیده و سریع تر سرویس بگیرند.

تعطیلی سرویس دهنده: در اکثر صف ها سرویس دهنده به دلایل گوناگونی برای مدت معینی سرویس را ترک می کند. در یک خط تولید برای تعمیر یا ارتقاء دستگاه، در باجه بانکی برای رفع خستگی کارمند و یا ارایه سرویس دیگر و یا صرف ناهار و از این قبیل سرویس تعطیل می شود.

بنابر تعریف های فوق، تعریف رسمی و علمی یک سامانه صف را بر اساس مفاهیم احتمال و فرایند های تصادفی می آوریم:

تعریف 1-2-1 برای هر $n=1,2,3,\ldots$ فرض می کنیم T_n و S_n متغیرهای تصادفی حقیقی و مثبت باشند به قسمی که $T_{n+1} > T_n$. دنباله $T=\{T_n\ ;\ n=1,2,..\}$ را فرآیند نقطه ای ورود و دنباله $S=\{S_n\ ;\ n=1,2,..\}$ را دنباله زمان های سرویس گویند. علاوه بر این عدد k به عنوان تعداد سرویس دهنده ها و عدد c به عنوان ظرفیت سامانه با شرط $k,c \in \aleph \cup \{\infty\}$ می باشند.

بالاخره خط مشی سرویس ß باید تعیین شود. این خط مشی می تواند اولین ورود، اولین سرویس؛ آخرین ورود، اولین سرویس؛ سرویس با اولویت و انواعی از این قبیل باشد.

معمولاً خط مشی (FCFS) مورد استفاده قرار می گیرد.

نماد گذاری یک سامانه صف بندی به صورت (T,S,k,c,ß) است که در آن T و S فرآیند ورود والگوی سرویس؛ k تعداد سرویس دهنده ها ؛ c ظرفیت سامانه و ß خط مشی سرویس را تعریف می کنند.

همچنین n امین زمان بین دو ورود با $z_1=T_1$ و $z_n=T_n-T_{n-1}$ به ازای n□2 را تعریف می کنیم. یک نوع نمایش برای تعیین یک صف **نمایش کندال**[8] است که به صورت T/S/k/c/ß نشان داده می شود. علاوه بر این قراردادهای زیر نیز وجود دارد:

اگر =c باشد چهارمین پارامتر و اگر ß برابر با FCFS باشد، پنجمین پارامتر نوشته نمی شود. برای دو پارامتر نخستین حرفG نشانگر توزیع کلی احتمال است؛ یعنی هیچ فرضی در مورد شکل دقیق توزیع نداریم. نماد M برای توزیع نمایی E_k برای توزیع ارلانگ از نوع k و H_k برای توزیع ابرنمایی ازنوع k به کار می روند.

رفتار سامانه بر اساس زمان t نقش مهمی در تحلیل آن دارد.

تعریف 2-2-1 تحلیل سامانه های صف بندی به کمک اندازه های مؤثر بودن مانند طول صف، زمان انتظاری که ممکن است متقاضی متحمل شود و... انجام می گیرد. در سامانه هایی که این اندازه ها و توزیع های آنها به زمان وابسته باشند، سامانه را **حالت گذرا**[8] گوییم.

هرگاه رفتار سامانه مستقل از زمان باشد آن را **حالت پایا**[9] گوییم.

7)Kendal notation 8)Transient state 9) Steady state

اگر $p_n(t)$ نشانگر احتمال وجود n متقاضی در سامانه در زمان t باشد، آنگاه در حالت پایا داریم:

$$\lim_{t\to\infty} p_n(t) = p_n \qquad (1\text{-}2\text{-}1)$$

در نتیجه

$$\lim_{t\to\infty} \frac{d}{dt} p_n(t) = 0 \qquad (2\text{-}2\text{-}1)$$

اگر نرخ ورود دارای توزیع پواسون باشد آنگاه توزیع زمان بین دو ورود متوالی نمایی است و بر عکس. این موضوع مضمون قضیه ای اساسی در فرایندهای تصادفی است که برای اثبات آن نیاز به مقدماتی داریم که به شرح زیر است:

فرض کنیم r ورودی بین زمان t و t+h به شرط ورود n متقاضی تا لحظه t داشته باشیم. اگر N(t) تعداد ورودها تا زمان t باشد، آنگاه احتمال رخداد مذکور برابر است با:

$$p_r(h) = P[N(t+h) = r \mid N(t) = n]$$

$$\begin{aligned} &= \lambda h + o(h) && r = 1 \\ &= o(h) && r \geq 2 \qquad (3\text{-}2\text{-}1) \\ &= 1 - \lambda h + o(h) && r = 0 \end{aligned}$$

در این صورت داریم:

$$p_n(t+h) = p_n(t)(1-\lambda h + o(h)) + p_{n-1}(t)(\lambda h) + o(h) \qquad n \geq 1 \qquad (4\text{-}2\text{-}1)$$

با آرایش این رابطه و تقسیم آن بر h و حد $h \to 0$ داریم:

$$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t) \qquad (5\text{-}2\text{-}1)$$

علاوه بر این

$$p_0'(t) = -\lambda p_0(t) \qquad (6\text{-}2\text{-}1)$$

قضیه 1-2-1 : تعداد پیشامدها تا لحظه t یعنی N(t) دارای توزیع پواسون است اگر و فقط اگر فواصل زمانی رخداد پیشامدها متغیرهای تصادفی مستقل و هم توزیع نمایی باشند.

اثبات: فرض کنیم تعداد پیشامد ها تا لحظه t دارای توزیع پواسون است .اگر X_1 زمان رخداد اولین پیشامد و به طور کلی X_n فاصله زمانی بین وقوع (n-1)امین تا n امین پیشامد باشد، آنگاه برای t>0 ،

$$P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}$$

پس

$$F_{X_1}(x) = 1 - e^{-\lambda t}$$

در نتیجه X_1 دارای توزیع نمایی است. حال داریم :

$$P(X_2 > t) = \int_0^{+\infty} P(X_2 > t \mid X_1 = x) f(x) dx$$

اما

$P(X_2>t$□$X_1=x)=P[$ورودی در بازه $(x,x+t]$ رخ ندهد$]=P[N(t)=0]=e^{-\ t}$

رابطه اخیر نشان می دهد که X_2 از X_1 مستقل است و دارای توزیع نمایی است. همچنین

$$P(X_2 > t) = \int_0^{+\infty} e^{-\lambda t} f(x) dx = e^{-\lambda t}$$

به همین ترتیب می توان دید که برای هر $n \in \aleph$ متغیر X_n از X_{n-1} مستقل و دارای توزیع نمایی با میانگین 1/λ است .

حال فرض کنیم X_i ها متغیرهای تصادفی مستقل ودارای توزیع نمایی اند. بامحاسبه تبدیل لاپلاس معادله (1-2-5) داریم:

$$L\{p_n'(t)\} = -\lambda L\{p_n(t)\} + \lambda L\{p_{n-1}(t)\}$$

در این صورت

$$sL\{p_n(t)\} - p_n(0) = -\lambda L\{p_n(t)\} + \lambda L\{p_{n-1}(t)\}$$

چون $p_0(0) = P[N(0) = 0] = 1$ و $p_n(0) = P[N(0) = n] = 0$ ، پس داریم:

$$(s+\lambda) L\{p_n(t)\} = \lambda L\{p_{n-1}(t)\}$$

این یک معادله بازگشتی است که جواب آن با شرایط اولیه فوق برابر است با: $L\{p_n(t)\} = \dfrac{\lambda^n}{(\lambda + s)^n}$

پس : $$p_n(t) = P[N(t) = n] = L^{-1}\left\{\frac{\lambda^n}{(\lambda + s)^n}\right\} = \lambda^n . e^{-\lambda t} \frac{t^n}{n!} = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

بنابراین N(t) دارای توزیع پواسون است.

نرخ سرویس: تعداد متوسط متقاضیانی که در واحد زمان سرویس دریافت می کنند، نرخ سرویس می گوییم و آنرا با μ نمایش می دهیم.

عامل بهره دهی یا شدت ترافیک: این مقدار را با ρ نمایش داده و آنرا به صورت زیر تعریف می کنیم:

ρ = نرخ ورود/نرخ سرویس (7-2-1)

اگر $\rho > 1$ ، آنگاه شدت ترافیک بالا بوده و در نتیجه زمان انتظار زیاد خواهد بود.

نرخ بیکاری: بنابر تعریف (1-2-1) نرخ بیکاری در صف M/M/1 برابر است با:

$$p_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}$$

در این صورت میانگین زمان بیکاری عبارتست از:

میانگین زمان بیکاری = نرخ بیکاری×زمان کل سرویس

3-1 صف M/M/1

این سامانه صف، ساده ترین و معمولی ترین سامانه است که حالت خاص اکثر سامانه های صف است. به علت اینکه ظرفیت سامانه نامتناهی و خط مشی سرویس FCFS است لذا پارامتر چهارم و پنجم قید نمی شوند.

مراجعات می توانند به عنوان زادهای سیستم تلقی شوند زیرا اگر تعداد در سیستم n باشد و مراجعه ای رخ دهد تعداد n+1 می شود از طرف دیگر وقتی تعداد در سیستم n است ، یک عزیمت، سیستم را به یکی کمتریعنی n-1 می رساند می توان آنرا به عنوان مرگ تلقی کرد. این حالت را فرایند زاد-مرگ می نامیم.

فرض می کنیم ورودها از فرایند پواسون با نرخ λ تبعیت می کنند و زمان سرویس تک سرویس دهنده سامانه دارای توزیع نمایی با نرخ μ است.

حال فرض می کنیم $p_n(t)$ احتمال وجود n متقاضی در سامانه در زمان t باشد. در این صورت برای باقی ماندن در حالت n تا لحظه t+□t حالت های زیر را داریم:

i) سامانه تا لحظه t در حالت n قرار دارد و هیچ ورودی یا سرویسی در فاصله زمانی t رخ نمی دهد.

ii) سامانه تا لحظه t در حالت n قرار دارد و فقط یک ورود و یک سرویس در فاصله زمانی t اتفاق می افتد.

iii) سامانه در زمان t در حالت n-1 قرار دارد و در فاصله زمانی (کوچک) t یک ورودی رخ داده و سرویسی تکمیل نشود.

iv) سامانه در زمان t در حالت n+1 قرار دارد و در طول بازه زمانی (کوچک) t یک سرویس تکمیل شده و هیچ ورودی رخ نمی دهد.

از این رو چون برای هر n□1 ورودی ها و سرویس ها از همدیگر مستقل اند، لذا داریم:

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)[1-\lambda\Delta t][1-\mu\Delta t] + p_n(t)[\lambda\Delta t][\mu\Delta t] \\ &\quad + p_{n+1}(t)[1-\lambda\Delta t][\mu\Delta t] + p_{n-1}(t)[\lambda\Delta t][1-\mu\Delta t] \\ &= p_n(t)[1-(\lambda+\mu)\Delta t] + p_{n-1}(t)[\lambda\Delta t][1-\mu\Delta t] + p_{n+1}(t)[1-\lambda\Delta t][\mu\Delta t] \end{aligned}$$

(از جملاتی با درجه 2 و بالاتر برحسب t صرفنظر کرده ایم) در نتیجه:

$$\lim_{\Delta t\to 0} \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} = -(\lambda+\mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t) \qquad (1\text{-}3\text{-}1)$$

حال چون به ازای n=0 احتمال $p_{n-1}(t)$ تعریف نشده است، لذا $p_0(t+□t)$ را به طور مجزا محاسبه می کنیم. سامانه می تواند در لحظه t+□t در حالت n=0 باشد اگر در زمان t درحالت n=0 بوده و هیچ ورودی در طول زمان t رخ ندهد (سرویس غیر ممکن است چون سامانه خالی است)؛ و یا سامانه در زمان t در حالت n=1 قرار داشته و هیچ ورودی رخ نداده ولی یک سرویس تکمیل شود. پس:

$$p_0(t+\Delta t) = p_0(t)[1-\lambda\Delta t] + p_1(t)[\mu\Delta t]][1-\lambda\Delta t]$$

با صرفنظر از جملات شامل $(\ t)^2$ داریم:

$$\lim_{\Delta t\to 0} \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) \qquad (2\text{-}3\text{-}1)$$

معادلات (1-3-1) و (2-3-1) را معادلات دیفرانسیلی تفاضلی گویند.

معادلات حالت پایا:

در حالت پایا با حد $t\to\infty$ داریم $p_n(t)\to p_n$ و $p'_n(t)\to 0$. جواب حالت پایا وجود دارد اگر $\frac{\lambda}{\mu}<1$.

اگر $\lambda=\mu$ ، آنگاه صفی وجود ندارد و اگر $\frac{\lambda}{\mu}>1$ ، آنگاه حالت انفجار رخ می دهد. با برقراری شرایط حالت پایا معادلات (1-3-1) و(2-3-1) به صورت زیر خواهند بود:

$$0=-(\lambda+\mu)p_n+\mu p_{n+1}+\lambda p_{n-1} \qquad (3\text{-}3\text{-}1)$$

$$0=-\lambda p_0+\mu p_1 \quad \Rightarrow \quad p_1=\frac{\lambda}{\mu}p_0 \qquad (4\text{-}3\text{-}1)$$

معادلات اخیر را **معادلات تعادلی**[10] گویند.

معادله (1-3-3) یک معادله تفاضلی مرتبه دو است که با شرط اولیه (1-3-4) جواب آن عبارتست از:

$$p_n=(\frac{\lambda}{\mu})^n p_0 \qquad n\geq 0 \qquad (5\text{-}3\text{-}1)$$

برای محاسبه p_0، ازشرط $\sum_{n=0}^{\infty}p_n=1$ و معادله (1-3-5) داریم:

$$p_0+(\frac{\lambda}{\mu})p_0+(\frac{\lambda}{\mu})^2p_0+...=1$$
$$\Rightarrow p_0[1+(\frac{\lambda}{\mu})+(\frac{\lambda}{\mu})^2+...]=1$$

چون $\frac{\lambda}{\mu}<1$ لذا سری فوق یک سری هندسی با قدر نسبت کمتر از یک است، لذا $p_0=1-\frac{\lambda}{\mu}=1-\rho$

که درآن $\rho=\frac{\lambda}{\mu}<1$. بنابراین:

$$p_n=(\frac{\lambda}{\mu})^n p_0=(\frac{\lambda}{\mu})^n(1-\frac{\lambda}{\mu}) \;=\rho^n(1-\rho) \qquad (6\text{-}3\text{-}1)$$

متوسط تعداد متقاضیان در سامانه برابر است با :

$$L=E(N)=\sum_{n=0}^{\infty}np_n=(1-\rho)\sum_{n=0}^{\infty}n\rho^n$$
$$=(1-\rho)\rho\sum_{n=0}^{\infty}n\rho^{n-1}=(1-\rho)\rho(1-\rho)^{-2}$$
$$=\frac{\rho}{1-\rho}=\frac{\lambda}{\mu-\lambda} \qquad (7\text{-}3\text{-}1)$$

10)Stochastic equations

متوسط تعداد متقاضیان درصف در حالت پایا نیز برابر است با :

$$L_Q = E(N_Q) = 0.p_0 + \sum_{n=1}^{\infty}(n-1)p_n = \sum_{n=1}^{\infty} np_n - \sum_{n=1}^{\infty} p_n$$

شبیه محاسبات بالا می توان دید که

$$L_q = E(N_q) = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} \qquad (8\text{-}3\text{-}1)$$

توزیع زمان انتظار: فرض کنیم متغیر تصادفی T_q "زمان انتظار در صف" و $W_Q(t)$ تابع توزیع تجمعی آن باشد. آنگاه داریم:

$$W_Q(0) = P[T_Q \le 0] = P[T_Q = 0] = P[\text{در بدو ورود سامانه تهی باشد}]$$
$$= p_{0.} = 1-\rho \qquad (9\text{-}3\text{-}1)$$

حال به محاسبه $W_Q(t)=P[T_Q \quad t]$ می پردازیم، یعنی احتمال اینکه زمان انتظار متقاضی برای سرویس کوچکتر یا مساوی t باشد. اگر در بدو ورود متقاضیی که بین 0و t وارد سرویس می شود، n متقاضی در سامانه وجود داشته باشد، آنگاه تمام این n نفر باید تازمان t سرویس گرفته باشند. چون توزیع سرویس بی حافظه است، توزیع زمان لازم برای تکمیل سرویس n متقاضی مستقل اززمان ورود متقاضی مفروض و برابر پیچش n متغیر تصادفی نمایی است، که ارلانگ نوع n است. به علاوه چون ورودی پواسون است، لحظات ورود به طور یکنواخت توزیع شده و از این رو احتمال اینکه مراجعه کننده n متقاضی را در سامانه ببیند برابر توزیع مانای اندازه سامانه است. پس می توان نوشت:

$$\begin{aligned}
W_Q(t) = P[T_Q \le t] &= \sum_{n=1}^{\infty} P(\text{تکمیل سرویس n متقاضی در زمان کمتر از t} \mid \text{مراجعه کننده n متقاضی را در سامانه ببیند})\, p_n + W_Q(0) \\
&= (1-\rho)\sum_{n=1}^{\infty}\rho^n \int_0^t \frac{\mu(\mu x)^{n-1}}{(n-1)!} e^{-\mu x}\,dx + (1-\rho) \\
&= (1-\rho)\rho\int_0^t \mu e^{-\mu x}\sum_{n=1}^{\infty}\frac{(\mu x\rho)^{n-1}}{(n-1)!}\,dx + (1-\rho) \\
&= (1-\rho)\rho\int_0^t \mu e^{-\mu x(1-\rho)}\,dx + (1-\rho) \\
&= 1-\rho e^{-\mu(1-\rho)t} \qquad t>0
\end{aligned}$$

بنابراین توزیع زمان انتظار در صف برابرست با:

$$W_Q(t) = \begin{cases} 1-\rho & t=0 \\ 1-\rho e^{-\mu(1-\rho)t} & t>0 \end{cases} \tag{10-3-1}$$

پس **امید ریاضی زمان انتظار** برابراست با

$$\begin{aligned}
W_Q = E(T_Q) &= \int_0^{+\infty} t\,dW_q(t) = 0(1-\frac{\lambda}{\mu}) + \int_0^{+\infty} t\frac{\lambda}{\mu}(\mu-\lambda)e^{-(\mu-\lambda)t}\,dt \\
&= \frac{\lambda}{\mu}\int_0^{+\infty} t(\mu-\lambda)e^{-(\mu-\lambda)t}\,dt
\end{aligned}$$

با محاسبه انتگرال داریم:

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)} \tag{11-3-1}$$

فرض کنیم متغیر تصافی T کل زمانی باشد که متقاضی باید تا زمان سرویس در سامانه صرف کند. اگر W(t) تابع توزیع آن و w(t) تابع چگالی آن باشد، آنگاه با روشی مشابه روش فوق می توان دید که :

$$w(t) = (\mu-\lambda)e^{-(\mu-\lambda)t} \qquad t>0 \tag{12-3-1}$$

و

$$W = E(T) = \frac{1}{\mu-\lambda} \tag{13-3-1}$$

روابط لیتل: روابط فوق نتایج مفید و جالبی به همراه دارند. یکی از این نتایج عبارتست از

$$W = W_Q + \frac{1}{\mu} \qquad (14\text{-}3\text{-}1)$$

اثبات این روابط چندان مشکل نیست. فرض کنیم "S" زمان سرویس باشد آنگاه $T=T_Q+S$ چون زمان سرویس و زمان انتظار در صف مستقل اند، پس:

$$E(T) = E(T_Q) + E(S)$$

در واقع $w(t)$ را در صورت وجود می توان به صورت پیچش $w_Q(t)$ و چگالی زمان سرویس به دست آورد.

بنابراین رابطه (14-3-1) نه تنها برای صف M/M/1 ، بلکه بطور کلی برای هر صفی کاربرد دارد.

یک رابطه دیگر بین L_Q و W_Q وجود دارد که از روابط (8-3-1) و(11-3-1) به دست می آید:

$$L_Q = \lambda W_Q$$

همچنین از (7-3-1) و(13-3-1) داریم:

$$L = \lambda W$$

معادله لیتل به دلیل کارهای لیتل در سال 1961 به فرمول لیتل معروف است.

در فصل دوم این تحقیق سامانه های صف بندی m/m/1 را، در حالتی که تعطیلی سرویس دهنده نیز در آن وجود داشته باشد، مورد بررسی قرار می دهیم و به کمک روابط لیتل، اندازه های مؤثر سامانه را به دست می آوریم.
در فصل سوم، سامانه های صف بندی با اولویت در سرویس دهی مورد بررسی قرار گرفته اند. در این حالت به علت اینکه مدل صف بندی به صورت m/m/1 در نظر گرفته نشده است، امکان استفاده از معادلات لیتل وجود ندارد وبه کمک تابع مولد احتمال اندازه سامانه، اندازه های مؤثر سامانه به دست آمده اند. همچنین با استفاده از فرضیات مدل یکسان، اندازه های مؤثر سامانه در حالت بدون اولویت نیز به دست آورده شده است و در فصل چهارم با استفاده از چند مثال عددی این دو مدل یعنی مدل با اولویت و مدل بدون اولویت ، از لحاظ تأثیر آنها بر طول صف، با یکدیگر مقایسه شده اند.

فصل دوم

صف با دو نوع ورودی، دو نوع سرویس و تعطیلی با شیوه برنولی

2-1مقدمه

خط مشی مورد نظر در این صف شامل دو نوع ورودی با نرخ های متفاوت است. هر دو ورودی فرایند پواسون با میانگین متفاوت در نظر گرفته می شوند.همچنین فرض می کنیم هر دو نوع سرویس دارای توزیع نمایی با میانگین متفاوت هستند. بدون کاستن از کلیت مسأله، فرض می کنیم که یک سرویس دهنده هردونوع سرویس را ارایه می دهد.هر متقاضی پس از ورود به سامانه در یک صف واحد برای دریافت سرویس منتظر می ماند. ارایه سرویس به ترتیب ورود (FCFS) است. پس از اتمام هر نوع سرویس، سرویس دهنده به شیوه برنولی وبا احتمال

معینی باجه را تعطیل می کند. دوره های تعطیلی، دارای توزیع نمایی بوده و پس ازاتمام دوره تعطیلی سرویس دهنده دوباره به سامانه برمی گردد واگر متقاضی در سامانه وجود داشته باشد با خط مشی فوق به ارایه سرویس می پردازد، در غیراین صورت جهت ارایه سرویس منتظر می ماند.

صف M/M/1 به عنوان ساده ترین نوع صف بطور کامل در [1] آمده است.همچنین در [2] مطالب مفیدی در موردصف M/G/1 آورده شده است که G به معنای سرویس با توزیع دلخواه است. در سه دهه اخیر توجه خاصی به این نوع صف ها با زمان تعطیلی سرویس دهنده صورت گرفته است.

در [1] به صف های با دو نوع ورودی ودو نوع سرویس اشاره شده است. درسال های اخیر کارهای ارزشمندی در مورد صف های با یک نو ع ورودی و دو نوع سرویس غیر متجانس و زمان تعطیلی سرویس دهنده صورت گرفته است که می توان به [3] و[4] اشاره کرد.

در این تحقیق به بررسی تعمیمی از حالت های فوق می پردازیم که درآن صف مورد نظر دارای دو نوع ورودی با نرخ های متفاوت و دارای توزیع پواسون؛ دو نوع سرویس غیرمتجانس با نرخ های متفاوت و دارای توزیع نمایی هستند. در این زمینه در منبع [5] نتایجی به دست آورده شده است.

در زندگی روزمره می توان نمونه هایی از این صف ها را در بیمارستان ها، خط تولید کارخانه ها، شبکه های رایانه ای، آرایشگاه ها و غیره یافت. ابتدا در بخش 2 به بیان تعریف های مقدماتی و مدل سازی ریاضی صف مورد نظر می پردازیم. سپس در بخش 3 معادلات حالت پایا را به دست می آوریم و به کمک آنها توابع مولد احتمال تعداد متقاضیان درسامانه وصف؛ همچنین متوسط زمان انتظار در سامانه و صف را محاسبه خواهیم کرد. در بخش 5 به بررسی حالت های خاصی می پردازیم که از نتایج این بحث حاصل می شوند.

2-2تعریف ها و مدل سازی ریاضی

یکی از مشکلات موجود در بررسی صف ها پیچیدگی محاسبات و ظهور فرمول های طولانی به واسطه تعداد پارامترهای موجود است. در صف مورد بررسی علیرغم وجود شرایط مختلف، با اتخاذ شیوه ای خاص سعی نموده ایم پارامتر های موجود را در ظاهر کاهش دهیم. این کار موجب آسان شدن محاسبات و بیان نسبتاً ساده فرمول ها و نتایج شده است. با این حال به طور معمول در کار با سامانه های صف انتظار محاسبات طولانی و طاقت فرسا داریم. براین اساس فرض های زیر را برای صف مورد نظر داریم:

1) دو نوع متقاضی متفاوت با فرایند پواسن و نرخ های $\alpha\lambda$ و λ $(1-\alpha)$ وارد سامانه می شوند. در صورتی که سرویس دهنده مشغول باشد، آنگاه صفی واحد با خط مشی سرویس به ترتیب ورود توسط این متقاضیان تشکیل می شود.