بسمه تعالی

دانشگاه علامه طباطبائی تهران

گروه آمار

عنوان

سـامـانه های صف بندی با ورودی گـروهی تحت محـدودیت سـرویس دهی

استاد راهنما : دکتر عبدالرحیم بادامچی زاده
استاد مشاور : دکتر محمد رضا صالحی راد

تهیه و تنظیم : سحر احمدی خرم

شهریور ۸۸

بسم الله الرحمن الرحیم

فهرست مطالب

فصل اول معرفی صف ها

فصل دوم سامانه صف M^X/G/۱ با تعطیلی

فصل سوم سامانه صف $M^X/(G_1,G_2)/1/V_S$ با محدودیت در پذیرش و زمان تنظیم تصادفی

فــصــل اول

معرفی صف ها

۱-۱ مقدمه

همه ما ناراحتی انتظار کشیدن در صف را تجربه کرده ایم. متاسفانه این پدیده با افزایش تراکم جمعیت و شهری شدن روز افزون جامعه بیش از پیش گسترش می یابد. با ازدیاد جمعیت جهان و مسایل گوناگون و پیچیده مرتبط با آن، برای موفقیت در انجام امور اجتماعی لازم است که همه افراد حقوق یکدیگر را محترم شمرده و برای گرفتن سرویس از جامعه، نوبت یکدیگر را رعایت نمایند.

اگر تابع این روش باشیم که در مقابل هر متقاضی فوراً سرویس لازم را انجام دهیم، آنگاه علاوه بر بروز مسائل متعدد، از نظر اقتصادی نیز به صرفه نخواهد بود. پس باید بپذیریم که انتظار جزئی از زندگی ما را تشکیل می دهد. ما همه روزه برای دریافت خدمات، در ترافیک، در فروشگاه های بزرگ، در آرایشگــاه ها و سالن های زیبایی، دراداره پست و در پمپ بنزین ها، برای پرداخت عوارض راه و ... به صف می ایستیم.

ما به عنوان متقاضی این انتظار کشیدن را دوست نداریم و مدیران موسساتی هم که ما در صف های آنها نوبت گرفته ایم، به این دلیل که انتظار ممکن است با اعمال هزینه هایی همراه باشد، انتظار کشیدن ما را دوست ندارند، پس چرا باید انتظار کشید؟

یک پاسخ ساده به این سوال اینست، "زیرا درخواست برای سرویس، بیش از امکانات موجود برای سرویس دهی است." ولی چرا این چنین است؟ برای پاسخ می توان دلایل بسیاری را شمرد. برای مثال ممکن است تعداد سرویس دهندگان موجود کم باشد، یا فراهم نمودن تعداد کافی سرویس دهنده برای جلوگیری از تشکیل صف، از نظراقتصادی به صرفه نباشد؛ یا ممکن است فضای لازم برای ارائه بهتر سرویس محدود باشد و ... .

نظریه صف بندی می کوشد تا به سوال هایی نظیر " چه مدت متقاضی باید در صف انتظار بکشد؟ " و " چند متقاضی در صف خواهند بود؟ " و ... بر اساس تحلیل ریاضی پاسخ مبسوط دهد. اینجاست که نظریه صف اهمیت ویژه ای می یابد. بر اساس این نظریه، مسئولین امر برای سرویس دهی به مشتریان خود می توانند به گونه ای برنامه ریزی کنند تا باعث ناراحتی و از دست دادن آن ها نگردند.

نظریه صف

مطالعه ریاضی صف های انتظار بر اساس اندازه صف، متوسط زمان انتظار در صف، متوسط زمان لازم برای انتظار در کل سامانه موضوعات اساسی نظریه صف است.

اصول نظریه صف را می توان از نظر ریاضی به دو گروه تقسیم نمود.

الف- صف هایی که دارای یک توزیع مشخص بوده و در آن ها خط مشی ورودی و سرویس دارای فرمول مشخصی می باشد.

ب- صف هایی که دارای توزیع کلی یا تجربی (فرضی) بوده که با استفاده از شبیه سازی مورد مطالعه قرار می گیرند.

در زندگی روزمره با تعداد زیادی ازصف ها مواجه هستیم که برخی از آن ها را می توان برنامه ریزی نمود و بعضی غیر قابل پیش بینی است. مدت زمانی که در این صف ها به انتظار سپری می شود گاهی ممکن است یک هفته، یک ماه، یا حتی سال ها باشد.

تاریخچه صف

اولین کاربرد واقعی و مساله ی اولیه ای که در زمینه صف مطرح شد مربوط به تراکم در خواست مکالمات تلفنی، طراحی وتحلیل شبکه های تلفن بود. محقق پیشگام این نظریه، ریاضیدان دانمارکی به نام ارلانگ[۱] بود که در سال ۱۹۰۹ نظریه احتمالها و مکالمات تلفنی را منتشر نمود. وی را باید به حق بانی نظریه صف دانست. ارلانگ کارمند شرکت مخابرات دانمارک بود و وظیفه اش استفاده از روش های ثابت شده احتمال در تعیین بهینه استفاده از خطوط تلفن بر اساس درخواست های مکالمه بود.

وی دریافت که یک سیستم تلفن عموماً به یکی از دو صورت زیر مشخص می شود :

۱) ورودی پواسون و زمان های اشتغال نمایی و باجه های چندگانه (سرویس دهنده ها)

۲) ورودی پواسون، زمان های اشتغال ثابت و یک باجه

پس از ارلانگ نیز تحقیقات مربوط به کاربرد این نظریه در مورد تلفن ادامه یافت. در سال ۱۹۲۷ مولینا[۲] کاربرد نظریه احتمال در مسائل صف را منتشر کرد. در اوایل سال ۱۹۳۰ فیلیکس پولاچک[۳] بعضی کارهای پیشینیان برای ورودی پواسون، خروجی دلخواه و مسائل یک باجه ای را دنبال کرد و در همان ایام کار دیگری در روسیه به وسیله کولموگروف[۴] و خین- چین[۵] و در سوئد به وسیله پالم[۶] انجام گرفت.

۱)Erlang
۲)Molina
۳)Pollaczek
٤)Kolmogorov
٥)Khintchine
٦)Palm

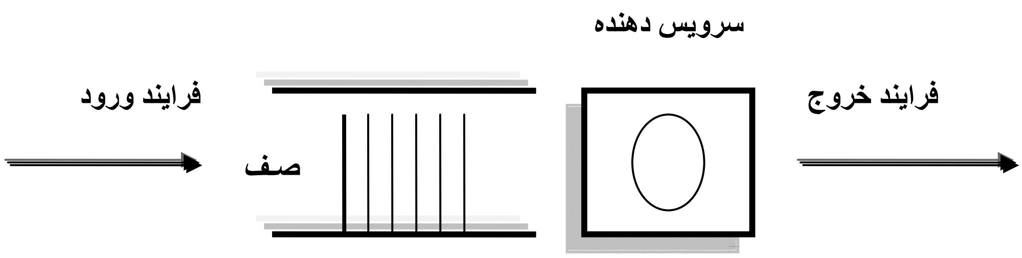
نخستین کارها در نظریه صف نسبتاً به کندی انجام می گرفت؛ اما از سال ۱۹۵۰ به بعد روند آن تغییر نموده و در این اواخر کارهای قابل توجهی در این زمینه انجام شده است.

۱-۲ مثال ها و مفاهیم مقدماتی

بسیاری از سامانه ها دارای مولفه هایی هستند که به کمک مفهوم صف، مدل سازی می شوند. بیشتر ایده های این مفهوم از تجربه های روزمره درصف باجه ی بانک ها، فرودگاه ها، سوپرمارکت ها، اسکله ها و موارد مشابه گرفته شده اند. یک صف به معنای علمی عبارتست ازسامانه ای که در آن جمعیتی از کاربران پیش از ترک سامانه ظرفیتی از آن را اشغال می کنند. این کاربران در سامانه توسط یک سرویس دهنده یا بیشتر سرویس می شوند. بنابراین یک سامانه صف را می توان به عنوان توصیفی از جمعیت ورودی و نوع تقاضای هر کاربر و همچنین خط مشی سرویس تعریف کرد. در ذیل با آوردن چند مثال به تعریف بهتر سامانه های صف می پردازیم.

مثال ۱-۲-۱: صف تک سرویس دهنده:

الف) صف تک سرویس دهنده با یک مرحله سرویس

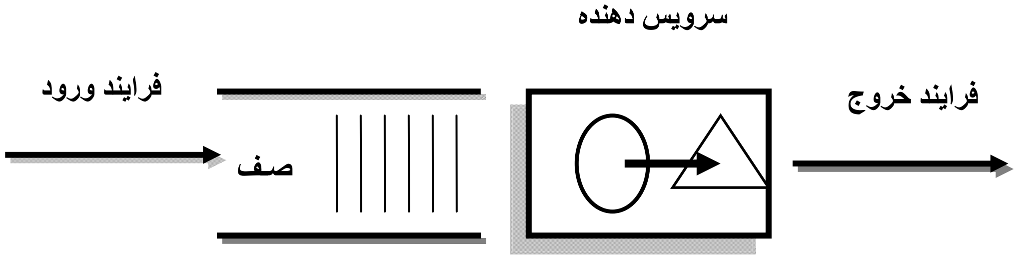


شکل ۱-۲-۱ صف تک سرویس دهنده با یک مرحله سرویس

صفی که در مقابل باجه پرداخت قبوض یک بانک تشکیل می شود، ساده ترین توصیف از یک سامانه صف است. در این سامانه یک جریان ورودی و یک سرویس دهنده بر اساس ورود متقاضیان به آن ها

سرویس می دهد. این خط مشی سرویس که در آن هیچ اولویتی برای سرویس دهی به متقاضیان وجود ندارد و متقاضیان به ترتیب ورود به سامانه، سرویس دهی می شوند (FCFS)[۷] نامـیده می شود.

ب) صف تک سرویس دهنده با دو مرحله سرویس متوالی و نامتجانس



شکل ۱-۲-۲ صف تک سرویس دهنده با دو مرحله سرویس متوالی و غیر متجانس

مثال ۱-۲-۲: کاربرد دیگر صف در ترافیک فرودگاه ها است. در این حالت باند فرود، سرویس دهنده و هواپیماهای در حال فرود، متقاضیان سامانه هستند.

مثال ۱-۲-۳: خطوط تولید در یک کارخانه مثال دیگری از یک سامانه صف است. ورودی شامل قطعاتی است که باید تعمیر و یا پردازش شوند. در همین سامانه صف، ممکن است به دلایلی سرویس دهنده خراب شده و از کار بیفتد؛ این اتفاق منجر به تغییراتی در صف ورودی خواهد شد. صف هایی با این ویژگی را صف با تعطیلی گویند؛ که یکی از موارد مدنظر این پایان نامه می باشد.

مثال ۱-۲-۴: از کاربرد های به روز نظریه صف در شبکـــه های رایانه ای و اینترنتی است. برای مثال دریک شبکه رایانه ای، سرویس دهنده تقاضاهایی از کاربران دریافت کرده و به آنها سرویس می دهد. تقاضا در اینجا به عنوان جریان ورودی به سامانه تعبیر می شود. زمان سرویس هر متقاضی در این حالت به تعداد متقاضیان حاضر وابسته است.

با این چند مثال می توان توصیفی از یک صف ارائه نمود.

یک سامانه صف عبارتست از یک یا چند سرویس دهنده، یک طرح ورودی از متقاضیان، طرح سرویس، و نظم صف. واژه "صف" اغلب برای توصیف کل سامانه به کار می رود، ولی در واقع بخشی از سامانه است

---

۷) First come first serve

که متقاضیان منتظر سرویس را در خود جای داده است. لذا تعداد متقاضیان در سامانه در هر زمان دلخواه برابر تعداد متقاضیان در صف و تعداد متقاضیان در سرویس است. البته این تعداد بر حسب زمان ورود و عزیمت متقاضیان تغییر می کند و در نتیجه یک متغیر تصادفی است. در زیر مفاهیم اساسی در یک سامانه صف را می آوریم:

طرح ورود : الگوی ورود به یک سامانه صف بندی را گویند که معمولا بر حسب متوسط تعداد ورود در واحد زمان یا میانگین نرخ ورود یا میانگـین فـاصله زمـانی دو ورود متوالـی بیان می شود و با λ نشان می دهیم. طرح ورود تصادفی است. مراجعات گاهی به صورت تکی و گاهی گروهی صورت می گیرد.

طرح سرویس : الگوی سرویس می تواند به وسیله ی نرخ سرویس (تعداد متقاضیانی که در واحد زمان سرویس می شوند) و یا به وسیله زمان لازم برای سرویس یک متقاضی توصیف شود. آن را با μ نشان می دهیم.

آرایش های سرویس: برای ارائه سرویس به متقاضیان همواره یک یا چند مکان برای سرویس در نظر گرفته می شود. تعداد بهینه سرویس دهندگان به تعداد متقاضیان، نرخ ورود و زمان لازم برای ارائه سرویس هر متقاضی بستگی دارد. بر حسب این متغیرها محیط سرویس تکی و یا چند تایی است.

زمان سرویس: زمان لازم برای اتمام سرویس یک متقاضی را گوییم. زمان سرویس ممکن است ثابت باشد و یا برحسب نوع تقاضای متقاضی تغییر کند. غالباً برای ساده شدن فرض می شود که زمان سرویس برای همه متقاضیان یکسان و ثابت باشد. از آنجا که طرح ورودی تصادفی است لذا زمان سرویس نیز تصادفی می باشد.

خط مشی سرویس: ترتیب انتخاب متقاضیان برای سرویس را گویند. اولین ورود ، اولین سرویس (FCFS) یا آخرین ورود اولین سرویس (LCFS[۸]) یا سرویس به ترتیب تصادفی و یا سرویس با اولویت از این دسته اند.

---

۸) Last come first serve

مراحل سرویس: گاهی سامانه صف تنها شامل یک مرحله سرویس است مانند آرایشگاه ها، و گاهی دو یا چند مرحله سرویس را در بر می گیرد مانند خطوط تولید کارخانه ها. بر این اساس سامانه های شامل چند مرحله سرویس نیز بر اساس تعداد سرویس دهنده ها تنوع دارند:

الف) سامانه های صف چند مرحله ای با تک سرویس دهنده که مورد نظر این پایان نامه است.

ب) سامانه های صف چند مرحله ای با C سرویس دهنده.

اندازه سامانه: تعداد متقاضیانی که در هر لحظه از زمان در سامانه هستند.

زمان انتظار در سامانه: مدت زمانی که یک متقاضی برای دریافت سرویس در سامانه منتظر می ماند.

دوره اشتغال : طول بازه زمانی که سرویس دهنده به طور پیوسته در سامانه مشغول به سرویس دهی می باشد.

دوره بیکاری : فاصله زمانی بین ۲ دوره اشتغال متوالی را گویند.

تعطیلی سرویس دهنده[۹] : در اکثر صف ها سرویس دهنده به دلایل گوناگون برای مدت تصادفی سرویس را تعطیل می کند. در یک خط تولید برای تعمیر و ارتقاء دستگاه، در باجه بانکی برای رفع خستگی کارمند و یا ارائه سرویس دیگر و یا صرف نهار و از این قبیل سرویس تعطیل می شود. واضح است که این اتفاق تاثیر اساسی در طول صف و زمان انتظار متقاضیان خواهد داشت.

پذیرش محدود[۱۰] : نوعی اعمال محدودیت است که طبق آن هر متقاضی در لحظه ورود اجازه اخذ سرویس را نخواهد داشت، بلکه متقاضیان در دوره اشتغال به سرویس دهی با احتمال C_1 اجازه ورود به سرویس را داشته و درزمان تعطیلی سرویس دهنده نیز اخذ سرویس با احتمال C_2 انجام می شود.

زمان تنظیم[۱۱] : زمان مورد نیاز برای آماده شدن سرویس دهنده برای امر سرویس دهی به متقاضیان را گویند که شامل مهیا شدن ابزار، ماشین و یا فرایند برای پذیرفتن متقاضیان است.

---

۹) Vacation time
۱۰)Restricted admissibility
۱۱)Setup time

عامل بهره دهی : این مقدار را با ρ نمایش و به صورت زیر تعریف می کنیم:

عامل بهره دهی(ρ) = متوسط نرخ سرویس / متوسط نرخ ورود

این مقدار در واقع احتمال وجود صف درسامانه است. اگر $\rho >1$ ، آنگاه شدت ترافیک بالا بوده و تقاضا برای دریافت سرویس بیشتر است، در نتیجه صف طولانی و زمان انتظار زیاد خواهد بود.

نرخ بیکاری: نرخ بیکاری در سامانه در هر زمان دلخواه از رابطه زیر حاصل می شود:

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

و بر این اساس میانگین زمان بیکاری عبارتست از:

$$\text{میانگین زمان بیکاری} = (1-\frac{\lambda}{\mu}) \times \text{زمان کل سرویس}$$

۳-۱ سامانه صف M/M/۱ [۱۲]

این صف از ساده ترین و معمول ترین مدل های صف است که حالت خاصی برای اکثر سامانه های صف می باشد. نتایج این مدل از صف، الهام بخش مطالعه سامانه های صف بندی دیگر و تاییدی بر روابط آنها است.

در این مدل تک باجه ای، توزیع ورودی پواسون و توزیع سرویس نمایی است. در این سامانه ورودی ها و خروجی ها در طول زمان تصادفی و از یکدیگر مستقل بوده و نظم صف FCFS است که حالت خاصی از فرایند زاد و مرگ می باشد. می توان ورودی ها را به عنوان زاد تعبیر کرد بدین صورت که اگر فرض کنیم سامانه در حالت E_n باشد (یعنی در لحظه ای خاص از زمان n متقاضی در سامانه حضور داشته باشد) و یک ورودی رخ دهد آنگاه سامانه به حالت E_{n+1} می رود. همچنین اگر عزیمتی رخ دهد حالت E_n به E_{n-1} تغییر می یابد، که به عنوان مرگ تعبیر می شود.

۱۲) مطالب این بخش و بخش های بعدی این فصل برگرفته از [۷] می باشد.

فرض می کنیم ورودی ها در این مدل از فرایند پواســـون با نرخ λ تبعیت می کند و زمان سرویس تک سرویس دهنده سامانه دارای توزیع نمایی با نرخ μ (میانگین $\frac{1}{\mu}$) 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 E_n تا لحظه $(t + \Delta t)$ حالت های زیر را داریم:

۱) سامانه تا لحظه t در حالت E_n قرار دارد و هیچ ورودی یا سرویس در فاصله زمانی Δt رخ نمی دهد.

۲) سامانه تا لحظه t در حالت E_n قرار دارد و فقط یک ورودی و یک سرویس در فاصله زمانی $(t, t + \Delta t)$ اتفاق می افتد.

۳) سامانه در لحظه t در حالت E_{n-1} قرار دارد و در طول بازه زمانی کوچک Δt یک ورودی رخ داده و سرویسی تکمیل نشود.

۴) سامانه در لحظه t درحالت E_{n+1} قرار دارد و در طول بازه زمانی Δt یک سرویس تکمیل شده و هیچ ورودی رخ ندهد.

چون برای هر $n \geq 1$ ورودی ها و سرویس ها از یکدیگر مستقل اند، لذا داریم:

$$p_n(t+\Delta t) = p_n(t)\, p[\Delta t \text{ یک ورودی در فاصله }]\; p[\Delta t \text{ یک سرویس در فاصله }]$$
$$+\, p_{n+1}(t)\, p[\Delta t \text{ هیچ ورودی در فاصله }]\; p[\Delta t \text{ یک سرویس در فاصله }]$$
$$+\, p_{n-1}(t)\, p[\Delta t \text{ يك ورودی در فاصله }]\; p[\Delta t \text{ هیچ سرویس در فاصله }]$$
$$= p_n(t)[1-\lambda\Delta t][1-\mu\Lambda t] + p_n(t)[\lambda\Delta t][\mu\Delta t] + p_{n+1}(t)[1-\lambda\Delta t][\mu\Delta t]$$
$$+\, p_{n-1}(t)[\lambda\Delta t][1-\mu\Lambda t]$$

با صرفنظر از جملات درجه دو و بالاتر بر حسب Δt داریم:

$$p_n(t)[1-(\lambda+\mu)\Delta t] + p_{n-1}(t)[\lambda\Delta t][1-\mu\Delta t] + p_{n+1}(t)[\mu\Delta t][1-\lambda\Delta t]$$

در نتیجه:

$$\lim_{t\to 0} \frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = -(\lambda+\mu)\, p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t) \qquad (۱-۳-۱)$$

حال چون به ازای $n=0$ احتمال $P_{n-1}(t)$ تعریف نشده است، لذا $P_0(t+\Delta t)$ را به طور مجزا محاسبه می کنیم. سامانه می تواند درلحظه $(t+\Delta t)$ در حالت E_0 باشد اگر در زمان t در حالت E_0 بوده و هیچ ورودی درطول زمان Δt رخ ندهد و یا سامانه در زمان t در حالت E_1 قرار داشته و هیچ ورودی رخ ندهد 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 $(\Delta t)^2$ داریم:

$$\lim_{\Delta t\to 0} \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) \qquad (۱-۳-۲)$$

معادلات (۱-۳-۱) و (۱-۳-۲) را معادلات دیفرانسیلی تفاضلی گویند.

تعریف:

گوییم یک سامانه در حالت پایاست هرگاه رفتار سامانه مستقل از زمان باشد. اگر $p_n(t)$ نشانگر احتمال وجود n متقاضی در سامانه در زمان t باشد، آنگاه در حالت پایا داریم:

$$\lim_{t\to\infty} p_n(t) = p_n$$

$$\lim_{t\to\infty} \frac{d}{dt} p_n(t) = 0$$

معادلات حالت پایا:

در صورتیکه $\frac{\lambda}{\mu}<1$ باشد سامانه به متقاضیانی که نرخ ورودشان از نرخ سرویس آن ها کوچکتر است سرویس می دهد و اگر $\lambda=\mu$ شود، آنگاه صفی وجود ندارد و اگر $\frac{\lambda}{\mu}>1$، آنگاه حالت انفجار برای سامانه رخ خواهد داد. با فرض برقراری حالت پایا، معادلات (۱-۳-۱) و (۱-۳-۲) به صورت زیر خواهند بود:

$$0 = -(\lambda+\mu)p_n + \mu p_{n+1} + \lambda p_{n-1} \qquad (۱-۳-۳)$$

$$0 = -\lambda p_0 + \mu p_1 \longrightarrow p_1 = \frac{\lambda}{\mu} p_0 \qquad (۱-۳-۴)$$

معادلات اخیر را معادلات تعادلی گویند.

معادله (۱-۳-۳) یک معادله تفاضلی مرتبه ۲ است که با شرط اولیـه (۱-۳-۴) جواب آن عبارتست از:

$$p_n = (\frac{\lambda}{\mu})^n p_0 \qquad n \geq 0 \qquad (۱-۳-۵)$$

برای محاسبه P_0 ، از شرط $\sum_{n=0}^{\infty} p_n = 1$ و معادله (۱-۳-۵) داریم:

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

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

چون $\frac{\lambda}{\mu} < ۱$، لذا سری فوق یک سری هندسی با قدر نسبت کمتر از یک است لذا

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

که در آن $\rho = \frac{\lambda}{\mu} < 1$. بنابراین

$$p_n = (\frac{\lambda}{\mu})^n p_0 = (\frac{\lambda}{\mu})^n (1 - \frac{\lambda}{\mu}) = \rho^n (1 - \rho) \qquad (۱-۳-۶)$$

اگر فرض کنیم N تعداد متقاضیان در سامانه باشد (شامل متقاضیان منتظر در صف و متقاضی در حال دریافت سرویس) در این صورت **متوسط این تعداد** برابر است با

$$L = E(N) = \sum_{n=0}^{\infty} np_n = (1-\rho)\sum_{n=0}^{\infty} n\rho^n$$

$$= (1-\rho)\rho \sum_{n=0}^{\infty} n\rho^{n-1} = (1-\rho)\rho(1-\rho)^{-2}$$

$$= \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda} \qquad (۱-۳-۷)$$

و اگر فرض کنیم N_Q تعداد متقاضیان در صف باشد در این صورت **متوسط این تعداد** 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

$$L_Q = E(N_Q) = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} \qquad (۸\text{-}۳\text{-}۱)$$

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

$$W_Q(0) = P[T_Q \le 0] = P[T_Q = 0]$$

$= P$ [بدو ورود سامانه تهی باشد] $= p_0 = 1-\rho$ (۹-۳-۱)

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

$$W_Q(t) = P[T_Q \le t]$$

$= \sum_{n=1}^{\infty} P$ [مراجعه کننده n متقاضی را در سامانه ببیند | تکمیل سرویس n متقاضی در زمان کمتر از t] p_n

$$+W_Q(0) = (1-\rho)\sum_{n=1}^{\infty} \rho^n \int_0^t \frac{\mu(\mu x)^{n-1}}{(n-1)!} e^{-\mu x} dx + (1-\rho)$$

$$= (۱-\rho)\ \rho \int_0^t \mu e^{-\mu t} \sum_{n=1}^{\infty} \frac{(\mu x \rho)^{n-1}}{(n-1)!} dx + (1-\rho)$$

$$=\rho\,(۱\text{-}\rho)\int_0^t \mu e^{-\mu x(1-\rho)}dx \;+\; (1-\rho)$$

$$= 1 - \rho e^{-\mu(1-\rho)t} \qquad\qquad (t>0)$$

بنابراین توزیع زمان انتظار در صف عبارتست از

$$W_Q(t) = \begin{cases} 1-\rho & t=0 \\ 1-\rho e^{-\mu(1-\rho)t} & t>0 \end{cases} \qquad (۱۰\text{-}۳\text{-}۱)$$

پس **امید ریاضی زمان انتظار** برابر است با

$$W_Q = E(T_Q) = \int_0^\infty t\, dW_Q(t)$$

$$= 0.(1-\frac{\lambda}{\mu}) + \int_0^\infty t\frac{\lambda}{\mu}(\mu-\lambda)e^{-(\mu-\lambda)t}dt = \frac{\lambda}{\mu}\int_0^\infty t(\mu-\lambda)e^{-(\mu-\lambda)t}\, dt$$

با محاسبه انتگرال داریم:

$$W_Q = \frac{\lambda}{\mu(\mu-\lambda)} \qquad (۱۱\text{-}۳\text{-}۱)$$

فرض می کنیم متغیر تصادفی 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 باشند، آنگاه با روشی مشابه روش فوق می توان دید که

$$w(t) = (\mu-\lambda)e^{-(\mu-\lambda)t} \qquad t>0 \qquad (۱۲\text{-}۳\text{-}۱)$$

و

$$W = E(T) = \frac{1}{\mu-\lambda} \qquad (۱۳\text{-}۳\text{-}۱)$$

روابط لیتل[۱۳] **:** روابط فوق نتایج مفید وجالبی به همراه دارند. یکی از این نتایج عبارتست از:

$$W = W_Q + \frac{1}{\mu} \qquad (۱۴\text{-}۳\text{-}۱)$$

اثبات این رابطه چندان مشکل نیست. فرض می کنیم S زمان سرویس باشد آنگاه $T = T_Q + S$، در این صورت:

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

---

۱۳)Little

ونتیجه مورد نظر به دست می آید. 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 (۱-۳-۱٤) نه تنها برای صف M/M/۱ ، بلکه به طور کلی برای هر صفی کاربرد دارد.

بستگی دیگر بین L_Q و W_Q وجود دارد که از روابط (۱-۳-۸) و (۱-۳-۱۱) به صورت زیر به دست می آید:

$$L_Q = \lambda W_Q \qquad (۱۵\text{-}۳\text{-}۱)$$

همچنین از (۱-۳-۷) و (۱-۳-۱۳) داریم:

$$L = \lambda W \qquad (۱٦\text{-}۳\text{-}۱)$$

معادله (۱-۳-۱٦) به دلیل کارهای لیتل در سال ۱۹۶۱ به فرمول لیتل معروف گردید. این رابطه را به صورت شهودی نیز می توان توجیه کرد. برای مثال متقاضیی را درنظر بگیرید که اکنون مراجعه کرده است . به طور متوسط او پس از زمان W_Q به سرویس می رسد. با ورود به سرویس، پشت سر خود نگاه کرده و افراد بعد از خود را می شمارد. این تعداد به طور متوسط برابر L_Q است. همچنین آمدن هر یک از L_Q نفر به طور متوسط $\frac{1}{\lambda}$ وقت گرفته است و کل زمانی که طول می کشد تا L_Q مراجعه کننده پشت سر او گرد آیند باید برابر زمان انتظارش باشد، بنابراین $L_Q(\frac{1}{\lambda}) = W_Q$. اثبات های دقیق و منطقی تری نیز برای روابط لیتل وجود دارند که از حوصله این نوشته خارج است.

۱-۴ سامانه صف M/G/۱

در بررسی سامانه صف M/M/۱ با فرایند مارکوف مواجه بودیم و نتایج زنجیرهای مارکوف در مطالعه این سامانه به کار رفت. در سامانه M/G/۱ توزیع زمان سرویس فاقد خاصیت بی حافظگی (توزیع نمایی) است. فرایند $N(t)$ به عنوان حالت یا اندازه سامانه در زمان t دیگر مارکوفی نیست. با این حال این

فرایند بر اساس فرایندی مشابه که مارکوفی باشد، تحلیل می شود. به طور کلی برای انجام این کار دو روش وجود دارد:

۱) روش منسوب به کـندال[۱۴] که به زنجیر مارکوف نشانده شده معروف است.

۲) روش متغیـر تکمیـلی که به کـیلسون[۱۵] و کوهـاویان[۱۶] منسوب است.

بر این اساس روش زنجـیرهای نشـانده شده مارکوف رایـج تر است که به توضیـــح آن می پردازیم.

سامانه صف تک سرویس دهنده ای را در نظر بگیرید که فرایند ورودی آن پواسون با نرخ λ است. زمان های سرویس را با B نمایش می دهیم که متغیرهای تصادفی مستقل و هم توزیع با میانگین $\frac{1}{\mu}$ و تابع توزیع F_B و تابع چگالی (در صورت وجود) f'_B است. خط مشی صف FCFS است. در حالت خاص اگر F_B را توزیع نمایی با پارامتر μ در نظر بگیریم آنگاه سامانه صف M/M/۱ به دست می آید و اگر زمان های سرویس ثابت در نظر گرفته شوند، آنگاه سامانه صف M/D/۱ به دست می آید.

به ازای n=۱،۲,... ، فرض می کنیم t_n نشانگر n امین دوره عزیمت باشد و $t_0 = 0$. به عبارت دیگر t_n زمانی است که n امین متقاضی سرویس خود را تکمیل کرده و سامانه را ترک می کند. نقاط t_n را نقاط تولید مجدد فرایند $N(t)$ گویند. فرض می کنیم X_n تعداد متقاضیان در سامانه در زمان t_n باشد آنگاه

$$X_n = N(t_n) \qquad n = 1,2,... \tag{۱-۴-۱}$$

فرایند تصادفی $\{X_n \quad : \quad n = 1,2,...\}$ یک زنجیر مارکوف گسسته پارامتر است که به زنجیر مارکوف نشانده شده ی فرایند تصادفی پیوسته پارامتر $\{N(t) \quad : t \quad \geq 0\}$ معروف است.

فرض کنیم متغیر تصادفی A_n برابر تعداد متقاضیان ورودی در طول زمان سرویس n امین متقاضی است آنگاه

---

۱٤)Kendall
۱٥)Keilson
۱٦)Koahauvian

$$X_{n+1} = \begin{cases} X_n - 1 + A_{n+1} & X_n \geq 1 \\ A_{n+1} & X_n = 0 \end{cases} \qquad (۲-۴-۱)$$

به عبارت دیگر اگر $X_n = 0$ و متقاضی بعدی $(n+1)$ امین باشد، آنگاه در طول مدت سرویس گرفتن او A_{n+1} متقاضی وارد می شوند و $(n+1)$ امین متقاضی در زمان t_{n+1} سامانه را ترک کرده و A_{n+1} متقاضی پشت سر او باقی می مانند. اگر $X_n > 0$ ، آنگاه تعداد متقاضیان باقی مانده پشت سر $(n+1)$ امین متقاضی برابر $X_n - 1 + A_{n+1}$ است. چون ورودی ها از فرایند پواسون تبعیت می کنند، لذا متغیر تصادفی A_{n+1} فقط به زمان سرویس وابسته است و به طول صف یا لحظه شروع سرویس بستگی ندارد، بنابراین آن را با A نمایش می دهیم. در نتیجه داریم:

$$P[A=r] = \int_0^{+\infty} P(A=r \,|\, \text{زمان سرویس یک متقاضی برابر } t \text{ باشد})\, dF_B(t)$$

که در آن

$$P(A=r \,|\, \text{زمان سرویس یک متقاضی برابر } t \text{ باشد}) = \frac{e^{-\lambda t}(\lambda t)^r}{r!}$$

لذا توزیع A یعنی تعداد ورودی ها در طول زمان سرویس یک متقاضی برابر است با :

$$a_r = P[A=r] = \int_0^{+\infty} \frac{e^{-\lambda t}(\lambda t)^r}{r!} f_B(t)\,dt \qquad (۳-٤-۱)$$

و واضح است که $\sum_{j=0}^{\infty} a_j = 1$. همچنین با توجه به احتمال های

$$p_{ij} = P[X_{n+1}=j \mid X_n=i] = \begin{cases} P(A = j-i+1) = a_{j-i+1} & i \geq 1, j \geq i-1 \\ P(A = j) = a_j & i = 0, j \geq 0 \\ 0 & o.w \end{cases} \qquad (۴-۴-۱)$$

به وضوح می توان دید که $\{X_n : n \geq 0\}$ یک زنجیر مارکوف با ماتریس احتمال تغییر وضعیت زیر است: