

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

مقاله مروری: درس ارزیابی سیستم‌های کامپیوتری

تجزیه و تحلیل ازدحامهای نظیر به نظیر از دیدگاه صف: توزیع‌های ایستا و حدهای مقیاس گذاری آنها

Queueing analysis of peer-to-peer swarms: stationary distributions and their scaling limits

نویسندگان:

آندرس فرگوت، فرناندو پاگانینی

مرور به فارسی:

امیرحسین کارگران خوزانی

استاد درس:

دکتر علی موقر

۱۳۹۹

فهرست مطالب

صفحه	عنوان
سه	فهرست مطالب
۱	چکیده
فصل اول: مقدمه	
۳	۱-۱ پیشینه تحقیق
۳	۲-۱ ساختار گزارش و اهداف تحقیق
فصل دوم: مدل سیستم	
۵	۱-۲ یک مدل صف برای سیدرهای ثابت و توزیع ایستای آن‌ها
۶	۲-۱-۱ حالت درخواست‌های کار عمومی
۷	۲-۱-۲ توصیف‌گر حالت نقطه پروسس
۸	۲-۲ توزیع حالت سکون تحت مقیاس‌گذاری بزرگ سیستم
۸	۲-۲-۱ حالت سیدر پایدار ($\rho < y_0$)
۸	۲-۲-۲ حالت پایداری عمومی ($\rho > y_0$)
۹	۳-۲ تقریب حد مرکزی دانلود نمایه
۹	۴-۲ سیدرهای متغیر و شبه ایستا
فصل سوم: شبیه‌سازی و نتیجه‌گیری	
۱۱	۱-۳ آزمایشات شبیه‌سازی
۱۱	۱-۱-۳ حالت سیدر پایدار
۱۲	۲-۱-۳ حالت پایداری عمومی
۱۳	۳-۱-۳ سیدرهای با تغییر آرام
۱۳	۲-۳ نتیجه‌گیری
۱۵	مراجع

چکیده

این مقاله پویایی یک ازدحام تبادل فایل نظیر به نظیر از دیدگاه صف را بررسی می‌کند. در چنین سیستم‌هایی، نرخ سرویس دریافت شده توسط یک نظیر به دو چیز وابسته است. مورد اول در بیشتر موارد یک مولفه ثابت مانند سرور یا سیدر است و مورد دوم تعداد نظیرهای در حال حاضر در سیستم است. در این مقاله یک کلاس از صف‌های M/G اشتراک پردازنده بررسی می‌شود. این کلاس جمعیت و بار کاری باقی‌مانده در این شرایط را توصیف می‌کند و رژیم ایستای آن را در حالت تعداد سرورهای ثابت مشخص می‌کند. نتایج نشان می‌دهد این به مانند ترکیب یک صف از نوع $M/G/1$ و صف‌هایی از نوع $M/G/\infty$ است. در ادامه حدهای مقیاسی بر روی این صف اعمال شده و دو رژیم محدودکننده، وابسته به اینکه مشارکت سرور یا نظیر، تبدیل به مشارکت غالب می‌شود یا خیر شناسایی می‌شود. برای مورد دوم توصیف بارگیری نمایه حد سیال با یک تقریب عملی گوسی اصلاح شده است. همچنین حالتی که تغییر آهسته جمعیت سرورها اتفاق می‌افتد نیز با بسط دادن حالت ثابت از طریق تحلیل شبه ایستا نیز تحلیل شده است. در این مقاله برای تایید عملی مقایسه شبیه‌سازی بسته‌ها پیشنهاد و به کار گرفته شده است.

کلمات کلیدی: نظیر به نظیر، صف‌های اشتراک پردازنده، اندازه‌های کار عمومی، حدهای سیال.

فصل اول

مقدمه

امروزه از سیستم^۱ های به اشتراک گذاری نظیر^۲ به نظیر فایل^۳ مانند بیت تورنت^۴ بیشتر استفاده می شود. این سرویس ها بخش زیادی از ترافیک اینترنت را به خود اختصاص داده اند. محتوا به قطعه های کوچک به نام چانک^۵ تقسیم می شود و نظیرها با استفاده از یک ارتباط دوطرفه این چانک ها را با یکدیگر به اشتراک می گذارند. قدرت اشتراک نظیر به نظیر در آن است که نظیرهای درحال بارگیری^۶، همزمان مشغول بارگذاری^۷ چانک ها به بقیه نظیرها نیز هستند. مجموعه نظیرها یک فایل مشخص که اغلب با عنوان ازدحام^۸ از آن یاد می شود را با یکدیگر تبادل می کنند. نظیرها به دو دسته کلاس سیدر^۹، که مانند سرور^{۱۰} عمل می کنند و لیچر^{۱۱}، که هم مانند سرور و هم مشتری^{۱۲} عمل می کنند تقسیم می شود. سیدرها همه فایل را به طور کامل در اختیار دارند و آن را برای بقیه بارگذاری می کنند و لیچرها همزمان هم بارگیری و هم بارگذاری فایل را انجام می دهند. این ازدحام ها در طی زمان و بر اساس رسیدن ها و خروج ها تکامل می یابند، بنابراین می توان که پویایی جمعیت آنها را با ابزارهای نظریه صف تحلیل کرد.

در مقاله ای که اولین مدل صف برای این مساله را ارائه کرده بود، نویسندگان یک زنجیره مارکوف^{۱۳} زمان پیوسته را برای دو صف M/M پشت سر هم ارائه دادند [۹]. این دو صف جمعیت لیچر و سیدر را نمایش می دهند. ورودی هایی از توزیع پواسون صف لیچر را تغذیه می کنند و هنگام تکمیل سرویس، نظیرها به صف سیدر هدایت می شوند و به مدت زمان نمایی در آن صف مانده و نهایتاً سیستم را ترک می کنند. ویژگی دیگر سیستم نظیر به نظیر در آن است که ظرفیت سرویسی که گذار^{۱۴} های درونی را کنترل می کند از

¹System

²Peer

³File

⁴BitTorrent

⁵Chunk

⁶Download

⁷Upload

⁸Swarm

⁹Seeder

¹⁰Server

¹¹Leecher

¹²Client

¹³Markov

¹⁴Transition

حالت سیستم مستقل است و با مجموع تعداد نظیرها رشد می‌کند. این مدل مارکوف به صورت ضمنی فرض بر نمایی و مستقل بودن نیازمندی‌های بارگیری برای نظیرها را دارد که یک فرض ناهنجار است چرا که نظیرها در حال اشتراک‌گذاری محتویات فایل مشابه هستند. با وجود ساده‌سازی در این تحقیق زنجیره خروجی یک راه حل تحلیلی^۱ را نتیجه نمی‌دهد و به صورت عددی مطالعه شده است. این کار و سایر مراجع مرتبط در بخش ۱-۱ بررسی شده‌اند.

۱-۱ پیشینه تحقیق

در [۹، ۲] زنجیره مارکوف به صورت ساده مدل شده است اما علی‌رغم سادگی نتایج تحلیلی به دلیل برگشت‌پذیر نبودن زنجیره مارکوف، به سادگی نتیجه نمی‌شود. در این زنجیره مارکوف ورود لیچرها یک فرآیند پوآسون با نرخ λ ، تعداد لیچرها و $y(t)$ تعداد سیدرها در سیستم در زمان t است. لیچرها با نرخ نمایی $\mu(\eta x + y)$ ، که η پارامتر بهره‌وری و μ نرخ بارگذاری یک نظیر به سیدرها تبدیل می‌شوند و سیدرها به مدت $\exp(\gamma)$ در سیستم خواهند ماند. این پیچیدگی تحلیلی موجب شد تا [۸] مدل معادله دیفرانسیلی‌ای با توصیف جمعیت‌ها به عنوان متغیرهای واقعی و ساختار مشابه قبل را ارائه کند. در این حالت، تعادل^۲ به سادگی تعیین می‌شود. همچنین در [۱] نرخ‌های ناهمگن بارگیری نیز در نظر گرفته شده است.

اما در مدل‌های با احتساب جمعیت‌ها به عنوان متغیرهای حالت، یک فرض ضمنی بدون حافظه بودن در نظر گرفته شده است که بدون فراهم کردن هیچ جزئیاتی در مورد بار کاری باقی‌مانده بر روی حالت توزیع اندازه کار عمومی اعمال می‌شود.

بنابراین در [۷] یک معادله دیفرانسیلی جزئی^۳ توسعه داده شده که جمعیت را به عنوان تابعی از کار باقی‌مانده برای یک ازدحام نظیر به نظیر با توزیع فایل عمومی دنبال می‌کند. در مقاله در حال مرور پویایی ازدحام شامل گذارها اما در فقط در جهت سیال توصیف می‌شود. در اینجا نتایج حالت پایداری پلی میان صف‌ها و حدهای سیال‌شان در تعادل ایجاد می‌کند و می‌توان این کار را مکمل کار [۷] دانست.

همچنین مقالات دیگری نیز وجود دارند که پیشرفت بارگیری را از طریق چندین صف پشت سر هم مدل کرده‌اند [۴، ۵]. با این وجود در این مقالات رابطه‌ای میان نرخ‌های سرویس با سیستم نظیر به نظیری که اشتراک‌گذاری فایل را میان همه کلاس‌های نظیرها انجام می‌دهد، وجود ندارد.

از طرفی دیگر مدل‌های مارکوفی ارائه شدند که جزئیات را در سطح نوشتن متغیرهای جمعیتی برای هر زیرمجموعه از تکه‌های تحت کنترل انجام می‌دهند [۳، ۶]. این موضوع موجب می‌شود تعداد متغیرهای حالت بزرگ و وابسته به تعداد تکه‌ها شود.

۲-۱ ساختار گزارش و اهداف تحقیق

در این مقاله نویسندگان به غلبه بر روی هر دو محدودیت عنوان شده در [۹] تمایل دارند. در این مقاله مشخصه‌های تحلیلی توزیع‌های ایستای نتیجه‌شده پیگیری شدند همچنین توزیع اندازه کار عمومی نیز برای نظیرها (دامنه صف‌های M/G) اجازه داده شده است. در حالت خاص این اجازه، با اندازه کارهای قطعی، تمام نظیرها می‌خواهند یک فایل اشتراک‌گذاری شده یکسان را بارگیری کنند.

¹Analytical

²Equilibrium

³Partial Differential Equation

در این مقاله مشخص می‌شود که هر دو هدف عنوان‌شده در صورت تمرکز بر روی مدل ازدحام لیچر با فرض ثابت بودن تعداد سیدرها، میسر خواهد بود، که این مدل نزدیک‌تری به واقعیت از سیستم‌های نظیر به نظیر است. در واقع در بیشتر موارد لیچرها بعد از بارگیری‌شان سیستم را ترک خواهند کرد و عملکرد سیدر توسط مجموعه‌ای از نظیرهای متعهد در نقش سرور پایدار می‌شود. در حالت M/M جمعیت لیچرها به یک فرآیند یک حالت تولد و مرگ تبدیل می‌شود که توزیع ایستای آن را می‌توان به صورت ترکیبی از صف‌های $M/M/1$ و $M/M/\infty$ در نظر گرفت. برای رفتن به حالت عمومی‌تر M/G نیازمند اطلاع از نحوه توزیع ظرفیت سرویس بین نظیرها در سیستم است. در نتیجه یک مدل اشتراک پردازنده‌ای برای سیستم‌هایی با پارامترهای نظیر همگن همانگونه که در ادامه در مورد آن توضیح داده شده به کار گرفته می‌شود. در این شرایط می‌توان ویژگی‌های غیرحساسی را برای مشخص کردن توزیع ایستای M/G بدست آورد. این مدل‌ها در بخش ۲-۱ بررسی شده‌اند.

هدف بعدی که در این مقاله به آن پرداخته می‌شود بدست آوردن توصیفات ساده‌ای از توزیع ایستا در محدودیت جمعیت زیاد است. در این مقاله حالت M/G از منظر اندازه نقطه‌ای روی خط واقعی مثبت مشخص می‌شود که جمعیت و کار باقی‌مانده را نشان می‌دهد. در بخش ۲-۲ با به کارگیری این تشریح‌کننده حالت، از طریق حد مقیاسی و با استفاده از ابزارهای تحلیلی، از پردازنده‌های نقطه‌ای برای اثبات همگرایی به قانون‌های مشخص حدی استفاده می‌شود. این به کارگیری موجب شناسایی دو حالت متفاوت در سیستم می‌شود. یک آن که مولفه سرور غالب باشد سیستم حتما دارای یک صف $M/G/1$ خواهد بود و دوم آن که اگر اشتراک‌گذاری فایل غالب باشد. در این حالت از طریق یک قانون مقیاس‌گذاری اعداد بزرگ، یک حد سیال قطعی که توصیف‌کننده حالت پایدار جمعیت نمایه^۱ لیچرهاست به عنوان یک تابع از باقی‌مانده کار آن‌ها بدست می‌آید. در بخش ۲-۳، یک مقیاس‌گذاری از نوع حد مرکزی اعمال می‌شود تا انحراف از حد نمایه با بدست آوردن یک مشخصه گوسی کاربردی از تغییر روند فرآیند مطالعه شود.

در بخش ۲-۴، به سناریو دو صف در [۹] باز می‌گردیم و اجازه می‌دهیم (در حالت M/M) تغییر آرام در جمعیت سیدر توسط یک نسخه نازک‌شده^۲ از فرآیند اتمام لیچر تغذیه شود. با آرام شدن این پروسس^۳ در سیدر یک تحلیل نیمه ایستا برای توزیع مشترک محدودکننده به عنوان یک ترکیب مناسب از قانون‌های سیدر ثابت که قبلا دیده شده بود در این بخش بدست می‌آید. شبیه‌سازی‌هایی که نتایج مربوط به شبکه بیت‌تورنت را تایید می‌کند در بخش ۳-۱ بیان شده‌اند. در آخر نتیجه‌گیری این مقاله در بخش ۳-۲ ارائه شده است.

¹ Profile

² Thinned

³ Process

فصل دوم

مدل سیستم

۱-۲ یک مدل صف برای سیدرهای ثابت و توزیع ایستای آنها

مقاله با مدل ارائه شده در [۹، ۲] و با تمرکز بر روی تکامل جمعیت لیچرها شروع می‌شود. به عنوان ساده‌ترین و گام‌های ابتدایی جمعیت سیدرها ثابت و برابر با y_0 در نظر گرفته می‌شود. همچنین فرض می‌شود که لیچرها خودخواه هستند و به محض اتمام بارگیری سیستم را ترک می‌کنند که اصلاً دور از واقعیت نیست. $x(t)$ برابر جمعیت لیچرها در این سیستم در نظر گرفته می‌شود و زمان‌های گذار نیز نمایی انتخاب می‌شود. ورود لیچرها یک فرآیند پواسون با نرخ λ و با نرخ بارگیری $R_{up} := \mu(x + y_0)$ است. μ در اینجا پهنای باند مشارکت هر نظیر در سیستم بر واحد فایل بر ثانیه است همچنین پهنای باند مشارکت انجام شده توسط لیچرها و سیدرها با یکدیگر برابر گرفته می‌شود. در نتیجه $x(t)$ یک زنجیره مارکوف زمان پیوسته با فضای حالت \mathbb{N} و نرخ‌های گذار q_{ij} خواهد بود. این زنجیره مارکوف به صورت معادله ۱-۲ مدل می‌شود که یک فرآیند تولد مرگ ساده است و می‌تواند به سادگی حل شود.

$$q_{x,x+1} = \lambda, \quad q_{x,x-1} = \mu(x + y_0), \quad x > 0 \quad (1-2)$$

گزاره ۱. اگر $\rho := \lambda/\mu$ باشد آنگاه توزیع تعادل برای تعداد لیچرها در فرآیند تولد مرگ ۱-۲ به صورت زیر است:

$$\pi(n) = \pi(0) \prod_{i=1}^n \frac{\lambda}{\mu(i + y_0)} = \left[\sum_{m=y_0}^{\infty} \frac{\rho^m}{m!} \right]^{-1} \frac{\rho^{n+y_0}}{(n + y_0)!} \quad \text{for } n \geq 0 \quad (2-2)$$

به طور خاص، سیستم برای هر λ ، μ و y_0 پایدار است.

توجه داشته باشید که این سیستم می‌تواند به عنوان ترکیبی از صف‌های $M/M/1$ و $M/M/\infty$ فکر شود. اگر از مشارکت لیچرها صرف نظر شود این سیستم به یک صف $M/M/1$ با بار $\lambda/(\mu y_0)$ کاهش می‌یابد و تنها برای حالت $\rho < y_0$ پایدار است. اگر از مشارکت سیدرها صرف نظر شود این سیستم به یک صف $M/M/\infty$ تبدیل می‌شود و سیستم برای تمام مقادیر ρ پایدار است. همچنین توجه داشته باشید در حالت $\rho > y_0$ نیز مشارکت لیچرها برای حفظ پایداری الزامی است.

همانگونه که ذکر شده بود در تحلیل بالا پهنای باند مشارکت انجام شده توسط لیچرها و سیدرها با یکدیگر برابر بودند. حال اگر تفاوتی وجود داشته باشد، مجموع پهنای باند موجود برابر $R_{up} = \mu x + \mu_0 + y_0$ خواهد بود که μ_0 مشارکت هر سیدر با امکان

برابر نبودن با μ است در مدل اصلی مطرح شده در [۹] رابطه $\mu \geq \mu_0$ نیز در نظر گرفته شده است که نابرابری عدم کارآمدی در تبادل لیچرها را بیان می‌کند.

اینجا یادآوری می‌شود که الگوریتم‌های اشتراک‌گذاری فایل در عمل بسیار کارآمد هستند و اگر یک نظیر پهنای باند اضافی داشته باشد کسی پیدا می‌شود که از آن عمل بارگیری را انجام دهد. بنابراین لیچرها بارگذاری بهینه‌ای دارند. در شبیه‌سازی‌های انجام شده در بخش ۳-۱ این مورد قابل رویت است

حالتی که $\mu \neq \mu_0$ اتفاق می‌افتد را در نظر بگیرید. در این حالت توزیع نامتغیر می‌تواند بر اساس روابطی مشابه معادله ۲-۲ مشخص شود:

$$\pi_\eta(n) = K \frac{\rho^n}{\Gamma(n + \frac{y_0}{\eta} + 1)} \quad \text{for } n \geq 0 \quad (3-2)$$

که در اینجا $\rho = \lambda/\mu$ همانند قبل و $\eta := \mu/\mu_0$ و K ثابت مناسب نرمال کردن است.

دلیل تغییرات بالا امکان صحیح نشدن عدد y_0/η است. در هر صورت می‌توان کران‌هایی را برای آن در گزاره ۲ نشان داد. گزاره ۲. فرض کنید $X \sim \pi_\eta$ نمایش‌دهنده جمعیت تصادفی لیچرها در سیستم در حالت تعادل باشد. X^u و X^l را جمعیت تعادل یک سیستم را به ترتیب با تعداد سیدرهای $\lceil y_0/\eta \rceil$ و $\lfloor y_0/\eta \rfloor$ در نظر بگیرید. در نتیجه:

$$X^l \leq_{st} X \leq_{st} X^u$$

که \leq_{st} بیانگر ترتیب تصادفی است. گزاره ۲ یک تقریب خوب از رفتار صف در حالت تعادل، زمانی که $\eta \neq 1$ است می‌دهد. در ادامه از تابع مولد احتمالی π^1 برای بدست آوردن چند سنجه استفاده می‌شود. پس قبل از استفاده از آن به صورت زیر تعریف می‌شود:

اگر $X \sim \pi$ باشد آنگاه تابع مولد احتمالاتی به صورت $G(z) = E[z^x]$ خواهد بود. تابع مولد احتمالاتی ۲-۲ برابر با:

$$G(z) = z^{-y_0} \frac{\sum_{m=y_0}^{\infty} (\rho z)^m / m!}{\sum_{m=y_0}^{\infty} \rho^m / m!} \quad (4-2)$$

خواهد بود.

رابطه ۲-۴ اجازه می‌دهد کارایی سیستم را از منظر تعداد میانگین لیچرها و میانگین زمان اتمام بارگیری تحلیل شود هر دو این سنجه‌ها از طریق گزاره ۳ بدست می‌آید.

گزاره ۳. برای $y_0 > 0$ ، میانگین تعداد لیچرها در سیستم به صورت زیر خواهد بود:

$$\bar{x} = \rho - y_0 + \frac{\rho^{y_0-1} / (y_0 - 1)!}{\sum_{m=y_0}^{\infty} \rho^m / m!} \quad (5-2)$$

اثبات گزاره ۳ از تکامل مستقیم $G'(1)$ دنبال می‌شود همچنین از طریق قانون لیتل^۲ می‌توان گفت میانگین زمان اتمام بارگیری برابر با $\frac{\bar{x}}{\lambda}$ است. همچنین می‌توان سنجه‌های دیگری نظیر $Var(x)$ را نیز توسط روش‌های مشابه محاسبه نمود.

۲-۱-۱ حالت درخواست‌های کار عمومی

فرض زمان سرویس نمایی در سیستم‌های نظیر به نظیر مناسب نیست چرا که درخواست‌های کار نمایی نیستند. بنابراین در نظر گرفته می‌شود که درخواست‌های کار یک توزیع عمومی را دنبال می‌کنند که توسط یک تابع توزیع تجمعی مکمل^۳ به نام $H(\sigma)$ معین می‌شود.

¹Probability Generating Function

²Little's Law

³Complementary Cumulative Distribution Function

همانند تابع توزیع تجمعی مکمل نرمال مساحت زیر نمودار این تابع نیز برابر ۱ تعیین می‌شود. در حالت M/G جمعیت $x(t)$ بدلیل غیرنمایی بودن گذارها زنجیره مارکوف نیست. برای توصیف کامل پویایی سیستم نیازمند در نظر گرفتن سرویس‌های باقی‌مانده از کار فعلی در هر زمان داده شده است. همانگونه که قبلاً مطرح شد فرض تساوی پهنای باند لیچرهای حاضر با یک نرخ سرویس برابر با $r = R_{up}/x$ نیز ادامه داده می‌شود. در ادامه سیستم صف تشریح می‌شود.

سیستم ۱. (صف نظیر به نظیر با تعداد ثابت سیدرها). سیستم صف این‌گونه تعریف می‌شود: یک فرآیند پواسون از رسیدن مشتری‌ها با نرخ λ که هر مشتری یک اندازه سرویس مستقل برابر $\sigma_n > 0$ با $H(\sigma)$ بی‌میانگین ۱ است. همچنین یک ترتیب اشتراک پردازنده برای حالت تصرف x به هر مشتری با نرخ $r = \mu \frac{x+y_0}{x}$ سرویس می‌دهد.

همچنین داریم که در حالت تعادل مجموع جمعیت سیستم نسبت به مشخصه‌های جزئی توزیع اندازه کار (به جز میانگین) غیر حساس است. این موضوع گزاره ۴ را نتیجه می‌دهد. گزاره ۴. در حالت تعادل جمعیت $x(t)$ سیستم نظیر به نظیر ۱ توزیع معادله ۲-۲ را دنبال می‌کند.

باید گفت که در این مقاله بر روی مشخصه توزیع پایدار صف معرفی شده در [۱۰] اتکا خواهد شد.

در اینجا توزیع \bar{H} از باقی‌مانده زندگی فرآیند تجدید^۱ با فاصله زمانی بین رسیدن‌ها که از H بدست می‌آید توسط رابطه زیر معرفی می‌شود:

$$\bar{H}(\sigma) = \int_{\sigma}^{\infty} H(s) ds \quad (۶-۲)$$

در [۱۰] اثبات می‌شود که توزیع تعادل یک صف اشتراک پردازنده‌ای M/G با نرخ‌های وابسته به حالت، از یک فرآیند دو گامه بدست می‌آید. اول مجموع تعداد کارهای x دنبال‌کننده π در حالت M/M و دوم آن که اگر $x = n$ باشد بار کاری باقی‌مانده کارهای حاضر در سیستم به عنوان n و از توزیع منحصر به فرد مستقل^۲ \bar{H} انتخاب می‌شوند. در ادامه به یک توصیف‌گر حالت نیاز است که هر دو جمعیت و کار باقی‌مانده را در یک جسم واحد ریاضی مدل کند.

۲-۱-۲ توصیف‌گر حالت نقطه پروسس

همانطور که قبلاً گفته شد در حالت اندازه کار عمومی جمعیت $x(t)$ یک فرآیند مارکوف نیست و برای بدست آوردن یک نمایش مارکوف از پویایی سیستم نیاز است تا از بار کاری باقی‌مانده استفاده کرد. یک انتخاب مناسب‌تر برای ذخیره حالت سیستم‌های اشتراک پردازنده یک سنج شمارشگر تصادفی بر روی \mathbb{R}^+ با توده نقطه‌ها بر روی زمان‌های باقی‌مانده سرویس است. در این حالت توصیف‌گر حالت نقطه پروسس به صورت زیر تعریف می‌شود:

$$\Phi_t = \sum_{i=1}^{x(t)} \delta_{\sigma_i(t)}$$

که $x(t)$ تعداد مشتریان و $\sigma_i(t)$ زمان سرویس باقی‌مانده از هر کدام از آن‌ها باشد. همچنین δ_{σ} سنجه دیراک^۳ متمرکز بر σ می‌باشد. برای تحلیل‌های بیشتر چنین سنج‌های تصادفی یک مشخصه مناسب تابع لاپلاس است که به فرم زیر تعریف می‌شود:

$$\mathcal{L}_{\Phi}[f] = E \left[e^{-\int_0^{\infty} f(\sigma) \Phi(d\sigma)} \right]$$

^۱Renewal

^۲Independent and Identically Distributed

^۳Dirac

که برای هر $f \geq 0$ و محدود به \mathbb{R}^+ برقرار است.

گزاره ۵. توزیع ایستای سنجه اندازه‌گیری شده فرآیند Φ_t که یک سنجه تصادفی در \mathbb{R}^+ است با تابع لاپلاس به صورت:

$$\mathcal{L}_{\Phi}[f] = G\left(\int_0^{\infty} e^{-f(\sigma)} \bar{H}(d\sigma)\right) \quad (۷-۲)$$

توصیف می‌شود. که برای هر $f \geq 0$ که $G(\cdot)$ یک تابع مولد احتمالی از π است برقرار است.

در نتیجه معادلات ۲-۴ و ۲-۷ توزیع حالت سکون مجموع تعداد کارها و بار کاری باقی‌مانده در سیستم را مشخص می‌کند.

۲-۲ توزیع حالت سکون تحت مقیاس‌گذاری بزرگ سیستم

سیستم‌های اشتراک‌گذاری معمول شامل تعداد زیادی نظیر است. بنابراین برای نزدیک بودن به واقعیت نیاز است که توزیع حالت سکون برای مقیاس‌گذاری بزرگ سیستم بررسی شود. در این قسمت سعی می‌شود حد رفتار سیستم با افزایش اندازه سیستم مشخص شود. برای محقق شدن این امر یک خانواده از سیستم‌ها را با توجه به تعریف سیستم ۱ و نرخ رسیدن $L\lambda$ در نظر بگیرید. $L > 0$ در این عبارت یک پارامتر مقیاس‌گذاری است. در ادامه اجازه داده خواهد شد که پارامتر L به بی‌نهایت نیز متمایل شود. در این صورت لیچری که پارامتر ρ را بارگذاری می‌کند نیز به صورت خطی با پارامتر $L\rho$ رشد می‌کند. با مقیاس‌گذاری تعداد لیچرهای رسیده به سیستم، تعداد سیدرهای ثابت نیز به Ly_0 افزایش داده می‌شود. فرض شود که π_L و G_L به ترتیب توزیع نامتغیر و توزیع مولد احتمالی آن برای سیستم مقیاس‌گذاری شده است. بسته به آن که سیدرها بتوانند به تنهایی با تقاضا مقابله کنند یا خیر دو حالت بوجود می‌آید که در ادامه این دو حالت بررسی می‌شود.

۱-۲-۲ حالت سیدر پایدار ($\rho < y_0$)

هر زمان که $\rho < y_0$ باشد سیدرها می‌توانند با تقاضا مقابله کنند. در حالت حدی زمانی که $L \rightarrow \infty$ مشارکت بارگذاری لیچر ناچیز شده و سیستم به یک صف $M/G/1$ از نوع اشتراک پردازنده تنها توسط سیدرها تبدیل می‌شود. این حالت را می‌توان به صورت قضیه زیر بیان کرد:

قضیه ۱. اگر $\rho < y_0$ باشد، توزیع تعادل سیستم مقیاس‌شده به توزیع تعادل یک صف اشتراک پردازنده $M/G/1$ با بار $\rho_0 = \frac{\rho}{y_0} < 1$ همگرا می‌شود.

و از قضیه ۱ می‌توان نتیجه زیر را گرفت:

نتیجه ۱. زمانی که $\rho < y_0$ باشد و $L \rightarrow \infty$ میل کند آنگاه میانگین تعداد لیچرها \bar{x}^L برابر با:

$$\bar{x}^L \xrightarrow{L \rightarrow \infty} \frac{\rho/y_0}{1 - \rho/y_0} = \frac{\rho}{y_0 - \rho}$$

است.

۲-۲-۲ حالت پایداری عمومی ($\rho > y_0$)

حال، حالت دوم در نظر گرفته می‌شود، در این حالت مشارکت لیچرها حیاتی است. اگر سیستم مانند حالت اول مدل شود، میانگین تعداد نظیرهای حاضر (معادله ۲-۵) نیز با L رشد خواهد کرد. در اینجا از قانون مقیاس‌گذاری اعداد بزرگ برای بدست آوردن یک حد غیر بدیهی استفاده می‌شود. خانواده پروسس‌ها را در نظر بگیرید:

$$\Phi^L = \frac{1}{L} \int_{i=1}^{x^L} \delta_{\sigma_i^L} \quad (۸-۲)$$

x^L تعداد کارها در سیستم اشتراک‌گذاری پردازنده با نرخ رسیدن $L\lambda$ و Ly_0 تعداد سیدرها و σ_i^L بار کاری باقی‌مانده از آن‌ها با در نظرگرفتن وجود x^L مشتری است. فاکتور $1/L$ تعداد لیچرها را نرمالایز می‌کند. می‌توان نتیجه گرفت:

قضیه ۲. اگر $\rho > y_0$ باشد، توزیع تعادل سیستم مقیاس‌گذاری شده $\bar{\Phi}^L$ (در صورت $L \rightarrow \infty$) به سنجه قطعی $\bar{H}(d\sigma)$ $(\rho - y_0)$ بر روی \mathbb{R}^+ همگرا می‌شود. ($\bar{H}(d\sigma)$ قبلاً در ۲-۶ معرفی شده است.)

با استناد به معادله ۲-۲ توزیع نامتغیر برای تعداد کارها در سیستم مقیاس‌گذاری شده می‌تواند این‌گونه بیان شود:

$$\pi_L(n) = P(Y = Ly_0 + n \mid Y \geq Ly_0)$$

که $Y \sim \text{Poisson}(L\rho)$ است. اگر $\rho > y_0$ باشد برای L های بزرگ $1 \approx P(Y \geq Ly_0)$ است و بنابراین x مانند یک متغیر تصادفی پواسون که به اندازه Ly_0 به چپ انتقال یافته رفتار می‌کند. بنابراین، میانگین تعداد مشتری‌ها در سیستم برابر $L(\rho - y_0)$ خواهد بود و بار کاری باقی‌مانده آنها دارای توزیع منحصر به فرد مستقل با توجه به $\bar{H}(d\sigma)$ است.

با افزایش L ، فرآیند دوباره مقیاس‌شده $\bar{\Phi}^L$ یک تخمین‌گر تجربی از توزیع $\bar{H}(d\sigma)$ است که با $\rho - y_0$ مقیاس شده است. به طور میانگین، فرآیند به عنوان یک صف $M/G/\infty$ با بار ρ رفتار می‌کند، که y_0 واحد به سمت چپ انتقال یافته است. مشارکت سیدرها کمتر از میانگین تعداد نظیرها در سیستم است، در نتیجه در میانگین زمان بارگیری هم مشارکت می‌کنند. در واقع می‌توان گفت:

نتیجه ۲. اگر $\rho > y_0$ باشد با $L \rightarrow \infty$ میانگین تعداد لیچرها و زمان بارگیری در سیستم مقیاس‌شده تایید می‌شود که:

$$x^L = L(\rho - y_0) + o(L); \quad \bar{T}^L = \frac{1}{L} \left(1 - \frac{y_0}{\rho} \right) + o(1)$$

توجه شود که واریانس تعداد لیچرها با انتقال عوض نمی‌شود.

۳-۲ تقریب حد مرکزی دانلود نمایه

نتایج بخش ۲-۲ یک قانون برای توصیف تعداد زیاد برای بارگیری پیشرفت در سیستم در حالتی که $\rho > y_0$ است فراهم می‌کند. می‌توان با استفاده از قضیه حد مرکزی تقریب فوق را اصلاح کرد. به این منظور قضیه ۳ مطرح می‌شود:

قضیه ۳. در سیستم عمومی پایدار نظیر به نظیر با اندازه‌های کار قطعی، فرآیند دوباره مقیاس‌شده بارگیری نمایه $v_L(\sigma)$ یک فرآیند گوسی بین $[0, 1]$ است، یعنی:

$$v_L(\sigma) \xrightarrow{w} \sqrt{\rho - y_0} B_0(\sigma) + \sqrt{\rho} Z(1 - \sigma)$$

که B_0 پل استاندارد برونین^۱ و Z یک متغیر استاندارد نرمال مستقل از B_0 است. همچنین \xrightarrow{w} به معنای همگرایی در توزیع در حالت حد قطعی است.

۴-۲ سیدرهای متغیر و شبه ایستا

بررسی صف بخش ۲-۱ موجب شد تا عبارت دقیقی برای توزیع تعادل تعداد لیچرها در زمانی که تعداد سیدرها ثابت است بدست بیاید. همچنین برای حالت تعداد سیدرهای متغیر تحلیلی وجود ندارد. اگرچه یک شبیه‌سازی از یک شرایط تولید سیدرهای درون‌زا می‌تواند با عبارت‌های صریح تقریب زده شود: زمانی که یک کسر کوچکی از لیچرها انتخاب می‌کنند که به سیدر تبدیل شوند و برای مدت طولانی‌ای در سیستم به عنوان سیدر بمانند. این شرایط در سناریوهای تبادل عملی معمول است.

¹ Standard Brownian Bridge

برای مدل کردن این شرایط شبکه صف شکل ۲-۱ را در نظر بگیرید که تغییر یافته مدل اصلی مطرح شده در [۹] است. در اینجا، $\alpha \in (0, 1]$ احتمال تبدیل شدن یک لیچر به سیدر بعد از اتمام بارگیری را نشان می‌دهد. حالت خاص $\alpha = 1$ مدل مطرح شده در [۹] است. برای ساده کردن بررسی حالت M/M را فرض کنید. همچنین اندازه‌های کار که به صورت نمایی با میانگین ۱ توزیع شده را در نظر بگیرید. سیدرها نیز برای مدت زمانی نمایی با پارامتر γ در سیستم می‌مانند.

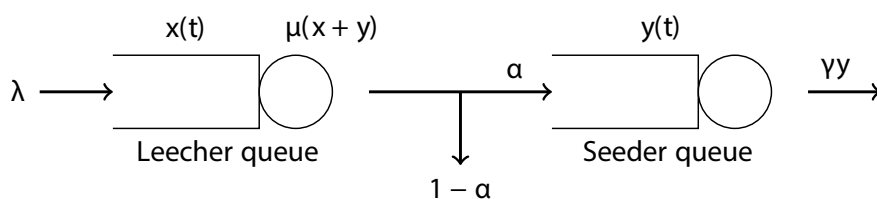
نرخ‌های گذار زنجیره مارکوف شکل ۲-۱ به صورت زیر محاسبه می‌شود:

$$q_{(x,y),(x+1,y)} = \lambda, \quad (2-9\text{آ})$$

$$q_{(x,y),(x-1,y+1)} = \alpha \mu(x+y) 1_{\{x>0\}}, \quad (2-9\text{ب})$$

$$q_{(x,y),(x-1,y)} = (1-\alpha) \mu(x+y) 1_{\{x>0\}}, \quad (2-9\text{ج})$$

$$q_{(x,y),(x,y-1)} = \gamma y \quad (2-9\text{د})$$



شکل ۲-۱: شبکه صف برای تحلیل شبه ایستا

با توجه به مفروضات مطرح شده α و γ کوچک خواهند بود. بنابراین دومین صف در مقیاس زمانی بزرگتری نسبت به صف اول فعالیت می‌کند سیستم مطرح شده یک فرآیند شبه تولد مرگ است. اگر α و γ به صفر میل کنند، صف اول مانند زنجیره مارکوف در بخش ۲-۱ با سیدرهای ثابت برای هر سطح y تکامل می‌یابد.

ثابت می‌شود که این تجزیه نیمه ایستا برقرار خواهد بود و حد دومین صف به عنوان صف $M/M/\infty$ با بار λ/v رفتار می‌کند. در نهایت قضیه ۴ نیز که مهمترین قضیه این بخش است به صورت زیر تعریف می‌شود:

قضیه ۴. خانواده توزیع‌های $\{\pi_\alpha\}_{\alpha \in (0,1]}$ یک حد $\pi_0(x,y)$ با $\alpha \rightarrow 0$ دارند که با توجه به تجزیه نیمه ایستا خواهیم داشت:

$$\pi_0(x,y) = \pi_0(y)\pi_0(x|y)$$

که $\pi_0(x|y)$ توزیع نامتغیر برای سیستمی با سیدرهای ثابت است:

$$\pi_0(x|y) = \frac{1}{\sum_{m=y}^{\infty} \rho_x^m / m!} \frac{\rho_x^{x+y}}{(x+y)!} \quad \text{with } \rho_x := \frac{\lambda}{\mu} \quad (2-10)$$

و

$$\pi_0(y) = e^{-\rho_y} \frac{\rho_y^y}{y!} \quad \text{with } \rho_y := \frac{\lambda}{\mu} \quad (2-11)$$

فصل سوم

شبیه‌سازی و نتیجه‌گیری

در این بخش چند سناریو شبیه‌سازی به منظور اعتبارسنجی مدل ارائه شده، برای یک سیستم واقعی تبادل فایل نظیر به نظیر، انجام می‌شود. سپس در بخش ۳-۲ نتایج جمع‌بندی می‌شود.

۳-۱ آزمایشات شبیه‌سازی

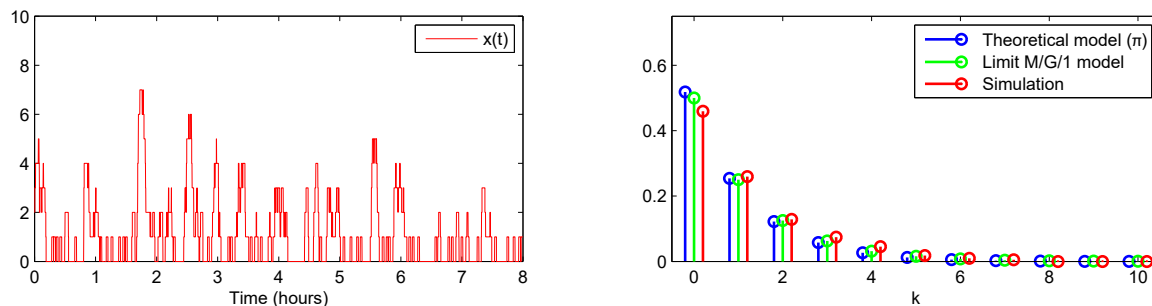
تمام شبیه‌سازی‌ها بوسیله شبیه‌ساز شبکه ns2 و با استفاده از کتابخانه بیت‌تورنت انجام شده است. این کتابخانه به استفاده‌کنندگان این اجازه را می‌دهد تا یک شبکه مشتری‌های نظیر به نظیر را بر روی یک توپولوژی شبکه ایجاد کنند. فایل در نظر گرفته شده یک فایل ۱۰۰ مگابایتی است که به ۴۰۰ چانک تقسیم شده است. تمام نظیرها ظرفیت بارگذاری برابر ۲۰۰ کیلوبیت بر ثانیه را دارند. در تجربه‌های شبیه‌سازی انجام شده بهره‌وری پهنای باند بارگذاری نظیرهایی به منظور اندازه‌گیری مقدار μ سنجیده شده است. در این آزمایش‌ها مشاهده می‌شود که لیچرها به سرعت و موثر به نرخ بارگذاری ۹۰٪ ظرفیت لینک فیزیکی می‌رسند. همچنین این نتیجه را می‌توان دریافت که نرخ موثر بارگذاری سیدرها نیز همانند لیچرها است که استفاده مقدار یکسان μ را توجیه می‌کند. مقدار صحیح μ تحت معادله زیر تعیین می‌شود:

$$\mu = \frac{\text{پهنای باند موثر بارگذاری}}{\text{اندازه فایل}} \approx 2.82 \times 10^{-4} \text{ دقیقه/فایل}$$

با مقدار فوق یک رونوشت کامل از فایل حدود ساعت $1 \approx 1/\mu$ زمان برای بارگذاری نیاز خواهد داشت. در ادامه به تحلیل ۳ حالت معرفی شده در قبل پرداخته می‌شود.

۳-۱-۱ حالت سیدر پایدار

در اولین سناریوی شبیه‌سازی سیدرهای ثابت می‌توانند با تقاضا مقابله کنند. مقدار $y_0 = 50$ تنظیم شده و لیچرهای جدید بدون هیچ محتوایی با فرآیند پوآسون به نرخ λ وارد سیستم می‌شوند. مقادیر به گونه‌ای انتخاب می‌شود که $\rho = \lambda/\mu = 25 < y_0$ برقرار باشد.



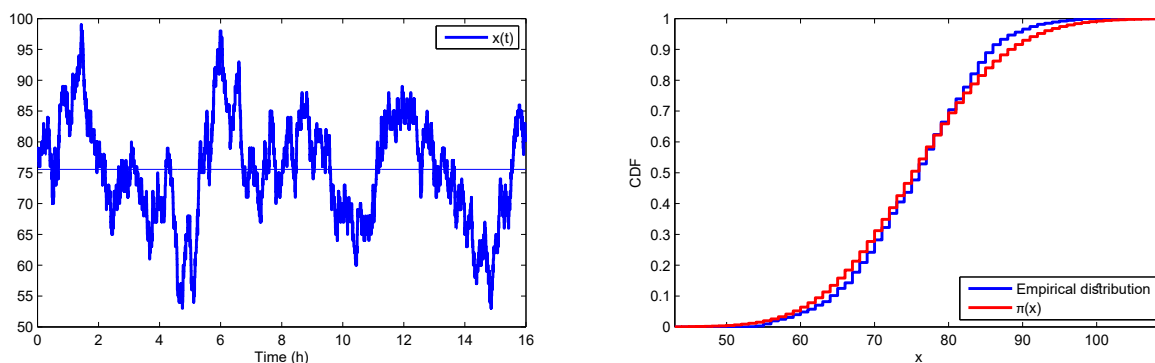
شکل ۳-۱: توزیع حالت بلند مدت پایداری (راست) - زمان تکامل تعداد لیچرها (چپ) - نتایج شبیه‌سازی برای حالت $\rho < y_0$

یک تکامل از سیستم را در شکل ۳-۱ مشاهده می‌کنید. در این تکامل صف برای مدت زمانی خالی است که نشان دهنده این حقیقت است که میانگین درخواست‌ها از ظرفیت سیدرها کمتر است. در این شکل همچنین توزیع نظری پیش‌بینی شده توسط مدل صف نیز نمایش داده شده است که توسط π در معادله ۲-۲ محاسبه شده است. توزیع حد پیش‌بینی شده عنوان شده در قضیه ۱ نیز برابر با $\text{Geometric}(\rho/y_0)$ همراه با $\rho/y_0 = 1/2$ نیز در این شکل رسم شده است. به طور خاص، میانگین تعداد نظیرها در سیستم برابر $\bar{x} \approx 1.1$ است که مشابه با مقدار پیش‌بینی شده $\frac{\rho}{y_0 - \rho} = 1$ است.

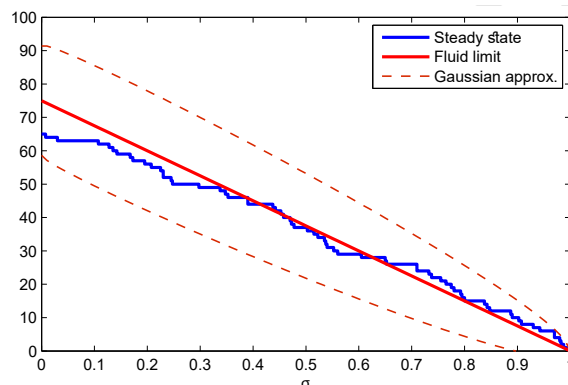
۳-۱-۲ حالت پایداری عمومی

در دومین سناریوی شبیه‌سازی مشارکت لیچرها ضروری است. مقدار $y_0 = 25$ تنظیم و ثابت نگه‌داشته می‌شود. لیچرها همانند حالت قبل بدون محتوا و با فرآیند پواسونی با نرخ λ وارد سیستم می‌شوند. مقادیر به گونه‌ای انتخاب می‌شود که $\rho = \lambda/\mu = 100 > y_0$ برقرار باشد.

یک تکامل از سیستم را در حالت سکون در شکل ۳-۲ مشاهده می‌کنید. جمعیت لیچرها، حول یک عدد غیر صفر نوسان می‌کند که با توجه به نتیجه ۲ باید برابر $\rho - y_0 = 75$ باشد. همچنین در این شکل توزیع تجربی جمعیت لیچرها بوسیله شبیه‌سازی نشان داده شده است و با توزیع نظری ۲-۲ و توزیع حد بدست آمده در بخش ۲-۲-۲ مقایسه شده است. در ادامه پیش‌بینی مدل‌های ارائه شده برای نمایه کار باقیمانده نشان داده شده است. در شکل ۳-۳ تابع توزیع تجمعی برا کار باقی‌مانده برای یک نمونه خاص حالت بلند مدت نشان داده شده است. حد سیال از قضیه ۲ و انحرافات پیش‌بینی شده در قضیه ۳ نیز به شکل مقایسه‌ای در این شکل نمایش داده شده اند.



شکل ۳-۲: توزیع حالت بلند مدت پایداری (راست) - زمان تکامل تعداد لیچرها (چپ) - نتایج شبیه‌سازی برای حالت $\rho > y_0$

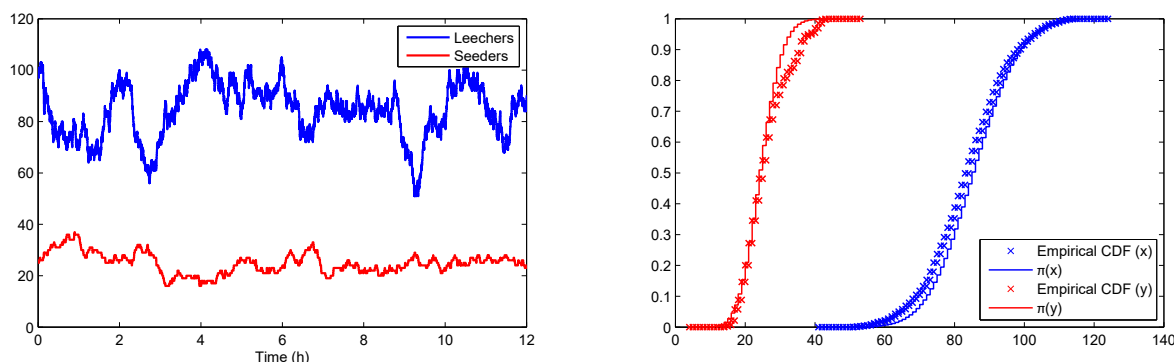


شکل ۳-۳: رفتار بلند مدت پیشرفت بارگذاری و حد سیال آن به همراه تقریب گوسی

۳-۱-۳ سیدرهای با تغییر آرام

در سومین سناریوی شبیه‌سازی بخشی از لیچرها به عنوان سیدر در سیستم باقی می‌مانند. برای تغییر آرام جمعیت سیدرها ظرفیت بارگذاری دو برابر، برابر با ۵۱۲ کیلو بیت بر ثانیه می‌شود و در نتیجه مدت زمان بارگذاری رونوشت فایل نصف و برابر ۳۰ دقیقه خواهد شد. همچنین نرخ ورود لیچرها نیز دو برابر می‌شود تا جمعیت تعادل در $\rho_x \approx 100$ باقی بماند. ده درصد لیچرها ($\alpha = 0.1$) به عنوان سیدر بعد از اتمام بارگیری در سیستم باقی می‌مانند. γ نیز به صورتی انتخاب می‌شود که میانگین تعداد سیدرهای حاضر در سیستم حدود $\rho_y = 25$ باشد.

یک تکامل از سیستم را در حالت سکون در شکل ۳-۴ مشاهده می‌کنید. در این شکل توزیع تجربی شبیه‌سازی با توزیع نظری (محاسبه شده توسط حد شبه ایستا) مقایسه شده است.



شکل ۳-۴: توزیع حالت بلند مدت پایداری (راست) - زمان تکامل تعداد لیچرها و سیدرها (چپ) - نتایج شبیه‌سازی برای حالت $\rho > y_0$

۳-۲ نتیجه‌گیری

در مقاله در حال مرور، یک مدل صف برای شبکه تبادل فایل نظیر به نظیر ارائه و تحلیل شده است. در این مقاله ابتدا به طور خاص بر روی سیستم‌های با تعداد ثابت سیدر تمرکز شده است. این تمرکز امکان تحلیل مدل‌های دیگر را نیز به نویسندگان می‌دهد. همچنین در این مقاله یک کلاس از صف‌های اشتراک پردازنده M/G قابل انعطاف، شناسایی و توزیع ایستای آن نیز مشخص شده است.

مشخص شد که توزیع ایستای آن به اندازه‌های کار غیرحساس است و در نتیجه برای کارهای غیرقطعی سیستم‌های نظیر به نظیر قابل استفاده است. در ادامه سیستم تحت یک شبکه بزرگ تحلیل شده است. در این رابطه نشان داده شده است که سیستم را می‌توان توسط یک صف $M/G/1$ یا یک صف انتقال داده شده $M/G/\infty$ بسته به غالب بودن سرور یا نظیر تقریب زد. در نهایت نتایج برای سیدرهای ثابت به حالتی که در آن جمعیت سیدرها به آرامی تغییر می‌کند تعمیم داده شده است. در آخر شبیه‌سازی‌ها همراه با نتایج نیز به طور دقیق اعتبارسنجی شدند.

References

- [1] CLÉVENOT-PERONNIN, F., NAIN, P., AND ROSS, K. W. Multiclass {P2P} networks: Static resource allocation for service differentiation and bandwidth diversity. *Performance Evaluation* 62, 1–4 (2005), 32 – 49. Performance 200524th International Symposium on Computer Performance, Modeling, Measurements and Evaluation.
- [2] DE VECIANA, G., AND YANG, X. Fairness, incentives and performance in peer-to-peer networks. *Seeds* 250, 300 (2003), 350.
- [3] KESIDIS, G., KONSTANTOPOULOS, T., AND SOUSI, P. Modeling file-sharing with bittorrent-like incentives. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (2007), vol. 4, IEEE, pp. IV–1333.
- [4] LEIBNITZ, K., HOFELD, T., WAKAMIYA, N., AND MURATA, M. *Modeling of Epidemic Diffusion in Peer-to-Peer File-Sharing Networks*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 322–329.
- [5] LESKELÄ, L., ROBERT, P., AND SIMATOS, F. Interacting branching processes and linear file-sharing networks. *Advances in Applied Probability* 42, 03 (2010), 834–854.
- [6] MASSOULIÉ, L., AND VOJNOVIĆ, M. Coupon replication systems. In *ACM SIGMETRICS Performance Evaluation Review* (2005), vol. 33, ACM, pp. 2–13.
- [7] PAGANINI, F., AND FERRAGUT, A. Pde models for population and residual work applied to peer-to-peer networks. In *2012 46th Annual Conference on Information Sciences and Systems (CISS)* (March 2012), pp. 1–6.
- [8] QIU, D., AND SRIKANT, R. Modeling and performance analysis of bittorrent-like peer-to-peer networks. *SIGCOMM Comput. Commun. Rev.* 34, 4 (Aug. 2004), 367–378.
- [9] YANG, X., AND DE VECIANA, G. Service capacity of peer to peer networks. In *IEEE INFOCOM 2004* (March 2004), vol. 4, pp. 2242–2252 vol.4.
- [10] ZACHARY, S. A note on insensitivity in stochastic networks. *Journal of applied probability* 44, 01 (2007), 238–248.