

## Prediction-based Mobility-aware Computation Offloading in Multi-access Edge Computing

Shokufeh Norouzi, Zeinab Movahedi\*

1- School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.

2\*- School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.

<sup>1</sup> shokufeh\_norouzi@comp.iust.ac.ir, <sup>2\*</sup> zmovahedi@iust.ac.ir

Corresponding author's address: Zeinab Movahedi, Faculty of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.

**Abstract-** Today, as the new generation of communication networks is implemented, we are witnessing a considerable change in IoT development and new programs in this context. Despite recent advancements in mobile networks and devices, the limitations of devices connected to this platform in terms of computational power and energy have resulted in severe challenges for running resource-intensive programs with exigent latency requirements. To address these challenges, the concept of computation offloading in Multi-access Edge Computing (MEC) has been recently developed, in which storage and computation resources are provided close to the user. However, due to the user mobility and changes in the profile of offloaded applications over time, the problem of assignment of edge servers to users with the aim of minimizing the overall offloading latency is a complicated task. In this regard, existing mobility-aware offloading approaches are not based on fine-grain offloading and use random and unrealistic mobility models. In this article, to address the aforementioned challenges, we propose a mobility-aware fine-grain computation offloading method to minimize the overall offloading delay. In the proposed approach, the user application is divided into several components and the offloading decision is made for each component according to the mobility and specifications of user components during the time slots defined in the system. In one hand, this latter results in more efficient offloading decision. In the other hand, it reduces the overhead of migration since the migration of a subset of program's components imposes lower cost compared to the migration of the entire program. Moreover, we use user profile and location prediction to optimize the offloading decisions considering the underlying context over time. According to the evaluation results, it is observed that the proposed method achieves significantly better performance compared to other alternatives while the complexity of offloading decision is kept very low.

**Keywords-** Computation Offloading, Mobility-aware, Multi-access Edge Computing, Internet of Things.

## تخلیه‌بار محاسباتی تحرک آگاه مبتنی بر پیش‌بینی در رایانش لبه با دسترسی چندگانه

شکوفه نوروزی<sup>۱</sup>، زینب موحدی<sup>۲\*</sup>

۱- دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

۲- دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

<sup>۱</sup>shokufeh\_norouzi@comp.iust.ac.ir, <sup>۲\*</sup>zmovahedi@iust.ac.ir

\* نشانی نویسنده مسئول: زینب موحدی، تهران، میدان رسالت، خیابان هنگام، دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر.

چکیده- امروزه با پیاده‌سازی نسل جدید شبکه‌های ارتباطی، شاهد تحولی عظیم در توسعه اینترنت اشیاء و ظهور برنامه‌های جدید در این بستر می‌باشیم. محدودیت در توان محاسباتی و انرژی دستگاه‌های متصل به این بستر موجب ایجاد چالش و عدم پشتیبانی این دستگاه‌ها برای اجرای برنامه‌ها با بار محاسباتی بالا و نیازمند تاخیر کم می‌شود. روش‌های تخلیه‌بار محاسباتی در رایانش لبه با دسترسی چندگانه، با فراهم آوردن منابع محاسباتی و ذخیره‌سازی در نزدیکی کاربر راهکاری کارآمد برای مقابله با چالش‌های ذکر شده است. با این وجود، به علت تحرک کاربر و تغییر در مشخصات برنامه‌های تخلیه‌شده در طول زمان، مسئله تخصیص خدمت‌گزاران لبه به کاربران با هدف کاهش تاخیر با چالش‌هایی مواجه است. رویکردهای فعلی تخلیه‌بار تحرک آگاه در این حوزه از مدل‌های تحرک تصادفی و غیرواقعی-گرایانه‌ای استفاده می‌کنند و همچنین اجرای تخلیه‌بار در آن‌ها به صورت درشت‌دانه صورت می‌گیرد. در این مقاله تخلیه‌بار به منظور بهره‌مندی از مزایای آن ریزدانه می‌باشد. بر این اساس برنامه کاربران به تعدادی مولفه تقسیم و اخذ تصمیم تخلیه با توجه به تحرک و مشخصات مولفه‌های کاربران در طول شکاف‌های زمانی تعریف شده در سیستم، انجام می‌گیرد. این تصمیم علاوه بر بهینه بودن در مورد هر مولفه به کاهش سربار ناشی از مهاجرت یک مولفه به نسبت کل برنامه نیز منجر می‌شود. همچنین، به منظور اخذ تصمیم بهینه در راستای نیل به هدف مسئله یعنی کمینه کردن برآیند زمان تخلیه‌بار، از پیش‌بینی مشخصات کاربران و موقعیت مکانی آن‌ها استفاده می‌کنیم. با توجه به نتایج به دست آمده از ارزیابی، مشاهده می‌شود که روش پیشنهادی به نسبت روش‌های مورد مقایسه دارای بهبود در تابع هدف مسئله و پیچیدگی محاسباتی اخذ تصمیم است.

واژه‌های کلیدی: تخلیه‌بار محاسباتی، تحرک آگاه، رایانش لبه با دسترسی چندگانه، اینترنت اشیاء

### ۱- مقدمه

محدودیت در توان محاسباتی و انرژی دستگاه‌های متصل بوده که موجب می‌شود آن‌ها توان اجرای برنامه‌هایی با ویژگی‌های فوق را نداشته باشند. تخلیه‌بار محاسباتی<sup>۶</sup> در رایانش لبه با دسترسی چندگانه<sup>۷</sup> (MEC) با در اختیار قرار دادن منابع محاسباتی قدرتمند، امکان جمع‌آوری داده، محاسبات و تحلیل هوشمند را در نزدیکی منابع تولید داده و محاسبات، فراهم می‌آورد. همچنین، در MEC برای به اتصال به شبکه می‌توان از فناوری‌های دسترسی متنوعی از جمله Wi-Fi و شبکه‌های سلولی<sup>۸</sup> استفاده کرد. بدین ترتیب علاوه

یکی از مأموریت‌های نسل پنجم شبکه‌های ارتباطی (5G) ارائه اتصالات گسترده به منظور پشتیبانی تطابق‌پذیر با توسعه روزافزون اینترنت اشیاء<sup>۲</sup> (IoT) می‌باشد. برنامه‌های جدید در بستر اینترنت اشیاء همچون واقعیت افزوده<sup>۳</sup>، بازی‌های تعاملی<sup>۴</sup> و پردازش زبان طبیعی<sup>۵</sup> برای اجرا در دستگاه‌های متصل به IoT به توان محاسباتی بالا و تاخیر کم نیاز دارند. از جمله چالش‌های اساسی در این بستر

محاسباتی ریزدانه و تحرک آگاه مبتنی بر پیش‌بینی ارائه می‌شود. بدین منظور، مسئله تخلیه‌بار ریزدانه تحرک آگاه با هدف کمینه کردن برآیند زمان تخلیه‌بار ریزدانه متشکل از زمان‌های تخلیه، پردازش و مهاجرت برای تمامی کاربران و با لحاظ قیود محدودیت منابع رایانشی و ارتباطی در خدمت‌گزاران مدل‌سازی شده و پیش‌بینی تحرک و مشخصات برنامه کاربران در طول زمان نیز در آن مورد نظر قرار گرفته است. روش پیشنهادی جهت پیش‌بینی اطلاعات مورد نظر از دو روش انحراف<sup>۱۴</sup> و میانگین وزن‌دار استفاده کرده و بر این اساس هزینه مورد انتظار تخلیه مولفه‌ها به ازای تمامی خدمت‌گزاران را برآورد می‌کند. در نهایت، با استفاده از هزینه برآورد شده، تخصیص هوشمندانه مولفه‌ها به خدمت‌گزاران لبه با روشی مبتنی بر برنامه‌ریزی پویا<sup>۱۵</sup> صورت می‌پذیرد. بر این اساس، این مقاله با پیش‌بینی تحرک کاربران و همچنین پیش‌بینی مشخصات برنامه‌های کاربردی آن‌ها، با نمود تغییرات پویا بر اساس زمان و مکان در حل مسئله، به افزایش دقت در تصمیم‌گیری و کاهش سربار دست یافته است. علاوه بر این با توجه به بهره‌مندی از تخلیه‌بار ریزدانه، از تخلیه‌بار کارآمدتر و همچنین کاهش هزینه و سربار مهاجرت یک مولفه به نسبت کل برنامه دست یافته است. ادامه این مقاله به صورت زیر ساختار بندی شده است. در بخش دوم به بررسی کارهای انجام شده در زمینه تخلیه‌بار محاسباتی تحرک-آگاه در MEC پرداخته می‌شود. سپس، در بخش سوم مدل سیستم و مدل‌سازی پیشنهادی تخلیه‌بار محاسباتی تحرک آگاه ریزدانه را ارائه خواهیم کرد. در بخش چهارم، الگوریتم حل مسئله مبتنی بر پیش‌بینی موقعیت مکانی کاربران و مشخصات مولفه‌های برنامه کاربردی آن‌ها بیان شده است در بخش پنجم، معیارهای ارزیابی و نتایج به دست آمده از ارزیابی روش پیشنهادی به نمایش در می‌آید. در نهایت، در بخش آخر نتیجه‌گیری و پیشنهادات برای کارهای آتی مطرح می‌شود.

## ۲- مروری بر کارهای گذشته

در این قسمت به بررسی کارهای اخیر و مرتبط در حوزه روش‌های تخلیه‌بار محاسباتی تحرک آگاه در MEC می‌پردازیم. در [۵]، یک روش تخلیه بار محاسباتی درشت‌دانه در محیطی شامل چندین کاربر متحرک و چندین گره مه با محدوده پوشش معین ارائه شده است. در صورتی که کاربری از محدوده پوشش گره مه متصل به خود خارج و محاسبات بار تخلیه‌شده خود در آن ناتمام باقی مانده باشد، مهاجرت بار میان دو گره مه رخ می‌دهد. به این منظور، زمان مورد نیاز برای پردازش بار محاسبه شده و با زمان اتصال کاربر به گره مه که از تابع توزیع احتمال به دست می‌آید،

بر مقابله با چالش‌های اینترنت اشیا، مزایایی همچون افزایش قابلیت محاسبه کاربران، ذخیره پهنای باند هسته شبکه<sup>۱۶</sup>، کنترل ازدحام<sup>۱۷</sup>، کاهش تاخیر و پشتیبانی از خدمات بلادرنگ<sup>۱۸</sup> فراهم می‌آید [۱-۲].

با وجود مزایای یادشده، تداوم ارائه خدمات و تخصیص بهینه خدمت‌گزاران لبه<sup>۱۲</sup> به کاربران، با وجود تحرک و تغییر مکان آن‌ها، یکی از چالش‌برانگیزترین موضوعات در این حیطه می‌باشد [۳-۴]. به‌علاوه، در اثر تحرک کاربران و نیز در طول زمان ممکن است مشخصات برنامه‌های کاربران تغییر کند که این موضوع نیز به پیچیدگی اخذ تصمیم و تخصیص بهینه خدمت‌گزاران به کاربران می‌افزاید. در واقع برنامه‌ی کاربر بسته به شرایط مکانی و زمانی اجرای آن، دارای بارکاری متفاوتی خواهد بود و به دنبال آن نیازمند سرعت پردازش متناسب با تغییر بار خود می‌باشد. به‌علاوه، در اثر تغییرات پویا همچون تحرک کاربر و تغییر بار در لبه، ممکن است خدمت‌گزار تخصیص یافته جهت تخلیه‌بار کاربر، در گذر زمان بهینه نباشد به طوری که مهاجرت بار محاسباتی میان خدمت‌گزاران لبه اجتناب ناپذیر باشد. با توجه به اینکه در صورت مهاجرت متعدد یک برنامه، شاهد افزایش سربار شبکه، افزایش تاخیر و کاهش کارایی می‌باشیم، اتخاذ تصمیم تخلیه‌بار بر اساس پیش‌بینی شرایط آتی و آگاهی از تحرک می‌تواند جهت بهبود کارایی و اجتناب از مهاجرت‌های متوالی موثر باشد.

هرچند تاکنون تحقیقاتی در زمینه تخلیه‌بار محاسباتی تحرک آگاه<sup>۱۳</sup> در MEC انجام شده است اما این کارها غالباً از مدل‌های تحرک تصادفی و غیرواقع‌بینانه‌ای استفاده می‌کنند که نمی‌توانند الگوهای تحرک و ویژگی‌های آن در طول زمان را منعکس نمایند. به‌علاوه، کارهای موجود از پیش‌بینی تحرک بهره نمی‌برند، این در حالیست که به علت بالا بودن نرخ پویایی سیستم می‌توان از پیش‌بینی تحرک جهت ارائه الگوهای تحرک پویا، افزایش دقت در تصمیم‌گیری و کاهش هزینه‌های ناشی از تخلیه و مهاجرت بار محاسباتی کاربران استفاده کرد. علاوه بر این، غالب کارها از تخلیه‌بار درشت‌دانه که در آن کل برنامه به سمت لبه تخلیه می‌شود استفاده می‌کنند، درحالی‌که با تخلیه‌بار ریزدانه بر اساس مولفه‌های تشکیل‌دهنده برنامه، امکان اخذ تصمیم تخلیه به ازای هر مولفه بر اساس شرایط شبکه و در نتیجه بهبود عملکرد تخلیه‌بار وجود دارد. همچنین، در صورت وقوع مهاجرت، در تخلیه‌بار ریزدانه تنها مولفه‌های تخلیه‌شده به خدمت‌گذار لبه مورد نظر تحت تاثیر قرار خواهند گرفت و لذا هزینه و سربار مهاجرت به نسبت مهاجرت کل برنامه کمتر خواهد بود.

به منظور حل این چالش‌ها، در این مقاله یک روش تخلیه‌بار

ارائه می‌شود. هدف این کار بهبود نرخ موفقیت تخلیه بار و کاهش مصرف انرژی بوده و برای رسیدن به این هدف قابلیت اطمینان هر نقطه دسترسی برآورد می‌شود. مسئله بر اساس الگوریتم ژنتیک مورد حل قرار گرفته است. از چالش‌های این کار می‌توان به عدم پیش‌بینی موقعیت مکانی کاربر و مشخصات برنامه کاربردی او در گذر زمان اشاره کرد. همچنین، پیش‌بینی نقاط دسترسی ارائه شده در آن از یافتن یک الگوی اتصال ثابت مبنی بر اتصالات پیشین پیروی می‌کند که انعکاسی از محیط جدید و شرایط پویا منطبق با آن را در خود نخواهد داشت.

در [۱۰]، تخلیه بار درشت‌دانه تحرک آگاه با هدف کمینه کردن مصرف انرژی کاربران مورد بررسی قرار گرفته است. در این کار، به منظور تعیین خدمت‌گزار مقصد برای مهاجرت بار در اثر تحرک کاربران از روش کوتاه‌ترین مسیر در گراف خدمت‌گزاران استفاده شده است. از چالش‌های این کار می‌توان به عدم لحاظ ریزدانه‌گی، عدم پیش‌بینی تحرک و مشخصات برنامه کاربران اشاره کرد.

در [۱۱]، مساله تخلیه بار درشت‌دانه تحرک آگاه با هدف کمینه کردن تاخیر مورد توجه قرار گرفته است. در این روش با استفاده از پیش‌بینی مشخصات برنامه و موقعیت مکانی کاربر، تاخیر مورد انتظار با لحاظ هزینه مهاجرت در صورت وقوع برآورد می‌شود. همچنین، تصمیم مناسب تخلیه بار بر اساس تاخیر مورد انتظار با استفاده از برنامه‌ریزی پویا اخذ می‌گردد. همانند سایر کارهای بررسی شده، در این کار از مزایای تخلیه بار ریزدانه بهره برده نشده است.

در [۱۲]، یک روش تخلیه بار درشت‌دانه تحرک آگاه با هدف کمینه کردن برآیند تاخیر کاربران در طول زمان با استفاده از یک روش حل مبتنی بر الگوریتم حریمانه<sup>۲۱</sup> ارائه شده است. همچنین، به منظور لحاظ پویایی محیط، مشخصات برنامه و موقعیت مکانی کاربر با استفاده از یک روش مبتنی بر یادگیری پیش‌بینی شده و در محاسبه تاخیر مورد استفاده قرار گرفته است. همانند سایر کارهای تخلیه بار تحرک آگاه، این کار نیز از مزایای تخلیه بار ریزدانه استفاده نکرده است.

بررسی کارهای مرتبط مشخص می‌کند که با وجود تحقیقات در زمینه تخلیه بار محاسباتی تحرک آگاه در MEC، این کارها غالباً از مدل‌های تحرک تصادفی و غیرواقعی بینانه‌ای استفاده می‌کنند که نمی‌توانند الگوهای تحرک واقع‌گرایانه و ویژگی‌های آن در طول زمان را انعکاس دهند. همچنین در غالب کارها از پیش‌بینی مشخصات برنامه کاربر در طول زمان و بر اساس تغییر مکان اجرای برنامه، به منظور لحاظ پویایی شرایط و افزایش دقت در تصمیم‌گیری استفاده نشده است. همچنین، کارها از تخلیه بار درشت‌دانه که در آن کل برنامه به سمت لبه تخلیه می‌شود استفاده می‌کنند. این

مقایسه می‌شود. در صورت بیشتر بودن زمان پردازش از زمان اتصال به گره مه، هزینه‌ای بابت مهاجرت بار از تابع هدف مسئله که به منظور پیشینه کردن انرژی و زمان تخلیه به نسبت اجرای محلی است، کاسته می‌شود. مسئله فرمول شده مبتنی بر الگوریتم ژنتیک مورد حل قرار می‌گیرد. این کار از پیش‌بینی تحرک کاربر استفاده نکرده و همچنین تخلیه بار به صورت درشت‌دانه پیاده شده است.

در [۶]، تخلیه بار محاسباتی درشت‌دانه با تاکید بر مسئله کارآمدی انرژی مورد بررسی قرار گرفته است. به این منظور، قابلیت برداشت انرژی<sup>۱۶</sup> برای دستگاه‌های اینترنت اشیا در نظر گرفته شده است. هدف این کار کمینه کردن هزینه متشکل از انرژی مصرفی، مهاجرت بار و عدم پذیرش بار بوده و مبتنی بر بهینه‌سازی لیاپانوف<sup>۱۷</sup> مورد حل قرار می‌گیرد. در این کار، اثر تحرک تنها به صورت هزینه مهاجرت در هدف مسئله در زمان‌های متوالی لحاظ شده است و هیچ‌گونه پیش‌بینی موقعیت مکانی یا در نظر گرفتن شرایط پویا ناشی از تحرک در آن مورد توجه قرار نگرفته است.

در [۷]، تخلیه بار محاسباتی درشت‌دانه در محیطی شامل چندین کاربر متحرک حول یک ایستگاه پایه<sup>۱۸</sup> مورد بررسی قرار گرفته است. در این کار، مسئله پیشینه کردن سودمندی متشکل از کاهش مصرف انرژی و تاخیر به صورت برنامه‌ریزی صحیح غیرخطی<sup>۱۹</sup> مدل‌سازی شده است. به منظور مدیریت تحرک، مسیر هر کاربر در یک مدت مشخص بر اساس تجربه تخمین زده می‌شود. بر اساس این تخمین، با تخصیص منابع پردازشی کافی به کاربر پیش از خروج او از محدوده پوشش ایستگاه پایه، از وقوع مهاجرت بار محاسباتی کاربر جلوگیری می‌شود. چنانچه با این وجود محدودیت تاخیر مورد تحمل کاربر ارضا نشود، برنامه کاربر به صورت محلی اجرا خواهد شد. در این کار، علاوه بر عدم استفاده از تخلیه بار ریزدانه، از پیش‌بینی تحرک و مشخصات برنامه کاربر استفاده نشده است.

در [۸]، یک روش تخلیه بار درشت‌دانه مبتنی بر ارتباطات دستگاه به دستگاه<sup>۲۰</sup> در محیطی متشکل از تعدادی دستگاه متحرک با منابع محدود محاسباتی و باتری ارائه شده است. به این منظور، مسئله تخلیه بار درشت‌دانه با هدف کمینه کردن تاخیر به صورت برنامه ریزی غیرخطی مدل‌سازی شده و با الگوریتم ژنتیک مورد حل قرار گرفته است. در این کار، با استفاده از تجربیات قبلی دستگاه‌های متحرک، فاصله هر دستگاه تا محل تخلیه محاسبه شده و در محاسبه بهره آنتن و در نهایت تعیین نرخ و تاخیر انتقال میان دو دستگاه مورد استفاده قرار می‌گیرد. این کار از تخلیه بار ریزدانه، پیش‌بینی تحرک و مشخصات برنامه کاربردی استفاده نکرده است.

در [۹]، یک روش تخلیه بار محاسباتی تحرک آگاه با قابلیت پیش‌بینی نقاط دسترسی مبتنی بر الگوی اتصال پیشین کاربران

هر کاربر  $u_i \in U$  دارای یک برنامه کاربردی بوده که شامل مولفه‌ها و روابط میان آن‌ها را می‌باشد. این برنامه را می‌توان توسط یک گراف  $G = (V, E, W_V, W_E)$  مدل کرد. در این گراف هر رأس  $v \in V$  نشان‌دهنده یک مولفه برنامه و هر یال  $e \in E$  بیانگر داده فراخوانی شده میان دو مولفه مجاور برنامه است که به ترتیب با  $W_v \in W_V$  و  $W_e \in W_E$  وزن دهی می‌شوند. وزن هر رأس در گراف برنامه کاربردی بر اساس بار کاری مولفه متناظر آن و وزن هر یال بر اساس مقدار داده فراخوانی شده میان دو مولفه مجاور مشخص می‌شوند. با این توضیح مجموعه  $V_i = \{v_{i1}, \dots, v_{iv}, \dots, v_{i|V|}\}$  مشخص کننده مولفه‌های کاربر  $u_i$  می‌باشد. هر مولفه از هر کاربر با توجه به شکاف زمانی واقع شده در آن دارای مشخصات متفاوتی است که با چندتایی  $v_{iv}^t = \{x_i^t, y_i^t, w_{iv}^t, p_{iv}^t, md_{iv}^t, b_i^t\}$  تعریف می‌شود. مشخصه‌های  $x_i^t$  و  $y_i^t$  تعیین کننده موقعیت مکانی کاربر  $u_i$  در زمان  $t$  بر اساس طول و عرض جغرافیایی می‌باشند.  $w_{iv}^t$  تعیین کننده بار کاری مولفه  $v$  کاربر  $u_i$  در زمان  $t$  بوده و  $p_{iv}^t$  سرعت پردازش مورد نیاز مولفه  $v$  کاربر  $u_i$  در زمان  $t$  را مشخص می‌کند. همچنین  $md_{iv}^t$  تعیین کننده کد و فراداده مهاجرت داده شده میان دو خدمت‌گزار برای مولفه  $v$  کاربر  $u_i$  در زمان  $t$  و  $b_i^t$  تعیین کننده پهنای باند مورد نیاز کاربر  $u_i$  در زمان  $t$  می‌باشد. همچنین، به عنوان ورودی داده‌ی مولفه  $v$  کاربر  $u_i$  مجموع وزن یال‌های ورودی به آن مولفه از سایر مولفه‌های برنامه، در نظر گرفته شده و با  $id_{iv}$  مشخص می‌شود.

هر خدمت‌گزار  $e_m \in E$  دارای مشخصاتی می‌باشد که به صورت چندتایی  $e_m = \{x_m, y_m, \rho_m, \beta_m\}$  نمایش داده می‌شود.  $x_m$  و  $y_m$  مشخص کننده موقعیت مکانی خدمت‌گزار لبه  $e_m$  می‌باشد. همچنین،  $\rho_m$  تعیین کننده سرعت پردازش خدمت‌گزار  $e_m$  و  $\beta_m$  تعیین کننده پهنای باند در دسترس آن می‌باشد. در جدول ۱، نمادهای مورد استفاده در مدل سیستم و مدل سازی مسئله و شرح مختصری از هر یک آورده شده است.

جدول ۱: فهرست نمادهای بکار رفته در روش پیشنهادی

نماد	توضیح
$t$	شاخص تعیین کننده شکاف زمانی
$E$	مجموعه خدمت‌گزاران لبه
$M$	تعداد خدمت‌گزاران لبه
$m$	شاخص تعیین کننده خدمت‌گزاران
$e_m$	تعیین کننده هر خدمت‌گزار لبه
$x_m, y_m$	تعیین کننده موقعیت مکانی خدمت‌گزار لبه $e_m$
$\rho_m$	سرعت پردازش خدمت‌گزار لبه $e_m$
$\beta_m$	پهنای باند خدمت‌گزار لبه $e_m$
$U$	مجموعه کاربران
$N$	تعداد کاربران
$i$	شاخص تعیین کننده کاربران

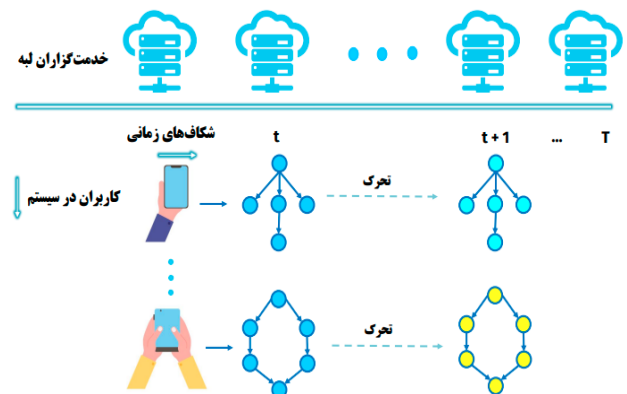
درحالیست که با تخلیه بار ریزدانه برنامه کاربردی بر اساس مولفه‌های تشکیل دهنده آن، علاوه بر امکان اخذ تصمیم تخلیه مناسب‌تر به ازای هر مولفه در شرایط پویای زمینه، می‌توان به کاهش در هزینه و سربار در صورت وقوع مهاجرت برای یک مولفه به نسبت کل برنامه دست یافت.

### ۳- روش تخلیه بار محاسباتی تحرک آگاه مبتنی بر پیش‌بینی

به منظور برطرف کردن چالش‌های اشاره شده در بخش پیشین، در ادامه به ارائه یک روش تخلیه بار تحرک آگاه مبتنی بر پیش‌بینی در رایانش لبه با دسترسی چندگانه برای اینترنت اشیاء می‌پردازیم. در این راستا در این بخش ابتدا، مدل و اجزای تشکیل دهنده سیستم را تشریح کرده و سپس به معرفی نحوه مدل سازی هدف مسئله یعنی کمینه کردن برآیند تاخیر تخلیه بار ریزدانه برای تمامی کاربران در طول زمان، تحت قیود مشخص می‌پردازیم. در ادامه در بخش بعد، به ارائه الگوریتم تخلیه بار تحرک آگاه ریزدانه مبتنی بر پیش‌بینی پرداخته و با استفاده از هزینه برآورد شده حاصل از این الگوریتم، تخصیص هوشمندانه خدمت‌گزاران لبه را به مولفه‌های برنامه کاربران رقم می‌زنیم.

#### ۳-۱- مدل سیستم

همانطور که در شکل ۱ نمایش داده شده است، در مدل سیستم پیشنهادی مجموعه‌ای از کاربران و خدمت‌گزاران لبه موجود می‌باشند. کاربران در سیستم با مجموعه  $U = \{u_1, \dots, u_i, \dots, u_N\}$  مشخص می‌شوند، بطوریکه  $|U| = N$ . مجموعه خدمت‌گزاران لبه با دسترسی چندگانه با مجموعه  $E = \{e_1, \dots, e_m, \dots, e_M\}$  مشخص شده و شامل  $M$  خدمت‌گزار می‌باشد. مدل سیستم مبتنی بر شکاف‌های زمانی  $1, 2, \dots, T$  می‌باشد.



شکل ۱: نمای کلی از اجزا سیستم

گزار  $e_{m'}$  مهاجرت داده می شود.

مسئله تخلیه بار به صورت یک مسئله برنامه ریزی خطی صحیح<sup>۲۳</sup> فرمول می شود که به شرح زیر است:

$$\min_{s_{ivm}^t, \alpha_{ivmm'}^t} \sum_{i=1}^N \sum_{v=1}^V \sum_{m=1}^M \sum_{t=1}^T \left( \frac{f_d^t(u_i, e_m)}{\theta} + \frac{id_{iv}^t}{b_i^t} + \frac{w_{iv}^t}{p_{iv}^t} \right) s_{ivm}^t + \sum_{i=1}^N \sum_{v=1}^V \sum_{m=1}^M \sum_{m'=1}^M \sum_{t=2}^T \left( \frac{f_d^t(e_m, e_{m'})}{\theta} + \frac{md_{iv}^t}{b_{mm'}^t} \right) \alpha_{ivmm'}^t \quad (2)$$

s. t:

$$\sum_{m=1}^M s_{ivm}^t = 1, \forall v \in V, u_i \in U, t \in T \quad (3)$$

$$\sum_{i=1}^N \sum_{v=1}^V p_{iv}^t s_{ivm}^t \leq \rho_m, \forall e_m \in E, t \in T \quad (4)$$

$$\sum_{i=1}^N \sum_{v=1}^V b_i^t s_{ivm}^t \leq \beta_m, \forall e_m \in E, t \in T \quad (5)$$

$$\alpha_{ivmm'}^t \geq s_{ivm}^{t-1} + s_{ivm'}^t - 1, \forall v \in V, u_i \in U, e_m, e_{m'} \in E, m \neq m', \forall t \in T \quad (6)$$

$$s_{ivm}^t \in \{0, 1\}, \forall v \in V, u_i \in U, e_m \in E, t \in T \quad (7)$$

$$\alpha_{ivmm'}^t \in \{0, 1\}, \forall v \in V, u_i \in U, e_m, e_{m'} \in E, m \neq m', \forall t \in T \quad (8)$$

فرمول (۲) برآیند زمان تخلیه بار را برای همه مولفه های کاربران در تمامی شکاف های زمانی کمینه می کند. بخش اول فرمول زمان تخلیه و پردازش و بخش دوم آن زمان مهاجرت را محاسبه می کند. زمان تخلیه و مهاجرت از جنس تاخیر ارتباطی بوده و طبق فرمول (۱) به زمان انتشار و انتقال تفکیک می شود. در کسر اول مربوط به زمان انتشار، تابع  $f_d^t(u_i, e_m)$  فاصله میان کاربر  $u_i$  با خدمت گزار  $e_m$  را در زمان  $t$  با  $\sqrt{(x_i^t - x_m)^2 + (y_i^t - y_m)^2}$  محاسبه می کند. کسر دوم مربوط به زمان انتقال بوده و از تقسیم ورودی داده ی هر مولفه  $v$  کاربر  $u_i$  بر پهنای باند مورد نیاز کاربر در زمان  $t$  به دست می آید. در نهایت، کسر سوم مربوط به زمان پردازش بوده و از تقسیم بار کاری بر سرعت پردازش مورد نیاز مولفه کاربر  $u_i$  در زمان  $t$  حاصل می شود.

در بخش دوم فرمول (۲)، زمان مهاجرت بر اساس زمان انتشار و انتقال داده جهت مهاجرت محاسبه می شود. زمان انتشار با تقسیم تابع  $f_d^t(e_m, e_{m'}) = \sqrt{(x_{m'} - x_m)^2 + (y_{m'} - y_m)^2}$  میان دو خدمت گزار لبه  $e_m$  و  $e_{m'}$  را محاسبه می کند بر سرعت انتشار ( $\theta$ ) حاصل می شود. در کسر دوم نیز برای محاسبه زمان انتقال، حجم کد و فراداده مربوط به مهاجرت یک مولفه در زمان  $t$  بر پهنای باند میان دو خدمت گزار  $e_m$  و  $e_{m'}$  در همان زمان تقسیم می شود. همچنین، از آنجایی که امکان مهاجرت در شکاف زمانی اول وجود ندارد، این محاسبه از شکاف زمانی دوم آغاز می گردد.

قید (۳) مشخص می کند که هر مولفه در هر زمان تنها باید به یک

$u_i$	تعیین کننده هر کاربر
$V_i$	مجموعه مولفه های کاربر $u_i$
$V$	تعداد مولفه های هر کاربر
$v$	شاخص تعیین کننده مولفه ها
$v_{iv}^t$	مجموعه مشخصات مولفه $v$ کاربر $u_i$ در زمان $t$
$x_i^t, y_i^t$	تعیین کننده موقعیت مکانی کاربر $u_i$ در زمان $t$
$w_{iv}^t$	تعیین کننده بار کاری مولفه $v$ کاربر $u_i$ در زمان $t$
$p_{iv}^t$	تعیین کننده سرعت پردازش مولفه $v$ کاربر $u_i$ در زمان $t$
$md_{iv}^t$	تعیین کننده کد و فراداده مولفه $v$ کاربر $u_i$ در زمان $t$
$b_i^t$	تعیین کننده پهنای باند کاربر $u_i$ در زمان $t$
$id_{iv}$	ورودی داده به مولفه $v$ کاربر $u_i$
$s_{ivm}^t$	تخصیص مولفه $v$ کاربر $u_i$ در زمان $t$ به خدمت گزار $e_m$
$\alpha_{ivmm'}^t$	مهاجرت مولفه $v$ کاربر $u_i$ در زمان $t$ از خدمت گزار $e_m$ به خدمت گزار $e_{m'}$
$b_{mm'}^t$	پهنای باند میان دو خدمت گزار $e_m$ و $e_{m'}$ در زمان $t$
$\theta$	سرعت انتشار

### ۳-۲- مدل سازی مسئله

به منظور رفع چالش های ذکر شده، ما مسئله تخلیه بار را بر اساس کمینه کردن برآیند زمان تخلیه بار محاسباتی برای همه مولفه های کاربران در طول تمامی شکاف های زمانی تعریف شده در سیستم مدل سازی نموده ایم و این زمان خود شامل سه زمان تخلیه، پردازش و مهاجرت می باشد.

زمان تخلیه و مهاجرت به عنوان تاخیر ارتباطی در هنگام ارسال داده از کاربر به خدمت گزار لبه و یا از خدمت گزار لبه به دیگری در نظر گرفته می شود. فرمول (۱) محاسبه کلی تاخیر را مشخص می کند:

$$l = \frac{d}{\theta} + \frac{s}{\beta} \quad (1)$$

در این فرمول، کسر اول زمان انتشار را بر اساس  $d$  که مشخص کننده فاصله میان کاربر و خدمت گزار لبه یا دو خدمت گزار لبه و  $\theta$  که نشان دهنده سرعت انتشار است، محاسبه می کند. در کسر دوم، زمان انتقال برحسب  $s$  به عنوان اندازه داده منتقل شده و  $\beta$  به عنوان پهنای باند محاسبه می شود.

مسئله تخلیه بار یافتن تخصیص بهینه خدمت گزاران لبه برای هر یک از مولفه های کاربران در هر یک از شکاف های زمانی می باشد به نحوی که برآیند زمان تخلیه بار کمینه گردد. برای این مسئله دو متغیر تصمیم گیری دودویی  $s_{ivm}^t$  و  $\alpha_{ivmm'}^t$  را تعریف نمودیم که به ترتیب جهت مشخص کردن تخصیص مولفه  $v$  کاربر  $u_i$  در زمان  $t$  به خدمت گزار لبه  $e_m$  و مهاجرت مولفه  $v$  کاربر  $u_i$  در زمان  $t$  از خدمت گزار  $e_m$  به خدمت گزار  $e_{m'}$  استفاده می شود. در صورتیکه  $s_{ivm}^t$  مقدار یک را به خود بگیرد، مشخص می کند مولفه  $v$  کاربر  $u_i$  در زمان  $t$  به خدمت گزار لبه  $e_m$  تخصیص داده شده است. همچنین، اگر متغیر تصمیم گیری  $\alpha_{ivmm'}^t$  مقدار یک را به خود بگیرد، مشخص می کند که مولفه  $v$  کاربر  $u_i$  در زمان  $t$  از خدمت گزار  $e_m$  به خدمت

```

8:   if  $t \bmod \omega = 0$  then
9:      $\{\bar{v}_{iv}^{t+1}, \dots, \bar{v}_{iv}^{t+\omega}\} = \text{PREDICT}(\hat{U}_{iv}^t, t, \omega)$ 
10:     $C_{iv}^t \leftarrow \infty$ 
11:    for all  $e_m \in E$  do
12:       $c_{ivm}^t = \frac{f_d^t(u_i, e_m)}{\theta} + \frac{id_{iv}}{b_i^t} + \frac{w_{iv}^t}{p_{iv}^t}$ 
13:       $c_{ivm}^t +$ 
14:         $= \sum_{\tau=1}^{\omega} \gamma^\tau \left( \frac{f_d^{t+\tau}(\bar{u}_i, e_m)}{\theta} + \frac{id_{iv}}{b_i^{t+\tau}} \right.$ 
15:         $\left. + \frac{\bar{w}_{iv}^{t+\tau}}{\bar{p}_{iv}^{t+\tau}} \right)$ 
16:      if  $t \geq 2$  and  $m \neq \hat{R}_{iv}^{t-1}$  then
17:         $m' = \hat{R}_{iv}^{t-1}$ 
18:         $c_{ivm}^t += \frac{f_d(e_m, e_{m'})}{\theta} + \frac{md_{iv}^t}{b_{mm'}^t}$ 
19:      if  $t \bmod \omega = 0$  then
20:        while  $\exists \hat{R}_{iv}^t \in \hat{R}^t$  which is not assigned do
21:           $S^t = \text{ASSIGN}(C^t, U^t, E)$ 
22:           $\hat{R}^t \leftarrow S^t$ 
23:          update  $C^t, U^t, E$ 
24:        else
25:          for all  $u_i \in U$  do
26:            for all  $v \in V$  do
27:               $\hat{R}_{iv}^t = \hat{R}_{iv}^{t-1}$ 
28:    return  $\bigcup_{t=1}^T \hat{R}^t$ 

```

این الگوریتم از  $\omega$  به عنوان اندازه پنجره برای تعیین تعداد شکاف زمانی بعدی برای پیش‌بینی مشخصات و لحاظ در تصمیم‌گیری تخلیه و مهاجرت استفاده می‌کند. همچنین  $\gamma$  نرخ تخفیفی میان صفر و یک است که از آن برای لحاظ تاثیر هر شکاف زمانی استفاده می‌کنیم. همچنین  $\hat{U}_{iv}^t$  سابقه مشخصات هر مولفه  $v$  کاربر  $u_i$  را از اولین شکاف زمانی تا شکاف  $t$  در خود حفظ می‌کند و به صورت  $\hat{U}_{iv}^t = \bigcup_{\delta=1}^t v_{iv}^\delta$  می‌باشد.  $\hat{R}_{iv}^t$  سابقه خدمت‌گزار نخصیص‌یافته مولفه  $v$  کاربر  $u_i$  در شکاف زمانی  $t$  را در خود نگه می‌دارد و چنانچه  $\hat{R}_{iv}^t = m$  باشد، خدمت‌گزار  $e_m$  به مولفه  $v$  کاربر  $u_i$  در زمان  $t$  تخصیص یافته است.

الگوریتم در پی این است که با استفاده از پیش‌بینی مشخصات مولفه به خصوص موقعیت مکانی کاربر در طول زمان، برآیند زمان تخلیه‌بار را کمینه کند. پیش‌بینی مشخصات در  $\omega$  زمان بعدی براساس سابقه مشخصات است و روش پیش‌بینی مورد استفاده توسط تابع PREDICT در الگوریتم نمود یافته است.

در ادامه الگوریتم ۱ بررسی می‌کند که شکاف زمانی فعلی مضرری از  $\omega$  باشد تا اطمینان حاصل شود که تخصیص به دست آمده، برای  $\omega$  زمان بعدی نیز معتبر است. از آنجایی که مشخصات پیش‌بینی-شده در شکاف‌های زمانی آینده تاثیر یکسانی بر تصمیم‌گیری ندارند، با استفاده از  $\gamma$  و وزن‌دهی، تاثیر آن مورد مدیریت قرار می‌گیرد. برای شکاف‌های زمانی نزدیک به شکاف  $t$  نرخ تخفیف دارای توان

خدمت‌گزار لبه تخصیص یابد. قیود (۴) و (۵) نیز تضمین می‌کنند که مجموع سرعت پردازش و پهنای باند تخصیص داده‌شده به کاربران برای هر خدمت‌گزار لبه در هر شکاف زمانی، از سرعت پردازش و پهنای باند در دسترس آن خدمت‌گزار بیشتر نباشد. قید (۶) اطمینان حاصل می‌کند که متغیر تصمیم‌گیری مهاجرت، تنها زمانی می‌تواند مقدار یک را به خود بگیرد که یک مولفه در شکاف زمانی قبلی به خدمت‌گزاری متفاوت از خدمت‌گزار شکاف فعلی تخصیص-یافته باشد. همچنین، قیود (۷) و (۸) نشان می‌دهند که متغیرهای تصمیم‌گیری بصورت دودویی تعریف شده‌اند.

با توجه به اینکه مسئله تخلیه‌بار مدل‌سازی شده یک مسئله NP-hard می‌باشد، جواب بهینه آن در مقیاس‌های بزرگ و متوسط در زمان منطقی قابل حصول نمی‌باشد. به منظور حل این مسئله در زمان قابل قبول، در بخش بعد ابتدا به ارائه الگوریتم تخلیه‌بار محاسباتی ریزدانه تحرک‌آگاه مبتنی بر پیش‌بینی مشخصات مولفه-های برنامه کاربران، به خصوص موقعیت مکانی آن‌ها پرداخته و سپس با یک روش مبتنی بر برنامه‌ریزی پویا جهت یافتن جواب نزدیک به بهینه مسئله گام برمی‌داریم. این روش با توجه به ارائه تخصیص نزدیک به بهینه در زمان منطقی در مقیاس بزرگ، به افزایش مقیاس‌پذیری تخلیه‌بار محاسباتی در بستر اینترنت اشیا منجر می‌شود.

#### ۴- الگوریتم حل مسئله

به منظور ارائه روش حل مسئله ابتدا مشخصات مولفه‌های هر کاربر به خصوص موقعیت مکانی آن‌ها در شکاف‌های زمانی بعدی براساس مشخصات در شکاف‌های زمانی قبلی با استفاده از روش انحراف و میانگین وزن‌دار، پیش‌بینی می‌شوند. در ادامه با توجه به مشخصات هر مولفه در زمان فعلی و در شکاف‌های زمانی آینده، هزینه تخلیه‌بار از نظر برآیند تاخیر مورد انتظار کاربران برآورد می‌شود. در نهایت با استفاده از هزینه برآوردشده، تخصیص هوشمندانه مولفه‌ها به خدمت‌گزاران لبه در جهت رسیدن به هدف با یک روش مبتنی بر برنامه‌ریزی پویا صورت می‌پذیرد. شبکه کد حل مسئله در الگوریتم ۱ ارائه شده است.

الگوریتم ۱. تخصیص برخط مولفه‌های کاربران به خدمت‌گزاران لبه با دسترسی چندگانه

```

1: Input:  $\omega$ : window size,  $\gamma$ : discount factor
2:  $\hat{U}_{iv}^0 = 0, \forall v \in V, u_i \in U$ 
3:  $\hat{R}_{iv}^0 = 0, \forall v \in V, u_i \in U$ 
4: for all  $t \in T$  do
5:   for all  $u_i \in U$  do
6:     for all  $v \in V$  do
7:        $\hat{U}_{iv}^t = \hat{U}_{iv}^{t-1} \cup v_{iv}^t$ 

```

$$md_{iv}^{t+h} = \frac{t \times md_{iv}^t + (t-1) \times md_{iv}^{t-1} + \dots + md_{iv}^1}{t + (t-1) + \dots + 1} \quad (10)$$

این شیوه پیش‌بینی برای مشخصات  $w_{iv}^t$ ،  $p_{iv}^t$  و  $b_i^t$  نیز مورد استفاده قرار می‌گیرد.

#### ۴-۲- تخصیص خدمت‌گزاران به مولفه‌های کاربران مبتنی بر برنامه‌ریزی پویا

به منظور تخصیص خدمت‌گزاران لبه به مولفه‌های کاربران در الگوریتم ۱ تابع ASSIGN را فراخوانی می‌کنیم. این تابع با یک مسئله تخصیص تعمیم‌یافته<sup>۲۳</sup> که از نوع NP-hard می‌باشد، مواجه است. برای حل مسائل از این نوع، الگوریتم‌های نزدیک به بهینه متعددی موجود است که در برخی از آن‌ها زمان اجرا نامایی می‌باشد. در این پژوهش از روش تطبیق‌یافته کار [۱۴] برای حل مسئله استفاده شده است. در این تابع هزینه برآورد شده، مشخصات مولفه‌های برنامه کاربر و مشخصات خدمت‌گزاران لبه دریافت شده و تصمیم تخصیص خدمت‌گزاران به مولفه‌های کاربران در زمان  $t$  را باز می‌گرداند.

#### ۵- ارزیابی روش پیشنهادی

به منظور ارزیابی و حل، روش پیشنهادی به زبان برنامه‌نویسی C++ نوشته شده است. مشخصات سیستمی که آزمایش‌ها در آن انجام گرفته، یک PC با پردازنده Corei7-6500، 2.50 GHz و 16 GB RAM می‌باشد. طبق آنچه در قسمت مدل‌سازی مسئله شرح داده شد، هدف مسئله کمینه‌کردن برآیند زمان تخلیه‌بار برای همه مولفه‌های کاربران در شکاف‌های زمانی می‌باشد. بنابراین معیارهای ارزیابی عملکرد مدل‌سازی پیشنهادی و مقایسه آن با کارهای مرتبط به دو دسته پیچیدگی تخلیه‌بار و عملکرد آن تقسیم می‌شود. در بررسی عملکرد برآیند زمان تخلیه‌بار، زمان تخلیه و زمان مهاجرت ارزیابی و مقایسه شده و برای بررسی پیچیدگی تخلیه‌بار نیز زمان اجرای الگوریتم مورد تحلیل قرار می‌گیرد.

به منظور ارزیابی عملکرد روش پیشنهادی، روش پیشنهادی ارائه شده در این مقاله با کارهای مرتبط اخیر OAMC<sup>۲۴</sup> [۱۱]، GOAMC<sup>۲۵</sup> [۱۲]، BFD-C<sup>۲۶</sup> [۱۳] و BFD-F<sup>۲۷</sup> که یک توسعه ریزدانه از کار [۱۳] می‌باشد مورد مقایسه قرار گرفته است. لازم به ذکر است روش حل مورد استفاده در BFD-F و BFD-C مبتنی بر بهترین کاهش تناسب<sup>۲۸</sup> به عنوان یکی از روش‌های اکتشافی<sup>۲۹</sup> می‌باشد و از پیش‌بینی موقعیت مکانی و مشخصات برنامه کاربران استفاده نمی‌کند.

برای بررسی معیارهای ارزیابی در ادامه به معرفی پارامترهای ارزیابی می‌پردازیم. خدمت‌گزاران لبه در آزمایش‌های صورت گرفته به تعداد

کمتری بوده و مقدار بیشتری دارد و با گذشت زمان و در شکاف‌های زمانی دورتر نسبت به زمان فعلی نرخ تخفیف دارای توان بیشتری بوده و مقدار کمتری خواهد داشت. الگوریتم پس از مرحله پیش‌بینی مشخصات، هزینه مورد انتظار یعنی برآیند زمان تخلیه‌بار را محاسبه می‌کند که با  $C_{iv}^t$  مشخص شده است و زمان تخلیه، پردازش و مهاجرت را پوشش می‌دهد. در ادامه در صورتی که شکاف زمانی فعلی مضربی از  $\omega$  باشد براساس هزینه برآورد شده، تخصیص خدمت‌گزاران موجود به مولفه‌های کاربران فراخوانی می‌شود. تابع ASSIGN هزینه در زمان فعلی به صورت  $C^t = U_{i=1}^N U_{v=1}^V U_{m=1}^M c_{ivm}^t$  و مشخصات کاربران و خدمت‌گزاران را به عنوان ورودی دریافت کرده و تصمیم تخصیص خدمت‌گزاران به مولفه‌های کاربران در زمان  $t$  را باز می‌گرداند. این تخصیص در خط ۲۰ الگوریتم، در سابقه تخصیص حفظ شده و مشخصات خدمت‌گزاران به منظور انعکاس منابع موجود آن‌ها بروزرسانی می‌شود. این تخصیص تا زمانی که مولفه متصل نشده به خدمت‌گزار باقی بماند ادامه یافته و تکرار می‌شود.

چنانچه شکاف زمانی فعلی مضربی از  $\omega$  نباشد، تمامی مولفه‌ها در آن زمان به خدمت‌گزار تخصیص یافته به آن‌ها در زمان قبلی متصل می‌شوند. در انتها نیز الگوریتم تخصیص تمام خدمت‌گزاران به تمامی مولفه‌های کاربران از شکاف زمانی اول تا شکاف زمانی انتهایی تعریف شده در سیستم را به عنوان خروجی، نتیجه می‌دهد.

#### ۴-۱- پیش‌بینی مشخصات مولفه‌های کاربران

در این زیربخش برای پیش‌بینی مشخصات که در تابع PREDICT الگوریتم ۱ نمود یافته است دو روش تعیین می‌شود. برای پیش‌بینی موقعیت مکانی کاربران از روش انحراف استفاده می‌کنیم. در این روش میزان تغییرات بستگی به میانگین تغییرات داده‌های پیشین داشته و فرض می‌شود طول و عرض جغرافیایی کاربران به میزان این میانگین به نسبت مقدار قبلی خود تغییر خواهند کرد. مقدار پیش‌بینی شده برای طول جغرافیایی به شرح زیر بوده و قابل تعمیم برای عرض نیز می‌باشد:

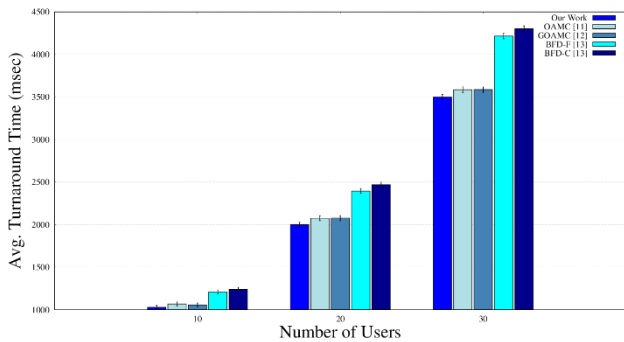
$$x_i^{t+h} = x_i^t + \frac{h}{t-1} \sum_{v=2}^t (x_i^v - x_i^{v-1}) = x_i^t + \frac{h}{t-1} (x_i^t - x_i^1) \quad (9)$$

مقدار پیش‌بینی شده در زمان  $t+h$  با  $x_i^{t+h}$  مشخص است.

برای پیش‌بینی سایر مشخصات هر مولفه نیز از روش میانگین وزن-دار استفاده می‌کنیم. این روش برای مشخصات در شکاف‌های اخیر و نزدیک به زمان فعلی ارزش بیشتری قائل بوده و ضرایب بزرگتری لحاظ می‌کنیم. پیش‌بینی مقدار کد و فراداده مولفه  $v$  کاربر  $i$  در زمان  $t+h$  به صورت زیر است:

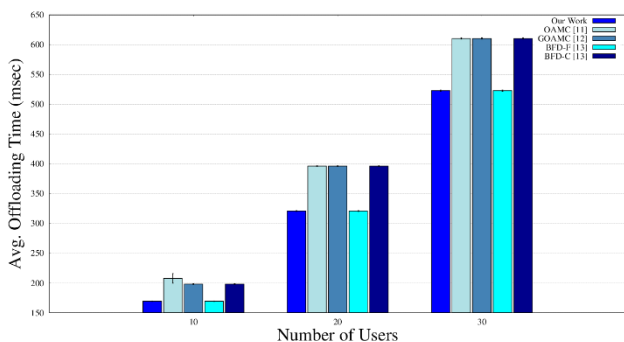


زمان تخلیه بار افزایش یافته و روش پیشنهادی به جهت استفاده از مزایای ریزدانگی دارای بهبود به نسبت کار [۱۱] و [۱۲] می باشد. همچنین، در روش های BFD-F و BFD-C [۱۳] به علت عدم استفاده از پیش بینی موقعیت مکانی و همچنین مشخصات مولفه ها یا برنامه کاربر، برآیند زمان تخلیه بار با افزایش قابل توجهی مواجه شده است که این افزایش در BFD-F به علت بهره مندی از مزایای ریزدانگی کمتر از BFD-C بوده است.



شکل ۲: تاثیر تعداد کاربران بر برآیند زمان تخلیه بار محاسباتی

همچنین، در شکل ۳ به بررسی تاثیر افزایش تعداد کاربران بر زمان تخلیه می پردازیم. مطابق نتایج بدست آمده، زمان تخلیه با افزایش تعداد کاربران افزایش یافته و با توجه به رویکرد اخذ تصمیم مناسب برای هر مولفه در روش پیشنهادی، دارای بهبود به نسبت کارهای مرتبط می باشد. همچنین BFD-F نیز به جهت اخذ تصمیم به ازای هر مولفه به نسبت کل برنامه، به نسبت BFD-C موفق تر عمل کرده است.



شکل ۳: تاثیر تعداد کاربران بر زمان تخلیه

در شکل ۴، تاثیر افزایش تعداد کاربران بر زمان مهاجرت مورد ارزیابی قرار گرفته است. مطابق این شکل، با افزایش تعداد کاربران زمان مهاجرت افزایش می یابد زیرا به ازای افزایش کاربران تعداد مولفه ها نیز افزایش یافته و به تبع، زمان مهاجرت برای آن ها نیز بیشتر خواهد شد. همچنین، به جهت سربار کمتر حاصل از مهاجرت یک مولفه به نسبت کل کار، روش پیشنهادی دارای بهبود نسبت به کارهای مورد مقایسه می باشد. لازم به ذکر است از آنجایی که زمان پردازش از تقسیم بار کاری هر مولفه به سرعت پردازش مورد نیاز آن حاصل می شود و فارغ از رویکرد تخصیص بوده از درج نمودار مربوط

۱۰، ۲۰ و ۳۵ می باشد. موقعیت مکانی خدمت گزاران به صورت تصادفی و حاصل از توزیع یکنواخت در موقعیت مکانی کاربران توزیع شده است. همچنین فرض شده هر خدمت گزار دارای سرعت پردازش و پهنای باند معادل مجموع  $U(1,20)$  تعداد ماشین مجازی است که هر ماشین دارای سرعت پردازش  $U(25000,50000)$  MHz و پهنای باند  $U(5000,10000)$  Mbps می باشد. همچنین، در آزمایش ها انجام گرفته تعداد کاربران ۱۰، ۲۰ و ۳۰ بوده برنامه هر برنامه کاربردی کاربر به صورت تصادفی بین ۴ تا ۹ مولفه را شامل می شود. مجموع تعداد مولفه ها به ازای تعداد کاربران تعیین شده در آزمایش ها به ترتیب ۶۰، ۱۲۰ و ۲۱۰ مولفه می باشد. موقعیت مکانی کاربران در هر شکاف زمانی با استفاده از ۳ مجموعه داده متفاوت با چندین الگوی تحرک از جمله پیاده روی، رانندگی و حرکت با اتوبوس و مترو به دست آمده است [۱۵-۱۶]. بار کاری در بازه  $MI [7000,600000]$  مشخص شده و سرعت پردازش مورد نیاز آن در بازه  $MIPS [90,100000]$  تعیین شده است. پهنای باند کاربر در بازه  $Mbps [20,120]$  تعیین شده و داده فراخوانی شده میان دو مولفه نیز در برد  $Mb [1,9]$  می باشد. تعداد شکاف زمانی تعریف شده در سیستم، ۲۰ شکاف بوده و هر آزمایش به تعداد ۵۰ بار تکرار شده است.

جدول ۲ برای معرفی پارامترهای ارزیابی ارائه شده است.

جدول ۲: پارامترهای ارزیابی

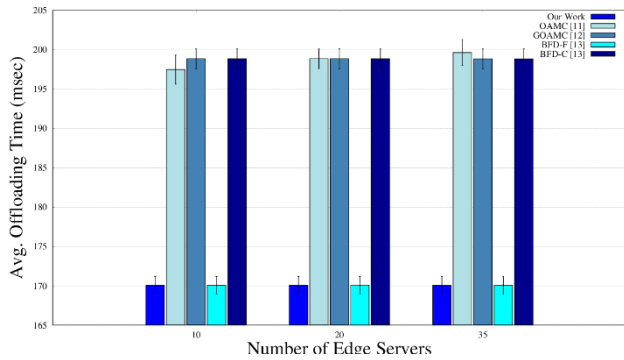
مقدار	پارامتر ارزیابی
10, 20, 35	تعداد خدمت گزاران
معادل $U(1,20)$ ماشین مجازی با سرعت پردازش $U(25000,50000)$	سرعت پردازش خدمت گزار
معادل $U(1,20)$ ماشین مجازی با پهنای باند $U(5000,10000)$	پهنای باند خدمت گزار
10, 20, 30	تعداد کاربران
60, 120, 210	تعداد مولفه ها
$[7000,600000]$ MI	بار کاری مولفه
$[90,100000]$ MIPS	سرعت پردازش مورد نیاز مولفه
$[1,9]$ Mb	داده فراخوانی شده بین دو مولفه
$[20,120]$ Mbps	پهنای باند کاربر
20	تعداد شکاف زمانی
50	تعداد تکرار

#### ۵-۱- بررسی تاثیر تعداد کاربران بر نتایج ارزیابی

در آزمایش های صورت گرفته در این قسمت تعداد خدمت گزاران لبه به تعداد ۳۵ نگاه داشته شده و تعداد کاربران به ترتیب ۱۰، ۲۰ و ۳۰ می باشد. در شکل ۲ تاثیر افزایش تعداد کاربران و به تبع آن تعداد مولفه ها در تابع هدف مسئله مورد ارزیابی قرار گرفته است. همانطور که نتایج نشان می دهد، با افزایش تعداد کاربران برآیند

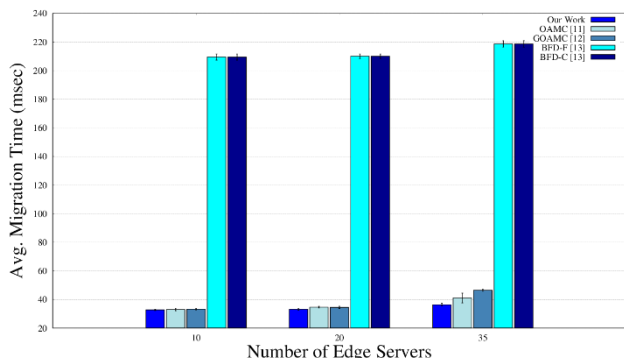
شکل ۶: تاثیر تعداد خدمت گزاران لبه بر برآیند زمان تخلیه بار

در شکل ۷، تاثیر افزایش تعداد خدمت گزاران لبه بر زمان تخلیه مورد مطالعه قرار گرفته است. مطابق این شکل، با افزایش تعداد خدمت گزاران زمان تخلیه روند تقریباً ثابتی را طی می کند اما روش پیشنهادی و روش BFD-F به جهت لحاظ ریزدانی و اخذ تصمیم بهینه به ازای هر مولفه به جای تصمیم برای کل برنامه، زمان تخلیه کمتری نسبت به روش های دیگر دارند.

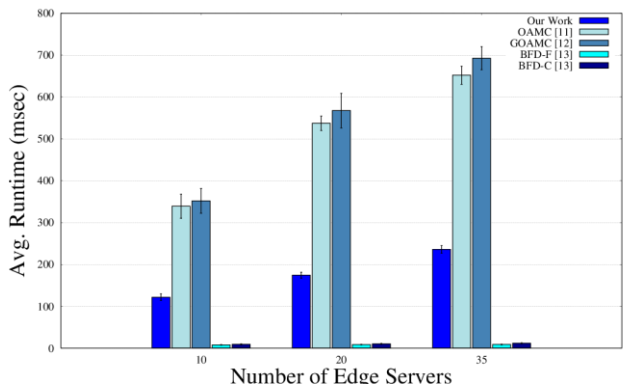


شکل ۷: تاثیر تعداد خدمت گزاران لبه بر زمان تخلیه

شکل ۸ تاثیر افزایش تعداد خدمت گزاران لبه را بر زمان مهاجرت نمایش می دهد. همانطور که در شکل نشان داده شده است، با افزایش تعداد خدمت گزاران تعریف شده در مسئله، امکان مهاجرت بیشتری مهیا شده و تعداد مهاجرت و به تبع آن زمان مهاجرت نیز افزایش می یابد.

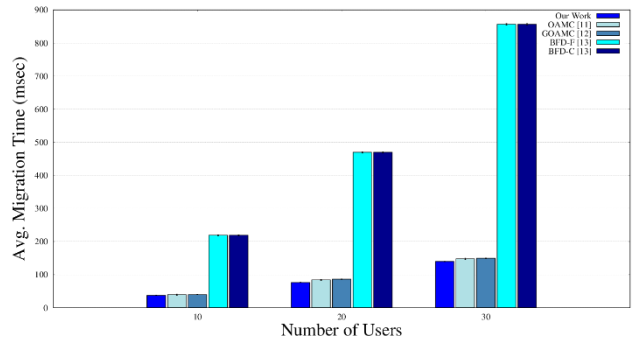


شکل ۸: تاثیر تعداد خدمت گزاران لبه بر زمان مهاجرت



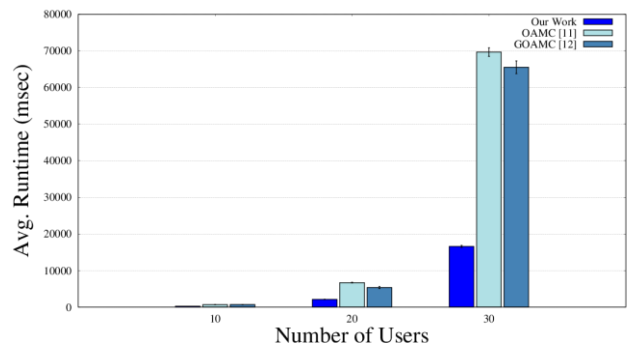
شکل ۹: تاثیر تعداد خدمت گزاران لبه بر زمان اجرای الگوریتم

به آن صرف نظر شده است.



شکل ۴: تاثیر تعداد کاربران بر زمان مهاجرت

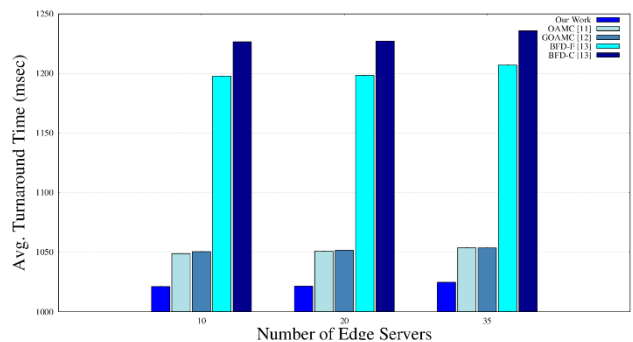
شکل ۵ بیانگر تاثیر افزایش تعداد کاربران بر زمان اجرا و پیچیدگی تخلیه بار می باشد. مطابق این شکل، با توجه به افزایش فضای مسئله با افزایش تعداد کاربران زمان اجرا نیز افزایش می یابد. همچنین، روش پیشنهادی به نسبت کارهای مرتبط از زمان اجرای کمتری برخوردار بوده و تصمیم برای هر مولفه از کاربر به نسبت کل کار با سرعت بیشتری اخذ می شود. همچنین، روش GOAMC به نسبت روش OAMC به علت استفاده از روش تخصیص حریمانه با سرعت بیشتری عمل کرده است.



شکل ۵: تاثیر تعداد کاربران بر زمان اجرای الگوریتم

## ۵-۲- بررسی تاثیر تعداد خدمت گزاران لبه بر نتایج ارزیابی

در آزمایش های صورت گرفته در این قسمت تعداد کاربران به تعداد ۱۰ نگاه داشته شده و تعداد خدمت گزاران به ترتیب ۱۰، ۲۰ و ۳۵ می باشد. در شکل ۶ مشاهده می شود که با افزایش تعداد خدمت گزاران برآیند زمان تخلیه بار افزایش کمی یافته و روش پیشنهادی دارای بهبود قابل توجه به نسبت کارهای مرتبط می باشد.



Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCCom/IoP/SmartWorld), pp.1144-1148, 2016.

[5] D. Wang, Z. Liu, X. Wang, and Y. Lan, "Mobility-aware task offloading and migration schemes in fog computing networks," IEEE Access, vol.7, pp.43356-43368, 2019.

[6] H. Hu, Q. Wang, R. Q. Hu, and H. Zhu, "Mobility-aware offloading and resource allocation in an MEC-enabled IoT network with energy harvesting," IEEE Internet of Things Journal, pp.1-1, 2021.

[7] W. Zhan, C. Luo, G. Min, C. Wang, Q. Zhu, and H. Duan, "Mobility-aware multi-user offloading optimization for mobile edge computing," IEEE Transactions on Vehicular Technology, vol.69, no.3, pp.3341-3356, 2020.

[8] U. Saleem, Y. Liu, S. Jangsher, Y. Li, and T. Jiang, "Mobility-aware joint task scheduling and resource allocation for cooperative mobile edge computing," IEEE Transactions on Wireless Communications, vol.20, no.1, pp.360-374, 2021.

[9] Y. Shi, S. Chen, and X. Xu, "Maga: A mobility-aware computation offloading decision for distributed mobile cloud computing," IEEE Internet of Things Journal, vol.5, no.1, pp.164-174, 2018.

[10] S. Thananjeyan, C. A. Chan, E. Wong, and A. Nirmalathas, "Mobility-aware energy optimization in hosts selection for computation offloading in multi-access edge computing," IEEE Open Journal of the Communications Society, vol.1, pp.1056-1065, 2020.

[11] A. E. F. Maleki and L. Mashayekhy, "Mobility-aware computation offloading in edge computing using prediction," in 2020 IEEE 4th International Conference on Fog and Edge Computing (ICFEC), pp.69-74, 2020.

[12] E. Farhangi Maleki, L. Mashayekhy, and S. M. Nabavinejad, "Mobility-aware computation offloading in edge computing using machine learning," IEEE Transactions on Mobile Computing, pp.1-1, 2021.

[13] K. Braiki, H. Youssef, "Fuzzy-logic-based multi-objective best-fit-decreasing virtual machine reallocation," The Journal of Supercomputing, vol. 76, pp. 427-454, 2020.

[14] R. Cohen, L. Katzir, and D. Raz, "An efficient approximation for the generalized assignment problem," Information Processing Letters, vol. 100, no. 4, pp. 162-166, 2006.

[15] S. Moosavi, B. Omidvar-Tehrani, and R. Ramnath, "Trajectory annotation by discovering driving patterns," in Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics, UrbanGIS'17, (New York, NY, USA), Association for Computing Machinery, 2017.

[16] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the levy-walk nature of human mobility," IEEE/ACM Transactions on Networking, vol.19, no.3, pp.630-643, 2011.

در شکل ۹، تاثیر افزایش تعداد خدمت‌گزاران بر زمان اجرای الگوریتم نمایان شده و با این افزایش، زمان اجرا افزایش می‌یابد. طبق آنچه در زیربخش قبلی بیان شد به نسبت کارهای مرتبط OAMC و GOAMC، این پژوهش در پیچیدگی تخلیه‌بار دارای بهبود می‌باشد. دو روش BFD-F و BFD-C نیز به علت عدم استفاده از پیش‌بینی در حل خود، زمان اجرای کمتری داشته‌اند که البته این امر منجر به تصمیم تخلیه‌بار غیربهبوده و ناکارآمدی بسیار پایین این روش‌ها از نظر معیارهای کیفیت سرویس شده است.

## ۶- نتیجه‌گیری و کارهای آینده

در این مقاله مدل‌سازی و حل یک مسئله تخلیه‌بار محاسباتی تحرک‌آگاه مبتنی بر پیش‌بینی در رایانش لبه با دسترسی چندگانه برای اینترنت اشیا ارائه شد. در این روش مشخصات مولفه‌های کاربر به خصوص موقعیت مکانی در شکاف‌های زمانی آینده مورد پیش‌بینی قرار گرفته و لحاظ پویایی مشخصات مولفه‌ها و تحرک کاربر در مسئله وارد شد. همچنین تخلیه‌بار برای بهره‌مندی از مزایای آن ریزدانه بود. با توجه به نتایج ارزیابی عملکرد و پیچیدگی روش پیشنهادی به نسبت کارهای مورد مقایسه نیز دارای بهبود بوده است. به عنوان کار آینده در نظر داریم از ظرفیت روش‌های مبتنی بر یادگیری ماشین برای پیش‌بینی مشخصات مولفه‌ها و موقعیت مکانی کاربر استفاده کنیم.

## مراجع

[1] N. Abbas, Y. Zhang, A. Taherkordi and T. Skeie, "Mobile Edge Computing: A Survey," in IEEE Internet of Things Journal, vol. 5, no. 1, pp. 450-465, Feb. 2018.

[2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, 2016.

[3] S.K.u Zaman, A.I Jehangiri, T. Maqsood et al. "Mobility-aware computational offloading in mobile edge networks: a survey," Cluster Computing 24, pp. 2735-2756, 2021.

[4] R. Roostaie, Z. Movahedi, "Mobility and Context-Aware Offloading in Mobile Cloud Computing," 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted

## پاورقی‌ها:

MEC Server<sup>۱۲</sup>  
 Mobility-aware<sup>۱۳</sup>  
 Drift<sup>۱۴</sup>  
 Dynamic Programming<sup>۱۵</sup>  
 Energy Harvesting<sup>۱۶</sup>  
 Lyapunov Optimization<sup>۱۷</sup>  
 Base Station<sup>۱۸</sup>  
 Non-Linear Integer Programming (NLIP)<sup>۱۹</sup>  
 Device-to-Device Communication (D2D)<sup>۲۰</sup>  
 Greedy<sup>۲۱</sup>

Fifth Generation Communication Networks<sup>۱</sup>  
 Internet of Things<sup>۲</sup>  
 Augmented Reality<sup>۳</sup>  
 Interactive Gaming<sup>۴</sup>  
 Natural Language Processing<sup>۵</sup>  
 Computation Offloading<sup>۶</sup>  
 Multi-Access Edge Computing<sup>۷</sup>  
 Cellular Networks<sup>۸</sup>  
 Core Network<sup>۹</sup>  
 Congestion Control<sup>۱۰</sup>  
 Realtime<sup>۱۱</sup>

Integer Linear Programming <sup>۲۲</sup>	Best-Fit Decreasing – Coarse-grain <sup>۲۶</sup>
Generalized Assignment Problem (GAP) <sup>۲۳</sup>	Best-Fit Decreasing – Fine-grain <sup>۲۷</sup>
Online Assignment of Mobile applications to Cloudlets <sup>۲۴</sup>	Best-Fit Decreasing <sup>۲۸</sup>
Greedy Online Assignment of Mobile applications to Cloudlets <sup>۲۵</sup>	Heuristic <sup>۲۹</sup>