



دانشگاه کاشان
University of Kashan

مجله محاسبات نرم

SOFT COMPUTING JOURNAL

تارنمای مجله: scj.kashanu.ac.ir



رویکرد تخصیص منبع کارا برای پردازش داده‌های حوزه سلامت در محیط رایانش ابری*

مصطفی قبائی آرانی^{۱*}، استادیار، فاطمه مهدی بابایی^۲، کارشناس ارشد

^۱ گروه مهندسی کامپیوتر، واحد قم، دانشگاه آزاد اسلامی، قم، ایران.

^۲ گروه مهندسی کامپیوتر، واحد محلات، دانشگاه آزاد اسلامی، مرکزی، ایران.

اطلاعات مقاله

چکیده

تاریخچه مقاله:

دریافت ۰۳ اردیبهشت ماه ۱۳۹۹

پذیرش ۲۷ مهر ماه ۱۳۹۹

کلمات کلیدی:

تخصیص منابع

رایانش ابری

کنفدراسیون ابر

کیفیت خدمات

انتخابگر فازی

داده‌های حوزه سلامت

امروزه، پردازش داده‌های مراقبت‌های بهداشتی رسانه‌های بزرگ در ابر به راه‌حلی مؤثر برای برآورده کردن تقاضاهای کیفیت سرویس کاربران پزشکی تبدیل شده است. اما فراهم ساختن مراقبت‌های بهداشتی برای جامعه فعالیتی پیچیده است که شدیداً متکی به پردازش اطلاعات است و پردازش داده‌های حوزه سلامت می‌تواند برای سازمان‌ها بسیار هزینه‌بر باشد. بر همین اساس امروزه، پردازش داده‌های حوزه سلامت رسانه‌های بزرگ در ابر به راه‌حلی مؤثر جهت مرتفع‌سازی تقاضاهای کیفیت سرویس کاربران حوزه بهداشت و سلامت تبدیل شده است. بر همین اساس در این مقاله، یک راهکار فازی برای تعیین ابر بهینه با استفاده از تکنیک پیش‌بینی منابع ارائه می‌شود. همچنین به منظور ایجاد توازن در حین پردازش وظایف از تکنیک مهاجرت ماشین‌مجازی مبتنی بر انتخابگر فازی جهت انتقال ماشین‌مجازی از یک سرور با بار زیاد به یک سرور کم‌بار استفاده می‌شود. ساختار معماری پیشنهادی شامل دو بخش محلی و سراسری است که برای تحویل درخواست به قسمت سراسری باید ابتدا قسمت محلی بررسی گردد، در صورت نداشتن شرایط، درخواست به قسمت سراسری تحویل داده می‌شود. در واقع معماری پیشنهادی به صورت سلسه مراتبی عمل می‌کند و در ابتدا لیست درخواست‌های رسیده ایجاد می‌گردد، سپس با استفاده از راهکار پیشنهادی، میزان منابع موجود تخمین زده می‌شود و بر اساس آن، تخصیص منبع جهت پردازش انجام می‌گیرد. جهت ارزیابی از شبیه‌ساز Cloudsim استفاده شده است و با توجه به پارامترهای مختلف، راهکار پیشنهادی با الگوریتم‌های FAHP و ICA-K-Means مورد مقایسه قرار گرفته است. نتایج حاصل از شبیه‌سازی نشان می‌دهد که راهکار پیشنهادی در مقایسه با الگوریتم FAHP به میزان ۱۰ درصد و نسبت به ICA-K-Means به میزان ۱۲ درصد کاهش هزینه بوجود آمده است. همچنین تعداد درخواست‌های رد شده نیز به همین صورت و به نسبت الگوریتم FAHP به میزان ۸ درصد و نهایتاً در مقایسه با ICA-K-Means به میزان ۷ درصد افزایش بهره‌وری منابع بوجود آمده است. © ۱۳۹۹ - مجله محاسبات نرم، کلیه حقوق محفوظ است.

۱. مقدمه

بیماران، نیازمند حجم زیادی از نیروی کار و محاسبات است. این فرایندها معمولاً مستعد خطا و دارای تأخیر زیادی هستند که مانع در دسترس بودن اطلاعات صحیح در زمان واقعی می‌شود. رایانش ابری به‌عنوان یک فناوری نوین، با زیرساخت اینترنتی و راهکارهای جدید، مزایای قابل توجهی را در ارائه خدمات پزشکی به‌صورت الکترونیکی به همراه داشته است. در کنار آن،

فرایندهای موجود برای جمع‌آوری و پردازش اطلاعات حیاتی

* نوع مقاله: پژوهشی

* نویسنده مسئول

پست‌های الکترونیک: m.ghobaei@qom-iau.ac.ir (قبائی آرانی)

babaee1369@gmail.com (مهدی بابایی)

محیط ابری ارائه می‌گردد. به کمک این روش درخواست‌ها به صورت پویا و متوازن در میان تمامی گره‌ها، با حصول اطمینان از تخصیص عادلانه و کارآمد هر منبع رایانشی، توزیع خواهد شد و در واقع هدف، یافتن راهکار تخصیص منابع کارآمد خواهد بود؛ به طوری که در هر پردازنده، مقدار تقریباً مساوی از وظایف اجرا شود و در نتیجه پاسخ و همچنین میزان تخطی از توافقنامه سطح سرویس به کمترین میزان خود می‌رسد. راهکار پیشنهادی از دو بخش محلی و سراسری^۱ تشکیل شده است که برای تحویل درخواست به قسمت سراسری باید ابتدا قسمت محلی بررسی گردد. در صورت نداشتن شرایط، درخواست به قسمت سراسری تحویل داده می‌شود. بر همین اساس، در ابتدا یک راهکار فازی برای تعیین ابر بهینه با استفاده از تکنیک پیش‌بینی منابع ارائه می‌شود. همچنین به منظور ایجاد توازن در حین پردازش وظایف از تکنیک مهاجرت ماشین مجازی مبتنی بر انتخابگر فازی [۶-۵] برای انتقال ماشین مجازی از یک سرور به بار زیاد به یک سرور کم‌بار استفاده می‌شود. دستاوردهای اصلی این پژوهش به شرح زیر است:

- ارائه یک راهکار مبتنی بر منطق فازی برای تعیین ابر بهینه در فدراسیون ابر با تقسیم‌بندی به دو بخش محلی و سراسری؛
 - ارائه ساختار واسطه‌گر مبتنی بر تخمین میزان منابع مورد نیاز با استفاده از تکنیک پیش‌بینی ARIMA^۲؛
 - کاهش هزینه و تعداد درخواست‌های رد شده و افزایش خاصیت کشسانی و بهره‌وری منابع با استفاده از رویکرد پیشنهادی؛
 - ارزیابی راهکار پیشنهادی با استفاده از مجموعه داده‌های حوزه سلامت واقعی تحت ساختار سه لایه برنامه کاربردی ابری؛
- ادامه مقاله بدین صورت سازمان‌دهی شده است: در بخش دوم، سوابق تحقیق در حوزه مدیریت منابع ابری را مرور می‌کنیم در بخش سوم، رویکرد پیشنهادی با جزئیات ارائه شده و بخش چهارم به ارزیابی کارایی رویکرد پیشنهادی اختصاص دارد و در

از طریق پیشرفت سریع فناوری‌های پردازشی و ذخیره‌سازی و نیز موفقیت اینترنت، منابع رایانشی ارزان‌تر، قوی‌تر و در دسترس‌تر از قبل شده‌اند و سازمان‌های دولتی برای ارائه خدمات و برآورده ساختن نیازهای زیرمجموعه خود، شروع به استفاده از معماری، بسترها و برنامه‌های ابری کرده‌اند. اما در این بین چالش‌های زیادی نیز وجود دارد که مدیریت بهینه منبع و اینکه هر داده به چه میزان منبع پردازشی نیاز دارد، از مهم‌ترین این چالش‌هاست [۳-۱]. سیستم مدیریت منبع در رایانش ابری مکانیزمی است که هدف آن تضمین تأمین نیازمندی‌های برنامه‌های کاربردی است. به علاوه مکانیزم مدیریت منبع باید وضعیت جاری هر منبع را در محیط ابر به منظور فراهم ساختن الگوریتم‌هایی برای تخصیص بهتر منابع فیزیکی یا منابع مجازی و بنابراین کاهش هزینه‌های عملیاتی در محیط ابر بررسی کند. واضح است که به دلیل مقیاس و پیچیدگی این سیستم‌ها انتساب متمرکز کارها به سرور مشخصی بدون در نظر گرفتن راهکارهای خاصی، در واقع غیرممکن است [۴] و همچنین با توجه به رشد روزافزون داده‌ها در مراکز داده‌ای پیشرفته و نیاز مبرم برای دستیابی به کیفیت سرویس مطلوب، ضرورت ارائه راهکارهایی جهت افزایش بهره‌وری سرویس‌دهنده‌های موجود در مرکز پردازش داده‌ها احساس می‌شود. امروزه با توسعه زیرساخت‌های ارتباطی، امکان استفاده از رایانش ابری در سایر حوزه‌های کاربردی به خصوص برای حوزه بهداشت و سلامت فراهم شده است. داده‌های حوزه سلامت در بیمارستان‌ها، کلینیک‌ها، مراکز آموزشی، دانشگاه‌ها و... تولید می‌شوند. امکان پردازش این داده‌ها از طریق بستر رایانش ابری فراهم آمده است. برای مثال اگر شخصی بیمار باشد و در شهر A زندگی کند، با مسافرت به شهر B می‌تواند با مراجعه به مراکز درمانی به اطلاعات خود دسترسی داشته باشد. مشتریان در این ساختار، ممکن است مراکز آموزشی درمانی و یا پژوهشی باشند. این مراکز، نگرشی به حجم بار کاری و کنترل بار ندارند، لذا وظیفه زیر ساخت ابری است که این امر را مدیریت کند. در این پژوهش، یک راهکار فازی به منظور مدیریت تخصیص منبع برای پردازش داده‌های حوزه سلامت در

1. Global
2. AutoRegressive Integrated Moving Average

نهایت در بخش پنجم نتیجه‌گیری و پیشنهادات ارائه می‌شود.

۲. کارهای گذشته

در این بخش، مروری بر تحقیقات انجام‌شده در حوزه مدیریت منابع برای پردازش داده‌ها خواهیم داشت:

در [۷]، مروری بر الگوریتم‌های متداول در زمینه ایجاد توازن بار انجام شده است. یکی از پرکاربردترین این الگوریتم‌ها، Min-Min می‌باشد. این الگوریتم با مجموعه‌ای از وظایف تخصیص داده‌نشده شروع می‌شود. اول از همه، زمان اتمام حداقل برای تمام وظایف یافت می‌شود. سپس در میان این زمان‌های حداقل، کمترین مقدار انتخاب می‌شود که حداقل زمان میان تمام وظایف روی هر منبع موجود است. سپس مطابق با زمان حداقل، وظیفه روی ماشین مربوط زمان‌بندی می‌شود. در کنار زمان‌بندی بهینه‌ای که این راهکار ارائه می‌دهد، یک اشکال اصلی دارد که می‌تواند منجر به قحطی شود. در [۸]، یک استراتژی زمان‌بندی ترکیبی برای مدیریت و پردازش داده‌های پزشکی در منابع ماشین‌های مجازی ابری ارائه شده و برای این منظور ابتدا به مبحث مهاجرت وظایف از یک سرور به سرور دیگر پرداخته شده است. در این تحقیق از الگوریتم ژنتیک در کنار بهینه‌سازی ازدحام ذرات استفاده شده که توازن بار را برقرار نموده و باعث می‌شود منابع موجود در شبکه به کارها تخصیص یابند؛ به طوری که این توزیع کار بین منابع به صورت عادلانه باشد و سرورها به طور هم‌زمان عمل پردازش را بر روی درخواست‌های مختلف کاربران انجام دهند. الگوریتم ارائه‌شده در این تحقیق با استفاده از نرم‌افزار کلودسیم^۱ شبیه‌سازی شده و نتایج به‌دست‌آمده در آن، بیانگر کارآمد بودن این راهکار در ایجاد توازن بار می‌باشد. در مقاله [۹] یک مدل زمان‌بندی و توازن بار بر اساس مفهوم تقسیم‌بندی ابر ارائه شده است. الگوریتمی ترکیبی ارائه شده است که در آن از الگوریتم زمان‌بندی RR^۲ و نظریه بازی‌ها استفاده شده است و سروری انتخاب می‌شود که بار کمتری را متحمل شده باشد و سرعت اجرای آن بیشتر باشد و برش زمانی در RR با توجه به انتخاب سرور بهتر صورت می‌پذیرد. بر اساس

این الگوریتم بیان شده است که توازن بار بهبود یافته است. در [۱۰]، برای اجرای رسیدگی به خواسته‌های مشتریان داده‌های IoT ناهمگون، یک مدل کنفدراسیون ابری را که انتخاب‌های ایدئال برای فراهم آوردن ابر هدف تعیین می‌کند، پیشنهاد کرده است. همچنین یک مدل بهینه چندمنظوره ارائه داده است. برای حل این مدل یک ساختار عمومی برای الگوریتم ژنتیک توسعه داده است. مدل پیشنهادی او از میان تست ارزیابی‌های متنوع آزمایش شده است. در [۱۱]، یک مدل واقعی مبتنی بر برنامه‌نویسی خطی برای شناسایی استراتژی‌ها و تصمیمات در فدراسیون ابری ارائه شده است. بر همین اساس درخواست‌های کاربر در تمام سطح فدراسیون ابری منتشر شده و کاربران به‌سهولت هدایت می‌شوند. در نتیجه سودهای به‌دست‌آمده باعث ایجاد انگیزه مشارکت ارائه‌دهنده‌ها به وسیله درون‌سپاری آزاد منابع برای پشتیبانی دیگر اعضای فدراسیون می‌شود. این مدل بسیار مقیاس‌پذیر است. در [۱۲]، تخصیص منبع چندلایه ارائه شده است. ایده اصلی مبتنی بر تخصیص منابع در یک رویکرد لایه‌ای است که تخصیص منبع را برای کارهای شخصی روی یک گره از کلاستر مبتنی بر سطوح بهره‌برداری منبع تنظیم می‌کند. نتایج نشان داده که راهکاری پیشنهادی آن‌ها عملکرد را تا ۱۸٪ و زمان اجرا را به‌طور متوسط تا ۱۰٪ بهبود می‌بخشد. در [۱۳]، به بررسی مکانیزم‌های پردازش داده با محوریت رایانش ابری پرداخته شده است. در ادامه، راهکارهایی مرور و مورد بحث قرار گرفته که بهره‌گیری و بررسی آن‌ها می‌تواند دید عمیق‌تری در مورد پردازش داده در رایانش ابر را فراهم آورد همچنین به روش‌های تخصیص بهینه منبع در زیرساخت ابری نیز اشاره شده تا به کمک آن بتوان داده‌های مختلف از جمله حوزه سلامت را به‌درستی مدیریت و پردازش کرد. در انتها مهم‌ترین چالش‌های پیش‌روی پردازش داده و توسعه آن در بهداشت، بررسی شده است تا از طریق آن بتوان در توسعه و بهبود کارایی پردازش داده در رایانش ابری گام‌هایی برداشته شود. در [۱۴]، یک راهکار تخصیص منبع به کمک رایانش ابری و اینترنت اشیا به‌منظور پردازش بلادرنگ و دسته‌ای داده‌های حوزه سلامت ارائه شده است. در این مقاله، یکی از مشکلات

1. Cloudsim
2. Round Robin

راهکاری است که علاوه بر ارائه یک رویکرد تخصیص بهینه منبع، بتواند وضعیت سرورها را پس از فرایند تخصیص نیز بررسی و کنترل کند. در نتیجه در این مقاله در مرحله اول به کمک راهکار فازی و تکنیک پیش‌بینی منبع، سرور پردازشی کاندید جهت تخصیص منبع انتخاب می‌شود. سپس در ادامه و به منظور ایجاد توازن در حین پردازش وظایف، از تکنیک مهاجرت ماشین مجازی مبتنی بر انتخابگر فازی برای انتقال ماشین مجازی از یک سرور با بار زیاد به یک سرور کم‌بار استفاده می‌شود.

۳. روش پیشنهادی

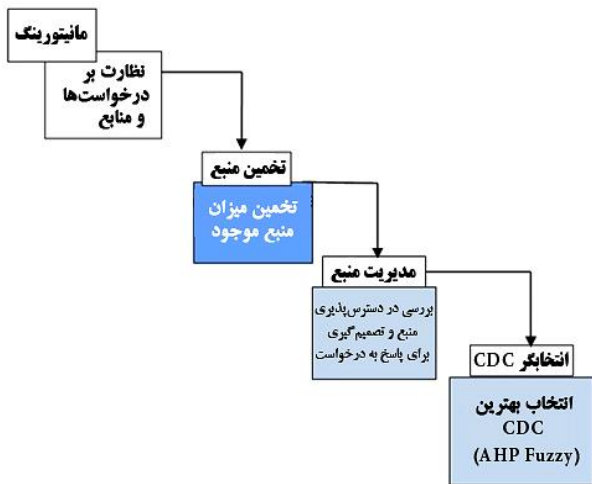
در این بخش، رویکردی مبتنی بر تکنیک AHP^۱ فازی برای تعیین ابر بهینه و همچنین تکنیک پیش‌بینی منابع با استفاده از الگوریتم ARIMA ارائه می‌شود. به‌کارگیری این راهکار با توجه به آنکه سبب بهینه‌سازی فرایند استفاده از منابع می‌شود و همچنین توازن بار بهینه‌ای که فراهم می‌آورد، پردازش سریع‌تر و در نتیجه کاهش زمان اجرا را ممکن می‌سازد. همچنین به‌منظور ایجاد توازن در حین پردازش وظایف از تکنیک مهاجرت ماشین مجازی مبتنی بر انتخابگر فازی برای انتقال ماشین مجازی از یک سرور با بار زیاد به یک سرور کم‌بار استفاده می‌شود. ساختار معماری پیشنهادی شامل دو بخش محلی و سراسری^۲ است که برای تحویل درخواست به قسمت سراسری باید ابتدا قسمت محلی بررسی گردد، در صورت نداشتن شرایط، درخواست به قسمت سراسری تحویل داده می‌شود. در واقع معماری پیشنهادی به‌صورت سلسله‌مراتبی است و درخواست کاربر به قسمت محلی داده می‌شود و توسط آنالیزگر محلی برای پیش‌بینی منابع تحلیل و مانیتور می‌شود. بر همین اساس در ابتدا لیست درخواست‌های رسیده ایجاد می‌شود سپس با استفاده از الگوریتم ARIMA میزان منابع موجود تخمین زده می‌شود. ساختار معماری پیشنهادی این گونه است که فرض می‌شود هر مرکز داده ابری (CDC^۳) یک واسطه‌گر^۴ دارد، که اولاً

اصلی برای پردازش بلادرنگ، شتاب تولید داده‌های مربوط به مراقبت‌های بهداشتی بیان شده است. در نتیجه روزبه‌روز نیاز به منابع پردازشی بیشتری برای این منظور است. بر همین اساس یک راهکار مبتنی بر اینترنت اشیا و رایانش ابری ارائه شده است به این صورت که از اینترنت اشیا به‌عنوان یک کلونی مه که به کاربر نزدیک‌تر است و وظیفه پردازش بلادرنگ داده‌ها حساس را بر عهده دارد، استفاده می‌شود و در صورت نیاز و به‌صورت دوره‌ای پردازش‌های سنگین به رایانش ابری انتقال داده می‌شود. در انتها راهکار پیشنهادی در وب‌سرویس آمازون مورد ارزیابی قرار گرفته و این نتیجه حاصل شده که از طریق به‌کارگیری آن می‌توان زمان پاسخ به درخواست‌ها را کاهش داد. در [۱۵]، خورسند و همکاران رویکردی مبتنی بر تکنیک سلسله‌مراتبی فازی برای تأمین منابع برنامه‌های کاربردی چندلایه در محیط رایانش ابری ارائه دادند. راهکار پیشنهادی آن‌ها بر اساس مدل رایانش خودمختار است که در فاز تحلیل از تکنیک ماشین بردار پشتیبان و رگرسیون خطی برای پیش‌بینی تعداد درخواست‌های کاربران استفاده کردند نتایج شبیه‌سازی آن‌ها تحت دو بار کاری مصنوعی و واقعی نشان می‌دهد که رویکرد پیشنهادی شان از دقت بالایی در پیش‌بینی نسبت به سایر روش‌ها برخوردار بوده و هزینه و زمان پاسخ را در مقایسه با رویکردهای مشابه به‌طور قابل توجهی کاهش می‌دهد.

در [۱۶]، نویسندگان رویکردی مبتنی بر تجزیه و تحلیل بارهای کاری برای تأمین کارای منابع استفاده کردند. رویکرد پیشنهادی آن‌ها از ترکیب الگوریتم رقابت استعماری و خوشه‌بندی K-Means فازی برای دسته‌بندی درخواست‌های کاربران استفاده می‌کند. همچنین، آن‌ها از درخت تصمیم برای تخصیص کارآمد منابع بر اساس دو معیار بهره‌وری و بار کاری استفاده کردند.

به‌طور کلی، همان‌طور که در این بخش بررسی شد، راهکارهای موجود در واقع فقط بر روی مرحله نگاهداشت و تخصیص منبع تمرکز کرده‌اند، درحالی‌که پس از فرایند تخصیص منبع نیز امکان دارد به دلایل مختلفی از جمله نقص سخت‌افزاری یا نرم‌افزاری یک سرور پردازشی دچار ازدحام شود و در نتیجه زمان اجرا بالا رود. بر همین اساس نیاز به

1. Analytic Hierarchy Process
2. Global
3. Cloud Data Center (CDC)
4. Broker



شکل (۱): ساختار واسطه‌گر به همراه اجزای آن

۳.۱. فرموله‌بندی مسئله

در این بخش، متغیرها و فرمول‌های مورد استفاده در روش پیشنهادی بررسی می‌شود. تعداد کاربران به صورت یک ارائه به شکل $\{u_1, u_2, \dots, u_n\}$ نشان داده می‌شود. حد مجاز، برای پاسخ‌گویی به درخواست i ام کاربر i با نام Req_i^A در نظر گرفته می‌شود. تعداد درخواست هر کاربر با $Req-count$ و میزان منابع درخواست با $Res-req$ نشان داده می‌شود. در خواست کاربران برای پاسخ‌گویی در صف $Q-req$ قرار می‌گیرد. در این راهکار از CDC که تعداد آن‌ها با $N-CDC$ نشان داده می‌شود، استفاده می‌گردد و ممکن است CDCهای محلی یعنی LCC جوابگوی نیاز نباشند و نیاز به کنفدراسیون باشد، در این حالت از GCC استفاده می‌شود. آستانه زمان سرویس‌دهی با tts نشان داده می‌شود. برای هر CDC ، هاست‌هایی در نظر گرفته می‌شود که به صورت $Host_j-cdc_i$ نشان داده می‌شود. هر هاست مجموعه‌ای از ماشین‌های مجازی است که به صورت VM_i-host_j تعریف می‌شود. میزان کیفیت خدمات (QoS) با توجه به سود (b) و هزینه (C_k) سنجیده می‌شود. علاوه بر این از فاصله بین CDC ها ($D_{i,j}$) و در دسترس بودن همان نوع ماشین مجازی ($STVM$)^۴ و مجموعه ارائه‌دهندگان ابری استفاده می‌کنیم. در جدول (۱) خلاصه‌ای از پارامترهای مورد استفاده بیان شده است. کیفیت ارائه سرویس (QoS) با توجه به میزان مصرف

درخواست‌های کاربران مراجعه کننده به این CDC را دریافت و تحلیل می‌کند و اگر میزان منابع مورد نیاز برای اجرای درخواست‌های کاربران موجود باشد، آن را اجرا می‌کند و نتیجه را به کاربران آن CDC برمی‌گرداند اما اگر واسطه‌گر نتواند درخواست‌های کاربران را با استفاده از منابع CDC جاری اجرا کند، باید با استفاده از الگوریتم AHP فازی که هر واسطه‌گر به آن دسترسی دارد، اقدام به انتخاب CDC مناسب از بین چندین CDC موجود که با آن‌ها در ارتباط است، می‌کند. بنابراین برای هر CDC یک واسطه‌گر وجود دارد که درون این واسطه‌گر مؤلفه‌های زیر موجود است:

۱. مؤلفه مانیتورینگ منبع^۱: وظیفه مانیتور کردن تعداد درخواست‌های کاربران و همچنین تعداد و ظرفیت میزان منابع موجود در یک CDC را دارد. بر همین اساس در این مؤلفه اطلاعات مربوط به تعداد درخواست‌ها، میزان منابع و همچنین تعداد منابع در دسترس ذخیره می‌شود.
 ۲. مؤلفه تخمین‌کننده منبع^۲: برای تخمین میزان منابع مورد نیاز درخواست‌های کاربران با استفاده از الگوریتم پیش‌بینی $ARIMA$ ، که از جمله روش‌های پیش‌بینی مقادیر مبتنی بر زمان است.
 ۳. مؤلفه انتخابگر مرکز داده ابری^۳ که با استفاده از AHP فازی برای انتخاب یک CDC در صورتی که واسطه‌گر نتواند با منابع همان CDC محلی درخواست‌ها را اجرا کند.
- در شکل (۱)، ساختار واسطه‌گر به همراه اجزای آن نشان داده شده است.

همچنین پس از انتخاب سرور جهت پردازش، از تکنیک مهاجرت ماشین مجازی مبتنی بر انتخابگر فازی استفاده می‌شود تا در صورتی که در حین پردازش و به دلایل نقض سخت‌افزاری یا نرم‌افزاری یک سرور دچار ازدحام گردید، بتوان ماشین مجازی آن را به یک سرور کم‌بار یا بدون بار در دسترس، مهاجرت داد و در نتیجه از ایجاد ازدحام و بالا رفتن زمان اجرا در سرورها جلوگیری به عمل آورد.

1. Request Monitoring
2. Resource Estimator
3. CDC Selector

4. Same Type Virtual Machine

می‌دهد. به دلیل اینکه هر کاربر تنوعی از درخواست را دارد، هر درخواست کاربر در صف قرار می‌گیرد (خط ۲ تا ۴). این عمل برای تمامی کاربران تکرار می‌شود (خط ۱). در ادامه مشخص می‌شود هر درخواست به کدام منبع نیاز دارد (خط ۶ تا ۸).

الگوریتم ۱: افزودن درخواست به صف

```

1: for  $i \leftarrow 1$  to  $n$  do
2:   for  $j \leftarrow 1$  to  $Req\_Count_j$  do
3:     ADD  $Req_j^i$  to  $Request\_queue$ 
4:   end for
5:   while ( $Request\_queue$  is not empty)
6:     foreach  $req$  in  $Request\_queue$  do
7:       Send  $req$  to  $LCC$ 
8:     end for
9:   end while
10: end for
    
```

۳.۳. فاز مانیتور

این فاز مسئول جمع‌آوری درخواست‌های رسیده و منابع در دسترس است. فاز مانیتور شامل دو زیرمجموعه مانیتورینگ درخواست و منابع است. مؤلفه مانیتور منبع، مسئول جمع‌آوری اطلاعات در مورد بهره‌وری منابع، ظرفیت منابع استفاده‌شده، بررسی حداکثر تعداد ماشین مجازی است. مؤلفه مانیتور درخواست، مسئول جمع‌آوری اطلاعات در مورد بار کاری درخواست‌های ارسال‌شده توسط کاربران شامل شناسه درخواست، مهلت زمانی برای اجرای درخواست، میزان هزینه قابل پرداخت توسط کاربر و طول درخواست برحسب میلیون دستورالعمل (MI) است. این اطلاعات مانیتور شده، توسط این دو زیر مؤلفه جمع‌آوری و یکپارچه شده و برای استفاده توسط سایر فازها در پایگاه دانش ذخیره می‌شود. الگوریتم (۲) برای پایش میزان خدمات هر منبع در LCC^1 (منبع محلی) یا GCC^3 (منبع عمومی) استفاده می‌شود. تا زمانی که منابع آزاد وجود داشته باشد (خط ۱) این الگوریتم اجرا می‌شود. این الگوریتم هر دو منبع محلی و عمومی را مورد بررسی قرار می‌دهد، اما اولویت اصلی پایش منابع محلی است. ابتدا میزان خدمات منابع محلی مورد بررسی قرار می‌گیرد؛ در صورتی که منابع محلی وجود نداشته باشد، (خط ۶) به منابع عمومی که

پهنای باند شبکه و میزان ترافیک تولیدشده از هر ماشین فیزیکی، به صورت زیر مدل شده است.

$$D_j = \sum_{\forall m \in V_j} \lambda(j, m) \sum_{i=1}^{\rho(j, m)} C_i$$

که D_j میزان ارتباطات ماشین فیزیکی j ام با دیگر ماشین‌های فیزیکی، $\lambda(j, m)$ بار ترافیکی میان ماشین فیزیکی j ام با ماشین فیزیکی m ام، V_j مجموعه ماشین‌های فیزیکی که با ماشین فیزیکی j ام در ارتباط هستند و C_i وزن لینک ارتباطی میان دو ماشین فیزیکی در سطح i ام و همچنین $\rho(j, m)$ سطح ارتباط میان ماشین‌های فیزیکی m با ماشین فیزیکی j ام است.

جدول (۱): پارامترهای مورد استفاده

کاربرد	نماد
پاسخ‌گویی به درخواست j ام کاربر i ام	Req_j^i
تعداد درخواست هر کاربر	Req-count
میزان منابع درخواستی	Res-req
صف درخواست کاربر	Q-req
تعداد CDC	N-CDC
CDCهای محلی	LCC
CDCهای سراسری	GCC
آستانه زمان سرویس‌دهی	tts
مجموعه ماشین‌های مجازی هر میزبان	VM_i-host_j
میزبان‌های هر CDC	$Host_j-cdc_i$
سودمندی QoS	b
هزینه تابع سودمندی	C_k
فاصله بین CDC، i و j	D_{ij}
در دسترس‌پذیری ماشین مجازی	STVM
تعداد کاربران	user
آستانه تأمین زیاد	Threshold
میزان منابع موجود	Resource

۳.۲. دریافت درخواست‌ها

در الگوریتم شماره (۱) ابتدا درخواست‌ها، در صف درخواست قرار می‌گیرند و به تناسب نیاز، به منبع مورد نظر ارجاع می‌شوند. این الگوریتم نحوه قرارگیری درخواست‌ها را در صف نشان

1. Million Instruction
2. Local Cloud Confederation
3. Global Cloud Confederation

آن خیلی بیشتر از منابع مورد نیاز درخواست نباشد تا از تخصیص زیاد جلوگیری شود.

الگوریتم ۲: پایش منابع

```

1: while (resource != 0) {
2:   for i ← 1 to n do
3:     for j ← 1 to m do
4:       if (cdc is available in our lcc)
5:          $y^i = x_1cdc_i + \dots + x_jcdc_i$ 
6:       else
7:         if (cdc is available in gcc)
8:           for i ← 1 to n do
9:             for j ← 1 to m do
10:               $G^i = x_1cdc_i + \dots + x_jcdc_i$ 
11:            else
12:              resource = 0
13: }
```

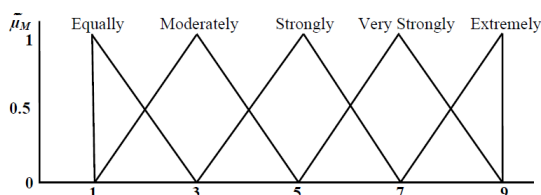
الگوریتم ۳: فاز پیش بینی (ARIMA PrediICA-K-Meansion)

```

1: while req > 0 do
2:   for i ← 1 to 3 do //num of tiers
3:     for j ← 1 to n do
4:       sum = sum + WL(j)
5:     end for
6:      $\mu = \text{sum}/n$ 
7:     for j ← 1 to n do
8:        $MA[i] = \mu + \theta[j] * WL[n - j] + \epsilon$ 
9:     end for
10:    for j ← 1 to n do
11:       $AR[i] = \phi[j] * WL[n - j] + \epsilon$ 
12:    end for
13:    for j ← 1 to n do
14:       $L = \text{sqrt}(\phi[j] * WL[n - j] + \epsilon)$ 
15:    end for
16:    Predict[i] = L +  $\alpha$  + AR[i] + MA[i]
17:  end for
18: return [Predict]
19: end while
```

۳.۶. فاز انتخابگر مرکز داده ابری

در فاز انتخاب مرکز داده ابری، تخصیص بهینه منابع برای هر درخواست وجود دارد. این مؤلفه بهینه‌ترین مرکز داده ابری را با استفاده از الگوریتم AHP فازی انتخاب می‌کند [۱۲]. برای نمایش معیارهای ارزیابی به متغیرهای مثلثی فازی نیاز است، شکل (۲) نمایش زبانی هریک از این متغیرها و بازه عددی آن‌ها را نشان می‌دهد. جدول (۲) نیز نمایش مثلثی و تفسیر زبانی آن‌ها را نشان می‌دهد.



شکل (۲): نمایش زبانی متغیرهای مثلثی و مقادیر آن‌ها

در کنفدراسیون ابر عمومی قرار دارند، مراجعه کرده و برای هر کدام از CDCهای موجود در GCC میزان خدمات را از طریق به کارگیری رابطه موجود در (خط ۱۰) بررسی می‌کند.

۳.۴. فاز پیش بینی

فاز پیش بینی کننده منبع مسئول تخمین میزان منابع موجود برای سرویس دهی به درخواست هاست. اگر منابع تخصیص داده شده نتوانند به صورت کامل تقاضا را برآورده سازند، تأمین کم و اگر منابع تخصیص داده شده بیش از نیازمندی‌های حقیقی باشند تأمین زیاد اتفاق می‌افتد. در اینجا از طریق بهره‌گیری از الگوریتم ARIMA می‌توان برآورد دقیقی از شناسایی نیازهای درخواست منبع را به دست آورد. در الگوریتم (۳) شبه کد ARIMA نشان داده شده است که برای پیش بینی منابع مورد نیاز هر درخواست استفاده می‌شود. در ابتدا درخواست‌های کاربران به صورت لایه‌ای بین سه لایه برنامه کاربردی ابری شامل لایه وب، لایه برنامه کاربردی و لایه پایگاه داده توزیع می‌شود (خطوط ۱ تا ۵). از آنجایی که مدل پیش بینی ARIMA، ترکیبی از مدل‌های پیش بینی خودبرگشتی (AR) و میانگین متحرک (MA) است، ابتدا میزان پیش بینی حاصل از تکنیک میانگین متحرک محاسبه شده (خطوط ۷ تا ۹) و سپس، پیش بینی حاصل از تکنیک خودبرگشتی محاسبه می‌شود (خطوط ۱۰ تا ۱۲). در نهایت، در ادامه مقدار نهایی پیش بینی نهایی بار کاری با استفاده از تکنیک ARIMA به دست می‌آید (خطوط ۱۳ تا ۱۷).

۳.۵. فاز مدیریت منبع

بعد از اندازه‌گیری منبع برای یک درخواست خاص، پیش بینی کننده منبع این درخواست را به مدیر منبع می‌فرستد. مدیر منبع در دسترس بودن منابع را بررسی می‌کند و در مورد سرویس دهی به درخواست جدید تصمیم‌گیری می‌کند. اگر هیچ منبع در دسترس در CDC میزبان وجود نداشته باشد، سعی می‌شود یک CDC متحد شده بهینه با استفاده از AHP پیدا گردد. در صورت موجود بودن منبع باید بررسی شود که اندازه

1. Auto-Regressive
2. Moving Average

جدول (۲): اعداد فازی مثالی

تفسیر زبانی	عدد فازی	تابع عضویت	بازه	تفسیر زبانی
Just Equal	$\tilde{1}$			(1,1,1)
Equally important		$\mu_M(x) = (3-x)/(3-1)$	$1 \leq x \leq 3$	(1,1,3)
Weakly important	$\tilde{3}$	$\mu_M(x) = (x-1)/(3-1)$	$1 \leq x \leq 3$	(1,3,5)
		$\mu_M(x) = (5-x)/(5-3)$	$3 \leq x \leq 5$	
Essential or Strongly important	$\tilde{5}$	$\mu_M(x) = (x-3)/(5-3)$	$3 \leq x \leq 5$	(3,5,7)
		$\mu_M(x) = (7-x)/(7-5)$	$5 \leq x \leq 7$	
Very strongly important	$\tilde{7}$	$\mu_M(x) = (x-5)/(7-5)$	$5 \leq x \leq 7$	(5,7,9)
		$\mu_M(x) = (9-x)/(9-7)$	$7 \leq x \leq 9$	
Extremely preferred	$\tilde{9}$	$\mu_M(x) = (x-7)/(9-7)$	$7 \leq x \leq 9$	(7,9,9)

$$W' = (d'(A_1), d'(A_2), \dots, d'(A_n))^T \quad (1)$$

در ادامه بردار W' را نرمال‌سازی می‌شود و نتیجه آن به صورت زیر است.

$$W = (d(A_1), d(A_2), \dots, d(A_n))^T \quad (2)$$

الگوریتم ۲: الگوریتم AHP فازی

- 1: Get monitoring info
- 2: for level 1 and 2 do
- 2-1: Load fuzzy rating of the relative importance of each pair of criteria (Table 1)
- 2-2: for criterion C1 to Cn (i in 1 to n) do
- 2-2-1: Calculate fuzzy synthetic extent as S_i (Eq. 1)
- 2-3: for i, j in 1, ..., n do
- 2-3-1: Calculate $V (S_i \geq S_j)$
- 2-4: Calculate un-normalized wight using Eq. 7-3
- 2-5: Convert un-normalized wight to normalized wight using (Eq. 2)
- 3: return final wight as priority of CDC

۳.۶. متوازن‌کننده فازی

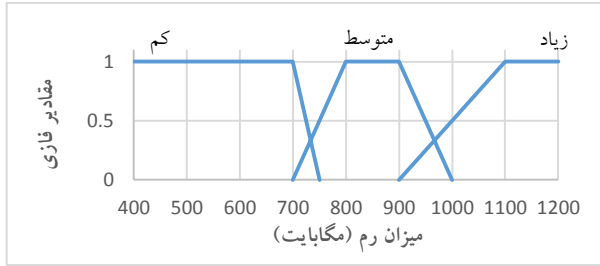
با وجود اقداماتی که به منظور انتخاب بهینه‌ترین منبع در دسترس انجام می‌گیرد امکان دارد در طی فرایند پردازش داده‌های حوزه سلامت، تعدادی از سرورها در اجرای دستورات ناموفق باشند و یا به دلایلی مانند تأخیر در پردازش دستورات قبلی و یا خطاهای سخت‌افزاری و نرم‌افزاری، دچار ازدحام^۱ و بار اضافی شوند در حالی که تعدادی از سرورها با اتمام دستورات، بیکار و یا کم‌بار باشند. در این شرایط نیاز به انتقال ماشین‌های مجازی از یک گره به گره دیگر است که برای این منظور از یک انتخابگر فازی به منظور تعیین و انتخاب

هدف اصلی، انتخاب CDC بهینه مبتنی بر AHP فازی از بین مجموعه CDCهای کنفدراسیون است. کنفدراسیون مربوطه بر اساس پارامترهای هزینه محاسباتی (C1)، فاصله CDC درخواست‌شده از CDC کنفدراسیون (C2)، انحراف از توافقات (C3)، منابع موجود (C4) سازمان‌دهی شده است.

الگوریتم (۴) روش پیشنهادی برای انتخاب CDC را نشان می‌دهد. این الگوریتم ساختاری سلسله‌مراتبی با سه سطح دارد. سطح صفر که دارای بالاترین سطح است و سطح اول و سطح دوم. خط دوم الگوریتم نشان می‌دهد که الگوریتم بر روی دو سطح اجرا می‌شود. در هر سطح، تعداد معیارها n در نظر گرفته شده است. در ادامه، اجرای الگوریتم برای سطح اول توضیح داده شده است. در سطح دوم نیز همین عملیات‌ها تکرار می‌شود. بر همین اساس ابتدا مقادیر ماتریس مقایسه فازی خوانده می‌شوند سپس در گام (خط ۲-۳ از الگوریتم) بعد همه زوج S_i ها با یکدیگر مقایسه می‌شود. در ادامه برای محاسبه اولویت هریک از معیارها از رابطه (۱) استفاده می‌شود. برای هر معیار یک وزن غیرنرمال به دست می‌آید. با استفاده از رابطه (۲) وزن‌ها نرمال‌سازی می‌شود. همین عملیات برای سطح دوم که CDCها قرار دارند نیز تکرار می‌شود. وزن به دست آمده اولویت هریک از CDCها را نشان می‌دهد و بر اساس آن می‌توان CDC با بالاترین اولویت را انتخاب کرد.

برای $i=k! : k=1,2,\dots,n$ بردار وزن به صورت زیر محاسبه

می‌شود.

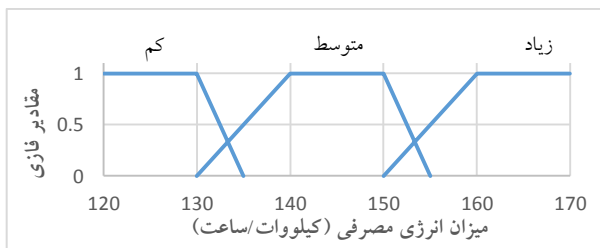


شکل (۳): توابع عضویت RAM

در ادامه و در جدول (۴) و شکل (۴) پارامترهای تابع عضویت مربوط به انرژی مصرفی نشان داده شده است:

جدول (۴): پارامترهای توابع عضویت انرژی مصرفی

پارامتر	میزان انرژی مصرفی
$a=0, b=0, c=120, d=135$	کم
$a=130, b=140, c=150, d=155$	متوسط
$a=150, b=160, c=170, d=170$	زیاد

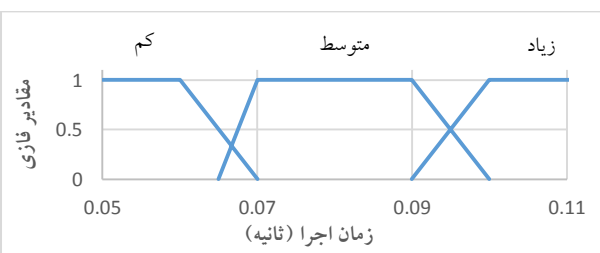


شکل (۴): توابع عضویت انرژی مصرفی

در انتها نیز در جدول (۵) و شکل (۵) پارامترهای تابع عضویت مربوط به زمان اجرا نشان داده شده است:

جدول (۵): پارامترهای توابع عضویت زمان اجرا

پارامتر	میزان زمان اجرا
$a=0, b=0, c=0.06, d=0.07$	کم
$a=0.065, b=0.07, c=0.09, d=0.1$	متوسط
$a=0.09, b=0.1, c=0.11, d=0.15$	زیاد



شکل (۵): توابع عضویت زمان اجرا

ماشین‌های مجازی برای مهاجرت استفاده می‌شود. بر همین اساس متغیرهای زبانی و توابع عضویت به شرح زیر تعریف می‌شوند.

۳.۶.۱. متغیرهای زبانی^۱

متغیرهای زبانی، متغیرهای ورودی و خروجی به شکل کلمات و یا جمله‌های ساده هستند. در این راهکار، پیشنهادی از سه متغیر (کم، متوسط و زیاد) برای بیان وضعیت (حالت) پارامتر مورد نظر استفاده می‌شود.

۳.۶.۲. توابع عضویت

بر اساس متغیرهای زبانی تعریف‌شده، در این بخش توابع عضویت بیان می‌شود. در واقع با استفاده از سه متغیر ورودی بیان‌شده یک انتخابگر فازی ماشین مجازی ارائه شده است. بر همین اساس، توابع عضویت مربوط به متغیرهای زبانی برای پارامترهای اصلی مربوط به منابع، به صورت زیر تعریف می‌شود:

- میزان حافظه در دسترس (T_{RAM}): (کم، متوسط، زیاد)
- زمان اجرا (T_{exe}): (کم، متوسط، زیاد)
- میزان مصرف انرژی (T_{energy}): (کم، متوسط، زیاد)

حال بر اساس تعریف سه پارامتر فوق، ماشین‌های مجازی در حالت‌های کم، متوسط، زیاد، خیلی زیاد دسته‌بندی می‌شود و ماشین‌هایی که در حالت بار خیلی زیاد و زیاد قرار داشته باشند، کاندید برای مهاجرت به ماشین‌های با بار کم و متوسط می‌باشند. در ادامه جداول و اشکال مربوط به پارامترهای توابع عضویت نشان داده شده است. ابتدا در جدول (۳) و شکل (۳) توابع عضویت مربوط به میزان RAM مصرفی برای تعیین حالت‌های کم، متوسط و زیاد نشان داده شده است.

جدول (۳): پارامترهای توابع عضویت RAM

پارامتر	میزان RAM
$a=0, b=0, c=650, d=750$	کم
$a=700, b=800, c=900, d=1000$	متوسط
$a=900, b=1100, c=1800, d=1800$	زیاد

۳.۶.۳. قوانین فازی

در این بخش، به منظور انتخاب ماشین‌های مجازی برای مهاجرت، قوانین فازی با توجه به توابع عضویت تشریح شده در مرحله قبل ارائه می‌شود. از این قوانین به منظور انتخاب ماشین‌های مجازی که به دلایل بار اضافی دچار ازدحام شده‌اند، مورد استفاده قرار می‌گیرد. قوانین مربوط در جدول (۶) نشان داده شده است.

۴. ارزیابی کارایی

در این بخش به بررسی شبیه‌سازی راهکار پیشنهادی پرداخته می‌شود. ابتدا تنظیمات شبیه‌سازی را شرح داده و سپس پارامترهای مورد ارزیابی معرفی خواهند شد. در نهایت، نتایج به دست آمده را مورد بحث و تحلیل قرار خواهیم داد.

جدول (۶): قوانین فازی برای انتخاب ماشین مجازی

Execution	RAM	Energy	VM Selection
High	High	High	Very High
High	High	Medium	Very High
High	High	Low	Very High
High	Medium	High	Very High
High	Medium	Medium	High
High	Medium	Low	High
High	Low	High	High
High	Low	Medium	Medium
High	Low	Low	Medium
Medium	High	High	Very High
Medium	High	Medium	High
Medium	High	Low	Medium
Medium	Medium	High	High
Medium	Medium	Medium	Medium
Medium	Medium	Low	Low
Medium	Low	High	Medium
Medium	Low	Medium	Low
Medium	Low	Low	Low
Low	High	High	High
Low	High	Medium	Medium
Low	High	Low	Medium
Low	Medium	High	Medium
Low	Medium	Medium	Low
Low	Medium	Low	Low
Low	Low	High	Low
Low	Low	Medium	Low
Low	Low	Low	Low

۴.۱. تنظیمات شبیه‌سازی

برای شبیه‌سازی و ارزیابی راهکار پیشنهادی، ابزار Cloudsim مورد استفاده قرار گرفته است [۱۸-۱۹]. مقیاس‌پذیری این شبیه‌ساز بالاست و می‌تواند ابرهای بسیار بزرگی را شبیه‌سازی کند. در حال حاضر، کلودسیم علاوه بر شبیه‌سازی یک ابر قادر به شبیه‌سازی محیط‌های ابری تشکیل شده از چندین ابر دیگر، نیز می‌باشد. در ادامه این بخش تنظیماتی که در پیکربندی شبیه‌سازی مورد استفاده قرار گرفته است، بیان خواهد شد. مجموعه داده‌های استفاده شده در این پژوهش داده‌های مربوط به مصارف دارویی مربوط به ایالت‌های مختلف آمریکا در سال‌های ۲۰۱۵ تا ۲۰۱۸ است. این مجموعه داده اطلاعات دارویی مربوط به ۶۸۰۰ دارو را در خود دارد. همچنین مصارف دارویی در ایالت‌های مختلف آمریکا و سایر اطلاعات مربوط به مصرف و برنامه‌های مصرفی هر ایالت را در خود دارد. جدول (۸) اطلاعات مربوط به این مجموعه داده‌ها را نشان می‌دهد. این مجموعه داده به صورت آنلاین در اختیار محققان قرار گرفته است [۲۰].

جدول (۸) اطلاعات مجموعه داده‌ها را نشان می‌دهد. این مجموعه داده، رکوردهای مربوط به برنامه‌های درمانی در هر منطقه جغرافیایی، سرویس‌های انجام شده، رکوردهایی که مصرف داروها را نشان می‌دهند و سایر اطلاعات را نشان می‌دهد. رکوردهایی مربوط به سلامت بیماران نیز در این مجموعه داده ثبت شده است. همان‌طور که در جدول (۸) دیده می‌شود این داده‌ها از نواحی جغرافیایی گسترده‌ای در آمریکا جمع‌آوری شده‌اند. لذا مدیریت منابع در رایانش ابری تأثیر زیادی در کارایی درخواست‌های رسیده دارد. در جدول، تعداد کل درخواست‌ها را می‌بینیم، ۶۷۳۰۵ درخواست ثبت شده است. با استفاده از بازه زمانی ۵ دقیقه‌ای این درخواست‌ها را گسسته‌سازی می‌کنیم. سپس درخواست‌ها را در بازه گسسته و بر اساس توزیع نرمال با میانگین ۰ و واریانس ۱ نمونه‌برداری می‌کنیم. همین درخواست‌ها را نیز با توزیع پواسون نیز نمونه‌برداری می‌کنیم. توزیع نرمال، توزیع غالب در اغلب

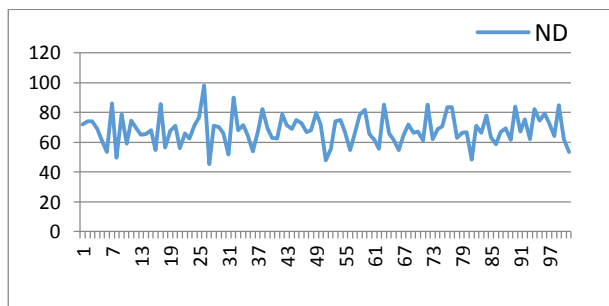
1. https://downloads.cms.gov/medicare/2018Med2000_flatfiles.zip

رویدادهای طبیعی است. از طرفی نیز توزیع پواسون نیز توزیع غالب برای مراجعه مشتری به مراکز خدماتی است.

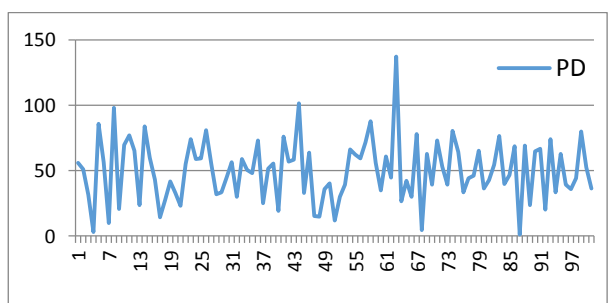
جدول ۸: ویژگی‌های مجموعه داده حوزه سلامت [۲۰]

ردیف	نام	تعداد
۱	نواحی جغرافیایی	۵۷
۲	خدمات درمانی	۱۶۵۵۷
۳	مصرف دارو در سرویس	۵۷۹۰
۴	دانش دارویی	۶۸۳۵۳
۵	درخواست‌ها	۶۷۳۰۵
۶	گزارش‌ها	۵۷

شکل (۶) درخواست‌های نمونه برداری شده بر اساس توزیع نرمال را نشان می‌دهد و شکل (۷) درخواست‌ها بر اساس توزیع پواسون را نشان می‌دهد.



شکل (۶): مجموعه داده‌های ND



شکل (۷): مجموعه داده PD

به منظور شبیه‌سازی راهکار پیشنهادی، یک برنامه سه لایه ابری در نظر گرفته‌ایم، که هر لایه دارای ویژگی‌های متفاوتی نسبت به سایر لایه‌هاست. برای مثال هر لایه نوع ماشین مجازی مجزا و نرخ ورود درخواست‌های متفاوت را داراست و مجموعه داده حوزه سلامت اشاره شده با استفاده از پلت فرم RUBIS بین این سه لایه توزیع می‌شوند. جدول (۷) مشخصات کلی لایه‌های مختلف معماری سه لایه‌ای ابری را

نشان می‌دهد. نرخ ورود درخواست‌ها در لایه‌های مختلف، متفاوت است. به طور دقیق‌تر تعداد درخواست‌های وارد شده در لایه اول، مطابق آنچه از بار کاری خوانده می‌شود، می‌باشد اما تعداد درخواست‌های لایه دوم و سوم به ترتیب به طور میانگین برابر با ۷۰ و ۴۰ درصد درخواست‌های لایه اول می‌باشد.

همان‌گونه که دیده می‌شود، ماشین مجازی بزرگ به لایه اول (وب) که بیشترین حجم کار را دارد، تخصیص داده شده است، درحالی‌که ماشین مجازی با اندازه متوسط به لایه دوم و در نهایت کوچک‌ترین و کم‌قدرت‌ترین نوع ماشین مجازی، به لایه آخر (پایگاه داده) که کمترین حجم درخواست‌ها را دریافت خواهد کرد، اختصاص یافته است. هرچند تعداد ماشین‌های مجازی اولیه هر سه لایه یکسان است، به دلیل تفاوت ظرفیت، قدرت و ماهیت ماشین‌های مجازی به کار گرفته شده در لایه‌های مختلف، قدرت محاسباتی لایه‌ها از همان ابتدا متفاوت خواهد بود. همچنین، مشخصات تمامی مراکز داده مورد استفاده در شبیه‌سازی در جدول (۹) مشخص شده است.

جدول (۷): مشخصات لایه‌های مختلف برنامه کاربردی ابری

لایه	متوسط نرخ ورود	نوع ماشین	تعداد ماشین‌های مجازی اولیه
Web	۱۰۰٪	Large	۵
Application	۷۰٪	Mediun	۵
Database	۴۰٪	Small	۵

جدول (۹): ساختار مراکز داده مورد استفاده در شبیه‌سازی

ساختار	مقدار
معماری	X64
سیستم عامل	Cloud Linux
مدیریت ماشین‌های مجازی	XEN
مصرف انرژی ماشین مجازی	50W = En_Stand آماده‌باش
	120W = En_Comp محاسبه

در نهایت، در شبیه‌سازی برای هر مرکز داده، ۲۷ میزبان فیزیکی تعریف شده که به صورت تصادفی یکی از هفت نوع ماشین زیر در آن‌ها جای می‌گیرد. ماشین‌های مجازی استفاده شده در شبیه‌سازی روش پیشنهادی، در سه گروه و از هر گروه، سه نمونه انتخاب شده است. جدول (۱۰)، مشخصات ماشین‌های مجازی را در هر میزبان نشان می‌دهد

(تمام آزمایش‌ها به زبان جاوا پیاده‌سازی شده و کد شبیه‌سازی از لینک <https://github.com/mostafa13651365/FAHP-ARIMA> قابل دانلود است).

پردازنده با رابطه (۴) محاسبه می‌شود.

$$U_i(\Delta t) = \frac{\sum_{i=1}^{Num_i(\Delta t)} VMi Requested MIPS}{\sum_{i=1}^{Num_i(\Delta t)} VMi Total MIPS} \quad (4)$$

▪ **هزینه (Total Cost):** هزینه، حاصل جمع هزینه ماشین‌های مجازی و هزینه پنالتی است که هزینه پنالتی، میزان کل هزینه جریمه برای تمام درخواست‌های مشتریان است و هزینه کلی ماشین‌های مجازی برای راه‌اندازی و اجرای همه ماشین‌های مجازی است. هزینه کل با رابطه (۵) محاسبه می‌شود.

$$Total Cost = VM Cost + Penalty Cost \quad (5)$$

دو معیار مهم دیگر که در ارزیابی خاصیت کشسانی رایانش ابری بسیار تأثیرگذار هستند، تعداد درخواست‌های رد شده و میزان مقیاس‌پذیری می‌باشند. مطلوب است الگوریتم پیشنهادی دارای خاصیت کشسانی مناسبی در مقیاس‌پذیری باشد تا بتواند تعداد کمتری از درخواست‌های رسیده را رد کند. توجه داشته باشید که معیارهای ذکر شده با هم در ارتباط هستند و بالا و پایین شدن مقادیر هریک بر دیگر معیارها تأثیرگذار خواهد بود. برای مثال اگر تعداد ماشین‌های مجازی زیاد شود، بهره‌وری کم می‌شود، هزینه بالا می‌رود و سود پایین می‌آید. اگر تعداد ماشین‌های مجازی کم شود، زمان پاسخ بالا می‌رود و ممکن است تخطی از توافقات سطح سرویس رخ دهد، که این امر می‌تواند باعث بالا رفتن هزینه شود.

۲.۴. معیارهای ارزیابی

معیارهای مورد ارزیابی در شبیه‌سازی به شرح زیر است:

▪ **تعداد ماشین‌های مجازی (VM allocated):** باید در نظر داشته باشیم که اگر تعداد ماشین‌های مجازی کمتر از میزان مورد نیاز باشد، این امر باعث بالا رفتن زمان پاسخ‌گویی به درخواست‌ها و تخطی از توافقات سطح سرویس می‌شود. در این موقعیت، موظف به پرداخت جریمه خواهیم بود و میزان سودمان کم خواهد شد. همچنین اگر تعداد ماشین‌های مجازی بیشتر از میزان مورد نیاز باشد، عملاً باید هزینه همه منابع را پردازیم ولی استفاده‌ای از همه آن‌ها نمی‌کنیم و این یک هزینه غیرضروری محسوب می‌شود. تعداد ماشین‌های مجازی با رابطه (۳) محاسبه می‌گردد.

$$VM = \{VM_1, VM_2, VM_3, \dots, VM_N\} \quad (3)$$

▪ **بهره‌وری (Utilization):** درصدی از زمان را که یک منبع در حال سرویس‌دهی به درخواست‌هاست، بهره‌وری می‌نامیم. بهره‌وری پردازنده نباید خیلی بالا یا خیلی پایین باشد. هرچه بهره‌وری بیشتر باشد، یعنی کار ماشین مجازی هم بیشتر است (یعنی نزدیک ۱۰۰٪). میزان بهره‌وری

جدول (۱۰): مشخصات ماشین‌های مجازی مراکز داده

Machin Name	MIPS	RAM(GB)	Storage	BW	Price (\$ per Hour)
t2.small	10200	2	1 GB - 16 TB	100Mbps	0.023
m3.Medium	12000	3.75	1×4 GB	1Gbps	0.070
m4.4Xlarge	15000	64	1 GB - 16 TB	1Gbps	0.862
r3.4Xlarge	80000	122	1×320 GB	10Gbps	1.330
m4.10Xlarge	97000	160	1 GB - 16 TB	10Gbps	2.155
d2.4Xlarge	105000	122	12×2000 GB(24 TB)	10Gbps	2.76
m4.16Xlarge	280000	256	1 GB - 16 TB	100Gbps	3.447
r4.16Xlarge	350000	488	1 GB - 16 TB	100Gbps	4.256
d2.8Xlarge	500000	244	24×2000 GB(48 TB)	100Gbps	5.52

۳.۴. تحلیل نتایج

همان‌طور که در بخش پیشین ذکر شد، معیارهای مختلفی برای ارزیابی کارایی روش پیشنهادی بر روی داده‌های حوزه سلامت ارائه گردید. در ادامه، بر اساس معیارهای ارائه‌شده سناریوهای مختلفی برای ارزیابی عملکرد روش پیشنهادی در سه لایه مختلف ارائه می‌دهیم. همچنین، برای ارزیابی مؤثر بودن راهکار پیشنهادی را با روش‌های FAHP [۱۶] و ICA-K-Means [۱۷] استفاده کرده‌ایم. در ادامه اشاره مختصری به روش‌های مورد مقایسه داریم:

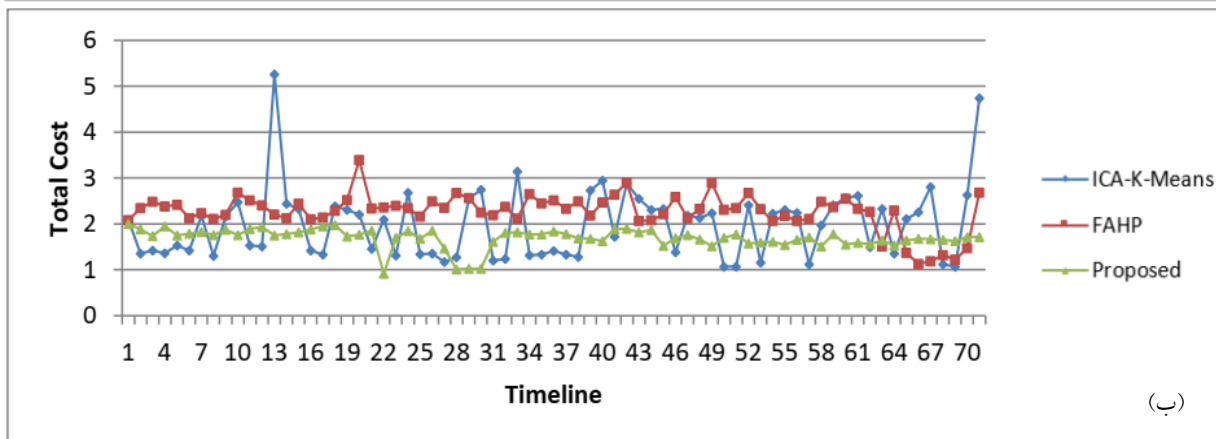
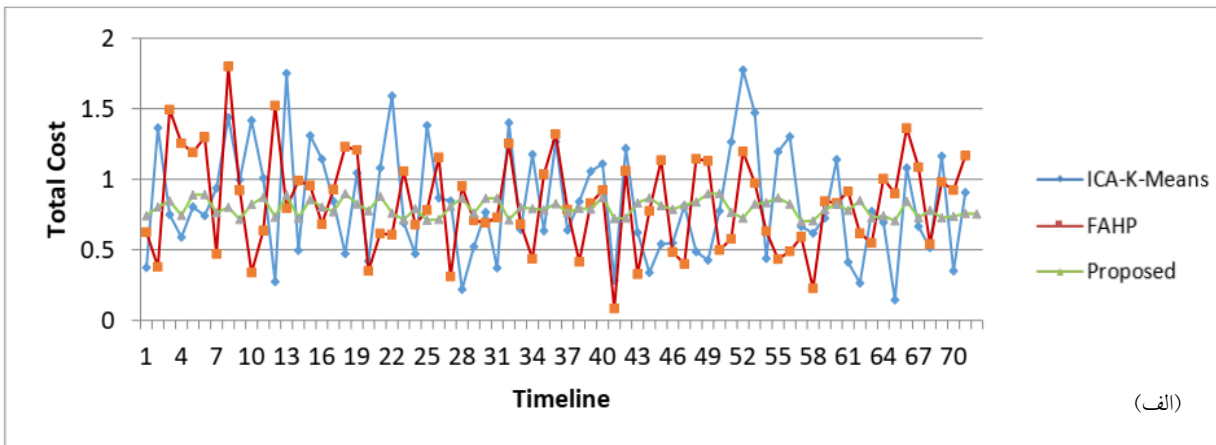
- FAHP: در این راهکار، از تکنیک ماشین بردار پشتیبان و رگرسیون خطی برای پیش‌بینی بار کاری استفاده می‌شود و از فازی سلسله‌مراتبی برای تصمیم‌گیری در مورد میزان افزایش یا کاهش منابع استفاده می‌شود.
- ICA-K-Means: در این راهکار، از ترکیب الگوریتم رقابت استعماری و خوشه‌بندی k-Means فازی برای

دسته‌بندی درخواست‌های کاربران و از درخت تصمیم برای تخصیص کارآمد منابع بر اساس دو معیار بهره‌وری و بار کاری استفاده می‌شود.

دلیل انتخاب این دو روش برای مقایسه این است که اولاً هر دو روش پیش‌کنشی هستند و از قبل پیش‌بینی بار کاری را برای تأمین منابع انجام می‌دهند؛ ثانیاً هر دو روش بر اساس ساختار چند لایه برنامه‌های کاربردی ابری هستند.

۳.۴.۱. ارزیابی هزینه

در این بخش، مجموعه داده‌های سلامت را بر روی دو گروه داده‌های PD و ND انجام می‌دهیم. مجموعه داده‌های سلامت در دو گروه داده‌های PD و ND، هریک ویژگی‌های توزیعی متفاوت دارند. لذا هریک از پارامترها را برای این سه مجموعه داده بیان می‌کنیم. شکل (۸) تا (۹) پارامتر هزینه را برای روش‌های مورد مقایسه در سه لایه وب، برنامه کاربردی و پایگاه داده نمایش می‌دهند.



شکل (۸): مقایسه هزینه کلی روش‌های پیشنهادی در لایه برنامه کاربردی: (الف) مجموعه داده ND، (ب) مجموعه داده PD

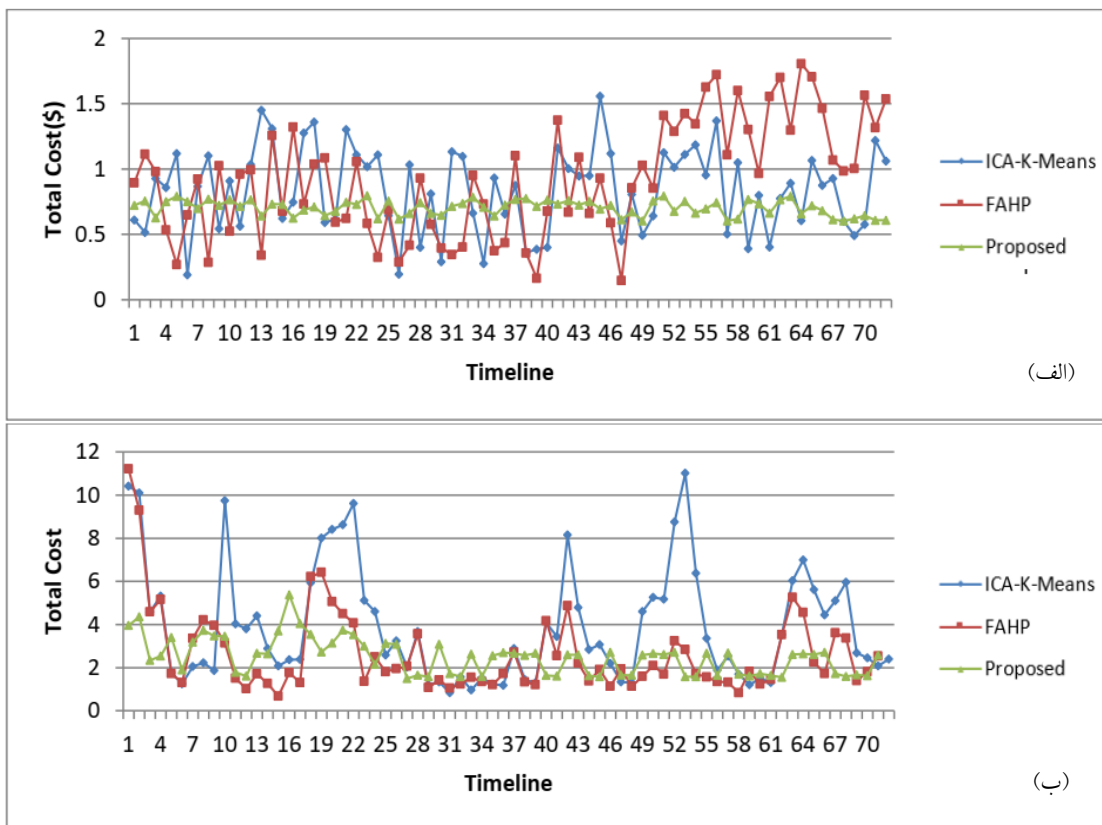
ابری، روش‌های مورد مقایسه رفتارهای متفاوتی با معیار هزینه دارند. هر اندازه ورود بار کاری دارای توزیع نامتوازنی باشد، پیچیدگی‌های تحلیل آن و تصمیم‌گیری توسط چارچوب‌های مورد بررسی نیز بیشتر می‌شود. لذا دیده می‌شود در اغلب موارد و به صورت متوسط روش پیشنهادی در این پژوهش کارایی بهتری در لایه برنامه کاربردی دارد.

شکل (۱۰) معیار هزینه در لایه پایگاه داده را برای روش‌های پیشنهادی بررسی می‌کند. این لایه با تعداد زیادی از تراکنش‌های موازی مواجه است، لذا استفاده از روش‌های پویا که قابلیت کشسانی را برای این لایه مهیا کنند از لزوماتی است که تأمین‌کننده خدمات باید به آن اهمیت فراوان دهد.

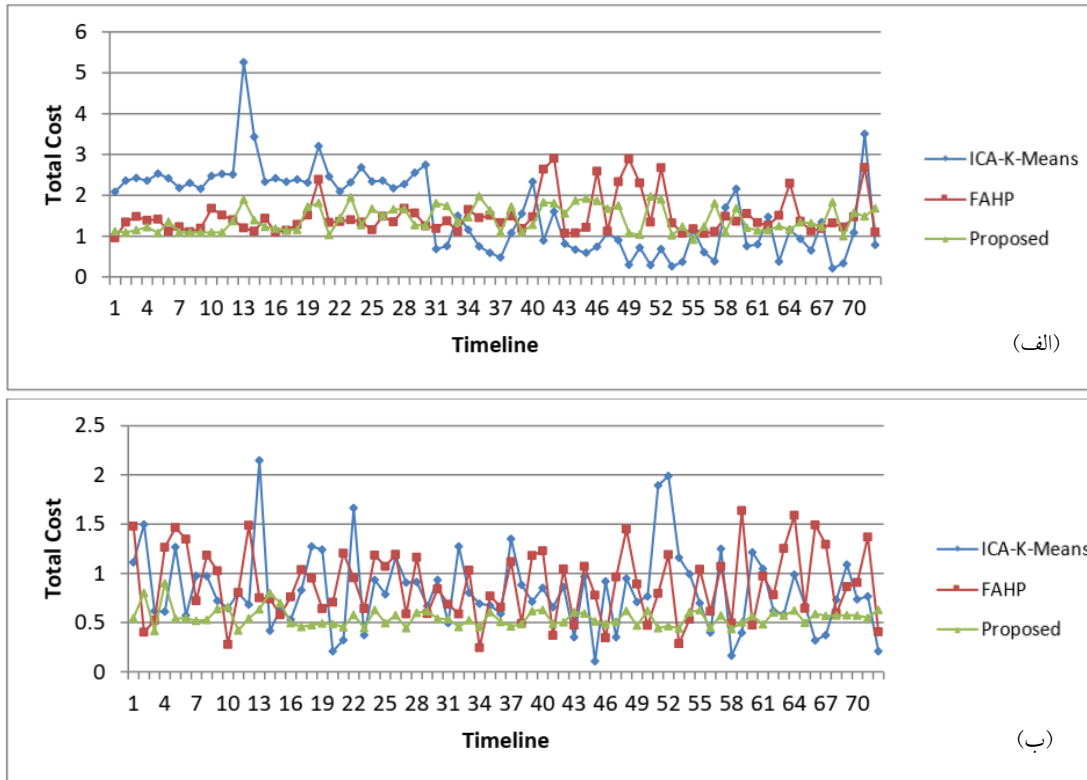
همان‌طور که در شکل (۱۰) مشاهده می‌کنید، رویکرد پیشنهادی در این پژوهش، از توابع تصمیم‌گیری استفاده می‌کند که تلاش دارند تعادل بار بین منابع را با استفاده از خاصیت کشسانی رایانش ابری تأمین کند. به همین دلیل نیز ثبات بیشتری در زمینه هزینه تأمین منابع دیده می‌شود.

شکل (۸ الف)، هزینه برای مجموعه داده ND را نشان می‌دهد. همان‌طور که دیده می‌شود، روش پیشنهادی قادر است در لایه برنامه کاربردی، به صورت یکنواختی خاصیت کشسانی رایانش ابری را تأمین کند. بخش (ب) از شکل (۸) نیز نتایج را برای مجموعه داده PD نمایش می‌دهد. مجموعه داده PD دارای واریانس بیشتری است. دیده می‌شود به دلیل رفتار بار کاری در این توزیع داده‌ای، اختلاف بین روش‌های مورد مقایسه بیشتر است. ضمن اینکه روش پیشنهادی در این پژوهش هزینه کل را بیشتر از سایر روش‌ها کاهش می‌دهد. به این نکته توجه داشته باشید که در مورد داده‌های سلامت، با توجه به حجم بسیار زیاد داده‌ها و اهمیت موضوع، کمیت هزینه بسیار حائز اهمیت است.

شکل (۹) هزینه در لایه وب و شکل (۱۰) نیز هزینه در لایه پایگاه داده را نشان می‌دهد. با توجه به سه لایه بودن معماری پیشنهادی برای داده‌های حجیم در حوزه سلامت، بررسی معیارهای ارزیابی در هر یک از سه لایه ضروری است. شکل (۹) نشان می‌دهد بر اساس توزیع بار کاری وارد شده به رایانش



شکل (۹): مقایسه هزینه کلی روش‌های پیشنهادی در لایه وب: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۰): مقایسه هزینه کلی روش‌های پیشنهادی در لایه پایگاه داده: (الف) مجموعه داده ND، (ب) مجموعه داده PD

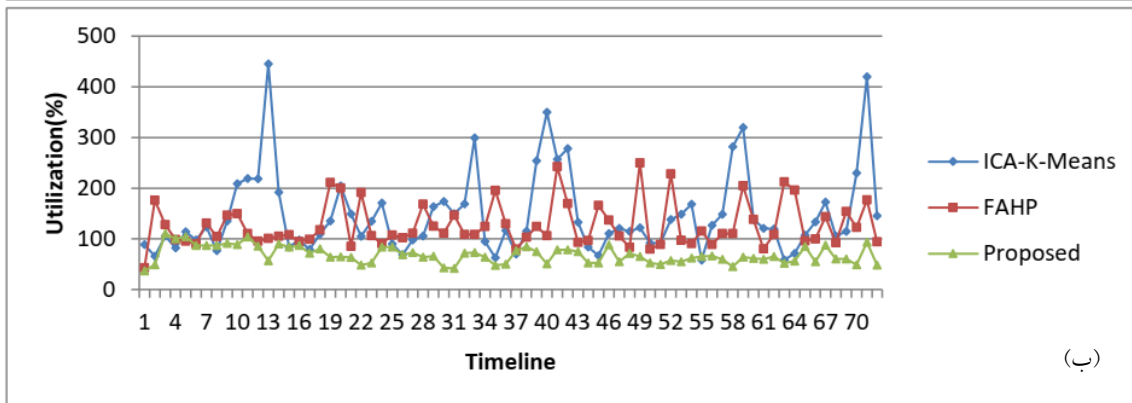
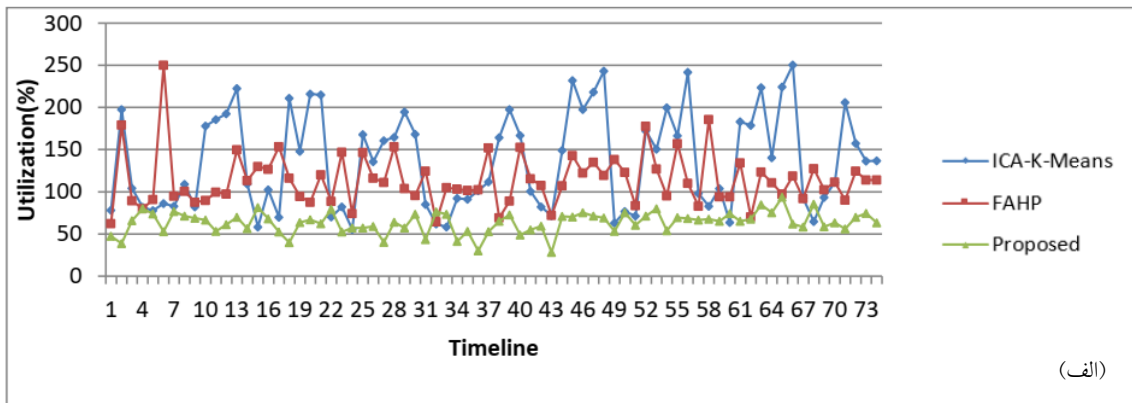
PD نسبت به ND می‌باشد.

۴.۳.۲. ارزیابی بهره‌وری منابع

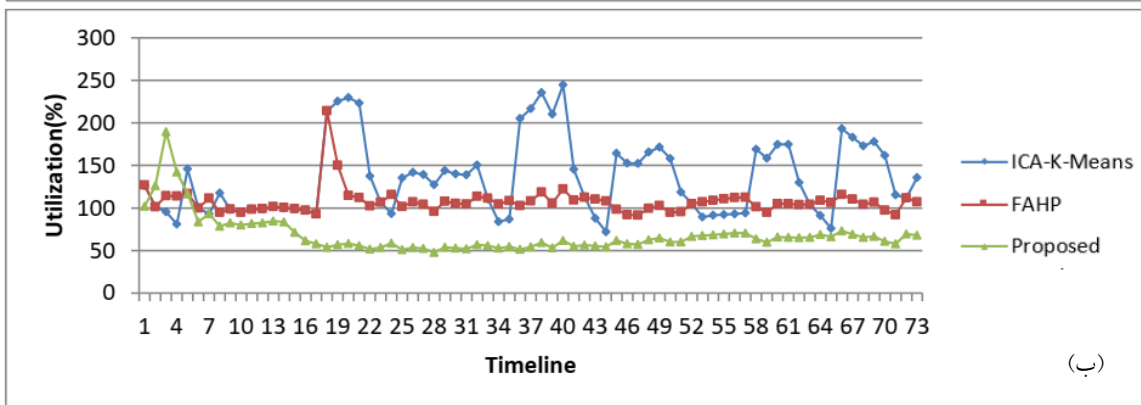
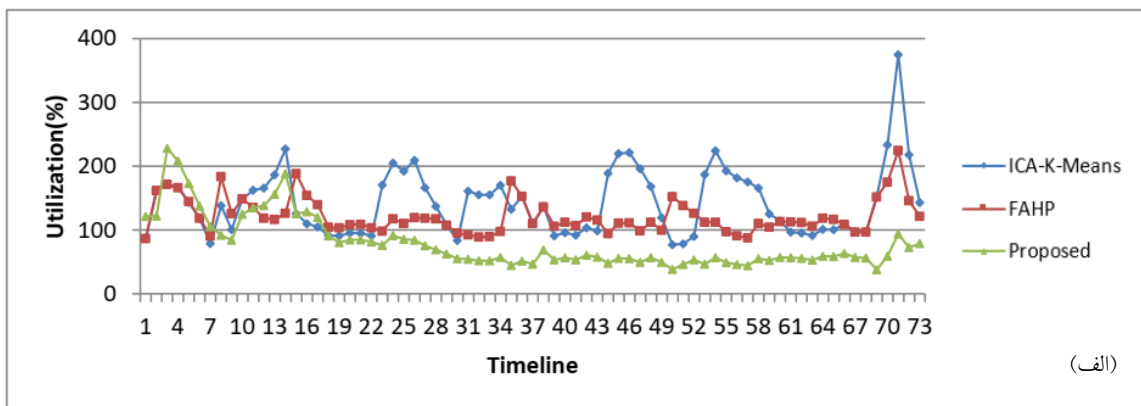
شکل‌های (۱۱)، (۱۲) و (۱۳) میزان بهره‌وری منابع را در سه لایه مختلف کاربرد، وب و پایگاه داده نشان می‌دهد. با بررسی شکل‌ها می‌توان دید روش ICA-K-Means از روش‌های مورد مقایسه، در هر سه لایه عملکرد بهتری دارد، به این دلیل که در این دسته روش‌ها از قواعد کنترلی استفاده می‌شود؛ به عبارتی دانش پیشین کمتر از سایر روش‌ها استفاده می‌شود. روش FAHP دارای ثبات بیشتری است چون از دانش پیشین استفاده می‌کند. البته روش پیشنهادی در این پژوهش با استفاده از دانش پیشین در مورد زیرساخت رابانش ابری، نتایج بهتری در مقایسه با دو روش دیگر دارد.

شکل (۱۲) معیار بهره‌وری در لایه وب را نشان می‌دهد. در این حالت نیز دیده می‌شود روش پیشنهادی در این پژوهش نتایج متعادلی را تولید می‌کند. شکل (۱۳) نیز نتایج بهره‌وری در لایه پایگاه داده را نشان می‌دهد.

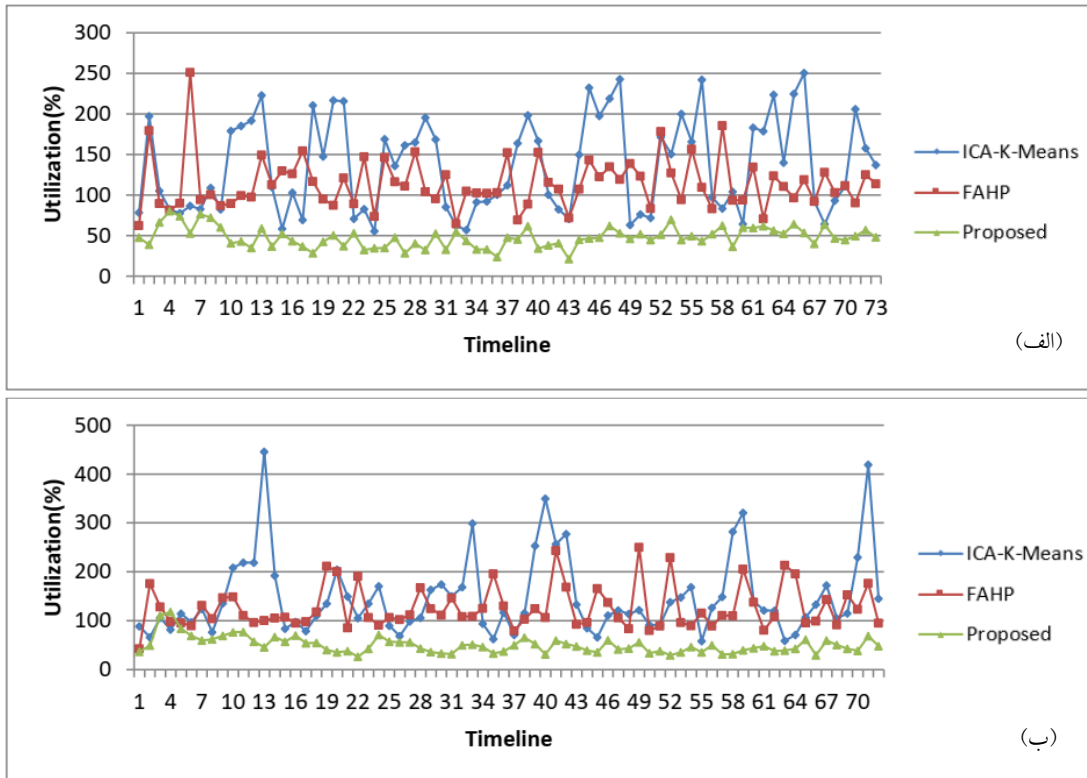
درباره نمودارهای هزینه از نظر نوع الگوی بار کاری سینوسی ND و PD، بر طبق تعریف معیار هزینه (رابطه ۵)، همان طور که اشاره شد، هزینه کل به تعداد ماشین‌های مجازی تخصیص داده شده و میزان تخطی از SLA بستگی دارد. به دلیل ذات پیچیده الگوی بار کاری PD نسبت به ND، و همچنین داشتن تعداد زیادتری از درخواست به صورت همزمان و ناگهانی در یک بازه زمانی، زمان پاسخ برای درخواست‌های ابری افزایش یافته و باعث افزایش میزان تخطی از SLA و متعاقب آن هزینه کل برای الگوی بار کاری PD نسبت به ND در هر سه لایه خواهد شد و از طرفی، الگوی بار کاری ND، به دلیل تکرار الگوهای منظم و قابل پیش‌بینی، هزینه کل کمتری نسبت به الگوی بار کاری PD خواهد داشت. همچنین از طرفی، همان طور که در بخش ارزیابی مقیاس‌پذیری ماشین‌های مجازی مشاهده می‌کنید، تعداد ماشین‌های مجازی تخصیص داده شده برای الگوی بار کاری PD نسبت به ND در هر سه لایه وب، برنامه کاربردی و پایگاه داده بیشتر است. این خود دلیل دیگری بر افزایش هزینه کل در الگوی بار کاری



شکل (۱۱): مقایسه بهره‌وری روش‌های پیشنهادی در لایه برنامه کاربردی: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۲): مقایسه بهره‌وری روش‌های پیشنهادی در لایه وب: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۳): مقایسه بهره‌وری روش‌های پیشنهادی در لایه پایگاه داده: (الف) مجموعه داده NP، (ب) مجموعه داده PD

خاص بر طبق تعریف بهره‌وری منابع (رابطه ۴) ایجاد نمی‌کند. در واقع، در محاسبه بهره‌وری منابع، نسبت دستورالعمل‌های درخواست که در یک بازه زمانی مشخص اجرا می‌شوند (نه همه دستورالعمل‌های درخواست که ممکن است در چندین بازه زمانی اجرا می‌شوند)، به عنوان میزان MIPS اختصاص داده شده برای محاسبه بهره‌وری منابع در یک بازه زمانی در نظر گرفته می‌شوند. توجه داشته باشید که روش‌هایی که بهره‌وری منبع بالای ۱۰۰ درصد دارند، برایشان حالت یا وضعیت *over-utilization* اتفاق می‌افتد و این وضعیت خوبی نیست؛ زیرا تأمین‌کننده ابری منابع به اندازه کافی برای تکمیل نیازهای مشتریان را در لحظات اوج بار کاری در اختیار ندارد و سیستم مدیریت تخصیص منابع ابر مجبور است بیش از حد از همین منابعی که در اختیار دارد استفاده کند و به همین دلیل است که بهره‌وری منابع برای دو رویکرد دیگر در برخی بازه‌های زمانی بالای ۱۰۰ رفته است. این وضعیت، وضعیت کسر تأمین (وضعیت *Under-Provisioning*)، نیز گفته می‌شود و به دلیل کمبود منابع در ساعات اوج بار کاری باعث ایجاد تخطی از

روش پیشنهادی در این پژوهش از دو ساختار بسیار انعطاف‌پذیر در کنار هم استفاده می‌کند، تصمیم‌گیری فازی در کنار تصمیم‌گیری سلسله‌مراتبی است. روش پیشنهادی FAHP قادر است با انعطاف بسیار زیاد از داده‌های پیشین برای تطبیق خود استفاده کند. با توجه به اینکه در حوزه سلامت با داده‌های بسیار حجیم مواجهیم، انعطاف‌پذیری الگوریتم می‌تواند تأثیر زیادی در به‌دست‌آوردن بهره‌وری داشته باشد.

کاهش هزینه‌ها امر ضروری برای تأمین و توسعه رایانش ابری در حوزه سلامت است؛ از طرفی ارائه سرویس‌های با کیفیت مستلزم بهره‌ خوب، برای تأمین‌کننده نیز می‌باشد. لذا پارامتر بهره‌وری از اهمیت بسیار زیادی برخوردار است. همان‌طور که در نمودارهای بهره‌وری منابع مشاهده می‌کنید، در برخی بازه‌های زمانی دو روش مورد مقایسه بهره‌وری بالاتر از ۱۰۰ دارند. علت این اتفاق این است که ممکن است یک درخواست ابری فراتر از چندین بازه زمانی اجرا شود و اجرای آن درخواست در یک بازه زمانی تمام نشود. با این حال، این واقعیت تضادی برای محاسبه بهره‌وری منابع در یک بازه زمانی

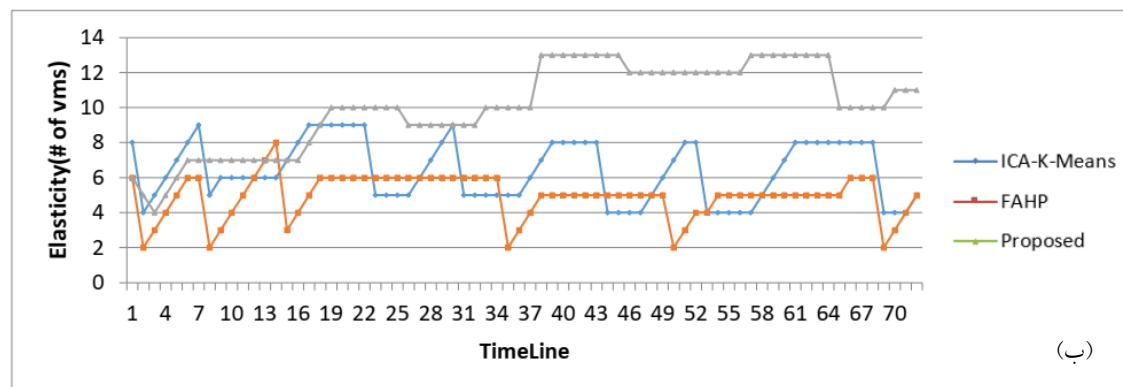
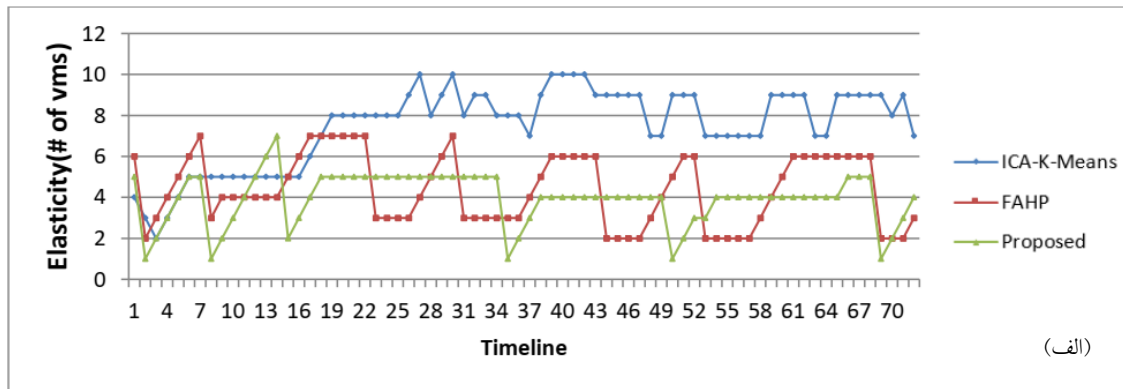
تعداد ماشین‌های مجازی شروع می‌کند، در ادامه از تعداد منابع کمتری برای رسیدن به اهداف استفاده می‌کند. در مقابل، روش ICA-K-Means بیشترین رشد در منابع را به همراه دارد و روش FAHP نسبت به ICA-K-Means عملکرد بهتری را از خود نشان می‌دهد. شکل (۱۵) خاصیت کشسانی روش‌های مورد مقایسه را در لایه وب و برای سه مجموعه داده مختلف بررسی می‌کند. نتایج نشان می‌دهد روش پیشنهادی در این پژوهش انعطاف بهتری نسبت به سایر روش‌های مورد مقایسه دارد. شکل (۱۶) نیز خاصیت کشسانی در لایه پایگاه داده و برای سه مجموعه داده مختلف را نشان می‌دهد.

حداقل نگر داشتن تعداد منابع برای کم کردن هزینه یک الزام است، از طرفی استفاده از منابع بیشتر برای تأمین توافقات با مشتری و جلوگیری از جریمه، نیز یک الزام است. بنابراین این معیارها دارای تعارض هستند، افزایش یکی منجر به کاهش دیگری می‌شود. الگوریتم مبتنی بر تصمیم‌گیری فازی سلسله‌مراتبی، انعطاف لازم برای تأمین این ملزومات به‌طور همزمان را دارد.

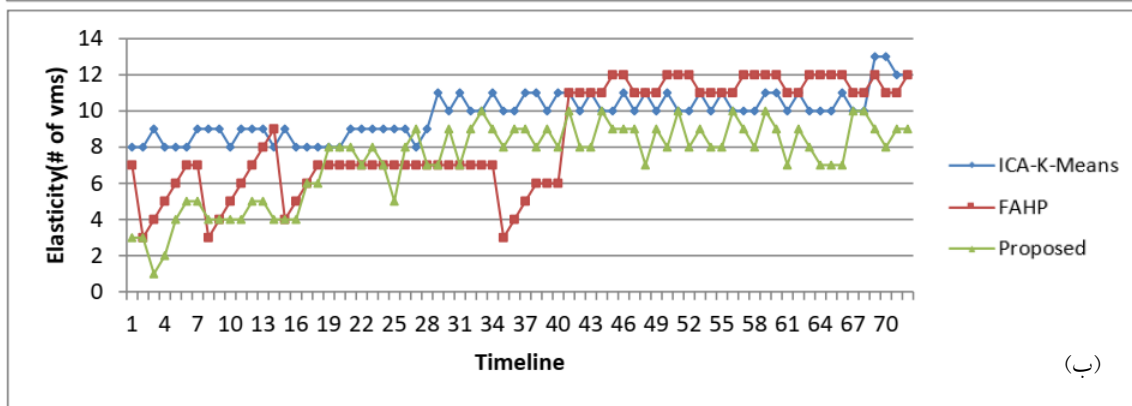
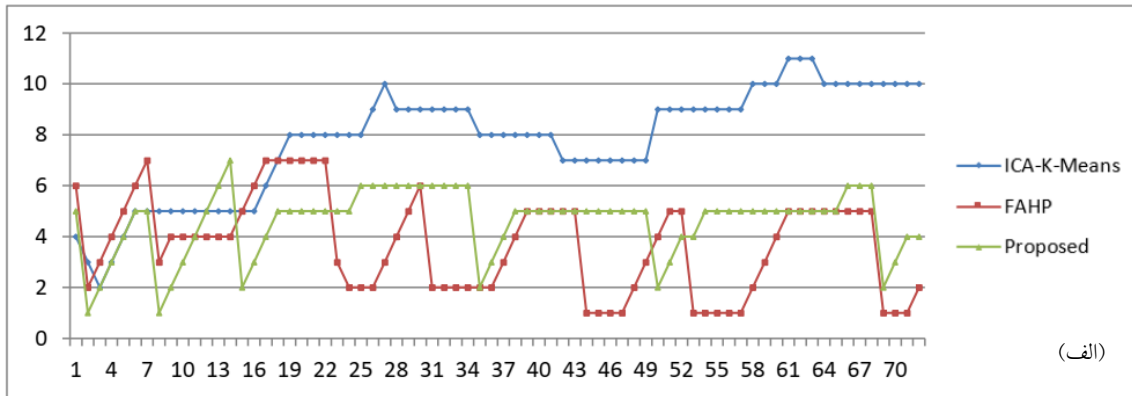
SLA شود، که این تأثیر منفی بر کیفیت سرویس کاربران گذاشته و به دنبال آن باعث از دست رفتن کاربران و درآمد حاصل از آن‌ها خواهد شد.

۳.۳.۴. ارزیابی مقیاس‌پذیری ماشین‌های مجازی

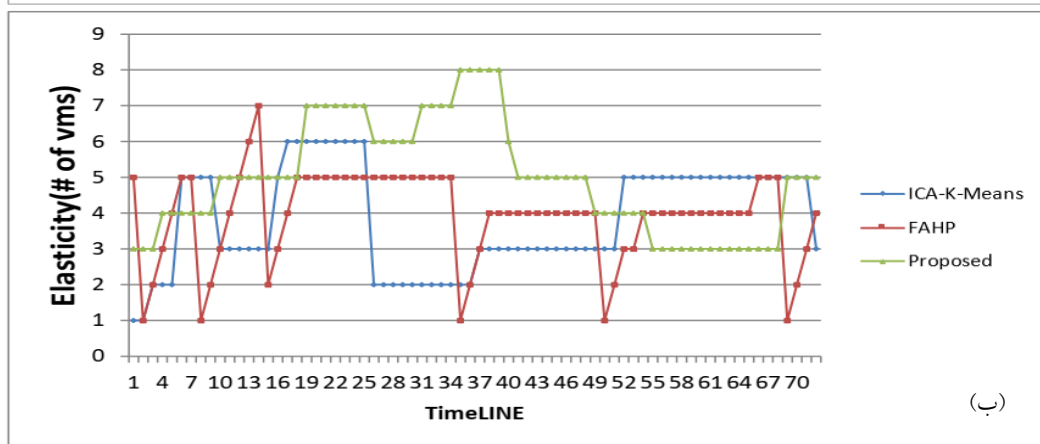
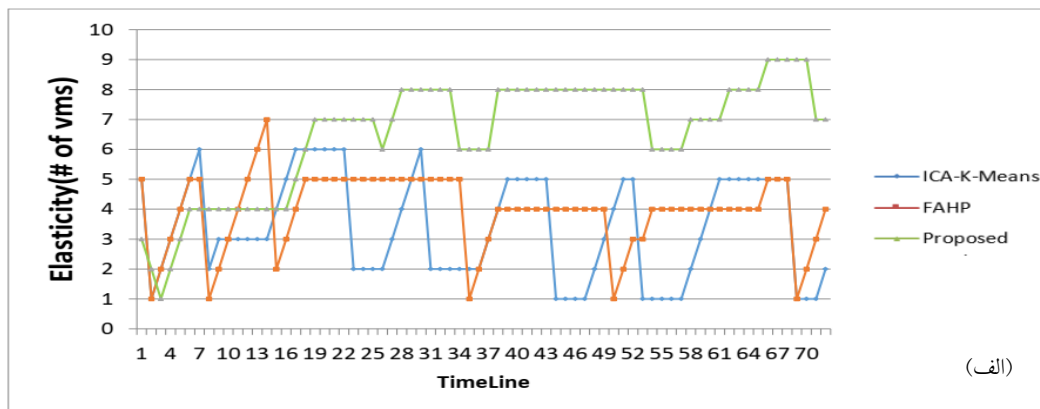
میزان مقیاس‌پذیری زیرساخت رایانش ابری در لایه‌های مختلف، معیاری برای ارزیابی خاصیت کشسانی رایانش ابری است. در ادامه به بررسی این معیار در لایه‌های مختلف می‌پردازیم. شکل‌های (۱۴)، (۱۵) و (۱۶) مقایسه روش‌های مورد بررسی در معیار مقیاس‌پذیری را نشان می‌دهد. در این بخش تعارض زیادی بین داده‌ها وجود دارد، در حالتی ممکن است نیاز باشد ماشین‌های مجازی فعال افزایش یابد، در غیر این صورت از توافقات سطح سرویس، تخطی می‌شود. البته لزوماً بحث ذکر شده، درست نیست چون ممکن است بتوان با تعداد ماشین‌های مجازی کمتری به نیازها پاسخ داد و بهره‌وری منابع را نیز در حد مناسبی حفظ کرد. شکل (۱۴) مقیاس‌پذیری در طول زمان و بر اساس تعداد ماشین‌های مجازی را نشان می‌دهد. روش پیشنهادی در این پژوهش در ابتدا با افزایش



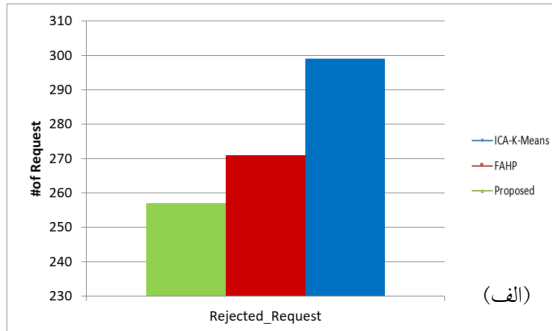
شکل (۱۴): مقایسه مقیاس‌پذیری روش‌های پیشنهادی در لایه برنامه کاربردی: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۵): مقایسه روش‌های مورد بررسی در پارامتر مقیاس‌پذیری لایه وب: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۶): مقایسه روش‌های مورد بررسی در پارامتر مقیاس‌پذیری لایه پایگاه داده: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۸): مقایسه روش‌های مورد بررسی در تعداد درخواست‌های رد شده در لایه وب: (الف) مجموعه داده ND، (ب) مجموعه داده PD



شکل (۱۹): مقایسه روش‌های مورد بررسی در تعداد درخواست‌های رد شده در لایه پایگاه داده: (الف) مجموعه داده ND، (ب) مجموعه داده PD

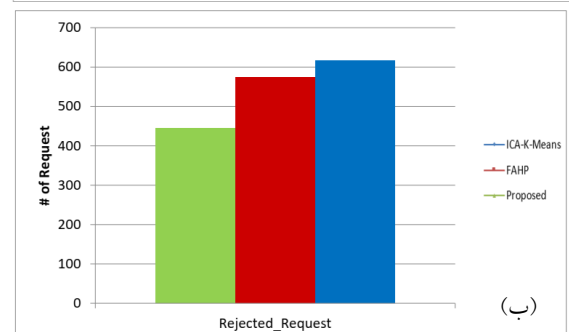
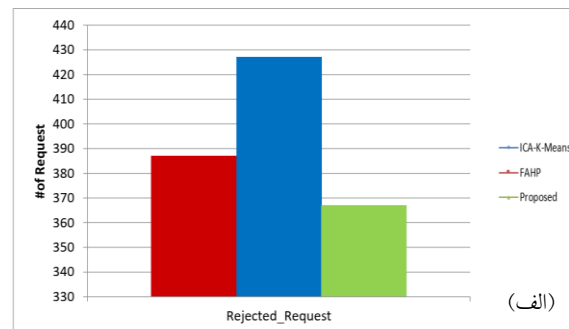
۴.۳.۵. تحلیل نهایی

شکل‌های (۲۰) و (۲۱) میانگین هزینه و بهره‌وری را برای لایه‌های مختلف نشان می‌دهند. مقادیر نمایش داده شده در این شکل‌ها میانگین هر سه مجموعه داده می‌باشند.

۴.۳.۴. ارزیابی میزان درخواست‌های رد شده

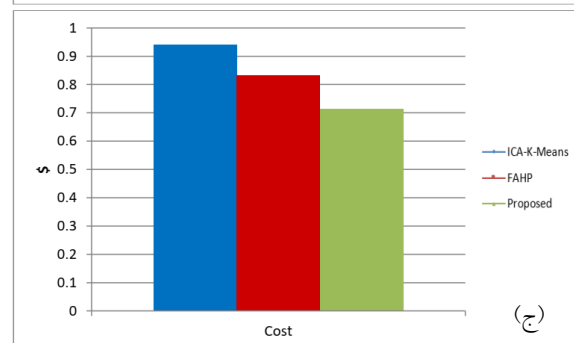
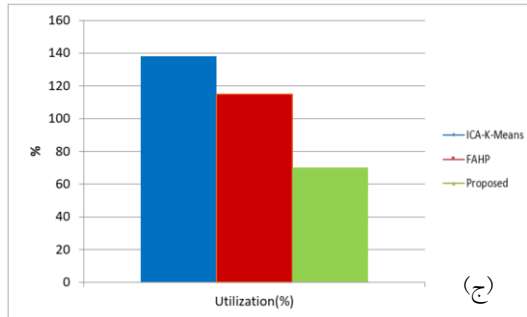
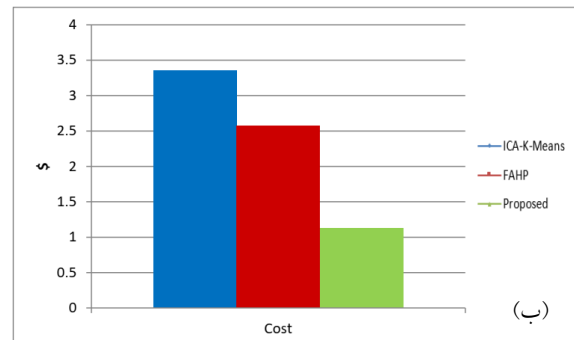
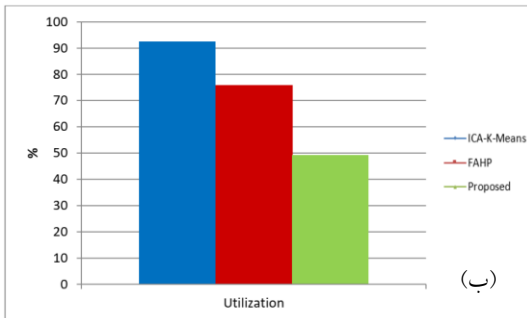
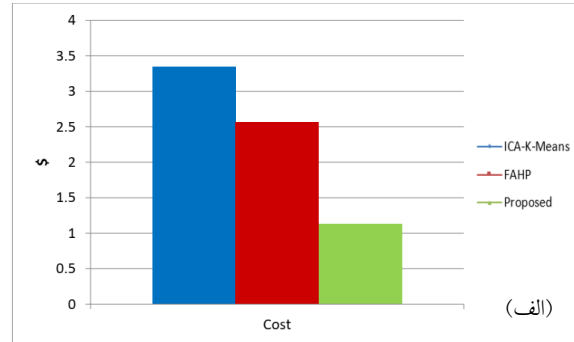
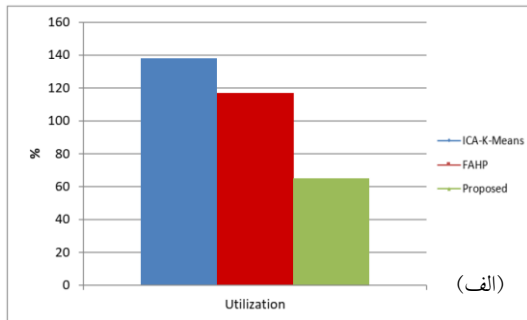
زیاد بودن تعداد درخواست‌های رد شده، باعث کاهش سود و از طرفی نیز ممکن است باعث افزایش هزینه شود. بنابراین روش پیشنهادی باید قادر باشد تعداد درخواست‌های کمتری را رد کند. به این طریق سود را افزایش دهد و کیفیت سرویس‌ها را در حد توافقات نیز نگه می‌دارد.

شکل (۱۷) درخواست‌های رد شده در لایه برنامه‌های کاربردی را برای سه مجموعه داده تست، شامل مجموعه داده‌های PD و ND را نشان می‌دهد. دیده می‌شود روش پیشنهادی در این پژوهش عملکرد بهتری نسبت به سایر روش‌ها دارد.



شکل (۱۷): مقایسه روش‌های مورد بررسی در تعداد درخواست‌های رد شده در لایه برنامه کاربردی: (الف) مجموعه داده ND، (ب) مجموعه داده PD

نمودارهای شکل (۱۸) نیز درخواست‌های رد شده در لایه وب را نشان می‌دهند. نمودارهای شکل (۱۹) نیز تعداد درخواست‌های رد شده در لایه پایگاه داده را نشان می‌دهند. از این نمودارها می‌توان دید روش پیشنهادی در این پژوهش عملکرد بهتری نسبت به روش‌های مورد مقایسه دارد.



شکل (۲۱): مقایسه متوسط بهره‌وری برای روش‌های مورد مقایسه: (الف) لایه کاربرد، (ب) لایه وب، (ج) لایه پایگاه داده

شکل (۲۰): مقایسه متوسط هزینه برای روش‌های مورد مقایسه: (الف) لایه کاربرد، (ب) لایه وب، (ج) لایه پایگاه داده

۵. نتیجه‌گیری و پیشنهادات

در سال‌های اخیر، تولید داده‌های حوزه سلامت به‌طور چشمگیری رشد داشته است؛ به طوری که بسیار حجیم بوده و پردازش آن‌ها نیاز به امکانات پردازشی قدرتمندی دارد. بر همین اساس، اخیراً به‌کارگیری قدرت پردازشی رایانش ابر جهت پردازش این نوع داده‌ها رایج شده است. یکی از مقوله‌های بسیار مهم، در رایانش ابر تأمین منابع به‌صورت دقیق است. به هر میزان تأمین منابع دقیق‌تر صورت گیرد، میزان تخطی از سرویس کاهش می‌یابد و به تناسب رضایت کاربر بالاتر می‌رود. بر همین اساس در این مقاله یک راهکار مدیریت منبع بهبودیافته برای پردازش داده‌های حوزه سلامت در محیط رایانش ابری ارائه شد. ساختار معماری پیشنهادی این گونه است که فرض می‌شود هر CDC یک واسطه‌گر یا واسطه‌گر

شکل (۲۰) متوسط هزینه برای سه لایه مختلف را نشان می‌دهد. با توجه به نمودارهای الف، ب و ج از شکل (۲۰) می‌توان دید که روش پیشنهادی در این پژوهش متوسط هزینه کمتری بر تأمین‌کننده تحمیل می‌کند. هر یک از اعداد نمایش داده شده در دو شکل (الف) و (ب) متوسط می‌باشند، برای مثال در شکل (۲۰) بخش الف، میانگین هزینه برای هر سه مجموعه داده را در لایه کاربرد محاسبه کردیم و نتیجه را در بخش الف نمایش دادیم. شکل (۲۱) نیز بهره‌وری را به‌طور متوسط نشان می‌دهد. همان‌طور که دیدیم روش پیشنهادی در این پژوهش به‌طور متوسط دارای عملکرد بهتری نسبت به سایر روش‌های مورد مقایسه است.

بهره‌وری و تعداد درخواست‌های رد شده با الگوریتم‌های FAHP و ICA-K-Means مورد بررسی و ارزیابی قرار گرفت که نتایج حاصل از آن بیانگر بهینه‌تر بودن راهکار در همه آزمایش‌های انجام گرفته است. از جمله مواردی که در آینده می‌توان در ادامه تحقیق انجام داد عبارت‌اند از: ترکیب سری‌های زمانی برای پیش‌بینی منابع به صورت بهینه، ترکیب روش یادگیری تقویتی با منطق فازی برای مقیاس‌بندی اتوماتیک و استفاده از ترکیب یادگیری تقویتی و شبکه عصبی در روش‌های مبتنی بر توافقات سطح سرویس که می‌تواند به‌طور چشمگیری باعث بهبود راهکار پیشنهادی شود.

دارد، که این بروکر مربوط به هر CDC اولاً درخواست‌های کاربران مراجعه‌کننده به این CDC را دریافت و تحلیل می‌کند و اگر میزان منابع مورد نیاز برای اجرای درخواست‌های کاربران موجود باشد، آن را اجرا می‌کند و نتیجه را به کاربران آن CDC برمی‌گرداند اما اگر واسطه‌گر نتواند درخواست‌های کاربران را با استفاده از منابع CDC جاری اجرا کند، باید با استفاده از الگوریتم AHP فازی که هر واسطه‌گر به آن دسترسی دارد، اقدام به انتخاب CDC مناسب از بین چندین CDC موجود کند که با آن‌ها در ارتباط است. در انتها راهکار پیشنهادی در محیط کلودسیم پیاده‌سازی و شبیه‌سازی گردید و سپس بر اساس پارامترهای مختلفی از جمله هزینه اجرا،

مراجع

- [1] Gill S. S., Garraghan P., Stankovski V., Casale G., Thulasiram R. K., Ghosh S. K., Ramamohanarao K., Buyya R., "Holistic resource management for sustainable and reliable cloud computing: An innovative solution to global challenge," Journal of Systems and Software, vol. 155, pp. 104-129, 2019.
- [2] اسداللهی، الناز، اصغری، سید امیر، «پیش‌بینی تعداد مناسب ماشین‌های مجازی بر اساس سری زمانی و روش‌های هوشمند مبتنی بر خوشه بندی ماشین‌های مجازی»، مجله محاسبات نرم، جلد ۶، شماره ۱، ص ۶۶-۷۷، ۱۳۹۶.
- [3] هراتیان اول، نفیسه، صفائی، علی اصغر، «کشف سرویس‌های ابری در زبان فارسی از طریق تکامل هستان‌شناسی»، مجله محاسبات نرم، جلد ۴، شماره ۲، ص ۸۴-۹۳، ۱۳۹۴.
- [4] Nzanywayingoma F., Yang Y., "Efficient resource management techniques in cloud computing environment: a review and discussion," International Journal of Computers and Applications, vol. 41, no. 3, pp. 165-182, 2019.
- [5] Ghobaei-Arani M., Jabbehdari S. and Pourmina M. A., "An autonomic resource provisioning approach for service-based cloud applications: A hybrid approach," Future Generation Computer Systems, vol. 78, pp.191-210, 2018.
- [6] قاسمی، رضا، محمدی، حمید رضا، طاهر، سید عباس، «کنترل فرکانس یک ریزشبکه جزیره‌ای با استفاده از کنترل هوشمند پاسخ‌گویی بار مبتنی بر منطق فازی و الگوریتم بهینه‌سازی ازدحام ذرات»، مجله محاسبات نرم، جلد ۶، شماره ۲، ص ۱۸-۳۱، ۱۳۹۶.
- [7] Kumar P., and Rakesh K., "Issues and challenges of load balancing techniques in cloud computing: a survey," ACM Computing Surveys (CSUR), vol. 51, no. 6, pp. 1-35, 2019.
- [8] Aghdashi A., and Mirtaheri S. L., "A Survey on Load Balancing in Cloud Systems for Big Data Applications," in International Congress on High-Performance Computing and Big Data Analysis, pp. 156-173, Springer, Cham, 2019.
- [9] Elhoseny M., Abdelaziz A., Salama A. S., Riad A. M., Muhammad K., and Sangaiah A. K., "A hybrid model of internet of things and cloud computing to manage big data in health services applications," Future generation computer systems, vol. 86, pp. 1383-1394, 2018.
- [10] Xu G., Pang J., and Fu X., "A load balancing model based on cloud partitioning for the public cloud," Tsinghua Science and Technology, vol. 18, no. 1, pp. 34-39, 2017.
- [11] Alelaiwi A., "A collaborative resource management for big IoT data processing in Cloud," Cluster Computing, vol. 20, pp. 1791-1799, 2017.
- [12] Hadji M., and Zeghlache D., "Mathematical programming approach for revenue maximization in cloud federations," IEEE trans ICA-K-Means on cloud computing, vol. 5, no. 1, pp. 99-111, 2015.
- [13] Ryan T., and Lee Y. C., "Multi-tier resource allocation for data-intensive computing," Big Data Research, vol. 2, no. 3, pp. 110-116, 2015.
- [14] Rajabion L., Shaltook A. A., Taghikhah M., Ghasemi A., and Badfar A., "Healthcare big data processing mechanisms: the role of cloud computing," International Journal of Information Management, vol. 49, pp. 271-289, 2019.
- [15] Taher N. C., Mallat I., Agoulmine N., and El-Mawass N., "An IoT-Cloud Based Solution for Real-Time and Batch Processing of Big Data: Application in Healthcare," in 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART), pp. 1-8, IEEE, 2019.
- [16] Khorsand R., Ghobaei-Arani M., and Ramezani M., "FAHP approach for autonomic resource provisioning of multitier applications in cloud computing environments," Software: PraICA-K-Means and Experience, vol. 48, no. 12, pp. 2147-2173, 2018.
- [17] Shahidinejad A., Ghobaei-Arani M. and Masdari M., "Resource provisioning using workload clustering in cloud computing environment: a hybrid approach," Cluster Computing, pp.1-24, 2020.
- [18] Barbierato E., Gribaudo M., Iacono M., and Jakóbič A., "Exploiting CloudSim in a multiformalism modeling approach for cloud based systems," Simulation Modelling PraICA-K-Means and Theory, vol. 93, pp. 133-147, 2019.
- [19] Calheiros R.N., Ranjan R., Beloglazov A., De Rose C. A. and Buyya R., "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," Software: PraICA-K-Means and experience, vol. 41, no. 1, pp. 23-50, 2011.
- [20] https://downloads.cms.gov/medicare/2018Med2000_flatfiles.zip