یک الگوریتم تخصیص منابع دو مرحله‌ای در رایانش ابری با استفاده از ترکیب Fuzzy-PSO

مصطفی وکیلی فرد[۱] امیر صحافی[۲] امیرمسعود رحمانی[۳] پیمان شیخ‌الحرم مشهدی[۴]

۱- دانشجوی دکتری، گروه مهندسی کامپیوتر، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
st_m_vakilifard@azad.ac.ir

۲- استادیار، گروه مهندسی کامپیوتر، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
sahafi@iau.ac.ir

۳- استاد، گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
rahmani@srbiau.ac.ir

۴- استادیار، گروه مهندسی کامپیوتر، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
p_sheikholharam@azad.ac.ir

چکیده: **رایانش ابری برپایه روش پرداخت به ازای مصرف، نوع جدیدی از پردازش را برای کاربران فراهم می‌نماید تا منابع مورد نیاز برنامه‌های کاربردی و سیستم‌ها را در سطوح مختلف تخصیص دهد. از اینرو تخصیص منابع در رایانش ابری، بسیار مهم می‌باشد. در این پژوهش، الگوریتمی برای تخصیص منابع برپایه منطق فازی ارائه شده است. میزان بهره‌وری منابع، رضایت کاربران و سود فراهم‌کنندگان سرویس به عنوان پارامترهای هدف در ارزیابی این الگوریتم در نظر گرفته شده‌اند. درمرحله اول تعداد ۱۰۰ درخواست در بین ۹ ماشین فیزیکی توزیع شده‌اند و توسط الگوریتم فازی تخصیص منابع انجام شده است. ۹۵ درصد درخواست‌ها پاسخ داده شده و بهره‌وری منابع نیز ۶۱,۶۱ درصد حاصل شده است. در مرحله بعد با ترکیب الگوریتم فازی با PSO، عملیات تخصیص منابع به درخواست‌ها انجام شده است. مقایسه نتایج نشان می‌دهد که بهره‌وری منابع ماشین‌های فیزیکی در الگوریتم ترکیبی فازی و PSO ۰,۵ درصد بهبود یافته است. همچنین در الگوریتم ترکیبی، یک درصد پاسخ‌گویی به درخواست‌های کاربران ارتقاء یافته است که سود بیشتر فراهم‌کنندگان سرویس و رضایت بیشتر مشتریان را سبب می‌شود.**

واژه‌های کلیدی: **رایانش ابری، تخصیص منایع، منطق فازی، الگوریتم بهینه‌سازی ازدحام ذرات (PSO)، بهره‌وری منابع.**

نوع مقاله: **پژوهشی**

DOI: 10.52547/jiaeee.20.1.43

تاریخ ارسال مقاله: ۱۳۹۹/۹/۲۳
تاریخ پذیرش مشروط مقاله: ۱۳۹۹/۱۱/۲۰
تاریخ پذیرش مقاله: ۱۴۰۰/۳/۲۹
نام نویسنده‌ی مسئول: دکتر امیر صحافی
نشانی نویسنده‌ی مسئول: ایران – تهران – خیابان ده‌حقی – دانشگاه آزاد اسلامی- واحد تهران جنوب –– گروه مهندسی کامپیوتر

۱- مقدمه

رایانش ابری برپایه مدل پرداخت به ازای مصرف، این قابلیت را فراهم می‌کند تا کاربران به طور همزمان منابع مورد نیاز را درخواست نمایند. این تکنولوژی سه مدل سرویس‌دهی زیرساخت به عنوان سرویس (IaaS)، سکو به عنوان سرویس (PaaS) و نرم‌افزار به عنوان سرویس (SaaS) را فراهم می‌نماید. همچنین چهار مدل پیاده‌سازی به صورت عمومی، خصوصی، اجتماعی و ترکیبی را دارا می‌باشد. مطابق تعریف موسسه ملی استاندارد و تکنولوژی امریکا، رایانش ابری عبارتست از یک مدل برای دسترسی بر بستر شبکه به مجموعه‌ای از منابع اشتراکی قابل پیکربندی که این منابع به سرعت قابل تامین و یا آزادسازی باشند [۱]. یک سیستم مدیریت منابع ابری، از بخش‌های مختلفی تشکیل شده است. هر بخش با همکاری با سایر بخش‌ها، کارایی سیستم ابری را مهیا می‌سازد. متداولترین واحدها در سیستم‌های مدیریت منابع عبارتند از : کشف منابع، زمانبندی منابع، پایش منابع، پیش‌بینی منابع، رزروسازی منابع، قیمت‌گذاری و محاسبه قیمت، راهبری و تخصیص منابع [۲].

واحد کشف منابع، وظیفه یافتن منابع برای سیتسم ابری را برعهده دارد. این منابع می‌توانند درون یک سیستم ابری باشند یا تحت عنوان ائتلافی از ابرها، در بین چند سیستم ابری، مشترک باشند.

واحد زمانبندی، درخواست‌ها را براساس پارامترهایی که راهبران سیستم ابری تعیین می‌کنند، زمانبندی می‌نماید.

واحد پایش، مسئولیت پایش منابع را برعهده دارد تا منابع دچار سربار یا کمبود بار نباشند و وضعیت منابع را در اختیار عملیات راهبری سیستم ابری قرار می‌دهد. از آنجایی که ممکن است کاربران با سیستم ابری دارای توافق سطح سرویس باشند، وضعیت منابع توسط این واحد پایش و ثبت می‌شود تا در صورت نقض قرارداد اطلاعات لازم فراهم شوند.

واحد پیش‌بینی، با استفاده از روابط ریاضی، نظریه صف، الگوریتم‌های تکاملی و موارد مشابه، رفتارها و سوابق درخواست‌ها را بررسی می‌کند و مدل‌هایی برای پیش‌بینی نیازهای آتی درخواست‌ها فراهم می‌کند.

واحد رزروسازی منابع، با توجه به اطلاعات واحد پایش و با توجه به توافق سطح سرویس درخواست‌ها، در صورت نیاز منابع مورد نیاز آنها را رزرو می‌نماید.

واحد قیمت‌گذاری، هزینه‌های درخواست‌ها و محاسبه هرینه منابع مصرف شده را برعهده دارد.

واحد راهبری سیستم ابری، به صورت سیستماتیک و قابل پیکربندی، مجموعه فرآیندها و سایر واحدها را راهبری می‌نماید و توسط عوامل انسانی نیز قابل تنظیم و هدایت است.

یکی از مهمترین بخش‌ها در یک سیستم مدیریت ابری، واحد تخصیص منابع می‌باشد. به فرآیند اختصاص دادن منابع موجود به درخواست‌های کاربران، تخصیص منابع گفته می‌شود. کارایی یک سیستم ابری، به چگونگی تخصیص منابع به کاربران، بستگی دارد. تاکنون الگوریتم‌های مختلفی برای تخصیص منابع در یک سیستم ابری توسط پژوهشگران این حوزه ارائه شده است که هر یک از این الگوریتم‌ها دارای نقاط قوت و ضعف هستند. در ادامه، تعدادی از الگوریتم‌های تخصیص منابع بررسی شده‌اند.

در [۳] یک مکانیزم تخصیص منابع براساس مزایده برخط برای منابع چند بعدی متغییر در زمان ارائه شده است. همچنین یک مدل برنامه‌نویسی برای این دسته از منابع ارائه شده است تا با مکانیزم مزایده‌ی برخط منصفانه، مساله تخصیص منابع را حل نمایند.

در [۴] یک مکانیزم تخصیص منابع برای میکروسرویس‌های متمرکز ارائه شده است و تخصیص منابع به عنوان یک مساله برنامه‌نویسی درجه دوم باینری تعریف شده است. در این پژوهش، به منظور کاهش تاخیر بر زمان پاسخ، یک مدل تعامل براساس نظریه گراف‌های غیرجهتدار، بین ماژول‌های میکروسرویس فراهم شده است. تمرکز این مکانیزم بر کاهش تعامل بین منابع فیزیکی است تا زمان پاسخ کاهش یابد اما به بررسی مواردی مثل ازدحام و کاهش مصرف انرژی نپرداخته است.

در [۵] یک مکانیزم تخصیص منابع رادیویی و غیربرخط‌سازی کارها برای MEC در C-RAN ارائه شده است. در این پژوهش، یک مساله برنامه‌نویسی غیرخطی ترکیبی تصادفی تعریف شده است و برمبنای بهینه‌سازی، مساله به چهار زیرمساله شکسته شده است.

در [۶] یک مکانیزم تخصیص منابع مستقل ارائه شده است که به صورت اتوماتیک منابع را مطابق درخواست‌های کاربران، تخصیص می‌دهد. مکانیزم ارائه شده که بر پایه روش‌های فرااکتشافی بنا شده است، دارای دو هدف بهبود بهره‌وری منابع و زمان پاسخ است که در قرارداد توافق سطح سرویس بین کاربر و فراهم‌کننده سرویس ابری توافق شده است. در این مکانیزم، با توجه به نسبت بین زمان پاسخ عملیاتی شده با زمان پاسخ آستانه، تعداد ماشین‌های مجازی کم یا زیاد می‌شود.

در [۷] یک الگوریتم برای تخصیص منابع در سیستم‌های ابری توزیع‌شده در مناطق مختلف ارائه شده است. مولفان یک حد بالا و پایینی را برای راه‌حل پیشنهادی در نظر گرفتند که نتایج این مکانیزم به آن وابسته می‌شود. این پژوهش بر سه مولفه مقادیر مطلوب، تنوع و همبستگی مرزهای مناطق، تمرکز یافته است.

در [۸] پژوهشی برای تخصیص ایمن و موثر ماشین‌های مجازی، انجام شده است تا از حملات بین ماشین‌های مجازی مشترک بر روی یک سرور فیزیکی جلوگیری نمایند. استراتژی ارائه شده، دارای اهداف ارتقاء امنیت، توازن بار و مصرف انرژی می‌باشد.

در [۹] یک مدل تخصیص منابع براساس تغییرات ماتریس هزینه الگوریتم Hungarian ارائه شده است. یک مدل ریاضی خطی برای توسعه مقیاس‌پذیری آن نیز فراهم شده است.

در [۱۰] یک مدل پیش‌تخصیص برای منابع شبکه‌ای در مراکز داده ارائه شده است. فراهم‌کنندگان سرویس چند استراتژی پیش-تخصیص را بر اساس مدل شبکه، نظریه گراف و پهنای‌باند اشتراکی تولید می‌کنند. افزایش اندازه شبکه، پیچیدگی و زمان اجرای زیاد آن از چالش‌های این الگوریتم می‌باشد.

در [۱۱] یک الگوریتم تخصیص پویا ماشین‌های مجازی براساس مرور پیوسته دوره‌ای در سطح SaaS ارائه شده است. مرور دوره‌ای براساس مدل مارکوف بوده و تعداد کاربران منتظر/ دریافت کننده سرویس و تعداد ماشین‌های مجازی برای درخواست مدلسازی شده‌اند.

در [۱۲] یک الگوریتم تخصیص ماشین‌های مجازی براساس الگوریتم Haizea و با استفاده از سیاست حریصانه در سطح IaaS ارائه شده است. منابع بر اساس ضوابطی همچون اولویت، ظرفیت و قیمت رده‌بندی می‌شوند و سپس لیست آنها مرتب شده و مناسب‌ترین منبع انتخاب می‌شود.

در [۱۳] یک الگوریتم تخصیص منابع چند متغییره برای سیستم‌های سرویس‌گرا ارائه شده است. در این الگوریتم، درخواست‌ها به سه دسته پردازنده محور، حافظه محور و نرمال دسته‌بندی شده‌اند. از اینرو برای واحد تخصیص منابع، اجزایی از قبیل انتقال، بازیابی و تطبیق را در نظر گرفته بودند تا در حلقه‌های متناوب عملیات را انجام دهند.

در [۱۴] یک مکانیزم تخصیص ماشین‌های مجازی با قابلیت رزروسازی منابع ارائه شده است به طوری که زمان ورود و زمان اجرای درخواست‌ها به عنوان متغییر در نظر گرفته شده اند.

در [۱۵] یک الگوریتم تخصیص منابع تطبیقی با استفاده از خوشه‌بندی و توازن بار ارائه شده است. برای پیش‌بینی وضعیت درخواست‌های ورودی نیز از رگرسیون استفاده شده است. زمان اجرا و تخصیص ماشین‌های مجازی به عنوان متغییرهای این الگوریتم در نظر گرفته شده است.

در [۱۶] چهار سیاست برای تخصیص منابع در نظر گرفته شده است به طوری که یکی از سیاست‌ها به صورت ایستا و بقیه پویا هستند. سود فراهم‌کنندگان سرویس ابری به عنوان متغییر در نظر گرفته شده که برای بیشینه‌سازی آن، از مدل مارکوف با زمانبندی بلادرنگ استفاده شده است.

در [۱۷] یک مکانیزم تخصیص منابع برای برنامه‌های محاسبه‌گرا بسیار حساس به تاخیر ارائه شده است. این مکانیزم از دو بخش زمانبندی و تطبیقی تشکیل شده است. از الگوریتم تابو برای انجام زمانبندی و از نظریه گرف‌ها برای بخش تطبیقی استفاده شده است. وضعیت شبکه، منابع و اهمیت فعالیت‌ها به عنوان متغییرها در نظر گرفته شده‌اند.

در [۱۸] یک ساختار برای تخصیص منابع در مراکز داده که در نقاط مختلف جغرافیایی توزیع شده‌اند، ارائه شده است. در ابتدا یک مدل سلسله مراتبی از مراکز داده و ارتباطات آن‌ها ایجاد شده و سپس با استفاده از برنامه‌نویسی خطی، یک مرکز داده مناسب برای پاسخگویی به درخواست انتخاب می‌شود.

در [۱۹] یک الگوریتم تخصیص منابع در فضای ابر اشیاء ارائه شده است. با مفروض داشتن همگنی منابع و کاربردها، اشیاء را در قالب یک مدل سه لایه‌ای از تجهیزات، گره‌های لبه و مراکز داده ابری در نظر گرفتند به طوری که گره‌هایی که بیشترین منابع را دارا بوده و در کمترین فاصله با در خواست کننده قرار دارند، به عنوان گره‌های رده بالا برای سرویس دهی انتخاب می‌شوند.

به دلیل اهمیت تخصیص منابع در کارایی یک سیستم ابری، پژوهش‌های متعددی در این زمینه انجام شده است به طوری که در هر سال، حداقل یک مقاله‌ی مروری برای بررسی و تحلیل الگوریتم-های تخصیص منابع منتشر شده است[۲].

به طور کلی مکانیزم‌های تخصیص منابع در قالب تخصیص ایستا و یا تخصیص پویا، منابع را به درخواست‌ها تخصیص می‌دهند. در دسته ایستا، پس از وصول درخواست‌ها، منابع تخصیص داده می‌شوند و در زمان تخصیص منابع، امکان وصول درخواست جدید و یا تغییر درخواست‌های پیشین وجود ندارد، حال آنکه در دسته پویا، در زمان عملیات تخصیص منابع، امکان وصول درخواست‌های جدید و تغییر درخواست‌های پیشین میسر می‌باشد.

از منظر یک دسته‌بندی دیگر، واحد تخصیص منابع، عملیات مورد نظر را یا به صورت متمرکز انجام می‌دهد و یا غیرمتمرکز. همچنین مکانیزم‌های تخصیص منابع را می‌توان از جنبه‌ی تعداد اهداف نیز مورد مقایسه قرار داد. برخی مکانیزم‌ها فقط یک هدف را دنبال می‌نمایند اما برخی دیگر از مکانیزم‌ها، دو یا چند هدف را مدنظر قرار می‌دهند.

پارامتر دیگر برای بررسی مکانیزم‌های تخصیص منابع، اهمیت لحاظ نمودن کیفیت سرویس در عملیات تخصیص منابع می‌باشد. یک مکانیزم تخصیص منابع می‌تواند متغییر/ متغییرهای مختلفی را برای بهبود کیفیت سرویس درنظر گیرد.

موضوع دیگر برای بررسی مکانیزم‌های تخصیص منابع، بهره‌گیری از هوش مصنوعی و الگوریتم‌های فرااکتشافی است. به عبارت دیگر آیا یک مکانیزم تخصیص منابع از الگویتم‌های هوش مصنوعی به منظور بهبود نتایج حاصله، بهره‌برداری می‌نماید؟

با توجه به موارد مذکور، در این پژوهش یک الگوریتم تخصیص منابع با قابلیت‌های

- تخصیص پویا
- چندهدفه
- لحاظ پارامترهای کیفیت سرویس
- بهره‌گیری از هوش مصنوعی و الگوریتم‌های فرااکتشافی

ارائه شده است به طوری‌که در مکانیزم‌های مشابه، به تمامی این موارد در قالب یک مکانیزم، پرداخته نشده است.

در ادامه ساختار مقاله بدین صورت است: تعریف مساله پژوهش در بخش دو ارائه شده است. الگوریتم‌های پیشنهادی برای تخصیص منابع، در بخش سه فراهم شده‌اند و نهایتا در بخش چهار نتیجه‌گیری و موضوعاتی برای پژوهش‌های آتی ارائه شده است.

۲- طرح مساله پژوهش

در این پژوهش، پردازنده، حافظه اصلی و حافظه جانبی به عنوان منابع در نظر گرفته شده‌اند. فرض شده است که یک کاربر می‌تواند حداکثر تا ۸ هسته پردازنده، ۱۶ گیگابایت حافظه اصلی و ۵۰۰ گیگابایت حافظه جانبی را درخواست نماید. همچنین حداقل میزان برای یک درخواست، ۱ هسته پردازنده، ۱ گیگابایت حافظه اصلی و ۵۰ گیگابایت حافظه جانبی در نظر گرفته شده است. تمامی سرورهای فیزیکی مشابه هستند و هر کدام ۳۲ هسته پردازنده، ۲۵۶ گیگابایت حافظه اصلی و ۴ ترابایت حافظه جانبی دارند. تابع هرینه برای هر درخواست مطابق معادله (۱) محاسبه می‌شود که در آن، Cost_CPU، Cost_RAM و Cost_Disk میزان هرینه برای دریافت هر واحد از این منابع می‌باشند و Req_CPU، Req_RaM و Req_Disk نیز مقدار منابع درخواست شده هستند.

$$Cost\ Function = Req_CPU * Cost_CPU + Req_RAM * Cost_RAM + Req_Disk * Cost_Disk \quad (1)$$

مقدار بهای هر یک از منابع مطابق بهای اقلام CPU Intel Xeon E5 V4 2660، RAM 16 GB DDR4 HP و Disk SATA 4TB Transcend در نظر گرفته شده‌اند. با توجه به هزینه هر یک از این منابع، با تقسیم بر تعداد واحد آنها، بهای هر واحد از منابع مشخص شده است. به طور مثال با تقسیم هزینه یک عدد حافظه جانبی Disk SATA 4TB Transcend بر ۴۰۰۰، میزان هزینه یک گیگابایت حافظه جانبی مورد درخواست کاربران بدست آمده است. سپس سهم هر واحد از یک منبع، از هزینه‌ی یک ماشین فیزیکی محاسبه و نرمال‌سازی شده است که مقادیر آن در جدول(۱) نشان داده شده است.

جدول(۱) : منابع ماشین‌های فیزیکی و هزینه هر واحد از منبع

منبع	ماشین فیزیکی	هزینه هر واحد از منبع
پردازنده (هسته)	۳۲	۰,۸۹
حافظه اصلی (گیگابایت)	۲۵۶	۰,۱۰۹۵
حافظه جانبی(گیگابایت)	۴۰۰۰	۰,۰۰۰۵

با توجه به موارد مذکور، الگوریتم پیشنهادی باید با در نظر گرفتن منابع ماشین‌های فیزیکی، تعداد ۱۰۰ درخواست را پاسخ و منابع مورد نیاز کاربران را تخصیص دهد تا نتایج آن مورد بررسی قرار گیرد.

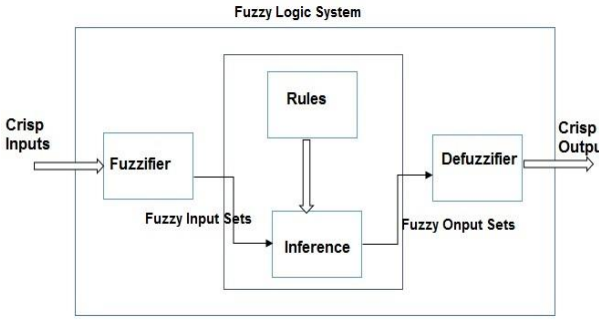
۳- الگوریتم فازی تخصیص منابع

۳-۱- طراحی الگوریتم فازی

از آنجایی که الگوریتم ارائه شده برپایه منطق فازی می‌باشد، در این بخش، توضیح مختصری در مورد منطق فازی نیز ارائه شده است.

منطق فازی برای اولین بار توسط پروفسور لطفعلی‌زاده در سال ۱۹۶۵ ارائه شده است. بنیان منطق فازی، بر پایه تئوری مجموعه‌های فازی است. در مجموعه‌های کلاسیک، یک عنصر یا متعلق به یک مجموعه است و یا نیست، اما در مجموعه‌های فازی، هر عضو می‌تواند تا حدودی عضو هر مجموعه‌ای باشد و سبب شود تا بر برخی محدودیت‌های محاسبات کلاسیک فائق آید. از اینرو در حوزه‌های مختلف ازجمله سیستم‌های کنترلی بکار گرفته شود. برای نمونه در [۲۰] یک الگوریتم مسیریابی در شبکه‌های حسگر بیسیم با استفاده از منطق فازی ارائه شده است. همچنین در [۲۱] برای تقطیع تصاویر و در [۲۲] برای شناسایی افراد از منطق فازی به منظور طراحی سیستم‌ها استفاده شده است.

همانطور که در شکل (۱) نشان داده شده است یک سیستم فازی از بخش‌های مختلفی تشکیل شده است. یک سیستم فازی دارای ورودی‌ها، قوانین فازی، موتور استنتاج فازی و تابع غیرفازی‌ساز می‌باشد تا عملیات پردازشی را انجام دهد و نهایتا خروجی‌های مورد نظر را تولید نماید. همانطور که در جدول (۲) نشان داده شده است سیستم فازی دارای دو ورودی و یک خروجی است. تابع عضویت هزینه برای هر درخواست و وضعیت بهره‌وری منابع به عنوان ورودی‌های سیستم‌فازی و میزان مطلوبیت ماشین فیزیکی برای پاسخ‌دهی به درخواست نیز به عنوان متغییر خروجی لحاظ شده است. به منظور تبدیل نتایج فازی حاصل از سیستم به مقادیر غیرفازی، از تابع متداول مرکز ثقل استفاده شده است.



شکل (۱): اجزاء سیستم فازی

جدول (۲): پارامترهای سیستم فازی

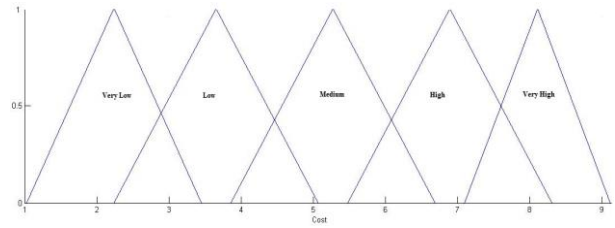
پارامتر	مقدار
Type	Mamdani
andMethod	Min
orMethod	Max
defuzzMethod	Centroid
impMethod	Min
aggMethod	Max
Input	[1 * 2 struct]
Output	[1 * 1 struct]

مطابق شکل (۲)، که شبه کد مکانیزم تخصیص منابع نمایش داده شده است، روند عملیات تخصیص منابع بدین صورت است که پس از تعیین ورودی‌ها و خروجی‌ها و ایجاد سیستم استنتاج فازی، برای درخواست (CPU, RAM, Disk) وارد شده، مطابق معادله (۱)، هزینه مربوطه محاسبه شده و تابع عضویت هزینه آن درخواست نگاشت می‌شود. سپس برای درخواست مذکور و با استفاده از اطلاعات حاصل از پایش منابع، وضعیت تمامی ماشین‌های فیزیکی به طور جداگانه مورد بررسی قرار می‌گیرد و ماشین‌های فیزیکی که می‌توانند منابع مورد نیاز درخواست مورد نظر را تامین نمایند، به عنوان ماشین‌های فیزیکی کاندیدای پاسخگویی به درخواست لیست می‌شوند. سپس برای هر یک از این ماشین‌های فیزیکی، عملیات استنتاج فازی انجام می‌شود و هر ماشین فیزیکی که میزان مطلوبیت بهتری دارد، به عنوان میزبان درخواست انتخاب و درخواست پاسخ‌دهی می‌شود. در ادامه وضعیت منابع ماشین فیزیکی انتخاب شده (با توجه به منابعی که به درخواست اختصاص داده است) بروزرسانی می‌شود تا ماشین فیزیکی مذکور در پاسخگویی به درخواست‌های جدید، با وضعیت بروزرسانی شده، بررسی شود. الگوریتم تخصیص منابع فازی به اختصار FRA نامگذاری شده است.

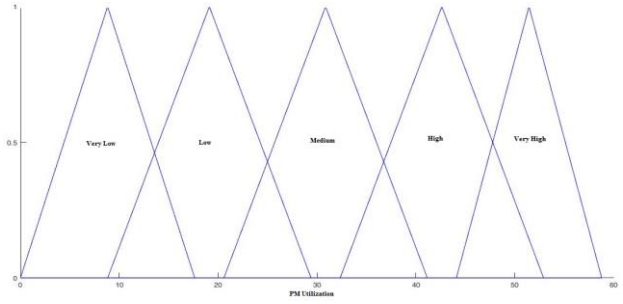
```
FRA Algorithm
Inputs: Requests, PM count
Output: Selected PM
        Function RA = FRA (requests, PMs count)
          Determine fuzzy input/ output MF, rules;
          Create FIS;
          LoadData();
          For I = 1: request numbers
            Get status of each PM;
             For j = 1: PM numbers
               List candidate PM for each request;
               Call FIS and determine selected PM by FIS;
               Update PM Information for selected PM;
            end
         end
        end
```

شکل (۲): شبه کد الگوریتم فازی تخصیص منابع

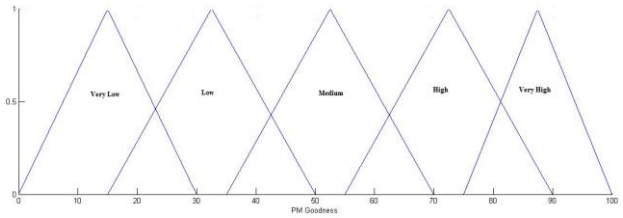
توابع عضویت مربوط به عضویت هزینه، بهره‌وری منابع و میزان مطلوبیت ماشین فیزیکی در شکل‌های (۳) الی (۵) ارائه شده‌اند.



شکل (۳): تابع عضویت هزینه

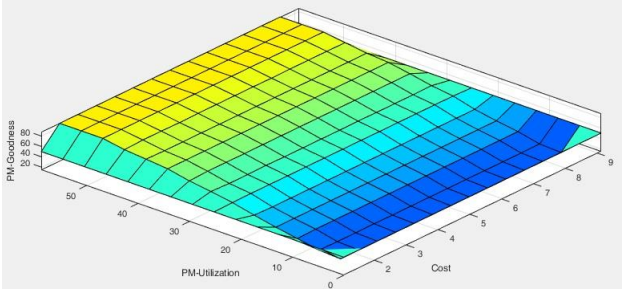


شکل (۴): تابع عضویت بهره‌وری منابع



شکل (۵): تابع عضویت مطلوبیت ماشین فیزیکی

وضعیت تاثیر متغییرهای ورودی و خروجی نسبت به یکدیگر، در شکل (۶) نمایش داده شده است.



شکل(۶): وضعیت تاثیر متغیرهای ورودی و خروجی

برای شبیه‌سازی از محیط متلب استفاده شده است. همانطور که در بخش دو ذکر شد، تعداد ۱۰۰ درخواست بین ۹ ماشین فیزیکی توسط الگوریتم ارائه شده، توزیع می‌شوند. برای اطمینان از نتایج، عملیات شبیه‌سازی ۵ مرتبه تکرار شده است. از بین ۱۰۰ درخواست، ۹۵ درخواست پذیرفته و ۵ درخواست رد شده‌اند. بهره‌وری منابع ماشین‌های فیزیکی، توسط الگوریتم FRA ۶۱,۶۱ درصد است.

۳-۲- ترکیب منطق فازی و PSO

الگوریتم بهینه‌سازی ازدحام ذرات Particle Swarm Optimization که به اختصار PSO نامیده می‌شود، یکی از روش‌های فرااکتشافی است که در آن تمامی ذرات برای رسیدن به بهترین مقادیر تلاش می‌کنند. در این الگوریتم اعضای جمعیت جواب‌های مساله، مستقیما با هم در ارتباط هستند و هر ذره در ارتباط با سایر ذرات، به تبادل اطلاعات می‌پردازد و همچنین با تکیه بر تجربیات مقادیر خوب پیشین، به سمت بهترین مقادیر تلاش می‌کند تا نهایتا بهینه‌ترین مقادیر برای یافتن بهترین جواب حاصل شود.

در این پژوهش الگوریتم فازی حاصل شده در مرحله قبل، با الگوریتم PSO ترکیب شده است تا نتایج حاصل از ترکیب عملکرد منطق فازی و بهینه‌سازی ازدحام ذرات با یکدیگر مورد مطالعه قرار گیرد. از اینرو الگوریتم ترکیبی FRA-PSO نامگذاری شده است. شبه کد الگوریتم و پارامترهای بکارگرفته شده، به ترتیب در شکل (۷) و جدول (۳) نشان داده شده است.

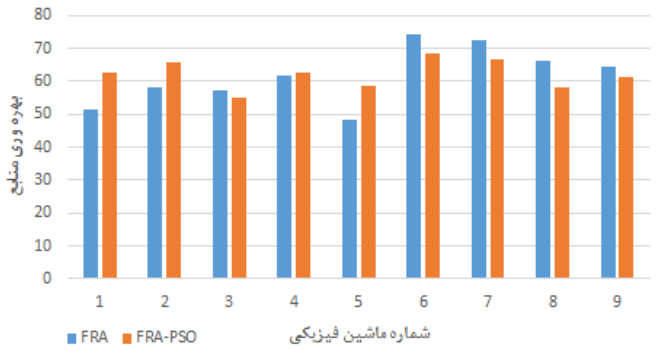
```
FRA-PSO Algorithm
Inputs: Requests, PM count
Output: Trained FIS
    Function RA = FRA-PSO (requests, PMs count)
      Get status of each PM;
      data = Load Data ();
      FIS = readfis('FRA.fis');
      FIS = TrainUsingPSO (FIS, data);
      TrainOutputs=evalfis (data. TrainInputs, FIS);
      TestOutputs=evalfis (data. TestInputs, FIS);
    end
```

شکل(۷): شبه‌کد الگوریتم FRA-PSO

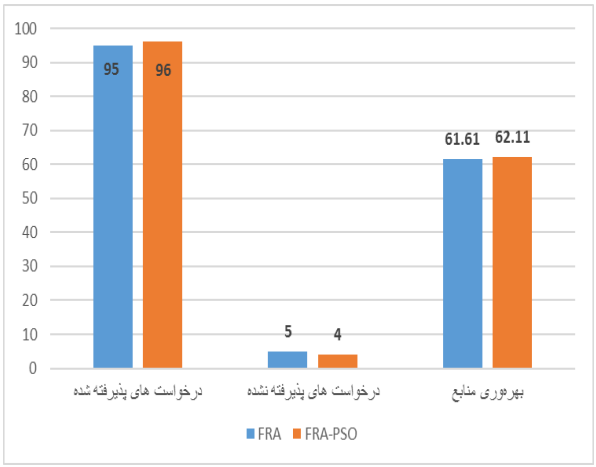
جدول (۳) : پارامترهای PSO

پارامتر	مقدار
MaxIt	۲۰۰
nPop	۵۰
Inertia Weight	۱
Inertia Weight Damping Ratio	۰,۹۹
Personal Learning Coefficient	۱
Global Learning Coefficient	۲

الگوریتم FRA-PSO نیز در شرایط مشابه با الگوریتم FRA شبیه‌سازی شده است. با بکارگیری این الگوریتم تعداد ۹۶ درخواست پذیرفته شده و ۴ درخواست رد شده است. همچنین میزان بهره‌وری منابع به ۶۲,۱۱ درصد ارتقاء پیدا یافته است. وضعیت ماشین‌های فیزیکی حاصل از اجرای الگوریتم‌های تخصیص منابع FRA و FRA-PSO در شکل (۸) نشان داده شده است. همانطور که در شکل(۹) نشان داده شده است، بهره‌وری منابع در الگوریتم ترکیبی FRA-PSO از الگوریتم FRA بیشتر است و تعداد درخواست بیشتری نیز پاسخ داده شده است.



شکل(۸): وضعیت ماشین‌های فیزیکی در FRA و FRA-PSO



شکل(۹) : وضعیت نتایج الگوریتم FRA با FRA-PSO

مطابق نتایج بدست آمده، ترکیب الگوریتم PSO با منطق فازی سبب شده است تا به میزان یک درصد درخواست‌های بیشتری پذیرفته شوند که رضایت‌مندی بیشتر کاربران را در پی دارد. از طرف دیگر به میزان حدود ۰,۵ نیز بهره‌وری منابع ارتقاء یافته است و علاوه بر استفاده موثرتر از منابع، سود بیشتری برای فراهم‌کنندگان سرویس ابری را سبب شده است. شایان ذکر است که بهره‌وری بیشتر منابع، از راه‌اندازی ماشین‌های فیزیکی بیشتر جلوگیری می‌نماید و کاهش مصرف انرژی و توسعه رایانش سبز را به همراه دارد.

۴- نتیجه‌گیری

رایانش ابری یک تکنولوژی پردازشی است که برمبنای "پرداخت به ازای مصرف" نیازهای کاربران را برآورده می‌نماید. یک سیستم ابری شامل واحدهای مختلفی است که این واحدها در تعامل با یکدیگر، عملیات سیستم ابری را انجام می‌دهند. یکی از مهمترین واحدها، تخصیص منابع است که عملیات آن، تاثیر زیادی در راندمان سیستم ابری دارد. در این مقاله یک الگوریتم تخصیص منابع بر اساس منطق فازی ارائه شد که برای بهبود کارایی آن، از الگوریتم PSO استفاده شد و ترکیب Fuzzy وPSO سبب شد تا درخواست‌های بیشتری توسط سیستم ابری پذیرفته شود و نهایتا بهره‌وری منابع در بین ماشین‌های

فیزیکی زیادتر شود. ترکیب سیستم فازی ارائه شده، با سایر الگوریتم-های فرااکتشافی و بررسی نتایج آن می‌تواند به عنوان موضوعی برای پژوهش‌های آتی در نظر گرفته شود.

مراجع

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing (Draft)," NIST Spec. Publ., vol. 800, p. 7, 2011.

[2] M. Vakili Fard, A. Sahafi, A. M. Rahmani, and P. S. Mashhadi, "Resource allocation mechanisms in cloud computing: a systematic literature review," IET Softw., vol. 14, no. 6, pp. 638 – 653, Oct. 2020.

[3] J. Zhang, X. Yang, N. Xie, X. Zhang, A. V. Vasilakos, and W. Li, "An online auction mechanism for time-varying multidimensional resource allocation in clouds," Futur. Gener. Comput. Syst., vol. 111, pp. 27–38, 2020.

[4] C. T. Joseph and P. K. Chandrasekaran, "IntMA : Dynamic Interaction-Aware Resource Allocation for Containerized Microservices in Cloud environments," J. Syst. Archit., vol. 111, 2020.

[5] Q. Zhang, L. Gui, F. Hou, J. Chen, S. Zhu, and F. Tian, "Dynamic Task Offloading and Resource Allocation for Mobile-Edge Computing in Dense Cloud RAN," IEEE Internet Things J., vol. 7, no. 4, pp. 3282–3299, 2020.

[6] T. Bhardwaj, H. Upadhyay, and S. C. Sharma, "Autonomic Resource Allocation Mechanism for Service-based Cloud Applications," in 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2020, pp. 183–187.

[7] E. Sherzer and H. Levy, "Resource allocation in the cloud with unreliable resources," Perform. Eval., vol. 137, pp. 366–373, 2020.

[8] H. Jia et al., "Security Strategy for Virtual Machine Allocation in Cloud Computing," Procedia Comput. Sci., vol. 147, pp. 140–144, 2019.

[9] T. R. Win, T. T. Yee, and E. C. Htoon, "Optimized Resource Allocation Model in Cloud Computing System," in 2019 International Conference on Advanced Information Technologies (ICAIT), 2019, pp. 49–54.

[10] Y. Guo, Z. Mi, Y. Yang, H. Ma, and M. S. Obaidat, "Efficient Network Resource Preallocation on Demand in Multitenant Cloud Systems," IEEE Syst. J., vol. 13, no. 4, pp. 4027–4038, 2019.

[11] B. Bouterse and H. Perros, "Performance analysis of the reserve capacity policy for dynamic VM allocation in a SaaS environment," Simul. Model. Pract. Theory, vol. 93, no. July, pp. 293–304, 2019.

[12] S. Alhassan and M. Abdulghani, "A Bio-Inspired Algorithm for Virtual Machines Allocation in Public Clouds," Procedia Comput. Sci., vol. 151, no. 2018, pp. 1072–1077, 2019.

[13] S. Gong, B. Yin, Z. Zheng, and K. Y. Cai, "Adaptive Multivariable Control for Multiple Resource Allocation of Service-Based Systems in Cloud Computing," IEEE Access, vol. 7, pp. 13817–13831, 2019.

[14] X. Zhang et al., "Energy-aware virtual machine allocation for cloud with resource reservation," J. Syst. Softw., vol. 147, pp. 147–161, 2019.

[15] C. Li, H. Sun, H. Tang, and Y. Luo, "Adaptive resource allocation based on the billing granularity in edge-cloud architecture," Comput. Commun., vol. 145, no. December 2018, pp. 29–42, 2019.

[16] J. Niño-Mora, "Resource allocation and routing in parallel multi-server queues with abandonments for cloud profit maximization," Comput. Oper. Res., vol. 103, pp. 221–236, 2019.

[17] H. Tang, C. Li, J. Bai, J. H. Tang, and Y. Luo, "Dynamic resource allocation strategy for latency-critical and computation-intensive applications in cloud–edge environment," Comput. Commun., vol. 134, pp. 70–82, 2019.

[18] H. Ziafat and S. M. Babamir, "A hierarchical structure for optimal resource allocation in geographically distributed clouds," Futur. Gener. Comput. Syst., vol. 90, pp. 539–568, 2019.

[19] T. C. S. Xavier et al., "Collaborative resource allocation for Cloud of Things systems," J. Netw. Comput. Appl., vol. 159, p. 102592, 2020.

[20] M. Vakili Fard, D. Ahmad Zadeh, and A. Jafarnezhad, "Introducing a Fault Tolerant Routing Protocol Based on Fuzzy Logic in Wireless Sensor Network s," J. Basic Appl. Sci. Res., vol. 3, pp. 200–204, 2013.

[21] P. Parsa and R. Safabakhsh, "A New Method for Image Segmentation based on Multi-Objective Differential Evolution Fuzzy Clustering," J. Iran. Assoc. Electr. Electron. Eng., vol. 13, no. 2, pp. 103–114, 2016.

[22] A. Ghanbari Sorkhi and H. Hassanpour, "Tracking and Re-identification of People in a Network of Cameras with Disjoint Views based on Fuzzy System in Indoor Environments," J. Iran. Assoc. Electr. Electron. Eng., vol. 13, no. 2, pp. 75–92, 2016.