دوفصلنامه محاسبات و سامانه‌های توزیع شده
سال هفتم، شماره ۱، شماره پیاپی ۱۳، صفحه ۱۲۴ تا ۱۳۶، سال ۱۴۰۳

بهینه‌سازی جانمایی چندمنظوره سرویس‌های اینترنت اشیا در رایانش ابری با استفاده از الگوریتم نهنگ و شبکه عصبی MLP

سید صادق حسینی(نویسنده مسئول)[۱]، محمد فرجی مهماندار[۲]
[۱] کارشناسی ارشد، دانشکده مهندسی فناوری اطلاعات، شبکه های کامپیوتری، دانشگاه آزاد، واحد پرند
[۲] استادیار، دانشکده مهندسی فناوری اطلاعات(گروه کامپیوتر)، دانشگاه آزاد، واحد پرند

چکیده

با رشد روزافزون کاربردهای اینترنت اشیا(IoT) ، استفاده از این فناوری به‌منظور تسهیل در انجام فعالیت‌های روزمره در زمینه‌های مختلف به‌طور قابل توجهی افزایش یافته و به بهبود کیفیت زندگی انسان‌ها کمک کرده است. در این راستا، رایانش مه به‌عنوان یک راه‌حل محاسباتی توزیع‌شده به منظور پشتیبانی از این برنامه‌ها با استفاده از گره‌های مه نزدیک به دستگاه‌های IoT به وجود آمده است. برنامه‌های مربوط به اینترنت اشیا به شکل چندین خدمت IoT با نیازهای مختلف کیفیت خدمات (QoS) توسعه یافته‌اند و می‌توانند بر روی گره‌های مه قرار گیرند. بنابراین، شناسایی یک طرح مناسب برای جانمایی خدمات و بهینه‌سازی توانمندی‌های منابع مختلف در اکوسیستم مه از جمله چالش‌های مهمی است که باید مدنظر قرار گیرد . این تحقیق، یک راه‌حل کارآمد برای جانمایی خدمات IoT ارایه کرده است که مبتنی بر شبکه عصبی پرسپترون چندلایه (MLP)و الگوریتم بهینه‌سازی نهنگ (WOA) می‌باشد و برای استقرار برنامه‌های اینترنت اشیا در زیرساخت مه طراحی شده است. این راه‌حل، الزامات QoS خدمات اینترنت اشیا و قابلیت‌های گره‌های مه موجود را برای تعیین یک سرویس کارآمد مورد بررسی قرار می‌دهد. در ابتدا، طرح جانمایی با استفاده از مدل MLP تخمین زده می‌شود و سپس با روش‌های بهینه‌سازی فراابتکاری بهینه‌سازی می‌شود .در این رویکرد، از پارامترهایی نظیر توان عملیاتی، مصرف انرژی و تاخیر به‌عنوان توابع هدف برای به دست آوردن طرح جانمایی بهینه استفاده شده است. نتایج شبیه‌سازی نشان می‌دهد که این راه‌حل پیشنهادی، موجب افزایش مصرف منابع و نرخ پذیرش خدمات می‌گردد و همچنین تاخیر خدمات و مصرف انرژی را نسبت به سایر روش‌های فراابتکاری کاهش می‌دهد

کلمات کلیدی: رایانش مه، برنامه‌های کاربردی IoT، جانمایی خدمات، الگوریتم بهینه‌ساز نهنگ (WOA)، شبکه عصبی پرسپترون چندلایه (MLP).

تاریخچه‌ی مقاله:
تاریخ ارسال: ۱۴۰۲/۱۲/۲۷
تاریخ اصلاحات: ۱۴۰۳/۰۵/۲۹
تاریخ پذیرش: ۱۴۰۳/۰۶/۱۵
تاریخ انتشار: ۱۴۰۳/۰۶/۳۰
ایمیل نویسنده مسئول:sadeghhosseini88@outlook.com

۱- مقدمه

در سال‌های اخیر رایانش مه[۱] به‌عنوان یک الگوی محاسباتی توزیع‌شده[۲] برای ارائه قابلیت‌های ذخیره‌سازی، محاسباتی و شبکه‌ای برای برنامه‌های کاربردی اینترنت اشیا (IoT)[۳] پدیدار شده است. رایانش مه به‌عنوان واسط بین ابر و دستگاه‌های پایانی عمل می‌کند و پردازش ذخیره‌سازی و خدمات شبکه را نزدیک‌تر به دستگاه‌های خود پلیان می‌دهد. این دستگاه‌ها گره‌های مه نامیده می‌شوند و آن‌ها می‌توانند در هر نقطه با اتصال به شبکه مستقر شوند. رایانش مه در یک محیط توزیع‌شده ظاهر می‌شود که محاسبات و امکانات ذخیره‌سازی را در لبه شبکه گسترش می‌دهد.

نتایج این تحقیق باعث دست‌یابی به یک راه حل محاسباتی توزیع‌شده برای استفاده از گره‌های موجود در نزدیکی دستگاه‌های اینترنت اشیا می‌گردد. از آنجایی که برنامه‌های اینترنت اشیا، در قالب چندین سرویس با الزامات کیفیت سرویس‌های مختلف توسعه یافته‌اند، می‌توانند بر روی گره‌های مه با قابلیت‌های منابع مختلف در اکوسیستم مه مستقر شوند. این موضوع از این لحاظ اهمیت دارد که با پیشرفت سریع اینترنت اشیا در سال‌های اخیر استفاده از آن‌ها

[1] Fog computing
[2] Distributed computing model
[3] Internet f Things (IoT)

برای تسهیل فعالیت‌های روزمره در حوزه‌های مختلف برای افزایش کیفیت زندگی انسان به‌طور چشمگیری افزایش یافته است و سیستم‌های اطلاعاتی و ارتباطی به‌طور نامتناهی در محیط اطراف ما جاسازی شده است و با استفاده از تجهیزات هوشمند به هم متصل شده‌اند. نتایج این پژوهش می‌تواند راه حل پیشنهادی مصرف منابع و نسبت پذیرش خدمات را افزایش داده و تاخیر سرویس و مصرف انرژی را در مقایسه با سایر مکانیسم‌های موجود را کاهش دهد. راه حل‌های پیشنهادی این پژوهش، خدمات اینترنت اشیا و قابلیت‌های در دسترس بودن گره‌های مه را برای تعیین یک طرح جانمایی سرویس کارآمد با استفاده از الگوریتم بهینه‌سازی نهنگ و شبکه عصبی پرسپترون چند لایه را بررسی می‌کند. نپرداختن به موضوع حاضر، الگوریتم ثابتی جهت بهره‌برداری کامل از تمام زوایای کاربردی جانمایی سرویس‌های اینترنت اشیا را نمایان نمی‌سازد و نمی‌توان به جریان داده‌ها با سرعت و به‌صورت بی‌درنگ سرویس‌دهی و رسیدگی نمود و پلتفرمی برای جمع‌آوری، پردازش و تحلیل ترافیک اینترنت اشیا بهینه‌سازی نمی‌گردد. علاوه بر این مکانیسم مبتنی بر تکامل، از توان عملیاتی، تاخیر و مصرف انرژی به‌عنوان توابع هدف برای یافتن طرح جانمایی اینترنت اشیا به‌صورت مطلوب استفاده نمی‌گردد؛ و الزامات کیفیت سیستم در هریک از سرویس‌های اینترنت اشیا برآورده نخواهد شد.

۲- اهداف پژوهش

هدف از ارائه این پژوهش، رسیدن به رویکردهای کارآمد برای جانمایی سرویس در محیط رایانش مه می‌باشد. همچنین می‌توانیم با استفاده از چهارچوب مفهومی رایانش مه به اهداف زیر دست‌یابیم:

کاهش تاخیر ارتباطی در شبکه با استفاده از الگوریتم فراابتکاری WOL و شبکه عصبی MLP

- کاهش هزینه و زمان پاسخ برنامه‌های کاربردی اینترنت اشیا
- افزایش بهره‌وری سیستم‌ها با استفاده از جانمایی مناسب سرویس‌ها بر روی منابع
- بهینه‌سازی مصرف انرژی، بهبود کارایی و افزایش رضایت کاربران

همچنین هدف کاربردی در این پژوهش رسیدن به کمترین زمان پاسخ برنامه‌های کاربردی می‌باشد. این فرایند برای برنامه‌هایی که به تاخیر حساس هستند مانند کارخانه‌های هوشمند، کنترل ترافیک، شبکه‌های خودرویی هوشمند، بهداشت هوشمند و بسیاری از ارگان‌های دولتی و خصوصی که دارای کاربردهای بلادرنگ و استفاده از اینترنت اشیا می‌باشند مفید واقع می‌شود.در این پژوهش یک راه حل کارآمد جانمایی سرویس اینترنت اشیا با الهام از استفاده از الگوی محاسباتی خودمختار برای استقرار برنامه‌های IoT بر روی زیرساخت مه مجازی ارائه کنیم. راه حل پیشنهادی شامل دو مرحله اصلی است: " پیش‌پردازش و تصمیم‌گیری"، با یک پایگاه دانش مشترک برای جانمایی خدمات اینترنت اشیا به‌طور خودکار بر روی گره‌های مه براساس نیازهای QOS. ابتدا، اطلاعات مورد نیاز درمورد الزامات QOS خدمات اینترنت اشیا و قابلیت‌های گره‌های مه موجود به‌طور مداوم در مرحله پیش‌پردازش توسط شبکه MLP جمع‌آوری می‌شود. اطلاعات جمع‌آوری شده برای اولویت بندی خدمات اینترنت اشیا براساس مهلت و زمان انتظار آن‌ها استفاده می‌شود. سپس در مرحله تصمیم‌گیری، یک راه حل مناسب برای قرارگیری با استفاده از روش فرا ابتکاری الگوریتم تعیین می‌شود. WOA به‌عنوان یک الگوریتم تکاملی برای حل مسائل بهینه‌سازی با توجه به مراحل اکتشاف و بهره‌برداری برای یافتن راه‌حل‌های مقرون به صرفه در فضای جستجو مناسب است. در نهایت راه حل قرارگیری به دست آمده با استفاده از تکنیک WOA در مرحله آخر اجرا می‌شود. این مراحل به‌طور مداوم در فواصل زمانی از پیش تعیین شده انجام می‌شود و خدمات اینترنت اشیا را بر روی گره‌های مه موجود مطابق با نیازهای QOS آن‌ها مستقر می‌کنند.

۲-۱: مطالعات پیشین

در جدول یک طبقه‌بندی از مطالعات پیشرفته با در نظر گرفتن استراتژی‌های بهینه‌سازی ارائه شده است. بر اساس بررسی‌های گسترده، چهار دسته اصلی از رویکردهای رایج مورد استفاده برای مقابله با مشکل قرارگیری برنامه IoT در رایانش مه شناسایی کردیم. مزنی و همکاران [۴۹] به چالش جانمایی خدمات در محیط‌های ابری پرداخته‌اند. جانمایی خدمات شامل استقرار و سازماندهی خدمات بر روی سرورهای محاسباتی/ذخیره‌سازی با کارایی بالا است، در حالی که محدودیت‌های مختلفی نظیر مصرف منابع، سطوح امنیتی، زمان انتقال داده و توافق‌نامه‌های سطح خدمات (SLA) را در نظر می‌گیرد. بسیاری از روش‌های جانمایی خدمات موجود به‌صورت ثابت عمل می‌کنند و به تغییرات محیط ابری تطبیق نمی‌یابند. برای حل این مسئله، این مقاله رویکردی مبتنی بر پیش‌بینی پیشنهاد می‌کند که شامل پیش‌بینی مصرف منابع در سرورهای ابری در مدت کوتاه و

پیش‌بینی ترافیک مهاجرت خدمات در آینده می‌شود. این اقدام با استفاده از یک شبکه عصبی بازگشتی با دروازه‌های ورودی انجام می‌شود، که نوعی از یادگیری عمیق است. تولی و همکاران [۵۰] به چالش استقرار شبکه‌های عصبی بزرگ در پلتفرم‌های محاسباتی محدود منابع مرتبط با حوزه‌های بحرانی نظیر نظارت و بهداشت پرداخته‌اند. شبکه‌های عصبی عمیق توانایی‌های برجسته‌ای در تشخیص الگوها نشان داده‌اند، اما نیازهای محاسباتی و حافظه‌ای آن‌ها چالش‌هایی را برای استقرار در محیط‌های لبه ایجاد می‌کند. برای حل این مشکل، مقاله راه حلی را پیشنهاد می‌کند که شامل تقسیم شبکه‌های عصبی با نیازهای منابع بالا به اجزای جداگانه و سبک برای پردازش توزیع شده می‌شود. برای حل این موضوع، تولی و همکاران SplitPlaceرا معرفی نمودند، SplitPlace از یادگیری تقویتی حساس به تصمیم برای جانمایی قطعات تقسیم شبکه عصبی روی دستگاه‌های لبه موبایل بهره‌مندی می‌کند و با محیط‌های ناپایدار سازگار می‌شود. در [۳۸] محمود و همکاران، یک روش ابتکاری آگاه از انرژی برای قرار دادن وظایف برنامه بر روی دستگاه‌های مه با هدف به حداقل رساندن مصرف انرژی پیشنهاد شده است. ایده آنها استفاده از فناوری مقیاس‌بندی ولتاژ دینامیکی و فرکانس برای تنظیم فرکانس CPU دستگاه‌های مه به‌گونه‌ای است که حداقل افزایش مصرف انرژی را تضمین کند. در [۴۴]، بیتام و همکاران (۲۰۱۸) یک روش جدید، یعنی الگوریتم زندگی زنبورها (BLA)[4]، برای رسیدگی به مشکل تخصیص کار در یک محیط محاسباتی مه پیشنهاد کردند. هدف از این مطالعه بهینه‌سازی توزیع وظایف بین گره‌های مه با دستیابی به مبادله بهینه بین زمان اجرای CPU و حافظه مورد استفاده توسط گره‌های مه است.

- برنامه‌ریزی ریاضی مانند ILP[5]
- ابتکاری مانند best-fit و استراتژی‌های جستجوی backtracking
- فراابتکاری مانند الگوریتم‌های GA و PSO
- رویکردهای مبتنی بر یادگیری ماشین مانند یادگیری تقویتی عمیق

از منظر مقیاس‌پذیری، رویکردهای جانمایی سرویس مبتنی بر ILP از زمان CPU محاسباتی بالایی رنج می‌برند. حل کننده‌های ILP نمی‌توانند تعداد زیادی از متغیرها را در یک بازه زمانی معقول مدیریت کنند. روش‌های مبتنی بر ILP انتخاب کارآمدی برای یک محیط مه با تعداد زیادی منابع و پارامترهای بهینه‌سازی نیستند. از سوی دیگر، رویکردهای ابتکاری و فرا ابتکاری در مقایسه با روش‌های مبتنی بر ILP، زمان اجرای بسیار کوتاه‌تری دارند. با این حال، راه حل‌های بهینه را تضمین نمی‌کنند. بنابراین، روش‌های مبتنی بر ILP کاندیدای خوبی برای مسائل با اندازه کوچک هستند، در حالی که روش‌های ابتکاری و فراابتکاری را می‌توان برای مسائل با اندازه بزرگ ترجیح داد. علاوه بر این، راه‌حل‌های بازگشتی با روش‌های مبتنی بر ILP را می‌توان برای ارزیابی ابتکاری و فراابتکاری استفاده کرد. در نهایت، رویکردهای مبتنی بر یادگیری ماشین نیز مانند استراتژی‌های یادگیری تقویتی عمیق از زمان طولانی مدت آموزش با مشکل فضای راه‌حل بزرگ برای تصمیم‌گیری رنج می‌برند.

جدول ۱: خلاصه‌ای از مطالعات انجام شده

کاربرد	بهینه‌سازی	معیار	محدودیت‌ها
مراقبت بیمار(۳۸)	ابتکاری	انرژی	تحرک پذیری نادیده گرفته شده
عمومی(۴۴)	فراابتکاری	تاخیر	تاخیر ارسال لحاظ نشده است
Cloud(۴۹)	یادگیری عمیق	QOS	پیچیدگی محاسباتی بالا
بهداشت و نظارت(۵۰)	شبکه عصبی - SP	توان عملیاتی	تاخیر ارسال لحاظ نشده است
*پژوهش حاضر	**MLP WOA**	**توان عملیاتی مصرف انرژی تاخیر**	-

۲-۲: مهماری کلی

رویکرد حاضر (MLP-WOA) با کمک شبکه عصبی MLP و بهینه‌ساز WOA بهترین جانمایی خدمات را در مقابله با رویکردهای پیشین به ارمغان آورده است. حرکت وال‌ها در WOA متناظر با کاوش ترکیب‌های مختلف وزن‌های گره‌های مه است و بر جانمایی خدمات IoT در گره‌های مه تأثیر می‌گذارد. این الگوریتم به دنبال پیدا کردن پیکربندی‌های وزن گره‌های مه است که معیارهای QoS مطلوب را بهینه می‌کنند. MLP به‌عنوان یک ارزیاب هوش مصنوعی

[4] Bees Life Algorithm (BLA)

[5] Integer Linear Programming (ILP)

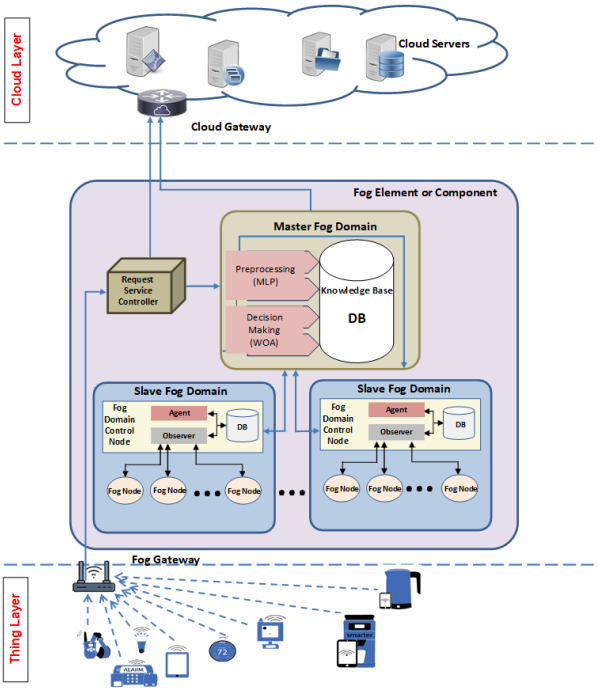
عمل می‌کند که معیارهای QoS را بر اساس وزن‌های گره‌های مه پیش‌بینی می‌کند. MLP با ارائه تخمین‌های سریع و دقیق از معیارهای QoS، به الگوریتم کمک می‌کند تا روی راه‌حل‌های ممکن تمرکز کند. فرآیند بهینه‌سازی تکرارپذیر یک حلقه بازخورد بین WOA و MLP ایجاد می‌کند. هرچه WOA راه‌حل‌ها را بهبود می‌دهد، MLP به دلیل آموزش بر روی داده‌های بهبود یافته دقیق‌تر می‌شود. پیش‌بینی‌های بهبودی که MLP ارائه می‌دهد، به WOA کمک می‌کند تا ترکیب‌های بهتری از وزن‌های گره‌های مه را کاوش کند.برای دستیابی به قابلیت‌های خود مدیریتی، واحد پیش‌پردازش و تصمیم‌گیری را به‌عنوان یک معماری منطقی مرجع برای تحقق مدیریت مستقل سیستم‌های کامپیوتری معرفی نموده‌ایم. در ابتدا، مرحله پیش‌پردازش، اطلاعات (به‌عنوان مثال، استفاده از شبکه، استفاده از CPU) را از طریق حسگرهای یک محیط پویا یا عنصر مدیریت‌شده (به‌عنوان مثال، یک شبکه کامپیوتری) ضبط می‌کند. سپس، اطلاعات را جمع‌آوری، فیلتر و تبدیل می‌کند تا یک مدل مطلوب به‌منظور تشخیص موقعیت‌های غیرعادی، پیش‌بینی موقعیت‌ها و توصیف موقعیت‌های مطلوب‌تر ایجاد کند. مرحله تصمیم‌گیری شامل تصمیم‌گیری با استفاده از قوانین رویداد- شرط- عمل[6] برای انتقال از وضعیت فعلی به وضعیت مطلوب است. در نهایت، تصمیمات تعیین شده توسط مرحله پیش‌پردازش را به‌منظور واکنش به تحولات محیطی از طریق تاثیرگذاری بر عنصر مدیریت شده اجرا می‌کند. علاوه بر این، یک پایگاه داده مشترک مسئول ذخیره دانش و تبادل رویدادها است.

۳-۲: چارچوب رویکرد پیشنهادی

یک چارچوب مستقل برای استقرار خدمات اینترنت اشیا در زیرساخت مه مجازی برای تحقق راه‌حل پیشنهادی شرح داده شده است، چارچوب پیشنهادی شامل سه لایه یعنی لایه اشیا، لایه مه و لایه ابر است.لایه اشیا به‌عنوان لایه زیرین شامل یک سری اشیا هوشمند مانند تبلت و تلفن همراه است تا اطلاعات مورد نیاز را از مه جمع‌آوری کند. لایه مه به‌عنوان لایه اصلی بین لایه‌های اشیا و ابر قرار دارد و شامل مجموعه‌ای از دامنه‌های مه برای رسیدگی به درخواست‌های کاربران اینترنت اشیا در مکان‌های مختلف جغرافیایی توزیع شده است. هر دامنه مه شامل چندین گره مه با قابلیت‌های منابع مختلف برای ارائه خدمات اینترنت اشیا با توجه به نیازهای QoS و یک گره کنترل دامنه است. حوزه مه مسئول تعیین استراتژی ارسال و استقرار خدمات اینترنت اشیا در میان حوزه‌های مه در بازه‌های زمانی خاص Δt با توجه به مدل محاسبات خودکار است؛ اطلاعات جمع‌آوری‌شده در مورد گره‌های مه را از طریق گره‌های کنترلی هر دامنه مه در پایگاه دانش مشترک نگه می‌دارد. از آنجایی که لایه مه به‌عنوان لایه اصلی مسئول تعیین گره‌های مه مناسب برای استقرار خدمات اینترنت اشیا است. طبق شکل۱، ردیف مه شامل سه جزء اصلی است: کنترل پذیرش[7]، مجموعه دامنه‌های مه فرعی[8] و مجموعه دامنه مه اصلی[9] است.

- **کنترل پذیرش:** این جز، مسئول پذیرش خدمات اینترنت اشیا از طریق دروازه مه یعنی از لایه اشیا و تحویل آنها به لایه مه یا لایه ابر است.
- **دامنه‌های مه فرعی:**. حوزه‌های مه، مراکز داده‌های خرد هستند که شامل مجموعه‌ای از گره‌های مه و یک گره کنترل دامنه مه هستند. هر گره مه به‌عنوان یک منبع محاسباتی، قدرت پردازش و قابلیت‌های ذخیره‌سازی را برای استقرار و ارائه خدمات اینترنت اشیا فراهم می‌کند. گره کنترل دامنه مه یک گره مه قدرتمند از نظر ظرفیت‌های محاسباتی با عملکرد گسترده برای نظارت بر وضعیت گره‌های مه از طریق مشاهده و ذخیره آنها در پایگاه داده محلی خود و بازگرداندن آنها به دامنه‌های مه اصلی است.
- **حوزه مه اصلی:** به‌عنوان مهم‌ترین مؤلفه در محیط مه، مسئولیت مدیریت و استقرار خدمات اینترنت اشیا دریافت شده از مؤلفه کنترل پذیرش را بر عهده دارد. اطلاعات مورد نیاز در مورد الزامات QoS خدمات را در پایگاه دانش خود نگهداری می‌کند.

[6] event-condition-action
[7] Admission Control
[8] Slave Fog Domains
[9] Master Fog Domain



شکل ۱. چارچوب رویکرد پیشنهادی

۴-۲: الگوریتم پیشنهادی

در این بخش، یک الگوریتم کارآمد جانمایی سرویس اینترنت اشیا برای استقرار برنامه‌های اینترنت اشیا در اکوسیستم مه ارائه شده است. گره کنترل مه در دامنه مه اصلی به‌طور دوره‌ای الگوریتم پیشنهادی را با توجه به مراحل "پیش‌پردازش، تصمیم‌گیری" در فواصل زمانی معین، اجرا و تکرار می‌کند. در ابتدا خدمات اینترنت اشیا دریافتی از طریق کنترل پذیرش و گره‌های مه موجود در مرحله پیش‌پردازش رصد می‌شوند و اطلاعات مورد نیاز در پایگاه دانش ذخیره می‌شود. سپس، اطلاعات ثبت شده را برای اولویت‌بندی برنامه‌های IoT برای استقرار در زیرساخت مه، پالایش و بررسی می‌کند. این کار به کمک شبکه عصبی MLP صورت می‌پذیرد و راه‌حل‌های کاندید برای جانمایی را در اختیار واحد تصمیم‌گیری (WOA) قرار می‌دهد. در واقع، به‌عنوان یک وظیفه پیش‌پردازش برای مرحله تصمیم‌گیری عمل می‌کند. پس از آن، مرحله تصمیم‌گیری از WOA برای یافتن یک طرح جانمایی سرویس کارآمد برای استقرار برنامه‌های IoT در گره‌های مه استفاده خواهد کرد. در نهایت، طرح قرارگیری به‌دست‌آمده از مرحله قبل را برای برآوردن نیازهای QoS مختلف برنامه‌های اینترنت اشیا انجام می‌دهد.

۵-۲: تابع هدف بهینه سازی

سه نوع تابع هدف یعنی توان عملیاتی، مصرف انرژی و تاخیر را در نظر می‌گیریم. هدف اول معیار توان عملیاتی یا نسبت پذیرش خدمات است، به معنی به حداکثر رساندن تعداد سرویس اینترنت اشیا در گره‌های مه (به جای سرورهای ابری) در حالی که الزامات QoS هر سرویس اینترنت اشیا را برآورده می‌کند، همانطور که در رابطه (۱-۱) بیان شده است.

$$T = \sum_{k=1}^{m}\sum_{l=1}^{r}\sum_{i=1}^{n}\sum_{j=1}^{p} X_{S_{k,l}}^{FN_{i,j}} \qquad (۱-۱)$$

هدف دوم تاخیر است. مقدار تاخیر کل برای هر سرویس IoT به شرح زیر محاسبه می‌شود.

$$D = (d_1 + d_2 + \ldots + d_n) / \lambda_{k,l}^{FN_{i,j},FN_{i',j'}}$$

که در آن D تاخیر، d_1 تاخیر انتقال داده از منبع به گره مه فرعی، d_2 تاخیر پردازش در گره مه فرعی، و d_n تاخیر انتقال داده از گره مه آخر (اصلی) به سمت مقصد است. همچنین $\lambda_{k,l}^{FN_{i,j},FN_{i,j}'}$ نشان دهنده نرخ انتقال داده بین دو گره مه $FN_{i,j}$ و $FN_{i',j'}$ برای مدیریت سرویس $S_{k,l}$ را نشان می‌دهد.

هدف سوم مصرف انرژی است و به محاسبات و انرژی ارتباطی گره‌های مه بستگی دارد. میزان مصرف انرژی محاسباتی یک گره مه به دو بخش ایستا و پویا تقسیم می‌شود. مصرف انرژی ایستا بر اساس زمانی که گره مه فعال است تعیین می‌شود. همچنین مصرف انرژی پویا با توجه به میزان منابع مورد نیاز خدمات اینترنت اشیا مستقر محاسبه خواهد شد. کل انرژی مصرفی گره مه $FN_{i,j}$ در بازه زمانی Δt به‌صورت زیر محاسبه می‌شود.

$$E_{FN_{i,j}}(\Delta t) = ES_{FN_{i,j}}(\Delta t) + ED_{FN_{i,j}}(\Delta t)$$

$$E_{FN_{i,j}}(\Delta t) = \begin{cases} \alpha & if\ FN_{i,j}\ is\ active \\ 0 & otherwise \end{cases}$$

مصرف انرژی پویای یک گره مه به‌صورت یک معادله خطی بر اساس استفاده از CPU گره مه و حداکثر مصرف انرژی پویای به شرح زیر تعیین می‌شود.

$$ED_{FN_{i,j}}(\Delta t) = FN_{i,j}^{CPU}(\Delta t).MDE_{FN_{i,j}}$$

که $MDE_{FN_{i,j}}$ نشان دهنده حداکثر مصرف انرژی پویای گره مه $FN_{i,j}$ در دامنه مه F_i است. اگر میزان استفاده از پردازنده گره مه ۱۰۰ درصد باشد، میزان مصرف انرژی پویا حداکثر خواهد بود، در غیر

این صورت، بر اساس میزان استفاده از پردازنده در بازه زمانی جاری محاسبه می‌شود.

بنابراین، مجموع انرژی مصرفی محاسباتی تمام گره‌های مه شامل انرژی مصرفی ایستای و پویای است که با رابطه (۲-۱) محاسبه می‌شود.

$$TE_{comp}(\Delta t) = \sum_{i=1}^{n}\sum_{j=1}^{p} E_{FN_{i,j}}(\Delta t)\ ;\ \forall_{i,j}\ ;\ \forall\ \Delta t \qquad (۱-۲)$$

ضمنا مصرف انرژی ارتباطی با تاخیر سرویس و موقعیت‌های جغرافیایی بین دو گره مه به شرح زیر مشخص می‌شود.

$$EC_{FN_{i,j}}(\Delta t) = \sum_{k=1}^{m}\sum_{l=1}^{r} X_{S_{k,l}}^{FN_{i,j}} \times \lambda_{k,l}^{FN_{i,j},FN_{i',j'}} \times UCE_{FN_{i,j}}\ ;\ \forall_{i,j}$$

که $\lambda_{k,l}^{FN_{i,j},FN_{i,j}'}$ نشان دهنده نرخ انتقال داده بین دو گره مه $FN_{i,j}$ و $FN_{i',j'}$ برای مدیریت سرویس $S_{k,l}$ و $UCE_{FN_{i,j}}$ واحد انرژی ارتباطی مصرف یک گره مه است. بنابراین، کل انرژی مصرفی ارتباط با رابطه (۱-۳) محاسبه می‌شود

(۱-۳)

$$TE_{comm} = \sum_{i=1}^{n}\sum_{j=1}^{p} EC_{FN_{i,j}}(\Delta t)\ ;\ \forall_{i,j}\ ;\ \forall\ \Delta t$$

در نهایت، کل مصرف انرژی به‌عنوان تابع هدف دوم، مجموع انرژی محاسباتی و ارتباطی برای همه گره‌های مه برای استقرار تمام خدمات اینترنت اشیا خواهد بود و با رابطه (۴-۱) محاسبه می‌شود.

$$E = TE_{comm} + TE_{comp} \qquad (۴-۱)$$

تابع تناسب برای قرارگیری خدمات اینترنت اشیا بر روی گره‌های مه مناسب با توجه به الزامات QoS آنها که توسط معادله زیر مشخص شده است، استفاده می‌شود.

$$Minimize: Z = [E + \frac{1}{T} + D]$$

علاوه بر این، باید ترجیحات کاربر اینترنت اشیا را برای استقرار برنامه‌های IoT در زیرساخت مه در نظر بگیریم. برای انجام این کار، فاکتورهای مقیاس‌بندی φ_E و φ_T و φ_D را به‌عنوان پارامترهای وزنی معرفی می‌کنیم تا ترجیحات کاربر اینترنت اشیا را منعکس کنند. بنابراین، تابع تناسب برای مشکل قرار دادن خدمات اینترنت اشیا با معادله زیر مشخص می‌شود.

$$Minimize: Z = [\varphi_E.E + \varphi_T.\frac{1}{T} + \varphi_D.D]$$

که $\varphi_E + \varphi_T + \varphi_D = 1\ ;\ \varphi_E, \varphi_T, \varphi_D \in [0,1]$. همچنین، از آنجایی که مقادیر تابع هدف دارای واحدهای اندازه گیری مختلفی هستند، باید مقادیر را در بازه [۰، ۱] نرمال کنیم. در این مطالعه، مقادیر تناسب نرمال شده توابع هدف مثبت (یعنی توان عملیاتی) و منفی (یعنی مصرف انرژی) را طبق رابطه (۵-۱) محاسبه می‌کنیم.

(۱-۵)

$$X_N = \frac{X - X_{min}}{X_{max} - X_{min}}$$

که X_N، X_{max} و X_{min} مقادیر نرمال شده، حداکثر و حداقل توابع هدف هستند.

الگوریتم جانمایی خدمات IoT

۱: شروع

۲: برای هر بازه زمانی Δt تا انجام بده

2- [Monitored_inputs] =Monitoring()
3- Sorted_IoTservices_List=Analysis (Monitored_inputs)
4- Placement_Plan=Planning (Sorted_ IoT services_List)
5- Execution (Placement_Plan)

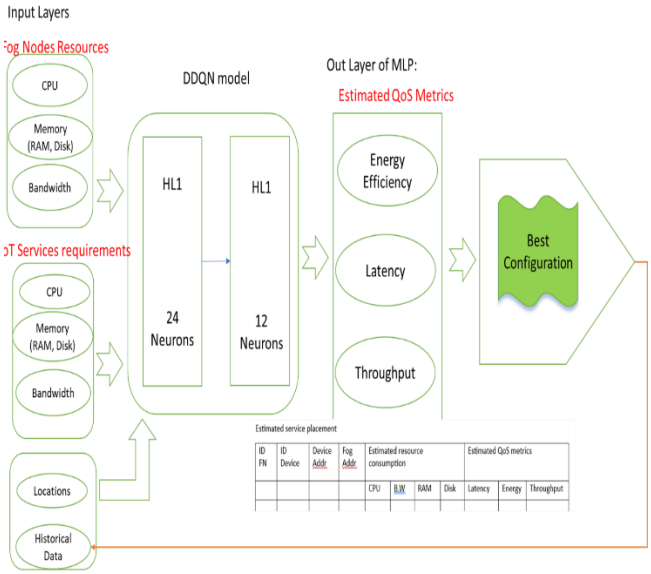
۷: پایان بازه

۸: پایان الگوریتم

۶-۲: مرحله پیش پردازش و تصمیم گیری

در مرحله پیش پردازش، دو نوع اطلاعات جمع‌آوری می‌شود: اول، اطلاعات مورد نیاز در مورد الزامات QoS برنامه‌های کاربردی اینترنت اشیا که توسط ماژول پذیرش ارسال می‌شود؛ و دوم، اطلاعات مورد نیاز در مورد قابلیت‌های محاسباتی منابع گره‌های مه موجود در دامنه‌های مه. مرحله تصمیم‌گیری به‌عنوان مؤلفه اصلی گره کنترل مه در دامنه مه اصلی، راه‌حل کارآمد جانمایی سرویس مه را با استفاده از MLP-WOA به‌عنوان یک رویکرد نوین برای استقرار خدمات اینترنت اشیا بر روی گره‌های مه با انواع مختلف قابلیت‌های منابع در حالی که نیازهای QoS مختلف خدمات اینترنت اشیا را برآورده می‌کند، پیدا می‌کند. در واقع، این مرحله تعیین می‌کند که سرویس اینترنت اشیا درخواست شده در کدام حوزه مه و کدام گره مه تحت آن دامنه مه کنترل می‌شود، مستقر شود.روندکلی این مرحله

درشکل ۲ نشان داده شده است. ورودی مدل پیشنهادی مجموعه‌ای از خدمات اینترنت اشیا و گره‌های مه موجود است و همچنین، خروجی نگاشت خدمات اینترنت اشیا به گره‌های مه به‌عنوان طرح جانمایی بهینه در اکوسیستم مه است.

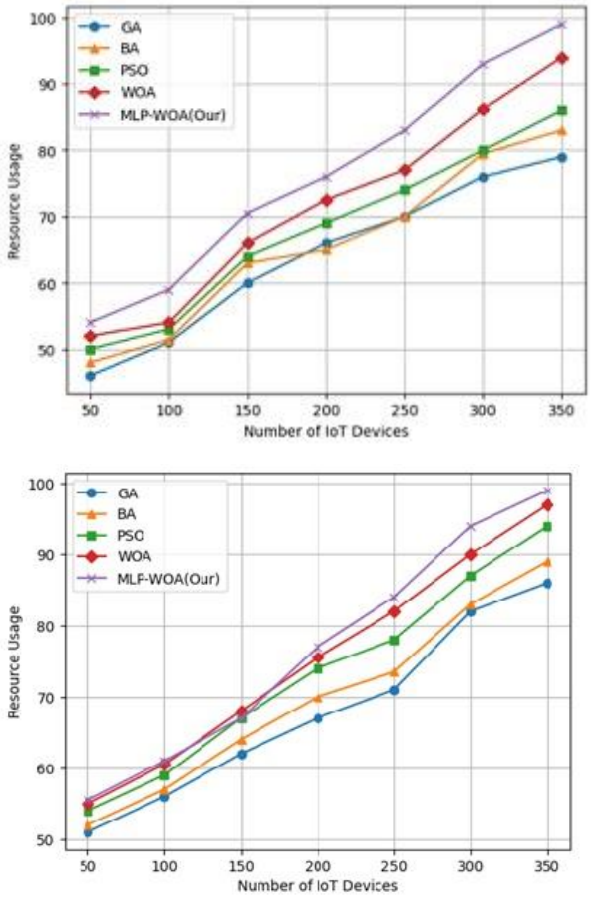


شکل ۲. روند اجرای مدل پیش‌پردازش و برنامه‌ریزی پیشنهادی

۲-۷: ارزیابی معیارهای کیفیت سرویس

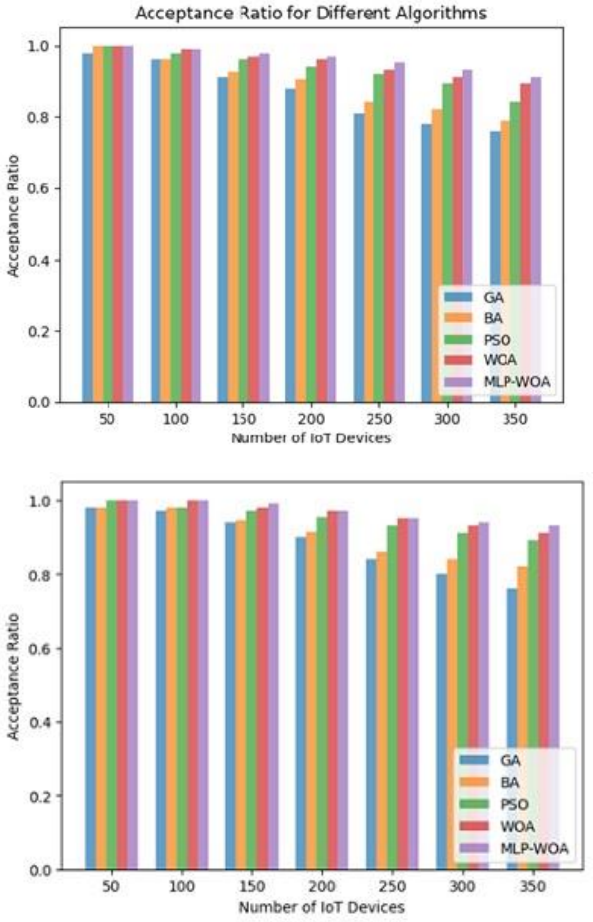
۲-۷-۱: استفاده از منبع

استفاده از منابع گره‌های مه تخصیص داده شده به درخواست‌های IoT دریافت شده از دستگاه‌های اینترنت اشیا تحت ۵ و ۱۰ تعداد انواع خدمات را برای تعداد متفاوتی از خدمات اینترنت اشیا نشان می‌دهد. با افزایش تعداد سرویس‌های اینترنت اشیا، استفاده از منابع برای همه رویکردها افزایش می‌یابد. همانطور که در شکل‌ها مشاهده می‌شود، MLP-WOA قادر است از منابع بیشتری نسبت به سایر روش‌ها استفاده کند، زیرا قادر است از بهینه محلی اجتناب کند و سریع‌تر از سایر رویکردها به راه‌حل بهینه سراسری دست یابد، همچنین تعادل بهتری را بین منافع متعارض به کمک MLP برقرار می‌کند. افزایش تعداد انواع خدمات از ۵ به ۱۰ منجر به افزایش استفاده از منابع می‌شود؛ زیرا انعطاف‌پذیری بیشتری برای تصمیم‌گیری بهتر در محل قرارگیری وجود خواهد داشت.



۲-۷-۲: نرخ پذیرش خدمت

نرخ پذیرش خدمات به‌عنوان نسبت بین تعداد خدمات اینترنت اشیا قرار داده شده در گره‌های مه تعداد کل خدمات اینترنت اشیا تعریف می‌شود. شکل ها مقدار نرخ پذیرش خدمات الگوریتم‌های مختلف را به ازای ۵ و ۱۰ نوع خدمات برای تعداد متفاوتی از خدمات اینترنت اشیا نشان می‌دهد. مشاهده می‌شود که با افزایش تعداد خدمات IoT درخواستی، نرخ پذیرش خدمات کاهش می‌یابد. با این حال، WOA و MLP-WOA نرخ پذیرش بالاتری نسبت به سایر رویکردهای مقایسه شده دارند. با افزایش تعداد انواع خدمات از ۵ به ۱۰، نرخ پذیرش خدمات افزایش می‌یابد، چرا که انعطاف‌پذیری بیشتری برای تصمیم‌گیری بهتر در محل کار وجود خواهد داشت. می‌توان نتیجه گرفت که رویکرد پیشنهادی به دلیل قابلیت جستجوی بهتر و برقراری توازن بهتر، در نتیجه استفاده بهتر از منابع در مقایسه با سایر الگوریتم‌ها، می‌تواند به‌طور موثر برای درخواست‌های خدمات متعدد در سناریوهای محدود به منابع مورد استفاده قرار گیرد.



۳-۷-۲: تاخیر سرویس

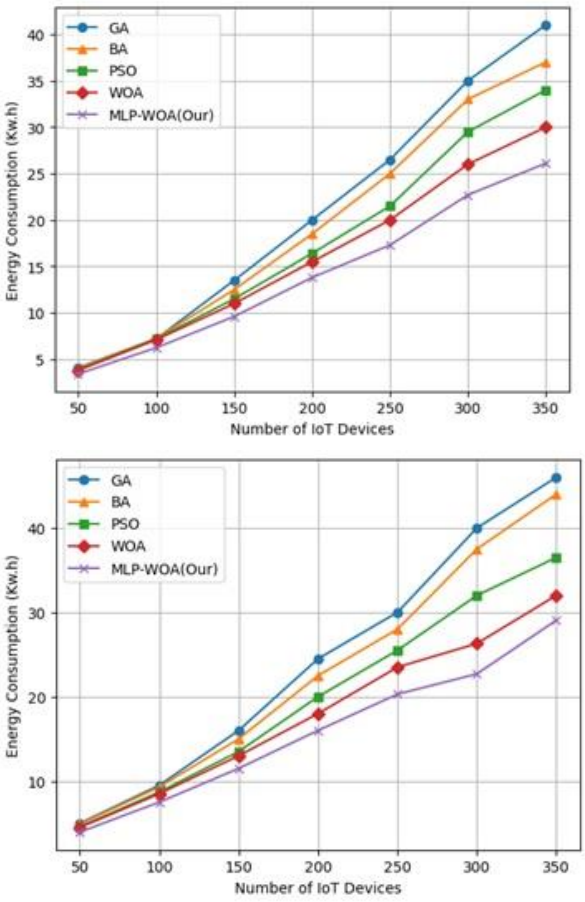
میانگین تاخیر سرویس برای تعداد متفاوتی از خدمات اینترنت اشیا در جدول (۱) نشان داده شده است. تحت پنج نوع سرویس، رویکرد MLP-WOA زمان اجرا را تا ۵.۲٪، ۴.۶٪،۴.۲٪ و ۳.۵٪ در مقایسه با GA، BA، PSO و WOA به ترتیب کاهش می‌دهد. علاوه بر این، برای ۱۰ نوع سرویس، روش پیشنهادی در مقایسه با GA، BA، PSO و WOA به ترتیب تا ۶.۶٪، ۵.۹٪،۴.۶٪ و ۳.۴٪ تاخیر سرویس را کاهش می‌دهد.

جدول۲: میانگین تاخیر سرویس مبتنی بر ۵ و ۱۰ نوع خدمات

تعداد سرویس	WOA [53]	GA [55]	PSO [56]	BA [57]	MLP-WOA
۵ نوع سرویس	1.21	1.58	1.37	1.52	1.11
۱۰ نوع سرویس	0.94	1.34	1.06	1.29	0.846

۴-۷-۲: مصرف انرژی

مصرف انرژی کل، مصرف انرژی مورد نیاز برای استقرار خدمات اینترنت اشیا است که شامل انرژی محاسباتی و ارتباطی برای همه منابع در یک شبکه مه است. شکل ها مصرف انرژی درخواست‌های دستگاه‌های اینترنت اشیا را در زیر ۵ و ۱۰ تعداد نوع خدمات برای تعداد متفاوت خدمات اینترنت اشیا نشان می‌دهد.



همانطور که در این شکل نشان داده شده است، برای هر دوی ۵ و ۱۰ نوع سرویس، WOA از سایر روش‌های جانمایی سرویس بهتر عمل می‌کند. میانگین مصرف انرژی برای تعداد متفاوتی از خدمات اینترنت اشیا در جدول ۲ نشان داده شده است. تحت پنج و ده نوع سرویس، رویکرد MLP-WOA پیشنهادی مصرف انرژی را در مقایسه با GA روش‌های مقایسه شده کاهش می‌دهد.

جدول۳: میانگین مصرف انرژی در ازای ۵ و ۱۰ نوع خدمات

تعداد سرویس	WOA [53]	GA [55]	PSO [56]	BA [57]	MLP-WOA (our)
۵ نوع سرویس	16.17	21.2	17.7	19.5	15.2
۱۰ نوع سرویس	19.2	24.5	20.5	23	17.1

۳- نتیجه‌گیری

رویکرد حاضر (MLP-WOA) با کمک شبکه عصبی MLP و بهینه‌ساز WOA بهترین جانمایی خدمات را در مقابله با رویکردهای پیشین به ارمغان آورده است. حرکت وال‌ها در WOA متناظر با کاوش ترکیب‌های مختلف وزن‌های گره‌های مه است و بر جانمایی خدمات IoT در گره‌های مه تأثیر می‌گذارد. این الگوریتم به دنبال پیدا کردن پیکربندی‌های وزن گره‌های مه است که معیارهای QoS مطلوب را بهینه می‌کنند. MLP به‌عنوان یک ارزیاب هوش مصنوعی عمل می‌کند که معیارهای QoS را بر اساس وزن‌های گره‌های مه پیش‌بینی می‌کند. MLP با ارائه تخمین‌های سریع و دقیق از معیارهای QoS، به الگوریتم کمک می‌کند تا روی راه‌حل‌های ممکن تمرکز کند. فرآیند بهینه‌سازی تکرارپذیر یک حلقه بازخورد بین WOA و MLP ایجاد می‌کند. هرچه WOA راه‌حل‌ها را بهبود می‌دهد، MLP به دلیل آموزش بر روی داده‌های بهبود یافته دقیق‌تر می‌شود. پیش‌بینی‌های بهبودی که MLP ارائه می‌دهد، به WOA کمک می‌کند تا ترکیب‌های بهتری از وزن‌های گره‌های مه را کاوش کند.با استفاده از یک روش خودکار برای استقرار برنامه‌های IoT در زیرساخت مه مجازی، به مشکل جانمایی سرویس پرداختیم. با توجه به تنوع در الزامات QoS خدمات اینترنت اشیا و ناهمگونی منابع و گره‌های مه محدود با منابع در اکوسیستم مه، یافتن یک راه‌حل کارآمد قرارگیری خدمات به‌عنوان یکی از مسائل چالش برانگیز را ضروری می‌کند. از تکنیک فراابتکاری WOA و شبکه عصبی MLP برای استقرار کارآمد برنامه‌های IoT در گره‌های مه موجود استفاده کردیم. رویکرد پیشنهادی تحت سرویس‌های اینترنت اشیا با نیازمندی‌های QoS مختلف و گره‌های مه با قابلیت‌های منابع مختلف طبق نتایج بدست‌آمده نشان می‌دهد که از نظر استفاده از منابع، نرخ پذیرش خدمات، مصرف انرژی و تاخیر خدمات نسبت به سایر مکانیزم‌های مبتنی بر روش فراابتکاری برتری دارد. ارزیابی های انجام شده به خوبی نشان می دهد که روش پیشنهادی در مقایسه با رویکردها و الگوریتم های موجود از منظر انرژی مصرفی حدود ۲ تا ۲.۲ درصد بهبود داشته، از نظر توان عملیاتی نسبت به بهترین عملکرد که مربوط به WOA می باشد حدود ۲ درصد و همچنین از نظر تاخیر حدود ۰.۱ تا ۰.۵ درصد بهبود داشته است. برای کارهای آینده، می‌توان مکانیزم پیشنهادی را با سیستم‌های مبتنی بر زنجیره عمومی گسترش داد تا از حفظ حریم خصوصی در بین برنامه‌های IoT اطمینان حاصل شود و مقیاس‌پذیری رویکرد پیشنهادی تایید شود. همچنین می‌توان از رویکرد یادگیری تقویتی جهت بهبود مدل یادگیری استفاده نمود.

۶- منابع

[1] *J acob B, Lanyon-Hogg R, Nadgir DK, Yassin AF. A practical guide to the IBM autonomic computing toolkit. IBM Redbooks. 2004;4(10):1–268*

[2] *Hazra, A., Rana, P., Adhikari, M., & Amgoth, T. (2023). Fog computing for next-generation internet of things: fundamental, state-of-the-art and research challenges. Computer Science Review, 48, 100549.*

[3] *Lin, C. C., Deng, D. J., Suwatcharachaitiwong, S., & Li, Y. S. (2020). Dynamic weighted fog computing device placement using a bat-inspired algorithm with dynamic local search selection. Mobile Networks and Applications, 25, 1805-1815.*

[4] *Lee, J. H., Chung, S. H., & Kim, W. S. (2019). Fog server deployment technique: An approach based on computing resource usage. International Journal of Distributed Sensor Networks, 15(1), 1550147718823994.*

[5] *Canali, C., & Lancellotti, R. (2019). Gasp: Genetic algorithms for service placement in fog computing systems. Algorithms, 12(10), 201.*

[6] *Djemai, T., Stolf, P., Monteil, T., & Pierson, J. M. (2019, June). A discrete particle swarm optimization approach for energy-efficient IoT services placement over fog infrastructures. In 2019 18th International Symposium on Parallel and Distributed Computing (ISPDC) (pp. 32-40). IEEE.*

[7] *Yousefpour, A., Patil, A., Ishigaki, G., Kim, I., Wang, X., Cankaya, H. C., ... & Jue, J. P. (2019). Fogplan: A lightweight qos-aware dynamic fog*

In Handbook of Research on the IoT, Cloud Computing, and Wireless Network Optimization (pp. 213-234). IGI Global.

[17] *Hong, C. H., & Varghese, B. (2019). Resource management in fog/edge computing: a survey on architectures, infrastructure, and algorithms. ACM Computing Surveys (CSUR), 52(5), 1-37.*

[18] *Zhang, J., Li, C., Yin, Y., Zhang, J., & Grzegorzek, M. (2023). Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. Artificial Intelligence Review, 56(2), 1013-1070.*

[19] *Rana, A., Rawat, A. S., Bijalwan, A., & Bahuguna, H. (2018, August). Application of multi layer (perceptron) artificial neural network in the diagnosis system: a systematic review. In 2018 International conference on research in intelligent and computing in engineering (RICE) (pp. 1-6). IEEE.*

[20] *Chen, Z., Li, X., Wang, W., Li, Y., Shi, L., & Li, Y. (2023). Residual strength prediction of corroded pipelines using multilayer perceptron and modified feedforward neural network. Reliability Engineering & System Safety, 231, 108980.*

[21] *Isabona, J., Imoize, A. L., Ojo, S., Karunwi, O., Kim, Y., Lee, C. C., & Li, C. T. (2022). Development of a multilayer perceptron neural network for optimal predictive modeling in urban microcellular radio environments. Applied Sciences, 12(11), 5713.*

[22] *Maloo, S., Kumar, M., & Lakshmi, N. 2020. A modified whale optimization algorithm based digital image watermarking approach. Sensing and Imaging, 21, 1-22.*

[23] *Mirjalili, S. and Lewis, A., 2016. The whale optimization algorithm. Advances in engineering software, 95, pp.51-67.*

[24] *Gasmi, K., Dilek, S., Tosun, S., & Ozdemir, S. (2022). A survey on computation offloading and service placement in fog computing-based IoT. The Journal of Supercomputing, 78(2), 1983-2014.*

[25] *Peter, N. (2015). Fog computing and its real time applications. Int. J. Emerg. Technol. Adv. Eng, 5(6), 266-269.*

[26] *Salaht, F. A., Desprez, F., & Lebre, A. (2020). An overview of service placement problem in fog and service provisioning framework. IEEE Internet of Things Journal, 6(3), 5080-5096.*

[8] *Rana, A., Rawat, A. S., Bijalwan, A., & Bahuguna, H. (2018, August). Application of multi layer (perceptron) artificial neural network in the diagnosis system: a systematic review. In 2018 International conference on research in intelligent and computing in engineering (RICE) (pp. 1-6). IEEE.*

[9] *Zhang, J., Li, C., [] Yin, Y., Zhang, J., & Grzegorzek, M. (2023). Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. Artificial Intelligence Review, 56(2), 1013-1070.*

[10] *Zhou, X., Li, S., Cao, L., Zhang, H., Jia, Z., Zhong, C., ... & Babar, M. A. (2023). Revisiting the practices and pains of microservice architecture in reality: An industrial inquiry. Journal of Systems and Software, 195, 111521.*

[11] *Roy, P., Sarker, S., Razzaque, M. A., Hassan, M. M., AlQahtani, S. A., Aloi, G., & Fortino, G. (2020). AI-enabled mobile multimedia service instance placement scheme in mobile edge computing. Computer Networks, 182, 107573.*

[12] *Guerrero, C., Lera, I., & Juiz, C. (2019). Evaluation and efficiency comparison of evolutionary algorithms for service placement optimization in fog architectures. Future Generation Computer Systems, 97, 131-144.*

[13] *Mehran, N., Kimovski, D., & Prodan, R. (2019, October). MAPO: a multi-objective model for IoT application placement in a fog environment. In Proceedings of the 9th International Conference on the Internet of Things (pp. 1-8).*

[14] *Chakraborty, S., Saha, A. K., Sharma, S., Chakraborty, R., & Debnath, S. (2023). A hybrid whale optimization algorithm for global optimization. Journal of Ambient Intelligence and Humanized Computing, 14(1), 431-467.*

[15] *Jamil, B., Ijaz, H., Shojafar, M., Munir, K., & Buyya, R. (2022). Resource allocation and task scheduling in fog computing and internet of everything environments: A taxonomy, review, and future directions. ACM Computing Surveys (CSUR), 54(11s), 1-38.*

[16] *Chowdhury, A., Karmakar, G., & Kamruzzaman, J. (2019). The co-evolution of cloud and IoT applications: Recent and future trends.*

software-defined embedded system. IEEE Transactions on Computers, 65(12), 3702-3712.

[36] *Gu, B., Chen, Y., Liao, H., Zhou, Z., & Zhang, D. (2018). A distributed and context-aware task assignment mechanism for collaborative mobile edge computing. Sensors, 18(8), 2423.*

[37] *Yousefpour, A., Ishigaki, G., & Jue, J. P. (2017, June). Fog computing: Towards minimizing delay in the internet of things. In 2017 IEEE international conference on edge computing (EDGE) (pp. 17-24). IEEE.*

[38] *Mahmoud, M. M., Rodrigues, J. J., Saleem, K., Al-Muhtadi, J., Kumar, N., & Korotaev, V. (2018). Towards energy-aware fog-enabled cloud of things for healthcare. Computers & Electrical Engineering, 67, 58-69.*

[39] *Lera, I., Guerrero, C., & Juiz, C. (2018). Availability-aware service placement policy in fog computing based on graph partitions. IEEE Internet of Things Journal, 6(2), 3641-3651.*

[40] *Mahmud, R., Ramamohanarao, K., & Buyya, R. (2018). Latency-aware application module management for fog computing environments. ACM Transactions on Internet Technology (TOIT), 19(1), 1-21.*

[41] *Mebrek, A., Merghem-Boulahia, L., & Esseghir, M. (2017, October). Efficient green solution for a balanced energy consumption and delay in the IoT-Fog-Cloud computing. In 2017 IEEE 16th International Symposium on Network Computing and Applications (NCA) (pp. 1-4). IEEE.*

[42] *Li, G., Liu, Y., Wu, J., Lin, D., & Zhao, S. (2019). Methods of resource scheduling based on optimized fuzzy clustering in fog computing. Sensors, 19(9), 2122.*

[43] *Tang, Z., Zhou, X., Zhang, F., Jia, W., & Zhao, W. (2018). Migration modeling and learning algorithms for containers in fog computing. IEEE Transactions on Services Computing, 12(5), 712-725.*

[44] *Bitam, S., Zeadally, S., & Mellouk, A. (2018). Fog computing job scheduling optimization based on bees swarm. Enterprise Information Systems, 12(4), 373-397.*

[45] *Li, H., Ota, K., & Dong, M. (2019). Deep reinforcement scheduling for mobile crowdsensing in fog computing. ACM Transactions on Internet Technology (TOIT), 19(2), 1-18.*

[46] *Yang, N., Zhang, H., Long, K., Hsieh, H. Y., & Liu, J. (2019). Deep neural network for resource edge computing. ACM Computing Surveys (CSUR), 53(3), 1-35.*

[27] *Ren, J., Zhang, D., He, S., Zhang, Y., & Li, T. (2019). A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. ACM Computing Surveys (CSUR), 52(6), 1-36.*

[28] *Taneja, M., & Davy, A. (2017, May). Resource aware placement of IoT application modules in Fog-Cloud Computing Paradigm. In 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM) (pp. 1222-1228). IEEE.*

[29] *Mahmood, Z., & Ramachandran, M. (2018). Fog computing: Concepts, principles and related paradigms. In Fog computing: concepts, frameworks and technologies (pp. 3-21). Cham: Springer International Publishing.*

[30] *C. Mouradian, D. Naboulsi, S. Yangui, R. H. Glitho, M. J. Morrow and P. A. Polakos, "A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges," in IEEE Communications Surveys & Tutorials, vol. 20, no. 1, pp. 416-464, Firstquarter 2018, doi: 10.1109/COMST.2017.2771153.*

[31] *Xia, Y., Etchevers, X., Letondeur, L., Coupaye, T., & Desprez, F. (2018, April). Combining hardware nodes and software components ordering-based heuristics for optimizing the placement of distributed IoT applications in the fog. In Proceedings of the 33rd annual ACM symposium on applied computing (pp. 751-760).*

[32] *Skarlat, O., Schulte, S., Borkowski, M., & Leitner, P. (2016, November). Resource provisioning for IoT services in the fog. In 2016 IEEE 9th international conference on service-oriented computing and applications (SOCA) (pp. 32-39). IEEE.*

[33] *Daneshfar, N., Pappas, N., Polishchuk, V., & Angelakis, V. (2018, December). Service allocation in a mobile fog infrastructure under availability and qos constraints. In 2018 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.*

[34] *Velasquez, K., Abreu, D. P., Curado, M., & Monteiro, E. (2017). Service placement for latency reduction in the internet of things. Annals of Telecommunications, 72, 105-115.*

[35] *Zeng, D., Gu, L., Guo, S., Cheng, Z., & Yu, S. (2016). Joint optimization of task scheduling and image placement in fog computing supported*

management in NOMA networks. IEEE Transactions on Vehicular Technology, 69(1), 876-886.

[47] *Liu, T., Ni, S., Li, X., Zhu, Y., Kong, L., & Yang, Y. (2022). Deep reinforcement learning based approach for online service placement and computation resource allocation in edge computing. IEEE Transactions on Mobile Computing.*

[48] *Talaat, F. M. (2022). Effective prediction and resource allocation method (EPRAM) in fog computing environment for smart healthcare system. Multimedia Tools and Applications, 81(6), 8235-8258.*

[49] *Mezni, H., Hamoud, F. S., & Charrada, F. B. (2023). Predictive service placement in cloud using deep learning and frequent subgraph mining. Journal of Ambient Intelligence and Humanized Computing, 14(9), 11497-11516.*

[50] *Tuli, S., Casale, G., & Jennings, N. R. (2022). Splitplace: AI augmented splitting and placement of large-scale neural networks in mobile edge environments. IEEE Transactions on Mobile Computing.*

[51] *Lalanda, P., McCann, J.A. and Diaconescu, A., 2013. Autonomic computing: principles, design and*

[52] *implementation. Springer Science & Business Media.*

[53] *Parashar, M. and Hariri, S. eds., 2018. Autonomic computing: concepts, infrastructure, and applications.*

[54] *CRC press.*

[55] *Ghobaei-Arani, M., & Shahidinejad, A. (2022). A cost-efficient IoT service placement approach using whale optimization algorithm in fog computing environment. Expert Systems with Applications, 200, 117012.*

[56] *Gupta, H., Vahid Dastjerdi, A., Ghosh, S. K., & Buyya, R. (2017). iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments. Software: Practice and Experience, 47(9), 1275-1296.*

[57] *Canali, C., & Lancellotti, R. (2019). Gasp: Genetic algorithms for service placement in fog computing systems. Algorithms, 12(10), 201.*

[58] *Roy, P., Sarker, S., Razzaque, M. A., Hassan, M. M., AlQahtani, S. A., Aloi, G., & Fortino, G. (2020). AI-enabled mobile multimedia service instance placement scheme in mobile edge computing. Computer Networks, 182, 107573.*

[59] *Lin, C. C., Deng, D. J., Suwatcharachaitiwong, S., & Li, Y. S. (2020). Dynamic weighted fog computing device placement using a bat-inspired algorithm with dynamic local search selection. Mobile Networks and Applications, 25, 1805-1815.*

سید صادق حسینی دانش آموخته مهندسی فناوری اطلاعات (IT) گرایش شبکه های کامپیوتری دانشگاه آزاد واحد پرند می باشد. مقطع کارشناسی در رشته مهندسی کامپیوتر گرایش نرم افزار در دانشگاه غیر انتفاعی علامه رفیعی قزوین به اتمام رساند.زمینه های موردعلاقه ایشان اینترنت اشیا، رایانش ابری و رایانش مه و یادگیری ماشین می باشد، یک مقاله از پایان نامه او استخراج شده و در حال انجام مقالاتی در کنفرانس ها و مجلات در حوزه اینترنت اشیا، رایانش مه و هوش مصنوعی می باشد.
نشانه رایانامه ایشان عبارتند از:
sadeghhosseini88@outlook.com

محمد فرجی مهماندار مدرک کارشناسی خود را از دانشگاه خوارزمی ایران در سال ۱۳۷۸ در رشته مهندسی کامپیوتر و مدرک کارشناسی ارشد در رشته علوم کامپیوتر از دانشگاه تهران در سال ۱۳۸۱دریافت کرده است. وی همچنین مدرک دکتری خود را در رشته مهندسی کامپیوتر گرایش سیستم های نرم افزاری از دانشگاه آزاد اسلامی (IAU) دریافت کرده است. استادیار گروه مهندسی کامپیوتر دانشگاه آزاد اسلامی واحد پرند بوده و علایق تحقیقاتی او شامل رایانش مه، رایانش محاسبات توزیع شده، اینترنت اشیا، داده کاوی و هوش مصنوعی است.
نشانه رایانامه ایشان عبارتند از:
mm.faraji@iau.ac.ir

روش ارجاع: ص. حسینی، م. فرجی مهماندار، بهینه‌سازی جانمایی چندمنظوره سرویس‌های اینترنت اشیا در رایانش ابری: با استفاده از الگوریتم نهنگ و شبکه عصبی MLP، دوفصلنامه محاسبات و سامانه‌های توزیع شده، سال هفتم، شماره۱، شماره پیاپی۱۳، صفحه ۱۲۴ تا ۱۳۶، سال ۱۴۰۳.

How to cite: S. Hosseini, M. Faraji Mehmandar, Multi-objective optimization of IoT service placement in cloud computig : Using Whale Optimization Algorithm And MLP Neural Network Journal of Distributed Computing and Systems (JDCS), Vol 7, Issue 1, Pages 124-136, 2024.

"Multi-objective Optimization of IoT Service Placement in Cloud Computing Using Whale Optimization Algorithm and MLP Neural Network"

S.Hosseini[1], M.Faraji Mehmandar[2],
[1]Azad University, Parand Branch,Tehran, Iran
[2]Azad University, Parand Branch,Tehran, Iran

Abstract

With the rapid growth of Internet of Things (IoT) applications, the use of this technology to facilitate daily activities across various fields has significantly increased and contributed to improving the quality of human life. In this context, fog computing has emerged as a distributed computational solution to support these applications using fog nodes located close to IoT devices. IoT-related programs have been developed in the form of multiple IoT services with different Quality of Service (QoS) requirements, which can be deployed on fog nodes. Therefore, identifying a suitable layout scheme for service placement and optimizing the capabilities of various resources in the fog ecosystem is one of the significant challenges that must be considered.

This research presents an efficient solution for IoT service placement based on a Multi-Layer Perceptron (MLP) neural network and the Whale Optimization Algorithm (WOA), designed for deploying IoT applications within fog infrastructure. The proposed solution examines the QoS requirements of IoT services and the capabilities of existing fog nodes to determine an efficient service placement. Initially, the placement scheme is estimated using the MLP model and then optimized using heuristic optimization methods. In this approach, parameters such as throughput, energy consumption, and delay are used as objective functions to achieve an optimal service placement scheme. Simulation results indicate that this proposed solution increases resource utilization and service acceptance rates while reducing service delay and energy consumption compared to other heuristic methods.