فصل اول

**کلیات تحقیق**

**اهداف فصل:**

* توصیف محیط محاسبات لبه و مه
* معرفی چالش‌های مهم در محاسبات لبه و مه
* معرفی مسئله زمانبندی در محاسبات لبه و مه محور و اهمیت آن
* معرفی ساختار پایان‌نامه

پژوهش پیشنهادی در پنج مرحله اجرا می شود. در فاز اول، بیان مسئله، مفاهیم محاسبات ابری، اینترنت اشیا و ساختار مه و لبه در آن مورد بررسی قرار گرفته در ادامه مسئله زمانبندی در این ساختارها به عنوان چالش مطالعه می شود. در بخش دوم از پژوهش، به بررسی و ارزیابی مفاهیم مربوط به زمانبندی در محیط مه و لبه، الگوها و راهکارهای ارائه شده توسط محققان در این زمینه پرداخته می شود و در ادامه مزایا و معایب هریک از رویکردهای بررسی شده توصف می شود. در فصل سوم، روش پیشنهادی توصیف شده و ساختار آن به صورت کلی مورد ارزیابی قرار می­گیرد. در در فصل چهارم، شبیه سازی و ارزیابی روش پیشنهادی اجرا شده و نتایج به دست آمده از شبیه سازی، با رویکردهای موجود در این زمینه مقایسه می شود. و در نهایت در فصل پنجم جمع بندی رویکرد درنظر گرفته انجام و پیشنهاداتی برای ادامه و بهبود تحقیق و فرایند تخصیص منابع و زمانبندی جهت آیندگان ارائه می شود.

# 2-1) بیان مسئله

در چند سال گذشته شاهد رشد سریع صنعت اینترنت اشیا [[1]](#footnote-1) بوده ایم که امکان اتصال مردم به اشیا و ساختارهای هوشمند فیزیکی را فراهم کرده و روند و دیجیتالی شدن دنیای فیزیکی را تسهیل می کند.

رایانش ابری با استفاده از قابلیت های محاسباتی قدرتمند و فناوری های ذخیره سازی پیشرفته، امنیت و قابلیت اطمینان اطلاعات ذخیره شده را تضمین می کند. با این حال، سرورها در ساختار رایانش ابری معمولاً در فاصله فیزیکی طولانی از دستگاه های اینترنت اشیا قرار دارند و تأخیر زیاد ناشی از فواصل طولانی نمی تواند به طور مؤثر نیازهای مربوط به برنامه های اینترنت اشیا با نیازهای پردازش آنی را برآورده کند[1]. به دلیل این مسائل، محاسبات لبه و مه به عنوان یک ساختار محاسباتی محبوب در زمینه اینترنت اشیا پدیدار شده اند.

ساختاری که از منابع لبه برای برنامه های کاربردی اینترنت اشیا به صورت پردازش بلادرنگ استفاده می کند، به عنوان محاسبات لبه در نظر گرفته شده و ساختاری که از لبه استفاده می کند و هر زمان که لازم باشد نیز از منابع ابری استفاده کند، به عنوان محاسبات مه در نظر گرفته شده است.

به عنوان مثال، در زمینه مراقبت های بهداشتی هوشمند، استقرار گره های محاسباتی لبه بر روی دستگاه های پزشکی می تواند پارامترهای فیزیولوژیکی بیماران را در زمان واقعی پایش کرده و داده ها را برای تجزیه و تحلیل و تشخیص، تحقق پزشکی از راه دور و پزشکی شخصی سازی شده به ابر منتقل کند.

که برنامه ها باید توسط بهترین سرور (یعنی قوی ترین و از نظر فیزیکی نزدیک ترین سرور) پردازش شوند. علاوه بر این، بار سیستم باید به طور ایده آل متعادل سازی و توزیع شود تا بر روی چندین واحد عملیاتی اجرا شود[1]. به عنوان مثال، با توزیع درخواست ها در چندین سرور به روشی یکپارچه (مانند محیط های محاسباتی بدون سرور)، تعادل بار می تواند از بارگذاری بیش از حد سرورهای موجود در شبکه جلوگیری کند و اطمینان حاصل کند که هر سرور بار پردازشی متوسطی را مدیریت می کند. این امر زمان پاسخ، عملکرد کلی سیستم و توان عملیاتی را بهبود می بخشد و همچنین به سرورها کمک می کند تا با ثبات تر کار کنند. بنابراین، بهبود سطح توازن بار سرورها (به عنوان مثال، کاهش حجم استفاده از منابع سرور) و در عین حال کاهش زمان پاسخ به یک مسئله مهم اما چالش برانگیز برای زمان بندی برنامه های اینترنت اشیا در سرورها در محیط های محاسبه لبه / مه تبدیل شده است[3].

پیچیدگی محاسبات لبه / مه منجر به چالش در تکمیل وظایف در یک چارچوب زمانی معین به دلیل حجم زیادی از داده هایی که باید پردازش شوند، می شود. بنابراین، زمان بندی وظایف نقش بسیار مهمی دارد. الگوریتم های زمان بندی با چالش زیادی برای مدیریت چنین وظایفی روبه رو هستند، زیرا تأخیر در چنین وظایفی غیرقابل قبول است

وظایف در محیط مه به دو دسته، داده فشرده یا محاسبات فشرده طبقه بندی می شوند. در طول زمان بندی، از منابع محاسباتی با کارایی بالا برای محاسبات فشرده و برای انتقال داده ها از منابع با کارایی پایین استفاده می شود که منجر به کاهش زمان اجرای کار می شود. تعداد انتقال داده ها در طول زمان بندی وظایفی که به شدت داده نیاز دارند به حداقل می رسد. این باعث کاهش زمان انتقال داده می شود

محققان از ساختارهایی با عنوان یادگیری ماشین و یادگیری عمیق به منظور حل مسئله زمانبندی در محیط لبه و مه استفاده کرده اند، در یادگیری عمیق، عامل به طور مداوم با محیط در تعامل است و تعداد زیادی از مسیرهای تجربی (به عنوان مثال، توالی حالت ها، اقدامات و پاداش ها) را ثبت می کند که در مرحله آموزش برای یادگیری سیاست های بهینه استفاده می شود.

ایجاد تعادل بین سادگی پیاده سازی، پیچیدگی نمونه و عملکرد راه حل، به یک چالش تحقیقاتی کلیدی در استفاده از یادگیری عمیق در محیط های محاسباتی لبه و مه در موقعیت های پیچیده است

برای مدیریت چالش ها، در این پژوهش، یک رویکرد زمانبندی کاربردی اینترنت اشیا مبتنی بر یادگیری عمیق پیشنهاد شد که از ساختار شبکه عصبی حافظه طولانی کوتاه مدت به منظور طبقه بندی منابع پردازشی و اجرای فرایند زمانبندی استفاده می کند.

# 3-1) اهمیت و ضرورت تحقیق

تکامل دستگاه های اینترنت اشیا منجر به تولید حجم عظیمی از داده ها شده است که نیاز به نظارت، پردازش و تجزیه و تحلیل دارند. ابر به مدلی شایسته برای اجرای برنامه های کاربردی کاربر تبدیل شده است در پردازش برنامه ها در مرکز داده ابری به دلیل موقعیت جغرافیایی دور از کاربران باعث تأخیر است است برای برنامه هایی با زمان پاسخ دهی سخت گیرانه، مانند مراقبت های بهداشتی هوشمند غیرقابل قبول می باشد.

معماری فقط ابری ممکن است برای برنامه های حساس به تأخیر مناسب نباشد، به دلیل مسافتی که در بالا مشخص شده است. محاسبات مه یک معماری متشکل از تعدادی دستگاه مه / مراکز داده میکرو، در مجاورت منابع تولید داده، مانند گوشی های هوشمند و حسگرها، ارائه می دهد. از این رو، داده های تولید شده را می توان در دستگاه های مه به موقع پردازش کرد[8], [9].

دستگاه های مه دارای قدرت محاسباتی و ظرفیت ذخیره سازی محدودی هستند. معماری کلی محاسبات مه را می توان به سه لایه تقسیم کرد. اولین لایه، لایه دستگاه های IoT است که شامل انواع مختلف دستگاه ها، مانند گوشی های هوشمند، وسایل نقلیه هوشمند، رایانه های لوحی و دستگاه های مختلف خانه هوشمند است. این لایه می تواند محیط اطراف را حس کرده و از طریق دستگاه های حسگر داده ها را جمع آوری کند و از طریق فناوری های 3G، 6G، 5G، WiFi و بلوتوث با لایه محاسباتی مه ارتباط برقرار کند. لایه دوم لایه میانی، لایه محاسباتی مه است که شامل مسیریاب ها، دروازه ها، ایستگاه های کاری، سوئیچ ها، نقاط دسترسی است. این لایه قابلیت محاسبات، شبکه و ذخیره سازی را دارد. در نهایت، لایه بالایی محاسبات ابری است که شامل سرورهای ابری با قدرت محاسباتی بالا است.

زمان بندی چالش اصلی در محاسبات مه است که ضرورت های خاص خود در این ساختارها را دارد،

# 1-3-1) اهمیت تحقیق

در محیط مه، وظایف به دو گروه تقسیم می شوند: وظایفی که به شدت محاسبات نیاز دارند و وظایفی که به شدت داده نیاز دارند. در حالی که زمان بندی کارهایی که نیاز به شدت محاسبات دارند، زمان بندی داده ها را به منبع بهره وری بالا منتقل می کند و از این رو، زمان اجرای کار کاهش می یابد. از سوی دیگر، ضمن زمان بندی وظایفی که نیاز به شدت داده دارند، سعی می شود تعداد مهاجرت داده ها کاهش یابد. در نتیجه زمان انتقال داده کاهش می یابد.

یک تکنیک کارآمد برای برنامه ریزی وظایف و مدیریت منابع محیط های مه و ابر نیاز دارد.

به طور کلی مشکلات زمان بندی نیز به پنج دسته اصلی زمان بندی کار، زمان بندی منابع، تخصیص منابع، زمان بندی تخصیص کار، و زمان بندی گردش کار تقسیم شده اند.

به طور کلی، اهمیت به کارگیری ساختارهای زمانبندی منابع در محیط لبه و مه را می توان از زوایای متفاوت مانند کاهش تأخیر، بهبود مصرف انرژی، هزینه استقرار، بهبود دقت و پردازش چند هدفه بررسی کرد. به طور کلی در اجرای کار به حداقل رساندن تأخیر یکی از اهداف اصلی در طول زمانبندی منابع است صرف نظر از نوع زمانبندی منابع، هدف کلی به کاهش کل تأخیر اجرای کار محدود می شود.

دومین هدف متداول در هنگام زمانبندی منابع، به حداقل رساندن مصرف انرژی است. مصرف انرژی معمولاً به دستگاه های پایانی اشاره دارد. دلیل آن این است که دستگاه های سیار و اینترنت اشیا معمولاً با باتری کار می کنند، بنابراین یک نگرانی عمده این است که چگونه می توان طول عمر باتری را با کاهش مصرف انرژی دستگاه به حداکثر رساند.

# 2-3-1) ضرورت تحقیق

مزایای زمانبندی در محیط لبه و ابر بسیار زیاد است، زیرا این رویکردها امکان افزایش کیفیت سرویس در برنامه ها را فراهم می کنند و در عین حال عمر باتری دستگاه های پایانی را افزایش می دهند.

زمانبندی منابع در محیط هنوز با چالش های قابل توجهی در سناریوی محاسبات ابری مواجه است. به عنوان مثال، تأخیر اضافی انتقال و مصرف انرژی زمانی رخ می دهد که تعداد کارهای بارگذاری زیاد باشد،

# 4-1) جنبه نوآوری پژوهش

محاسبات لبه / مه، به عنوان یک الگوی محاسباتی توزیع شده، الزاماتی شامل تأخیر پایین و قابلیت پردازش بلادرنگ را برای تعداد روزافزونی از برنامه های اینترنت اشیا برآورده می کند

تعداد زیادی از برنامه های اینترنت اشیا نیاز به اجرا در منابع لبه / مه دارند، ممکن است سرورها بیش از حد بارگذاری شوند. از این رو، ممکن است سرورهای لبه / مه را مختل کند

محیط های محاسباتی لبه / مه و برنامه های اینترنت اشیا ذاتاً پویا و تصادفی هستند.

منابع محاسباتی محدود سرورهای لبه و مه بار اضافی را برای اعمال تکنیک های بهینه اما محاسباتی تحمیل می کنند. برای غلبه بر این چالش ها، در این پژوهش یک رویکرد زمانبندی بهینه مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی LSTM پیشنهاد می شود تا زمان پاسخ را به طور سازگار و کارآمد بهینه کند.

به طور کلی، جنبه نوآوری پژوهش به شرح زیر است:

* در این پژوهش از یک رویکرد پیش پردازش و حذف افزونگی اطلاعات مبتنی بر مدل پیرسون به منظور مدیریت منابع با استفاده از پارامترهای کیفیت سرویس ارائه می شود. هنگامی که محیط محاسباتی تغییر می کند (به عنوان مثال، درخواست از برنامه های مختلف اینترنت اشیا، منابع محاسباتی سرور، تعداد سرورها)، این راهکار می تواند به طور تطبیقی خط مشی مربوط به پارامترهای کیفیت سرویس و زمان بندی را با سرعت همگرایی بالا به روز کند.
* در ادامه به منظور زمانبندی وظایف، یک مدل هزینه مبتنی بر شبکه عصبی LSTM در محیط های محاسباتی لبه / مه پیشنهاد می شود تا سطح تعادل بار سرورها را بهبود بخشد و زمان پاسخ برنامه را به حداقل برساند.

# ) اهداف پژوهش

* بهینه‌سازی تعادل بار سیستم در محیط‌های محاسباتی لبه و مه
* کمینه‌سازی زمان پاسخ سیستم در محیط‌های محاسباتی لبه و مه
* بهبود سربار محاسباتی در محیط‌های محاسباتی لبه و مه

# 6-1) سؤالات پژوهش

* راهکار زمانبندی پیشنهادی با استفاده از چه فرایندی باعث بهینه‌سازی تعادل بار سیستم در محیط‌های محاسباتی لبه و مه می شود؟
* رویکرد پیشنهادی زمانبندی با استفاده از چه راهکاری باعث بهینه‌سازی زمان پاسخ سیستم در محیط‌های محاسباتی لبه و مه می شود؟
* راهکار زمانبندی پیشنهادی با استفاده از چه فرایندی موجب بهینه‌سازی باعث بهبود سربار محاسباتی در محیط‌های محاسباتی لبه و مه می شود؟
* رویکرد پیشنهادی زمانبندی با استفاده از چه راهکاری باعث بهبود نرخ همگرایی در محیط‌های محاسباتی لبه و مه می شود؟

فصل دوم

**پیشینه پژوهش**

**اهداف فصل:**

* توصیف ادبیات پژوهش
* توصیف راهکارهای موجود و پیشینه پژوهش
* مقایسه روش‌های موجود

# 1-2) مقدمه

این فصل جهت آشنایی بیشتر با روش های ایجاد تعادل بار در شبکه های نرم افزار محور تنظیم شده است. در ابتدا روش های موجود مورد مطالعه و بررسی قرار می گیرند. سپس، مزایا و معایب نمونه هایی از آن ها بررسی شده و با یکدیگر مقایسه می شوند.

# 1-2-2) فرایندهای زمانبندی در محاسبات مه

از آنجایی که قرار است گره های مه در نزدیکی لایه دستگاه قرار گیرند، درخواست های دارای الزامات تأخیر بالا را می توان در لایه مه پردازش کرد و سایر درخواست ها را می توان به ابر ارسال کرد.

لایه مه، درخواست ها و داده های دریافتی را فیلتر، تجزیه و تحلیل و پیش پردازش می کند. اگر لایه مه منابع کافی در دسترس داشته باشد و بتواند وظیفه ای را اجرا کند، کار در لایه مه برنامه ریزی می شود. در غیر این صورت، اگر منابع کافی برای اجرای کار وجود نداشته باشد، به لایه ابری ارسال می شود.

اگر هیچ یک از آنها (یعنی لبه و مه) نتوانند در زمان مشخص شده پاسخ دهند، کار رد می شود.

لایه مه ممکن است به طور همگن یا ناهمگن سازماندهی شود[14].

در ساختار همگن، منابع مه دارای قابلیت های پردازش، ذخیره سازی و پهنای باند یکسانی هستند، اما در حالت ناهمگن، منابع دارای قابلیت های متفاوتی هستند. در مورد دستگاه های مه ناهمگن، زمان بندی وظایف از پیچیدگی بیشتری برخوردار است، زیرا منابع دارای قابلیت های متفاوتی از نظر سرعت پردازش، جابه جایی و غیره هستند که منجر به ترکیب های مختلفی (وظیفه و منبع) می شود.

هدف محاسبات مه کاهش تأخیر شبکه با ارائه منابع مورد نیاز در لبه شبکه نزدیک به دستگاه های پایانی است، بنابراین به کاهش تأخیر، بهبود کیفیت خدمات و تحرک پشتیبانی کمک می کند. با این حال، منابع مه محدود هستند و نمی توانند حجم عظیمی از داده های تولید شده توسط دستگاه های IoT را به صورت محلی پردازش کنند. بنابراین، انتخاب مؤثر پردازش وظایف بین سرور ابری و سرور مه برای برآوردن نیازهای روزافزون و سفارشی برنامه های اینترنت اشیا مهم است..

سه جزء اصلی در یک موتور اجرای جریان کار، زمانبندی جریان کار، اجرای جریان کار انتقال داده ها و مدیریت خطاها هست. زمانبند جریان کار منابع را کشف کرده و وظایف را به منابع مناسب اختصاص می دهد. انتقال داده ها برای ارتباط منابع داده و مدیریت خطا برای مدیریت خطاها در طی اجرا به کار می رود.

زمانبندی در محیط لبه و مه به معنی انتخاب بهترین منبع برای یک وظیفه است یا اختصاص ماشین های مجازی به وظایف به گونه ای که زمان اجرای کل به حداقل برسد. به طور کلی در الگوریتم های زمانبندی لیستی از وظایف وجود دارد که به ترتیب اولویت برای همه ی وظایف ساخته شده است . وظایف بر اساس اولویت انتخاب می شوند و به پردازنده ها تخصیص داده شده

روش های مختلفی برای مشکل زمانبندی در سیستم های لبه و مه شده وجود دارد که در ادامه تعدادی از این رویکردها ارائه می شود[1]:

* **الگوریتم های زمانبدی ایستا و پویا:** دو نوع الگوریتم زمانبندی با نامهای ایستا و پویا وجود دارد. در زمان بندهای ایستا اطلاعات لازم در مورد همه منابع در لحظه ی زمانبندی موجود است به همین دلیل زمان اجرا را می توان تخمین زد، در مقابل آن در زمان بندهای پویا ایده اصلی تخصیص وظایف در زمان اجرا می باشد.
* **الگوریتم های زمانبندی سراسری و محلی** : در زمانبندی محلی در مورد تخصیص فرایندها به یک پردازنده ولی در زمانبندی سراسری تخصیص فرایندها به چندین پردازنده تصمیم گیری می شود.
* **الگوریتم های توزیع شده و متمرکز**: در زمانبندی توزیع شده، مدیر مرکزیوجود ندارد که مسئول اجرای باشد و بلکه هر سیستم بر اساس الگوریتم توزیع شده به فعالیت می پردازد، ولی در زمان بندیهای متمرکز یک کنترل مرکزی برای تخصیص تمامی وظایف وجود دارد.
* **الگوریتم های زمانبندی تقریبی و اکتشافی** : الگوریتم های تقریبی راه حل هایی شبه بهینه همراه با ضریب برای میزان تقویت جواب واقعی ارائه می دهند ولی الگوریتم های اکتشافی با فرض های واقعی در مورد سیستم و جست و جوی فضای مسئله ، سعی در پیدا کردن جوابی نزدیک به جواب بهینه را دارد.
* **الگوریتم های زمانبندی همکار و غیر همکار:** الگوریتم های زمانبندی همکار با اجزای توزیع شده سروکار دارند و در ارتباط هستند و در زمان بندیهای غیر همکار هر پردازنده مستقل از سایر پردازنده ها است
* **الگوریتم های زمانبندی بهینه و نیمه بهینه:** روش بهینه سعی دارد هدف را به بهترین شکل برآورده سازد در صورتی که روش های نیمه بهینه تنها به یک هدف نمی پردازد بلکه سعی بر بهبود نسبی چندین هدف را به صورت همزمان دارند.

# 2-2-2) چالش‌های محاسبات مه و لبه

# 3-2) پیشینه پژوهش

# 1-3-2) زمان‌بندی کار عظیم انرژی و تأخیری در سیستم محاسباتی لبه و مه

در سیستم های محاسباتی لبه و مه، داده ها بسیار فشرده هستند و بر روی نزدیک ترین دستگاه های مه یا پلتفرم های ابری توزیع شده اجرا می شوند

که در آن هر کار زمان کوتاهی (چند میلی ثانیه) برای اجرا نیاز دارد. زمان اجرای یک کار در دستگاه مه یا پلتفرم ابری به تعداد کارهایی که در همان دستگاه در صف هستند بستگی دارد.

وظایف در ابتدا توسط نزدیک ترین دستگاه های مه شکن جمع آوری می شوند و سپس برخی از آنها به دلیل ظرفیت محدود دستگاه های محاسباتی مه به پلتفرم های ابری مناسب تخلیه می شوند. از آنجایی که این وظایف داده فشرده هستند، تأخیرهای ارتباطی از مه به ابر قابل چشم پوشی نیست.

# 2-3-2) زمان‌بندی وظایف در سیستم‌های محاسباتی مه ابری

. این برنامه ها دارای مهلت های پردازشی مرتبط هستند و اعتبار خروجی آنها به زمانی که در دسترس می شوند بستگی دارد. اگر یک برنامه بلادرنگ نتواند ضرب الاجل خود را برآورده کند،

به طور کلی چندین مسئله در شبکه های اینترنت اشیا و مه ابری در اصل مسائل بهینه سازی هستند و یکی از این موارد، مسئله زمانبندی است. هنگامی که یک برنامه کاربردی به لبه شبکه می رسد، وظایفی که این برنامه کاربردی را تشکیل می دهند باید بر روی عناصر پردازشی برنامه ریزی شوند. زمانبند مسئول نگاشت وظایف یک برنامه کاربردی بر روی منابع موجود است تا بتوان به اهداف خاصی دست یافت.

# 3-3-2) زمان‌بندی وظایف چند هدفه در محاسبات ابری و مه با استفاده از رویکرد برنامه نویسی هدف

رویکرد محاسبات مه نقش واسطه ای را در اتصال دستگاه های اینترنت اشیا به مرکز ابر ایفا می کند

دستگاه های اینترنت اشیا در یک شبکه محلی به گره های مه متصل می شوند و تمام اتصالات آنها به مرکز ابر از طریق گره های مه برقرار می شود.

گره های مه می توانند تمام اطلاعات نزدیک اینترنت اشیا را دریافت کنند و سپس فقط اطلاعات قابل استفاده در مورد دستگاه های اینترنت اشیا را به ابر ارسال کنند و از این طریق ترافیک شبکه را کاهش داده و عملکرد کلی شبکه را افزایش دهند. از سوی دیگر، دستگاه های اینترنت اشیا اغلب قدرت پردازش ضعیفی دارند و می توانند از قدرت پردازش گره های مه استفاده کنند،

# 4-3-2) مشخص کردن زمان‌بندی برنامه ها در منابع محاسباتی لبه، مه و ابر

# 5-3-2) زمان‌بندی کار آگاه از هزینه در محیط مه-ابر

# 4-2) مقایسه روش‌های موجود

رویکردهایی در زمینه بهبود فرایند زمانبندی در محیط محاسباتی مه و لبه مورد بررسی قرار گرفتند و نقاط قوت و ضعف آنها ارائه شد. در ادامه و در جدول (2-1) مقایسه کلی از رویکردهای بررسی شده ارائه می شود و با درنظر گرفتن نتایج مقایسه، پیشنهاد کار آتی در فصل بعدی بیان می شود.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **مؤلف** | **روش پیشنهادی** | **سرعت** | **تأخیر** | **مصرف انرژی** | **کیفیت سرویس** | **طول عمر شبکه** |
| **وانگ و همکارانش[2]** | یک الگوریتم زمانبندی مبتنی بر یادگیری تقویتی را پیشنهاد کردند که تا زمان پاسخ‌دهی برنامه های ناهمگن اینترنت اشیا را به‌طور سازگار و کارآمد بهینه کند و بار سرورهای لبه / مه را متعادل کند. | بالا | پایین | متوسط | پایین | متوسط |
| **هوسین و همکارانش[21]** | یک تکنیک زمان‌بندی کار به نام زمان‌بندی وظایف اولویت‌دار آگاه از منابع را در یک محیط محاسباتی مه ناهمگن پیشنهاد کردند. | متوسط | بالا | پایین | متوسط | متوسط |

# 5-2) خلاصه و نتیجه گیری

در این فصل از پژوهش، یک مرور کلی از رویکردهای زمابندی در معماری لبه و مه به همراه مزایا و معایب هر یک از رویکردهای بررسی شده ارائه شد. با توجه به مطالب بیان‌شده، در فصل آتی ضمن ارائه جمع بندی کلی، پیشنهادات آتی در این زمینه ارائه می شود.

3) فصل سوم

**راهکار پیشنهادی**

**اهداف فصل:**

* توصیف رهیافت و روش پژوهش
* ارائه طرح کلی و مدل مفهومی پژوهش
* ارائه مدل پیشنهادی پژوهش

# 1-3) مقدمه

محققان برای حل مسئله زمانبندی در محاسبات ابری و ساختارهای مه، از ساختارهایی با عنوان یادگیری ماشین و یادگیری عمیق به منظور حل مسئله زمانبندی در محیط لبه و مه استفاده کرده اند، در یادگیری عمیق، عامل به طور مداوم با محیط در تعامل است و تعداد زیادی از مسیرهای تجربی (به عنوان مثال، توالی حالت ها، اقدامات و پاداش ها) را ثبت می کند که در مرحله آموزش برای یادگیری سیاست های بهینه استفاده می شود.

در این پژوهش رویکرد زمانبندی کاربردی اینترنت اشیا مبتنی بر یادگیری عمیق پیشنهاد شد که از ساختار شبکه عصبی حافظه طولانی کوتاه مدت به منظور طبقه بندی منابع پردازشی و اجرای فرایند زمانبندی استفاده می کند

روش پیشنهادی می تواند به طور موثر هزینه متعادل سازی بار سرورها، هزینه زمان پاسخ برنامه های اینترنت اشیا و هزینه وزنی آنها را بهینه کند. علاوه بر این، با استفاده از روش پیشنهادی برای محدود کردن دفعات به روز رسانی پارامترهای کیفیت سرویس در هر تکرار، انتظار می رود سرعت همگرایی الگوریتم بهبود یابد.

# 2-3) روش شناسی پژوهش

# 3-3) مدل مفهومی پژوهش

پژوهش پیشنهادی در پنج مرحله اجرا می شود. در فاز اول، بیان مسئله، مفاهیم محاسبات ابری، اینترنت اشیا و ساختار مه و لبه در آن مورد بررسی قرار گرفته در ادامه مسئله زمانبندی در این ساختارها به عنوان چالش مطالعه می شود. در بخش دوم از پژوهش، به بررسی و ارزیابی مفاهیم مربوط به زمانبندی در محیط مه و لبه، الگوها و راهکارهای ارائه شده توسط محققان در این زمینه پرداخته می شود و در ادامه مزایا و معایب هریک از رویکردهای بررسی شده توصف می شود. در فصل سوم، روش پیشنهادی توصیف شده و ساختار آن به صورت کلی مورد ارزیابی قرار می­گیرد. در در فصل چهارم، شبیه سازی و ارزیابی روش پیشنهادی اجرا شده و نتایج به دست آمده از شبیه سازی، با رویکردهای موجود در این زمینه مقایسه می شود. و در نهایت در فصل پنجم جمع بندی رویکرد درنظر گرفته انجام و پیشنهاداتی برای ادامه و بهبود تحقیق و فرایند تخصیص منابع و زمانبندی جهت آیندگان ارائه می شود.

.

یکی از کاراترین روش ها برای حل مسائل غیرقطعی استفاده از الگوریتم های تکاملی است. اما مسئله مدیریت اطلاعات در ساختارهای تکاملی یک چالش مهم است. در همین راستا به منظور مدیریت مسئله و بهینه سازی فرایند زمانبندی، از راهکارهای یادگیری عمیق استفاده می شود.

در روش پیشنهادی یک رویکرد زمانبندی ارائه می شود که وابسته به پارامترهای مختلف کیفیت سرویس است، لذا می توان از این پارامترها عنوان مولفه­های ارزیابی در راهکار پیشنهادی استفاده نمود

می توان از پارامترهای کیفیت سرویس مختلفی از قبیل پهنای باند، نرخ گذردهی، نرخ خطای بیتی، قابلیت اطمینان، قابلیت اعتماد، هزینه و موارد مشابه دیگر برای بهینه سازی انتخاب منابع جهت زمانبندی کارا در محیط لبه و مه استفاده کرد، ولی باید دقت داشت که استفاده از پارامترهای بیشتر، ­می­تواند منجر به افزایش پیچیدگی محاسباتی و در نتیجه افت کارایی کلی شبکه شود.

با توجه به این که در این روش از پارامترهای کیفیت سرویس چندگانه استفاده می شود. نیاز به نرمال سازی این پارامترها و همچنین حذف افزونگی اطلاعات وجود دارد. در راهکار پیشنهادی برای نرمال سازی پارامترهای کیفیت سرویس، از فرایند نرمال­سازی استاندارد استفاده شده و به منظور حذف افزونگی از مدل پیرسون استفاده می شود.

در شکل زیر ساختار راهکار پیشنهادی شرح داده شده است.

شکل 3- 1) مدل مفهومی رویکرد پیشنهادی

# 4-3) الگوریتم‌ها و راهکارها

در این بخش از پژوهش، الگوریتم­ها و راهکارای ارائه شده در پژوهش شامل پیش پردازش و نرمال سازی داده های کیفیت سرویس، توصیف خواهد شد.

# 1-4-3) پارامترهای کیفیت سرویس

در این بخش از پژوهش، پارامترهای کیفیت منابع پردازشی که به منظور اجرای فرایند تخلیه منابع مورد استفاده قرار می­گیرند، توصیف خواهند شد در این پژوهش، از پارامترهای کیفیت سرویس، نرخ خطای بیتی منابع پردازشی، قابلیت اطمینان، پهنای باند مصرفی و سرعت پردازشی منابع استفاده می شود که در ادامه هریک از این پارامترها توصیف می شوند.

**پهنای­باند**: پهنای­باند یکی از پارامترهای اولیه برای زمانبندی کارها، در تعیین اولویت واگذاری منابع در سیستم­های رایانش ابری است که علاوه بر قابلیت های محاسباتی منابع، قابلیت های ارتباطی را نیز در نظر می­گیرد. در الگوریتم های زمانبندی کار، منابع بر اساس نوع الگوریتم باتوجه به مقدار پهنای باند می­توانند به کارهای مختلف واگذار شوند. پهنای باند برابر نرخ انتقال داده در واحد زمان توسط یک اتصال شبکه است. واحد اندازه­گیری این پارامتر بیت بر ثانیه در نظر گرفته می شود.

* در یک شبکه کامپیوتر، حداکثر پهنای باند از رابطه 3-1 به دست می آید[34]:

|  |  |
| --- | --- |
| (3-1) |  |

که در این رابطه، RWINپنجره دریافت TCPو RTT زمان رفت و برگشت[[2]](#footnote-2) در منبع پردازشی است.

**قابلیت اطمینان**: هنگام طراحی و توسعه یک محیط رایانش ابری، ارائه روشی با قابلیت اطمینان بالا در این شبکه ها برای ارسال دقیق و سریع پیام ها بسیار مهم است

به منظور بهبود قابلیت اطمینان در شبکه، تکنیک های مختلفی به عنوان راه حلی برای بهبود تحمل خطا در این سیستم­ها معرفی شده است. تحمل خطا شامل شناسایی خطاها و خرابی ها در منبع پردازشی و بازیابی خرابی ها برای ادامه ارسال داده ها است.

تعریف کلی قابلیت اطمینان را می توان به صورت زیر در نظر گرفت: قابلیت اطمینان یک منبع پردازشی در ساختار ابر و مه این امکان را دارد که در آن منبع پردازشی برای مدت معینی در شرایط موجود و از پیش تعیین شده به صورت سالم و بدون نقص فعالیت کند.

تابع قابلیت اطمینان R(t) نشان دهنده احتمال عملکرد بهینه منبع پردازشی در بازه (0, T) است.

* تابع قابلیت اطمینان مکمل تابع توزیع احتمال شکست است، بنابراین می توان آن را با تابع چگالی احتمال تعریف کرد[34].

از آنجایی که تابع قابلیت اطمینان نشان دهنده احتمال شکست است، این توابع به عنوان توابع پایداری نیز شناخته می شوند. زیرا این توابع احتمال عمر باقیمانده تجهیزات را تعیین می­کنند.

قابلیت اطمینان یک منبع پردازشی R برابر با احتمال این است که این منبع در زمان معین و در شرایط خاص ویژگی های مورد نیاز را داشته باشد و در غیر این صورت، عدم اطمینان F برابر است با احتمال اینکه موجودیت مورد نظر نتواند ویژگی های خواسته شده را در بازه زمانی معین تحت شرایط کاری به دست آورد. عدم اطمینان و قابلیت اطمینان هردو وابسته به زمان هستند. در زمان صفر، قابلیت اطمینان موجودیتی که آغاز به فعالیت می کند برابر یک است، پس از مدت زمانی این مقدار به 0.5 و سپس زمانی که از کار بیفتد به صفر می رسد. به بیان دیگر، عدم اطمینان از مقدار صفر شروع شده و به طور صعودی عمل کرده و زمانی که سیستم خراب می شود به مقدار یک می رسد. به طورکلی در هر لحظه مجموع قابلیت اطمینان و عدم اطمینان برابر با یک خواهد بود. این مسئله را می­توان با استفاده از رابطه 3-2 بیان نمود [38].

|  |  |
| --- | --- |
| (3-2) |  |

قابل قبول ترین تعریف برای قابلیت اطمینان برابر است با توانایی یک موجودیت (محصول، سیستم، و غیره) در انجام وظایف تحت شرایط عملیاتی طراحی شده برای مدت زمان معین.

بنابراین قابلیت اطمینان را می توان با استفاده از رابطه 3-3 تعریف نمود [38]:

|  |  |
| --- | --- |
| (3-3) |  |

که در این رابطه نشان دهنده مدت زمان معین یا سیکل کاری موجودیت است و نشان دهنده زمان شکست است و نشان دهنده قابلیت اطمینان موجودیت است. در این رابطه نشان دهنده شرایط تعیین شده همچون شرایط محیطی هستند، معمولاً این شرایط در آنالیز قابلیت اطمینان نادیده گرفته می شوند و معادله 3-3 را می توان به صورت معادله 3-4 بازنویسی نمود[38].

|  |  |
| --- | --- |
| (3-4) |  |

**نرخ پردازشی منابع:** با توسعه فناوری نیاز به سرعت پردازشی افزایش یافته است. پردازنده های رایانه و سرعت ساعت آن ها ویژگی است که معمولاً با فناوری با عملکرد بالا و سریع مرتبط هستند. سرعت پردازنده یک منبع یکی از مهمترین عناصری است که باید در هنگام مقایسه منابع پردازشی در نظر گرفت.

**نرخ گذردهی:** نرخ گذردهی در شبکه های نرم افزار محور، برابر متوسط نرخ تحویل موفق پیام در یک کانال ارتباطی است. این داده ها ممکن است از یک پیوند فیزیکی، منطقی یا با عبور از طریق [گره های](https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D9%87_%D8%B4%D8%A8%DA%A9%D9%87) خاص شبکه، تحویل داده شوند. نرخ گذردهی معمولاً به وسیله [بیت](https://fa.wikipedia.org/wiki/%D8%A8%DB%8C%D8%AA_(%D8%B1%D8%A7%DB%8C%D8%A7%D9%86%D9%87)) بر ثانیه و در برخی موارد [بسته های داده](https://fa.wikipedia.org/w/index.php?title=%D8%A8%D8%B3%D8%AA%D9%87%E2%80%8C%D9%87%D8%A7%DB%8C_%D8%AF%D8%A7%D8%AF%D9%87&action=edit&redlink=1) بر ثانیه در [بازه های مختلف زمانی](https://fa.wikipedia.org/w/index.php?title=%D8%A8%D8%A7%D8%B2%D9%87%E2%80%8C%D9%87%D8%A7%DB%8C_%D8%B2%D9%85%D8%A7%D9%86%DB%8C&action=edit&redlink=1)، [اندازه گیری](https://fa.wikipedia.org/wiki/%D8%A7%D9%86%D8%AF%D8%A7%D8%B2%D9%87%E2%80%8C%DA%AF%DB%8C%D8%B1%DB%8C) می شود. در کل نرخ گذردهی عبارت است از مجموع نرخ داده هایی که به همه پایانه ها که در یک شبکه تحویل داده شده است.

**فاصله ارتباطی اشیا:** در ساختار شبکه های ارتباطی، توابع زیادی برای اندازه گیری فاصله بین گره ها در شبکه وجود دارد.

در این کار، تابع فاصله اقلیدسی برای محاسبه فاصله بین دو گره در یک شبکه استفاده شد. با استفاده از فاصله اقلیدسی، کوتاه ترین فاصله بین دو نقطه بر اساس رابطه فیثاغورث محاسبه می شود.

* اگر برابر نقطه قرار گیری سرخوشه و برابر نقطه قرارگیری چاهک باشد، فاصله اقلیدسی بین این دو نقطه را می توان با استفاده از رابطه 3-5 محاسبه کرد[28].

|  |  |
| --- | --- |
| *(3-5)* |  |

*که در آن شروط زیر برقرار هستند:*

* *فاصله اقلیدسی باید نامنفی باشد: مشخص است که این رابطه مقداری نامنفی دارد زیرا از مجموع مربعات تفاضل ها ساخته شده است.*
* *وجود رابطه انعکاسی برای فاصله اقلیدسی: برای نقطه های و با مقدار مولفه های یکسان، فاصله اقلیدسی برابر با صفر خواهد بود. همچنین زمانی که فاصله اقلیدسی صفر باشد باید همه مربعات تفاضل ها صفر باشند، در نتیجه دو نقطه بر هم منطبق هستند.*

# 2-4-3) پیش پردازش اطلاعات

برای دستیابی به داده های قابل استفاده جهت تصمیم گیری بر اساس پارامترهای کیفیت سرویس بیان شده در بخش قبل، ابتدا باید داده های مختلف ارائه شده همبسته سازی شوند تا بتوان داده های مفید را در قالب یک مدل کلی مشخص نموده و داده های زائد که نقشی در تصمیم گیری ندارند را حذف نمود. به منظور همبسته سازی داده ها، روش های بسیار متنوعی وجود دارد؛ اساس تمامی این روش ها تعیین ضریب همبستگی بین پارامترهای مختلف است. در این پژوهش برای تعیین ضریب همبستگی داده ها از آزمون همبستگی پیرسون[[3]](#footnote-3) استفاده خواهد شد

نتایج به دست آمده از این آزمون به عنوان پارامترهای ورودی برای تعیین منابع بهینه استفاده می شوند.

ضریب همبستگی[[4]](#footnote-4) ابزاری آماری برای تعیین نوع و درجه رابطه یک متغیر کمی با متغیر کمی دیگر است.

ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می دهد. این ضریب بین ۱ تا ۱- است و در عدم وجود رابطه بین دو متغیر، برابر صفر است.

علامت اختصاری آن است. ضریب پیرسون نشان می دهد که تا چه اندازه بین متغیرهای کمّی رابطه خطی وجود دارد.

درحالت کلی ضریب همبستگی پیرسون بین دو متغیر تصادفی برابر با [کوواریانس](https://fa.wikipedia.org/wiki/%DA%A9%D9%88%D9%88%D8%A7%D8%B1%DB%8C%D8%A7%D9%86%D8%B3) آن ها تقسیم بر [انحراف معیار](https://fa.wikipedia.org/wiki/%D8%A7%D9%86%D8%AD%D8%B1%D8%A7%D9%81_%D9%85%D8%B9%DB%8C%D8%A7%D8%B1) آن ها است. برای یک جامعه آماری، ضریب همبستگی جامعه به صورت رابطه (3-6) قابل تعریف خواهد بود[35].

|  |  |
| --- | --- |
| (6-3) |  |

که در این رابطه cov، نشان دهنده کوواریانس است و نشان دهنده انحراف معیار متغیر X است و نشان دهنده انحراف معیار متغیر Y است. نشان دهنده میانگین متغیر X است و نشان دهند میانگین متغیر Y است. نهایت E نشان دهنده امید به ریاضی است. در حالت کلی ضریب همبستگی پیرسون برای یک نمونه آماری با n زوج داده به صورت با استفاده از رابطه (3-7) تعریف می شود[35].

|  |  |
| --- | --- |
| (7-3) |  |

رابطه بالا را می توان به صورت خلاصه تر با استفاده از رابطه (3-8) بیان نمود[35].

|  |  |
| --- | --- |
| (8-3) |  |

که در آن هر یک از مقادیر،،و به ترتیب با استفاده از روابط (3-9)، (3-10)، (3-11) و (3-12) تعریف می شوند[35].

|  |  |
| --- | --- |
| (9-3) |  |
| (10-3) |  |
| (11-3) |  |
| (12-3) |  |

بعد از تعیین معنی داری و جهت رابطه، باید شدت رابطه ارزیابی شود. برای تفسیر شدت رابطه دو متغیر، بسته به کاربردهای موردنظر، تقسیم بندی های گوناگونی ارائه شده­ است. از این تقسیم بندی ها به منظور همبسته سازی داده ها و حذف داده های غیرمفید از بین انبوهی از داده ها استفاده می شود. با استفاده از همبسته نمودن داده ها می توان بخش بسیار بزرگی از ارزیابی های غیرمفید برای داده های غیرضروری را حذف نمود. تقسیم بندی ارائه شده در جدول 3-1 نمونه­ای از این تقسیم­بندی است[35].

**جدول 3- 1) شیوه تفسیر شدت رابطه در همبستگی پیرسون**

|  |  |
| --- | --- |
| شدّت رابطه | تفسیر |
| 8/. تا 1 | **رابطه بسیار قوی** |
| 6/. تا 8/. | **رابطه قوی** |
| 4/. تا 6/. | **رابطه متوسط** |
| 2/. تا 4/. | **رابطه کم (یا ضعیف)** |
| صفر تا 2/. | **فقدان رابطه یا رابطه ناچیز** |

# 3-4-3) نرمال‌سازی اطلاعات

نرمال سازی داده ها روشی برای یکنواخت کردن بازه مقادیر مربوط به متغیرهای مختلف پژوهش است و به بی مقیاس سازی داده ها نیز معروف است. اگر واحد سنجش متغیرهای مورد مطالعه متنوع باشد با استفاده از روش های نرمال سازی می توان داده ها را بی مقیاس کرد. در این پژوهش به منظور بی مقیاس کردن پارامترهای کیفیت سرویس از روش نرمال سازی استفاده می شود.

. در این روش کافی است هر عدد در یک مجموعه بر مجموع عناصر آن مجموعه تقسیم شود. در این صورت جمع کل عناصر پس از نرمال سازی یک خواهد بود. در رابطه 3-13 فرایند نرمال سازی اطلاعات به منظور بی مقیاس کردن داده ها ارائه شده است.

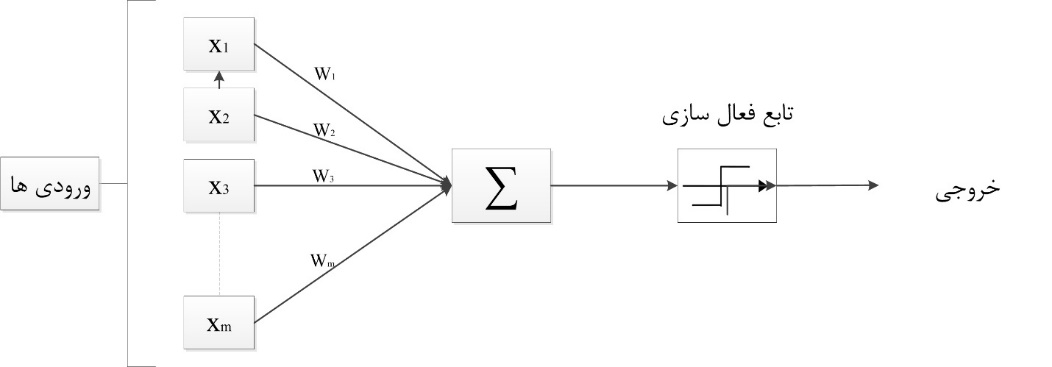
|  |  |
| --- | --- |
| (13-3) |  |

*که در این رابطه نشان دهنده پارامتر کیفیت سرویس i ام برای گره jام و نشان دهنده مجموع مقادیر پارامترهای کیفیت سرویس برای گره j ام است هست.*

# 4-4-3) شبکه عصبی LSTM

به طورکلی، شبکه عصبی یک سامانه پردازشی داده است که از مغز انسان الهام گرفته و پردازش داده ها را بر عهده پردازنده های کوچک و بسیار زیادی سپرده که به صورت شبکه ای به هم پیوسته و موازی با یکدیگر رفتار می کنند تا یک مسئله را حل نمایند

در يك نگاه ساده، مدل يك عصب بايد شامل ورودی هایی باشد كه در نقش سيناپس انجام وظیفه كنند. اين ورودی ها در وزن نهايي ضرب می شوند تا قدرت سيگنال را تعيين كنند. نهايتاً يك عملگر رياضي، تصميم­گيري می کند كه آيا نورون فعال شود يا خير و اگر جواب مثبت باشد، ميزان خروجي را مشخص می سازد. بنابراين شبكه عصبي مصنوعي با استفاده از مدل ساده شده عصب واقعي به پردازش اطلاعات می پردازد. باتوجه به اين توضيحات، می توان مدل ساده­اي براي توصيف يك نورون (يك گره در شبكه عصبي مصنوعي) پيشنهاد كرد. اين مدل در شكل 3-3 نشان داده شده است.



**شکل 3- 2) مدل آموزش در شبکه عصبی**

همان طور که مشاهده می شود ورودی ها که شامل مولفه های مربوط به ساختار شبکه عصبی هستند، در تصویر بالا با مقادیر نمایش داده می شوند. در مدل شبکه عصبی مصنوعی به هر ورودی یک وزن اختصاص داده می شود. این وزن ها در واقع اهمیت ورودی ها را تعیین می کنند، یعنی هر چه وزن بیشتر باشد، ورودی برای آموزش شبکه مهم تر است. سپس تمامی ورودی ها با هم جمع شده و به صورت یک لایه به آکسون وارد می شوند. در مرحله بعد تابع فعال سازی[[5]](#footnote-5) بر روی داده ها اعمال می شود. تابع فعال سازی در واقع نسبت به نیاز مسئله و نوع شبکه عصبی تعریف می شود. این تابع شامل یک فرمول ریاضی برای به روزرسانی وزن ها در شبکه است که می توان با استفاده از رابطه 3-14 بیان نمود.

|  |  |
| --- | --- |
| (3-14) |  |

که در آن پارامتر تنظیم کننده در ساختار شبکه اینترنت اشیا مبتنی بر مدیریت ساختار مصنوعی است. پس از انجام محاسبات در این مرحله اطلاعات از طریق سیناپس های خروجی وارد نورون دیگر می شوند و این مرحله تا جایی ادامه پیدا می کند که شبکه اصطلاحاً آموزش کامل ببیند.

وزن های گره های موجود در اینترنت اشیا )که به صورت نودهای پردازشی شبکه در نظر گرفته می شوند،( می توانند مثبت يا منفي باشند.

، توابع مورد استفاده می توانند بسيار متنوع باشند. از جمله مشهورترين اين توابع مي­توان به تابع هایی نظير arcsin. arctan. sigmoid اشاره كرد

همچنين تعداد گره های ورودی می تواند متغير باشند. زیاد شدن تعداد اين گره ها، به وضوح تعيين وزن ها را با مشكل روبرو می کند

لذا بايد به دنبال روش هایی براي حل اين موضوع بود. روند تعين وزن های بهينه و تنظيم مقادير انجام می شود.

شبكه را با استفاده از قواعد و داده ها آموزش داده و با استفاده از قابليت يادگيري شبكه، و الگوریتم هایی متنوعي همگي سعي در نزدیک کردن خروجي تولید شده توسط شبكه به خروجي ایده آل و مورد انتظار دارند.

بايد توجه داشت كه توپولوژي مغز انسان آن قدر پيچيده است كه نمی توان از آن به عنوان مدلي براي اعمال به شبكه عصبي استفاده نمود، چرا كه مدلي كه استفاده می شود، يك مدل ساده شده است درحالی که چيدمان مغز از المان های بسيار زيادي استفاده می کند.

شبکه های عصبی LSTM به طور خاص برای غلبه بر مشکل وابستگی طولانی مدت با شبکه های عصبی گذشته طراحی شده اند. LSTM ها دارای اتصالات بازخوردی هستند که آنها را با شبکه های عصبی پیش خور سنتی متفاوت می کند. این ویژگی LSTM ها را قادر می­سازد تا اطلاعات مفیدی در مورد داده های قبلی در توالی برای کمک به پردازش نقاط داده جدید حفظ کنند. در نتیجه، LSTM ها به ویژه در پردازش توالی داده ها مانند متن، گفتار و سری های زمانی عمومی خوب هستند.

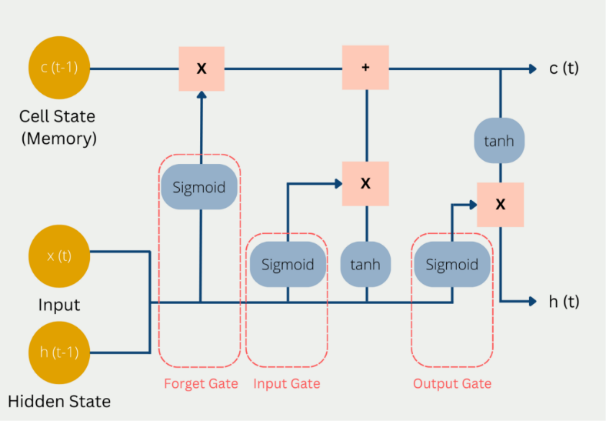
بر اساس این الگوریتم، در سطح پایه، خروجی یک LSTM در یک نقطه زمانی خاص به سه چیز وابسته است:

* حافظه بلند مدت فعلی شبکه - که به صورت وضعیت سلولی شناخته می شود
* خروجی در نقطه قبلی در زمان - به صورت حالت پنهان قبلی شناخته می شود
* داده های ورودی در مرحله زمانی فعلی

LSTM ها از یک سری "دروازه" استفاده می کنند که نحوه ورود، ذخیره و خروج اطلاعات در یک دنباله داده از شبکه را کنترل می کند. سه دروازه در یک LSTM معمولی وجود دارد. گیت فراموشی، گیت ورودی و گیت خروجی. این گیت ها را می توان فیلتر در نظر گرفت و هر کدام شبکه عصبی خودشان هستند.

مشکل شبکه های عصبی بازگشتی این است که آنها به سادگی داده های قبلی را که در "حافظه کوتاه مدت" خود ذخیره می کنند در هنگامی که حافظه موجود در آن تمام شد، به سادگی طولانی ترین اطلاعات حفظ شده را حذف می کند و آن را با داده های جدید جایگزین می کند. مدل LSTM سعی می کند با حفظ اطلاعات انتخاب شده در حافظه بلندمدت این مشکل را رفع می­کند.

در روش پیشنهادی برای ارائه یک راهکار کیفیت سرویس، در هر مرحله محاسباتی، ورودی فعلی ، که شامل پارامترهای موثر در کیفیت سرویس است، حالت قبلی حافظه کوتاه مدت برابر و حالت قبلی حالت پنهان استفاده می شود. ساختار اولویت بندی مبتنی بر رویکرد پیشنهادی در شکل 3-3 نشان داده شده است:

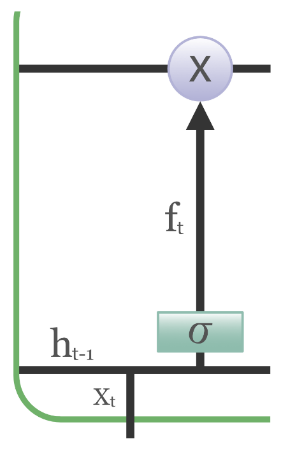


**شکل 3- 3) معماری اولویت‌بندی پارامترهای کیفیت سرویس مبتنی بر شبکه عصبی LSTM**

در راهکار پیشنهادی، به منظور طبقه­بندی توصیه­ها به کاربران سه پارامتر ، و در مسیر خود از گیت­های زیر عبور می کنند:

در گیت فراموشی[[6]](#footnote-6)، تصمیم گرفته می شود که کدام اطلاعات فعلی و قبلی در مورد پارامترهای کیفیت سرویس برای هر توصیه نگهداری شود و کدام یک به بیرون استخراج شود. این شامل وضعیت پنهان از پاس قبلی و ورودی فعلی است. این مقادیر به یک تابع سیگموئید[[7]](#footnote-7) منتقل می شوند که فقط می تواند مقادیر بین 0 و 1 را خروجی دهد. مقدار 0 به این معنی است که اطلاعات قبلی را می توان فراموش کرد؛ زیرا احتمالاً اطلاعات جدید و مهم تری وجود دارد. عدد یک به این معنی است که اطلاعات قبلی حفظ می شود.

* ساختار گیت فراموشی باتوجه به شکل 4-3 قابل توصیف است:



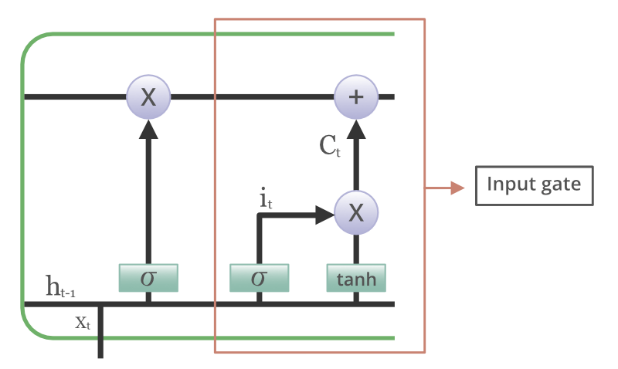
**شکل 3- 4) ساختار گیت فراموشی مبتنی بر شبکه عصبی LSTM**

نتایج حاصل از این در حالت سلولی فعلی ضرب می شود، بنابراین دانشی که دیگر موردنیاز نیست فراموش می شود، زیرا در 0 ضرب می شود و در نتیجه حذف می شود. رابطه مربوط به گیت فراموشی بر اساس رابطه 3-15 قابل توصیف است:

|  |  |
| --- | --- |
| (3-15) |  |

جایی که در آن:

* نشان دهنده ماتریس وزن مرتبط با دروازه فراموشی است.
* نشان دهنده پیوستگی ورودی فعلی و حالت پنهان قبلی است.
* سوگیری با گیت فراموشی است.
* σ تابع فعال سازی سیگموئید است.
* در گیت ورودی[[8]](#footnote-8)، تصمیم گرفته می شود که ورودی فعلی چقدر برای حل کار ارزشمند است. برای این، ورودی فعلی در حالت پنهان و ماتریس وزن آخرین اجرا ضرب می شود. سپس تمام اطلاعاتی که در ورودی مهم به نظر می رسند به حالت سلول اضافه می شوند و حالت سلولی جدید را تشکیل می دهند. این وضعیت سلولی جدید اکنون وضعیت فعلی حافظه بلند مدت است و در اجرای بعدی مورد استفاده قرار خواهد گرفت. همچنین، افزودن اطلاعات مفید به حالت سلول توسط گیت ورودی انجام می شود. ساختار گیت ورودی در شکل 3-5 نشان داده شده است:



**شکل 3- 5) ساختار گیت ورودی مبتنی بر شبکه عصبی ­LSTM**

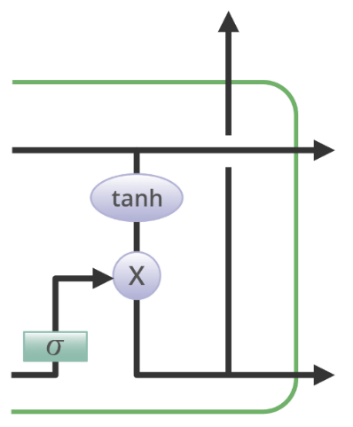
در این گیت، ابتدا، اطلاعات با استفاده از تابع sigmoid تنظیم می شود و مقادیری که باید به خاطر سپرده شوند، مشابه دروازه فراموشی با استفاده از ورودی های و فیلتر می شوند. سپس، یک بردار با استفاده از تابع tanh ایجاد می شود که خروجی از 1- تا 1+ می دهد، که شامل تمام مقادیر ممکن از و است. در نهایت، مقادیر بردار و مقادیر تنظیم شده برای به دست آوردن اطلاعات مفید ضرب می شوند. معادله گیت ورودی بر اساس رابطه3-16 قابل توصیف است:

|  |  |
| --- | --- |
| (3-16) | = tanh ()  = ⊙ C)t-1) + ⊙ |

جایی که در آن:

* ⊙ نشان دهنده ضرب عنصر است
* tanh تابع فعال سازی tanh است

گیت خروجی[[9]](#footnote-9)، خروجی مدل LSTM در حالت پنهان محاسبه می شود. بسته به کاربرد، می تواند؛ مثلاً کلمه ای باشد که معنای جمله را تکمیل کند. برای انجام این کار، تابع سیگموئید تصمیم می گیرد که چه اطلاعاتی می تواند وارد گیت خروجی شود و سپس حالت سلول پس از فعال شدن به تابع tanh ضرب می شود. همچنین وظیفه استخراج اطلاعات مفید از وضعیت سلول فعلی برای ارائه به عنوان خروجی توسط گیت خروجی انجام می شود. ابتدا یک بردار با اعمال تابع tanh بر روی سلول تولید می شود. سپس، اطلاعات با استفاده از تابع sigmoid و فیلتر توسط مقادیری که باید با استفاده از ورودی های و به خاطر سپرده شوند، تنظیم می شوند. ساختار گیت خروجی در شکل 6-3 نشان داده شده است.



**شکل 3- 6) ساختار گیت خروجی مبتنی بر شبکه عصبی ­LSTM**

در نهایت، مقادیر بردار و مقادیر تنظیم شده ضرب می شوند تا به عنوان خروجی و ورودی به سلول بعدی ارسال شوند. معادله گیت خروجی به صورت رابطه 3-17است:

|  |  |
| --- | --- |
| (3-17) | = σ() |

در نهایت وقتی به روزرسانی ها در حافظه بلند مدت شبکه کامل شده است، می توان به مرحله نهایی که نشان دهنده تصمیم گیری در مورد وضعیت پنهان جدید است حرکت کرد. برای تصمیم گیری در این مورد، از سه مؤلفه استفاده می شود. وضعیت سلولی که به تازگی به روز شده است، وضعیت پنهان قبلی و داده های ورودی جدید.

# 5-3) زمانبندی وظایف با استفاده از روش پیشنهادی

حال با نرمال سازی و تعیین همبستگی مقادیر پارامترهای کیفیت سرویس هر یک از منابع، از مدل شبکه عصبی LSTM به منظور طبقه بندی منابع برای حل مسئله استفاده می شود.

. با توجه به اینکه هر یک از پارامترهای کیفیت بیان شده در بالا، دارای واحدهای مختلفی بوده و نمی توان آن ها را با یکدیگر مقایسه کرد، در نتیجه باید از یک روش نرمال سازی برای نرمال نمودن مقادیر و در نتیجه قابل مقایسه کردن این پارامترها استفاده شود.

در نهایت یک ماتریس m در n به منظور ارزیابی هریک از منابع به دست می آید. این ماتریس شامل یک ترتیب نزولی از منابع مناسب (بر اساس پارامترهای کیفیت سرویس) برای تخلیه بار در شبکه است.

گزینش، منبعی با پارامترهای کیفیت سرویس مناسب، که دارای قابلیت اطمینان، نرخ خطای بیتی و نرخ گذردهی بالا و فاصله بین گره های کمتری هستند، می تواند منجر به ایجاد تعادل بار در شبکه شده و کارایی را افزایش دهد.

این روش پیشنهادی با کاهش ترافیک در شبکه نرخ تحویل موفق بسته را افزایش داده و تأخیر را به صورت چشمگیری کاهش می­دهد.

بعد از اتمام هر دور از تخصیص وظایف در شبکه، مقادیر هر یک از پارامترهای بیان شده برای منابع به­روزرسانی شده و مراحل فوق برای دور بعدی زمانبندی وظایف تکرار می شوند.

# 6-3) فلوچارت روش پیشنهادی

به‌طور کلی فلوچارت روش پیشنهادی در شکل 3-7 نشان داده‌شده‌است.



شکل 3- 7) فلوچارت رویکرد پیشنهادی

# 7-3) الگوریتم نهایی

در این بخش از پژوهش، الگوریتم نهایی رویکرد پیشنهادی به صورت گام به گام شرح داده شده و در ادامه فلوچارت روش پیشنهادی بیان می شود.

گام اول) شروع

گام دوم) تعیین جمعیت اولیه منابع پردازشی

گام سوم) مقدار دهی اولیه پارامترهای کیفیت سرویس برای منابع موجود

گام چهارم) ارزیابی پارامترهای کیفیت سرویس برای هریک از درخواست­ها و منابع به منظور بارگذاری وظایف

گام پنجم) نرمال سازی پارامترهای کیفیت سرویس به منظور یکسان کردن مقیاس داده ها

گام ششم) ورود داده ها در گیت فراموشی شبکه عصبی LSTM

گام هفتم) انتخاب داده های مورد نظر با استفاده از گیت فراموشی و انتقال آنها به گیت شبکه عصبی LSTM

گام هشتم) اجرای تابع سیگموئید به منظور طبقه بندی منابع پردازشی بر اساس ورودی های کیفیت سرویس

گام نهم) ارسال داده ها به گیت خروجی شبکه عصبی LSTM به منظور تعیین طبقه منابع پردازشی

گام دهم) اجرای فرایند زمانبندی و تخصیص منابع

گام یازدهم) به روز رسانی پارامترهای کیفیت سرویس پس از بارگذرای منابع پردازشی

گام دوازدهم) تکرار مراحل چهارم تا دهم تا پایان بارگذاری وظایف

گام سیزدهم) پایان

# 8-3) جمع بندی فصل

در این فصل از پژوهش، رویکرد پیشنهادی زمانبندی در محیط های محاسباتی لبه و مه مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی حافظه طولانی کوتاه مدت برای بهینه سازی بار سیستم و زمان پاسخ ارائه شده و الگوریتم های بکارگرفته شده در آن به همراه ساختار کلی روش پیشنهادی توصیف گردید. در ادامه، فلوچارت و الگوریتم روش پیشنهادی نیز توصیف گردید. در فصل آتی به بررسی و ارزیابی کارایی روش پیشنهادی در مقایسه با رویکردهای مشابه بررسی شده و نتایج ارزیابی به صورت جدول و نمودار توصیف می شوند.

4) فصل چهارم

**ارزیابی نتایج**

**اهداف فصل:**

* توصیف نتایج شبیه سازی
* ارزیابی نتایج شبیه سازی
* مقایسه روش پیشنهادی با رویکردهای موجود

# 1-4) مقدمه

در این فصل از پژوهش، شبیه سازی ها برای ارزیابی عملکرد روش پیشنهادی انجام شده است. در ابتدا، محیط شبیه سازی، روش مدیریت اطلاعات، برنامه های مرجع و سناریوهای شبیه سازی شرح داده شده و در ادامه روش پیشنهادی با رویکردهای مشابه در این زمینه مقایسه خواهد شد. سپس نتایج ارزیابی از دیدگاه‌های مختلف ارائه و مورد بحث قرار می گیرد. در ادامه توضیح مختصری در مورد محیط شبیه سازی با عنوان ifogsim ارائه شده و در ادامه فرایند شبیه سازی و نتایج به‌دست‌آمده بیان می شوند.

# 2-4) معرفی محیط شبیه سازی

. با در نظر گرفتن شبیه سازی به عنوان یک کلید برای محدودیت تست و اجرا در دنیای واقعی، نرم افزارهای مختلفی توسعه یافته اند که می‌توانند رفتار فیزیکی محیط های محاسباتی مه یا لبه را تقلید کنند. با این وجود، شبیه سازهای موجود اغلب به دلیل معماری یکپارچه، فقدان مجموعه داده واقعی و محدوده کوچک برای به روزرسانی دوره ای، از ویژگی های مدیریت خدمات پیشرفته پشتیبانی نمی‌کنند.

شبیه ساز IfogSim، که یک شبیه ساز پرکاربرد .شبیه ساز IfogSim یک کتابخانه با زبان برنامه نویسی جاوا می باشد. پروژه های مختلف مدیریت منابع در محیط‌های IfogSim، در آن قابل پیاده سازی است. سهولت پذیرش IfogSim و سیاست‌های مدیریت سرویس داخلی آن با استفاده از سه سناریو مورد استفاده ارزیابی شده و با شبیه سازهای موجود و سیاست‌های معیار در تنظیمات مختلف مقایسه شده‌اند. نتایج نشان‌می دهد که شبیه ساز ما حافظه کمتری مصرف می‌کند و در مقایسه با شبیه سازهای دیگر، زمان شبیه سازی را به طور متوسط ۲۸ درصد به حداقل می رساند. از جمله مزایای این شبیه ساز می‌توان به موارد زیر اشاره نمود:

* در نظر گرفتن تأثیرات زمان
* این شبیه ساز کاملاً منعطف و کاربردی است.
* تست سیاست های مورد نظر در یک محیط قابل کنترل و قابل تکرار.
* تنظیم گلوگاه های سیستم قبل از استقرار روی ابر واقعی.
* پیکربندی قراردادهای SLA و چک کردن میزان تخطی روش های مختلف از آن.
* ارزیابی سیاست‌هایی مانند بهینه سازی روی سرویس دهنده های مورد استفاده.

با استفاده از جعبه ابزار IfogSim، بسیاری از عملیات مربوط به محاسبات ابری و لبه را می‌توان، شبیه سازی کرد. برخی از کاربردهای مهم این ابزار به شرح زیر می باشند:

* مدل سازی و شبیه سازی مراکز داده
* شبیه سازی ماشین های مجازی مبتنی بر مه همراه با تخصیص سیاست های مختلف بر روی آن ها
* درج پویای عناصر شبیه سازی، متوقف کردن و راه اندازی مجدد یک شبیه سازی
* شبیه سازی و مدل سازی توپولوژی های شبکه ای استفاده شده در مراکز مدیریت داده
* مدل سازی محاسبات مه و لبه
* نمایش چگونگی ساخت یک فضای مه و لبه
* چگونگی ساخت یک فضای مه و لبه
* چگونگی ساخت و پیکربندی محیط های مه و لبه
* نمایش گردش کار در مراکز داده ها و اختصاص دادن ساختار لبه به ماشین های مجازی و....

شبیه ساز IfogSim، متشکل از لایه هایی می‌باشد که با استفاده از این لایه ها، عملیات مربوط به شبیه سازی و مدل سازی را انجام می دهد.

# 3-4) متغیرهای شبیه سازی

همان طور که بیان شد، برای ارزیابی و بررسی روش پیشنهادی و مقایسه آن با سایر رویکردها، از میحط شبیه سازی IfogSim و ویرایش گر NetBeans با زبان برنامه نویسی مبتنی بر جاوا و ماشین مجازی جاوا JDK استفاده‌شده است. در این ارزیابی ها، از یک کامپیوتر با سیستم عامل ویندوز 10 شرکت مایکروسافت نسخه 64 بیتی با پردازنده اینتل 4 هسته‌ای و رم پردازشی 8 گیگابایتی استفاده شد. جدول 4-1 نشان دهنده پارامترهای شبیه سازی برای روش پیشنهادی است. مجموعه داده مورد استفاده در پژوهش، مجموعه داده IfogSim می باشد که به صورت عمومی در اختیار محققان قرار داده شده است.

**جدول 4- 1) متغیرهای شبیه سازی**

|  |  |
| --- | --- |
| پارامترهای ارزیابی | مقادیر پارامتر |
| تعداد سرویس ها | 500 |
| اندازه بسته ورودی – خروجی | 50 کلیوبایت |
| بیشترین توان ارتباطی سرویس ها | 50 متر |
| توان ارتباطی معمول در سرویس ها | 20 متر |
| زمان مورد نیاز برای ارتباط | 3 تا 30 ثانیه برای هر ارتباط |
| بیشترین سطح ارتباطی | 50 ارتباط |
| پردازنه مرکزی (MIPS) | 200 تا 500 |
| مجموعه داده استفاده شده | مجموعه داده fog-computing-edge-servers |
| اندازه حافظه رم | 8 گیگابایت |
| حداکثر حجم داده | 900 مگابایت |
| زمان بند ابری | RoundRobin |
| تعداد کلودلت | 150 |
| تعداد هسته ها در هر ماشین مجازی | 2 هسته |

در این پژوهش از مجموعه داده ای استفاده می شود که به صورت آنلاین در اختیار محققان قرار داده شده است. این مخزن مجموعه ای از مجموعه داده های EUA را که محققان از منابع داده های دنیای واقعی جمع آوری کرده و نگهداری می‌کنند. مجموعه داده ها به‌صورت عمومی منتشر شده است. ساختار کلی مجموعه داده به شرح زیر است:

* سرورهای ابر: شامل مجموعه داده های مکان های سرور ابر و پارامترهای کیفیت سرویس سرورها است.
* کاربران: شامل مجموعه داده های کاربر در محیط ابری است.
* درخواست ها: شامل وظایف درخواست شده از طریق کاربران برای پردازش در محیط ابر است.

**همان طور که بیان شد، در این بخش از پژوهش، چارچوب شبیه سازی توسعه داده شده و الگوریتم پیشنهادی در محیط IfogSim ارزیابی شد.**

# 4-4) معیارهای عملکرد در فرایند شبیه سازی

برای ارزیابی عملکرد روش پیشنهادی برای زمانبندی در محیط‌های محاسباتی لبه و مه مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی حافظه طولانی کوتاه مدت برای بهینه سازی بار سیستم و زمان پاسخ، پارامترها و معیارهای زیر درنظر گرفته می شوند.

* **میانگین تأخیر زمان:** میانگین تأخیر زمان بیانگر زمان مورد نیاز برای انتقال داده در پردازنده است.
* **درجه تفاوت سرویس دهنده جهت انتقال داده**: این متریک عملکرد برای اندازه گیری تعادل سرویس تخصیص داده شده استفاده می شود. تعادل مصرف انرژی و توازن بار ارتباط نزدیکی با این متریک عملکرد دارد.
* **کیفیت سرویس و زمان بهینه:** ویژگی خوب متریک عملکرد QoS است. این مؤلفه به صورت تعدادی از داده های ارتباطی است که با موفقیت در مقصد در زمان ارتباط موردنیاز دریافت می شود.
* **توان**: میزان کارایی میزان تحویل پیام موفق از طریق کانال بی سیم ارتباطی، در محیط محاسباتی مه است.
* **زمان بارگیری از سرور (DLTS):** زمان بارگیری از سرور در شبیه سازی ها اندازه گیری می شود تا عملکرد تأخیر زمانی از سرور در چارچوب های مختلف محاسباتی را در محیط محاسباتی مه نشان دهد.

# 5-4) نتایج شبیه سازی

در این بخش از پژوهش نتایج مربوط به ارزیابی های ارائه شده توسط روش پیشنهادی و رویکردهای پیشنهادی گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] مورد بررسی قرار گرفته و در قالب نمودار مقایسه می شوند.

. اولین مورد ارزیابی شامل تأخیر متوسط برای روش پیشنهادی و رویکرد مقایسه شده‌است. به طور کلی، تأخیر متوسط برابر میانگین تأخیر زمانی متفاوتی برای بهترین عملکرد از نظر تأخیر زمانی و ساختار روش پیشنهادی نشان می دهد میانگین تأخیر زمانی با افزایش مقدار داده برای همه روش ها افزایش می یابد. مقادیر مربوط به تأخیر برای روش پیشنهادی و گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] در شکل 4-1 نشان داده شده است. این شکل نشان می دهد که با افزایش مقدار داده های ارتباطی برای هر دو، میانگین تأخیر زمان افزایش می‌یابد. اما افزایش تأخیر برای روش پیشنهادی نسبت به رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] سطح پایین تری دارد، با توجه به اینکه روش پیشنهادی قبل از انجام محاسبات روی ساختار خود، از یک فرایند یادگیری عمیق به منظور پیش بینی سطح ترافیک و پرهیز از ایجاد ازدحام در سطح شبکه استفاده می کند، این ساختار موجب بهبود تعادل بار در شبکه شده و میانگین زمان انتظار در شبکه را کاهش می دهد. در نتیجه میانگین تأخیر برای روش پیشنهادی نسبت به رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] سطح بهتری خواهد داشت.

**شکل 4- 1) مقایسه تأخیر میانگین روش پیشهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

عملکرد سربار محاسباتی وظایف به روش های مختلف و برای تعداد متفاوتی از گره ها و مراکز داده هایی که توسط ساختار مدل پیشنهادی ارزیابی می شوند، در شکل 4-2 ارائه شده است. به طور کلی، نرخ سربار محاسباتی از هر استراتژی با افزایش مهلت و تعداد مراکز داده افزایش می یابد. با توجه به ارزیابی ها می‌توان دو مسئله را در نظر گرفت، مورد اول ارسال و تحویل اطلاعات بدون در نظر گرفتن مهلت زمانی پردازنده ها و مورد دوم با در نظر گرفتن مهلت زمانی است،

بر اساس شکل 4-2 (الف) و بدون در نظر گرفتن مهلت زمانی برای پردازش، ساختار پیشنهادی، نرخ سربار محاسباتی کمتری در مقایسه با رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] دارد. با کاهش سربار محاسباتی، سرعت اجرای فرایند زمانبندی بهینه تر می شود. باتوجه به اینکه مهلت زمانی تأثیر بالایی بر سربار محاسباتی منابع دارد، به نظر می‌رسد با افزایش مهلت زمانی، این مقدار در روش پیشنهادی و رویکردهای مقایسه شده نیز سطح بالاتری به خود اختصاص داده و حجم پردازش‌های قابل اجرا در شبکه افزایش پیدا می‌کند.

**شکل 4- 2) مقایسه سربار محاسباتی روش پیشهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

رویکرد پیشنهادی سربار محاسباتی بهتری نشان‌می دهد و در عین حال، رویکردهای مقایسه شده نیز سربار محاسباتی بالاتری ایجاد می‌کنند. با توجه به اینکه روش پیشنهادی از اطلاعات پالایش شده و پیش پردازش شده به منظور اجرای فرایند زمانبندی استفاده می‌کند، به نظر می‌رسد، در هر دو حالت استفاده از مهلت زمانی و عدم وجود مهلت زمانی، کارایی قابل توجهی نسبت به رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] با محاسبه مه و لبه کارایی بالاتری را نشان‌می دهد.

در ادامه فرایند شبیه سازی، مقادیر مربوط به عملکرد روش پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] در زمینه تعادل بار ارزیابی می شود. تعادل بار در زمانبدی وظایف، یک متریک عملکرد برای شبکه بوده و شاخصی مفید برای ارزیابی کارایی است. این فرایند نشان می دهد که تعادل توزیع داده های دریافت شده از منبع تا چه اندازه موفق است. نتایج ارزیابی عملکرد تعادل بار برای روش پیشنهادی رویکرد مقایسه شده در شکل 4-3 نشان داده شده است.

**شکل 4- 3) مقایسه تعادل بار روش پیشهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

بدیهی است، وقتی حجم داده به حد متفاوتی می‌رسد، مقادیر تعادل بار نیز متفاوت خواهد بود. با توجه به اینکه روش پیشنهادی مؤثرترین منبع برای زمانبندی را انتخاب خواهد کرد، در نتیجه مقادیر مربوط به عملکرد نرخ تعادل بار به دلیل توزیع بهینه منابع به وظایف نیز برای روش پیشنهادی بهتر از رویکردهای مقایسه شده خواهد بود. با توجه به اینکه در روش پیشنهادی از فرایند یادگیری عمیق برای بهبود ترافیک شبکه و از رویکرد محاسبات ابری برای توزیع بار ترافیکی شبکه استفاده می شود، لذا فرایند انتخاب منابع بهینه برای تخصیص درخواست در شبکه انجام شده و نتایج به صورت قابل توجهی بهبود پیدا می‌کنند. با بهبود تعادل بار، عملکرد کیفیت سرویس نیز نسبت به رویکرد مقایسه شده نیز بهبود پیدا می‌کند.

در ادامه فرایند شبیه سازی ها، مقادیر مربوط به انرژی مصرفی برای رویکرد پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] با محاسبات ابری مورد بررسی و ارزیابی قرار می‌گیرد. انرژی مصرفی در مقادیر داده های مختلف برای روش پیشنهادی و رویکرد مقایسه شده در شکل 4-4 برای 5 مقیاس متفاوت از داده ها ارائه شده است. نتایج ارزیابی نشان می دهد مصرف انرژی برای روش پیشنهادی نسبت به رویکرد مقایسه شده گروه تحقیقاتی هوسین و همکارانش[21] سطح بهتری دارد ولی گروه تحقیقاتی وانگ و همکارانش[2] نسبت به روش پیشنهادی سطح پایین تری دارد.

**شکل 4- 4) مقایسه انرژی مصرفی روش پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

با توجه به اینکه روش پیشنهادی از ساختار ارزیابی منابع مبتنی بر یادگیری ماشین در داخل شبکه استفاده می‌کند، خروجی حاصل از این فرایند، یک ساختار بهینه شده از نظر ترافیک شبکه می‌باشد. همچنین با استفاده از محاسبات ابری، روش پیشنهادی سعی در ایجاد تعادل بار ترافیکی در شبکه داشته و موجب افزایش پهنای باند قابل استفاده در شبکه می‌گردد، این فرایند به نوبه خود تأثیر بسزایی در انرژی مصرفی سیستم دارد. بر اساس شکل 4-4 مشاهده می شود که انرژی مصرفی برای روش پیشنهادی در تمامی حجم داده‌ای ارزیابی‌شده، نسبت به رویکرد مقایسه شده گروه تحقیقاتی وانگ و همکارانش[2] سطح بهتری دارد.

مقایسه مصرف انرژی برای روش پیشنهادی رویکرد گروه تحقیقاتی وانگ و همکارانش[2] نیز به طور میانگین بهبود 1.98 درصدی داشت اما نسبت به گروه تحقیقاتی هوسین و همکارانش[21] 2.32 درصد روش پیشنهادی عملکرد ضعیف تر را نشان می دهد.

در ادامه مقادیر مربوط به نرخ همگرایی به جواب بهینه برای روش پیشنهادی رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] مورد بررسی و ارزیابی قرار گرفته و نتایج ارزیابی در قالب نمودار در شکل 4-5 ارائه می شود.

**شکل 4- 5) مقایسه نرخ همگرایی به جواب بهینه برای روش پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

نتایج ارزیابی نشان‌می دهد که روش پیشنهادی با زمانبدی وظایف بهینه در شبکه همگرایی بهتری به جواب بهینه دارد. همچنین با توجه به اینکه مسئله مدیریت ترافیک در روش پیشنهادی با استفاده از راهکار مبتنی بر یادگیری ماشین بهینه می شود. می توان توجیه کرد که راهکار پیشنهادی قادر به کمینه سازی تأخیر و در نهایت بهبود نرخ همگرایی به جواب بهینه است.

در نهایت مقادیر مربوط به زمان پاسخ برای روش پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] در شکل 4-6 مورد بررسی قرار گرفته و نتایج نشان دهنده کارایی بالاتر روش پیشنهادی نسبت به رویکردهای مقایسه شده در مولفه زمان پاسخ است. با توجه به اینکه روش پیشنهادی با ایجاد تعادل بار در توزیع منابع پردازشی، باعث بهبود زمان اجرا و نرخ همگرایی به جواب بهینه می شود، از این رو می‌توان انتظار داشت که زمان پاسخ برای روش پیشنهادی به منظور اجرای وظایف پردازشی در مقایسه با رویکردهای گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2] کارایی بالاتری داشته‌باشد.

**شکل 4- 6) مقایسه زمان پاسخ برای روش پیشنهادی و رویکرد گروه تحقیقاتی هوسین و همکارانش[21] و گروه تحقیقاتی وانگ و همکارانش[2]**

# 6-4) ارزیابی نتایج

در این فصل از پژوهش، رویکرد جدیدی در زمینه زمانبندی در محیط‌های محاسباتی لبه و مه مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی حافظه طولانی کوتاه مدت برای بهینه سازی بار سیستم و زمان پاسخ ارائه و در محیط شبیه سازی IfogSim ارزیابی‌شده و نتایج حاصل از شبیه سازی، به صورت نمودار ارائه شدند. بر اساس نتایج به دست آمده از شبیه سازی مشخص شد که روش پیشنهادی نسبت به رویکرد مقایسه شده دارای تأخیر میانگین و زمان پاسخ بهتری است،

وجود تعادل بار در روش پیشنهادی عاملی بر بهینگی روش پیشنهادی در تأخیر میانگین است. همچنین نتایج ارزیابی نشان داد که روش پیشنهادی تعادل بار بیشتری داشته و در نهایت قادر به تخصیص منابع بیشتر و در نتیجه تعادل بار بهتر در زمانبندی است.

همچنین مقادیر مربوط به عملکرد بار محاسباتی برای روش پیشنهادی و رویکردهای مقایسه شده ارزیابی شدند و نتایج نشان دهنده کارایی بالای روش پیشنهادی نسبت به راهکارهای مقایسه شده بود.

در ادامه نرخ همگرایی به جواب بهینه هر سه روش بر اساس مقادیر داده متفاوت ارزیابی شد و نتایج نشان داده که روش پیشنهادی به دلیل ارزیابی مبتنی بر ترافیک و تعادل بار بهینه، کارایی بالاتری نسبت به رویکرد مقایسه شده دارد. همچنین در فرایند شبیه سازی با ارزیابی انرژی مصرفی سرور مشاهده شد که روش پیشنهادی نست به رویکرد مقایسه شده، انرژی کمتری صرف تخصیص منابع مورد نظر دارد

و در نتیجه کارایی روش پیشنهادی نسبت به رویکردهای مقایسه شده سطح بالاتری خواهد داشت.

5) فصل پنجم

# جمع‌بندی و پیشنهادات

**اهداف فصل:**

* ارائه جمع‌بندی کلی
* پیشنهادات آتی

# 1-5) جمع‌بندی

از نقطه نظر فناوری، ماهیت غیرمتمرکز محاسبات ابری با ساختار نسبتاً متمرکز ابر مطابقت ندارد. امروزه داده های محاسبات ابری عمدتاً به روش توزیع شده تولید می شوند، برای پردازش به یک ابر متمرکز ارسال می شوند و سپس به سهام داران توزیع شده یا سایر دستگاه های محاسبات ابری، که اغلب نزدیک به منابع داده اولیه قرار دارند، تحویل داده می شوند.

این رویکرد پردازش متمرکز منجر به تأخیرهای ارتباطی بالا و نرخ انتقال داده پایین محاسبات ابری و همچنین ساختارهای مجازی و کاربران بالقوه می شود.

پشتیبانی از پردازش غیرمتمرکز داده ها در دستگاه های محاسبات ابری در ترکیب با مزایای فناوری های ابری و مجازی سازی به عنوان یک رویکرد امیدوارکننده برای کاهش هزینه های ارتباطی و زمان های انتقال داده در محاسبات ابری شناسایی شده است. برای تحقق پردازش غیر متمرکز داده، لازم است بخش هایی از منابع محاسباتی و ذخیره سازی مورد نیاز برای پردازش داده های موجود به منابع داده و مصرف کنندگان خدمات (یعنی کاربران نهایی یا مخزن های داده) نزدیک تر شوند.

# 2-5) پیشنهادات آتی

عنوان پیشنهاد برای کارهای آتی می‌توان به موارد زیر اشاره نمود.

* به منظور بالا بردن فرایند زمانبندی وظایف می‌توان از روش های چندمعیاره با ایجاد وزن بر روی پارامترهای مربوط به انتخاب گره استفاده‌کرد.
* استفاده از رویکردهای بهینه سازی ریاضی و همچنین رویکردهای ترکیبی جهت زمانبندی وظایف در محیط ابر
* ارزیابی و بررسی کارایی راهکار پیشنهادی در مقیاس های بزرگ تر و واقعی تر
* ارزیابی روش پیشنهادی براساس پارامترهای متفاوت دیگر از کیفیت سرویس منابع پردازشی
* استفاده از رویکردهای متفاوت بهینه سازی مصرف انرژی و کاهش هزینه مدیریت منابع پردازشی

1. IoT [↑](#footnote-ref-1)
2. Round-Trip Time [↑](#footnote-ref-2)
3. Pearson's Correlation Coefficient [↑](#footnote-ref-3)
4. Correlation Coefficient [↑](#footnote-ref-4)
5. Multi Layer Perceptron [↑](#footnote-ref-5)
6. Forget Gate [↑](#footnote-ref-6)
7. Sigmoid function [↑](#footnote-ref-7)
8. Input Gate [↑](#footnote-ref-8)
9. Output Gate [↑](#footnote-ref-9)