دوفصلنامه محاسبات و سامانه های توزیع شده
سال چهارم، شماره اول، شماره پیاپی ۷، صفحه ۳۹ تا ۵۰، سال ۱۴۰۰

بهینه‌سازی مصرف انرژی در سیستم‌های توزیع شده براساس زمان‌بندی وظیفه با استفاده از یادگیری تقویتی و روش عملگر-نقاد

آرش علی بخشی‌فرد*[۱]، لیدا ندرلو[۲]، زهرا طیبی قصبه[۳]
[۱]دانشکده فنی و مهندسی، گروه کامپیوتر، دانشگاه آزاد اسلامی واحد زنجان، زنجان، ایران.
[۲]دانشکده فنی و مهندسی، گروه کامپیوتر گرایش هوش مصنوعی، موسسه آموزش عالی روزبه واحد زنجان، زنجان، ایران.
[۳]مدرس، دانشکده مهندسی کامپیوتر، دانشگاه پیام‌نور گیلان، رشت.

چکیده

مصرف انرژی در مراکز و سیستم‌های داده‌ای به شدت در حال افزایش است که بررسی این موضوع یک مساله اساسی در عصر کنون است. از جمله مزایای مهم سیستم‌های توزیعی صرفه‌جویی در هزینه‌ها است زیرا نیاز به نصب و راه‌اندازی اولیه منابع نداشته و مقیاس‌پذیری و انعطاف‌پذیر هستند، اما مبحث توازن بار و زمان‌بندی مناسب در سیستم‌های توزیعی یک چالش می‌باشد. در این مقاله روشی برای زمان‌بندی وظایف بر روی منابع در دسترس پویا ارایه شده و سیستم از یادگیری مداوم برای بهترین عملکرد استفاده می‌کند. در روش پیشنهادی از عملگر نقاد برای بهبود تصمیم‌گیری در یادگیری تقویتی استفاده شده تا استخراج قوانین توزیع و استفاده آن‌ها در یادگیری تقویتی سبب بهبود شود و اهداف بهبود مصرف انرژی را میسر سازد. روش پیشنهادی از نظر معیارهای "زمان اتمام تمامی کارها" و "میزان مصرف انرژی" با روش ارایه شده در کار مشابه مقایسه شد و در ارزیابی‌ها روش پیشنهادی به نسبت روش مورد مقایسه میزان مصرف انرژی مناسب‌تری داشته است. البته در محیط‌هایی که طول صف تشکیل شده و منابع و درخواست‌ها سریع تغییر می‌کند به دلیل افزایش تعداد حالات و یادگیری مستمر این مصرف انرژی کمی افزایش می‌یابد. در کل روش پیشنهادی مناسب محیط‌های پایدار، تغیرات کم و یا با فواصل زمانی متعادل‌تر است زیرا پروسه یادگیری مدت زمانی را به طول می‌انجامد.

کلمات کلیدی: سیستم‌های توزیعی، مصرف انرژی، عملگر نقاد، زمان‌بندی وظیفه، یادگیری تقویتی.

تاریخچه مقاله:

تاریخ ارسال: ۱۴۰۰/۰۳/۲۵
تاریخ اصلاحات: ۱۴۰۰/۰۵/۰۹
تاریخ پذیرش: ۱۴۰۰/۰۶/۰۲
تاریخ انتشار: ۱۴۰۰/۰۶/۳۱

Keywords:

Distribution Systems,
Energy Efficiency,
Actor-Critic,
Task Scheduling,
Reinforcement Learning

*ایمیل نویسنده مسئول:
arashalibakhshifard@gmail.com

Energy Efficiency in Distribution Systems Based on Task Scheduling using Reinforcement Learning and Actor-Critic Method

Arash Ali Bakhshifard*[1], Lida Naderloo[2] and Zahra Tayyebi Qasabeh[3]
[1]Department Engineering computer,Islamic Azad university of Zanjan Branch, Zanjan,Iran.
[2]Department computer, non-profit higher education institutions Rouzbeh of Zanjan Branch, Zanjan, Iran
[3]Supervisor, Department Engineering computer, Payame noor university of Guilan,Iran.

Abstract

Energy consumption in data centers and systems is increasing rapidly, which is a fundamental issue in the present age. An important advantage of distribution systems is cost savings because they do not require the initial installation and commissioning of resources and are scalable and flexible, but Load balance and scheduling are a challenge in distribution systems. This paper presents a method for scheduling tasks on dynamically available resources and the system uses continuous learning for best performance. In the proposed method, the Actor-Critic is used to improve decision making in reinforcement learning to extract the rules of distribution and use them in reinforcement learning to improve and facilitate energy efficiency goals. The proposed method was compared with the method presented in the same work in terms of "Completion time of all tasks " and "energy consumption" criteria. In th e evaluations, the energy consumption of the proposed method was more appropriate than the compared method. In environments where queue length is formed and resources and requests change rapidly, this energy consumption increases slightly due to the increasing number of scenarios and continuous learning. In general, the proposed method is suitable for stable environments, low changes or more balanced time intervals Because the learning process takes time.

۱ – مقدمه

عملکرد سیستم‌های کامپیوتری امروزه دارای اهمیت به‌سزایی می‌باشد. گاهی مدارات مجتمع و اجزای کامپیوتر روزبه‌روز گسترش یافته و افزایش می‌یابد. هم‌چنین مسأله افزایش اجزای کامپیوتر و سیستم‌های توزیع منجر به بالا رفتن مصرف انرژی شده است. مراکز داده گزارش داده‌اند که زیرساخت پایه برای تولید امکانات کاربردی در مقیاس اینترنت حدود ۱/۳ دقیقه از برق دنیا را در سال ۲۰۱۱ مصرف می‌کنند و این مقدار در سال ۲۰۲۰ به ۸٪ رسید. چندین روش برای بهینه‌سازی مصرف انرژی در سیستم‌های توزیع پیشنها شده است. یکی از روش‌های رایج در این زمینه تنظیم پویایی قدرت و توان اجزای کامپیوتر می‌باشد. روش جدیدتر پیمایش سرعت پویا است که فرکانس‌های اجزای کامپیوتر را در دامنه خاص مانند زمان خواب کامپیوتر تنظیم می‌کند[۱]. روش دیگری که برای بهینه-سازی مصرف انرژی در سیستم‌های توزیع استفاده می‌شود، زمان-بندی وظیفه در این سیستم‌ها می‌باشد. در این پژوهش بهینه‌سازی جامع جهت مصرف انرژی لایه کاربردی و نرم‌افزاری در سیستم‌های توزیع‌شده استفاده می‌گردد.

اغلب روش‌های پیشنهادی موجود که در زمینه کیفیت سیستم پژوهش انجام داده‌اند، بیشتر در پیش‌بینی درخواست ورود یا ویژگی‌های افزایش کیفیت سیستم کار کرده‌اند. در واقع هیچ‌کدام از روش‌ها دقت پیش‌بینی را تضمین نمی‌کنند. هم‌چنین با رشد جمعیت سرویس‌ها در اینترنت، تکنیک‌های بهینه‌سازی متمرکز سنتی مانند پویا و بهینه‌سازی ترکیبی نمی‌توانند در سیستم‌ها با ابعاد بالای سرویس‌دهی که دارای پیچیدگی زمانی زیادی هستند، پاسخ‌گو باشند. یکی از راه‌حل‌ها برای رفع چالش‌های بیان شده، زمان‌بندی برای بهینه‌سازی انرژی در سیستم‌های توزیع می‌باشد. ایجاد یک مکانیزم مدیریت خدمات بدون نیاز به اطلاعات آماری درخواست ورود می‌تواند در بهینه‌سازی مؤثر باشد[۲].

در این پژوهش برای زمان‌بندی وظیفه در سیستم‌های توزیع زنجیره مارکوف ایجاد شده و سپس یادگیری تقویتی که توانایی عمل در فضاها با مقیاس بزرگ و پیوسته را دارد، اعمال می‌گردد. به‌منظور تعیین توابع ارزشی و سیاست‌گذاری‌ها در یادگیری تقویتی و انتخاب پارامترهای مناسب آن از روش عملگر-نقاد[1] استفاده می-شود.

بخش‌های این مقاله شامل کارهای مرتبط، روش پیشنهادی و در انتها نتایج تجربی که شامل پیاده‌سازی روش پیشنهادی، ارزیابی‌ها و تجزیه و تحلیل داده‌ها، در نهایت بخش آخر نتیجه‌گیری ارایه شده است.

۲ – کارهای مرتبط

مومنی[2] و همکارانش(۲۰۲۱) مقاله ای تحت عنوان "رویکرد زمانبندی وظیفه در زمان واقعی با آگاهی از انرژی در محیط رایانش ابری " ارائه دادند. آن‌ها معتقدند علاقه به رایانش ابری طی سالهای اخیر بطور قابل ملاحظه ای افزایش یافته است، دلیل اصلی این امر منابع مجازی مقیاس پذیر است. بنابراین، رایانش ابری به پیشرفت برنامه‌های در زمان واقعی مانند پردازش سیگنال، نظارت بر محیط و پیش بینی آب و هوا که در آن ملاحظات زمان و انرژی برای انجام کارها بسیار مهم است، کمک کرده است. در برنامه‌های زمان واقعی، از دست دادن مهلت تعیین شده برای کارها عواقب فاجعه باری ایجاد می‌کند، بنابراین، برنامه‌ریزی زمان واقعی کار در محیط رایانش ابری یک مسئله مهم و اساسی است. علاوه‌براین، صرفه‌جویی در مصرف انرژی در مرکز داده‌های ابری، با توجه به مزایایی مانند کاهش هزینه‌های عملیاتی سیستم و حفاظت از محیط زیست، مسئله مهمی است که طی سالهای اخیر مورد توجه قرار گرفته و با برنامه-ریزی مناسب کار قابل کاهش است. در این مقاله ، یک رویکرد برنامه ریزی وظیفه آگاه از انرژی، یعنی EaRT[3] برای برنامه‌های زمان واقعی ارائه شده است . این روش از تکنیک مجازی‌سازی و تلفیق استفاده می‌کند تا بتواند مصرف انرژی را به حداقل برساند، استفاده از منابع را بهبود بخشد و مهلت مقررات را رعایت کند. در روش تلفیق، بالا و پایین آوردن منابع مجازی‌سازی می‌تواند عملکرد اجرای کار را بهبود بخشد. روش پیشنهادی شامل چهار الگوریتم برنامه‌ریزی وظیفه آگاه از انرژی در رایانش ابری[4] (ETC)، ماشین مجازی مقیاس عمودی[5] (V2S) ، ماشین مجازی مقیاس افقی[6] (HVS)و مقیاس پایین فیزیکی ماشین[7] (PSD) است. برای اثبات دقیقاً ویژگی زمانبندی و صحت EaRTها ، مدل رسمی روش پیشنهادی با استفاده از زمانبندی خودکار ارائه شده است. نتایج نشان می‌دهد که روش پیشنهادی در استفاده از منابع و مصرف انرژی در مقایسه با سایر الگوریتم‌های برنامه‌ریزی کارهای در زمان واقعی آگاه از انرژی کارآمدتر است[۱۱].

مه[8] و همکارانش (۲۰۲۱) مقاله‌ای تحت عنوان "برنامه ریزی وظیفه آگاه از انرژی با مهلت محدودیت در خوشه های ناهمگن با قابلیت DVFS" ارائه دادند. آن‌ها معتقدند که صرفه‌جویی در انرژی در مراکز داده بزرگ برای بارهای محاسباتی با کارایی بالا، مانند

[1]Actor – Critic Algorithm
[2]Momeni
[3]Energy-Aware Real-Time Task
[4]Energy-Aware Task Scheduling in Cloud Computing
[5]Vertical VM Scale Up
[6]Horizontal VM Scale up
[7]Physical Machine Scale Down
[8]Mei

یادگیری عمیق با داده های بزرگ، از اهمیت ویژه‌ای برخوردار است ، جایی که کاهش چند درصد برق به صرفه‌جویی میلیونی تبدیل می‌شود. این مقاله مطالعه صرفه‌جویی در مصرف انرژی در خوشه های ترکیبی CPU-GPU از طریق اندازه‌گیری ولتاژ و فرکانس پویا[1] (DVFS) است. هدف این مقاله این است که کل انرژی مصرفی پردازش دسته‌ای از کارهای آفلاین یا دنباله‌ای از کارهای زمان واقعی تحت محدودیت‌های مهلت به حداقل برسد. در این مدل از مدل تحلیلی سریع و دقیق برای محاسبه تنظیم ولتاژ / فرکانس مناسب برای هر کار استفاده می‌شود و چندین کار با الگوریتم‌های برنامه‌ریزی ابتکاری به خوشه اختصاص داده می‌شود. به طور خاص ، این مدل بر رابطه غیر خطی بین زمان اجرای کار و سرعت پردازنده برای برنامه‌های شتاب دهنده GPU تأکید می‌کند، تا با دقت بیشتری مصرف انرژی GPU در دنیای واقعی را به دست آورد. در ارزیابی عملکرد مبتنی بر ردیابی‌های اندازه‌گیری توان در دنیای واقعی، این الگوریتم زمان‌بندی صرفه‌جویی در انرژی قابل مقایسه با حد بالای نظری را نشان می‌دهد با فاصله مقیاس‌گذاری GPU که از نظر تحلیلی حداکثر ۳۶٪ انرژی می‌تواند صرفه‌جویی کند، ۳۳-۳۵٪ پس‌انداز انرژی را به ثبت رساند. نتایج این مقاله برای مدیریت انرژی بر روی خوشه‌های ناهمگن مدرن قابل استفاده است. [۱۲].

پژمان[2] و همکارانش (۲۰۲۰) مقاله ای تحت عنوان "رویکرد جدیدی برای برنامه‌ریزی وظایف آگاه از انرژی در محاسبات مه با استفاده از الگوریتم ترکیبی فرا ابتکاری" ارائه دادند. آن‌ها معتقدند که در سال‌های اخیر ، مشکلات محاسباتی بزرگی توسط محیط توزیع شده‌ای که برنامه‌ها به طور موازی در آن اجرا می‌شوند، حل شده است. همچنین، اخیراً، محاسبات مه یا محاسبات لبه‌ای به عنوان یک محیط جدید برای جمع‌آوری داده‌ها از دستگاه‌ها اعمال می‌شود و قبل از ارسال برای پردازش اصلی در محاسبات ابری، پیش پردازش انجام می شود. از آنجا که یکی از موضوعات مهم در چنین سیستم‌هایی، برنامه‌ریزی وظایف است، با در نظر گرفتن کاهش مصرف انرژی، این مسئله حل می‌شود. در این مقاله، یک روش آگاه از انرژی با استفاده از تکنیک مقیاس‌گذاری ولتاژ و فرکانس پویا برای کاهش مصرف انرژی معرفی شده است. علاوه‌براین، به منظور ساخت توالی‌های کار معتبر، از الگوریتم تکاملی بهینه‌سازی و کشت علف‌های هرز مهاجم استفاده شده است. نتایج تجربی نشان داد که الگوریتم پیشنهادی برخی الگوریتم‌های فعلی را از نظر مصرف انرژی بهبود می‌بخشد[۳].

ژیهائو[3] و همکارانش (۲۰۲۰) مقاله‌ای تحت عنوان "برنامه ریزی آگاه از انرژی برای گردش کار با استفاده از یک روش ابتکاری در ابر سبز " ارائه دادند. آن‌ها معتقدند که مصرف انرژی یکی از اصلی‌ترین نگرانی‌ها برای حمایت از رشد سریع مراکز داده ابری بوده است، زیرا نه تنها هزینه برق را برای تأمین کنندگان خدمات افزایش می‌دهد بلکه در افزایش انتشار گازهای گلخانه‌ای و در نتیجه آلودگی محیط زیست نیز نقش مهمی دارد و دارای تأثیر منفی بر قابلیت اطمینان و در دسترس بودن سیستم دارد. در نتیجه، معیارهای مصرف انرژی و کارایی به یک مسئله حیاتی برای برنامه‌های موازی برنامه ریزی مبتنی بر وظایف انجام شده در مراکز داده ابری تبدیل شده است. در این مقاله، یک الگوریتم زمان‌بندی دو فازی با زمان و انرژی به نام بهترین برنامه‌ریزی اکتشافی برای برنامه‌ریزی نمودار چرخش مستقیم در پردازنده‌های مرکز داده ابر ارائه شده است. در مرحله اول، الگوریتم بر اساس چهار روش ابتکاری و الگوریتم ملخ، منابع را به ترتیب با مرتب‌سازی به کارها اختصاص می‌دهد. سپس برای دستیابی به راه حل طراحی شده در زمان مناسب با توجه به فاکتور مهمی که توسط کاربر نهایی یا ارائه دهنده خدمات تعیین می‌شود و مناسب‌ترین روش را برای انجام هر کار انتخاب می‌کند. در مرحله دوم ، بهترین برنامه‌ریزی ابتکاری[4] با توجه به فاکتور مهمی که توسط کاربر نهایی یا ارائه دهنده خدمات تعیین می‌شود و با در نظر گرفتن زمان شروع، زمان راه‌اندازی، زمان پایان و مشخصات انرژی ماشین‌های مجازی، میزان مصرف و انرژی را به حداقل می‌رساند. سرانجام، یک مجموعه داده آزمایشی برای ارزیابی الگوریتم پیشنهادی BHS در مقایسه با الگوریتم تخصیص منابع چندمرحله ای[5] ساخته شد. نتایج نشان می‌دهد که الگوریتم پیشنهادی ۱۹.۷۱٪ ذخیره انرژی بیشتر از الگوریتم MHRA را داشته است. علاوه‌براین، انرژی در محیط های ناهمگون ۵۶/۱۲ درصد کاهش می‌یابد[۴].

القمدی[6] و همکاران (۲۰۱۷)، یک تکنیک زمان‌بندی دوفاز برای سیستم‌های توزیع زمان واقعی پیشنهاد داده‌اند. این روش در دو فاز کاملاً جداگانه بررسی می‌شود. فاز اول، مسئول تولید یک توالی برنامه‌ریزی و فاز دوم، ارسال وظایف برای گره‌های محاسبه‌گر می‌باشد. هم‌چنین فاز دوم زمان شروع هر وظیفه را نیز تعیین می‌کند. یکی از ویژگی‌های مهم این روش پیشنهادی انعطاف‌پذیری بالای آن می‌باشد که توسعه‌دهندگان برنامه را در سیاست‌گذاری‌های چندگانه آزاد می‌گذارد. دو فاز نام برده شده در این روش، کاملاً مجزا بوده و به یکدیگر وابسته نمی‌باشند. بنابراین با تغییر یک سیاست در یک فاز، نیاز به ایجاد تغییر در فاز دیگر نیست. در این مقاله سه سیاست در فاز اول و دو سیاست برای زمان‌بندی در فاز

[1]Dynamic Voltage and Frequency Scaling
[2]Pejman
[3]Zhihao
[4]Best Heuristic Scheduling
[5]Multiheuristic Resource Allocation Algorithm
[6]Almaghdi

دوم در نظر گرفته شده است. نتایج نشان‌گر عملکرد بهتر با استفاده از این روش می‌باشند[۵].

تیان[1] و همکاران (۲۰۱۷)، یک روش ارسال پیام میان گره‌های انرژی را پیشنهاد داده‌اند. این روش، احتمال مؤفقیت در ارسال پیام را قبل از ارسال، پیش‌بینی کرده و استراتژی آنلاین را برای نرخ انرژی ورودی تنظیم می‌نماید. نتایج نشان می‌دهد که مصرف انرژی در این روش نسبت به روش‌های سنتی در سیستم‌های توزیع که توان ورودی غیرپایدار دارند، کمتر بوده‌است[۶].

شیائو[2] و همکاران (۲۰۱۶)، روشی را بر مبنای تکنیک طراحی ولتاژ پویا به‌منظور کاهش مصرف انرژی و مقیاس‌گذاری فرکانس ارائه داده‌اند. این مقاله برای کاهش طول زمان‌بندی مصرف انرژی طراحی شده است. مسأله بهینه‌سازی انرژی به دو زیرمسأله تقسیم می‌شود که شامل کاهش مصرف انرژی در تجهیزات و هم کاهش پیچیدگی زمانی برای هر وظیفه می‌شود. نتایج با استفاده از تبدیل فوریه سریع نشان داد که کاهش طول زمان‌بندی و مصرف انرژی با روش این مقاله، اتفاق افتاده‌است[۷].

تاناوانیش[3] و همکاران (۲۰۱۶)، برای کاهش مصرف انرژی منابع محاسباتی در سیستم‌های توزیع با استفاده از نرخ تأثیر[4] و تابع هزینه که آستانه کاهش منابع می‌باشد، الگوریتمی را پیشنهاد داده‌اند. در این روش، ابتدا به بهینه‌سازی پردازنده‌های غیر بهینه پرداخته‌شده و سپس جریان کار دوباره زمان‌بندی را انجام می‌دهد. عملکرد الگوریتم پیشنهادی با استفاده از تخمین جامع پردازشی بررسی شده و مشاهده گردید که نرخ مصرف انرژی کاهش داشته‌است[۸].

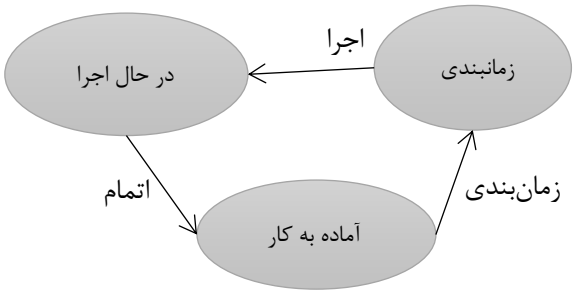
چن[5] و همکاران (۲۰۱۵)، پس از بررسی معیارهای متناقض عملکرد سیستم‌ها نظیر ازدحام در صف و مصرف انرژی به ارائه یک روش در مدیریت و زمان‌بندی آنلاین سیستم‌های توزیع پرداخته‌اند. این روش نیازی به اطلاعات آماری و ورود درخواست ندارد. تحلیل ریاضی در این روش می‌تواند تعادل میان عملکرد بهتر و کاهش مصرف انرژی در سیستم‌های توزیع ایجاد کند. اطلاعات ردیابی عددی و واقعی براساس تجربه در بهینه‌سازی مصرف انرژی در سیستم پیشنهادی مؤثر بوده‌است[۲].

جی‌وای[6] و همکاران (۲۰۱۵)، برای کاهش مصرف انرژی با استفاده از پیمایش سرعت پویا و زمان‌بندی وظیفه روشی را پیشنهاد نموده‌اند. مدل‌های مارکوف برای سیستم‌های محاسباتی توزیع‌شده پیشنهاد و جزئیات پردازش ارائه شده است. در این مقاله فرآیندهای تصمیم مارکوف برای فرمول‌سازی استفاده شده و روش‌های حل مسأله بهینه‌سازی از این طریق بیان شده است[۱].

۳ – تحلیل روش پیشنهادی

این مقاله به‌دنبال یافتن قوانین زمان‌بندی مؤثر در سرعت بخشیدن و افزایش کارایی سیستم می‌باشد. برای این منظور از روش یادگیری تقویتی و روش عملگر-نقاد برای یافتن بهینه‌ترین قوانین استفاده می‌شود. از آنجایی که یادگیری تقویتی یک روش یادگیری برای پیش‌بینی ارائه می‌دهد، این مدل، تکنیکی است که می‌تواند با استفاده از آموزشی که می‌بیند اطلاعات آینده را پیش‌بینی نماید و همچنین قادر است در فضاهای بزرگ محاسباتی قدرتمند عمل نماید یادگیری تقویتی یعنی آموزش آنچه باید انجام شود، که تصمیم گیرنده را عامل هوشمند و محیط که عامل با آن تعامل دارد. از طرفی با استفاده از روش عملگر-نقاد توابع سیاست‌گذاری و ارزش‌ها آموزش داده می‌شود. زیرا از آنجائیکه معماری عملگر نقاد از اصول یادگیری-تقویتی تفاضل موقتی استفاده می‌نماید، قابلیت پیاده‌سازی به صورت زمان حقیقی در طی مسیر سیستم را دار می‌باشد. در این معماری ساختار سیاست به عنوان عملگر شناخته می‌شود زیرا از آن برای تولید عمل استفاده شده و ساختار تابع ارزش به عنوان نقاد شناخته می‌شود زیرا برای نقد اعمال انجام گرفته توسط عملگر بکار گرفته می‌شود. لذا می‌توان از این طریق هم زمان‌بندی وظایف در سیستم‌های توزیع را انجام داده و هم از نظر مصرف انرژی، آن را بهینه نمود.

زمان‌بندی تطبیقی براساس یادگیری تقویتی: وضعیت سیستم شبکه‌ای را می‌توان یک فرآیند تصادفی توصیف کرد که وضعیت سیستم را به‌عنوان تابعی از زمان مشخص می‌کند. فضای وضعیت سیستم مشخص است (در حال فعالیت، آماده به کار و در حال زمانبندی) و زمان‌بندی وظایف پیشنهاد شده براساس وضعیت فعلی سیستم انجام می‌گیرد. به‌طورکلی، فرآیند زمان‌بندی وظیفه را می‌توان یک MDP در نظر گرفت، که در آن می‌توان انواع مختلفی از زمان‌بندی کنترل مبتنی بر RL را برای سیستم به‌کار برد.



(شکل-۱): چارچوب کلی روش پیشنهادی

[1]Tian
[2]Xiao
[3]Thanavanish
[4]Rate of Effectiveness (RE)
[5]Chen
[6]Jiwei

یادگیری تقویتی بدون مدل: مسئلهٔ یادگیری تقویتی به‌صورت فرآیند تصمیم‌گیری مارکوف بیان می‌شود، که با چندتایی S, A,T, r توصیف می‌گردد، که در آن S عبارت است از فضای وضعیت محیط، A مجموعه‌ای از اقدامات احتمالی است. T تابع گذار[1] می‌باشد که احتمال گذار در وضعیت s با اقدام a را مشخص می‌کند، و r تابع پاداش است که پاداش فوری پس از گذار به یک وضعیت جدید از وضعیت s با اقدام a را تعیین می‌کند[۹].

به‌طورکلی، هدف یادگیرنده برای یادگیری سیاست بهینه این است که پاداش کلی را به حداکثر برساند. یادگیری–Q، یک الگوریتم یادگیری تقویتی است که در هر مرحله از یادگیری، ارزش Q– توسط (رابطه-۱) به‌روزرسانی می‌شود[۹]:

$$Q(s_t.a_t) \leftarrow Q(s_t.a_t) + \alpha[r_{t+1} + \gamma max_a Q(s_{t+1}.a) - Q(s_t.a_t)]$$

(رابطه-۱)

که در آن $\alpha \in [0.1]$ عبارت است از نرخ یادگیری، که حدود یادگیری اطلاعات جدید توسط یادگیرنده‌ها را تعیین می‌کند، و $\gamma \in [0.1]$ عامل تخفیف[2] است که وزن پاداش‌های آینده در به‌روزرسانی و r پاداش فوری است [۹]. مقدار ۱ برای نرخ یادگیری یعنی عامل جدیدترین اطلاعات را در نظر می‌گیرد، اما مقدار صفر برای فاکتور تخفیف نشان می‌دهد که عامل صرفا پاداش فعلی را در نظر می‌گیرد، درحالی‌که نزدیک شدن این مقدار به ۱ باعث خواهد شد که عامل سعی کند پاداش بلندمدت بالایی کسب کند.

زمان‌بندی تطبیقی دو مرحله‌ای براساس عملگر نقاد و یادگیری تقویتی: کاهش هزینه‌های دسترسی به داده‌ها (زمان مورد نیاز برای دسترسی به داده‌ها جهت پردازش وظایف) در شبکهٔ داده‌ای، نقش مهمی در کاهش زمان تکمیل وظایف و بهبود عملکرد زمان‌بندی دارد. مدل محاسبهٔ هزینهٔ دسترسی به داده‌ها در شبکهٔ داده‌ای مبتنی بر خوشه به صورت زیر تعریف می‌شود[۹]:

اگر وظیفهٔ Ti بر روی گرهٔ Sj زمان‌بندی شده باشد، هزینهٔ ارتباط داده برای دسترسی به داده‌های مورد نیاز Ti از Sj توسط (رابطه-۲) تعیین می‌شود [۹]:

$$DCC_{Sj}^{Tj} = \sum_{For\ all\ Flk\ in\ Rj} |Fl_k|/B_{jk}$$

(رابطه-۲)

که در آن Ri عبارت است از فهرست فایل‌های رونوشت/دادهٔ مورد نیاز جهت پردازش وظیفه، FlK عبارت است از رونوشت دادهٔ kth در Ri، $|Fl_k|$ اندازهٔ رونوشت است و BjK پهنای باند شبکه بین گره S و گرهٔ منبع رونوشت kth می‌باشد. بر طبق توپولوژی مبتنی بر خوشهٔ شبکهٔ داده‌ای، داده‌های مورد نیاز ممکن است از گره‌های مختلف در سایر خوشه‌ها یا از همان خوشه به گره Sj منتقل شوند. مدل هزینهٔ ارتباط داده را می‌توان به‌صورت (رابطه-۳) تعریف کرد:

$$DCC_{Sj}^{Ti} = Inter_C_{Ti,Sj} + Intra_C_{Ti,Sj}$$

(رابطه-۳)

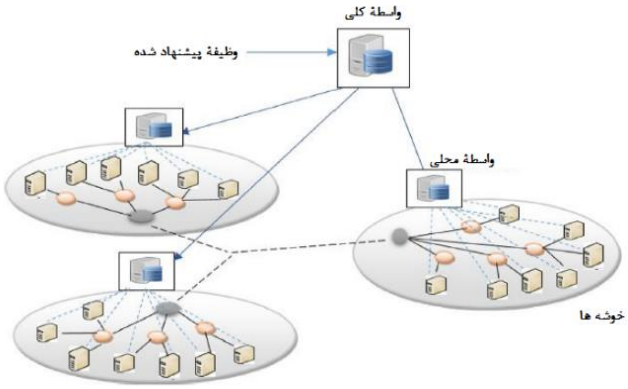
که در آن، $Inter_C_{Ti.Sj}$ عبارت است از هزینهٔ ارتباط داده برای دسترسی به رونوشت‌های مورد نیاز باقیمانده در خوشه‌های مختلف از خوشهٔ اصلی Ti و $Intra_C_{Ti.Sj}$ عبارت است از هزینهٔ ارتباط داده برای رونوشت‌های مورد نیازی که در خوشهٔ محلی باقی مانده‌اند. ناحیهٔ وسیع بین خوشه‌ها معمولا بسیار کُندتر از شبکه‌های محلی در داخل یک خوشه است. بنابراین، هزینهٔ ارتباط داده بین خوشه‌ها بیشتر از هزینهٔ ارتباط داده در داخل خوشه می‌باشد. در نتیجه، کاهش تعداد ارتباط داده بین خوشه‌ها برای دسترسی به داده‌های توزیع‌یافته، برای زمان‌بندی وظیفهٔ داده-فشرده در سیستم‌های شبکه‌ای مبتنی بر خوشه بسیار مهم است.

الگوریتم پیشنهادی، یک زمان‌بندی بی‌واسطهٔ وظیفه براساس یک فرآیند تصمیم‌گیری دو مرحله‌ای می‌باشد، که اولین مرحلهٔ آن انتخاب خوشه‌ای است که شامل گره‌ای (گره‌هایی) با کمترین هزینهٔ ارتباط داده باشد. عامل تعیین‌کنندهٔ کل هزینهٔ ارتباط داده، هزینهٔ ارتباط داده‌ها بین خوشه‌های مختلف است. در مرحلهٔ بعد، برای افزایش تطابق زمان‌بندی با این محیط پویا که شامل حجم کارهای پویا با فواصل زمانی متغیر وظیفه، و نامتجانسی شدید وظایف و منابع پیشنهاد شده می‌باشد، به‌منظور انتخاب گرهٔ صحیح در خوشهٔ انتخاب شده‌ای که حداقل هزینهٔ ارتباط داده را دارد، با استفاده از یادگیری–Q از یک سیاست تخصیص وظیفهٔ مبتنی بر یادگیری تقویتی تطبیقی با عملگر نقاد استفاده می‌شود.

این زمان‌بندی تطبیقی پیشنهادی برای زمان‌بندی پردازش از یک سیستم سلسله‌مراتبی چندعاملی استفاده می‌کند و شامل یک عامل واسطهٔ کلی در سطح اول و چندین عامل واسطهٔ محلی در سطح دوم در داخل خوشه می‌باشد. واسطهٔ کلی، خوشهٔ صحیحی را انتخاب می‌کند که دارای حداقل هزینهٔ ارتباط داده باشد. سپس، واسطهٔ محلی براساس یادگیری–Q برای انتخاب گرهٔ پردازش صحیح، در داخل خوشهٔ انتخاب شده به استخراج یک سیاست تخصیص وظیفهٔ تطبیقی می‌پردازد. (شکل-۲) بیانگر ساختار کلی زمان‌بندی پیشنهادی است. واسطهٔ محلی وضعیت فعلی را مشاهده کرده، یک گرهٔ صحیح را به‌عنوان اقدام احتمالی انتخاب نموده، سپس پاداش دریافت می‌کند و ارزش‌های–Q را به‌روزرسانی می‌نماید.

[1] Transition Function

[2] Discount Factor



(شکل-۲): ساختار زمان‌بندی پیشنهادی [۹]

وضعیت محیط با چندتایی $n_{s1}, n_{s2}, \ldots, n_{sn}$ تعیین می‌شود، که در آن n_{sn} بیانگر تعداد وظایف (در حال انتظار و درحال پردازش) در گرهٔ s_n خوشهٔ محلی است. اقدامات احتمالی با مجموعه‌ای از $\{a_{s1}. a_{s2}. \ldots . a_{sn}\}$ تعیین می‌شوند، که در آن a_{sn} بیانگر اقدام انتخاب گرهٔ s_n می‌باشد. تابع پاداش به‌صورت (رابطه-۴) تعریف می‌شود [۹]:

$$reward = 1/Completiontime \qquad \text{(رابطه-۴)}$$

که در آن Completetime عبارت است از زمان مورد نیاز برای تکمیل یک وظیفهٔ برنامه‌ریزی شده. در شیوهٔ مبتنی بر یادگیری-Q، می‌توان از استراتژی‌های مختلف انتخاب اقدام استفاده کرد. زمان تکمیل (زمان واکنش) وظیفهٔ T_i بر روی گرهٔ S_j را می‌توان با استفاده از (رابطه-۵) محاسبه نمود [۹]:

$$CT_{Sj}^{Ti} = W_{Ti} + DAC_{Sj}^{Ti} + P_{Ti} \qquad \text{(رابطه-۵)}$$

که در آن W_{Ti} زمان انتظار در صف (تاخیر صف‌بندی) برای رسیدن به خدمات پردازش، DAC_{Sj}^{Ti} هزینهٔ دسترسی به داده‌ها برای وظیفهٔ T_i، و P_{Ti} زمان پردازش فایل داده‌ها می‌باشند. در محیط شبکه، به‌منظور اجتناب از برخورد داده‌ها، باید جداسازی عملکرد انتقال داده‌ها از طریق لینک‌های ارتباطی تضمین شود. هزینهٔ دسترسی به داده‌ها بر حسب زمان انتظار برای دریافت مجوز انتقال داده‌ها به صورت (رابطه-۶) تعریف می‌شود [۹]:

$$DAC_{Sj}^{Ti} = DCC_{Sj}^{Ti} + Delay_DT_{Ti} \qquad \text{(رابطه-۶)}$$

که در آن $Delay_DT_{Ti}$ عبارت است از تاخیر انتقال داده‌ها در لینک‌های شبکه و DCC_{Sj}^{Ti} عبارت است از هزینهٔ ارتباط داده [۹].

تابع هدف: تابع هدف در نظر گرفته شده در این روش با توجه به مسئله زمان‌بندی وظایف در چنین سیستم‌هایی باید کاهش میزان زمان انتظار برای انجام کارها و افزایش توان عملیاتی باشد که بدین ترتیب یک تابع با دو پارامتر زمان و وظایف وجود دارد که با کاهش زمان، تعداد وظایف و بازدهی افزایش می‌یابد. بنابراین تابع هدف برابر (رابطه-۷) می‌باشد.

$$G = \alpha TP \qquad \text{(رابطه-۷)}$$

که در آن α ضریب زمان‌بندی، T میانگین زمان انتظار و P میانگین تعداد وظایف به اجرا رسیده در واحد زمان است.

۴- نتایج تجربی

در این قسمت نرم‌افزار و سخت افزار شبیه‌سازی معرفی می‌شود که از نرم‌افزار Matlab برای رسیدن به نتایج روش پیشنهادی استفاده شده است. همچنین سیستم استفاده شده در ارزیابی دارای پردازنده Intel Pentium T4400 @2.2 GHz Processor، حافظه Ram 3.0 GB، Hard Disk 320 GB و سیستم عامل Windows 7 بوده است.

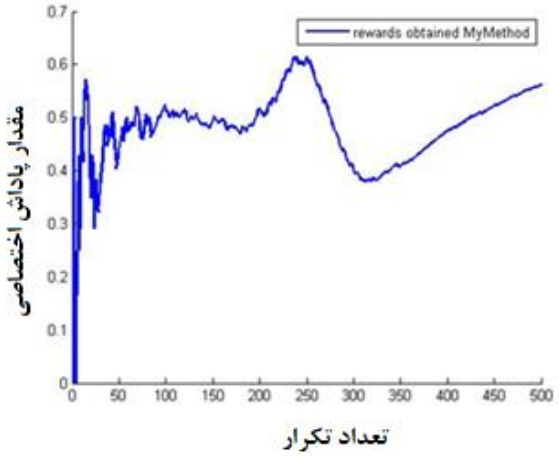
ارزیابی اول: روش پیشنهادی بعد از محاسبات اولیه و آماده‌سازی داده‌ها برای تصمیم‌گیری از یادگیری تقویتی به همراه عمل نقاد استفاده می‌نماید. ورودی مربوط به این مسئله وظایف درخواست شده و تعداد منابع اجرا و خروجی نیز ترتیب و اختصاص هر یک از وظایف به یکی از منابع و یا انتظار وظایف در صف یکی از منابع است. هدف نیز اتمام و اجرای سریع‌تر همه وظایف و کاهش مصرف انرژی و صرفه‌جویی می‌باشد.

الگوریتم یادگیری تقویتی سعی در چینش مناسب وظایف در منابع مورد نظر و در دسترس را برعهده دارد طوریکه که صرفه جویی انرژی هم حاصل گردد. پارامترهای در نظر گرفته شده برای این الگوریتم در (جدول-۱) نشان داده شده است.

(جدول-۱): پارامترهای الگوریتم یادگیری تقویتی

ردیف	نام پارامتر	توضیحات	مقدار
1	n	حداکثر تکرار	500
2	G	ماتریس پاداش	G = [2 -1 ; 0 1]
3	eta	مقدار حد اتا	0.0402
4	beta	مقدار حد بتا	0.402
5	maxRewards	حداکثر مقدار پاداش	1

مقادیر ثابت استفاده شده در الگوریتم از تکرارهای مکرر به دست آمده است و نشان دهنده بهترین خروجی با این مقادیر بوده است. نتیجه پاداش‌های اختصاص داده شده در این الگوریتم در (شکل-۳) نشان داده شده است.

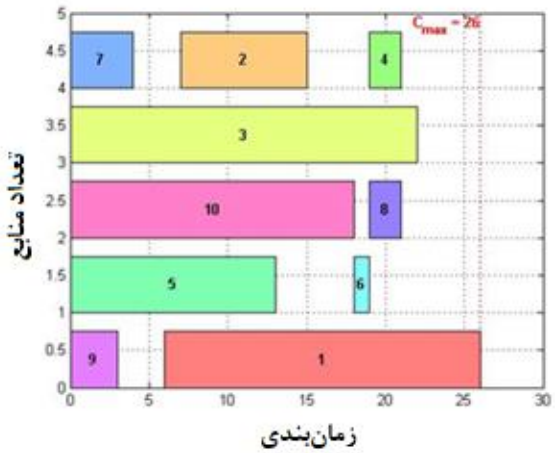


(شکل-۳): مقدار پاداش اختصاصی هر تکرار الگوریتم

همانگونه که در (شکل-۳) نشان داده شده است در ابتدا شروع اجرای الگوریتم مقدار پاداش اختصاصی پایین بوده و الگوریتم سعی در به حداکثر رسانی و میل آن به سمت عدد یک دارد که با پیشروی در تعداد تکرار به این امر نزدیک می‌شود.

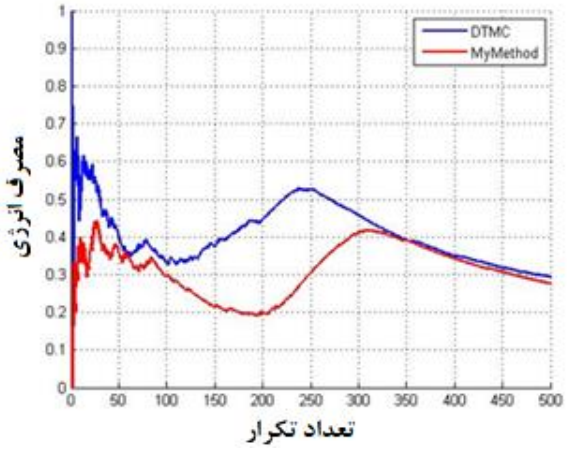
در گام بعدی کل الگوریتم با محیطی شامل ۵ منبع مختلف با توانایی یکسان اجرا شد. برای اجرا ده وظیفه با زمان‌های بارگذاری شده از مجموعه داده، به صورت تصادفی و با وابستگی‌های وظایف تولید شده استفاده شده است. نتیجه اجرا این سناریو روی روش پیشنهادی در (شکل-۴) نشان داده شده است. در کلیه ارزیابی‌ها روش پیشنهادی با روش DTMC[1] (زنجیره مارکوف با زمان‌بندی گسسته) [۱۰] مورد مقایسه قرار می‌گیرد.

همانگونه که در شکل ۴ نشان داده شده است زمان کامل شدن کلیه کارها ۲۶ میلی‌ثانیه است و طبق این خروجی وظایف ۱ و ۹ برروی منبع ۱، وظایف ۵ و ۶ برروی منبع ۲، وظایف ۸ و ۱۰ برروی منبع ۳، وظیفه ۳ برروی منبع ۴ و وظایف ۲،۴ و۷ بر روی منبع شماره ۵ اجرا می‌شود. ترتیب اجرا وظایف از سمت چپ بوده و دلیل این ترتیب اولویت و وابستگی کارهایی است که در نظر گرفته شده است.



(شکل-۴): زمان‌بندی ۱۰ وظیفه روی ۵ منبع روش پیشنهادی

در تکرارهای مختلف و استفاده از الگوریتم پیشنهادی نمودار مقایسه مصرف انرژی آن در شکل ۵ نشان داده شده است.

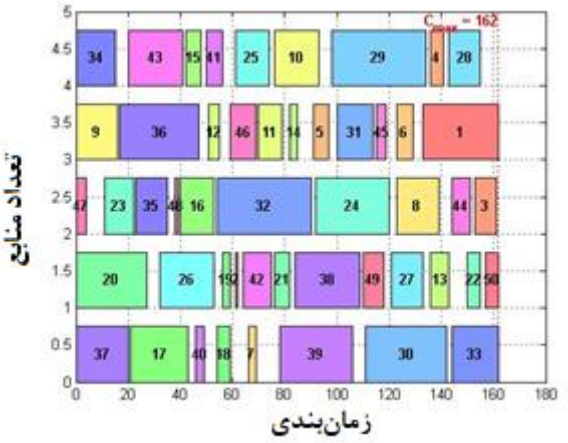


(شکل-۵): مقایسه مصرف انرژی روش پیشنهادی و روش DTMC برروی پنج منبع و ده وظیفه

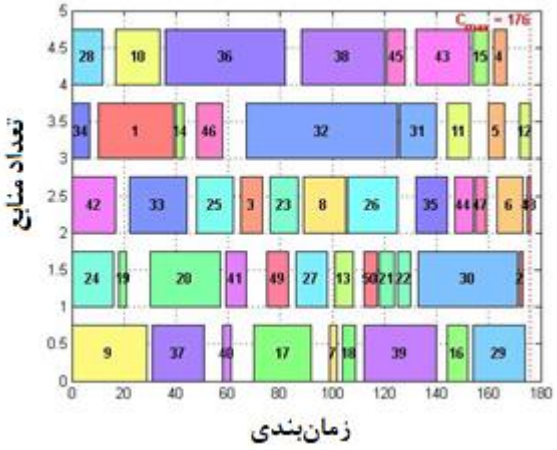
همانگونه که در (شکل-۵) مشخص است روش پیشنهادی از نظر بهبود بهینه‌سازی به نسبت روش مورد مقایسه دارای مصرف انرژی مناسب‌تری است.

در ادامه برای ارزیابی‌های بیشتر سناریوهای مختلفی را در نظر گرفته و مورد بررسی قرار گرفت. سه آزمایش مختلف با تعداد منابع ۵،۱۰ و ۲۰ انجام شد که در هر یک از آزمایش‌ها تعداد مختلفی وظیفه انتخاب شده و تولید شده به منابع جهت زمان‌بندی داده شده و از نظر زمان اتمام کار الگوریتم‌ها با یکدیگر مورد مقایسه قرار خواهند گرفت همچنین در این میان میزان مصرف انرژی دو الگوریتم نیز مقایسه می‌شود. نمونه‌ای از اجرای دو الگوریتم در شکل ۶ نشان داده شده است.

[1] Discrete-Time Markov Chain



الف) روش پیشنهادی

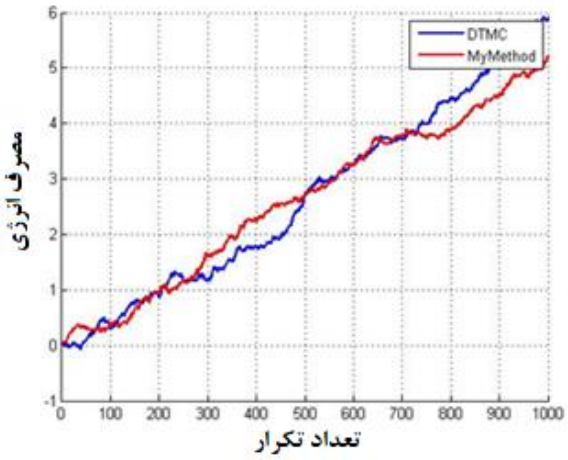


ب) روش DTMC

(شکل-۶): نتیجه اجرای روش‌های مورد مقایسه با پنج منبع و پنجاه وظیفه

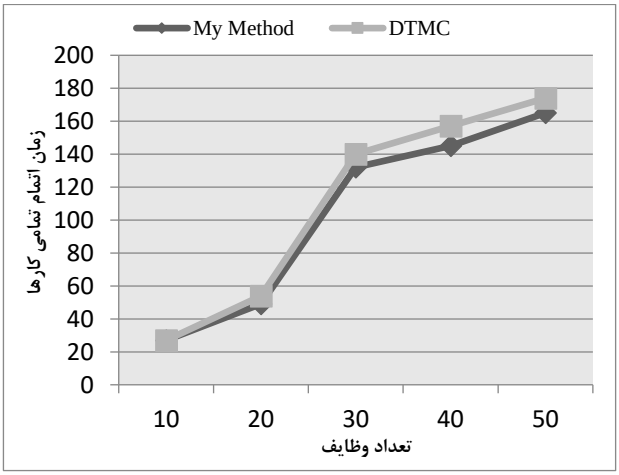
همانگونه که در (شکل-۶) نشان داده شده است روش پیشنهادی اجرای ۵۰ وظیفه در ۵ منبع را با ۱۶۲ میلی ثانیه به اتمام رسانده که ترتیب اجرای وظایف در منابع مختلف نیز در شکل مشخص است ولی روش DTMC همان تعداد وظایف و منبع با زمان‌ها و اولویت‌های برابر را در ۱۷۶ میلی ثانیه به اتمام رسانده است. نمودار مقایسه میزان مصرف انرژی دو روش در شکل ۷ نشان داده شده است.

همانگونه که در (شکل-۷) نشان داده شده است میزان مصرف انرژی روش DTMC و روش پیشنهادی نزدیک به یکدیگر بوده ولی در انتها این میزان مصرف انرژی در روش پیشنهادی کمتر بوده و مقدار بهتری را نشان می‌دهد البته دلیل این امر استفاده همزمان از یادگیری تقویتی و عملگر نقاد در روش پیشنهادی بوده که مقدار میزان مصرف انرژی را در انتها افزایش می‌دهد. همچنین تا به مرحله پایدار و یادگیری مناسب برسد کمی از نظر زمانی طول می‌کشد.



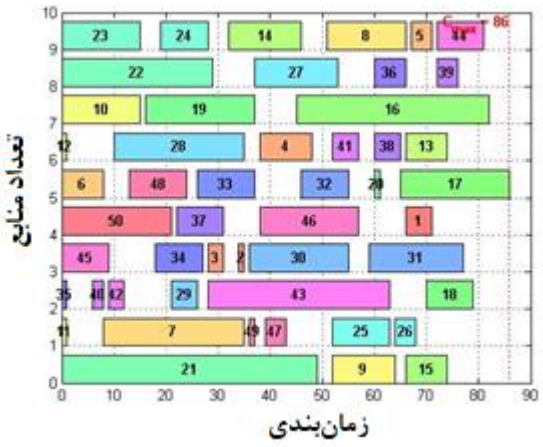
(شکل-۷): میزان مصرف انرژی روش‌های مورد مقایسه با پنج منبع و پنجاه وظیفه

نتایج مقایسه دو روش با تعداد وظایف مختلف در (شکل-۸) نشان داده شده است. همانگونه که در (شکل-۸) نشان داده شده است هر دو روش تعداد وظایف ۱۰ را با زمان یکسانی در ۵ منبع به اتمام رسانده‌اند ولی زمانی که تعداد وظایف افزایش پیدا می‌کند روش پیشنهادی زمان اتمام کار پایین‌تری را از خود نشان می‌دهد و این بهبود با افزایش تعداد وظایف بیشتر می‌شود. اختلاف دو روش در تعداد ۳۰ وظیفه ۱۱ میلی ثانیه در ۴۰ وظیفه ۱۳ و در ۵۰ وظیفه به ۱۴ میلی ثانیه رسیده است.

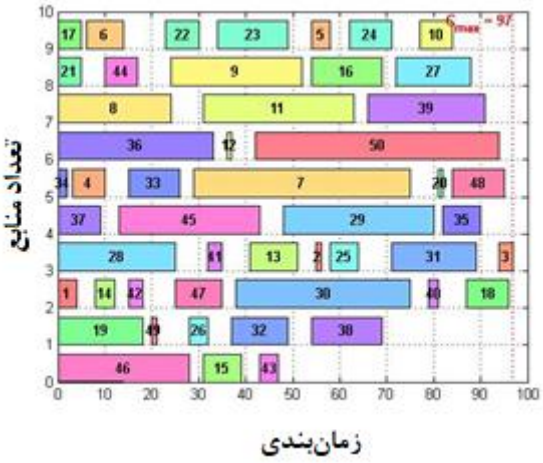


(شکل-۸):مقایسه روش پیشنهادی و روش DTMC با تعداد وظایف مختلف بر روی ۵ منبع

ارزیابی دوم: تعداد منابع به ۱۰ عدد افزایش داده شده است. برای تولید مجموعه درخواست‌ها از مجموعه داده ابتدایی استفاده شده است و مدل استخراجی به عنوان ورودی روش پیشنهادی مورد استفاده قرار گرفته است.



الف) روش پیشنهادی

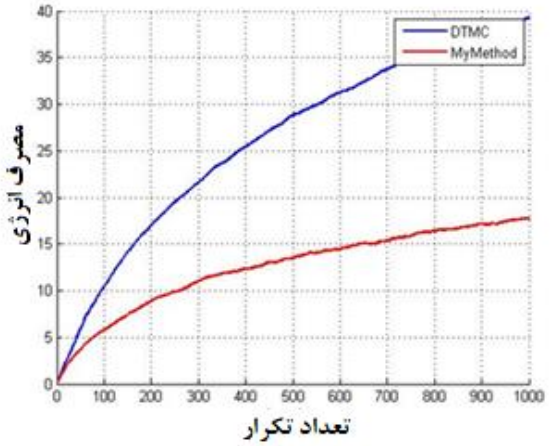


ب) روش DTMC

(شکل-۹): نتیجه اجرای روش‌های مورد مقایسه با ده منبع و پنجاه وظیفه

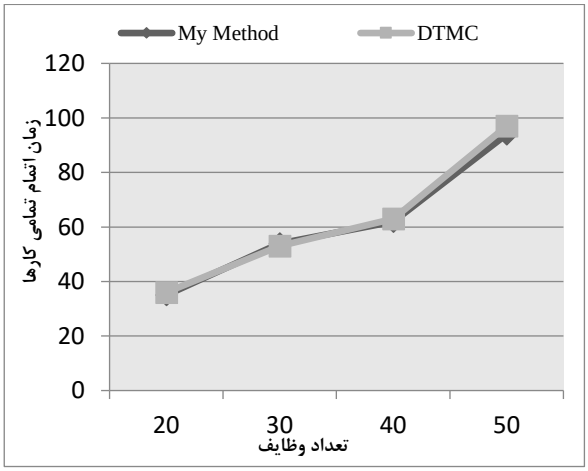
نمونه‌ای از اجرای دو الگوریتم در (شکل-۹) نشان داده شده و همانگونه که در شکل مشخص است روش پیشنهادی اجرای ۵۰ وظیفه در ۱۰ منبع را با ۸۶ میلی ثانیه به اتمام رسانده که ترتیب اجرای وظایف در منابع مختلف نیز در شکل مشخص است (به عنوان مثال وظایف ۲۲،۲۷،۳۶ و ۳۹ بر روی منبع ۹ اجرا می‌شود) ولی روش DTMC همان تعداد وظایف و منبع با زمان‌ها و اولویت‌های برابر را در ۹۷ میلی ثانیه به اتمام رسانده است. نمودار مقایسه میزان مصرف انرژی دو روش در (شکل-۱۰) نشان داده شده است.

همانگونه که در (شکل-۱۰) نشان داده شده است میزان مصرف انرژی روش DTMC از ابتدا بالاتر از روش پیشنهادی بوده است. در انتها نیز این میزان مصرف انرژی بالاتر و با اختلاف بیشتری بوده و مقدار بهتری را روش پیشنهادی نشان می‌دهد.



(شکل-۱۰): میزان مصرف انرژی روش‌های مورد مقایسه با ده منبع و پنجاه وظیفه

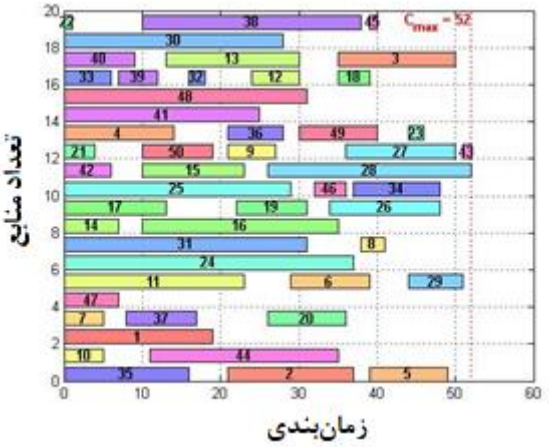
نتایج مقایسه دو روش با تعداد وظایف مختلف در (شکل-۱۱) نشان داده شده است.



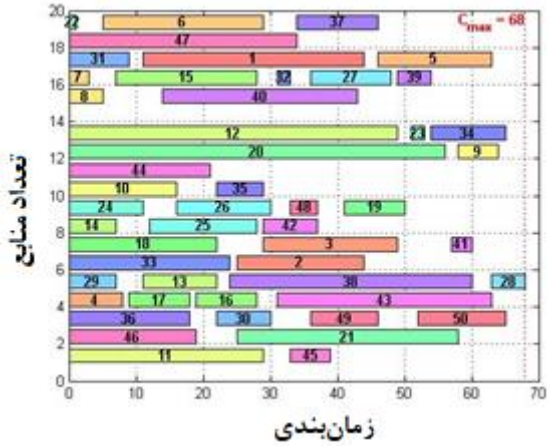
(شکل-۱۱): مقایسه روش پیشنهادی و روش DTMC با تعداد وظایف مختلف بر روی ده منبع

همانگونه که در (شکل-۱۱) نشان داده شده است اختلاف در تعداد وظایف پایین کمتر بوده ولی زمانی که تعداد وظایف افزایش پیدا می‌کند روش پیشنهادی زمان اتمام کار پایین‌تری را از خود نشان می‌دهد و این بهبود با افزایش تعداد وظایف بیشتر می‌شود. اختلاف دو روش در تعداد ۳۰ وظیفه ۵ میلی ثانیه در ۴۰ وظیفه ۴ و در ۵۰ وظیفه به ۱۱ میلی ثانیه رسیده است.

ارزیابی سوم: در ارزیابی بعدی تعداد منابع به ۲۰ عدد افزایش داده شده است. نمونه‌ای از اجرای دو الگوریتم در (شکل-۱۲) نشان داده شده است.



الف) روش پیشنهادی

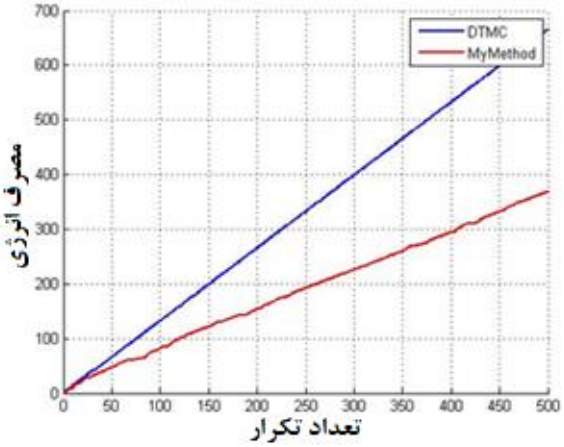


ب) روش DTMC

(شکل-۱۲): نتیجه اجرای روش‌های مورد مقایسه با بیست منبع و پنجاه وظیفه

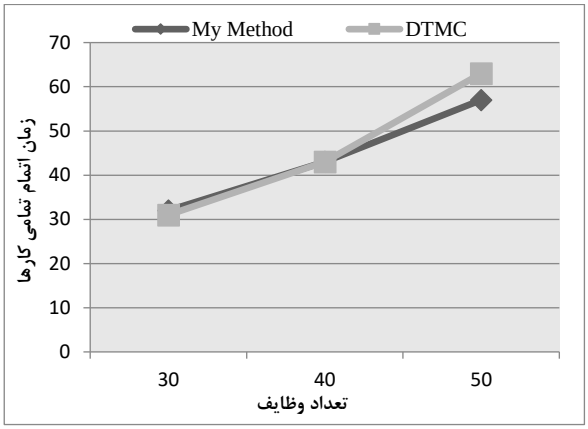
همانگونه که در (شکل-۱۲) نشان داده شده است روش پیشنهادی اجرای ۵۰ وظیفه در ۲۰ منبع را با ۵۲ میلی ثانیه به اتمام رسانده که ترتیب اجرای وظایف در منابع مختلف نیز در شکل مشخص است (به عنوان مثال وظایف ۱۷،۱۹و۲۶ برروی منبع ۱۰ اجرا می‌شود) ولی روش DTMC همان تعداد وظایف و منبع با زمان‌ها و اولویت‌های برابر را در ۶۸ میلی ثانیه به اتمام رسانده است. نمودار مقایسه میزان مصرف انرژی دو روش در شکل ۱۳ نشان داده شده است.

همانگونه که در شکل ۱۳ نشان داده شده است میزان مصرف انرژی روش DTMC در ابتدا بالاتر از روش پیشنهادی بوده و در انتها نیز این میزان مصرف انرژی بالاتر و با اختلاف بیشتری بوده و مقدار بهتری را روش پیشنهادی نشان می‌دهد البته دلیل این امر استفاده همزمان از یادگیری تقویتی و عملگر نقاد در روش پیشنهادی بوده که مقدار میزان مصرف انرژی را در انتها بهبود مناسبی می‌دهد. هدف از ارایه روش پیشنهادی کاهش مقدار زمان اتمام کار و صرفه‌جویی در مصرف انرژی و به اصطلاح بهینه سازی آن است.



(شکل-۱۳): میزان مصرف انرژی روش‌های مورد مقایسه با بیست منبع و پنجاه وظیفه

نتایج مقایسه دو روش با تعداد وظایف مختلف در (شکل-۱۴) نشان داده شده است.



(شکل-۱۴):مقایسه روش پیشنهادی و روش DTMC با تعداد وظایف مختلف بر روی ده منبع

همانگونه که در (شکل-۱۴) نشان داده شده است اختلاف در تعداد وظایف پایین‌تر مقدار کمتری نتیجه داده است طوریکه در تعداد ۳۰ وظیفه هر دو روش در ۳۰ میلی ثانیه تمامی کارها را به پایان رسانده‌اند، ولی زمانی که تعداد وظایف افزایش پیدا می‌کند روش پیشنهادی زمان اتمام کار پایین‌تری را از خود نشان می‌دهد که این بهبود روش پیشنهادی در افزایش تعداد وظایف را نشان می دهد . اختلاف دو روش در تعداد ۴۰ وظیفه ۱ میلی ثانیه و در ۵۰ وظیفه به ۱۶ میلی ثانیه رسیده است.

۵- نتیجه‌گیری و پژوهش‌های آتی

با توجه به توسعه فناوری اطلاعات و ارتباطات بحث‌هایی مانند کاهش هزینه، کارایی و افزایش سرعت باعث ایجاد تحولی روزافزون در دنیای اینترنت شده است. محیط‌های توزیع و انجام وظایف در

آن‌ها شیوه‌ای از محاسبات کامپیوتری است که قابلیت‌های مرتبط با فناوری اطلاعات را به عنوان سرویس، به کاربران عرضه می‌کند و به آن‌ها این امکان را می‌دهد تا در بسترهای مختلف، بدون داشتن اطلاعات تخصصی یا کنترل کردن زیرساخت، به این سرویس‌ها دسترسی داشته باشند. این فناوری مزایای زیادی را فراهم کرده و مورد استقبال کاربران زیادی قرار گرفته است؛ اما با معایب و چالش‌هایی نیز روبرو است. یکی از این چالش‌ها، بهره‌وری منابع و واگذاری کارهاست که همان مسئله زمان‌بندی است. در این مقاله روشی جهت بهینه‌سازی مصرف انرژی در سیستم‌های توزیع براساس زمان‌بندی وظیفه ارائه شد که مبتنی بر ترکیب یادگیری تقویتی و عملگر نقاد بوده است که می‌تواند گزینه‌ای قابل رقابت در برنامه‌ریزی در محیط‌های توزیع تلقی گردد. تحلیل و تعدادی از نتایج نشان می دهند که این روش برای برنامه‌ریزی وظایف نه تنها شرط کیفیت لازم برای انجام وظیفه درخواستی را تامین می‌نماید بلکه تضمین کننده بالاترین سود برای ارائه کنندگان خدمات است. همچنین به زمان اجرای واقعی وظیفه در سیستمی متفاوت و نیز مهلت انجام و الویت و وابستگی کارها در الگوریتم توجه می‌نماید. در مقایسه روش پیشنهادی با روش مشابه که مبتنی بر الگوریتم زمان‌بندی کنترل سرعت بوده است نتایج بهبود اتمام زمان کار را از خود نشان داده است به نحوی که در تعداد منبع ۵ عدد با ۵۰ وظیفه اختلاف دو روش به ۱۴ میلی ثانیه، با تعداد ۱۰ منبع و ۵۰ وظیفه به ۱۱ میلی ثانیه و در ۲۰ عدد منبع و همان تعداد وظیفه به ۱۶ میلی ثانیه رسیده است. البته روش پیشنهادی در مقایسه با روش مبتنی بر الگوریتم DCMT دارای مقدار مناسب میزان مصرف انرژی و بهبود مناسبی در این زمینه بوده که به دلیل استفاده همزمان از یادگیری تقویتی و عملگر نقاد است. این بهینه‌سازی در تعداد منابع و درخواست‌های بالاتر مناسب بوده که می‌تواند در اجراهای دنیای واقعی مناسب باشد.

در مورد کارهای آتی می‌توان گفت که تحقیقات نشان می‌دهد استفاده از الگوریتم‌های اکتشافی عملکرد بهتری نسبت به سایر الگوریتم‌ها دارد در نتیجه از کارهای آینده قابل اشاره می‌توان اجرا و پیاده‌سازی روش ارایه شده و دستیابی به مزایا و معایب آن در دنیای واقعی، همچنین مطالعه و ارایه روش‌های ترکیبی دیگر با الگوریتم‌های بهینه‌سازی مختلف مانند PSO، ژنتیک و قورباغه جهنده را نام برد.

۶ – مراجع

[1] Jiwei. H, Chuang. L, Bo. Ch, (2015), "Energy Efficient Speed Scaling and Task Scheduling for Distributed Computing Systems ", IEEE, Chinese Journal of Electronics Vol.24, No.3, PP. 468-473.

[2] Chen. Y, Lin. Ch, Jiwei. H, Xiang. X, Shen. X, (2015), "Energy Efficient Scheduling and Management for Large-Scale Services Computing Systems ", IEEE, TRANSACTIONS ON SERVICES COMPUTING, MANUSCRIPT ID, PP. 1-13.

[3] Hosseinioun, Pejman, Maryam Kheirabadi, Seyed Reza Kamel Tabbakh, and Reza Ghaemi. "A new energy-aware tasks scheduling approach in fog computing using hybrid meta-heuristic algorithm." Journal of Parallel and Distributed Computing 143 (2020): 88-96.

[4] Peng, Zhihao, Behnam Barzegar, Maryam Yarahmadi, Homayun Motameni, and Poria Pirouzmand. "Energy-Aware Scheduling of Workflow Using a Heuristic Method on Green Cloud." Scientific Programming 2020 (2020).

[5] Alghamdi.A, Jiang. X, Zhang. J, Zhang. J, Jiang. M, Qin. X, (2017), "Towards Two-Phase Scheduling of Real-time Applications in Distributed Systems ", IEEE, Journal of Network and Computer Applications, PP. 1-29.

[6] Tian. Y, Xu. Q, Xue. J, (2017), "On Efficient Message Passing in Energy Harvesting Based Distributed System ", IEEE, Department of Computer Science & Engineering, PP. 139-144.

[7] Xiao. X, Xie. G, Li. R, Li. K, (2016), "Minimizing Schedule Length of Energy Consumption Constrained Parallel Applications on Heterogeneous Distributed Systems ", IEEE, TrustCom-BigDataSE-ISPA, PP.1471-1476.

[8] Thanavanich. Th, Siri. A, Uthayopas. P, (2016), "Energy-Aware Scheduling of Multiple Workflows Application on Distributed Systems ", IEEE, 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), PP. 1-6.

[9] Moghadam, Mahshid Helali, and Seyed Morteza Babamir. "Makespan reduction for dynamic workloads in cluster-based data grids using reinforcement-learning based scheduling." Journal of computational science 24 (2018): 402-412.

[10] Huang, Jiwei, Chuang Lin, and Bo Cheng. "Energy efficient speed scaling and task scheduling for distributed computing

systems." Chinese Journal of Electronics 24, no. 3 (2015): 468-473.

[11] Momeni, Hossein, and Nahid Mabhoot. "An Energy-aware Real-time Task Scheduling Approach in a Cloud Computing Environment." Journal of AI and Data Mining 9, no. 2 (2021): 213-226.

[12] Mei, Xinxin, Qiang Wang, Xiaowen Chu, Hai Liu, Yiu-Wing Leung, and Zongpeng Li. "Energy-aware Task Scheduling with Deadline Constraint in DVFS-enabled Heterogeneous Clusters." arXiv preprint arXiv:2104.00486 (2021).

آرش علی بخشی فرد مدرک کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۹۶ از دانشگاه آزاد هیدج، کارشناسی ارشد خود را در رشته کامپیوتر گرایش نرم‌افزار در سال ۱۳۹۸ از دانشگاه آزاد زنجان اخذ کرده است. نشانه رایا نامه ایشان عبارتند از :

arashalibakhshifard@gmail.com

لیدا ندرلو مدرک کارشناسی خود را در رشته مهندسی فناوری اطلاعات گرایش خدمات رایانه در شهرداری در سال ۱۳۹۴ از دانشگاه علمی کاربردی زنجان اخذ نموده و اکنون (سال۱۴۰۰) دانشجوی کارشناسی ارشد رشته مهندسی کامپیوتر گرایش هوش مصنوعی در موسسه آموزش عالی روزبه زنجان می‌باشد. زمینه‌های پژوهشی مورد علاقه ایشان عبارتند از: داده‌کاوی، سیستم پزشک‌یار، الگوریتم‌های بهینه‌سازی و احراز هویت. نشانه رایانامه ایشان عبارتند از :

lidanaderlou@gmail.com

زهرا طیبی قصبه مدرک کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال۱۳۹۰ از موسسه آموزش عالی احرار رشت و کارشناسی ارشد خود را در رشته کامپیوتر گرایش نرم‌افزار در سال ۱۳۹۲ از دانشگاه پیام‌نور تهران واحد شمیرانات اخذ کرده است. ایشان در حال حاضر به عنوان مدرس دانشگاه پیام نور گیلان، رشت در رشته های تخصصی کامپیوتر، ریاضی و آمار مشغول به کار است. زمینه‌های پژوهشی مورد علاقه ایشان: رایانش‌ابری، شبکه‌های حسگر، داده‌کاوی و امنیت داده‌های همگن، امنیت و مجازی‌سازی، الگوریتم‌های بهینه-سازی، امنیت در بانکداری، هوش مصنوعی در تشخیص نفوذ شبکه. نشانه رایانامه ایشان عبارتند از :

Tayyebi.shiva@gmail.com

روش ارجاع به مقاله : آ. بخشی فرد، ل. ندرلو، ز. طیبی قصبه ، بهینه‌سازی مصرف انرژی در سیستم‌های توزیع شده براساس زمان‌بندی وظیفه با استفاده از یادگیری تقویتی و روش عملگر-نقاد، دوفصلنامه محاسبات و سامانه های توزیع شده، سال چهارم، شماره اول، شماره پیاپی ۷، صفحه ۳۹ تا ۵۰، سال ۱۴۰۰.

How to cite: Arash Ali Bakhshifard, Lida Naderloo and Zahra Tayyebi Qasabeh. Energy Efficiency in Distribution Systems Based on Task Scheduling using Reinforcement Learning and Actor-Critic Method, Journal of Distributed Computing and Systems(JDCS), Vol 4, Issue 1, Page 39-50, 2021.