

الگوریتمی مبتنی بر تلفیق اطلاعات چند بانده با استفاده از شبکه GMDH برای پیشبینی بار ترافیک NFV در زیرساخت ابری

سیما جدی خواجه^۱ سعید شریفیان^۲

۱- دانش آموخته کارشناسی ارشد- دانشکده مهندسی برق- دانشگاه صنعتی امیرکبیر- تهران- ایران

simajeddi@aut.ac.ir

۲- استادیار- دانشکده مهندسی برق- دانشگاه صنعتی امیرکبیر- تهران- ایران

Sharifian_s@aut.ac.ir

چکیده: با فراگیر شدن استفاده از عملگرهای مجازی شبکه (NFV) که سرویس‌های شبکه نظیر مسیریابی، دیواره آتش و... را به صورت نرم‌افزاری در قالب ماشین‌های مجازی روی ساختار ابری توزیع شده ارائه می‌دهند، نیاز به تضمین کیفیت سرویس و عدم تخطی از قرارداد SLA کاربر در کنار مصرف انرژی بهینه مرکز داده ابری و کاهش تلفات انرژی از جمله مشکلات مطرح در این حوزه می‌باشند. برای حل این مشکلات نیاز به الگوریتمی برای مقیاس‌دهی و تخصیص پویای منابع ابری با توجه به پیشبینی نرخ بارکاری ابر می‌باشد. از آنجا که نرخ بارکاری ورودی به ابر دارای تغییرات زیادی در بازه‌های کوتاه مدت است، الگوریتم‌های فعلی برای پیشبینی نرخ بارکاری دقت لازم را ندارند. در این مقاله روش Wavelet-GMDH با دقت بالاتر در پیشبینی حجم بارکاری در مقایسه با کارهای قبلی ارائه شده است. ایده اصلی روش ارایه شده مدل‌سازی زیرباندهای مستقل زمان-فرکانس بارکاری با نگاشت تبدیل موجک و استفاده از شبکه GMDH با میزان غیرخطی‌نگی قابل کنترل در هر یک از زیر باندها برای پیشبینی بارکاری و در نهایت تلفیق خروجی‌های زیرباندها برای بدست آوردن مقدار نهایی پیشبینی بارکاری می‌باشد. ارزیابی عملکرد الگوریتم پیشنهادی با دو بارکاری استاندارد ابری Intel و TSDL، به سبب مدل‌سازی بهتر تغییرات شدید بارکاری در هر یک از زیرباندهای زمان-فرکانس نشان داد که درصد میانگین خطای مطلق پیشبینی ۷٫۲ درصد نسبت به بهترین روش موجود در مقالات مرتبط که مبتنی بر شبکه عصبی می‌باشد کاهش داشته است.

واژه‌های کلیدی: پیشبینی بارکاری ابر، مقیاس‌پذیری ابر، عملگرهای مجازی شبکه، تجزیه موجک، شبکه GMDH

نوع مقاله: پژوهشی

DOI: 10.29252/jiaeee.18.4.191

تاریخ ارسال مقاله: ۱۳۹۷/۶/۱۵

تاریخ پذیرش مشروط مقاله: ۱۳۹۸/۰۳/۰۱

تاریخ پذیرش مقاله: ۱۳۹۸/۷/۱

نام نویسنده‌ی مسئول: دکتر سعید شریفیان

نشانی نویسنده‌ی مسئول: ایران - تهران - خیابان حافظ - پلاک ۴۲۴ - دانشگاه صنعتی امیر کبیر - دانشکده برق

۱- مقدمه

مجازی‌سازی عملگرهای شبکه (NFV^1) یکی از تکنولوژی‌های در حال توسعه می‌باشد که بر روی زیرساخت ابری سرویس‌دهی می‌شود. در این تکنولوژی از مجازی‌سازی برای ارائه سرویس‌های مورد استفاده در شبکه به صورت نرم‌افزاری به صورت جایگزین بلوک‌های سخت‌افزاری استفاده می‌شود. این سرویس‌های نرم‌افزاری می‌توانند به صورت زنجیره‌ای از عملگرهای مختلف نظیر، مسیریاب، دیواره آتش، IDS/IPS و غیره به دنبال هم قرار گیرند [۲، ۱].

با استفاده از این تکنولوژی عملگرهای مختلف شبکه می‌توانند متناسب با نیاز و حجم بارکاری در هر کجای شبکه در ابرهای توزیع شده ایجاد شده و منابع مصرفی آنها که در ماشین مجازی اجرا می‌شوند به صورت پویا توسط زیرساخت ابری مدیریت شوند. اینکار نیاز به سخت‌افزار شبکه خاصی نداشته و بر روی پردازنده‌های همه منظوره ابری ارائه می‌شود. لذا هزینه تمام شده ($CAPEX^2$) کاهش یافته و سرعت ارائه خدمات نیز افزایش می‌یابد. همچنین هزینه نگهداری ($OPEX^3$) و انرژی الکتریکی مصرفی ابرها نیز کاهش می‌یابد [۳].

استفاده از ساختار ابری برای ارائه سرویس NFV به سبب خاصیت کشسانی ابر و مقیاس پذیری خودکار، اجازه می‌دهد که مدیریت منابع و افزایش و کاهش منابع شبکه به راحتی و متناسب با حجم ترافیک ورودی در هر بخش به صورت پویا انجام گیرد [۵، ۴]. اینکار سبب می‌شود که کیفیت سرویس تضمین شده در قرار داد سرویس فراهم شده و در عین حال مصرف منابع و انرژی ابرها کمینه باشد [۶].

برای اینکه مدیریت مقیاس‌پذیری خودکار ابر به درستی عمل نماید باید بتواند نرخ ترافیک و بارکاری شبکه را برای بازه زمانی بعدی (آینده) به خوبی پیش‌بینی نموده و با پیش‌دستی منابع مجازی مورد نیاز آنرا در ابر تخصیص دهد. پیش‌بینی دقیق بارکاری یعنی تعیین تعداد درخواست‌های ورودی به ابر برای استفاده از ماشین مجازی سرویس NFV که تضمین کننده تخصیص درست منابع و در نتیجه ارائه کیفیت سرویس توافق شده با کاربر است [۲]. بارکاری در شبکه ماهیت شدیداً غیرخطی با تغییرات ناگهانی زیاد (واریانس بسیار بالا) در بازه های کوتاه است که اصطلاحاً به آن ($highly\ volatile$) می‌گویند [۲]. بسیاری از روش‌های فعلی پیش‌بینی فرض می‌کنند که ترافیک ورودی دارای تغییرات اندکی می‌باشند [۸، ۷] و یا دارای مولفه‌های فصلی^۴ و تکرارشونده هستند که این فرض باعث به وجود آمدن خطای زیادی در پیش‌بینی بارکاری بخصوص در کاربردهای NFV می‌شود [۹].

در این مقاله یک روش جدید برای پیش‌بینی بارکاری یا ترافیک شبکه در کاربردهای NFV ارائه شده است که با تجزیه سری زمانی بارکاری شبکه توسط تبدیل موجک گسسته به چهار زیر فضای زمان-فرکانس مجزا سبب می‌شود که داده های سری زمانی هر زیر فضا شباهت بیشتری به هم داشته باشند. در هر زیرفضا از یک شبکه $GMDH$

[۱۰-۱۲] مجزا استفاده شده است تا با استفاده از ویژگی‌های مستخرج از هر زیرفضای زمان-فرکانس پیش‌بینی دقیق‌تری برای سری زمانی آن زیرفضا ارائه نماید. لذا پیش‌بینی در هر زیر فضا با دقت بالاتر و شبکه $GMDH$ با پیچیدگی کمتر انجام می‌پذیرد. برای پیش‌بینی بارکاری نهایی در بازه زمانی اتی اطلاعات خروجی شبکه‌های $GMDH$ با هم تلفیق شده و پیش‌بینی نهایی را می‌سازند. روش ارائه شده با استفاده از نگاشت اطلاعات به فضاهای مختلف و استفاده از یک شبکه $GMDH$ بسیار غیرخطی توانسته است درصد میانگین خطای مطلق پیش‌بینی را ۷،۲ درصد برای بارهای کاری استاندارد [۱۳] Intel و [14] TSDL نسبت به بهترین روش موجود در مقالات مرتبط که مبتنی بر شبکه عصبی می‌باشد [۱۵] کاهش دهد.

در ادامه در بخش ۲ کارهای مشابه مرور شده‌اند. در بخش ۳ مساله پیش‌بینی بارکاری برای تخصیص منابع در NFV مطرح شده و روش پیشنهادی برای حل آن مطرح شده است. در بخش ۴ نتایج ارزیابی الگوریتم با دو بار کاری استاندارد مطرح شده و در نهایت با روش‌های مشابه در مقالات مقایسه گردیده. در بخش ۵ نیز نتیجه‌گیری و جمع‌بندی ارائه شده است.

۲- بررسی کارهای مشابه

در مرجع [۱۶] که یک مقاله تحقیقی است روش‌های مختلف پیش بینی سری‌های زمانی در سال‌های جدید به صورت دسته‌بندی شده و مقایسه‌ای ارائه شده است که می‌تواند برای آشنایی خواننده با این موضوع بسیار مفید باشد. در ادامه به بررسی تعدادی از روش‌های ارائه شده برای پیش‌بینی در سال‌های اخیر که به کار ارائه شده در این مقاله نزدیک است می‌پردازیم.

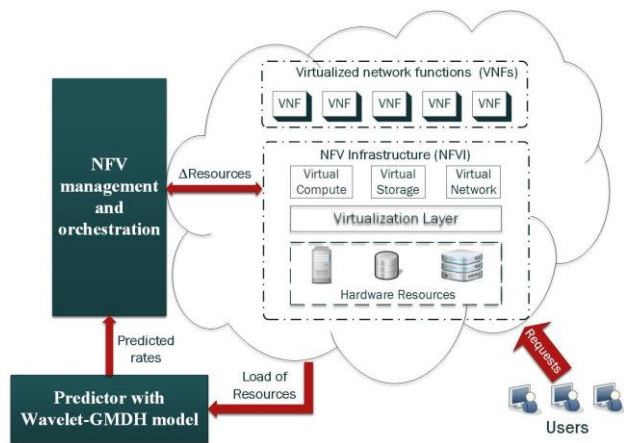
در مرجع [۱۷] مسئله پیش‌بینی بارکاری مرکز داده ابری را به یک مسئله کلاس‌بندی سنتی با استفاده از الگوی تقسیم‌بندی نمایی تبدیل کرده است. در این الگو بازه پیش‌بینی به قطعاتی که طولشان به صورت نمایی افزایش می‌یابند، تقسیم می‌شود و سپس با پیش‌بینی میانگین بار برای هر قطعه مقدار بار در کل بازه پیش‌بینی به دست می‌آید سپس با استفاده از یک تبدیل، به یک مسئله کلاس‌بندی تبدیل می‌شود. نتایجی که با استفاده از این روش به دست می‌آیند در بازه زمانی طولانی صادق هستند و بنابراین در دسته روش‌های پیش‌بینی بلندمدت قرار می‌گیرد. این روش برای دقت خوب نیاز به تعداد زیادی از نمونه‌ها دارد و همچنین حجم محاسبات زیادی دارد. اگرچه به طور کلی پیش‌بینی میانگین بار به این دلیل که میانگین بار نمی‌تواند نوسانات بار را در بازه‌های زمانی بلندمدت به خوبی نشان دهد، یک بحث چالش برانگیز است [۲].

در مرجع [۱۸] روش‌های آماری سنتی میانگین متحرک و فیلتر نمایی و همچنین روش جدید تفکیک مدل برای پیش‌بینی توان تابشی خورشید استفاده شده است. این روش‌ها قادر به پیش‌بینی در بازه کوتاه مدت یک روزه هستند. دو روش میانگین متحرک و فیلتر نمایی

بارکاری و تعداد کاربران در یک سرویس اینترنت اشیا استفاده نمودند که نتایج نشان دهنده دقت بیشتر شبکه عصبی در پیش‌بینی بارکاری بود. در مرجع [۲۵] نویسندگان با استفاده از یک مدل SVR که با روش‌های فراابتکاری پارامترهای آن تنظیم شده است بارکاری را برای محیط ابری پیش‌بینی نمودند که به خطای پیش‌بینی زیر ده درصد در بهترین حالت دست یافتند. در مرجع [۹] روش ترکیبی EMD-SVR برای پیش‌بینی بارهای الکتریکی در شبکه‌های توزیع برق ارائه شده که در آن از متد Empirical mode decomposition برای حذف تغییرات زیاد روی داده آموزشی استفاده می‌کند. EMD یک روش تحلیل داده ایستان کارآمد است [۹]. در این مقاله خروجی فیلتر شده EMD توسط روش SVR برای پیش‌بینی استفاده شده است.

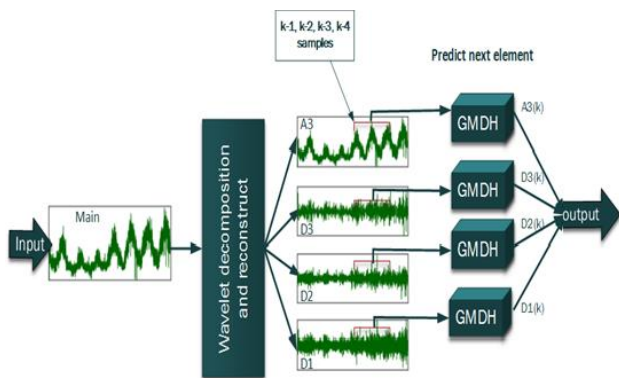
۳- مساله پیش‌بینی بارکاری در سرویس NFV و الگوریتم پیشنهادی برای حل آن

در شکل (۱) زیرساخت مدیریت پویای منابع در سرویس NFV بر روی زیرساخت ابری نشان داده شده است [۱]. کاربران شبکه درخواستهای خود را به مرکز داده ابری می‌فرستند این درخواستها در یک صف درخواست ذخیره شده و تعداد آنها در هر واحد زمانی به واحد پیش‌بینی کننده فرستاده می‌شود تا نرخ درخواستهای هر ماژول مجازی عملکرد شبکه در آینده را پیش‌بینی نماید. این واحد خود درون واحد هماهنگی و مدیریت NFV قرار دارد. این واحد مدیریت براساس بارکاری پیش‌بینی شده و بازخوردی که از میزان بار فعلی منابع ابری از ساختار NFV می‌گیرد تغییر مورد نیاز (کاهش یا افزایش) را در منابع مجازی هر VNF در ابر مشخص می‌نماید. از شکل (۱) مشاهده می‌شود که ساختار NFV در داخل زیرساخت ابری پیاده‌سازی شده است. منابع سخت‌افزاری موردنیاز از جمله پردازنده، حافظه و شبکه در ابر فراهم می‌شوند. سپس این منابع توسط لایه ی مجازی‌سازی به منابع مجازی تبدیل شده و برای پیاده‌سازی توابع مجازی شبکه (VNF) مورد استفاده قرار می‌گیرند [۲].



شکل (۱): زیرساخت مدیریت پویای منابع در NFV [1]

پیشنهادی این مقاله برای بارکاری دارای تغییرات کم نتایج خوبی ارائه می‌کنند در حالی که همانطور در مقدمه اشاره شد یکی از مولفه‌های بارکاری ابری تغییرات زیاد آنهاست. روش تفکیکی ارائه شده نیز برای بارهای دارای تغییرات کم و فصلی پیش‌بینی‌های مناسبی ارائه می‌کند در حالی که بارهای کاری ابری دارای تغییرات زیادی بوده و تغییرات فصلی محسوسی ندارند [۲]. در مرجع [۱۹] برای پیش‌بینی همین مساله از شبکه عصبی استفاده نموده است که به پیش‌بینی با دقت بالاتری منجر شده است. در روشی دیگر که در مرجع [۲۰] ارائه شده ابتدا با استفاده از تبدیل موجک آشوبناک مولفه‌های تقریب و جزئیات سری زمانی به دست آمده، سپس با استفاده از مدل‌های SVR و ARIMA خروجی نهایی را پیش‌بینی می‌نمایند که این روش باعث کاهش خطای پیش‌بینی شده است. به تازگی در مرجع [۲۱] نیز یک مدل بهبود داده شده SVR با الگوریتم بهینه‌سازی زنبور عسل مصنوعی برای پیش‌بینی مقدار انرژی باقیمانده در باتری‌های لیتیوم یون ارائه شده است که بر اساس نتایج بدست آمده در این مقاله پیش‌بینی با دقت بالا نسبت به روش‌های مشابه بدست آمده است. در مرجع [۲۲] نویسندگان نرخ بارکاری ابر در بازه زمانی طولانی مدت را با یک مدل بر پایه تئوری بیز پیش‌بینی می‌کنند. در این مرجع پیش‌بینی میانگین بارکاری در بازه‌های زمانی با طول متفاوت با استفاده از ۱۰ ویژگی مهم که نشان دهنده روند، الگو و پایداری بارکاری هستند به کمک تئوری بیز انجام می‌شود. پیش‌بینی میانگین بارکاری ابری آن هم در بازه‌های زمانی طولانی مدت به سبب تغییرات زیاد بارکاری ابری دارای خطای پیش‌بینی کوتاه مدت زیادی می‌باشد [۲۲]. در مرجع [۲۳] نیز یک مدل بر پایه تئوری بیز برای پیش‌بینی منابع مجازی پردازنده و حافظه مورد نیاز براساس الگوی بارکاری در مرکز داده ابری ارائه شده است. تئوری بیز ساده بوده و اطلاعات اولیه مدیریت منابع را به صورت احتمال اولیه در نظر می‌گیرد. ویژگی‌های در نظر گرفته شده برای توصیف بارکاری باید مستقل از هم باشند و برای انطباق با تغییرات بارکاری به صورت منظم بروزرسانی شوند که این باعث افزایش بار محاسباتی می‌شود [۲۳]. در مرجع [۷] در ابتدا بارکاری مرکز داده ابری براساس ویژگی‌هایشان کلاس‌بندی شده و سپس دو مدل ARMA و شبکه عصبی برای پیش‌بینی بارکاری برای هر کلاس استفاده می‌شوند. مشکلی که در این روش وجود دارد این است که برای مشخص کردن مدل مورد استفاده برای پیش‌بینی ابتدا باید کلاس آن تشخیص داده شود و برای اینکار نیاز به داده‌های بلندمدت می‌باشد و لذا به کمک این روش پیش‌بینی بارکاری با تغییرات زیاد در دوره زمانی کوتاه مدت دارای خطای زیادی می‌باشد. یک شبکه عصبی بازگشتی که وزن‌های آن با استفاده از الگوریتم ایمنی مصنوعی انتخاب شده است در مرجع [۱۵] برای پیش‌بینی سری‌های زمانی مالی ارائه شده است. شبکه‌های عصبی به خوبی قادر به مدل‌سازی رفتار غیرخطی هستند. به تازگی در مرجع [۲۴] نویسندگان از مدل ARIMA و SVR و شبکه عصبی برای تخمین



شکل (۲): مدل ارائه شده Wavelet-GMDH

در الگوریتم Wavelet-GMDH برای آموزش شبکه‌های GMDH از ۷۵٪ داده‌های بارکاری استفاده می‌شود و برای پیش‌بینی از ۲۵٪ باقی‌مانده داده‌ها استفاده می‌شود. در ادامه هر یک از بلوک‌های الگوریتم Wavelet-GMDH با جزئیات بیشتر ارائه شده‌اند.

۳-۲- تجزیه ورودی به چهار زیر باند زمان-فرکانسی با استفاده از تبدیل موجک گسسته

تبدیل موجک به عنوان یک روش قدرتمند در آنالیز و رفع نوفه سریهای زمانی دارای تغییرات زیاد و ناپیوسته شناخته شده است [۲۵، ۲۴]. تبدیل موجک، محتوای فرکانسی یک سیگنال و تغییرات موقتی زمانی آن را توانمند آنالیز می‌کند [۲۷]. از تبدیل موجک در این مقاله برای تجزیه و نگاشت سری زمانی بارکاری به زیرباند های زمان-فرکانسی مجزا که هر یک از این زیرباندها مشخصات و ویژگیهای مشابهی دارند استفاده شده است. استفاده از این زیرباندهای مجزا کمک می‌نماید که مدل GMDH مربوط به هر یک از این زیرباندها پیچیدگی کمتری داشته و منجر به دقت بالاتری در پیش‌بینی گردد. در تبدیل موجک گسسته (DWT) نگاشت زمان-فرکانس از سیگنال گسسته با استفاده از فیلترهای دیجیتال انجام می‌گردد. در واقع با عبور سیگنال از فیلترهای بالاگذر و پایین‌گذر با فرکانس قطعهای مختلف یا به بیان دیگر کانولوشن ورودی $x[n]$ و پاسخ ضربه فیلتر $g[n]$ مطابق رابطه (۱) نگاشت می‌شود.

$$x[n] * g[n] = \sum_{k=-\infty}^{\infty} x[k].g[n-k] \quad (1)$$

در نتیجه این عمل فیلترینگ، تمام مولفه‌های فرکانسی که بیشتر از نصف بزرگترین فرکانس موجود در سیگنال باشند حذف می‌شوند. از آنجا که بیشترین فرکانس موجود در سیگنال خروجی فیلتر نصف بزرگترین فرکانس موجود در سیگنال اصلی است و از طرفی دیگر بر اساس اصل نایکوئیست نرخ نمونه‌برداری باید دو برابر بزرگترین فرکانس موجود در سیگنال باشد، نصف نمونه‌های سیگنال خروجی قابل حذف هستند [۲۷]. لذا با حذف یکی در میان نمونه‌ها

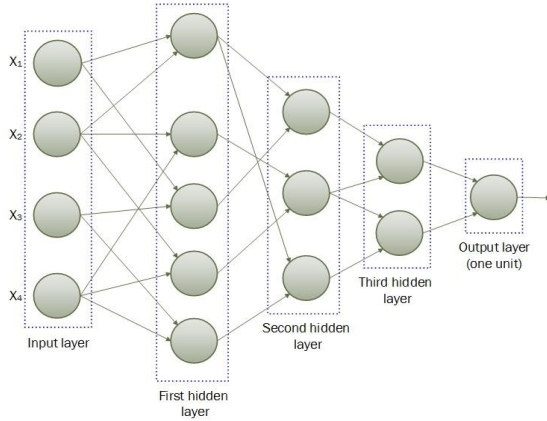
چارچوب سرویس‌دهی NFV استاندارد [۱] شامل اجزای زیر است:

- زیرساخت (NFVI): مجموعه‌ای از تمام اجزای سخت‌افزاری و نرم‌افزاری است که محیطی را که در آن VNF ها مستقر هستند ایجاد می‌کند. NFVI می‌تواند در چندین مرکز داده ابری به صورت توزیع شده پخش شود. شبکه ارتباطی بین این مراکز بخشی از NFVI است.
- مدیریت مجازی‌سازی توابع شبکه و چارچوب معماری هماهنگ‌سازی (NFV-MANO): مجموعه‌ای از تمام بلوک‌های عملکردی، مخازن داده استفاده شده توسط این بلوک‌های کاربردی و نقاط مرجع و رابطها است که از طریق آن این بلوک‌های کاربردی مبادله اطلاعات را با هدف مدیریت و هماهنگ کردن NFVI و VNF ها انجام می‌دهند. در این مقاله بر روی واحد پیش‌بینی بارکاری تمرکز شده و روش Wavelet-GMDH برای پیش‌بینی نرخ درخواستهای ورودی به زیرساخت ابری ارائه شده است که در ادامه با جزئیات بیشتر تشریح می‌گردد.

۳-۱- الگوریتم پیشنهادی Wavelet-GMDH

همانطور که در مقدمه اشاره شد، ویژگی مهم بارکاری ابری فرآوردن یا تغییرات شدید دامنه در نمونه‌های متوالی می‌باشد [۲]. این ویژگی سبب می‌شود که تبدیل لازم برای مدل‌سازی آن بسیار غیرخطی و پیچیده باشد. ایده‌ای که در این مقاله ارائه شده است بر این اساس می‌باشد که در ابتدا به کمک تبدیل موجک سری زمانی مربوط به بارکاری ابری را مطابق شکل ۲ به چهار زیرفضای زمان-فرکانسی مستقل شامل یک زیر فضای تقریب A3 و سه زیرفضای جزئیات D1، D2 و D3 نگاشت می‌کنیم. در هر یک از این زیرفضاها سری زمانی بدست آمده به سبب محدودیت باند فرکانسی اعمالی به نوعی فیلتر شده و مشخصات فرکانسی یکدست‌تری دارد. این ویژگی کمک می‌کند که برای مدل‌سازی آن بتوان از مدلی با پیچیدگی کمتر استفاده نمود. ایده بعدی این مقاله مدل‌سازی هر یک از زیرباندها با یک شبکه GMDH مستقل است. علت انتخاب شبکه GMDH این است که میزان غیرخطی بودن و به تبع آن پیچیدگی شبکه توسط کاربر به سادگی قابل تنظیم می‌باشد [۲۶]. همانطور که در شکل ۲ نشان داده شده است به ورودی هر یک از ۴ شبکه GMDH (که هر کدام شبیه شکل ۴ هستند) چهار نمونه قبلی از هر زیر باند اعمال می‌شود تا نتیجه پیش‌بینی مقدار آن زیر باند برای آینده بدست آید. در انتها مقادیر پیش‌بینی شده هر زیر باند با هم جمع می‌شوند تا مقدار نهایی پیش‌بینی بارکاری برای لحظه‌ای را بدست دهند.

GMDH با ایجاد یک تابع تحلیلی در یک شبکه جلوسو براساس یک تابع انتقال درجه دوم در هر گره خود که در آن ثوابت چند جمله‌ای توسط روش کمترین مربعات به دست می‌آیند عمل می‌کند [۱۲]. در این شبکه با ترکیب چندجمله‌ای‌های خیلی ساده به دلیل پیچیدگی در ترکیب و کثرت رفتارهای پیچیده را می‌توان به سادگی مدل کرد. در شکل (۴) ساختار یک شبکه GMDH با چهار ورودی آمده است.



شکل (۴): ساختار GMDH نمونه [۲۶]

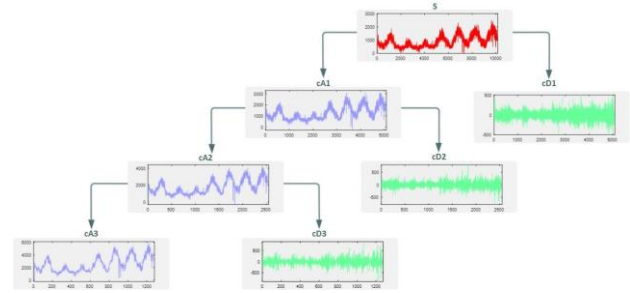
در GMDH هدف کلی مدل کردن سیستمی با ورودی بردار x و خروجی y به صورت زیر است.

$$y = a_0 + \sum_{i=1}^m a_i f_i \quad (2)$$

در رابطه (۲) a_0 ضریب ثابت به عنوان بایاس‌ها، a_i ضرایب وزنی و f_i توابع پایه را تشکیل می‌دهند. توابع f_i تابعی از ورودیهای بردار x هستند. این مدل شباهت بسیاری به مدل‌های استفاده شده در شبکه عصبی، سیستمهای فازی و... دارد [۲۶]. تفاوتی که GMDH را از این روشها مجزا می‌کند نحوه به دست آوردن این مدل و توصیف آن است. در GMDH عملکرد کلی به این شکل است یک سری مدل‌های جزئی را از طریق الگوریتم حداقل مربعات ایجاد می‌کنیم در هر مرحله از این مدل‌های جزئی، تعدادی را که کارآمد هستند انتخاب کرده و با ترکیب این مدل‌ها در نهایت به مدل کلی سیستم در چند لایه می‌رسیم [۱۲]. ایده‌ای که در GMDH وجود دارد این است که مانند الگوریتم‌های تکاملی اجازه می‌دهیم که تعداد زیادی از مدل‌های جزئی و پایه‌ای از چند جمله‌ای‌ها ایجاد شوند. سپس همانند روش انتخاب طبیعی تعدادی مدل جزئی با تابع هزینه (حداقل مربعات) کمتر از بین آنها انتخاب می‌کنیم و دوباره اجازه بهبود و ترکیب به این مدل‌ها را می‌دهیم تا در ساخت چند جمله‌ای لایه بعدی مورد استفاده قرار گیرند. در واقع یک فرآیند بازگشتی وجود دارد، فرآیند انتخاب و ساخت مدل‌های جزئی چندبار تکرار می‌شود تا نهایتاً به مدل نهایی می‌رسیم. مدل پایه‌ای که در رابطه (۲) آمده را می‌توان به صورت زیر نوشت [۸-۱۲]:

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} x_i x_j x_k + \dots \quad (3)$$

طول سیگنال نصف خواهد شد بدون اینکه اطلاعاتی را از دست داده باشیم.



شکل (۳): تجزیه بانک فیلتر سه مرحله‌ای

روند مشابهی نیز با استفاده از یک فیلتر دیجیتال بالاگذر نیم باند با پاسخ ضربه $h[n]$ انجام می‌پذیرد. در نتیجه در خروجی اولین مرحله از اعمال تبدیل موجک، دو نسخه، یکی بالاگذر و دیگری پایین‌گذر با طول کاهش یافته از سیگنال اولیه به دست می‌آید. این عمل تفکیک زمانی نصف شده و در مقابل تفکیک فرکانسی دو برابر می‌شود. این روند را می‌توان مجدداً بر روی نسخه پایین‌گذر بدست آمده اعمال نمود و در هر مرحله، با کاهش تفکیک زمانی به میزان نصف مرحله قبل، تفکیک فرکانسی را دو برابر نمود. این ایده برای محاسبه تبدیل موجک گسسته، به روش بانک فیلتری مشهور است [۲۷]. بلوک دیاگرام این روش در شکل (۳) آمده است.

ضرایب خروجی فیلتر پایین‌گذر شکل اولیه سیگنال را دنبال می‌کنند به همین دلیل به این ضرایب تقریباً گفته می‌شود. ضرایب خروجی فیلتر بالاگذر، جزئیات فرکانس بالای سیگنال را دربردارند به همین دلیل به این ضرایب، جزئیات گفته می‌شود.

۳-۳- پیشبینی بارکاری با استفاده از شبکه GMDH برای چهار زیرباند زمان-فرکانسی تبدیل موجک

شبکه عصبی GMDH نوعی شبکه عصبی است که با استفاده از یک تکنیک خود سازمان ده با عمیق شدن شبکه براساس تکرار یک سری مدل‌های چندجمله‌ای جزئی تر به تدریج پیچیده‌تر می‌شود [۱۱، ۱۰]. انواع مختلفی از شبکه‌های GMDH وجود دارد اما در اینجا مدل GMDH پایه‌ای که توسط Ivakhnenko [۲۶] ارائه شده و در پیاده سازی این مقاله نیز از این مدل استفاده شده است تشریح می‌گردد. GMDH، یک شبکه شناخته شده برای بهینه‌سازی و مدل‌سازی سیستمهای پیچیده و غیرخطی است و جهت تخمین و پیشبینی رفتار این سیستمها استفاده می‌شود [۸، ۱۲]. در روش ارائه شده پایه توسط Ivakhnenko مدل سازی بر اساس نهایتاً چندجمله‌ای‌های درجه ۲ انجام شده است که حجم محاسباتی کمتری نسبت به انواع پیچیده‌تر این شبکه دارد [۲۶].

$$ec = \alpha * L(1).RMSE2 + (1 - \alpha) * L(end).RMSE2 \quad (7)$$

در رابطه (7) مقدار $L(1)$ نشان دهنده نورون با کمترین خطا و $L(end)$ نشان دهنده نورون با بیشترین خطا در یک لایه است. نورونهایی که میزان خطای بیشتر از ec دارند طبق رابطه (8) حذف می شوند و نورونهای با خطای کمتر از ec در آن لایه باقی می ماندند و برای تشکیل ورودیهای لایه بعدی استفاده می شوند. از مقدار α به 0.6 برای α در پیاده سازی استفاده شده است [26]. هر چه مقدار α نزدیک به 1 باشد نورونهای با مقدار خطای کمتر انتخاب می شوند و در نتیجه تعداد نورونهای باقیمانده کمتر شده و به تبع آن تعداد نورونهای لایه بعدی کمتر می شود و زودتر به لایه انتهایی شبکه می رسیم. از طرفی هر چه α نزدیکتر به صفر شود میزان خطای نورونها زیاد بوده و تعداد بیشتری از نورونها باقی می ماندند که باعث می شود تعداد نورونهای لایه بعد زیاد شده و حجم محاسبات بالا رود. پس در انتخاب α یک هزینه-فایده برقرار است و باید به گونه ای انتخاب شود که از یک طرف میزان محاسبات بیش از حد نشود و از طرفی دیگر تعداد نورونهای هر لایه خیلی کم نشود.

$$L = L([L.RMSE2] \leq ec) \quad (8)$$

در رابطه (8) L مجموعه نورونهای موجود در یک لایه را نشان می دهد. در صورتی که یکی از شرایط زیر رخ دهد ساخت شبکه عصبی GMDH به اتمام می رسد: یکی اینکه تعداد لایه های شبکه به ماکزیمم تعداد لایه های تعیین شده برسد و دیگری اگر تنها یک نورون در لایه باقی مانده باشد. تفاوت میان شبکه های مختلف GMDH در نحوه انتخاب بهترین نورونها باشد.

4- نتایج ارزیابی الگوریتم Wavelet-GMDH

به منظور ارزیابی دقت پیش بینی روش پیشنهادی Wavelet-GMDH از رد دوبارکاری واقعی که از مراکز داده ابری بدست آمده اند و در مقالات مشابه نیز به کار رفته اند استفاده شده است [30, 29]. یکی از آنها بارکاری Intel Netbatch logs است [13] که به بارکاری Intel pool B نیز مشهور است و بارکاری دیگر داده های مربوط به ترافیک اینترنت می باشد که به Time Series Data Library (TSDL) در مقالات به آن اشاره می شود [14].

در مقاله های [30, 29] از بارکاری اینتل برای پیش بینی زمان اجرای درخواست ها، تخمین زمان انتظار درخواست ها در صف و بهبود عملکرد صف بهره برده اند. در [31] نیز از بارکاری اینتل به منظور کمینه کردن زمان مشغول بودن در صف درخواست های یک ابر استفاده شده است. در مراجع [24, 23, 22] از داده های TSDL برای پیاده سازی الگوریتم های مختلف و تست عملکرد آنها در پیش بینی درخواست های کاری مربوط به ابر استفاده شده است. داده های اینتل شامل یک ماه بررسی شبکه Netbatch اینتل است. این شبکه از چندین مرکز داده ابری، که هر کدام شامل ده ها هزار گره

در این رابطه مولفه های درجه 1 و 2 و 3 ... با هم جمع می شوند. به رابطه (3) اصطلاحاً چندجمله ای Kolmogrov-Gabor گفته می شود [28]. تعداد ضرایب یا جملات این چندجمله ای با فرض n ورودی 2^n است. به عنوان مثال در سیستمی با 30 عدد ورودی تقریباً $10^{10} \times (2^3)^{10} = 2^{30}$ ضریب خواهیم داشت که محاسبه آنها بسیار زمان بر است. برای اجتناب از این حجم محاسبات برای حل رابطه (3) از GMDH استفاده می شود. معمولاً در مسائلی از روش GMDH استفاده می شود که سیستم غیر خطی بوده و پیچیدگی بالایی داشته باشد. در این سیستمها به دلیل تعداد زیاد ورودیها، تعداد ضرایب نیز بسیار زیاد می شود. مزیت GMDH این است که مسئله پیدا کردن 2^n مجهول در مسئله بزرگ را به چند مسئله کوچک با تعداد مجهول کمتر تبدیل می کند. در لایه اول به تعداد همه ترکیبهای غیر تکراری ورودیها، نورون در نظر گرفته می شود که تعداد آنها از رابطه زیر به دست می آید.

$$N = \frac{n(n-1)}{2} \quad (4)$$

در رابطه (4) n تعداد ورودیها و N تعداد نورونهای لایه اول است. این فرمول را می توان گسترش داد و برای دیگر لایه ها نیز استفاده کرد:

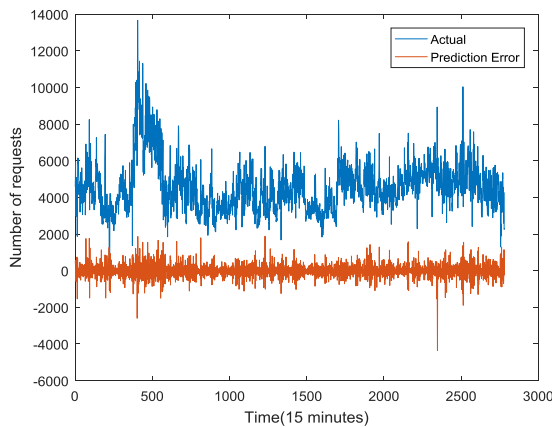
$$N_{k+1} = \frac{N_k(N_k - 1)}{2} \quad (5)$$

در رابطه (5) N_k تعداد نورونهای لایه k ام و N_{k+1} تعداد نورونهای لایه $k+1$ است. البته در پیاده سازی مقداری برای بیشینه تعداد نورونها در یک لایه در نظر گرفته می شود در صورتی که تعداد نورونهای محاسبه شده برای یک لایه بیشتر از تعداد بیشینه باشد، نورونهای اضافی را تا رسیدن به این حد بالا با الویت بندی بر اساس تابع هزینه حذف می کنیم. در شبکه عصبی چندجمله ای هر نورون براساس چندجمله ای های درجه 2 از ورودیهای مدل می شود. در صورتی که خروجی نورون با ورودیهای i و j را z_{ij} در نظر بگیریم داریم [12]:

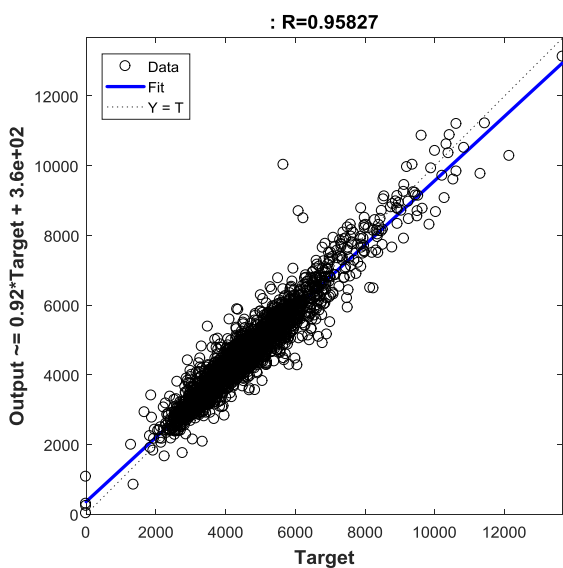
$$z_{ij} = c_1 + c_2 x_i + c_3 x_j + c_4 x_i^2 + c_5 x_j^2 + c_6 x_i x_j \quad (6)$$

در رابطه (6) x_i و x_j ورودیهای i و j ام و c ها نیز ضرایبی هستند که باید مشخص شوند. برای محاسبه مقادیر بهینه این ضرایب از روش کمترین مربعات خطا استفاده می شود. به این ترتیب مقدار 6 ضریب c_1 تا c_6 برای هر نورون در لایه اول به گونه ای محاسبه می شوند که خروجیهای z_{ij} کمترین اختلاف را با خروجی نهایی y داشته باشند.

برای انتخاب بهترین نورونها خطای حاصل از چندجمله ای هر نورون را روی دادگان اعتبار سنجی به دست می آوریم. خطای به دست آمده برای این داده ها را با $RMSE2$ نشان می دهیم. برای انتخاب نورونها از رتبه بندی استفاده می شود. نورونها بر اساس مقدار $RMSE2$ به ترتیب از کوچک به بزرگ مرتب می شوند. سپس پارامتر فشار انتخاب α که عددی بین صفر و یک است برای مشخص کردن مرز انتخاب نورونها با رابطه زیر استفاده می شود [8]:



شکل (۵): بارکاری Intel pool B و خطای پیشبینی بار Intel Wavelet-GMDH با الگوریتم



شکل (۶): پراکندگی خطای پیشبینی بار Intel pool B با الگوریتم Wavelet-GMDH

در شکل ۷ سری زمانی نرخ درخواست های بارکاری TSDL و نتایج خطای پیشبینی الگوریتم Wavelet-GMDH برای این بارکاری نشان داده شده است. قابل ذکر است که همه مقادیر ترافیک اینترنت که برحسب بیت ارائه شده بود، برای پیشبینی پس از تقسیم بر ۱۰۲۴ بر حسب کیلوبیت استفاده شده اند [۱۴]. با توجه به شکل ۷ که در آن دامنه تغییرات خطا مشخص است و همچنین مقدار درصد مطلق خطای ۱٫۳ درصد می‌رسیم.

در شکل ۸ پراکندگی خطای پیشبینی آورده شده است که نشان می‌دهد خطای پیشبینی تقریباً به صورت یکسان برای کلیه مقادیر دامنه ورودی بدست آمده است. این به دلیل این است که تغییرات لحظه این بارکاری کمتر بوده و تغییرات دامنه با شیب مناسبی انجام می‌گیرد و لذا مدل‌سازی غیرخطی آن با دقت مناسب و هزینه محاسباتی قابل قبولی توسط الگوریتم Wavelet-GMDH انجام می‌گیرد.

است، در مکانهای مختلف در سراسر جهان تشکیل شده است [۱۳]. در مرجع [۳۵] از داده های درخواست‌های Intel pool B برای تست عملکرد الگوریتم پیش‌بینی استفاده شده است. داده‌های ترافیک اینترنت مربوط به TSDL از یک ISP خصوصی با ۱۱ مرکز در شهرهای اروپایی در بازه ساعت ۶:۵۷ روز هفتم June تا ساعت ۱۱:۱۷ روز ۳۱ July سال ۲۰۰۵ جمع‌آوری شده است [۱۴]. در این مقاله به منظور مقایسه روش پیشنهادی Wavelet-GMDH با روشهای متداول پیش‌بینی در مقالات جدید نظیر روش‌های برپایه [7, 15, 19, 32] ANN، [7] ARMA و [21, 25, 33, 36] SVR این سه روش نیز پیاده‌سازی شده و نتایج آنها براساس درصد مطلق خطای پیشبینی (MAPE) مقایسه شده که در ادامه آمده است. تمام ارزیابی‌ها در نرم افزار MATLAB R2017 بر روی سخت‌افزار با پردازنده core i7 و حافظه ۸ گیگابایت و در سیستم عامل Windows 10 پیاده‌سازی شده‌اند.

۴-۱- عملکرد الگوریتم Wavelet-GMDH درپیش‌بینی دقیق بارکاری ابری

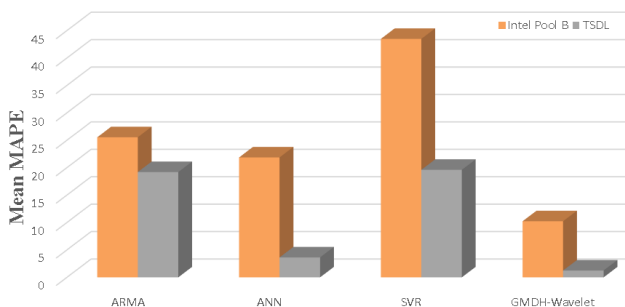
در این بخش با استفاده از الگوریتم wavelet-GMDH سعی داریم تا دقت پیش‌بینی را برای دو بار کاری ابری Intel pool B و TSDL که در بخش قبل به تفصیل به آن پرداختیم را بررسی نماییم. سری زمانی نرخ درخواست‌ها و همچنین نتایج خطای پیش‌بینی الگوریتم wavelet-GMDH بر روی بارکاری Intel pool B در شکل (۵) نشان داده شده است.

در شکل (۵) با مقایسه مقیاس خطا با مقیاس داده اصلی و همچنین محاسبه مقادیر خطای پیش‌بینی به این جمع‌بندی می‌رسیم که مدل ارائه شده Wavelet-GMDH با درصد خطای مطلق ۸٫۲ درصد توانسته بار کاری Intel pool B را پیش‌بینی کند. در شکل ۶ نمودار پراکندگی خطای مقادیر پیش‌بینی شده توسط الگوریتم-Wavelet GMDH نشان داده شده است. همانطور که در شکل ۶ نشان داده شده است، پیش‌بینی برای دامنه عددی کم و زیاد بارکاری بعضاً دارای خطاهای قابل توجهی می‌باشد. علت این پدیده تغییرات دامنه ناگهانی و شدید بارکاری (از دامنه زیاد به دامنه کم و برعکس) است که بسیار غیرخطی بوده و مدل نمودن آنها دشوار است و همانطور که در بخش دو اشاره شد از ویژگیهای بارکاری ابری می‌باشد که پیش‌بینی دقیق آنها دشوار است. در روش Wavelet-GMDH با وجود تمرکز بر روی مدل کردن این مولفه ها و بهبود در مدل‌سازی نسبت به روش‌های قبلی پیش‌بینی (مقایسه نتایج در شکل (۶)) به طور کامل نتوانسته‌ایم منشا این خطا را حذف نماییم. هر چند با افزایش تعداد زیرباندهای تبدیل موجک و همچنین افزایش لایه‌های شبکه GMDH تاثیر این خطا کاهش می‌یابد اما حجم محاسبات بسیار بالا رفته و در مقایسه با بهبود اندکی که در دقت حاصل می‌شود به صرفه نمی‌باشد.

نام بار کاری و y_i نماینده عنصر i ام پیش‌بینی شده است. n نیز نشان دهنده تعداد عناصر بار کاری است [۳۶].

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{f_i - y_i}{f_i} \right| \quad (9)$$

در شکل ۹ نتایج حاصل از پیاده‌سازی الگوریتم‌های [7] ARMA، [21,25,33,36] ANN، [7,15,19,32] SVR و Wavelet-GMDH بر روی دو رد بار کاری واقعی ابری Intel pool B و TSDL برحسب MAPE نشان داده شده است. با توجه به شکل (۹) می‌توان به این نتیجه رسید که الگوریتم جدید Wavelet-GMDH در پیش‌بینی هر دو بار کاری خطای کمتری از دیگر الگوریتم‌های ARMA، ANN و SVR داشته است. در حالی که خطای میانگین روش پیشنهادی برای بار کاری Intel pool B ۸٫۲٪ است این خطا برای الگوریتم‌های ARMA، ANN و SVR به ترتیب ۲۴٫۵٪، ۲۰٫۱٪ و ۴۲٪ است که بهبود قابل توجه روش پیشنهادی را در مقابل سایر روش‌ها نشان می‌دهد.

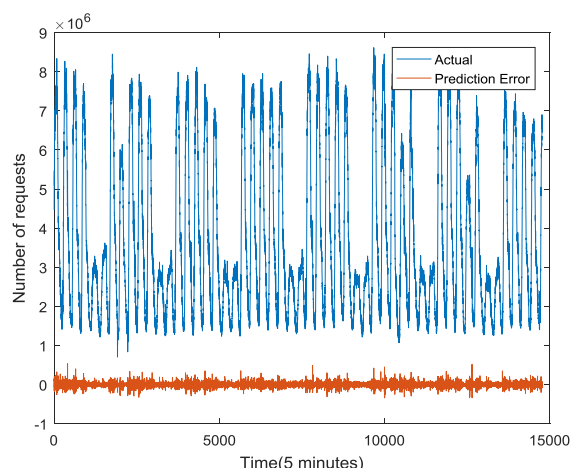


شکل (۹): مقایسه MAPE الگوریتم‌های پیش‌بینی مختلف

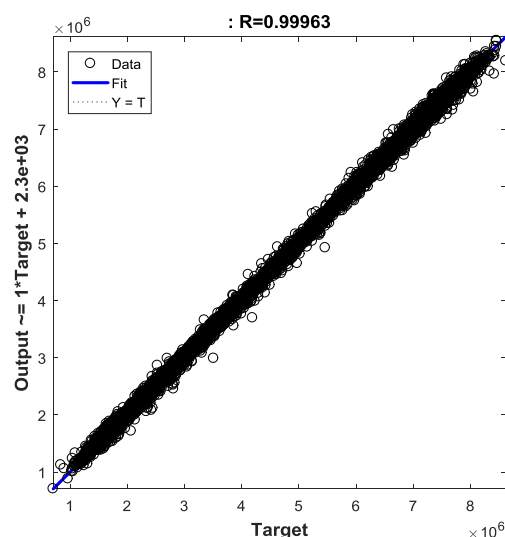
تکرار آزمایشات با رد بار کاری TSDL به خطای ۱٫۳٪ برای Wavelet-GMDH در پیش‌بینی منجر می‌شود. این در حالی است که این خطا برای الگوریتم‌های ARMA، ANN و SVR به ترتیب ۱۶٫۸٪، ۴٫۲٪ و ۱۷٫۳٪ می‌باشد. با توجه به نتایج بدست آمده از این آزمایشات و از آنجایی که هر یک از بارهای کاری رد واقعی یک سرویس ابری هستند، می‌توان نتیجه گرفت که الگوریتم Wavelet-GMDH ارائه شده به سبب مدل‌سازی دقیق‌تر و بهره‌مندی ویژگی‌های غیرخطی در هر یک از زیرباند‌های زمان-فرکانسی منجر به افزایش دقت پیش‌بینی و میانگین مطلق درصد خطای متوسط ۴٫۷۵ در پیش‌بینی رد دو بار کاری ابری رسیده است. این در حالی است که این مقدار برای نزدیک‌ترین رقیب این روش که ANN می‌باشد، ۱۲٫۱۵ می‌باشد. در نتیجه روش ارائه شده در حدود ۷٫۲ درصد دقت پیش‌بینی را نسبت به بهترین روش موجود افزایش داده است.

۵- نتیجه‌گیری

در این مقاله، روش جدیدی به نام Wavelet-GMDH برای پیش‌بینی بار کاری سرویس‌های مجازی‌سازی شده شبکه بر روی زیرساخت ابری



شکل (۷): بار کاری TSDL و خطای پیش‌بینی TSDL با الگوریتم Wavelet-GMDH



شکل (۸): پراکندگی پیش‌بینی بار کاری TSDL با الگوریتم Wavelet-GMDH

۴-۲- مقایسه نتایج پیش‌بینی با الگوریتم‌های مختلف

در این بخش برای نشان دادن برتری در دقت پیش‌بینی روش پیشنهادی Wavelet-GMDH با روش‌های مشابه موجود در مقالات نتایج آزمایشات با یکدیگر مقایسه شده است. برای ارزیابی دقت روش‌های پیش‌بینی معیارهای مختلفی وجود دارد که شرح جزئیات آنها در مرجع [۳۶] آورده شده است. در این مقاله ما از معیار میانگین مطلق درصد خطا یا MAPE که در رابطه (۹) نشان داده شده است، برای مقایسه دقت پیش‌بینی الگوریتم‌های مختلف استفاده کرده‌ایم. علت انتخاب این معیار این است که مستقل از مقیاس داده‌های ورودی می‌باشد و بدین ترتیب مقایسه مقادیر خطا بر روی بارهای کاری مختلف که محدوده تغییرات متفاوتی دارند به راحتی و مستقل از مقدار داده انجام می‌پذیرد. در رابطه (۹) مقدار f_i نشان دهنده مقدار عنصر

applications hosted in the cloud by means of weighted multi-objective linear programming." Proceedings of the 21st Brazilian Symposium on Multimedia and the Web. ACM, 2015.

- [7] Babu, C.N., Reddy, B.E. A moving-average filter-based hybrid ARIMA-ANN model for forecasting time series data. *Applied Soft Computing Journal*, 23, pp. 27-38. 2014. doi: 10.1016/j.asoc.2014.05.028.
- [8] Atashrouz, Saeid, Gholamreza Pazuki, and Younes Alimoradi. "Estimation of the viscosity of nine nanofluids using a hybrid GMDH-type neural network system." *Fluid Phase Equilibria* 372 (2014): 43-48.
- [9] Yaslan, Yusuf, and Bahadır Bican. "Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting." *Measurement* 103 (2017): 52-61.
- [10] Sohani, Ali, Hoseyn Sayyaadi, and Sina Hoseinpoori. "Modeling and multi-objective optimization of an M-cycle cross-flow indirect evaporative cooler using the GMDH type neural network." *International Journal of Refrigeration* 69 (2016): 186-204.
- [11] Jafarian, Hamoon, Hoseyn Sayyaadi, and Farschad Torabi. "Modeling and optimization of dew-point evaporative coolers based on a developed GMDH-type neural network." *Energy Conversion and Management* 143 (2017): 49-65.
- [12] Dargahi-Zarandi, Atefeh, et al. "Modeling gas/vapor viscosity of hydrocarbon fluids using a hybrid GMDH-type neural network system." *Journal of Molecular Liquids* 236 (2017): 162-171.
- [13] http://www.cs.huji.ac.il/labs/parallel/workload/1_int_el_netbatch/index.html
- [14] <https://datamarket.com/data/list/>
- [15] Hussain, Abir Jaafar, et al. "Regularized dynamic self-organized neural network inspired by the immune algorithm for financial time series prediction." *Neurocomputing* 188 (2016): 23-30.
- [16] Amiri, Maryam, and Leyli Mohammad-Khanli. "Survey on prediction models of applications for resources provisioning in cloud." *Journal of Network and Computer Applications* (2017).
- [17] Tong, Jun-jie, et al. "Host load prediction in cloud based on classification methods." *The Journal of China Universities of Posts and Telecommunications* 21.4 (2014): 40-46.
- [18] Prema, V., and K. Uma Rao. "Development of statistical time series models for solar power prediction." *Renewable Energy* 83 (2015): 100-109.
- [19] L. C. Parra; A. Romero; M. Martinez, "Solar energy production forecasting through artificial neuronal networks, considering the Föhn, north and south winds in San Juan, Argentina", *The Journal of Engineering*. Vol:2019, Issue:18, Page(s): 4824 - 4829, 2019.
- [20] Zhongda, Tian, et al. "A prediction method based on wavelet transform and multiple models' fusion for chaotic time series." *Chaos, Solitons & Fractals* 98 (2017): 158-172.
- [21] Y. Wang; Y. Ni; S. Lu; J. Wang; X. Zhang, "Remaining Useful Life Prediction of Lithium-Ion Batteries Using Support Vector Regression Optimized by Artificial Bee Colony". *IEEE Transactions on Vehicular Technology* (Early Access), 2019. DOI: 10.1109/TVT.2019.2932605.
- [22] Di, Sheng, Derrick Kondo, and Walfredo Cirne. "Google hostload prediction based on Bayesian model with optimized feature combination." *Journal of Parallel and Distributed Computing* 74.1 (2014): 1820-1832.

ارایه شده است. از آنجا که بارکاری واقعی ابر دارای تغییرات لحظه‌ای دامنه شدیدی است، به شدت غیرخطی بوده و پیش‌بینی آن به کمک روش‌های موجود پیش‌بینی سری زمانی با خطای زیادی همراه است. در این مقاله از ایده نگاشت سری زمانی مربوط به بارکاری ابری به زیرفضاهای مختلف زمان-فرکانس به کمک تبدیل موجک استفاده شده است. اینکار سبب می‌شود که هر زیرفضا را بتوان با مدلی با پیچیدگی کمتر بر داده‌ها افزا نمود. برای مدل کردن داده‌ها در هر زیرفضا از یک شبکه GMDH استفاده شده است که میزان غیرخطی بودن آنرا می‌توان با توجه به داده ورودی تنظیم نمود. از این مدل‌های GMDH مستقل برای پیش‌بینی سری زمانی در هر زیرباند استفاده شده است و درنهایت با برهم نهدی مقادیر پیش‌بینی هر زیر باند، مقدار پیش‌بینی شده برای بارکاری بازه بعدی را بدست می‌آوریم. برای ارزیابی الگوریتم Wavelet-GMDH دو رد بارکاری ابری که در مقالات قبلی استفاده شده بودند را به همراه روش‌های مشابه از خانواده‌های ARMA، ANN، SVR و ارزیابی نمودیم. نتایج ارزیابی حاکی از این است که روش Wavelet-GMDH دقت متوسط پیش‌بینی را اندکی بیش از ۷ درصد نسبت به بهترین روش قبلی افزایش داده است. علت این امر مدل‌سازی با خطینگی قابل کنترل GMDH و استفاده از مدل‌سازی چندباند زمان-فرکانسی به کمک تبدیل موجک می‌باشد. برای ادامه کار پیشنهاد می‌گردد که تعداد مناسب گره‌ها و لایه‌های شبکه GMDH با استفاده از الگوریتم‌های بهینه‌سازی نظیر آنچه در مرجع [۳۷] مطرح شده انتخاب گردد. همچنین الگوریتم فوق را می‌توان به منظور بهبود استفاده از منابع و انرژی در کاربردهای دیگری نظیر شبکه‌های ناهمگون بیسیم [۳۸] و یا شبکه‌های حسگر بیسیم [۳۹] و اینترنت اشیا نیز استفاده نمود.

مراجع

- [1] Jain, Raj, and Subharthi Paul. "Network virtualization and software defined networking for cloud computing: a survey." *IEEE Communications Magazine* 51.11 (2013): 24-31.
- [2] Mijumbi, Rashid, et al. "Network function virtualization: State-of-the-art and research challenges." *IEEE Communications Surveys & Tutorials* 18.1 (2016): 236-262.
- [3] Hawilo, Hassan, et al. "NFV: state of the art, challenges, and implementation in next generation mobile networks (vEPC)." *IEEE Network* 28.6 (2014): 18-26.
- [4] Messias, Valter Rogério, et al. "Combining time series prediction models using genetic algorithm to autoscaling Web applications hosted in the cloud infrastructure." *Neural Computing and Applications* 27.8 (2016): 2383-2406.
- [5] Messias, Valter Rogério, Julio Cezar Estrella, and Ricardo Ehlers. "Efficient resource allocation for web applications hosted in the cloud by means of weighted multi-objective linear programming." *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web*. ACM, 2015.
- [6] Messias, Valter Rogério, Julio Cezar Estrella, and Ricardo Ehlers. "Efficient resource allocation for web

آشوب"، مجله مهندسی برق و الکترونیک ایران، جلد ۱۶ شماره ۳، صفحات ۱۴۷-۱۵۵، ۱۳۹۸.

- [۳۸] خرسندی، حقیقت دوست، "ارائه مدل اقتصادی رقابت انحصاری چندجانبه برای تخصیص پهنای باند و قیمت‌گذاری عادلانه در شبکه‌های دسترسی ناهمگون بیسیم"، مجله مهندسی برق و الکترونیک ایران - جلد ۱۶ شماره ۴، ۱۳۹۷.
- [۳۹] عباس‌زاده، حسینی سنو، "ساخت توپولوژی انرژی آگاه با مکانیزم نگهداری در شبکه‌های حسگر بی‌سیم"، مجله مهندسی برق و الکترونیک ایران، جلد ۱۴ شماره ۲، صفحات ۳۶-۳۳، ۱۳۹۶.

¹ Network function virtualization

² capital expenditures

³ operating expenditures

⁴ seasonal

⁵ Approximation

⁶ Details

- [23] Shyam, Gopal Kirshna, and Sunilkumar S. Manvi. "Virtual resource prediction in cloud environment: A Bayesian approach." *Journal of Network and Computer Applications* 65 (2016): 144-154.
- [24] Doucoure, Boubacar, Kodjo Agbossou, and Alben Cardenas. "Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data." *Renewable Energy* 92 (2016): 202-211.
- [25] Amezcuita-Sanchez, Juan P., and Hojjat Adeli. "A new music-empirical wavelet transform methodology for time-frequency analysis of noisy nonlinear and non-stationary signals." *Digital Signal Processing* 45 (2015): 55-68.
- [26] Ivakhnenko, A.C. Polynomial theory of complex systems. *IEEE Trans. Syst. Man Cybern.* 1971, 1, 364-378.
- [27] Barzegar, Rahim, Jan Adamowski, and Asghar Asghari Moghaddam. "Application of wavelet-artificial intelligence hybrid models for water quality prediction: a case study in Aji-Chay River, Iran." *Stochastic Environmental Research and Risk Assessment* 30.7 (2016): 1797-1819.
- [28] Li, Qiumin, Yixiang Tian, and Gaoxun Zhang. "The k-nearest neighbour-based GMDH prediction model and its applications." *International Journal of Systems Science* 45.11 (2014): 2301-2308.
- [29] Song, Baiyi, Carsten Ernemann, and Ramin Yahyapour. "Parallel computer workload modeling with markov chains." *Workshop on Job Scheduling Strategies for Parallel Processing*. Springer Berlin Heidelberg, 2004.
- [30] Smith, Warren, Valerie Taylor, and Ian Foster. "Using run-time predictions to estimate queue wait times and improve scheduler performance." *Workshop on Job Scheduling Strategies for Parallel Processing*. Springer Berlin Heidelberg, 1999.
- [31] Tian, Wenhong, and Chee Shin Yeo. "Minimizing total busy time in offline parallel scheduling with application to energy efficiency in cloud computing." *Concurrency and Computation: Practice and Experience* 27.9 (2015): 2470-2488.
- [32] Donate, Juan Peralta, et al. "Time series forecasting by evolving artificial neural networks with genetic algorithms, differential evolution and estimation of distribution algorithm." *Neural Computing and Applications* 22.1 (2013): 11-20.
- [33] Cortez, Paulo, and Juan Peralta Donate. "Global and decomposition evolutionary support vector machine approaches for time series forecasting." *Neural Computing and Applications* 25.5 (2014): 1053-1062.
- [34] Rajashree Dash. Performance analysis of a higher order neural network with an improved shuffled frog leaping algorithm for currency exchange rate prediction *Applied Soft Computing*. Vol 67, Pages 215-231, 2018. doi: 10.1016/j.asoc.2018.02.043
- [35] Ohad Shai, Edi Shmueli, and Dror G. Feitelson, "Heuristics for resource matching in Intel's compute farm". In *Job Scheduling Strategies for Parallel Processing*, Walfredo Cirne and Narayan Desai, (ed.), Springer-Verlag, 2013.
- [36] Barati, Masoud, and Saeed Sharifian. "A hybrid heuristic-based tuned support vector regression model for cloud load prediction." *Journal of Supercomputing* 71.11 (2015).

[۳۷] کلانتری، سهرابی، رشیدی کنعان، " الگوریتم بهینه‌سازی ترکیبی با استفاده از الگوریتم ژنتیک و جستجوی ابرکروی مبتنی بر نظریه

