

Thermal Prediction in Multicore processors using Support Vector Regression

Javad Mohebbi Najm Abad ¹, Ali Soleimani ²

¹ Dept. of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

² Dept. of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

Abstract:

Increasing the number of processor cores leads to increasing the density of the computing power processor and also raising the temperature. Temperature management is very important in these processors. Thermal management methods are introduced to reduce the CPU temperature. Reactive and proactive approaches are two sets of these schemes. Unlike the reactive techniques, proactive methods predict the temperature using thermal prediction model before reaching its threshold. In this paper, a hybrid model of several SVR models is proposed for predicting temperature. An appropriate dataset is created for training proposed model that includes a high diversity of processor temperature variations. Some features of dataset are measured using temperature sensors and system performance counters. Other features, with historical and control names are calculated with the proposed processes to increase the accuracy of thermal model. Two SVR models are used in the proposed thermal model to reduce its operational overhead. The proper features for each SVR model are selected by the feature selection algorithm based on mutual information. The proposed model is evaluated for temperature prediction for 2 to 5 time distances. The results show that with a selection of 11 features for thermal prediction model of the next 2 seconds, the mean absolute error is about 0.5 °C.

Keywords: Mutual Information, Feature Selection, Thermal Prediction, Support Vector Regression, Dynamic Thermal Management.

پیش‌بینی دما در پردازنده‌های چند هسته‌ای با استفاده از رگرسیون بردار پشتیبان

جواد محبی نجم‌آباد^۱، علی سلیمانی^۲

۱- دانشجوی دکتری، دانشکده مهندسی کامپیوتر - دانشگاه صنعتی شاهرود - شاهرود - ایران

javad.mohebi@shahroodut.ac.ir

۲- دانشیار، دانشکده مهندسی کامپیوتر - دانشگاه صنعتی شاهرود - شاهرود - ایران

solimani_ali@shahroodut.ac.ir

چکیده: افزایش تعداد هسته‌های پردازنده به افزایش توان محاسباتی و به موجب آن، افزایش دمای پردازنده منجر می‌شود. کنترل و مدیریت دما در این پردازنده‌ها اهمیت زیادی دارد. مدیریت دما به دو رویکرد واکنشی و فعال تقسیم می‌شود. برخلاف رویکرد واکنشی، در روش‌های فعال، دما پیش از رسیدن به حد آستانه با استفاده از مدل دمایی، پیش‌بینی و کنترل می‌شود. در این مقاله، برای پیش‌بینی دما، مدلی براساس SVR پیشنهاد شده است. برای آموزش مدل، با استفاده از حسگرهای دمایی و شمارنده‌های کارایی موجود در داخل پردازنده، مجموعه داده‌ای شامل تنوع زیادی از تغییرات دمایی جمع‌آوری شده است. برای افزایش دقت مدل، ویژگی‌های دیگری با نام‌های سابقه‌ای و کنترلی از ویژگی‌های موجود استخراج شده‌اند. برای کاهش سربار محاسباتی دو راهکار پیشنهاد شده است؛ یکی استفاده از دو مدل SVR و دیگری انتخاب ویژگی مناسب بر پایه اطلاعات متقابل. در انتها، مدل پیشنهادی برای پیش‌بینی دما برای فاصله‌های زمانی ۲ تا ۵ ثانیه، در شرایط کاری مختلف ارزیابی شده است. نتایج نشان می‌دهند با انتخاب ۱۱ ویژگی دمایی ۲ ثانیه آینده با میانگین قدر مطلق خطای ۰/۵ درجه سانتی‌گراد پیش‌بینی می‌شود.

واژه‌های کلیدی: اطلاعات متقابل، انتخاب ویژگی، پیش‌بینی دما، رگرسیون بردار پشتیبان، مدیریت دمای پویا.

۱- مقدمه

انجام عملیات زیادی دارند. برای افزایش توان محاسباتی در این نوع پردازنده‌ها، افزایش هسته‌ها در ناحیه محدودی از سطح نیمه‌هادی صورت گرفته است. این موضوع به افزایش توان مصرفی و به موجب آن، افزایش دما در پردازنده منجر می‌شود. دمای زیاد باعث کاهش قابلیت اطمینان و کاهش طول عمر پردازنده می‌شود [۱].

یکی از رویکردهای کنترل دمای پردازنده، مدیریت دمای پویا^۱ (DTM) نام دارد. در این رویکرد، در زمان اجرای سیستم از روش‌های مناسب برای کاهش دما استفاده می‌شود [۱]. از دسته‌بندی‌های DTM، تقسیم‌بندی آن به دو رویکرد واکنشی^۲ و فعال^۳ است. در رویکرد واکنشی، بعد از رسیدن دما به حد آستانه پردازنده، از روش‌هایی بهره می‌برند که سریع‌تر به کاهش دما منجر می‌شوند. این روش‌ها بیشتر باعث کاهش عملکرد سیستم می‌شوند. در

برای افزایش عملکرد پردازنده‌ها از روش‌های مختلفی برای افزایش توان محاسباتی آنها بهره برده‌اند. نخست، از افزایش فرکانس اعمالی به پردازنده، برای افزایش توان محاسباتی آن استفاده کردند. افزایش فرکانس پردازنده علاوه بر افزایش هزینه تولید، به افزایش گرمای تولیدی پردازنده‌ها منجر می‌شود. بدین منظور، پردازنده‌های چند هسته‌ای ارائه شدند که با وجود استفاده از فرکانس‌های پایین‌تر، سرعت

^۱ تاریخ ارسال مقاله: ۱۳۹۶/۰۱/۲۰

تاریخ پذیرش مقاله: ۱۳۹۷/۰۱/۲۲

نام نویسنده مسئول: علی سلیمانی

نشانی نویسنده مسئول: ایران - شاهرود - خیابان دانشگاه -

دانشگاه صنعتی شاهرود - دانشکده برق

در این مقاله، یک مدل پیش‌بینی دما پیشنهاد شده است که ترکیبی از دو مدل SVR است. ورودی‌های مدل پیشنهادی شامل فرکانس پردازنده، سرعت فن، دمای هسته‌های پردازنده و تعدادی از شمارنده‌های کارایی هستند. همچنین، با هدف افزایش دقت مدل پیش‌بینی دما، تعدادی ویژگی با استفاده از روش‌های پیشنهادی ارائه می‌شوند. برای هر یک از مدل‌های SVR به صورت مجزا، ویژگی‌های مناسب انتخاب می‌شوند. همچنین، درباره تنظیم مدل برای پیش‌بینی دما برای فاصله‌های زمانی مختلف، مطالعاتی انجام شده است. برای بیان واضح‌تر، به صورت مختصر نوآوری‌های این مقاله عبارت‌اند از:

- برای مدل پیش‌بینی دمای پیشنهادی، ویژگی‌هایی جدید با پردازش‌های پیشنهادی ایجاد شده است که تأثیر چشمگیری در دقت مدل دارد.
 - در مدل پیشنهادی از دو مدل SVR استفاده شده است که برای آموزش هر یک از آنها مجموعه داده مناسب ایجاد می‌شود. برای افزایش عملکرد مدل، برای هر یک از مدل‌های SVR، ویژگی‌های ورودی مناسب انتخاب می‌شوند.
 - دقت مدل پیشنهادی، برای فاصله‌های مختلف زمانی با توجه به ویژگی‌های انتخابی در سیستم چند هسته‌ای بررسی شده است.
- در ادامه، این مقاله به صورت زیر سازمان‌دهی شده است: در بخش دوم مدل‌های پیش‌بینی دمای پردازنده بررسی شده‌اند. در بخش سوم، روش SVR و انتخاب ویژگی بر پایه اطلاعات متقابل معرفی شده‌اند. در بخش چهارم، مدل پیش‌بینی دمای پیشنهادی معرفی می‌شود. ارزیابی مدل پیشنهادی برای فاصله‌های مختلف زمانی در بخش پنجم ارائه می‌شود. در نهایت، در بخش ششم، نتیجه‌گیری کار بیان می‌شود.

۲- ادبیات پژوهش

بسیاری از مدل‌های پیش‌بینی دمای معرفی شده، با استفاده از آمارهای جمع‌آوری شده از اجرای برنامه‌ها بر روی پردازنده، ایجاد شده و آموزش دیده‌اند [۶،۷]. بیشتر از این آمارها، در روش رگرسیون، برای تعیین ضریب‌های

مقایسه با روش مذکور، رویکردهای فعال DTM، پیش از رسیدن دمای پردازنده به مقدار نامطلوب، از آن جلوگیری می‌کنند. پیش‌بینی دقیق دمای پردازنده برای فاصله‌های زمانی مطلوب از بخش‌های مهم رویکرد فعال DTM است. از طرفی، در هر دو رویکرد، پیش‌بینی دما پیش از استفاده از روش‌های کنترل دما، این امکان را ایجاد می‌کند تا روشی مناسب را برای افزایش عملکرد سیستم انتخاب کنند.

برخی از مدل‌های پیش‌بینی دما، از روش‌های ریاضی مانند روش‌های درونیابی بر مجموعه داده‌ای از رکوردهای فعالیت سیستم استفاده می‌کنند [۵-۲]. این رکوردها، دما و اطلاعات وضعیت پردازنده را شامل می‌شود که در زمان اجرای بارکاری، با نظارت حسگرهای دمایی و شمارنده‌های کارایی پردازنده، جمع‌آوری می‌شوند. همچنین از روش‌های هوش مصنوعی و روش‌های یادگیری ماشین نیز برای ایجاد این مدل‌ها استفاده شده است [۱۰-۶]. مدل رگرسیون بردار پشتیبان (SVR) مدلی است که به دلیل سرعت و دقت زیاد، در سال‌های اخیر برای پیش‌بینی دما استفاده می‌شده است [۸،۶]. این دو ویژگی مدل SVR، از دلایل استفاده از آن در مدل دمایی پیشنهادی در این مقاله است.

برای اینکه مدل‌های پیش‌بینی دما در رویکردهای DTM، بهره‌وری مناسب داشته باشند، باید ویژگی‌هایی داشته باشند. از ویژگی‌های کلیدی آن، سربرار محاسباتی است. سربرار پایین مدل، به روش‌های DTM این امکان را می‌دهد تا در بازه‌های زمانی کوتاه‌تری از آن استفاده کنند. به علاوه، مدل‌ها باید توانایی پذیرفتنی در پیش‌بینی حالت‌هایی داشته باشند که با آنها آموزش داده نشده‌اند. باید توجه داشت فراهم کردن مجموعه داده همه حالت‌های ممکن سیستم، غیرممکن است. از دیگر ویژگی‌های مدل این است که بتواند برای پیش‌بینی دما در فاصله‌های مختلف زمانی استفاده شود. هرچه پیش‌بینی برای فاصله زمانی طولانی‌تر انجام شود، قابلیت انتخاب روش‌های متنوع برای مدیریت دما فراهم می‌شود. در نهایت، دقت، مهم‌ترین ویژگی مدل است. پیش‌بینی دقیق دما به رویکردهای DTM این توانایی را می‌دهند که در اتخاذ تصمیم‌های کنترلی بهینه، علاوه بر حفظ قید دمایی پردازنده، حداقل کاهش سطح کارایی برنامه را موجب شوند.

نمونه‌های مشاهده شده مشخص می‌شود. p و q به ترتیب اندازه پنجره AR و MA را مشخص می‌کنند. با توجه به تأثیر چشمگیر تغییرات دمای همسایه‌های یک هسته بر دمای آن، در تعدادی از مدل‌های پیشنهادی، برای پیش‌بینی دمای یک هسته، به مقادیر دمای دیگر هسته‌ها نیز توجه شده است [۱۱، ۱۲، ۱۳]. در نمونه‌ای از کارهای ارائه شده، از رابطه ۲ برای پیش‌بینی دما استفاده شده است [۴].

$$T_i^{pred} = \alpha_i T_i^{curr} + \beta_i I_i^{in} + \gamma_i I_i^{nb} \quad (2)$$

T_i^{curr} دمای فعلی هسته، I_i^{in} گرادیان دمایی هسته و I_i^{nb} میانگین گرادیان دمایی هسته‌های همسایه هستند. مقدار ضریب‌های α_i ، β_i و γ_i به صورت برون‌خط، با استفاده از رگرسیون خطی، بر مجموعه داده جمع‌آوری شده از دمای هسته‌های پردازنده محاسبه می‌شوند. در این روش به این نکته توجه نشده است که در سیستم عملی، تأثیر هسته‌های مجاور یک هسته در دمای آن متفاوت است.

از مهم‌ترین مشخصه‌های یک مدل پیش‌بینی دما، توجه آن به رفتار بارکاری در حال اجرا است. در بیشتر مدل‌های ارائه شده، برای بررسی رفتار برنامه از ابزارهای سنجش فراهم شده برای پردازنده مانند شمارنده‌های کارایی و نشانگرهای توان مصرفی بهره برده‌اند [۱۳، ۱۴، ۱۵]. در این راستا، برای تخمین دما از رابطه ۳ استفاده شده است که ترکیبی از شمارنده‌های کارایی است [۵].

$$T_{overall} = w_{const} + \sum_{i=1}^{22} w_i \frac{u_i}{t_{total}} \quad (3)$$

u_i مقدار شمارنده کارایی نام و t_{total} تعداد چرخه‌های پردازنده است. w_i ها ضریب‌های شمارنده‌های کارایی هستند. برای تعیین ضریب‌ها، ابتدا با اجرای برنامه‌های مختلف روی پردازنده، مجموعه داده لازم ایجاد می‌شود؛ آنگاه با درون‌یابی خطی روی مجموعه داده، ضریب‌ها محاسبه می‌شوند. همچنین، w_{const} دمای اندازه‌گیری شده پردازنده در حالت بیکاری است.

از دیگر کارهای انجام شده برای ایجاد مدل پیش‌بینی دما با توجه به رفتار برنامه‌ها، کار کچران و همکاران است [۱۱]. آنها با استفاده از شمارنده‌های کارایی، فازهای رفتاری بارکاری را شناسایی کرده و برای هر فاز، از مدل جداگانه

رابطه دمایی چندجمله‌ای استفاده می‌شوند. در یک تقسیم‌بندی، مدل‌های پیش‌بینی دما به دو دسته برون‌خط و برخط متمایز می‌شوند. مدل‌های برون‌خط با استفاده از مجموعه داده‌های جمع‌آوری شده، آموزش می‌بینند. سپس، در زمان اجرا، از مدل‌های فراهم شده برای پیش‌بینی برخط دما استفاده می‌شود [۱۱، ۶]. این نوع الگوها سربار آموزش را در زمان اجرا ندارند؛ ولی در صورتی که با تغییرهای ناگهانی دما برخورد کنند، ممکن است دقت مناسب را در پیش‌بینی ارائه نکنند. در مقابل این رویکرد، استفاده از روش‌های برخط برای پیش‌بینی دمای پردازنده، با هدف افزایش دقت مدل، شایان توجه قرار گرفته است [۳، ۲]. این نوع مدل‌های پیش‌بینی دما، در زمان اجرا روی پنجره‌ای از آخرین مشاهدات، آموزش مدل را به صورت برخط انجام می‌دهند؛ برای مثال، اگر از رابطه‌ای چندجمله‌ای برای مدل پیش‌بینی استفاده شود، ضرایب رابطه با استفاده از آخرین پنجره مشاهدات سیستم تعیین می‌شوند. در این حالت، در صورت استفاده از پنجره‌ای با اندازه m ، نمونه m ام، با اجرای تابع $f(t_{n-1}, t_{n-2}, \dots, t_{n-m})$ روی m نمونه قبلی، پیش‌بینی دما را انجام می‌دهد. برای تابع f که در آن t_i دمای نمونه i ام است، از تابع‌های میانگین ساده، وزن‌دار و نمایی برای پیش‌بینی t_n استفاده می‌شود [۶].

روش‌های تحلیل سری زمانی برای پیش‌بینی برخط دما استفاده شده‌اند. در این راستا، مدل‌های خود رگرسیون میانگین متحرک^۰ (ARMA) که همبستگی را در سری‌های زمانی شناسایی می‌کنند، در تعدادی از کارها برای پیش‌بینی دما استفاده شده‌اند [۳، ۲]. معادله ARMA شامل دو بخش خود رگرسیون (AR) و میانگین متحرک (MA) است که در رابطه ۱ نمایش داده شده است.

$$y_t + \sum_{i=1}^p (a_i y_{t-i}) = e_t + \sum_{i=1}^q (c_i e_{t-i}) \quad (1)$$

y_t مقدار دما در زمان t و a_i ها ضریب مقادارهای پیشین دما بوده است که در بخش AR معادله قرار گرفته‌اند. e_{t-i} خطاهای به دست آمده و c_i ها ضریب مقدار خطاها هستند. این بخش از معادله، بخش MA در معادله معرفی می‌شود. مقدار e_t نیز در زمان تعیین ضریب‌ها روی پنجره

از ۳۴ شمارنده کارایی برای پیش‌بینی دمای آینده استفاده شده است. باید توجه داشت ویژگی‌های مهمی از سیستم، از جمله دمای فعلی هسته‌ها در آن لحاظ نشده است. برای افزایش دقت مدل پیش‌بینی دما، انتخاب ویژگی‌های مناسب از مهم‌ترین نیازهای مدل است. در کارهای متعددی، پارامترهای مدل‌های پیش‌بینی دما، به صورت محدود و با توجه به بینش و مهارت‌های فردی محقق انتخاب شده‌اند. در برخی دیگر از مدل‌ها نیز، از تعداد زیادی پارامتر، بدون توجه به وابستگی آنها با دمای آینده و البته وابستگی بین خود پارامترها استفاده شده است. تعداد ویژگی‌های زیاد باعث کاهش دقت مدل و همچنین افزایش سرریز آموزش و اجرای آن می‌شود [۱۶، ۱۷]. انتخاب مجموعه ویژگی‌های مؤثر با حداقل تعداد لازم، تأثیر چشمگیری در افزایش دقت مدل‌های پیش‌بینی دما دارند.

۳- مبانی نظریه

در این بخش، مدل SVR و روش کاهش ویژگی بر پایه اطلاعات متقابل معرفی می‌شوند.

۳-۱- رگرسیون بردار پشتیبان (SVR)

رگرسیون بردار پشتیبان بر پایه مدل طبقه‌بندی ماشین‌های بردار پشتیبان ارائه شده است [۱۸]. SVR برای حل مسائل رگرسیون غیرخطی استفاده می‌شود [۱۸-۲۰]. برای توضیح مختصر SVR، فرموله‌سازی آن انجام می‌شود. در این راستا، مجموعه داده M که با رابطه ۵ نمایش داده شده، شامل بردارهای ورودی x_i و خروجی متناظر y_i است.

$$M = \left\{ (x_i, y_i) \mid i = 1, 2, \dots, n \right. \\ \left. , x_i \in R^N, y_i \in R \right\} \quad (5)$$

n تعداد نمونه‌های مجموعه داده را نشان می‌دهد. هدف تحلیل رگرسیون، تعیین تابع $f(x)$ به گونه‌ای است که خروجی پیش‌بینی آن حداقل خطا را نسبت به خروجی مطلوب داشته باشد. تابع رگرسیون با رابطه ۶ معرفی می‌شود که δ خطای تصادفی با توزیع $N(0, \sigma^2)$ است.

$$y_i = f(x_i) + \delta \quad (6)$$

بهره برده‌اند. نخست، مجموعه داده‌ای شامل تعداد زیادی شمارنده کارایی حاصل از اجرای برنامه‌ها روی سیستم تهیه کرده‌اند؛ آنگاه با استفاده از تحلیل مؤلفه‌های اصلی^۶ (PCA)، خروجی این شمارنده‌ها را به دو جزء اصلی کاهش داده‌اند. سپس، برای تعیین فازهای رفتاری مختلف از خوشه‌بندی KMeans روی نمونه‌های جمع‌آوری شده استفاده کرده‌اند. در نهایت، برای هر فاز با رابطه ۴، مدل مجزایی برای پیش‌بینی دما تنظیم کرده‌اند.

$$g_m[i] = T_m[i] - \sum_{n=1}^N a_{mn} T_n[i-1] - a_m^{idle} \quad (4)$$

$T_m[i]$ دمای هسته m در زمان i و a_{mn} و a_m^{idle} ضریب‌های مدل از دید هسته m هستند. مقدار g_m اثر بارکاری را بر دمای هسته m تعیین می‌کند. به سبب اینکه در این روش، برای هر یک از حالت‌های ممکن سیستم، به ضریب‌های جداگانه‌ای برای مدل دما نیاز است، تعداد مدل‌های موردنیاز این رویکرد، افزایش می‌یابد.

برای ایجاد مدل‌های پیش‌بینی دما، روش‌های یادگیری ماشین نیز به کار برده شده‌اند. از شبکه عصبی پرسپترون چندلایه‌ای^۷ (MLP) برای پیش‌بینی دمای آینده بهره گرفته‌اند [۷، ۹، ۱۰]. در این کارها، برای پیش‌بینی دمای آینده، از دمای فعلی و توان مصرفی بهره گرفته‌اند. در مدل‌های پیشنهادی، به صورت مستقیم به فرکانس پردازنده، سرعت فن و بارکاری اجرایی روی هر یک از هسته‌های پردازنده توجه نمی‌شود.

یکی از کارایی‌ترین روش‌های پیشنهادی برای پیش‌بینی دما، استفاده از روش SVR است [۶، ۸]. در یکی مهم‌ترین پژوهش‌هایی که به‌تازگی درباره این موضوع ارائه شده است، از روش SVR و ابزار HotSpot [۱۵] بهره برده‌اند [۶]. در این راستا، آموزش SVR برای پیش‌بینی رابطه توان مصرفی با شمارنده‌های کارایی، با استفاده از داده‌های جمع‌آوری شده انجام می‌شود؛ آنگاه در زمان اجرا، توان پیش‌بینی شده با استفاده از SVR به ابزار پیکربندی شده HotSpot داده می‌شود تا دمای هسته را پیش‌بینی کند. داشتن دانش معماری پردازنده از مهم‌ترین نیازهای پیکربندی ابزار HotSpot است. از پژوهش‌هایی دیگر انجام شده با روش SVR می‌توان به [۸] اشاره کرد که در آن

برای ایجاد شکل دوگان رابطه ۱۴، برای هر یک از قیدها از ضریب لاگرانژ، استفاده و سپس ساده‌سازی انجام می‌شود. در صورتی که α_i^+ و α_i^- ، به ترتیب ضریب قیدهای اول و دوم رابطه ۱۴ باشند، بعد از ساده‌سازی، رابطه ۱۵ حاصل می‌شود [۱۸].

$$\begin{aligned} \text{minimize } & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i^+ - \alpha_i^-) (\alpha_j^+ - \alpha_j^-) < \varphi(x_i) \cdot \varphi(x_j) > - \sum_{i=1}^n (\alpha_i^+ - \alpha_i^-) y_i + \quad (15) \\ \text{S.t. } & \begin{cases} \sum_{i=1}^n (\alpha_i^+ - \alpha_i^-) = 0 \\ \alpha_i \cdot \alpha_i^* \in [0, C] \end{cases} \end{aligned}$$

برای مسائل غیرخطی ضرب داخلی دو تابع $\varphi(x_i)$ و $\varphi(x_j)$ با تابع کرنل گوسی ارائه‌شده در رابطه ۱۶ جایگزین می‌شود [۱۸].

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (16)$$

در نهایت، تابع f با رابطه ۱۷ محاسبه می‌شود.

$$f(x) = \sum_{i=1}^n (\alpha_i^+ - \alpha_i^-) K(x_i, x) + b \quad (17)$$

در این رابطه، b با رابطه ۱۸ محاسبه می‌شود. SV بردار پشتیبان است.

$$\begin{aligned} b = \frac{1}{n} \left\{ \sum_{0 < \alpha_i^+ < C} \left[y_i - \sum_{x_j \in SV} (\alpha_j^+ - \alpha_j^-) K(x_i, x_j) - \varepsilon \right] \right. \\ \left. + \sum_{0 < \alpha_i^- < C} \left[y_i - \sum_{x_j \in SV} (\alpha_j^+ - \alpha_j^-) K(x_i, x_j) + \varepsilon \right] \right\} \quad (18) \end{aligned}$$

۳-۲- انتخاب ویژگی بر پایه اطلاعات متقابل

انتخاب مجموعه ویژگی‌ها با هدف بیشینه‌سازی معیار وابستگی آماری آنها با خروجی مدل و کمینه‌کردن اطلاعات متقابل MI^A بین آنها انجام می‌شود [۱۶، ۱۷، ۲۱، ۲۲]. در این راستا رابطه‌های مربوطه معرفی می‌شود. مقدار MI بین دو ویژگی x و y با رابطه ۱۹ محاسبه می‌شود [۱۶].

$$I(x; y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (19)$$

برای حل مسئله رگرسیون غیرخطی با SVR ، نخست، ورودی‌ها به صورت غیرخطی به فضای ویژگی f با ابعاد زیاد نگاشت می‌شود که به صورت خطی با خروجی وابستگی دارند؛ بدین منظور از رابطه ۷ استفاده می‌شود [۱۸].

$$f(x_i) = w\varphi(x_i) + b \quad |w \in R^{N \times h}, b \in R \quad (7)$$

w بردار وزن، b مقدار بایاس و $\varphi(x)$ تابعی است که ورودی‌ها را از فضای R به فضای $R^{N \times h}$ تصویر می‌کند. در SVR هدف پیدا کردن تابع f به گونه‌ای است که به اندازه ε از مقادیر واقعی انحراف داشته باشد و با وجود این، خطی نیز باشد. برای داده‌های خارج از باند از یک تابع جریمه با رابطه ۸ تعریف می‌شود [۱۸].

$$L_\varepsilon(y_i, f(x_i)) = \begin{cases} 0 & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{otherwise} \end{cases} \quad (8)$$

از طرفی ریسک عملیاتی تابع f با رابطه ۹ محاسبه می‌شود.

$$R_{emp}[f] = \sum_{i=1}^n L_\varepsilon(y_i, f(x_i)) \quad (9)$$

کاهش ریسک عملیاتی برای همه داده‌ها با توجه به بهینه‌سازی تابع رگرسیون با رابطه ۱۰ انجام می‌شود [۱۸].

$$J = \frac{1}{2} \|w\|^2 + C R_{emp}[f] \quad (10)$$

C ضریب ثابت تابع ریسک است. برای داده‌هایی که مقدار $|y - f(x_i)|$ آنها بیشتر از ε باشد، مقدار ξ_i^+ و یا ξ_i^- که مقدار تخطی را نشان می‌دهد، با رابطه‌های ۱۱ و ۱۲ محاسبه می‌شود.

$$\xi_i^+ = y - f(x_i) - \varepsilon \quad (11)$$

$$\xi_i^- = \varepsilon - y - f(x_i) \quad (12)$$

از طرفی رابطه ۱۳ بین تابع جریمه و مقدارهای تخطی برقرار است.

$$L_\varepsilon(y_i, f(x_i)) = \xi_i^+ + \xi_i^- \quad (13)$$

در نهایت، تابع هدف کلی برای تخمین تابع f با رابطه ۱۴ محاسبه می‌شود [۱۸].

$$\begin{aligned} \text{minimize } & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i^+ + \xi_i^-) \\ \text{S.t. } & \forall i \quad (14) \end{aligned}$$

$$\begin{aligned} -y_i + f(x_i) + \varepsilon + \xi_i^+ & \geq 0 \\ y_i - f(x_i) + \varepsilon + \xi_i^- & \geq 0 \\ \xi_i^+, \xi_i^- & \geq 0 \end{aligned}$$

$$\max_{i \in \varphi(S)} [I(i, h) - \frac{1}{|S|} \sum_{j \in S} I(i, j)] \quad (23)$$

$$\max_{i \in \varphi(S)} \{I(i, h) / [\frac{1}{|S|} \sum_{j \in S} I(i, j)]\} \quad (24)$$

۴- مدل دمایی پیشنهادی

در این بخش روال توسعه مدل پیش‌بینی دمای پیشنهادی معرفی می‌شود. برای توسعه مدل، سه مرحله آماده‌سازی مجموعه داده، پالایش مجموعه ویژگی‌ها و آموزش و آزمایش SVR انجام می‌شود. این مراحل به ترتیب در بخش‌های بعدی معرفی می‌شوند.

۴-۱- ایجاد مجموعه داده

در یک پردازنده چند هسته‌ای، علاوه بر وجود منابع مشترک بین هسته‌ها، هریک از هسته‌ها اجزای اختصاصی نیز دارند. در بسیاری از پردازنده‌ها از فرکانس ساعت یکسان برای همه هسته‌ها استفاده می‌شود. برای خنک کردن مکانیکی، از فن خنک‌کننده به همراه حرارت گیر^{۱۲} بهره می‌برند. پردازنده‌های کنونی از تعدادی شمارنده کارایی برای شمارش رخداد‌های مربوط به سیستم عامل، برنامه‌ها و سرویس‌های در حال اجرا روی پردازنده استفاده می‌کنند. بیشتر پردازنده‌ها مجموعه‌ای اختصاصی از شمارنده‌های کارایی برای هر هسته دارند. همچنین، برای فراهم کردن اطلاعات تکمیلی، از حسگرهایی مانند حسگر دما برای هر هسته پردازنده بهره می‌برند.

مدل پیش‌بینی دمای پیشنهادی روی سیستم چهار هسته‌ای Intel's Core i7-950 بررسی می‌شود. با استفاده از ابزار Jm-sensor اطلاعات دمایی از حسگرها دریافت می‌شود. تنظیم پویای ولتاژ و فرکانس، با استفاده از ابزار cpufreq انجام می‌شود. همچنین، انجام تغییرات سرعت فن با استفاده از اسکریپت هسته‌ای fancontrol است. شمارنده‌های کارایی با استفاده از زیرسیستم perf موجود در سیستم عامل لینوکس خوانده می‌شوند. لیست شمارنده‌های کارایی استفاده شده در مدل پیشنهادی در جدول (۱) نمایش داده شده است. همچنین، از برنامه‌های محک SPEC CPU2006، که برنامه‌های تهیه شده برای ایجاد تغییرات

$p(x, y)$ و $p(y)$ ، $p(x)$ به ترتیب توابع چگالی احتمال متغیرهای x ، y و وقوع هم‌زمان آنها هستند. با استفاده از رابطه ۱۹، وابستگی همه زیرمجموعه‌های ممکن از مجموعه ویژگی‌های اولیه با خروجی مدل محاسبه می‌شود. با استفاده از رابطه ۲۰، زیرمجموعه با بیشترین مقدار وابستگی انتخاب می‌شود [۱۶].

$$\max V_I(S, h), \quad V_I = \frac{1}{|S|} \sum_{i \in S} I(i, h) \quad (20)$$

S زیرمجموعه‌ای از ویژگی‌های اولیه و V_I مقدار وابستگی S را با خروجی مدل h نشان می‌دهد.

در صورتی که مجموعه انتخابی شامل ویژگی‌هایی باشد که همبستگی زیادی با هم داشته باشند، این امکان وجود دارد که ویژگی‌های انتخابی، بهترین انتخاب برای شناسایی سیستم نباشند. در این راستا، از شرط حداقل افزونگی برای یافتن زیرمجموعه‌ای از ویژگی‌ها با کمترین همبستگی با یکدیگر استفاده می‌شود. حداقل افزونگی با رابطه ۲۱ محاسبه می‌شود [۱۶].

$$\min W_I(S), \quad W_I = \frac{1}{|S|^2} \sum_{i, j \in S} I(i, j) \quad (21)$$

W_I میانگین MI بین ویژگی‌ها در زیرمجموعه S است. با ترکیب رابطه‌های ۲۰ و ۲۱، مفهومی با عنوان کمینه افزونگی بیشینه وابستگی^۹ (mRMR) معرفی شده است. اساس mRMR با رابطه ۲۲ تعریف می‌شود [۱۶].

$$\max \varphi(V, W), \quad \varphi = V - W \quad (22)$$

$\varphi(V, W)$ عملگری برای ترکیب روابط بیشینه وابستگی و کمینه افزونگی است.

با افزایش تعداد اعضای مجموعه ویژگی‌های مدل، تعداد زیرمجموعه‌ها به صورت نمایی افزایش می‌یابد. در این حالت، محاسبه‌های لازم برای یافتن حداکثر مقدار $\varphi(V, W)$ ، غیرعملی خواهد بود؛ به همین دلیل، روش‌هایی برای تخمین mRMR معرفی شده‌اند که پیچیدگی محاسباتی بسیار کمتری نسبت به آن دارند. از متداول‌ترین روش‌های تخمین mRMR، دو روش تفاضل اطلاعات متقابل^{۱۰} (MID) و خارج قسمت اطلاعات متقابل^{۱۱} (MIQ) است [۱۷]. MID و MIQ به ترتیب با رابطه‌های ۲۳ و ۲۴ محاسبه می‌شوند.

دو، سه و چهار برنامه به صورت تصادفی انتخاب شده است که یکی از برنامه‌ها به هسته مدنظر و باقی برنامه‌ها به صورت تصادفی به دیگر هسته‌های پردازنده تخصیص داده می‌شوند. در هنگام اجرای این الگوریتم، در فواصل زمانی یک‌ثانیه‌ای حالت سیستم ثبت می‌شود. در هر مرحله، مقادیر سرعت فن، با توجه به لیست آنها و به صورت صعودی افزایش می‌یابند. برای تغییرات فرکانس، ابتدا لیست مقادیر فرکانس با استفاده از تابع *Shuffle* به هم ریخته می‌شود. این امر باعث می‌شود حالات متنوعی در میزان تغییرات فرکانس در پردازنده ایجاد شود. بعد از هر تغییر فرکانس پردازنده، به مدت ۱۰ ثانیه، تغییری در میزان فرکانس و سرعت فن ایجاد نمی‌شود.

۴-۲- آماده‌سازی مجموعه داده

تغییر دما به رویدادی وابسته است که موجب می‌شود به تدریج و با فاصله‌ای زمانی صورت گیرد. این امر سبب شده است ترتیب نمونه‌ها در مجموعه داده جمع‌آوری شده، مدل معنی‌دار مشخصی داشته باشند. برای بهبود عملکرد مدل پیش‌بینی دما، می‌توان بر وابستگی بین نمونه‌های متوالی تکیه کرد. در این راستا تلاش شده است ویژگی‌های جدیدی از مقادیر آماری اندازه‌گیری شده ایجاد شود تا شامل تاریخچه حالت‌های سیستم برای نمونه باشد.

نخستین بخش از پارامترهای پیشنهادی، پارامترهای سابقه‌ای برای هر نمونه است. این پارامترها، سابقه تغییرهای سخت‌افزاری و نرم‌افزاری را برای هر نمونه نگهداری

```

ψ ← {fan speed levels}
χ ← {processor frequency levels};
Start system profiling;
for η = 1 to #cores do:
begin
Dispatch a benchmark for each of η cores;
for i = 1 to |ψ| do:
begin
Set_fan(ψi);
Shuffle(χ);
for j = 1 to |χ| do:
begin
Set_frequency(χj);
Wait(10 seconds);
end
end
Terminate all dispatched benchmarks;
end
Stop system profiling;

```

شکل (۱): شبه‌کد ثبت کردن حالت سیستم

دمایی هستند، به صورت موازی روی هسته‌های پردازنده استفاده می‌شود.

به سبب اینکه مدل پیش‌بینی دمای پیشنهادی مبتنی بر مشاهده عملکرد سیستم است، برای افزایش دقت مدل، باید تنوع بیشتری از حالت‌های سیستم را ایجاد کرد. در این راستا، برای ایجاد حالت‌های متنوع در سیستم، از ترکیب‌های مختلف سرعت فن، فرکانس پردازنده و بارکاری، آزمایشی انجام شده است. در شکل (۱) الگوریتم استفاده شده در این آزمایش ارائه می‌شود.

در الگوریتم ارائه شده، در ابتدا سطوح مختلف سرعت فن و فرکانس پردازنده انتخاب می‌شوند. در آزمایش انجام شده، با توجه به محدوده سرعت فن، هفت گام شامل مقادیر {۵۰، ۸۴، ۱۱۰، ۱۴۰، ۱۸۴، ۲۲۴، ۲۵۵} مدولاسیون پهنای پالس^{۱۳} (PWM) برای سرعت فن انتخاب می‌شوند. برای تغییرات فرکانس پردازنده نیز از هفت گام شامل مقادیر {۲۶۶۰۰۰، ۲۹۲۶۰۰۰، ۳۰۶۰۰۰۰} گیگاهرتز بهره برده شده است. سپس، یک نخ^{۱۴} فراخوانی شده که به صورت هم‌زمان با الگوریتم اصلی اجرا می‌شود، از آن، برای خواندن و ثبت حالت سیستم استفاده می‌شود. حالت سیستم شامل دمای هسته‌های پردازنده، فرکانس پردازنده، سرعت فن و شمارنده‌های کارایی هسته مد نظر برای پیش‌بینی دما است. در الگوریتم، از حلقه‌ای استفاده شده است که برابر با تعداد هسته‌های پردازنده اجرا می‌شود. در مرحله نخست، برنامه محک به صورت تصادفی، انتخاب و روی هسته مدنظر اجرا می‌شود. در مرحله دوم، سوم و چهارم به ترتیب

جدول (۱): لیست شمارنده‌های کارایی

CONTEXT-SWITCHES
CPU-MIGRATIONS
PERF_COUNT_SW_PAGE_FAULTS
CYCLES
PERF_COUNT_STALLED_CYCLES_FRONTEND
PERF_COUNT_STALLED_CYCLES_BACKEND
INSTRUCTIONS
BRANCHES
BRANCH-MISSES
CACHE-REFERENCES
IO_TRANSACTIONS
L1D_ALL_REF
CACHE-MISSES

است. خروجی مدل، پیش‌بینی دما را با فاصله زمانی $d \geq 1$ انجام می‌دهد.

در این مقاله، برای بررسی دقت مدل پیشنهادی، از معیارهایی مبتنی بر ارزیابی خطا، همچون میانگین قدر مطلق خطا^{۱۵} (MAE)، انحراف معیار قدر مطلق خطا^{۱۶} (SDAE) و احتمال خطای بیش از ۳ درجه سانتی‌گراد^{۱۷} (POE) استفاده می‌شود که به ترتیب با رابطه‌های ۲۷، ۲۸ و ۲۹ محاسبه می‌شود.

$$MAE = \frac{1}{N} \sum_{i=1}^N |t_i - y_i| \quad (27)$$

$$SDAE = \sqrt{\frac{1}{N} \sum_{i=1}^N (|t_i - y_i| - MAE)^2} \quad (28)$$

$$POE = \frac{|\{t_i - y_i \geq 3 \mid \forall t_i \in S, \forall y_i \in Y, i \in \{1, 2, \dots\}\}|}{|S|} \quad (29)$$

N تعداد نمونه‌ها، t_i دمای هدف و y_i دمای پیش‌بینی شده برای نمونه i ام است. S مجموعه مورد آزمون و Y مجموعه پیش‌بینی‌های مدل است.

درخور ذکر است اگر دمای پیش‌بینی شده پردازنده، بیش از مقدار درست آن باشد، روش‌های مدیریت دمای استفاده شده به کاهش بیش از نیاز دمای سیستم منجر می‌شوند. بیشتر روش‌های کاهش دما، به کاهش عملکرد سیستم منجر می‌شوند. از طرفی، در صورتی که دمای پیش‌بینی شده کمتر از مقداری باشد که در سیستم ایجاد می‌شود، احتمال افزایش دما از دمای حد آستانه پردازنده وجود دارد. هرچه دقت مدل کمتر باشد، احتمال ایجاد نقاط داغ در پردازنده و خرابی آن افزایش می‌یابد. در این راستا، برای محاسبه دقت مدل دمایی، از پارامترهای MAE و SDAE استفاده می‌شود. همچنین، از پارامتر POE برای محاسبه احتمال خطاهای بالای ۳ درجه سانتی‌گراد بهره گرفته شده است. از این پارامتر برای محاسبه میزان خطاهای بالای مدل پیش‌بینی دما استفاده می‌شود. این خطاها باعث آسیب جدی برای پردازنده می‌شوند.

می‌کند. از جمله پارامترهای سابقه‌ای، مقادیر دمای هسته‌ها و شمارنده‌های کارایی چند نمونه قبلی هر نمونه است. همچنین، با استفاده از گرادیان، تعدادی دیگر از پارامترهای سابقه‌ای ایجاد شده است. گرادیان پارامتر p با فاصله d با رابطه ۲۵ محاسبه می‌شود.

$$\nabla p_t(d) = \frac{p_t - p_{t-d}}{d} \quad (25)$$

p_t مقدار پارامتر در نمونه t و p_{t-d} مقدار همان پارامتر در d نمونه قبلی است. برای هر نمونه تعدادی از گرادیان‌های اول و تعدادی از گرادیان‌های دوم تهیه شده است. گرادیان اول، میانگین سرعت تغییرات در طی d نمونه متوالی است؛ در حالی که گرادیان دوم، میانگین شتاب تغییرات در همان بازه است.

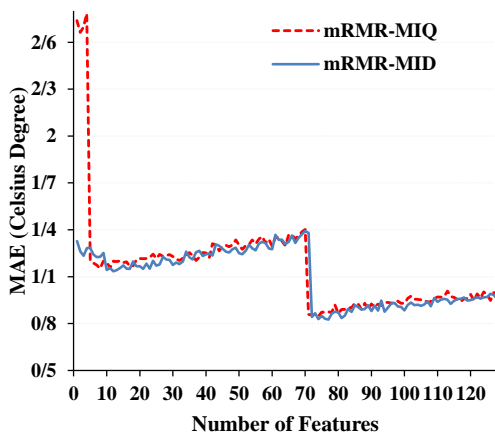
همچنین پارامترهایی را با نام پارامترهای کنترلی معرفی کرده است که به صورت مستقیم زیر نظر سیستم عامل هستند. بیشتر، مقدار پارامترهای کنترلی زمانی تغییر می‌کند که زمان بند سیستم عامل، واکنشی مانند تغییر فرکانس انجام می‌دهد. برای محاسبه پارامترهای کنترلی، مفهوم گرادیان‌های آینده را پیشنهاد می‌کند که با رابطه ۲۶ محاسبه می‌شود.

$$\nabla p_t = p_{t+1} - p_t \quad (26)$$

برای پارامتر کنترلی p ، مقدار پارامتر در نمونه t و p_{t+1} مقدار همان پارامتر در نمونه بعدی است. در این مقاله میزان تغییر فرکانس پردازنده و سرعت فن به عنوان پارامترهای کنترلی به مدل پیش‌بینی دما داده می‌شود.

۴-۳- توسعه و ارزیابی مدل

مدل پیش‌بینی دمای پیشنهادی، برای هسته در سیستم چند هسته‌ای ارزیابی می‌شود. هسته مدنظر با نام هسته هدف معرفی می‌شود. ورودی‌های مدل شامل مقادیر پارامترهای سابقه‌ای دمای هسته، دمای هسته‌های مجاور، فرکانس پردازنده، سرعت فن خنک‌کننده و مقدار شمارنده‌های کارایی و گرادیان آنها در فاصله‌های زمانی $\{1, 2, 3\}$ ثانیه است. همچنین، از پارامترهای کنترلی تغییر فرکانس پردازنده و سرعت فن نیز به عنوان ورودی‌های مدل، بهره گرفته شده



شکل (۲): اثر افزایش تعداد ویژگی‌های انتخابی با استفاده از روش‌های MID و MIQ در دقت مدل پیش‌بینی دما

در شکل (۲)، ویژگی که باعث کاهش محسوس مقدار خطا می‌شود، میزان تغییر فرکانس است. به سبب اینکه در آزمایش‌ها بازه زمانی تغییر فرکانس ۱۰ ثانیه است، تعداد نمونه‌های کمتری شامل اثر تغییر فرکانس بر دما هستند. بر اساس این، تغییر فرکانس با اینکه اهمیت زیادی در پیش‌بینی دمای تعداد زیادی از نمونه‌ها دارد، همبستگی بیشتری با مقدار دما ندارد. درخور ذکر است تغییر فرکانس به سرعت باعث تغییر رفتار دمایی پردازنده می‌شود و در صورت توجه نکردن به آن، پیش‌بینی دما با خطای زیادی مواجه می‌شود. با بررسی مجموعه داده جمع‌آوری شده، به جز در موارد مربوط به تغییر فرکانس، تغییر رفتار دمایی به سرعت اتفاق نمی‌افتد. در این راستا، داده‌های فراهم شده برای پیش‌بینی دمای ۲ ثانیه آینده به دو مجموعه داده تقسیم شده است. بخش نخست شامل نمونه‌هایی است که در آن میزان تغییر فرکانس و به صورت کلی، گرادبان آینده صفر باشد و بخش دوم شامل نمونه‌های غیر صفر است. برای داده‌های بخش نخست، حداکثر تغییرات دما ۷ درجه سانتی‌گراد است؛ در حالی که در بخش دوم ۲۱ درجه سانتی‌گراد است. از طرفی واریانس تغییرات دما برای بخش اول و دوم به ترتیب حدود ۱.۵ و ۴.۴ درجه سانتی‌گراد است. این تفاوت رفتاری دما برای نمونه‌های این دو بخش، دلیلی برای ارائه مدل ترکیبی برای پیش‌بینی دما شد. مدل پیش‌بینی دمای پیشنهادی از دو مدل SVR، برای این دو حالت تشکیل شده است. با توجه به شکل (۳)، برای نمونه‌هایی که

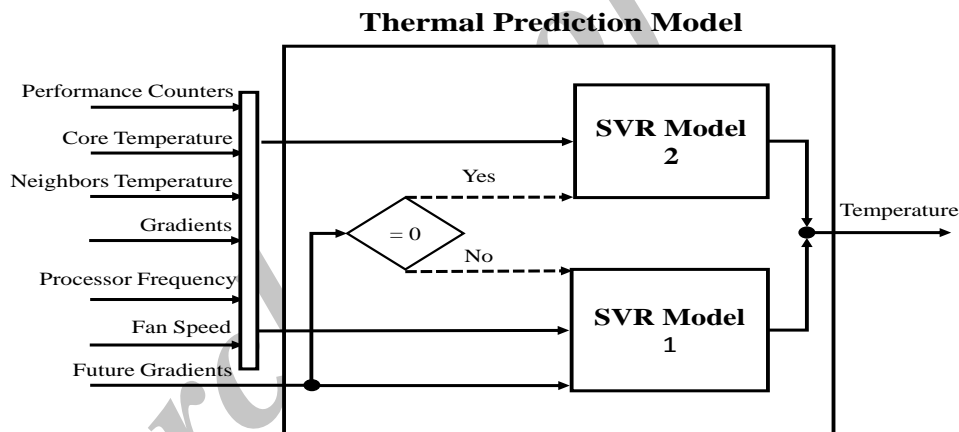
برای توسعه مدل، ابتدا به مقایسه دو حالت پرداخته شده است. در یک حالت، تنها از پارامترهای نشان‌دهنده حالت فعلی سیستم برای پیش‌بینی دما استفاده می‌شود. در حالت دوم، علاوه بر استفاده از پارامترهای حالت فعلی سیستم، از پارامترهای پیشنهادی در این مقاله استفاده می‌شود که شامل پارامترهای سابقه‌ای و کنترلی هستند. مجموع پارامترها در این حالت ۱۲۸ عدد است. نتیجه مقایسه در جدول (۲) نمایش داده شده است. استفاده از پارامترهای پیشنهادی، سبب کاهش مقدار MAE مدل پیش‌بینی دما از ۲/۲ به ۰/۹۷ درجه سانتی‌گراد می‌شود. با این وجود، زمان اجرای مدل برای پیش‌بینی نمونه حدود ۱۰ برابر افزایش می‌یابد.

اما به سبب اینکه ممکن است نیازی به استفاده از همه این ۱۲۸ پارامتر نباشد، تأثیر پارامترهای مختلف بر دقت مدل دمایی بررسی می‌شود. در این راستا با الگوریتم‌های MID و MIQ، ویژگی‌های موجود در مجموعه پارامترهای ورودی، در بردار $\vec{\gamma}$ بر اساس میزان وابستگی داده‌ای با مقدار هدف مدل‌ها اولویت‌بندی می‌شود. بر اساس این، در $|\zeta|$ گام، زیرمجموعه‌هایی شامل $1 \leq j \leq |\zeta|$ عنصر ابتدایی از $\vec{\gamma}$ را با مدل SVR، ارزیابی و میزان خطای پیش‌بینی برای هر یک اندازه‌گیری می‌شود. نتیجه در شکل (۲) نمایش داده شده است. با توجه به شکل، به استفاده از ۱۲۸ پارامتر برای مدل پیش‌بینی دما نیاز نیست و با انتخاب حدود ۷۲ ویژگی با MID و MIQ حداقل خطای مدل به دست می‌آید. از آنجا که بعد با افزایش تعداد ویژگی‌ها، خطا در مدل افزایش می‌یابد. از طرفی کاهش تعداد پارامترهای ورودی، باعث کاهش چشمگیر زمان اجرای مدل می‌شود. درخور ذکر است در آزمایش‌های انجام شده، زمان اجرای مدل SVR برای پیش‌بینی دما برای نمونه، در حالتی که از ۷۲ پارامتر بهره می‌برد، نسبت به ۱۲۸ پارامتر بیش از ۵ برابر کاهش می‌یابد.

دو مدل SVR، ابتدا با افزایش تعداد ویژگی‌ها مقدار خطای مدل‌ها کاهش می‌یابد و از نقطه‌ای به بعد افزایش تعداد ویژگی‌ها به افزایش خطا منجر می‌شود. براساس نتیجه به دست آمده، برای رسیدن به بهترین دقت مدل‌ها، باید تعدادی از ویژگی‌های پیشنهادی را به‌عنوان ورودی انتخاب کرد. با توجه به شکل (۴)، با انتخاب ۱۱ ویژگی برای هر دو مدل می‌توان به دقت مناسب مدل پیش‌بینی دمای ۲ ثانیه آینده دست یافت. در این حالت، میانگین خطای MAE مدل، ۰٫۵۶، درجه سانتی‌گراد است. در جدول (۲) مقایسه استفاده از پارامترهای مختلف ورودی برای مدل پیش‌بینی دمای ۲ ثانیه آینده ارائه شده است. مدل پیشنهادی علاوه بر کاهش خطای درخور توجه مدل، زمان اجرای کمتری نسبت به دیگر رویکردها دارد.

در آنها تغییری در فرکانس و سرعت فن نسبت به نقطه پیش‌بینی هدف ایجاد می‌شود از مدل SVR دوم استفاده می‌شود. از طرفی، برای دیگر نمونه‌ها، مدل SVR اول استفاده می‌شود. برای بررسی رویکرد ارائه‌شده، نخست مجموعه داده جمع‌آوری شده، با توجه به نمونه‌ها به دو بخش تقسیم می‌شوند. یک بخش، مناسب برای آموزش مدل SVR اول و بخش دیگر، مناسب برای آموزش مدل SVR دوم است. برای هر یک از این مدل‌ها، با استفاده از مجموعه داده مربوطه، انتخاب پارامترهای مناسب با استفاده از الگوریتم MID انجام می‌شود. در شکل (۴)، اثر افزایش تعداد ویژگی‌ها در مقدار MAE مدل پیش‌بینی دما با فاصله دو ثانیه، برای هر دو مدل SVR نمایش داده شده است. در هر

شکل (۳): مدل پیش‌بینی دمای پیشنهادی



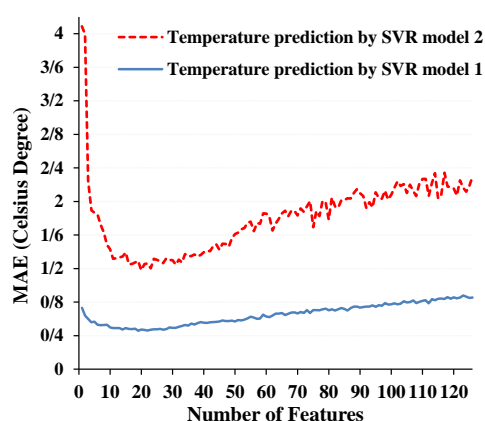
جدول (۲): پیش‌بینی دمای ۲ ثانیه آینده با پارامترهای مختلف

نوع ورودی‌های مدل پیش‌بینی دما	تعداد ورودی‌ها	MAE	SDAE	POE(%)	زمان اجرای نمونه (S)
حالت فعلی سیستم	۱۹	۲/۲	۲/۹	۸/۲۴	$۸/۶۵۶۸ \times ۱۰^{-۵}$
حالت فعلی سیستم + پارامترهای سابقه‌ای و کنترلی	۱۲۸	۰/۹۷	۱/۳۷	۴	$۷/۷۸۵۲ \times ۱۰^{-۴}$
پارامترهای انتخابی برای مدل پیشنهادی	۱۱	۰/۵۶	۰/۶۷	۰/۵۴	$۲/۶۰۲۷ \times ۱۰^{-۵}$

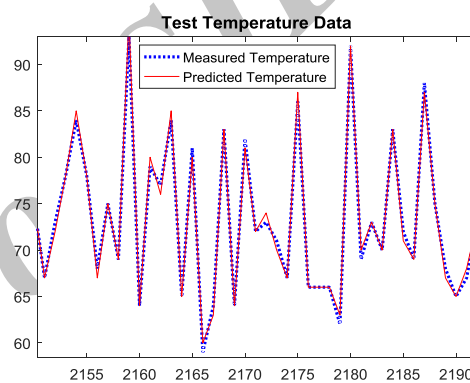
۵- ارزیابی مدل پیشنهادی

مدل پیشنهادی برای پیش‌بینی دمای هسته، برای فاصله‌های زمانی ۲ تا ۵ ثانیه آینده، ارزیابی شده است. برای هر فاصله زمانی، مدل پیش‌بینی مجزا فراهم می‌شود. برای هر مدل SVR استفاده‌شده در این مدل‌های پیش‌بینی، مجموعه داده مناسب برای آموزش فراهم شده و ویژگی‌های مناسب انتخاب می‌شوند. ویژگی‌های انتخابی برای هر یک از مدل‌های SVR مربوط به فاصله‌های مختلف زمانی تفاوت زیادی با یکدیگر دارند. در جدول (۳) برای هر یک از مدل‌های پیش‌بینی، تعداد ویژگی‌های مناسب به همراه مقدار MAE، SDAE و POE مربوطه نمایش داده شده است.

بعد از انتخاب ویژگی‌های مناسب و به‌عبارت‌دیگر، ورودی‌های مناسب، ارزیابی مدل‌های پیش‌بینی دما انجام شده است. باید توجه داشت داده‌های آزمایش به‌صورت به‌هم‌ریخته به مدل‌ها داده می‌شوند تا سنجش دقیق‌تری از کارایی مدل‌ها انجام شود. در شکل (۵) برای نمایش دقت مدل پیش‌بینی دمای ۲ ثانیه آینده، مقدار پیش‌بینی‌شده و مشاهده‌شده بخشی از داده‌ها ارائه شده است. نتایج حاکی از دقت زیاد مدل پیشنهادی است.



شکل (۴): اثر تعداد ویژگی‌ها در دقت مدل‌های SVR استفاده‌شده در مدل پیش‌بینی دمای ۲ ثانیه آینده



شکل (۵): مقایسه دمای پیش‌بینی و مشاهده‌شده ۲ ثانیه آینده برای تعدادی از نمونه‌های آزمایشی

جدول (۳): بهترین نتیجه به‌دست‌آمده از تعداد ویژگی‌های انتخابی برای مدل‌های دمایی برای فاصله‌های زمانی مختلف

مدل	فاصله زمانی	تعداد ویژگی‌ها	MAE (°C)	SDAE (°C)	POE (%)
SVR model 1	۲	۱۱	۰/۴۷	۰/۶	۰/۱
SVR model 1	۳	۱۱	۰/۵۳	۰/۶۵	۰/۱۱
SVR model 1	۴	۱۰	۰/۵۷	۰/۶۷	۰/۱۳
SVR model 1	۵	۱۰	۰/۶۲	۰/۷۱	۰/۱۶
SVR model 2	۲	۱۱	۱/۲	۱/۲	۳/۱۱
SVR model 2	۳	۱۰	۱/۲۹	۱/۲۸	۴/۱
SVR model 2	۴	۱۲	۱/۳۷	۱/۴۸	۵/۱
SVR model 2	۵	۱۱	۱/۴۳	۱/۳۸	۵/۵

دما در جدول (۴) نمایش داده شده است. از آنجایی که میزان تغییر فرکانس در پردازنده، کمتر از زمان‌هایی است که تغییر فرکانس صورت نمی‌گیرد، میانگین خطای کلی مدل نزدیک به خطای مدل اول SVR است. درخور ذکر است خطای کمتر از ۰/۸ درجه سانتی‌گراد در پیش‌بینی دما برای

با توجه به نتایج نمایش داده‌شده در جدول (۳)، با انتخاب حدود ۱۱ ویژگی برای هر یک از مدل‌های SVR موجود در مدل‌های پیش‌بینی دما، می‌توان به دقت بیشتری در پیش‌بینی دما برای فاصله‌های مختلف زمانی دست یافت. نتایج کلی استفاده از مدل‌های SVR در قالب مدل پیش‌بینی

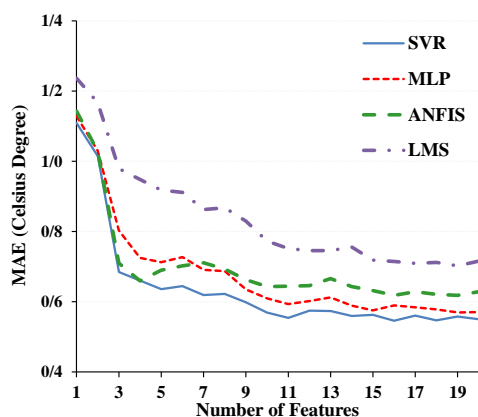
پیشنهادی نسبت به کچران است. از طرفی، زمان پیش‌بینی دمای ۲ ثانیه آینده برای هر نمونه در روش کچران $10 \times 151 \times 7$ است. زمان اجرای مدل کچران، ۲٫۵ برابر بیشتر از مدل پیشنهادی روی مجموعه داده یکسان است. در ادامه، مقایسه‌ای بین مدل پیشنهادی با مدل استاکمن از مدل SVR برای پیش‌بینی دما استفاده کرده است. ورودی‌های مدل ایشان، تنها شامل شماره‌های کارایی است. نتیجه استفاده از مدل ایشان روی مجموعه داده یکسان با مدل پیشنهادی نیز در جدول (۴) نمایش داده شده است. مدل ارائه شده از استاکمن، به دلیل انتخاب نشدن ویژگی‌های مناسب برای مدل دمایی، خطای زیادی دارد.

فاصله‌های زمانی ۲ تا ۵ ثانیه، نشان از دقت زیاد مدل پیشنهادی دارد. گفتنی است برای پیش‌بینی دمای ۲ ثانیه آینده، در ۹۶ درصد مواقع، حداکثر خطای یک درجه سانتی‌گراد دیده می‌شود. همچنین برای پیش‌بینی ۳، ۴ و ۵ ثانیه آینده، به ترتیب ۹۴، ۸۸ و ۸۴ درصد داده‌ها دارای حداکثر خطای یک درجه سانتی‌گراد هستند.

در ادامه ارزیابی، مدل پیشنهادی با مدل پیش‌بینی دمای ارائه شده کچران مقایسه می‌شود. مدل ایشان در بخش ۲ معرفی شده است [۱۱]. همچنین در جدول (۴)، نتایج حاصل از پیش‌بینی دما با استفاده از رویکرد ارائه شده کچران برای فاصله‌های مختلف زمانی نشان داده شده‌اند. مقایسه نتایج این دو رویکرد، نشان‌دهنده دقت بسیار بهتر رویکرد

جدول (۴): معیارهای ارزیابی خطای مدل‌های پیش‌بینی دمای پیشنهادی و کچران برای فاصله‌های مختلف زمانی

مدل	فاصله زمانی (S)	MAE (°C)	SDAE (°C)	POE (%)
پیشنهادی	۲	۰/۵۶	۰/۶۷	۰/۵۴
کچران	۲	۱/۴	۲/۲	۸
استاکمن	۲	۳/۷	۳/۱	۴۰
پیشنهادی	۳	۰/۶۴	۰/۷۲	۰/۹
کچران	۳	۱/۷	۲/۵	۱۱
استاکمن	۳	۳/۹	۳/۲۱	۴۱
پیشنهادی	۴	۰/۷	۰/۷۷	۱/۰۶
کچران	۴	۲/۱	۲/۸	۱۵
استاکمن	۴	۴/۰۱	۳/۳۴	۴۱
پیشنهادی	۵	۰/۷۶	۰/۸۲	۱/۲
کچران	۵	۲/۶	۳	۲۰
استاکمن	۵	۴/۳	۳/۴	۴۴



شکل (۶): مقایسه دقت استفاده از مدل‌های MLP، SVR، ANFIS و LMS در مدل دمایی پیشنهادی برای پیش‌بینی دمای ۲ ثانیه آینده

از دیگر آزمایش‌های انجام شده، مقایسه مدل‌های SVR، MLP، سیستم استنتاج عصبی-فازی و فقی^{۱۸} (ANFIS) و الگوریتم حداقل میانگین مربع^{۱۹} (LMS) است. از LMS برای یافتن ضریب‌های فیلتر خطی مناسب برای مدل پیش‌بینی دما استفاده می‌شود. به بیانی دیگر، LMS، نماینده‌ای از مدل‌های خطی مناسب برای پیش‌بینی دما، در این مقایسه شرکت دارد. نتیجه استفاده از مدل‌های مذکور با ویژگی‌های ورودی یکسان برای پیش‌بینی دمای ۲ ثانیه آینده در شکل (۶) ارائه شده است. در این مقایسه نیز، SVR نتیجه بهتری نسبت به دیگر مدل‌ها ارائه می‌دهد.

- approach for dynamic thermal management", Proceedings of the 43rd annual Design Automation Conference, pp. 548-553, 2006.
- [6] M. Zaman, A. Ahmadi, and Y. Makris, "Workload characterization and prediction: A pathway to reliable multi-core systems", On-Line Testing Symposium (IOLTS), pp. 116-121, 2015.
- [7] G. Yang, Q. Qiu, and Q. Wu, "A multi-agent framework for thermal aware task migration in manycore systems", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 20, No. 10, pp. 1758-1771, 2012.
- [8] M. Stockman, M. Awad, H. Akkary, R. Khanna, "Thermal status and workload prediction using support vector regression", Energy Aware Computing, International Conference, pp. 1-5, 2012.
- [9] P. Kumar, D. Atienza, "Neural network based on-chip thermal simulator", Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium, pp. 1599-1602, 2010.
- [10] A. Vincenzi, A. Sridhar, M. Ruggiero, D. Atienza, "Fast thermal simulation of 2D/3D integrated circuits exploiting neural networks and GPUs", Proceedings of the 17th IEEE/ACM international symposium on low-power electronics and design, pp. 151-156, 2011.
- [11] R. Cochran, and S. Reda, "Thermal prediction and adaptive control through workload phase detection", ACM Transactions on Design Automation of Electronic Systems (TODAES), Vol. 18, No. 1, pp. 7-26, 2013.
- [12] Z. Liu, T. Xu, S. X. D Tan, H. Wang, "Dynamic thermal management for multi-core micro-processors considering transient thermal effects", Design Automation Conference (ASP-DAC), pp. 473-478, 2013.
- [13] K. J. Lee, K. Skadron, "Using performance counters for runtime temperature sensing in high-performance processors," Parallel and Distributed Processing Symposium, 2005.
- [14] S. J. Lu, R. Tessier, W. Burleson, "Dynamic on-chip thermal sensor calibration using performance counters", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 33, No. 6, pp. 853-866, 2014.
- [15] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, D. Tarjan, "Temperature-aware microarchitecture: Extended discussion and results," University of Virginia, Dept. of Computer Science, Technical Report, 2003.
- [16] H. Peng, F. Long, C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and minredundancy", IEEE Transactions on pattern analysis and machine intelligence, Vol. 27, No. 8, pp. 1226-1238, 2005.
- [17] M. Mandal, A. Mukhopadhyay, "An improved minimum redundancy maximum relevance approach for feature selection in gene expression data", Procedia Technology 10, pp. 20-27, 2013.
- [18] N. Cristianini, J. Shawe-Taylor, "An introduction to support vector machines and other kernel-based learning methods", Cambridge university press, 2000.

۶- نتیجه گیری

در این مقاله، مدل دمایی شامل دو مدل SVR برای پیش‌بینی دما ارائه شده است. برای هر مدل SVR، ویژگی‌ها و مجموعه داده مناسب برای آموزش و آزمایش آن ایجاد شده است. تعدادی از ویژگی‌های انتخابی، حالت فعلی سیستم را نشان می‌دهد و شامل دمای هسته‌ها، فرکانس پردازنده، سرعت فن و شمارنده‌های کارایی هستند. علاوه بر این، برای هر نمونه در مجموعه داده از چند مقدار پیشین این ویژگی‌ها و گرادین‌های اول و دوم آنها تا ۳ ثانیه قبل نیز به عنوان دیگر ویژگی‌های مدل پیش‌بینی دما استفاده می‌شود. از طرفی، مقدار تغییر فرکانس پردازنده و سرعت فن با عنوان گرادین‌های آینده به مدل داده می‌شوند. به دلیل تعداد زیاد این ویژگی‌ها، برای افزایش عملکرد مدل پیش‌بینی دما، تعدادی از بهترین ویژگی‌ها با استفاده از الگوریتم mRMR-MID، برای هر مدل SVR انتخاب می‌شوند. از معیارهای میانگین قدرمطلق خطا، انحراف معیار قدرمطلق خطا و درصد خطای پیش‌بینی بالای ۳ درجه سانتی‌گراد، در بررسی دقت مدل پیش‌بینی دمای پیشنهادی استفاده شده است. مدل پیشنهادی برای پیش‌بینی دما برای فاصله‌های زمانی ۲ تا ۵ ثانیه آینده ارزیابی شده است. نتایج به‌دست‌آمده نشان‌دهنده دقت زیاد مدل پیش‌بینی دمای پیشنهادی هستند.

مراجع

- [1] J. Kong, S. W. Chung, K. Skadron, "Recent thermal management techniques for micro-processors", ACM Computing Surveys (CSUR), Vol. 44, No. 3, 42 pages, 2012.
- [2] A. K. Coskun, T. S. Rosing, and K. C. Gross, "Proactive temperature balancing for low cost thermal management in MPSoCs", Computer-Aided Design, IEEE/ACM International Conference, pp. 250-257, 2008.
- [3] A. K. Coskun, T. S. Rosing, and K. C. Gross, "Utilizing predictors for efficient thermal management in multiprocessor SoCs", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 28, No. 10, pp. 1503-1516, 2009.
- [4] G. Liu, M. Fan, and G. Quan, "Neighbor-aware dynamic thermal management for multi-core platform", Proceedings of the Conference on Design, Automation and Test in Europe, pp. 187-192, 2012.
- [5] A. Kumar, L. Shang, L. S. Peh, and N. K. Jha, "HybDTM: a coordinated hardware-software

- [21] F. Li, D. Miao, W. Pedrycz, "Granular multi-label feature selection based on mutual information", *Pattern Recognition*, Vol. 67, No. 1, pp. 410-423, 2017.
- [22] N. Hoque, H. A. Ahmed, D. K. Bhattacharyya, J. K. Kalita, "A Fuzzy mutual Information-based feature selection method for classification", *Fuzzy Information and Engineering*, Vol. 8, No. 3, pp. 355-384, 2016.
- [19] R.K. Jain, K. M. Smith, P. J. Culligan, J.E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy", *Applied Energy*, Vol. 123, No. 15, pp. 168-178, 2014.
- [20] L. Ge, J. Shi, P. Zhu, "Melt index prediction by support vector regression", In *Control, Automation and Information Sciences (ICCAIS)*, pp. 60-63, 2016.

-
- ¹ Dynamic Thermal Management
² Reactive
³ Proactive
⁴ Support Vector Regression
⁵ AutoRegressive–Moving-Average
⁶ Principal Component Analysis
⁷ Multi-Layer Perceptron
⁸ Mutual Information
⁹ Minimum Redundancy Maximum Relevance
¹⁰ Mutual Information Difference
¹¹ Mutual Information Quotient
¹² Heat sink
¹³ Pulse Width Modulation
¹⁴ Thread
¹⁵ Mean Absolute Errors
¹⁶ Standard Deviation of Absolute Errors
¹⁷ Percentage of Outliers Errors
¹⁸ Adaptive Neuro-Fuzzy Inference System
¹⁹ Least Mean Square

Archive of SID