

文章编号: 2095-2163(2023)11-0161-05

中图分类号: TP311.1

文献标志码: A

基于 SVR 数据中心空调系统瞬态热参数预测

黄金森, 朱兵, 张一鸣, 殷佳辉, 苗益川

(贵州大学 电气工程学院, 贵阳 550025)

摘要: 数据中心空调系统是维持数据中心的关键设备,直接影响到数据中心的安全运行,目前对空调系统的研究大多集中在节能降耗以及气流优化等领域。在空调故障影响等瞬态变化领域的研究仍然较少。因此有必要探究空调系统故障对机房气流组织的影响,建立针对空调失效极端工况下的快速温度预测模型,为能效控制系统及运行系统提供参考。本文根据空调系统故障实验分别建立了空调冷冻水泵失效及风机失效情况下的关键位置的温度变化时间序列预测模型,模型基于线性核函数支持向量回归机。研究表明相较于非线性核函数支持向量机,线性核函数支持向量机更适合进行冷冻水泵失效时的热参数预测。

关键词: 数据中心; 机器学习; 热参数预测; 数据驱动

Transient thermal parameter prediction based on SVR data center air conditioning system

HUANG Jinsen, ZHU Bing, ZHANG Yiming, YIN Jiahui, Miao Yichuan

(School of Electrical Engineering, Guizhou University, Guiyang 550025, China)

Abstract: The data center air conditioning system is a key equipment to maintain the data center, which directly affects the safe operation of the data center. Most of the current research on air conditioning system is focused on energy saving and consumption reduction as well as airflow optimization. Fewer studies have emerged in the area of transient changes such as the impact of air conditioning failures. Therefore, it is necessary to explore the impact of air conditioning system failure on the airflow organization of the server room and establish a fast temperature prediction model for the extreme operating conditions of air conditioning failure to provide a reference for energy-efficient control systems and operation systems. In this paper, time series prediction models of temperature changes at critical locations under air conditioner chilled water pump failure and fan failure are established based on linear kernel function support vector regression machine. The research demonstrates that the linear kernel function support vector machine is more suitable for the prediction of thermal parameters in case of chilled water pump failure than the nonlinear kernel function support vector machine.

Key words: data center; machine learning; thermal parameter prediction; data driven

0 引言

截至 2021 年,中国数据中心机架总规模超过 520 万,在用服务器规模达到 1 900 万台,预计未来国内数据中心装机容量将继续保持快速增长^[1]。庞大的装机容量使得数据中心行业用电量飞速增长,预计到 2025 年国内数据中心用电量将达到 3 950 亿千瓦时,占全社会用电量的 5.8%^[2]。Ni 等学者^[3]调查了 100 个数据中心能耗情况,暖通空调系统能耗占总能耗平均值为 38%,其中,在调查的

能耗最低占比为 21%,最高占比为 61%,可见数据中心空调系统节能潜力巨大。仅通过改造机房结构进行气流组织优化的传统措施已不能满节能需求,在此基础上开发动态优化冷源供应控制系统是目前行业的迫切需要,而建立动态优化供冷控制系统首要任务是实现数据中心气流组织热参数的快速预测。Athavale 等学者^[4]评估了人工神经网络(ANN)、高斯过程回归(GPR)和支持向量回归(SVR)3 种机器学习方法对数据中心稳态工况下机架入口温度分布的预测性能,研究表明三种方法所

基金项目: 贵州省科技支撑计划项目(2017YFB0902100)。

作者简介: 黄金森(1998-),男,硕士研究生,主要研究方向:建筑节能、能源系统优化。

通讯作者: 朱兵(1967-),女,高级工程师,主要研究方向:分布式能源系统、建筑节能技术。Email:452179224@qq.com

收稿日期: 2023-06-10

得到的稳态模型具有相当的准确性,对于冷却故障情况下的瞬态预测,核函数的选择决定了 SVR 和 GPR 所开发模型的外推能力。本文使用带线性核函数的支持向量回归机对空调故障情况下热参数进行预测。

1 实验数据分析

空调系统冷冻水泵故障是数据中心日常运行中常见的故障情况之一,在实验室 RL^[5] 中进行空调冷冻水泵故障实验和风机故障实验。实验过程中,在每个模拟机箱的入口设置了 9 个测温点,10 块高架地板每块上方设置一个测温点,空调回风口设置 6 个测温点,空调出风口设置 4 个测温点,每个测温点布置 1 个热电偶测温,取每个位置所有热电偶的平均值,实验温度测量的不确定度约为 0.5~1 K。

1.1 冷冻水泵失效实验

冷冻水泵失效实验研究了冷冻水泵失效后 6 min 内高架地板入口,空调回风口、空调进风口、每个模拟机箱入口的温度变化,实验中为了避免实验设备损坏,在 R1、R2、R3 负荷均为 20% 的条件下进行实验,在服务器及空调稳定运行过程中关闭,冷冻水回水阀,从关闭时开始约 60 s 冷冻水回水阀完全关闭,冷冻水泵停止运行时间总计 6 min。在此期间,测量了空调回风温度、空调出风温度、高架地板入口温度等参数。机房及静压箱内的空气循环如图 1 所示。

高架地板的出风温度曲线与空调出风温度曲线出现交叉,在 40 s 左右空调出风温度超过高架地板出风温度,空调出风温度上升速度高于高架地板出风口的温度,这是因为静压箱内混凝土底板、侧墙等建筑材料比热容比空气大,在静压箱内空气温度迅速升高时起到了冷却作用。模拟机箱 R1C4 进风温度变化趋势与空调回风温度变化趋势大致相同,在 0~50 s 区间内温度变化较小,这是因为冷源失效的初始时刻静压箱内储备了一定的冷空气,因此仍可以维持一定时间的制冷效果;50 s 后基本呈线性增长趋势,此时由于静压箱内储备的冷空气逐渐耗尽,在服务器的加热作用下,机房内空气循环的平均温度逐渐升高。

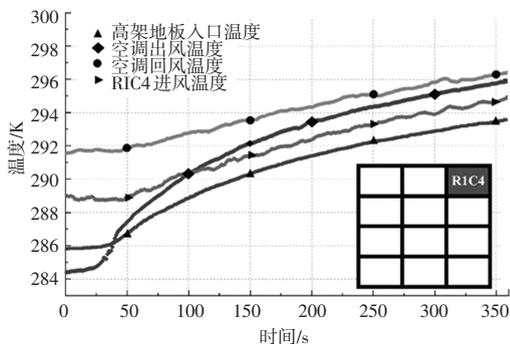


图 2 水泵失效期间机房部分位置温度变化图

Fig. 2 Temperature variation of some parts of the equipment room during water pump failure

1.2 空调风机失效实验

空调系统冷风扇故障同样是数据中心日常运行中常见的故障情况之一,为了避免实验设备损坏,风机失效实验在 R1、R2、R3 负荷均在 30% 的条件下进行,在服务器及空调稳定运行过程中关闭空调风机停止运行时间总计 6 min。在此期间没有新风进入机房,模拟机箱入口循环吸入自身排气。实验测量了 R3C4、R2C4、R1C4、R1C3 共 4 个模拟机箱的入口温度变化,4 个模拟机箱的入口温度变化情况相差不大。图 3 展示了在空调风机失效的 6 min 内模拟机箱 R3C4 和 R1C4 的入口温度变化情况。在风机失效后的 360 s 内,模拟机箱 R1C4 与 R3C4 的变化趋势大致相同:在风机失效后的 0~180 s 区间内,机柜入口温度约上升了 12 K;在 180~360 s 区间内,机柜入口温度约上升了 5 K。温度变化呈现出先快后慢的趋势,这是由于风机失效后,空调出风流量逐渐减小,静压箱与机房之间的压差逐渐消失,冷风无法穿过高架地板进入机房。此时由于冷却不足,模拟机箱不断循环吸入自身排气,使得机柜内部温度不断升高,入口温度迅速上升。与此同时,机柜入口空

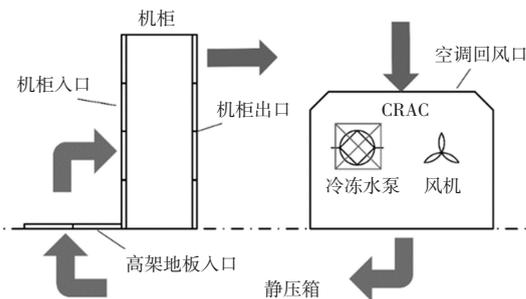


图 1 机房及静压箱内的空气循环示意图

Fig. 1 Schematic diagram of air circulation in the equipment room and plenum

在冷冻水泵失效的 6 min 内空调回风温度、空调出风温度、高架地板入口温度与模拟机箱 R1C4 的入口温度变化情况如图 2 所示。冷冻水泵停止运行意味着从机房进入空调的热风无法被循环水冷却,只能与滞留在换热器内的冷冻水进行热交换,由于此时冷冻水处于静止状态,换热系数较低,因此空调的时间常数比正常运行情况下更长、约为 70 s。由于风机正常运行,所以机房内流场未发生较大变化,空调的出风温度逐渐上升,最终接近于空调回风温

气流速降低导致边界层厚度增加, 湍流程度降低, 从而使空气对流换热系数降低, 随着温度的上升空气与服务器之间的温差降低, 温度上升速度减小。

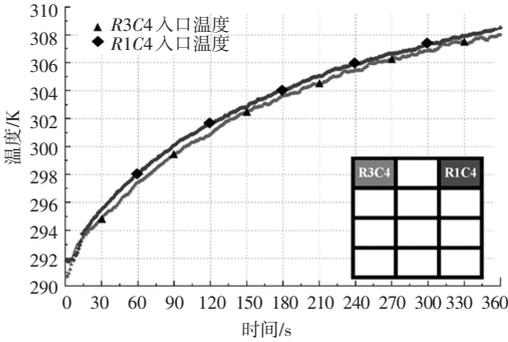


图 3 风机失效期间 R3C4 及 R1C4 入口温度变化图

Fig. 3 Temperature variation of R3C4 and R1C4 inlet during fan failure

2 时间序列预测简介

时间序列预测原理是根据某个变量的历史变化情况对该变量在未来某段时间或某个时刻的变化情况做出预测。时间序列预测不局限于对未来值的预测, 还包括异常检测、时间序列分类等领域。时间序列也叫时间数列、历史复数或动态数列, 是在某一时间区间内按顺序记录下的相同指标的数据集合, 具有可比性。常见的时间序列预测模型有时间分解、自回归 (AR) 模型、移动平均 (MA) 模型、自回归滑动平均 (ARMA) 模型、差分自回归移动平均 (ARIMA) 模型、带输入差分自回归平均移动 (ARIMAX) 模型等, 然而由于 AR、MA、ARMA 模型特性, 三者只适合对平稳时间序列进行线性拟合, 当处理的时间序列呈现上升或下降的趋势时预测效果不尽人意。因此 ARIMA 模型应运而生, 该模型通过对不平稳的时间序列进行差分或对数化运算, 将其转化为平稳时间序列后对其进行自回归滑动平均分析^[6]。ARIMAX 则是在 ARIMA 的基础上实现了基于输入变量时间序列预测。

随着计算机技术的飞速发展推动了机器学习和深度学习等建模方法的迅速崛起, 相较于传统时间序列预测模型, 机器学习模型非线性拟合能力更强, 可以对超大维度及变化复杂的时间序列数据进行处理。支持向量回归机在处理小样本数据时相较于其他算法更有优势, 且空调冷冻水泵停机后温度变化接近线性, 因此与文献[4]使用的非线性核函数支持向量机不同, 本文使用基于线性核函数的支持向量回归机对空调系统失效时机房内温度变化情况进行预

测, 实验样本数量为 360 组, 是空调失效 360 s 内关键位置的温度变化情况, 每秒记录一次。在空调失效的 360 s 内, 前 180 s 的数据用于模型训练, 后 180 s 的数据进行外推预测。时间步长设置为 15 s, 单时间步预测, 即使用某点前 15 个数据对其进行预测, 对实验数据进行滑窗处理, 将训练集输入整理为 15 个数据为一组, 时间序列的实现过程见图 4。

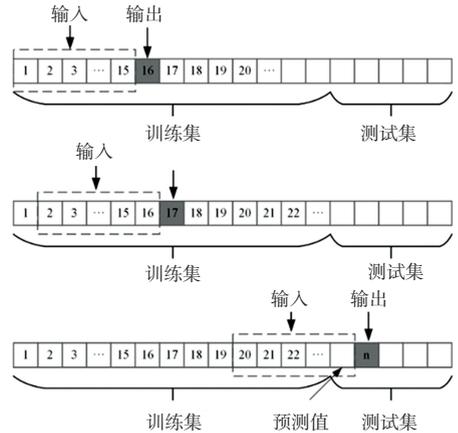


图 4 时间序列外推预测示意图

Fig. 4 Schematic diagram of time series extrapolation prediction

3 基于 SVR 的时间序列预测

3.1 冷冻水泵失效参数预测

如前文所述, 前 180 s 的数据用于训练模型, 后 180 s 的数据用于测试外推精度, 检验支持向量回归模型对时间序列的预测精度。使用带线性核函数的支持向量回归机对空调回风温度及 R1C4 模拟机箱入口温度进行预测。

3.1.1 空调回风温度预测

冷冻水泵停机后, 空调回风温度实验值与预测值对比结果如图 5 所示。预测值与观测值在 360 个观测点的误差分布如图 6 所示。

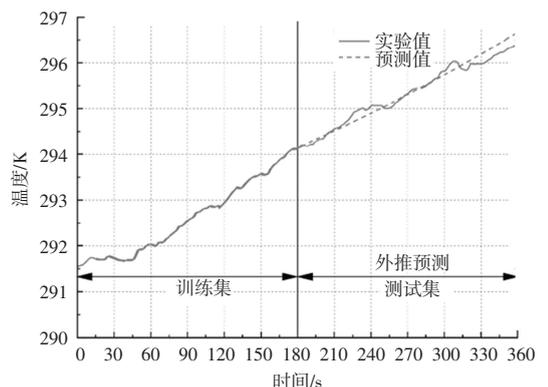


图 5 空调回风温度时间序列预测结果

Fig. 5 Time series prediction results of air conditioner return air temperature

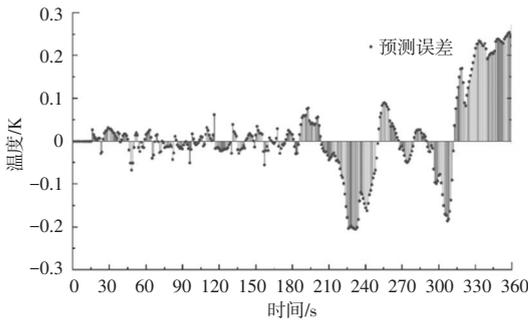


图6 空调回风温度时间序列预测误差

Fig. 6 Time series prediction error of air conditioner return air temperature

从图5和图6中可以明显看到,在16~180s的

表1 空调回风温度预测评价指标

Tab. 1 Evaluation index of air conditioner return air temperature prediction

数据集	ME	MAE	MSE	RMSE	MAPE	R	R ²
Train	0.067 46	0.016 59	0.000 41	0.020 26	0.006	0.999 66	0.999 33
Test	0.254 48	0.045 38	0.003 93	0.062 67	0.035	0.995 47	0.990 96

3.1.2 R1C4 模拟机箱入口温度预测

冷冻水泵停机后模拟机箱 R1C4 入口温度实验值与预测值对比结果如图7所示。预测值与观测值在360个观测点的误差分布如图8所示。

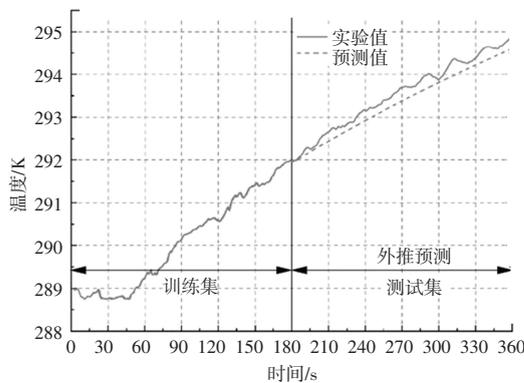


图7 模拟机箱 R1C4 入口温度时间序列预测结果

Fig. 7 Time series prediction results of inlet temperature of simulated chassis R1C4

区间内(由于时间序列的滑动窗口长度(时间步长)为15,因此前15s的数据不作为参考),时间序列训练集吻合效果较好,最大绝对误差小于0.1K。在180~360s区间内,预测结果可以准确反映温度变化的总体趋势,但线性模型无法反映实际测量结果的轻微波动,空调回风温度预测模型评价指标见表1。

从评价指标看,外推预测的最大绝对预测误差为0.255K,均方根误差为0.0627K,均远小于空调回风温度的测量不确定度1K。模型训练时间约为1s,可以认为线性核函数支持向量机在180~360s外推区间内实现了对空调回风温度的快速准确预测。

从图7和图8中可以明显看到,在前180s训练集内支持向量回归机预测值与观测值总体拟合良好,误差随着曲线的波动小范围内变化。在180~360s区间内,预测值均小于实验值,外推预测误差随曲线波动变化,平均误差逐渐增大。从变化趋势看,在360s之后的预测值精度将难以保证,模拟机箱 R1C4 入口温度预测模型评价指标见表2。

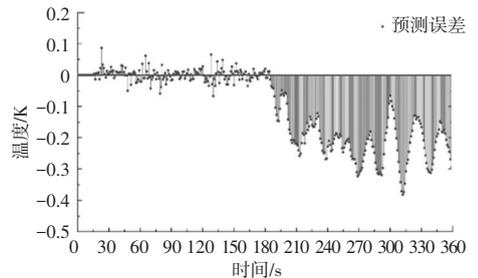


图8 模拟机箱 R1C4 入口温度时间序列预测误差

Fig. 8 Prediction error of inlet temperature time series of simulated chassis R1C4

表2 模拟机箱 R1C4 入口温度预测评价指标

Tab. 2 Evaluation indexes of inlet temperature prediction of simulated chassis R1C4

数据集	ME	MAE	MSE	RMSE	MAPE	R	R ²
Train	0.086 21	0.015 29	0.000 43	0.020 81	0.005	0.999 81	0.999 62
Test	0.381 57	0.203 17	0.047 48	0.217 89	0.069	0.963 21	0.927 77

从评价指标来看测试集的各项误差均大于训练集,测试集均方根误差为0.218K,最大误差为0.382K,均小于实验测量的不确定度,且模型训练时间小于1s,因此可认为实现了较为准确的快速温度预测。测试集相关性相较于训练集明显下降,这主要是实验值曲线波动较大造成的。在气流组织变化较快的位置,预测模型准确性会有所降低。

根据本文中空调回风温度预测模型、模拟机箱 R1C4 入口温度预测模型的评价指标对比文献[4]中研究成果,可以发现与非线性核函数支持向量机相比,线性核函数支持向量机更适用于冷冻水泵失效时的热参数预测,因为在冷冻水泵失效后的短期内机柜入口温度变化趋势及空调回风温度变化趋势接近线性。

3.2 风机失效参数预测

冷冻水泵停机后空调回风温度实验值与预测值对比结果如图 9 所示。预测值与观测值在 360 个观测点的误差分布如图 10 所示。

从图 9 和图 10 中可以明显看到,在前 180 s 训练集内支持向量回归机预测值与观测值总体拟合良好,误差随着曲线的波动小范围内变化。在 180~310 s 外推预测区间实验值并非严格线性变化,并且实验数据有一定的噪声,造成预测误差较大。预测值在 180~345 s 内预测值均低估了实验值,随着实际温度上升速度下降,预测值与实验值出现相交,可以看出在 360 s 之后预测值将不断增长,预测结果将不再可信,模拟机箱 R1C4 入口温度预测模型评价指标见表 3。

从评价指标来看,测试集的误差远大于训练集,最大误差达到 0.862 K,均方根误差为 0.496 K,且测试集的相关性不高, R^2 约为 0.85。此外,由于模型的惩罚因子较大,训练时间约为 6 s。这是因为当风机失效后,机柜入口温度变化趋势与冷冻水泵失效后不同,温度上升速度先快后慢,机柜入口温度变化趋势呈非线性,因此线性核函数支持向量机的预测结果不够理想。

表 3 模拟机箱 R3C4 入口温度预测评价指标

Tab. 3 Evaluation indexes of inlet temperature prediction of simulated chassis R3C4

数据集	ME	MAE	MSE	RMSE	MAPE/%	R	R^2
Train	0.100 94	0.027 56	0.001 14	0.033 86	0.010	0.999 93	0.999 86
Test	0.862 07	0.435 20	0.246 48	0.496 47	0.142	0.921 79	0.849 69

4 结束语

根据 Erden^[5] 在 RL 实验室进行的空调系统故障实验分别建立了空调冷冻水泵失效时的空调回风温度预测模型与模拟机箱 R1C4 入口温度预测模型及风机失效情况下模拟机箱 R3C4 入口温度预测模型。3 个模型均基于线性核函数支持向量回归机,冷冻水泵失效时的空调回风温度预测模型和 R1C4 入口温度预测模型,在 180~360 s 区间内均实现了较为准确的快速外推预测,均方根误差分别为 0.063 K、0.218 K,但风机失效时的 R3C4 入口温度预测模型表现不够理想,外推预测均方根误差为 0.049 6 K,但最大预测误差达到 0.862 K,其原因是风机失效后的机柜入口温度上升速度逐渐减小,变化趋势呈非线性特征,如果有关于温度变化趋势的先验知识,通过选择一个线性增长的核函数来提高基于 SVR 的模型的外推能力。研究表明了相较于

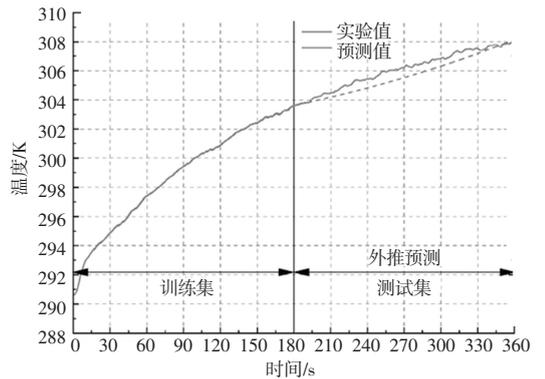


图 9 模拟机箱 R3C4 入口温度时间序列预测误差

Fig. 9 Prediction error of temperature time series at the entrance of the simulated chassis R3C4

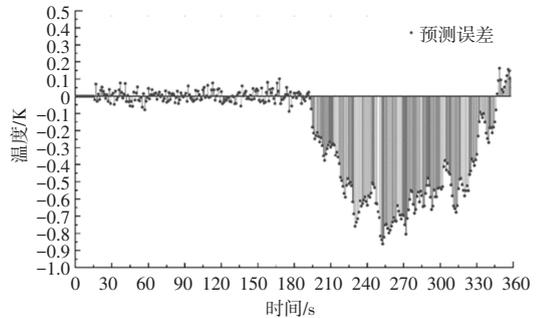


图 10 模拟机箱 R3C4 入口温度时间序列预测误差

Fig. 10 Prediction error of temperature time series at the entrance of the simulated chassis R3C4

非线性核函数支持向量机,线性核函数支持向量机更适合进行冷冻水泵失效时的热参数预测。

参考文献

- [1] 中国信息通信研究院. 中国算力发展指数白皮书 [R]. 北京:中国信息通信研究院, 2022.
- [2] CY315. 2019-2025 年中国数据中心用电需求规模及占全社会用电量比值预测 [EB/OL]. (2019-12-6). [2021-11-01]. <https://www.chyxx.com/industry/201912/818362.html>.
- [3] NI Jiacheng, BAI Xuelian. A review of air conditioning energy performance in data centers[J]. Renewable and Sustainable Energy Reviews, 2017, 67: 625-640.
- [4] ATHAVALLE J, YODA M, JOSHI Y. Comparison of data driven modeling approaches for temperature prediction in data centers[J]. International Journal of Heat and Mass Transfer, 2019, 135: 1039-1052.
- [5] ERDEN H S. Experimental and analytical investigation of the transient thermal response of air cooled data centers [D]. New York: Syracuse University, 2013.
- [6] 许锐. 基于 EEMD 的金融时间序列多尺度分析 [D]. 合肥: 中国科学技术大学, 2016.