

基于云计算平台实现电网短期负荷预测算法的研究

王 帅, 赵建平, 王志远, 谢 广
(国网乌鲁木齐供电公司, 新疆 乌鲁木齐 830011)

摘 要: 短期电网负荷预测在任何电力系统的运行决策中都起着重要作用。由于负荷短期预测的实时性要求与在大型数据集上执行复杂的计算过程所需的高计算能力相冲突, 提出了基于支持向量回归的短期电网负荷预测算法, 并基于云计算平台实现了该算法。将所提预测算法分别在云计算平台和单机计算平台上进行对比实验, 结果表明基于云计算平台的实现有效提高了算法执行效率。

关键词: 负荷预测; SVR; 云计算; Azure ML

中图分类号: TM769 文献标志码: A 文章编号: 1003-6954(2019)01-0029-04

DOI: 10.16527/j.cnki.cn51-1315/tm.2019.01.006

Research on Short-term Load Forecasting Algorithm Based on Cloud Computing Platform

Wang Shuai, Zhao Jianping, Wang Zhiyuan, Xie Guang
(State Grid Urumqi Power Supply Company, Urumqi 830011, Xinjiang, China)

Abstract: Short-term grid load forecasting plays an important role in the operational decision for any power system. Because the real-time requirements for short-term load forecasting conflict with the high computing capability which is required to perform complex processes on large data sets, a short-term grid load forecasting algorithm based on support vector regression is proposed, which is implemented based on cloud computing platform. The proposed forecasting algorithm is compared between the cloud computing platform and the stand-alone computing platform, and the results show that the implementation based on cloud computing platform effectively improves the execution efficiency of the algorithm.

Key words: load forecasting; SVR; cloud computing; Azure ML

0 引 言

电力工业的根本目标是通过科学的规划最大限度地提高发电和用电效率, 提高电力系统运行的经济性。而电力负荷预测一直是电力系统有效规划和经济运行的重要组成部分。

短期电网负荷预测是电力系统运行决策的关键, 对电力系统的机组组合、减少发电备用、提高经济调度、保持系统可靠性以及维护调度都有重要意义^[1-2]。近年来, 准确的短期电网负荷预测在电力市场管制中得到了更大的重视和面临更大的挑战。

短期电网负荷预测算法需要处理从一个小时到几天的负载预测。由于短期负载具有极大的随机性, 因此短期电网负荷预测的复杂性对电力运营者

是一个极大挑战。在过去的几十年里, 人们提出了许多预测模型。这些预测模型可以分为传统模型或基于机器学习的模型。前者包括时间序列预测器, 如自回归移动平均外生变量模型^[3]。这些传统模型是基于线性回归模型, 不能准确代表复杂负载的非线性特征^[4]。不同的机器学习技术也被用于短期电网负荷预测, 如人工神经网络(artificial neural network, ANN)^[5]、径向基函数(radial basis function, RBF)^[6]、模糊神经模型^[7]和支持向量回归^[8]。

预测过程依赖于对某一国家或地区电力需求历史数据的分析, 还可以考虑许多其他因素, 如天气预报和商业计划。因此, 需要整个历史数据来训练预测模型, 但是这样的方法其缺点是, 如果考虑到新的信息则所有参数都可能需要重新训练。此外, 这种庞大的数据量和预期的复杂预测过程导致需要大量

的计算能力。研究人员试图找到近似方法来最小化这一数据量,并将所需的计算能力降到最低。这些研究试图尽量减少数据采样量等。许多方法解决了这些回溯问题,其中一个采用了中的局部预测器^[9-10]。

经济约束在任何算法研究中都起着主要作用。云计算的出现解决了研究人员和开发人员面临的许多经济问题。在云计算技术之前,超级计算机是获得巨大计算能力的唯一合适选择,显然这是一个非常昂贵的选择。有许多计算系统可以提供巨大的计算能力,如分布式系统、网格计算、互联网计算以及量子计算等,但云计算技术是最具性价比的选择,获得了广泛的商业应用。为此,一些研究试图借助云计算技术为短期负荷预测提供足够的计算能力^[11-13]。

1 基于支持向量回归的短期负荷预测算法

一般来说,电力负荷由不同的消耗单元组成。各种因素都影响着电力负荷的变化,如天气、重大事件、经济因素和随机因素。短期负荷预测可以被认为是一个多变量预测问题。它可以作为回归问题的函数来求解。次日负荷为回归模型的输出,历史负荷数据及其影响因素为回归模型的输入。历史数据库提供训练数据。该回归问题的最终目标是从历史负荷数据及其影响因素中找到一个具有良好泛化能力的预测负荷映射函数。历史负荷数据被分为两个不同的数据集:一个是训练数据集用来训练回归模型;另一个是测试数据集用来评估训练后的回归模型^[14]。

基于统计学习理论提出的支持向量回归(support vector regression, SVR)^[15]已被研究作为一种有前景的电力负荷预测方法。其优势主要来自于采用结构风险最小化原理,并作为经验风险最小化原则的替代方案,它可以通过求解二次问题来获得最优的全局解。

SVR的执行有两个主要特征:二次规划和核函数。二次规划问题将用线性约束求解得到SVR的参数。核函数的灵活性使该技术能够搜索宽范围的解空间。SVR的主要目标是通过非线性映射将数据x映射到高维特征空间,并在该特征空间中执行线性回归^[16-17]:

$$f(x) = \langle w, x \rangle + b \tag{1}$$

式中: $\langle \cdot, \cdot \rangle$ 表示点积; w 包含必须从数据中估计的系数; b 是一个实常数。使用 Vapnik 的 ϵ -不敏感损失函数^[18] 将整体优化为

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^N (\xi_i + \xi_i^*) \tag{2}$$

约束条件为

$$\begin{cases} y_i - (w^T \phi(x_i) + b) \leq \epsilon + \xi_i^* \\ (w^T \phi(x_i) + b) - y_i \leq \epsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0 \quad i = 1, \dots, N \end{cases} \tag{3}$$

式中: x_i 通过函数 ϕ 映射到更高维空间; ϵ 是实常数; ξ_i 和 ξ_i^* 是受 ϵ 不敏感区域影响的松弛变量; C 为正 regularization 参数,其确定 $f(x)$ 的平坦度和训练误差之间的折衷。

引入拉格朗日乘数 α_i 和 α_i^* , 这两个参数满足 $\alpha_i \alpha_i^* = 0$ 和 $\alpha_i, \alpha_i^* = 0 (i = 1, \dots, N)$ 。并且根据 Karush-Kuhn-Tucker 最优性条件^[18], SVR 训练过程相当于求解凸二次曲线问题:

$$\min_{\alpha, \alpha^*} \frac{1}{2} \sum_{i,j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) + \epsilon \sum_{i=1}^N (\alpha_i - \alpha_i^*) - \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) \tag{4}$$

约束条件为

$$\begin{cases} 0 \leq \alpha_i, \alpha_i^* \leq C \\ \sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0 \quad i = 1, \dots, N \end{cases} \tag{5}$$

输出是一个独特的全局优化结果,其形式如下:

$$\hat{y} = \hat{f}(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) Q(x_i, x) + b \tag{6}$$

式中, $Q(x_i, x)$ 是核函数。在 SVR 中采用核函数,所有必要的计算可以直接在输入空间中计算。核函数存在各种各样的内核,如线性、双曲正切、高斯径向基函数、多项式等^[19]。在这里,使用常用的 RBF 内核:

$$Q(x_i, x) = e^{-\gamma \|x_i - x\|^2} \tag{7}$$

SVR 的参数 C, γ 和 ϵ 在 SVR 的性能中起着至关重要的作用。因此,选择这些参数的正确值可以最大限度地减少预测误差。基于核的 SVRs 需要计算数据集中每个点之间的距离函数

$$O(n_{\text{features}} * n_{\text{observations}}^2) \tag{8}$$

基于 SVR 解决短期电网负荷预测问题可以归纳为以下步骤:

- 1) 加载历史数据并将其分为训练集和验证集;
- 2) 准确确定 SVR 的参数;

- 3) 使用定义的参数训练 SVR 以获得支持向量和相应的系数;
- 4) 利用式(6) 得到预测的负载。

2 实验验证

为检验云计算平台在电力负荷预测领域的影响,实验旨在测试两点:一是使用 Azure ML 实现负荷预测技术的准确性;二是测试执行时间的改进。

2.1 数据集

该数据集收集了从 2016 年 1 月至 2017 年 12 月的 0.5 h 电力负荷、2016 年至 2017 年的日平均温度以及 2016 年至 2017 年的假期信息。目标是预测 2018 年 1 月的每日最大负荷,并用 2018 年 1 月的负荷实际值验证计算预测值。

2.2 实验平台

设计实验的实现有两个选择:本地实现和基于云的实现。对于本地实现,可使用台式计算机及使用 MATLAB 软件实现所提出的算法。台式计算机具有以下规格: Microsoft Windows 10, Intel Core i7 2.7 GHz, RAM 16 GB。对于基于云的实现,则用 Azure ML^[20]。

2.3 性能指标

实验考虑了两个主要的性能指标:第 1 个也是最重要的一个是执行时间($T_{Execution}$);第 2 个是预测准确性。所有实验将使用 4 个度量来评估预测准确度:平均绝对百分比误差(MAPE)、最大误差幅度(MAX)、平均绝对误差(MAE)和归一化均方误差(NMSE)。这些值由式(9)至式(13)定义:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|A_i - P_i|}{A_i} \times 100\% \quad (9)$$

$$MAX = \max(|A_i - P_i|) \quad (10)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |A_i - P_i| \quad (11)$$

$$\Delta^2 = \frac{1}{N-1} \sum_{i=1}^N (A_i - \bar{A})^2 \quad (12)$$

$$NMSE = \frac{1}{\Delta^2} \sum_{i=1}^N (A_i - P_i)^2 \quad (13)$$

式中 A_i 、 P_i 、 \bar{A} 和 N 分别为实际值、预测值、实际值的平均值和测试数据集大小。

2.4 结果与讨论

实验包括了 4 种不同的实验方案,用以比较和评估所提出的基于 SVR 算法的短期电网负荷模型

在云计算平台和单机计算平台上的性能差异。

所有这些实验方案的目标是使用训练集训练上面所述的预测模型,再使用这个预测模型来预测 2018 年 1 月的 31 天内的最大日负荷,并与实际负荷进行比较。所提出的各实验方案的训练数据有所不同,如表 1 所示。

表 1 实验方案的数据集

实验方案	训练数据集		测试数据	
	数据设定标准	数量	数据设定标准	数量
1	2016 至 2017 年的每日负荷	723	2018 年 1 月实际负荷值	31
2	2016 至 2017 年的每日负荷和温度	723	2018 年 1 月实际负荷值	31
3	2016 至 2017 年的冬季(10、11、12、1、2、3 月)的每日负荷	357	2018 年 1 月实际负荷值	31
4	2016 至 2017 年的冬季(10、11、12、1、2、3 月)的负荷和温度	357	2018 年 1 月实际负荷值	31

4 种实验方案的执行结果如表 2、图 1 和图 2 所示。结果表明,使用云计算平台的执行时间远远少于使用基于本地机器的单机计算平台的执行时间。特别是在训练数据相对较大的在实验方案 2 中,云计算平台与单机计算平台的执行时间相比,时间执行的改进在 10 倍以上。值得指出的是,在云计算平台上进行的第 2 次实验时比第 1 次运行更快,例如,实验方案 2 的第 2 次执行时间等于 9 s,远低于第 1 次的 72 s。此外,从预测精度来看,两者几乎相同,在实验方案 4 的情况下实现了最佳预测精度为 2.04%。

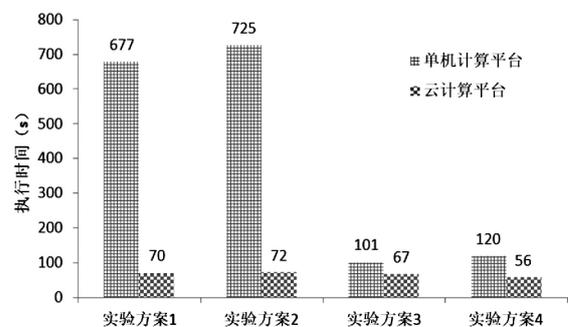


图 1 单机平台和云平台执行时间对比

3 结论

前面提出了一种基于支持向量回归预测(SVR)

表2 实验结果

实验方案	$T_{Execution} /s$		MAPE %	MAE	MAX	NMSE	SVR 参数	
	单机	云					γ	C
1	677	70	2.21	16.45	42.8	0.346	0.5	1
2	725	72	2.41	17.65	60.6	0.421	0.25	2
3	101	67	2.08	15.55	36.35	0.295	0.25	2
4	120	56	2.04	15.2	40.4	0.297	0.125	2

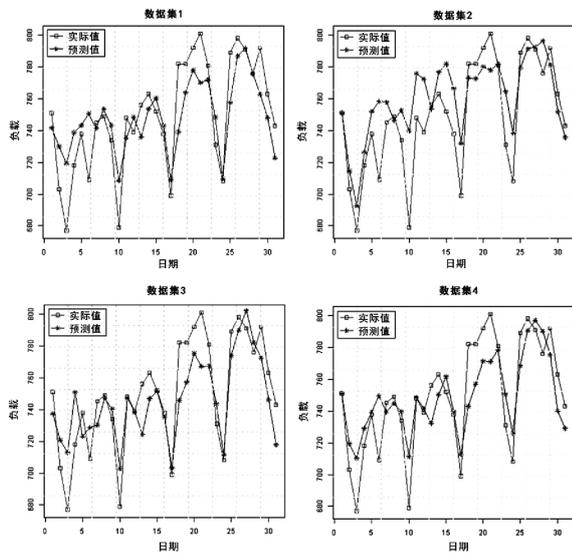


图2 各个实验方案下预测值和实际值的偏差

算法短期电网负荷模型,并通过实验分别在云计算平台和单机计算平台上实现了该模型。在两个平台的对比实验中,选择执行时间和预测精度作为性能指标。对于所使用的实验数据集,单机计算平台和云计算平台实现的算法预测精度是相同的;但是在云计算平台实现的算法执行时间显著减小。一般来说,SVR并不推荐用于大型数据集,因为它的计算成本很高,而前面采用云计算平台体现了较好的计算效率,因此所进行的实验结果可为相关研究提供借鉴。

参考文献

[1] 王雁凌,吴梦凯.经济新常态下基于偏最小二乘回归的中长期负荷预测模型[J].电力自动化设备,2018,38(3):133-139.
 [2] 王萍,李磊,胡聪,等.云计算环境下智能电网短期负荷预测方法[J].科学技术与工程,2018,18(7):153-158.
 [3] 苏小林,刘孝杰,阎晓霞,等.计及需求响应的主动配电网短期负荷预测[J].电力系统自动化,2018,42(10):60-66.
 [4] 赵芝璞,高超,沈艳霞,等.基于关联模糊神经网络和

改进型蜂群算法的负荷预测方法[J].中国电力,2018,51(2):54-60.
 [5] 吴润泽,包正睿,宋雪莹,等.基于深度学习的电网短期负荷预测方法研究[J].现代电力,2018,35(2):43-48.
 [6] 林涛,董栅,秦冬阳,等.基于支持向量回归的风电场短期功率预测[J].中南民族大学学报(自然科学版),2017,36(4):95-99.
 [7] 吴润泽,包正睿,王文韬,等.Hadoop架构下基于模式匹配的短期电力负荷预测方法[J].电工技术学报,2018,33(7):1542-1551.
 [8] 张智晟,于道林.考虑需求响应综合影响因素的RBF-NN短期负荷预测模型[J].中国电机工程学报,2018,38(6):1631-1638.
 [9] 韩中合,李秋菊,苑一鸣,等.基于相关向量机的短期风速预测模型[J].电力科学与技术学报,2017,32(3):38-42.
 [10] 汤强,谢明中,罗元盛.基于SVR的用电负荷特征三维回归模型[J].计算机工程,2017,43(9):300-303.
 [11] 李滨,黄佳,吴茵,等.基于分形特性修正气象相似日的节假日短期负荷预测方法[J].电网技术,2017,41(6):1949-1955.
 [12] 肖白,刘庆永,房龙江,等.基于模糊粗糙集理论和时空信息的空间负荷预测[J].电力建设,2017,38(1):58-67.
 [13] 苏学能,刘天琪,曹鸿谦,等.基于Hadoop架构的多重分布式BP神经网络的短期负荷预测方法[J].中国电机工程学报,2017,37(17):4966-4973.
 [14] 钱志.基于改进型SVR的电网短期负荷预测[J].中国电力,2016,49(8):54-58.
 [15] 马天男,牛东晓,黄雅莉,等.基于Spark平台和多变量L2-Boosting回归模型的分布式能源系统短期负荷预测[J].电网技术,2016,40(6):1642-1649.
 [16] 孟安波,胡函武,刘向东.基于纵横交叉算法优化神经网络的负荷预测模型[J].电力系统保护与控制,2016,44(7):102-106.
 [17] 罗滇生,王新坤.基于云计算的智能电网负荷预测平台架构研究[J].湖南大学学报(自然科学版),2016,43(2):101-108.

(下转第56页)

其他参数保持不变,仅将模型的硬质管母线导热率由 $200 \text{ W}/(\text{m} \cdot \text{K})$ 增加到 $250 \text{ W}/(\text{m} \cdot \text{K})$ 即可。改变后硬质管母线在电流 800 A 、螺栓力矩 $50 \text{ N} \cdot \text{m}$ 时的温度分布如图 10 所示。

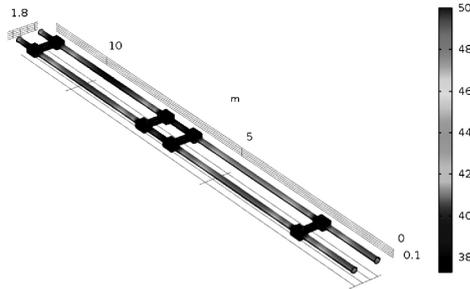


图 10 增加热通量后硬质管母线温度分布

比较热通量改变前后的温度变化可以发现:提升硬质管母线本身的散热性能,即增大其表面导热率,可以加快硬质管母线与空气的热交换,温升由 32.3 K 降低到 20.2 K ,下降了 12.1 K 。说明选取合适的铝合金制品提升硬质管母线的散热能力,可以有效缓解硬质管母线的发热问题。

4 结 论

1) 通过对硬质管母线进行的大电流试验以及仿真结果的分析,其过热的主要原因是由于接触点处铝管接触不紧密,使硬质管母线自身产生的热量远大于其与空气进行热交换时的散热量。

2) 通过仿真分析可以发现减小温升主要有两种途径:一是通过打磨以及涂抹电力脂等方法增加管母线连接处光滑程度以减小接触电阻,减少发热量;二是在铝制品中增加一定量金属元素做成合金,增加硬质管母线本身的导热率,从而减缓温度的上升。

参考文献

[1] 包健康,高波,张血琴. 500kV 输电线路耐张线夹电

(上接第 32 页)

[18] 王宁,谢敏,邓佳梁,等. 基于支持向量机回归组合模型的中长期降温负荷预测[J]. 电力系统保护与控制, 2016, 44(3): 92-97.

[19] 沈沉,秦建,盛万兴,等. 基于小波聚类的配变短期负荷预测方法研究[J]. 电网技术, 2016, 40(2): 521-526.

接触计算及其发热研究[J]. 中国电力, 2016, 49(11): 25-30.

[2] 丁永福,王祖力,张燕乘,等. $\pm 800 \text{ kV}$ 特高压直流换流站阀厅金具的结构特点[J]. 高压电器, 2013, 49(9): 13-18.

[3] 徐元哲,王乐天,刘雪冬,等. 电力电缆接头测温系统的设计[J]. 高电压技术, 2009, 35(12): 2977-2982.

[4] 胡艳. 弓网系统在电滑动过程中材料转移特性的试验研究[D]. 成都: 西南交通大学, 2016.

[5] 韩国栋. 焊缝缺陷对管母线导电性能和发热的影响研究[D]. 北京: 华北电力大学, 2015.

[6] 韩国栋,吴章勤,万书亭,等. 管母线损耗发热状况多物理场分析计算[J]. 科学技术与工程, 2014, 14(28): 214-219.

[7] 陈森昌,迟彦惠,张平. 电接触现象和接触电阻模型的研究进展[J]. 广东技术师范学院学报, 2015, 36(2): 40-44.

[8] 潘良胜. 接触电阻影响因素的研究和分析[J]. 电气制造, 2014(4): 58-60.

[9] 何旭岩,黄旭骏,郑宏伟,等. 降低电气连接点过热的方法探讨[J]. 电子制作, 2013(19): 222.

[10] 龚炳林,刘银河,黄龙林. 浅析导电膏在电气连接中的作用[J]. 盐业与化工, 2013, 42(3): 41-42.

[11] VDI 2230-1-2014, Systematic Calculation Of Highly Stressed Bolted Joints with One Cylindrical Bolt [S].

[12] 产品集合技术规范(GPS)表面结构轮廓法表面粗糙度参数及其数值: GB/T 1031-2009 [S], 2009.

[13] 张文峰,彭向阳,陈锐民,等. 基于无人机红外视频的输电线路发热缺陷智能诊断技术[J]. 电网技术, 2014, 38(5): 1334-1338.

作者简介:

赵 强(1975) 高级工程师,从事输电线路运维检修工作;
刘 意(1982) 高级工程师,从事输电线路运维检修工作;
邱中华(1986) 高级工程师,从事输电线路带电作业工作。

(收稿日期: 2018-09-13)

[20] 卢朝晖. 基于多用户的 Windows Azure 平台下的访问控制模型的设计[J]. 科技通报, 2015, 31(5): 241-245.

作者简介:

王 帅(1988) 本科、工程师,研究方向为电气工程及其自动化。

(收稿日期: 2018-08-24)