

基于支持向量机的纯电动公交车充 / 换电站日负荷预测

刘文霞,徐晓波,周 樾

(华北电力大学 电气与电子工程学院,北京 102206)

摘要: 讨论了基于相似日选取的支持向量机电动汽车日负荷预测方法。通过对北京现有纯电动公交车充 / 换电站充电负荷的大量调研,分析了公交车充电站充电负荷的数据特征,采用关联分析方法提取了影响纯电动公交站充电负荷的因素,基于相关因素应用灰色关联理论构建相似日的小样本集合,而后建立多输入单输出的支持向量机预测模型。针对支持向量机预测模型,提出了两阶段确定模型参数的方法,首先直接确定不敏感损失参数 ε ,再通过遗传算法寻找最优核参数 p 和正则化参数 C ,以提高参数 ε 选取范围设置较大时的预测精度。实例测试结果表明,日负荷预测的均方根误差为 10.85%,能基本满足有序控制的要求;与其他预测方法相比,改进方法具有较高的预测精度和稳定性。

关键词: 电动汽车; 负荷预测; 支持向量机; 参数选择; 充电; 关联理论; 相似日

中图分类号: TM 715;U 469.72

文献标识码: A

DOI: 10.3969/j.issn.1006-6047.2014.11.007

0 引言

随着环境、能源问题的日益突出,电动汽车 EV (Electric Vehicle) 受到了各国政府的重视,电动汽车的大规模使用可以大量减少温室气体的排放,促进能源结构的调整^[1-2]。文献[3-4]讨论了大量电动汽车负荷的接入对电网产生的影响,如大量电动汽车在负荷高峰期间充电,将加剧电网负荷的峰谷差从而增加调峰的难度,在配电网中会造成局部过负荷从而加重配电网的负担等,将对电网的安全、调度和经济运行产生不利影响。所以需要较精确的电动汽车负荷预测作为依据,进行电动汽车的有序充电控制,参与电力系统的调峰甚至调频。

根据使用能源和动力驱动系统的不同,电动汽车可以分为纯电动汽车 PEV(Pure Electric Vehicles)、插电式混合动力汽车 PHEV(Plug-in Hybrid Electric Vehicles)及燃料电池电动汽车。其中,纯电动汽车完全靠电能驱动;由于受电池容量限制,尚未大规模普及,但已成为中国电动汽车产业发展的重点,且通过科研和试点取得了一定的成绩。目前,在国内已有一些一定规模的纯电动公交车和出租车的试点。北京在奥运会后逐步增加了电动公交汽车的数量,采用换电方式,公交站也作为充换电站。目前的运行已经有了一定的规模,充电站的负荷也初步体现出了统计特征,在此背景下,开展充电站的负荷预测,为后续电动汽车有序充电管理系统提供有力支撑。

本文研究对象是公交汽车充电站日负荷预测方法。由于试点较少,国内外对纯电动汽车充电站负荷

预测研究的文献不多^[5-6]。文献[5]提出了模糊聚类 and BP 神经网络相结合的方法预测换电站的短期负荷。文献[6]采用蒙特卡洛的方法预测电动汽车充电负荷的时空分布特性。本文为预测充电站的日充电负荷曲线,借鉴了电力系统日负荷预测的方法。目前,电力系统日负荷预测方法主要分为 2 种:基于历史负荷数据的传统经典方法和基于历史负荷数据及其影响因素的机器学习方法。为了在预测模型中能更好地结合各种因素影响,提高预测精度,基于人工智能技术的神经网络预测模型和模糊预测模型得到了广泛的应用^[7-9]。神经网络等机器学习方法是基于最小化经验风险的机器学习方法,但是在实际应用中,样本的数量总是有限的,此时经验风险最小并不意味着期望风险最小,如神经网络的“过学习”问题,在某些情况下,训练误差过小会导致推广能力的下降,即泛化能力变差。由 Vapnik 等人提出的基于结构风险最小化原理(SRM)的支持向量机 SVM(Support Vector Machine)算法,以有限的样本信息在模型的复杂性和学习能力之间寻求最佳折中,以获得最佳的泛化能力。

由于 SVM 的参数对其性能影响很大,所以 SVM 用于预测的研究主要集中于预测模型参数的选择,目前确定参数的方法主要有 2 种,一种是经验或实验法选取^[10-12],另一种是使用遗传算法(GA)^[13-14]、粒子群优化(PSO)算法^[15]、微分进化^[16]等优化算法确定。文献[10]的参数分析表明核参数 p 和正则化参数 C 对 SVM 的性能影响较大,当 p 和 C 固定于某一合适值时,SVM 性能对不敏感损失参数 ε 不敏感。本文采用两阶段确定参数的方法:第一阶段对参数 ε 取定值;第二阶段采用 GA 对 p 和 C 寻优。

本文对北京市某电动公交车站的充电负荷特性进行了分析,确定了影响电动公交车充电负荷的主要

收稿日期:2013-12-25;修回日期:2014-09-24

基金项目:国家高技术研究发展计划(863 计划)资助项目(2011-AA05A109)

Project supported by the National High Technology Research and Development Program of China(863 Program)(2011AA05A109)

因素,并根据这些因素,通过计算灰色关联度选取相似日建立小样本作为 SVM 模型的输入数据;介绍了 SVM 模型的回归算法、确定 SVM 预测模型的 3 个参数的方法和整个预测流程;使用公交换电站的充电负荷数据对本文的方法进行了验证。

1 纯电动汽车负荷特性分析

电动汽车的类型目前主要分为公交车、出租车、公共事业车、微型车和私人乘用车。电池的更换方式有整车充电和更换电池 2 种。充放电方式有单向无序充电、单向有序充电和双向有序充放电 3 种。电动汽车充电负荷的形成是由用户的行为引发,在充电设施处产生,因此充电负荷的分布具有时空特性,必须明确汽车在特定地点、特定时间的充电行为。影响电动汽车充电负荷的主要因素包括:用户的行为特性、电动汽车电池的充放电特性、电动汽车充电方式和电动汽车规模等。作为电动汽车充电负荷的触发侧,用户的行为特性对于确定电动汽车充电负荷至关重要。用户的行为特性包含电动汽车的类型、行驶里程、接入电网的时间、离开电网的时间、停放的场所等。天气情况、温度、节假日等外在因素也会影响用户的行为特性。

1.1 电动公交车充电站负荷特性分析

本文采集了北京某公交车充电站 2012 年 7、8、9 月的充电负荷数据,公交车规模 30 辆左右,目前,9 kW 充电桩有 240 个,75 kW 应急充电桩 4 个,充电方式为单向无序换电充电。

图 1 是该充电站 2012 年 7 月 6 日、7 月 7 日和 7 月 8 日的日充电负荷曲线图。从该日负荷曲线可以看出,负荷峰值出现在 13:00—14:00 和 17:00—18:00,所有充电在 23:00 之前基本完成,此时开始没有负荷,一直持续到次日 09:00 左右。

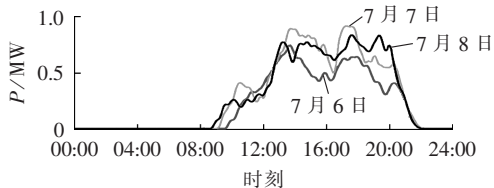


图 1 电动公交车充电站日负荷曲线

Fig.1 Daily load curves of electric bus charging station

通过对 7、8、9 月的充电负荷曲线(见图 2)的观察,可以看出它们的变化趋势具有一定的相似性,表明了选取相似日进行负荷预测的可行性。

图 3 是 2012 年 7 月该充电站充电功率日最大峰谷差直方图。其中,最大峰谷差为 1.065 MW,最小峰谷差为 0.64 MW。峰谷差围绕 0.8 MW 上下波动,并且波动幅度较明显。图 4 是 7 月 6 日每 5 min 的短时波动量曲线。从图 4 中可以看出,在有充电负荷

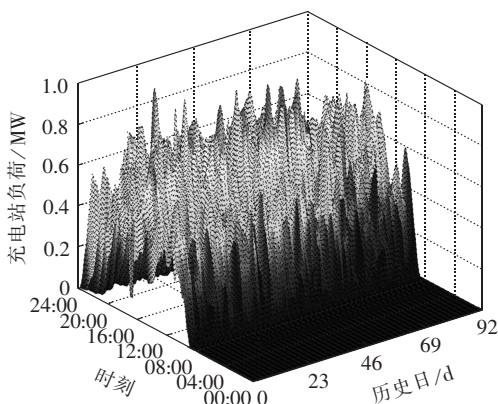


图 2 电动公交车充电站 3 个月的日负荷曲线

Fig.2 Daily load curves of electric bus charging station for three months

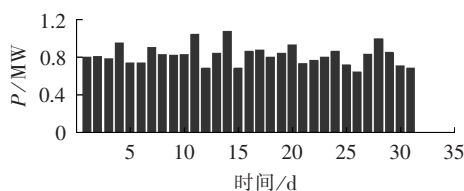


图 3 日最大峰谷差

Fig.3 Maximum daily peak-valley differences

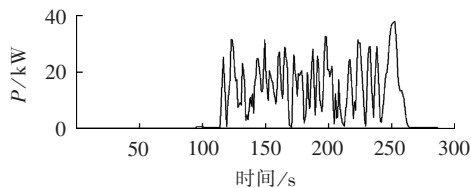


图 4 5 min 负荷波动量

Fig.4 Load fluctuations for 5 minutes

期间,短时波动量较大,最大值为 37 kW,最小波动量接近为 0,最小波动量和最大波动量之间的差值比较大。充电站充电负荷的短时波动较为明显。

1.2 影响因素分析

电动公交车充电站的充电负荷是由电池充电产生,影响电池消耗速度的因素有很多,公交车所用电池型号一致,所以这里不考虑电池本身的特性。由于公交车行驶线路的固定,影响电池消耗的速度主要与当时的车速有关,而车速又与当时的车流量有关。文献[17]中叙述了影响车流量的因素,如温度、天气情况(晴、下雨等)等。结合现有数据,本文着重分析温度、日类型、天气情况、充电开始时间及充电结束时间对公交充电站充电负荷的影响。

图 5 是 7 月 1 日至 9 月 30 日的日最高温度(归一化值)和日平均负荷的散点图,为了证明日最高温度和日平均负荷之间的相关性,本文采用计算线性相关系数的方法衡量 2 个随机变量间的相关程度。计算公式如下:

$$r = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (1)$$

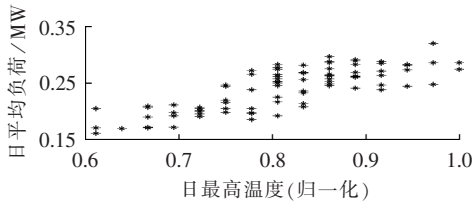


图 5 日平均负荷与日最高温度散点图

Fig.5 Scatter plot of daily average load and maximum temperature

其中, r 为相关系数; X_i, Y_i 分别为第 i 日的最高温度和日平均负荷; \bar{X}, \bar{Y} 分别为历史日最高温度和日平均负荷的平均值; N 为历史日天数。当 $|r| > 0.8$ 时, 认为最高温度和日平均负荷有强相关性。

通过计算, 日最高温度和日平均负荷的相关系数为 0.800 2, 表明最高温度和日平均负荷之间存在强相关关系。夏天当温度升高时, 充电负荷也增加, 其原因是天气炎热、公交车开空调, 加快了电能的消耗, 最终导致充电负荷的增加。

日类型和天气情况会影响用户的行为特性, 进而影响充电站的充电负荷。由表 1 可以看出日类型对充电站负荷的影响不显著, 这是由于公交车在工作日和非工作日的调度安排差别不大。但是, 今后若调度安排改变, 对于不同日类型, 公交车的出行规律也不同, 进而影响充电站的充电负荷曲线。为了使本文的模型更具泛用性, 因此本文考虑日类型的影响。由表 2 可以看出下雨和无雨对充电站充电负荷有一定的影响。雨天的充电站充电负荷小于无雨天。此外, 由图 1 可以看出充电站充电开始时间、充电结束时间不同, 日负荷曲线白天从零开始上升、夜里下降为零的时刻也不相同, 但是对于预测日, 充电开始时间和充电结束时间无法得知。所以本文考虑的因素有最高温度、日类型和天气情况。

本文以最高温度的归一化值和日类型、天气情

表 1 不同日类型负荷分析

Table 1 Load analysis for different day types

日类型	3 个月日最大负荷均值/MW	3 个月日平均负荷均值/MW
星期一	0.799 84	0.247 48
星期二	0.810 65	0.235 31
星期三	0.830 51	0.244 16
星期四	0.761 87	0.236 19
星期五	0.802 56	0.238 53
星期六	0.813 14	0.236 28
星期日	0.775 49	0.236 24

表 2 不同天气情况负荷分析

Table 2 Load analysis for different weathers

天气情况	3 个月日最大负荷均值/MW	3 个月日平均负荷均值/MW
有雨	0.764 83	0.228 57
无雨	0.803 69	0.242 96

况的映射值作为特征量选取相似日。归一化公式如式(2)所示, 日类型和天气情况的映射值分别如表 3、4 所示。

$$x = \frac{x'}{x'_{\max}} \quad (2)$$

其中, x' 为某一天的最高温度; x'_{\max} 为 7、8、9 月中最高温度; x 为归一化后的值。

表 3 日类型映射值

Table 3 Mapping value for different day types

日类型	映射值
星期一	0.1
星期二、三、四	0.2
星期五	0.3
星期六	0.7
星期日	1.0

表 4 天气情况映射值

Table 4 Mapping value for different weathers

天气情况	映射值
晴	0.1
多云	0.2
阴	0.3
小雨/小雪	0.7
中雨/中雪	1.0
大雨/大雪/极端天气	1.5

1.3 基于灰色关联度的样本筛选

通过选取相似日的方法, 可以提高预测精度, 也可以降低 SVM 样本训练时间。传统选取相似日的方法是基于人工经验的选取, 常常会引入不良样本, 增大预测误差。目前选取相似日的方法有证据理论^[12]、聚类分析^[18]、趋势相似度法^[19]、灰色关联法^[20]等。本文采用计算灰色关联度的方法选取相似日。

采用灰色关联度分析方法选取相似日, 本文计入的因素为最高温度、日类型和天气情况, 使得训练样本和预测日之间在气象特征上具有较高相似性, 可以提高预测精度。计算历史日 n 天的因素序列 $X_m = (x_m(1), \dots, x_m(t))$ 与待预测日的因素序列 $X_0 = (x_0(1), \dots, x_0(t))$ 之间的相关度。其中, $m=1, 2, \dots, n; t$ 是计入因素个数。

首先求取待预测日因素序列与历史日因素序列的差序列矩阵 $\Delta = [\Delta_{10}, \dots, \Delta_{m0}, \dots, \Delta_{n0}]'$, 其中 $\Delta_{m0} = (|x_m(1) - x_0(1)|, \dots, |x_m(t) - x_0(t)|)$, 再求取两级最大差和最小差:

$$\Delta_{\min} = \min \min (\Delta)$$

$$\Delta_{\max} = \max \max (\Delta)$$

最后计算 X_0 对 X_m 的灰色关联度:

$$r(X_0, X_m) = \frac{1}{t} \sum_{k=1}^t \frac{\Delta_{\min} + \rho \Delta_{\max}}{\Delta_{m0}(k) + \rho \Delta_{\max}}$$

其中, $\Delta_{m0}(k) = |x_m(k) - x_0(k)|$; ρ 为分辨系数, 取 0.5。

2 SVM 回归模型及参数选择

2.1 SVM 回归算法

SVM 回归算法目前主要使用的是 ν -SVM、 ε -SVM 和 LS-SVR 等算法。本文采用 ε -SVM 算法。

对于非线性负荷,通过非线性映射 ϕ 将每个样本点映射到高维空间,在高维空间中作线性回归。样本点如下:

$$G \{(\mathbf{x}_i, y_i)\}_{i=1}^l \quad (3)$$

其中, \mathbf{x}_i 和 y_i 分别为输入和输出量,其中 \mathbf{x}_i 为包括最高温度归一化值、日类型,天气情况映射值和相似日负荷值的向量, y_i 为真实值; l 为样本点个数。

则非线性负荷的预测模型为:

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b \quad \mathbf{w} \in \mathbf{R}^n \quad (4)$$

其中, $f(\mathbf{x})$ 为模型输出的预测值; \mathbf{w} 为权向量; b 为偏置。

根据 Vapnik 的最小化结构风险原则,其结构风险定义为:

$$R_{\text{reg}}(f) = \frac{1}{2} \|\mathbf{w}\|^2 + C \frac{1}{l} \sum_{i=1}^l e_i \quad (5)$$

其中, $\frac{1}{2} \|\mathbf{w}\|^2$ 是正则化部分,反映的是函数 $f(\mathbf{x})$ 的泛化能力; C 是正则化系数,决定正则化部分和经验风险之间的平衡; $e_i = f(\mathbf{x}_i) - y_i$, $C \frac{1}{l} \sum_{i=1}^l e_i$ 是经验风险,由 ε 不敏感损失函数来度量,如式(6)所示。

$$e_i = \begin{cases} 0 & |y_i - f(\mathbf{x}_i)| \leq \varepsilon \\ |y_i - f(\mathbf{x}_i) - \varepsilon| & |y_i - f(\mathbf{x}_i)| > \varepsilon \end{cases} \quad (6)$$

因此, ε -SVR 的目标函数为:

$$\begin{aligned} \min & \frac{1}{2} \|\mathbf{w}\|^2 + C \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{s.t.} & y_i - \mathbf{w}\phi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i \quad \xi_i \geq 0 \\ & \mathbf{w}\phi(\mathbf{x}_i) + b - y_i \leq \varepsilon + \xi_i^* \quad \xi_i^* \geq 0 \end{aligned} \quad (7)$$

其中, ξ_i, ξ_i^* 为松弛变量。为求解该问题,引入拉格朗日乘子和核函数,根据对偶理论,把该问题变为:

$$\begin{aligned} \max & -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(\mathbf{x}_i, \mathbf{x}_j) - \\ & \varepsilon \sum_{i=1}^l (\alpha_i^* + \alpha_i) + \sum_{i=1}^l y_i (\alpha_i^* - \alpha_i) \\ \text{s.t.} & \sum_{i=1}^l (\alpha_i^* - \alpha_i) = 0 \quad 0 \leq \alpha_i^*, \alpha_i \leq C/l \end{aligned} \quad (8)$$

解得:

$$f(\mathbf{x}) = \sum_{n_{\text{sv}}} (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b \quad (9)$$

其中, n_{sv} 为支持向量的个数。由式(9)可以看出,只要知道核函数的形式即可进行预测,而不需要知道映射函数 $\phi(\mathbf{x})$ 和高维空间 \mathbf{R}^n 。由于目前尚无理论指导选取核函数,常用的核函数有线性函数、多项式函数、径向基函数、多层感知器函数,因此,本文使用的是径向基核函数,如下式所示:

$$K(\mathbf{x}_i, \mathbf{x}) = \exp[-\|\mathbf{x}_i - \mathbf{x}\|^2 / (2p^2)] \quad (10)$$

由式(8)可知, ε -SVR 模型需要选择正则化系数 C 、参数 ε 和核参数 p 。

2.2 参数自适应

以 9 月 1 日至 9 月 7 日的相似日作为训练样本,以 9 月 8 日为测试样本,则当 p 在 0.001~1 范围内以步长 0.05 变化、 C 在 0.1~40 范围内以步长为 2 变化、 ε 在 0.001~0.01 范围内以步长为 0.001 变化时,试验 1 结果如表 5 所示。当 C 和 p 在上述范围内变化、 $\varepsilon=0.001$ 时,试验 2 结果如图 6 所示。采用均方根误差(RMSE)进行评价:

$$\text{RMSE} = \frac{1}{\max y_i} \sqrt{\frac{1}{T} \sum_{i=1}^T (y_i - \hat{y}_i)^2}$$

其中, $T=1, 2, \dots, 96$ 为日负荷预测总点数。

表 5 随 ε 变化的平均 RMSE 值
Table 5 Average RMSE for different ε values

ε	平均 RMSE	ε	平均 RMSE
0.001	0.0820	0.006	0.0828
0.002	0.0821	0.007	0.0830
0.003	0.0822	0.008	0.0828
0.004	0.0823	0.009	0.0828
0.005	0.0824	0.010	0.0830

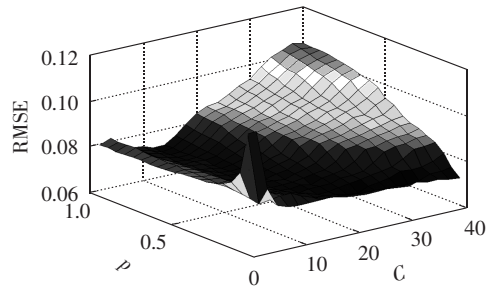


图 6 RMSE 随 p 和 C 的变化值
Fig.6 RMSE varying along with p and C

试验 1 的 RMSE 均处于 0.07~0.11 之间,从表 5 可以看出,平均 RMSE 均处于 0.082~0.083 之间, RMSE 变化较稳定。从图 6 中可以看出,当 $\varepsilon=0.001$ 、 C 和 p 在上述范围内变化时, RMSE 都处于 0.07~0.11 之间,变化不大,较稳定。可见参数 ε 对模型性能的影响不显著。因此,本文对 ε 取定值,由于参数 ε 在 ε -SVR 模型中控制着支持向量的稀疏性, ε 越大,支持向量的个数就越少,当大于某一值时,就会出现“欠学习”现象,增大预测误差。本文令 $\varepsilon=0.001$ 。

试验发现选取不合适的 C 和 p ,会对 SVR 的性能产生较大影响。所以本文使用 GA 确定参数 C 和 p 。GA 不依赖于初始种群,通过复制、交叉和变异作用,具有较好的全局优化特性。步骤如下:

a. 随机产生初始种群,个体数目取 20,根据实际数据,分别从 (0.1, 40) 和 (0.001, 1) 范围内随机选取 C 和 p 的值;

b. 计算适应度,适应度函数为结构风险倒数,即 $1/R_{\text{reg}}(f) = 1 / \left(\frac{1}{2} \|\mathbf{w}\|^2 + C \frac{1}{l} \sum_{n_{\text{sv}}} |f(\mathbf{x}_i) - y_i - \varepsilon| \right)$;

c. 选择轮盘赌策略,适应度高的个体被选中遗传到下一代的概率高,适应度低的个体可能被淘汰,然后对下一代的个体进行交叉和变异操作;

d. 由新一代的种群返回步骤 b,经过一定数量的迭代后,就会得到最优的 C 和 p ,即所选的 C 和 p 使得结构风险最小。

表 6 选取了 3 组训练样本进行收敛性分析(训练样本的形成见算例验证部分),每组数据重复上述 4 个步骤 300 次,记录下每次最终的适应度作统计,如数据 1,除去一些明显的异值,剩下的数据的个数占总的数据个数的 93%,其均值为 1.9392,方差为 0.00072。由表 6 可知,GA 具有较好的收敛性。

表 6 收敛性分析
Table 6 Convergence analysis

训练样本	适应度均值	适应度方差	占百分比/%
数据 1	1.9392	0.00072	93
数据 2	2.1374	0.00079	94
数据 3	1.7295	0.00071	94

2.3 预测流程

预测流程如下:

a. 采用五点滑动平均^[21]的方法对充电负荷数据进行预处理,然后再对负荷数据、影响因素数据进行归一化;

b. 采用计算灰色关联度的方法选取关联度最大的 3 天为待预测日的相似日,相似日的负荷数据和因素数据共同构成训练样本和测试样本,作为 ε -SVR 模型的输入;

c. 使用 GA 选取最佳正则化参数 C 和核参数 p ;

d. 利用 ε -SVR 模型进行电动汽车充电站的日充电负荷预测。

具体流程图如图 7 所示。

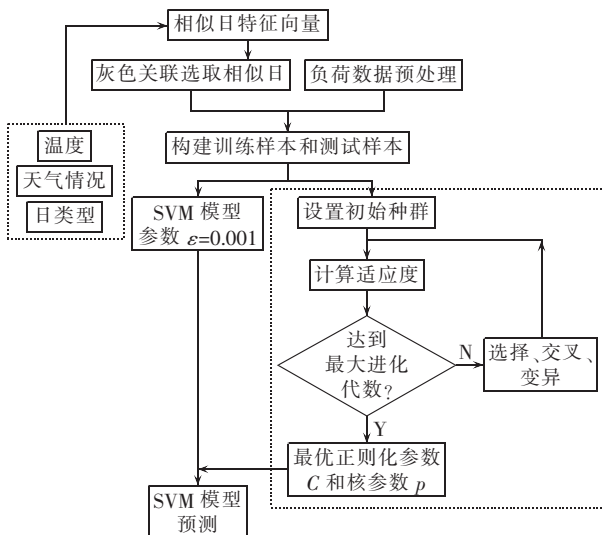


图 7 充电负荷预测流程图

Fig.7 Flowchart of charging load forecasting

3 算例验证

对北京市某公交换电站进行 96 点的日充电负荷预测,以 2012 年 7 月 1 日至 8 月 31 日的充电负荷数据为历史数据,以 9 月 1 日至 9 月 7 日为训练样本,9 月 8 日为测试样本,再以 7 月 2 日至 9 月 1 日为历史数据,9 月 2 日至 9 月 8 日为训练样本,9 月 9 日为测试样本,依此类推,将最新的信息放入模型中。采用 RMSE 进行评价,所得的预测结果如表 7 所示。

表 7 预测结果
Table 7 Results of forecasting

日期	RMSE / %					一元线性回归
	标准 SVM	改进 SVM	遗传算法 ($\varepsilon \in [0.001, 0.01]$)	遗传算法 ($\varepsilon \in [0.001, 0.1]$)	ARMA	
09-08	12.02	6.95	7.04	9.60	20.12	18.29
09-09	13.28	14.45	14.30	14.12	13.70	9.75
09-10	11.28	11.16	11.14	12.58	20.80	12.73
09-11	9.34	7.58	7.92	8.72	14.47	12.24
09-12	13.05	11.83	12.29	13.33	14.53	12.66
09-13	13.82	13.00	13.79	14.80	14.59	10.52
09-14	13.26	12.39	11.23	11.69	20.08	11.63
09-15	12.08	11.92	11.52	12.11	12.03	8.44
09-16	9.26	7.22	7.55	9.67	13.34	11.16
09-17	7.87	7.36	7.42	7.85	11.67	6.41
09-18	14.45	12.27	12.16	13.15	17.32	12.63
09-19	11.04	10.75	10.74	10.79	12.28	12.12
09-20	13.65	13.50	13.40	13.74	14.56	11.17
09-21	12.67	12.55	12.35	12.39	13.48	9.86
09-22	8.50	8.31	7.20	8.45	14.38	12.58
09-23	8.65	7.23	8.34	8.65	18.13	13.97
09-24	13.64	11.22	12.40	13.34	26.12	16.87
09-25	17.87	11.26	10.78	14.29	15.20	13.12
09-26	11.13	10.68	10.85	10.54	10.04	9.88
09-27	11.23	10.26	10.36	11.30	19.03	20.36
09-28	16.67	15.61	16.00	16.42	23.14	12.75
09-29	13.39	12.13	12.57	13.08	15.28	13.07
09-30	16.28	9.86	9.98	15.13	15.26	12.58
平均值	12.37	10.85	10.93	11.99	16.00	12.27

从表 7 可以看到,使用标准 SVM 得到的预测误差,最大为 17.87%,最小为 7.87%,平均为 12.37%,采用本文的方法,得到的预测误差最大为 15.61%,最小为 6.95%,平均为 10.85%,基本满足预测要求。相比于标准 SVM 得到的预测结果,本文方法的精度提高了 1.52%。

当 C 在 0.1~40 之间, p 在 0.001~1 之间, ε 在 0.001~0.01 之间时,使用 GA 对以上 3 个参数寻优时,得到的预测误差最大为 16%,最小为 7.04%,平均值为 10.93%,由表 7 可以看出预测结果和本文提出的方法接近。

当 C 在 0.1~40 之间, p 在 0.001~1 之间, ε 在 0.001~0.1 之间时,使用 GA 对以上 3 个参数寻优时,得到的预测误差最大为 16.42%,最小为 7.85%,平均

误差为 11.99%。由表 7 可以看出当参数 ε 选取的寻优范围不合理时,会降低预测的精度。

表 7 的最后 2 列是采用传统的 ARMA 和一元线性回归方法得到的预测结果。由于充电站充电负荷的波动性较大,使得 ARMA 和一元线性回归的预测误差较大。在这 23 天中,ARMA 的 RMSE 最大为 26.12%,最小为 10.04%,平均值为 16%,波动范围为 16.1%;一元线性回归的 RMSE 最大为 20.36%,最小为 6.41%,平均值为 12.27%,波动范围为 13.95%;而本文的改进 SVM 方法的 RMSE 最大为 15.61%,最小为 6.95%,平均值为 10.85%,波动范围为 8.66%;可见本文的方法具有较好的预测精度和稳定性。

图 8 为采用本文方法预测的 9 月 8 日和 9 月 28 日的充电站充电负荷曲线图。9 月 8 日的预测误差最小, RMSE 为 6.95%;9 月 28 日的预测误差最大, RMSE 为 15.61%。

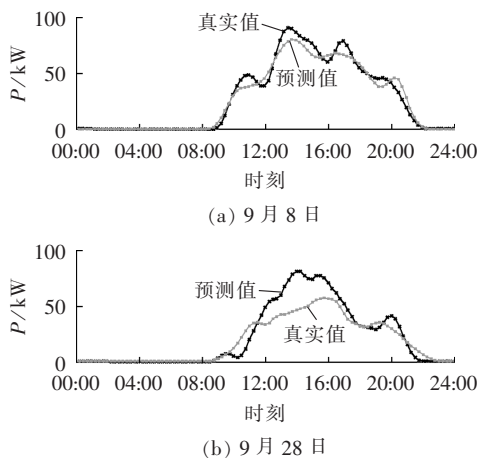


图 8 电动公交车充电站日负荷预测误差最小/最大日图

Fig.8 Actual and forecast daily load curves of electric bus charging station for minimum and maximum forecasting error days

4 结论

a. 本文结合现有数据,分析了影响充电站公交车充电负荷的主要因素。对原始数据进行了预处理,减少了坏数据对预测结果的影响,并采用灰色关联方法选取相似日,增加了数据的相关性,采用基于结构风险最小的 SVM 方法进行建模,具有全局最优的优点和更好的泛化性能。通过实例表明了公交车充电站的日负荷预测精度平均可达 10.85%,预测时间在 5 min 之内,能基本满足电动汽车充电有序控制的要求。

b. 对于 SVM 模型,本文采用了两阶段确定模型参数的方法,与标准的 SVM 模型相比,预测精度提高了 1.52%;当学习参数 ε 的选取范围较大时,改进方法的预测精度与 3 个参数同时寻优的方法相比,提高了 1.14%。同时与传统的 ARMA 和一元线性回

归方法相比,采用本文的方法具有较好的预测精度和稳定性。

c. 目前公交电动车数量不多,公交电动充换电站的充电负荷数据的统计特性不明显,预测误差较大。随着国家对电动汽车的推广,充换电站的充电负荷数据的统计特性会越来越明显,未来本文方法的预测精度将会进一步提高,为电动汽车的有序控制提供强有力的支撑。

参考文献:

- [1] GUO Ping, LIU Peng. Research on development of electric vehicles in China[C]//International Conference on Future Information Technology and Management Engineering. Changzhou, China: IEEE, 2010:94-96.
- [2] SITU Lixin. Electric vehicle development: the past, present & future[C]//3rd International Conference on Power Electronics Systems and Applications. Hong Kong, China: IEEE, 2009:1-3.
- [3] 胡泽春, 宋永华, 徐智威, 等. 电动汽车接入电网的影响与利用[J]. 中国电机工程学报, 2012, 32(4):1-10.
HU Zechun, SONG Yonghua, XU Zhiwei, et al. Impacts and utilization of electric vehicles integration into power system [J]. Proceedings of the CSEE, 2012, 32(4):1-10.
- [4] 王建, 吴奎华, 刘志珍, 等. 电动汽车充电对配电网负荷的影响及有序控制研究[J]. 电力自动化设备, 2013, 33(8):47-52.
WANG Jian, WU Kuihua, LIU Zhizhen, et al. Impact of electric vehicle charging on distribution network load and coordinated control [J]. Electric Power Automation Equipment, 2013, 33(8):47-52.
- [5] 张维戈, 颀飞翔, 黄梅, 等. 快换式公交充电站短期负荷预测方法的研究[J]. 电力系统保护与控制, 2013, 41(4):61-66.
ZHANG Weige, XIE Feixiang, HUANG Mei, et al. Research on short-term load forecasting methods of electric buses charging station [J]. Power System Protection and Control, 2013, 41(4):61-66.
- [6] 张洪财, 胡泽春, 宋永华, 等. 考虑时空分布的电动汽车充电负荷预测方法[J]. 电力系统自动化, 2014, 38(1):13-20.
ZHANG Hongcai, HU Zechun, SONG Yonghua, et al. A prediction method for electric vehicle charging load considering spatial and temporal distribution [J]. Automation of Electric Power Systems, 2014, 38(1):13-20.
- [7] 谢宏, 牛东晓, 张国立, 等. 一种模糊模型的混合建模方法及在短期负荷预测中的应用[J]. 中国电机工程学报, 2005, 25(8):17-22.
XIE Hong, NIU Dongxiao, ZHANG Guoli, et al. A hybrid fuzzy modeling method and its application in short-term load forecast [J]. Proceedings of the CSEE, 2005, 25(8):17-22.
- [8] 葛少云, 贾欧莎, 刘洪. 基于遗传灰色神经网络模型的实时电价条件下短期电力负荷预测[J]. 电网技术, 2012, 36(1):224-229.
GE Shaoyun, JIA Ousha, LIU Hong. A gray neural network model improved by genetic algorithm for short-term load forecasting in price-sensitive environment [J]. Power System Technology, 2012, 36(1):224-229.
- [9] HINOJOSA V H, HOESE A. Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms [J]. IEEE Transactions on Power Systems, 2010, 25(1):565-574.

- [10] 李元诚,方廷健,于尔铿. 短期负荷预测的支持向量机方法[J]. 中国电机工程学报,2003,23(6):55-59.
LI Yuancheng,FANG Tingjian,YU Erkeng. Study of support vector machines for short-term load forecasting[J]. Proceedings of the CSEE,2003,23(6):55-59.
- [11] 牛东晓,谷志红,邢棉,等. 基于数据挖掘的 SVM 短期负荷预测方法研究[J]. 中国电机工程学报,2006,26(18):6-12.
NIU Dongxiao,GU Zhihong,XING Mian,et al. Study of forecasting approach to short-term load of SVM based on data mining[J]. Proceedings of the CSEE,2006,26(18):6-12.
- [12] 李鑫滨,张娟,张岩,等. 基于 D-S 证据理论的相似日支持向量机短期负荷预测[J]. 电网技术,2010,34(7):143-147.
LI Xinbin,ZHANG Juan,ZHANG Yan,et al. Short-term load forecasting for similar days based on support vector machine and Dempster-Shafer theory[J]. Power System Technology,2010,34(7):143-147.
- [13] 牛东晓,刘达,陈广娟,等. 基于遗传优化的支持向量机小时负荷滚动预测[J]. 电工技术学报,2007,22(6):148-153.
NIU Dongxiao,LIU Da,CHEN Guangjuan,et al. Support vector machine models optimized by genetic algorithm for hourly load rolling forecasting[J]. Transactions of China Electrotechnical Society,2007,22(6):148-153.
- [14] WU Chihhung,TZENG Gwohshung,LIN Rongho. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression[J]. Expert Systems with Application,2009,36(3):4725-4735.
- [15] WU Qi. A hybrid-forecasting model based on Gaussian support vector machine and chaotic particle swarm optimization [J]. Expert Systems with Application,2010,37(3):2388-2394.
- [16] 彭春华,刘刚,孙惠娟. 基于小波分解和微分进化支持向量机的风电场风速预测[J]. 电力自动化设备,2012,32(1):9-13.
PENG Chunhua,LIU Gang,SUN Huijuan. Wind speed forecasting based on wavelet decomposition and differential evolution-support vector machine for wind farms[J]. Electric Power Automation Equipment,2012,32(1):9-13.
- [17] 宫晓燕,汤淑明. 基于非参数回归的短时交通流量预测与事件检测综合算法[J]. 中国公路学报,2003,16(1):82-86.
GONG Xiaoyan,TANG Shuming. Integrated traffic flow forecasting and traffic incident detection algorithm based on non-parametric regression[J]. China Journal of Highway and Transport,2003,16(1):82-86.
- [18] 张平,潘学萍,薛文超. 基于小波分解灰色聚类 and BP 神经网络的短期符合预测[J]. 电力自动化设备,2012,32(11):121-125.
ZHANG Ping,PAN Xueping,XUE Wenchao. Short-term load forecasting based on wavelet decomposition,fuzzy gray correlation clustering and BP neural network[J]. Electric Power Automation Equipment,2012,32(11):121-125.
- [19] 莫维仁,张伯明,孙宏斌,等. 短期负荷预测中选择相似日的探讨[J]. 清华大学学报:自然科学版,2004,44(1):106-109.
MO Weiren,ZHANG Boming,SUN Hongbin,et al. Select similar-days for short-term load forecasting[J]. Journal of Tsinghua University:National Science Edition,2004,44(1):106-109.
- [20] 刘梦良,刘晓华,高荣. 基于相似日小波支持向量机的短期电力负荷预测[J]. 电工技术学报,2006,21(11):59-64.
LIU Mengliang,LIU Xiaohua,GAO Rong. Short term load forecasting using wavelet transform and SVM based on similar-days [J]. Transactions of China Electrotechnical Society,2006,21(11):59-64.
- [21] 孙苗钟. 基于 MATLAB 的震动信号平滑处理方法[J]. 电子测量技术,2007,30(6):55-57.
SUN Miaozhong. Smooth processing methods of vibration signal based on MATLAB [J]. Electronic Measurement Technology, 2007,30(6):55-57.

作者简介:



刘文霞

刘文霞(1967—),女,北京人,副教授,博士,研究方向为电力系统智能规划、电力系统风险评估、电力系统通信(E-mail:lwxia@ncepu.edu.cn);

徐晓波(1990—),男,江苏苏州人,硕士研究生,研究方向为电力系统分析;

周 禛(1989—),男,上海人,硕士研究生,研究方向为电力系统分析。

Daily load forecasting based on SVM for electric bus charging station

LIU Wenxia,XU Xiaobo,ZHOU Xi

(North China Electric Power University,Beijing 102206,China)

Abstract: The method of daily electric vehicles load forecasting based on the similar days by the SVM (Support Vector Machine) is discussed. The charging load data of electric bus charging stations in Beijing are researched and their characteristics are analyzed. The correlation analysis is applied to extract their influencing factors and the gray relational analysis is applied to establish the small sample of similar days, based on which a multi-input single-output SVM forecasting model is built. Two steps are proposed to determine the model parameters;the insensitive loss parameter ε is directly set in the first step,while the optimal kernel parameter p and the regularization parameter C are determined by the genetic algorithm to improve the forecast accuracy when the selection range of ε is larger. Results of case test show that,the RMSE(Root Mean Square Error) of daily load forecast is 10.85%,basically meeting the requirement of coordinated control. Compared with other forecasting methods,the proposed method has better accuracy and stability.

Key words: electric vehicles; electric load forecasting; support vector machines; parameter selection; charging; correlation theory; similar days