

基于粒子群优化和支持向量机的电力负荷预测

曾 勍 炜¹, 徐 知 海¹, 吴 键²

(1 南昌大学 网络中心, 江西 南昌 330031; 2 江西省电力公司, 江西 南昌 330077)

摘 要: 提出支持向量机的粒子群优化算法的用电量预测方法. 其中, 采用粒子群优化算法选取较优的支持向量机训练参数组合. 以江西省 2008 年 7 月~10 月的用电量数据以及相关特征数据作为实验数据, 实验结果表明该算法电量负荷预测精度高于 BP 神经网络.

关键词: 负荷预测; 支持向量机; 粒子群; BP 网络

中图分类号: TP183

文献标识码: A

文章编号: 1000-7180(2011)01-0147-03

Forecasting of Electricity Load Based on Particle Swarm Optimization and Support Vector Machine

ZENG Qing-wei¹, XU Zhi-hai¹, WU Jian²

(1 Network Center, Nanchang University, Nanchang 330031, China;

2 Dispatching and Communication Center, Jiangxi Electric Power Company, Nanchang 330077, China)

Abstract: The values of training parameters of support vector machine have close contact with its forecasting accuracy. Therefore, particle swarm optimization algorithm and support vector machine (PSVM) is proposed to predict electricity consumption in the study. In the model, particle swarm optimization algorithm is used to select the optimal training parameters of support vector machine. The electricity consumption data and relevant features data of Jiangxi province from July to October in 2008 are used as the experimental data. The experimental results indicate that the PSVM model has higher prediction accuracy than BP neural network in the forecasting of electricity load.

Key words: forecasting of electricity load; support vector machine; particle swarm; BP network

1 引言

我国电力市场在长期运行中经常出现严重的电能需求和供给之间的不平衡现象^[1]. 为了避免用电量过多或过少带来的经济损失, 需对用电量进行预测, 并以此对用电量做出科学的规划和指导, 从而使电能供需接近平衡. 由于用电量受温度、气候等诸多因素的影响, 其预测模型难以用数学模型进行表示. 人工神经网络(ANN)具有强大的非线性映射能力、具有自适应、自学习、容错性和并行处理等性质, 人工神经网络理论在用电量预测中具有较好的应用, 取得了一定的效果^[2]. 然后, 人工神经网络尤其是 BP 神经网络具有过拟合, 容易陷入局部极值等缺点, 影响了其应用效果.

支持向量机(Support Vector Machine, SVM)是由 Vanpik 领导的 AT&T Bell 实验室研究小组提出的一种新的机器学习方法^[3]. 它能有效解决小样本、非线性及高维等问题. 支持向量机的预测精度与其训练参数的取值有着较大关系^[4]. 对此, 提出基于粒子群优化算法与支持向量机(PSVM)的用电量预测方法. 其中, 采用粒子群优化算法(Particle Swarm Optimization, PSO)^[5]选取较优的支持向量机训练参数组合. 粒子群优化算法是一种基于群体的智能寻优算法^[6], 通过它能得到较优的支持向量机电量预测模型. 实验结果表明基于粒子群优化算法与支持向量机的用电量预测具有很高的预测精度, 非常适合于用电量预测.

2 支持向量机回归模型

对于训练样本集 $(x_i, y_i)_{i=1}^n, x_i \in R^n$, 为输入矢量, $y_i \in R$ 为相应的输出值. 支持向量机回归理论的基本思想是寻找一个输入空间到输出空间的非线性映射 θ , 将数据 x 映射到一个高维特征空间 F , 并在此特征空间通过函数 $f(x) = [w \cdot \theta(x)] + b$ 进行线性回归, 其中 b 为阈值, w 为权值矢量.

根据统计理论, 支持向量机通过对以下目标函数极小化确定回归函数:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (1)$$

$$\text{s. t. } \begin{cases} y_i - w \cdot \varphi(x) - b \leq \epsilon + \xi_i \\ w \cdot \varphi(x) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

式中, ξ 和 ξ^* 为非负松弛变量; C 为惩罚因子, 其作用是在经验风险和模型复杂度之间的折中; ϵ 为不敏感损失函数参数, 具体定义为:

$$|f(x) - y|_\epsilon = \begin{cases} |f(x) - y| - \epsilon & |f(x) - y| \geq \epsilon \\ 0 & |f(x) - y| < \epsilon \end{cases} \quad (2)$$

该问题转化为对偶问题:

$$\max -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i - a_i^*)(a_j - a_j^*) K(x_i, x_j) + \sum_{i=1}^n a_i^*(y_i - \epsilon) - \sum_{i=1}^n a_i(y_i - \epsilon) \quad (3)$$

$$\text{s. t. } \sum_{i=1}^n (a_i - a_i^*) = 0 \quad 0 \leq a_i, a_i^* \leq C$$

式中, $K(x_i, x_j) = \langle \theta(x_i) \cdot \theta(x_j) \rangle$ 为支持向量机的核函数, 核函数的引入使得高维空间中的内积运算可以通过输入空间中的函数来实现, 避免了维数灾难. 得到 SVM 回归函数

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) K(x_i, x) + b \quad (4)$$

其中, 当 $a_i - a_i^*$ 非零时对应的训练样本就是支持向量, b 通过 Karush-Kuhn-Tucker 条件计算出:

$$\begin{cases} b = y_i - \epsilon - \langle w \cdot \theta(x_i) \rangle, a_i, a_i^* \in [0, C] \\ b = y_i + \epsilon - \langle w \cdot \theta(x_i) \rangle, a_i, a_i^* \in [0, C] \end{cases} \quad (5)$$

文中选择径向基核函数, 其中径向基核参数 σ 、惩罚参数 C 、不敏感损失参数 ϵ 将是 SVM 训练参数.

3 用电量预测模型

3.1 构建训练数据

用电量影响因素主要有最高温度、最低温度、风

力、湿度、气象情况等. 训练样本中的输入数据为最高温度、最低温度、风力、湿度、气象情况, 输出数据为用电量. 由此训练样本训练的 SVM 预测模型描述为:

$$y_i = f(t_{hi}, t_{li}, f_i, s_i, q_i) \quad (6)$$

其中 t_{hi} 为最高温度, t_{li} 为最低温度, f_i 为风力, s_i 为湿度, q_i 为气象情况, y_i 为用电量.

3.2 粒子群优化支持向量机

支持向量机的预测精度与其惩罚参数 C 、RBF 核参数 σ 、不敏感损失参数 ϵ 的取值有着较大关系. 因此, 需采用优化算法在一定区域内搜索支持向量机各参数的最优组合, 进而获得具有较好预测性能的支持向量机.

粒子群优化算法是一种基于群体的智能寻优算法, 其基本思想源于对鸟类捕食行为的研究. PSO 初始化一群随机粒子(随机解), 然后通过迭代寻找最优解. 在每一次迭代中, 粒子通过跟踪个体极值 pbest 和全局极值 gbest 来更新自己, 个体极值 pbest 为粒子本身所找到的最优解; 全局极值 gbest 是指整个种群目前找到的最优解(图 1 所示). 粒子群优化算法的这些参数优化过程描述为:

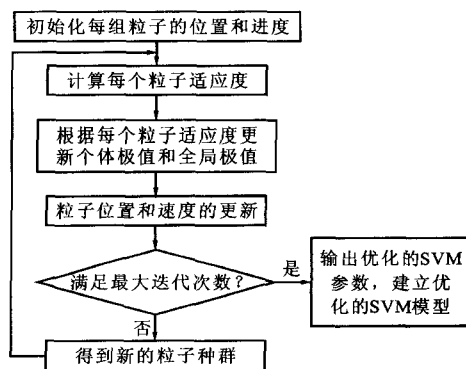


图 1 基于 PSO 的 SVM 参数优化流程

步骤 1 初始化粒子种群. 设定迭代次数、粒子个数, 并随机产生一组粒子的初始位置和速度.

步骤 2 适应度评估. 采用适应度函数 $\frac{1}{m} \sum_{i=1}^m \left| \frac{y_i - \hat{y}_i}{y_i} \right|$ 进行适应度评估. 其中, $y_i, \hat{y}_i$ 分别为实际值和预测值.

步骤 3 若粒子的当前适应度值优于先前的最好位置 pbest, 则将其替换为当前的最好位置 pbest; 若粒子的当前适应度值优于全局所经历的最好位置 gbest, 则将 gbest 替换为最好粒子的当前位置.

步骤 4 根据下式更新粒子的速度和位置.

$$v_{i+1} = w \cdot v_i + c_1 \cdot \text{rand} \cdot (p_i - x_i) + c_2 \cdot \text{rand} \cdot (p_{\text{gt}} - x_i)$$

(7)

$$x_{i+1} = x_i + v_{i+1}$$

(8)

式中,rand 为 0 到 1 之间的随机数; w 为惯性权值系数,用于控制前面速度对当前速度的影响; c_1 和 c_2 是加速因子. x_i 为粒子位置, v_i 为粒子速度; p_i 为个体最优解 pbest 和 p_{gt} 为全局最优解 gbest.

步骤 5 若达到迭代次数,则停止迭代,输出最优解;否则转至步骤 2.

4 实例分析

按照以上步骤,以江西省 2008 年 7 月~10 月的全省用电量数据以及相关气象特征数据,作为本文的实验数据,如表 1 所示.

表 1 江西省 2008 年 7 月~10 月的用电量数据

日期	最高温度/℃	最低温度/℃	风力/级	湿度/ g/m ³	气象情况	电量/万度
2008-7-3	35	27	3	10	晴	575.89
2008-7-4	35	27	2	40	晴	606.80
2008-7-5	35	28	3	10	晴	609.84
2008-7-7	35	26	2	40	多云	542.46
2008-7-8	33	26	3	70	多云	513.97
2008-7-9	30	25	2	70	中雨	494.85
2008-10-22	27	18	4	60	阴	517.52
2008-10-23	21	16	4	60	阴	494.07
2008-10-24	21	16	4	10	多云	496.42
2008-10-25	18	15	0	60	小雨	497.32
2008-10-26	17	14	3	90	小雨	486.03
2008-10-27	22	15	3	20	多云	495.10
2008-10-28	24	16	2	10	晴	496.52

为了消除量纲的影响,分别将各特征数据及用电量数据进行归一化,构建训练样本,并以此建立 PSVM 用电量预测模型.采用 BP 神经网络与 PSVM 进行对比分析,此 BP 神经网络具有 5 个输入节点、20 个隐节点、1 个输出节点. PSVM 与 BP 的用电量预测数据如图 2 所示.实验结果表明在用电量预测中,PSVM 比 BP 神经网络有着更高的预测

精度,能更有效的通过最高温度、最低温度、风力、湿度、气象情况等特征数据进行用电量的预测.

5 结束语

文中提出基于粒子群优化算法与支持向量机(PSVM)的用电量预测的新方法.其中,采用粒子群优化算法选取较优的支持向量机训练参数组合,得到性能较好的 SVM 用电量预测模型.采用某省 2008 年 7 月~10 月的用电量数据以及相关特征数据作为实验数据,并以最高温度、最低温度、风力、湿度、气象情况作为输入的特征数据,用电量为输出数据,构建 PSVM 实验样本.实验结果表明基于粒子群优化算法与支持向量机的用电量预测具有很高的预测精度,非常适合于用电量预测.

参考文献:

[1] 陈国初,刘军. 基于神经网络的中长期用电量预测模型[J]. 上海电机学院学报,2009,12(1):20-24.

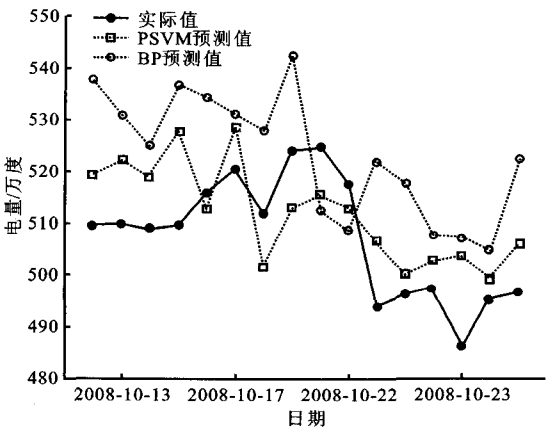


图 2 PSVM 与 BP 的用电量预测数据

(3,6,7));

(3) 最大资源利用率为 0.7838.

表 1 任务表

任务	参数	资源需求					计划开	计划结	前置	后置	位置
		r_1	r_2	r_3	r_4	r_5	始时间	束时间	任务	任务	(x,y)
1		3	0	2	2	1	1	3	—	2,3	(0,0)
2		0	0	4	4	0	3	6	1	4	(0,1)
3		1	4	0	0	0	3	7	1	5	(2,2)
4		1	0	4	0	3	6	9	2	6	(3,4)
5		0	4	7	0	0	7	10	3	6	(3,6)
6		0	6	0	12	0	10	16	—	4,5	(4,2)

表 2 资源表

节点	参数	资源能力					机动	执行任务	初始
		r_1	r_2	r_3	r_4	r_5	速度	的时间	位置
1		0	0	4	0	5	2	1	(3,4)
2		2	4	0	0	1	2	2	(3,5)
3		0	2	0	4	0	2	1	(0,2)
4		0	4	5	3	0	1	2	(5,6)
5		1	5	0	2	0	1	1	(2,2)
6		0	6	0	0	2	2	2	(4,2)
7		0	0	0	8	0	1	1	(4,5)
8		0	0	5	0	0	2	1	(3,0)
9		0	0	0	5	5	1	2	(6,8)
10		3	0	2	4	2	2	1	(6,0)

在基本遗传算法下执行需要迭代 97 次,因此可以看出该算法大大缩小了搜索空间,加快了收敛速度. 由于该算法综合考虑了协作资源和任务之间的通信量,任务的执行时间,资源利用率等参数,具有实际的应用价值.

(上接第 149 页)

[2] 田有文,王文涛,王思宇. 基于 BP 神经网络的农村用电量预测的研究[J]. 沈阳农业大学学报,2005,36(4): 506—507.

[3] Bo Jin, Tang Y C, Zhang Yanqing. Support vector machines with genetic fuzzy feature transformation for bio-medical data classification[J]. Information Sciences, 2007, 177(2):476—489.

[4] Matteo Pardo, Giorgio Sberveglieri. Classification of e-lectronic nose data with support vector machines[J]. Sensors and Actuators B: Chemical, 2005, 107(2):730—737.

6 结束语

文中在进行资源分配时结合网格资源的特点,将所有异构资源统一描述和对待,扩展了资源的概念,并针对基本遗传算法在进行资源优化配置时的不足,对遗传算法进行了改进. 从测试算例中可以看出,该算法能在获得最优解的条件下提高搜索能力、加快收敛速度,可以有效的解决网格资源优化配置问题.

参考文献:

[1] Fo ster I, Kesselman C, Tuecke S. The anatomy of the grid: enabling scalable virtual organizations[J]. International Journal of Supercomputer Applications, 2001, 15(3): 200—222.

[2] 罗宇平. 基于 Min2Min 改进后的网格调度算法[J]. 微电子学与计算机,2009,26(3):86—88.

[3] Alhusaini A H, Prasanna V K, Raghavendra C S. A framework for mapping with resource coallocation in heterogeneous computing systems[C]// Proc. of 9th Heterogeneous Computing Workshop. Mexico, Cancun, 2000:273—286.

[4] 王小平,曹立明. 遗传算法理论、应用与软件实现[M]. 西安:西安交通大学出版社,2002.

[5] 何琨,赵勇. 网格环境下资源调度问题的统一建模与分析[J]. 华中科技大学学报:自然科学版,2006,34(3): 35—38.

作者简介:

杨庆娜 女,(1985—),硕士研究生. 研究方向为网格计算.

李定主 男,(1965—),研究生导师,研究员级高级工程师,教授. 研究方向为多媒体与计算机仿真、网络技术应用.

[5] 孙向军,赵斯强,严宗睿. 基于粒子群优化的反潜搜索研究[J]. 微电子学与计算机,2008,25(10):91—93.

[6] 韦杏琼,周永权. 基于粒子群算法的数值积分方法研究[J]. 微电子学与计算机,2009,26(7):117—119.

作者简介:

曾勍炜 男,(1973—),硕士,副研究员. 研究方向为应用软件开发.

徐知海 男,(1975—),硕士,助理研究员. 研究方向为应用软件开发.

吴 键 男,(1973—),高级工程师. 研究方向为电力调度.