

基于支持向量机的网格负载信息预测模型

黄 谦,肖 侬,刘 波

(国防科学技术大学 并行与分布重点实验室,湖南 长沙 410073)

摘 要:提出了采用小波分析和最小二乘支持向量机(LS-SVM)混合模型对网格负载信息进行预测。该模型首先基于小波多分辨率分析对非平稳的网格负载样本做序列分解,得到不同尺度下的负载分量,然后利用 LS-SVM 对不同尺度的分量进行预测,最后通过对各分量预测信息进行重构得到相应的预测值。实验结果表明,使用本模型进行短期负荷预测比传统小波神经网络方法可以获得更好的预测精度。

关键词:网格预测;最小二乘支持向量机;多分辨率分析;小波变换

中图分类号:TP18

文献标识码:A

文章编号:1673-629X(2007)06-0032-04

Grid Load Forecasting Based on Least Squares Support Vector Machine

HUANG Qian, XIAO Nong, LIU Bo

(National Lab. for Parallel and Distributed Procession, National Univ. of Defense Techn., Changsha 410073, China)

Abstract: An algorithm for grid load forecasting based on wavelet analysis and least squares support vector machine, was introduced. Begin with discussion of decomposition of serial signal of grid load and then get the forecasts of each sub-signals by LS-SVM. The third step is combination of these forecasts. This method was successfully achieved on forecasting of memory load. The experiment result shows that it can get better forecasting accuracy to traditional wavelet neural network in short-term load forecasting.

Key words: grid forecasting; LS-SVM; multi-scale prediction; wavelet transform

0 引 言

在网格计算^[1]系统中,任务调度与负载平衡是十分重要的功能。由于资源的共享造成系统行为和系统性能经常变化,调度器通常会对实时网格负载做出反应。在调度器分派任务之前,会根据资源状态信息来确定运行该任务合适的机器。网格的基础设施平台比如 Globus,提供了资源目录,该目录记录了资源的静态信息(如 CPU 的类型、时钟频率等)。因此,选择最好机器的最简单的方法就是去查询所有有效机器的资源目录,从中选择当前负载最低的一个。但是,加入网格的资源具有自治性和动态性,资源所有者可以进行加入、撤出和改变其参数等操作。在这样的环境下,需要调度器对资源的有效性、内存^[2]的利用率、网络带宽及延迟等性能做出比较精确的预测。因此,计算网格体系结构中,对有效资源性能的准确预测是实施高效的

调度工作、提高系统的执行性能的保证。

通过对单机、集群以及网格的性能数据观测,表明单机的预测意义不是太明显,但是对于集群以及作为信息基础设施的网格,监控的更为有效。因为集群以及网格,性能数据更有规律性及稳定性。造成这种现象的原因是单机在接受任务及完成任务时性能数据变化较大,同时系统管理的噪声不可忽略。这种现象在网格环境中不会出现。网格作为基础设施,计算资源相当丰富,网格中的任务也多以大任务为主,因而其性能数据变化更具有周期性和稳定性,所以对网格进行预测比对单机进行预测更加准确有效。

常用的预测方法有相关分析法、时间序列法、人工神经网络等^[3-5]。人工神经网络具有良好的非线性逼近能力和自适应自学习的功能,在短期负荷预测上得到了广泛的应用^[4,6,7],但是存在需要确定网络结构、要求的训练样本大、过学习和在非线性优化问题中的局部极小等问题。基于小样本统计学习理论的支持向量机(SVM)在小样本、非线性、高维问题及泛化能力方面表现突出,而且具有成熟的理论基础,它在形式上类似于多层前向网络,可自动解决网络结构问题。从 SVM 引申出的支持向量回归(SVR)已被广泛使用在

收稿日期:2006-08-31

基金项目:国家自然科学基金资助项目(60573135,0-5223-22);

国家重大基础研究 973 资助项目(2003CB317008)

作者简介:黄 谦(1979-),男,硕士研究生,研究方向为网格监控;肖 侬,博士,教授,研究方向为网格计算、普适计算、海量信息处理。

时间序列预测领域,并取得了较好的预测效果^[8]。

小波分析是处理非平稳信号的有力工具,具有良好的时频局域化性质。与小波分析相结合的神经网络(小波网络)也广泛地用于时间序列预测^[6],如水文数据预测^[5]、短期电力负荷预测^[8],但这些小波网络采用的是传统的人工神经网络技术,其预测精度存在人工神经网络固有的缺点,而且效率较低。网络系统负荷序列既具有波动性又具有特殊的周期性,可看作多个不同频率分量的叠加。每个分量呈近似周期变化,具有相似的频率特性和一致的变化规律,也具有更强的可预测性。因此,在对负荷进行频域分解的基础上,根据各频率分量特点进行建模是提高预测精度的有效途径。完成此任务的算法首选小波变换。小波变换的时频窗口具有可调性,当尺度较大时,时频窗口的时域部分较宽、分析频率低,适合做概貌观察;当尺度较小时窗口的时域部分较窄而分析频率高,适合做细节观察。这种逐级分析方法称为多分辨率分析。文中提出了一种基于小波支持向量机的预测模型,在网格负载信息预测中获得了较为理想的预测效果。

1 小波变换及多分辨率分析

小波变换^[4]的主要思想就是选择合适的小波基函数 $\phi(t)$,通过在时间轴上的位移、缩放和幅度的变化产生一系列的派生小波,如式(1)所示。用系列小波与要分析的信号进行时间轴上的平移比较,获得用以表征信号与小波相似程度的小波系数。

$$\phi_{a,b}(t) = |a|^{-\frac{1}{2}} \phi\left(\frac{t-b}{a}\right) \quad (1)$$

式中 a 为时间尺度; b 为时间上的位移。对于信号函数 $f(t)$,其连续小波变换定义为:

$$W_f(a,b) = \langle f, \phi_{a,b} \rangle = |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(t) \overline{\phi\left(\frac{t-b}{a}\right)} dt \quad (2)$$

其逆变换方程为:

$$f(t) = \frac{1}{C_\phi} \iint_{\mathbb{R}^2} W_f(a,b) \phi_{a,b}(x) \frac{da db}{a^2} \quad (3)$$

其中 C_ϕ 满足下式:

$$\iint_{\mathbb{R}^2} W_f(a,b) \overline{W_f(a,b)} \frac{da db}{a^2} = C_\phi \langle f, f \rangle \quad (4)$$

式中 $W_f(a,b)$ 为原信号 $f(t)$ 在时间 b 处包含具有尺度 a 的小波函数分量的数目,表示信号 $f(t)$ 与小波函数 $\phi(\frac{t-b}{a})$ 相关联的程度。

小波多分辨率分析的主要思想是:用不同的分辨率来逐级逼近待分析的函数 $f(t)$,利用多分辨率分析可以将函数 $f(t)$ 投影到相互正交的函数空间中。尺度

相同的小波函数或尺度函数构成了一个函数空间的基,而不同函数空间的基的尺度不同,不同的尺度则表征了不同的频率特性。通过不断地进行投影,可逐步得到原函数在其任意子空间中的分量,从而可以通过区分粗、细信号来分析信号。利用滤波器组的思想方法可以方便地实现信号的分解,这里把信号分为近似部分和细节部分,近似部分是指信号中可以投影到大尺度(即尺度值较大)空间中的低频分量;而细节部分是指信号中与小尺度空间对应的高频分量。

负荷信号 S 经过抽样后得到近似分量小波系数和细节分量小波系数; $h(n), l(n)$ 分别为高通滤波器和低通滤波器。这种分解过程可迭代进行,即对分解后的近似分量再进行分解,从中抽取出细节分量。经过几次迭代后原序列就被分解为由一个近似分量和一组细节分量组成的树状结构。

综上所述,可得到基于小波及最小二乘支持向量机的网格负载信息预测算法的流程如图1所示:负载信号 S 经过分解后得到近似分量小波系数 c_1 和细节分量小波系数 d_1 ;这种分解可迭代进行,即对分解后的近似分量 c_1 再进行分解,从中抽取出细节分量。数次迭代后原信号序列就被分解为由一个近似分量和一组细节分量组成的树状结构。

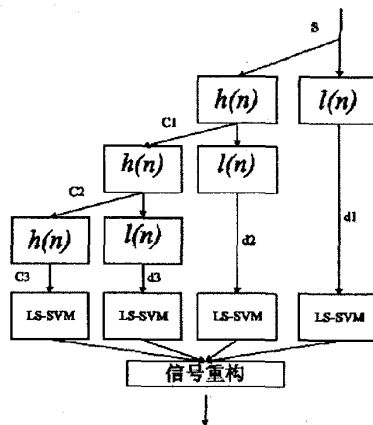


图1 基于最小二乘支持向量机的
网格负载预测算法流程

在此需注意两个问题:

(1)分解级数既不能太小也不能太大。太小则不能有效地将原信号中具有不同频率特征的分量分离出来;过大则需要较多的模型对分解后的各分量进行预测,各个模型都会引入一定的误差,从而导致最终预测误差变大。

(2)小波基的选择。选择不同的小波基将得到不同的分解分量。通过选用不同的小波基和分解级数进行实验,笔者最终采用Db3小波基及相应的滤波器对

系统负荷序列进行 4 级分解。

2 LS-SVM

支持向量机^[9]是根据统计学习理论提出的一种新的机器学习方法,在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势。支持向量机从本质上讲是一种前向神经网络,根据结构风险最小化准则,在使训练样本分类误差极小化的前提下,尽量提高分类器的泛化推广能力。从实施的角度,训练支持向量机的核心思想等价于求解一个线性约束的二次规划问题,从而构造一个超平面作为决策平面,使得特征空间中两类模式之间的距离最大,而且它能保证得到的解为全局最优解,并巧妙地利用原空间的核函数取代高维特征空间中的点积运算^[8]。

SVM 的基本思想是:通过一个非线性映射,把输入空间的数据映射到一个高维特征空间中去,然后在此空间中应用结构风险最小化原则构造最优决策函数,并巧妙地利用原空间的核函数取代高维特征空间中的点积运算^[8]。

假设给定训练样本 $\{x_i, y_i\} (i = 1, 2, \dots, n)$, n 为样本总数, $x_i \in R^d$, d 为 R 空间的维数, $y_i \in R$ 。用非线性映射 $\varphi(\cdot)$ 将样本从原空间映射到高维特征空间 F 中, F 为 K 维空间(K 可能是无穷大)。在这个空间中构造如下最优线性回归函数:

$$f(x) = w^T \varphi(x) + b \quad (5)$$

式中 w 为权向量, $w \in R^k$; b 为常数, $b \in R$ 。

根据 SRM 准则,优化目标和约束条件分别为:

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

$$\text{s.t. } y_i = w^T \varphi(x_i) + b + \xi_i \quad (i = 1, 2, \dots, n)$$

式中 $\xi_i \geq 0$ 为松弛因子; $C > 0$ 为惩罚参数,控制对超出误差的样本的惩罚程度。

将上述问题表示成拉格朗日乘子式:

$$L = \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_i \xi_i^2 - \sum_i \alpha_i [w^T \varphi(x_i) + b + \xi_i - y_i] \quad (6)$$

由 Kuhn-Tucker 条件有:

$$\nabla_w = w - \sum_i \alpha_i \varphi(x_i) = 0, \quad \frac{\partial L}{\partial b} = \sum_i \alpha_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = C \xi_i - \alpha_i = 0, \quad \frac{\partial L}{\partial \alpha_i} = w^T \varphi(x_i) + b + \xi_i - y_i = 0$$

对于 $i = 1, 2, \dots, n$, 消去 w 和 ξ 可得到如下线性方程组:

$$\begin{bmatrix} 0 & e \\ e & K + C_e^{-1} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (7)$$

式中:

$$Y = [d_1, d_2, \dots, d_n]^T; \quad \alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T;$$

$e = [1, 1, \dots, 1]^T$; $K(x_i, x_j) = K_{ij} = \varphi(x_i) \varphi(x_j)^T$
原回归问题可以通过最小二乘法求解, 所得回归函数为:

$$y = \sum_i \alpha_i K(x_i, x_j) + b \quad (8)$$

3 网格负载信息预测模型

使用小波 LS-SVM 方法进行网格负载信息预测通常分为以下三个阶段:

(1) 首先利用小波多分辨率分析将历史样本序列进行小波多尺度分解, 得到小波系数和尺度系数 d_1, d_2, d_3, c_3 四个分量。

(2) 将 t 时刻的小波分解序列系数作为 LS-SVM 的输入, $t + T$ (T 为预测步) 时刻的某个小波分解系数作为输出来建立不同的模型, 并构造相应的支持向量机模型对 $t + T$ 时刻的各小波分解系数进行预测。

(3) 利用小波重构获得 $t + T$ 时刻的预测值。

本节的样本数据是从一个中国国家网格监控系统 GridEye 中采集网格负载信息, 每三分钟采集一个样本, 持续一天, 总共 480 个样本。选取第 150 ~ 250 个数据点为历史数据进行小波多尺度分解, 作为下一步预测的输入, 预测步 $T = 30$, 进行实验模拟。

图 2 为所选取信号的分层结果, 这段信号 4 层细节系数中, 第一层的细节系数波动非常小, 可以把它看作是由于网格系统管理与偶发事件所引起的干扰。第二层细节系数中, 绝对值较高的系数给出了中部波峰的起始和结束位置。它可以定位出局部最小值点的位置。第三层细节系数与原信号中形状吻合的很好。第四层细节系数没有给出什么有用的信息。因此第四层细节系数提供了最重要的信息。无论从数量或质量的角度, 它给出了我们关心的范围内的信号的总体形状信息。

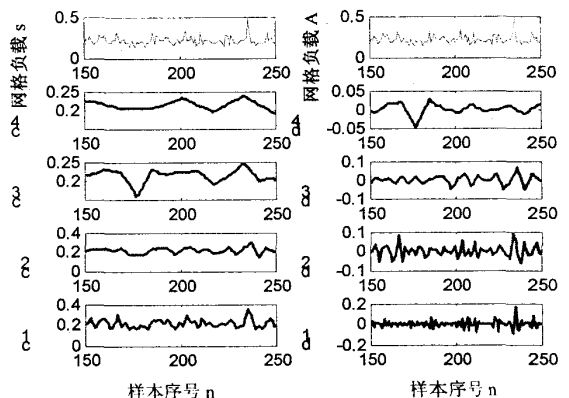
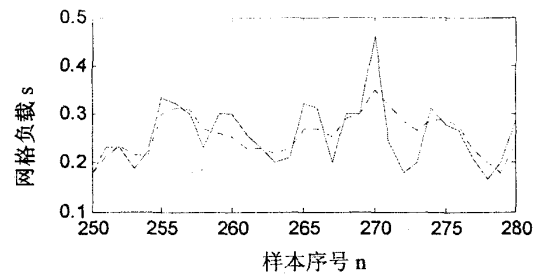


图 2 CNGRID 环境下某节点平均负载数据的小波分解

图 3 为预测步 $T = 30$ 时的预测结果曲线。



(实线为实际采集的网络负载数据, 虚线为预测结果)

图 3 $T = 30$ 步预测结果曲线

4 总 结

基于小波多分辨率分析对中国国家网络负载信号进行分解,然后对其各个主要分量建立基于最小二乘支持向量机的预测模型,最后由预测信号重构负荷信号,充分利用了 LS-SVM 求解速度快、模型参数确定方便的优点。实验结果表明该方法能得到有效的预测结果,而且提高了学习精度,在网格监控系统所获得的信息的基础上,对网格资源如 cpu、memory 等资源的性能值做出预测,实现网格资源负载信息的实时采集、处理和分析、快速高效整合和协调决策,有效地提高了网格中调度组件的调度性能。

参考文献:

[1] Foster I, Steven C K, Ueche T. The Anatomy of the Grid – Enabling Scalable Virtual Organizations[J]. Internet Journal of High Performance Computing Applications, 2001, 15: 220 – 222.

[2] Chu Rui, Xiao Nong, Liu Yuhao, et al. A Distributed Paging RAM Grid Systemfor Wide – area Memory Sharing[C]// In 20th International Parallel and Distributed Processing Symposium. 2006. (IEEE IPDPS). Greece:[s. n.],2006.

[3] Yang Yanxi, Zheng Gang, Liu Ding. BP – GA mixed algorithms for short – term load forecasting[C]// Proceedings of 2001 International Conference on Information Technology and Information Network. Beijing, China:[s. n.],2001.

[4] 徐军华,刘天琪. 基于小波分解和人工神经网络的短期负荷预测[J]. 电网技术,2004,28(8):30 – 33.

[5] 李贤彬,丁 晶,李后强. 基于子波变换序列的人工神经网络组合预测[J]. 水利学报,1992(2):1 – 4.

[6] 张国忠. 应用人工神经网络预电力负荷[J]. 电力自动化设备,2002(5):20 – 21.

[7] 刘 涵,刘 丁,郑 岗,等. 基于最小二乘支持向量机的天然气负荷预测[J]. 化工学报,2004,55(5):828 – 832.

[8] 张伏生,汪 鸿,韩 伟,等. 基于偏最小二乘回归分析的短期负荷预测[J]. 电网技术,2003,27(3):36 – 40.

[9] 李元诚,方廷健,于尔铿. 短期负荷预测的支持向量机方法研究[J]. 中国电机工程学报,2003,23(6):55 – 59.

(上接第 31 页)

words= words.substring(0,l)+“...”; //用 substring 获取整个字符串的子串

return words; //返回 words

(Ⅲ) 用户网页的生成与显示。

基于开发考虑,在数据库的设计中为每个用户赋予一个不同的 t_id,用来唯一标识每位注册用户。通过 JSP 的 request 语句中的 getParameter,从前台页面获得该用户独一无二的 t_id。图 4 中,给出了‘网页动态生成系统’中的数据库节选字段图。站点中为每位用户提供了中、英文两种用户网页版块,两个 name 字段(t_name 和 t_ename)分别映射着一位用户的两列信息,一列存储中文信息,一列英文信息。在数据库中定义了一个布尔值字段 t_type,它的值只有 0 或者 1。0 代表此用户的中文网页,1 代表英文网页。当 t_type 字段值为 1 时,表示使用的是英文版块,若值是 0,则表示使用中文版块。有了布尔值 t_type 字段的设置,能方便地在中英文版块间进行切换。

络编程中也显得越来越重要。由于 JSP 基于强大的 Java 语言,具有极强的扩展能力,良好的收缩性,以及与平台无关的开发特性,在根据 Java 平台构建动态商务网站成为主流的今天,JSP 有着其他技术所不具备的优势。

t_name	t_ename	t_type
用户 a	Jin	1
用户 b	Huang	1
用户 c	Shu	0
用户 d	Chen	0

图 4 数据库节选表

参考文献:

[1] Duffey K. JSP 站点设计编程指南[M]. 王 军等译. 北京: 电子工业出版社,2002.

[2] 高 成. JSP 动态网站建设[M]. 北京:国防工业出版社, 2002.

[3] Horstmann C S, Cornell G. Core Java2[M]. 5th ed. 北京:机械工业出版社,2002.

[4] j2se1. 4.2 API 帮助文件[EB/OL]. 2003 – 07 – 22. http:// www.java. sun. com.

[5] Wutka M. JSP 和 Servlet 程序设计使用专辑[M]. 程显华等译. 北京:机械工业出版社,2002.

4 结束语

随着 Java 技术的日益成熟和流行,JSP 技术在网