

1-4-2019

Modeling and Simulation on Delay of Central Heating System for Temperature Prediction

Sun Tao

Dalian University of Technology, Dalian 116024, China;

Shaoxiong Ji

Dalian University of Technology, Dalian 116024, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Modeling and Simulation on Delay of Central Heating System for Temperature Prediction

Abstract

Abstract: Heating prediction is essential for energy saving. However, it is difficult to some extent because the central heating system (CHS) is nonlinear with large-scaled delay and other characteristics. Based on physical laws of the CHS, *a novel model combining pruning algorithm with the least mean square* is proposed to solve the problem of delay and predict the temperature. The model with the delay taking into consideration *conforms to the law of thermal transmission*. The pruning algorithm *lowers the computational complexity*. Simulations using gradient descent least squared regression (LSR), support vector machine (SVM), least-squared support vector machine (LS-SVM) and BP neural network are put forward to evaluate the model and the results show that the model has low average relative error and mean squared error; and the model is simple and easy to calculate.

Keywords

central heating system, prediction of secondary network, least mean square, pruning algorithm

Recommended Citation

Sun Tao, Ji Shaoxiong. Modeling and Simulation on Delay of Central Heating System for Temperature Prediction[J]. Journal of System Simulation, 2018, 30(4): 1328-1336.

基于集中供热时延的温度预测模型及仿真

孙焘, 季少雄

(大连理工大学, 辽宁 大连 116024)

摘要: 供热预测对实现供热节能有重要意义, 然而由于供热系统具有非线性、复杂性和大时滞性等特征, 预测存在一定难度, 为了解决供热预测中的时延建模、求解和温度预测问题, 基于供热系统的传热物理规律建立优化模型, 提出最小二乘意义上的时延求解剪枝算法, 并进行温度预测。将供热系统的时延考虑在内的模型符合热传递的物理规律, 剪枝降低了时延求解的计算复杂度。通过仿真实验对比与分析证明, 与梯度下降最小二乘回归、支持向量机、最小二乘支持向量机和BP神经网络等模型相比, 本模型具有较低的平均相对误差和均方误差, 且模型简单, 计算便捷。

关键词: 集中供热系统; 二次网预测; 最小二乘原理; 剪枝算法

中图分类号: TP391.9 文献标识码: A 文章编号: 1004-731X(2018)04-1328-09

DOI: 10.16182/j.issn1004731x.joss.201804015

Modeling and Simulation on Delay of Central Heating System for Temperature Prediction

Sun Tao, Ji Shaoxiong

(Dalian University of Technology, Dalian 116024, China)

Abstract: Heating prediction is essential for energy saving. However, it is difficult to some extent because the central heating system (CHS) is nonlinear with large-scaled delay and other characteristics. Based on physical laws of the CHS, a novel model combining pruning algorithm with the least mean square is proposed to solve the problem of delay and predict the temperature. The model with the delay taking into consideration conforms to the law of thermal transmission. The pruning algorithm lowers the computational complexity. Simulations using gradient descent least squared regression (LSR), support vector machine (SVM), least-squared support vector machine (LS-SVM) and BP neural network are put forward to evaluate the model and the results show that the model has low average relative error and mean squared error; and the model is simple and easy to calculate.

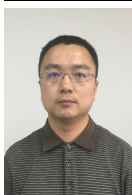
Keywords: central heating system; prediction of secondary network; least mean square; pruning algorithm

引言

集中供热是我国北方冬季供热的最主要的一种形式, 由于供热系统的复杂性, 供暖的质量取决于多个因素, 为实现供热节能, 需要对供热系

统进行预测来指导热网的运行与调节^{[1]-[3]}。随着集中供热的推广与发展, 供热公司和学者对供热预测开展一定的研究, 然而由于供热系统本身具有非线性、内部关联性强和影响因素多等特质, 长距离传输管网中有大时延, 以及供热数据采集存在噪声与误差, 供热预测工作具有一定难度。

目前, 供热预测研究主要集中于热负荷和温度。前者指在某一室外温度下, 达到热用户所需室内温度的前提下, 供热系统在单位时间内提供



收稿日期: 2015-05-19 修回日期: 2016-12-20;
基金项目: 辽宁省科学技术计划(2014104011);
作者简介: 孙焘(1975-), 男, 辽宁大连, 博士, 副教授, 研究方向为数据挖掘; 季少雄(1991-), 男, 福建南平, 硕士生, 研究方向为数据挖掘。

<http://www.china-simulation.com>

• 1328 •

的热量^[1]。温度预测包括室内温度和二次网回水温度预测。

传统的预测方法认为负荷仅与室外温度有关, 文翰以室外温度为自变量, 供热量为因变量, 对供热量和室外温度进行最小二乘拟合^[4], 采用线性回归, 忽略了供热系统的时延特性、复杂性与非线性。随机优化和机器学习等方法被应用于供热系统, 李琦和李梅提出一种基于 PSO 算法的非线性模型, 应用于热网系统中一次网热量产生和热量分配上, 进行热网工况的预测和控制^[5]。高丙坤、李阳和许明子将 PSO 算法与组合预测结合进行供热预测, 获得了高于单项预测 40% 精准确度的结果^[6]。张永明、陈烈和齐维贵利用支持向量机模型进行区间回归, 采取了一种不同于点预报的策略^[7]。

人工神经网络以其强大的非线性映射能力, 也被应用于供热预测领域。王美萍、田琦和张佼采用模糊耦合神经网络对热网的热流量进行预测, 获得了比小波神经网络更小的误差^[8]。冯敬芳、谢慕君和步伟明等人以一、二次网供回水温度、流量和室外温度作为模型输入, 采用 LM、Elman^[9]和 RBF^[10]三种神经网络预测回水温度^[11]。神经网络尽管具有高度的非线性, 但是神经网络模型结构复杂, 且存在对于学习收敛慢和容易过拟合等缺点。

供热预测的方法可分为基于物理模型的预测方法和基于数据模型的方法, 其中, 数据驱动的预测方法采用黑箱模型, 运用抽象的数学模型以及机器学习算法, 很少建立在供热物理模型的基础上。实际热负荷可能还与房屋的围护结构特性、传热规律和管网流量分布等物理规律相关^[1], 此外, 实际的供热系统存在时延, 供热站传输线路中的时延通常为 1~3 个小时。本文提出一种基于热网传热规律与供热时延的最小二乘剪枝求解算法, 将传热物理规律和供热系统时延纳入模型, 以最小化预测误差为优化目标进行时延求解, 通过剪枝降低求解的次数, 降低了算法的时间复杂度, 最后对室内温度和二次网回水温度进行回归预测。

1 热传递物理模型

热网中热量耗散主要集中于供暖系统的末端装置, 末端因为存在温差主要有两个散热过程: 散热器向室内的热耗散和室内到室外热耗散。无遮挡、敞开设的散热器有形如 $Q = f(\Delta t)$ 的传热关系式^[1], 即:

$$Q_{in} = k_1 \left(\frac{T_g + T_h}{2} - T_{in} \right)^{\alpha_1} \quad (1)$$

其中, Q_{in} 为单位时间内在一定散热面积下的散热器传递到室内空气的散热量, k_1 为实验确定的散热器传热的系数, 与散热器的传热系数、散热面积有关, T_g 为二次网的供水温度, T_h 为二次网的回水温度, T_{in} 为室内温度, α_1 为系数。根据比热容公式:

$$Q = cm\Delta t$$

其中, Q 为放出的热量, c 为比热容, m 为质量, Δt 为温度变化量, 由比热容公式导出散热器向室内空气的热量变化公式:

$$Q_{in} = gc_1(T_g - T_h) \quad (2)$$

其中 g 为二次网的流量(即单位时间内流经管道截面的热水的质量), c_1 为水的比热容, 常压下水的定压比热容为 $4.2 \times 10^3 \text{ J}/(\text{kg} \cdot ^\circ\text{C})$ 。

室内与室外存在温度差, 室内向室外传热, 建筑及房间存在围护结构传热耗热量, 其关系式为:

$$Q_{out} = k_2(T_{in} - T_{out})^{\alpha_2} \quad (3)$$

其中, Q_{out} 为单位时间内室内向室外传递的热量, 即围护结构传热耗热量, k_2 为实验确定的房间传热的系数, 与房屋的传热系数、房屋散热面积和房屋材料有关, T_{out} 为室外温度, α_2 为系数。此外, 建筑及房屋内外还存在以下热量关系:

$$Q_{in} - Q_{out} = k_3 Q_{in} \quad (4)$$

其中, k_3 为实验确定的系数, 与围护结构的比热容、围护结构及散热器的散热面积等因素有关。

2 基于时延的最小二乘预测模型

2.1 模型建立

在集中供热系统的二次网模型中, 输入数据为:

室内温度 $T_{in} = [T_{in}^{(1)}, \dots, T_{in}^{(n)}] \in R^n$ ，室外温度 $T_{out} = [T_{out}^{(1)}, \dots, T_{out}^{(n)}] \in R^n$ ，供水温度 $T_g = [T_g^{(1)}, \dots, T_g^{(n)}] \in R^n$ ，回水温度 $T_h = [T_h^{(1)}, \dots, T_h^{(n)}] \in R^n$ ，瞬时流量 $G = [G^{(1)}, \dots, G^{(n)}] \in R^n$ ， n 为向量长度。传输过程中的存在时延，时延系数为 $\tau \in N^+$ ，二次网供水温度以 τ 为参数，记作 $T_g(\tau) = [T_g^{(1)}(\tau), \dots, T_g^{(n)}(\tau)] \in R^n$ 。根据热传递物理公式以最小二乘准则优化参数，分别对室内温度和二次网回水温度进行回归预测。

在室内温度预测中，根据热传递物理模型，室内温度与二次网供水温度、回水温度和流量有关，存在以下函数关系：

$$T_{in}^{(i)} = f(T_g^{(i)}, T_h^{(i)}, G^{(i)}; \alpha_1, k_1, c_1, \tau), \quad i=1, 2, \dots, n$$

其中，散热器传热系数 k_1 、实验系数 α_1 、水的比热容 c_1 和时延系数 τ 待确定。求参数 α_1, k_1, c_1, τ ，使其为式(6)损失函数极小化问题的解：

$$\min_{\alpha_1, k_1, c_1, \tau} J_{in}(\alpha_1, k_1, c_1, \tau) = \sum_{i=1}^n (T_{in}^{(i)} - \hat{T}_{in}^{(i)})^2 = \sum_{i=1}^n (T_{in}^{(i)} - f(T_g^{(i)}, T_h^{(i)}, G^{(i)}; \alpha_1, k_1, c_1, \tau))^2 \quad (5)$$

其中， $T_{in}^{(i)}, \hat{T}_{in}^{(i)}, T_g^{(i)}, T_h^{(i)}, G^{(i)}$ 为对应向量的第 i 个元素。根据式(1)(2)联立，化简并整理即得：

$$\hat{T}_{in}^{(i)} = f(T_g^{(i)}, T_h^{(i)}, G^{(i)}; \alpha_1, k_1, c_1, \tau) = \frac{T_g^{(i)}(\tau) + T_h^{(i)}}{2} - \exp\left(\frac{\ln(c_1 G^{(i)}(T_g^{(i)}(\tau) - T_h^{(i)})/k_1)}{\alpha_1}\right) \quad (6)$$

其中， $i=1, 2, \dots, n$ 。供热系统具有非线性，直接拟合参数比较困难，通过线性变换可降低拟合难度。将式(6)变换为等价的线性形式：

$$Y_{in}^{(i)}(\tau) = a_1 X_{in}^{(i)}(\tau) + b_1, \quad i=1, 2, \dots, n \quad (7)$$

其中， $X_{in}^{(i)}(\tau), Y_{in}^{(i)}(\tau), i=1, \dots, n$ 分别为室温线性预测模型的自变量和应变量，均以时延系数 τ 为参变量，参数 $a_1 = \alpha_1$ ， $b_1 = \ln k_1 - \ln c_1$ 。数据变换为：

$$\ln\left(\frac{T_g^{(i)}(\tau) + T_h^{(i)}}{2} - T_{in}^{(i)}\right) = X_{in}^{(i)}(\tau), \quad i=1, 2, \dots, n \quad (8)$$

$$\ln(G^{(i)}(T_g^{(i)}(\tau) - T_h^{(i)})) = Y_{in}^{(i)}(\tau), \quad i=1, 2, \dots, n \quad (9)$$

经变换后，室内温度预测的损失函数即式(5)可等价求参数 a_1, b_1, τ ，使其为式(10)损失函数极小化

问题的解：

$$\min_{a_1, b_1, \tau} J_{in}(a_1, b_1, \tau) = \sum_{i=1}^n \left(Y_{in}^{(i)}(\tau) - \hat{Y}_{in}^{(i)}(\tau) \right)^2 = \sum_{i=1}^n \left(Y_{in}^{(i)}(\tau) - a_1 X_{in}^{(i)}(\tau) - b_1 \right)^2 \quad (10)$$

在二次网回水温度预测中，根据传热模型，二次网回水温度与二次网供水温度、室内外温度和流量有关，存在以下函数关系：

$$T_h^{(i)} = g(T_g^{(i)}, T_{in}^{(i)}, T_{out}^{(i)}, G^{(i)}; c_1, \alpha_2, k_2, k_3, \tau), \quad i=1, 2, \dots, n$$

其中，水的比热容 c_1 、实验系数 α_2 、房间传热系数 k_2 、实验系数 k_3 和时延系数 τ 待确定。求参数 $c_1, \alpha_2, k_2, k_3, \tau$ ，使其为式(11)损失函数最小化问题的解：

$$\min_{c_1, \alpha_2, k_2, k_3, \tau} J_h(c_1, \alpha_2, k_2, k_3, \tau) = \sum_{i=1}^n \left(T_h^{(i)} - \hat{T}_h^{(i)} \right)^2 = \sum_{i=1}^n \left(T_h^{(i)} - g(T_g^{(i)}, T_{in}^{(i)}, T_{out}^{(i)}, G^{(i)}; c_1, \alpha_2, k_2, k_3, \tau) \right)^2 \quad (11)$$

其中， $T_h^{(i)}, \hat{T}_h^{(i)}, T_{out}^{(i)}$ 表示对应向量的第 i 个元素。

根据式(2)(3)(4)联立，化简并整理即得：

$$\hat{T}_h^{(i)} = g(T_g^{(i)}, T_{in}^{(i)}, T_{out}^{(i)}, G^{(i)}; c_1, \alpha_2, k_2, k_3, \tau) = T_g^{(i)}(\tau) - \exp\left(\ln\left(\frac{k_2(T_{in}^{(i)} - T_{out}^{(i)})\alpha_2}{c_1(1-k_3)}\right)\right) / G^{(i)} \quad (12)$$

同理，将式(12)可变换为等价的线性形式：

$$Y_h^{(i)}(\tau) = a_2 X_h^{(i)} + b_2, \quad i=1, 2, \dots, n \quad (13)$$

其中， $X_h^{(i)}, Y_h^{(i)}(\tau), i=1, 2, \dots, n$ 分别为回水温度预测自变量和应变量， $Y_h^{(i)}(\tau)$ 以时延系数 τ 为参变量，参数 $a_2 = \alpha_2$ ， $b_2 = \ln k_2 - \ln(c_1(1-k_3))$ 。数据变换如下：

$$\ln(T_{in}^{(i)} - T_{out}^{(i)}) = X_h^{(i)}, \quad i=1, 2, \dots, n \quad (14)$$

$$\ln(G^{(i)}(T_g^{(i)}(\tau) - T_h^{(i)})) = Y_h^{(i)}(\tau), \quad i=1, 2, \dots, n \quad (15)$$

损失函数即式(11)经变换等价求参数 a_2, b_2, τ ，使其为式(16)损失函数极小化问题的解：

$$\min_{a_2, b_2, \tau} J_h(a_2, b_2, \tau) = \sum_{i=1}^n \left(Y_h^{(i)}(\tau) - \hat{Y}_h^{(i)}(\tau) \right)^2 = \sum_{i=1}^n \left(Y_h^{(i)}(\tau) - a_2 X_h^{(i)} - b_2 \right)^2 \quad (16)$$

参数 a_1, b_1 和 a_2, b_2 分别利用线性最小二乘原理可求得其估计值 $\hat{a}_1, \hat{b}_1, \hat{a}_2, \hat{b}_2$ 。最佳时延系数 $\tau_{opt} \in N^+$ 是表示供热系统的传输线路中延迟实际大小的参数, 取误差平方最小时的时延系数, 通过穷举搜索的方式求解, 运用剪枝规则减少搜索的次数。用求解得的最佳时延系数对采集的供水温度进行反向时延变换, 得到最小二乘意义上的供水温度, 即: $T_{g-opt} = [T_{g-opt}^{(1)}, T_{g-opt}^{(2)}, \dots, T_{g-opt}^{(n)}] \in \mathbf{R}^n$, 其中, $i=1, 2, \dots, n$, n 为向量长度。将观测数据和利用最小二乘原理训练的参数代入, 进行回归预测。得室内温度预测值即:

$$T_{in}^{(i)} = \frac{T_{g-opt}^{(i)} + T_h^{(i)}}{2} - \exp\left(\frac{\ln(G^{(i)}(T_{g-opt}^{(i)} - T_h^{(i)})) - \hat{b}_1}{\hat{a}_1}\right) \quad (17)$$

二次网回水温度预测值即:

$$T_h^{(i)} = T_{g-opt}^{(i)} - \exp\left(\hat{a}_2 \ln\left(T_{in}^{(i)} - T_{out}^{(i)}\right) + \hat{b}_2\right) / G^{(i)} \quad (18)$$

其中, $i=1, 2, \dots, n$ 。

2.2 模型求解

2.2.1 最小二乘参数估计

对线性最小二乘模型进行参数估计, 得参数的估计值 $\hat{a}_1, \hat{b}_1, \hat{a}_2, \hat{b}_2$ 。在室内温度预测中, 分别对损失函数的参数 a_1 和 b_1 求偏导, 并另偏导数为零, 即:

$$\begin{cases} \frac{\partial J_{in}(a_1, b_1, \tau)}{\partial b_1} = -2 \sum_{i=1}^n (Y_{in}^{(i)}(\tau) - b_1 - a_1 X_{in}^{(i)}(\tau)) = 0 \\ \frac{\partial J_{in}(a_1, b_1, \tau)}{\partial a_1} = -2 \sum_{i=1}^n (X_{in}^{(i)}(\tau) (Y_{in}^{(i)}(\tau) - b_1 - a_1 X_{in}^{(i)}(\tau))) = 0 \end{cases}$$

得参数的估计值为:

$$\begin{cases} \hat{a}_1 = \frac{\sum_{i=1}^n X_{in}^{(i)}(\tau) Y_{in}^{(i)}(\tau) - n \overline{X_{in}} \overline{Y_{in}}}{\sum_{i=1}^n (X_{in}^{(i)}(\tau))^2 - n \overline{X_{in}}^2} \\ \hat{b}_1 = \overline{Y_{in}} - \hat{a}_1 \overline{X_{in}} \end{cases}$$

其中, $\overline{X_{in}} = \frac{1}{n} \sum_{i=1}^n X_{in}^{(i)}(\tau)$, $\overline{Y_{in}} = \frac{1}{n} \sum_{i=1}^n Y_{in}^{(i)}(\tau)$ 。

同理, 回水温度预测中, 分别对目标函数的参数

a_2 和 b_2 求偏导, 另偏导数为零, 即:

$$\begin{cases} \frac{\partial J_h(a_2, b_2, \tau)}{\partial b_2} = -2 \sum_{i=1}^n (Y_h^{(i)}(\tau) - b_2 - a_2 X_h^{(i)}(\tau)) = 0 \\ \frac{\partial J_h(a_2, b_2, \tau)}{\partial a_2} = -2 \sum_{i=1}^n ((Y_h^{(i)}(\tau) - b_2 - a_2 X_h^{(i)}(\tau)) (-X_h^{(i)}(\tau))) = 0 \end{cases}$$

得参数估计值为:

$$\begin{cases} \hat{a}_2 = \frac{\sum_{i=1}^n X_h^{(i)} Y_h^{(i)}(\tau) - n \overline{X_h} \overline{Y_h}}{\sum_{i=1}^n (X_h^{(i)})^2 - n \overline{X_h}^2} \\ \hat{b}_2 = \overline{Y_h} - \hat{a}_2 \overline{X_h} \end{cases}$$

其中, $\overline{X_h} = \frac{1}{n} \sum_{i=1}^n X_h^{(i)}$, $\overline{Y_h} = \frac{1}{n} \sum_{i=1}^n Y_h^{(i)}(\tau)$ 。

2.2.2 基于剪枝的时延求解

优化时延系数 τ 时, 若采用穷举搜索法对每一个时延系数进行一次最小二乘参数估计, 从中选取最小误差平方和对应的时延系数为最佳时延系数 τ_{opt} , 但是穷举搜索法的计算量大。

本文提出剪枝算法降低时延求解的计算量。先将穷举区间划分为若干个等距子区间, 求解各个区间的误差平方和的极大值 Max_i 与极小值 Min_i , 再运用剪枝减少穷举搜索的子区间个数。

将时延系数区间 $[0, \tau_{max}]$ 划分为 l 个等距区间, 即: $w_i: [\frac{(i-1)\tau_{max}}{l}, \frac{i \cdot \tau_{max}}{l}]$, $i=1, 2, \dots, l$, 其中 τ_{max}

为设定的时延系数最大值。采用逼近法通过解不等式求解区间内误差平方极小值 Min_i 。选定时延系数区间内的时延系数 τ_0 作为初始值, 进行一次最小二乘回归, 令误差平方和为: $r_0^2 = \sum_{i=1}^n d_i^2$, 其中,

$d_i \in \mathbf{R}, i=1, 2, \dots, n$ 为估计值到实际值的欧氏距离, 如图 1 所示。对于 τ_0 所在时延区间内任意时延系数均存在以下关系:

$$\sum_{i=1}^n (d_i + h_i)^2 \geq r_0^2 \quad (19)$$

其中, $h_i \in \mathbf{R}, i=1, 2, \dots, n$ 为初始点到时延区间内最近待估计点的欧氏距离, 如图 1 所示。式(19)完全平方展开可得:

$$\sum_{i=1}^n d_i^2 + 2 \sum_{i=1}^n d_i h_i + \sum_{i=1}^n h_i^2 \geq r_0^2 \quad (20)$$

令 $r^2 = \sum_{i=1}^n d_i^2$ ，显然 $r > d_i$ ，不等式(20)转换为：

$$r^2 + 2r \sum_{i=1}^N h_i + \sum_{i=1}^N h_i^2 \geq r_0^2 \quad (21)$$

解式(21)的一元二次方程，取不等式解的右半区间 $[r_{\min}, +\infty)$ ，误差平方和的极小值即 $\min_i = r_{\min}^2$ 。误差平方和极大值 \max_i 即：

$$\max_i = r_{\max}^2 = \sum_{i=1}^N (d_i + h_i)^2$$

其中， $h_i \in R, i = 1, 2, \dots, n$ 取初始点到时延区间内最远待估计点的欧氏距离。

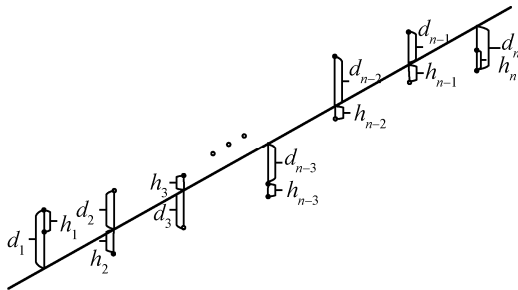


图1 延时时子区间内误差平方的极大值和极小值的求解原理
Fig. 1 Illustration of the calculation on maxima and minima of squared error in the intervals of delay

比较时延系数区间的极大值和极小值，应用以下规则进行剪枝：

(1) 当区间 w_i 的 Min_i 大于区间 w_j 的 Max_j ，误差平方和全局最小值不在 w_i 区间，则区间 w_i 发生剪枝，搜索最佳时延时舍弃该区间。

(2) 当区间 w_i 的 Max_i 小于区间 w_j 的 Min_j ，误差平方和全局最小值不在 w_j 区间，则区间 w_j 发生剪枝，搜索最佳时延时舍弃该区间。

应用剪枝规则，可去除误差过大的区间，减少穷举搜索的次数。在剪枝后的区间内进行最小二乘回归，搜索得到误差平方和最小的时延系数 τ_{opt} 。

2.3 模型流程

本文模型输入输出如下：

Input: $T_g, T_h, T_{in}, T_{out}, G \in R^n$ ， n 为输入数据长度；Output: 最小二乘估计参数 a_1, b_1, a_2, b_2 ，最

佳时延系数 τ ，预测值 $T_{in}, T_h \in R^m$ ， m 为预测长度。

本文模型步骤如下：

Step 1: 数据预处理，去除异常值；通过数据变换，将非线性回归预测问题化为线性问题。

Step 2: 设置时延系数上限 τ_{\max} ，将时延系数区间 $[0, \tau_{\max}]$ 划分为 l 个等距区间，即：

$w_i : [\frac{(i-1)\tau_{\max}}{l}, \frac{i \cdot \tau_{\max}}{l}], i = 1, 2, \dots, l$ ，求解对应区间内的误差平方和最小值 \min_i 和最大值 \max_i ， $i = 1, 2, \dots, l$ 。

Step 3: 比较各区间内误差平方和的极值，若存在 $\min_i > \max_j, i, j = 1, 2, \dots, l$ ，应用剪枝规则进行剪枝，否则，增加时延区间数，返回 Step 2。

Step 4: 在剪枝后的时延系数区间内进行穷举搜索，取误差平方和最小的时延系数为最佳时延系数 τ_{opt} ，并反向时移得最小二乘最优的供水温度 T_{g-opt} 。

Step 5: 进行最小二乘参数估计，利用式(17)预测室内温度，利用式(18)预测二次网回水温度。

3 仿真与分析

3.1 数据采集与预处理

使用某供热站的实测数据作为样本，数据每隔3分钟采集一次，数据24小时采集。在供热系统的二次网中，采集的数据有：二次网供水温度 $T_g = [T_g^{(1)}, \dots, T_g^{(n)}] \in R^n$ ，回水温度 $T_h = [T_h^{(1)}, \dots, T_h^{(n)}] \in R^n$ ，瞬时流量 $G = [G^{(1)}, \dots, G^{(n)}] \in R^n$ ， $i = 1, 2, \dots, n$ ， n 为数据长度。由于采集设备的本身的局限，容易产生一些异常值，这些异常值对预测结果会产生较大影响，因此需要在进行预报前去除异常数据。采用3 sigma 原则，去除异常数据，并用相邻的两个数据的平均值替代。室内温度 $T_{in} = [T_{in}^{(1)}, \dots, T_{in}^{(n)}] \in R^n$ 和室外温度 $T_{out} = [T_{out}^{(1)}, \dots, T_{out}^{(n)}] \in R^n$ 尚未布设传感器进行采集，室外温度采用天气预报温度，室内温度利用状态空间法建立房间室温模型^[12]仿真得到。

3.2 仿真实验与效果

将 $T_g, T_h, T_{in}, T_{out}, G$ 作为输入数据，长度为300，

热网中的时延通常是大时延, 设置时延系数上限为40, 将时延区间划分为8个等距区间, 运用本文算法, 得到每个区间的方差最大值和最小值, 见表1。

表1 各个时延子区间内误差平方最大值与最小值
Tab. 1 The maxima and minima of squared errors in the intervals of delay

时延区间	误差平方最小值	误差平方最大值
[1,5]	1.072 7	1.461 5
[6,10]	1.036 8	1.221 0
[11,15]	1.033 2	1.155 0
[16,20]	0.891 1	1.076 8
[21,25]	0.765 4	1.140 5
[26,30]	0.882 0	1.030 5
[31,35]	0.827 5	1.070 7
[36,40]	0.808 3	0.997 5

根据剪枝规则, 区间1、区间2和区间3可被剪枝, 将时延区间重新选定为区间4至区间8, 在剪枝后的区间中逐次进行最小二乘回归, 得最佳时延系数为33, 求得实验参数如下: α_1 为0.787825, \hat{b}_1 为2.247 025, α_2 为0.609 737, \hat{b}_2 为0.592 641。运用求解得的最佳时延系数再进行最小二乘回归得到室内温度的预测值如图2, 平均相对误差绝对值为0.071 003, 均方误差为2.909 970。回水温度预测结果见图3, 平均相对误差绝对值为0.029083, 均方误差为1.618 852。误差产生的原因有传感器采集数据时产生的噪声, 数据采集间隔以及时延系数的分辨率也可能导致误差的产生。

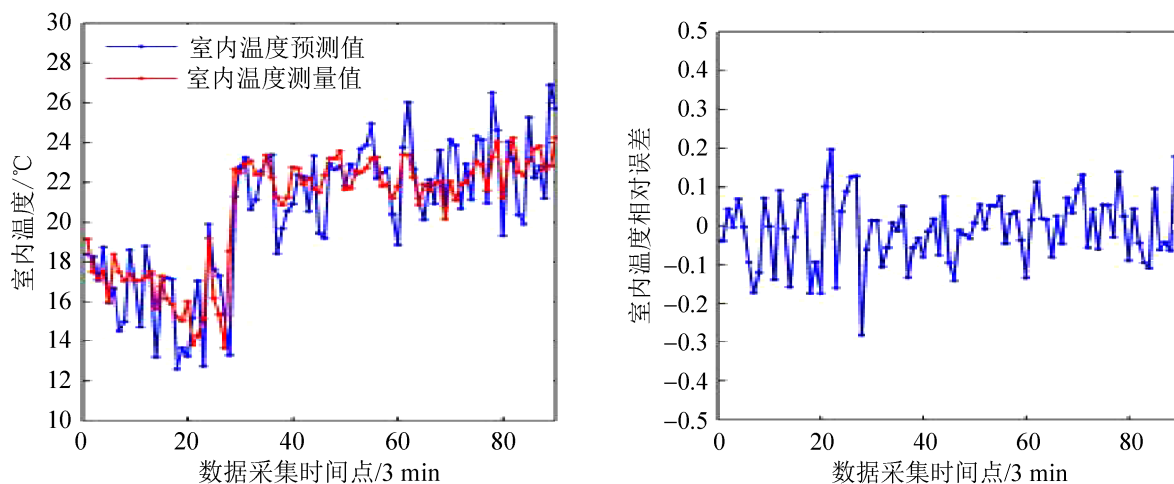


图2 室内温度预测结果与相对误差

Fig. 2 The prediction results and relative errors of indoor temperature

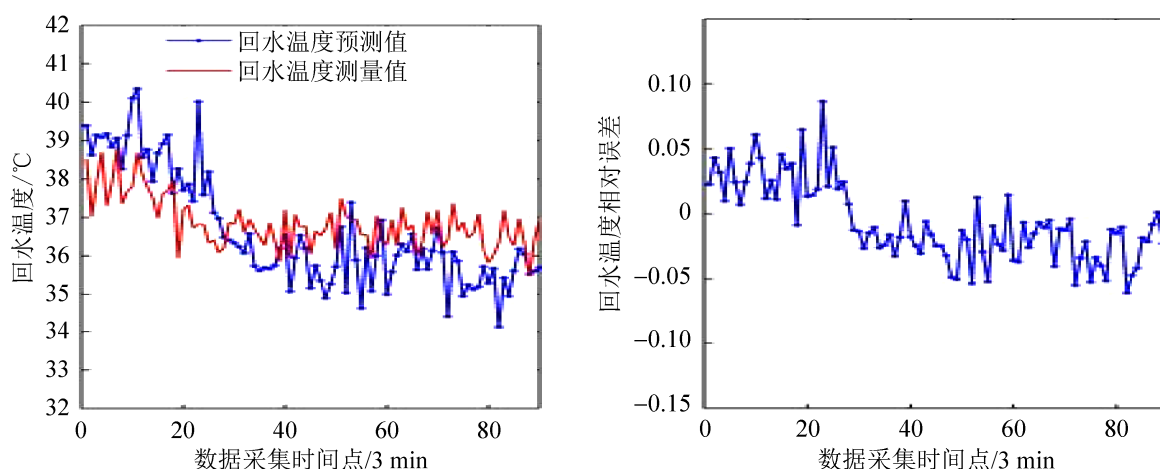
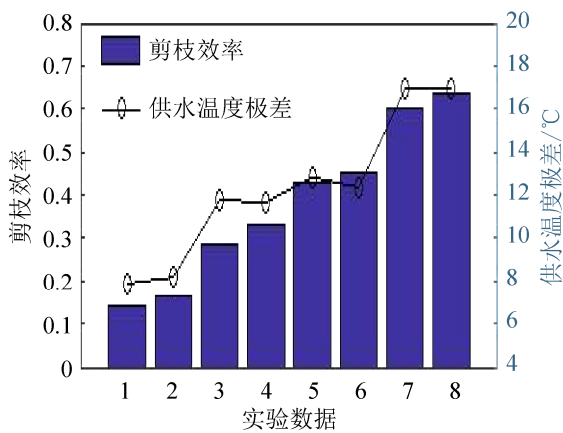


图3 二次网回水温度预测结果与相对误差

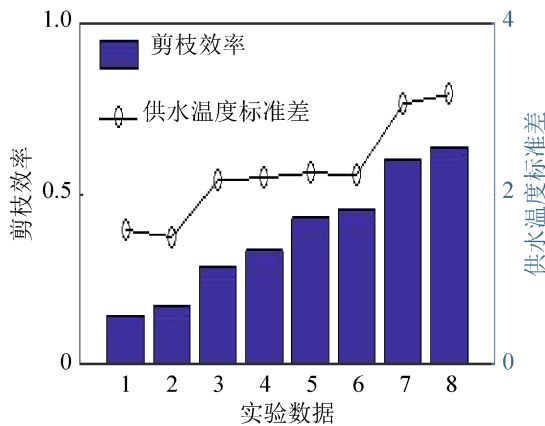
Fig. 3 The prediction results and relative errors of return water temperature in the secondary network

为了验证剪枝算法的有效性,进行了剪枝效率(发生剪枝的区间长度与数据长度之比)验证的仿真实验,剪枝效率与供水温度极差和供水温度标准差如图4所示。当温度波动较大(极差)和温度偏离程度较大(标准差)时,剪枝效果明显;在温度变化平缓的数据中,时延效果相对不明显,剪枝效率较低,甚至不发生剪枝。

经仿真实验,本文方法的预测误差在满意的范围内,剪枝算法能够在温度变化剧烈和间歇供热的情况下取得较高的效率,降低计算量。



(a) 不同数据集的供水温度极差对剪枝效率的影响



(b) 不同数据集的供水温度标准差对剪枝效率的影响

图4 不同数据集中供水温度极差与标准差对剪枝效率的影响

Fig. 4 The effect of range and standard deviation of supply water on the pruning power in different datasets

3.3 仿真实验对比与分析

为了验证本文算法的效果,分别利用梯度下降

最小二乘(GDLS)曲线拟合、标准支持向量机(SVM)、最小二乘支持向量机(LS-SVM)和人工神经网络(ANN)对二次网回水温度预测进行仿真实验对比。SVM仿真利用 libsvm 软件包进行支持向量回归^[13],采用其中的非线性函数估计形式。LS-SVM^[14]采用 Suykens 等人提出的 SVM 改进算法。ANN 采用三层 BP 神经网络,使用 Levenberg-Marquardt 算法进行最优求解^[15]。

直接对式(6)和式(12)进行非线性多元最小二乘回归的复杂度为 $O(mnp^2)$,其中, m 为时延系数上限值, n 为序列长度, p 为输入元素的个数。经本文算法的线性变换,穷举搜索的线性一元最小二乘算法复杂度为 $O(mn)$ 。引入剪枝后,设发生剪枝而舍弃的时延系数区间长度为 k ,且 $k < m$,算法复杂度为 $O((m-k)n)$ 。SVM 的复杂度主要是训练阶段的复杂度,即解二次规划问题的复杂度 $O(n^3)$ 。LS-SVM 的复杂度主要是求解 $n \times n$ 矩阵的逆,复杂度为 $O(n^2)$ 。采用不同 SVM 的优化算法的复杂度不同,本文不再作具体分析。类比于信息论中不完全和不精确求解,ANN 计算的复杂度即为:

$$P \propto \left(\frac{1}{\varepsilon}\right)^{n/R}$$

其中, ε 为允许误差, R 为样本复相关系数。同理,不同优化算法的 ANN 模型复杂度也不同。从以上复杂度分析可以看出,由于预测序列的长度 n 通常较长,本文算法剪枝后的复杂度将低于平方复杂度,理论上时间复杂度低于非线性多元最小二乘回归、LS-SVM、SVM 和 ANN。通过仿真实验进行验证,实验在戴尔 Optiplex 9020 台式机上进行,使用 Matlab R2014b 实现所有算法的仿真,分别进行预测误差对比和程序运行时间对比。

程序运行时间实验分为两组:等长度数据(长度为 300)预测时间和变长度数据(长度分别为 300、600、900、1200、1500、1800、2100、2400 和 2700)预测程序运行时间室内温度和回水温度预测均以

平均相对误差(MAPE)和均方误差(MSE)进行对比, 计算公式分别为:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

其中, y_i, \hat{y}_i 分别表示变量的第 i 个实测值和预测值, 结果如表 2 所示。程序运行时间对比结果见图 5。从仿真结果可以看到, GDLS 预测误差较大, 因其没有

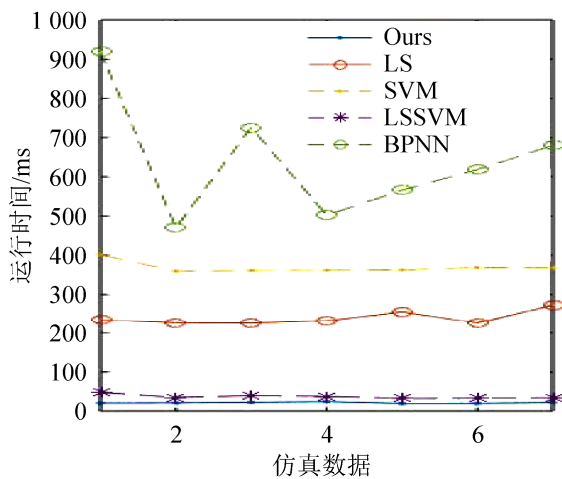
对时延建模和求解, 且容易得到局部最优。BP 神经网络在预测室内温度时在训练集中误差很小, 但在测试集上均方误差很大, 产生了过度训练。

本文方法与另外四种方法相比, 在室内温度预测和回水温度预测上均获得了相对较低的误差。此外, 根据程序算法复杂度分析和程序运行时间比较, 相比复杂的机器学习模型, 本文模型时间复杂度更低, 消耗更少的计算时间。

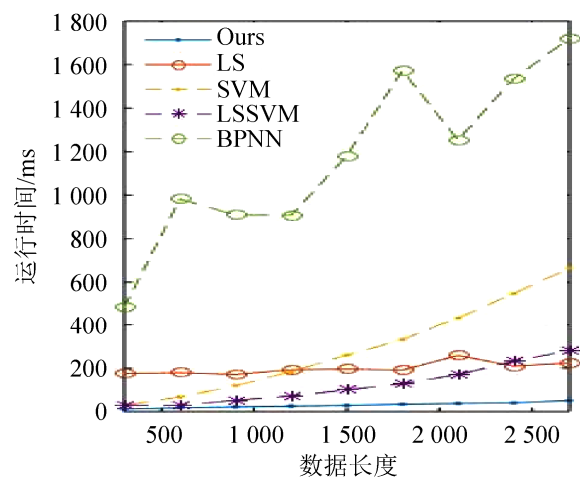
表 2 本文算法、GDLS、SVM、LS-SVM 和神经网络预测平均相对误差绝对值与均方误差对比

Tab. 2 The comparison of mean absolute percentage error and mean squared error among our algorithm, GDLS, SVM, LS-SVM and BPNN

方法	室内温度 MAPE	室内温度 MSE	回水温度 MAPE	回水温度 MSE
本文方法	0.071 003	2.909 970	0.029 083	1.618 852
GDLS	0.095 573	4.437 335	0.035 133	2.689 138
SVM	0.087 516	3.569 746	0.015 482	0.494 675
LS-SVM	0.062 243	2.614 216	0.037 915	2.398 414
BP NN	0.066 172	5.604 367	0.027 439	1.829 617



(a) 序列长度为 300 的不同实验数据集的程序运行时间对比



(b) 不同序列长度下的程序运行时间对比

图 5 本文算法、GDLS、SVM、LS-SVM 和神经网络的程序运行时间对比

Fig. 5 The comparison of time consumption among our algorithm, GDLS, SVM, LS-SVM and BPNN

4 结论

供热系统具有非线性和时滞性等特征, 针对供热预测的时延问题, 本文提出了基于时延的最小二乘预测剪枝算法, 运用于供热系统的短时预测。

本文工作主要集中在以下三方面: (1)运用对数变换解决了供热系统的非线性问题。供热系统非线性是进行供热预测的难点, 对数变换比复杂的机

器学习模型计算更为简单。(2)根据供热系统的传热规律, 提出了符合物理规律的时延模型, 在求解时延系数过程中, 采用剪枝算法, 降低了计算的复杂度。(3)本文同时完成了回水温度和室内温度的预测, 模型简便, 且计算量少, 适用于工程计算。经仿真实验, 本文方法验证有效, 能够为工程实践提供了一定的参考价值。

参考文献:

- [1] 贺平, 孙刚. 供热工程[M]. 4版. 北京: 中国建筑工业出版社, 2009.
Ping He, Gang Sun. Heating Engineering[M]. 4th edition. Beijing: China Architecture & Building Press, 2009.
- [2] 刘满平. 供热工程[M]. 北京: 中国建材工业出版社, 2013.
Manping Liu. Heating Engineering[M]. Beijing: China Architecture & Building Press, 2013.
- [3] Bass M S, Batukhtin A G. An Integrated Approach for Optimizing the Operation of Modern Heat Supply Systems [J]. Thermal Engineering(S0040-6015), 2011, 58(8): 678-679.
- [4] 文翰. 基于最小二乘法的热负荷预测的研究与应用[D]. 大连海事大学, 2014: 12-17.
Wen Han. The research and application of heating load forecast based on the least square method[D]. Dalian Maritime University, 2014:12-17.
- [5] 李琦, 李梅. PSO 优化在供热网络控制中的应用研究[J]. 计算机仿真, 2013, 30(12): 294-297.
Li Qi, Mei Li. Study and Application of PSO in Heating Network Control[J]. Computer Simulation, 2013, 30(12): 294-297.
- [6] 高丙坤, 李阳, 许明子. 优化粒子群算法在组合供热负荷预测中的应用[J]. 信息与电子工程, 2012, 9(5): 655-659.
Gao Bingkun, Li Yang, Xu Mingzi. Application of particle swarm optimization algorithm in the heating load combination forecasting[J]. Information and Electronic Engineering, 2012, 9(5): 655-659.
- [7] 张永明, 陈烈, 齐维贵. 基于支持向量区间回归的供热负荷概率预报[J]. 吉林大学学报(工学版), 2010, 40(6): 1693-1697.
Zhang Yongming, Chen Lie, Qi Weigui. Probabilistic prediction approach of heat load based on support vector interval regression[J]. Journal of Jilin University (Engineering and Technology Edition), 2010, 40(6): 1693-1697.
- [8] 王美萍, 田琦, 张佼. 模糊耦合神经网络的系统供热预测研究[J]. 太原理工大学学报, 2015, 45(4): 485-489.
Wang Meiping, Tian Qi, Zhang Jiao. Research on Heat Forecasting Based on Fuzy Neural Network[J]. Journal of Taiyuan University of Technology, 2015, 45(4): 485-489.
- [9] 崔高健, 凡东生, 曲永利. 基于 Elman 型神经网络集中供热负荷预测模型的研究[J]. 建筑节能, 2011(3): 9-11.
Cui Gaojian, Fan Dongsheng, Qu Yongli. Load Predication Model of District Heating System Based on Elman Neural Network[J]. Building Energy Efficiency, 2011(3): 9-11.
- [10] 谢慕君, 冯敬芳, 姜长泓. 集中供热二次网回水温度的预测和控制研究[J]. 控制工程, 2015, 22(2): 291-295.
Xie Mujun, Feng Jingfang, Jiang Changhong. Research of Secondary Network Backwater Temperature Forecast and Control for Centralized Heat-supply System[J]. Control Engineering of China, 2015, 22(2):291-295.
- [11] 冯敬芳, 谢慕君, 步伟明, 等. 基于神经网络集中供热二次回水温度预测研究[J]. 计算机仿真, 2014, 31(3): 351-354.
Feng Jingfang, Xie Mujun, Bu Weiming, et al. Research on Secondary Network Backwater Temperature Forecast for Centralized Heat-Supply System Based on Neural Network[J]. Computer Simulation, 2014, 31(3):351-354.
- [12] 刘兰斌, 江亿, 付林, 等. 供暖末端通断调节系统仿真模型介绍[J]. 建筑科学, 2010, 26(10): 310-317.
Liu Lanbin, Jiang Yi, Fu Lin, et al. An introduction to the system simulation models of on-off control in the end of heating system[J]. Building Science, 2010, 26(10): 310-317.
- [13] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems & Technology(S2157-6904), 2011, 2(3): 389-396.
- [14] Suykens J A K, Vandewalle J. Least Squares Support Vector Machine Classifiers[J]. Neural Processing Letters (S1370-4621), 1999, 9(3): 293-300.
- [15] Ding S, Chang X H, Wu Q H. A Study on Approximation Performances of Improved BP Neural Networks Based on LM Algorithms[J]. Applied Mechanics & Materials(S1662-7482), 2013(411/412/413/414): 1935-1938.