

DOI: 10.11992/tis.201712003

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180515.1116.004.html>

基于 GABP-KF 的 WSN 数据漂移盲校准算法

武加文^{1,2}, 李光辉^{1,2,3}

(1. 江南大学 物联网工程学院, 江苏 无锡 214122; 2. 物联网技术应用教育部工程技术研究中心, 江苏 无锡 214122; 3. 江苏省模式识别与计算智能工程实验室, 江苏 无锡 214122)

摘要: 针对无线传感器网络节点容易产生数据漂移的问题, 提出了一种新型的跟踪和校准节点数据流漂移的算法。首先使用基于遗传算法优化的 BP 神经网络对目标节点和其邻居节点间的时空相关性进行建模, 以获得目标节点的预测值, 再使用卡尔曼滤波器跟踪和校准该节点的数据漂移。针对不同的真实数据集进行仿真实验显示, 该方法相较于其他对比方法模型预测精度更高, 漂移校准性能更好。实验结果表明, 该算法可以精确地校准传感器节点的数据漂移, 提高节点数据的可靠性。

关键词: 无线传感器网络; 数据漂移; 盲校准; BP 神经网络; 遗传算法; 卡尔曼滤波器; 去噪; 模型构建

中图分类号: TP274.2 **文献标志码:** A **文章编号:** 1673-4785(2019)02-0254-09

中文引用格式: 武加文, 李光辉. 基于 GABP-KF 的 WSN 数据漂移盲校准算法[J]. 智能系统学报, 2019, 14(2): 254-262.

英文引用格式: WU Jiawen, LI Guanghui. GABP-KF-based blind calibration algorithm of data drift in wireless sensor networks[J]. CAAI transactions on intelligent systems, 2019, 14(2): 254-262.

GABP-KF-based blind calibration algorithm of data drift in wireless sensor networks

WU Jiawen^{1,2}, LI Guanghui^{1,2,3}

(1. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; 2. Research Center of IoT Technology Application Engineering (MOE), Wuxi 214122, China; 3. Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Wuxi 214122, China)

Abstract: Data drifts easily occur in wireless sensor network. To solve this problem, we propose a novel algorithm for tracking and calibrating the drifts of sensor data stream. First, backpropagation (BP) neural network optimized by genetic algorithm is applied to model the spatio-temporal correlations between the target node and its neighbor nodes to predict the value of the node, and then, the data drift of the node is tracked and calibrated by a Kalman filter. The simulation results using different datasets demonstrate that this method has superior prediction accuracy and calibration performance, compared with other methods. The experimental results show that this method can accurately calibrate the sensor drift and improve the reliability of node data.

Keywords: wireless sensor network; data drift; blind correction; BP neural network; genetic algorithm; Kalman filter; denoising; model building

近年来, 随着无线通信, 计算机网络和传感器技术的巨大进步, 无线传感器网络 (wireless sensor

networks, WSN) 已成为重要的研究对象并被广泛应用于医疗、军事、环境监测等众多领域^[1]。在环境监测领域, 大规模的无线传感器网络通常被部署于无人监管、气候复杂的野外环境中。若长期使用, 节点读数容易产生数据漂移。这对于终端用户来说是一个严重的问题, 因为在无线传感器网络的大规模应用场景中必须使用精确数据, 才

收稿日期: 2017-12-02. 网络出版日期: 2018-05-17.

基金项目: 国家自然科学基金项目 (61472368)、江苏省重点研发计划项目 (BE2016627)、中央高校基本科研业务费项目 (JUSRP51635B)、无锡市国际科技研发合作项目 (CZE02H1706)、江苏省普通高校专业学位论文研究生实践创新计划项目 (SJLX16_0499).

通信作者: 李光辉. E-mail: ghli@jiangnan.edu.cn.

能制定科学决策。因此,对节点数据进行漂移校准至关重要。然而,在许多大规模无线传感器网络的应用中,数百到数千个节点被分散部署到气候复杂的野外环境中,手动校准这些节点是非常困难的^[2]。因此,需要自动跟踪和校准节点数据漂移的方法,现有的方法可以分为两类:盲校准和非盲校准。

非盲校准方法是基于传感器节点读数和已知的参考信息作为输入来调整参数^[3]。一种简单的非盲校准方法是将已知的激励作用于传感器网络中并测量其响应,将网络响应与参考信息(网络预期值)进行比较,然后相应地调整参数。参考信息一般为出厂标准值或用户手动校准的结果。例如,Ramanathan等^[4]采用额外的高精度节点读数作为参考信息。另一种类型的方法是手动校准一组节点,然后未校准节点组基于该组来调整参数^[5]。非盲校准方法适用于特定空间(例如室内)小规模网络的场景下^[6],然而在大规模无线传感器网络中,使用上述方法需要在每个阶段进行大量的工作,这是不切实际的且成本高昂。

鉴于非盲校准方法存在的局限性,盲校准方法得到了广泛的关注。许多盲校准方法是基于节点被密集部署的假设,因此当网络中未产生数据漂移时,节点间应具有相关的读数^[7]。然而,在许多现有大型网络中,由于环境因素和成本因素,可能无法实现节点的密集部署。Balzano等^[8]首先提出了一种无需密集部署的校准方法,假定节点的测量位于测量空间的较低维信号子空间中,然后通过信号子空间投影将校准过程转换为求解线性方程。来自清华大学的杨华中等^[9,10]对这种方法进行了利用和扩充,取得了非常良好的实用价值。

基于相邻节点间具有相关测量值的假设,Takruri等^[11]提出了一种分布式递归贝叶斯(distributed recursive Bayesian, DRB)盲校准算法,其使用邻居节点读数的平均值作为漂移节点的预测值。为了更好地拟合传感器节点间的时空相关性,Takruri等^[12]又提出了一种基于支持向量回归(support vector regression, SVR)来预测真实值的方法。针对SVR可能存在的训练精度较低的情况,Kumar等^[13]采用克里金插值作为预测方法,以避免SVR的训练阶段。这些方法基于节点间的时空相关性以预测节点数据的漂移,并都使用了卡尔曼滤波器(Kalman filter, KF)以跟踪漂移。然而,一旦预测值不精确,偏差较大的预测值也会被用来进一步校准漂移值,从而导致最终校准值

的精度受累积误差的限制。

针对上述问题,本文提出了一种基于遗传算法(genetic algorithm, GA)优化的BP神经网络(BP neural network optimized by GA, GABP)与卡尔曼滤波器相结合的无线传感器网络数据漂移盲校准算法(GABP-KF)。该算法首先对节点读数进行二次采集和降噪处理。然后使用基于遗传算法优化的神经网络对目标节点和其邻居节点间的时空相关性进行数学建模,从而得到目标节点的预测值。最后将目标节点的预测值和测量值反馈入卡尔曼滤波器中以跟踪并校准其数据漂移。仿真结果表明,GABP-KF具有良好的漂移校准性能。

1 预备知识

1.1 遗传算法

遗传算法是模拟自然界遗传机制的生物进化模型,是一种启发式随机搜索最优化方法。它将“物竞天择,适者生存”的进化原理和遗传变异理论引入编码串联群体中,即在每一次的进化过程中,按照所构造的适应度函数,通过选择操作、交叉操作和变异操作来对个体进行筛选,把适应度值好的个体传给下一代,反复循环,直到最终个体满足特定的条件表示进化结束^[14]。遗传算法具备良好的全局搜索能力,不受求导和函数连续性的限定,能自动获取待优化的搜索空间,从而自适应地控制搜索过程^[15]。

1.2 卡尔曼滤波器

卡尔曼滤波器是由R.E.Kalman提出的线性最小方差递推估算的方法,是一种离散数据线性滤波问题的递归解决方案^[16]。它是利用当前时刻的观测值和前一时刻的估计值来更新当前时刻的最佳估计,其所需数据存储量较小,便于实时处理^[17]。

2 基于GABP-KF的WSN数据漂移校准方法

2.1 定义测量环境

假设在某一区域内的传感器网络中,随机分布多个节点。其中每个节点采集该区域内的环境参数(温度、湿度或其他类型参数)。随着时间的推移,一些节点产生数据漂移,降低了数据质量,从而影响了传感器网络的可靠性。为了更好地模拟实际场景,本文对传感器网络的测量环境做了如下定义。

定义1 传感器节点的数据漂移指的是节点测量值受其内部固有偏差或外界环境影响所产生

的缓慢、单向、长期的变化。在本文中,节点的数据漂移表示如下:

$$X = T + d + W \quad (1)$$

式中: X 表示节点的测量值; T 表示所测环境的真实值; d 表示漂移值; W 表示测量高斯白噪声。其中, 漂移 d 是平滑的, 因为漂移通常是一个缓慢的线性或指数的变化过程, 没有突变, 激增或尖峰现象^[18]。同时, 漂移的产生是随机的, 与其内部构造以及环境因素密切相关。因此, 不同节点是否产生漂移以及漂移值的大小不具有相关性。由于节点在部署前都经过预先校准, 以确保它们处于工作状态。因此, 节点部署后的短时间内, 漂移值应为零。在本文中, 设定节点不会同时产生数据漂移, 且产生的漂移值不同。

定义 2 在以目标节点为圆心的测量区域内, 该节点向其周围节点传输信号, 该信号的强度会随距离的增加而减弱, 当信号强度小于阈值时, 接收节点无法接收到数据。因此, 通信半径指的是信号衰减至阈值的距离, 而邻居节点指目标节点通信半径内的其他节点。在本文中, 假设目标节点 i 共有 $m-1$ 个邻居节点, 分别标号为 $1, 2, \dots, m, i-1, i+1, \dots, m$ 。

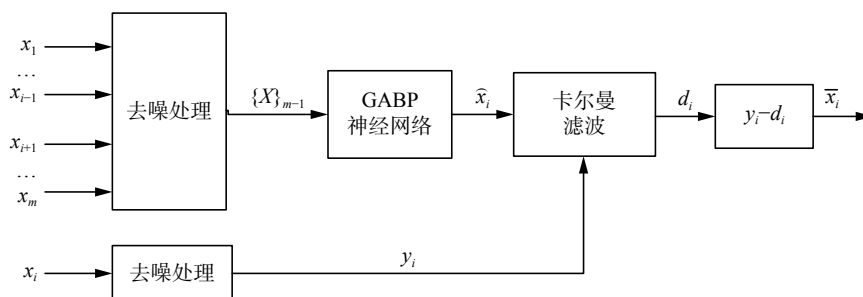


图 1 GABP-KF 算法的数据漂移校准框架

Fig. 1 Data drift calibration framework of GABP-KF algorithm

2.3 基于 GABP 神经网络的建模过程

2.3.1 数据预处理

在传感器网络的实际部署中, 由于受到传感器节点内部固有偏差或其他环境因素, 节点在测量过程中会受到噪声的干扰。在节点校准之前, 必须考虑消除噪声的影响。本文采用小波去噪^[19]方法来降低噪声的影响, 首先使用阈值去噪法对含噪数据进行小波分解, 再对小波系数进行阈值处理, 最后利用处理后结果重构原信号。实验证明该方法在信号保护和噪声抑制之间达到了良好的平衡。

2.3.2 GABP 神经网络的建模

传统的 BP 神经网络具有良好的自学习、自适应能力, 但同时也存在容易陷入局部最优解、

定义 3 目标节点 i 与其邻居节点的时空相关性指的是同一时段内的节点测量数据的变化趋势是一致的。在本文中, 假设目标节点 i 与其邻居节点间具有时空相关性, 因此, 可以用来预测目标节点的测量值, 即:

$$\hat{x}_i = f(\text{neighbours_data}) \quad (2)$$

式中: \hat{x}_i 表示目标节点 i 的测量数据的预测值, $f(\text{neighbours_data})$ 表示相邻节点之间的时空相关性函数。

2.2 算法框架

GABP-KF 算法的设计目标是在随机误差 (噪声) 和系统误差 (漂移) 的干扰下精确地校准目标节点 i 的测量值 x_i 。该算法有两个运行阶段: 训练阶段和测试阶段。在训练阶段, 首先对所有节点数据进行去噪处理。然后使用来自邻居节点的去噪后测量值 $\{X\}_{m-1}$ 作为 GABP 神经网络的输入值, 并输出节点 i 的预测值 \hat{x}_i 。在测试阶段, 将预测值 \hat{x}_i 和节点 i 的去噪后测量值 y_i 输入到卡尔曼滤波器中以跟踪其数据漂移 d_i 。最后通过测量值 y_i 减去估计漂移值 d_i 以得到最终校准后的读数 \bar{x}_i 。图 1 描述了本文算法的具体框架。

收敛速度差等问题。为了克服这些缺点, 本文使用遗传算法来优化 BP 神经网络, 将遗传算法的全局搜索能力与 BP 神经网络的局部搜索能力有效地结合起来, 从而提高了网络的预测精度和泛化能力。具体方法如下: 首先, BP 神经网络根据数据集的划分情况确定网络结构, 从而确定遗传算法的个体长度。其次, 算法通过选择操作、交叉操作和变异操作来优化 BP 神经网络的参数。最后, BP 神经网络使用最优参数来预测目标节点的测量值。GABP 神经网络的模型建立方法如图 2 所示。

本文使用在传感器网络部署初期 (一般为前五天) 所采集到的节点数据作为训练数据集, 训练数据集划分为: $\text{Train} = \{\text{TY}; \text{TX}\}$ 。TY 表示数据

测试集, TX 表示训练集。TY = $\{x_{i,t} : t = 1, 2, \dots, n\}$, TX = $\{x_{j,t} : j = 1, 2, \dots, i-1, i+1, \dots, m, t = 1, 2, \dots, n\}$, 其中, j 表示节点序号, t 表示时序。

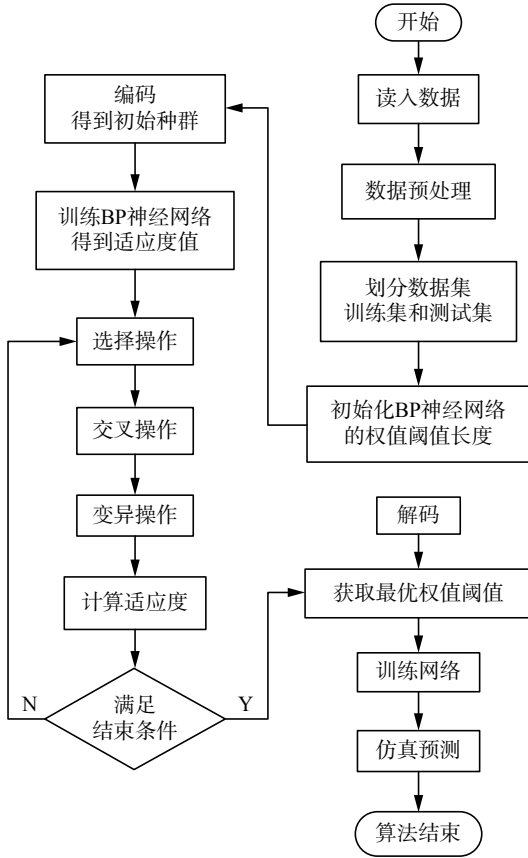


图2 GABP神经网络的建模算法

Fig. 2 Modeling algorithm of GABP neural network

2.4 基于卡尔曼滤波的漂移跟踪算法

本文设定 t 时刻节点 i 的平滑漂移值为 $d_{i,t}$, t 时刻节点 i 所测量的真实温度值为 $T_{i,t}$, t 时刻节点 i 的测量值为 $r_{i,t}$ 。其中 i 表示节点序号, t 表示时间序列。为了预测漂移值 $d_{i,t}$, 首先需要对节点的数据漂移进行数学建模, 见式 (3):

$$d_{i,t} = d_{i,t-1} + w_{i,t}, w_{i,t} \sim N(0, Q_{i,t}) \quad (3)$$

式中: $w_{i,t}$ 表示高斯噪声, 而 $Q_{i,t}$ 表示状态噪声协方差。

式 (3) 表示节点数据漂移的状态跟踪方程, 用来跟踪漂移值随时间的变化情况。除了状态方程, 还需要建立一个观测方程来模拟节点的观测结果。假设现实中存在一种测量仪器, 可以测量出节点的数据漂移值, 同时一定存在与节点测量时相关联的误差, 见式 (4):

$$z_{i,t} = d_{i,t} + v_{i,t}, v_{i,t} \sim N(0, R_{i,t}) \quad (4)$$

式中: $v_{i,t}$ 表示高斯噪声, 而 $R_{i,t}$ 表示观测噪声协方差。

实际上并不存在可以直接测量节点数据漂移的仪器。在这种情况下, 理想的数据漂移值 $d_{i,t}$ 等

于测量值减去实际值, 见式 (5):

$$d_{i,t} = r_{i,t} - T_{i,t} \quad (5)$$

然而在实际场景中, 实际值是无法得知的, 只能使用其预测值 $\hat{x}_{i,t}$ 。因此:

$$d_{i,t} = r_{i,t} - \hat{x}_{i,t} \quad (6)$$

综上所述, 式 (3) 和式 (4) 建立了一个传感器节点的状态-观测模型, 使用卡尔曼滤波器以分散迭代的方式进行跟踪漂移, 具体算法参见图 3。

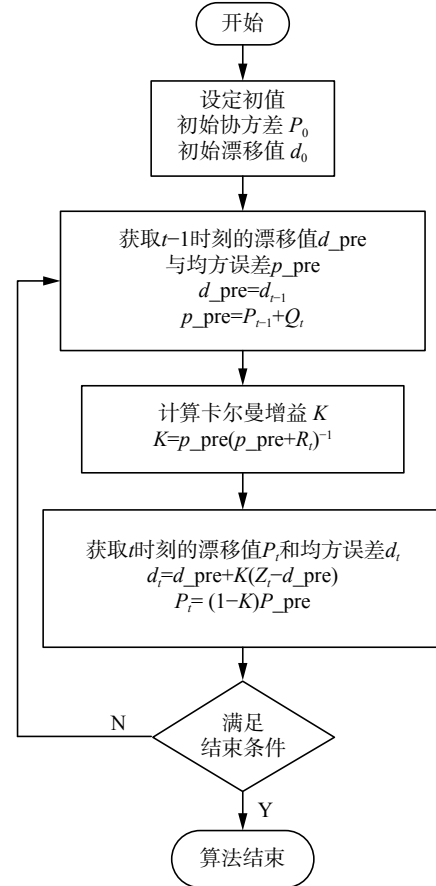


图3 针对目标节点的漂移跟踪算法

Fig. 3 Drift tracking algorithm for target node

如图 3 所示, 其中 Z_t 、 Q_t 、 R_t 分别表示 t 时刻目标节点的漂移测量值、过程噪声和测量噪声。由于目标节点 i 共测量了 n 个数据, 因此算法需要迭代 n 次, 即, 算法结束条件为: $t > n$ 。在进行第 t 次迭代时, 首先基于 $t-1$ 时刻的漂移估计值 d_{t-1} 和均方误差 p_{t-1} 预测 t 时刻的漂移估计值 d_pre 和均方误差 p_pre , 然后计算 t 时刻的卡尔曼滤波增益 K , 最后更新 t 时刻的漂移值 d_t 和均方误差 P_t 。算法迭代完成后, 基于滤波后漂移值 d_t 来校准节点数据, 见式 (7):

$$\bar{x}_{i,t} = r_{i,t} - d_{i,t} \quad (7)$$

3 实验结果及分析

为了验证 GABP-KF 的算法性能, 本文在 mat-

lab 2013b 的实验环境下针对国际通用的 Intel Berkeley(LBRL) 无线传感器网络数据集和 LUCE 无线传感器网络数据集分别实现了 GABP-KF 算法、未优化的 BP-KF 算法、Takaruri 提出的 SVR-KF 算法^[12]以及 DRB-KF 算法^[11], 并进行了实验结果的比较。

3.1 数据集

3.1.1 IBRL 数据集

IBRL 数据集来自于部署在 Intel Berkeley 实验室内的无线传感器网络, 包含 54 个节点, 用于监控实验室环境 (参见图 4)。

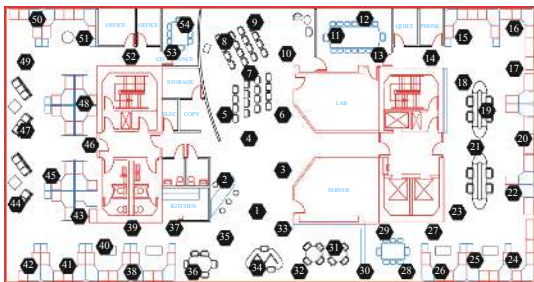


图 4 IBRL 的节点分布

Fig. 4 Deployment scheme of IBRL sensor network

选取由 IBRL 部署中的两组传感器节点的数据子集进行实验。第 1 组数据子集 (IBRL_1) 包含的节点 ID 分别是 1、2、3、4、6、7、10、33。第 2 组数据子集 (IBRL_2) 包含的节点 ID 分别是 17、18、19、20、21、22、23。两组数据集都对应于 2004-02-28 ~ 2004-03-07 9 天内所收集的数据。

3.1.2 LUCE 数据集

LUCE 数据集 (洛桑城市冠层实验) 来自于 2006 年 7 月以来部署在洛桑联邦理工学院内的无线传感器网络。该网络共包含 97 个节点, 根据节点之间的时空相关性分为 10 组传感器节点集。

选取 LUCE 数据集的两组传感器节点的数据子集作为实验对象, 第 1 组数据子集 (LUCE_1) 包含的节点 ID 分别是 10、14、15、17、18、19。第 2 组数据子集 (LUCE_2) 包含的节点 ID 分别是 21、23、24、25、26、27、28。两组数据子集都对应于 2016-10-10 ~ 2016-10-13 4 天内所收集的数据。

由于上述数据集中节点的测量数据量庞大, 其中存在缺漏值, 且采样周期均为 31 s。因此, 对每个节点分别以 5 min 的时间间隔进行重新采样, 并均使用温度作为评估数据。为了精确评估 GABP-KF 算法的性能, 排除实验数据由于缺漏现象造成的影响, 人工建立了一组具有时空相关性的测试数据集 (DR_d)。表 1 列出了所使用的所有

数据集。

表 1 实验所用数据集

Table 1 Experimental data sets

序号	数据集	训练样本数	测试样本数
I	IBRL_1	1 200	1 200
II	IBRL_2	1 000	1 000
III	LUCE_1	1 100	1 100
IV	LUCE_2	1 100	1 100
V	DR_d	1 100	1 100

3.2 去噪处理

针对上述数据集中的所有节点数据, 采用硬阈值去噪的方法进行处理, 以减少噪声对 GABP-KF 算法评估所造成的影响。由于本文所采用的数据集中噪声干扰较小, 为进行去噪算法的评估, 故选择在数据集 III 中 10 号节点处对节点数据手动添加高斯白噪声, 去噪结果参见图 5。如图 5 所示, 原始数据与去噪后数据几近重合, 这表明本方法对噪声有很好的抑制作用。

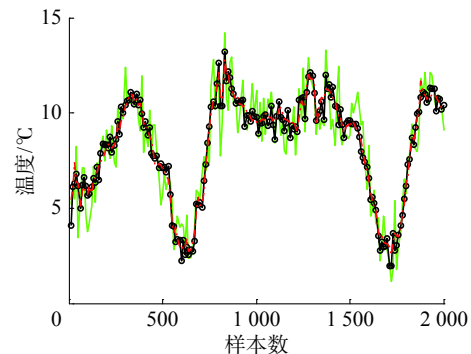


图 5 小波去噪实验结果

Fig. 5 Experimental results of wavelet denoising

3.3 模型对比实验

为比较 GABP-KF 算法与其他同类算法的模型拟合程度, 基于表 1 中的 5 个数据集分别随机选择其中一个节点的数据针对 4 种算法进行了 5 次对比实验, 并选取平均值作为最终数据, 表 2 给出了 4 种算法的模型拟合程度对比结果。本文选用均方误差 (式 (8)) 与决定系数 (式 (9)) 作为判断依据。

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (8)$$

$$R^2 = \frac{\left(n \sum_{i=1}^n \hat{x}_i x_i - \sum_{i=1}^n \hat{x}_i \sum_{i=1}^n x_i \right)^2}{\left(n \sum_{i=1}^n \hat{x}_i^2 - \left(\sum_{i=1}^n \hat{x}_i \right)^2 \right) \left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right)} \quad (9)$$

式中: n 表示训练样本数目, x_i 与 \hat{x}_i 分别为第 i 个样本的真实值与预测值。一般情况下, 决定系数越高, 均方误差越小, 表明模型的拟合程度越好。

表2 4种算法的模型拟合程度对比结果

Table 2 Comparison of the fitting degree among four algorithms

数据集	DRB		BP		SVR		GABP	
	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²
I	0.249 3	98.57%	0.161 2	98.13%	0.409 8	96.31%	0.038 7	99.53%
II	0.719 2	92.05%	0.640 4	92.88%	1.044 4	86.67%	0.105 1	98.73%
III	0.098 7	98.66%	0.297 1	95.30%	0.289 9	95.36%	0.321 5	98.62%
IV	0.251 8	97.89%	0.211 7	96.67%	0.271 4	96.11%	0.334 2	98.17%
V	0.040 6	99.11%	0.100 0	99.24%	0.037 7	99.38%	0.006 0	99.94%
平均值	0.271 9	97.26%	0.282 1	96.44%	0.410 6	94.77%	0.161 1	99.00%

从表2可以看出,基于本文所采用的数据集,GABP算法的均方误差比DRB算法、BP、SVR算法平均减少了0.110 8、0.121 0、0.249 5;GABP算法的决定系数比DRB、BP、SVR算法平均提高了1.74%、2.56%、4.23%。其中,4种算法在数据集V下的拟合程度均好于其余数据集,这是由于数据集V为人工建立,不受外界误差影响,其余4个数据集选自真实数据集,均受到数据缺失,数据突变等误差的影响。DRB、BP、SVR算法受这些误差的影响较大,而GABP神经网络通过GA优化有效地避免了局部误差,因此,GABP神经网络的模型拟合程度要优于其他算法。

图6表示4种算法在数据集I下针对1号节点所建立的预测模型,其中real value描述的曲线表示真实值,其余4条曲线分别表示GABP、SVR、BP和DRB算法所获得的预测结果。从图6中可以直观地看出,GABP神经网络的拟合程度更好,这与表2的结果是一致的。综上所述,GABP算法具有更好的模型拟合程度,所输出的预测值与真实值的误差更小。

3.4 基于GABP-KF的漂移校准对比实验

在漂移校准过程中,于不同节点测试数据的

随机位置处引入不同的指数漂移值与测量噪声。其中,当观测噪声协方差 R 值或状态噪声协方差 Q 值设置过高时,预测值会产生较大误差。因此,采用试凑法对参数值进行调整,在数据集I下,基于网格搜索的原理将卡尔曼滤波器的参数 Q 和 R 分别设置为0.1和0.01。

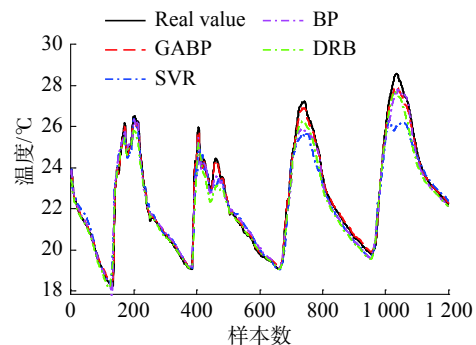


图6 4种算法的模型对比结果

Fig. 6 Comparison among the training models of four algorithms

为比较GABP-KF算法与其他同类算法的漂移校准性能,基于表1中的5个数据集分别随机选择其中一个节点的数据针对4种算法进行了3次对比实验,并选取平均值作为最终数据,表3给出了实验结果。

表3 4种算法的漂移校准性能对比结果

Table 3 Comparison of the drift calibration performance among four algorithms

数据集	DRB-KF		BP-KF		SVR-KF		GABP-KF	
	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²
I	0.230 5	98.10%	0.171 1	97.93%	0.403 0	96.39%	0.017 1	99.35%
II	0.898 2	93.44%	0.401 4	93.76%	1.051 2	91.77%	0.021 7	98.59%
III	0.099 8	98.71%	0.246 7	96.10%	0.188 7	97.24%	0.107 5	98.11%
IV	0.376 1	97.81%	0.310 6	98.72%	0.379 2	97.53%	0.058 4	98.28%
V	0.080 6	99.12%	0.089 0	99.17%	0.037 7	99.38%	0.004 4	99.98%
平均值	0.337 0	97.44%	0.243 8	97.14%	0.412 0	96.46%	0.041 8	98.86%

从表3中可以看出,基于本文所采用的数据集,GABP-KF算法的均方误差比DRB-KF、BP-KF、SVR-KF算法平均减少了0.2952、0.2020、0.3702;GABP-KF算法的决定系数比DRB-KF、BP-KF、SVR-KF算法平均提高了1.43%、1.72%、2.40%。实验表明,基于本文所采用的数据集,GABP-KF算法具有更好的漂移校准性能。

图7表示在数据集Ⅲ下针对14号节点的数据使用GABP-KF算法进行漂移校准后的结果。从图7中可以看出,节点真实值曲线与滤波后数值曲线基本保持一致。这表明本算法成功地消除了引入的漂移值和测量噪声。

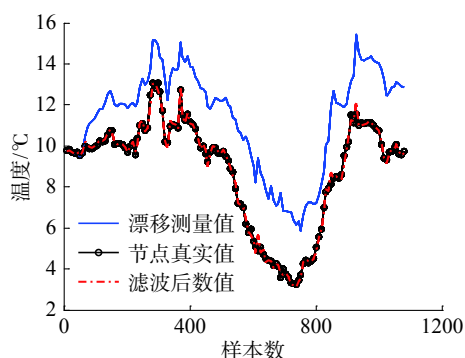


图7 卡尔曼滤波后的校准示例

Fig. 7 Example of data drifts calibration after Kalman filter

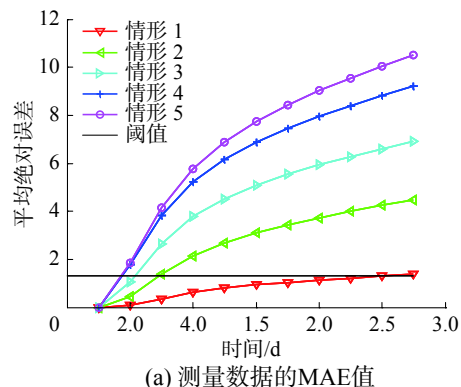
3.5 整体评估

在进行整体评估时,考虑以下两种情形。第1种情形为所有节点数据均未添加漂移值,称为N-data,第2种情形为在指定的节点数据中加入漂移值,称为D-data。两组数据N-data和D-data在经过GABP-KF算法的校准后分别称为CN-data和CD-data。其中D-data中在不同节点测试数据的随机位置处引入不同的指数漂移,即每个节点开始产生漂移的时刻不同,产生漂移的大小不同,是否产生漂移互不相关。

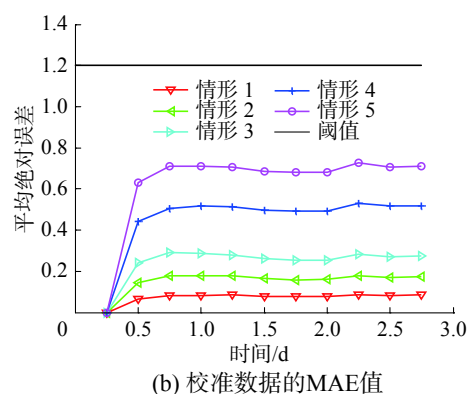
本文比较了5种情形下节点的平均绝对误差MAE值(见式(10))。如图8所示,情形1、2、3、4、5分别表示仅有一个节点漂移、两个节点同时漂移、3个节点同时漂移、4个节点同时漂移、5个节点同时漂移等5种情形。具体计算方法如下:在每个时刻分别计算每个节点的测量数据N-data与D-data之间的MAE值和校准后数据CN-data和CD-data之间的MAE值,然后计算网络中5种情形下的整体MAE值。本文选择在数据集IV下21号节点的数据上进行上述对比实验。

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}_i| \quad (10)$$

式中: n 为训练样本数目, X_i 与 \hat{X}_i 分别为第 i 个样本的真实值和预测值。



(a) 测量数据的MAE值



(b) 校准数据的MAE值

图8 5种情形下MAE值的变化情况

Fig. 8 Comparison of MAE among five cases

图8(a)表示5种情形下测量数据N-data和D-data之间的平均绝对误差。图8(b)表示5种情形下校准数据CN-data和CD-data之间的平均绝对误差。比较图8(a)与图8(b),显然,应用GABP-KF算法导致所有情形下的MAE值较小。本文假设无线传感器网络中阈值MAE值为1.2,超过阈值则表示该网络失效,阈值的选择取决于实际网络中允许的容误差。在图8(a)中,情景2~5的曲线在第一天时就跨越了阈值线。相比之下,图8(b)中的所有情形的曲线在实验的整个周期中均未跨越阈值线。这证明使用本算法可以允许多个节点同时产生数据漂移的情况,这更符合实际情况。

4 结束语

本文提出了一种使用基于遗传算法优化的BP神经网络和卡尔曼滤波器相结合的无线传感器网络数据漂移盲校准算法。仿真实验表明,与以往同类算法相比,GABP-KF算法对无线传感器网络节点数据流漂移有更好的跟踪和校准性能,使用该算法有效地提高了传感器网络数据的可靠性。在大规模无线传感器网络的应用中,节点可以同

时测量多种属性的数据,本算法只考虑了单一属性(温度)下的漂移校准情况。在今后的研究将考虑节点同时采集多种属性数据时产生或未产生漂移的各种情况,以便更好地模拟实际情况。

参考文献:

- [1] WU Mou, TAN Liansheng, XIONG Naixue. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications[J]. *Information sciences*, 2016, 329: 800–818.
- [2] KUMAR D, RAJASEGARAR S, PALANISWAMI M. Geospatial estimation-based auto drift correction in wireless sensor networks[J]. *ACM transactions on sensor networks*, 2015, 11(3): 50.
- [3] TAN Rui, XING Guoliang, YUAN Zhaohui, et al. System-level calibration for data fusion in wireless sensor networks[J]. *ACM transactions on sensor networks*, 2013, 9(3): 28.
- [4] RAMANATNAN N, BALZANO L, BURT M, et al. Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks[R]. Technical Report CENS TR 62, Center for Embedded Networked Sensing, 2006.
- [5] MILUZZO E, LANE N D, CAMPBELL A T, et al. CaliBree: A self-calibration system for mobile sensor networks [C]//*Proceedings of the 4th IEEE International Conference on Distributed Computing in Sensor Systems*. Berlin, Heidelberg, Germany: Springer-Verlag, 2008: 314–331.
- [6] TAYLOR C, RAHIMI A, BACHRACH J, et al. Simultaneous localization, calibration, and tracking in an ad hoc sensor network[C]//*Proceedings of the 5th International Conference on Information Processing in Sensor Networks*. Nashville, TN, USA: IEEE, 2006: 27–33.
- [7] LEE B T, SON S C, KANG K. A blind calibration scheme exploiting mutual calibration relationships for a dense mobile sensor network[J]. *IEEE sensors journal*, 2014, 14(5): 1518–1526.
- [8] BALZANO L, NOWAK R. Blind calibration of sensor networks[C]//*Proceedings of the 6th International Conference on Information Processing in Sensor Networks*. Cambridge, MA, USA: IEEE, 2007: 79–88.
- [9] LI Zhan, WANG Yuzhi, YANG Anqi, et al. Drift detection and calibration of sensor networks[C]//*Proceedings of 2015 International Conference on Wireless Communications & Signal Processing*. Nanjing, China: IEEE, 2015: 1–6.
- [10] WANG Yuzhi, YANG Anqi, LI Zhan, et al. Blind drift calibration of sensor networks using signal space projection and Kalman filter[C]//*Proceedings of 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*. Singapore, Singapore: IEEE, 2015: 1–6.
- [11] TAKRURI M, ABOURA K, CHALLA S. Distributed recursive algorithm for auto calibration in drift aware wireless sensor networks[M]//ELLEITHY K. *Innovations and Advanced Techniques in Systems, Computing Sciences and Software Engineering*. Netherlands: Springer, 2008: 21–25.
- [12] TAKRURI M, RAJASEGARAR S, CHALLA S, et al. Online drift correction in wireless sensor networks using spatio-temporal modeling[C]//*Proceedings of 2008 11th International Conference on Information Fusion*. Cologne, Germany: IEEE, 2008: 1–8.
- [13] KUMAR D, RAJASEGARAR S, PALANISWAMI M. Automatic sensor drift detection and correction using spatial Kriging and Kalman filtering[C]//*Proceedings of 2013 IEEE International Conference on Distributed Computing in Sensor Systems*. Cambridge, MA, USA: IEEE, 2013: 183–190.
- [14] DING Shifei, SU Chunyang, YU Junzhao. An optimizing BP neural network algorithm based on genetic algorithm[J]. *Artificial intelligence review*, 2011, 36(2): 153–162.
- [15] 丁建立, 陈增强, 袁著祉. 遗传算法与蚂蚁算法的融合[J]. *计算机研究与发展*, 2003, 40(9): 1351–1356.
DING Jianli, CHEN Zhenqiang, YUAN Zhuozhi. Genetic algorithm and ant algorithm fusion[J]. *Journal of computer research and development*, 2003, 40(9): 1351–1356.
- [16] WELCH G, BISHOP G. An introduction to the Kalman filter[R]. Chapel Hill, NC, USA: University of North Carolina at Chapel Hill, 1995.
- [17] 钟伟才, 刘静, 刘芳, 等. 二阶卡尔曼滤波分布估计算法[J]. *计算机学报*, 2004, 27(9): 1272–1277.
ZHONG Weicai, LIU Jin, LIU Fang, et al. Second order estimation of distribution algorithms based on kalman filter[J]. *Chinese journal of computers*, 2004, 27(9): 1272–1277.
- [18] TAKRURI M, RAJASEGARAR S, CHALLA S, et al. Spatio-temporal modelling-based drift-aware wireless sensor networks[J]. *IET wireless sensor systems*, 2011, 1(2): 110–122.
- [19] CHANG S G, YU Bin, VETTERLI M. Adaptive wavelet thresholding for image denoising and compression[J]. *IEEE transactions on image processing*, 2002, 9(9): 1532–1546.

作者简介:



武加文, 男, 1994 年生, 硕士研究生, 主要研究方向为无线传感器网络数据漂移的检测与校准。



李光辉, 男, 1970 年生, 教授, 博士生导师, 博士, 主要研究方向为无线传感器网络、无损检测技术等。主持国家 863 计划、国家自然科学基金重大研究项目、国家自然科学基金面上项目、科技部中小企业创新基金及省部级科研项目 15 项。获得浙江省科技进步二等奖 2 项、省部级奖励 3 项、浙江省农机科技奖一等奖及浙江省“科技兴林”奖一等奖各 1 项, 发表学术论文 60 余篇。

2019 年中国计算机学会人工智能会议征文通知

2019 CCF Conference on Artificial Intelligence (CCFAI 2019 Call for Papers)

主办: 中国计算机学会

协办: 中国计算机学会人工智能与模式识别专业委员会

承办: 中国矿业大学计算机科学与技术学院

中国计算机学会人工智能会议 (CCFAI, CCF Conference on Artificial Intelligence) 是由中国计算机学会主办, 中国计算机学会人工智能与模式识别专业委员会协办的人工智能领域盛会。该系列会议每两年举行一次, 现已成为国内人工智能领域最主要的学术活动之一, 旨在为人工智能领域从业者提供一个互动交流平台, 使参会者了解最前沿的学术动态, 分享人工智能领域的最新研究成果、创新思想和科学方法, 提高国内人工智能领域的研究水平。

2019 年中国计算机学会人工智能会议将于 2019 年 8 月 19—22 日在江苏省徐州市举行, 由中国矿业大学计算机科学与技术学院承办, 淮海工学院计算机学院协助承办。会议期间还将举行多场研讨会 (Workshop), 以及与 ACM 数据挖掘中国分会联合举行的“CCFAI 工业高峰论坛暨 ACM 数据挖掘中国分会年会”。本次会议将汇聚从事人工智能理论与应用研究的人员, 广泛开展学术交流, 研讨发展战略, 共同促进人工智能相关理论、技术及应用的发展, 并将邀请国内外人工智能领域的著名学者做精彩大会报告。

征文范围 (包括但不限于)

人工智能理论基础	人工智能应用	智能机器人
Agent 理论及应用	智能控制与智能管理	机器感知与虚拟现实
生物信息学与人工生命	机器学习	数据挖掘
时空知识表示、推理与挖掘	社会网络分析及应用	神经网络与计算智能
人工免疫	粗糙集与软计算	图像和语音处理
模式识别	知识科学与知识工程	自然语言处理和机器翻译
深度学习及其应用	面向大数据的人工智能技术	

重要日期

征文截止日期: 2019 年 3 月 31 日

录用通知日期: 2019 年 5 月 31 日

会议召开日期: 2019 年 8 月 19—22 日

论文要求

1、论文必须未公开发表过, 请勿一稿多投! 会议仅接受中文论文, 采用《计算机研究与发展》格式排版, 一般不超过 6 000 字。

2、论文应包括题目、作者姓名、作者单位、摘要、关键词、正文和参考文献。另附作者通讯地址、邮编、电话及 E-mail 地址。

3、学生 (不包括博士后和在职博士生) 第一作者的论文稿件请在首页脚注中注明, 否则将不具有参选“优秀学位论文”的资格。

4、会议网址为: <http://ccfai2019.cumt.edu.cn>

5、本会议采用在线投稿方式, 投稿地址: <https://easychair.org/conferences/?conf=ccfai2019>

论文出版

本次会议录用的论文将推荐到《中国科学》、《计算机学报》、《计算机研究与发展》、《模式识别与人工智能》、《计算机科学与探索》、《智能系统学报》、《计算机科学》、《计算机应用》、《南京大学学报》(自然科学版)、《山东大学学报》(工学版)、《山西大学学报》等期刊的正刊发表。

会议联系方式

1、联系人: 许新征 (手机: 15150023022) 袁力 (手机: 15252019560)

2、通讯地址: 江苏省徐州市大学路 1 号中国矿业大学计算机学院 221116

3、邮件联系: ccfai2019@163.com