

基于 ICA 独立成分和加权依赖贝叶斯的传感器节点故障诊断*

张新华

(太原大学 外语师范学院, 太原 030012)

摘要:传统的传感器节点故障诊断模型通常采用分布式模型或集中式模型,因此具有诊断效率低和扩展性差的缺点,为此,提出了一种基于分簇路由协议并结合集中式故障诊断和分布式故障诊断的混合式故障诊断模型。首先,在监测区域建立层次分簇路由协议和 WSN 节点故障诊断模型,然后,采用 ICA 独立成分分析法(Independent component analysis, ICA)对特征向量数据属性提取独立成分,以降低数据维数,从而获得最小属性集样本数据。最后,通过赋予各属性权值对朴素贝叶斯分类模型进行改进,得到加权依赖贝叶斯分类模型,并采用此模型实现节点故障诊断。仿真实验证明该模型能有效地进行故障诊断,与其他方法相比,具有故障诊断效率高和诊断精度高的优点,具有很强的可行性。

关键词:传感器节点;故障诊断;独立成分分析;贝叶斯

中图分类号:TP319

文献标志码:A

文章编号:1672-6693(2015)02-0138-05

无线传感器网络(Wireless sensor network, WSN)^[1-3]是由布置在某一监测区域实现数据采集和监测的传感器节点(Sensor node, SN)通过无线自组织的方式形成,节点在采集数据后通过 WSN 将数据发送到远程监控中心,目前已经成功应用于医疗服务、矿井安防、智能家居、环境监测等领域^[4-8]。

目前已有的对 WSN 节点进行故障诊断的主要方法是粗糙集提取规则诊断法、神经网络故障诊断方法、支持向量机故障诊断方法和数据融合比较故障诊断方法。文献[9]设计了一种基于粗糙集的故障诊断方法,通过粗糙集对样本数据进行规则提取,然后在汇聚节点处利用提取的规则在线判断节点是否发生故障,从而实现节点故障诊断,其缺点是粗糙集本身学习能力较弱,因此泛化能力不强。文献[10]设计了一种基于 Hamming 网的 WSN 节点故障诊断方法,首先将数据采用粗糙集进行离散化和属性约简,然后再采用 Hamming 网络进行故障诊断和分类,但 Hamming 网络存在易于早熟收敛的缺点。文献[11]建立了一种基于支持向量机的 WSN 节点故障诊断方法,在采用 ARIMA 对传感器故障征兆数据进行预测和估计的基础上采用多个二分类的支持向量机对传感器节点进行故障诊断。文献[12]建立了一种数据融合比较的节点故障诊断方法,通过收集邻居节点广播数据包并自己收集数据包比较,以判断节点是否发生故障。

上述工作具有重要意义,但仍具有样本需求量大、故障诊断精度不高和诊断时间长的缺点,为了解决上述问题,文中提出了一种基于独立成分分析和贝叶斯分类模型的故障诊断方法,并通过实验证明了文中方法的正确性。

1 WSN 节点故障诊断模型

以往的故障诊断模型主要采用分布式故障诊断模型和集中式故障诊断模型,分布式故障诊断模型是在所有节点处进行故障诊断,而集中式故障模型是在基站收集了所有节点的数据后进行故障诊断。分布式故障诊断方法具有扩展性强和鲁棒性高的优点,但其诊断效率低,而集中式故障诊断法方法诊断效率高但其扩展性较差。

针对集中式故障诊断和分布式故障诊断的缺点,文中设计了一种在簇头节点处进行簇内节点的集中式故障诊断和各簇头节点的分布式故障诊断的混合故障诊断模型。

故障诊断原理可以描述为:将监测区域的节点采用簇的形式进行组织,每个簇都选举出簇头进行簇内数据收集和集中式故障诊断,同时簇头节点之间相互传递数据进行分布式故障诊断,簇头在收集了簇内成员节点和

* 收稿日期:2014-03-27 修回日期:2014-12-19 网络出版时间:2015-01-22 11:56

作者简介:张新华,女,讲师,研究方向为计算机应用技术,E-mail: q100320115@126.com

网络出版地址: <http://www.cnki.net/kcms/detail/50.1165.N.20150122.1156.022.html>

其他簇头的数据后,首先对数据采用独立成分分析法进行属性约简,降低数据的维数,然后通过贝叶斯网络进行故障分类。

文中的诊断模型如图 1 所示。

2 ICA 独立成分分析获取最小属性集

为了减少故障诊断时间,需要减少故障特征向量的数据维数,采用 ICA 独立成分分析法(Independent component analysis, ICA)^[13]降低特征向量维数从而获得最小属性集,提高诊断效率。

ICA 独立成分分析方法是在 Kirby 等人于 1990 年首次提出的主成分分析法 PCA(Principal component analysis, PCA)的基础上发展而来,其主要原理是在不丢失数据的主要表征信息的前提下,将高维空间中的故障样本空间投影到独立的低维空间,采用一组独立基函数表示一系列随机变量,其模型如(1)式

$$\mathbf{X}=\mathbf{A}\mathbf{S}+\eta, \tag{1}$$

其中, \mathbf{X} 表示 m 维观察信号向量, \mathbf{A} 为 $m \times n$ 的未知混合矩阵, \mathbf{S} 为各分量独立的未知信号源, η 为噪声。因此, ICA 对数据降维要解决的问题就是在 \mathbf{A} 和 \mathbf{S} 未知的情况下,求取分离矩阵 $\mathbf{W}=\mathbf{A}^{-1}$,使得输出 \mathbf{Y} 是 s 的最优近似估计,如图 2 所示。

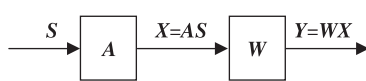


图 2 ICA 主成分分析模型

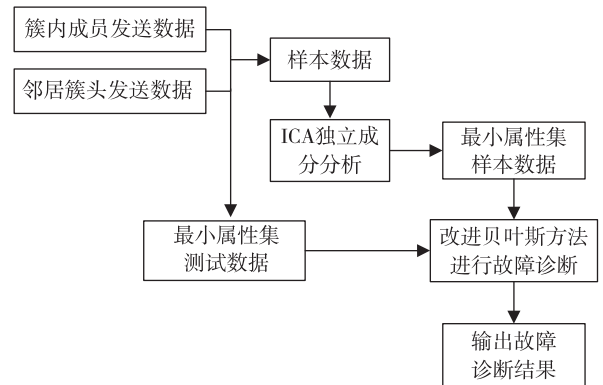


图 1 WSN 节点故障诊断模型

$$\mathbf{Y}=\mathbf{W}\mathbf{x} \tag{2}$$

根据快速 ICA 方法分离矩阵 \mathbf{W} 的第 i 行 ω_i 可以表示为

$$\omega_i = E\{z \tan h(\omega_i^T z)\} - E\{\tan h(\omega_i^T z)\} \omega_i, \tag{3}$$

其中, ω_i^T 为 ω_i 的行分量的转置。

因此,采用 ICA 独立成分法对故障诊断征兆数据进行降维的流程可以描述为:

1) 将原始样本数据排成输入矩阵 \mathbf{X} ,其中矩阵的行数对应了样本数据的个数,矩阵的列数对应了样本数据的维数;

2) 求取输入矩阵 \mathbf{X} 的特征值矩阵 \mathbf{D} 和特征向量矩阵 \mathbf{E} ;

3) 对数据进行中心处理:

$$\mathbf{X}' = \mathbf{X} - \mathbf{E}(\mathbf{X}). \tag{4}$$

在此基础上对 \mathbf{X}' 进行白化处理,得到矩阵 \mathbf{Z}

$$\mathbf{Z} = \mathbf{D}^{-1/2} \mathbf{E}^T \mathbf{X}, \tag{5}$$

其中, \mathbf{Z} 对应了主元空间;

4) 将代入(1)式作 ICA 映射,即 $\mathbf{Y}=\mathbf{W}\mathbf{Z}$,从而可以得到故障征兆特征数据在主元空间的映射, \mathbf{Y} 中的各行即为相互独立的故障征兆特征数据。

3 基于改进贝叶斯的故障诊断

3.1 朴素贝叶斯分类模型

贝叶斯网络是一种通过数据变量的依赖关系来进行推理的一种图形模式,目前已经在人工智能、概率理论、知识发现和决策分析等领域获取高度关注,是处理不确定知识表示和推理的有效理论和分析模型,非常适合从不完全、不确定和不确定的信息中进行表示和推理^[14-15]。

朴素贝叶斯分类器在基本贝叶斯理论的基础上,通过加入条件独立假设对未知数据样本预测,将其后验概率最高的类作为最终分类。

假设 $A = \{A_1, A_2, \dots, A_n\}$ 为属性变量集,其中, n 为样本数据的属性维数, $C = \{C_1, C_2, \dots, C_m\}$, m 为类别数,因此,样本数据可以表示为 $x_i = \{a_1, a_2, \dots, a_n, C_j\}$,即故障征兆数据为 $a_1, a_2, \dots, a_n, a_i$ 对应了属性 A_i 的取值,则故障征兆数据 a_1, a_2, \dots, a_n 属于分类 $C_j \in C$ 的后验概率可以表示为

$$P(C_j | a_1, a_2, \dots, a_n) = \frac{P(a_1, a_2, \dots, a_n | C_j) \cdot P(C_j)}{P(a_1, a_2, \dots, a_n)}, \tag{6}$$

其中,先验概率 $P(C_j)$ 可以表示为

$$P(C_j) = N_{C_j} / N, \tag{7}$$

其中, N_{C_j} 为分类 $C_j \in C$ 的样本数, N 为总样本数。

$P(a_1, a_2, \dots, a_n | C_j)$ 在假设条件独立的情况下,即各属性 A_1, A_2, \dots, A_n 相互独立时,可以表示为

$$P(a_1, a_2, \dots, a_n | C_j) = \prod_{i=1}^n P(a_i | C_j), \tag{8}$$

其中, $P(a_i | C_j)$ 为样本数据中类别为 C_j 时属性 A_i 的值为 a_i 的样本的概率,可以通过下式求取。

$$P(a_i | C_j) = \frac{1/N}{N_{C_j} + N_{C_j} / N}, \tag{9}$$

其中, N 为样本总数, N_{C_j} 为样本数据中类别为 C_j 时属性 A_i 的值为 a_i 的样本个数。

因此,采用朴素贝叶斯模型进行故障诊断即为对每个样本数据根据(6)式求取其属于各个分类的后验概率,将后验概率最大的分类作为故障类别。

3.2 加权依赖贝叶斯分类模型

加权依赖贝叶斯分类模型是在朴素贝叶斯模型的基础上,根据样本数据实际所属类与朴素贝叶斯模型分类结果的相符合程度对不同的属性赋权,其权值定义为

$$\omega_i = \frac{Diff(A_i)}{\sum_{k=1}^n Diff(A_k)}, \tag{10}$$

其中, $Diff(A_i)$ 表示样本数据属性值分类结果与实际分类的差异程序,可以通过(11)式求取。

$$Diff(A_i) = \frac{1-\alpha}{1-\beta}, \tag{11}$$

其中, α 为属性值对应的分类结果与朴素贝叶斯模型分类结果偶然一致的概率, β 表示属性值对应的分类结果与朴素贝叶斯模型分类结果完全一致的概率。

因此,通过将每个属性对应的贝叶斯模型分类结果进行加权,可以得到加权依赖的贝叶斯分类模型为

$$P(C_j | X) = \begin{cases} \frac{\sum_{s=1}^n \omega_s P(C_j, a_s) \prod_{k=1, k \neq s}^n P(a_k | C_j, a_s)}{P(X)}, & \text{当 } \sum_{s=1}^n \omega_s \neq 0 \text{ 时} \\ \frac{P(C_j) \prod_{k=1, k \neq s}^n P(a_k | C_j, a_s)}{P(X)}, & \text{其他} \end{cases}. \tag{12}$$

4 仿真实验

传感器节点通常是由能量供应模块、微控制器模块、存储器模块、通信模块等 4 个部分组成。因此,传感器节点的故障类别为能量供应模块故障、微控制器模块故障、存储器模块故障和通信模块等 4 类。

采用网络仿真软件 NS2 对文中方法进行仿真,在簇头节点处收到来自簇成员节点和其他簇头的 200 组数据后,将其中 100 组作为训练数据,另外 100 组作为测试数据。部分样本故障诊断征兆特征向量如表 1 所示。

表 1 故障诊断征兆特征向量

| 故障类型 | 样本数据 | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|
| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | x_8 |
| 无故障 | 0.81 | 0.56 | 0.68 | 0.56 | 0.72 | 0.68 | 0.63 | 0.62 |
| 能量供应模块故障 | 0.72 | 0.32 | 0.19 | 0.24 | 0.25 | 0.29 | 0.35 | 0.26 |
| 微控制器模块故障 | 0.65 | 0.30 | 0.28 | 0.21 | 0.31 | 0.60 | 0.37 | 0.64 |
| 存储器模块故障 | 0.37 | 0.35 | 0.68 | 0.23 | 0.29 | 0.29 | 0.62 | 0.28 |
| 通信模块故障 | 0.44 | 0.49 | 0.75 | 0.26 | 0.24 | 0.31 | 0.65 | 0.31 |

首先,采用 ICA 独立成分分析法对故障诊断征兆特征向量进行属性约简,得到属性简约子集为 $\{a[1], a[2], a[4], a[6], a[8]\}$,此时得到的经过属性约简后的最小属性集特征向量如表 2 所示。

表 2 约简的故障特征向量

| 故障类型 | 样本数据 | | | | |
|--------|-------|-------|-------|-------|-------|
| | x_1 | x_2 | x_4 | x_6 | x_8 |
| 无故障 | 0.81 | 0.56 | 0.56 | 0.68 | 0.62 |
| 电源故障 | 0.72 | 0.32 | 0.24 | 0.29 | 0.26 |
| 存储器故障 | 0.65 | 0.30 | 0.21 | 0.60 | 0.64 |
| 通信模块故障 | 0.37 | 0.35 | 0.23 | 0.29 | 0.28 |
| 控制器故障 | 0.44 | 0.49 | 0.26 | 0.31 | 0.31 |

表 3 节点故障诊断结果

| 故障类别 | 实际诊断故障数 | 文中方法诊断故障个数 | 文中方法诊断误差 | 文献[10]诊断故障个数 | 文献[10]诊断误差 | 文献[11]诊断故障个数 | 文献[11]诊断误差 |
|--------|---------|------------|----------|--------------|------------|--------------|------------|
| 无故障 | 10 | 10 | 0 | 9 | 1 | 8 | 2 |
| 电源故障 | 18 | 17 | 1 | 17 | 1 | 16 | 2 |
| 存储器故障 | 23 | 23 | 0 | 24 | 1 | 22 | 1 |
| 通信模块故障 | 17 | 18 | 1 | 15 | 2 | 19 | 2 |
| 控制器故障 | 19 | 19 | 0 | 21 | 2 | 21 | 2 |
| 错误数据 | 13 | 13 | 0 | 14 | 1 | 14 | 1 |

在获取了最小属性集的故障征兆特征向量测试数据后,采用改进的加权依赖朴素贝叶斯模型进行故障诊断,并与文献[10]和[11]进行对比,结果如表 3。

从表 3 中可以看出,文中方法实现了 WSN 节点故障诊断,故障诊断仿真总时间为 300 ms,而文献[10]和文献[11]所耗费的总诊断时间分别为 1200 ms 和 1400 ms,同时文中方法的故障诊断精确率为 98%,而文献[10]和文献[11]对应的故障诊断精度分别为 92%和 90%,较文中方法分别低 6%和 8%,因此,文中方法在具有较少诊断时间的同时具有较高的仿真精度,

这是因为文中故障诊断方法结合了分布式故障诊断和集中式故障诊断法方法的优点,在采用 ICA 独立成分分析的基础上,采用改进的加权依赖贝叶斯模型进行故障诊断和分类,因此,具有较小的诊断误差。

5 结语

为了实现监测区域的数据采集和有效监控,需要对监测区域中的节点进行故障诊断,以保证采集数据的可靠性。为此,设计了一种基于 ICA 独立成分分析和改进贝叶斯分类模型的 WSN 节点故障诊断方法。首先采用 ICA 独立成分分析法获取样本数据的最小属性子集,然后采用改进的加权依赖贝叶斯分类模型对样本数据进行故障诊断。在 NS2 仿真环境下进行实验,结果证明了文中方法能较为精确地实现传感器节点故障诊断,具有很强的可行性。

参考文献:

[1] Hong L, Xu S. Energy-efficient node placement in linear wireless sensor networks[C]//Proceedings of international conference on measuring technology and mechatronics automation. Changsha:[S. l.],China,2010.

[2] 顾晶晶,陈松灿,庄毅. 基于无线传感器网络拓扑结构的物联网定位模型[J]. 计算机学报,2010,33(9):1548-1555.

Gu J J, Chen S C, Zhuang Y. Wireless sensor network-based Topology structure for the Internet of things localization[J]. Chinese Journal of computer,2010,33(9):1548-1555.

[3] Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey[J]. Computer Networks,2008,52(12):2292-2330.

[4] 钟森,周小佳,闫斌. 基于移动锚节点的无线传感器网络三边质心定位[J]. 计算机测量与控制,2009,17(7):1438-1440.

Zhong S,Zhong X J,Yan B. Trilateral-centroid-localization in wireless sensor network with mobile anchor node[J]. Computer Measurement & Control,2009,17(7):1438-1440.

[5] 王粉花,年忻,郝国梁,等. 物联网技术在生命状态监测系统中的应用[J]. 计算机应用研究,2010,27(9):3375-3380.

Wang F H,Nian X,Hao G L,et al. Application of network of things technology on system of status detector[J]. Application Research of Computers,2010,27(9):3375-3380.

[6] 严萍,张兴敢,柏业超,等. 基于物联网技术的智能家居系统[J]. 南京大学学报:自然科学,2012,48(1):26-32.

Yan P,Zhang X G,Bai Y C,et al. Smart home based on Internet of things[J]. Journal of Nanjing University:Natural Science,2012,48(1):26-32.

[7] 周纯杰,黄雄峰,秦元庆,等. 无线传感器网络隧道施工监控系统设计与实现[J]. 计算机工程与设计,2011,32(7):2501-2504.

- Zhou C J, Huang X F, Qin Y Q, et al. Design and implementation of tunnel construction supervisory system based on wireless sensor network [J]. *Computer Engineering and Design*, 2011, 32(7): 2501-2504.
- [8] 吴学军, 雷雄. 基于低功耗器件的无线环境监测器设计 [J]. *电子技术应用*, 2011, 37(5): 138-145.
Wu X J, Lei X. Design of wireless environment monitor based on low power devices [J]. *Computer Technology and its Application*, 2011, 37(5): 138-145.
- [9] 沈敏, 张荣标, 冯友兵. 基于粗糙集的温室无线传感网络节点故障诊断 [J]. *江苏大学学报: 自然科学版*, 2009, 30(5): 510-513.
Shen m, Zhang R B, Feng Y B. Rough set based fault diagnosis of node in greenhouse wireless sensor network [J]. *Journal of Jiangsu University: Natural Science*, 2009, 30(5): 510-513.
- [10] 雷霖, 代传龙, 王厚军, 等. 粗糙集-神经网络集成的 WSN 节点故障诊断 [J]. *电子科技大学学报*, 2008, 37(4): 565-568.
Lei L, Dai C L, Wang H J, et al. Fault diagnosis of node in WSN based on integration of rough sets and neural network [J]. *Journal of University of Electronic Science and Technology of China*, 2008, 37(4): 565-568.
- [11] 李强. 一种基于诊断融合技术的传感器节点故障诊断 [J]. *科技通报*, 2013, 29(4): 79-81.
Li Q. A fault diagnosis of sensor node diagnosis based on data fusion technique [J]. *Bulletin of Science and Technology*, 2013, 29(4): 79-81.
- [12] 刘凯, 彭力. 分簇式无线传感器网络节点故障诊断算法研究 [J]. *传感器与微系统*, 2011, 30(4): 37-40.
Liu K, Peng L. Research on fault diagnosis algorithm for clustering node in wireless sensor network [J]. *Transducer and Microsystem Technologies*, 2011, 30(4): 37-40.
- [13] 刘嵩, 罗敏, 向军, 等. 基于小波变换和独立成分分析的人脸识别 [J]. *华中师范大学学报: 自然科学版*, 2012, 46(2): 166-169.
Liu S, Luo M, Xiang J, et al. Face recognition based on Wavlet transform and independant component analysis [J]. *Journal of Huazhong Normal University: Natural Science*, 2012, 46(2): 166-169.
- [14] 周忠宝, 马超群, 董豆豆, 等. 具有自学习功能的贝叶斯融合故障诊断方法 [J]. *计算机应用研究*, 2010, 5(27): 1764-1766.
Zhou Z B, Ma C Q, Dong D D, et al. Auto-study diagnosis method based on Bayesian fusion [J]. *Application Research of Computer*, 2010, 5(27): 1764-1766.
- [15] 李玲俐. 数据挖掘中分类算法综述 [J]. *重庆师范大学学报: 自然科学版*, 2011, 28(4): 44-47.
Li L L. A review on classification algorithm in data mining [J]. *Journal of Chongqing Normal University: Natural Science*, 2011, 28(4): 44-47.

Fault Diagnosis for Sensor Node in Wireless Sensor Network Based on Independent Component Analysis and Weighted Dependant Bayes

ZHANG Xinhua

(Foreign Language Normal School, Taiyuan College, Taiyuan 030012, China)

Abstract: The traditional diagnosis model for Sensor node usually using the distributed model or centralized model having the defects such as low efficiency and extension ability, therefore, a compound model combing the centralized fault diagnosis and distributed fault diagnosis based on clustering routing protocol is proposed. Firstly, the hierarchical clustering routing protocol and WSN node diagnosis model are built. Then the ICA (Independent component analysis, ICA) model is used to reduce the size of the data dimensions to get the sample data with lower dimensions. Finally, the Bayes model was improved by adding the attribute weight to get the weighted dependent Bayes classifying model, and the model is used to diagnose the nodes. The simulation in NS2 environment shows the method in this paper can diagnose the node diagnosis effectively, and compared with the other methods, it has the high diagnosis efficiency and high diagnosis accuracy, so it has strong feasibility.

Key words: sensor; fault diagnosis; independent component analysis; Bias

(责任编辑 游中胜)