

基于物联网的地下矿井空气质量智能预测

李湘文, 周辅杰, 崔 巍, 邓琴秀, 张辉雨
(成都理工大学 工程技术学院, 四川 乐山 614000)

摘 要: 为了提高安全性,全世界都在寻求实施无线传感器网络(WSNs)来监测复杂的、动态的和环境恶劣的地下煤矿。文中引入了一种可靠的物联网(IoT)空气质量监测系统,该系统由传感器模块、通信协议和基站组成。基于STM32的传感器模块具有八个不同的参数,安装在可操作的地下煤矿的不同位置。基于感知数据,该系统用煤矿环境指数(MEI)对地下煤矿矿井空气质量进行评价。采用主成分分析法确定了 CH_4 、 CO 、 SO_2 和 H_2S 是影响矿井空气质量最主要的气体。将主成分分析的结果输入到RNN神经网络模型中,实现了MEI的预测。结果表明,基于主成分分析的神经网络在MEI预测中具有较好的性能,主成分分析+RNN预测模型的性能指标 R^2 和RMSE值分别为0.489 0和0.120 4,提高了线性回归模型对矿井大气污染物的预测精度。因此,提出的基于STM32和Tensorflow平台的人工神经网络可以快速评估和预测矿井空气质量,提高矿井环境安全性。

关键词: 物联网; 矿井环境指数; RNN; 人工神经网络; 机器学习; TensorFlow

中图分类号: TD712

文献标识码: A

文章编号: 1673-629X(2020)08-0115-05

doi: 10.3969/j.issn.1673-629X.2020.08.019

Intelligent Prediction of Mine Air Quality Based on Internet of Things

LI Xiang-wen, ZHOU Fu-jie, CUI Wei, DENG Qin-xiu, ZHANG Hui-yu
(Engineering & Technical College, Chengdu University of Technology, Leshan 614000, China)

Abstract: In order to improve security, the world is seeking to implement wireless sensor network (WSNs) to monitor complex, dynamic and harsh underground coal mines. We introduce a reliable Internet of Things (IoT) air quality monitoring system which is composed of sensor module, communication protocol and base station. The sensor module based on STM32 has eight different parameters and is installed in different positions of operable underground coal mine. This system evaluates the air quality of underground coal mines with the coal mine environmental index (MEI). The principal component analysis method is used to determine that CH_4 , CO , SO_2 and H_2S are the main gases affecting mine air quality. The system inputs the results of principal component analysis into the RNN neural network model and realizes the prediction of MEI. The results show that the neural network based on principal component analysis has better performance in MEI prediction. Principal component analysis combined with RNN prediction model method reduces the error. Its error indices R^2 and RMSE are 0.489 0 and 0.120 4 respectively, which improve the prediction accuracy of linear regression model for mine air pollutants. Therefore, the proposed neural network based on STM32 and Tensorflow platform can quickly evaluate and predict mine air quality and improve mine environmental safety.

Key words: IoT (Internet of Things); MEI; RNN; ANNs (Artificial Neural Networks); machine learning; TensorFlow

0 引言

地下煤矿的恶劣的和狭窄的工作条件,导致采矿业被列为最危险的行业。中国是产煤大国,存在煤层储藏条件差,地下矿井安全监测设施不足和安全技术落后等问题^[1],因此,对地下煤矿环境进行持续、准确的监测显得非常重要。

近年来,无线传感器网络(WSNs)、射频识别(RFID)和云计算等领域的发展,引领了物联网(IoT)在智能电网、电子健康服务、家庭自动化、环境监测等领域的发展。基于IoT的实时监测能够实现准确、经济和可靠的环境监测。Jo和Khan^[2]在地下煤矿中引入了一个基于STM32的开源、经济高效的物联网系

收稿日期: 2019-08-31

修回日期: 2019-12-31

基金项目: 四川省科技计划项目(2019JDKP0020); 四川省教育重点科研项目(18ZA0066, 18ZA0068); 乐山市科技计划项目(18JCXY011); 成理工程科研基金项目(C122016002, C122018009, C122017021); 教育部高等教育司产学研合作协同育人项目(201802022027, 201802047128)

作者简介: 李湘文(1971-),男,副教授,硕士,CCF会员(C7076M),研究方向为物联网、人工智能。

统,用于早期预警和事件报告。Dixon 等^[3]将神经网络应用于瓦斯监测数据中,预测矿井环境中甲烷气体的浓度。这些文献侧重于煤矿环境监测的无线传感器的实施。

矿井空气质量评价与预测对提高矿井安全起着至关重要的作用。这不仅可以减少矿工与恶劣空气的接触,还可以有效控制矿井通风。然而,空气质量变量的复杂和非线性行为超出了简单的数学预测公式的能力。在这方面,各种统计工具,如 ARIMA、多重线性回归 (MLR)、人工神经网络 (ANN)、模糊时间序列和主成分分析 (PCA) 在准确评价空气质量和预测污染物浓度方面都取得了令人满意的结果。有毒气体浓度是影响空气质量的主要因素,因此,确定特定环境中存在的污染物的类型和数量是非常重要的。在这方面,统计工具主成分分析显示了识别影响污染物的高能力,而人工神经网络使预测更加准确。因此,在地下煤矿的恶劣环境中,主成分分析和人工神经网络的混合方法可以有效地识别污染物,并给出准确的预测。

近几十年来,各种科学研究使用多变量统计方法,如聚类分析 (CA)、主成分分析、因子分析 (FA)、判别分析等来解决环境和空气质量问题。然而,基于特征

值求解的主成分分析法是最常用和最简单的方法。具体来说,在空气质量问题上,它已被单独使用或与其他方法结合使用。在目前的研究中,采用主成分分析对矿井环境中存在的主要污染源进行识别。

近年来,神经网络在工程、工业过程控制、医药、计算机、风险管理和营销等领域显示出巨大的潜力,循环神经网络在复杂环境下显示了高精度预测能力^[4]。文中旨在开发一个可靠、具成本效益及高效的物联网系统,利用机器学习进行空气质量预测。

1 系统设计

1.1 系统硬件架构

该系统专为地下矿井空气质量监测和评估而设计。系统基本架构如图 1 所示。主要框架包括数据采集、数据传输、空气质量评价与预测的数据处理。传感节点单元连接到基于 STM32 处理器的传感器模块,传感节点捕获空气质量数据,并通过 LoRa 无线技术将数据通过基站传输到服务器分析处理数据, Tensorflow 作为神经网络平台进行矿井空气质量预测。

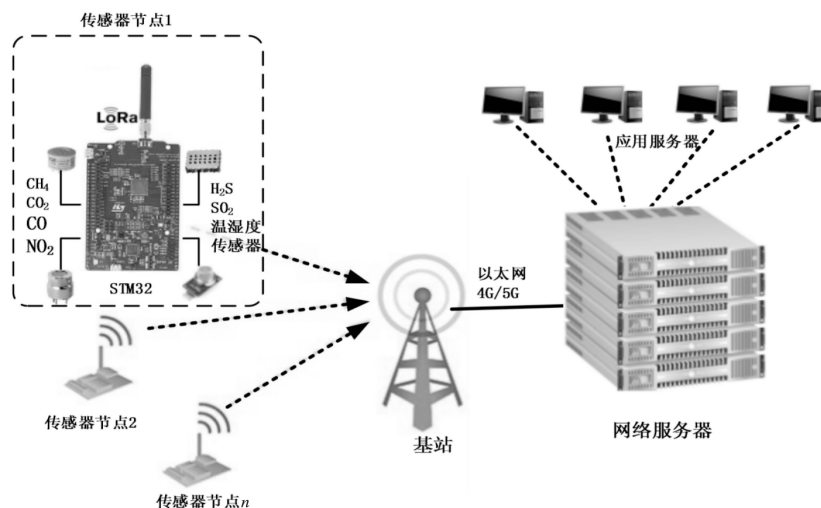


图 1 系统硬件架构

传感器节点:传感器的基本功能是测量矿井环境中的空气参数。传感器节点由传感器模块、微控制器和无线发射机组成。选择合适的传感器来监测矿井环境是一个相对复杂的问题,它需要综合考虑测量范围、精度和灵敏度等因素。DHT11 是一款有已校准数字信号输出的温湿度传感器。其精度湿度 $\pm 5\% \text{ RH}$, 温度 $\pm 2^\circ\text{C}$, 量程湿度 $20\% \sim 90\% \text{ RH}$, 温度 $0^\circ\text{C} \sim 50^\circ\text{C}$ 。通常情况下,煤矿的工作温度在 15°C 到 45°C 之间变化,湿度在 DHT11 的特定范围内。这些电位使得 DHT11 非常适合在地下煤矿中使用。地下煤矿中常见的气体有 CH_4 、 CO_2 、 CO 、 NO_2 、 H_2S 和 SO_2 。文中利

用 MQ-4、MQ-9、MQ-811、MQ-136 和 MiCS-2714 传感器模块来监测各种气体的浓度。其中,大部分传感器模块为金属氧化物 (SnO_2),对挥发性气体分子反应良好;因此,它们对于气体监测更加可靠和有效。此外,传感器模块,无论是用于气体监测还是用于温度测量,都是经济有效、低功耗、稳定的。文中以内嵌 LoRaWAN 的 STM32 单片机为基础,用 MQ-4、MQ-9、MQ-811、MQ-136、MiCS-2714 和 DHT11 传感器模块来监测煤矿中常见的气体 CH_4 、 CO_2 、 CO 、 NO_2 、 H_2S 和 SO_2 以及温湿度。

通信协议:LoRa 无线通信协议已被证明在矿井中

具有高透射率、稳定性、超低功耗和高通信性能。LoRaWAN 是为 LoRa 远距离通信网络设计的一套通讯协议和系统架构。一个 LoRaWAN 网络架构中包含了终端、基站、NS(网络服务器)、应用服务器这四个部分。由若干终端 end nodes 采集底层数据,通过(RF 射频技术/LoRaWAN 技术)把数据发送到网关(Gateway)基站,再通过(4G/5G/以太网通信技术)发送到 NS(网络服务器),网络服务器(NS)^[5]再把数据发给应用程序端(APP)。基站则对 NS 和终端之间的 LoRaWAN 协议数据做转发处理,将 LoRaWAN 数据

分别承载在了 LoRa 射频传输和 Tcp/IP 上。网络服务器再把数据发给应用程序端 APP。

1.2 机器学习 Tensorflow 平台

文中将 Tensorflow 应用于矿井空气质量循环神经网络模型的开发、训练和测试上^[6]。Tensorflow 是用于数值计算的开源软件库,可以在创建深度学习网络时对数据进行数值和图形计算。用 Tensorflow 建立循环神经网络(RNN)模型^[7],煤矿环境指数(MEI)的多层感知循环神经网络的网络架构^[8]如图2所示。

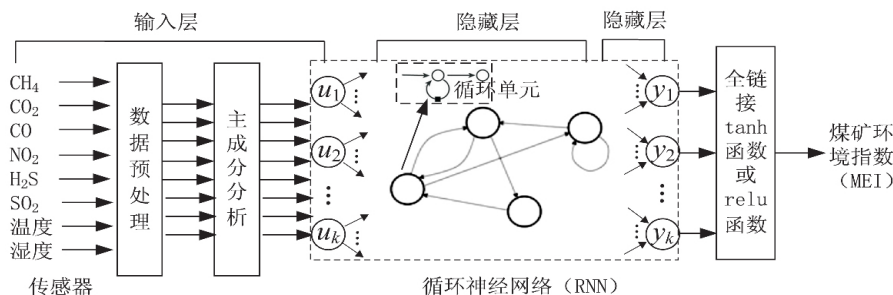


图2 煤矿环境指数(MEI)的多层感知循环神经网络的网络架构

2 方法和模型

2.1 矿井环境指数(MEI)

由于开采过程中煤层中排放的 CH_4 、 CO_2 、 CO 、 NO_2 、 H_2S 和 SO_2 等气体的存在^[9],地下煤矿的环境相对恶劣。露天或室内空气质量指标不足以完全代表井下空气质量。应该有一个能真实反映矿井环境并能轻易评估矿井空气质量的可用指标。为了快速评估和简

化矿井空气质量的解释,文中引入了矿井环境指数(mine environmental index, MEI)。

$$\text{MEI}_p = (\text{MEI}_{\text{Hi}} - \text{MEI}_{\text{Lo}}) \times C_p \quad (1)$$

其中, MEI_p 为污染物 p 的指标值, MEI 值分为五类:非常好,良好,中等,差和非常差; C_p 为给定污染物 p 的输入浓度。

各种气体对矿山空气质量指数(MEI)如表1所示。

表1 各种气体对矿山空气质量指数(MEI)(气体浓度(ppm))

状态	NO_2	CO	SO_2	H_2S	CH_4	CO_2
优级	0-1.0	0-13	0-2.5	0-3.0	0-1 000	0-2 000
良好	1.0-2.0	13-23	2.5-4.0	3.0-5.0	1 000-2 000	2 000-3 000
中等	2.0-3.0	23-30	4.0-6.0	5.0-13	2 000-4 000	3 000-4 000
差	3.0-4.0	30-49	6.0-8.0	13-20	4 000-5 000	4 000-5 000
非常差	>4.0	>49	>8.0	>20	>5 000	>5 000

2.2 数据预处理

在实施任何统计方法之前,有必要对收集到的数据进行预处理,以便提取传感器响应的实质特征,从而产生用于进一步处理的特征。文中将转换作为预处理的一个初始步骤,使用平均值为0,标准差为1的Z变换进行,Z变换计算公式为:

$$Z_{ij} = (X_{ij} - \mu) \times \sigma \quad (2)$$

其中, Z_{ij} 为变量 i 的第 j 个值, X_{ij} 为变量 i 的第 j 个观测值, μ 为平均值, σ 为标准差。

Z尺度变换确保任何统计过程中变量的权重相等。这种转换均匀了分布方差,降低了由于数据集大小不同而产生任何错误的概率。

2.3 主成分分析建模

主成分分析(principal component analysis, PCA)^[10]是最有价值的统计方法之一,它将 N 维数据压缩并转换成一个新的 N 维数据集。PCA 通过线性变换将原始数据变换为一组各维度线性无关的表示,可用于提取数据的主要特征分量,用于高维数据的降维^[11]。这项研究使用主成分分析方法来确定矿井环境中最重要的污染物。

2.4 循环神经网络(RNN)建模

人工神经网络是人工智能中应用最广泛的信息处理系统,神经网络由几个相互连接的神经元组成,能够根据内部或外部数据改变其结构。循环神经网络可以

训练非线性和复杂数据^[12],可以在时间序列上有延拓,便于分析连续的数据,主成分分析的输出被用作循环神经网络模型的输入^[13],以准确地预测空气质量。将主成分分析结果与人工神经网络相结合,识别出对环境影响显著的污染源。在循环神经网络的网络中,第一层是输入层,负责信息的收集、误差的消除,并将数据传输到循环神经网络结构^[14-15]。第二层是隐藏层,有任意数量的神经元和若干层。神经元在前馈和后向传播的过程中被激活,从而提供各层之间的连接。隐藏层在输入层和输出层之间接收和传输数据^[16]。图 2 显示了使用主成分分析输入来确定 MEI 的循环神经网络体系结构。

循环神经网络模型的准确性和训练能力高度依赖于最佳神经元数量的选择。在本例中,RNN 是用 TensorFlow 设计的,其权重由随机化函数决定。模型使用 80% 的整个数据集,剩余 20% 的数据集用于测试。

3 实验结果和讨论

3.1 空气污染源识别

课题组实验场所在山西石泉煤业某矿井正在运营地下煤矿的主要巷道内,拱形直巷道的尺寸为 1.8 米×2.2 米,传感器模块单元安装到矿井隧道洞口的屋顶中心,所有传感器模块都通过 LoRa 协议连接到基站,巷道共安装 32 个传感器,数据采集率设定为 5 分钟。每个传感器每天有 288 组数据,15 天收集了 138 240 组样本,在这些数据集中,110 592 组数据被用来训练模型,并且对 27 548 组数据进行测试。在完成训练后,对超过一个月的 MEI 数据集进行了预测。

主成分分析的输入是 8 个不同变量的传感器读数。主成分分析提取给出了主要成分。

文中将绝对值大于 0.85 的变量因子(表示相对较强的荷载)设定为阈值限制值。主成分分析提取给出的主要成分分别为:湿度: -0.746; CH₄: 0.912; CO₂: 0.835; CO: 0.059; NO₂: 0.768; SO₂: 0.872; H₂S: 0.926。这表明了满足 0.85 阈值限制条件的成分。这些成分 CH₄、CO₂、SO₂和 CO 以及温度是该矿空气中的主要污染因子。在这些污染物中,对温度、CH₄、CO₂、CO 和 SO₂的负荷因子贡献率为 65.16%。这些影响因素主要与地下煤矿矿井环境中常见的气体有关。在这些气体中,CH₄和 SO₂通常以气穴的形式被煤层所限制。在煤炭开采过程中,这些矿坑爆裂,并向矿井空气中释放气体。矿井环境中 CO₂的主要原因可能是工人的呼吸活动,以及任何柴油机械的排气。

3.2 预测结果

为了建立循环神经网络模型,以原始成分和主成分分析提取成分为输入,测试了不同数量神经元的网

络结构。在这些循环神经网络结构中,神经元的最佳数目是根据 MAE(平均绝对误差)、RMSE(均方根误差)、RSE(残差的标准误差)和 R²(确定系数)来确定的。神经元的数量逐渐增加,用主成分分析的情况下,原始和主成分分析提取成分作为输入的神经元数量不同的验证模型如表 2 所示。对于原始数据集,给出最小误差的最佳神经元的数目为 20,而对于主成分分析提取的数据集,神经元的最佳数目为 8。

表 2 原始数据集为输入,主成分分析提取成分为输入,对不同神经元进行模型验证

模型特征	隐藏节点数量	MAE	RMSE	RSE	R ²
原始数据集	12	0.208 4	0.285 4	0.685 4	0.681
	14	0.189 7	0.254 6	0.625 6	0.624 8
	18	0.167 2	0.242 1	0.558 7	0.564 8
	20	0.173 4	0.247 8	0.612 7	0.524 7
	22	0.187 4	0.252 3	0.623 8	0.538 5
主成分分析	6	0.216 9	0.256 6	0.681 2	0.726 4
	7	0.186 5	0.217 8	0.648 9	0.615 9
	8	0.161 4	0.185 1	0.587 1	0.520 4
	9	0.174 6	0.213 4	0.602 4	0.515 6
	10	0.182 7	0.224 8	0.658 9	0.522 3

将该模型与多元线性回归(MLR)、主成分分析、循环神经网络定义的模型进行了性能比较。对于所有这些模型,MAE、RMSE、RAE 和 RSE 都进行了比较,因为接近零的错误值表示一个更好的模型。另一方面,通过计算确定系数(R²)来检验模型的准确性。在这些试验中,模型的高精度值接近 1。因此,模型的精度可能因预测所需的时间间隔而异。对于地下矿井,存在一些局限性;因此有必要预测瓦斯浓度。最佳预测模型为主成分分析-循环神经网络(RNN),MAE、RMSE 和 RSE 分别为 0.085 2、0.120 4 和 0.510 8。此外,测定系数 R² 为 0.489 0; MEI 预测模型的性能指标如表 3 所示。

表 3 MEI 预测模型的性能指标

神经网络模型	MAE	RMSE	RSE	R ²
LR 线性回归	0.229 1	0.275 4	0.575 8	0.544 9
主成分分析	0.174 6	0.185 1	0.587 1	0.520 4
循环神经网络(RNN)	0.112 5	0.143 8	0.624 1	0.527 8
主成分分析+RNN	0.085 2	0.120 4	0.510 8	0.489 0

主成分分析将线性回归模型的准确度提高了约 2.23%,在循环神经网络(RNN)的情况下提高了 17.8%。结果表明,主成分分析与 RNN 相结合,提高了矿井大气污染物的预测精度。因此,主成分分析的

应用有助于与空气质量有关的预测研究。

4 结束语

近年来,人工智能计算能力彻底改变了复杂和非线性数据问题的解决方案。地下煤矿中存在的空气污染物总是呈现非线性,最终导致预测不确定,预警可靠性降低。因此,提出的基于物联网的矿井空气质量监测、评估和预报系统,利用 Tensorflow 预测空气质量,并有可能通过预警广泛提高地下矿井安全性。该系统安装在一个运行中的煤矿井下,并提出空气质量评估模型来确定 MEI,提高地下矿井安全性。IoT 监测煤矿环境参数,主成分分析结果表明,四种矿井气体(CH_4 、 SO_2 、 CO 和 H_2S)对矿井空气质量影响最大。与单线性回归和 RNN 模型相比,所提出的 RNN+主成分分析模型预测 MEI 更准确。该研究表明,经过适当训练的 RNN 模型可以有效地预测 MEI。

该系统通过预处理传感器,有效缓解了矿井环境变量的非线性行为,如温度、湿度、 CO_2 、 CH_4 、 CO 、 SO_2 和 H_2S 数据。数据作为人工神经网络模型的输入数据集上传到 Tensorflow 计算。主成分分析结果表明,四种矿井气体(CH_4 、 SO_2 、 CO 和 H_2S)对矿井空气质量影响最大。神经网络的多层感知器能准确地预测 MEI。由于人工神经网络模型的准确性和有效性与输入参数和隐藏层神经元数目高度相关,因此可以通过观测最小误差来确定隐藏层神经元的最佳数目。

尽管实验结果表明预测准确,但仍存在一些局限性。这些限制包括地下煤矿的恶劣环境、数据隐私以及多传感器输出的集成。此外,文中只考虑了八个空气质量参数,忽略了可能对矿山环境造成更严重影响的参数。文中依赖于随时间变化的空气污染物浓度,而其他与预测效果相关的因素则被忽略。未来还需要在以下方向继续探究:首先,确定污染物的复杂非线性行为;浓度需要一个更精确的混合模型,以增强早期预警。其次,高污染物浓度是影响空气质量的主要因素,因此,需要建立一个预测大气污染物峰值浓度的模型。

参考文献:

[1] 李隘优.基于物联网的煤矿移动安全监控系统设计[J].闽

西职业技术学院学报,2017,19(4):114-116.

- [2] 金文,姚凯学.基于物联网的农业墒情监测系统的设计与实现[J].计算机应用与软件,2018,35(3):84-88.
- [3] 《中国公路学报》编辑部.中国交通工程学术研究综述·2016[J].中国公路学报,2016,29(6):1-161.
- [4] ZHOU L, LI X W, ZHANG S, et al. Power balance and control in optically controlled phased array antenna based on optical true time delay[J]. Open Automation & Control Systems Journal, 2015, 7(1): 987-990.
- [5] LI X, LI L, ZHOU S, et al. Model design of indoor visible light MIMO communication system based on space shift keying[J]. The Open Cybernetics & Systemics Journal, 2015, 9(1): 1051-1054.
- [6] 刘蕾.基于 Tensorflow 的循环神经网络模型在上海市空气质量预测中的应用[D].上海:上海师范大学,2019.
- [7] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.
- [8] ZHU D, XIONG K, XIAO H, et al. Variation characteristics of rainfall erosivity in Guizhou Province and the correlation with the El Niño Southern Oscillation[J]. Science of the Total Environment, 2019, 691: 835-847.
- [9] WANG X, DAN B, MU C. Study on formation mechanism of gas emission partition in hydraulic flushing coal seam[J]. Journal of China Coal Society, 2012, 37(3): 467-471.
- [10] PÉREZ-ARRIBAS L V, LEÓN-GONZÁLEZ M E, ROSALES-CONRADO N. Learning principal component analysis by using data from air quality networks[J]. Journal of Chemical Education, 2017, 94(4): 458-464.
- [11] 郎江涛.面向高维数据的 PCA-Hub 聚类方法研究[D].重庆:重庆大学,2017.
- [12] 艾玲梅,叶雪娜.基于循环卷积神经网络的目标检测与分类[J].计算机技术与发展,2018,28(2):31-35.
- [13] 张宪峰,魏久传,张延飞,等.基于主成分分析与 BP 神经网络的矿井涌水量预测研究[J].煤炭技术,2018,37(6):201-203.
- [14] 段豪.基于 Android 的健康监测应用研究与实现[D].成都:电子科技大学,2017.
- [15] BALLUFF S, BENDFELD J, KRAUTER S. Meteorological data forecast using RNN[J]. International Journal of Grid and High Performance Computing, 2017, 9(1): 61-62.
- [16] 宋利红.基于深度学习的雾霾预测方法研究[D].成都:电子科技大学,2018.