

基于 ZigBee 的室内空气质量感知系统的设计

叶 贵, 张林静

(安徽警官职业学院 信息管理系, 合肥 230031)

摘要: 物联网技术的快速发展, 给人们的生活带来越来越多的应用; 因此, 对无线传感器网络节点部署到室内, 实时采集室内环境数据, 并对数据进行相应的融合和处理, 分析出室内环境的动态变化情况, 并且能够预测每个时间段的环境状况, 并将其分析结果实时反馈给用户, 提醒当前的室内环境问题; 首先, 采集室内的温度、湿度、PM2.5 和 VOCs, 利用 ZigBee 无线通信技术作为传感器网络的通信基础, 同时利用各传感器实时获取室内环境物理数据, 通过多元数据的多元线性回归算法对室内环境进行有效监控和评估, 通过实验对数据的分析最终在多元线性回归算法中拟合出的均方根误差 RMSE 为 0.049 5, 能够预测出室内各项因素的变化情况。

关键词: ZigBee; 感知环境质量; 传感器; 多元数据回归

Design of Indoor Air Quality Perception System Based on ZigBee

YE Gui, ZHANG Linjing

(Department of information Management, Anhui Vocational College of Police Officers, Hefei 230031, China)

Abstract: With the rapid development of the Internet of Things technology, more and more applications are brought to people's lives. Therefore, wireless sensor network nodes are deployed indoors, real-time collection of indoor environmental data, and corresponding fusion and processing of the data to analyze the dynamic changes of the indoor environment, and can predict and analyze the environmental conditions in each time period. The results are fed back to the user in real time, reminding the current indoor environmental issues to be studied. First, collect indoor temperature, humidity, PM2.5 and VOCs, utilize ZigBee wireless communication technology as the communication basis of the sensor network. At the same time, each sensor is used to obtain physical data of the indoor environment in real time. For effective monitoring and evaluation, the data is analyzed through experiments and the root mean square error RMSE is finally fitted in the multiple linear regression algorithm to be 0.049 5, which can predict the changes of various indoor factors.

Keywords: ZigBee; perceived environmental quality; sensor; multiple data regression

0 引言

随着物联网、大数据、人工智能技术的快速发展, 诞生了各种应用产品应用在人们的日常生活中。现在社会中出现了很多关于智能家居的产品, 其中包括了各种家电, 常见的有智能电视、音响、空调、冰箱、电灯等, 都会给我们带来很多的便利。根据调查显示, 人们每天在室内的生活时间, 大概占到一天总时间的 90% 以上。根据世界卫生组织报告, 人类感染疾病与室内环境质量有密切相关。文献 [1] 研究, 发现了关于室内环境的污染源预测, 室内环境对人们身体健康影响极大, 因此对室内环境质量的研究具有重要意义。该系统是一款智能硬件设备, 采集数据、数据传输、数据的融合与分析, 利用无线传感器网络节点的部署到室内, 并通过传感器采集室内的环境物理信息 [2-3], 进而对数据进行相应的处理和分析, 最后可以实时

得出室内环境的动态变化情况, 并且可以预测出一天中每个时间段的环境信息的变化情况, 并将其分析结果实时反馈给用户, 提醒用户当前的室内环境状况。很好地给用户掌握室内环境的各物理信息的指标, 从而极大地提高了室内的环境质量 [4-7], 保证了人们的身体健康。

实时获取室内环境信息, 并对得到的多元数据进行融合处理并通过多元线性回归算法分析和预测, 将当前的室内环境空气质量实时显示给用户, 从而大大降低了室内环境对人们造成的身体危害程度, 保证了人们在室内有一个舒适健康的环境指标。因此, 对该课题的研究具有重要的意义。

1 系统架构与原理

该系统的设计主要包括 3 部分: 无线传感器节点、协调器和多元数据融合算法模块。该系统的设计系统架构如

收稿日期: 2021-04-14; 修回日期: 2021-06-16。

基金项目: 2020 年安徽省自然科学基金重点项目 (KJ2020A1058)。

作者简介: 叶 贵 (1992-), 女, 安徽肥西人, 硕士研究生, 助教, 主要从事物联网、软件应用等方向的研究。

引用格式: 叶 贵, 张林静. 基于 ZigBee 的室内空气质量感知系统的设计 [J]. 计算机测量与控制, 2021, 29(12): 161-165.

图 1 所示。该系统采用 ZigBee 无线通信技术用于系统的组网和通信基础，并利用各类传感器进行室内环境数据的采集，将各类数据经过无线传感器网络节点进行数据传输到协调器，最后传回到云端，在云端进行数据处理和多元数据的融合，将几种数据进行融合分析，最终得到对室内环境的评价指标。

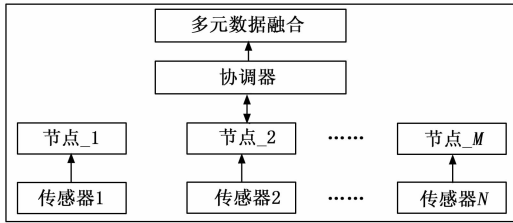


图 1 系统结构图

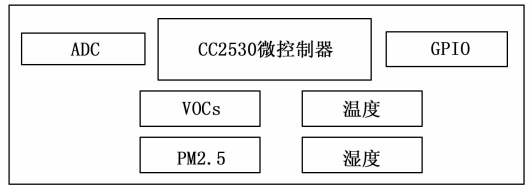


图 2 节点硬件示意图

通过与蓝牙、wifi 等相关的无线通信技术对比，选择 ZigBee 做为本系统的设计基础，因为其具有以下优点：

- 1) 抗干扰力强：采用基于 IEEE802.15.4 的通信协议，ZigBee 收发模块使用的是 2.4G 直序扩频技术，具有更好的抗干扰能力；
- 2) 传输距离：传输距离相对于蓝牙和 wifi 具有较远点的距离；
- 3) 低功耗：无线模块自动每隔一段时间，如果没有收到网关指令，就进入休眠状态；
- 4) 组网灵活：ZigBee 是自组网方式进行组网，可以是树状、网状、星状不同形式；
- 5) 传输速度快：ZigBee 传输数据主要是采用短帧传送方式，因此，传输数据速度快，实时性强；
- 6) 鲁棒性高：ZigBee 组网容易且灵活，如果某一个节点断电或损坏，整个网络可以自动恢复组网方式。

该系统主要分为无线传感器节点、协调器、数据融合 3 部分。

1) 无线传感器终端节点：该无线传感器节点采用开源的 ZigBee 通信技术^[8]，并使用德州仪器的 Z-Stack 半开源的协议栈^[12-14]，对其应用层进行相应的开发，我们知道 ZigBee 具有一些优点，其中包括传输距离远、低功耗、节点的组网灵活。其中传感器包括了 PM2.5 传感器、VOCs 传感器及 DHT11 温湿度传感器。这些传感器都与终端相连，并由终端的 CC2530 芯片对其进行控制。PM2.5 传感器采集室内的细微颗粒的浓度。

VOCs 传感器是采集室内有机气体的浓度；温湿度采集室内温度和湿度情况。该 CC2530 芯片对传感器的连接方式^[9-11]如图 2 所示。

2) 协调器：协调器是系统的无线通信网络模块的核心，主要是传输数据和组网的纽带，其组网采用了星型网络拓扑结构，有 4 个终端与一个协调器进行组网。并将获取到的数据通过协调器和串口线与电脑端的用户软件通信，

而无线传感器节点的终端都是采用 ZigBee 无线通信技术进行通信。在这里，协调器启动必须首先启动，进而各个终端部分开始启动，只用这种方式，系统的组网才是成功的，不然会失败，这是由 ZigBee 通信协议所决定。

多元数据融合：将 4 类传感器采集的数据进行融合得到结果。本文考虑到系统的实时性和有效性采用多元线性回归算法对数据进行建模，并对数据进行预测和判别不同显著性因素的影响。

另一方面，将数据进行有效分析处理的用户软件界面应用。其中驱动应用软件，是对 CC2530 芯片与传感器的相关应用，其中包括对传感器的控制和数据采集的逻辑业务软件。还包括对 Z-Stack 协议栈的开发，主要是负责对数据的发送和接受相关业务应用的软件。

本系统主要是采集室内 4 个地方的环境，分别是客厅、卧室、厨房和卫生间，每个场地部署若干个无线传感器终端节点进行室内环境数据的采集。

2 系统硬件与软件算法设计

本章主要对系统的硬件和软件以及在多元数据融合 3 个方面进行详细地说明。首先，系统的硬件设计主要描述该系统采用的硬件类型，何如设计和工作；其次，软件设计是实现该工程的应用，包括通信协议栈的使用和传感器的程序设计；最后，多元数据融合算法本文从多元线性回归算法^[15-19]详细介绍。

2.1 硬件模块设计

系统为了组网灵活和低功耗等特点，选择使用 ZigBee 无线通信技术作为该系统的通信基础平台，由于目前 ZigBee 应用在很多无线传感器网络的应用领域中，有着很好的稳定性和鲁棒性。该系统是采用德州仪器公司生产的 CC2530 的微控制器作为该系统的主芯片。同时，利用 Z-Stack 半开源协议栈，进行有效地开发相应的应用程序代码，并且可以熟练掌握该通信技术的方式。根据 CC2530 芯片手册，对其进行开发和外围电路的设计。该芯片是 8-bit 微控制器采用的类型是 CC2530F256，它是 8051 内核，具有 flash 存储为 256 KB，RAM 为 8 KB 的容量。

该无线传感器节点的硬件设计如图 3 所示，其中包括了主控制单元 CC2530、时钟电路、电源电路、外围设备电路（提供传感器使用的 IO 端口）等。

室内环境的物理信息采集通过不同种类的传感器，每一种传感器都是连接到 CC2530 控制板的 GPIO (general-

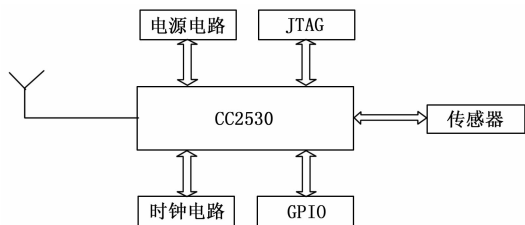


图 3 硬件设计示意图

purpose input/output)。通过主控芯片来控制各类传感器的工作、数据采集、传输。对于采集到的不同数据进行对应的编码格式,通过 Z-stack 协议进行传输。JTAG (joint test action group) 接口模块是为了给主控芯片下载和调试程序使用。时钟电路模块是为了让单片机按照一定频率进行工作。电源电路则是为了让每一个模块提供电力,不同模块需要的电路是不一样的,有的需要 3.3 V 的电压供电,有的需要 5 V 电压供电。这些电压都是通过稳压芯片进行稳压后才为后续的几个模块使用。传感器模块这部门是按照不同类型的传感器进行与主控板的 GPIO 端口进行连接,以及根据具体需要提供不同的电源。

2.2 软件模块设计

利用 ZigBee 无线通信技术原理和传感器技术,对室内的环境信息进行采集,并将获取的数据进行有效的分析。软件设计主要是利用 ZigBee 通信技术的 Z-stack 通信协议栈完成的,在该协议栈中,开发出属于本系统的应用功能模块,就可以完成整个系统的功能。主要是对不同传感器的初始化和采集数据以及控制其采集编码方式等。软件的整个伪代码如下所示:

相关软件实现的伪代码

```

PROCEDURE
Begin
HW_INIT();
Sensor_Init();
While(true) do
Sample_APP();
Z_stack_Func();
AF_DataRequest();
zb_SendDataRequest();
End while;
End;

```

CC2530 主控板单片机的通信应用类软件是采用 IAR 工具进行程序编写、编译、调试、烧录的。主要是在 Z-stack 通信协议栈进行的应用程序的开发,其中就包括各种传感器的初始化、数据采集的相关代码的实现,最终由协议栈完成数据的路由与发送到协调器,然后协调器连接到电脑端,通过串口进行传输采集到的数据。

此外,还有另外一部分软件是对数据处理、分析以及

训练的代码,该部分的代码在后序进行详细介绍。

2.3 多元线性回归算法

对于 n 维特征的样本数据,使用多元线性回归,那么对应的模型可以用公式 (1) 的形式表示:

$$H_{\theta}(x_1, x_2, \dots, x_n) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \quad (1)$$

可以把公式 (1) 简写成下公式 (2):

$$H_{\theta}(x_0, x_1, x_2, \dots, x_n) = \sum_{i=0}^n \theta_i x_i \quad (2)$$

其中: $x_0 = 1$ 。对于样本采集的数据可以用矩阵 \mathbf{X} 表示维度为 $m \times n$, 参数 θ 表示维度为 $n \times 1$, 则公式 (1) 可以采用 $H_{\theta}(\mathbf{X}) = \mathbf{X} * \theta$ 代替。在使用该模型的时候,其损失函数一般采用均方误差作为损失函数,这样可以有效地求解出模型的最优解。

损失函数可以用公式 (3) 表示:

$$L(\theta_0, \theta_1, \dots, \theta_n) = \sum_{i=0}^m (H_{\theta}(x_0, x_1, x_2, \dots, x_n) - y_i)^2 \quad (3)$$

同理,可以将公式 (3) 表示为矩阵的形式,如公式 (4) 所示:

$$L(\theta) = \frac{1}{2} (\mathbf{X}\theta - \mathbf{Y})^T (\mathbf{X}\theta - \mathbf{Y}) \quad (4)$$

在计算最优解采用梯度下降法或者采用最小二乘法来求解多元线性回归的解,梯度下降法其优化函数如公式 (5) 表示:

$$\theta_{t+1} = \theta_t - \alpha \mathbf{X}^T (\theta \mathbf{X} - \mathbf{Y}) \quad (5)$$

其中: α 为学习步长也被称为学习率。最小二乘法优化函数如公式 (6) 表示:

$$\theta = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T - \mathbf{Y}) \quad (6)$$

3 实验结果与分析

3.1 实验平台部署和数据采集

通过 4 类不同的传感器对室内环境数据采集,分别采集温度、湿度、PM2.5、VOCs 共 4 种不同的室内空气环境的数据信息。实验地点分别在客厅、卧室、厨房和卫生间,各放置 3 个无线传感器终端节点,每一个终端节点都具有 4 种物理信息的采集、发送功能,经过一个月时间的采集收集,并对数据进行有效的分析处理。同时,采用多元线性回归算法对数据进行拟合,求出最优解,最终可以得出多元线性回归模型并预测得到室内环境的信息。

对于数据的采集如图 4 所示,分别是客厅、卧室、厨房、卫生间的平均值的分布,每 20 分钟采样记录一次当前的数据。图中显示的是计算的是每小时内采集的 3 次的平均值,再对室内的每个位置的 3 个终端节点采集的温度、湿度、PM2.5、VOCs 的数值平均。

由图 4 可以观察到,其中厨房的 VOCs 的指标变化幅度比较大,对于卫生间的湿度变化且平均值较高,卧室的温度较稳平稳,客厅的各项指标都有所变化。这些指标也正说明了人们生活在室内的生活习惯。做饭时厨房的 VOCs

和 PM2.5 指标会明显增加,使用卫生间时其湿度变化就会较为明显。

3.2 实验平台部署和数据采集

本文通过设计一套无线传感器网络的系统,对室内的各个位置的室内物理信息的采集进行数据分析,通过多元线性回归算法对室内的环境变化进行模型的搭建和训练,得到的模型可以预测出室内各项指标的情况。

在数据分析过程中,系统主要是采用机器学习框架 Scikit-learn 包^[20-22],并使用 python 语言实现代码。其中对于室内不同的位置进行了不同的多元线性回归的模型构建,最后将整个室内的数据进行多元线性回归算法的模型构建。

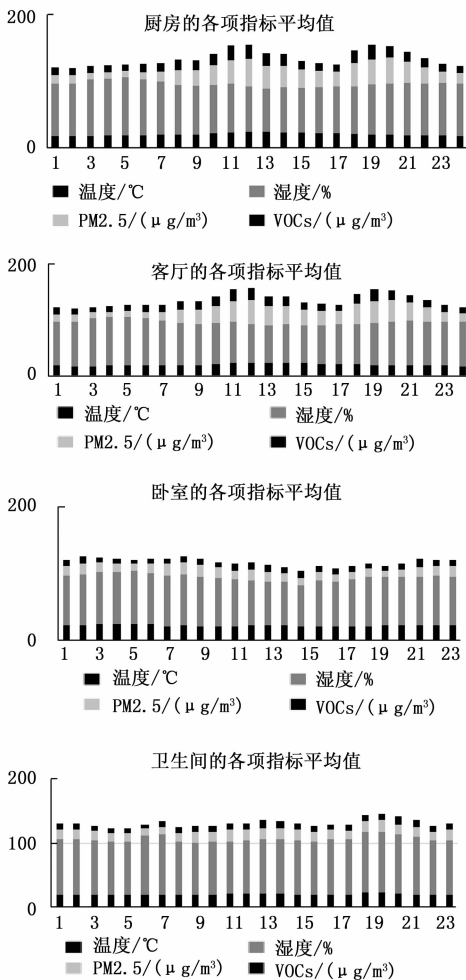


图 4 不同场所的各项指标的均值分布

多元线性回归的模型构建过程,将不同时间段作为预测用 Y 表示,各项指标的在不同时间段内的平均值用 X 表示,这样就可以采用多元线性回归模型进行有效的构建,其中在一个月中的数据很多是噪声数据和冗余数据,因此在数据处理之前做一些数据预处理的过程采用归一化处理,这里采用 Z -score 标准化的方法,可以使得模型学习的收敛速度更快,并提高精度。在使用机器学习框架 Scikit-learn 包中,用到的主要工具包接口进行简单介绍如表 1

所示。

表 1 Scikit-learn 框架接口

Scikit-learn 框架在训练数据的主要接口:	
1	from sklearn.model_selection import train_test_split
2	from sklearn.linear_model import LinearRegression
3	X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.1)
4	linereg = LinearRegression()
5	model = linereg.fit(X_train, Y_train)
6	y_pred = linereg.predict(X_test)

由表 1 可以看到,在使用 Scikit-learn 工具包使用主要用到了两个 API (application programming interface)。其中,用到的一个是 train_test_split 接口,这是将数据分割为训练集和测试集两个部分,这里的 test_size 参数是测试的所占的比例,这里设为 0.1,则说明测试集数据占 10%。另一个接口是 LinearRegression 线性回归,是为了后面数据进行模型的学习。linereg.fit() 接口是用来模型训练的。linereg.predict() 接口是用来模型测试。

无线传感器节点的部署在 4 个位置的数据进行多元线性回归拟合,采集得到的不同时间的各项指标的显著性权重和 RMSE 均方根误差,如表 2 所示。

表 2 不同位置的各項指标通过多元线性回归得到的显著性参数和拟合出的 RMSE

实验场所	客厅	厨房	卧室	卫生间
温度	-0.379 5	-0.427 2	-0.290 7	0.659 1
湿度	-0.149 6	-0.125 8	-0.185 1	-0.085 5
PM2.5	1.127 7	0.992 1	-0.329 0	-0.003 6
VOCs	-0.418 5	-0.018 6	-0.023 5	0.053 4
截距	0.001 9	0.005 1	0.000 16	0.000 91
RMSE	0.049 5	0.035 1	0.043 1	0.042 5

根据采集到的数据进程数据处理、训练,得到对应的模型,用模型进行相应的测试,采用均方根误差作为模型的评价指标。

根据多元线性回归拟合得到的结果在表 3 中可以看到,各项均方根误差都是 5% 以下,同时也得到各项指标的显著度因子权重大小的值。

整个室内环境的变化可以将所有数据合并一起再进行多元线性回归的拟合得到如图 5 所示,每项指标所得到的分类决策平面。

实验最终得到的结果如图 6 所示,通过一个月时间的系统部署所采集的数据。将无线传感器节点部署在室内的 4 个地方进行采集和实验,对数据进行清洗、整理、处理、训练、最终得到的多元线性回归模型用来做预测,得到的预测值与真实值的拟合结果是比较接近,因此本系统的方

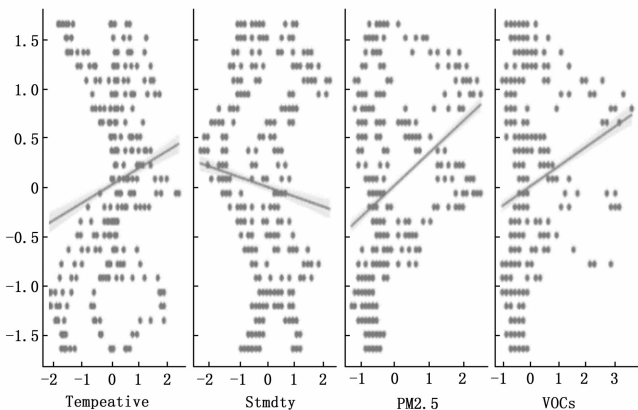


图 5 各项指标的分类决策面

法适用于室内环境指标的预测。

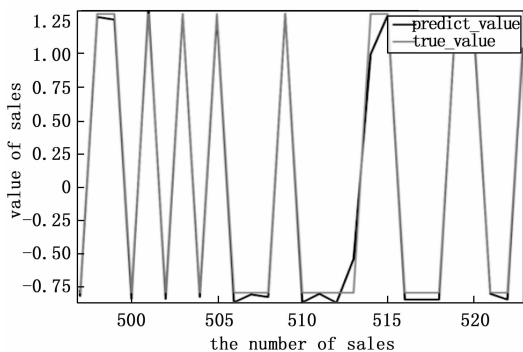


图 6 拟合得到的预测值结果

通过得到的模型应用于该系统。将无线传感器节点部署在室内 4 个场所, 并采集的空气环境的各项物理信息进行预测处理, 得到的结果与实际采集到的数值进行对比, 能够较好地对室内各项空气指标的预测。

4 结束语

该系统采用无传感器节点对室内空气环境的四类重要指标进行数据采集, 并采用多元线性回归算法对数据进行分析和拟合。该系统是自组网, 低功耗, 同时具有较好的稳定性和鲁棒性, 并且系统的组网方式多样, 具有自组网能力, 传输距离远、体积小、方便使用的优点。

该系统可以实时获取室内多种物理信息, 并能够可视化地展现出来, 能够智能化的提醒用户, 某项环境信息不达标。通过多元线性回归算法拟合得到的 RMSE 最大为 0.049 5, 所有采用该方法能够很好地拟合出室内的各项因素的预测。

参考文献:

[1] FANG B, XU Q, PARK T, et al. AirSense: an intelligent home-based sensing system for indoor air quality analytics [C] //Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2016: 109

-119.

[2] 孙金祥, 曹银杰. 基于多传感器的室内环境质量监测系统 [J]. 电子科技, 2016, 29 (8): 127-129.

[3] 黄激之. 基于 ZigBee 传感器网络的室内环境监测系统设计 [D]. 哈尔滨: 哈尔滨理工大学, 2012.

[4] 祝学礼, 刘颖, 尚琪, 等. 空调对室内环境质量与健康的影响 [J]. 卫生研究, 2001, 30 (1): 62-63.

[5] 中国室内装饰协会. 室内环境质量及检测标准汇编 [M]. 北京: 中国标准出版社, 2003.

[6] 王瑶瑶. 基于 ARM 的智能室内空气质量检测与自动化处理系统 [D]. 合肥: 安徽大学, 2020.

[7] 万幸. 采用 ZigBee 技术改善家居环境中空气质量系统方案硬件设计 [J]. 电子测试, 2020 (5): 79-81.

[8] 赵玲云. 基于 ZigBee 通信技术的机房环境监控系统设计 [J]. 甘肃高师学报, 2014, 19 (5): 42-45.

[9] 汪振中. 基于 CC2530 芯片的 ZIGBEE 无线数据传输模块设计 [J]. 中国设备工程, 2017 (14): 133-134.

[10] 王素青, 吴超. 基于 CC2530 的环境监测系统的设计与实现 [J]. 计算机测量与控制, 2015, 23 (8): 2650-2653.

[11] 李正, 张兴伟, 柳宏川. 基于 CC2530 的温湿度监测系统的设计与实现 [J]. 测控技术, 2013, 32 (5): 25-28, 39.

[12] 曾宝国. Z-STACK 协议栈应用开发分析 [J]. 物联网技术, 2011 (3): 77-79.

[13] 高翔, 邓永莉, 吕愿愿, 等. 基于 Z-Stack 协议栈的 ZigBee 网络节能算法的研究 [J]. 传感技术学报, 2014 (11): 1534-1538.

[14] QST 青软实训. ZigBee 技术开发: Z-Stack 协议栈原理及应用 [M]. 北京: 清华大学出版社, 2016.

[15] 王振友, 陈莉娥. 多元线性回归统计预测模型的应用 [J]. 统计与决策, 2008 (5): 46-47.

[16] 封铁英, 王毅敏, 段兴民. 企业资本结构及其影响因素的关系研究——多元线性回归模型与神经网络模型比较与应用 [J]. 系统工程, 2005, 23 (1): 42-48.

[17] 武继峰. 多元信息融合的交通干线地质灾害快速应急与监测技术研究 [J]. 测绘学报, 2017 (8): 1070.

[18] 王惠文, 孟洁. 多元线性回归的预测建模方法 [J]. 北京航空航天大学学报, 2007, 33 (4): 500-504.

[19] 张晓, 李晓戈. 基于多元线性回归的学生成绩分析 [J]. 计算机与数字工程, 2020, 48 (9): 2089-2092, 2151.

[20] PAPER D, PAPER D. Scikit-learn regression tuning [J]. Hands-on Scikit-Learn for Machine Learning Applications: Data Science Fundamentals with Python, 2020: 189-213.

[21] EMAASIT D. Pymc-learn: practical probabilistic machine learning in python [J]. arXiv preprint arXiv: 1811.00542, 2018.

[22] 雷大江, 杜萌, 李智星, 等. 稀疏多元逻辑回归问题优化算法研究 [J]. 重庆邮电大学学报 (自然科学版), 2019, 31 (3): 354-366.