Loosely Coupled Environmental IoT Architecture and Application

Zhengyi Zhao\textsuperscript{1, a}, Liang Li\textsuperscript{2, b}, Xiaxi Han\textsuperscript{1, c} and Shaoyang Zhang\textsuperscript{1, d}

\textsuperscript{1}Chang’an University, Xi’an, China  
\textsuperscript{2}Xi’an TuXun Info Tech Co., Ltd, Xi’an, China

\textsuperscript{a}736155329@qq.com, \textsuperscript{b}18508210@qq.com, \textsuperscript{c}1578857866@qq.com, \textsuperscript{d}zhhsy@chd.edu.cn

Abstract. In the traditional environmental protection tax collection, the discharge caliber of taxable pollutants is inconsistent, the authenticity of sensor data is difficult to guarantee, and information islands are common. Aiming at these problems, a loosely coupled environmental IoT architecture is proposed and implemented. The sensor data access layer achieves compatibility with mainstream sensors in the market, reduces the situation that can only rely on the original sensor manufacturers, and uses the lightweight text data format JSON to deliver messages to improve the transmission rate; the sensor data management layer utilizes big data technology, integration and management of various types of sensor data, laying the foundation for data analysis and data fusion; the integrated application layer uses the C4.5 algorithm to construct a decision tree model, and senses the state of the front-end sensing device to ensure the authenticity of the sensor data. The experimental results show that the proposed method can sense the real state of the front-end sensing device and realize the accurate collection of environmentally-friendly IoT taxable pollutants.

1. Introduction

Since the reform and opening up, with the rapid development of China's economy, environmental pollution problems have gradually become prominent. The 25th meeting of the Standing Committee of the 12th National People's Congress passed the "Environmental Protection Tax Law of the People's Republic of China" on December 25, 2016. This is the "Three Principles of the Taxation Principles" proposed by the Third Plenary Session of the 18th CPC Central Committee. The first single-line tax law reviewed and approved by the Standing Committee of the National People's Congress is also the first single-line tax law in China that specifically reflects the "green tax system." The law will be implemented on January 1, 2018. From the date of implementation of the law, environmental protection tax will be levied in accordance with the law, and sewage charges will no longer be levied. The overall idea of the environmental protection tax law is to change the “tax” from “fee”, that is, to achieve a smooth transfer of the sewage charge system to the environmental tax system in accordance with the principle of “transfer of tax burden”. The law will “protect and improve the environment, reduce pollutant emissions, and promote the construction of ecological civilization” into the legislative purpose, and clarify the “enterprise units and other production and business operators that directly discharge...
taxable pollutants into the environment” as taxpayers to determine air pollution. Matter, water pollutants, solid waste and noise are taxable pollutants.

There are still many problems that need to be solved in the traditional environmental protection tax collection. If the measured emissions of taxable pollutants are inconsistent, the sensor data is not timely and unreal, and the phenomenon of “information islands” is prone to occur. How to implement the tax law, embody fairness, realize the use of tax system adjustment, achieve environmental protection, and promote the goal of ecological civilization construction is a difficult problem facing the tax authorities.

Relying on traditional means, it is difficult to truly implement the environmental tax law and reflect the purpose of the tax law. Technological and managerial changes must be made.

The proposal of the Internet of Things provides a technical solution to solve the above problems. The Internet of Things has better solved a series of problems in traditional collection. The analysis, detection and management of data through the Internet of Things technology has effectively promoted the further implementation of the environmental protection tax law.

The exploration of domestic environmental protection Internet of Things began at the end of the 20th century. After nearly 20 years of development, China has initially established an environmentally-friendly Internet of Things, which focuses on monitoring key pollution sources. At the same time, it has built and applied the “Wuxi Taihu Environmental Monitoring Internet of Things System and Early Warning Platform”. Some provincial and municipal environmentally-friendly IoT systems, such as the Inner Mongolia Trinity Environmental Monitoring System, have achieved environmental quality monitoring functions to a certain extent. However, the current environmental IoT system has many types of sensors, multiple sources of data, and no uniformity. The characteristics of the data platform, the "information island" phenomenon is more common, the current system is directly applied to the direction of environmental taxation implementation is more difficult and complex. To this end, based on the above problems and solutions, this paper uses the Internet of Things technology to design and implement a loosely coupled environmental IoT architecture and application.

2. Environmentally friendly Internet of Things

2.1. Problem Analysis

The Internet of Things is another revolutionary development in the field of information technology after computers, the Internet and mobile communications. At present, the application of the Internet of Things has covered the fields of environmental protection, medical care, transportation and industrial control, which have a tremendous impact on the development of human society.

Industrial sensor networks are the dominant way to achieve infrastructure state awareness. However, in a typical industrial sensor network architecture, various types of sensors and data acquisition terminals are generally produced by multiple vendors and are controlled by corresponding monitoring systems. The main features are: data communication interface, data content and data format. The current main data communication interfaces are RS-232, RS-422, RS-423, RS-485, PLC, HFC, CAN, Industrial Ethernet and other wired standards and IEEE 802.15.3, IEEE 802.15.4, Bluetooth, CDMA, WLAN. And other wireless standards. The main data content may include audio data, video data, image data, text data, etc. The encoded and compressed data formats include MPEG, JPEG, ASN.1, XML, and the like. Therefore, it is difficult to integrate and aggregate data, thus forming a single sensor information island.

2.2. Loosely Coupled Environmental IoT Architecture

To this end, this paper proposes a loosely coupled IoT architecture, the architecture of which is shown in Fig. 1. Its core is divided into two parts: Agent-based intelligent sensor gateway and sensor data management platform.
The functions of each level are:

1) Sensing access and control layer: The access control devices in the traditional control system are integrated and standardized to form a sensing access and control layer. The core device at this level is an intelligent sensor gateway. The gateway should be compatible with the mainstream interface, realize any expansion of the sensor, realize the high scalability of the sensor network, and break the system expansion can only rely on the drawbacks of the original manufacturer.

2) Sensor data management: Integrate the data layers of isolated monitoring software to form a sensor data management layer. The level is realized by the sensor data management platform. The platform utilizes big data technology to realize the integration and management of various sensor data, and lays a foundation for sensor data fusion, emergency event linkage and disposal.

3) Integrated application layer: Based on the sensor data management layer, a comprehensive application layer is formed. This level enables the integrated application of sensor data. Including front-end device state awareness, control, and sensor data fusion analysis.

2.3. Agent-based intelligent sensing gateway

The technical architecture of the intelligent sensor gateway based on Agent is shown in Fig. 2.
Among them, the sensor adapter (SensorApp) can collect current, voltage, RS-485, RS-232, LoRa and other types of sensor data, use the filtering operation to improve the accuracy of the analog signal data, and obtain the number through the internal analog signal to the digital signal circuit. Signal; embedded with MCU microprocessor with programming function, has the function of parsing, processing data and interacting with Agent.

The Agent is the medium for communication between the sensor adapter and the sensor data management platform. On the one hand, the sensor data collected by different sensor adapters can be packaged and delivered to the sensor data management platform; on the other hand, the Agent will sense the data. The message sent by the management platform is parsed and transmitted to the sensor adapter to implement configuration and control of the sensor adapter. The Agent temporarily stores the sensor data in the SQLite database, and utilizes the advantages of SQLite lightweight, fast read/write speed, and efficient C language API to effectively prevent conflicts and loss during data transmission. Using the sensor adapter under the universal interface, the sensing agent can be easily compatible with the data transmission protocol and data format of different sensors, and the data is stored in the database of the application service system through the network in a unified data format.

Intelligent sensor gateway messaging uses a lightweight text data format JSON (JavaScript Object Notation) that is independent of the programming language. Its simple and clear hierarchy is easy to machine parsing and generating, which effectively improves network transmission efficiency.

2.4. Sensor data management platform
The sensor data management platform provides a convenient and feature-rich application interface API, enabling sensor system manufacturers to use the transmission channels provided by the system to achieve their business goals, and even use multiple sensor fusion to achieve comprehensive analysis. Its main features include:

1) Intelligent device management
Users can register their accounts through the sensor data platform and add their own smart devices and configure related information about the device, such as device type, device ID, device description and device status.

2) Data service
The sensor data management platform can display the data collected by various sensors in the form of graphs and curves, which is convenient for users to detect the sensor data in real time. Users can remotely manage the sensors and data points through the device management console page. Parameter configuration; to ensure data security and privacy, the iThings cloud platform uses the API Key mechanism for authorization management of third-party applications.

2.5. Environmentally friendly IoT solution
The agent-based environmental IoT logical architecture is shown in Fig. 3.
First, place sensors and collection terminals at the enterprise's sewage outlets or noise sources, and collect the emissions of four kinds of taxables, including air pollutants, water pollutants, solid waste and noise, and then further collect the data through different SensorApps. Encapsulation and aggregation to the intelligent sensor gateway, the intelligent sensor gateway performs scheduling processing on different SensorApps, encrypts the data by SSL (Secure Sockets Layer) protocol, and finally transmits and stores the data to the sensor data management platform.

This architecture provides the following benefits:
1) The pollutant data is monitored in real time, and the data comes from the sensor, which makes the tax accurate, pays more taxes, and guarantees fairness.
2) The statistical caliber is consistent.
3) Taxation is simple and rapid, just connect the cloud platform statistics.

By supplementing the operational plan, the following advantages can be further provided:
1) Specialized companies carry out equipment investment and operation, taxpayers and tax authorities purchase services to solve the problem of insufficient taxpayer specialization and difficulty in reviewing.
2) Reducing the burden on enterprises while providing data services to enterprises.
3) Realize full coverage of taxation. Specialized operations can quickly deploy sensors, quickly integrate market objects without monitoring equipment into the collection system, increase the scope of taxation, and solve business problems for the tax department.

3. State analysis based on multi-sensor devices

3.1. Overview
For the environmentally friendly Internet of Things, there may be equipment failure or human interference during the use of the sensor, and it is impossible to ensure accurate data collection, resulting in uploading unreal data. This section proposes a solution to determine the accuracy of sensor data and avoid tax losses.

This paper selects air pollutants as an example for comprehensive analysis. Firstly, the C4.5 decision tree algorithm is used to construct the air quality assessment decision tree. Secondly, the accuracy of the factory PM2.5 sensor data is judged by analyzing the air pollutant sensor data of the factory and the air quality level of the area.
3.2. Dataset selection
There are six main types of pollutants that affect air quality: fine particulate matter (PM2.5), respirable particulate matter (PM10), sulfur dioxide (SO2), carbon monoxide (CO), nitrogen dioxide (NO2), and ozone (O3). As shown in Table 1, this paper selects the Beijing Air Quality Index as a sample of the research data set in recent years, and the numerical unit is $\mu g/m^3$ (CO is mg/m³).

Table 1. Beijing Air Quality Index

| ID | PM2.5 | PM10 | SO2 | CO  | NO2 | O3  | LEVEL |
|----|-------|------|-----|-----|-----|-----|-------|
| 1  | 145   | 250  | 47  | 2.5 | 85  | 23  | IV    |
| 2  | 206   | 314  | 48  | 2.6 | 90  | 16  | V     |
| 3  | 79    | 141  | 30  | 2.6 | 81  | 34  | III   |
| 4  | 274   | 358  | 32  | 3.2 | 80  | 13  | VI    |
| 5  | 12    | 46   | 2   | 0.3 | 19  | 93  | I     |
| ...| ...   | ...  | ... | ... | ... | ... | ...   |
| 100| 19    | 69   | 3   | 0.3 | 21  | 102 | II    |

In this experiment, 70% of the data samples were randomly selected as the training data set, the decision tree prediction model was constructed, 30% of the data samples were selected as the test data set, and the correct rate was calculated.

3.3. Information gain rate
Information entropy was first proposed by Shannon and used to describe the uncertainty of the source. The smaller the entropy, the higher the purity of the information. The calculation formula of information entropy is as follows:

$$Ent(D) = -\sum_{k=1}^{\mid y \mid} p_k \log p_k$$

Where $p_k$ is the proportion of the kth sample in the current sample set D (k=1, 2, 3... $\mid y \mid$).

The information gain refers to the reduced value of the information entropy after dividing the sample by a certain feature and when it is not divided, that is:

$$Gain(D,a_i) = Ent(D) - \sum_{j=1}^{\mid y \mid} \frac{D_{ij}}{D} Ent(D_{ij})$$

The value of the attribute $a_i$ is {a$_{i1}$, a$_{i2}$, ..., a$_{ij}$}. When determining the attribute $a_i$, all samples with the value a$_{ij}$ are recorded as $D_{ij}$.

$$Gain\_Ratio(D,a_i) = \frac{Gain(D,a_i)}{Split\_Info(a_i)}$$

The information gain rate is calculated by dividing the information gain by the split information metric ($Split\_Info(a_i)$), where the split information metric is defined as (the split information is used to measure the breadth and uniformity of the attribute split data):

$$Split\_Info(a_i) = \sum_{j=1}^{\mid y \mid} \frac{D_{ij}}{D} \log_2 \frac{D_{ij}}{D}$$

$Split\_Info$ is proportional to the number of possible values of $a_i$, and the information gain rate is inversely proportional to the number of possible values of $a_i$. 
3.4. Decision tree model

The decision tree algorithm processes the data through a top-down recursive method, and generalizes the unordered and irregular instance sets into a set of classification rules represented by the tree structure. At present, the algorithms for generating decision trees mainly include: CART algorithm, ID3 algorithm, and C4.5 algorithm. The C4.5 algorithm inherits the advantages of the ID3 algorithm and improves the deficiencies of the ID3 algorithm: the information gain rate is used instead of the information gain for attribute selection, which has the characteristics of fast classification and high precision. Therefore, this paper uses C4.5 algorithm to construct decision tree.

As shown in Fig. 4, the model generated using the C4.5 algorithm is:

![Decision tree model](image)

Figure 4. Air quality assessment prediction decision tree model

The test data set of 30% of the test data sets was tested and classified, and the correct rate reached 96.67%. This result indicates that the rule information obtained by the model can make the maximum possible prediction of air quality to some extent.

3.5. Sensing state judgment

This paper assumes that the data collected by the five sensors of PM10, SO2, CO2, NO2, and O3 are all true data, and the fine particle (PM2.5) sensor is the sensor to be judged.

First, use the generated decision tree to predict the air quality level of the above six sensors collected by the factory, and then compare the air quality level released by the factory area to determine whether the PM2.5 sensor status is normal. The method can determine whether the PM2.5 sensor is caused by human error or equipment failure, so that the data is mistaken, so that it can be found and processed in time to reduce the national tax loss.

### Table 2. Modify PM2.5 Air Quality Index

| ID | PM2.5 | PM10 | SO2 | CO  | NO2 | O3  | LEVEL |
|----|-------|------|-----|-----|-----|-----|-------|
| 1  | 74    | 358  | 32  | 3.2 | 80  | 13  | VI    |
| 2  | 12    | 365  | 48  | 3.7 | 105 | 26  | VI    |
| 3  | 93    | 449  | 42  | 4   | 105 | 18  | VI    |
| 4  | 98    | 314  | 48  | 2.6 | 90  | 16  | V     |
| 5  | 66    | 231  | 40  | 2.6 | 94  | 17  | V     |
| ... | ...   | ...  | ... | ... | ... | ... | ...   |
| 50 | 30    | 181  | 45  | 2.4 | 72  | 19  | IV    |
As shown in Table 2, some normal data are selected and the PM2.5 sensor data is modified to be close to the standard value (China's PM2.5 standard value is less than 75 μg/m³ for 24 hours), simulating human tampering. Then it is used as the input of the decision tree, and the obtained output is compared with the LEVEL of the data itself. Finally, the abnormal state detection rate of the PM2.5 sensor is 98%.

This result shows that the decision tree obtained by C4.5 algorithm for data mining can meet the high requirements for the abnormal state detection rate of PM2.5 sensor. Using this model, the abnormality of sensor data can be accurately captured. Achieve the purpose of judging the state of the sensor.

4. Conclusion

In this paper, the measurement of the taxable pollutants in the traditional environmental protection tax is inconsistent, the sensor data is not timely, the real problem, and the multi-source heterogeneous, data communication interface, data content and data format of the traditional industrial sensor network. The different problems are first analyzed. The root cause of these problems is that the individual access control devices in the traditional control system are scattered, independent, and have no unified data management platform. Then, specific solutions are given, and the ideas are the access control equipment is integrated and standardized, and the data layer is integrated to form a comprehensive application layer. The paper proposes and designs a loosely coupled environmental IoT architecture and application, which solves the above problems of current environmental taxation. On this basis, the loosely coupled environmental Internet of Things can intelligently sense sensor status by generating decision trees, avoiding sensors uploading erroneous data or stopping work for a long period of time, and escorting the implementation of environmental taxation.

Acknowledgments

This article was completed under the careful guidance of the instructor Professor Shaoyang Zhang. The rigorous work style and profound academic attainment of the instructor are the strong backing that I can successfully complete. During the study, Professor Shaoyang Zhang’s strict requirements, enthusiasm and enthusiasm for me in terms of ideology, study, and life have enabled me to regain my confidence in the face of setbacks and to maintain my small achievements. Sober, seeing the lack. The teachings given to me by the mentor will surely benefit me for the rest of my life. It is the guide and motivation for my life. The corresponding author of this article is Professor Shaoyang Zhang.

References

[1] Ji Wei. Analysis and Improvement Suggestions on Tax Elements of Environmental Protection Tax [J]. Southern Entrepreneur, 2018 (01): 76.
[2] Shi Xuejun, Yang Bowen. Institutional Coupling and Cohesive Mechanism of Environmental Protection Tax and Pollution Discharge Permit Management in China [J]. Tax Economic Research, 2019, 24 (01): 17-24.
[3] Huang Zhuowen, Zhang Suhua. Implementation and Improvement of Environmental Protection Tax in China [J]. Taxation, 2019, 13 (09): 4-6.
[4] Xiao Tianbai. Study on the Construction of Environmental Tax Legal System in China [J]. Journal of Huangshan University, 2018, 20 (02): 27-31.
[5] Fleisch E. What is the internet of things? An economic perspective [J]. Economics, Management, and Financial Markets, 2010, 5 (2): 125-157.
[6] Soldatos J, Kefalakis N, Hauswirth M, et al. Openiot: Open source internet-of-things in the cloud [M] //Interoperability and open-source solutions for the internet of things. Springer, Cham, 2015: 13-25.
[7] Hou L, Zhao S, Xiong X, et al. Internet of Things Cloud: Architecture and Implementation [J]. IEEE Communications Magazine, 2016, 54 (12): 32-39.
[8] M. H. Syed and E. B. Fernandez, "Cloud Ecosystems Support for Internet of Things and DevOps Using Patterns," 2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI), Berlin, 2016, pp. 301-304. doi: 10.1109/IoTDI.2015.31
[9] The key technology of Zhan Bin Internet of Things and the application research of computer Internet of Things [J] Electronic World, 2017 (13): 15 + 17.

[10] Z. Xiang and L. Zhang, "Research on an Optimized C4.5 Algorithm Based on Rough Set Theory," 2012 International Conference on Management of e-Commerce and e-Government, Beijing, 2012, pp. 272-274. doi: 10.1109/ICMeCG.2012.74

[11] S. Das, S. Dahiya and A. Bharadwaj, "An online software for decision tree classification and visualization using c4.5 algorithm (ODTC)," 2014 International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2014, pp. 962-965. doi: 10.1109/IndiaCom.2014.6828107

[12] Huang Ming, Niu Wenying and Liang Xu, "An improved Decision Tree classification algorithm based on ID3 and the application in score analysis," 2009 Chinese Control and Decision Conference, Guilin, 2009, pp. 1876-1879. doi: 10.1109/CCDC.2009.5192865

[13] Z. Yuan and C. Wang, "An improved network traffic classification algorithm based on Hadoop decision tree," 2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS), Chongqing, 2016, pp. 53-56. doi: 10.1109/ICOACS.2016.7563047