Massive data processing techniques and model in Internet of things

Yuebin Wang [1, a], Quanfeng Yan [1,b] and Liqiang Liu [1,c]

[1]School of Computer Science, Hunan Institute of Science and Technology

[a]yuebingw@163.com, [b]acmfcc@163.com, [c]liuliqiang2@163.com

Keywords: Massive data, processing techniques, Internet of things

Abstract:with the further development of Internet,rapid scales. Technology the expansion of Internet use is growing due to such rapid growth in this industry, the amount of data dealt by Internet of things becomes the feature of massive amount. The key problems that should be considered with, is how to efficiently processing these data, from which can get useful information, and then to provide intelligent decision. In connection with the new inquiry of massive data processing in the era of Internet of things. To understand the massive data processing technique is used in the Internet of things, through the analysis of massive, heterogeneous, Multi-dimension and dynamic network data. To explore from 4 aspects of the network data acquisition, data transmission, data model and storage, data processing applications, based on the data processing model of Internet of things technology system , construct a general network data storage, analysis and presentation model.

Introduction

Internet of Things (IOT) is the new information industrial wave after the computer, Internet and mobile communication network, The goal is to keep global scope of the object, the systems of information technology and people organically connected through various information sensing equipment and intelligent communication system, in order to better aware and manage the physical world. At present, the internet of things has been widely used in intelligent residence, intelligent traffic, intelligent fire control, environmental monitor, logistics and express delivery, food traceability, industry monitor, hotel management, health care and other fields. The literature [1,2] developed a series of prototype system of Internet of things which can be applied in everyday life, and explored the new way of the human-computer interaction in the kitchen, laundry, entertainment, sports and other fields. The literature[3] designed a systematic construction of Internet of things, which takes environment of sensor nodes, energy monitor system and RFID label object as the prototype, put into practice based on the RESTful principle, gather data and analyze platform through RESTful interface interaction.

The literature[4] developed an infrastructure based on basic construction of network, used for storing, sharing, searching, visualizing and analyzing data from various types of equipment's data. The sensing devices in the Internet of things is various. It's necessary for processing and storage of massive heterogeneous data from different network, different massive isomery data in subsystems, and thus requires the processing mechanism can melt multi-network fusion, multi-source, heterogeneous and deal with the data efficiently, and get valuable information from them, and then provide intelligent decision. At the same time, the task of completing the mass data procession requires a variety of resources, including computing resources, storage resources, cyber source and so on, The unified regulation and schedule of these resources can effectively accelerate the speed of massive data procession. For the feature of massive data in Internet things, the data processing technology has become one of the most important technology in the realization of internet large-scale popularization and application.

Data Figure

The Internet of things is to obtain the goods and the environment through the embedded in the object tag or all two forms of the sensor information, the real-time information is sent to the network data

analysis platform through the wireless network mode, and the information system can be interconnected to form a huge network, so as tracking, environmental monitoring intelligent management of goods. A large number of real world objects, in various forms, constant movement and change, distribution in different locations, and is easily affected by the environment, which leads to the IOT obtained through various sensing equipment data has the following characteristics:

The massive feature of data

Each network system usually has one of more wireless sensor network, and the network contains a large number of sensor nodes. Each node will constantly produce new data, these data types include sampling data simple numeric types like temperature and humidity, pressure, GPS locator sensor generated, and numeral data including video data acquisition camera and remote sensing data and other complex multimedia data. At the same time, the system also will be collecting historical data storage for a certain period of time to meet the needs of processing and data mining analysis.

The data feature of heterogeneity

Sensors gather and form different types of data types, data accuracy and data amount in accordance with their physical feature, product characteristics, and system design. This leads to IOT data cannot be unified into a single mode. For example, It usually contains the function of the vehicle GPS location, special vehicle information gathering, RFID license recognition, illegal electronic photography, traffic information, real-time traffic data collection and so on in the intelligent traffic. To achieve the above function involves dozens of sensor, which produce a variety of multimedia video types of heterogeneous data from simple to complex numeric types.

The data feature of Multi-dimension and association

The data feature of Multi-dimension is an important and essential characteristic of network data, which is important distinction that different from the ordinary Internet. As shown in Figure 1 the original data acquisition system of the Internet of things has three properties of time , space, and devicestamp under default situation. For example, in a forest fire or environmental monitoring application, each sensor equipment will record and locate the monitored data of specific place and time; the intelligent transportation logistics company can understand the specific time of wagon position as well as the truck driver information multivariate data through the GPS equipment. In the food traceability applications, including processing time, processing locations and food manufacturers and other multivariate data information through radio frequency label record. In addition, the physical object of the Internet of things are not independent existence, existence or bright or dark complex associated attributes between them. For example in the smart grid, the relative position of electricity users in the physical network will affect their relationship and degree of correlation.

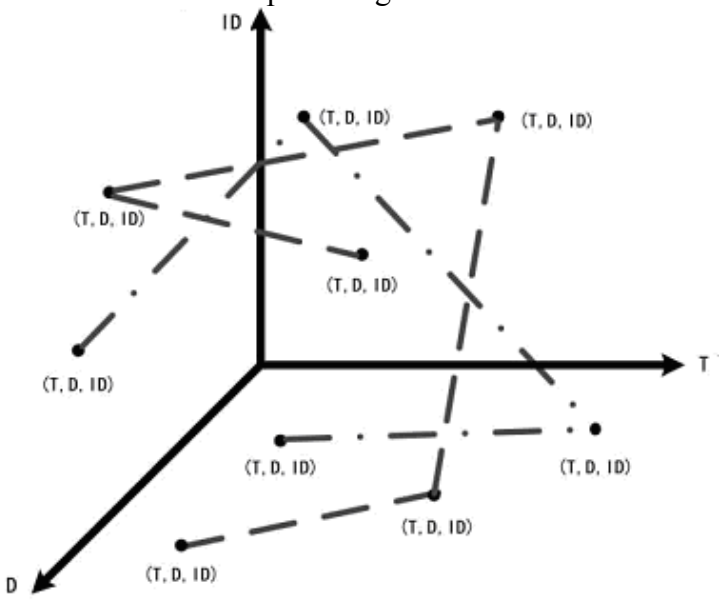

Figure 1: The data feature of Multi-dimension and association

Data of real-time and dynamic

The application of the give prominence to timeliness,Whether the RFID system or WSN system,they acquis real-time data,Every other cycle to send a data to the server. In the Internet of things system,to query a monitoring object in a moment of physical state is not simply based on a keyword matching to complete. In order to effectively query processing,we need to be the same as a monitored object previous data into a sequence of data to calculate the monitoring object in the specified time physical state. With the acquisition of data update,Data acquisition sequences are also constantly changes.

Technical system

In order to implement the data to the specific processing stages, according to the principle of technology accumulation, in accordance with the data acquisition, data transmission, data modeling, and the internal logic order of storage and data processing applications, build the system that Internet of things technology in the service of the data processing, mainly divided into four layers.

Awareness and control layer

The function of this layer is proper processing, transmission gateway to pass data to the network layer of the sensor by sensors and other devices to collect production and operation status information; At the same time can also be informated, receiving gateway receives control commands locally is passed to the controller to achieve control of production, operation, etc. In this level, the perception and control device of the management, the local data and signal processing is an important technical field; Sensors, controllers and other components of miniaturization and intelligence is the developing direction.

Network layer

The function of this layer is that data has been collected in the end by public or private network with wireless or wired communication data between perception and platform services layer and control layer for transmission through the wireless sensor network, field bus technology, such as receiving perception and control layer. Among them, we especially need to manage the safety and quality of transport services for avoiding the loss of data, out-of-order and delay problems. Efficient and reliable access protocol transmission technology is the developing direction.

Platform service layer

Function is to get the data through the perception of the layer and control layer, network layer, after routing and processing of data necessary for storage. However data processing logic depends on the server, operating system, equipment and application type, and it produces high quality as well as the integration of data will be transmitted to the data analysis module for further data mining processing. Analyzing module to data model and early warning mechanism is used to transform the data associated with the physical environment, equipment and application, based on the current data and historical data, evaluate and predict the current state of the system, and risk factors. In this layer, data modeling, mining depth, intelligent and analysis is the development direction of [5].

Application service layer

According to the needs of enterprise business, the platform establishment related Internet application service layer. The application integration of perception form of a business process , service layer, control layer, network layer and platform to achieve timely awareness, timely analysis, response time and the Internet of things intelligent management business model raise operation efficiency, promote the business model innovation and reduce the cost of operation and management. In this layer, the first application of advanced technology use the paramount language and the configuration to the first three layer provides the data to integrate and develop function and realize the Internet of things in the enterprise internal or local implementation and application of industry; Secondly, with the aid of M2M, MDA, RSA implementation architecture technologies such as iot application of the industry, and can improve service efficiency through the platform integration; Finally with the help of data mining, edge computing, cloud computing achieve the implement multiple integration of iot are widely used in industry and innovations.

The model of data processing

To construct data processing model based on technology system of Internet of things, in accordance with the principle of technology mapping As shown in figure 2.

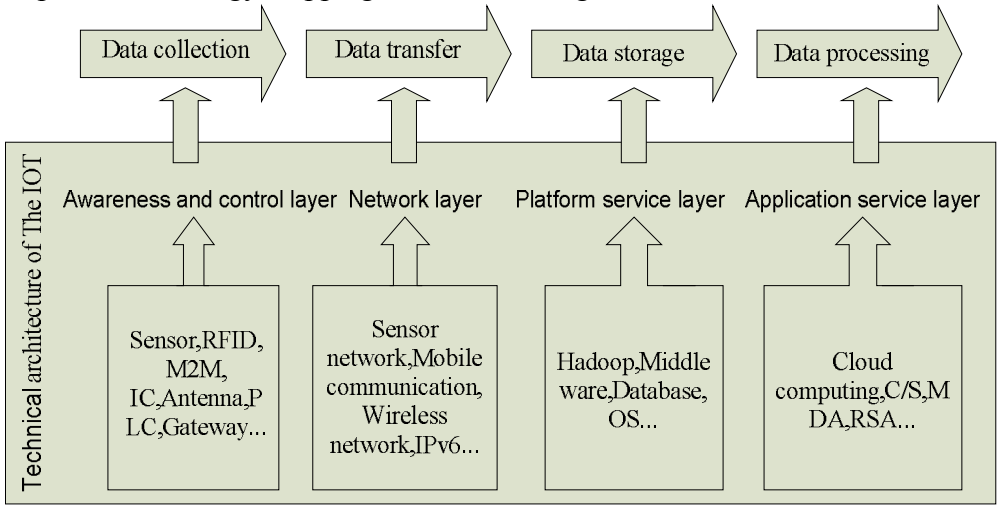

Figure 2. Data model of The IOT

As the important development direction of next generation Internet,, The Internet of things to attract industry and academia. They produce huge amounts of data and the typical of data exists,so the data processing has become a challenging task. . Based on the technology of the Internet of things system data processing model, To practitioners for guidance in data processing the thread, But for the implementation of specific technical details, Especially the technology of the basic data of intelligent modeling, data mining technology has not been discussed.. Future research can focus on data modeling method based on network, data mining model and corresponding algorithm.

References

[1] International Telecommunication Union. Internet reports 2005: the Internet of Things . Geneva: ITU,2005.

[2] Kranz M,Holleis P,Schmidt A. Embedded Interaction Interacting with the Internet of Things [J] . IEEE Computer Society,2009,14(2):46-53.

[3] Guinard D,Trifa V,Mattern F,et al. Architecting the Internet of Things. Berlin Heidelberg:Springer,2011.

[4] Gupta V,Poursohi A,Udupi P. Sensor. Network: An Open Data Exchange for the Web of Things,2010. Mannheim:IEEE,2010.

[5] Paul C Zikopoulous,Chris Eaton,Dirk deRoos,Thomas Deutsch, George Lapis. Uderstanding Big Data. NewYork: Mc Graw Hill,2012.