DEVELOPING A DISASTER SURVEILLANCE SYSTEM BASED ON WIRELESS SENSOR NETWORK AND CLOUD PLATFORM

Jianhu Cen[1], Tao Yu[1], Zhe Li[1], Song Jin[1], Shaohua Liu[2]

1 International School, Beijing University of Posts and Telecommunications, Beijing 102209, China
2 School of Electronic Engineering,
Beijing University of Posts and Telecommunications, Beijing 100876, China
ee08b187@bupt.edu.cn

Abstract

Typically, today's WSN disaster surveillance system operates poorly in its accuracy and timeliness and can only detect a single type of disasters. This owes largely to the congestion brought about by excessive communication traffic and the processing limitation of the server which dramatically restrict the development of disaster surveillance systems. In order to solve these problems, this paper proposes a new scheme in improving the traditional disaster surveillance systems. At the data collection and transmission layer, orthogonal neural network algorithm, which is based on the wavelet transform, is introduced to promote the surveillance accuracy and reduce the network congestion. At the data storage and computing layer, cloud storage and distributed parallel computing are used to overcome the limitation of the previous storage and computation. At last, the paper gives the concrete implementation plan and verifies the superiority of the system.

Keywords: WSN; Network congestion; Cloud platform.

1 Introduction

As people pay more and more attention to the disaster, accurate surveillance as well as timely feedback plays an extremely important role in reducing the loss of lives and property in the disaster. In order to provide a real-time surveillance of different types of disasters simultaneously, the system has to solve network congestion problem with large scale of communication traffic and also overcome the limitation of storage and computation. This paper solves the congestion problem and increases the accuracy of surveillance by improving the WSN algorithm and overcomes the limitation of storage and computation by using the cloud computing servers instead of the previous single server, which leads to the competence of the whole system.

Generally, approaches for data transmission and congestion control at home and abroad are: 1) CODA[1]: congestion detection based on the receiving end， which uses post-compressing mechanism with hop by hop in the open loop congestion and multiple sources adjusting in the closed loop, and can regulate the local congestion nodes. However, this algorithm makes the nodes nearer the sink node send more packets; 2) Adaptive resources control[2]: by creating multiple paths, it can adaptively adjust the resource supply and increase accuracy of the grouping of transmission, but if the destination end is too far away from the source end, transmission delay and energy consumption will increase; 3) Fusion[3]: process large amount of raw data collected by the sensors in several ways to eliminate the redundancy and send just the useful information to the gathering node. In this way, it can effectively reduce the communication traffic, but the BER will increase accordingly and the transmission accuracy needs further improvement; 4) many-to-one routing congestion control[4]: determine the minimum allowed sending rate of the child nodes downstream to reduce congestion, which is simple and applicable to the node with limited resources, but the recursive process of eliminating congestion will greatly influence the transmission across depths.

The storage and computing approaches for network of WSN are: 1) the use of wireless sensor network node module to storage and compute; 2) store and operation in the gateway node; 3) increase of the outer self-built server. However, after using these approaches, the storage and computation capabilities of the system are still too limited, which may result in further power consumption and hardware costs.
As for the previous algorithm problems, this paper proposes a new disaster surveillance scheme. The characteristics are shown as below: 1) based on the

existing fusion algorithm, this paper introduces the orthogonal neural network algorithm based on wavelet analysis to improve the accuracy of regional monitoring and to reduce data throughput as well as the network node power consumption; 2) use cloud platform to analyze large scale of data and forecast disasters. Meanwhile use exception catching to get real time environmental data to largely reduce the network maintenance costs.

2 Analysis of the system structure

The system uses three-layer structure described in Figure 1. The first layer is data collection and transmission, which consists of WSN and gateways. Each kind of sensor module is directly connected to communication module and collects data for corresponding coordinator nodes. The network coordinator, which is in charge of collecting and transmitting the data from sensor nodes, transfers data with the gateway.

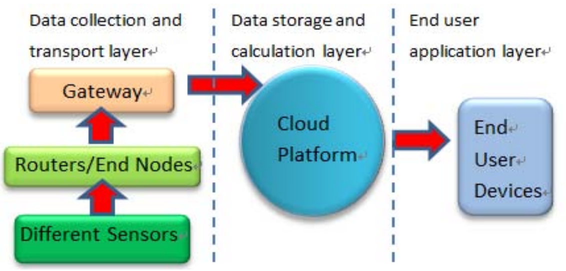

Figure 1 System structure

The second layer is data storage and processing (Cloud Platform). The gateway sends data to the cloud server through the Internet. The cloud uses distributed storage to store sensors' data, and do the analysis by smart data parallelism to predict the form and severity of disaster. The third layer is for end user application. The user interface is implemented on the cloud server. Users can use network devices, such as PCs or mobile phones, to login for local environment information.

3 Models design and analysis

3.1 WSN congestion control model

This model mainly adopts the fusion algorithm to reduce communication traffic, and improve the forecast accuracy at the same time. Therefore, we put forward the orthogonal neural network (WONN) algorithm based on the wavelet analysis.

The complete network includes input layer, first hidden layer, second hidden layer and output layer. First carry on the wavelet processing [5]for the input vector $X=[x_1,x_2,\cdots,x_m]^T$:

$$Y=\sum_{j=1}^{k} p_j h\left(\sum_{i=1}^{m}\frac{x_i-d_{ij}}{c_{ij}}\right) \quad (1)$$

where $p_1,p_2,\cdots,p_k$ are the weights of the second hidden layer in the wavelet process, d_{ij},c_{ij} are shift factor and scale factor of wavelet base respectively. k is the number of wavelet base, $h(*)$ is the wavelet base function.

In this way, the velocity of convergence and effects of approaching are strengthened, and we can get vector $Y=[y_1,y_2,\cdots,y_n]$. Then replace the variables for every single element y_i

$$t_i=\frac{\pi}{b-a}(y_i-a), y_i\in[a,b] \quad (2)$$

where a,b are the minimum and maximum of y_i respectively.

Use the orthogonal primary function set $1,\cos(t),\cos(2t),\cdots,\cos(nt)$ as neural network excitation function. Taking advantage of its ability to approach arbitrary non-linearly, we operate with t_i and get the output of the network z. Assume set of samples of network training in neural network is $\{x_i,\overline{x}\mid i=0,1,2,\cdots,n\}$, and

$$X=\frac{1}{n-1}\sum_{i=0}^{n}x_i \quad (3)$$

After orthogonal transformation, the weight vector from neural network to output process is $W=[w_0,w_1,\cdots,w_k]^T$. The second hidden layer neurons excitation function vector is $S=[1,\cos(\frac{\pi}{b-a}(y_i-a)),\cdots,\cos(\frac{N\pi}{b-a}(y_i-a))]$. Then the above process finally reduce to

$$z_i=SW \quad (4)$$

Assume error function and performance indicators are

$$e_i=\overline{x}-z_i \quad (5)$$

$$J=0.5\sum_{k=0}^{n}e_i^2 \quad (6)$$

When J is the minimum, W then is what we need. Le $\frac{\partial J}{\partial W}=0$, using the least square method to evaluate and get

$$W^{k+1}=W^k+Q^k[\overline{x}-SW^k] \quad (7)$$

$$Q^k = \frac{P^k S^T}{1 + SP^k S^T} \quad (8)$$

$$P^{k+1} = (1 - Q^k S) P^k \quad (9)$$

Randomly generated initial weight vector $W^0 = rand(N+1,1)$, $P^0 = \alpha I$ (α is positive number big enough, $I \in R^{(N+1)(N+1)}$ is unit matrix) through training sample data, we can obtain neural network weight vector W .

3.2 Model of data storage and operation on cloud platform

3.2.1 Storage and management of data

Cloud computing uses distributed storage to store massive data in the large scale of cluster of cheap servers, which make its total storage capacity increase greatly. This paper utilized the framework of Hadoop to realize distributed system. In the HDFS, the file is divided into fixed size of the block of data stored in a group of data nodes. HDFS is a kind of low-level storage, which can be divided into three parts: Client, NameNode and DataNode. Client is an application based on the HDFS; NameNode is the manager of the distributed file system, mainly responsible for namespace of file system, configuration information the cluster, coping information of data block and storing metadata of file system in memory. The DataNode is the actual storage path of files. It will store metadata of block in the local file system, periodically send all block information to NameNode.

3.2.2 Model of map-reduce operating

Map-reduce divides data into small blocks for multiple numbers of computers in the network to operate, and then gather the operating results to get a conclusion[6]. Its main steps are: 1)User program calls the library of Map-reduce and divide the big data inputted into M fragments of data; 2)Master distributing program is responsible of distributing. M Map tasks and R Reduce tasks will be distributed to an idle work station; 3)The work station allocated with Map tasks reads a block of data and then resolves key/value pair. The user-defined Map function accepts value of key/value pair and then produces a set of middle value of key/value pair; 4)Map-reduce library collects all sets of values which have the same key and delivery them to the reduce function (The flowchart is shown in Figure 2).

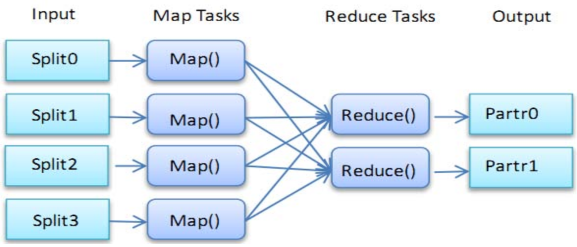

Figure 2Operating flowchart of Map-reduce

4 System implementation

4.1 WSN

System uses temperature, humidity and smoke sensors to implement data collection. All DS18B20 temperature sensor, DHT11 humidity sensor and MQ-2 smoke sensor have the advantage of low cost and long working time, which enhance the ability of anti-interference and field measurement.

ZigbeeCC2530 is a good choice to be used as the communication module, because it provides excellent receiver sensibility and anti-inference. As for software, IAR Embedded Workbench provides an integrated development environment for module code and chip downloading. Each communication module (coordinator, router and end) is connected with corresponding sensor in a layered topology. The coordinator, which is connected with gateway, is in charge of data collection from each node.

4.2 Gateway

Considering cost limit and power consumption, the gateway only has the basic Internet access function and external storage, in order to run Linux. Our system uses Intel Atom N270 as CPU, which has low cost and power consumption. It also uses X86 instructional set, which has the most abundant software resources.

The software design consists of two aspects: cloud server and network coordinator communication. For former one, system uses open-source CURL class library under POST in HTTP. For the latter one, the serial port file in Linux is used. The system can filter data information from the uploading one, then make HTTP package, finally send to cloud server for storage and processing.

4.3 Cloud platform

Cloud platform is implemented with the help of Hadoop, which is an open-source distributed software frame. It consists of Hadoop Common, ZooKeeper, HBase, HDFS and Mapreduce, each one is in charge of distributed data storage, processing model and parallel access control. Hadoop depends on community servers, so the cost is low, and everyone can access it. Hadoop uses Java to build the frame; therefore it is practical running on Linux. For external interface, our system provides two interfaces (Figure3). One is for gateway; the other is for end user. Both interfaces are packaged in the form of HTTPService, which is convenient for adding other system into ours.

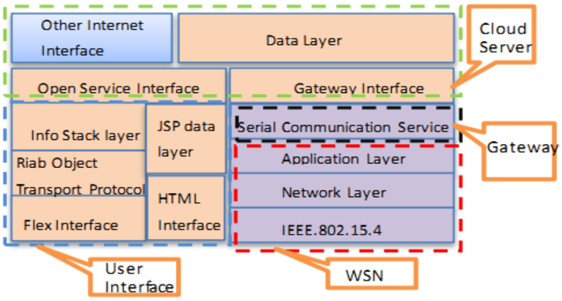

Figure 3 System layer and interface diagram

4.4 End user terminal

The terminal devices can be either PC or mobile phone. User only needs to visit certain website to login the system, instead of installing and configuration. Due to this, for PCs, our system uses Flex to build a functional UI. For mobile phone, we build another simple UI made by JSP and HTML.

5 Simulation and evaluation

Fusion algorithm combines M groups of data of sensor nodes into a packet. Simply using this algorithm will relieve pressure on network so as to make congestion packet loss rate and data queue time reduced by M times. But for the accuracy of the data integration analysis, we take data collected by the temperature sensor as an example, utilize the literature data[7] seven temperature sensors to measure the temperature of the same heat source at the same time for 10 times, and use temperature tester of high precision to get real temperature value of the heat source. Results are showed in Table 1:

Table 1 Temperature of different sensors

Sensor	1	2	3	4	5
T/°C	846.2	852.0	848.1	851.9	849.3
Sensor	6	7	8	9	Real value
T/°C	856.6	849.9	848.3	851.2	850.0

The values of sensor 1 and 6 are error data. The rest of other 7 values has an arithmetic mean of x_0 =849.6°C. Compared with the actual temperature of 850.0°C of the heat source, it has a measurement error of -0.4°C. Now using WONN algorithm, assume that α =10^5, structure of neural network topology is $1\times3\times1$, randomly generated vector of second hidden weight layer $P_0 = rand(3,1)$, and let a= 848. 1, b= 852. 0.

Send the rest seven groups of data of temperature sensor to the WONN network for training, and the data get a output mean of $\bar{z}$ = 850.03. Compared with temperature of 850.00 of heat source, the measurement error is 0.03, which is showed in Figure 4.

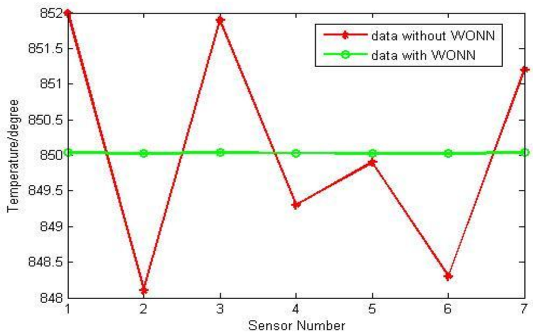

Figure 4 Data analysis with VS without WONN

Table 2 shows the comparison of storage, computing abilities and price of traditional servers and the cloud server. Obviously, with the help of distributed storage and computing model, the cloud server has better storage and computing ability.

Table 2 Single server VS cloud server

Type	Single Server	Cloud Server
Name	Sun SPARC Enterprise M500	IBM System x3850 X5
Number	1	28
CPU	8	2 × 28
RAM	256G	16G × 28
Hard Disk	146G × 4	>>200G
Price	￥1629000	￥58000 × 28

6 Conclusions

The congestion control algorithm for WSN not only prevents network congestion caused by data booming during the real time monitoring of the disaster, but also improves the accuracy of regional disaster monitoring. The introduction of the cloud platform on one hand effectively improves the ability of mass data storage and operation so as to make disaster data be accurately gradient analyzed and finally achieve a disaster prediction; On the other hand it reduces the system server construction and maintenance cost, make the system more practically implemented. Hence, under such a situation, the WSN system based on cloud platform will effectively solve the compatibility issue between real time monitoring and non-real-time prediction of disasters.

Acknowledgments

This research is partly supported by China Important National Science and Technology Specific Project with No.2009ZX03007-003-01, the National Natural Science Foundation of China under Grant No.60703036 and 410440040, Chinese Universities Scientific Fund (2009RC0310). Also, thanks to "International research laboratory on electromagnetic theory and application-wireless telecommunication engineering" jointly set up by Beijing University of Posts and Telecommunication and Queen Marry, University of London.

References

[1] C Wan, S B Eisenman, A T Camp bell. DODA: Congestion Detection and Avoidance in Sensor Networks [C]. First ACM Conference on Embedded Net worked Sensor Systems, 2003.

[2] Jaewon Kang. Adaptive Resource Control Scheme to Alleviage Congestion in Sensor Networks [C]. Basenets, 2004.

[3] Ping Wang, Mingchuan Cheng and Wenzao Shi. Research on Data Fusion Algorithm in ZigBee Protocol. Computer Science and Computational Technology, 2008.

[4] Cheng TienEe, RuzenaBajcsy. Congestion Control and Fairness for Many to One Routing in Sensor Networks[C]. ACM SenSys, 2004.

[5] Delyon B, Juditsky A, Benveniste A. Accuracy analysis for wavelet approximations [J]. IEEE Trans on NN, 1995, 6(2):332-348.

[6] Shi Hengliang, Bai Guangyi and Gui Wenming. An Implementation of Novel Cloud Computing Model - Improved Map-reduce Model. International Conference on Grid and Cooperative Computing. 2010.

[7] Xinjian Xiang. Multi Sensor Fusion Method based on Fuzzy and statistical theory. Journal of Transducer Technology. 2004.