Journal of Computing and Electronic Information Management

ISSN: 2413-1660

Design and implementation of data management system for hazardous chemicals transportation based on SOM

Zhuyao Liu, Lu Zhang [a, *]

School of Electrical and Information Engineering, Quzhou University, Quzhou, 324000, P.R. China

[a]shanwen1122@outlook.com

Abstract: The research on transportation process monitoring and risk assessment are both important parts of the transportation safety of hazardous chemicals, and the respective research and application have also achieved quite good results. As a safety auxiliary system, this subsystem will participate in the transportation risk assessment and early warning before the transportation of hazardous chemicals, and will definitely become the development trend of the hazardous chemical transportation safety application system. During the development of the hazardous chemicals transportation data management subsystem, the hazardous chemicals transportation data involves the problem of mobile risks. There are major difficulties in quantitative analysis of transportation risks. Currently, many risk analysis models have been established, but different analysis models The results obtained vary widely. Therefore, when the system conducts transportation risk assessment, according to our existing actual situation, the data model of the data management subsystem is built based on the SOM algorithm in machine learning.

Keywords: machine learning, data modeling, risk assessment.

1. Introduction

The losses caused by hazardous chemical transportation accidents are often difficult to estimate, which makes governments and researchers in various countries realize the importance of hazardous chemical management early. Hazardous chemicals transportation risk assessment is an important part of the management of hazardous chemicals, so it pays great attention to the research of transportation risks.

In the 1970s, the National Transportation Safety Administration promulgated regulations governing the transportation of hazardous chemicals to assess and analyze the risks of nuclear fuel transportation. However, the research on the risk assessment of hazardous chemicals transportation started late in the country, and there was less relevant research data in the beginning. However, with the development of China's industrialization, the safety of hazardous chemicals transportation has been valued, and risk assessment is also one of the precautions to measure safety. It has attracted much attention and many valuable results have emerged.

It was not until the end of the 1990s that China began to carry out comparative research on the risk model of hazardous chemicals transportation. Some scholars have proposed a risk balance model, which analyzes the individual risks and transportation costs of hazardous chemicals transportation under time-varying conditions. At the same time, some people have proposed a general framework for road transport risk assessment of hazardous chemicals, and in-depth research on the minimum transport accident rate and personnel risk model. Guo Xiaolin added accident classification to the original model when researching the risk measurement model, which improved the accuracy of the assessment results. On the basis of analytic hierarchy process and fuzzy comprehensive evaluation method, some people assign values to the relevant influence factors in the transportation of hazardous chemicals, and carry out risk assessment on the transportation routes of hazardous chemicals.

Therefore, this system stands on the shoulders of the predecessors and developed a hazardous chemicals transportation data management subsystem. As a safety auxiliary system, this system will participate in the transportation risk assessment and early warning before the transportation of hazardous chemicals, and will certainly become the development trend of the hazardous chemical transportation safety application system. During the development of the hazardous chemicals transportation data management subsystem, the hazardous chemicals transportation data involves the problem of mobile risks. There are major difficulties in quantitative analysis of transportation risks. Currently, many risk analysis models have been established, but different analysis models The results obtained vary widely.

Therefore, when the system conducts transportation risk assessment, according to our existing actual situation, the data model of the data management subsystem is built based on the SOM algorithm in machine learning.

2. Establish system infrastructure

Authentication and authorization: It is implemented by OAuth protocol. This function module filters illegal users and illegal operations to ensure data security and business correctness.

System management: Manage and maintain basic data supporting the operation of the system, including:

Rights management:

Role management
Resource management
Authorization
User management: manage and maintain user information for operating the system;
System operation parameter management: system basic data management, such as remote data port, analysis mode setting, etc.
System log: record various detailed information during system operation;
Vehicle management: manage vehicle information data of special materials transportation operations
Vehicle group management: group vehicles according to business type (for example, transportation direction, different materials)
Vehicle basic information management: manage the basic information of the vehicle, such as: license plate, vehicle type, vehicle condition, etc.
Driver and passenger management:
Personnel group management: group personnel according to their division of labor, such as driver group, flight attendant group, etc.
Basic personnel information management: such as age, gender, ID number, driver's license information, whether to enable it, etc.
Transportation plan management:
Customer information management: manage basic customer information that needs to be transported, such as: address, contact person, etc .;
Transportation plan formulation: develop transportation plans, such as: destination, customer, departure time, planned arrival time, material information, driver and passenger arrangement, route data, backup route data, plan implementation status, etc.
Vehicle real-time information monitoring:
Remote data analysis: The remote real-time data obtained through the interface is parsed into a data structure that meets the requirements of this system for use by other modules.
Real-time vehicle information display: Display the parsed vehicle real-time data to users in a friendly manner.
Abnormal information alarm: compare real-time information with planned information. If an abnormality occurs, the user will be proactively alerted. Anomaly monitoring indicators include:
Index of abnormal stop time during transportation;
Vehicle GPS online information abnormal indicators;
Vehicle GPS power failure information abnormal indicators;
GPS positioning matching index for vehicle weighing time;

Query statistics: (specific query statistics indicators need to be given)
Query related data of the system;
Statistics of relevant data of the system;
System message bus: The business logic between each module of the system interacts through the system message bus
System data bus: All business data of the system are transmitted through the data bus
Remote data acquisition: acquire remote data according to system policies, such as regular extraction;
Database: Stores the data required or produced by the above modules.
Establish a system infrastructure based on the analysis of the needs for early warning of dangerous goods transportation risk assessment. Dangerous vehicle transportation involves a lot of content, such as the management of the vehicle itself, the management of transportation items, the planning of the transportation route, the management of the crew and passengers, and the interaction with other third-party systems. Each piece of content can constitute an independent application, and each application requires large-scale data support. In order to achieve the goal of this project, a set of infrastructure is required to effectively combine the functions of each system and organize the data of each system.
This project intends to use the microservice architecture as the project's infrastructure. Microservices use multiple services to build an application. Various services are independently deployed in different processes. Different services communicate through some lightweight interaction mechanisms. This project will deploy the management of vehicles involved in the transportation of dangerous vehicles, the management of transported items, the planning of transport routes, the management of crew and passengers, and the interaction with other third-party systems as independent services to deploy and provide various applications. The functions are organized to jointly accomplish the objectives of this project.

3. Data management subsystem architecture

Based on risk assessment and early warning demand analysis of dangerous goods transportation, the system infrastructure is established. There are many contents involved in the transportation of dangerous vehicles, such as the management of the vehicles themselves, the management of transportation articles, the planning of transportation routes, the management of the personnel of the company and the interaction with other third-party systems. Each piece of content can constitute an independent application in which large-scale data support is required. To achieve the objectives of this project, a set of infrastructures are required to effectively combine the system functions and organize the system data. The solution infrastructure is shown in figure 2.

3.1 Data modeling

Format the data processed by the ETL tool to build a transport data model. mainly include:

Data Node = {vehicle information, location information, item information, customer information, route information, data acquisition time}

Among them vehicle information = {license plate, service life, driver information, vehicle length, vehicle weight}

Location information = {start longitude, start latitude, start place name, plan end longitude, plan end latitude, end place name, real time longitude, real time latitude, real time place name}

Item information = {Item number, item type, danger level, item form (solid, liquid, gaseous), weight, volume}

Customer information = {customer number, customer address, customer name, contract period}, route information = {planned route}.

3.2 Classification of transport data

The amount of data obtained from the remote end is huge, and when the system is established, there is no demarcation of which type of data is normal and which type is abnormal. Therefore, on the basis of data structure and taking historical data as samples, an unsupervised learning algorithm is designed to cluster the data before distinguishing normal and abnormal data by manual calibration.

The SOM algorithm is adopted without supervised learning algorithm. The idea is simple, essentially a neural network with only the input layer, the hidden layer. A node in the hidden layer represents a class that needs to be clustered. Training takes the form of "competitive learning," where each input sample finds its best match in the hidden layer, called its activation node, or "Winning Neuron." The parameters of the activated node were updated by stochastic gradient descent method. At the same time, the points adjacent to the activation node are updated appropriately according to their distance to the activation node.

Therefore, one of the characteristics of SOM is that the nodes of hidden layers are topologically related. This topological relation requires us to determine, if we want a one-dimensional model, then the hidden nodes are connected in order to form a line; If you want a two-dimensional topological relationship, you can make a plane, as shown in figure 1.

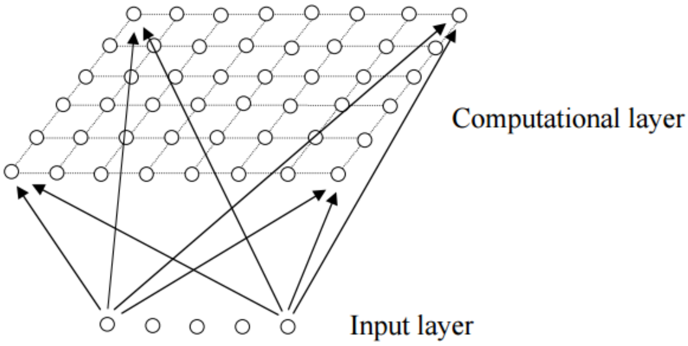

Figure 1 Korhonen Network

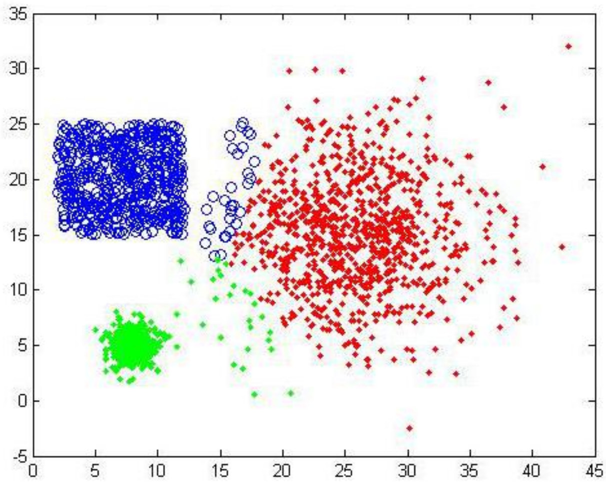

Figure 4 SOM convergence effect

Since the hidden layer has a topological relation, we can also say that SOM can discretize the input of any dimension into a discrete space of one or two dimensions (uncommon for higher dimensions). Nodes in the Computation layer are fully connected to nodes in the Input layer.

After the topology is determined, the calculation process is started, which is divided into several parts:

1) Initialization: each node randomly initializes its own parameters. Each node has the same number of parameters as the Input dimension.

2) For each input data, find the node that best matches it. Suppose D dimension is input, namely X={x_i, I =1... D}, then the discriminant function can be Euclidean distance:

$$d_j(\mathbf{x}) = \sum_{i=1}^{D}(x_i - w_{ji})^2$$

3) After the activation node I(x) is found, we also want to update the nodes adjacent to it. Let S_ij represent the distance between node I and j, and for nodes adjacent to I(x), assign them an update weight:

$$T_{j,I(\mathbf{x})} = \exp(-S_{j,I(\mathbf{x})}^2 / 2\sigma^2)$$

Simply put, nearby nodes are less updated based on distance.

4) The parameters of the node are then updated. Update according to gradient descent method:

$$\Delta w_{ji} = \eta(t) \,.\, T_{j,I(\mathbf{x})}(t) \,.\, (x_i - w_{ji})$$

Iterate until it converges. SOM can effectively distinguish data categories, and the convergence effect is shown in figure 6.

3.3 Classifier training

1) Risk assessment classifier

After data classification, manual calibration of normal data and abnormal data. Using the calibrated data as positive and negative samples, the supervised learning classifier was trained.

There is a supervised learning classifier, using SVM, support vector machine. The marked abnormal data was taken as the positive sample and the normal data as the negative sample, and the SVM was trained. The trained SVM was used to test the abnormal data and obtain the data samples of wrong classification. This part of the sample was taken as a difficult example and put into the sample for retraining, so as to eventually generate the SVM for transporting the abnormal data. Provided for use by other business modules.

2) Detection and warning model

Because the transportation process is a continuous process, in which vehicles, transportation goods, personnel and road conditions are all changing, the hidden markov model is adopted to build the detection and warning model.

The hidden markov model is a probabilistic model of time series, which describes the process of generating unobservable state random sequences from a hidden markov chain and then generating observation random sequences from each state. The sequence of state generated randomly by hidden markov chains is called state sequence. Each state generates an observation, and the resulting sequence of observations is called an observation sequence. Each position of the sequence can be regarded as a moment.

Given that Q is the set of all possible states, V is the set of all possible observations, that is, data such as vehicle condition, road condition, personnel status and so on collected through information means such as Internet of things sensor.

$$Q = \{q_1, q_2, \cdots, q_n\}, V = \{v_1, v_2, \cdots, v_m\}$$

Where n is the number of possible states and m is the number of possible observations. H is the state sequence with length T, and O is the corresponding observation sequence.

$$H = \{h_1, h_2, \cdots, h_T\}, O = \{o_1, o_2, \cdots, o_T\}$$

A is the state transition matrix: $A = [a_{ij}]_{n \times n}$ where,

$$a_{ij} = P(i_{t+1} = q_j | i_t = q_i), i = 1, 2, \cdots, n; j = 1, 2, \cdots, n$$

Is the probability that t+1 is transferred to state qj at time t is in state qi.
B is the probability $B = [b_j(k)]_{n \times m}$ that is the observation probability matrix is the probability that the observation vk is generated under the condition that t is in state qj.
So we say that is the initial state probability vector.

4. Conclusion

This system application information means such as the Internet of things, to establish a relatively perfect special vehicles supervisory system, the special vehicle can be real-time monitoring, to strengthen the tracking of dangerous goods transport relief and scheduling management, relevant departments can effectively supervise the driver violations, such as speeding, fatigue driving is speeding, anchor, captured, and so on and so forth of the transport vehicle alarm and automatically shut down measures, reduce the incidence of traffic accidents, increase safety factor for special vehicles, improve the operation efficiency of enterprises, to safeguard the driver person property safety and transportation safety also has the positive significance.
The above scheme can effectively improve the efficiency of emergency treatment and rescue work after an accident. However, this kind of scheme is unable to complete the task of risk assessment, detection, prediction and early warning before the departure of dangerous vehicles.
This project aims to establish a special vehicle transportation risk assessment, detection, prediction, warning and real-time monitoring system on the basis of Internet of things, big data, cloud computing and other new technologies.

Acknowledgments

This work was partially supported by Science and Technology Innovation Program for College Students in Zhejiang Province (New Miao Talents Program)(No.2018R435011).

References

[1] Yang qiang, li yanlai. Risk analysis of dangerous goods transportation based on quality function allocation method [J]. Journal of Chinese safety science, 2017(3): p65~70
[2] Yang na. Analysis and research on transportation safety strategies of dangerous goods [J]. People's transportation, 2018(6): p74~75
[3] Chen xiaoyong, shi shiliang, ren jingzhou, li runqiu, you bo. Statistical analysis and

countermeasures of road traffic accidents of hazardous chemicals in China from 2013 to 2014 [J]. Journal of hunan university of science and technology (natural science edition), 2017(9): p91~95

[4]Journal of east China university of science and technology (natural science edition), 2017(8): 591-596. (in Chinese with English abstract

[5] wang yunpeng, zhao hantao, wang junxi et al. Research on the integration framework of expressway system based on the common information platform [J], highway traffic science and technology,2016,23(2): p128~132.