Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 7, No. 1, 2023

Train Brake System Pipe Leakage Detection and Early Warning Method Based on Bayesian Networks

Jiaqiao Hu, Liming Zhou, Qiang Wu, Jin Hua

Crrc Nanjingpuzhen Co., Ltd., China

Abstract: This paper proposes a method for detecting and warning about leaks in train braking system pipelines based on Bayesian networks. Firstly, a detection model for pipeline leaks is established through the learning and inference of Bayesian networks. In the anomaly detection phase, the Bayesian network model is trained using historical data to monitor brake pressure abnormalities in real-time. Secondly, in the parameter regression calibration phase, the location and severity of the pipeline leaks are estimated based on the current brake pressure and relevant parameters. Finally, in the fault inference phase, the Bayesian network model is used to infer the possible causes of the leaks. The effectiveness and reliability of this method are verified through simulation design and actual data analysis. Compared to existing methods, this method can provide accurate leak detection and warning, thereby contributing to the safety of train operation. This research provides an effective method for detecting and warning about leaks in brake system pipelines and has practical application value.

Keywords: Bayesian Networks, Train Braking System, Pipe Leak Detection, Early Warning Methods.

1. Introduction

With the rapid development of modern transportation systems, train transportation has gained increasing attention as an efficient and convenient mode of transportation. Among them, the train braking system is an important component to ensure the safety of train operation. However, the issue of leaks in the brake system pipelines poses potential safety hazards to train operation. Therefore, it is crucial to develop an efficient and reliable method for detecting and warning about leaks in train braking system pipelines.

Bayesian networks, as powerful probabilistic inference tools, have been widely applied in various fields. They can accurately model the dependencies and uncertainties among variables, making them an effective approach to solving complex problems. This paper aims to combine the Bayesian network algorithm and propose a method for detecting and warning about leaks in train braking system pipelines based on Bayesian networks, aiming to enhance the safety of the train braking system [Chi Z,2023].

This study will focus on several key issues: Firstly, through the study of the Bayesian network algorithm, the definition, learning, and inference methods of Bayesian networks will be thoroughly explored. Secondly, a detection model for leaks in train braking system pipelines will be established based on Bayesian networks, including the stages of anomaly detection, parameter regression calibration, and fault inference, in order to improve the accuracy of leak detection and the precision of prediction. Finally, through simulation design and analysis of actual data verification, the effectiveness and feasibility of the proposed method will be evaluated.

This research aims to provide a reliable and efficient solution for detecting and warning about leaks in train braking system pipelines, thus providing strong support for ensuring the safety of train operation. Additionally, this study will deepen the understanding of the application of the Bayesian network algorithm in the field of engineering, providing valuable references and insights, and possessing broad practical application value and scientific research significance.

2. A Study of Bayesian Network Algorithms

2.1. Definition of Bayesian network

Bayesian networks are a probabilistic graphical model (as shown in Figure 1) used to describe the conditional dependencies between random variables. They are represented as directed acyclic graphs, where nodes represent random variables and edges represent the conditional dependencies between variables. Bayesian networks utilize conditional probability tables to represent the dependencies and uncertainties between variables, enabling probabilistic inference and prediction. In a Bayesian network, each node represents a random variable, and the parent nodes represent the direct influencing factors of that variable. The state of each node is determined by its parent nodes and the conditional probability tables, which reflect the probability distribution of the node's state given the states of its parent nodes. Through Bayesian network inference, it is possible to infer the posterior probability distribution of other variables based on the observed variables, thus facilitating reasoning and prediction in the system.

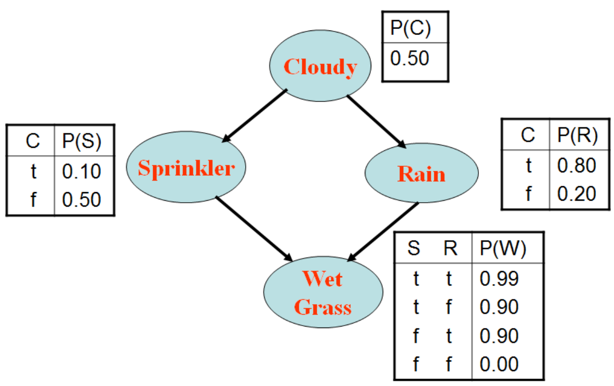

Figure 1. Bayesian network

Bayesian networks offer flexibility and interpretability, enabling the handling of incomplete information and uncertainty problems. They are widely applied in various fields, such as decision analysis, risk assessment, and medical diagnosis. By understanding the definition and principles of Bayesian networks, we can better establish probabilistic models, conduct inference, and make decisions, providing effective tools and methods for the analysis and resolution of complex problems [Yang Z,2023].

2.2. Learning in Bayesian networks

Learning in Bayesian networks refers to the process of inferring the structure and parameters of the Bayesian network from observed data. The goal of learning Bayesian networks is to obtain the conditional dependencies between variables based on the observed data, thus establishing an accurate probabilistic model.

The process of learning Bayesian networks can be divided into two parts: structure learning and parameter learning. The objective of structure learning is to determine the topology of the Bayesian network, i.e., the dependencies between variables and the connections of directed edges. The flowchart of structure learning is depicted in Figure 2.

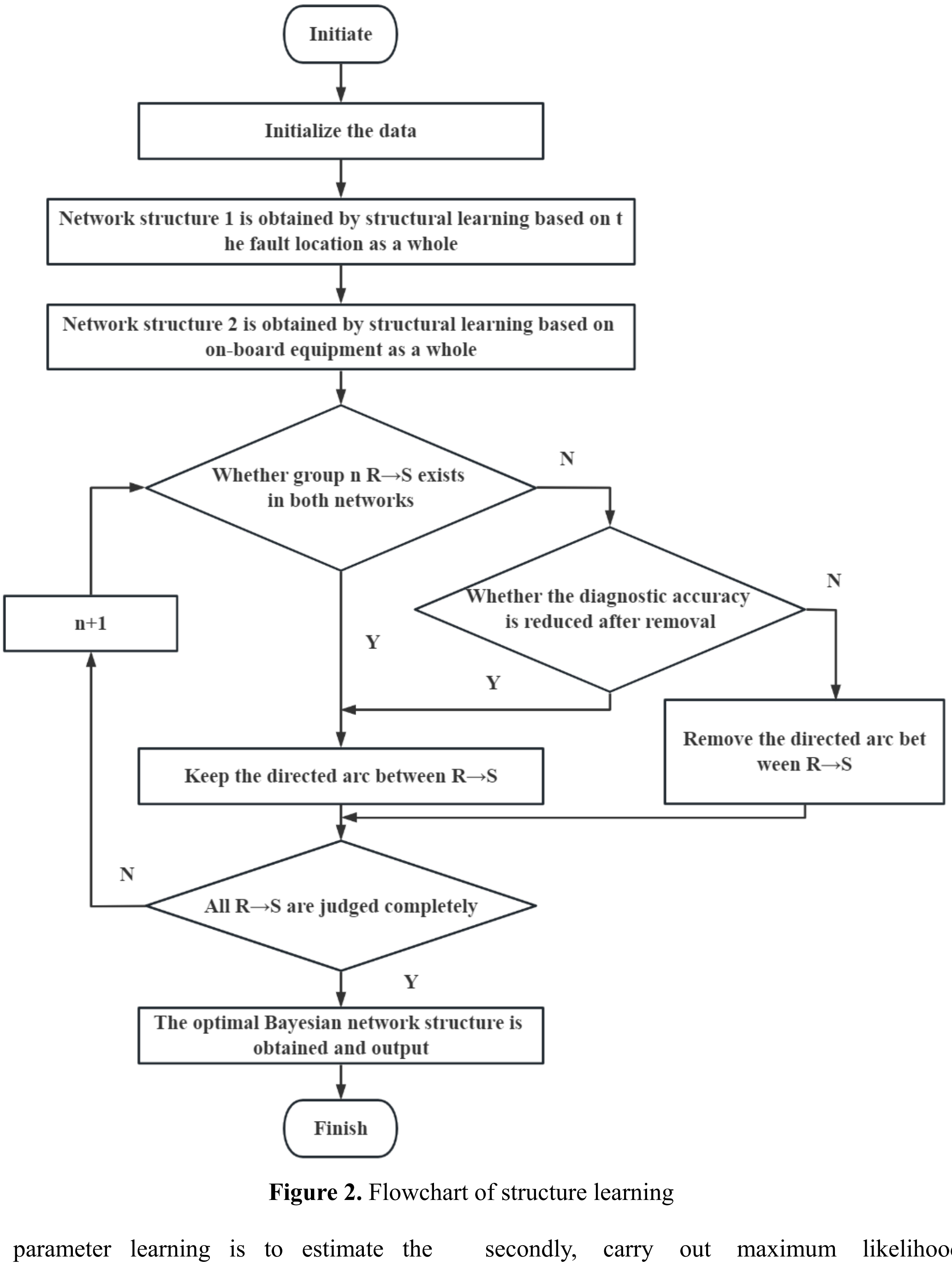

Figure 2. Flowchart of structure learning

The goal of parameter learning is to estimate the conditional probability table in a Bayesian network based on the observed data. In Figure 2, parameter learning can be carried out after the optimal Bayesian network structure is obtained. At this time, the model is known and the parameters are not determined, so this paper adopts the method of great likelihood estimation for parameter learning.

Firstly, assume a certain but unknown parameter 0, and secondly, carry out maximum likelihood parameter estimation on the sample set. Remember the known sample set is represented as shown in equation (1):

$$D = \{x_1, x_2, \cdots, x_n\} \tag{1}$$

Then the corresponding likelihood function is expressed as shown in Equation (2):

$$L(\theta) = P(x_1, x_2, \cdots, x_n \mid \theta) = \prod_{i=1}^{n} P(x_i \mid \theta) \tag{2}$$

Then solve for the parameter that maximizes the likelihood function, which is expressed as shown in Equation (3):

$$\hat{\theta} = \arg\max_{\theta} L(\theta) = \arg\max_{\theta} \prod_{i=1}^{n} P(x_i \mid \theta) \tag{3}$$

The log-likelihood function is expressed as shown in Equation (4):

$$H(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln P(x_i \mid \theta) \tag{4}$$

Its results after parameter learning are shown in Table 1:

Table 1. Results after parameter learning

Node number	1(F)	2(T)
R1	0.65	0.35
R2	0.74	0.26
R3	0.73	0.27
R4	0.87	0.13
R5	0.61	0.39
R6	0.63	0.37
R7	0.76	0.24

The learning process of Bayesian networks depends on reliable datasets and appropriate learning algorithms. The quality of the dataset is crucial for the accuracy of the learning results. Additionally, the choice of learning algorithm also affects the effectiveness and speed of learning. Different learning algorithms are suitable for different learning scenarios and data attributes. Learning Bayesian networks is an iterative process that requires continuous optimization and adjustment of the network. By learning Bayesian networks, we can capture the dependencies between variables from data and utilize these dependencies for probabilistic inference and prediction. The learning of Bayesian networks holds significant applications and research value in fields such as data mining, data analysis, and artificial intelligence [Doyeob Y,2020].

2.3. Inference in Bayesian networks

Inference in Bayesian networks refers to the process of calculating the probability distribution of unknown variables based on known observed data and the model. Through inference in Bayesian networks, we can use the observed variables to infer the posterior probability distribution of other variables, enabling probabilistic inference and prediction. In this paper, the variable elimination algorithm (as shown in Figure 3) is adopted for the fault diagnosis and warning inference of the subway braking system.

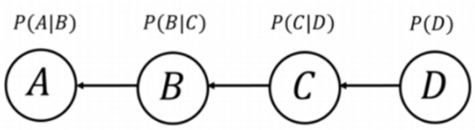

Figure 3. Example of variable elimination method

As shown in the figure, assuming that the goal of reasoning in this network is computation, it is specified as shown in Equation (5):

$$\begin{aligned} P(A) &= \sum_{B,C,D} P(A,B,C,D) \\ &= \sum_{B,C,D} P(D)P(C \mid D)P(B \mid C)P(A \mid B) \end{aligned} \tag{5}$$

Equation (5) can be rewritten as Equation (6) if the sequential calculations used sum the results of the factorization associated with each variable separately:

$$\begin{aligned} P(A) &= \sum_{C} P(A \mid B) \sum_{C} P(B \mid C) \sum_{D} P(D)P(C \mid D) \\ &= \sum_{B} P(A \mid B) \sum_{C} P(B \mid C) m_1(C) \\ &= \sum_{A} P(A \mid B) m_2(B) \\ &= m_3(A) \end{aligned} \tag{6}$$

Where, denotes the intermediate result obtained in the first summation calculation, which is only related to the variables. When the above formula is used, the number of product and summation calculations are reduced, and the reasoning process is effectively simplified.

3. Train Brake System Pipe Leakage Detection Model Based on Bayesian Networks

Bayesian network based pipe leakage detection model for train braking system is a method of modeling and reasoning using Bayesian network for detecting pipe leakage problem in train braking system. The simplified model diagram of the subway train braking system is shown in Figure 4:

The model first establishes a Bayesian network, where nodes represent different components or sensors in the train braking system, and edges represent the dependencies between components. For example, nodes can represent the brake pipeline, pressure sensors, and brake pistons.

During the training phase of the model, the structure and parameters of the Bayesian network need to be learned from the observed data. The objective of structure learning is to determine the network's topology, i.e., the dependencies and connections between nodes. The objective of parameter learning is to estimate the conditional probability tables between nodes, representing the state distribution of nodes given the states of their parent nodes. Once the model is trained, inference can be performed using the observed data.

When the train braking system is operational, the model receives real-time data from sensors, such as the pressure values in different pipelines. The model, utilizing Bayesian inference, calculates the posterior probability distribution of the nodes based on this data. By analyzing the posterior probabilities, it is possible to determine the presence of pipeline leaks, as well as their location and severity.

The advantage of this model is its ability to accurately infer the existence of pipeline leaks in the train braking system using existing knowledge and observed data. By monitoring and diagnosing in real-time, appropriate measures can be taken promptly to repair any leaks, ensuring the normal operation of the train braking system and the safety of passengers. However, the accuracy and performance of the model are also influenced by factors such as data quality, the scale of model training, and the complexity of the model. It requires sufficient validation and fine-tuning in real-world applications [Feng H D,2018].

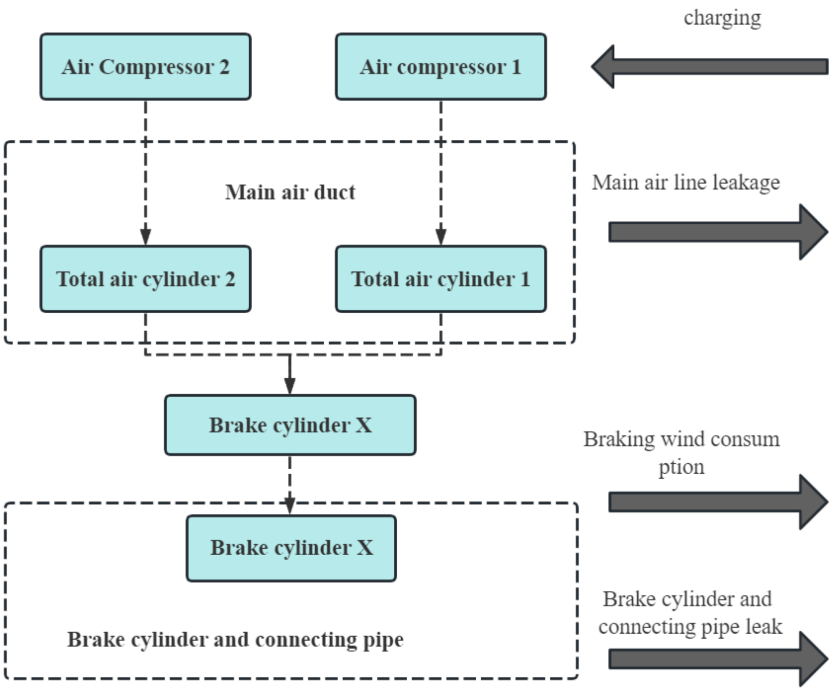

Figure 4. Simplified model of braking system of subway train

4. Validation Analysis

In order to validate the effectiveness of our proposed method, the following analysis has been performed:

4.1. Simulation Design

Simulation design refers to the process of using computer models and algorithms to simulate and mimic the behavior and performance of actual systems or processes. In the field of engineering, simulation design is widely applied in system design, performance evaluation, and optimization. In simulation design, it is necessary to first establish a mathematical model of the system that describes its structure, characteristics, and operating rules. The model can be based on equations derived from physical principles, stochastic processes based on statistics, or simulation models based on logical rules. In practical engineering, simulation design plays a vital role. It can be used to verify system design solutions, predict and analyze system performance before product development, and reduce the cost and risks associated with physical experiments. For example, in the field of transportation, simulation design can be used to simulate urban traffic flow and evaluate the effectiveness of traffic planning and control strategies. In the manufacturing industry, simulation design can optimize production line layouts and process flows, thereby improving production efficiency and quality.

4.2. Data set description and confidence analysis

Data set description and plausibility analysis is the process of describing a data set in detail and assessing its plausibility. This is a very important step before data analysis and modeling to ensure the quality and applicability of the data.

In order to verify the reliability of the optimal Bayesian network fault diagnosis model proposed in this paper, its calculation principle is shown in Equation (7):

$$F_1 = 2PR/(P+R) \tag{7}$$

Data set description and confidence analysis is the process of describing a data set in detail and assessing its confidence prior to data analysis and modeling. Data description includes information such as data source, data collection process, data content, data format, data volume and dimensionality. Credibility analysis includes methods such as data quality checking, data source validation, sampling method assessment, data consistency checking, and domain expert assessment. Through dataset description and plausibility analysis, we are able to gain a comprehensive understanding of the dataset and ensure the quality and applicability of the data. This helps to improve the accuracy and reliability of data analysis and modeling, ensures that the conclusions and decisions we obtain are credible, and provides a reliable basis for further data processing and analysis.

4.3. Diagnostic inference

Diagnostic inference results are diagnostic conclusions obtained by analyzing and extrapolating given observations through inference methods. These reasoning results can be used to determine the state, problem, or abnormality of a system or process. Through diagnostic reasoning results, we are able to perform accurate system analysis and scenario prediction using existing knowledge and observational data. This helps to identify and solve problems in a timely manner, improve system efficiency and reliability, and provide effective support for decision making and problem solving. However, the accuracy and reliability of the inference results are affected by factors such as data quality, accuracy of the model and choice of inference method, and need to be fully validated and evaluated [Quan W,2023].

5. Conclusion

By conducting research on a Bayesian network-based method for detecting and predicting leaks in train braking system pipelines, this paper effectively addresses the safety risks associated with such leaks. This method combines the characteristics and advantages of Bayesian network algorithms. Through simulation design and validation using actual datasets, the results demonstrate that the method can reliably detect pipeline leaks in different scenarios and provide accurate localization and fault inference. Compared to traditional methods, the Bayesian network-based approach exhibits higher accuracy and reliability, enabling early detection of pipeline leaks and timely repair measures, thereby enhancing the safety and reliability of train braking systems. In summary, the Bayesian network-based method for detecting and predicting leaks in train braking system pipelines is of significant importance in improving train operation safety. Future research can expand the application of this method to broader domains and combine it with other fault diagnosis and prediction methods, making greater contributions to ensuring the safety and reliability of transportation.

References

[1] Chi Z,J. B A,L. M S, et al. A convolutional neural network for pipe crack and leak detection in smart water network[J]. Structural Health Monitoring,2023(1):22.

[2] Yang Z,Huilin P,Yang Z, et al. Efficient visual fault detection for freight train braking system via heterogeneous self distillation in the wild[J]. Advanced Engineering Informatics,2023:57.

[3] Doyeob Y,Giyoung L,Cheol J L. Pipe Leak Detection System using Wireless Acoustic Sensor Module and Deep Auto-Encoder[J]. Journal of the Korea Society of Computer and Information,2020(2):25.

[4] Feng H D. Accuracy and Sensitivity Evaluation of TFR Method for Leak Detection in Multiple-Pipeline Water Supply Systems[J]. Water Resources Management,2018(6):32.

[5] Quan W,Zhiwei W,Jiliang M, et al. Modelling and stability analysis of a high-speed train braking system[J]. International Journal of Mechanical Sciences,2023:250.