

Article

Research on the Causes and Transmission Mechanisms of Railway Engineering Safety Risks

Tongyu Zhang ¹, Xuewei Li ^{2,*} and Xueyan Li ³

¹ College of Urban Rail Transit and Logistics, Beijing Union University, Beijing 100101, China; zhangtongyu0430@163.com

² School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China

³ Management College, Beijing Union University, Beijing 100101, China; gongye1632006@163.com

* Correspondence: xwli@bjtu.edu.cn

Abstract: In recent years, railway safety accidents have repeatedly occurred. Any omission in the process of management or operation can easily have very serious consequences. This study aimed to examine the causes and transmission mechanisms of safety risks in railway engineering departments. First, the multi-objective particle swarm optimization algorithm was employed to determine the key risk factors, allowing for indicator screening that was in line with the requirements of practical applications. Then, Bayesian networks were used, and their structure was optimized to analyze the propagation diagnosis and probability of key risk indicators, obtaining the causal logic chain that produces accidents and, from that, the four aspects (human, machine, environment, management) of the corresponding prevention of risk recommendations. Finally, in this article, it is shown that combining the indicators and Bayesian networks can improve the accuracy of risk prediction and provide more accurate results than using existing research and, hence, it can fill the gap in research on railway safety risks in risk transmission mechanisms.

Keywords: security risk; index screening; conduction mechanism; multi-objective optimization; Bayesian network



Citation: Zhang, T.; Li, X.; Li, X. Research on the Causes and Transmission Mechanisms of Railway Engineering Safety Risks. *Appl. Sci.* **2024**, *14*, 2739. <https://doi.org/10.3390/app14072739>

Academic Editor: Diogo Ribeiro

Received: 10 February 2024

Revised: 20 March 2024

Accepted: 22 March 2024

Published: 25 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of railway-related technologies and the continuous maturing of scientific theories, railway safety risks have attracted increasing research attention. Railway operation safety has become one of the key research directions in railway-related research. Various analysis methods have been used to conduct a large number of safety risk assessment analyses and experiments in different scenarios, such as car, machine, industrial, electric, and railway vehicles, aiming to develop scientific methods and decision-making supports for railway operation management departments to avoid or reduce the occurrence of high-speed railway safety accidents in practical scenarios.

Yankun Zhang [1] divided and analyzed the factors of high-speed railway vehicle safety into four aspects: human, machine, environment, and management. Following the principle of establishing evaluation indicators, these main factors were selected as evaluation indicators, and an AHP-fuzzy comprehensive evaluation model was established. Many examples were provided for the purpose of verification and application, and safety evaluation results were obtained. Similarly, Chengrui Li et al. [2] analyzed the safety-influencing factors of high-speed railway passenger stations. They established a three-layer safety evaluation index system, constructed a safety evaluation model using a fuzzy comprehensive evaluation method, and conducted a safety evaluation on the example of a high-speed railway passenger station in Northeast China. Honglei Yao [3] used the analytic hierarchy process to scientifically design a risk assessment index system for railways' key information infrastructure. Using weight data as the input data for the backpropagation neural network, the risk assessment results were predicted and the

risk level was determined. Yaoru Zhang [4] designed a risk assessment index system for the cooperation of work and electricity with skylight operations and used the analytic hierarchy process to calculate the combined weights of various safety risk factors. The impact of each risk factor on the safety of work and electricity with skylight operations was determined, and corresponding measures were proposed concerning four aspects: personnel, management, equipment, and environment. Xiangyu Sun [5] used relevant theories and modeling methods, such as the human factors classification and analysis system, logit model, and interpretive structural model, to identify and analyze the main human factors of five railway engineering liability accidents; they also analyzed the hierarchical relationship of the causal factors for the main problems. Kurz S.L. and Milius B. [6] focused on the risks that could not be ignored in risk assessment by conducting a risk assessment of European railways. Reinach S. and Viale A. [7] applied the Human Factors Analysis and Classification System (HFACS) accident analysis method to railway accident research and conducted a comprehensive study of human factors in accidents using data on British railway accidents, thus expanding the application of the HFACS to different accident fields. Zhan Q. et al. [8] developed an improved method for analyzing human factors in accidents based on the structural system of the HFACS accident analysis method, aiming to refine and explore the human and organizational factors in accidents. Lower M. et al. [9] combined the HFACS accident analysis method with the STAMP error classification method and then, based on the structure of the human factors model, separately controlled the causal factors at each level to identify potential hazards and proposed corresponding solutions.

The railway engineering department is one of the important parts of the railway system since it is responsible for the maintenance and repair of railway lines and related equipment. Its activities mainly include large and medium-sized maintenance and regular maintenance of bridges, tunnels, culverts, roadbeds, rails, turnouts, sleepers, and ballast, thus mostly mechanical operations. With the development of technology, various new technologies and equipment have been steadily introduced into railway engineering operations and management. However, due to various constraining factors, such as uneven worker quality, limited management level, and sudden and severe natural disasters, a large number of railway traffic accidents occur every year. According to the China Railway Yearbook (2021) [10], there were 138 responsible train accidents, including occupational injuries, in various engineering departments of the China Railway Group in 2020, including 3 major accidents, 8 general Class B accidents, 13 general Class C accidents, and 114 general Class D accidents. The total number of responsible accidents in 2020 increased by 21 compared to 2019.

Accidents occur repeatedly, and any negligence in the management and operation process can easily have severe consequences. Therefore, it is crucial to analyze the accident data, explore the internal mechanisms, potential laws, leading factors, and transmission mechanisms of accidents, and apply these experiences and laws to actual operation management, to improve the overall safety level of the railway operation management system. At the same time, considering the key defects, common deficiencies, frequent emergencies, equipment prone to malfunctions, and operational processes prone to problems, corresponding accident prevention measures should be taken, and efficient response strategies and emergency management measures should be implemented to reduce the occurrence of accidents and risk events. Thus, research in this area is of great significance for improving the overall level of safety management, reducing the number of accidents, ensuring personal and property safety, enhancing passenger satisfaction, and improving the efficiency of train operation organization [11].

Many existing studies have adopted a single method to analyze safety risks in a single scenario using a single indicator, without considering the current situation of various risk-related factors in a railway system, optimization under multi-objective conditions, and coordination problems with multiple different dimensions of demand [12–21]. There has been relatively little research on optimizing the indicator combinations, and some of the existing methods involve subjective evaluations, leaving them unable to determine the logical relationship between indicators and outputs.

To address the aforementioned shortcomings, this paper improves and optimizes the index screening analysis method, and it considers risk correlation factors and optimization under multi-objective conditions. First, the multi-objective particle swarm optimization algorithm is used to determine key risk factors, bringing the indicator screening more in line with the requirements of practical applications. Then, the Bayesian networks are selected, and their structure is optimized to analyze the propagation diagnosis and probability of key risk indicators and obtain the causal logic chain that produces accident results. Finally, the combination of the selected indicators and Bayesian networks is used to obtain more accurate results than those provided by the existing research and fill the gap in research on railway safety risks in risk transmission mechanisms.

2. Model Design

This section mainly introduces the indicators used in the process of model construction, including multi-objective optimization models, Pareto optimal solutions, multi-objective particle swarm optimization algorithms, and Bayesian networks.

2.1. Index Screening Multi-Objective Optimization Model and Pareto Optimal Solution

The principle of selecting key indicators is to remove redundant indicators while ensuring that the accuracy of railway engineering safety status recognition meets the requirements. In this study, the combination of indicators is defined as an independent variable, and x_k^R represents the binary code of the indicator combination. This principle can be expressed in the form of multi-objective programming as [22]

$$\left\{ \begin{array}{l} f_1 = \min_{x_k^R=0.1} \sum_{k=1}^m x_k^R \\ f_2 = \max_{x_k^R=0.1} G(x_1^R, \dots, x_k^R, \dots, x_m^R) \\ s.t. \sum_{k=1}^m x_k^R \leq m \end{array} \right. \quad (1)$$

Further, Equation (1) can be expressed in the standard form as follows:

$$\left\{ \begin{array}{l} \min F(x) = [f'_1(x), f'_2(x)]^T \\ s.t. x \in S \end{array} \right. \quad (2)$$

In Equations (1) and (2), G represents the accuracy of a machine-learning-based state recognition (classification) algorithm, and f'_1 and f'_2 denote the standard forms of f_1 and f_2 , respectively, where f_2 is the state recognition error. Based on the characteristics of machine-learning-based algorithms, there is a contradiction between various sub-objectives of the multi-objective optimization problem mentioned above. Therefore, this study adopts the Pareto optimal solution idea to balance the removal of redundant indicators with ensuring the accuracy of railway engineering safety state recognition.

Definition 1 (Dominant relationship of key indicator combinations). For an indicator combination $x_A, x_B \in S$, if there is $f'_i(x_A) \leq f'_i(x_B) (i = 1, 2, \dots, n)$ and at least one $j \in \{1, 2, \dots, n\}$, such that $f'_j(x_A) < f'_j(x_B)$, this combination is called the combination of key indicators x_A and the combination of dominant key indicators x_B , denoted as $x_A \prec x_B$.

Definition 2 (Combination of non-dominant key indicators). If the key indicator combination $x_C \in S$ is not dominated by any other key indicators, then combination x_C is called a non-dominated key indicator combination.

Definition 3 (Pareto frontier for key indicator screening). *The representation of the objective function value set calculated by combining all non-dominated key indicators in the solution space is called the Pareto front of key indicator screening (represented as PF), which is expressed as follows:*

$$PF = \{[f_1(x), f_2(x)] | x \in \text{Non dominated key indicator combination}\} \tag{3}$$

2.2. Multi-Objective Particle Swarm Optimization Algorithm

This study uses the multi-objective particle swarm optimization algorithm to optimize indicator selection, and the execution process of the algorithm is as follows [23,24].

Step 1: Establish the initial population. The solution set used to store individual historical solutions within the population is denoted by TY_i , and the solution set used to store the entire population is denoted by TY_j . The algorithm iteration time step number is set to the total number of particles, which is denoted by N' . Binary encoding is performed on a particle i in the population, where $x_i = [x_{i1}, x_{i2}, \dots, x_{ik}, \dots, x_{im}]$, $x_{i2} = 0.1$, and the corresponding solution to x_i is $F(x_i) = [F_1(x_i), F_2(x_i)]$;

Step 2: Use the random forest algorithm to train the dataset composed of indicator combinations. The random forest algorithm is selected because it performs well in handling imbalanced classification datasets and has been widely used for state recognition;

Step 3: Use the Euclidean distance to calculate the crowding distance in the entire population solution set TY_j , and use the crowding distance to determine the uniformity of the distribution of feasible solutions on the Pareto front. Then, the operations are performed:

- (1) Initialize the crowding distance of all individuals in set TY_j ;
- (2) For each objective function, sort the individuals in set TY_j according to the size of objective function m , that is, $Y = \text{sort}(F_i, m)$;
- (3) For individuals in set TY_j , which are at the boundary, set the crowding distance to infinity, that is, $TY(d_1) = \infty, TY(d_2) = \infty$;
- (4) For individuals outside the boundary, from $k = 2$ to $k = n - 1$, let

$$TY(d_k) = \sum_m \frac{TY(k+1) \cdot m - TY(k-1) \cdot m}{f_m^{max} - f_m^{min}}, \tag{4}$$

where $TY(k) \cdot m$ represents the value of the m th objective function of the k th individual in set Y .

Equation (4) indicates that after sorting according to the objective function value, the crowding distance of individuals in set TY_j represents the sum of differences between the objective function values of the previous and subsequent individuals.

Step 4: Select a risk-critical causal combination. Based on the crowding distance, the solution in set TY_i represents the individual, historical, optimal combination of key causes of risks. In set TY_j , the global, optimal combination of key causes of risks is selected, and the selection operator uses the widely used roulette wheel method in optimization algorithms. It should be noted that $\sum_{s=1}^K d_s$ is the sum of the crowded distances in set TY_j . According to the roulette wheel method, the probability of selecting the particles corresponding to group s of risk causation combinations is expressed by $\frac{d_s}{\sum d_s}$. The individual corresponding to the key causal combination of the individual historical risk is denoted by i' , and the selected individual in set TY_j is denoted by i'' ;

Step 5: Cross-operation. Convert the binary code $x_{i'}$ corresponding to an individual i' in set TY_i , individual $x_{i''}$ selected from TY_j , and the current binary code x_i of particle i to real code and denote their real encodings by $x_{i'}^R, x_{i''}^R$, and x_i^R , respectively; p_c is the probability of crossing. The cross-operation refers to the update of individuals in a population from their parents to their children, and the update process is expressed as follows:

$$\mu_i^R = (1 - p_c) \cdot x_{i'}^R + p_c \cdot x_{i''}^R \tag{5}$$

$$x_i^{R*} = (1 - p_c) \cdot x_i^R + p_c \cdot \mu_i^R \tag{6}$$

Step 6: Mutation operation. Assume that p_m represents the probability of variation and select particles in a particle swarm with a probability p_m . For a selected particle i , randomly select one of its risk-causing combinations corresponding to the binary code. If the bit is zero, it becomes one and vice versa;

Step 7: It can be seen that the newly generated offspring individuals and contemporary individuals jointly form a new population. Continuing to run the iterative algorithm will retain all excellent previous and contemporary particles in the current population. When the risk causal combination screening process meets the termination condition (i.e., the Pareto optimal solution no longer changes), the algorithm ends; otherwise, the algorithm returns to Step 2.

2.3. Bayesian Network

The Bayesian network is a directed acyclic graph with probability annotations, which is a knowledge visualization method based on probability analysis and graph theory. The Bayesian formula, also known as the inverse probability formula or a posterior probability formula, can be regarded as a causal search. Assume that B_1, B_2, \dots, B_n is a complete event with a positive probability, $P(B_i)$ is the prior probability, and $P(A|B_i)$ is the new additional information; then, the posterior probability is expressed by

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^n P(B_i)P(A|B_i)} \quad (7)$$

In the Bayesian network, nodes represent random variables abstracted from practical problems, and a directed arc between nodes reflects the correlation between them. The emitting node of a directed arc denotes a parent node, the pointed node represents a child node, the node without a parent node is a root node, and the node without a child node is a leaf node. The probability parameter value of the root node is a prior probability, which expresses the objective occurrence of the variable corresponding to the root node. The probability parameters of the other nodes in the network are taken as conditional probability tables, which express the correlation and influence with their parent nodes' variables. The Bayesian network can be utilized to conduct a qualitative analysis using its network structure (i.e., the causal relationships between events), as well as quantitative calculations using the prior or conditional probability parameters.

Bayesian network structure learning is to process data using certain algorithms to obtain the Bayesian network structure that best fits the given data. In other words, if there is a potential causal relationship between nodes x_1 and x_2 in the data, there must be a directed edge between x_1 and x_2 in the Bayesian network structure. Among numerous scoring algorithms for Bayesian network learning, the K2 algorithm proposed by Cooper G.F. in Ref. [25] is a greedy search algorithm that uses a scoring function to measure the degree of matching between a dataset and the network structure. The sample size of data used in this study is not particularly large, and according to [26], the K2 algorithm can effectively learn Bayesian network structures for small datasets in the case of ordered variables. Therefore, this study employs the K2 algorithm to learn the Bayesian network structure of railway engineering safety risks.

Similarly, Bayesian network parameter learning refers to obtaining the network nodes' parameters, namely the conditional probability table, for a particular Bayesian network structure. The conditional probability table, as a parameter of Bayesian networks, can quantify the strength of nodes' mutual influence. The research object of this study is the railway engineering safety risk, and the data source represents railway engineering accident cases from 2005 to 2013. By analyzing each accident case and extracting the causes of the risk, the "0-1" matrix is organized as parameter learning data. Since the data related to accident cases can have missing records, and the sample size of the collected data is small, the EM algorithm is selected in this study as a parameter learning algorithm. The EM algorithm can effectively compensate for the impact of missing accident records.

3. Railway Engineering Safety Risk Cause Analysis

This article analyzes 233 accident cases in China's railway engineering department from various channels between 2005 and 2013, and it explores the causes of risks from four aspects: risk cause types, responsible parties, specific manifestations of risk causes, and accident levels.

In the existing literature, the human–machine–environmental–management safety management system has been widely used in the safety management of various enterprises, and some railway safety fields have seen risk research using this approach. Accordingly, in terms of the types of risk causes, this article classifies and explores those from four aspects: personnel risk causes, facility and equipment risk causes, environmental risk causes, and management risk causes.

3.1. Personnel Risk Cause Analysis

In the railway engineering department, various personnel and scenarios are highly complex. By summarizing and analyzing the 233 collected real cases and accurately distinguishing the objects responsible for accident risk causes, those can be roughly divided into two types: internal railway personnel responsibility (A_1) and external railway personnel responsibility (A_2).

In terms of internal railway personnel responsibility, the specific causes of risks are mainly related to five aspects: internal railway personnel violate regulations during construction, loading and unloading, and maintenance processes (B_1); internal railway personnel have not followed standard specifications for construction, loading and unloading, and maintenance (B_2); lack of in-depth learning and understanding of railway-related rules and regulations and inadequate use of mechanical facilities and equipment by internal railway personnel (B_3); inadequate supervision and management by responsible internal railway persons or regulatory units (B_4); and the responsible internal person or regulatory unit of the railway failed to supervise and manage the system in accordance with standard specifications (B_5).

Finally, in terms of external railway personnel's responsibility, the specific causes of risks are mainly related to six aspects: irregular operations by external railway personnel during construction, loading and unloading, and maintenance processes (B_6); external railway personnel have not followed standard specifications for construction, loading and unloading, and maintenance (B_7); the learning and understanding of railway-related rules and regulations, mechanical facilities, and equipment usage methods by external railway personnel are not sufficient (B_8); inadequate supervision and management by responsible external persons or supervisory units of a railway (B_9); responsible external persons or supervisory units of railways fail to supervise and manage the system in accordance with standard specifications (B_{10}); and non-staff outside the railway affect the normal operation of a railway (B_{11}).

3.2. Facility and Equipment Risk Cause Analysis

Considering that the railway engineering department is specifically responsible for the maintenance and upkeep of railway lines and related equipment, including large and medium maintenance and upkeep of bridges, tunnels, culverts, roadbeds, rails, turnouts, sleepers, and tracks, the risk responsibility objects and specific manifestations of facilities and equipment are relatively complex. By analyzing the collected data on 233 real cases, the responsible risk-causing objects may be divided into six types: responsibility for rail line facilities and equipment (A_3); responsibility for locomotive and rolling stock facilities and equipment (A_4); responsibility for construction and other mechanical facilities and equipment (A_5); responsibility for bridge and tunnel facilities and equipment (A_6); responsibility for safety protection facilities and equipment (A_7); and, finally, responsibility for railway external facilities and equipment (A_8).

In terms of responsibility for rail line facilities and equipment, since the track line facilities and equipment mainly include gauge poles, roadbeds, steel rails, road studs,

sleepers, turnouts, and signal equipment, for the convenience of summarization and subsequent research, this study selects the status of the main facilities and equipment as a specific manifestation of their risk causes, which mainly includes ten aspects: the construction quality of track line facilities and equipment is poor, and the basic condition is poor (B₁₂); the track rod breaks off or falls off, the number is insufficient, the geometric size exceeds the limit, the height and the curve positive vector exceed the standard, and the curve track is not smooth and round (B₁₃); damage to the roadbed slope, insufficient quantity of ballast, insufficient width of ballast shoulder, and a lack of crushed and loose soil in the roadbed (B₁₄); the strength of the steel rail decreases, cracks appear, fractures occur, and the joint height is misaligned (B₁₅); road studs are floating, skewed, and out of joint, insufficient pressure and torque of fasteners, loosening and failure, weak nail holding force, and missing fasteners (B₁₆); sleeper decay and failure, empty suspension with exposed head, loose upright bolts, and damaged shoulder (B₁₇); loose connecting parts of the turnout, insufficient strength of the structural framework, and failure to pry along properly, which results in geometric dimensions exceeding the limit (B₁₈); rail line facilities and equipment exceeded their service period and were not replaced in a timely manner (B₁₉); due to the faults and other reasons, the track line facilities and equipment violate the limit, affecting the normal operation of trains (B₂₀); and failures of signal facilities and equipment on railway lines, which affect normal train operation (B₂₁).

In terms of responsibility for locomotive and rolling stock facilities and equipment, the specific causes of risks are mainly related to two aspects: the operational status of locomotive and vehicle facilities and equipment is unsatisfactory, and no safety hazards have been detected (B₂₂); and violation of limits by locomotive and vehicle facilities and equipment, which results in damage to track line facilities and equipment (B₂₃).

In addition, in terms of responsibility for construction and other mechanical facilities and equipment, the specific causes of risks are mainly trifold: mechanical facilities and equipment for construction are aging and in unsatisfactory operational conditions (B₂₄); construction and other mechanical facilities and equipment exceeded their service period and were not sent for inspection on time (B₂₅); and construction and other mechanical facilities and equipment encroach on limits, which affects the normal operation of trains (B₂₆).

Further, the responsibility for bridge and tunnel facilities and equipment is related to two types of specific causes of risk: low construction quality of bridge and tunnel facilities and equipment and poor foundation condition (B₂₇); and bridge and tunnel facilities and equipment are aging and are in poor operational conditions (B₂₈).

Beyond this, in terms of responsibility for safety protection facilities and equipment, the specific causes of risk include five aspects: failures in installing the line safety protection facilities and measures according to the specified regulatory requirements (B₂₉); failures in setting up personnel safety protection facilities and measures according to the regulations (B₃₀); safety protection equipment invading the line, which affects the normal operation of trains (B₃₁); aging of safety protection facilities and equipment and poor operational condition (B₃₂); and security monitoring facilities and equipment age and are in a poor operating condition (B₃₃).

Finally, the responsibility for railway external facilities and equipment relates to the following types of risk causes: railway external facilities and equipment invade a line and affect the normal operation of trains (B₃₄); and there is a low quality of construction of external facilities and equipment and a poor basic condition (B₃₅).

3.3. Environmental Risk Cause Analysis

Considering common meteorological factors and the natural disaster types and natural environment involved in railway engineering, and summarizing the data on 233 real cases, the objects responsible for risk causes can be roughly divided into three types: natural weather responsibility (A₉); natural disaster responsibility (A₁₀); and natural environmental responsibility (A₁₁).

In terms of the natural weather responsibility, the specific risk causes include thunder, heavy rain, and rainstorm (B₃₆); gale, tornado, and typhoon (B₃₇); heavy fog and haze (B₃₈); heavy snow and blizzard (B₃₉); and high temperature and freezing (B₄₀).

Next, the specific causes of risk related to the natural disaster responsibility include earthquake (B₄₁); floods, mudslides, and landslides (B₄₂); and collapse and falling rocks (B₄₃).

Finally, the main specific risk cause of the natural environmental responsibility is a poor natural environment foundation near the rail line (B₄₄).

3.4. Management Risk Cause Analysis

According to the settings of the railway engineering department in personnel management, facility and equipment management, environmental management, institutional and technical management, and safety risk management, and considering the collected data on 233 real cases, the responsibility for risk causes is divided into five types: personnel management responsibility (A₁₂), facility and equipment management responsibility (A₁₃), environmental management responsibility (A₁₄), institutional and technical management responsibility (A₁₅), and safety risk management responsibility (A₁₆).

First, in terms of the personnel management responsibility, the specific causes of risk are mainly manifested in three aspects: inadequate management, education, and training of internal railway staff (B₄₅); inadequate management, education, and training of external railway employees (B₄₆); and inadequate safety promotion for external railway personnel (B₄₇).

Second, the facility and equipment management responsibility has nine types of specific risk causes: inadequate daily troubleshooting and inspection of track lines and their ancillary facilities and equipment (B₄₈); incomplete or lagging maintenance, upkeep, and acceptance of rail line facilities and equipment, which affects the normal operation of a line (B₄₉); inadequate daily troubleshooting and inspection of locomotive and vehicle facilities and equipment (B₅₀); incomplete or delayed maintenance of locomotive and vehicle facilities and equipment, which affects the normal operation of locomotives and vehicles (B₅₁); inadequate inspection, repair, and maintenance of construction machinery and equipment (B₅₂); failure in supervision and management according to the safety supervision and management system standards for construction machinery, facilities, and equipment (B₅₃); incomplete or delayed maintenance of facilities and equipment, such as bridges and tunnels, which affects the normal operation of a line (B₅₄); inadequate daily inspection and inspection of safety protection facilities and equipment for railway lines (B₅₅); and, finally, inadequate inspection of railways' external facilities and equipment (B₅₆).

Third, in terms of the environmental management responsibility, the specific causes of risks are mostly manifested in two aspects: insufficient investigation and inspection of the natural environment near a rail line (B₅₇); and problems detected in the natural environment near a rail line are not promptly addressed and effective measures are not taken (B₅₈).

Fourth, for the institutional and technical management responsibility, the specific causes of risks relate to three aspects: relevant departments or personnel have not conscientiously implemented relevant management regulations or systems (B₅₉); inadequate professional technical management in the construction and usage of facilities and equipment, fault handling, and accident rescue (B₆₀); and mismatch between the construction organization plan, construction effect, and actual demand (B₆₁).

Finally, in terms of the safety risk management responsibility, the specific causes of risks include five aspects: the safety supervision and management system is not sound (B₆₂); failures in performing the supervision and management according to the safety supervision and management system standards (B₆₃); the safety risk prevention and management system is not sound (B₆₄); insufficient attention to potential security risks (B₆₅); and a failure in drawing inferences about similar safety risk issues (B₆₆).

3.5. Accident Level Indicators' Extraction

The accident level index is divided into four levels according to the suggestions given in Chapter 2 of the "Rules for Investigation and Handling of Railway Traffic Accidents in China" (Order No. 30 of the Ministry of Railways in 2007): particularly serious accidents (T_1), serious accidents (T_2), major accidents (T_3), and general accidents. The general accidents are further divided into general class A accidents (T_4), general class B accidents (T_5), general class C accidents (T_6), and general class D accidents (T_7).

4. Data Processing and Data Table Construction

This section processes the collected data on the selected real railway engineering accident cases and forms a data table for subsequent research.

4.1. Data Processing of Railway Engineering Safety Accident Cases

The railway engineering accident cases considered in this study were all in the form of text descriptions. Therefore, to construct a data table of railway engineering safety risk causes based on the various risk cause responsibility objects, specific manifestations of risk causes, and accident level indicators presented in Section 3, text mining methods were used to extract relevant information on each of the selected cases. A case extraction example is shown in Table 1.

Table 1. Example of raw data processing and information extraction for real railway engineering safety accident cases.

Accident Case Number and Description	Summary of Accident Case Content	Risk Causal Responsibility Object	Specific Manifestations of Risk Causes	Accident Level
44. Accident of collision between operator and freight train No. 13019 on the Jiaoliu Line	(1) The foreman violated the regulations of Guangzhou Railway An Dian [2007] No. 575 telegram during homework	(1) Personnel risk causes	(1) Inadequate supervision and management by responsible internal railway personnel and regulatory units (B_4)	General Class A accidents (T_4)
	(2) The work area was covered in heavy fog, with visibility of less than 10 m	(2) Environmental risk causes	(2) Heavy fog and haze (B_{38})	
	(3) Blindly acting without taking any special safety measures	(3) Facilities and equipment risk causes	(3) A failure in setting up personnel safety protection facilities and measures according to the specified regulations (B_{30})	
	(4) Commanding workers to work on the road under the violation of regulations has been an important cause of accidents	(4) Personnel risk causes	(4) Internal railway personnel violate regulations during the construction, loading and unloading, and maintenance processes (B_1)	
	(5) Inadequate implementation of rules and regulations	(5) Management risk causes	(5) Relevant departments or personnel have not conscientiously implemented relevant management regulations or systems (B_{59})	
	(6) The workshop paid no attention to the problems in the work area, which resulted in personal safety hazards that could not be eliminated in a timely manner	(6) Management risk causes	(6) A failure in the system of supervision and management according to the standards for the safety supervision and management system (B_{63})	

Table 1. Cont.

Accident Case Number and Description	Summary of Accident Case Content	Risk Causal Responsibility Object	Specific Manifestations of Risk Causes	Accident Level
54. The collision and derailment accident of the 40980 freight train on the Guangda Line	(1) The cutting slope in this area was steep, with sandstone on the slope surface and severely weathered and developed cracks. The weathered soil contained a large number of partially weathered large stones	(1) Environmental risk causes	(1) Poor natural environment foundation near a rail line (B ₄₄)	General Class B accidents (T ₅)
	(2) From October 27 to 29, there were a total of seven earthquakes with a magnitude of 2.0 or higher, with a maximum magnitude of 3.4	(2) Environmental risk causes	(2) Earthquake (B ₄₁)	
	(3) Since October 24, the region has experienced continuous rainfall of 178.8 mm, causing saturation of the slope soil with water content	(3) Environmental risk causes	(3) Thunder, heavy rain, and rainstorm (B ₃₆)	
	(4) During the construction of the Guangda Line, the standard level of equipment for lines, bridges, and roads was relatively low, which resulted in congenital structural deficiencies	(4) Facilities and equipment risk causes	(4) The construction quality of track line facilities and equipment is low, and the basic condition is poor (B ₁₂)	
	(5) The technical and safety management capabilities of Yunnan Railway Construction Engineering Company, which is an equipment maintenance agency, require further improvement	(5) Management risk causes	(5) Inadequate professional technical management in the construction and usage of facilities and equipment, fault handling, and accident rescue (B ₆₀)	
	(6) The Guangtong Engineering Section has promptly issued a notice requiring the maintenance unit to conduct inspections on the line, bridge, and road equipment, but the supervision and implementation of the equipment maintenance unit are not in place	(6) Personnel risk causes	(6) Inadequate supervision and management by responsible internal railway personnel and regulatory units (B ₄)	

4.2. Data Table Construction of Railway Engineering Safety Risk Causes

For the convenience of subsequent analysis, we represent the processed raw data in the form of a data table. The specific manifestation of risk causes (B₁–B₆₆) was used as a conditional attribute, and the accident level was regarded as a decision-making attribute. The attribute value corresponding to the specific manifestation of risk causes that appeared in each case was set to one, and the attribute value corresponding to the specific manifestation of risk causes that did not appear in any case was set to zero. The decision attribute values of each case corresponded to its accident level. The smaller the value of the decision attribute, the higher the accident level and the greater the consequences. The specific corresponding rules were as follows: the decision attribute value corresponding to a particularly serious accident (T₁) was one, the decision attribute value corresponding to a major accident (T₂) was two, the decision attribute value corresponding to a major accident (T₃) was three, and the decision attribute value corresponding to a general Class A accident (T₄) was four. Further, the decision attribute value corresponding to a Class B accident (T₅) was five, the decision attribute value corresponding to a Class C accident (T₆) was six, and

the decision attribute value corresponding to a Class D accident (T_7) was seven. The final data table of the railway engineering safety accidents is shown in Table 2.

Table 2. Data table of the railway engineering safety risk causes.

U	C														D
	B ₁	B ₂	B ₃	B ₄	B ₅	B ₆	...	B ₆₁	B ₆₂	B ₆₃	B ₆₄	B ₆₅	B ₆₆		
X ₁	0	0	0	0	0	0	...	0	0	0	1	1	0	2	
X ₂	1	1	0	1	0	0	...	0	0	0	0	0	0	2	
X ₃	0	1	0	0	0	0	...	1	0	0	0	0	0	2	
...	
X ₂₃₁	0	1	1	0	0	0	...	0	0	0	1	1	1	6	
X ₂₃₂	0	0	0	0	0	0	...	0	1	0	0	0	0	6	
X ₂₃₃	0	0	0	1	0	1	...	0	0	0	0	0	0	6	

5. Railway Engineering Safety Risk Key Causes' Analysis

As there were 66 specific performance indicators of railway engineering safety risks, directly establishing a Bayesian network would not have been conducive to diagnosis and inference as the network would have been too complex and contained too many nodes. In view of that, this study employed the multi-objective particle swarm optimization algorithm to explore specific manifestations of key causes of railway engineering safety risks. The results obtained by introducing the data table for railway engineering safety risk formed in Section 4.2 to the algorithm in Section 2.2 are shown in Figure 1, where F_1 represents the error value, and the smaller the error value, the higher the accuracy, while F_2 represents the number of key causes.

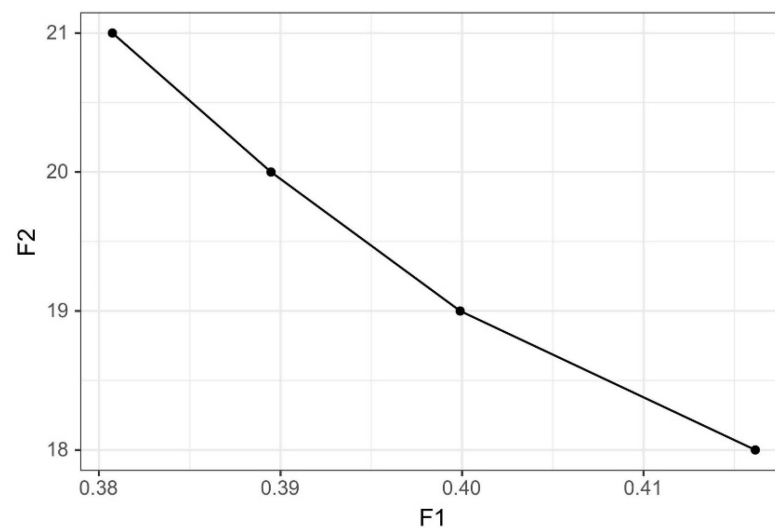


Figure 1. Results obtained by the algorithm using the data table.

When the maximum number of key indicators in the Pareto front of the key indicator combination F_2 was less than the specific number of indicators that caused the risk, there was a certain redundancy in the existing indicator system data. Through algorithmic iterations, this study selected the Pareto front corresponding to the combination of indicators with the highest accuracy (i.e., the smallest error margin) and the largest number of key indicators (the F_2 value). This combination of indicators was used as a key causal indicator. The 21 key causes of railway engineering safety risks are shown in Table 3.

Table 3. List of key causes of railway engineering safety risks.

Serial Number	Key Risk Factors
1	Internal railway personnel violated regulations during the construction, loading and unloading, and maintenance processes (B ₁)
2	Internal railway personnel did not follow standard specifications for construction, loading and unloading, and maintenance (B ₂)
3	Inadequate supervision and management by responsible external personnel and supervisory units of the railway (B ₉)
4	The track rod breaks off or falls off, the number is insufficient, the geometric size exceeds the limit, the height and the curve positive vector exceed the standard, and the curve track is not smooth and round (B ₁₃)
5	Sleeper decay and failure, empty suspension with exposed head, loose upright bolts, and damaged shoulder (B ₁₇)
6	Rail line facilities and equipment have exceeded their service period and have not been replaced in a timely manner (B ₁₉)
7	Violation of limits by locomotive and vehicle facilities and equipment, which results in damage to track line facilities and equipment (B ₂₃)
8	Construction and other mechanical facilities and equipment have exceeded their service period and have not been sent for inspection on time (B ₂₅)
9	Construction and other mechanical facilities and equipment encroach on limits, which affects the normal operation of trains (B ₂₆)
10	Failure to install the line safety protection facilities and measures according to the regulatory requirements (B ₂₉)
11	Security monitoring facilities and equipment are aging and in poor operating condition (B ₃₃)
12	Low quality of external facilities and equipment construction and poor basic condition (B ₃₅)
13	Gale, tornado, and typhoon (B ₃₇)
14	Earthquake (B ₄₁)
15	Floods, mudslides, and landslides (B ₄₂)
16	Inadequate safety promotion for external railway personnel (B ₄₇)
17	Inadequate daily troubleshooting and inspection of locomotive and vehicle facilities and equipment (B ₅₀)
18	Problems detected in the natural environment near a rail line were not promptly addressed, and effective measures were not taken (B ₅₈)
19	The construction organization plan, construction effect, and actual demand do not match (B ₆₁)
20	The safety supervision and management system is not sound (B ₆₂)
21	Failure to perform the supervision and management according to the standards for the safety supervision and management system (B ₆₃)

6. Risk Analysis of Railway Engineering Safety Based on Bayesian Network

6.1. Bayesian Network Structure Learning

This study used the K2 algorithm to learn the Bayesian network structure since it is effective for small datasets. The GeNIe 4.0 Academic software was used to import the railway engineering safety risk-critical cause data matrix, and the Greedy Think Thinning algorithm that comes with the software was employed to learn the Bayesian network structure from that matrix. Due to the exponential growth of the conditional probability of a node with the number of parent nodes, the maximum number of parent nodes was set to eight. In addition, by limiting the maximum number of parent nodes, we ensured that the algorithm would not search for the solution infinitely. Finally, one invalid node was removed, and the initial Bayesian network structure for risk-critical factors of railway engineering safety was output, as shown in Figure 2.

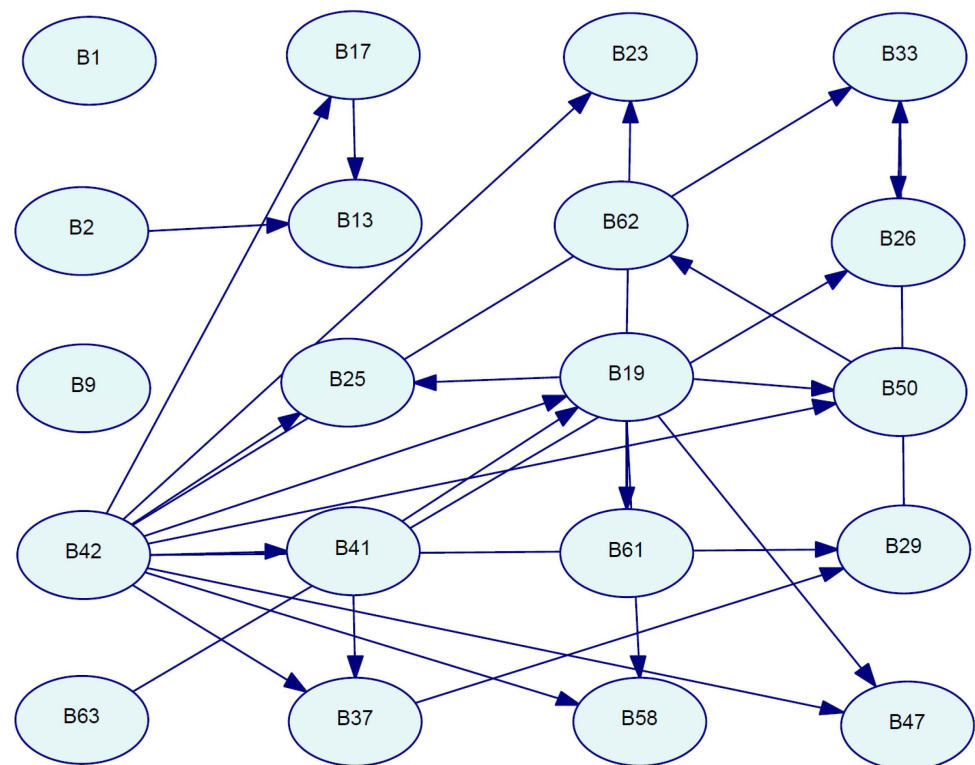


Figure 2. Bayesian network structure for the key causes of railway engineering safety risks.

6.2. Bayesian Network Parameter Learning

The data samples used in this study were sourced from real accident reports, but due to various different factors, the information contained in the accident reports was not necessarily uniform, and some information might have been missing. Therefore, the EM algorithm was employed to compensate for a small number of missing fault reports, improve the convergence speed for the small sample size, and reflect the probability values between nodes accurately. By using the parameter learning algorithm provided by the GeNIe 4.0 Academic software, each network node was matched with the corresponding column in the fault data matrix; the values of zero and one in the data matrix corresponded to States 0 and 1 of a node, respectively. State 0 indicated that the risk cause did not exist, and State 1 indicated that the risk cause existed. By setting the EM algorithm and selecting randomized initial parameters as the E-step starting parameters of the EM algorithm, the conditional probability table of the Bayesian network was obtained. The EM algorithm was selected because it has good applicability to data that might contain hidden variables. For example, when learning this type of data, using randomized parameters can effectively avoid the problem of local maximum values compared to using uniformly distributed parameters. The final generated probability map of risks' key causal conditions in railway engineering safety is shown in Figure 3.

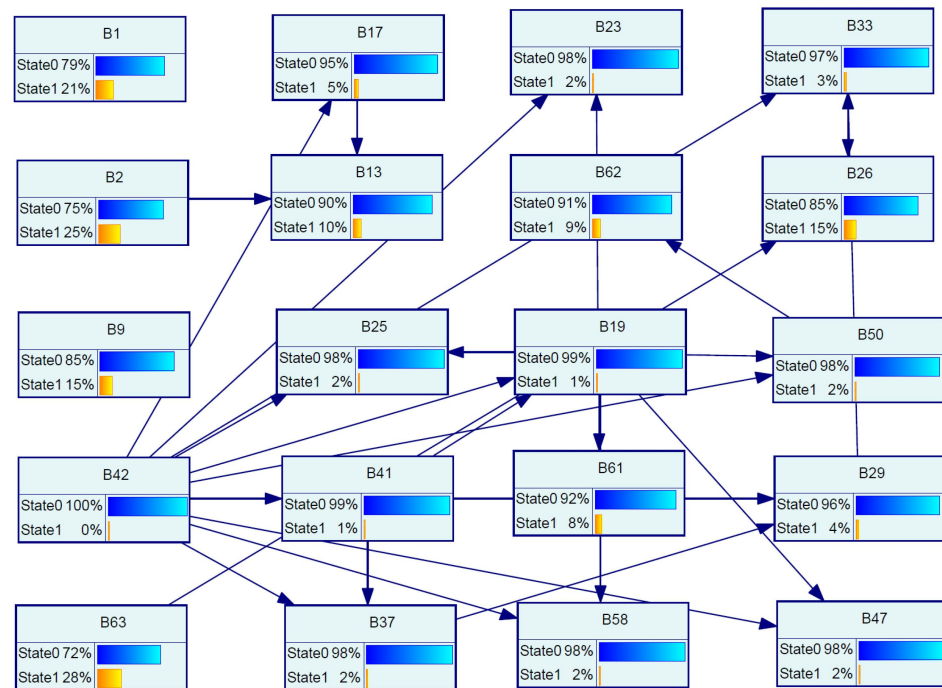


Figure 3. Conditional probability diagram for the key causes of railway engineering safety risks.

6.3. Occurrence Probability and Ranking of Risks' Key Causes

Based on the Bayesian network structure for the key causes of railway engineering safety risks established in Section 6.1, the edge probabilities of 20 risk-critical factors were calculated. The edge probability calculation results were extracted and sorted from high to low according to the probability of occurrence of risk-critical factors, as shown in Table 4.

Table 4. Statistical results for the probability of occurrence of risks' key causes.

Serial Number	Key Risk Factors	Probability of Occurrence
1	Failure to perform supervision and management according to the standards for the safety supervision and management system (B ₆₃)	0.278970
2	Internal railway personnel did not follow standard specifications for construction, loading and unloading, and maintenance (B ₂)	0.253219
3	Internal railway personnel violated regulations during the construction, loading and unloading, and maintenance processes (B ₁)	0.210300
4	Construction and other mechanical facilities and equipment encroach on limits, which affects the normal operation of trains (B ₂₆)	0.147996
5	Inadequate supervision and management by responsible external personnel and supervisory units of the railway (B ₉)	0.145923
6	The track rod breaks off or falls off, the number is insufficient, the geometric size exceeds the limit, the height and the curve positive vector exceed the standard, and the curve track is not smooth and round (B ₁₃)	0.095410
7	The safety supervision and management system is not sound (B ₆₂)	0.091093
8	The construction organization plan, construction effect, and actual demand do not match (B ₆₁)	0.078435
9	Sleeper decay and failure, empty suspension with exposed head, loose upright bolts, and damaged shoulder (B ₁₇)	0.046137
10	Failure to install the line safety protection facilities and measures according to the regulatory requirements (B ₂₉)	0.040342
11	Security monitoring facilities and equipment are aging and in poor operational condition (B ₃₃)	0.031529

Table 4. Cont.

Serial Number	Key Risk Factors	Probability of Occurrence
12	Inadequate safety promotion for external railway personnel (B ₄₇)	0.024505
13	Inadequate daily troubleshooting and inspection of locomotive and vehicle facilities and equipment (B ₅₀)	0.024505
14	Gale, tornado, and typhoon (B ₃₇)	0.019939
15	Violation of limits by locomotive and vehicle facilities and equipment, which results in damage to track line facilities and equipment (B ₂₃)	0.015934
16	Construction and other mechanical facilities and equipment have exceeded their service period but have not been sent for inspection on time (B ₂₅)	0.015934
17	Problems detected in the natural environment near a rail line were not promptly addressed, and effective measures were not taken (B ₅₈)	0.015934
18	Rail line facilities and equipment have exceeded their service period and have not been replaced in a timely manner (B ₁₉)	0.011356
19	Earthquake (B ₄₁)	0.009657
20	Floods, mudslides, and landslides (B ₄₂)	0.004292

According to the sorted results, the probability of failure in supervising and managing the system in accordance with the defined standards for a safety supervision and management system (27.90%) was the highest among all the risk causes. Thus, the railway engineering department should enhance safety prevention awareness, improve safety supervision and management, increase the supervision and management level, and avoid safety accidents caused by non-standard management.

In addition, it is necessary to pay additional attention to four aspects. First, internal railway personnel did not follow standard specifications for construction, loading and unloading, and maintenance (25.32%). Second, internal railway personnel violated the regulations during the construction, loading and unloading, and maintenance processes (21.03%). Third, the construction and other mechanical facilities and equipment encroached on limits, which affected the normal operation of trains (14.80%). Fourth, the supervision and management operations performed by the responsible external personnel and supervisory units of the railway were inadequate (14.59%). Therefore, more attention should be paid to the training process of internal and external personnel of the railway to enhance their awareness of safety precautions, strengthen the management of professional skills, and show them how to perform standardized operations according to the relevant regulations, thus preventing violations of personnel operations and poor facilities and equipment.

6.4. Posterior Probability Analysis of Risks' Key Causes

By using the Bayesian network diagnostic reasoning function of the GeNIe 4.0 Academic software, the posterior probability of the parent node was calculated for various risk-critical causes. The calculation results are shown in Table 5.

The calculation results obtained by reverse reasoning provided information on the probability of the specific manifestation of each risk cause occurring at its parent node risk cause. Based on this information, the most likely cause of the specific manifestation of the risk cause was identified. This will enable the railway engineering department to learn from lessons, draw analogies, and achieve precise management.

When the gauge rod was broken, falling off, or insufficient in quantity, the geometric size was beyond the limit, the height and curve vector were also beyond the limit, or the curve track was not smooth (B₁₃), then, consequently, the probability of the railway internal personnel not following the standard specifications for construction, loading and unloading vehicles, and maintenance (B₂) was high, reaching 73.41%. Based on the results, the status of facilities and equipment was closely related to whether internal personnel operated

in a standardized or non-standardized manner. Therefore, it is necessary to improve the training process of internal railway personnel, focus on inspecting various facilities and equipment in accordance with the standard specifications, and prevent the occurrence of safety accidents caused by non-standard operations, which can lead to damage to various facilities and equipment.

Table 5. Posterior probability calculation results of the parent node of the risks' key causes.

Specific Manifestations of Risks' Causes	Parent Node	Posterior Probability	Specific Manifestations of Risks' Causes	Parent Node	Posterior Probability
B ₁₃	B ₂	0.734051	B ₃₃	B ₂₉	0.296744
	B ₁₇	0.142758		B ₄₂	0.045197
B ₁₇	B ₄₂	0.069767	B ₃₇	B ₄₁	0.098655
B ₁₉	B ₄₁	0.173228		B ₄₂	0.067265
	B ₄₂	0.118110	B ₄₁	B ₄₂	0.111111
B ₂₃	B ₁₉	0.146832	B ₄₇	B ₁₉	0.095475
	B ₄₂	0.088380		B ₄₂	0.057467
B ₂₅	B ₁₉	0.146832	B ₅₀	B ₁₉	0.095475
	B ₄₂	0.088380		B ₄₂	0.057467
B ₂₆	B ₃₃	0.102072	B ₅₈	B ₁₉	0.146832
	B ₆₃	0.560470		B ₄₂	0.088380
B ₂₉	B ₃₇	0.131874	B ₆₁	B ₁₉	0.072388
	B ₄₂	0.034908	B ₆₂	B ₅₀	0.107606

Further, when construction and other mechanical facilities and equipment violated the limit, which affected the normal operation of trains (B₂₆), the probability of failure to perform the supervision and management process in accordance with the safety supervision and management system standards (B₆₃) was 56.05%. Thus, the status of facilities and equipment was also affected by management. Therefore, the railway engineering safety supervision and management system standards should be strictly implemented, and possible problems in facilities and equipment should be carefully investigated to prevent faults that might affect the normal operation of trains.

Finally, when some specific manifestations of risk causes occurred, there was a certain probability that the rail line facilities and equipment had exceeded their service period and were not replaced in a timely manner (B₁₉), and earthquakes (B₄₁), floods, mudslides, and landslides (B₄₂) could occur. Therefore, railway engineering departments should pay special attention to the risk prevention process under the conditions of natural disasters and adverse weather.

6.5. Transmission Mechanism and Probability Analysis of Risks' Key Causes

According to the Bayesian network structure of the railway engineering safety risk-critical causes designed in Section 6.1, this study defined B₂ and B₄₂ as the initial risk-critical causes with State 1, indicating the occurrence of risk-critical causes. Their sub-nodes were studied separately, the transmission path of risk key causes was determined, the probability of the occurrence of the final risk key cause was calculated, and the impact of the occurrence of the initial risk key cause on the final risk key cause was evaluated and quantified. The results of the positive transmission path and probability of occurrence are shown in Table 6.

Table 6. Positive transmission path and probability of occurrence.

Key Causes of Initial Risk	Risk-Critical Causal Transmission Pathways	Final Risk's Key Cause	Probability of Occurrence
B ₂	B ₂ → B ₁₃	B ₁₃	0.276582
	B ₄₂ → B ₁₇ → B ₁₃	B ₁₃	0.242850
	B ₄₂ → B ₂₃	B ₂₃	0.328125
	B ₄₂ → B ₄₁ → B ₁₉ → B ₂₃		
	B ₄₂ → B ₂₅	B ₂₅	0.328125
	B ₄₂ → B ₄₁ → B ₁₉ → B ₂₅		
	B ₄₂ → B ₃₃ → B ₂₆	B ₂₆	0.250736
	B ₄₂ → B ₃₃	B ₃₃	0.332031
	B ₄₂ → B ₄₁ → B ₃₇ → B ₂₉ → B ₃₃		
	B ₄₂ → B ₄₇	B ₄₇	0.328125
	B ₄₂ → B ₄₁ → B ₁₉ → B ₄₇		
	B ₄₂ → B ₅₈	B ₅₈	0.328125
	B ₄₂ → B ₄₁ → B ₁₉ → B ₅₈		
B ₄₂ → B ₄₁ → B ₁₉ → B ₆₁	B ₆₁	0.206845	
B ₄₂ → B ₅₀ → B ₆₂	B ₆₂	0.187240	
B ₄₂ → B ₄₁ → B ₁₉ → B ₅₀ → B ₆₂			
B ₆₃	B ₆₃ → B ₂₆	B ₂₆	0.297333

When natural disasters, such as floods, mudslides, and landslides (B₄₂), occurred, the probability of the security monitoring facilities and equipment being aged and in a poor operating condition (B₃₃) was the highest, having a value of 33.20%.

In addition, the following factors also had a significant impact: the violation of limits by locomotive and vehicle facilities and equipment resulted in damage to track line facilities and equipment (B₂₃); construction and other mechanical facilities, and equipment that exceeded their service period but were not sent for inspection on time (B₂₅); there was inadequate safety publicity for external railway personnel (B₄₇); and there was a failure to take timely and effective measures for problems detected in the natural environment near the track line (B₅₈).

Further, the violation of limits on locomotive and vehicle facilities and equipment, which could cause damage to track line facilities and equipment (B₂₃), overdue servicing of construction and other mechanical facilities and equipment, or failure to submit for inspection on time (B₂₅), inadequate safety promotion for external railway personnel (B₄₇), and problems detected in the natural environment near the rail line that were not promptly addressed or met with effective measures (B₅₈) each had a significant impact.

Finally, when the internal railway personnel did not follow standard specifications for construction, loading and unloading, and maintenance (B₂), floods, mudslides, landslides (B₄₂), and failure to supervise and manage the system according to safety supervision and management system standards (B₆₃), this caused multiple facility and equipment risk factors and management risk factors. In particular, natural disasters, such as floods, mudslides, and landslides, had a strong impact on facilities, equipment, the environment, and management. In recent years, several major railway safety accidents in China have occurred under natural disaster conditions, such as floods, mudslides, and landslides. Therefore, railway engineering departments should put additional efforts into preventing various safety risks and managing facilities and equipment during the rainy season and flood season.

7. Discussion

Based on a review of existing studies in the field of railway safety, it seems that most scholars have adopted a single method for analyzing safety risks under a single scenario or indicator, without considering the current situation of various risk correlation factors in the railway system, or considering optimization under multi-objectives and the coordination of multiple different dimensions of demand. Moreover, there have been few studies on the optimization of the index combination itself. And some methods used in subjective evaluations cannot judge the logical relationship between indicators and outputs.

Liu Yang [12] proposed and improved relevant text feature extraction methods, established an expert database, and considered text correlation as a means to determine key accident cause factors. They also identified the accident chain in the accident-cause-related network, uncovered the internal relationships between the factors, and found the key accident chain and cause factors. Zong Zhou [13] applied and improved the Domino model in railway operation safety risk identification, calculated the similarity of a railway operation risk network based on Vague set theory, explored and visualized association rules for the causes of railway operation risks in the UK, and determined the scope of railway operation risks in the UK from 2010 to 2019. By improving the FRAM model, Yifei Xu [14] proposed an accident cause analysis method for the railway transport system based on the ISM-IFRAM model, analyzed the causes of China's "7.23" Yongwen Line (especially major railway traffic accidents), and offered relevant improvement suggestions on a resource allocation basis for the railway transport system based on the analysis results. Shanshan Wang [15] quantitatively analyzed nodes and community risks in the network, built a railway accident-cause-related network model, identified and analyzed the accident cause chain in the network, and analyzed the railway accident cause network under dynamic weights. Xinsheng Shan [16] built an HFACS model suitable for safe production in power supply scheduling based on the statistical data of power supply scheduling safety information of the China Railway Guangzhou Bureau in the past three years, and designed a method for expert evaluation and fuzzy comprehensive calculation to be used to conduct a quantitative analysis of the membership degrees of factors at each level of the model. Weijian Cui [17], utilizing part of China's railway network security assessment data from 2019 to 2020, adopted a railway network security risk analysis method based on an improved Apriori algorithm, and they conducted application research on the whole process of risk analysis such as data collection, preprocessing, mining, and calculation. Yuchen Luo [18] established the HFACS model framework for improving the personal safety accidents of Chinese railway public servants, classified and elaborated the factors that caused the personal safety risks of public servants, analyzed the factors that caused the risks, and evaluated the accident model data using the set pair analysis theory model. Yu Gao [19] applied the constructed LTE-R risk assessment system and ranking model to the ranking field of LTE-R system risk factors, and they verified the effectiveness of the model in the field of rail transit by testing the sensitivity of the MAIRCA algorithm under different weights and comparing the ranking results of various algorithms. Based on the SPA-TOPSIS model, Xiaoxu Shen [20] proposed a multi-perspective evaluation method for railway transport safety, conducted case evaluations of railway transport safety from regional perspectives and line-specific perspectives, analyzed and calculated the values and fluctuations of evaluation indicators under different perspectives, analyzed the evaluation results, and gave corresponding decision-making suggestions. Shi Zhang [21] used the fuzzy interpretation structure model to build the fuzzy ISM model of the accident risks of the train service system, and they used the D-S evidence theory and Bayesian network to build an accident prediction model of the railway train service system, and combined with the actual train service system data, they conducted accident prediction and analysis for the train service system. In the above studies, most of the analysis methods of railway safety risks used by scholars were relatively simple. Although some scholars combined multiple analysis methods and put forward their own innovations in the research, they did not analyze the correlations between safety risks, resulting in certain deficiencies.

With the development of science and technology, new technologies and new equipment are constantly introduced into railway operation and management. Nonetheless, the railway safety situation will be related to various factors such as the uneven quality of workers, the limited management level, and sudden and severe natural disasters. How to study the existing railway accident data, excavate the internal mechanisms, potential laws, leading factors, and transmission mechanisms of the accidents, and how to apply the subsequent discoveries and experience to actual operation management, so as to improve the overall safety level of the railway operation management system, has been the content of this paper.

Firstly, the method of multi-objective index screening was used in this paper to make index screening more suitable for practical applications. Secondly, a Bayesian network was used, the structure of the Bayesian network was optimized, and the causal logic chain of the accident result was obtained. Then, based on the research results of this paper, we analyzed the impacts of personnel risk, facility and equipment risk, environmental risk, and management risk, respectively, and those risks combined. More precisely, we analyzed the causes of accidents and the relationships between them, and, in doing so, we made up for the lack of previous research on railway safety risks in the study of risk transmission mechanisms.

8. Conclusions

This study has analyzed real cases of railway engineering accidents in China, using the multi-objective particle swarm optimization algorithm to determine the key risk factors. Bayesian networks have been used to analyze the transmission mechanisms and probabilities of key risk factors and obtain the causal logic chain that produces accident results. Finally, it has been shown that combining the indicators and Bayesian networks can improve the accuracy of risk prediction and provide more accurate results than using existing research, and it can fill the gap in research on railway safety risks in risk transmission mechanisms.

Based on the results, the following conclusions are drawn.

From the perspective of personnel risk causes, the probabilities of internal railway personnel not following standard specifications for construction and violating regulations were as high as 25.3% and 21.0%, respectively. It is necessary to ensure effective training is given for various types of railway personnel and that we enhance their safety awareness, focus on the various operations of internal railway personnel, and avoid the occurrence of facility equipment failures and safety accidents caused by illegal and non-standard operations.

In terms of facilities and equipment risk causes, failures of facilities and equipment are closely related to the non-standard operation and supervision/management of personnel, and the probability of their occurrence reached 73.41%. We also found that these risk causes can easily serve as an intermediate node in the transmission path of key accident causes. Therefore, it is necessary to improve the operation and supervision/management of various railway personnel. Moreover, additional attention should also be paid to the usage status of facilities and equipment to avoid the problems of overdue services or lack of timely replacement.

In addition, regarding environmental risk causes, natural disasters, such as floods, mudslides, and landslides, are prone to causing nine chain reactions, including equipment failures in railway lines and other facilities and an inadequate safety supervision and management system, which can easily lead to safety accidents. Therefore, railway-related departments should improve relevant safety management measures according to the actual situation to reduce the probability of occurrence and transmission of related risks during severe natural disasters.

Further, in terms of management risk causes, note that the probability of failure to perform supervision and management according to the standards for the safety supervision and management system was found to be 27.90%, and this was shown to have a significant impact on the other risk causes. Thus, railway-related departments should give high importance to the management of personnel, facilities, equipment, and the environ-

ment to improve relevant safety supervision and management systems and promote the management to play its due role.

The limitation of this study is that due to the influence of relevant policies, actual accident cases have relatively long occurrence times, which reduces the reference value for the current railway engineering department. Moreover, in the process of extracting risk indicators, due to an insufficient understanding of the railway engineering department, the classification of individual risks might not have been precise enough, which could have impacted the screening of key causes. To overcome this, future research could refine risk classification, enhance the correlation between the risk causation responsibility objects and specific manifestations of risk causation, and promote more accurate transmission diagnosis and probability analysis, stronger causal logic, and practicality.

Author Contributions: Conceptualization, T.Z.; methodology, T.Z.; software, X.L. (Xuewei Li) and X.L. (Xueyan Li); validation, X.L. (Xuewei Li) and X.L. (Xueyan Li); formal analysis, T.Z.; investigation, T.Z.; resources, X.L. (Xuewei Li) and X.L. (Xueyan Li); data curation, T.Z.; writing—original draft preparation, T.Z.; writing—review and editing, X.L. (Xuewei Li) and X.L. (Xueyan Li); supervision, X.L. (Xuewei Li) and X.L. (Xueyan Li); project administration, X.L. (Xuewei Li) and X.L. (Xueyan Li); funding acquisition, X.L. (Xuewei Li) and X.L. (Xueyan Li) All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No. 72103019); Beijing Municipal Education Commission Science and Technology Plan General Project (Grant No. KM202211417009); Science and Technology Research and Development Plan of China National Railway Group Corporation Limited (Grant No. K2019Z006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available in the books “Cases of Railway Running Accidents 2006.1~2007.8”, “Cases of Railway Traffic Accidents 2007.9~2008.12”, and “Cases of Railway Traffic Accidents 2009” published by the China Railway Publishing House.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Zhang, Y.K. Station Safety of High Speed Railway Research and Implementation of Management Evaluation Model. Master’s Thesis, Southwest Jiaotong University, Chengdu, China, 6 June 2018.
- Li, C.R.; Zhao, M.; Yin, P.P.; Liu, M.C.; Ma, L.T. Safety evaluation of high-speed railway passenger station. *J. China Saf. Sci. J.* **2019**, *29*, 149–153.
- Yao, H.L. Risk assessment of railway critical information infrastructure based on AHP and BP models. *J. Netinfo Secur.* **2020**, *S2*, 36–39.
- Zhang, Y.R. Research on the Safety Risk Management of the Operation of Power and Electricity Combined with Skylights in the Common Speed Railway. Master’s Thesis, Lanzhou Jiaotong University, Lanzhou, China, 5 December 2021.
- Sun, X.Y. Research on Safety Risk Management of Construction and Singal Cooperation with Operation in Ordinary Speed Railway. Master’s Thesis, Shandong Jianzhu University, Jinan, China, 20 June 2023.
- Kurz, S.L.; Milius, B. Negligible Risk for European Railway Risk Assessments. In Proceedings of the 2nd IET International Conference on System Safety 2007, London, UK, 22–24 October 2007.
- Reinach, S.; Viale, A. Application of a human error framework to conduct train accident/incident investigations. *J. Accid. Anal. Prev.* **2006**, *38*, 396–406. [[CrossRef](#)] [[PubMed](#)]
- Zhan, Q.; Zheng, W.; Zhao, B. A hybrid human and organizational analysis method for railway accidents based on HFACS-Railway Accidents (HFACS-RAAs). *J. Saf. Sci.* **2017**, *91*, 232–250. [[CrossRef](#)]
- Lower, M.; Magott, J.; Skorupski, J. A System-Theoretic Accident Model and Process with Human Factors Analysis and Classification System taxonomy. *J. Saf. Sci.* **2018**, *110*, 393–410. [[CrossRef](#)]
- China State Railway Group Co., Ltd. *China Railway Yearbook 2021*; China Railway Publishing House: Beijing, China, 2022.
- Zhou, J. Study on Accident Causation and Safety Risk Analysis in Railway Transportation. Ph.D. Thesis, Beijing Jiaotong University, Beijing, China, 1 June 2018.
- Yang, L. Causes Analysis of Rail Transit Accidents and Risk Research Based on Text Data. Ph.D. Thesis, Beijing Jiaotong University, Beijing, China, 31 May 2021.

13. Zhou, Z. Cause Analysis and Risk Assessment of Railway Operation Accidents. Master's Thesis, Wuhan Polytechnic University, Wuhan, China, 10 June 2021.
14. Xu, Y.F. Research on Accident Causes Analysis Method of Railway Transportation System Based on Improved FRAM. Master's Thesis, Southwest Jiaotong University, Chengdu, China, 1 June 2021.
15. Wang, S.S. Research on the Causation of Railway Accident and Risk Based on Complex Network. Master's Thesis, Beijing Jiaotong University, Beijing, China, 16 March 2021.
16. Shan, X.S. Research on Safety Management of Guangzhou Railway (Group) Corporation Power Supply Dispatching Based on Accident Causation Theory. Master's Thesis, China Academy of Railway Sciences, Beijing, China, 30 June 2023.
17. Cui, W.J. Research on Railway Network Security Risk Analysis Method Based on Improved Apriori Algorithm. Master's Thesis, China Academy of Railway Sciences, Beijing, China, 30 June 2021.
18. Luo, Y.C. Personal Safety Risk Assessment of Railway Workers Based on Improved HFACS-SPA. Master's Thesis, Dalian Jiaotong University, Dalian, China, 11 June 2021.
19. Gao, Y. Research on LTE-R System Risk Assessment Method Based on Improved Evidence Theory. Master's Thesis, Lanzhou Jiaotong University, Lanzhou, China, 12 June 2023.
20. Shen, X.X. Multi-Perspective Assessment of Railway Transportation Safety Based on Accident Attribute Characteristics. Master's Thesis, Southwest Jiaotong University, Chengdu, China, 31 May 2022.
21. Zhang, S. Research on Accident Probability Prediction of Railway Train Service System Based on Bayesian Network Model. Master's Thesis, China Academy of Railway Sciences, Beijing, China, 30 June 2022.
22. Li, J.; Pan, S.X.; Li, X.Y.; Jia, L.J.; Zhao, Y.Z. Screening Critical Patients with Optimized Classifier Based on Multi Objective Quantum. *J. Data Anal. Knowl. Discov.* **2019**, *3*, 101–112.
23. Heckerman, D.; Geiger, D.; Chickering, D.M. Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *J. Tech. Rep.* **1995**, *20*, 197–243. [[CrossRef](#)]
24. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data Via the EM algorithm. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1977**, *39*, 1–38. [[CrossRef](#)]
25. Cooper, G.F.; Herskovits, E. A Bayesian Method for the Induction of Probabilistic Networks from Data. *J. Mach. Learn.* **1992**, *9*, 309–347. [[CrossRef](#)]
26. Wang, S.C.; Liu, X.H.; Zhang, P.Q. Learning Bayesian Networks Structure from Small Data Set in Operational Risk Analysis. *J. J. Syst. Manag.* **2008**, *17*, 448–454.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.