International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 4819-4830

https://doi.org/10.70517/ijhsa463412

Safety Risk Assessment Model of Dual Prevention Mechanism Based on Big Data Analysis Technology and Optimization Scheme for Hidden Trouble Screening

Fang Yan[1,*]
[1] Training and Education Department, Hunan Vocational Institute of Safety Technology, Changsha, Hunan, 410151, China
Corresponding authors: (e-mail: yanfang20242024@163.com).

Abstract Because the theory of dual prevention mechanism is put forward for a relatively short period of time and lacks the support of corresponding regulations and standards, it is still difficult to see the effectiveness of real and effective operation of dual prevention mechanism to prevent accidents in enterprises, and there are some professional and technical obstacles that have not been overcome. This paper shifts the focus of the establishment of the dual prevention mechanism to focus on safety risk control and develop corresponding emergency measures. The enterprise safety index system is established from the two parts of risk grading and control and hidden danger investigation and management. According to the definition of Bayesian formula network, determine the conditional probability of Bayesian network, and construct the safety risk assessment model based on Bayesian network. The risk reachable probability of the model constructed in this paper indicates that after a security event occurs in the S9 indicator, the reachable probability of each indicator in the experimental network shows an upward trend, and the overall security risk is rising, and at this time, the a posteriori reachable probability of S1, S4, and S6 is significantly higher than that of the other indicators, which is 0.85, 0.78, and 0.76, respectively, and it is very likely that there is a security risk in these three indicators. Comparing the a priori reachable probabilities of the indicator nodes given by the three methods, the a posteriori reachable probabilities of the indicator nodes of this paper's method for S5, S7, and S8 are 0.46, 0.32, and 0.14, respectively, and there is no underestimation of the real security risk.

Index Terms dual prevention mechanism, safety risk control, hidden danger investigation, Bayesian network, risk reachable probability

I. Introduction

With the continuous development of safety science and technology, although China's safety production state currently maintains a relatively stable and benign development, the total number of accidents is still maintained at a relatively high level [1]. At present, the main reason for the occurrence of accidents is that there is no effective double prevention mechanism. The way to effectively solve the various problems in safety management is to build a dual prevention mechanism of risk classification and control and hidden danger investigation and management [2], [3]. Since the 2015 Tianjin Port “8.12” particularly significant explosion accident, the state has begun to re-recognize and accurately position the current mode of supervision and management of production safety at the policy level, and its level of management of emergency preventive measures in various types of production safety accidents occurring in enterprises [4]-[6].

In view of the frequent occurrence of some safety production accidents at home and abroad, each enterprise pays more and more attention to the work of safety production, in order to improve the status quo of safety production, each enterprise, after summarizing and analyzing many years of management experience, constantly reforms and innovates the safety management system, so as to gradually improve the situation of safety production in the enterprise [7]. Christensen, I believes that the focus of safety management should be on the construction of a safety culture, and companies should pay attention to the personal safety of employees and encourage them to reduce unsafe behaviors [8]. The theory provides a new way of thinking for safety management, which is to take employees as the focus of safety management and establish a more efficient and systematic safety management system. Choudhry, R. M believes that it is necessary to formulate the enterprise safety observation table and behavioral observation plan, to supervise the production behavior of the enterprise, and when there is a hazardous operation, it is necessary to make behavioral corrections in a timely manner, to avoid the risk of expanding, and ultimately to gradually form a complete safety management system [9].

At present, the rapid development of information technology has triggered the rapid growth of data, which has become a key technical means that all industries and even countries are competing for research and development [10]. The emergence of big data technology, completely change the traditional movement, rough safety management means and methods, the application of big data can timely and accurately find hidden dangers, and greatly improve the ability of hidden danger investigation [11], [12]. Xie, K et al. utilized big data technologies to advance the potential for pedestrian safety hazard identification, including investigating factors contributing to pedestrian crashes and identifying high-risk locations in the city [13]. Huang, L et al. proposed a new conceptual model for accident investigation based on safety big data (SRBD), i.e., a hierarchical pyramid structure of safety-related big data, safety information, safety regulations, and safety knowledge using big data technology [14]. Latif, S et al. used big data technology to analyze the safety risks of human development so as to identify the root causes of accidents, develop targeted prevention programs, improve source governance, and reduce the incidence of safety accidents [15]. The use of big data technology is needed to promote the in-depth development of the dual prevention mechanism [16]. Because the dual prevention mechanism is put forward for a relatively short period of time, the research of scholars on the relationship between big data technology and the dual prevention mechanism is relatively small.

This paper establishes the dual prevention mechanism of enterprises from two parts of risk classification and control and hidden danger investigation and management, and establishes the intelligent enterprise safety index system after clarifying the construction procedures and specific contents. Quantitative operation of the indicators in the system, the use of hierarchical analysis to determine the relevant weights, and the construction of the indicator judgment matrix, through the CI value to control the indicator consistency test error. Define the Bayesian network, construct the Bayesian security risk assessment model, and determine the Bayesian network conditional probability. Design the Bayesian network risk assessment model simulation experiment to analyze the Bayesian network model's assessment of security risk under the dual prevention mechanism. At the same time, according to the information provided by the dual prevention mechanism, construct the dynamic Bayesian attack graph, and assess the current security risk reachable probability.

II. Dual defense mechanism construction

II. A.Dual Defense Mechanism Construction

The construction of dual prevention mechanism should have the awareness of “early warning” and “prevention and control”. Different from the previous focus on hidden danger investigation and management, now it is more important to focus on safety risk management and control [17]. For modern enterprises to establish scientific risk classification standards as soon as possible, risk assessment of the production process, the development of corresponding emergency measures. It is necessary to clarify the risk prevention and control responsibilities of each position and each person. Based on the results of risk classification, carry out targeted hidden danger investigation, so that the enterprise can effectively control operational risks and manage potential safety hazards.

Enterprise dual prevention mechanism mainly contains risk grading control and hidden danger investigation and management of two parts, in this regard, the first need to clarify the construction of the program and specific content, and on this basis to establish an intelligent enterprise safety indicator system. The construction process of the dual prevention mechanism is shown in Figure 1.

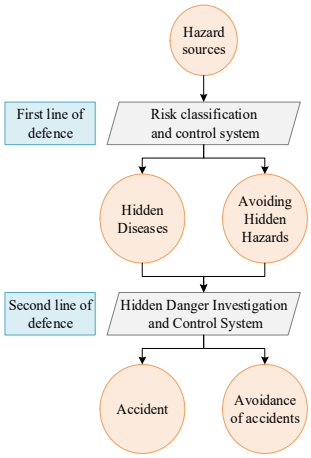


Figure 1: Dual prevention mechanism

II. A. 1) Risk classification and control

Risk classification and control refers to the classification of risks into several different levels according to the difficulty of controlling them. The general principle of risk classification is: according to the different degree of the consequences of risk to adjust the level of risk control, and in the enterprise corresponding to the management at all levels to implement measures to monitor and control the risk, the risk classification and control program mainly contains four major parts of the content.

(1) Identification of risk points

Before identifying the risk points, the level of the risk points should be determined and classified according to the risk consequences of the risk level classification guidelines. It should be noted that, in the process of identifying the risk points, it can also be combined with the layout of the enterprise production plant and other content to be divided, in order to establish a relatively independent risk identification and control unit.

(2) Hazard identification

The common methods for identifying the sources of danger include questionnaire interview method, research method, record retrieval method and information organization method. When identifying the scope of hazardous sources, the main contents include production activities, personnel activities, equipment and facilities, raw materials and products.

(3) Risk assessment and grading

For intelligent enterprises, the commonly used risk assessment and grading methods include hazard assessment and analysis of operating conditions, risk matrix analysis and hierarchical analysis.

The quantitative risk model of typical accidents is designed based on the risk function, and the risk comprehensive evaluation method is used to determine the risk value of accidents. In this paper, the risk quantitative model of typical accidents is constructed as shown in the following equation [18]:

$$R = \sum_{i=1}^{n} w_i D_i \tag{1}$$

where R is the total evaluation score, w_i is the weight of evaluation indicator i, D_i is the score of evaluation indicator i, and n is the number of evaluation indicators.

II. A. 2) Mechanisms for investigating and managing hidden dangers

For the four types of human-caused hazards, physical-caused hazards, environmental hazards, and management hazards, different hazard evaluation methods are used.

(1) For the human-caused hazards category evaluation, the LEC method is used.

The LEC method consists of three factors, where L denotes the likelihood of an accident, E denotes the frequency of human exposure to the hazardous environment, and C denotes the possible consequences of an accident. Based on the product of the three, the hidden hazards are graded.

(2) For the assessment of physical hazards category, the rubric method is used.

Through the literature review and analysis, the researcher found that there are many methods for risk evaluation, mainly including the assessment point method, the safety checklist method and the comprehensive evaluation method of hazard factors. Combined with the actual situation, the assessment of physical hazards should be prioritized to use the assessment point method.

Point-of-care method is generally applicable to complete and complex systems, which mainly considers the degree of danger of hidden hazards from five dimensions, i.e., the degree of consequences, the degree of system impact, the probability of occurrence, the difficulty of preventing failures, and whether it is a newly-designed system or not, of which the dimensions are divided into precise ones. In addition to this, the rating method is easy to operate and the comprehensive degree of danger is determined by the product (see equation 2):

$$Cs = \Pi Ci\left(i = 1,2,3,4,5\right) \tag{2}$$

where: Cs is the total number of points assessed, $0 < Cs < 10$, and Ci is the number of points assessed each, $0 < Ci < 10$.

II. B. Construction of Risk Grading Indicator System

II. B. 1) Design of risk evaluation indicators

The theoretical basis for the risk grading method of the dual prevention mechanism of enterprise safety production is the RBS theory, i.e., risk-based supervision theory.The primary theoretical basis of RBS is the safety degree function. Safety itself is an abstract qualitative concept, but in the research process, safety needs to be

quantitatively studied. Generally speaking, the concept of quantitatively describing safety is “safety” or “safety degree”, with the following mathematical expression:

$$S = F(R) = 1 - R(P, L, S) \tag{3}$$

where R is the systematic risk value, P is the probability of accident, L is the severity of accident, and S is the sensitivity of accident hazard.

Risk indicator system is a complete system composed of multiple risk indicators, in which the indicators, as the key factors for risk evaluation, should follow the principles of objectivity, scientificity, systematicity and operability in their selection. Based on the principle of indicator selection, this paper adopts the hierarchical analysis method to organize and summarize the risk grading management indicators after the field investigation of intelligent enterprises, and classifies them, as shown in Table 1.

Table 1: Risk classification indicators for intelligent factories

Target layer	Criterion layer	Scheme layer
Risk management	Resources	Personnel (S1)
		Equipment (S2)
		Material (S3)
	Technology	Instrument (S4)
		Process (S5)
		System (S6)
	Management	Equipment maintenance (S7)
		Personnel allocation (S8)
		Working hours (S9)
		Training learning (S10)
	Environment	Noise factor (S11)
		Plant layout (S12)

The traditional risk assessment usually includes two dimensions, accident severity and accident likelihood, and the dimension of accident sensitivity has been added after the demonstration and research of domestic experts and scholars. These three dimensions interact with each other and together determine the overall risk of an accident. Risk and safety are opposites and complementary relationships, the sum of safety and risk constitutes the overall state of the research object, that is, $S + R = 1$.

Accident is the product of safety risk” is the basic axiom of safety and the second theoretical basis of RBS. Since the existence of safety risk is the premise leading to safety accidents, the first issue to control the occurrence of safety accidents is to identify the risk, evaluate the risk, and finally control the risk.

On the basis of the safety function, RBS theory involves the following four basic functions:

Risk function: $MAX(Ri) = F(P, L, S) = P \times L \times S$.

Probability function: $P = F(4M) = F\begin{pmatrix}\text{Human Factors, Physical Factors,}\\ \text{Environment, Management}\end{pmatrix}$.

Consequence function: $L = F\begin{pmatrix}\text{Human impact, property impact,}\\ \text{environmental impact, social impact}\end{pmatrix}$.

Context function: $S = F(\text{Time sensitive, space sensitive})$.

It can be seen that the design of risk evaluation indexes should contain three dimensions P, L and S.

II. B. 2) Quantification and weighting of indicators

The quantification of indicators is divided into two stages, namely, the initial selection of indicators and the optimization of indicators. The initial selection of indicators can be realized by consulting relevant norms, standards, literature and expert meeting methods. SMART principle, KPI principle, 4M element theory and other methods are often needed in the process of initial selection of indicators. On the basis of the initial selection of indicators, it is

necessary to remove the indicators that are less important or not easy to quantify. The optimization of indicators is divided into three steps:

(1) Determine the importance of the evaluation indicators, and the grading of the importance of the evaluation indicators and the principle of assigning points are shown in Table 2:

Table 2: Evaluate the importance of indicators

Importance level	Extremely important	It doesn't matter	Ordinary	Important	Extremely important
Score	1	2	3	4	5

(2) Making and distributing questionnaires and calculating the importance coefficient of the evaluation indicators and the coefficient of variation of the indicators based on the results of the questionnaires:

$$RF_i = \frac{\sum_{j=1}^{m} a_{ij}}{m} \left(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots m\right) \tag{4}$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{m} \left(a_{ij} - RF_i\right)^2}{m-1}} \left(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots m\right) \tag{5}$$

$$CV_i = \frac{\sigma_i}{RF_i} \left(i = 1, 2, \cdots, n\right) \tag{6}$$

where m is the total number of experts participating in the questionnaire survey, n is the total number of indicators in the primary selection, a_{ij} is the importance assignment of the j th expert to the i th indicator, RF_i is the importance coefficient of the i th indicator, CV_i is the coefficient of variation of the i th indicator, and σ_i is the standard deviation of the i th indicator.

(3) Screening evaluation indicators. Remove the indicators with RF_i less than 2.5 or CV_i greater than 0.25, and the remaining indicators are the evaluation indicators of the risk warning model.

When the screening of indicators is completed, the hierarchical structure of the risk evaluation model is established according to the design principles of risk evaluation indicators. The risk grading model indicator system is divided into two levels, the first level indicators are likelihood, severity and sensitivity, and the second level indicators are the specific evaluation indicators.

II. B. 3) Indicator judgment matrix

(1) Constructing a judgment matrix

Two-by-two comparisons are to be made between the various indicators at the same level, and the results are expressed using the degree of importance. Specific comparison principles are shown in Table 3:

Table 3: Criteria for judging the importance of indicators

Importance	Judgment standard
1	Index A is as important as index B
3	Index A is slightly more important than index B
5	Index A is more important than index B
7	Index A is important to index B
9	Index A and index B are extremely important
2,4,6,8	The middle value of two important degrees

(2) Calculate the weight vector

Normalize the judgment matrix by columns:

$$b_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} (i, j = 1, 2, 3 \cdots, n) \tag{7}$$

Sum the normalized rows:

$$W = \sum_{j=1}^{n} a_{ij} \tag{8}$$

The feature vectors are obtained by normalizing the above calculations:

$$\bar{W} = \frac{W_i}{\sum_{i=1}^{n} W_i} \tag{9}$$

(3) Consistency test

The purpose of the consistency test is to control the error so that the results of weight calculation are in an acceptable range [19]. The consistency index is expressed by CI, and the smaller the CI, the higher the degree of consistency. Using λ to denote the maximum characteristic root of the judgment matrix, then:

$$CI = \frac{\lambda - n}{n - 1} \tag{10}$$

III. Bayesian cybersecurity risk assessment modeling

III. A. Bayesian network modeling approach

III. A. 1) Bayes' formula

(1) Conditional Probability

Let the events A, B be two randomly occurring events, and if $P(B) > 0$, the probability that the event A occurs under the condition that another random event B occurs is called the conditional probability and is denoted as $P(B \mid A)$.

(2) Joint Probability

Let A, B be two random events and $P(B) > 0$, $P(A) > 0$, the probability of the event A and the event B occurring at the same time is called joint probability, denoted as $P(A, B)$.

(3) A priori probability

The a priori probability $P(A)$ denotes: the probability of judging the occurrence of the event A before the occurrence of the event B, the a priori probability is also known as the marginal probability, in the joint probability, the final result of the unwanted those events are very likely to merge into the other events and disappear, which is the marginalization.

(4) Posterior Probability

The posterior probability is the probability of reassessing the occurrence of the event B after the occurrence of the event A, denoted as $P(A \mid B)$, and similarly, the probability of reassessing the occurrence of the event B after the occurrence of the event A, which is called the posterior probability of B, denoted as $P(B \mid A)$.

Bayesian formula:

According to the definition of conditional probability, the probability of an event A occurring conditional on the occurrence of the event B is:

$$P(A \mid B) = P(A, B) / P(B) \tag{11}$$

Similarly the probability of event B occurring under the condition that event A occurs is:

$$P(B \mid A) = P(A, B) / P(A) \tag{12}$$

Integrating (11) and (12) yields the Bayesian formula:

$$P(A \mid B) = P(B \mid A) P(A) / P(B) \tag{13}$$

(5) The density functional form of the Bayesian formulation

$P(X; \theta)$ represents the density function of the overall dependence on the parameter θ, which indicates that different θ correspond to different distributions in the parameter space. In Bayesian statistics it is denoted as

$P(X|\theta)$, which represents the conditional distribution of the overall indicator X when the random variable takes a given value.

The prior distribution θ can be determined from the prior information generated by the parameter (θ), and from the point of view expressed by Bayes, the sample $x=(x_1,x_2,\cdots,x_n)$ proceeds in two main steps: in the first step, a prior sample parameter θ_{0_0} is generated from a set of prior sample distributions $\pi(\theta)$ starting with $\pi(\theta)$, and in the second step, another set of samples is generated from $P(X|\theta_0)$. The joint conditional probability function of the samples at this point is:

$$P(X|\theta_0)=\prod P(X|\theta_0) \tag{14}$$

θ_0 is an unknown which arises from the distribution of a priori functions $\pi(\theta)$. In order to synthesize and generalize the a priori information into it, it is not only necessary to synthesize the constant θ_0, but it is also necessary to synthesize other possible values that can be generated, which can be synthesized once with the constant $\pi(\theta)$. In this way, a joint parametric distribution of the sample parameter $x=(x_1,x_2,\cdots,x_n)$ and the sample parameter θ is shown below:

$$h(x;\theta)=p(x|\theta)\pi(\theta) \tag{15}$$

This joint distribution contains three kinds of available information: aggregate information, sample information, and a priori information. When information about the a priori sample distribution is missing, one can usually only make inferences about the sample based on a priori sample score $\pi(\theta)$. When sample observations $x=(x_1,x_2,\cdots,x_n)$ are available, then one should make inferences about θ based on $h(x,\theta)$. As:

$$h(x;\theta)=\pi(\theta|x)m(x) \tag{16}$$

where $m(x)$ is the edge density function of x:

$$m(x)=\int_\Theta h(x,\theta)d\theta=\int_\Theta p(x|\theta)\pi(\theta)d\theta \tag{17}$$

It is independent of θ and contains no information about θ. Therefore the only thing that can be used to make inferences about θ is the conditional distribution $\pi(\theta|x)$ which is formulated as:

$$\begin{aligned}\pi(\theta|x)&=h(x,\theta)/m(x)\\&=p(x|\theta)\pi(\theta)\Big/\int_\Theta p(x|\theta)\pi(\theta)d\theta\end{aligned} \tag{18}$$

This conditional distribution is called the posterior distribution of θ, and it concentrates all the information about θ in the aggregate, the sample, and the prior, but is the result obtained by eliminating all the information that is not relevant to θ.

The formula for the posterior distribution $\pi(\theta|x)$ is the Bayesian formula expressed as a density function. It is the result of adjusting the prior distribution $\pi(\theta)$ with totals and samples, and all inferences in Bayesian statistics are made on the basis of the posterior distribution.

In the case where θ is a discrete random variable, the prior distribution can be represented by the prior distribution column $\pi(\theta_i)$, $i=1,2,\cdots$. At this point the posterior distribution is also in discrete form:

$$\pi(\theta_i|x)=h(x,\theta)/m(x)=p(x|\theta_i)\pi(\theta)\Big/\sum_i p(x|\theta_i)\pi(\theta_i) \tag{19}$$

III. A. 2) Bayesian network definition

Let $G=(I,R)$ denote a directed acyclic graph (DAG), where I then represents the set of every node in a graph region, R represents the set of directed connected line segments, and let $X=(X_i)i\in I$ be the random variable represented by a particular node i in the directed acyclic graph if the joint probability of node X is denoted by:

$$P(X)=\prod P\left(X_i \mid X_{pa(i)}\right) \tag{20}$$

Since A leads to B, A and B lead to C, which leads to:

$$P(A,B,C)=P(C \mid A,B)P(B \mid A)P(A) \tag{21}$$

III. B. Bayesian network conditional probability table determination

III. B. 1) Noisy-or Gate Modeling

The Noisy-or Gate model can be used to represent an intrinsic relationship between n variables $X_1,X_2,\cdots,X_n$ and their effect Y, and each variable is a binary variable, i.e., each variable has two specific states, assuming that one is true (1) and the other is false (0). Bayesian networks based on the Noisy-or Gate model need to fulfill two more conditions:

If the parent of any node Y is $X_1,X_2,\cdots,X_n$, then $X_1,X_2,\cdots,X_n$ should be independent of each other.

Each variable is able to cause the outcome Y to occur when all other variables are false, when only X_i is 1 and all other parents are 0. The probability that node Y takes the value 1 is $P_i=P\left(Y \mid X_1,X_2,\cdots,X_i,\cdots,X_n\right)$, and then the other terms of the conditional probability table for node Y are determined to be X by $P_1,\cdots,P_i,\cdots,P_n$ then the expression is as follows:

$$P(Y \mid X_P)=1-\prod_{i:x_i\in x_n}(1-P_i) \tag{22}$$

III. B. 2) Leaky Noisy-or Gate Modeling

The occurrence of a child node is not necessarily due to the occurrence of the parent node, but may also be due to the presence of some unpredictable or unknown factors that lead to accidents. For example, a child node has a total of 3 parent nodes, so even if all 3 parent nodes do not occur at the same time, the probability value are 0, the child node is still very likely to occur,i.e., the probability of the parent nodes are 0, but the probability of the child node is not 0 occurs, this case can be used in the Bayesian network in the Leaky Noisy-or Gate model to determine the conditional probability of its table.

Next, the values of P_i and P_1 can be computationally solved for according to the model of Leaky Noisy-or Gate. Suppose Y has only two parents: X_i and X_{all}. Then correspondingly, the sum of factors other than X_i is X_{all}, and P_i and P_{all} are the connection probabilities of X_i and X_{all}, respectively. From the theorem, X_{all} is always true and according to equation (23) there is:

$$P(Y \mid X_i)=P_i+P_{all}+P_iP_{all} \tag{23}$$

$$P(Y \mid \overline{X_i})=P_{all} \tag{24}$$

The association of Eq. (23) and Eq. (24) gives P_i as:

$$P_i=P(Y \mid X_i)-P(Y \mid \overline{X_i})\big/1-P(Y \mid \overline{X_i}) \tag{25}$$

From Eq. (25), we can calculate the connection probabilities $P_1,P_i,\cdots,P_n$ of all the parents of the node Y, and combine them with the unknown factor X_L, and its connection probability P_L, which results in the conditional probability of the node Y as:

$$P\left(Y=T\right)=1-\left(1-P_i\right)\prod_{i:x_i\in X_p}\left(1-P_i\right) \tag{26}$$

IV. Analysis of the results of the security assessment of the dual-defence mechanism

IV. A. Simulation Experiments on Bayesian Network Risk Assessment Models

IV. A. 1) Parameterized tests for Bayesian networks

In order to study the method in more detail, the study set some key parameters in the experiment. For the Bayesian network part of the study, the maximum number of iterations is set to 50, the learning rate is set to 0.05, and the training objective is set to 0.0002. The parameters of the Bayesian part are set to 3. The maximum abandonment probability is set to 0.5, the minimum abandonment probability is set to 0.001, and the maximum step length is set

to 0.5, and the minimum step length is set to 0.01. Then, respectively, from the maximum number of iterations of the scale effect of the Bayesian network algorithm was parametrized and the experimental results are shown in Fig. 2.

The accuracy of the experimental results gradually increases as the scale increases. A similar trend is observed for the maximum number of iterations, but the improvement associated with the number is small. Meanwhile, the execution time of the algorithm gradually increases with the increase of the scale and the maximum number of iterations. The results show that when the size exceeds 60 and the maximum number of iterations exceeds 30, the improvement of the algorithm accuracy is no longer obvious, and after the number of iterations 30, the error remains between 2.2 and 2.25 on average for all sizes. Therefore, in further experiments, a scale of 60 and a maximum number of iterations of 30 were chosen as parameter settings. On the other hand, the expansion of the number of iterations and the scale will make the time spent on Bayesian network learning increase and show a gradient upward trend, and the mean value of the time spent is increased from 2.29 min to 5.84, which is 155.02%.

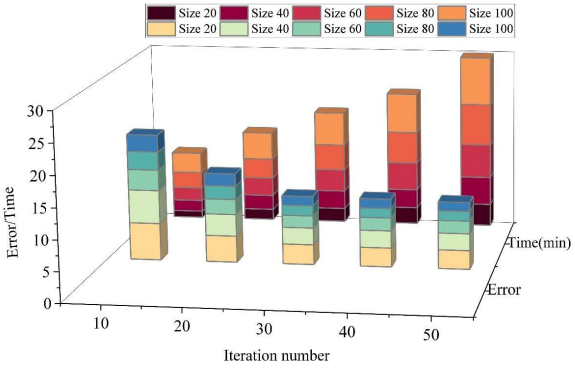
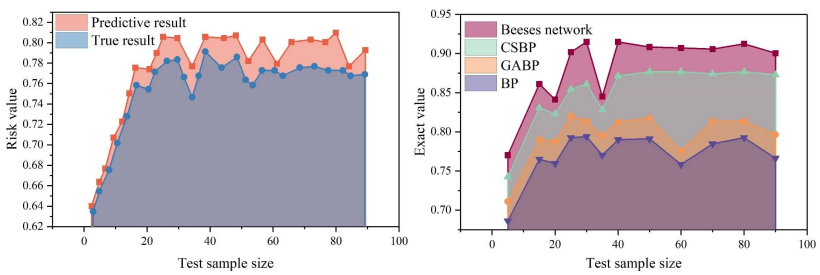


Figure 2: The parameterization test of bayesian network algorithm

IV. A. 2) Comparison of experimental results

In order to assess the stability of the proposed Bayesian network algorithm, the study conducted comparative experiments with other algorithms including BP neural network, Cuckoo's algorithm optimized BP neural network (CSBP), and Genetic Algorithm optimized BP neural network (GABP). These algorithms were tested on the same data for a total of ten experiments, and the average of the experiments was used as the final result as shown in Fig. 3, where Fig. (a) compares the true values with the predicted results, and Fig. (b) compares the results of the experiments with different algorithms.

Figure 3(a) compares the average and raw risk values of different test results. It can be seen that the mean value of the deviation between the predicted results of the simulation test and the real risk test results is 1.73%, which indicates that the model is close to the real value in predicting the results. Observing Figure 3(b), the Bayesian network algorithm achieves an accuracy of about 88.2%, which is significantly better than GABP and BP, and slightly better than CSBP.The BP neural network performs poorly in predicting the results of the risk assessment, with a large error in the prediction of the results, and an average accuracy of only 77.09%. Similar to BP neural network, CSBP also showed relatively large result prediction errors and unstable results during the experiment. Overall, the Bayesian network performs best in risk assessment, indicating that the algorithm proposed in the study is highly adaptable and robust in risk assessment.



(a)Comparison between the real result and the predicted result

(b)Compare the experimental results of different algorithms

Figure 3: The test results of multiple independent experiments

IV. B. Dynamic risk analysis

IV. B. 1) Risk reachability probability

Based on the intelligence provided by the dual prevention mechanism, a dynamic Bayesian attack graph can be constructed to assess the current dynamic security risk. It contains the updated calculation of the reachable probability of the security risk indicators, and the results are shown in Fig. 4.

It can be seen that after the security event of S9 indicator, the reachable probability of each indicator in the experimental network shows a rising trend, indicating that the overall security risk of the network is rising. Among them, the a posteriori reachable probabilities of S1, S4 and S6 are significantly higher than those of other indicators, which are 0.85, 0.78 and 0.76, respectively, indicating that they are likely to have security risks at this time and need to take remedial measures as soon as possible. The a posteriori reachable probabilities of S10, S11 and S12 are significantly higher than the a priori reachable probabilities, which are 0.28, 0.28 and 0.29, respectively, indicating that they are likely to have safety accidents next. Therefore, in the real enterprise environment, the accuracy of dynamic safety risk assessment is significantly higher than that of static safety risk assessment, which can provide effective support for risk management.

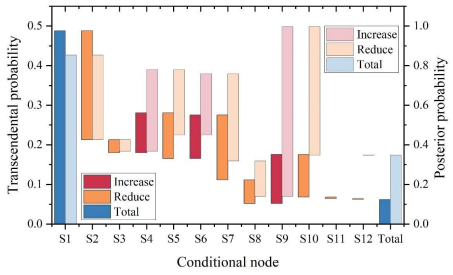


Figure 4: Index probability

IV. B. 2) Comparison of methods

Figure 5 shows the a priori reachable probability of the indicator nodes given by the three methods.CSBP and GABP also utilize algorithms to describe the causal relationship of the attack behavior, but their assessment indexes for the success probability of the atomic attack are too single, which results in the given success probability of the atomic attack and the reachable probability can not truly reflect the risk of the enterprise's asset security. In contrast, the method in this paper evaluates the success probability of atomic attack from multiple dimensions, which can better reflect the enterprise asset security risk. Moreover, both CSBP and GABP methods do not consider the vulnerability re-exploitation, so the risks of S11 and S12 are considered to be significantly smaller than that of S10, but in fact, the risks of the three are similar, and the a priori reachable probability is around 0.06. In the case that the a priori reachable probability of S10,S11 and S12 are close to each other,the standard deviation of the a priori reachable probability of this paper's method is still larger than that of CSBP and GABP,which are 0.124,0.121 and 0.132 respectively.This indicates that the method of this paper has a greater differentiation of the security risk of each node,and is favorable for distinguishing and delineating the security level of each node.

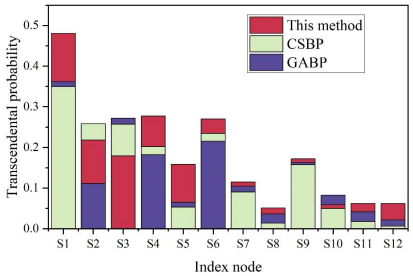


Figure 5: The probability of the prior accessibility of the index node given by three methods

Fig. 6 shows the a posteriori reachable probabilities of conditional nodes of the three methods in the event of a security event.The CSBP and GABP methods do not integrate the backward update with the forward update, and only the ancestor nodes of the attacked node are considered in the backward update.They update the reachable probabilities of S4 and S6 when S9 is attacked, but ignore S5, S7, and S8, and give results that are significantly lower than those of this paper's method, which gives significantly lower results than the present one, which gives

significantly lower results than the present method, which gives significantly lower results than this paper's method in the event of an attack on these three indicator nodes, which is 0.46, 0.32, and 0.14. The a posteriori reachable probabilities of the three indicator nodes are 0.46, 0.32, and 0.14, respectively, and the remaining two methods underestimate the real security risk.

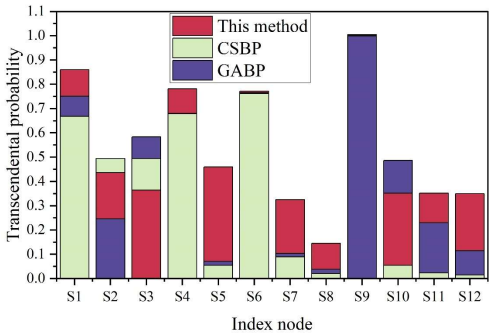


Figure 6: The probability of the three methods of the condition node after the safety event

V. Conclusion

In this paper, the risk is divided into different levels based on the difficulty of controlling the risk, and the comprehensive degree of danger is determined by the product to establish a hidden danger investigation and management mechanism, and build a dual defense mechanism for the enterprise. The Bayesian network security risk assessment model is utilized to represent the probability of occurrence of security risks. Design simulation experiments to assess the security risk situation of the dual defense mechanism. With the expansion of the number of iterations and scale, the time spent on Bayesian network learning increases, and the average value of the time spent is raised from 2.29min to 5.84, with an increase of 155.02%. Comparing the stability of other algorithms with the algorithm in this paper, the Bayesian network algorithm achieves an accuracy of about 88.2%, which is significantly better than GABP and BP, and slightly better than CSBP, and the Bayesian network performs the best in risk assessment. Based on the intelligence provided by the Dual Prevention Mechanism, a dynamic Bayesian attack graphical representation is constructed to assess the current dynamic security risk. The Bayesian for S10, S11, S12 three indicators of the a priori reachable probability has a substantial increase,respectively, 0.28, 0.28, 0.29, indicating that they are likely to be followed by a security incident.

References

[1] Dai, Y., Tong, X., & Wang, L. (2022). Workplace safety accident, employee treatment, and firm value: Evidence from China. Economic Modelling, 115, 105960.

[2] Shanfeng, W. E. I. (2023). Study on relationship between dual prevention mechanisms of safety production. China Safety Science Journal, 33(S1), 64.

[3] Yan, F. (2025). Research on Safety Risk Grading and Hidden Trouble Identification and Management Strategy of Dual Prevention Mechanism Based on Multi-Level Analysis Approach. J. COMBIN. MATH. COMBIN. COMPUT, 127, 2895-2909.

[4] Zhou, L., Fu, G., & Xue, Y. (2018). Human and organizational factors in Chinese hazardous chemical accidents: A case study of the '8.12'Tianjin Port fire and explosion using the HFACS-HC. International journal of occupational safety and ergonomics, 24(3), 329-340.

[5] Bai, T., & Zhou, Z. (2025, January). Study on the Risk Factors of the Dual Prevention Mechanism in Port Area. In 2024 7th International Conference on Civil Architecture, Hydropower and Engineering Management (CAHEM 2024) (pp. 280-285). Atlantis Press.

[6] Binbin, B. I. A. N., Yuqing, H. U. A. N. G., Zhijian, L. I. U., Xianmin, T. I. A. N., & Kai, Z. H. A. N. G. (2022). Construction of safety evaluation index system based on double prevention mechanism. China Safety Science Journal, 32(S1), 45.

[7] Dyreborg, J., Lipscomb, H. J., Nielsen, K., Törner, M., Rasmussen, K., Frydendall, K. B., ... & Kines, P. (2022). Safety interventions for the prevention of accidents at work: A systematic review. Campbell systematic reviews, 18(2), e1234.

[8] Christensen, I. (2017). The elements of a commercial human spaceflight safety reporting system. Acta Astronautica, 139, 228-232.

[9] Choudhry, R. M. (2014). Behavior-based safety on construction sites: A case study. Accident analysis & prevention, 70, 14-23.

[10] Choi, T. M., & Lambert, J. H. (2017). Advances in risk analysis with big data. Risk Analysis, 37(8), 1435-1442.

[11] Xiaorong, F., Shizhun, J., & Songtao, M. (2018, March). The research on industrial big data information security risks. In 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA) (pp. 19-23). IEEE.

[12] Zhang, Y., Geng, P., Sivaparthipan, C. B., & Muthu, B. A. (2021). Big data and artificial intelligence based early risk warning system of fire hazard for smart cities. Sustainable Energy Technologies and Assessments, 45, 100986.

[13] Xie, K., Ozbay, K., Kurkcu, A., & Yang, H. (2017). Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. Risk analysis, 37(8), 1459-1476.

[14] Huang, L., Wu, C., Wang, B., & Ouyang, Q. (2018). A new paradigm for accident investigation and analysis in the era of big data. Process safety progress, 37(1), 42-48.

[15] Latif, S., Qayyum, A., Usama, M., Qadir, J., Zwitter, A., & Shahzad, M. (2019). Caveat emptor: the risks of using big data for human development. IEEE technology and society magazine, 38(3), 82-90.

[16] Sai, Y. A. N. G., Chiyu, J. I. A. O., Ziyan, Z. H. A. O., & Jie, G. U. O. (2023). Research and practice on double prevention mechanism of university laboratory safety management system under the background of big data. Experimental Technology and Management, 40(11), 240-245.

[17] Quanlong Liu,Jianping Shang,Jingzhi Wang,Mengqi Li & Tongtong Li. (2024). Research on the evaluation of the operating effectiveness of the safety double prevention mechanism of coal mine enterprises based on matter-element extension. Process Safety and Environmental Protection,185,899-909.

[18] Wang Kun,Tian Guixian & Tao Yongchao. (2023). Quantitative model of financial risk management of forestry enterprises based on nonlinear differential equation. Journal of Computational Methods in Sciences and Engineering,23(2),809-823.

[19] Jian Li,Li li Niu,Qiongxia Chen & Mei Li. (2025). Consistency and consensus checking and improving methods for group decision-making with hesitant fuzzy preference relations. Journal of the Operational Research Society,76(1),137-154.