MDPI

Article

Risk Analysis of Laboratory Fire Accidents in Chinese Universities by Combining Association Rule Learning and Fuzzy Bayesian Networks

Fuqiang Yang [1,2,*], Xin Li [1], Shuaiqi Yuan [3,*] and Genserik Reniers [3,4,5]

1   College of Environment and Safety Engineering, Fuzhou University, Fuzhou 350116, China
2   Fujian Provincial Key Laboratory of Remote Sensing of Soil Erosion and Disaster Prevention,
    Fuzhou University, Fuzhou 350116, China
3   Safety and Security Science Section, Faculty of Technology, Policy and Management,
    Delft University of Technology, 2628 BX Delft, The Netherlands
4   Antwerp Research Group on Safety and Security (ARGoSS), Faculty of Applied Economics,
    University of Antwerp, 2000 Antwerp, Belgium
5   Center for Economics and Corporate Sustainability, Katholieke Universiteit Leuven, 1000 Brussels, Belgium
*   Correspondence: yangfuqiang@fzu.edu.cn (F.Y.); s.yuan-2@tudelft.nl (S.Y.)

Abstract: Targeting the challenges in the risk analysis of laboratory fire accidents, particularly considering fire accidents in Chinese universities, an integrated approach is proposed with the combination of association rule learning, a Bayesian network (BN), and fuzzy set theory in this study. The proposed approach has the main advantages of deriving conditional probabilities of BN nodes based on historical accident data and association rules (ARs) and making good use of expert elicitation by using an augmented fuzzy set method. In the proposed approach, prior probabilities of the cause nodes are determined based on expert elicitation with the help of an augmented fuzzy set method. The augmented fuzzy set method enables the effective aggregation of expert opinions and helps to reduce subjective bias in expert elicitations. Additionally, an AR algorithm is applied to determine the probabilistic dependency between the BN nodes based on the historical accident data of Chinese universities and further derive conditional probability tables. Finally, the developed fuzzy Bayesian network (FBN) model was employed to identify critical causal factors with respect to laboratory fire accidents in Chinese universities. The obtained results show that H4 (bad safety awareness), O1 (improper storage of hazardous chemicals), E1 (environment with hazardous materials), and M4 (inadequate safety checks) are the four most critical factors inducing laboratory fire accidents.

Keywords: laboratory fire accidents; Bayesian network; association rules; fuzzy set theory; fire safety

1. Introduction

University laboratories are important sites for educational and scientific activities, and meanwhile, some of them inevitably accommodate hazardous and flammable materials. Experimenters are usually threatened by physical, chemical, or biological threats in university labs [1], and this should be given enough attention from the public and academia. Previous research shows that university laboratories normally have laxer safety management and lower safety investment compared to industrial laboratories [2,3]. As a result, safety risks in some university laboratories are not well evaluated and treated. The frequent occurrence of university laboratory accidents, particularly the happening of lab fires in Chinese universities, has induced severe disastrous consequences. For example, a fire and explosion happened in a laboratory of the Beijing Jiao Tong University on 26 December 2018, and three people died in this accident. A deflagration happened in a laboratory of the Nanjing University of Aeronautics and Astronautics on 24 October 2021, causing two fatalities and nine injuries. Safety risk assessment and safety risk management are effective

tools that can be implemented to prevent undesired accidents and mitigate the corresponding consequences [4,5]. However, current research on the risk analysis of laboratory fire accidents is still lacking to boost the safety risk management of university laboratories. The development of new approaches for risk analysis of laboratory fire accidents and further improvement of safety risk management of university laboratories is urgently needed.

Previous studies have developed different risk analysis methods to investigate safety issues in laboratories [6–8]. However, those studies mainly focused on specific experiments [9], and a generic risk analysis model for university laboratory fire accidents is lacking. In previous studies, risky behaviors [10], safety policies [11], and chemicals [12] were considered influencing factors in the safety risk analysis of laboratory accidents. However, the dependency between those influencing factors has not been analyzed to serve a comprehensive and systematic risk assessment. Thus, a thorough identification and evaluation of the causal factors leading to university laboratory fire accidents should be performed, and the interdependency between those factors should be well addressed in the risk analysis. The Bayesian network (BN) is one of the widely used methods for safety risk analysis. BNs are able to combine probability theory and graph theory for uncertain event analysis and inference [13,14]. At present, a large number of studies have used BN models for risk analysis and casual factor identification [15–18]. For instance, Aliabadi et al. [19] assessed the gas leakage risks of storage tanks by using a BN model, and the results show that human factors are the most critical influencing factors. Wang et al. [20] constructed a risk analysis model for a phased task system based on BNs. Li et al. [21] evaluated explosion risks with respect to the aluminum production process by employing BNs, and five main causes were identified afterward. Hao and Hadjisophocleous [22] developed a BN model to estimate the probabilities of fire spreading and evaluated the corresponding fire risks. Li et al. [23] performed a risk analysis of hazardous chemical explosions by integrating BNs and association rules.

Additionally, the treatment of subjectivity and uncertainties should be well addressed in the risk analysis because it significantly influences the risk assessment results. Typically, the implementation of fuzzy set theory is able to facilitate the expert elicitation process by representing probabilities in the form of fuzzy languages. Additionally, the subjectivities and uncertainties in risk analysis may also be reduced by incorporating data-driven approaches. For instance, Nhat et al. [24] created a BN model combining a data-driven approach to improve the accuracy of the model's prediction with the help of historical data analysis. Lieng [25] combined a data-driven approach and BNs to reduce the uncertainties in risk modeling. Previous studies also show that the implementation of association rules (ARs) helps to identify and depict the probabilistic relationships between associated events/factors [26–29]. As a result, ARs are widely used for the data mining of accidental factors, particularly in the maritime and construction sectors. For instance, ARs were applied to investigate the causal factors of vessel navigation accidents [30], tugboat accidents [31], and fishing vessel accidents [32]. In construction industries, Wang et al. [28] employed ARs to improve the effectiveness of hazard identification in workplaces. Cabello et al. [29] used ARs to identify the main influencing factors leading to accidents in the construction phases. Shao et al. [33] employed ARs to identify the causes of building collapse accidents. Those studies show that the application of ARs has the capability to identify and describe probabilistic relationships among associated factors, and further, it helps to develop BN models.
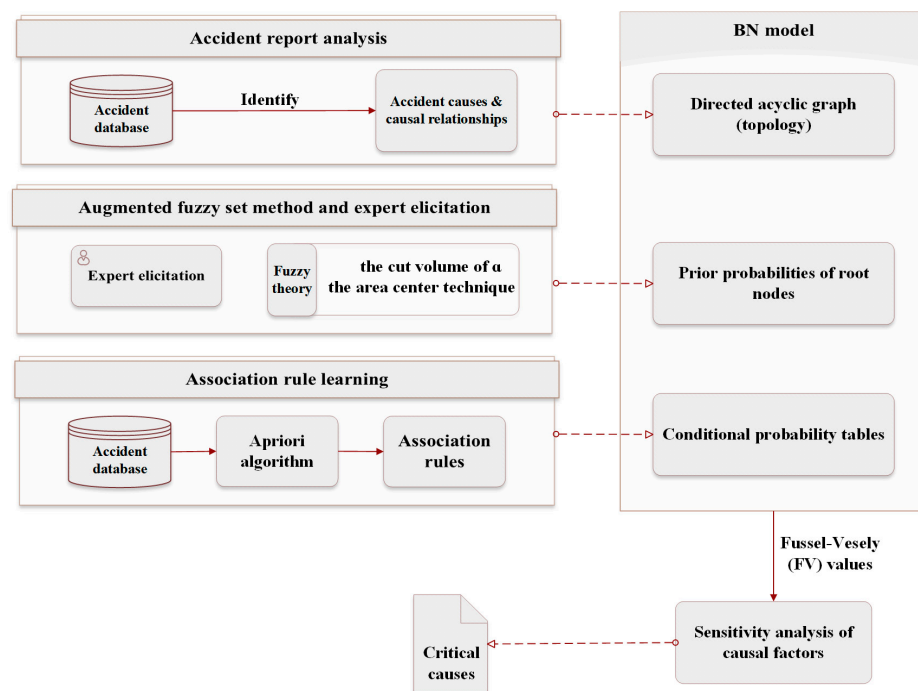
In order to fill the above-mentioned gaps in the risk analysis of university lab fire accidents, an integrated approach combining association rule learning and a fuzzy Bayesian network (FBN) is developed to perform a risk analysis of Chinese university laboratory fire accidents. The remainder of this paper is organized as follows: Section 2 demonstrates the overview of the proposed approach before illustrating each part (association rule learning, augmented fuzzy set theory, and the BN) of the proposed approach in detail. Section 3 demonstrates the detailed procedures of the application of the proposed approach to the risk analysis of a Chinese university laboratory regarding fire accidents. The risk assessment

results are discussed, and the limitations and recommendations for future works are given in Section 4. Finally, our conclusions are summarized in Section 5.

2. Overview of the Methodology

2.1. Overall Framework

A hybrid model is developed by this study to perform a risk analysis of laboratory fire accidents and to answer two research questions: What factors may lead to laboratory fire accidents in Chinese universities? and What is the criticality of each factor? Figure 1 presents the overview of the proposed approach. Firstly, the causal factors of accidents are identified through an accident report analysis of accidents. Then, a BN model topology (directed acyclic graphs (DAGs)) should be developed to describe the causal relationships among those factors, which consists of three types of nodes, including cause category nodes (human, object, environment, and management), cause nodes, and the accident node. Furthermore, the prior probabilities of the cause nodes (root nodes of the BN model) are determined by using an augmented fuzzy theory method with the help of expert elicitation. Specifically, the augmented fuzzy theory is developed by combining the cut volume of α and the area center technique. The conditional probability tables of the BN nodes are determined based on the association rule learning approach with the help of historical accident data. Finally, a complete BN model is developed to perform a quantitative risk analysis of accidents, and the calculation of Fussel–Vesely (FV) values helps to identify the critical causal factors.

Figure 1. Flowchart of the proposed methodology.

2.2. Methodology

2.2.1. Bayesian Networks

BNs are probabilistic networks consisting of DAGs and conditional probability tables (CPTs). The joint probability distribution of a set of nodes $V = \{X_1, X_2, \cdots X_n\}$ can be expressed as Equation (1) [34].

$$P(X1, X_2 \ldots Xn) = \prod_{i=1}^{n} P(Xi|Pa(Xi))$$

(1)

where $\{X_1, X_2, \cdots X_n\}$ means all nodes of a BN, and $Pa(X_i)$ denotes the parent set of node X_i. BNs employ the Bayes theorem to update the prior probabilities, called evidence, thus generating the posterior probabilities [35]:

$$P(X_j|X_i) = \frac{P(X_i, X_j)}{P(X_i)} = \frac{P\left(X_i|X_j\right) \cdot P(X_i)}{\sum_j P(X_i|X_i)P(X_i)} \tag{2}$$

It is critical to identify the important critical root nodes that cause the occurrence of the top event. Fussel–Vesely describes the contribution of the root event to the top event [36]. For a root event, X_i, the *FV* can be calculated as [35,36]:

$$FV = \frac{P(TE = occur) - P(TE = occur|X_i = 0)}{P(TE = occur)} \tag{3}$$

where *P*(*TE* = occur) refers to the probability of occurrence of the top event, T; *P*(*T* = occur | X_i = 0) refers to the probability of occurrence of the top event, T, when the X_i event does not occur.

2.2.2. Augmented Fuzzy Set Theory Method

The fuzzy set theory was proposed by Zadeh [37] to deal with imprecision and fuzzy problems. Generally, each fuzzy set is regulated by a membership function, $u(x)$, with a value domain, $U \in [0, 1]$. The value solved by the membership function is called the degree of membership, i.e., the degree to which the elements belong to the fuzzy set [38]. A fuzzy number can be defined in different forms. The concept of linguistic variables is particularly helpful when handling situations that cannot be described quantitatively. In risk assessment, fuzzy quantifiers are employed to deal with the "probability of failure/occurrence". For instance, *very low* (VL), *low* (L), *fairly low* (FL), *medium* (M), *fairly high* (FH), *high* (H), and *very high* (VH). The graphical representation of these linguistic values is presented in Figure 2.
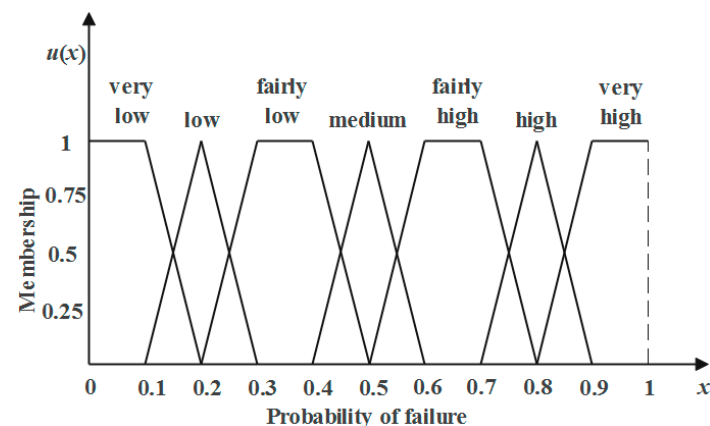

Figure 2. Fuzzy numbers representing linguistic variables.

Basically, increasing the number of experts is an important way that helps to alleviate the bias in expert opinions. Meanwhile, the application of an augmented fuzzy theory method may also help to reduce bias in expert opinions with no need to increase the number of experts. To reduce the subjective biases in expert elicitation, the cut volume of α [39] and the area center technique [40] are used for expert evaluation aggregation in this study. The cut volume of α is designed to aggregate expert opinions to a specific region, and the area center technique is developed to acquire the center of this region. By using those two approaches, expert judgments/opinions can be synthesized twice. The combination of the two methods enables the aggregation of expert opinions and, therefore, reduces the influence of subjective biases in expert elicitations. If the linguistic levels of expert *i* and

expert *j* are *medium* (M) and *fairly high* (FH), respectively, the fuzzy language corresponding to the membership function is presented as follows:

$$u_M(x) = \begin{cases} (x - 0.4)/0.1, & 0.4 < x \le 0.5 \\ (0.6 - x)/0.1, & 0.5 < x \le 0.6 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$u_{FH}(y) = \begin{cases} (y - 0.5)/0.1, & 0.5 < y \le 0.6 \\ 1, & 0.6 < y \le 0.7 \\ (0.8 - y)/0.1, & 0.7 < y \le 0.8 \\ 0, & \textit{otherwise} \end{cases} \quad (5)$$

Under the assumption that $\alpha \in u_M(x)$, then $\alpha = (x - 0.4)/0.1$ can be derived, i.e., $x = 0.1\alpha + 0.4$. x is called the left cut volume of α. Similarly, the right cut volume of α can be calculated as $-0.1\alpha + 0.6$. The cut volume of α of the fuzzy language M is $[0.1\alpha + 0.4, -0.1\alpha + 0.6]$. Table 1 displays the cut volume of α corresponding to the fuzzy numbers. L = (0.1, 0.2, 0.2, 0.3)

Table 1. Fuzzy number and cut volume of α.

Fuzzy Number	Cut Volume of α	Fuzzy Number	Cut Volume of α
VL = (0, 0, 0.1, 0.2)	$u_{VL}^{\alpha} = [0, -0.1\alpha + 0.2]$	L = (0.1, 0.2, 0.2, 0.3)	$u_L^{\alpha} = [0.1\alpha + 0.1, -0.1\alpha + 0.3]$
FL = (0.2, 0.3, 0.4, 0.5)	$u_{FL}^{\alpha} = [0.1\alpha + 0.2, -0.1\alpha + 0.5]$	M = (0.4, 0.5, 0.5, 0.6)	$u_M^{\alpha} = [0.1\alpha + 0.4, -0.1\alpha + 0.6]$
FH = (0.5, 0.6, 0.7, 0.8)	$u_{FH}^{\alpha} = [0.1\alpha + 0.5, -0.1\alpha + 0.8]$	H = (0.7, 0.8, 0.8, 0.9)	$u_H^{\alpha} = [0.1\alpha + 0.7, -0.1\alpha + 0.9]$
VH = (0.8, 0.9, 1, 1)	$u_{VH}^{\alpha} = [0.1\alpha + 0.8, 1]$	/	/

If the weight of expert *i* is 0.6 and the weight of expert *j* is 0.4, the set of fuzzy evaluations can be represented as follows:

$$W_\alpha = 0.6u_M^\alpha + 0.4u_{FH}^\alpha = [0.6(0.1\alpha + 0.4) + 0.4(0.1\alpha + 0.5), 0.6(-0.1\alpha + 0.6) + 0.4(-0.1\alpha + 0.8)] =$$

$$[0.1\alpha + 0.44, -0.1\alpha + 0.68]$$

From the fuzzy set theory point of view, W_α is also a fuzzy set. Let $W_\alpha = [C_1, C_2] = [0.1\alpha + 0.44, -0.1\alpha + 0.68]$, then $\alpha^L = (C_1 - 0.44)/0.1$ and $\alpha^R = (0.68 - C_2)/0.1$, where α^L is the left cut volume of α and α^R is the right cut volume of α. Therefore, the membership function of the new fuzzy number, W, is presented as follows:

$$u_W(C) = \begin{cases} (C - 0.44)/0.1, & 0.44 < C \le 0.54 \\ 1, & 0.54 < C \le 0.58 \\ (0.68 - C)/0.1, & 0.58 < C \le 0.68 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

Figure 3 shows the function region of the fuzzy number, W. x_0 denotes the center of the region in the x-coordinate, and it can be presented by using Equation (7).

$$x_0 = \frac{\int_X xf(x)dx}{\int_X f(x)dx} \quad (7)$$

Accordingly, the fuzzy possibility score (FPS), considering the opinions from both expert *i* and expert *j*, is x_0. Eventually, Equation (8) can be used to transform the fuzzy possibility scores into fuzzy probabilities [41,42].

$$P_f = \begin{cases} 1/10^K, & FPS \ne 0 \\ 0, & FPS = 0 \\ K = 2.301 \times [(1 - FPS)/FPS]^{1/3} \end{cases} \quad (8)$$

Additionally, this paper uses a similarity aggregation method to decide the weight of each expert [43]. It is assumed that expert i and expert j choose fuzzy sets $A = (A_1, A_2, \ldots, A_n)$ and $B = (B_1, B_2, \ldots, B_n)$, respectively. Then, the similarity between i and j is measured by Equation (9):
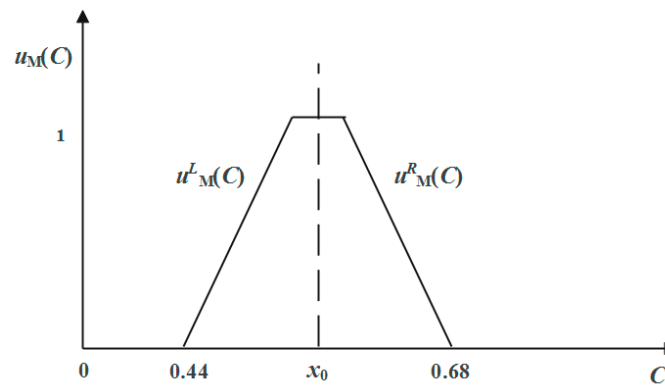
$$S_{ij} = 1 - \frac{\sum_{k=1}^{n} |Ak - Bk|}{4} \tag{9}$$

The average agreement ($AA(E_{xi})$) of each expert is presented by Equation (10):

$$AA(E_{xi}) = \frac{\sum_{\substack{i \neq j \\ j=1}}^{M} S_{ij}}{M - 1} \tag{10}$$

where M is the number of experts. $RAD(E_{xi})$ is defined as the ith relative agreement degree of expert weights, which can be calculated by using Equation (11).

$$RAD(E_{xi}) = \frac{AA(E_{xi})}{\sum_{j=1}^{M} AA(E_{xj})} \tag{11}$$

Figure 3. The function of fuzzy number W.

2.2.3. Association Rules (ARs) Method

The ARs method is a data mining method that is able to identify correlations between different factors associated with similar events and extract a subset of frequent factors associated with an event [44]. Typically, an apriori algorithm is one of the widely used algorithms for association rule learning. An apriori algorithm is able to express causal relationships in a quantitative/probabilistic manner. Therefore, this study employs an apriori algorithm to quantify the conditional probabilities of the BN nodes. In ARs, the algorithm usually covers three indexes (support, confidence, and lift) for presenting the correlations between different factors. They can be calculated as follows [26]:

$$S(A \Rightarrow B) = P(A \cup B) = \frac{|A \cup B|}{|D|} \tag{12}$$

where $S (A \Rightarrow B)$ is the support of the A to B association. A and B are two different item sets. P is the probability that A and B item sets appear simultaneously in the D transaction set. $|D|$ is the transaction set. $|A \cup B|$ is the number of times that the A and B item sets appear simultaneously in the transaction set. The support of the A to B association refers to the probability that A and B appear simultaneously, and the correlation is strong if the two appear together frequently.

$$C(A \Rightarrow B) = P(B|A) = \frac{|A \cup B|}{|A|} \tag{13}$$

where $C(A \Rightarrow B)$ is the confidence of the *A* to *B* association, which refers to the probability of B occurring if A has already occurred. $|A|$ is the number of occurrences of *A* in the transaction set.

$$L(A \Rightarrow B) = \frac{P(BA)}{P(B)} = \frac{P(A \cup B)}{P(A)P(B)} \tag{14}$$

where $L(A \Rightarrow B)$ is the lift of the *A* to *B* association, which refers to the correlation between *A* and *B*. When $L = 1$, A is uncorrelated with B. When $L < 1$, *A* and *B* are negatively correlated. When $L > 1$, *A* and *B* are positively correlated.

Figure 4 shows the AR algorithm procedures, where the confidence index is used to present the conditional probability of the BN nodes. The role of the AR is to determine the association degree between different BN nodes. The accident cause types and the accident are considered to be associated only when $L > 1$. Consequently, the conditional probabilities can be calculated for the interrelated factors; otherwise, the conditional probabilities are zero. Because the existing literature and practices on AR technologies have no standardized guidelines for determining the $support_{min}$ and $confidence_{min}$ thresholds [26,29,45,46], this study obtains the appropriate $support_{min}$ and $confidence_{min}$ by comparing the effectiveness of the threshold values continuously. For instance, the $support_{min}$ and $confidence_{min}$ were initially set as 0.2 and 0.5, respectively. However, the generated association rules cannot be fully corresponded to the conditional probabilities of the BN nodes. Until the $support_{min}$ and $confidence_{min}$ were set as 0.1 and 0.3, respectively, the conditional probabilities of the BN nodes were all derived. Detailed procedures for association rule learning can be found in Figure 4, in which the node "association rule" represents the probabilistic association between the BN nodes. The calculation of the association rules was performed by using the software IBM SPSS Modeler 18.0.
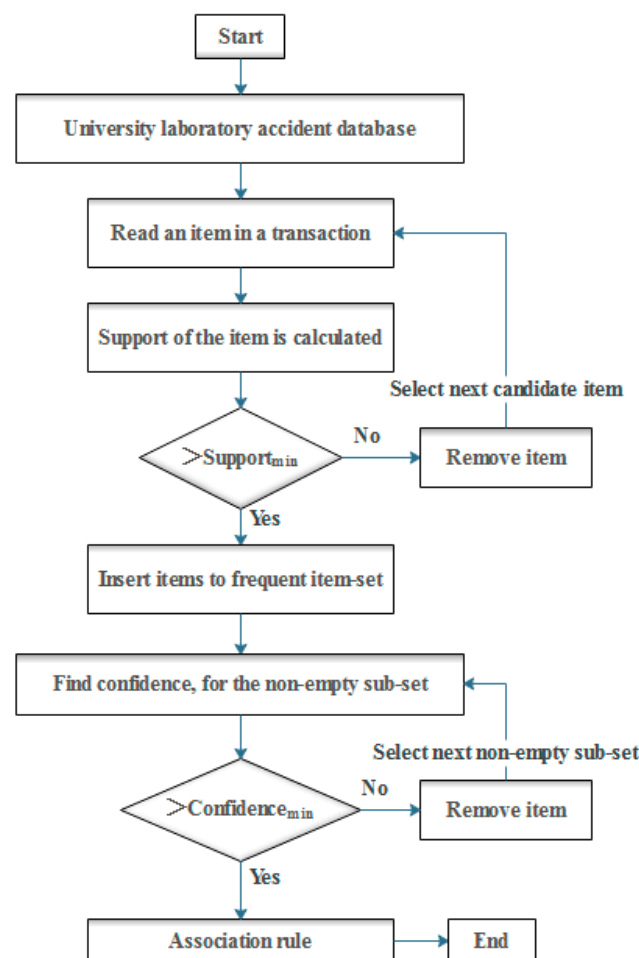

Figure 4. Flowchart of the apriori algorithm.

3. Example Analysis of Fire Accidents in Laboratories

3.1. Data Collection

The confidence ($A \Rightarrow B$) is considered the conditional probability of the happening of B given A. However, according to Equation (2), when the AR analyzes only one accident, $|A \cup B| = |A|$ and C ($A \Rightarrow B$) = 1. In other words, if only one type of accident is analyzed, the confidence is always 1. The confidence cannot be converted into the conditional probability of the BN in this situation. Accordingly, the statistical data need to include all types of accidents. The relationship between accident causes and fire accidents is explored from all types of accidents.

Accordingly, 121 cases of laboratory accidents (about 121 laboratory accidents were publicized in Chinese universities from 2000–2022) are used to identify the main causes of laboratory accidents. These accidents can be found on the official websites of some colleges and universities, while accident reports need to be retrieved. Fires, explosions, electric shocks, poisoning, and other accidents are included in the accident database. The identified causes are classified into four categories: human (human factors), object (object conditions), environment (environmental factors), and management (organizational and management factors). Among these 121 accidents, only the type of accident and the type of cause are collected. This is to explore the probabilistic relationships between fire accidents and the human, object, environment, and management.

With respect to fire incidents, detailed accident causes need to be identified and collected based on the 121 accident reports. Table 2 presents the description of the identified accident causes with their symbols. The names of the accident causes are defined based on the statements in the accident reports. The description of the data format and data processing with software can be found in Section 3.3. Additionally, some issues need to be noticed during the data collection process. For example, the time span of the data should be chosen as no less than 10 years to ensure the annual regularity of the data. Meanwhile, the number of accident cases should be as large as possible to ensure data generality. In addition, some accident reports with incomplete information should not be selected to safeguard the completeness of the results.
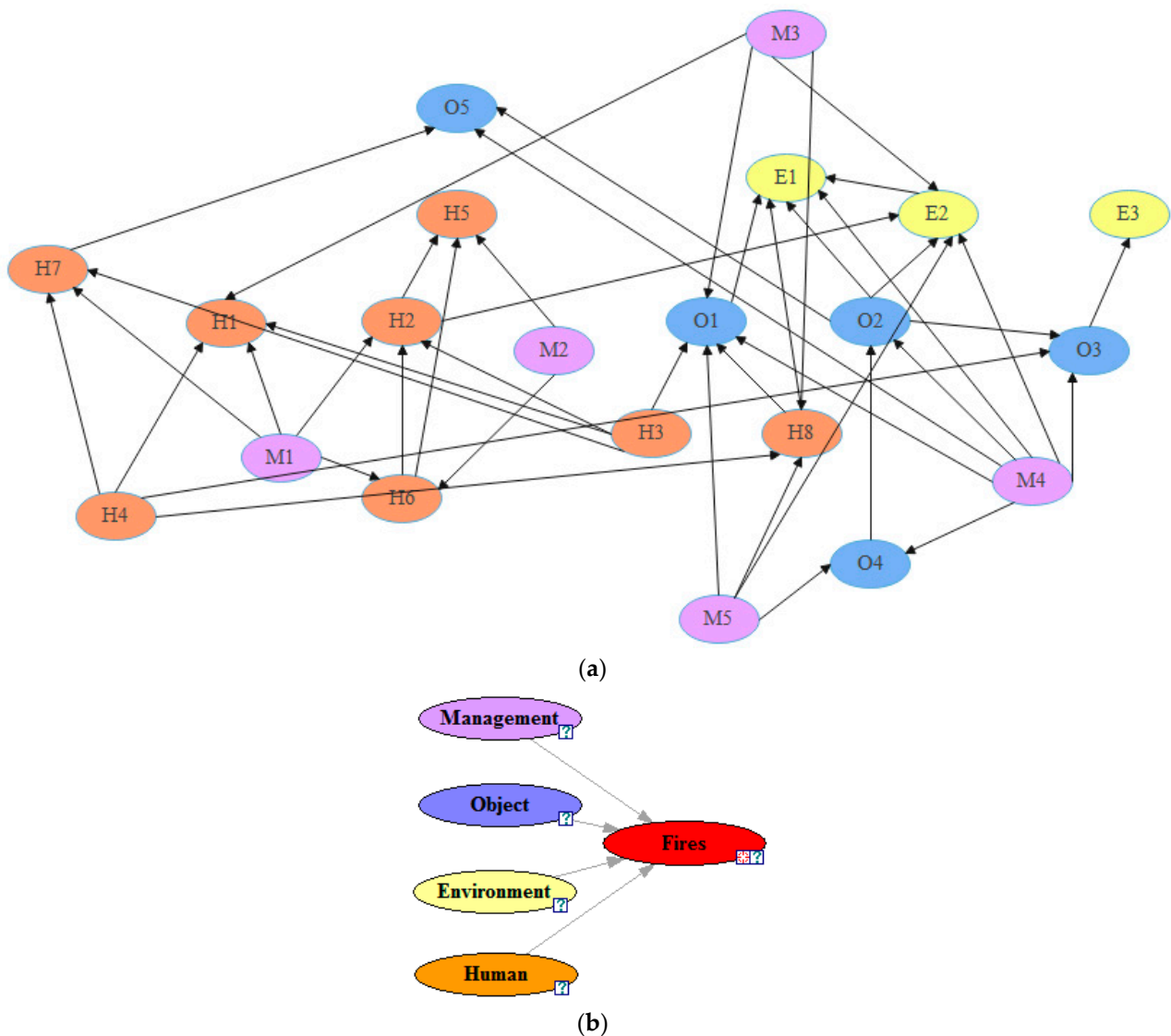
Table 2. Descriptions of the identified causes/causal factors with their symbols.

Cause Categories	Cause of the Fire Accidents	Marks
Human	Leaving the lab for too long of a time during the experiment	H1
	Violation of operating procedures (including improper operation)	H2
	Violation of the laboratory management regulations	H3
	Bad safety awareness	H4
	Improper configuration of experimental conditions	H5
	Lack of specialized knowledge	H6
	Failure to turn off the power of the instrument after the experiment (including electrical equipment)	H7
	Cluttered placement of chemicals	H8
Object	Improper storage of hazardous chemicals	O1
	Equipment and facilities problems (e.g., equipment maintenance problems)	O2
	Insufficient protection equipment (e.g., fire-fighting equipment or personal protection equipment)	O3
	Equipment exceeds its service life	O4
	Circuit failure (such as circuit aging)	O5
Environment	Environment with hazardous materials (including flammable and explosive gases, liquids, and solids)	E1
	Leak or retention of the experimental substance	E2
	Presence of undesirable environment (e.g., high temperature, noise, etc.)	E3
Management	Lack of supervision and guidance	M1
	Inadequate safety practice and education	M2
	Safety management system problems	M3
	Inadequate safety checks (including performing risk assessments)	M4
	Safety responsibility division problems	M5

3.2. Construction of Bayesian Network Topology

After the identification of the accident causes/causal factors, a BN topology structure was developed, considering the dependency between those BN nodes, as presented in Figure 5. Since the whole diagram is complex, Figure 5 divides the topology into two parts:

(a) presents the relationships among specific accident causes, and (b) shows the relationship between the cause category nodes and the fire accident node. In the BN model, the cause nodes "H1" to "H8" in Figure 5a should be pointed to the "Human" node in Figure 5b. Similarly, other cause nodes should be linked to the corresponding cause categories. Then, the prior probabilities of the root nodes in Figure 5a were determined by using fuzzy set theory and expert elicitations. The conditional probabilities of the BN nodes were obtained from the accident database with the help of AR learning. The determination of the prior probabilities and conditional probabilities of the BN nodes is illustrated in the next sections.

(**a**)

(**b**)

Figure 5. The topology of the developed BN model for university laboratory fire risk analysis. (**a**) indicates the relationship between the specific causes; (**b**) shows the relationship between the cause categories and the fire accidents. (different node colors correspond to different node categories).

3.3. Association Rule Learning

To clearly explain the methodology, the data format during the data collection process is described. Some of the accident data analysis results are presented in Table 3 for demonstrative purposes. In Table 3, "√" indicates the accident types (column header). "1" indicates that the causes of the specific accident case belong to the corresponding cause category (the column header); by contrast, "0" presents that the accident causes do not belong to the corresponding category. For example, the first case is a fire accident, which is

mainly caused by humans, objects, and the environment. By following the data format in Table 3, the data were collected in an Excel spreadsheet.

Table 3. Data expressions required by the apriori algorithm.

Case Number	Fire	Explosion	Electric Shock	Poisoning	Other	Human	Object	Environment	Management
1	√	0	0	0	0	1	1	1	0
2	√	0	0	0	0	1	0	0	0
3	0	0	0	√	0	1	0	1	0

Table 4 shows a part of the association rule learning results, with a total of 15 rules. As shown in Table 4, because the lift index value of the human and management factors in relation to a fire accident is less than one, the corresponding conditional probability is considered zero. Similarly, the lift index value of environmental factors to fire incidents is less than one, so the corresponding conditional probability is also zero. The obtained conditional probability table (CPT) for the fire accident node is presented in Table 5. The conditional probability tables for other BN child nodes were also obtained in the same way by using association rule learning.

Table 4. The results of association rules.

Rule No.	Front Items	Subsequent Item	Support	Confidence	Lift
1	Environment	Fire	0.292	0.543	1.229
2	Management	Fire	0.683	0.476	1.077
3	Object	Fire	0.367	0.614	1.389
4	Human	Fire	0.725	0.391	1.885
5	Environment Object	Fire	0.133	0.750	1.698
6	Environment Management	Fire	0.183	0.636	1.441
7	Object Management	Fire	0.258	0.645	1.461
8	Environment Human	Fire	0.142	0.353	0.799
9	Human Management	Fire	0.467	0.446	1.011
10	Object Human	Fire	0.175	0.524	1.186
11	Environment Object Management	Fire	0.092	0.727	1.647
12	Environment Management Human	Fire	0.075	0.556	1.258
13	Environment Object Human	Fire	0.042	0.400	0.906
14	Object Management Human	Fire	0.108	0.615	1.393
15	Human Object Management Environment	Fire	0.025	0.333	0.755

Table 5. The conditional probability table of the fire accident node derived by ARs.

Human	Environment	Object	Management	Y(Fire)	N(Fire)
Y	Y	Y	Y	0.667	0.333
			N	0.667	0.333
		N	Y	0.692	0.308
			N	0.577	0.423
	N	Y	Y	0.615	0.385
			N	0.524	0.476
		N	Y	0	1
			N	0.404	0.596
N	Y	Y	Y	0.786	0.214
			N	0.789	0.211
		N	Y	0.692	0.308
			N	0	1
	N	Y	Y	0.645	0.355
			N	0.605	0.395
		N	Y	0.456	0.544
			N	0	1

Y indicates the event occurs; N means the event does not occur.

3.4. Determination of Prior Probabilities

In this study, five experts with rich lab fire safety and lab experiment experience were invited to score the basic events related to laboratory fires concerning the safety engineering laboratory at Fuzhou University. This section demonstrates the determination of the prior probabilities of the root nodes (basic events) by using the BN node, "M1", as an example. The evaluation results from the experts on the "M1" node are "M", "FH", "FH", "VH", and "M". The weight of each expert was obtained according to Equations (9)–(11), as presented in Table 6.

Table 6. Results of fuzzy opinion aggregation.

Experts	$\sum_{\substack{i \neq j \\ j=1}}^{M} S_{ij}$	AA (Average Agreement)	RAD (Relative Agreement Degree)
1	3.275	0.819	0.205
2	3.425	0.856	0.214
3	3.425	0.856	0.214
4	2.6	0.65	0.162
5	3.275	0.819	0.205

The corresponding fuzzy set is constructed by using the cut volume of α and the expert weights. The calculation process is as follows:

$$W_\alpha = 0.205u_M^\alpha + 0.214u_{FH}^\alpha + 0.214u_{FH}^\alpha + 0.162u_{VH}^\alpha + 0.205u_M^\alpha$$

$$= [0.205(0.1\alpha + 0.4) + 0.214(0.1\alpha + 0.5) + 0.214(0.1\alpha + 0.5) + 0.162(0.1\alpha + 0.8) + 0.205(0.1\alpha + 0.4),$$

$$0.205(-0.1\alpha + 0.6) + 0.214(-0.1\alpha + 0.8) + 0.214(-0.1\alpha + 0.8) + 0.162 + 0.205(-0.1\alpha + 0.6)]$$

$$= [0.1\alpha + 0.5076, -0.0838\alpha + 0.7504]$$

Consequently, the membership function of the fuzzy set is as follows:

$$u_W(C) = \begin{cases} (C - 0.5076)/0.1, & 0.5076 < C \le 0.6076 \\ 1, & 0.6076 < C \le 0.6666 \\ (0.7504 - C)/0.0838, & 0.6666 < C \le 0.7504 \\ 0, & \text{Otherwisetherwise} \end{cases}$$

The fuzzy set of x_0 is calculated as 0.6325, according to Equation (7). Then, the fuzzy probability of the happening of "M1" is calculated as 0.0120 with Equation (8). The prior probabilities of the remaining root nodes were also determined by using the same method. Table 7 presents the obtained expert evaluations for the remaining root nodes. Correspondingly, Table 8 provides the FPS and the fuzzy prior probabilities for the root nodes.

Table 7. The results of the expert elicitation.

Experts / Events	Ex_1	Ex_2	Ex_3	Ex_4	Ex_5
M2	H	M	FL	VH	H
M3	FH	FL	H	M	L
M4	VH	H	M	VH	FH
M5	M	M	H	FH	M
H3	M	VH	H	FL	H
H4	H	FH	H	M	H

Table 8. The FPS and fuzzy probabilities of the root nodes.

Events	FPS (Fuzzy Possibility Score)	Prior Probabilities	Events	FPS (Fuzzy Possibility Score)	Prior Probabilities
M1	0.6325	1.20×10^{-2}	M5	0.5811	8.65×10^{-3}
M2	0.6890	1.72×10^{-2}	H3	0.6913	1.74×10^{-2}
M3	0.4998	4.99×10^{-3}	H4	0.7174	2.06×10^{-2}
M4	0.7677	2.85×10^{-2}	-	-	-

4. Results and Discussions

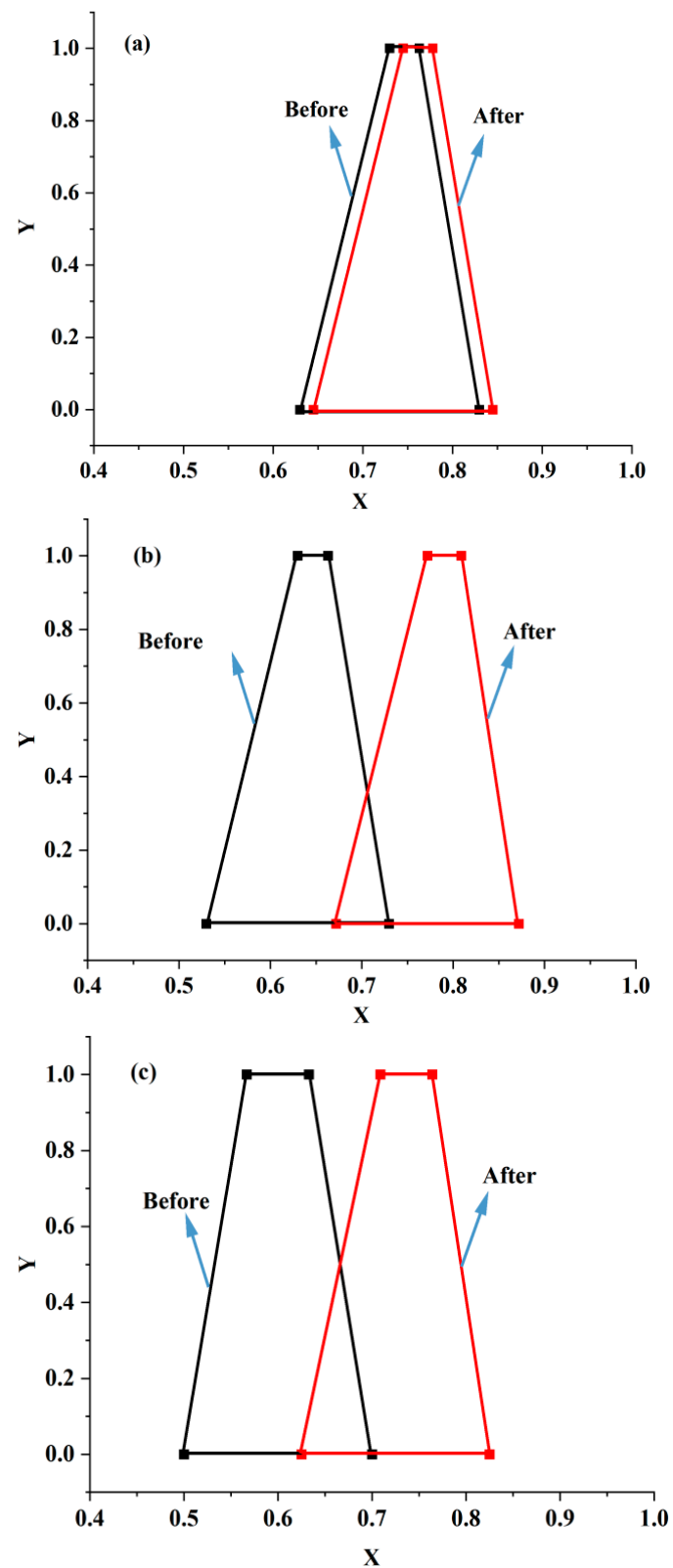### 4.1. Validation of the Augmented Fuzzy Set Method

This section discusses the effectiveness of the augmented fuzzy set method in aggregating expert opinions with possible biases. Table 9 shows some examples of the expert evaluation results. Cases 1 to case 3 are used to demonstrate the aggregation of expert opinions in the case of a low evaluation existing among the evaluation results from other experts being relatively higher. By contrast, cases 4 to case 6 are used to demonstrate the aggregation of expert opinions in the case of a high evaluation existing among the evaluation results from other experts being relatively lower.

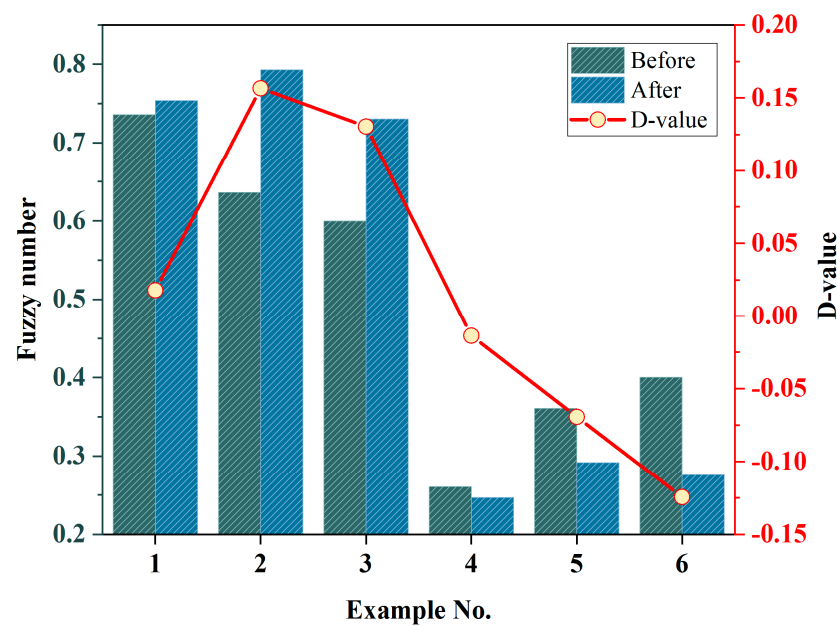**Table 9.** Examples of the evaluation results from three experts.

Example No.	Evaluation Results	Example No.	Evaluation Result
1	(M, H, VH)	4	(VL, L, M)
2	(L, H, VH)	5	(VL, L, H)
3	(VL, H, VH)	6	(VL, L, VH)

Figure 6 compares the fuzzy regions before and after using the similarity aggregation method (the cut volume of α). It can be observed from Figure 6 that the fuzzy regions are strongly reshaped by using the similarity aggregation methods when there are obvious differences/biases in the experts' opinions. The range of the reshaped fuzzy regions is more consistent with the expert opinions with similarities. Additionally, the fuzzy value of each example is calculated and presented in Figure 7 to demonstrate the effectiveness of the similarity aggregation method. "Before" indicates the calculated fuzzy number without using the similarity aggregation method, and "After" indicates the calculated fuzzy number using the similar aggregation method. "D-value" means the difference between the calculated fuzzy numbers with and without the implementation of the similarity aggregation method. It can be observed from Figure 7 that the aggregation effects of the similarity aggregation method become more obvious when the difference in the expert

evaluations is larger. In conclusion, the augmented fuzzy set method is able to aggregate expert opinions, especially when there are obviously differences/biases in the expert opinions.

Figure 6. Fuzzy regions before and after using the similarity aggregation method. (**a**) Example No. 1. (**b**) Example No. 2. (**c**) Example No. 3.

Figure 7. Changes in the fuzzy numbers before and after using the similarity aggregation method.

4.2. Impact Level Analysis of the Cause Categories

The impact of each cause category on the happening of lab fires is compared in this section through a sensitivity analysis of the cause category nodes. Figure 8 shows the FV value of each cause category node. It is observed that management factors are the most important causes of the occurrence of lab fire accidents. In fact, the habits, behaviors, and safety awareness of laboratory staff are significantly influenced by management factors [47]. A survey shows that the use of safety recording management systems in laboratories can reduce operators' risky behaviors and habits effectively [10]. Therefore, management factors should be given enough attention by the laboratory managers and the university safety management teams. The improvement of the safety management systems helps to reduce the risky behaviors and habits of lab operators and, meanwhile, enhances the safety awareness of laboratory staff to reduce the risks of the happening of lab fire accidents.

Figure 8. The FV values of different cause categories.

4.3. Impact Level Analysis of Basic Events

This section discusses the sensitivity of each basic event (causal factors) on the happening of lab fire accidents. Figure 9 shows the FV value of each basic event with respect to the

happening of lab fire accidents. In terms of human factors, H4 (bad safety awareness) is the most critical cause. Ozsahin et al. [48] and Walters et al. [49] emphasized the importance of the safety awareness of students to the safety of university laboratories, which is consistent with this finding. The establishment of high-quality safety training and education programs may help to enhance the safety awareness of laboratory staff and further reduce laboratory accident risks. Diverse learning methods can better attract the attention of students and improve learning efficiency [50]. Universities may provide diverse approaches to improve the learning efficiency of workers in safety education, for example, slides, videos, news, and the implementation of electronic games and VR techniques. In terms of object factors, O1 (improper storage of hazardous chemicals) is the most critical causal factor. Some existing accident cases also demonstrate this point. For instance, a fire and explosion happened in a laboratory of Beijing Jiao Tong University on 26 December 2018. According to the accident investigation report, an important cause of this accident was the improper storage of hazardous chemicals. In the management of hazardous chemical storage, the safety responsibility system should be implemented, and the information on hazardous chemicals should be recorded in detail. At the same time, colleges and universities should set up special safety inspection programs for hazardous chemical storage tanks. In terms of environmental factors, E1 (environment with hazardous materials) is the most critical factor. In university laboratories, experiment environments may have huge interventions to the happening of laboratory accidents. For instance, conducting experiments in a space with excessive concentrations of dangerous gases (e.g., methane and other flammable gases) may trigger undesired lab accidents. Zhang et al. [15] developed a BN model to investigate the evolution process of a gas leak in laboratories. The results obtained in the same study show that gas accumulation is one of the critical factors affecting accident evolutions. Therefore, the monitoring of the experiment environment is important to the safety of university laboratories. For instance, gas detection devices and hazardous alarm systems can be installed to monitor experiments with flammable and explosive gases. In terms of management factors, M4 (inadequate safety checks) is the most critical influence factor. This result is consistent with the result obtained by Ma et al. [16]. There are many ways that can be employed to improve the quality of safety checks/inspections in university laboratories. For example, the utilization of artificial intelligence tools to help with some of the safety check tasks. Zhang and Guo [51] designed a sentry robot that can take appropriate emergency measures based on the detection of dangerous situations. In addition, universities can also use various risk assessment methods to provide a reference basis for safety inspections.
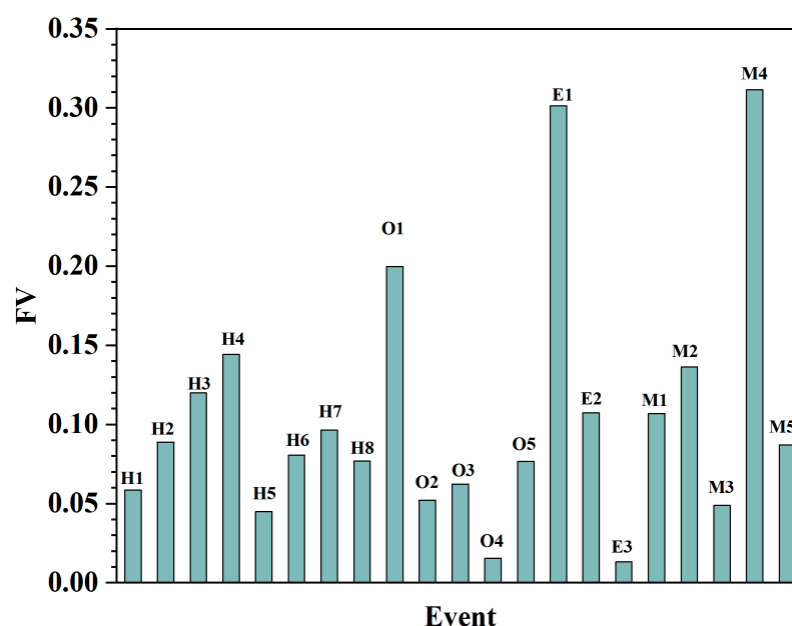

Figure 9. The FV values of basic events.

4.4. Limitations and Future Work

Inevitably, the risk analysis results of the proposed approach rely on the quality and amount of accident data, which is also a limitation for all similar data-driven risk analysis approaches. Therefore, with more historical accident data becoming available, the practicality and feasibility of the proposed model may be improved to generate more reasonable risk estimations. Furthermore, it should be noted that subjective bias may also be induced by an inappropriate selection of experts, the unreasonable formulation of questions, bad interpretation of evaluation results, and so on. Therefore, when applying the proposed approach in practice, those factors that may introduce bias in expert elicitation results cannot be ignored either. Future studies may focus on developing systematic approaches/guidelines for the appropriate selection of experts, framing questions, and interpreting results to avoid or reduce biases in expert elicitation.

In the future, the proposed methodology may be integrated with risk-based decision-making approaches [52] to achieve risk-based fire safety management for university laboratories. For example, it is possible to consider both the safety investment and risk-reduction performance of the candidate safety management strategies [53] and then achieve cost-effective safety management of the university laboratory regarding fire accidents. Moreover, it should be noted that it is also possible to apply the proposed approach to the risk analysis of laboratory fire accidents in other countries or regions. This can be performed by replacing the accident database with the accident statistics associated with the investigated university laboratory. Also, an expert evaluation team can help to determine the prior probabilities of the basic events/factors regarding fire accidents in a specific university laboratory.

5. Conclusions

In this study, by combining association rules, a BN, and an augmented fuzzy set method, an integrated risk analysis method is proposed for the risk analysis of laboratory fire accidents. A case study was employed to demonstrate the feasibility of the proposed approach in the risk analysis of lab fire accidents in Chinese universities. Expert elicitation was used to determine prior probabilities for the BN model with the help of an augmented fuzzy set method. The effectiveness of the augmented fuzzy set method in expert opinion aggregation with possible biases was demonstrated. Additionally, a database with 121 cases of laboratory accidents that happened in Chinese universities was used to derive conditional probabilities for the BN model with the help of association rule learning. With the application of the proposed model in the risk assessment of Chinese university lab fires, the results indicate that H4 (bad safety awareness), O1 (improper storage of hazardous chemicals), E1 (environment with hazardous materials), and M4 (inadequate safety checks) are the four most critical events for the occurrence of fire accidents. In addition, management factors have the most significant impact on the happening of laboratory fires. With more historical accident data becoming available, the practicality and feasibility of the proposed model may be improved to generate more reasonable risk estimations.

Abbreviations

AA	Average agreement
ARs	Association rules
BN	Bayesian network
CPT	Conditional probability tables
DAG	Directed acyclic graph
FBN	Fuzzy Bayesian network
FPS	Fuzzy possibility scores
FV	Fussell–Vesely
RAD	Relative agreement degree

References

1. Yang, Y.F.; Reniers, G.; Chen, G.H.; Goerlandt, F. A bibliometric review of laboratory safety in universities. *Saf. Sci.* **2019**, *120*, 14–24. [CrossRef]
2. Marendaz, J.-L.; Suard, J.-C.; Meyer, T. A systematic tool for Assessment and Classification of Hazards in Laboratories (ACHiL). *Saf. Sci.* **2013**, *53*, 168–176. [CrossRef]
3. Schröder, I.; Huang, D.Y.Q.; Ellis, O.; Gibson, J.H.; Wayne, N.L. Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers. *ACS Chem. Health Saf.* **2016**, *23*, 12–23. [CrossRef]
4. Omidvari, M.; Mansouri, N.; Nouri, J. A pattern of fire risk assessment and emergency management in educational center laboratories. *Saf. Sci.* **2015**, *73*, 34–42. [CrossRef]
5. Yuan, S.; Yang, M.; Reniers, G.; Chen, C.; Wu, J. Safety barriers in the chemical process industries: A state-of-the-art review on their classification, assessment, and management. *Saf. Sci.* **2022**, *148*, 105647. [CrossRef]
6. Kremer, G.G.; Ryan, T.J.; Switzer, S. *A Risk Assessment Method and Safety Plan for a University Research Lab*; The American Society of Mechanical Engineers: New York, NY, USA, 2008; pp. 263–268. [CrossRef]
7. Langerman, N. Lab-scale process safety management. *J. Chem. Health Saf.* **2009**, *16*, 22–28. [CrossRef]
8. Leggett, D.J. Identifying hazards in the chemical research laboratory. *Process Saf. Prog.* **2012**, *31*, 393–397. [CrossRef]
9. Pluess, D.; Meyer, T.; Masin, J.; Mikulasek, P.; Ferjencik, M. Joint applicability test of software for laboratory assessment and risk analysis. *J. Loss Prev. Process Ind.* **2016**, *40*, 234–240. [CrossRef]
10. Shariff, A.M.; Norazahar, N. At-risk behaviour analysis and improvement study in an academic laboratory. *Saf. Sci.* **2012**, *50*, 29–38. [CrossRef]
11. Wirth, O.; Foreman, A.M.; Friedel, J.E.; Andrew, M.E. Two discrete choice experiments on laboratory safety decisions and practices. *J. Saf. Res.* **2020**, *75*, 99–110. [CrossRef]
12. Fatemi, F.; Dehdashti, A.; Jannati, M. Implementation of Chemical Health, Safety, and Environmental Risk Assessment in Laboratories: A Case-Series Study. *Front. Public Health* **2022**, *10*, 898826. [CrossRef]
13. Hsu, W.H. Genetic wrappers for feature selection in decision tree induction and variable ordering in Bayesian network structure learning. *Inf. Sci.* **2004**, *163*, 103–122. [CrossRef]
14. Garg, A.; Pavlovic, V.; Rehg, J.M. Boosted learning in dynamic Bayesian networks for multimodal speaker detection. *Proc. IEEE* **2003**, *91*, 1355–1369. [CrossRef]
15. Zhang, X.; Hu, X.; Bai, Y.; Wu, J. Risk Assessment of Gas Leakage from School Laboratories Based on the Bayesian Network. *Int. J. Environ. Res. Public Health* **2020**, *17*, 426. [CrossRef]
16. Ma, L.; Ma, X.; Xing, P.; Yu, F. A hybrid approach based on the HFACS-FBN for identifying and analysing human factors for fire and explosion accidents in the laboratory. *J. Loss Prev. Process Ind.* **2022**, *75*, 104675. [CrossRef]
17. Chen, X.-L.; Lin, W.-D.; Liu, C.-X.; Yang, F.-Q.; Guo, Y.; Li, X.; Yuan, S.-Q.; Reniers, G. An integrated EDIB model for probabilistic risk analysis of natural gas pipeline leakage accidents. *J. Loss Prev. Process Ind.* **2023**, *83*, 105027. [CrossRef]
18. Wu, J.; Fang, W.; Tong, X.; Yuan, S.; Guo, W. Bayesian analysis of school bus accidents: A case study of China. *Nat. Hazards* **2019**, *95*, 463–483. [CrossRef]
19. Aliabadi, M.M.; Pourhasan, A.; Mohammadfam, I. Risk modelling of a hydrogen gasholder using Fuzzy Bayesian Network (FBN). *Int. J. Hydrogen Energy* **2019**, *45*, 1177–1186. [CrossRef]
20. Wang, Z.; Zeng, S.; Guo, J.; Che, H. A Bayesian network for reliability assessment of man-machine phased-mission system considering the phase dependencies of human cognitive error. *Reliab. Eng. Syst. Saf.* **2021**, *207*, 107385. [CrossRef]
21. Li, L.; Xu, K.; Yao, X.; Chen, S. Probabilistic analysis of aluminium production explosion accidents based on a fuzzy Bayesian network. *J. Loss Prev. Process Ind.* **2021**, *73*, 104618. [CrossRef]
22. Hao, C.; Hadjisophocleous, G.V. The modeling of fire spread in buildings by Bayesian network. *Fire Saf. J.* **2009**, *44*, 901–908.
23. Li, X.; Chen, C.; Hong, Y.-D.; Yang, F.-Q. Exploring hazardous chemical explosion accidents with association rules and Bayesian networks. *Reliab. Eng. Syst. Saf.* **2023**, *233*, 109099. [CrossRef]
24. Nhat, D.M.; Venkatesan, R.; Khan, F. Data-driven Bayesian network model for early kick detection in industrial drilling process. *Process Saf. Environ. Prot.* **2020**, *138*, 130–138. [CrossRef]
25. Lieng, H. A probabilistic framework for component-based vector graphics. *Comput. Graph. Forum* **2017**, *36*, 195–205. [CrossRef]

26. Hong, J.; Tamakloe, R.; Park, D. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. *Accid. Anal. Prev.* **2020**, *142*, 105497. [CrossRef]

27. Li, H.; Li, X.; Luo, X.; Siebert, J. Investigation of the causality patterns of non-helmet use behavior of construction workers. *Autom. Constr.* **2017**, *80*, 95–103. [CrossRef]

28. Wang, X.; Huang, X.; Luo, Y.; Pei, J.; Xu, M. Improving Workplace Hazard Identification Performance Using Data Mining. *J. Constr. Eng. Manag.* **2018**, *144*, 04018068. [CrossRef]

29. Cabello, A.T.; Martínez-Rojas, M.; Carrillo-Castrillo, J.A.; Rubio-Romero, J.C. Occupational accident analysis according to professionals of different construction phases using association rules. *Saf. Sci.* **2021**, *144*, 105457. [CrossRef]

30. Chen, D.J.; Pei, Y.L.; Xia, Q. Research on human factors cause chain of ship accidents based on multidimensional association rules. *Ocean Eng.* **2020**, *218*, 107717. [CrossRef]

31. Çakır, E.; Fışkın, R.; Sevgili, C. Investigation of tugboat accidents severity: An application of association rule mining algorithms. *Reliab. Eng. Syst. Saf.* **2021**, *209*, 107470. [CrossRef]

32. Özaydın, E.; Fışkın, R.; Uğurlu, Ö.; Wang, J. A hybrid model for marine accident analysis based on Bayesian Network (BN) and Association Rule Mining (ARM). *Ocean Eng.* **2022**, *247*, 110705. [CrossRef]

33. Shao, L.; Guo, S.; Dong, Y.; Niu, H.; Zhang, P. Cause analysis of construction collapse accidents using association rule mining. *Eng. Constr. Arch. Manag.* **2022**; *ahead-of-print*. [CrossRef]

34. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*; Morgan Kaufmann: Burlington, MA, USA, 1988.

35. Guo, X.; Ji, J.; Khan, F.; Ding, L.; Yang, Y. Fuzzy bayesian network based on an improved similarity aggregation method for risk assessment of storage tank accident. *Process Saf. Environ. Prot.* **2021**, *144*, 242–252. [CrossRef]

36. Yan, F.; Xu, K.; Yao, X.; Li, Y. Fuzzy bayesian network-bow-Tie analysis of gas leakage during biomass gasification. *PLoS ONE* **2016**, *11*, e0160045. [CrossRef]

37. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]

38. Zhou, Q.; Wong, Y.D.; Loh, H.S.; Yuen, K.F. A fuzzy and Bayesian network CREAM model for human reliability analysis—The case of tanker shipping. *Saf. Sci.* **2018**, *105*, 149–157. [CrossRef]

39. Lin, C.-T.; Wang, M.-J.J. Hybrid fault tree analysis using fuzzy sets. *Reliab. Eng. Syst. Saf.* **1997**, *58*, 205–213. [CrossRef]

40. Purba, J.H. A fuzzy-based reliability approach to evaluate basic events of fault tree analysis for nuclear power plant probabilistic safety assessment. *Ann. Nucl. Energy* **2014**, *70*, 21–29. [CrossRef]

41. Onisawa, T. An approach to human reliability in man-machine systems using error possibility. *Fuzzy Sets Syst.* **1988**, *27*, 87–103. [CrossRef]

42. Onisawa, T. An application of fuzzy concepts to modelling of reliability analysis. *Fuzzy Sets Syst.* **1990**, *37*, 267–286. [CrossRef]

43. Rajakarunakaran, S.; Kumar, A.M.; Prabhu, V.A. Applications of fuzzy faulty tree analysis and expert elicitation for evaluation of risks in LPG refuelling station. *J. Loss Prev. Process Ind.* **2015**, *33*, 109–123. [CrossRef]

44. Xu, R.; Luo, F. Risk prediction and early warning for air traffic controllers' unsafe acts using association rule mining and random forest. *Saf. Sci.* **2021**, *135*, 105125. [CrossRef]

45. Xu, C.; Bao, J.; Wang, C.; Liu, P. Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. *J. Saf. Res.* **2018**, *67*, 65–75. [CrossRef] [PubMed]

46. Weng, J.; Zhu, J.-Z.; Yan, X.; Liu, Z. Investigation of work zone crash casualty patterns using association rules. *Accid. Anal. Prev.* **2016**, *92*, 43–52. [CrossRef]

47. Escoboza, M.; Laborin-Alvarez, J.F.; Alvarez-Chavez, C.R.; Noriega-Orozco, L.; Borbon-Morales, C. Safety climate perceived by users of academic labora-tories in higher education institutes. *Saf. Sci.* **2019**, *121*, 93–99. [CrossRef]

48. Özsahin, A.; Demir, M.; Zencir, M.; Demir, S.; Kaleli, I. Safety awareness among laboratory workers. *Adv. Ther.* **2006**, *23*, 414–420. [CrossRef]

49. Walters, A.U.; Lawrence, W.; Jalsa, N.K. Chemical laboratory safety awareness, attitudes and practices of tertiary students. *Saf. Sci.* **2017**, *96*, 161–171. [CrossRef]

50. Vignoli, M.; Nielsen, K.M.; Guglielmi, D.; Mariani, M.G.; Patras, L.; Peirò, J.M. Design of a safety training package for migrant workers in the construction industry. *Saf. Sci.* **2021**, *136*, 105124. [CrossRef]

51. Zhang, D.; Guo, Z. Mobile Sentry Robot for Laboratory Safety Inspection Based on Machine Vision and Infrared Thermal Imaging Detection. *Secur. Commun. Networks* **2021**, *2021*, 6612438. [CrossRef]

52. Yuan, S.; Reniers, G.; Yang, M. Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. *J. Loss Prev. Process Ind.* **2023**, *83*, 105034. [CrossRef]

53. Yuan, S.; Reniers, G.; Yang, M.; Bai, Y. Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. *Process Saf. Environ. Prot.* **2023**, *170*, 356–371. [CrossRef]