Research Article

Critical Factors Analysis of Severe Traffic Accidents Based on Bayesian Network in China

Hong Chen, Yang Zhao, and Xiaotong Ma

College of Transportation Engineering, Chang’an University, Xi’an 710064, China

Correspondence should be addressed to Hong Chen; glch@chd.edu.cn

Received 26 June 2020; Revised 29 September 2020; Accepted 27 October 2020; Published 16 November 2020

Academic Editor: Kun Wang

Copyright © 2020 Hong Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of this study is to minimize the negative influences of the severe traffic accidents in China by profoundly analyzing the complex coupling relations among accident factors contributing to the single-vehicle and multivehicle traffic accidents with the Bayesian network (BN) crash severity model. The BN model was established by taking the critical factors identified with the improved grey correlation analysis method as node variables. The severe traffic accident data collected from accident reports published in China were used to validate this model. The model’s efficiency was validated objectively by comparing the conditional probability obtained by this model with the actual value. The result shows that the BN model can reflect the real relations among factors and can be seen as the target network for the severe traffic accidents in China. Besides, based on BN’s junction tree engine, five-factor combination sequences for the number of deaths and three-factor combination sequences for the number of injuries were ranked according to the severity degree to reveal the critical reasons and reduce the massive traffic accidents damage.

1. Introduction

Severe traffic accidents occur in random form regardless of time and space [1]. Mass casualties and high risk are two main distinctive features that can quickly differentiate severe traffic accidents from general accidents [2]. Besides, the enormous negative impacts of severe traffic accidents on public opinion and personal property security also need to be noticed by the traffic administration and scholars in the field of traffic safety [3–5]. However, the studies of severe traffic accidents are lacking, no matter at home or abroad. Although several relevant researches and policies have been carried out because of the high frequency of severe traffic accidents in foreign countries [6], there are still deficiencies in understanding the critical contributing factors and mechanisms of extraordinarily severe traffic accidents in China [7, 8]. Besides, the tool of early accident prevention and emergency rescue is not perfect, which weakens the prevention ability. Therefore, it is necessary to research severe serious traffic accidents deeply [9–11]. Exploring the occurrence law of serious accidents and taking effective prevention measures play an important role in reducing the severity and improving road safety in China [12].

Drivers’ behaviors are considered as the main factors causing traffic accidents in early studies. Some scholars believe that drivers’ illegal behaviors significantly impact road traffic safety [13], and drivers themselves are related to accidents [14–16]. For example, Shinar in Israel has verified that the use of seat belts is positively correlated with the age and education level of the drivers [17]. Vehicle conditions, road conditions, environmental conditions, and more and more social and economic factors have been gradually taken into account to study the impacts of fatal traffic accidents [18–20]. Researchers from Japan have evaluated the traffic safety of 46 prefectures in Japan and concluded that natural binding force, such as social rules and social capital, could reduce the dangerous driving behaviors [21]. Researchers from several European countries have deeply integrated police investigation data and accident reports and established an accident information collection and analysis system to analyze numerous traffic accidents in Europe [22–26]. It is found that the proportion of accidents caused by
“vehicles driving off the roads” is up to 70% [27]. Peng and Boyle et al. have studied traffic accidents based on Washington’s accident database and found that speeding, fatigue, driving, distraction, and driving without seat belts would affect the occurrence rate of accidents by prediction of the logical model [28]. Theofilatos et al. have analyzed the influencing factors of road accidents in urban and suburban areas based on Greece’s accident data and found the influencing factors of road accidents in urban areas are the drivers’ age, location of the intersection, and bicycle parking, while the main influencing factor of road accidents in suburban areas is weather condition [7]. Aidoo et al. have studied the relationship between road condition and accident frequency and found that lighting conditions at night, road alignment, and weather conditions significantly affect traffic accidents [6, 29]. Zhao and Deng have studied the characteristics and development trend of expressway accidents based on the annual statistical accident reports from 1995 to 2010. The results show that the factors of weather, region, time, and vehicle type contribute to the traffic accident [30]. The Bayesian network (BN) is widely used in sample learning methods, network structure construction, reasoning mechanism learning, and so on because of the powerful reasoning function [31]. In traffic accident safety analysis, BN is widely used to analyze the causes of maritime traffic accidents and road traffic accidents [32–34]. In algorithm solving, the genetic algorithm is introduced into BN’s incremental learning, which alleviates the local extremum problem in the searching process [35]. A loop deletion algorithm considering KL spacing is also used to learn the structure of BN, eliminating the dependency on node order in the modeling process [36]. When establishing the Bayesian network model, the researchers have comprehensively considered the decision variables of solving the problem and the relationship among various factors. They have used the reasoning ability of BN to analyze the multiattribute decision-making problem in an uncertain environment [37].

Accidents’ research has been transformed from the initial single-factor analysis to multifactor analysis for a long time [38, 39]. However, several systematic reviews of the iteration among the influencing factors only consider the polymorphism of the consequences of accidents. Few in-depth discussions have been conducted on the mechanism by the objective data [40, 41]. This paper aims to identify the critical factors contributing to severe single-vehicle and multivehicle traffic accidents separately and explore the inherent relationships among different factors based on objective data. Through a comprehensive comparison of these factors, some recommendations can be made in this paper for active precaution system construction. Hence, an improved grey correlation analysis method and BN traffic severity model were constructed in this paper. Firstly, the weighted grey relational degree was used to determine the critical factors contributing to single-vehicle and multivehicle traffic accidents, respectively. Secondly, the BN model was constructed, taking the critical factors as the nodes and the inherent correlations as the links. Thirdly, the sample data was trained based on the continuous condition solved by the CH score learning theory solved with the K2 algorithm. Finally, the conditional probability based on Bayesian estimation was used to validate the model’s efficiency.

2. Data Description

2.1. Data Sources. The investigation and disposal report of accident reports in the production process has been published in China annually to record the accidents accurately and timely, whose transparency was required since 2014. According to the property loss and casualties, four traffic accident categories are shown in Table 1. The standards and collected accident data define this paper’s research objects, namely, road traffic accidents with ten or more deaths, including serious and extremely serious traffic accidents.

The data of 142 investigation reports were collected from investigation and disposal reports of accident reports in the production process from 2010 to 2016, available on the State and Provincial Work Safety administrations website in China. Besides, traffic accidents were divided into two categories: single-vehicle accidents and multivehicle accidents as the distributions of “occurrence time,” “occurrence location,” “vehicle type,” and “accident characteristics” are quite different [42]. Table 2 shows the raw data of some samples.

2.2. Data Virtualization and Discretization for the Sample Set. It is necessary to select the influencing factors before factor analysis to improve computing efficiency and highlight the correlation degree among factors. According to the 4M systematic theory principle, humans, facilities, environments, and management are regarded as the direct factors that play the dominant role in the accident occurrence. In this paper, these surveyed reports are taken as research objects to sort out the critical accident data, which are the basis for fatal traffic accidents study in China, mainly including four aspects and 35 items.

According to the BN model, construction requirements, classification, and coding need to be processed to virtualize and discretize the nodes’ attribute variables. The variables’ virtualization is an assignment of each attribute. The discretization is to map the assignment of continuous variables to the several mutually disjoint ranges. Referenced by the model construction experience of the investigation and disposal report of accident reports in the production process, the assignment result of node variables identified by the improved grey correlation is shown in Table 3.

3. Methodology

The basic idea of the traffic crash severity analysis model based on BN is firstly, determining the critical factors based on the improved grey correlation for network construction; secondly, clarifying the potential interconnectivity among network nodes and expressing directly through the network graph and the structure learning process based on the CH score adapted in this paper; thirdly, using node probability learning based on Bayesian estimation (BE) for validating
tion theory can be described as follows: applying weights. The application of deviation maximization theory is applied to enhance the traditional grey analysis, and relative greyness analysis. Deviation maximization method based on a grey theory using the correlation degree distinguishes the evaluation objects. Traditional grey correlation analysis, absolute grey correlation method and then to overcome the limitation of Deng’s correlation degree, absolute greyness degree, and relative greyness degree, respectively.

(iii) Determination of weight coefficient.

\[
\begin{align*}
\alpha_{ij} &= Q_{t_{ij}} / (Q_{t_{ij}} + Q_{p_{ij}}), \\
\beta_{ij} &= Q_{c_{ij}} / (Q_{c_{ij}} + Q_{p_{ij}}), \\
y_{ij} &= Q_{p_{ij}} / (Q_{c_{ij}} + Q_{p_{ij}}).
\end{align*}
\]

where \(Q_{t_{ij}}, Q_{c_{ij}},\) and \(Q_{p_{ij}}\) are the deviation value of \(r[Y_i, X_j], \epsilon[Y_i, X_j],\) and \(\mu[Y_i, X_j]\) compared with other indexes, respectively. The calculation method is

\[
\begin{align*}
Q_{t_{ij}} = \sum_{i=1}^{n} |r[Y_i, X_j] - r[Y_i, X_j]|, \\
Q_{c_{ij}} = \sum_{i=1}^{n} |\epsilon[Y_i, X_j] - \epsilon[Y_i, X_j]|, \\
Q_{p_{ij}} = \sum_{i=1}^{n} |\mu[Y_i, X_j] - \mu[Y_i, X_j]|.
\end{align*}
\]

The number of accidents \(Y_1,\) the number of deaths \(Y_3,\) and the number of injuries \(Y_2\) are usually selected to describe a traffic accident [43]. Hence, these three indexes were used as the reference sequence in the improved grey correlation method. 

\[
\tau [Y_i, X_j] = \alpha_{ij} r [Y_i, X_j] + \beta_{ij} \epsilon [Y_i, X_j] + y_{ij} \mu [Y_i, X_j].
\]

(1)
Table 3: Network node variable assignment standard for single-vehicle and multivehicle traffic accidents.

| Variable classification | Variables             | Symbol | Type of variable | Variables assignment | Single-vehicle Frequency | Multivehicle Frequency |
|-------------------------|-----------------------|--------|------------------|-----------------------|--------------------------|------------------------|
| Accident feature        | Number of deaths      | De     | Category variable| 1 = 10–19 death, 2 = 20–29 death, 3 = 30 death, and more | 1.435                    | *                      |
|                         | Number of injuries    | In     | Category variable| 1 = 10 injury and less, 2 = 10–29 injury, 3 = 30 injury, and more | 1.597                    | *                      |
| Driver behavior         | Overspeed or not      | Os     | Category variable| 1 = yes and 2 = no   | 1.62                     | *                      |
|                         | Misoperation or not   | Mo     | Category variable| 1 = yes and 2 = no   | 1.72                     | *                      |
|                         | Mislane use or not    | Iu     | Category variable| 1 = yes and 2 = no   | 1.88                     | *                      |
|                         | Fatigue driving or not| Fd     | Category variable| 1 = yes and 2 = no   | 1.98                     | *                      |
| Vehicle                 | Vehicle types         | Vt     | Category variable| 1 = commuter bus and 2 = others | 0.418                   | 0                      |
|                         | Driving status        | Tc     | Category variable| 1 = turning and 2 = others | 1.516                   | *                      |
|                         | Vehicle driving status| Cl     | Category variable| 1 = cross line driving and 2 = others | *                      | 1.58                   |
|                         | Vehicle safety status | Bb     | Category variable| 1 = poor braking and 2 = others | *                      | 1.62                   |
| Road                    | Pavement alignment    | Sr     | Category variable| 1 = straight segment and 2 = nonstraight segment | 1.694                   | 1.30                   |
|                         | Physical separation   | Ps     | Category variable| 1 = nonseparation and 2 = separation | 1.210                   | 1.38                   |
| Environment             | Accident occurring period| Dt    | Category variable| 1 = daytime and 2 = night | 1.290                   | 1.40                   |
|                         | Workday or not        | Wd     | Category variable| 1 = yes and 2 = no | *                       | 1.32                   |
|                         | Weather condition     | Aw     | Category variable| 1 = bad weather and 2 = good weather | 1.710                   | 1.72                   |
|                         | Visibility            | Lv     | Category variable| 1 = low visibility and 2 = good visibility | 1.677                   | 1.50                   |

Determination of the critical factors based on the improved greyness correlation method

Data virtualization and discretization

Determination of node variables

Determination of network structure based on data learning and experts' knowledge

Sample data of extremely serious and serious traffic accidents in China

Determination of node probability by data learning

Construction of BN model

Factors combination ranking based on BN junction tree engine

Validation of the model

Figure 1: The flowchart of BN model construction.
model for the accidents’ feature. Moreover, the factors in Table 2 were inputted into the model also as the reference sequence. The flowchart of critical factor identification is shown in Figure 2.

### 3.2. BN Modeling

The interdependence of multiple factors in severe traffic accidents in network graphics can be studied with BN based on probability theory. BN is mainly composed of the Directed Acyclic Graph (DAG) and the Conditional Probability Table (CPT).

Utilizing the conditional relation among variables, the joint probability distribution can be formed with BN to reduce the complexity. Supposing that the random variable represented by node $i$ is $X = (X_i)_{n=1}^i$, then the joint probability of node $i$ is

$$P(I) = \prod_{i=1}^n P(X_i | X_1, X_2, \ldots, X_{i-1}) = \prod_{i=1}^n P(X_i | X_{\text{node}(i)}),$$

(4)

where $X_{\text{node}(i)}$ is the parent node of node $i$. With the probability value of the input variable (evidence variable), the probability distribution of the output variable (query variable) can be calculated according to the existing network structure and CPT. Therefore, the logical relationship between node variables in the network model is manifested in the propagation of conditional probability, which makes it possible to analyze the network’s inference.

### 3.3. Structure Learning Based on the CH Score Method and K2 Algorithm

Structure learning is a data mining process, aiming to clarify the potential interconnectivity among network nodes and express directly through network graph. The principle is to construct the network structure according to certain grading criteria and searching strategies. Although the most optimized network structure is not always available, the accuracy, complexity, and robust model can be evaluated thoroughly. The model is expressed as

$$\left\{ \begin{array}{l} \max, \ f(N, D), \\ s. t. \ N \in \Phi, |N| = C, \end{array} \right.$$

(5)

where $\Phi$ is the possible network structure; $f$ is the evaluation score; $|N| = C$ represents that structure $N$ meets the limitation of constrained requirements $C$. Since the evaluation function used in this paper is based on BN, the most-optimized network structure is

$$N_{\text{opt}} = \arg \max_N P(N|D) = \arg \max_N \frac{P(D|N)P(N)}{P(D)},$$

(6)

where $P(N|D)$ is the posterior probability of structure $N$ under a given training data set $D$, $P(N)$ is the corresponding prior probability.

The iteration steps for network construction are shown as follows:

- **Step 1**: The factors are selected as the initial network nodes.
- **Step 2**: An empty network is provided, and the node sequence of $c = \{x_1, x_2, \ldots, x_n\}$ is supposed.
- **Step 3**: The score function is calculated and the parent nodes are updated by the nodes with more significant posterior probability and connecting.
- **Step 4**: Judging the number of parent nodes. If $|\text{node}(x_j)| < 2$, continue search. Moreover, give priority to the other nodes without corresponding parent nodes, which must meet the requirement that the maximization of the new CH score function $V_{\text{new}} = CH(x_i, \text{node}(x_i) \cup \{x_j\} | D)$. If $V_{\text{new}} > V_{\text{old}} = CH(x_i, \text{node}(x_i) | D)$, then select the $x_j$ as the new parent node; else stop search.
- **Step 5**: The node variables and the parent nodes are connected to form the directional edge of the network.

### 3.4. Node Probability Learning Based on BE

Node probability learning is searching the parameters’ variables through data mining when the network structure is known. The parameter learning method of this paper is the BE, which can combine the prior knowledge and training data set to improve the model’s accuracy. The fundamental mechanism is as follows.

Supposing that the prior probability of network parameters is $p(\lambda)$, this paper searches the parameters with maximum posterior probability through the training data set $D = \{x_1, x_2, \ldots, x_n\}$. Then, the posterior probability is calculated as

$$P(\lambda|D) = \frac{P(D|\lambda)p(\lambda)}{P(D)},$$

(7)

According to the law of total probability, $p(D) = \int P(D|\lambda)p(\lambda)d\lambda$. Supposing that the samples are independent of each other, $p(D|\lambda) = \prod_{i=1}^n p(x_i|\lambda)$; then,

$$P(\lambda|D) = \frac{\left(\prod_{i=1}^n p(x_i|\lambda)\right)p(\lambda)}{\int \left(\prod_{i=1}^n p(x_i|\lambda)\right)p(\lambda)d\lambda}$$

(8)

$$\lambda = E(\lambda|D) = \int \lambda P(\lambda|D)d\lambda.$$

Because of the conjugate nature of Dirichlet distribution, the calculation complexity of this network model can be reduced significantly. Therefore, the Dirichlet distribution is usually used to improve the efficiency of $P(\lambda)$.

When the network structure is determined, the probability relation among variables can be described by the conditional probability. Supposing that the prior distribution of each node variable is Dirichlet distribution, the Full-BNT toolbox in Matlab was used to learn the conditional probability under different contributing factors with the BN estimation method. Then, the junction tree engine in Matlab
was used to combine the factor links. The model’s effectiveness was validated by comparing learning results and the actual results.

4. Result

4.1. Critical Factors Identification Result for Single-Vehicle and Multivehicle Accidents. The critical factor identification results based on the improved grey correlation analysis are shown in Tables 4 and 5.

According to Pearson’s correlation analysis principle, the factor with a coefficient of more than 0.75 is considered to have a significant effect [44]. Hence, the factors shown above were classified and organized according to an average weighted correlation degree of more than 0.75, which is the standard to build the set of key influencing factors for single-vehicle and multivehicle traffic accidents, respectively. In conclusion, the number of accidents is the dominant feature of the system, followed by the number of injuries and deaths, indicating that various factors have a considerable correlation with the number of accidents. Similarly, the number of accidents is the dominant feature of multivehicle traffic accidents, followed by the number of injuries and deaths, indicating that various factors correlate with the number of accidents. The critical factor set for single-vehicle and multivehicle traffic accidents was established through the weighted grey correlation degree, as shown in Figures 3 and 4.

4.2. BN Network for Single-Vehicle and Multivehicle Accidents’ Severity. Based on the analysis result of the improved grey correlation method, two categories of accident variables as network nodes were obtained in this paper. The first category is the primary variables, including drivers’ behavior, vehicles, road, and the environment. The second category is the result variables, including the number of deaths and the number of injuries. As for single-vehicle accidents, 13 nodes variables were selected in this paper, shown in the preliminary learning result in Figure 5. Besides, Figure 6 shows the preliminary learning result of a multivehicle traffic accident network with 16 nodes’ variables.

Figure 5 shows that the node variables 4, 6, and 10 are not connected to nodes 1 and 2. Figure 6 shows that the node variables of overspeed, fatigue driving, driving cross the line, and low visibility have a low correlation degree with the two features of the accident in the multivehicle network. Therefore, the node variables of misoperation, driving status, and safety protection facility were deleted from the single-vehicle network. These four factors in multivehicle preliminary BN were deleted as well. Continuing to be trained by the K2 algorithm for BN structure searching iteration, the final learning result is split into two networks, shown in Figures 7 and 8, respectively, because of the independence of the number of deaths and the number of injuries.

In Figure 7, ten factors contribute to the occurrence of single-vehicle traffic accidents with two accident feature
Table 4: The weighted grey correlation degree of single-vehicle traffic accidents’ factors.

| Categories   | Accident index       | Number of accidents | Number of deaths | Number of injuries | Mean value |
|--------------|----------------------|---------------------|------------------|--------------------|------------|
| Driver factors | Overspeed           | 0.788               | 0.822            | 0.870              | 0.827      |
|              | Overload             | 0.663               | 0.752            | 0.734              | 0.716      |
|              | Misoperation         | 0.814               | 0.735            | 0.753              | 0.767      |
|              | Driving without license | 0.735        | 0.707            | 0.723              | 0.722      |
| Vehicle factors | Poor braking         | 0.735               | 0.678            | 0.754              | 0.722      |
|              | Commuter bus         | 0.839               | 0.823            | 0.848              | 0.837      |
|              | Light track          | 0.641               | 0.710            | 0.720              | 0.690      |
|              | Straight driving     | 0.691               | 0.745            | 0.784              | 0.740      |
|              | Parking              | 0.827               | 0.807            | 0.798              | 0.811      |
| Road factors | Incomplete safety facility | 0.837          | 0.721            | 0.733              | 0.764      |
|              | Nonstraight road     | 0.872               | 0.857            | 0.849              | 0.860      |
|              | Slippery pavement    | 0.827               | 0.759            | 0.752              | 0.780      |
|              | Road without physical separation | 0.928   | 0.799            | 0.798              | 0.841      |
| Environment factors | Daytime            | 0.924               | 0.766            | 0.757              | 0.816      |
|              | Workday              | 0.830               | 0.692            | 0.717              | 0.746      |
|              | Bad weather          | 0.723               | 0.786            | 0.852              | 0.787      |
|              | Low visibility       | 0.752               | 0.775            | 0.773              | 0.767      |

Table 5: The weighted grey correlation degree of multivehicle traffic accidents’ factors.

| Categories   | Accident index       | Number of accidents | Number of deaths | Number of injuries | Mean value |
|--------------|----------------------|---------------------|------------------|--------------------|------------|
| Driver factors | Overspeed           | 0.847               | 0.669            | 0.709              | 0.741      |
|              | Overload             | 0.617               | 0.721            | 0.750              | 0.702      |
|              | Misoperation         | 0.637               | 0.721            | 0.750              | 0.619      |
|              | Lane use underserved | 0.814               | 0.659            | 0.641              | 0.705      |
|              | Fatigue driving      | 0.617               | 0.761            | 0.809              | 0.729      |
| Vehicle factors | Poor braking         | 0.828               | 0.665            | 0.701              | 0.731      |
|              | Commuter bus         | 0.900               | 0.727            | 0.783              | 0.803      |
|              | Normal bus           | 0.837               | 0.687            | 0.718              | 0.747      |
|              | Heavy track          | 0.794               | 0.779            | 0.807              | 0.793      |
|              | Cross the line moving | 0.873          | 0.658            | 0.707              | 0.746      |
|              | Parking              | 0.520               | 0.547            | 0.569              | 0.545      |
| Road factors | Located in the construction section | 0.639 | 0.706 | 0.754 | 0.677 |
|              | Located in the nonstraight sections | 0.735 | 0.730 | 0.756 | 0.715 |
|              | Road without physical separation | 0.924 | 0.691 | 0.748 | 0.772 |
| Environment factors | Daytime            | 0.775               | 0.908            | 0.807              | 0.914      |
|              | Workday              | 0.794               | 0.686            | 0.706              | 0.678      |
|              | Bad weather          | 0.820               | 0.746            | 0.725              | 0.715      |
|              | Low visibility       | 0.797               | 0.780            | 0.746              | 0.769      |
Figure 4: Critical factor set for the multivehicle traffic accident.

1. De: number of deaths
2. In: number of injuries
3. Os: overspeed or not
4. Ol: misoperation or not
5. Iu: mislane use or not
6. Fd: fatigue driving or not
7. Vt: commuter bus involved or not
8. Hv: heavy truck involved or not
9. Cl: vehicle driving status
10. Bb: vehicle safety status
11. Sr: pavement alignment
12. Ps: physical separation
13. Dt: accident occurring period
14. Wd: workday or not
15. Aw: weather condition

Figure 5: Preliminary learning result of single-vehicle traffic accident network.

Figure 6: Preliminary learning result of multivehicle traffic accident network.
variables (the number of deaths and the number of injuries) included. The factor variables are as follows:

(i) Direct factors contributing to the number of deaths: overspeed and vehicle type.

(ii) Indirect acting factors contributing to the number of deaths: road alignment, physical separation, weather condition, and accident period.

Figure 7: Final BN result of single-vehicle traffic accident, (a) factors contributing to the number of deaths and (b) factors contributing to the number of injuries.

Figure 8: Final BN result of multivehicle traffic accident, (a) factors contributing to the number of deaths and (b) factors contributing to the number of injuries.
pavement condition (wet or dry), accident period, weather condition, and visibility.

Indirect acting factors contributing to the number of injuries: road alignment, physical separation, visibility, and weather condition.

Besides, it can be found in Figure 7 that five-factor set sequences contribute to the number of deaths, and three-factor sequences contribute to the number of injuries. For example, one of the most extended sequences shown in Figure 7(a) is {accident occurring period ⟷ visibility ⟷ pavement alignment ⟷ physical separation ⟷ vehicle types ⟷ number of deaths}. One of the most extended sequences shown in Figure 7(b) is {weather condition ⟷ visibility ⟷ pavement alignment ⟷ physical separation ⟷ vehicle type ⟷ number of injuries}.

Figure 8 shows ten factors contributing to the number of deaths and five factors contributing to the number of injuries of multivehicle severe traffic accidents:

(i) Direct factors contributing to the number of deaths: mislane use, road alignment, and workday.

Direct factors contributing to the number of injuries: overload.

(ii) Indirect acting factors contributing to the number of deaths: bus, vehicle safety condition, physical separation, accident period, and weather condition.

Indirect acting factors contributing to the number of injuries: bus, heavy truck, road alignment, and weather condition.

Similarly, there are six factor sequences contributing to the number of deaths and three factor sequences contributing to the number of injuries in the multivehicle severe traffic accidents. The most extended sequence in Figure 8(a) is {accident occurring period ⟷ physical separation ⟷ vehicle safety status ⟷ mislane use ⟷ the number of deaths}. Besides, the most extended sequence in Figure 8(b) is {pavement alignment ⟷ heavy trunk involved or not ⟷ misoperation or not ⟷ the number of injuries}. Although the multivehicle traffic accidents usually lead to a more severe effect, the factor sequence of multivehicle traffic accidents is slightly shorter than single-vehicles', indicating that the multivehicle accidents can be prevented more quickly because of fewer causes.

4.3. BN Conditional Probability of Learning Result. The conditional probability of influencing factors of the single-vehicle and multivehicle traffic accident was calculated and shown in Tables 6 and 7, respectively.

It can be seen from Table 6 that overspeed, commuter bus involved, straight alignment, physical separation, slippery pavement, night, bad weather, and low visibility have significant impacts on single-vehicle accidents. As shown in Table 7, the significant factors contributing to multivehicle accidents are overload, mislane use, commuter bus involved, heavy truck involved, poor braking, straight alignment, no physical separation, night, weekend, and bad weather. The corresponding variable state is then selected to analyze the risk degree of each factor by the interval sorting theory of the BN network. The risk degree ranking of factors contributing to single-vehicle and multivehicle traffic accidents is shown in Tables 8 and 9.

As shown in Tables 8 and 9, different factors contribute to different degrees of severity in terms of the number of deaths and the number of injuries in single-vehicle and multivehicle traffic accidents. The results show that the risk factors that most significantly influence the number of deaths and injuries in single-vehicle accidents are bad weather and commuter bus. The risk factors that have the most significant influence on the number of deaths and injuries in multivehicle accidents are commuter bus and night.

4.4. Severity Ranking Result of Factors Combination. Since accidents result from multiple factors, it is necessary to study the probability distribution of the number of deaths and injuries under the combination of multiple factors based on the analysis of a single factor. As for the analysis of each factor’s effect on the severe traffic accident, the posterior probability of death and injury was deduced with BN’s interval theory. Then, the inherent logical relations among these factors were ranked by the severity degree in terms of the number of deaths and the number of injuries, as shown in Tables 10 and 11.

The study on factor sequences can reduce the accident damage and help safety managers propose effective measures. The key reasons for the enormous damage caused by bad weather conditions are overspeed, overload, and mislane use. Therefore, countermeasures of adverse weather conditions, reasonable control of vehicle speed, and proper lane use should be focused on to minimize severe traffic accidents.

4.5. Model Validation Test Result. Using the mathematical statistics, the conditional probability accuracy of this BN accident severity model is validated. The model’s efficiency can be tested by the MSE and RMSE calculated by the conditional probability of actual value and learning value shown in Table 12.

From Table 12, as for single-vehicle accidents, the model’s accuracy for the number of deaths is slightly lower than the number of injuries since the maximum absolute error is 0.0027, and the mean relative error is 0.3390. That is why the sample distribution between these two types is unbalanced. Hence, the crash severity for single-vehicle traffic accidents can be analyzed using BN model with a greater prediction accuracy when a more randomly distributed accident data sample is provided. As for multivehicle accidents, although the model’s accuracy for the number of deaths is slightly lower than the number of injuries, the model still meets the requirement of prediction. Hence, the BN model can be used to analyze the crash severity for multivehicle traffic accidents.
**Table 6:** The conditional probability of the number of deaths for the single-vehicle traffic accident.

| Factor variables | Result variables | Number of deaths | Number of injuries |
|------------------|------------------|------------------|-------------------|
|                  | 10 ≤ number of deaths ≤ 19 | 20 ≤ number of deaths ≤ 29 | Number of deaths ≥ 30 | 10 ≤ number of deaths ≤ 19 | 20 ≤ number of deaths ≤ 29 | Number of deaths ≥ 30 |
| Overspeed        | Yes | 0.5368 | 0.3880 | 0.0751 | 0.2022 | 0.6356 | 0.1622 |
|                  | No  | 0.3334 | 0.6265 | 0.0401 | 0.2045 | 0.6312 | 0.1643 |
| Commuter bus involved | Yes | 0.3996 | 0.9497 | 0.1016 | 0.1943 | 0.1992 | 0.0605 |
|                  | No  | 0.4756 | 0.3657 | 0.1587 | 0.6383 | 0.3037 | 0.0543 |
| Straight alignment | Yes | 0.5384 | 0.2224 | 0.2392 | 0.2009 | 0.6411 | 0.1579 |
|                  | No  | 0.5573 | 0.2133 | 0.2294 | 0.2049 | 0.6418 | 0.1533 |
| Physical separation | Yes | 0.4808 | 0.2240 | 0.2693 | 0.1882 | 0.6376 | 0.1742 |
|                  | No  | 0.5345 | 0.2499 | 0.2416 | 0.1986 | 0.6320 | 0.1694 |
| Slippery pavement | Yes | 0.5367 | 0.2226 | 0.2407 | 0.1964 | 0.6098 | 0.1937 |
|                  | No  | 0.5269 | 0.2264 | 0.2467 | 0.1991 | 0.6212 | 0.1796 |
| Occurring period | Day  | 0.5663 | 0.2108 | 0.2229 | 0.3567 | 0.4974 | 0.1459 |
|                  | Night | 0.4571 | 0.2588 | 0.2840 | 0.1341 | 0.4595 | 0.4064 |
| Poor weather | Yes | 0.4469 | 0.2583 | 0.2949 | 0.1890 | 0.6042 | 0.2068 |
|                  | No  | 0.5656 | 0.2117 | 0.2227 | 0.2254 | 0.7225 | 0.0521 |
| Visibility well | Yes | 0.5078 | 0.2363 | 0.2558 | 0.1889 | 0.6202 | 0.1910 |
|                  | No  | 0.5512 | 0.2173 | 0.2315 | 0.2088 | 0.6761 | 0.1150 |

**Table 7:** The conditional probability of the number of injuries for the single-vehicle traffic accident.

| Factor variables | Result variables | Number of deaths | Number of injuries |
|------------------|------------------|------------------|-------------------|
|                  | 10 ≤ number of deaths ≤ 19 | 20 ≤ number of deaths ≤ 29 | Number of deaths ≥ 30 | 10 ≤ number of deaths ≤ 19 | 20 ≤ number of deaths ≤ 29 | Number of deaths ≥ 30 |
| Overload         | Yes | 0.3442 | 0.3057 | 0.3502 | 0.3502 | 0.1787 | 0.6425 | 0.1787 |
|                  | No  | 0.3541 | 0.2998 | 0.3461 | 0.5439 | 0.1634 | 0.2415 |
| Mislane use      | Yes | 0.2687 | 0.4753 | 0.2560 | 0.5926 | 0.1659 | 0.2151 |
|                  | No  | 0.3321 | 0.4179 | 0.2500 | 0.5769 | 0.1680 | 0.2551 |
| Commuter bus involved | Yes | 0.3493 | 0.3029 | 0.3478 | 0.6114 | 0.1634 | 0.2252 |
|                  | No  | 0.3581 | 0.2980 | 0.3439 | 0.6118 | 0.1633 | 0.2249 |
| Heavy truck involved | Yes | 0.3630 | 0.2944 | 0.3425 | 0.5859 | 0.1668 | 0.2473 |
|                  | No  | 0.2914 | 0.3287 | 0.3799 | 0.5788 | 0.1677 | 0.2535 |
| Poor braking     | Yes | 0.3581 | 0.2977 | 0.3443 | 0.5825 | 0.1672 | 0.2502 |
|                  | No  | 0.3534 | 0.3000 | 0.3466 | 0.5769 | 0.1680 | 0.2551 |
| Straight alignment | Day  | 0.4052 | 0.4509 | 0.1439 | 0.5775 | 0.1679 | 0.2546 |
|                  | Night | 0.3656 | 0.6266 | 0.0078 | 0.5787 | 0.1676 | 0.2525 |
| Physical separation | Yes | 0.3641 | 0.2931 | 0.3427 | 0.5799 | 0.1683 | 0.2571 |
|                  | No  | 0.3602 | 0.2954 | 0.3444 | 0.5746 | 0.7225 | 0.0521 |
| Occurring period | Day  | 0.3646 | 0.2931 | 0.3424 | 0.5845 | 0.1670 | 0.2485 |
|                  | Night | 0.3608 | 0.2951 | 0.3441 | 0.5726 | 0.1686 | 0.2588 |
| Workday          | Yes | 0.3991 | 0.3481 | 0.2528 | 0.5805 | 0.1675 | 0.2520 |
|                  | No  | 0.2614 | 0.6209 | 0.1177 | 0.5784 | 0.1678 | 0.2538 |
| Poor weather     | Yes | 0.3678 | 0.2917 | 0.3405 | 0.5945 | 0.1656 | 0.2399 |
|                  | No  | 0.3637 | 0.2945 | 0.3418 | 0.5981 | 0.1652 | 0.2368 |
5. Conclusions and Discussions

Most previous researches have studied the relations between various factors and accident indexes from a particular perspective. This paper studies the critical factors contributing to severe traffic accidents in China from single vehicle and multiple vehicles with the BN crash severity model. From the case application result, the following conclusions can be drawn:

1. The direct factors contributing to the single-vehicle traffic accidents are commuter bus involved, mislane use, and night.

2. The commuter bus involved, overload, and night are the direct factors contributing to the multivehicle traffic accidents.

3. The severity ranking of single-vehicle traffic accidents is as follows:
   - Night → no physical separation → low braking → mislane use
   - Commuter bus involved → poor braking → mislane use
   - Poor weather → mislane use

4. The severity ranking of multivehicle traffic accidents is as follows:
   - Overload → overload → commuter bus involved
   - Mislane use → poor braking → mislane use
   - Commuter bus involved → poor braking → mislane use

5. The single-vehicle accident comparison result for the number of deaths shows that the maximum absolute error is 0.0059 and the maximum relative error is 1.00.

6. The single-vehicle accident comparison result for the number of injuries shows that the maximum absolute error is 0.0068 and the maximum relative error is 1.00.

7. The multivehicle accident comparison result for the number of deaths shows that the maximum absolute error is 0.0392 and the maximum relative error is 1.00.

8. The multivehicle accident comparison result for the number of injuries shows that the maximum absolute error is 0.0045 and the maximum relative error is 0.0129.
use, and night. As for the multivehicle traffic accidents, the direct factors are pavement alignment, mislane use, weekend, and overload.

(2) Under the influence of a single factor, the risk factors that have the most significant influence on the number of deaths and injuries in single-vehicle accidents are bad weather and commuter bus, while the risk factors that have the most significant influence on the number of deaths and injuries in multivehicle accidents are commuter bus and night. Single-vehicle accidents and multivehicle accidents involving commuter buses are more serious.

(3) There are different inherent hierarchical correlations among variables in different types of accidents. Overloading of the commuter bus may cause serious accidents. Bad braking of commuter buses on the improper lane is more likely to cause a serious accident. Single-vehicle accidents on separated highways are more likely to result in serious injuries. Under the conditions of low visibility, such as wet road and foggy days, the injuries caused by accidents are relatively low because of drivers’ high vigilance.

(4) One of the factor combination sequences for single-vehicle traffic accidents, namely, {bad weather ⟷ slippery pavement ⟷ overspeed}, can be speculated through the importance degree ranking of these critical factors. Moreover, the possible factors’ combination links of multivehicle traffic accident can be speculated as {night ⟷ nonphysical separation ⟷ poor braking ⟷ mislane use}. These links could provide some theoretical support for active precaution management of severe traffic accidents.

5.1. Limitation. Only several factor links were concluded in this study with a rough calculation of sample data collected from the investigation and disposal report of accident reports in the production process. The structure of the BN crash severity model was relatively simple, as only a few factors were analyzed. More accurate factors were needed to increase the complexity of the model’s structure, which could reveal factor coupling relation and explore the factors links more deeply. The improved grey correlation method and BN crash severity model were adopted to analyze the severity and contributing factors of severe traffic accidents.

This paper aims to construct a targeted model for traffic accidents to identify the contributing factors precisely. However, the data and sample distribution available are limited, which decreases the accuracy of BN model. This paper studies the crash severity and contributing factors from a superficial perspective, lacking the “drivers-vehicles-road” coupling analysis for traffic safety. Therefore, it recommended that the following research should focus more on the comprehensive factor analysis with the BN model constructed in this paper, such as “driving behavior + driving environment” analysis for road safety of the autodriving system, if the data from simulation experiments can be obtained.

Data Availability

The data of participants responding to the investigation and disposal report of accident reports are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the participants for their cooperation. This research was supported by the Natural Science Foundation of Zhejiang Province (no. LQ19E080003).

References

[1] D. Mohan, G. Tiwari, and S. Mukherjee, “Urban traffic safety assessment: a case study of six Indian cities,” IATSS Research, vol. 39, no. 2, pp. 95–101, 2016.
[2] E. Fitzgerald and B. Landfeldt, “Increasing road traffic throughput through dynamic traffic accident risk mitigation,” Journal of Transportation Technologies, vol. 5, no. 4, pp. 223–239, 2015.
[3] T. Zhihong, C. Jiejun, L. Qiong, Z. Jiyun, L. Xuezong, and Y. Yong, “The regional-scale atmospheric dispersion analysis and environmental radiation impact the hypothetical accident in Haiyang nuclear power plant,” Progress in Nuclear Energy, vol. 125, Article ID 103362, 2020.
[4] A. García-Ferrer, A. De-Juan, and P. Poncela, “Forecasting traffic accidents using disaggregated data,” International Journal of Forecasting, vol. 22, no. 2, pp. 203–222, 2006.
[5] J. L. Machado-León, J. De Oña, R. De Oña, L. Eboli, and G. Mazzulla, “Socio-economic and driving experience factors affecting drivers’ perceptions of traffic crash risk,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 37, no. Supplement C, pp. 41–51, 2016.
[6] J. Zhao and W. Deng, “Traffic accidents on expressways: new threat to China,” Traffic Injury Prevention, vol. 213, no. 13, pp. 230–238, 2012.
[7] E. N. Aidoo, R. Amoh-Gyimah, and W. Ackaah, “The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana,” Accident Analysis & Prevention, vol. 53, no. 53C, pp. 23–27, 2013.
[8] C. Gabaude, J.-C. Marquié, and F. Obriot-Claudel, “Self-regulatory driving behaviour in the elderly: relationships with aberrant driving behaviours and perceived abilities,” Le Travail Humain, vol. 73, no. 1, pp. 31–52, 2010.
[9] Z. Wei, Z. Shuangna, Z. Xiao, and Z. Tingsheng, “Identification of critical causes of construction accidents in China using a model based on system thinking and case analysis,” Safety Science, vol. 121, pp. 606–618, 2020.
[10] J. Hybiak, K. Jankowska, F. Machaj et al., “Reprogramming and transdifferentiation—two key processes for regenerative medicine,” European Journal of Pharmacology, vol. 882, Article ID 173202, 2020.
[11] Y. Wang, L. Li, L. Feng, and H. Peng, “Professional drivers’ views on risky driving behaviors and accident liability: a questionnaire survey in Xining, China,” Transportation Letters, vol. 6, no. 3, pp. 126–135, 2014.
[12] D. Parker, J. T. Reason, A. S. R. Manstead, and S. G. Stradling, "Driving errors, driving violations and accident involvement," Ergonomics, vol. 38, no. 5, pp. 1036–1048, 1995.

[13] G. Zhang, K. W. Yau, and G. Chen, "Risk factors associated with traffic violations and accident severity in China," Accident Analysis & Prevention, vol. 59, no. 10, pp. 18–25, 2013.

[14] B. Wang, C. Wu, L. Huang, L. Zhang, L. Kang, and K. Gao, "Prevention and control of major accidents (MAS) and particularly serious accidents (PSAs) in the industrial domain in China: current status, recent efforts and future prospects," Process Safety and Environmental Protection, vol. 117, pp. 254–266, 2018.

[15] D. Shinar, E. Schechtman, and R. Compton, "Self-reports of driving behavior in relationship to sex, age, education and income in the US adult driving population," Accident Analysis & Prevention, vol. 33, no. 1, pp. 111–116, 2001.

[16] D.-Y. Zhao and N. Bai-Sheng, "Statistical analysis of China's coal mine particularly serious accidents," Procedia Engineering, vol. 26, pp. 2213–2221, 2011.

[17] E. Yamamura, "Impact of formal and informal deterrents on driving behavior," The Journal of Socio-Economics, vol. 37, no. 6, pp. 2505–2512, 2008.

[18] S. Afqur, A. Melot, A. Ndiaye, E. Hammad, J.-L. Martin, and P.-J. Arnoux, "Descriptive analysis of the effect of back protector on the prevention of vertebral and thoracolumbar injuries in serious motorcycle accident," Accident Analysis & Prevention, vol. 135, Article ID 105331, 2020.

[19] L. J. Bellamy, M. Damen, H. J. Manuel et al., "Risk horoscopes: predicting the number and type of serious occupational accidents in The Netherlands for sectors and jobs," Reliability Engineering & System Safety, vol. 133, pp. 106–118, 2015.

[20] F. H. Hedlund and M. Madsen, "Incomplete understanding of biogas chemical hazards—serious gas poisoning accident while unloading food waste at biogas plant," Journal of Chemical Health and Safety, vol. 25, pp. 13–21, 2018.

[21] S. Reed and A. Morris, "Characteristics of fatal single-vehicle crashes in Europe," International Journal of Crashworthiness, vol. 17, no. 6, pp. 655–664, 2012.

[22] F. Abdat, S. Leclercq, X. Cuny, and C. Tissot, "Extracting recurrent scenarios from narrative texts using a Bayesian network: application to serious occupational accidents with movement disturbance," Accident Analysis & Prevention, vol. 70, pp. 155–166, 2014.

[23] Y. Chinniah, "Analysis and prevention of serious and fatal accidents related to moving parts of machinery," Safety Science, vol. 75, pp. 163–173, 2015.

[24] D. D. Clarke, P. Ward, C. Bartle, and W. Truman, "The role of motorcyclist and other driver behaviour in two types of serious accident in the UK," Accident Analysis & Prevention, vol. 39, no. 5, pp. 974–981, 2007.

[25] P. Strawiński and D. Celińska-Kopczyńska, "Occupational injury risk wage premium," Safety Science, vol. 118, pp. 337–344, 2019.

[26] D. S. Kirk, N. Cavalli, and N. Brazil, "The implications of ridehailing for risky driving and road accident injuries and fatalities," Social Science & Medicine, vol. 250, Article ID 112793, 2020.

[27] Y. Peng and L. Boyle, "Commercial driver factors in run-off-road crashes. Transportation research record," Journal of the Transportation Research Board, vol. 2281, no. 2281, pp. 128–132, 2012.

[28] A. Theoﬁlatos, D. Graham, and G. Yannis, "Factors affecting accident severity inside and outside urban areas in Greece," Traffic Injury Prevention, vol. 13, no. 5, pp. 458–467, 2012.

[29] J. Vahedi, A. Shariat Mohaymany, Z. Tabibi, and M. Mehdizadeh, "Aberrant driving behaviour, risk involvement, and their related factors among taxi drivers," International Journal of Environmental Research and Public Health, vol. 15, no. 8, p. 1626, 2018.

[30] B. Y. S. Jing, "Driving behaviour and influential factors study in Beijing," Journal of Transport Information and Safety, vol. 28, no. 2, pp. 114–119, 2010.

[31] Y. Jingxuan and Y. Yuntao, "The effect of image recognition traffic prediction method under deep learning and naive Bayes algorithm on freeway traffic safety," Image and Vision Computing, vol. 103, Article ID 103971, 2020.

[32] J. De Oña, R. O. Mujalli, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks," Accident Analysis & Prevention, vol. 43, no. 1, pp. 402–411, 2011.

[33] M. Hänninen, "Bayesian networks for maritime traffic accident prevention: benefits and challenges," Accident Analysis & Prevention, vol. 73, pp. 305–312, 2014.

[34] C. Cong, Z. Guohui, R. Tarefder, M. Jianming, W. Heng, and G. Hongzhi, "A multimonomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes," Accident Analysis & Prevention, vol. 80, pp. 76–88, 2015.

[35] M. B. Kamel, T. Sayed, and O. Ahmed, "Accounting for mediation in cyclist-vehicle crash models: a Bayesian mediation analysis approach," Accident Analysis & Prevention, vol. 131, pp. 122–130, 2019.

[36] Q. Zeng, H. Wen, H. Huang, and M. Abdel-Atty, "A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments," Accident Analysis & Prevention, vol. 100, pp. 37–43, 2017.

[37] W. L. Soro and D. Wayoro, "A Bayesian analysis of the impact of post-crash care on road mortality in Sub-Saharan African countries," IATSS Research, vol. 41, no. 3, pp. 140–146, 2017.

[38] C. Ma, J. Zhou, and D. Yang, "Causation analysis of hazardous material road transportation accidents based on the ordered logit regression model," International Journal of Environmental Research and Public Health, vol. 17, no. 4, pp. 1–25, 2020.

[39] S. Dong, J. Zhou, and C. Ma, "Design of a network optimization platform for the multivehicle transportation of hazardous materials," International Journal of Environmental Research and Public Health, vol. 17, no. 3, p. 1104, 2020.

[40] L. Guo, J. Zhou, S. Dong, and S. C. Zhang, "Analysis of urban road traffic accidents based on improved K-means algorithm," China J. Highway Trans, vol. 31, no. 4, pp. 270–279, 2018.

[41] L. Hanchen, T. Lili, and C. Xuewu, "Experimental study on hydrogen behavior and possible risk with different injection conditions in local compartment," Nuclear Engineering and Technology, vol. 52, no. 8, pp. 1650–1660, 2020.

[42] L. Xiaohei, T. Junqing, H. Xiaojiao, and W. Wei, "Assessing intercity multimodal choice behavior in a Touristy City: a factor analysis," Journal of Transport Geography, vol. 86, Article ID 102776, 2020.

[43] A. P. Afgahri, M. M. Haque, S. Washington, and T. Smyth, "Effects of globally obtained informative priors on bayesian factor analysis," International Journal of Environmental Research and Public Health, vol. 15, no. 8, pp. 1626–1636, 2018.