Article

Unsafe Behaviors Analysis of Sideswipe Collision on Urban Expressways Based on Bayesian Network

Huajing Ning 1,2, Yunyan Yu 1,* and Lu Bai 3,4

Abstract: The causes of crashes on urban expressways are mostly related to the unsafe behaviors of drivers before the crash. This study focuses on sideswipe collisions on urban expressways. Through real and visual crash data, 17 unsafe behaviors were identified for the analysis of sideswipe collisions on an urban expressway. The chains of high-risk and unsafe behaviors were then revealed to investigate the relationship between drivers’ unsafe behaviors and sideswipe collisions. A Bayesian network diagram of unsafe behaviors was used to obtain the correlation between unsafe behaviors and their influence. A topology diagram of unsafe behaviors was then constructed, and relational reasoning of typical behavioral chains was conducted. Finally, the unsafe behaviors and behavior chains that were likely to cause sideswipe collisions on the urban expressway were determined. The possibility of each behavior chain was quantified through the reasoning of variable structures constructed by the Bayesian network. The result shows that the significant influential single unsafe behavior leading to sideswipe collision on urban expressways was lane change without checking the rearview mirror or not scanning the road around and queue-jumping; moreover, based on unsafe behavior chains analysis, the most influential chains leading to sideswipe collision were: improper driving behavior in an emergency—failure to turn on signal when changing lanes—distracted and inattentive driving. Some safety precautions and countermeasures aimed at unsafe behaviors could be taken before the crash. The results of the study can be used to reduce the number of sideswipe collisions, thereby improving traffic safety on urban expressways.

Keywords: traffic casualty crash; traffic violation behavior; generalized additive model; multilevel model

1. Introduction

Vehicle crashes are a major concern globally. Each year, crashes lead to about 1.24 million deaths globally [1]. In 2020, 211,074 vehicle crashes in China led to 55,950 deaths and 21,442 injuries [2]. Sideswipe crashes are the second most frequent crash type on urban expressways, accounting for approximately 25.19% of all crashes; for rear-end crashes, the most common crash type, the proportion is 59.89% of all crashes [3]. Many factors influence the occurrence of sideswipe crashes, including drivers’ unsafe behaviors [4–7], traffic-flow conditions [7,8], speed [9,10], occupancy features [10,11], and roadway geometric features (e.g., lane width, shoulder width, and horizontal and vertical alignments) [12–14], are significant in their correlation with crash occurrence. Some external factors also affect the occurrence of crashes, for example, time of day, season, year [15,16], etc. Drivers’ unsafe behaviors are considered a major factor that contributes to crash occurrence [4–7]. Approximately 88% of crashes are related to drivers’ unsafe behaviors [17]. Therefore,
it is worth investigating the influence of drivers’ unsafe behaviors on the occurrence of crashes; the results can help illuminate the process of occurrence of crashes and reduce the occurrence of crashes.

Various data have been used for the analysis of unsafe behavior in previous studies [18–23]. Some data used in previous research were derived from questionnaires [24] or online questionnaires [25]; some were from professional agencies, such as the Highway Safety and Information System in the United States [26] and the airlines aviation industry in Nigeria [27], while others were from video event recorders [28]. For example, Van and Calvert [29] studied distraction by means of simulation, which provided some ideas and methods for the analysis of unsafe behaviors. Bell et al. [30] used a more advanced in-vehicle monitoring system, which was helpful to observe various behaviors of drivers. In addition to the abovementioned data, crash data can be obtained from government or official departments; however, there has been little research on drivers’ unsafe behaviors using official crash data, and moreover, the process of the occurrence of crashes still needs to be presented.

Many methods have been used for studying the relationship between traffic behaviors and crashes, such as the crash tree method [31–33] classification decision making [34–37], BP neural network [38–40], Petri net model [41,42], rough set [43,44], structural equation [45–47], etc. For the above analysis method, analyzing the cause of a crash is a strong point, but applying the possibility of the cause of an accident is a weak point; moreover, it is difficult to get accurate results when the sample data are sparse. The crash video is difficult to obtain and belongs to the category of small samples, so the Bayesian method is more suitable.

The Bayesian network, known as the reliability network, is a statistical reasoning method developed based on the Bayesian decision method in multivariate statistical analysis technology. It is widely used in the field of traffic safety for crash analysis and prevention by combining qualitative and quantitative methods [48–53]. Chen et al. [54] analyzed the complex coupling relations among accident factors contributing to the single-vehicle and multivehicle traffic accidents with the Bayesian network (BN) crash severity model. Ye et al. [55] analyzed the factors affecting the LOS (level of service) of non-motorized vehicles crossing the signalized intersection and aimed to construct an appropriate method to evaluate the LOS. Chen et al. [56] combined the Bayesian method with multiple regression to study the severity analysis of driver casualties in rear-end crashes. The results demonstrated Bayesian method can effectively discover the interdependency among variables. At present, research on traffic safety problems using the Bayesian method is mostly carried out from the perspective of various influencing factors of collisions. There is no special research on drivers’ behavior factors using the Bayesian method using crash videos; besides, this paper focused on the influence of drivers’ behavior on the sideswipe collisions on urban expressways, not only focusing on unsafe behaviors, but also the chain combination among unsafe behaviors.

Real and visual crash data have been rarely used in studies associated with the relationship between crashes and drivers’ unsafe behaviors. This study makes some contributions to the relationship between crashes and drivers’ unsafe behaviors before the crash based on real and visual crash data. Crash and drivers’ unsafe behavior data were extracted from the video of cameras installed in the car. The research results can help drivers correct their unsafe behaviors, so that the crash occurrence and crash rate can be reduced.

This paper was organized as follows. In Section 2, we extracted the data on crashes and drivers’ unsafe behaviors from the videos and presented a clear definition of each of the unsafe behaviors. In Section 3, we showed the methodology used to conduct the analysis. In Section 4, we discussed the results of the model in detail. The research conclusion was presented in Section 5.
2. Data Sources

Crash videos were obtained from the public security traffic police department of Hefei in China. The crash videos came from the driving data recorder installed in the vehicle by the driver himself, which was predominantly used to record the driving process. Once a collision occurred, and the crash result was disputed, it was convenient for the traffic police to determine the responsibility of the collision. All video data were the natural driving data of the driver in the normal driving process, excluding purposeful video selection. Crash videos were selected based on certain requirements. According to the crash video, the selected crashes occurred in the main road and ramp area of several urban expressways of Hefei during the period from 2016 to 2020. Selected crash videos could reflect the whole process of the crash, including drivers’ unsafe behaviors before the crash and the occurrence of the crash. Images in the video were clear and effective, and above all, they could identify unsafe behaviors during the whole process of a crash. For example, for a sideswipe collision, the research process is the period from the moment when a vehicle shows the lane change trend to the moment when a crash occurs, which is an independent research sample. For non-crash videos, the research sample starts when a vehicle has a tendency to change lanes, and it ends when a sideswipe collision is likely to occur, that is, the moment when the transverse distance between vehicles is minimum.

We selected 605 videos, including 70 sideswipe collision videos and 535 non-crash videos. The sideswipe collision videos represented videos in which drivers’ unsafe behaviors were clearly seen and led to a sideswipe collision. The non-crash videos represented videos in which drivers’ unsafe behaviors were clearly seen but did not lead to a collision. Members of the research team spent a year sorting videos related to unsafe behaviors. One of the advantages of the crash video was that we can watch it repeatedly to check the correct parameter values. The data sets in this paper were checked repeatedly to ensure the correctness and rationality of the analysis.

2.1. Definition of Unsafe Behaviors

According to the dynamic process of the crash shown in the video, 17 unsafe behaviors were identified in this study. Descriptions of unsafe behavior are shown in Table 1. The 17 unsafe behaviors are identified as follows.

Table 1. Unsafe behavior description.

<table>
<thead>
<tr>
<th>Number</th>
<th>Unsafe Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Speeding</td>
</tr>
<tr>
<td>C2</td>
<td>Improper parking</td>
</tr>
<tr>
<td>C3</td>
<td>Straddling lanes without changing lanes</td>
</tr>
<tr>
<td>C4</td>
<td>Dangerous driving</td>
</tr>
<tr>
<td>C5</td>
<td>Continuous changing more than one lane at a time</td>
</tr>
<tr>
<td>C6</td>
<td>Driving into a forbidden area</td>
</tr>
<tr>
<td>C7</td>
<td>Close following</td>
</tr>
<tr>
<td>C8</td>
<td>Unsafe passing</td>
</tr>
<tr>
<td>C9</td>
<td>Unsafe merging</td>
</tr>
<tr>
<td>C10</td>
<td>Failure to turn on hazard warning lights</td>
</tr>
<tr>
<td>C11</td>
<td>Failure to turn on signal when changing lanes</td>
</tr>
<tr>
<td>C12</td>
<td>Queue-jumping</td>
</tr>
<tr>
<td>C13</td>
<td>Distracted and inattentive driving</td>
</tr>
<tr>
<td>C14</td>
<td>Tailgating and forcing another vehicle to stop</td>
</tr>
<tr>
<td>C15</td>
<td>Failure to reduce speed in time</td>
</tr>
<tr>
<td>C16</td>
<td>Improper driving behavior in an emergency</td>
</tr>
<tr>
<td>C17</td>
<td>Lane change without checking the rearview mirror or not scanning the road around</td>
</tr>
</tbody>
</table>

C1, speeding. The design speed of the urban expressway is 80 km/h [57]. A vehicle is identified as speeding if its operating speed exceeds 80 km/h.
C2, improper parking. Improper parking is identified as parking unrealistically in places where stopping/parking is prohibited.

C3, straddling lanes without changing lanes. Based on the Road Traffic Safety Law of the People’s Republic of China, it is illegal to drive on traffic markings or the shoulders of the road for a long time. In a non-urgent situation, the limit of the time to collision (TTC), which is widely used to evaluate safety for sideswipe collisions, is 5.5 s [58]. When the span exceeds 5.5 s, in a non-urgent situation, the vehicle is going straight without lane-changing behavior, and the turn signal is not turned on; this behavior is identified as straddling lanes.

C4, dangerous driving. Drivers drive in a dangerous way. For example, they speed up to prevent another vehicle from overtaking, flash their high beams inappropriately, honk inappropriately, or make rude gestures to other drivers [59].

C5, continuously changing more than one lane at a time. It is illegal to change lanes more than one lane at a time.

C6, driving into a forbidden area. Vehicles are not allowed to drive into some traffic channelized zones or emergency vehicle lanes in the scope of urban expressways.

C7, close following. Under uncongested road conditions, this condition applies when a vehicle follows another vehicle with a distance of less than a vehicle length, and it lasts for more than 5.5 [59].

C8, unsafe passing. In the off-ramp driving process, a vehicle accelerates off the main line and into the ramp suddenly, slams on the brakes, or backs up.

C9, unsafe merging. In the on-ramp driving process, a driver fails to observe the surroundings to make sure it is safe for merging.

C10, failure to turn on hazard warning lights. A driver fails to turn on the hazard warning lights in emergency situations, for example, when the visibility is less than 100 m, the vehicle malfunctions, a crash occurs, or the vehicle temporarily parks on the roadside.

C11, failure to turn on signal when changing lanes. A driver fails to turn on the turn signal in advance when he makes a turn, changes a lane, prepares to overtake, or leaves a parking place.

C12, queue-jumping. To define the queue-jumping behavior, the following rules should be met. The initial horizontal distance between a lane-changing vehicle and a straight-moving vehicle is less than 2.2 m. The maximum lateral acceleration of the straight-moving vehicle is less than 0.07 g, and the lane offset is less than 1.0 m. The maximum length of lane-change process is no more than 75 m. The velocity of both lane-changing and straight-moving vehicles should be more than 1 m/s [59].

C13, distracted and inattentive driving. Distracted and inattentive driving refers to behavior leading to missed observations of some kind, which in turn leads to a critical event of ‘timing’ (premature, late action, or no action) or (incorrect) ‘direction’. When the driver lacks the motivation to carry out their task in the best way possible, an object or sequence of events diverts the driver’s attention, or the driver is used to ascertain the environment makes it difficult to discover changes [60].

C14, tailgating and forcing another vehicle to stop. A vehicle driver tailgates another vehicle and forces it to stop and get out of the way.

C15, failure to reduce speed in time. Accounting for possible unfavorable driving situation, the vehicle maximum deceleration is set for 10 m/s² and 8 m/s² for dry and wet roads, respectively. When the vehicle starts to decelerate until it stops, the maximum deceleration does not reach the value [61]. The vehicle is considered to fails to reduce speed in time.

C16, improper driving behavior in an emergency. For example, a driver improperly operates the vehicle when a vehicle at front changes a lane, which leads to more vehicles being involved in a crash.

C17, lane change without checking the rearview mirror or not scanning the road around. Before changing a lane, the driver should observe the situation from the rearview mirror in advance to understand the situation on both sides of the car and behind. Making
lane changes without checking the rearview mirror or not scanning the road around is an unsafe behavior.

2.2. The Analysis of Unsafe Behavior

The analysis data included 605 videos, including 70 sideswipe collision videos and 535 non-crash videos. There were 179 unsafe behaviors in the 70 crash videos; moreover, there were 663 unsafe behaviors in the 535 non-crash videos.

In all the videos of sideswipe collision, the average number of unsafe behaviors in crash videos was 2.56. The distribution of unsafe behaviors was obtained by statistical analysis of video samples, as shown in Figure 1. Most sideswipe collisions involve two or three unsafe behaviors, which account for 71.4% of all collisions. The sideswipe collisions involving only one type of unsafe behavior account for 14.3% of all collisions. The proportion was the same for those involving more than four types of unsafe behaviors. The results indicate that a sideswipe collision is usually caused by two or three types of unsafe behaviors. Therefore, this study identifies a behavior chain composed of two or three types of unsafe behaviors.

Figure 1. The distribution diagram of unsafe behaviors on crash videos.

In all the videos of non-crash collision, the average number of unsafe behaviors in non-crash videos was 1.31. The distribution of unsafe behaviors on non-crash is shown in Figure 2. The majority of non-crash incidents involve one or two types of unsafe behaviors, which account for 97.7% of all non-crash incidents. Only one type of unsafe behavior has the largest proportion, which accounts for up to 79.4% of all non-crash incidents. Incidents involving more than two types of unsafe behaviors account for merely 2.2%. The results indicate that the probability of occurrence of sideswipe collisions decreases with the decrease in the number of unsafe behaviors.

Figure 2. The distribution diagram of unsafe behaviors on non-crash videos.
3. Method

This study explores the influence of various unsafe behaviors on the occurrence of sideswipe collisions on the urban expressway based on the Bayesian network. The occurrence of sideswipe collisions was considered to be the target variable, various unsafe behaviors were considered to be factors. The probability graphical models (PGMs) were determined. Every variable was a node of PGMs, expressed by binary variables. The aim of Bayesian network modeling is to examine the probabilistic conditional dependency between the nodes, which can be theoretically specified or identified with a data-driven exploration [62].

The modeling process of Bayesian networks (BN) consists of two important parts: structure learning and parameter learning [63]. Structure learning is the process of discovering the graph relationship between variables. The main methods of structural learning are the Tree Augmented Naive (TAN) and Markov methods [64,65]. Parameter learning is the process of determining the quantitative relationship between variables. The main methods of parameter learning methods are the Bayesian method and the Maximum Likelihood (ML) approach [66]. Two types of BN models were built for crash analysis. Model 1 uses the structural learning method of TAN, and the parameter learning method is the ML approach. Model 2 uses the structural learning method of Markov, and the parameter learning method is the ML approach.

Take the TAN as an example, explain the steps of structure learning algorithm [63] are as follows:

1. Calculate the conditional mutual information of all input variables \(X_i\) and \(X_j\). The conditional mutual information is between 0 and 1, 0 means independent variables and no correlation; if the interaction information relationship is strong, it tends to 1.

\[
I(X_i; X_j|Y) = \sum_{m,n,k} P(x_i^m, x_j^n, y_n) \ln \frac{P(x_i^m, x_j^n|y_n)}{P(x_i^m|y_n)P(x_j^n|y_n)}
\] (1)

2. Find the variable with the maximum interaction information for each variable and connect it with an undirected arc.

3. Transform an undirected arc into a directed arc.

4. Output variables are connected to all input variables.

In the parameter learning of Bayesian method, the variables of each node are binary variables. Each parameter \(\theta\) in the parameter set of nodes is the probability of “success”, and the prior probability distribution of parameters \(\theta\) should choose the binomial conjugate distribution. Using sample data to modify prior probability, the final value \(\theta\) is the expectation of parameter \(\theta\) posterior distribution.

Through the constructed Bayesian network, the inference of relevant nodes can be conducted, such as cause-related reasoning, predictive reasoning, and diagnostic reasoning [67].

Cause-related reasoning refers to the relationship between different nodes that produce the same result; this method focuses on the influence relationship and conduction path between variables; it can be used to sort out the relationship between nodes and analyze the mechanism of traffic crashes.

Diagnostic reasoning deduces the cause from the conclusion. That is, given the results, according to the reasoning of Bayesian network, all of the causes and the paths of variables with greater influence can be found; this reasoning is often used in crash mechanisms to diagnose the cause of the crash, and this method can be used to analyze the causes of traffic crashes.

Predictive reasoning makes conclusions from causes. According to the situation of a known node occurrence, the impact of a subsequent node occurrence can be predicted, along with the impact of the overall crash; thus, it is often called predictive reasoning; this method can be used to predict the overall situation of traffic crashes.
Based on the Bayesian network, the correlation among unsafe behaviors that led to sideswipe collisions on urban expressways was sorted. The various behavior chains that caused crashes were constituted and quantified, which could help to conduct cause-related reasoning, predictive reasoning, and diagnostic reasoning of side collision. It yielded a more comprehensive understanding to grasp the crash mechanism, influence factors, behavioral transmission path, and combined risk.

According to Bayesian network analysis, we could obtain the confusion matrix for the analysis target, as shown in Table 2. Based on the confusion matrix, several metrics could be used to evaluate the effectiveness of different models. According to the parameter information from the table, we could get the overall accuracy metric from Equation (2). The false alarm rate could be obtained from Equation (3). The precision was obtained through Equation (4) and the recall was obtained through Equation (5) [56].

\[
\text{Overall accuracy} = \frac{T_{\text{crash}} + T_{\text{non\_crash}}}{T_{\text{crash}} + F_{\text{crash}} + T_{\text{non\_crash}} + F_{\text{non\_crash}}} \tag{2}
\]

\[
\text{The false alarm rate} = \frac{T_{\text{non\_crash}}}{F_{\text{crash}} + T_{\text{non\_crash}}} \tag{3}
\]

\[
\text{Precision} = \frac{T_{\text{crash}}}{T_{\text{crash}} + F_{\text{crash}}} \tag{4}
\]

\[
\text{Recall} = \frac{T_{\text{crash}}}{T_{\text{crash}} + F_{\text{non\_crash}}} \tag{5}
\]

Table 2. Confusion matrix.

<table>
<thead>
<tr>
<th>Predicted Crashes</th>
<th>Predicted Non-Crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real crashes</td>
<td>(T_{\text{crash}})</td>
</tr>
<tr>
<td>(F_{\text{crash}})</td>
<td>(T_{\text{non_crash}})</td>
</tr>
</tbody>
</table>

4. Analysis
4.1. The Model Training and Evaluation

A total of 605 videos were observed. Of the 605 videos, a training subset of 484 videos and a testing subset of 121 videos were assigned randomly for the training and evaluation of BN models. Both training and testing subsets included crash data and non-crash data.

Based on the 484 videos from which data were collected on the urban expressway to test, two Bayesian networks models were developed. 17 unsafe behaviors were the influencing factors, and whether sideswipe collisions occurred was the output result. According to the comparison between the prediction results of the BN model and the statistical calculation results of the original testing datasets, the learning error of the BN model could be obtained to evaluate the accuracy of the model.

The analysis result is shown in Table 3. The overall accuracy of both models is above 0.9, and the overall accuracy of the first model reaches up to 0.958. By comparing the accuracy of the two models, it can be found that model 1 performs better than model 2 in all indicators; thus, model 1 has better applicability for analysis to sideswipe collision on urban expressways.

Table 3. Results of two Bayesian networks models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall Accuracy</th>
<th>False Alarm Rate</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 to train</td>
<td>0.958</td>
<td>0.974</td>
<td>0.813</td>
<td>0.842</td>
</tr>
<tr>
<td>Model 2 to train</td>
<td>0.946</td>
<td>0.970</td>
<td>0.772</td>
<td>0.772</td>
</tr>
</tbody>
</table>

Here, based on 121 samples in the testing subset, we used model 1 to predict the collision occurrence of each case. As shown in Table 4, model 1 performs better in prediction,
and the overall accuracy rate is 0.983, which meets the prediction accuracy. Therefore, the proposed BN model can be used to analyze the occurrence of sideswipe collisions by unsafe behaviors.

### Table 4. Results of various Bayesian networks models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall Accuracy</th>
<th>False Alarm Rate</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 to test</td>
<td>0.983</td>
<td>0.991</td>
<td>0.923</td>
<td>0.923</td>
</tr>
</tbody>
</table>

#### 4.2. Identification Results of Single Unsafe Behavior

Analysis of unsafe behaviors and Bayesian networking modeling was conducted using SPSS Modeler 18.0. The correlation of each unsafe behavior and sideswipe collision is shown in Figure 3. Although most sideswipe collisions occur in the form of a collision chain, the single unsafe behavior has a certain influence on sideswipe collisions, and the influence cannot be ignored. Of all the single factors, C17 (lane change without checking the rearview mirror or not scanning the road around), C12 (queue-jumping), C8 (unsafe passing), C16 (improper driving behavior in an emergency), C11 (failure to turn on signal when changing lanes), and C9 (unsafe merging) affect the occurrence of sideswipe collision relatively obviously. And other unsafe behaviors have a certain influence on the occurrence of sideswipe collision on urban expressways, and their impacts are smaller than those of the top six unsafe behaviors. Reducing every unsafe behavior will play a positive role in reducing collisions; thus, regulating drivers’ driving behaviors is particularly important in avoiding crashes.

#### 4.3. Cause-Related Reasoning

Figure 4 shows the correlation between the variables. The sideswipe collision is the parent node of all nodes. If there is a logical association relationship between other nodes, the nodes are connected by directed lines. If nodes are independent, there are no lines to connect. The variables mostly interact with each other instead of remaining isolated, which proves that crashes are mostly caused by the joint action of multiple behavior variables. The enumeration method is used to tease out every unsafe behavior chain in the Bayesian network, which may induce sideswipe collision on urban expressways.

![Figure 3. The variable importance bar chart.](image)

![Figure 4. The correlation between the variables.](image)
Figure 4. All of the unsafe behaviors in Bayesian network. C1—speeding; C2—improper parking; C3—straddling lanes without changing lanes; C4—dangerous driving; C5—continuous changing more than one lane at a time; C6—driving into a forbidden area; C7—close following; C8—unsafe passing; C9—unsafe merging; C10—failure to turn on hazard warning lights; C11—failure to turn on signal when changing lanes; C12—queue-jumping; C13—distracted and inattentive driving; C14—tailgating and forcing another vehicle to stop; C15—failure to reduce speed in time; C16—improper driving behavior in an emergency; C17—lane change without checking the rearview mirror or not scanning the road around; A18—sideswipe collision.

Several nodes can be selected in the diagram to explore the relationship among variables. According to Figure 5, C4 (dangerous driving), C11 (failure to turn on signal when changing lanes), C14 (tailgating and forcing another vehicle to stop), C17 (lane change without checking the rearview mirror or not scanning the road around), and A18 (sideswipe collision) are composed of 4 chains of unsafe behaviors, and three in four chains are two unsafe behaviors; another chain is composed of 3 unsafe behaviors. Take one chain consisting of 2 unsafe behaviors as an example: C11 (failure to turn on signal when changing lanes), C17 (lane change without checking the rearview mirror or not scanning the road around) and A18 (sideswipe collision) compose a chain, causing a sideswipe collision together. When the number of unsafe behaviors increases, the correlation between variables is strengthened. C17 (lane change without checking the rearview mirror or not scanning the road around) is the parent node of C11 (failure to turn on signal when changing lanes). Whether C11 (failure to turn on signal when changing lanes) occurs is not entirely dependent on this one unsafe behavior, but also affected by C17 (lane change without checking the rearview mirror or not scanning the road around). In addition, the influence of C11 (failure to turn on signal when changing lanes) on the occurrence of sideswipe collision is also affected by C17 (lane change without checking the rearview mirror or not scanning the road around). In other words, C11 (failure to turn on signal when changing lanes) is important, but it also depends on whether there are errors in C17 (lane change without checking the rearview mirror or not scanning the road around).
4.4. Diagnosis Reasoning

Diagnostic reasoning was used to determine the chains of behaviors that had the greatest impact. Using Netica software, the unsafe behaviors of sideswipe collision and their probability values were used to construct a topology diagram. We identified each chain of behaviors in Bayesian networks and eliminated some unreasonable behavior chains. After removing unreasonable unsafe behaviors, 16 behavior chains containing 2 unsafe behaviors and 11 behavior chains containing 3 unsafe behaviors were retained. Two cases of diagnosis reasoning were made based on the chain reaction of unsafe behaviors. One case was the chain of unsafe behaviors that had the greatest influence when a sideswipe collision was confirmed to occur; the other was the chain of unsafe behavior that was less likely to appear when a sideswipe collision was determined not to occur.

When a sideswipe collision occurs, according to the probability values shown in Figure 6, the top unsafe behaviors in the chain are C17 (lane change without checking the rearview mirror or not scanning the road around) and C11 (failure to turn on signal when changing lanes). The reason a driver forgetting to signal when changing lanes has a great influence on the occurrence of sideswipe collisions is that a driver forgetting to signal when changing lanes and other unsafe behaviors work together. When a driver forgets to signal when changing lanes (failure to turn on signal when changing lanes = yes), and changes lane without checking the rearview mirror or not scanning the road around (lane change without checking the rearview mirror or not scanning the road around = yes), the combination of these two unsafe behaviors has a high probability of occurrence (probability > 50%). If the sideswipe collision is confirmed to occur, the most likely unsafe behavior chain is sideswipe collision, lane change without checking the rearview mirror or not scanning the road around, and failure to turn on signal when changing lanes.

Figure 6. The topology of sideswipe collision occurrence. C1—speeding; C2—improper parking; C3—straddling lanes without changing lanes; C4—dangerous driving; C5—continuous changing more than one lane at a time; C6—driving into a forbidden area; C7—close following; C8—unsafe passing; C9—unsafe merging; C10—failure to turn on hazard warning lights; C11—failure to turn on signal when changing lanes; C12—queue-jumping; C13—distracted and inattentive driving; C14—tailgating and forcing another vehicle to stop; C15—failure to reduce speed in time; C16—improper driving behavior in an emergency; C17—lane change without checking the rearview mirror or not scanning the road around; A18—sideswipe collision.

When a sideswipe collision is determined not to occur, according to the probability values shown in Figure 7, C1 (speeding), C12 (queue-jumping), and C5 (continuous
changing more than one lane at a time) are at a low risk. In this situation, C1 (speeding) and C12 (queue-jumping) almost never appear in the same crash; thus, chain composed of these three kinds of unsafe behaviors is extremely unlikely to occur. In other words, when driving at a normal speed, without straddling lanes without changing lanes and queue-jumping behavior, the probability of sideswipe collision crashes is low.

In Figure 7, C11 (failure to turn on signal when changing lanes) should been paid attention to, that is because the percentage of C11 (failure to turn on signal when changing lanes) occupies a relatively major proportion. We delve into the causes and find the source of this situation. The occurrence of this behavior will not necessarily lead to the occurrence of sideswipe collisions. The occurrence of collisions is not only related to the driver’s own unsafe behaviors, but also affected by the surrounding vehicles and the surrounding traffic environment; it means the C11 (failure to turn on signal when changing lanes) is a very common phenomenon, a behavioral action that drivers often ignore while driving. It does not necessarily cause collisions when it appears alone; however, when combined with other unsafe behaviors, it is one of the factors promoting the occurrence of collisions. Therefore, the driver should develop good driving habits, not ignore every unsafe behavior, avoid the unsafe behavior chain, has a significant impact on reducing collisions.

4.5. Predictive Inference

Predictive reasoning could be used to determine the probability of each behavior chain. According to each chain, we could predict the possible value when a variety of unsafe behaviors combination (2 or 3 unsafe behaviors) were determined to occur. It is convenient to sort out all the possibilities, and the top seven chains (probability > 0.5) are shown in Table 5; however, the chain of unsafe behaviors in actual crashes is also dominated by two or three unsafe behaviors, while more than four unsafe behaviors do not often appear in actual crashes, so our data are in good agreement with actual crashes.
Table 5. The probability chain of unsafe behaviors.

<table>
<thead>
<tr>
<th>Number</th>
<th>The Chain of Unsafe Behavior</th>
<th>The Probability of Sideswipe Collision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>improper driving behavior in an emergency = yes; failure to turn on signal</td>
<td>0.977</td>
</tr>
<tr>
<td></td>
<td>when changing lanes = yes; distracted and inattentive driving = yes;</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>improper driving behavior in an emergency = yes; failure to turn on signal</td>
<td>0.800</td>
</tr>
<tr>
<td></td>
<td>when changing lanes = yes;</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>improper driving behavior in an emergency = yes; dangerous driving = yes;</td>
<td>0.705</td>
</tr>
<tr>
<td></td>
<td>failure to turn on signal when changing lanes = yes;</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>improper driving behavior in an emergency = yes;</td>
<td>0.657</td>
</tr>
<tr>
<td></td>
<td>speeding = yes; queue-jumping = yes; straddling lanes without changing lanes = yes;</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>speeding = yes; queue-jumping = yes;</td>
<td>0.635</td>
</tr>
<tr>
<td>6</td>
<td>improper driving behavior in an emergency = yes;</td>
<td>0.635</td>
</tr>
</tbody>
</table>

With the increase of unsafe behaviors, the possibility of crashes increases. When two unsafe behaviors of C16 (improper driving behavior in an emergency) and C11 (failure to turn on signal when changing lanes) are confirmed to occur, the percentage of crashes is 0.8. Based on this, C16 (improper driving behavior in an emergency) and C11 (failure to turn on signal when changing lanes) with C13 (distracted and inattentive driving), the crash percentage is 0.977; this is because C16 (improper driving behavior in an emergency) is accompanied by C11 (failure to turn on signal when changing lanes). Both of these behaviors are caused by not obeying the traffic rules in the process of changing lanes; this kind of mistake combination is common on urban expressways. When C16 (improper driving behavior in an emergency), C11 (failure to turn on signal when changing lanes) is combined with C13 (distracted and inattentive driving), the probability of sideswipe collision is high (the probability is more than 0.95).

5. Summary and Conclusions

Sideswipe collisions on urban expressways occur frequently, and unsafe behaviors are key factors related to crashes. If unsafe behaviors are controlled, crashes can be avoided, which is an active risk aversion strategy that is effective for crash control. In this study, 17 common unsafe behaviors of urban expressway crashes were identified. Bayesian network modeling was used to identify the correlation between variables and constitute the chains of unsafe behaviors. Based on the correlation and probability value of variables obtained after modeling, the influence effect was determined. The cause of correlation reasoning, prediction reasoning, and diagnostic reasoning of sideswipe collisions on the urban expressway were conducted. Through the analysis of common sideswipe collisions on urban expressways, including the situation of only one unsafe behavior and a chain of many unsafe behaviors, which was analyzed separately, the corresponding risk assessment of various cases could be obtained. By controlling and avoiding high-probability and high-risk behaviors, crashes can be avoided. Only by constantly improving people’s traffic safety awareness and strengthening traffic safety education, all kinds of unsafe behaviors will be gradually reduced and the crash rate will be reduced accordingly, which can effectively improve traffic safety.

We drew several key conclusions from the study. Expressway crashes are associated with multiple driver unsafe behaviors. Unsafe behaviors averagely exist in each urban expressway sideswipe collision at a rate of 2.56, and the unsafe behaviors in most sideswipe collisions are 2 or 3. Risk unsafe behaviors are found to have varying degrees of effects on crashes on urban expressways; those unsafe behavior chains can reveal hidden risks of drivers’ behaviors. When unsafe behaviors increase, the probability of sideswipe collision also increases, which is related to the frequency and combination of unsafe behaviors in actual crashes.

Using the powerful reasoning of the Bayesian network, we integrated unsafe behaviors into our analysis to make inferences about crash risk chains. In future work, we will look
deeper into the collaborative factors between Bayesian network knowledge inferences and the dynamic knowledge graphs. Other types of crash behavior chains will also be discussed in depth in the future.

**Author Contributions:** Conceptualization, H.N. and Y.Y.; methodology, H.N.; software, H.N.; formal analysis, H.N.; investigation, L.B.; writing—original draft preparation, H.N.; writing—review and editing, H.N. and L.B.; supervision, Y.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was not applicable for studies not involving humans or animals.

**Informed Consent Statement:** The data were obtained from the public security traffic police department of Hefei in China.

**Acknowledgments:** We are grateful to anonymous referees and the editor for very useful comments and suggestions, which greatly improved this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


40. Yang, W.; Zhang, X.; Lei, Q.; Cheng, X. Research on longitudinal active collision avoidance of autonomous emergency braking pedestrian system (AEB-P). *Sensors* **2019**, *19*, 4671. [CrossRef]


