WILEY | Hindawi

Research Article

Accident Diagnosis and Evaluation System in Parking Lots Using Multisource Data Based on Bayesian Networks

Yijing Zhang [ID],[1] Zhan Zhang [ID],[2] Zhenyu Wang,[3] Tongtong Lyu,[1] Xianing Wang,[1] and Linjun Lu[1]

[1]*School of Naval Architecture Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[2]*School of Design, Shanghai Jiao Tong University, Shanghai 200240, China*
[3]*Center for Urban Transportation Research, University of South Florida, Tampa 33620, USA*

Correspondence should be addressed to Zhan Zhang; zhanzhang@sjtu.edu.cn

Parking lots have many complex structures, diverse functions, and plentiful elements. The frequent flow of vehicles with narrow and dim spaces increases the probability of various traffic accidents. Due to the low severity and lack of relevant data, there is limited understanding of safety analyses for parking lot accidents. This study integrates multisource data to establish a Bayesian diagnostic model for parking lot accidents. The mutual information method is used to screen the possible influencing factors before modeling to reduce the subjectivity of Bayesian networks. Studying the cause and effect analysis of accidents provides diagnosis and prediction for property damage and event causes. This provides valuable correlation information between factors and accident characteristics, as well as consequences under the influence of multiple factor chains. As the developed model has good accuracy, this study proposes a parking lot safety evaluation system with a library of countermeasures based on the model results to ensure rigorous conclusions. The combination with ITS technology gives the system high scalability and adaptability in multiple scenarios.

1. Introduction

Automobiles have brought convenience and comfort to people's lives but have also caused many significant urban problems, including increased parking demand and parking space constraints [1]. Compared with on-street parking, parking lots have become an important facility to effectively relieve parking pressure due to their zoning, advanced management methods, and user-friendly experience. With the rapid emergence of the intelligent transportation system (ITS), research on parking lot systems has gradually intensified, such as through the parking lot guidance algorithm [1, 2], space allocation [3], and dynamic parking pricing [4]. These works concentrate on parking lot efficiency and mobility, but fewer studies have emphasized safety inside parking lots, which is another critical consideration in ITS besides their efficiency [5].

Numerous studies have examined risk factors and prevention strategies for typical road components (i.e., intersections [6] and curved roadways [7]) and crash types (i.e., hazmat transportation [8] and rear-end crashes [9]). However, there is sparse research regarding accident safety in parking lots. Parking lots have large spaces, complex internal structures, and changing illumination conditions. With the advancement of ITS, parking lots will accommodate multiple types of elements, such as human vehicles (HV), automated vehicles (AV), electric vehicles (EV), and pedestrians, in diverse interactions. Existing studies have found that AVs are more prone to malfunction in parking lots due to the complexity of the environment [10]. In addition, improving the ability of AVs to sense their surroundings in complex scenarios such as parking lots is also a focus of research. For example, collision avoidance by additional signaling devices to sense a car in front of pedestrians [11], and global collision-free path planning for AVs

in underground parking spaces [12]. Meanwhile, despite the low vehicle speeds in the parking lots, frequent blind spots and narrow entrances, and exit ramps raise safety concerns and could cause property damage and casualties. Research on pedestrian tracking under parking lot occlusion [13], real-time positioning of vehicles in underground parking lots based on high-precision maps, and multisensor fusion positioning have also been highly discussed [14]. The dual impact of ITS and complex internal environments may create new parking safety issues. Therefore, there is an urgent need for parking lot safety analyses.

The innovation point of this study lies in analyzing the causes and key factors related to parking lot safety by proposing a quantitative model. This will not only provide a better understanding of the causes of traffic accidents in parking lots but also contribute to the theoretical basis for the application of intelligent transportation in parking lots. The intelligent deployment of parking lots relies on the mature advancement of V2X technology, Internet of Things (IoT) technology, and AVs. The key factors and findings from this study could provide inspiration for the design focus and control strategies of these technologies, such as the environmental elements (e.g., light, ground surface) that need to be sensed by the smart infrastructure and the factors that require additional attention by AVs in the parking lot environment so as to better identify the surroundings and avoid collisions, thereby improving the efficiency of decision-making and avoiding wasted resources.

This study combines accident-based multisource traffic data and a Bayesian network analysis approach. Focusing on parking lot accidents, we implement a proposed mutual information approach to reduce the subjective influence of expert knowledge on model performance and to filter out critical indicators affecting parking lot incidents. Based on the Bayesian network, a multidimensional analysis is conducted for the dual objectives of cause (incident cause) and effect (property damage), and the influence of each factor on the two objectives is obtained. The influence of a single factor on the accident consequences is then considered from five aspects: human, vehicle, road, natural, and social environment, and a logical chain of factors that affect parking lot accidents most is innovatively proposed using mathematical expectations. Based on the model analysis, corresponding strategic measures are established, and a parking lot safety evaluation system is proposed, which can provide technical support for its development to cope with the emergence of new technologies and has specific guidance value for intelligent parking lot design and management strategies.

2. Literature Review

2.1. Feature Selection and Accident Analysis.
Over the past few decades, the analysis of feature selection and accidents has become a hot topic of interest in many studies [15]. Feature selection can reduce the dimensionality of the dataset and improve the performance of the model while helping to better understand the causes of accidents and improve road safety. Mutual information is commonly used to select feature and discuss the degree of correlation

between information. Gu et al. [16] proposed a feature selection algorithm based on conditional mutual information with conditional correlation minimum redundancy (CMI-MRMR) to reduce the data dimensionality while retaining the maximum degree of information. Chen et al. [17] proposed a new Maximin Conditional Mutual Information (MMCMI) ranking method, which can obtain better classification results in less time.

Since 2009, various statistical models have been developed to analyze traffic accidents [18]. With the further expansion of data volume, more data-driven methods are applied to accident analysis, such as support vector machines, artificial neural networks, the XGBoost model [19], and deep learning. However, algorithms such as deep learning are highly dependent on large amounts of data, and the "black box" characteristics do not allow us to know the inner logic of the causes of accidents. Bayesian networks, on the other hand, are widely used because of their advantages in not requiring large amounts of data and their ability to reflect the correlation and causal logic among multiple factors. Meanwhile, Bayesian can be flexibly combined with other methods to target some specific problems [20–22], such as quality reliability analysis under Monte Carlo simulation [23] and risk ranking after random forest prediction results combined with Bayesian [21]. The parking lot data sample is not particularly sufficient to fully reflect the advantages of machine learning analysis, and we are aiming to understand the logical relationships between the factors. Therefore, we will use mutual information combined with Bayesian networks for modeling analysis, and the combination of these two methods has some advantages in accident analysis [22, 24].

2.2. Safety Analysis on Parking Lots.
Given the characteristics of low vehicle speeds and complex surroundings, safety assessments and collision analyses on parking lots differ from ordinary road compositions. Some studies have focused on the simulation and safety assessment of evacuation in parking lots to develop dynamic optimization models [25, 26]. Others have studied the impact of accidents in terms of unique environmental factors in parking lots, such as the illumination system design on safety and perception [27]. This study summarizes the factors and studies related to parking collisions, as presented in Table 1. These studies contribute greatly to the analysis of parking lot safety assessment.

3. Methodology

In this study, we developed a Bayesian network model of parking lot safety, which utilizes multisource data as input to analyze the two targets of parking lot accidents: the first event cause (causation) and property damage (impact). The whole modeling process is divided into three steps. (1) First step. The data are preprocessed and discretized after obtaining the multisource data, and the indicators are preliminarily screened based on the mutual information between the variables and the target nodes. (2) Second step.

TABLE 1: Summary of potential factors affecting parking safety.

Variable	Research	Conclusion
Time	Yahaya et al., 2021	Off-peak hours are one of the influencing factors in fatal injury crashes compared to peak hours [28]
Light conditions	Liu. et al., 2019	Low illumination conditions at night cause frequent and serious traffic accidents [29]
	Bullough et al., 2019	The setting of lighting conditions can partially predict parking lot safety [30]
Weather conditions	Xue et al., 2022	Different weather conditions can affect driving behavior with different levels of road familiarity [31]
	Tu et al., 2015	Driving behavior in hazy weather are different from those in clear conditions [32]
Urban size	Cabrera-Arnau et al., 2020	Urban areas have a higher frequency of minor and serious accidents, while rural areas are more likely to have fatal accidents [33]
Vehicle movement	Afshar et al., 2022	Vehicle movements like moving forward, and left turn are of great importance in accident analysis [34]
Vehicle type	Duan and He 2022	There are similarities and differences between the factors influencing the severity of traffic accidents of different vehicle types [35]
Vehicle speed	Lin et al., 2022	Speed limit is an important risk factor for traffic accidents [36]
Shoulder type	Kordani 2018	Setting proper shoulders helps improve road safety on two-lane roads [37]
Road surface conditions	Chen et al., 2018	Road surface conditions may have a significant impact on the likelihood of a crash on a highway, even on typical highways that do not experience frequent adverse weather [38]

The initial Bayesian network structure is established based on the previously screened variables. Using the network simplification principle proposed in this paper, we deleted the redundant nodes and arcs of the initial network to obtain the optimal model, followed by parameter learning to establish the Bayesian network model of the parking lot with two objectives. (3) In the third step, model validation and analysis are performed. The analysis includes node relationship analysis based on mutual information and strength of influence, sensitivity analysis to determine the impact of each variable on the two objectives, and identification of the most dangerous scenarios using multifactor conditional probability analysis. The implementation process of modeling and analysis is shown in Figure 1, and the specific methods used in each step will be described in the following sections.

3.1. Mutual Information (MI).

Mutual information (MI) is one of the most widely used measures of correlation between a pair of variables in machine learning and is often used to select feature variables [39]. In this study, the critical indicators were firstly screened by using the MI relationship among the variables in the first step. In addition, the MI relationship was also utilized in the network simplification principle in the second step. The MI between variables A and B is defined as,

$$I(A, B) = H(B) - H(B|A). \tag{1}$$

This is equivalent to,

$$I(A, B) = \sum_{a \in A} \sum_{b \in B} P(A, B) \log_2 \frac{P(A|B)}{P(A)P(B)} = \sum_{b \in B} P(B) \sum_{a \in A} P(A|B) \log_2 \frac{P(A|B)}{P(A)}, \tag{2}$$

where B denotes the target node (property damage and accident cause), and A is the set of factors that affect the target. $H(B)$ denotes the uncertainty of the target, and $H(B|A)$ denotes the uncertainty when investigating based on factor A. $P(A, B)$ describes the joint probability function of the target and the associated factors. $P(A)$ and $P(B)$ denote the prior probability functions of A and B, respectively.

3.2. Data Discretization Methods.

In this paper, we use an algorithm based on class attribute contingency coefficient (CACC) for discrete data [40], which is suitable for cases with small data samples and uneven sample distribution and can fully consider all sample distributions to avoid overfitting. The dependency between variables is measured using equations (3) and (4) as scoring functions. Where M is the total sample number; n is the number of classifications; q_{ir} is the number of samples of class i in the interval $[d_{r-1}, d_r]$; M_{i+} is the total sample number of class i; and M_{+r} is the total sample number in the interval $[d_{r-1}, d_r]$. The algorithm process is as follows. Step 1. Determine a data set with a sample size of M, l discrete variables, and S target classifications. Step 2. For each to-be-discrete variable X_l, find the maximum and minimum values as initialization interval boundaries. Step 3. Arrange the values in the initial interval in ascending order and calculate the endpoints of adjacent values. Step 4. Iteratively calculate the CACC value and know that the value is no longer increasing, then the interval is optimal.
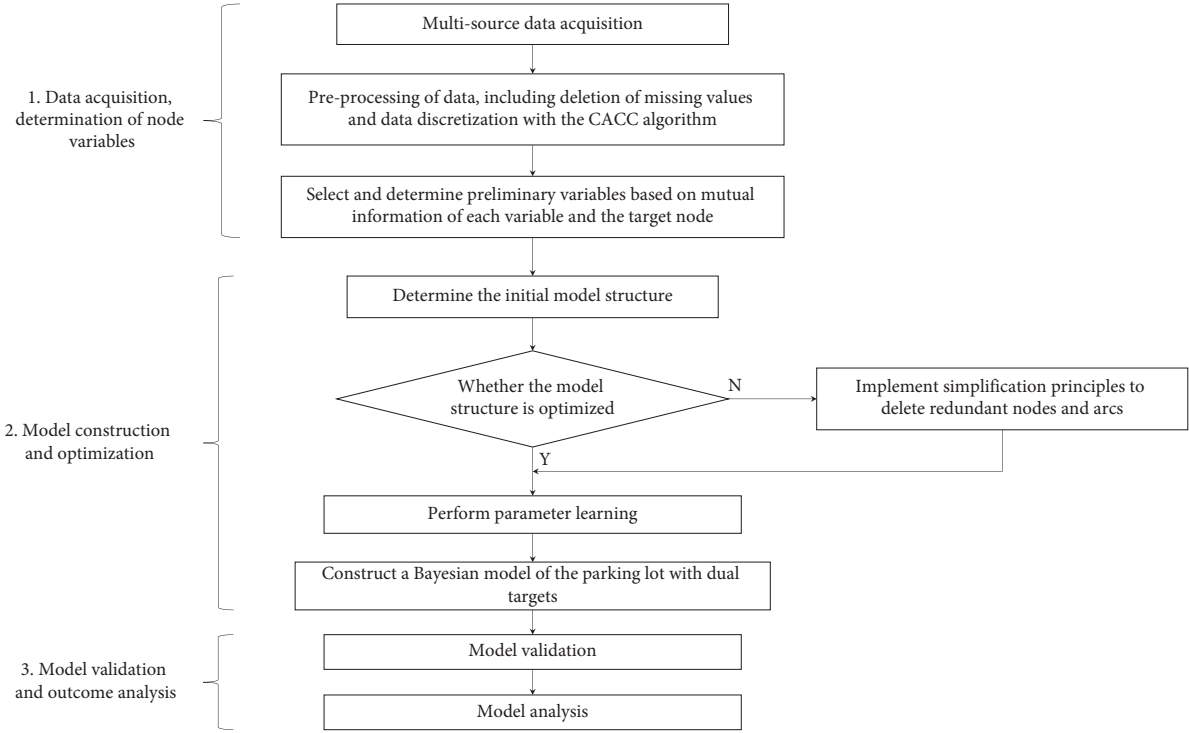
FIGURE 1: Bayesian network modeling process for parking lot safety.

$$\mathrm{cacc} = \sqrt{\frac{y}{y + M}},$$

$$y = \frac{M\left[\left(\sum_{i=1}^{S} \sum_{r=1}^{n} q_{ir}^2 / M_{i+} M_{+r}\right) - 1\right]}{\log_2^n}, \quad i = 1, 2, \ldots, S, \quad r = 1, 2, \ldots, n. \tag{3}$$

3.3. Bayesian Networks and Structure Learning.

Bayesian classification can be classified into naive Bayes, seminaive Bayes, and Bayesian networks. Although the Tree Augmented Naive Bayes (TAN) method [41] in seminaive Bayes also uses a mutual information approach to build the network structure, this study still uses Bayesian networks for analysis. The seminaive Bayes method assumes that each feature depends and only depends on one other feature other than it. However, when actually considering parking lot safety, we cannot be sure that each feature depends on only one feature, and some information may be lost by such an oversimplified assumption. Therefore, the seminaive Bayes approach is less suitable than the Bayesian networks. A Bayesian network is a probabilistic graphical model that represents a set of variables and their conditional dependencies through a directed acyclic graph (DAG) [42]. It can provide a convenient framework to represent causal relationships, making inference uncertainty more logically evident. The nodes in the DAG model represent random variables. Variables or propositions that are considered causal (or unconditionally independent) are connected by arrows. The node from which the arrow departs is the "parent node," and the other is the "child node." The strength of association or

confidence coefficient between variables is described by conditional probability tables (CPT), which can perform tasks such as prediction, diagnosis, and classification. The learning of Bayesian networks mainly includes structure learning, parameter learning, and network inference.

Structure learning is a process that obtains the network structure from a dataset D. In this study, the initial network structure is established using the greedy thick thinning (GTT) algorithm [43]. It has a relatively simple learning process, large search space, and high learning efficiency, which searches the optimal network structure by utilizing the scoring function. The specific implementation steps are as follows [44]. (1) Network edge addition: generate an initial network with an empty structure to search and add directed edges that benefit the network structure score until it no longer increases. (2) Network edge deletion: edges that have no positive effect on the network structure are continuously searched and deleted until the structure score no longer increases. The network structure is scored using the K_2 function as follows, where $P(G)$ denotes the prior probability of the network, n is the number of random variables X_i, r_i is the number of states of the node X_i, and q_i is the total number of states of the parent node X_i. The N_{ij} is the number of nodes X_i that take the k th combination of states. That is, $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$.

$$K_2(G, D) = \sum_{i=1}^{n} \sum_{j=1}^{q_i} \left[\log\left(\frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \right) + \sum_{k=1}^{r_i} \log(N_{ij}!) \right] + \log(P(G)). \tag{4}$$

3.4. Parameter Learning. Parameter learning of the Bayesian network is performed based on its constructed structure G and data D to obtain the conditional probability distribution of the node variables. In this study, the expectation-maximum (EM) algorithm [45] is used for parameter estimation. This learns the parameters using heuristics when there are unobserved implicit data (missing values) in the dataset. The EM algorithm is divided into two steps at each iteration: the E-step and the M-step. For a given data set $A = (A_1, \ldots, A_m)$, there exists unobserved implicit data $z = (z_1, \ldots, z_m)$. The maximum number of iterations is specified as J. Then, (1) randomly generate the initial value θ^0 of the initialized model parameters θ and (2) perform the jth iteration. This process defines the E-step. The conditional probability expectation of the joint distribution is then computed.

$$Q_i(z_i) = P(z_i | A_i, \theta^j),$$
$$L(\theta, \theta^j) = \sum_{i=1}^{m} \sum_{z_i} Q_i(z_i) \log P(A_i, z_i; \theta), \tag{5}$$

where $P(z_i | A_i, \theta^j)$ is the conditional distribution and $Q_i(z_i)$ is a newly introduced unknown distribution with $\sum_z Q_i(z_i) = 1$. $P(A_i, z_i; \theta)$ is the joint distribution of A_i and z_i, and $L(\theta, \theta^j)$ is the likelihood estimation function. This process defined the M-step. $L(\theta, \theta^j)$ is maximized to attain θ^{j+1}. If θ^{j+1} has converged, the iterations are stopped and the model parameters are taken. Otherwise, the E-step continues to iterate. θ^{j+1} is computed as follows:

$$\theta^{j+1} = \arg\max_{\theta} L(\theta, \theta^j). \tag{6}$$

4. Modeling and Analysis

The implementation and analysis of the Bayesian model for parking lot accident safety are presented in this section. The data source, data preprocessing, model analysis, and performance results are introduced.

4.1. Data Source. The crash database from the Florida Department of Highway Safety and Motor Vehicles (DHSMV) is used to model the contributing factors to parking lot accidents using the Bayesian network approach. The data in the DHSMV have two sources, the Florida Department of Transportation (FDOT) Safety Office and the Crash Analysis Reporting (CAR) System. These have recorded Florida traffic accident data since 2007. More than 300 attributes are used to describe and report each accident record, which covers four aspects of the accident: person (driver's gender, age, etc.), vehicle (damage, impact area, etc.), road (accident location, road type, road markings, etc.), and environment (weather, light, date, etc.). We concentrate on property damage (primary target) and event causes (secondary target)

in parking lot accidents. Such accidents are those where the location (SITE LOCATION) is in a parking lot, parking aisle, or stall. Parking types include surface parking lots, underground garages, above-ground parking lots, and multiuse parking lots. We screened the accident data that occurred in these places for subsequent analysis.

4.2. Data Preprocess and Variable Selection. Missing value processing, as well as continuous variable discretization, will be utilized as data preprocessing prior to modeling and data analysis. Data with incomplete information (more than 60% missing) and variables without information changes (i.e., variable values are of the same class) are first removed from the training data set to ensure the required accuracy. As parking lot accident data have no recorded fatalities and most people were only slightly injured, the casualties were unevenly distributed and not targeted for subsequent modeling. Ultimately, 5.92% of the parking lot incident records were missing. Thus, 5,599 events were selected, and 30 variables were retained for model generation. The names of the 30 variables involved and their abbreviation codes and explanations are shown in Appendix.

Before the preliminary screening of variables, continuous variables are discretized using an algorithm based on CACC to retain more knowledge. Variables to be discrete in the original data include continuous variables (MEDWIDTH and PROPERTY) and variables with more than ten categories to be reclassified (POINTIMP and VEHMOVE). The number of discrete categories was tried from four to eight categories, and the result with logical and evenly distributed classification results was selected. For example, POINTIMP is divided into four categories, just corresponding to the collision location in the four directions of the vehicle.

The mutual information-based variable screening was implemented to simplify the model complexity. This study selected and eliminated variables based on the following principles. (i) To retain as much data as possible, we performed variable selection with an MI of 0.5 as the criterion. First, for the preliminary target node (PTN), variables with an MI greater than 0.5 with PROPERTY were kept and denoted as the first-level related nodes (FLRN). (ii) We examine variables with an MI above 0.5 using FLRN one by one. We kept these nodes and denoted them as the second-level related nodes (SLRN). The implementation results are shown in Table 2, where only variables that comply with the requirements for the first time are presented. After filtering, 20 variables remained, and their MI correlation matrix is shown in Figure 2.

4.3. Bayesian Network Modeling. The GeNIe 3.0 Academic was employed to build the Bayesian model based on property damage in parking lot accidents. Prior to structure

TABLE 2: Node filtering process associated with the preliminary target node (PROPERTY).

Preliminary target node (PTN)	First-level related nodes (FLRN)	MI of FLRN	Second-level related nodes (SLRN)	MI of SLRN
			RDSURFCD	0.904344
			WEATCOND	0.904344
	RTESGNCD	1	TYPESHLD	0.904344
			VEHSPEED	0.584314
	LANDUSE	0.827324	Included	
	LGHTCOND	0.798956	FRST_EVNT_CAUS_CD	0.668544
PROPERTY*			VEHTYPE	0.676928
	MAXSPEEDLMT	0.65439	VEHUSE	0.676928
			DAYOWEEK	0.555487
			TOT_OF_PERS_NUM	0.564955
	VEHMOVE	0.622137	TRAVDIR	0.580224
			VEHDMTYP	0.53781
	POINTIMP*	0.564509	URBSIZE	0.518511
	TIME	0.52413	Included	

The *sign indicates variables that have been discretely processed by CACC.
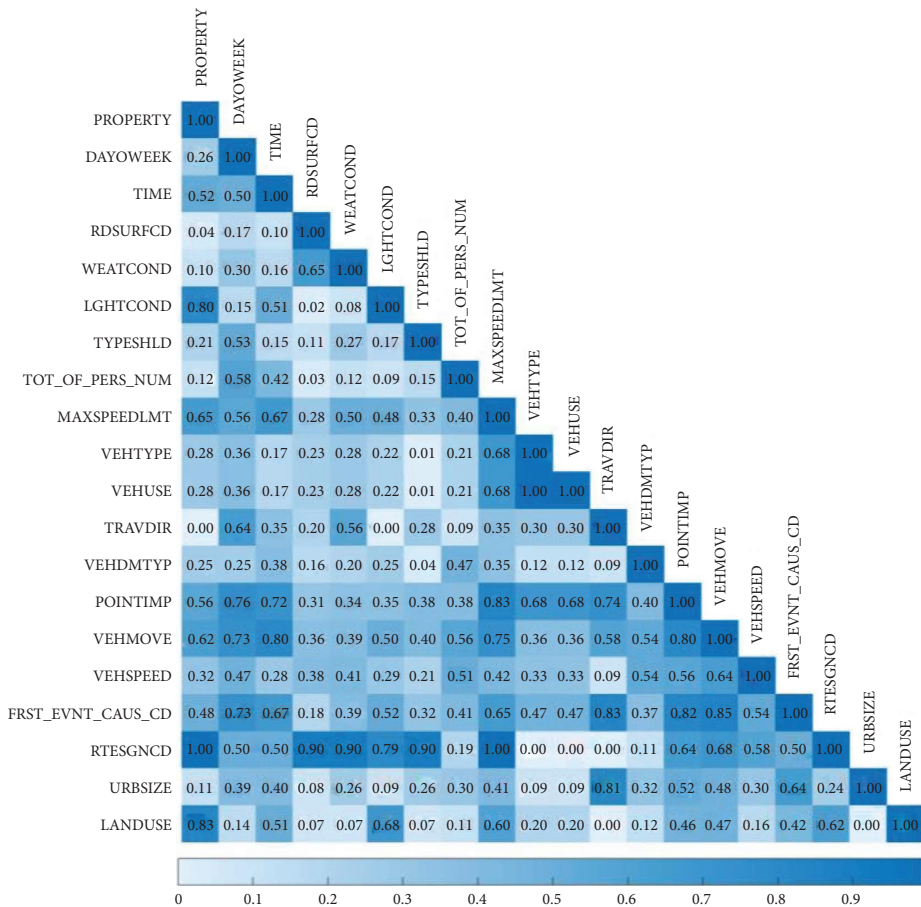

FIGURE 2: The MI matrix of the 20 variables for the preliminary selection.

learning, we incorporated temporal tier rules to specify the temporal order between variables. We considered that DAYOWEEK, WEATCOND, LGHTCOND, TYPESHLD, URBSIZE, LANDUSE, RTESGNCD, and TIME are in higher tiers (later in time), and FRST_EVNT_CAUS_CD and PROPERTY are in lower tiers (earlier in time). No arcs point from variables in higher tiers to lower tiers, which avoids unrealistic logical occurrences.

The structure learning and parameter learning of the network were performed using the GTT algorithm and EM algorithm, respectively. We applied the following simplification principles to simplify redundant arcs and optimize the model structure. (i) Remove nodes that do not connect any arcs as they are irrelevant to model causality. (ii) Remove arcs that are irrelevant and logically incorrect, such as cases where subjective factors influence objective ones. (iii) Arcs

with MI < 0.7 were deleted. This condition was tested to reduce the computational complexity while retaining more information. (iv) When there is more than one logical chain between two nodes, the logical chain with the larger MI is retained based on the MI between connected nodes. (v) Each time an arc is deleted, parameter learning is performed, and the accuracy and area under curve (AUC) values are compared using the "with and without comparison method" to ensure the accuracy improved at each step.

The resulting structure of the Bayesian network for parking accidents is shown in Figure 3. Thus, the network contains a total of 17 variable nodes. The lighting conditions and vehicle movement directly affect property damage, with weather conditions, time of day, and maximum parking lot speed limits being the first tier of indirect factors. Meanwhile, the point of impact and vehicle movement directly determine the event cause. After parameter learning, the Bayesian network model for the causality analysis of parking lot accidents is determined, as shown in Figure 4.

4.4. MI and Strength of Influence (SI) Analysis. Together with the MI, the strength of influence (SI) between the arcs in the network also indicates the influence magnitude between variables. The SI is calculated from the conditional probabilities of the child nodes. The MI and SI between all arcs and corresponding nodes are shown in Table 3.

With the integrated attention to the MI and SI analysis between variables, the lighting conditions (MI = 0.80, SI = 0.62) and vehicle movement (MI = 0.62, SI = 0.51) have the strongest correlation with the primary target (PROPERTY). The point of impact (MI = 0.22, SI = 0.34) and vehicle movement (MI = 0.85, SI = 0.32) have a relatively significant impact on the secondary target (FRST_EVNT_CAUS_CD). In addition, the time has a strong relevance to both lighting conditions and vehicle movement, and the high influence of the strength value ensures it has a potential relationship with the land use (SI = 0.56) and route signing qualifier (SI = 0.36). The maximum speed limit and vehicle type are the most likely to affect the point of impact and have some influence on the accident event cause.

4.5. Sensitivity Analysis. A sensitivity analysis [46] is a method to investigate the effect of small changes in probability on target nodes. Higher sensitivity values indicate that the node is more likely to cause interference to the target. The results of sensitivity for each variable to two target nodes are shown in Figure 5. The day of the week and maximum speed limit had the strongest influence on both the property damage and the event cause. The urban size and vehicle movement also had greater impacts on property damage, while vehicle movement placed significance on event causes. Moreover, lighting and weather conditions only affected property damage, implying that poor lighting and weather may lead to more severe accident losses. The vehicle movement, type, speed, point of impact, type of vehicle damage, and number of people involved all affected the cause of the accident to different extents. Most of these factors are related to the properties of the vehicle and can shed some light on accident prevention for drivers.

4.6. Conditional Probability Analysis. This subsection explores the impact of a single-factor change on the two targets and investigates the impact of a multifactor combination change on property damage. A series of factors and corresponding states that have the most significant influences are given. The conditional probability of the PROPERTY and FRST_EVNT_CAUS _CD influencing factors is shown in Figures 6(a) and 6(b), respectively. The following analysis discusses the impact of each factor on the targets in terms of the environment, people, vehicles, roads, and accident consequences.

4.6.1. Natural Environment. Mondays and Fridays are more sensitive to changes in property damage with accidents occurring on Mondays being more likely to have medium damage or collisions with gates. Poor streetlights significantly increase the likelihood of collisions with parked vehicles and fixed objects, which results in high losses. Early morning to midday is more likely to have vehicle-to-vehicle collisions with lower property damage, while more costly collisions with gates usually occur in the afternoon through midnight.

4.6.2. Social Environment. In high-density business centers and large cities, the probability of a backed-in collision increases slightly, but the risk of property damage is low. Parking lot accidents in residential or urban areas are usually vehicle collisions with gates. This implies that comprehensive and robust safety management measures and gate designs should be major considerations in parking lot planning and design.

4.6.3. People. The probability of higher damage loss increases in line with the number of people involved in parking lot accidents. This gives a greater likelihood of rear-end and backed-in collisions.

4.6.4. Vehicles. The probability of property damage greater than \$750 increases by 13–26% at moderate speed limits (between 30 and 40 mph), implying that high- or low-speed limits can increase driver safety awareness. Medium speed limits do not restrain drivers with careless driving causing more property damage. Higher speed limits in parking lots decrease the probability of rear-end, backed-in, and collision with gates. Vehicles are more likely to collide with gates when going straight and with fixed objects when turning left. There is an increased probability of backed-in collisions when backing up. Parked and slow-moving vehicles are safest in parking lots because the probability of all accidents is lower. Angular collisions are more likely to occur in left turn and driverless situations. As speeds increase, vehicles are more likely to have high-loss collisions with other vehicles and objects in parking
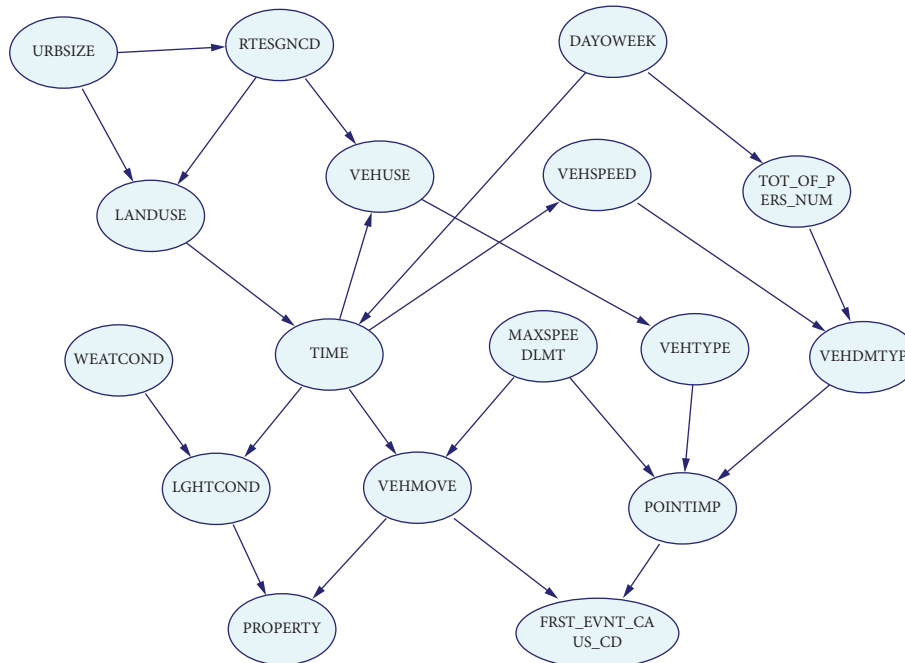
Figure 3: Bayesian network model for parking lot accidents.
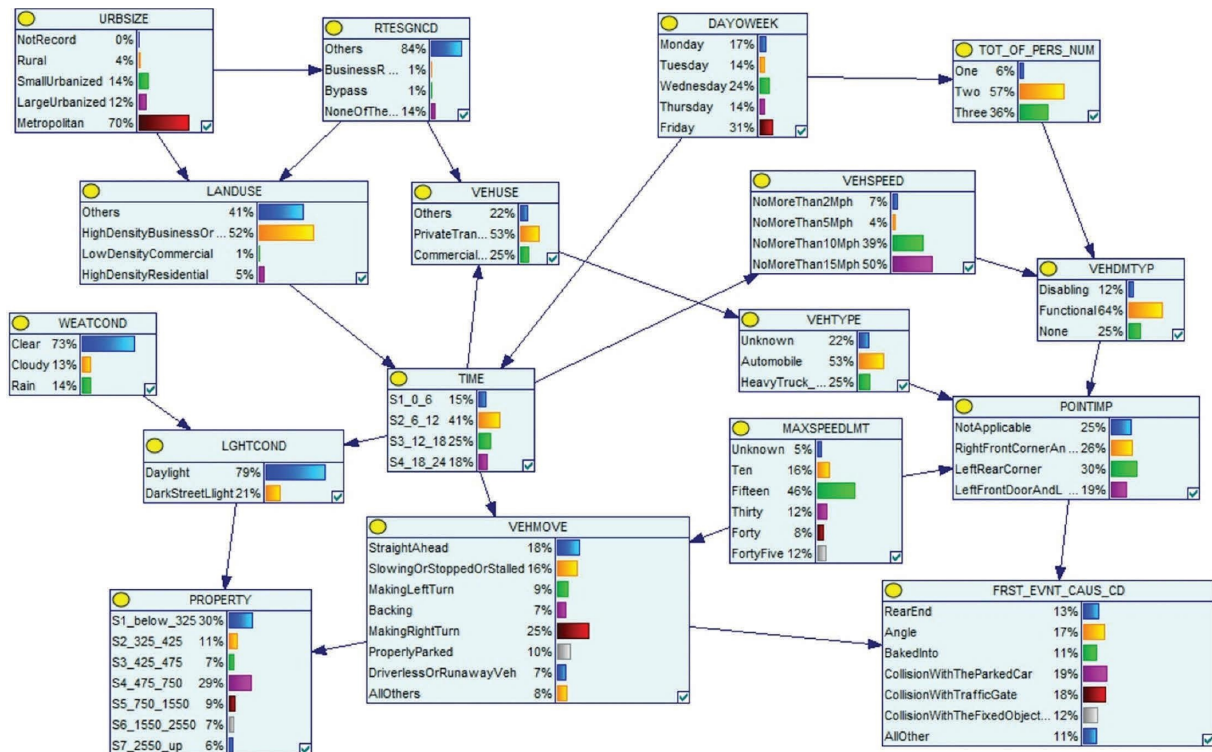
Figure 4: The Bayesian network model after parameter learning in the GeNIe 3.0 Academic software.

lots. This means that the complex environment of parking lots continues to have a greater impact on driving safety. Heavy trucks and commercial cargo are prone to collisions with gates, while commercial cargo may suffer moderate to high damage.

4.6.5. Roads. The probability that damage at a bypass exceeds $750 will increase with more traffic flow, implying that high-risk accidents are more likely to occur. Rear-end and backed-in collisions are the most common types of accidents as they have the highest probabilities in most cases.

TABLE 3: The MI and SI in the parking lot Bayesian network.

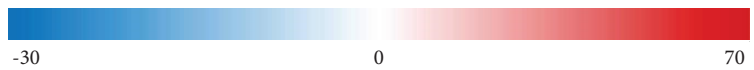| Parent node	Child node	MI	SI
VEHUSE	VEHTYPE	1.0000	0.9994
VEHMOVE	FRST_EVNT_CAUS_CD	0.8530	0.3181
MAXSPEEDLMT	POINTIMP	0.8294	0.1569
POINTIMP	FRST_EVNT_CAUS_CD	0.8154	0.3435
LGHTCOND	PROPERTY	0.7990	0.6197
TIME	VEHMOVE	0.7968	0.3869
MAXSPEEDLMT	VEHMOVE	0.7451	0.3444
VEHTYPE	POINTIMP	0.6828	0.1716
RTESGNCD	LANDUSE	0.6223	0.3601
VEHMOVE	PROPERTY	0.6221	0.5098
DAYOWEEK	TOT_OF_PERS_NUM	0.5828	0.5573
VEHSPEED	VEHDMTYP	0.5383	0.4229
TIME	LGHTCOND	0.5146	0.4436
LANDUSE	TIME	0.5115	0.3717
DAYOWEEK	TIME	0.4956	0.5597
TOT_OF_PERS_NUM	VEHDMTYP	0.4699	0.4542
VEHDMTYP	POINTIMP	0.4004	0.1716
TIME	VEHSPEED	0.2842	0.494
URBSIZE	RTESGNCD	0.2420	0.113
TIME	VEHUSE	0.1728	0.3661
WEATCOND	LGHTCOND	0.0790	0.3329
RTESGNCD	VEHUSE	0.0001	0.3267
URBSIZE	LANDUSE	0.0001	0.2752
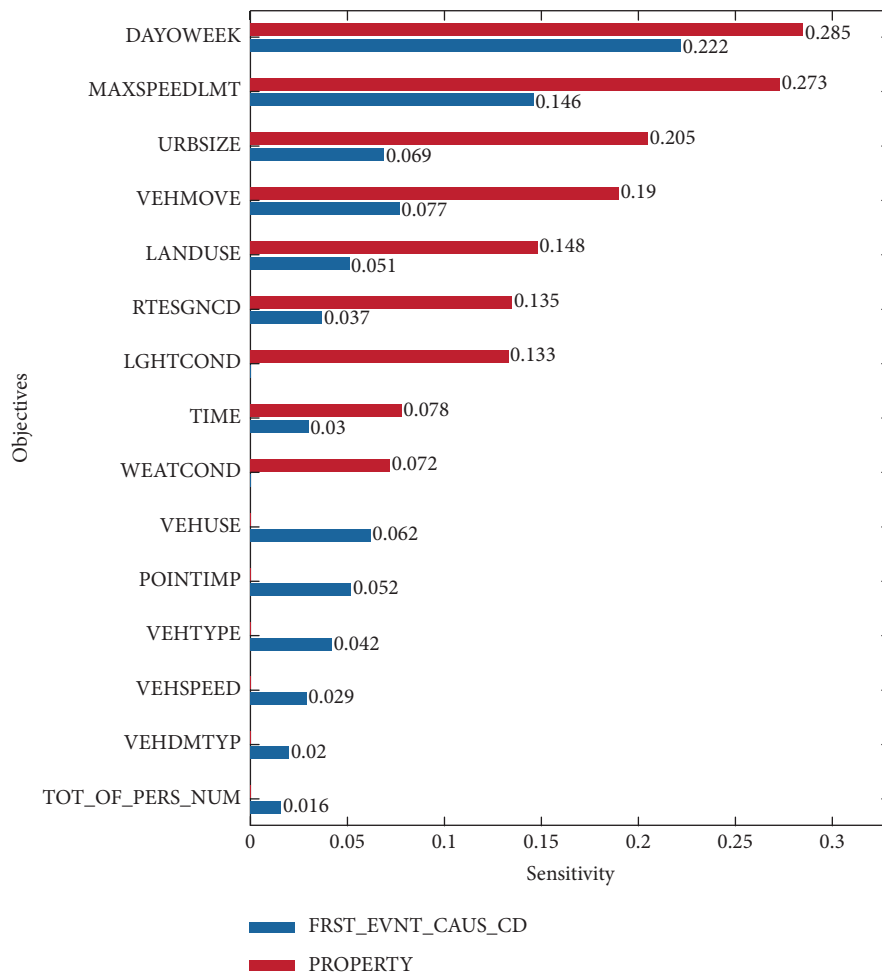
FIGURE 5: Sensitivity of each variable to two target nodes (PROPERTY and FRST_EVNT_CAUS_CD).

Variables	Code	\[0, 325]	(325, 425]	(425, 475]	(475, 750]	(750, 1550]	(1550, 2550]	(2550, +∞)
		Probability variation-PROPERTY (%)						
DAYOWEEK	Monday	-22	9	-5	23	4	-5	-4
	Friday	4	-3	2	-12	-1	8	4
LGHTCOND	Daylight	6	1	-1	5	0	-5	-4
	Dark (street light)	-20	-1	3	-19	1	23	15
TIME	0:00-6:00	-11	3	0	1	5	0	3
	6:00-12:00	27	-6	2	-8	-4	-6	-5
	12:00-18:00	-23	9	-5	24	4	-5	-4
	18:00-24:00	-19	-1	2	-17	1	22	12
WEATCOND	Cloudy	1	-1	2	-7	0	2	3
	Rain	1	1	0	1	1	-1	-1
LANDUSE	High density business center	14	-2	1	-2	-2	-4	-4
	High density residential	-13	2	0	-1	2	7	4
	Rural	-14	3	-1	2	3	5	3
URBSIZE	Small Urbanized	-14	3	-1	2	3	5	3
	Large Urbanized	-14	3	-1	2	3	5	3
	Metropolitan	7	-1	0	-1	-1	-1	-2
TOT_OF_PERS_NUM	Two	-2	2	-1	6	1	-3	-2
	Three	4	-3	2	-10	-1	6	3
MAXSPEEDLMT	10	0	17	3	-12	-4	-2	-1
	15	9	-6	-1	6	-4	-2	-1
	30	-11	2	-2	-4	4	13	-2
	40	-15	0	-2	-9	26	0	1
	45	-11	3	0	2	5	1	3
VEHMOVE	Straight ahead	-28	-9	-5	57	-7	-5	-4
	Slowing/stopped/stalled	-23	-4	0	-22	-2	48	1
	Making left turn	-20	-1	3	-19	1	3	36
	Backing	26	-10	9	-6	-8	-6	-5
	Making right turn	-27	69	-4	-26	-6	-4	-3
	Properly parked	-26	-7	-3	-25	68	-3	-2
	Driverless or runaway veh.	58	-9	-5	-27	-7	-5	-4
VEHSPEED	<=2 mph	27	-6	2	-8	-4	-6	-5
	<=5 mph	5	1	-1	6	0	-5	-4
	<=10 mph	-11	3	0	1	5	0	3
VEHTYPE	Heavy truck	3	1	-1	7	0	-5	-4
VEHUSE	Private transportation	-5	1	0	-2	1	5	2
	Commercial cargo	3	1	-1	7	0	-5	-4
RTESGNCD	Bypass	-12	2	0	-1	2	7	4
	None of the above	14	-2	1	-2	-2	-4	-4
POINTIMP	Right front & rear corner	2	2	0	-4	0	0	0
	Left rear corner	-2	3	0	-4	1	3	0
	Left front door & left front corner	0	-3	-1	6	1	-1	-1

-30   0   70

(a)

FIGURE 6: Continued.

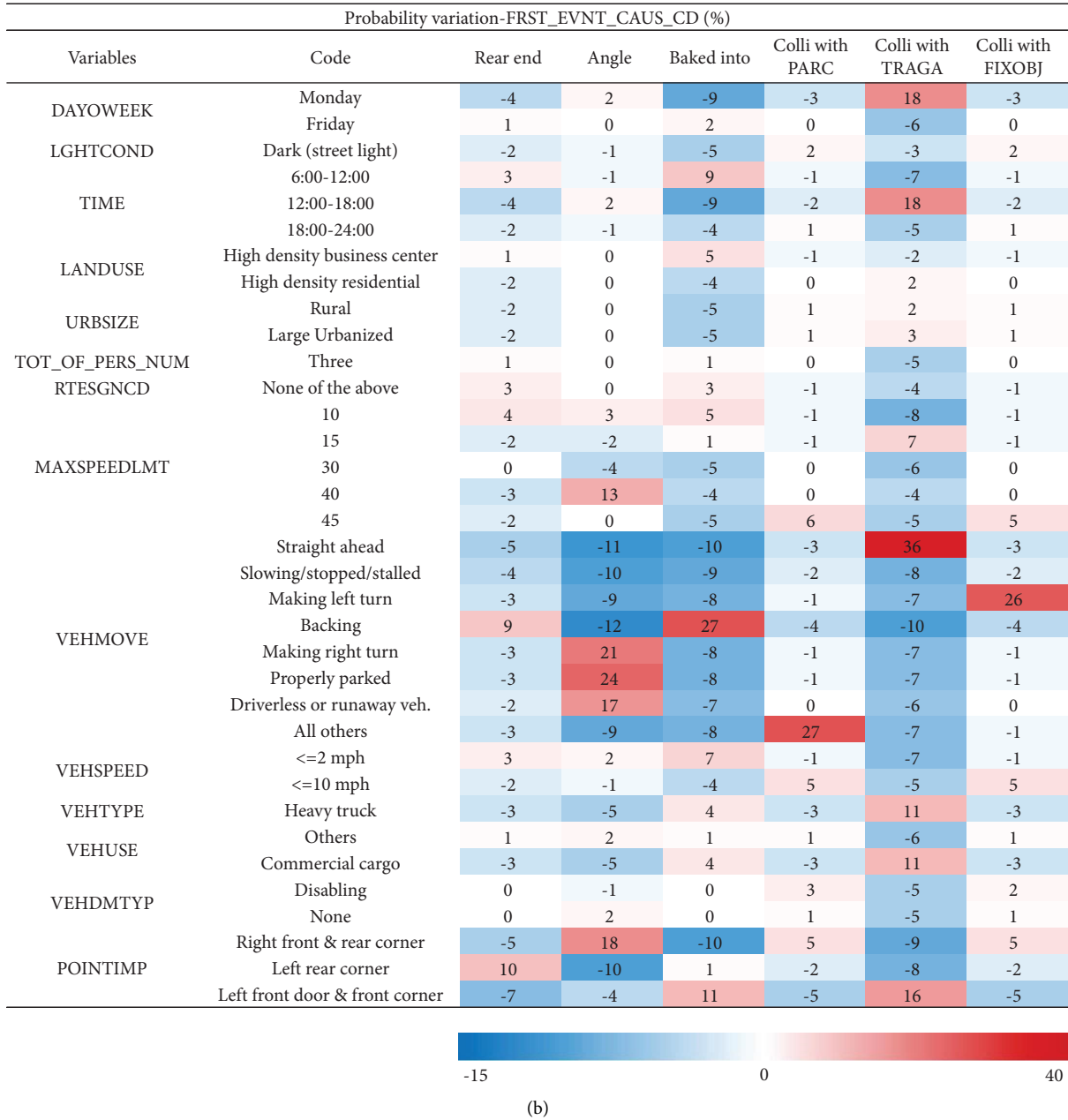| Probability variation-FRST_EVNT_CAUS_CD (%)							
Variables	Code	Rear end	Angle	Baked into	Colli with PARC	Colli with TRAGA	Colli with FIXOBJ
DAYOWEEK	Monday	-4	2	-9	-3	18	-3
	Friday	1	0	2	0	-6	0
LGHTCOND	Dark (street light)	-2	-1	-5	2	-3	2
TIME	6:00-12:00	3	-1	9	-1	-7	-1
	12:00-18:00	-4	2	-9	-2	18	-2
	18:00-24:00	-2	-1	-4	1	-5	1
LANDUSE	High density business center	1	0	5	-1	-2	-1
	High density residential	-2	0	-4	0	2	0
URBSIZE	Rural	-2	0	-5	1	2	1
	Large Urbanized	-2	0	-5	1	3	1
TOT_OF_PERS_NUM	Three	1	0	1	0	-5	0
RTESGNCD	None of the above	3	0	3	-1	-4	-1
MAXSPEEDLMT	10	4	3	5	-1	-8	-1
	15	-2	-2	1	-1	7	-1
	30	0	-4	-5	0	-6	0
	40	-3	13	-4	0	-4	0
	45	-2	0	-5	6	-5	5
VEHMOVE	Straight ahead	-5	-11	-10	-3	36	-3
	Slowing/stopped/stalled	-4	-10	-9	-2	-8	-2
	Making left turn	-3	-9	-8	-1	-7	26
	Backing	9	-12	27	-4	-10	-4
	Making right turn	-3	21	-8	-1	-7	-1
	Properly parked	-3	24	-8	-1	-7	-1
	Driverless or runaway veh.	-2	17	-7	0	-6	0
	All others	-3	-9	-8	27	-7	-1
VEHSPEED	<=2 mph	3	2	7	-1	-7	-1
	<=10 mph	-2	-1	-4	5	-5	5
VEHTYPE	Heavy truck	-3	-5	4	-3	11	-3
VEHUSE	Others	1	2	1	1	-6	1
	Commercial cargo	-3	-5	4	-3	11	-3
VEHDMTYP	Disabling	0	-1	0	3	-5	2
	None	0	2	0	1	-5	1
POINTIMP	Right front & rear corner	-5	18	-10	5	-9	5
	Left rear corner	10	-10	1	-2	-8	-2
	Left front door & front corner	-7	-4	11	-5	16	-5

-15   0   40

(b)

FIGURE 6: Conditional probability of target nodes for parking lot accidents considering (a) property damage and (b) event cause. The "Colli with PARC," "Colli with TRAGA," and "Colli with FIXOBJ" are short for "Collision with the parked car," "Collision with traffic gate," and "Collision with a fixed object above the road."

4.6.6. Accident Consequences.

Collisions with gates are unlikely to cause disabling damage to vehicles. Right front and rear-end crashes are more likely to result from angular collisions. The left front door corner may be presumed to be a backed-in collision or a collision with gates. The overall property damage is higher when the left side of the vehicle is damaged.

As multiple factors contribute to accidents, it is necessary to study the probability distribution of multiple factors under various condition combinations. Multifactor conditional probability analysis was performed for property damage because event causes cannot be ranked

quantitatively. We used expectations to characterize the amount of property damage. The expectation of property damage is calculated as follows:

$$E(PD_{ik}) = \sum_{j=1}^{J} p_{ik} \cdot \overline{I_j}, \qquad (7)$$

where $E(PD_{ik})$ is the expectation of property damage corresponding to the k th state of the i th variable, p_{ik} represents the conditional probability corresponding to the k th state of the i th variable, J represents the number of target classification nodes, and $\overline{I_j}$ represents the average of

the two ends of the j th interval. The factor conditions with an expected loss of more than \$900 are selected and ranked in Table 4.

Based on the causal logic of nodes in the network, we formed a complete logical chain by finding the preceding order nodes in Table 5 starting from the target node (PROPERTY). We list all scenarios and rank the expectations of property damage for each case. For instance, one of the predecessor nodes of PROPERTY is VEHMOVE, and one of the predecessor nodes of VEHMOVE is MAX-SPEEDLMT. Then, the logic chain is MAXSPEEDLMT $\longrightarrow$ VEHMOVE $\longrightarrow$ PROPERTY. If we consider VEHMOVE = Making a left turn, MAX-SPEEDLMT $\leq$ 30 mph, the expected property damage is calculated as \$1151.25. A total of nine combinations are considered in this paper, and the top three rankings are shown in Table 5.

The highest property damage in Table 5 was caused by slow-moving vehicles at speed limits below 30 mph, \$1,643.50. Although previous findings confirm the low probability of slow-moving vehicle accidents, slow-moving vehicles still carry a high risk of property damage under these speed limits. Left-turning vehicles under speed limits of less than 40 mph will also have great losses (\$1,527) in accidents, and operators need to take precautions. It is speculated that parking lot speed limits may increase the risk of vehicle movement that would not otherwise exist, which affects accident damage. Moreover, high-loss parking accidents are most likely to occur at dimly lit bypasses at night near high-density residential areas in large cities, resulting in \$1,479 in damage. Managers should consider factors such as night, lighting, and bypasses.

4.7. Model Validation. The accuracy and generalization ability of the model is verified using the leave-one-out (LOO) [47] method built in GeNIe. LOO is a special case of k-fold cross-validation, where k equals the number of samples in the dataset M. The method requires the network to be trained on $M - 1$ samples, tested on the remaining samples, and repeatedly implemented M times. LOO was shown to be more suitable for this study than k-fold cross-validation. Overall, the accuracy of the final prediction of property damage and accident causes reached 85.71% and 64.42%, respectively. The mean values of the AUC reached 0.978 for PROPERTY and 0.830 for FRST_EVNT_-CAUS_CD. The results of their confusion matrix are shown in Tables 6 and 7, where the notation in Table 7 is consistent with Figure 6.

The prediction accuracy is low when the property damage is between \$425 and \$475 or over \$2550. The rest of the states have high prediction accuracy, which are above 80%. In terms of accident causation, the prediction accuracy is lower for front-end collisions because the model confuses them with angle collisions. This may be due to the fact that the states of the remaining factors do not differ much under these two collision types. The angle collision has the highest prediction accuracy, and the other accident causes have

Table 4: Expectation of property damage in parking lot accidents (more than \$900).

Ranking	Variables	State	Expectation (USD)
1	VEHMOVE	Making left turn	1645.50
2	VEHMOVE	Slowing/stopped/stalled	1514.25
3	LGHTCOND	Dark (street light)	1472.75
4	TIME	18:00–24:00	1378.38
5	VEHMOVE	Properly parked	1142.50
6	RTESGNCD	Bypass	972.00
7	LANDUSE	High-density residential	970.37
8	MAXSPEEDLMT	40	950.87
9	URBSIZE	Large urbanized	929.13
10	URBSIZE	Small urbanized	929.13
11	URBSIZE	Rural	929.13
12	MAXSPEEDLMT	30	925.75
13	DAYOWEEK	Friday	906.88

about equal accuracy. The uneven data distribution may lead to some inaccuracy in the model during the training and recognition process. In the future, the study will enrich the data and expand the dataset volume to further explore the model performance.

To compare the performance of different methods, we perform further analysis for the primary target of the model (property damage). The result is verified by comparing the property damage of four different models under the same modeling process. They are the models proposed in this paper (with 20 variables), the expert knowledge model (with 10 variables summarized in Table 1), and the models under equal-width discretization and equal-volume discretization, respectively. The data sets used in the latter two of these models are the same as those used in the model proposed in this paper.

According to Figure 7, the accuracy and AUC of the proposed model outperformed the other models in predicting different states, with a greater advantage in AUC (significantly higher than the other models) when the property loss was below \$1550. By comparing the ROC curves of different models in predicting property damage (Figure 8), it can be found that the training effect of the proposed model is better than other model settings. The next better results are models under equal-width discretization and equal-volume discretization, and the worst results are obtained by the model using only expert knowledge. Thus, it is confirmed that the proposed model has good adaptability and generalization ability.

4.8. Evaluation System. The Bayesian modeling analysis of parking lot accidents reveals variations in the influence of different factors on parking lot safety. Thus, a parking lot safety evaluation system is developed to present guidance for planning and management. The system is divided into three main phases: multisource data fusion, parking lot Bayesian network modeling, and safety evaluation and management (see Figure 9).

TABLE 5: Factor combination sequence result for property damage.

Ranking	Variable sequences	Expectation (USD)
1	Maximum speed limit ≤ 30 mph ⟶ slowing/stopped/stalled	1643.50
2	Maximum speed limit ≤ 40 mph ⟶ making left turn	1527.00
3	Bypass ⟶ large urbanized ⟶ high-density residential ⟶ 18:00–24:00 ⟶ dark (street light)	1479.00

TABLE 6: The confusion matrix for property damage.

Property damage (USD)		Predict state						
		[0, 325]	(325, 425]	(425, 475]	(475, 750]	(750, 1550]	(1550, 2550]	(2550, +∞)
Actual state	[0, 325]	2062	0	0	40	0	0	0
	(325, 425]	0	473	1	2	1	0	0
	(425, 475]	0	274	7	3	4	0	0
	(475, 750]	216	34	0	1021	4	0	0
	(750, 1550]	0	34	0	0	423	0	0
	(1550, 2550]	0	0	0	0	0	693	1
	(2550, +∞)	0	0	0	0	0	186	120

TABLE 7: The confusion matrix for event cause.

Event cause		Predict state					
		Rear end	Angle	Baked into	Colli with PARC	Colli with TRAGA	Colli with FIXOBJ
Actual state	Rear end	7	335	7	16	2	59
	Angle	0	1670	29	65	5	216
	Baked into	0	326	767	36	26	184
	Colli with PARC	0	77	6	246	1	31
	Colli with TRAGA	0	59	120	7	286	16
	Colli with FIXOBJ	0	315	14	30	3	638

(1) In the multisource data fusion phase, various sensors, image recognition, and deep learning technologies are used to collect various types of information, including people, vehicles, roads, and aspects of the environment. Through database systems and cloud storage services, data from various departments (meteorological department, traffic department, etc.) are fused with historical accident data to collect comprehensive information related to parking lot safety. The integrated dataset can be used to build subsequent Bayesian models.

(2) In the parking lot Bayesian network modeling phase, machine learning methods are used to build a parking lot safety diagnosis model. The Bayesian approach based on MI preprocessing can reduce the interference of subjective experience to the model and provide more realistic conclusions. The associated accident diagnosis model uncovers potential risks and key factors that influence parking lot safety.

(3) Safety evaluation and management. Managers and decision-makers can adjust their parking management approach based on the measurement library by responding to the highly sensitive indicators fed back by the model. The model conclusions have a guiding role for safety and security management. The proposed library of measures can significantly enhance the safety control of parking lots. These measures will be combined with emerging ITS technology to improve the self-governance and intelligence of parking lots (see Table 8). Factors with intrinsic characteristics do not propose corresponding measures as they require external planning that cannot be improved quickly.

With the emergence of diverse ITS technologies, data-driven modeling approaches are increasingly being applied to safety management. Data-driven parking safety evaluation systems based on the Bayesian approach are the primary achievements of parking safety management. This approach can be well integrated with advanced technologies and has good adaptability and scalability in diverse situations. The perception and recognition of artificial intelligence technology inside parking lots help develop guidance systems and optimization algorithms to reduce the risks of sensitive indicators in advance (first phase). The IoT technologies ensure rapid information interoperability between vehicles and facilities, which allows for the safety of autonomous and unmanned vehicles in complex environments such as parking lots. [48] enhance vehicle communications and boost V2X technologies, which ensures that real-time information monitoring in parking lots is up-to-date and precise. This process can significantly enhance parking lot safety control in large cities and high-
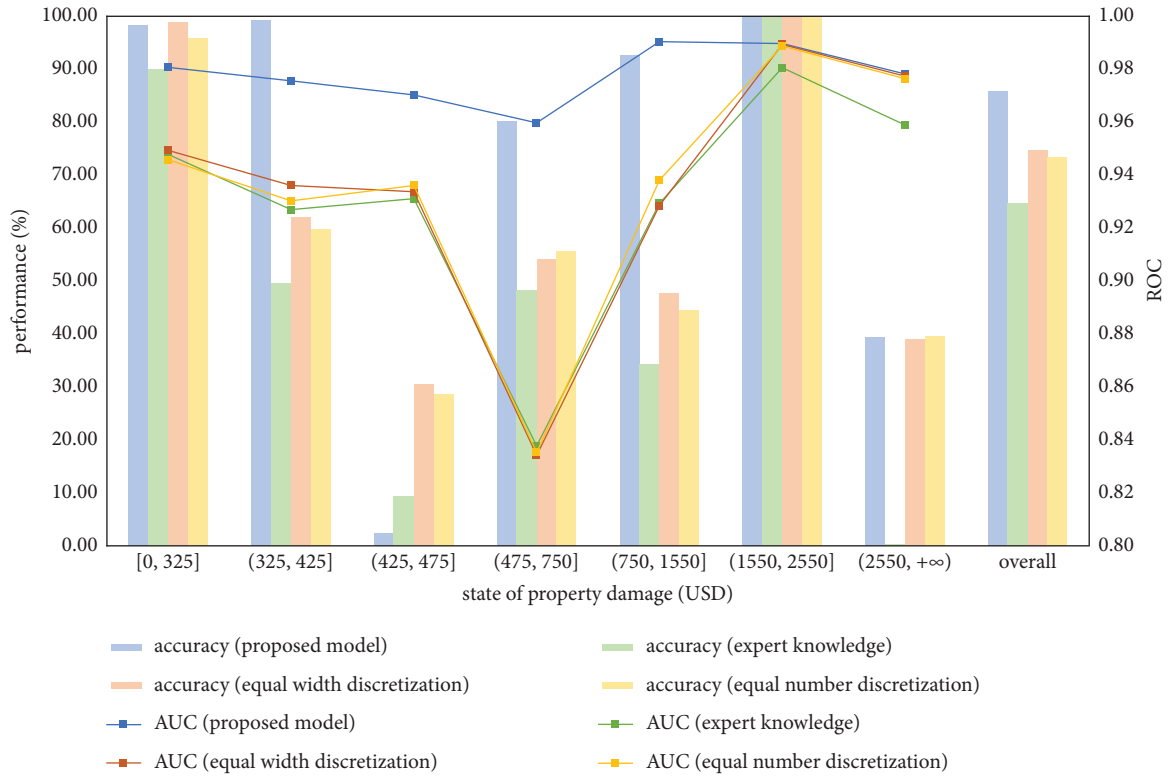
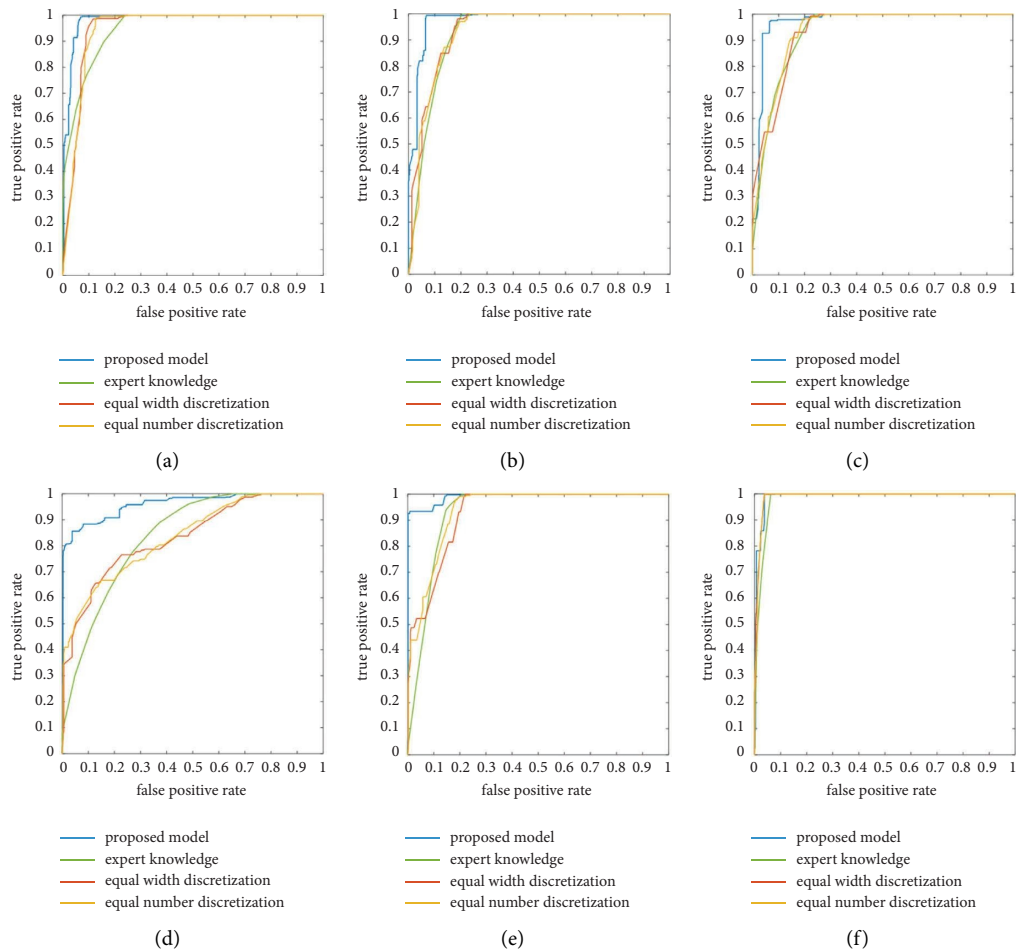FIGURE 7: Comparison of accuracy and AUC results of different models.
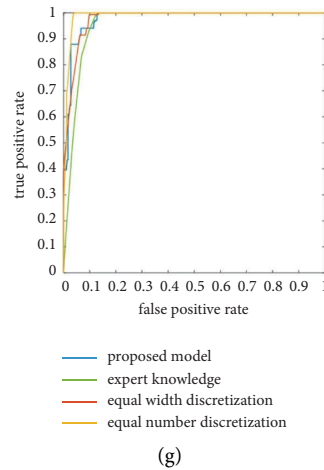

FIGURE 8: Continued.

(g)

FIGURE 8: Comparison of ROC curves of different models in predicting property damage. (a) [0, 325], (b) (325, 425], (c) (425, 475], (d) (475, 750], (e) (750, 1550], (f) (1550, 2550], (g) (2550, +∞).
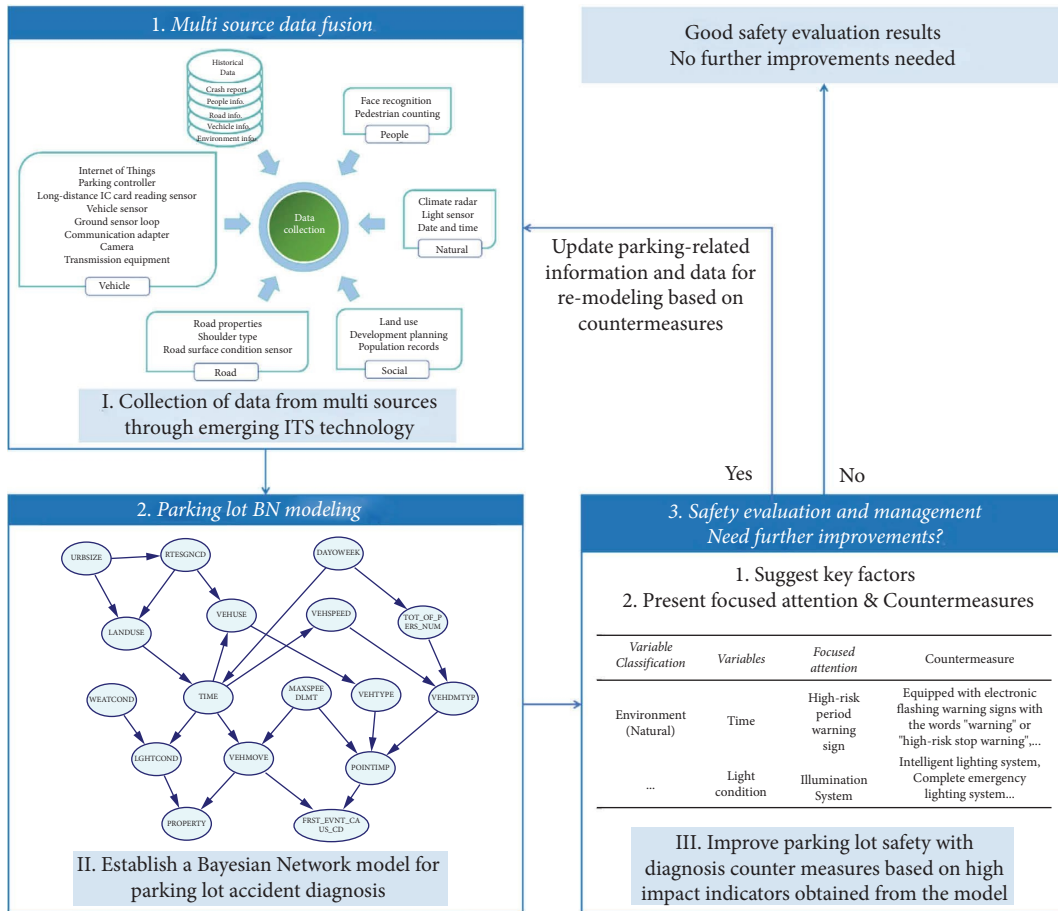

FIGURE 9: A Bayesian network-based parking lot safety evaluation system.

density areas. Focused monitoring of attention indicators allows for the rational and efficient use of resources for allocation. With the impending arrival of new technologies, the Bayesian network-based parking lot safety evaluation system is promising in more future scenarios. This provides efficient and profound assistance and insight to decision-makers.

TABLE 8: Library of focused attention on controllable variables and corresponding measures.

Variable classification	Variable	Focused attention	Countermeasure
Environment (natural)	Time	High-risk period warning sign	Equip electronic flashing warning signs that can automatically identify high-risk periods and flash to remind drivers to drive carefully
	Light condition	Illumination system	Intelligent adaptive lighting system; induction lighting system installed on the roads around the parking lot; complete emergency lighting system
	Weather condition	Real-time monitoring of hazardous weather	Dynamic collection of dynamic weather changes; timely flashing of active warning signals in adverse weather conditions; adaptive fog light system and ground moisture warning signs
	Day of week	Peak time management	Real-time monitoring of parking lot activity, automatic identification of peak hours, strengthening equipment monitoring; safety management during peak hours
People	Total number of people involved	Crowd management	Combine facial recognition and crowd counting for crowd management; focus on the number of people in the car; control the number of vehicles entering the parking lot at one time during high-risk hours
Vehicle	Maximum speed limit	Speed limit regulation setting	Automatic speed measurements; periodic check speed limit settings; set dynamic speed limit policies if necessary
	Vehicle movement	Vehicle route guidance	Installation of dynamic guidance algorithms to avoid potential vehicle congestion and collisions; equip automatic guidance systems for dynamic traffic flow and indications of safe uncongested roadways; installation of permanent warning signs for lane changes and entry with night-time reflective material
	Vehicle speed	Speed monitoring and warning	Real-time speed detection; dynamic information alert and overspeed warning; corresponding penalty measures
	Vehicle use and type	Vehicle type identification	Automatic identification of vehicle types and sizes at the entrance; separate parking for different vehicle types; exclusive management of high-risk vehicles

TABLE 9: Variable names and abbreviations in the database.

Variable name (abbreviation)	Explanation
DHSMV day of week (DAYOWEEK)	—
First cause of accident event (FRST_EVNT_CAUS_CD)	Describes the primary cause of a vehicle accident, such as a vehicle colliding while backing up or a collision with a parked car, traffic gate, or the fixed object
Federal functional classification (FUNCLASS)	—
Location code (HWYLOCAL)	The area in which the accident occurred (within the city, urban or rural area)
The prevailing type of land use (LANDUSE)	Described as commercial or residential land use, with the density situation
Light condition (LGHTCOND)	—
Maximum speed limit (mph) (MAXSPEEDLMT)	The maximum speed limit in the parking area, divided into 10, 15, 30, 40, and 45 five cases
Highway median width* (MEDWIDTH)	—
Point of impact** (POINTIMP)	Describe the part of the vehicle affected in an accident, including 23 cases such as the right front corner, right rear corner, and left front door
Total damage amount* (PROPERTY)	—
Road surface condition (RDSURFCD)	—
Route signing qualifier (RTESGNCD)	Describe the signing and use of the road, divided into nine cases: alternate, business route, bypass, spur, loop, proposed, temporary, and truck route
Number of people using safety equipment (SAF_EQUIP_USED_NUM)	—
Highway shoulder type, adjacent to outside travel lane (SHLDTYPE)	—
Time that accident has occurred (TIME)	Divide the 24-hour day into four time periods
Total number of drivers involved (TOT_OF_DR_NUM)	—
Total number of people involved (TOT_OF_PERS_NUM)	—
Total number of vehicles involved (TOT_OF_VHCL_NUM)	—

TABLE 9: Continued.

Variable name (abbreviation)	Explanation
Direction of travel (TRAVDIR)	—
Type of roadway parking (TYPEPARK)	Types of parking spaces, divided into ten categories, including whether there are curbs on both sides, with or without guardrails, and so on
Type of road (TYPEROAD)	—
Shoulder type (TYPESHLD)	Inside shoulder type
Urban size (URBSIZE)	Classification of urban types by different population size
Type of vehicle damage (VEHDMTYP)	—
Vehicle fault condition (VEHFLTCD)	—
Vehicle movement* (VEHMOVE)	Describe the movement state of the vehicle when an accident occurs, including fourteen states such as straight ahead, slow down, making left/right turn, and backward
Vehicle speed at the time of the crash as estimated (VEHSPEED)	The speed range at the time of the accident was divided into four categories, respectively
Vehicle type (VEHTYPE)	—
Vehicle use (VEHUSE)	—
Weather condition (WEATCOND)	—

5. Conclusion

In this study, a Bayesian network diagnosis model is developed to analyze property damage in parking lot accidents using multisource data. The mutual information approach for factor selection before modeling provides an objective basis for a priori network construction and variable screening, which reduces the subjectivity characteristic of Bayesian methods. The accuracy of the final prediction for property damage and accident causes reached 85.71% and 64.42%, respectively.

The direct influences on property damage are the light conditions and vehicle movement. Bad weather, dim lighting, and late afternoon also cause higher amounts of damage. Parking lot speed limits can raise the risk of vehicle movement that would otherwise not exist, which increases accident losses. Drivers should be aware of left turns, backing, and straight-ahead behaviors, as well as bypass areas with high traffic volumes. Commercial cargo and heavy truck drivers should pay extra attention to their surroundings. The number of people in the vehicle is positively correlated with property damage, and all other vehicle-related factors should be considered. The combination of factors with the most property damage is the maximum speed limit ≤ 30 mph $\longrightarrow$ slowing/stopped/stalled, and the most complex factors involved are bypass $\longrightarrow$ large urbanized $\longrightarrow$ high-density residential $\longrightarrow$ 18:00–24:00 $\longrightarrow$ dark (streetlights).

The conclusions also indicate that the point of impact and vehicle movement directly determine the cause of accidents. High-density residential and commercial areas require extra attention to collisions between vehicles and gates, especially for heavy trucks and commercial cargo. Gate geometry and layout design should be major considerations in parking lot planning and designs. Inadequate lighting and increased vehicle speeds can lead to collisions between vehicles and parking facilities.

A safety evaluation system applicable to parking lots is proposed based on the model. The system consists of three components: multisource data collection, data-driven Bayesian modeling of key factors, and safety management and measure implementation. The incorporation of emerging technologies allows the system to have a good adaptability and application prospects in various scenarios. However, the study has some shortcomings. With the emergence of new technologies, data collection and acquisition will become increasingly smart and accurate. Better uniform and reasonable data distributions give additional data dimensions, which will be examined and discussed in future studies.

Appendix

The definition of the variables in the database and their abbreviations is shown in Table 9. The continuous data are labeled with*, and those with too many categories requiring reclassification are labeled with**.

Data Availability

The data and record of accident reports are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J.-H. Shin and H.-B. Jun, "A study on smart parking guidance algorithm," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 299–317, 2014.

[2] J.-H. Shin, N. Kim, H.-b. Jun, and D. Y. Kim, "A dynamic information-based parking guidance for megacities considering both public and private parking," *Journal of Advanced Transportation*, vol. 2017, Article ID 9452506, 19 pages, 2017.

[3] V. Hassija, V. Saxena, V. Chamola, and F. R. Yu, "A parking slot allocation framework based on virtual voting and adaptive pricing algorithm," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5945–5957, 2020.

[4] M. Jakob, M. Menendez, and J. Cao, "A dynamic macroscopic parking pricing and decision model," *Transportation Business: Transport Dynamics*, vol. 8, no. 1, pp. 307–331, 2020.

[5] P. Papadimitratos, A. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 84–95, 2009.

[6] M. Essa and T. Sayed, "Full Bayesian conflict-based models for real time safety evaluation of signalized intersections," *Accident Analysis & Prevention*, vol. 129, pp. 367–381, 2019.

[7] L. Zhu, L. Lu, W. Zhang, Y. Zhao, and M. Song, "Analysis of accident severity for curved roadways based on bayesian networks," *Sustainability*, vol. 11, no. 8, p. 2223, 2019.

[8] M. Sun, R. Zhou, C. Jiao, and X. Sun, "Severity analysis of hazardous material road transportation crashes with a bayesian network using Highway safety information system data," *International Journal of Environmental Research and Public Health*, vol. 19, no. 7, p. 4002, 2022.

[9] C. Chen, X. Liu, H.-H. Chen, M. Li, and L. Zhao, "A rear-end collision risk evaluation and control scheme using a bayesian network model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 264–284, 2019.

[10] B. Kutela, R. E. Avelar, and P. Bansal, "Modeling automated vehicle crashes with a focus on vehicle at-fault, collision type, and injury outcome," *Journal of Transportation Engineering, Part A: Systems*, vol. 148, no. 6, Article ID 04022024, 2022.

[11] T. Singer, J. Kobbert, B. Zandi, and T. Q. Khanh, "Displaying the driving state of automated vehicles to other road users: an international, virtual reality-based study as a first step for the harmonized regulations of novel signaling devices," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 2904–2918, 2022.

[12] J. K. Yin and W. P. Fu, "A safety navigation method for integrating global path planning and local obstacle avoidance for self-driving cars in a dynamic environment," *Scientia Iranica*, vol. 28, no. 3, pp. 1318–1328, 2021.

[13] Z. Zhou, S. Yamada, Y. Watanabe, and H. Takada, "Tracking pedestrians under occlusion in parking space," *Computer Systems Science and Engineering*, vol. 44, no. 3, pp. 2109–2127, 2023.

[14] H. Yao, F. Feng, and J. Chen, "Practical application of vehicle-road cooperation based on high precision map and multisensor fusion positioning," *Journal of Geomatics*, vol. 47, no. 3, pp. 65–69, 2022.

[15] Y. He, J. Cui, F. Liu, and C. Zhu, "Risk-based quality accident ranking approach using failure mechanism and Axiomatic domain mapping," *Total Quality Management and Business Excellence*, vol. 31, no. 7-8, pp. 847–868, 2020.

[16] X. Gu, J. Guo, L. Xiao, and C. Li, "Conditional mutual information-based feature selection algorithm for maximal relevance minimal redundancy," *Applied Intelligence*, vol. 52, no. 2, pp. 1436–1447, 2022.

[17] S. Chen, X. Ma, L. Liu, and L. Wang, "Selective AnDE based on attributes ranking by Maximin conditional mutual information (MMCMI)," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 35, pp. 151–170, 2022.

[18] C. M. Ferreira-Vanegas, J. I. Vélez, and G. A. García-Llinás, "Analytical methods and determinants of frequency and severity of road accidents: a 20-year systematic literature review," *Journal of Advanced Transportation*, vol. 2022, Article ID 7239464, 17 pages, 2022.

[19] S. Wu, Q. Yuan, Z. Yan, and Q. Xu, "Analyzing accident injury severity via an extreme gradient boosting (XGBoost) model," *Journal of Advanced Transportation*, vol. 2021, Article ID 3771640, 11 pages, 2021.

[20] X. Liu, J. Lu, Z. Cheng, and X. Ma, "A dynamic bayesian network-based real-time crash prediction model for urban elevated expressway," *Journal of Advanced Transportation*, vol. 2021, Article ID 5569143, 12 pages, 2021.

[21] D. Li and M. F. M. Al-Mahamda, "Collective risk ranking of Highway segments on the basis of severity-weighted crash rates," *Journal of Advanced Transportation*, vol. 2020, Article ID 8837762, 9 pages, 2020.

[22] B. Wu, T. L. Yip, X. Yan, and Z. Mao, "A mutual information-based bayesian network model for consequence estimation of navigational accidents in the yangtze river," *Journal of Navigation*, vol. 73, no. 3, pp. 559–580, 2020.

[23] J. Yang, H. Len, J. Wang, and Y. Zhao, "Reliability assessment model and simulation of journal bearing of railway freight cars based on bayesian method under small sample sizes," *Journal of Advanced Transportation*, vol. 2022, Article ID 5275843, 10 pages, 2022.

[24] T. Lyu, Z. Zhang, L. Lu, and Z. Yanmeng, "An analysis of traffic accident severity based on mutual-information bayesian network," *Journal of Transport Information and Safety*, vol. 39, no. 6, pp. 36–43, 2021.

[25] X. Mao, C. Yuan, J. Gan, and J. Zhou, "Optimal evacuation strategy for parking lots considering the dynamic background traffic flows," *International Journal of Environmental Research and Public Health*, vol. 16, no. 12, p. 2194, 2019.

[26] J. Gan, X. Mao, and J. Zhao, "Optimal evacuation model of parking vehicles in dynamic traffic flows," *Journal of Southwest Jiaotong University*, vol. 56, no. 1, pp. 123–130, 2021.

[27] R. Bhagavathula and R. B. Gibbons, "Light levels for parking facilities based on empirical evaluation of visual performance and user perceptions," *Leukos*, vol. 16, no. 2, pp. 115–136, 2020.

[28] M. Yahaya, R. Guo, W. Fan et al., "Bayesian networks for imbalance data to investigate the contributing factors to fatal injury crashes on the ghanaian highways," *Accident Analysis & Prevention*, vol. 150, Article ID 105936, 2021.

[29] J. Liu, J. Li, K. Wang, J. Zhao, H. Cong, and P. He, "Exploring factors affecting the severity of night-time vehicle accidents under low illumination conditions," *Advances in Mechanical Engineering*, vol. 11, no. 4, Article ID 168781401984094, 2019.

[30] J. D. Bullough, J. D. Snyder, and K. Kiefer, "Impacts of average illuminance, spectral distribution, and uniformity on brightness and safety perceptions under parking lot lighting," *Lighting Research and Technology*, vol. 52, no. 5, pp. 626–640, 2019.

[31] G. Xue and H. Wen, "Pedestrian-injury severity analysis in pedestrian-vehicle crashes with familiar and unfamiliar drivers," *Transportmetrica a-Transport Science*, pp. 1–28, 2022.

[32] H. Tu, Z. Li, H. Li, K. Zhang, and L. Sun, "Driving simulator fidelity and emergency driving behavior," *Transportation Research Record*, vol. 2518, no. 1, pp. 113–121, 2015.

[33] C. Cabrera-Arnau, R. Prieto Curiel, and S. R. Bishop, "Uncovering the behaviour of road accidents in urban areas," *Royal Society Open Science*, vol. 7, no. 4, Article ID 191739, 2020.

[34] F. Afshar, S. Seyedabrishami, and S. Moridpour, "Application of Extremely Randomised Trees for exploring influential

factors on variant crash severity data," *Scientific Reports*, vol. 12, no. 1, Article ID 11476, 2022.

[35] Y. Duan and Y. He, "Causes analysis on severity of pedestrian collision accidents with different vehicle types based on Logistic," *Journal of Safety Science and Technology*, vol. 18, no. 1, pp. 175–181, 2022.

[36] D. J. Lin, M. Y. Chen, H. S. Chiang, and P. K. Sharma, "Intelligent traffic accident prediction model for Internet of vehicles with deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2340–2349, 2022.

[37] A. Abdi Kordani, S. Javadi, and A. Fallah, "The effect of shoulder on safety of highways in horizontal curves: with focus on roll angle," *KSCE Journal of Civil Engineering*, vol. 22, no. 8, pp. 3153–3161, 2018.

[38] F. Chen, S. Chen, and X. Ma, "Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data," *Journal of Safety Research*, vol. 65, pp. 153–159, 2018.

[39] H. Liu, J. Sun, L. Liu, and H. Zhang, "Feature selection with dynamic mutual information," *Pattern Recognition*, vol. 42, no. 7, pp. 1330–1339, 2009.

[40] C. J. Tsai, C. I. Lee, and W. P. Yang, "A discretization algorithm based on Class-Attribute Contingency Coefficient," *Information Sciences*, vol. 178, no. 3, pp. 714–731, 2008.

[41] L. Yu, L. Jiang, D. Wang, and L. Zhang, "Attribute value weighted average of one-dependence estimators," *Entropy*, vol. 19, no. 9, p. 501, 2017.

[42] H. Chen, Y. Zhao, and X. Ma, "Critical factors analysis of severe traffic accidents based on bayesian network in China," *Journal of Advanced Transportation*, vol. 2020, no. 2, Article ID 8878265, 14 pages, 2020.

[43] C. Jie, D. A. Bell, and W. Liu, "An algorithm for bayesian belief network construction from data," in *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics*, pp. 83–90, Istanbul, Turkey, November 1997.

[44] J. Wu and M. Yang, "Assessment of passenger satisfaction with urban bus service quality using Bayesian networks," *Journal of Southeast University (Natural Science Edition)*, vol. 47, no. 5, pp. 1042–1047, 2017.

[45] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–22, 1977.

[46] U. Kjærulff and L. C. van der Gaag, "Making sensitivity analysis computationally efficient," 2013, https://arxiv.org/abs/1301.3868.

[47] Y. Zhang and Y. Yang, "Cross-validation for selecting a model selection procedure," *Journal of Econometrics*, vol. 187, no. 1, pp. 95–112, 2015.

[48] X. Cheng, Z. Huang, and S. Chen, "Vehicular communication channel measurement, modelling, and application for beyond 5G and 6G," *IET Communications*, vol. 14, no. 19, pp. 3303–3311, 2020.