

A review of the research status of the cause analysis and severity prediction of traffic accidents

Hanhan Wang

School of transportation, Chongqing Jiaotong University, Chongqing 400000, China

Abstract

With the continuous increase of the number of motor vehicles, traffic accidents have become a social problem threatening public property and life. The causes of traffic accidents are mainly divided into four major factors: people, vehicles, roads and environment. This topic aims at the characteristics of existing traffic accident data and the key problems in data analysis and application. Based on relevant theories and methods of data mining such as classification and regression, a traffic accident analysis system is established. Through the construction of traffic accident causation analysis model and traffic accident severity classification and recognition model, the causality relationship between traffic accidents and drivers, vehicles, roads, environment and other factors is studied from multiple perspectives and levels. It provides theoretical support for reducing the incidence of traffic accidents, reducing the severity of accidents and improving traffic safety.

Keywords

Causes analysis; Key factors; Classification prediction.

1. Preface

As a kind of clothing, food, housing and transportation, transportation has a very significant impact on people's lives, especially today with the development of transportation. Due to the importance of traffic in life, it becomes a very necessary research to analyze the influencing factors of traffic accidents. The fact is that many scholars at home and abroad have carried out relevant analysis and research on traffic accidents, and have achieved fruitful results. Some characteristics of people and vehicles are often subjective factors in the occurrence of a traffic accident, and the severity of an accident generally determines how much property damage an accident will cause, as well as serious casualties.

2. Research background and significance

2.1. Research background

With the development of the world economy, car ownership is gradually increasing, and road mileage is also increasing. By studying the statistics of the World Automobile Organization in 2015, it can be seen that by the end of 2015, the number of cars in the world was 1,282,269,600, an increase of 47,382,900 over the previous year, a year-on-year increase of 3.8%, and the growth rate was higher than that of the previous year. It slowed by 0.4 percentage points in the previous year. Among them, the top three car ownerships are the United States, China and Japan, while the UK ranks ninth in terms of car ownership, with 38.22 million vehicles. The United Kingdom started to build highways in 1958, and it was not too late, but due to its slow construction speed, its overall construction speed and construction level are now more backward than other developed countries. The UK transport network consists of road, air, rail and water transport. The total length of British roads is 46,904 kilometers, of which 3,497 kilometers are highways and 344,000 kilometers are ordinary roads. Due to the occurrence of

traffic accidents, the number of deaths is as high as 100,000 people every year, and the probability of death due to traffic accidents is decreasing on a marked rise.

2.2. Research background

With the continuous increase in the number of motor vehicles, traffic accidents have become a social problem that threatens public property and life. Studying data analysis methods and forecasting methods, developed countries such as the United States, Japan, and Canada are at the forefront of the world in road transportation and informatization construction. The causes of traffic accidents are mainly divided into four major factors: people, vehicles, roads, and the environment. And today's traffic accidents are characterized by a large amount of data and complex relationships. The analysis of traffic accidents has begun as early as the 19th century. The main methods are divided into three categories: methods based on statistics, methods based on data mining and methods based on data simulation. Methods, the currently widely used and mature research ideas are to use the relevant theories and methods of data mining such as classification and regression to establish a traffic accident analysis system based on the characteristics of existing traffic accident data and key issues in data analysis and application. Due to the analysis model and the classification and identification model of the severity of traffic accidents, the causal relationship between traffic accidents and factors such as drivers, vehicles, roads, and the environment is studied from multiple angles and levels, in order to reduce the incidence of traffic accidents and reduce the severity of accidents. It provides theoretical support for improving driving safety, etc.

3. Research status at home and abroad

So far, scholars at home and abroad have carried out a series of studies on car accidents, analyzing the influence of different factors on car accidents. The research mainly includes the following aspects: optimization of the on-site data acquisition system and processing system, processing the on-site collected data with various models, judging the impact of different factors on the car accident, and on-site reproduction from the data. The following is a partial literature review and comment on the current research status at home and abroad.

3.1. Optimization of accident scene data acquisition system

When the data collected is more in line with the requirements, it is more convenient for later processing. Wu Wenke, through research on the accident scene investigation process, and through statistical and comparative analysis of traditional positioning methods, proposes a vehicle-led improved comprehensive positioning method, and use the improved comprehensive positioning method to parameterize the vehicle accident scene information [1]. On the one hand, the improved method guides the surveyors to complete the survey, on the other hand, because the data is directly imported into the Crashview software, it can be used for the later accident simulation reproduction, which makes the on-site identification work simpler and more convenient and more beneficial for the later data processing.

3.2. Analysis of the influence of various factors on traffic accidents

Car accidents are caused by a variety of factors, and many scholars at home and abroad have used various methods to analyze the factors and study the influence of each factor. Yang Wenchen et al. studied the impact of adverse weather on highway traffic safety based on empirical research in the past 30 years [2]; Liu Xinyue et al. excavated the correlation between traffic accidents and factors such as weather, time, and roads [3]; Zhang Zhuoshuang analyzed the Guangzhou City 2015 Carry out research on bus accidents that occurred in 2018, and analyze the factors that affect bus safety [4]; Chen Hui uses binary and multivariate Logistic regression models to analyze the influencing factors of major accidents [5]; Huang Wei uses

accident correlation rules to The causes of road traffic accidents in Xi'an are analyzed and the countermeasures are studied [6]; Tang Zuogan considers the heterogeneity and uses the mixed logit model and the latent category logit model to deal with the influencing factors of the rollover accident [7]; Jiang Shan et al. Apriori algorithm conducts data mining to analyze the causes of traffic accidents [8]; Zhang Cuiping uses statistical negative binomial distribution regression model and quasi-Poisson model to process data from three levels of intersection, road section and traffic area. Safety problems were analyzed, accident risks were analyzed, and future situations were predicted [9]; Yan Zhangcun analyzed the applicability of discrete choice models with solid collision accidents combined with accident characteristics, and summed up the ordered Logit model and the multinomial logit model in the definite variance Logit model. The Logit model is used to model, solve and test the theory of accident severity analysis, and build a complete analysis process [10]; Li Fu uses the fault tree analysis method for typical traffic accidents, and uses the cause analysis theory to determine the source of danger [11]; Chen Jingxi According to the requirements of different analysis levels, PCA feature extraction algorithm, Apriori association rule algorithm and K-Means clustering algorithm were selected to analyze the causes of accidents [12]; Thomas et al. applied linear and nonlinear multivariate statistical analysis to determine the severity of accidents and The relationship between traffic flow, weather and ambient lighting conditions [13]; Zhaoyou et al. conducted research on 8 representative road tunnels, revealing potential factors affecting tunnel traffic safety [14]; Yamazaki et al. Public data, urban population data and meteorological data, other findings about traffic accidents were found [15]; ZHANG et al. established the total duration and clearing time based on 9 variables including time, space, environment, traffic and accident details, respectively. Multiple linear regression model [16]; Jian et al. analyzed the causes of accidents, looked for weak links and mistakes in safety management from the aspects of people, vehicles, road environment, and management, and improved relevant policy systems to provide support for strengthening safety management in the transportation industry [17].

3.3. Field reproduction from data

Compared with graphs of data, using data to build a model to reproduce the scene allows people to perform accident analysis more intuitively. According to the characteristics of vehicle mechanics, Qin Yan uses C++ as a platform to build a dynamic model of the collision site between motorcycles and cars, and analyzes the collision accidents between cars and motorcycles [18]; The main component is to use the model construction platform with 3dsMax software as the core to carry out 3D modeling and analysis of the vehicle operating environment [19]; Lv Jihua, through the study of accident mechanics and deformation, uses CAD and HyperMesh to construct a vehicle-rigid wall finite element model. Model, using LS-DYNA to solve the model, using HyperView to reproduce the three-dimensional animation of the car crash deformation process to reproduce the car accident [20]; Zhao Chen compared the operation process of PC-CRASH and Crashview, vehicle loading, scene Drawing, parameter input and output results finally determine the 2D process reconstruction and 3D simulation reproduction of Crashview for 6 typical accident form cases [21]; Zhang Wen-hui obtained the changes of various factors at the accident scene through analysis and system simulation. relationship with changes in security [22].

3.4. Field reproduction from data

When data is available, in addition to studying the main influencing factors, factors can also be linked to analyze the possibility of future accidents and carry out risk assessment. Fang is a data mining framework for driver traffic safety evaluation based on drivers' personal information, traffic illegal behaviors and traffic accident data records, which can be used to make drivers avoid some dangerous behaviors and reduce the possibility of accidents to a certain extent [23].

3.5. Summary of domestic and international status

So far, the research on traffic accidents has never stopped, including the analysis of factors affecting traffic accidents, the optimization of on-site data collection methods, and the reconstruction of on-site scenes by using existing data. With the further improvement of motorization, car accidents have become a major killer threatening the safety of human life, health and property. Therefore, how to reduce the possibility of accidents by analyzing the existing accident data is also the direction of scholars at home and abroad. As Internet + and the rise of big data, the base of data analysis is more and more big, the model and algorithm is in constant updates, after extensive study, I found that for traffic accident data processing research, is usually combined with the already existing data, using statistical methods to the preliminary analysis and data mining methods, to explore the correlation between the single factor, Then according to the characteristics of the data and the correlation between the selection of appropriate model, the Logit model and factor analysis is more suitable for the index of the single factor analysis of the severity of the accident, to carry out further data mining, it is necessary to analyze the correlation between each index, reduce the redundant value of the Apriori classic algorithm, At present, the main methods to reduce redundant values are association rules and clustering rules.

4. Data processing ideas

There are many mature processing methods for data processing of traffic accidents. Many predecessors in related fields have established various models, and in the research process, these models have been continuously improved. The general flow of processing is shown in figure 1 below:

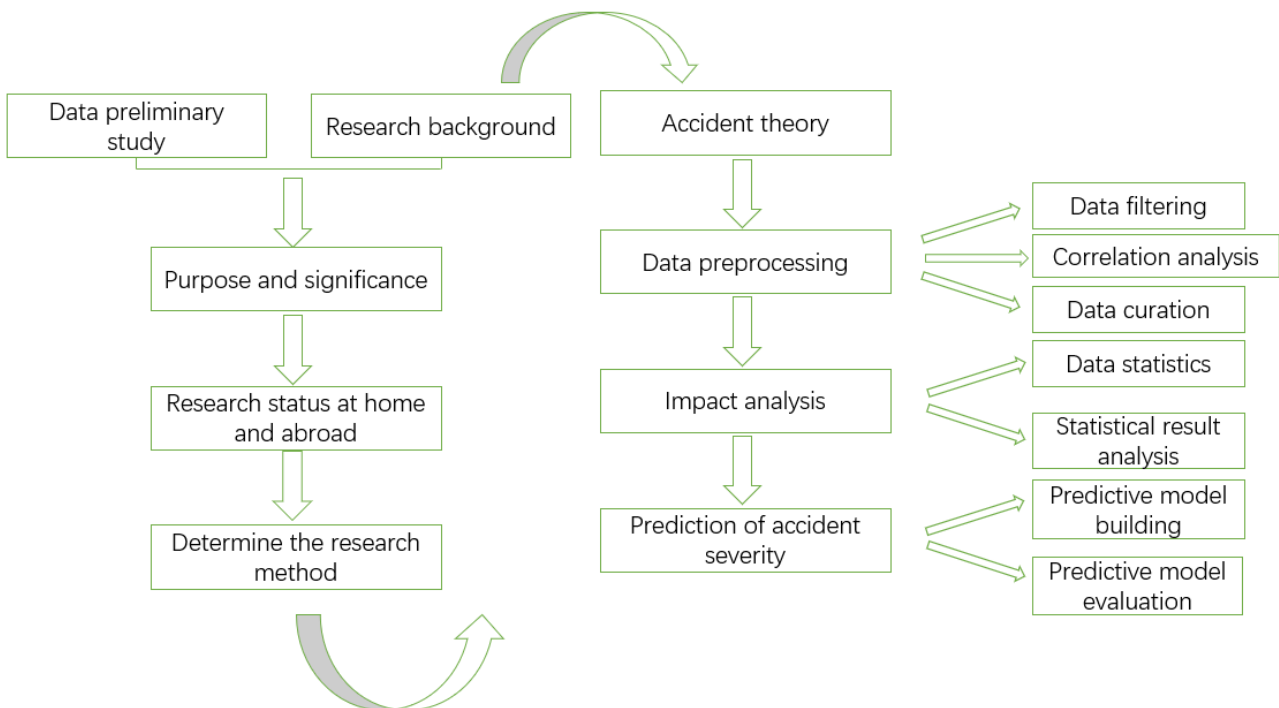


Figure 1: Data processing flow chart

5. Questions to discuss

From the birth of traffic accident analysis to today, various analysis methods and models have been widely used in accident causation and prediction. Models and Structural Equation Models, Decision Trees, Bayesian Nets. Each of these methods has its own scope of application. Among

them, the count data model is usually used to calculate the influencing factors of the accident frequency of a certain section or road, that is, to explore the main influencing factors of the accident frequency; the discrete selection model is mainly used for accidents. Consequence level division analysis, accident cause analysis and some unreasonable legal behavior judgment analysis; path model is mainly used to analyze the causal relationship between driving path conditions and driver psychology; structural equation model is usually used to analyze the relationship between psychological factors and driving behavior. Causal relationship; decision tree can be used to analyze the main factors causing accidents; Bayesian network can perform better data analysis for some cases where accident data is missing.

In today's increasingly diverse and sophisticated causal analysis methods, our country's road safety hazard detectors should not remain in the past counting data analysis model methods. The adaptive analysis method conducts anatomical research on the specific research object. In the actual situation on the spot, two or more analysis methods are used as much as possible to analyze the traffic accident in combination, or through various methods to compare, let the analysis The results are more precise and standardized to make my country's road traffic safety management more convincing.

Acknowledgements

Natural Science Foundation.

References

- [1] Wu Wenke. Development of Parametric Reconstruction System for Vehicle Accident Scene Information [D]. Chang 'an University,2017.
- [2] Yang Wen-chen, Li Li, Hu Cheng-yu, Zhang Lun, Fang Rui. Journal of Wuhan University of Technology (Traffic Science and Engineering),2019,43(05):843-849+854. (in Chinese)
- [3] Liu Xinyue, Xie Wenjun, Yin Chengsheng, Chen Jinguang, Liu Lu, Luo Yuetong. A Visual Analysis Method for Local Correlation of Urban Traffic Accidents [J]. Journal of Graphics,2019,40(05):843-851.843-851.
- [4] Zhang Zhuoshuang. Research on Safety Management of Bus Operation in Guangzhou [D]. South China University of Technology,2019.
- [5] Chen Hui. Influence Mechanism of Human and Vehicle Physical Behavior in Heavy Casualty Accidents [D]. Beijing Jiaotong University,2019.
- [6] Huang Wei. Research on Causes and Countermeasures of Road Traffic Accidents in Xi 'an [D]. Xi 'an University of Science and Technology,2017.
- [7] Tang Zugan. Comparative Study on Model of Injury Degree of Rollover Accident Considering Heterogeneity Effect [D]. South China University of Technology,2019.
- [8] Jiang Shan, Song Ke, Xie Weicheng, Pan Chengwei. Analysis of Road Traffic Accident Data Based on Grey Relation and Apriori Algorithm [J]. Highway Engineering,2019,44(04):67-73.
- [9] Zhang Cuiping. Study on Risk Analysis and Prediction Method of Road Traffic Accidents Based on Historical Data [D]. Beijing Jiaotong University,2017.
- [10] Yan Zhangcun. Analysis of Influencing Factors of Accident Severity Based on Variable Variance Ordered Logit Model [D]. Southwest Jiaotong University,2018.
- [11] Li Fu. Research on Vehicle Operation Risk Based on Typical Traffic Accident Analysis [D]. Jilin University,2013.
- [12] Chen Jingxi. Traffic Accident Analysis Based on K-means [D]. Chang 'an University,2018.
- [13] Thomas F. Golob, Wilfred W. Recker. Relationships Among Urban Freeway Accidents, Traffic Flow, Weather, and Lighting Conditions [J]. JOURNAL OF TRANSPORTATION ENGINEERING, 2003 (342): 342-353.

- [14] M. Zhaoyou, W. Changjun, F. Shouen and L. Shuo, Comparative Analysis and Control Strategy for Traffic Accidents in Different Types of Tunnels, [J] .Transportation Information and Safety,2019(5):1132-1136.
- [15] T. Yamazaki, "Analysis of Traffic Accident Occurrence in Niigata Prefecture of Japan using Open Data," [J] .2019 IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 2019, pp. 1-4.
- [16] J. ZHANG, J. WANG and S. FANG, "Prediction of urban expressway total traffic accident duration based on multiple linear regression and artificial neural network," [J] .2019 5th International Conference on Transportation Information and Safety (ICTIS), Liverpool, United Kingdom, 2019, pp. 503-510.
- [17] T. Jian, L. Zhi-qiang, X. Jian-feng and G. Hong-yu, "Cause Analysis and Countermeasures of fatal Traffic Accidents on Road Passenger Transportation Based on Typical Cases," [J] .2019 5th International Conference on Transportation Information and Safety (ICTIS), Liverpool, United Kingdom, 2019, pp. 951-955.
- [18] Qin Yan. Research on Analysis System of Motorcycle and Automobile Crash [D]. Chang 'an University,2018.
- [19] Lian Shikun. Research on 3D Modeling Technology of Vehicle Running Environment [D]. Chang 'an University,2011.
- [20] Lu Jihua. Research on Reproduction of Automobile Crash Accident Based on Shape Change [D]. Jilin University,2018.
- [21] Zhao Chen. Simulation Analysis of Road Traffic Accidents Based on CrashView [D]. Chang 'an University,2012.
- [22] Zhang Wen-hui, Deng Hong-xing and Wang Xian-bin,Safety factor analysis for traffic accident scene based on computer simulation [J].International Conference On Computer Design and Applications, 2010, pp: 89-91.
- [23] A. Fang, C. Qiu, L. Zhao and Y. Jin, "Driver Risk Assessment Using Traffic Violation and Accident Data by Machine Learning Approaches," [J] .2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE), Singapore, 2018, pp. 291-295.