Research on time and spatial characteristics based on GIS and the big data of city traffic accident

Zhanyong Fan[1], Dongjian Cai[2],Shun Yue[3], Yaofei Chang[4]
[1]Tongji University, Shanghai, China, 200092
Email: zhanyongfan@163.com
[2]Suzhou Industrial Park Surveying Mapping and Geoinformation Co.,Ltd. Suzhou, China, 215027
Email: caidongjian@dpark.com.cn
[3]Suzhou Industrial Park Surveying Mapping and Geoinformation Co.,Ltd. Suzhou, China, 215027
Email:yue_shun@163.com
[4]Tongji University, Shanghai, China, 200092
Email: cyf0923@tongji.edu.cn

Key words: GIS; Traffic accident; big data; black spot; time and space characteristics

Abstract: With the rapid growth of the national economy, the number of city motor vehicles and traffic accidents have been increasing. More and more people have paid attention to the road traffic safety. How to reduce the incidence of traffic accidents has become a serious problem which is the city traffic relevant departments urgently need to solve. This paper integrated the big traffic data with related multi-source data and constructed spatial correlation relationship by making use of the advantages of GIS based on multi-year traffic accident data of Suzhou Industrial Park. The paper improved the traffic accident black spot identification algorithm by machine learning methods, constructed the black spot model, and proposed improvement measures for different black spots. Then the paper made visual analysis for the traffic accident's distribution on time and space, the data structure and topological relationship. At the same time, the paper constructed the spatial evolution characteristics of road congestion and dissipation before and after the traffic accidents. The paper also revealed the implicit information from the traffic accidents data and constructed the system of road safety analysis and countermeasure library of the accident reasons. Finally, the paper developed analysis and management platform for the city traffic accident based on GIS, which provided preventive and decision-making technical support for traffic accident.

1 Introduction

With the continuous construction of urban road traffic network, people's demand for convenient transportation has driven the continuous increase of car ownership. Along with the high number of traffic accidents, road traffic safety has become a key issue related to the safety of people's lives and property, the quality and efficiency of economic and social development, which has attracted attention of the national security. How to minimize the incidence of traffic accidents and reduce the harm caused by traffic accidents has become an urgent problem, which is solved by the relevant departments of urban traffic [Li Shuqing,2014]. However, the traditional traffic infrastructure supply and the relatively lagging traffic accident management system have been unable to meet the requirements of modern transportation, which bring great obstacles to the road traffic planning and traffic control traffic planning [Niu Huiyong,2017].

With the extensive application of new technologies such as computer, internet, sensor and so on, the amount of data which the people get is increasing in the order of TB, GB, ZB, PB and so on. People's life begins to enter the "Big data" age. The traffic accident data also belong to it [Ji Qianqian,2015]. The rapid development of geographic information processing technology, information mining technology, spatial analysis, expression technology and information processing technology have provided a good platform for many researchers at home and abroad to study the big data of traffic accident [Xu Xin,2013]. Based on traffic accident data of Shanghai from 2007 to 2008, Shanghai Municipal Public Security Bureau realized semantic discrimination, identified the Shanghai road network accident black spots by analyzing Shanghai Road traffic the distribution and characteristics of safety accidents in 2010[Chen lin,2009]. Tongji University finished the investigation of four dangerous sections in Zhejiang, Yunnan and Guizhou provinces, investigated more than 200 incidents of accident-prone locations and accidents, analyzed the reason in depth, improve the countermeasures, and reported to the Ministry of Public Security investigation in 2013. Wanghai determined the accident-prone points, road segment and area by using buffer method and the optimized kernel density clustering method in 2014. Then, he established a complete road traffic accident system based on GIS platform, which included determining the accident-prone points and the analyzing of traffic accident reasons [Wang hai,2016]. In this paper, based on the big traffic accident data of Suzhou industrial park, it merged and pretreated the traffic accident data, put forward data entry standards, constructed the missing data processing method, found out the accident-prone locations by using spatial correlation analysis and other methods of GIS at first. Then, it studied the spatial evolution characteristics of traffic accident by combining with urban road network, township division, street lights, traffic flow information, accident-prone location driving behavior and other

geographic information data. Finally, it developed the urban traffic accident analysis and management platform which provided measures support for the traffic control department. At the same time, the traffic management person can control and analyze the traffic in time and space based on the platform, which provided reference value for the traffic plans of the road administration and traffic management department.

2 Black spot recognition model

Suzhou industrial park is located in the east of Suzhou, north latitude N31°17'34.39", east longitude E120°39'28.33". Its' area is about 278 square kilometers. The traffic accident data of this paper were from the traffic police brigade of Suzhou Industrial Park. The road network, road section and intersection road of the city were from Suzhou Industrial Park Survey and Mapping Geographic Information Co., Ltd. The historical traffic accident data were manually entered by the various detachment traffic policemen. So the quality of the data is uneven. We must deal with the data by geocoding, coordinate conversion, elimination of redundancy and other pretreatment before analyzing it.

The occurrence of traffic accidents has a certain chance. But the spatial distribution of traffic accidents is not random. So it is necessary to find out the number or characteristics of road traffic accidents in the longer period of time, which is obviously different from other normal locations. The people called it black spots. The nuclear density method is a commonly used black spots recognition algorithm [Tang Luliang,2017]. The calculation formula for nuclear density is

$$f(s)=\sum_{i=1}^{n}\frac{1}{h^2}k\left(\frac{s-c_i}{h}\right) \qquad (1)$$

In the formula (1), $f(s)$ is the kernel density calculation function at the spatial position s . h is the distance attenuation threshold. n is the number of points less than or equal to the distance h . The function k represents the spatial weight function. The kernel density value is the largest at each core element c_i . And the kernel density value is reduced to zero when the distance is h which is away from c_i . The traditional kernel density is based on the European linear distance, which ignores the conduction factors that the central factors depend on. It is difficult to objectively reflect the actual spatial feature distribution. And the network kernel density algorithm introduces the path distance of urban road network[Yu Wenhao,2015]. It can more accurately express the accident point along the urban street network distribution of the details[OKABE A,2009]. The traditional kernel density analysis and network kernel density analysis results were shown in Figure 1,2.

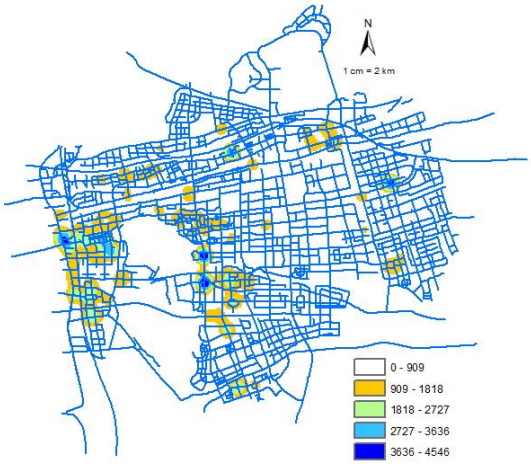

Figure 1 Traditional kernel density analysis result

Figure 2 Network kernel density analysis result

It can be seen from figure 1 and 2 that the general kernel density analysis and the network kernel density can identify the black spots on the road. However, the result of the traditional kernel density analysis is too vague. The positioning accuracy is low. The spatial distribution of the black spots cannot be identified. But the black spots identified by the network kernel density were accurately distributed in the city roads. Because the network kernel density had taken into account the road network and other geographic information. So the network kernel density can better identify the traffic accident black spots distributed than the traditional kernel density.

3 Spatial and Temporal Analysis of Traffic Accidents

In order to study the distribution of traffic accidents inside the Suzhou Industrial park, this paper dealt with the traffic accident data from 2010 to 2015.Then it distributed the traffic accident points into the road intersection and road sections by using the spatial connection function of GIS. The results were shown in Figure 3,4,5,6,7,8.

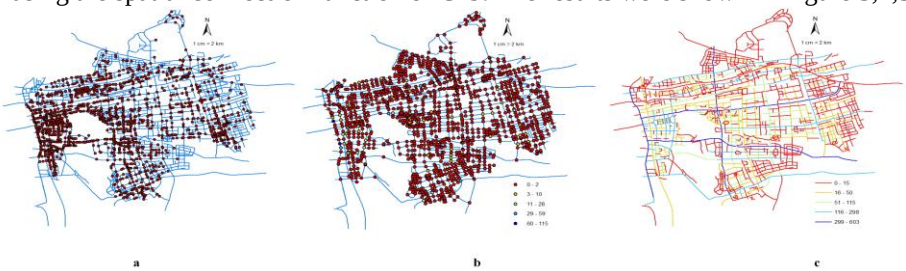

Figure 3 Spatial distribution of traffic accidents in 2010

Figure 4 Spatial distribution of traffic accidents in 2011

Figure 5 Spatial distribution of traffic accidents in 2012

Figure 6 Spatial distribution of traffic accidents in 2013

Figure 7 Spatial distribution of traffic accidents in 2014

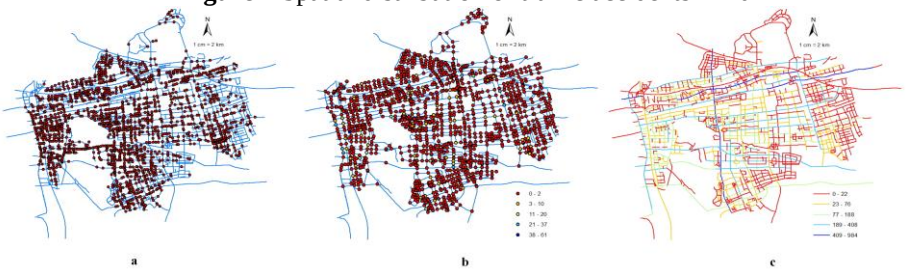

Figure 8 Spatial distribution of traffic accidents in 2015

It can be seen from figure 3 to figure 8 that figure (a), figure (b), figure (c), respectively represented the location of the accident occurred, the intersection and the road. The results showed that the spatial location of the annual traffic accident is mostly distributed in the eastern and western parts of Jinji Lake, Jinji Lake Avenue, Hyundai Avenue, Xinghu Street and other urban main roads. The overall traffic accident spatial distribution had change trend from west to east with time. The reason may be the western part of the Suzhou Industrial Park is the old city district. The urban facilities in the eastern part of the park will be improved from 2010 to 2015. In order to analyze the daily distribution of traffic accidents, this paper counted the number of traffic accidents per hour in six years. The results were shown in Figure 9.

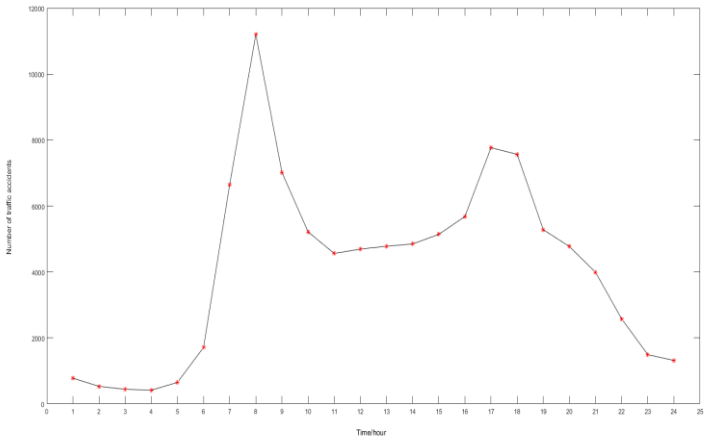

Figure 9 Counts of the number of accidents per hour in the day

It can be seen from Figure 9, traffic accidents occurred mainly in the morning 7:00 to 9:00 and 17:00 to 19:00 which was the peak of commuting. The traffic policemen should strengthen the traffic management of the key sections during this period.

4 Urban Traffic Accident Analysis and Management Platform Based on GIS

Geographical information system transportation (GIS-T) is a branch of the geographic information system, which is the specific application and extension of GIS in transportation field [Wang Xiaoyan, 2016]. It is the collecting, storing, analyzing and processing related to traffic geographical information system, which is the organic integration of GIS and traffic system [Qin Liyan, 2004]. The paper developed the urban traffic accident analysis and management platform based on the visual characteristics and powerful spatial analysis of GIS. The platform built the specialized system of traffic modeling based on the linear and network characteristics of traffic phenomena. The platform can effectively integrate, manage and analyze the traffic accident data which were so complex and isolated [Munyendo J N,2015].

The design of the platform conformed to the hierarchical design concept of service orientation and object. It maintained a low level of coupling between layers and layers. Parallel modules are developed and designed with the principle of high cohesion and low coupling among modules within modules. Between layers and modules, the unified information exchanging rules were used to invoke the services provided by other modules.

The realization of the platform using Browser/Server model, to achieve the purpose of cross platform and systems, solve the traffic accident kinds of archives management and analysis of the user terminal operation problems. The user's presentation layer in the browser environment invokes the business logic request and USES the TCP/IP transport protocol to perform data transmission and service invocation through a dedicated network. The architecture diagram and interface of the platform were shown in figure 10 and 11.

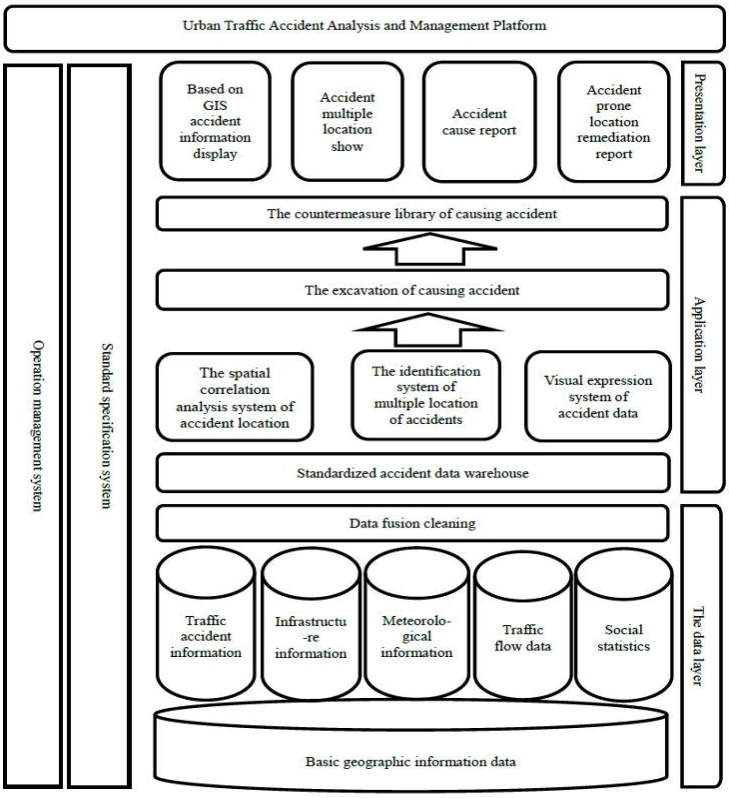

Figure 10 Urban traffic accident analysis and management platform's architecture diagram

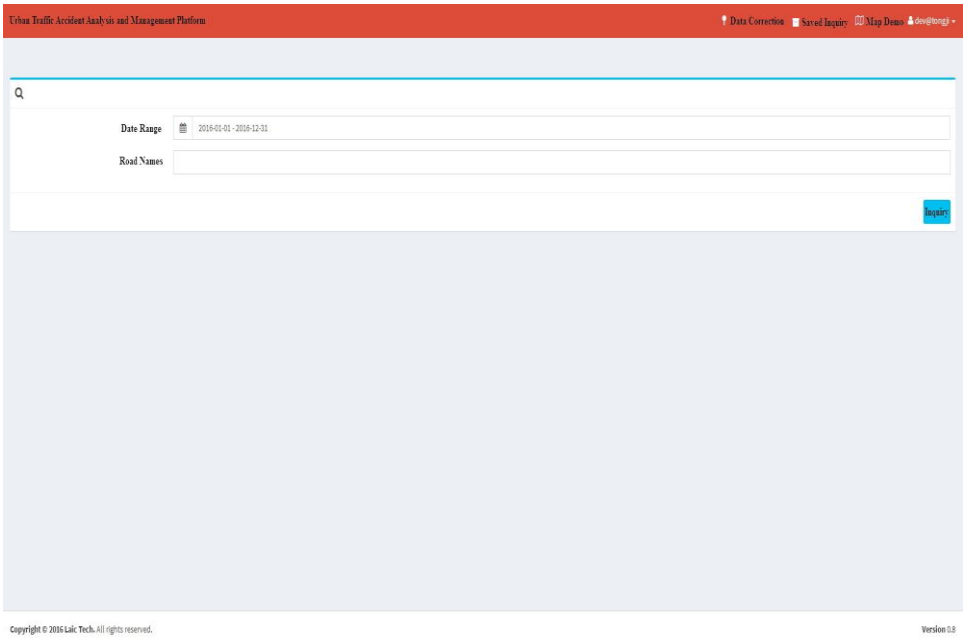

Figure 11 The platform's interface

As can be seen from figure 10, the platform was divided into three layers: data layer, application layer and presentation layer. The data layer was mainly integrated traffic accidents and traffic accidents related information, including the traffic accident information based on fundamental geographic information platform, infrastructure information, weather information, traffic flow data and social statistics data. The paper put together these data information, put forward data entry standard, cleaned data fusion, and finally formed standardized data warehouse. The application layer was mainly responsible for the business logic of the system and related calculation work, which found out the accidents in the city area, mined accident reasons by combined the feature of accident location's driving ways, established the accident reasons of library, put forward rectification opinions through analyzing the accident data. The presentation layer mainly showed traffic accidents information by GIS platform. It submitted the accident reason by the accident reason report, and provided some suggestions in the accident location for related traffic administrative department by the report of the remediation suggestions in multiple accidents locations.

5 Conclusion

Traffic accidents frequently brought a lot of inconvenience to people's lives. How to quickly and effectively mine out the effective information of traffic accidents, reduce the incidence of accidents is the urban traffic problems. This paper extracted the effective information of traffic accident by mining the big traffic accident data based on the data of traffic accident, spatial correlation data, road network, intersection and other data mining large data mining. The conclusions were as following:

(1) According to the traditional nuclear density is based on the lack of European distance, this paper identified urban traffic accident black spots quickly and accurately by using the method of network density which took into account the urban road network and other information;

(2) The paper concluded that the traffic accidents are mainly distributed in the western and eastern of Jinji Lake, Modern Avenue, Start Lake Street and other main road in the city by analyzing the traffic accidents from 2010 to 2015. Then the paper counted every hour in every day the number of the accident. The result showed that the traffic accidents occurred mainly in the 7:00-9:00, 17:00-19:00.

(3) Based on the powerful spatial analysis ability of GIS, the paper developed the urban traffic accident analysis and management platform. The platform can realize the effective integration of GIS and traffic, display the traffic accident information quickly and efficiently, and provide a making decision reference for transport agency.

References

Li Shuqing，Peng Youlang，Xiao Liying，Meng Simeng，Wang Peng，Xu Jipeng, Analysis of the mechanism of the road traffic accidents in-situ and the future research trends, Journal of Safety and Environment,2014,14(03): 14-19.

Niu Huiyong, Basic Characteristics and Preventive Measures of Road Traffic Accident in China, 2006,(07): 87-91.

Ji Qianqian, Wen Haoyu, Big Data Platform Structure in Public Transportation [J],Electronic Sci.&Tech. 2015, 28(2):127-130.

Xu Xin, Analysis of Road Traffic Accidents in China and Measures to Prevent Them, China Safety Science Journal, 2013,23(11):120-125.

Chen Lin, Wang Shenghe, Research on the Application of Association Rules in Road Traffic Accident, Journal of Chinese People's Public Security University, 2009,15(01):84-87.

Wang Hai, Li Ruimin. Buffer Analysis Method In The Identification of The Accident Black Point Application Research [J], Highway Engineering, 2016, 41(1):103-107.

Tang Luliang, Kan Zihan, Liu Huihui, etal．A Kernel Density Estimation Method for Linear Features in Network Space [J]．Acta Geodaetica et Cartographica Sinica,2017,46(1):107-113.

Yu Wenhao, Ai Tinghua．The Visualization and Analysis of POI Features under Network Space Supported by Kernel Density Estimation[J]．Acta Geodaeticaet Cartographica Sinica, 2015,44(1):82-90.

OKABE A, SATOH T, FURUTA T．A Kernel Density Estimation Method for Networks, Its Computational Method and a GIS-based Tool[J]．International Journal of Geographical Information Science,2009, 23(1):7-32.

Wang Xiaoyan, Zou Jianmin,Qiu Chenlu, Deng Yiping, Comprehensive Statistics and Analysis System of Traffic Accident Information System Implementation Based on Data Mining, China Public Security, 2016,(04): 57-62.

Qin Liyan, Shao Chunfu, Study on Road traffic Safety Management System Based on GIS, China Safety Science Journal, 2004,(02):37-39.

Munyendo J N, Kuria D, Mubea K. Development of AWebGIS Supported Road Traffic Accident Data Management System: A Case Study Of Nairobi County[J]. Journal of Veterinary Internal Medicine, 2015, 1(4):152-7.