Big Data Analytics and Visualization with Spatio-Temporal Correlations for Traffic Accidents

Xiaoliang Fan[1], Baoqin He[1], Cheng Wang[1], Jonathan Li[1](✉), Ming Cheng[1], Huaqiang Huang[1], and Xiao Liu[2,3]

[1] Fujian Key Laboratory of Sensing and Computing for Smart City, School of Information Science and Engineering, Xiamen University, Xiamen 361000, China
xfb_fxl@xm.gov.cn, {315441777, 514187809}@qq.com, {cwang, junli, chm99}@xmu.edu.cn
[2] Shanghai Key Laboratory of Trustworthy Computing, East China Normal University, Shanghai 200062, China
xliu@sei.ecnu.edu.cn
[3] School of Information Technology, Deakin University, Melbourne, VIC 3125, Australia

Abstract. Big data analytics for traffic accidents is a hot topic and has significant values for a smart and safe traffic in the city. Based on the massive traffic accident data from October 2014 to March 2015 in Xiamen, China, we propose a novel accident occurrences analytics method in both spatial and temporal dimensions to predict when and where an accident with a specific crash type will occur consequentially by whom. Firstly, we analyze and visualize accident occurrences in both temporal and spatial view. Second, we illustrate spatio-temporal visualization results through two case studies in multiple road segments, and the impact of weather on crash types. These findings of accident occurrences analysis and visualization would not only help traffic police department implement instant personnel assignments among simultaneous accidents, but also inform individual drivers about accident-prone sections and the time span which requires their most attention.

Keywords: Big data analytics · Accident occurrence analysis · Crash type analysis · Spatio-temporal correlations · Visualization

1 Introduction

Big Data analytics and visualization have been driving nearly every aspect of society, including mobile services, intelligent transportation, manufacturing, financial services, life sciences, and physical sciences [1, 2]. Despite significant development and advancement in vehicle technology and transportation engineering over the last 50 years, traffic accidents are still one of the major accidental causes of deaths and injuries worldwide [3]. For example, there are 700 $\sim$ 800 traffic accidents daily in Xiamen, a middle size city with over 3 million populations in China. Accident data seems not to be a big data scenario due to its “small” volume. However, accident data analyses face

G. Wang et al. (Eds.): ICA3PP 2015, Part II, LNCS 9529, pp. 255–268, 2015.
DOI: 10.1007/978-3-319-27122-4_18

typical big data challenges as well: (1) it is necessary to analyze accidents data very rapidly (i.e., “velocity” feature of big data). The occurrence of an accident would generate both potential traffic congestions and secondary accidents. For example, even a very small accident may cause big traffic congestion in rush hours at crowded segments; and (2) it is quite difficult to mine the value of accident data (i.e., “value” feature of big data), especially in the spatio-temporal view precisely. For example, when and where to occur which type of accidents is nearly unpredictable.

The analysis and prediction of accident occurrences are research areas of considerable interest for a long time [4, 5]. Comprehensive works have been investigated to find the most important risks and variables that might contribute to accident occurrences [10]. However, the limitations of previous works are [16–19]: (1) they considered temporal and spatial view separately; and (2) it is rare to analyze accident occurrences quantitatively in spatio-temporal visualization.

In this paper, we propose a novel accident occurrences analytics and visualization method in spatio-temporal dimension to foresee when and where an accident with a certain crash type will happen consequentially by whom. The abundant sources of 38,674 accident records come from Xiamen City in China from October 2014 to March 2015. First, we analyze the spatial and temporal view of accident occurrences. Second, we illustrate spatio-temporal visualization results through two case studies in multiple road segments, and the impact of weather on crash types.

The remainder of this paper is organized as follows. Section 2 introduces the related works. Section 3 analyzes accident occurrences in spatio-temporal view. Section 4 illustrates the accident big data analytics results and visualization in two case studies. Finally, the general conclusion is drawn in Sect. 5.

2 Related Works

2.1 Accident Occurrences Analysis

The assessment and prediction of the occurrence of accidents as well as how to deal with this risk are research areas of considerable interest for a long time. Comprehensive surveys could be found to investigate the most important risks and variables that might contribute to accident occurrences [4, 5, 17]. Later, several models focus on the impact of temporal and spatial patterns on accident occurrences separately. Lord and Geedipally [6] studied the effects of modeling single- and multivehicle crashes, separately and jointly. Classification and regression tree has been used to perform variable selection [7]. Furthermore, Park and Harghani [14] investigated a primary incident’s impact on secondary incidents. However, previous works mainly considered temporal or spatial view separately. Thus the analysis of spatio-temporal correlations on accidents is still an open issue.

2.2 Crash Types Analysis

In order to reduce accident occurrences and minimize the severity of crashes, various crash type models have been studied to reveal the mechanisms of crash occurrences.

Qin et al. [8] implemented a Bayesian framework to predict crash occurrences in relation to the hourly exposure by crash types. Four crash types were analyzed: (1) single-vehicle, (2) multivehicle same direction, (3) multivehicle opposite directions, and (4) multi-vehicle intersecting directions. Moreover, several other works [9] have also focused on propensity of crash types by developing safety performance functions for highway intersections. Furthermore, four features are widely considered as main factors that cause accidents in different crash types [10]: *geometric factors* (i.e., number of lanes, grade, road segments); *weather* (i.e., rains, snow), *traffic* (i.e., speed, volume), and *driver characteristics* (i.e., age, sex, driving years). However, existing works mainly discuss the features above in the qualitative dimension, while it is rare to analyze the spatio-temporal view quantitatively. In this paper, we focus on the impact of weather on crash types in terms of spatio-temporal visualization.

2.3 Accident Visualization

Wang et al. [11, 15] presented an interactive system for visual analysis of urban traffic congestion based on GPS trajectories. Pack et al. [12] studied incidents visualization from sensors data. Specifically, they designed a linked view interface to visualize the spatial, temporal and multi-dimensional aspects of accidents. Piringer et al. [13] proposed an automatic method to detect and prioritize different types of events by surveillance videos in a tunnel and marked them in space and time. We focus on accident occurrences and crash types analysis in the spatio-temporal dimensions, considering the impact of weather on the spatio-temporal visualization. No such accident data analysis has been studied before in the accident visualization community.

3 Accident Occurrences Analytics in Temporal and Spatial View

3.1 Datasets and Data Pre-processing

When a traffic accident was reported to Xiamen Intelligent Transport Control Center (ITCC), Traffic Accident Management System (TAMS) would initialize an accident record, including several initial information such as accident location (road name, longitude, latitude, and zone ID), number of vehicles, and accident time. In addition, during the accident response period, new information about this accident was continually added to the previous accident record, mostly details such as vehicle type and driver characteristics (i.e., age, sex, driving years). In this paper, we use the following four datasets for accident occurrences and crash-type analysis:

- *Dataset_1* consists six months of **accident data** from October 2014 to March 2015 in Xiamen, totally 38,674 records. Each record has 112 fields, indicating when, where and how the crash occurred.
- *Dataset_2* describes **accident vehicles and drivers** in Xiamen. Each record includes fields such as who were involved in this accidents as well as information on driver characteristics (e.g., age, sex, driving years). The timespan of this dataset is also from October 2014 to March 2015.

- *Dataset_3* is the **weather forecasting data** in Xiamen from October 2014 to March 2015, crawled from the website[1].
- *Dataset_4* is the **road networks** of Xiamen. It contains administrative region, expressway, urban main roads, branch roads, etc.

The pre-processing phase consists of two steps:

- **Road Network Processing.** We first improved the road network quality by filtering out irrelevant roads, merging and splitting ways, and correcting errors. Second, we extracted road segments, crossings and regions from road networks. Specifically, urban arterial roads are divided into 144 segments by traffic lights and road crossings. In addition, we marked 98 road crossings based on the road segments, in which the crossing is a circle with 150 meters. Therefore, we constructed three layers: road segments, road crossings and regions.
- **Accident Data Cleaning and Pre-processing.** Three raw datasets (Dataset_1, Dataset_2 and Dataset_3) are fused that fields with little correlations to accident occurrences are deleted. Therefore, we extracted 15 fields from three datasets. Other preprocessing methods include data integration, conversion and reduction.

Table 1 shows the summary descriptive statistics for part of variables in the above three datasets. Various variables are described in Table 1, such as weather (rain, no rain), rush hour, daytime (day, night), crash position (front, right, left, rear, right front, left front, right rear, left rear), accident liability (full responsibility, without fault), sex, age, driving years, crash type (single-vehicle, side-wipe, rear-end), and location type (segment, crossing, others).

3.2 Temporal View

The accident occurrences are widely considered to represent the temporal periodical pattern. For example, few hours may have more accidents than other time span. Figure 1 shows the temporal view of accident occurrences in the periodical pattern. Specifically, Fig. 1(a) and (b) describe the peridical pattern in weekdays and weekends in separate charts. First, weekdays are shown to have obvious morning and evening peak, while such peaks are not repeatable in the weekend (e.g., less accidents happened during the weekend). In addition, this phenomenon is highly corresponding to the traffic congestion. For instance, when the traffic flow reached the early peak in the morning, accident occurrences were at the early peak accordingly. Second, Fig. 1(c) is the overlay chart of the periodical pattern, and the supplementary comment is that the early peak in the weekend is both smaller and more postponed than those in the weekdays.

In summary, the early and evening peak of accidents is strongly correlated to the peak of traffic flows in urban arterial roads. For example, the early and evening peak of traffic flows is also similar in weekdays, while weakened and postponed in weekends.

[1] Weather forecasting website, http://lishi.tianqi.com/xiamen/index.html.

Table 1. Descriptive statistics for the variables in three datasets

Variables	Type	Coding	Descriptive Statistics (29425)
Weather	Binary	1 = Rainy day 0 = Non-rainy day	21.1 % (6222) 78.9 % (23203)
Rush hour	Binary	1 = Rush-hour traffic 0 = Non-rush-hour traffic	28.2 % (8310) 71.8 % (21115)
Daytime	Binary	1 = Daytime 0 = Night	86.5 % (25438) 13.5 % (3987)
Crash position	Nominal	1 = front 2 = right 3 = left 4 = rear 5 = right front 6 = left front 7 = right rear 8 = left rear	21.0 % (6181) 13.9 % (4095) 14.6 % (4302) 15.8 % (4648) 12.7 % (3727) 10.4 % (3067) 4.8 % (1423) 6.7 % (1982)
Accident liability	Binary	1 = full responsibility 0 = without fault	54.3 % (15988) 45.7 % (13437)
Sex	Binary	1 = male 0 = female	17.1 % (5031) 82.6 % (24316)
Age	Continuous	Driver's age	Mean = 36.18; Std Dev. = 8.981
Driving years Crash type Location type	Continuous Nominal Nominal	0 = Single-vehicle crash 1 = Side-wipe crash 2 = Rear-end crash 0 = Road segment 1 = Road crossing 2 = Others	Mean = 2.46; Std Dev. = 1.817 3.4 % (1009) 69.4 % (20413) 27.2 % (8003) 58.2 % (17133) 31.9 % (9388) 9.9 % (2904)

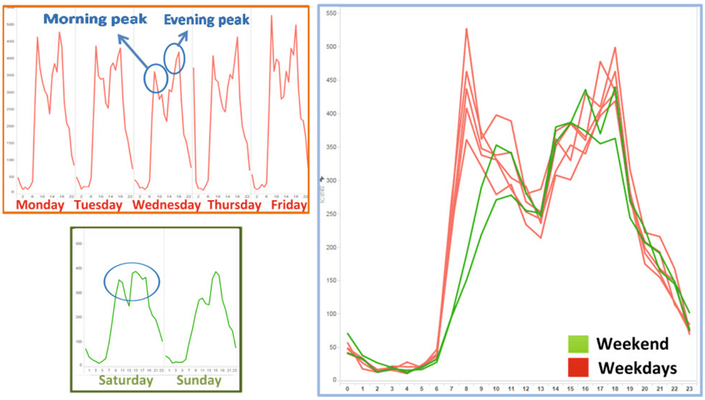

Fig. 1. Temporal view of accident occurrences: the periodical pattern in six months: (a) separate charts by weekdays; (b) separate charts by weekend; and (c) the overlay chart.

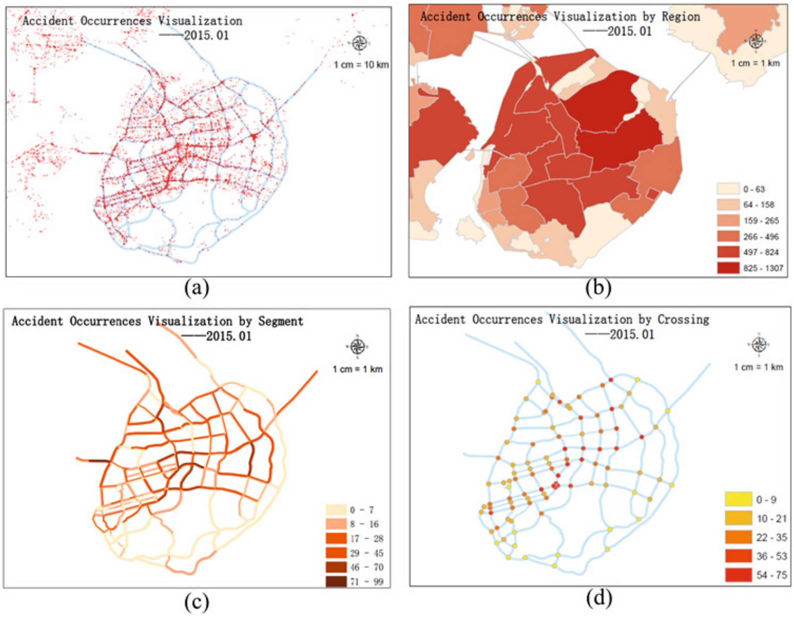

Fig. 2. Spatial view and visualization of accident occurrences for one month with different granularities: (a) coordinates of each accident; (b) by region; (c) by road segment; and (d) by crossing.

3.3 Spatial View

The spatial view and visualization of accident occurrences in January 2015 is presented with different granularities in Fig. 2. Figure 2(a) shows the latitude and longitude coordinates of each accident, and totally 5682 accident occurred in January. Figure 2(b) describes the spatial view of accident occurrences by administration region of the traffic police (the deeper the color, the more the accident in the region). Figure 2(c) shows where accidents occurred on the totally 144 road segments in Xiamen Island, including the four bridges and one tunnel which are connected with the Island. Finally, Fig. 2(d) illustrates the accident occurrences in the totally 99 road crossings in Xiamen Island.

There is an indication that the spatial view of accident occurrences meets the condition of long-tail distribution. For instance, a few road segments and crossings have many accidents while most of others have few accidents spatially. The result will help traffic polices implement instant personnel assignments among simultaneous accidents. However, the spatial view could only visualize where accidents happened, but don’t know when exactly accident happened. We will introduce the spatio-temporal view in Sect. 4.1.

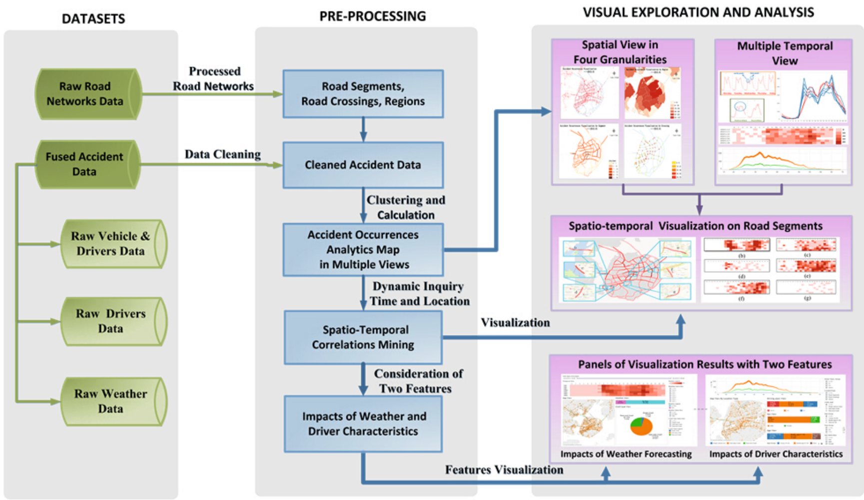

Fig. 3. The design methodology of accident occurrences analytics and visualization with spatio-temporal correlations. We use four datasets (left); in the preprocessing step (center), we extract road segments, crossing and regions from road networks, then clean the data from fused accidents data; and in the visual exploration and analysis step (right), we visualize the results in spatio-temporal view, integrated with the impact of two features on crash types.

3.4 Design of Accident Occurrences Analytics with Spatio-Temporal Correlations

Figure 3 shows the design methodology of accident occurrences analytics and visualization. Our visual analysis work consists of three phases. The first phase is data collection and fusion (left in Fig. 3). Second, in the preprocessing phase (center in Fig. 3), we start from the input data, and extract road segments, crossing and regions from road networks that fit our model. Finally in the third phase of visual exploration and analysis (right in Fig. 3), we visualize the results in spatio-temporal view, integrated with the impact of two features on crash types. The methodology will be implemented in Sect. 4 as two case studies. Plus, we use the Tableau Desktop V8.3[2] software for visualization.

3.5 Hypotheses

Based on the review of the state of the art in Sect. 2, we propose the following two hypotheses, which will be explained and testified in Sect. 4:

- The first hypothesis is that there are various spatio-temporal patterns for accident occurrences. For example, the temporal periodical pattern indicates that the accident occurrences would accumulate on certain time period(s) while decrease during other

[2] Tableau Desktop 8.3, www.tableau.com.

time period(s) in hour of the day, day of the week, and weekday/weekend. While the spatial long-tail pattern means several road segments and crossings have more accidents than most of others, which have very few accidents. The first hypothesis will be testified and further explained in Sect. 4.1.

- The second hypothesis is that there are spatio-temporal correlations of features on crash types Features that might have implicit or explicit impacts on crash types include geometric features, weather, traffic flow, driver characteristics, etc. For example, sideswipe crashes are more likely to occur in multiple-lane road with a heavy traffic flow, while rear-end crashes often happen in steep grade roads and if it has been raining for two hours. The second hypothesis will be testified and further explained in Sect. 4.2.

4 Case Studies and Visualization Results

We have proposed a novel accident occurrences analytics method for accident big data in multiple views. The following two cases will demonstrate the spatio-temporal visualization results in multiple road segments (Sect. 4.1), and the impact of weather on crash types (Sect. 4.2).

4.1 Case I: Accident Occurrences Visualization on Multiple Road Segments

On the road segment level, as shown in Fig. 4(b), we use a table to represent temporal accidents occurrences analytics for six month. The table contains seven rows representing days of the week from Monday to Sunday, and each row contains 24 columns (i.e., 24 h a day). Thus, each cell represent number of accidents occurred at an hour of the day. Specifically, the darker color indicates the more accidents at the hour of the day.

Through road segments classification, Fig. 4 shows the distinguished spatio-temporal patterns on six selected road segments in Xiamen Island. First, as shown in Fig. 4(a), each segment presents its distinguished temporal pattern in accident occurrences visualization. We believe that such differences are mainly due to the following factors correlated to the segment:

- **Commuting hours:** Fig. 4(b) shows a main bridge which links the Xiamen Island to outsides on weekdays, with both a morning peak when people go to work from home, and an evening peak when people go back home after work.
- **Tourist attractions:** Fig. 4(c) shows on segments next to two biggest tourist attractions in Xiamen (i.e., "Gulangyu Island" and "ZhongShan Road", accidents often happen at weekends. In fact, Xiamen is one of the famous tourist destinations nationwide on weekends and holiday.
- **Schools and Shopping malls:** Fig. 4(d) and (f) show two road segments with very different behaviors: as shown in Fig. 4(d), morning accident peaks often come earlier near primary schools, because parents bring children to school before they drive to work on weekdays; while as described in Fig. 4(f), evening accident peaks

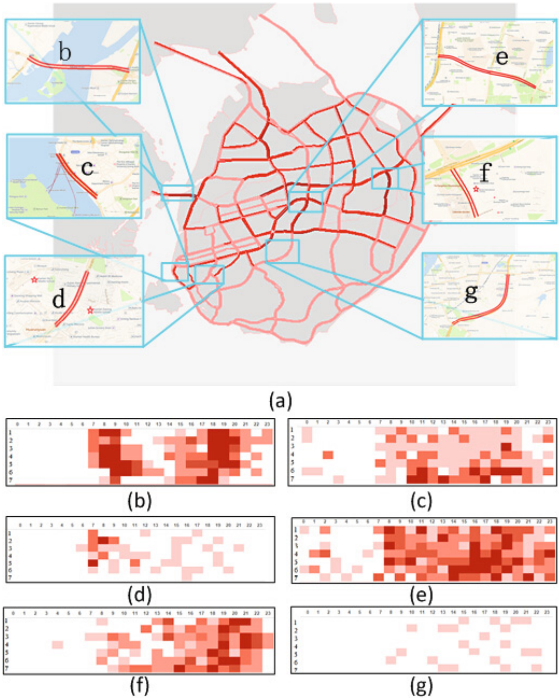

Fig. 4. Distinguished spatio-temporal patterns on six selected road segments in Xiamen Island: (a) each of six road segments has its distinguished temporal patterns in a global view; (b) accidents on Haicang bridge connecting Xiamen island with outside often occur at commuting hours on weekdays; (c) accidents near tourist attractions often happen on weekends; (d) accidents next to primary schools and accident-prone segments also are apt to happen on weekdays with earlier morning peak; (e) accidents could occur at almost any time span in the city downtown and most serious accident-prone segments; (f) accidents near shopping mall are likely to happen during evening peaks; and (g) accidents occurred very occasionally on expressways such as "Chenggong expressway".

were always postponed close to shopping malls, because people are likely to go shopping or have a dinner after work.

- **Accident-prone segments:** Fig. 4(e) shows that on arterial roads in the city downtown, which are also accident-prone segments, accidents could occur at almost any time span. This finding is coherent to the traffic flows at congestion-prone segments.
- **Expressways:** as shown in Fig. 4(g), accidents occurred very occasionally on expressways such as "Chenggong expressway". Because expressways usually have smooth traffic condition.

Furthermore, we could further analyze crash types with spatio-temporal view based on classification results from Fig. 4. Figure 5 shows three pie charts, which indicates different portions of crash types on three selected road segments (see Fig. 4(b), (d), and

(e) respectively). Specifically, the green portion represents rear-end crashes, blue represents single-vehicle crashes and orange represents side-wipe crashes:

- Figure 5(a) shows that on the Haicang Bridge (see Fig. 4(b)), rear-end crashes takes the largest portion while there is few single-vehicle crashes, because this type of road segments are normally straight without forks.
- Figure 5(b) indicates that it is likely to occur single-vehicle crashes on this segment (see Fig. 4(d)). The characteristics of these road segments are both a narrow road and there are not isolated strips between vehicles lanes and pedestrian/bicycle lanes.
- Figure 5(c) shows that in the busy arterial road segments, it is opt to occur side-wipe crashes (see Fig. 4(e)). The complicated road networks, multiple crossings and zebra crossing collectively contribute to the difficulties in driving through this type of segments. Especially, it is difficult for the driver to avoid side-wipe crashes at a turning. Furthermore, heavy traffic flows at early and evening peaks make the risk of side-wipe crashes even worse, because of frequent lane changes and cut-into-line in this circumstance.

Figure 5 collectively suggests that it is likely to occur side-wipe crashes at crossings, while rear-end crashes often happen at road segments. In summary, through the first case on spatio-temporal view of accident occurrences analysis, both traffic polices and drivers could explore different accident occurrence patterns with our visualization results.

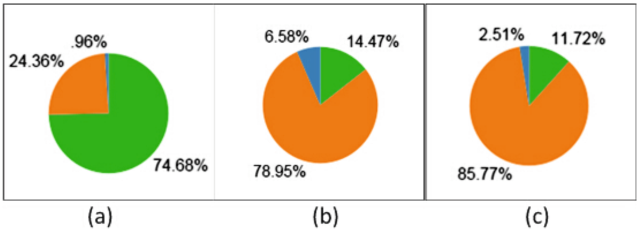

Fig. 5. The percentage of different crash types in three segments (see Fig. 4(b), (d), and (e) respectively) (Color figure online)

4.2 Case II: Crash Types Visualization with the Feature of Weather

Figure 6 presents visualization results of crash-type analysis with the feature of weather. The panel contains four views: map view (Fig. 6(a)), temporal view (Fig. 6(b)), weather view (Fig. 6(c)) and crash-type view (Fig. 6(d) and (e)). In addition, Fig. 6(f) shows various filters, which provide dynamic visual statistics and fast combination queries on demand. Data filters include spatial and temporal queries, and the selection of weather and crash-type. **From the spatial aspect,** three selection shapes could be used: circle, rectangular and lasso (Fig. 6(g)). We could also select road segments, crossings and regions through shape filters. **From the temporal aspect,** a three-level temporal filter could be used from date, week and time in Fig. 6(b). Therefore, we could build complex scenarios by the combination of atomic queries. For example, we could analyze and

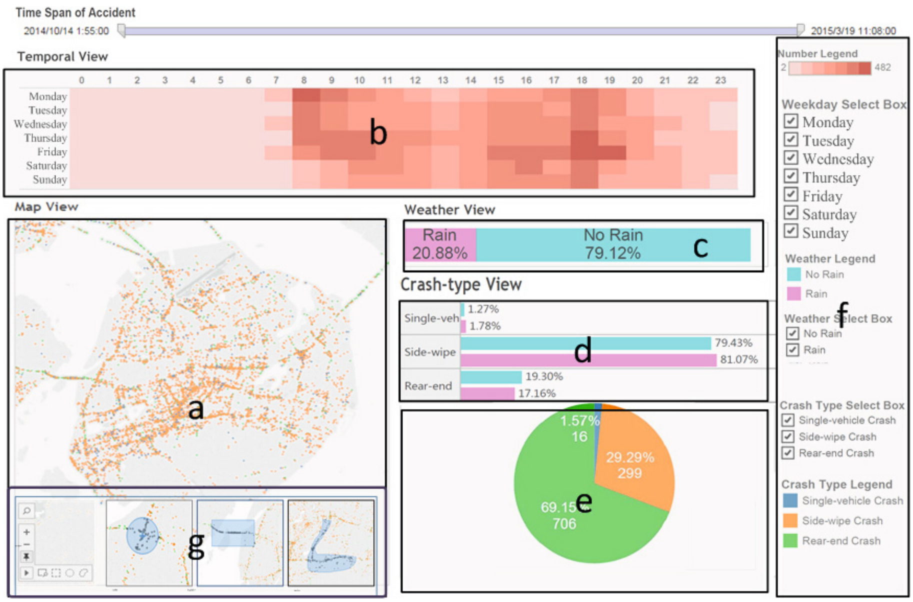

Fig. 6. Crash Types Analysis and Spatio-temporal Visualization on Weather: (a) map view of accident occurrences; (b) temporal view of accident occurrence; (c) percentage distribution on weathers; (d) percentage distribution on crash-type by weathers; (e) percentage distribution on crash-type; (f) (g) selection shapes on data filter.

visualize the results of when and where are likely to occur which type of crash in the rainy day.

To further illustrate the impact of weather on crash types, we select three segment types: four bridges (left picture in Fig. 7(a)), an arterial road (center picture in Fig. 7(a)) and several branch roads (right picture in Fig. 7(a)). Specifically, there is no intersection or ramp on the bridges, while there are zebra crossings and crossings on arterial roads. It seems that there is no connection between weather and accident occurrences, but it's meaningful when we consider temporal feature and cash-types.

The visualization results have the following indications:

- Most accidents are rear-end crashes on the bridges (65.30 %, 81.67 %). On the contrary, most accidents are side-wipe crashes on arterial roads (79.43 %, 81.07 %) and branch roads (78.26 %, 88.81 %). The main reasons for these results are: (1) the traffic condition is normally fluent with high speed on the bridge, which might result in rear-end crashes due to speeding behaviors; (2) there is always congestion in urban arterial roads, which increases the risk for side-wipe crashes if there are frequent lane changes at the crossings and turns; and (3) it is unavoidable that narrow branch roads could induce side-wipe crashes.
- Weather is likely to have more impact on accidents for bridges, next for branch roads, but doesn't have obvious influence on arterial roads. Because percentages of side-wipe crashes and rear-end crashes change more than 15 % when it rains on

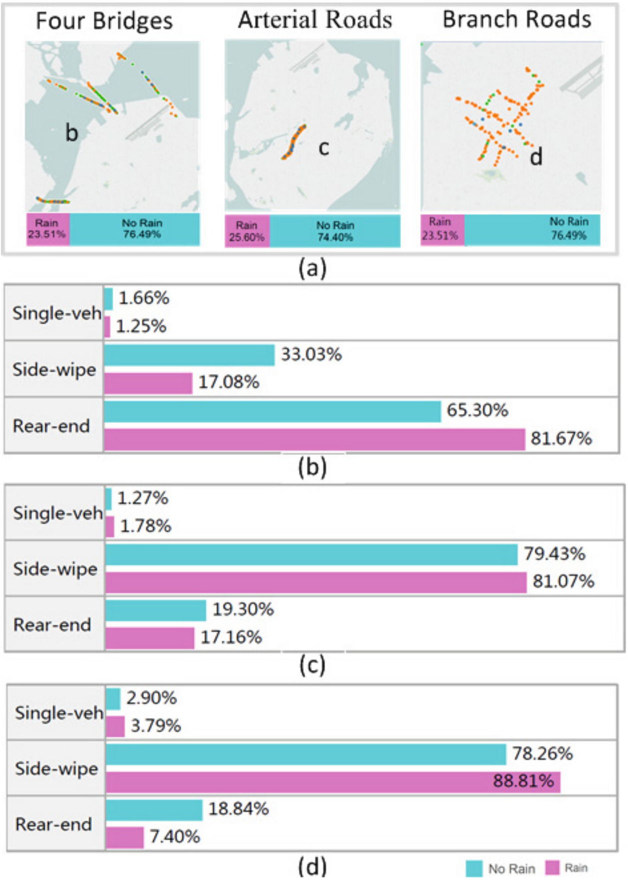

Fig. 7. The impact of weather on crash types, which has different influence on crash types: (a) three selected typical types of road segments; (b) the influence of weather on bridges; (c) the influence of weather on arterial roads; (d) the influence of weather on branch roads.

bridges, about 10 % on branch roads, and less than 3 % on the arterial roads. When we focus on the rear-end crashes, it seems that a heavy and long-lasting rain would decrease the rear-end crashes on the bridges (65.30 %, 81.67 %), but increase the possibility of side-wipe crashes on the branch roads (18.84 %, 7.40 %). On the rainy days, there are slippery ground and speed cars on bridges.

5 Conclusion

Given the massive traffic accident records in urban arterial roads, accident occurrence analysis is known to identify the main factors that contribute to crash type, crash position and severity. However, due to heterogeneous case-by-case nature of traffic accidents, the difficulty to analyze such traffic big data quantitatively has led to minimal effects in previous works. Especially, previous works have a major shortage in the lack of considering spatio-temporal correlations in accident occurrences analytics.

By analyzing the traffic accident data from October 2014 and March 2015 in Xiamen, China, we propose a novel accident occurrence analysis and visualization method in both spatial and temporal dimensions, in order to predict when and where an accident with a certain crash type will happen sequentially by whom. Despite the significant progress of spatio-temporal visualization in accident occurrences, there still remain numerous avenues to explore. Our future works include the consideration of geometrics features (number of lanes, grade, etc.) in spatio-temporal correlations mining in crash types analysis. For example, sideswipe crashes are likely to happen on multiple lanes, while rear-end crashes are often seen in road segments with steep grade.

Acknowledgments. This work is supported by the grants from Natural Science Foundation of China (No. 61300232, No. 61300042); Ministry of Education of China “Chunhui Plan” Cooperation and Research Project (No. Z2012114, Z2014141); Funds of State Key Laboratory for Novel Software Technology, Nanjing University (KFKT2014B09); Fundamental Research Funds for the Central Universities (lzujbky-2015-100); and China Telecom Corp. Gansu Branch Cuiying Funds (lzudxcy-2014-6). The authors acknowledge Xiamen Intelligent Transport Control Center (ITCC) for providing the data.

References

1. Jagadish, H.V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J., Ramakrishnan, R.: Big data and its technical challenges. Commun. ACM **57**(7), 86–94 (2014)
2. Wu, X., Zhu, X., Wu, G., Ding, W.: Data mining with big data. IEEE Trans. Knowl. Data Eng. **26**(1), 97–107 (2014)
3. WHO: World Report on Road Traffic Injury Prevention. World Health Organization, Geneva (2004)
4. Hauer, E.: Speed and safety. Transp. Res. Rec. **2103**(2103), 10–17 (2009)
5. Ma, J., Kockelman, K.M., Damien, P.: A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accid. Anal. Prev. **40**(3), 964–975 (2008)
6. Lord, D., Geedipally, S.: Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson–gamma models. Accid. Anal. Prev. **42**(4), 1273–1282 (2010)
7. Yu, R., Abdel-Aty, M.: Utilizing support vector machine in real-time crash risk evaluation. Accid. Anal. Prev. **51**(2), 252–259 (2013)
8. Qin, X., Ivan, J., Ravishanker, N., Liu, J., Tepas, D.: Bayesian estimation of hourly exposure functions by crash type and time of day. Accid. Anal. Prev. **38**(6), 1071–1080 (2006)
9. Jonsson, T., Lyon, C., Ivan, J., Washington, S., van Schalkwyk, I., Lord, D.: Investigating differences in the performance of safety performance functions estimated for total crash count and for crash count by crash type. Transp. Res. Rec. J Transp. Res. Board **137**(2102), 115–123 (2009)

10. Yu, R., Abdel-Aty, M.A., Ahmed, M.M., Wang, X.: Utilizing microscopic traffic and weather data to analyze real-time crash patterns in the context of active traffic management. IEEE Trans. Intell. Transp. Syst. **15**(1), 205–213 (2014)
11. Wang, Z., Lu, M., Yuan, X., Zhang, J., de Wetering, H.: Visual traffic jam analysis based on trajectory data. IEEE Trans. Vis. Comput. Graph **19**(12), 2159–2168 (2013)
12. Pack, M., Wongsuphasawat, K., VanDaniker, M., Filippova, D.: Ice–visual analytics for transportation incident datasets. In: IEEE International Conference on Information Reuse Integration, pp. 200–205. IEEE Press, New York (2009)
13. Piringer, H., Buchetics, M., Benedik, R.: Alvis: situation awareness in the surveillance of road tunnels. In: IEEE Conference VAST, pp. 153–162. IEEE Press, New York (2012)
14. Park, H., Haghani, A: Real-time prediction of secondary incident occurrences using vehicle probe data. Transp. Res. Part C Emerg. Technol. (2015). doi:10.1016/j.trc.2015.03.018
15. Wang, Z., Yuan, X.: Urban trajectory timeline visualization. In: IEEE International Conference on BigComp. IEEE Press, New York (2014)
16. Zhang, G., Yau, K.W., Zhang, X.: Analyzing fault and severity in pedestrian–motor vehicle accidents in China. Accid. Anal. Prev. **73**(73), 141–150 (2014)
17. Zhang, G., Yau, K.W., Zhang, X.: Risk factors associated with traffic violations and accident severity in China. Accid. Anal. Prev. **59**(10), 18–25 (2013)
18. Tak, S., Kim, S., Yeo, H.: Development of a Deceleration-Based Surrogate Safety Measure for Rear-End Collision Risk. IEEE Trans. Intell. Transp. Syst. **16**(5), 2435–2445 (2015)
19. Lin, L., Wang, Q., Sadek, A.W.: A novel variable selection method based on frequent pattern tree for real-time traffic accident risk prediction. Transp. Res. Part C Emerg. Technol. **55**, 444–459 (2015)