IEREK PRESS Academic Research Community

http://www.press.ierek.com

ARChive

CrossMark

ISSN (Print: 2537-0154, online: 2537-0162)

International Journal on:

The Academic Research Community Publication

DOI: 10.21625/archive.v4i2.752

Visualization System for Traffic Accident Data

Su Chenlu[1], Hong Yating[1], Liu Tianyi[1], Raja Majid Mehmood[1*]

[1] *School of Electrical and Computer Engineering, Information & Communication Department, Xiamen University Malaysia*

**Corresponding author: Raja Majid Mehmood (rmeex07@gmail.com)*

Abstract

At present, the traffic problem is a problem that the government attaches great importance to. Many papers also put forward their own visualization models for traffic problems. This research focused on the Map-matching and Spatial-temporal Visualization of Expressway Traffic Accident Information and improves the original two-dimensional visual model of accident rate into a three-dimensional model. The goal is to represent more attributes in a visual model and make them easier to compare, so as to provide users with more intuitive visual information.

Keywords

Information visuliztion, Visualization system, Three-dimensional model,Traffic accident

1. Introducation

Since the beginning of the 21st century, there is a large amount of traffic accidents over the world because of increasing number of vehicles and complex urban construction.

Previous data analysis models were more planar and in the form of digital tables, which made it difficult to analyze all the data information.

Considering the seriousness of this problem, it is necessary to design a traffic accident data visualization system, which allows the traffic departments to visually analyze accident data so as to better improve the traffic and road planning.

Based on that, this study will based on the 2D modeling of traffic accident data from *Research on the Map-matching and Spatial-temporal Visualization of Expressway Traffic Accident Information*, and improved it into a three-dimensional model that can represent more attributes in a visual model and make them easier to compare, so as to provide users with more intuitive visual information.

In this regard, the transportation department can separately analyze the occurrence of accidents at different time points according to the results of the visualization, and can clearly see the accident trends of working days and non-working days, so as to formulate more feasible transportation law. In addition, people can also plan the travel arrangements by looking at the visual results to avoid peak accidents and avoid high-risk roads.

2. Literature review

In *Research on the Map-matching and Spatial-temporal Visualization of Expressway Traffic Accident Information*, This paper proposes a method based on accident collection data and GIS roadmap data to realize rapid location, map

matching and verification of highway accidents. Through visual analysis, traffic management departments can be helped to improve accident prevention capabilities.

Specific implementation: First, collect data using map collection tools (not only can collect street view information of roads, but also collect road data). Next, it can achieve automatic road data extraction of panoramic images.

The visualization of accident time-space data helps to enhance the understanding of time and space factors and accident data changes, including three aspects:

a. Time series visulization

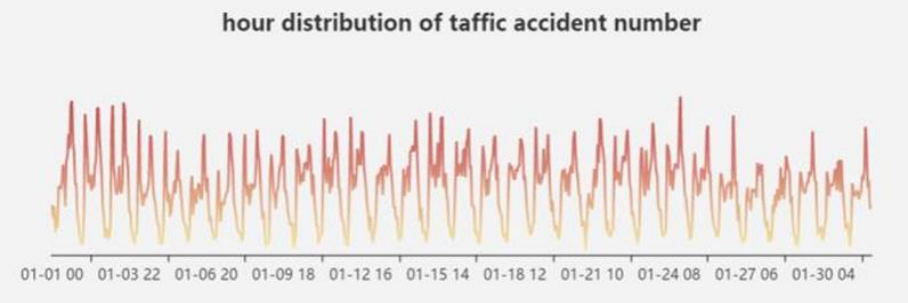

Figure 1 Time series visualization using the line graph

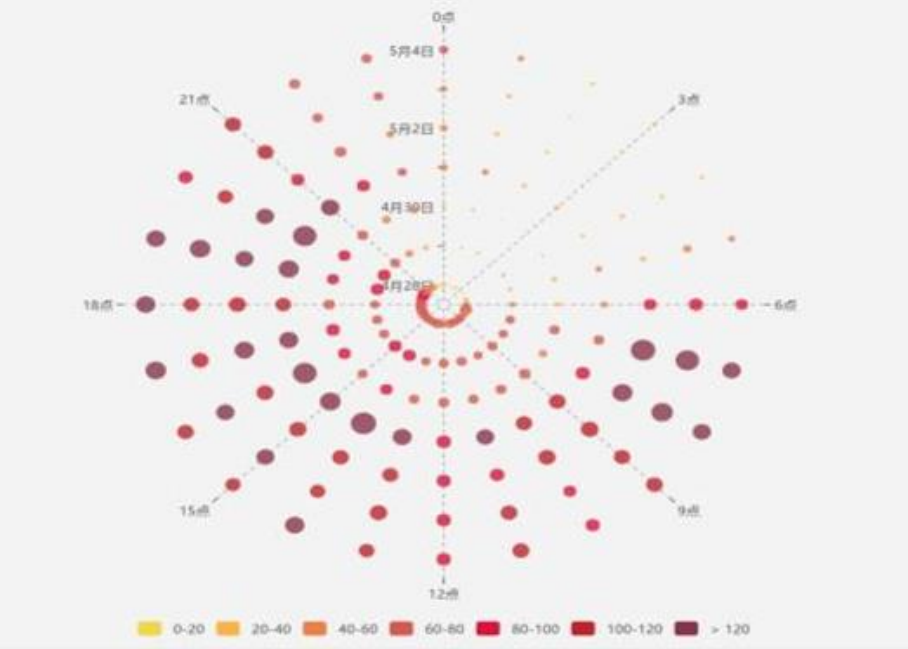

Figure 2 Time series visualization using the polar graph

b. Spatial distribution visualization

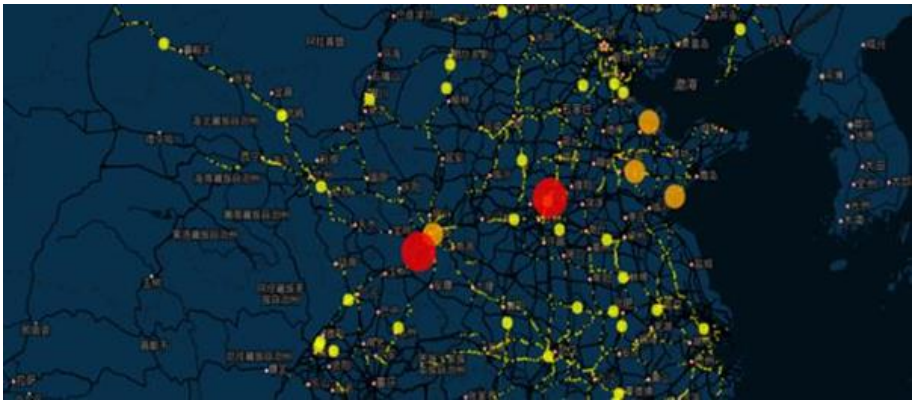

Figure 3 Spatial distribution visualization using map and scatter plot (simulating data)

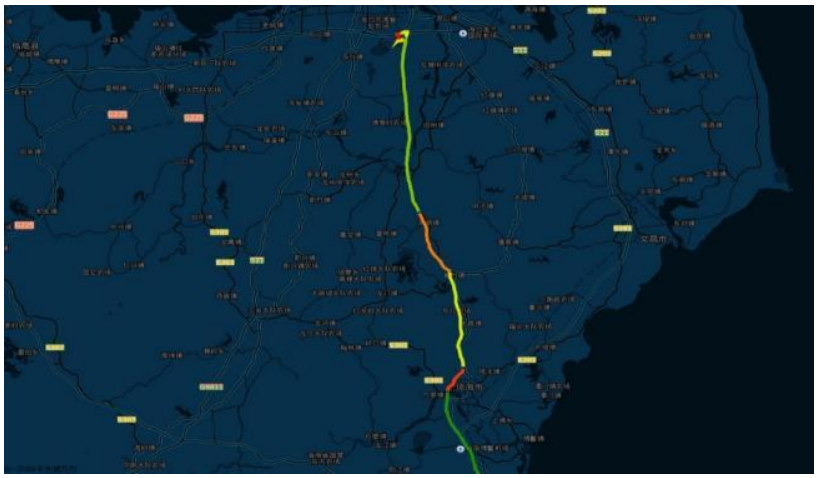

Figure 4 Accidents frequently-occurring road segment visualization

Figure 5 Spatial distribution visualization using reginal hierarchical rendering and scatter plot.

c. Space-time related visualization, considering the time and space of traffic accidents Distribution characteristics, specifically expressing the regularity of accidents in a certain time and space.

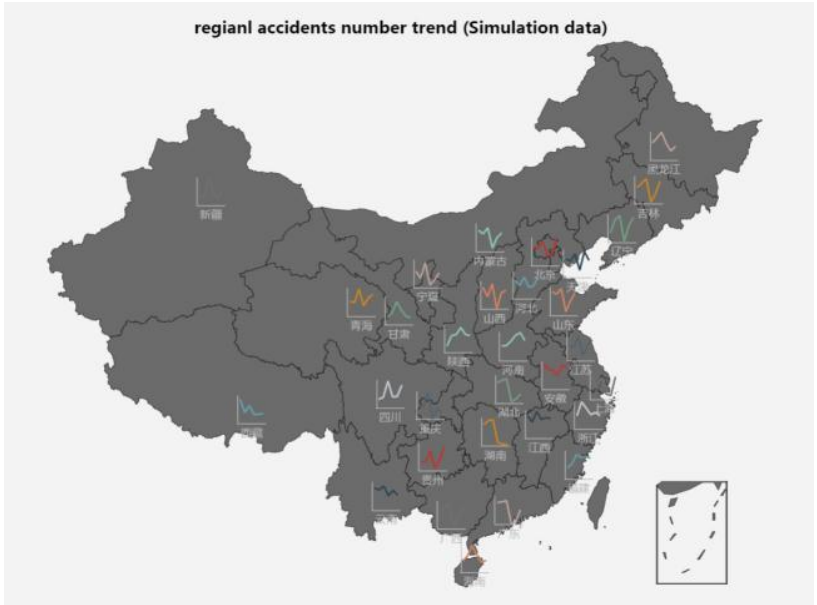

Figure 6 superimposing the time-line chart for different regions on the map. (simulating data)

Among those visualization model, time series visualization using line graph and Time series visualization using polar graph are compared. (M1 means Method 1: Time series visualization using line graph; M2 means Method2: Time series visualization using polar graph.)

Table 1. Type of visualization

	M1	M2
Type	Line chart	Polar chart
Musk	Points, lines and color	Circle, clock and coloe
Target Audience	Transportaion department	

Table 2. Displayed information

	M1	M2
Number of acccidents in each hour	Displayed	Displayed
Number of accidents every day	Displayed	Displayed

Table 3. Scalability

	M1	M2
Scalability	High	Medium
Reason	Because the number of accidents can refer to different value in y-axis, if the value is high, still can be represented.	Because the size and color of circle represent the quality of accident, and it is not really precise，but it still can represent large number.

Table 4. Accuracy

	M1	M2
Accuracy	Medium	Low
Reason	Although its visualization is based on exact value to represent the number of accidents, but it is not easy to get the precise quantity of accidents in each timestamp (any point) on x axis.	It is hard to determine the accurate value from size and color shade of circle., because each color refer to a range of value.

Table 5. Reliability

	M1	M2
Reliability	High	High
Reason	The source of the data for the three methods is derived from the accident record and use a double-check model to verify whether the data is reliable, so the data is highly authentic. That means the visualization has high reliability.	

Table 6. A degree of discrimination (when the data is ver close)

	M1	M2
Degree of discrimination	High	Low
Reason	It is differentiable from exact value in y axis, so it is really clear to see discrimination.	When the data is really close, it is not easy to distinguish them between color and size of circle.

3. Proposed Method

Based on the original data visualization model, we improved that visualization system for the accident data in the road.

			0:00-2:00	2:00-4:00	4:00-6:00	6:00-8:00	8:00-10:00	10:00-12:00	12:00-14:00	14:00-16:00	16:00-18:00	18:00-20:00	20:00-22:00	22:00-00:00
week1	Monday	1-Apr	1	0	0	7	4	5	7	5	7	6	7	3
	Tuesday	2-Apr	0	0	0	6	3	6	5	4	7	9	6	2
	Wednesday	3-Apr	2	0	0	8	4	6	5	5	8	10	7	3
	Thursday	4-Apr	1	1	1	9	5	5	7	6	7	12	9	2
	Friday	5-Apr	2	0	0	12	6	5	8	7	6	13	10	2
	Saturday	6-Apr	2	1	2	5	7	6	9	7	9	12	12	4
	Sunday	7-Apr	3	2	0	6	4	4	12	6	9	9	8	0
week2	Monday	8-Apr	2	1	2	8	3	3	7	6	11	12	7	3
	Tuesday	9-Apr	1	0	2	10	3	6	7	8	13	9	7	2
	Wednesday	10-Apr	0	0	0	9	5	3	8	6	9	8	6	4
	Thursday	11-Apr	0	0	1	7	8	3	6	5	7	10	12	5
	Friday	12-Apr	1	0	2	7	8	5	3	4	7	12	8	3
	Saturday	13-Apr	3	0	2	6	6	5	11	7	6	11	7	3
	Sunday	14-Apr	2	0	4	7	6	6	10	8	9	10	7	2
week3	Monday	15-Apr	2	0	3	10	8	5	9	11	7	8	6	4
	Tuesday	16-Apr	1	1	1	8	8	7	6	3	7	8	8	4
	Wednesday	17-Apr	0	0	3	7	5	3	5	3	6	8	9	2
	Thursday	18-Apr	1	0	4	7	5	3	3	6	8	8	8	1
	Friday	19-Apr	1	1	4	7	7	5	7	7	10	8	7	0
	Saturday	20-Apr	1	0	3	4	6	4	8	8	12	6	6	1
	Sunday	21-Apr	2	0	6	6	4	6	13	3	12	7	5	2
week4	Monday	22-Apr	1	0	5	6	4	3	9	8	13	4	2	3
	Tuesday	23-Apr	0	1	4	8	3	6	7	8	15	7	12	4
	Wednesday	24-Apr	1	0	1	9	2	4	3	9	14	3	7	4
	Thursday	25-Apr	1	0	2	10	3	3	5	8	13	3	2	3
	Friday	26-Apr	2	0	2	9	3	5	6	7	11	4	4	2
	Saturday	27-Apr	0	1	3	6	3	3	6	7	9	6	6	2
	Sunday	28-Apr	1	0	3	11	5	2	7	5	8	12	7	1

Figure 7 Data set for section2 (the value is number of accidents)

3.1. Version 1

A calender map is proposed to visualize the traffic accident data. In this map, days are arranged in a calendar form. All the days will be showed as a grid surrounded by red edges with 12 smaller grids inside (4 lines and 3 rows). Each smaller grid represents two hours in a day. They are arranged in chronological order from left to right, from top to bottom. Colors change from light green to dark green will show the probability of accidents from low to high.

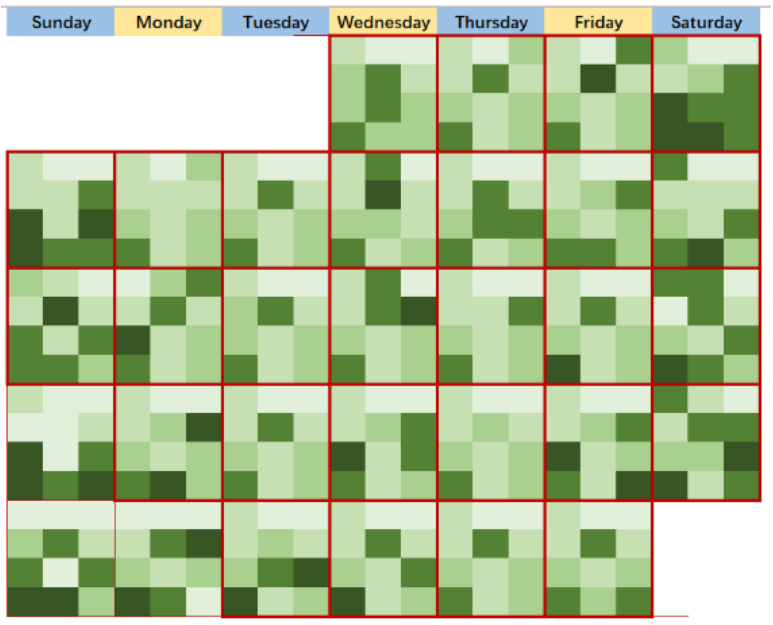

Figure 8 Calendar map-1

Figure 9 Calendar map-2

3.2. Version 2

Since the version 1 cannot give direct comparison between different attributes in more perceptible sequence, in this version, the proposed model is designed to deliver the information of different hours in a day and different days' information in a month. It is a three-dimensional visualization model that x axis represents hours in a day, y axis represents days in a week (from Monday to Sunday) and z axis represents the same weekdays in a month (four Mondays in a month will be assigned to the same z axis).

Table 7. The explanation of information in different axis

Axis	Meaning
X-axis	Hours in a day (from 0:00-24:00, each bar is 2 hours)
Y-axis	Days in a week (from Monday to Sunday)
Z-axis	The number of accidents in four same weekdays in one month.

The reason why we used days in a week (Monday to Sunday) instead of specific date (May 1st or May 2nd) is because weekday and weekend are more comparative, have more important directive to target audience (Transportation department).For example, it is more meaningful for user to compare different accident rates between weekdays and weekend at the same time period (such as the evening rush from 18:00 to 20:00).

In this model, readers can get different contrasts by comparing different dimension. In z axis, there are four weeks (first week, second week, third week, fourth week) in a month. Each of them is a column which filled with different colors: green, red, yellow and blue respectively.

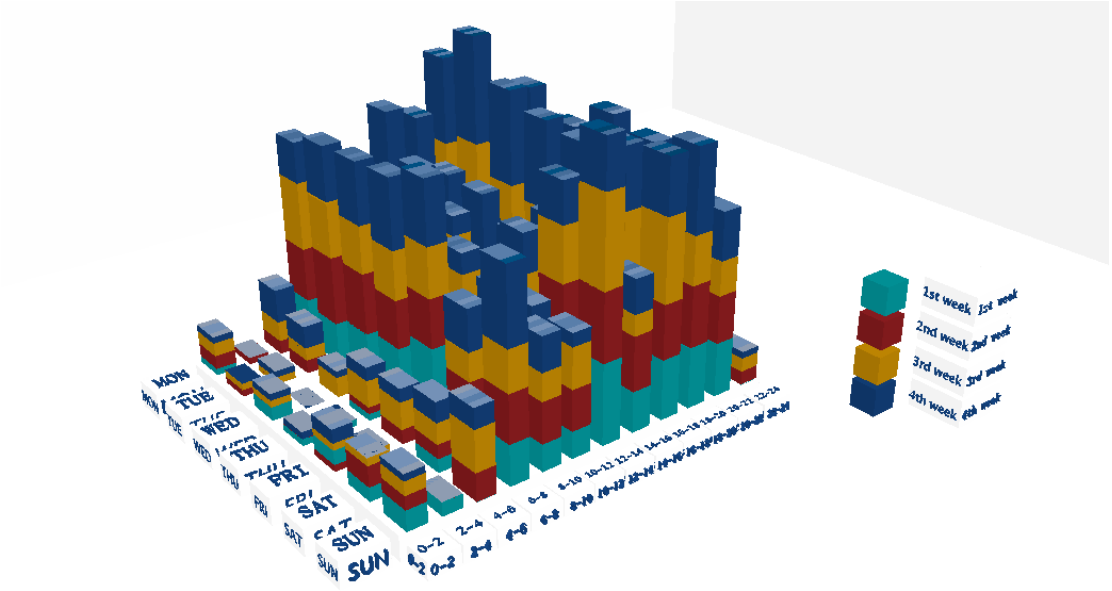

Figure 10 Three-dimensional method

When the x axis and z axis is displayed, it can show accident data in every 2 hours on specific day in one week. (Figure 7 four Mondays etc.). Two hours is a group.

Figure 11 The explanation of block with different colors

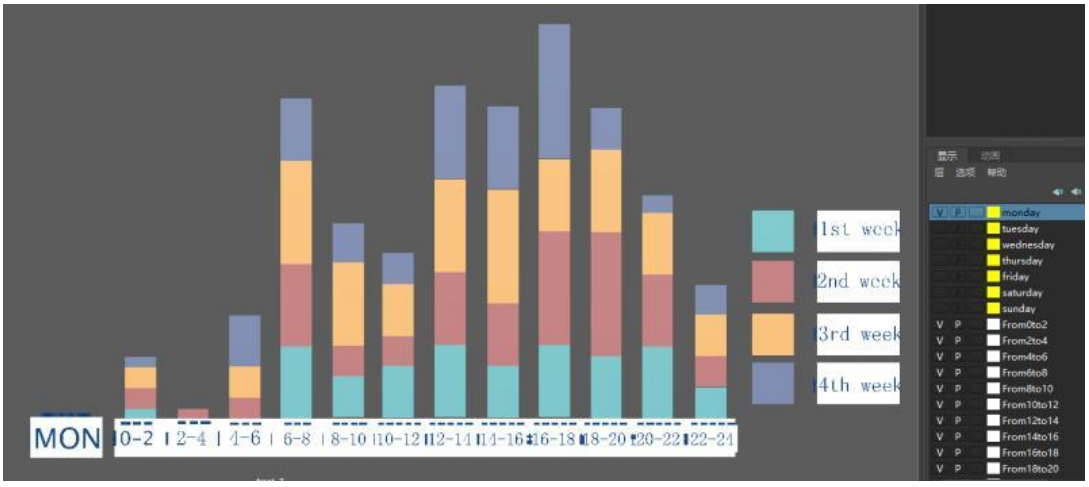

Figure 12 Monday (four weeks)

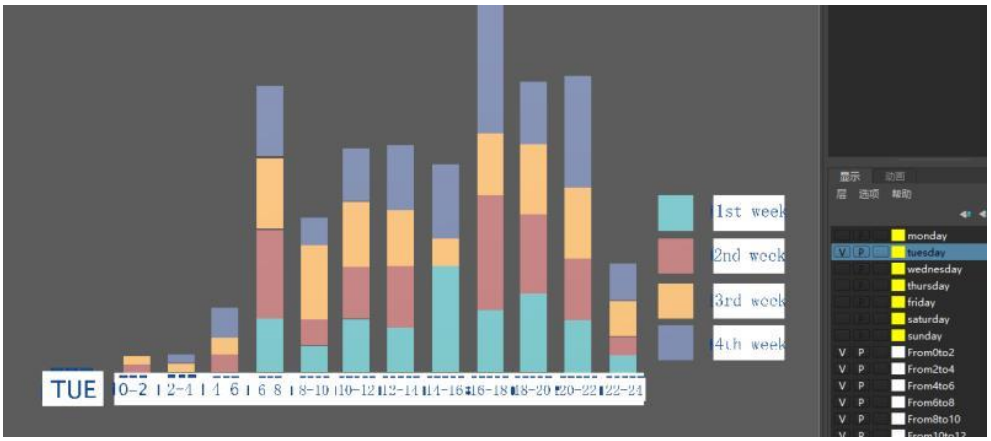

Figure 13 Tuesday (four weeks)

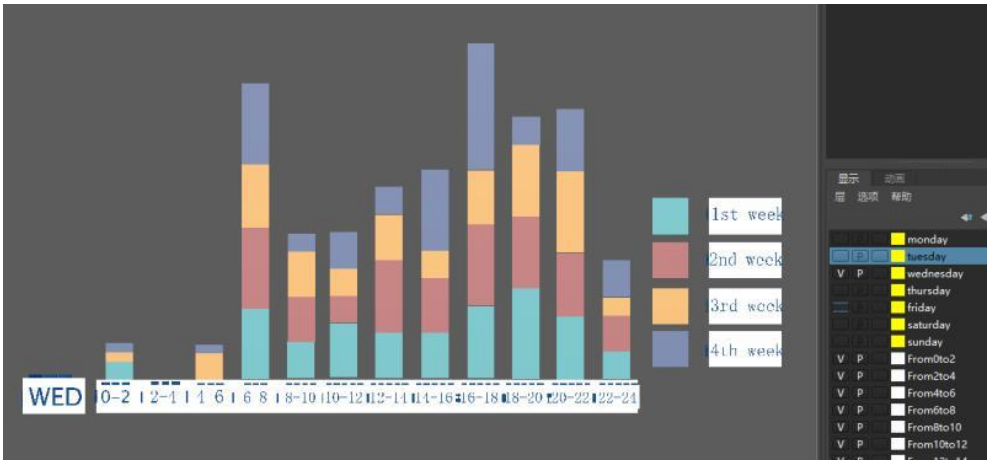

Figure 14 Wednesday (four weeks)

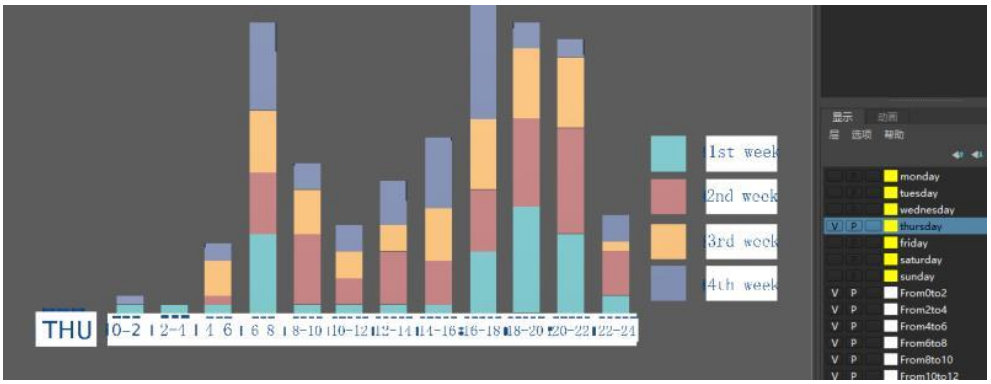

Figure 15 Thursday (four weeks)

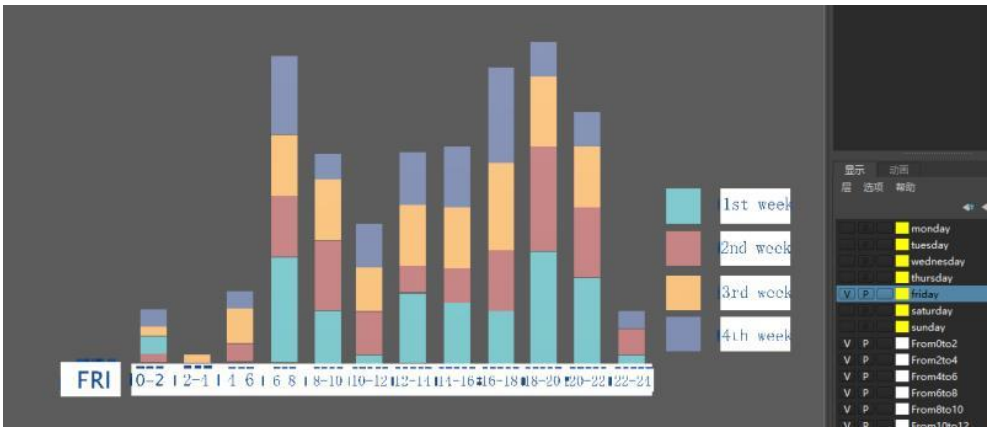

Figure 16 Friday (four weeks)

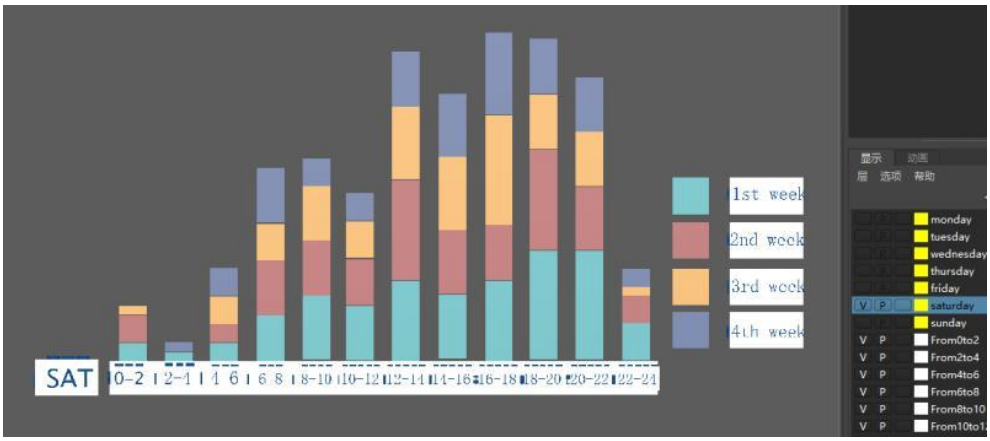

Figure 17 Saturday (four weeks)

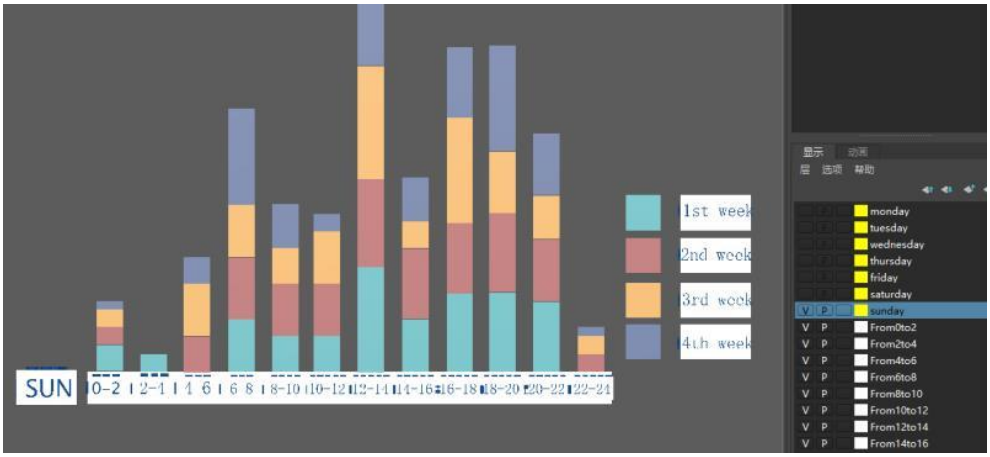

Figure 18 Sunday (four weeks)

When y-axis and z-axis is displayed, it can show accident data at specific 2 hours t (as Figure 41 shown is 0:00-2:00) on every weekday in one month.

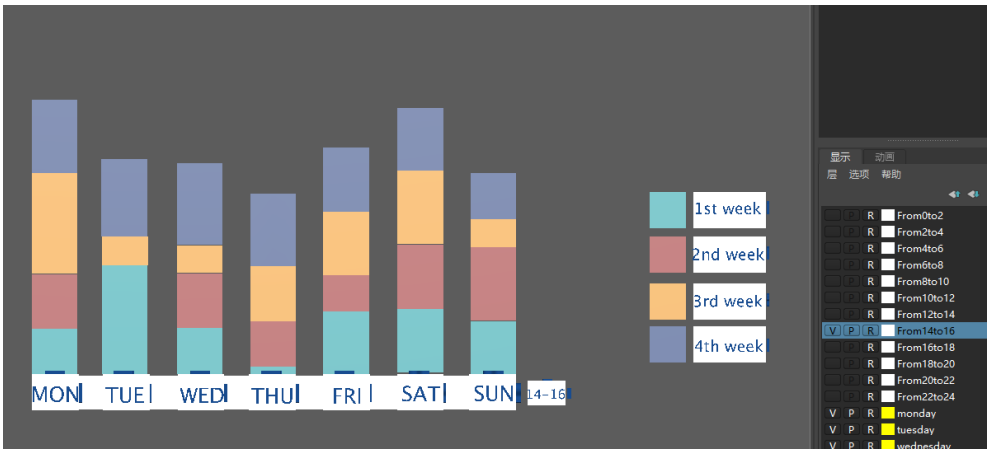

Figure 19 y and x axis

Other than this, the model can also be displayed in x axis and y axis. Blue color from light to dark will give information according different number of accidents. There are four smaller grids in one rectangle. Each of a smaller grid represents week in sequence (First row of girds represents the information of first week, and so on). All the information can be showed in the x and y axis.

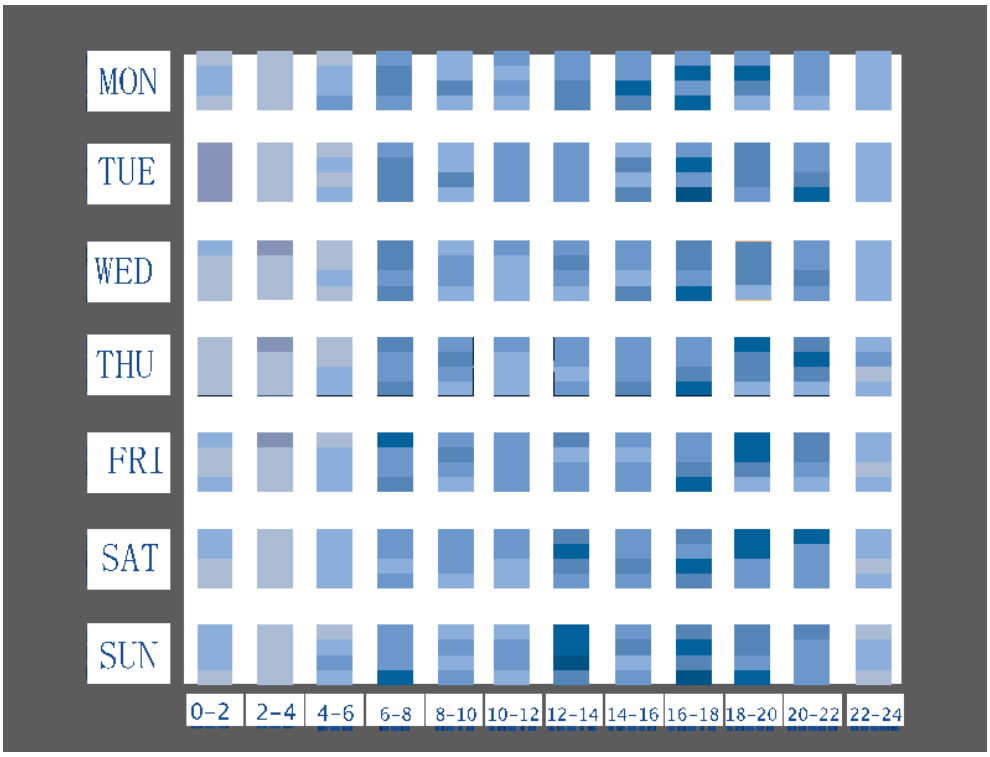

Figure 20 x and y axis

3.3. Version 3

Although the proposed version 2 has been greatly improved, for example, it can represent more data and make clearer comparisons. However, the color classification is so few that the visualization of each set of data is not clear enough. Therefore, in the proposed version 3, we use more color systems to represent the incidence of traffic accidents.

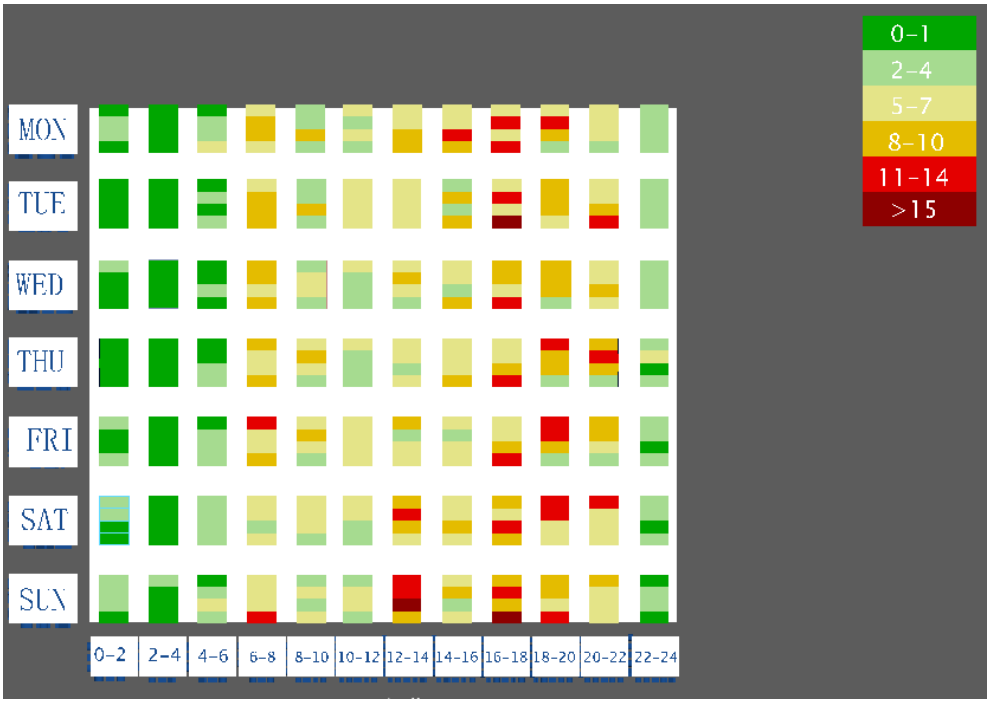

Figure 21 Proposed method version 3- whole month [x-y axis]

4. Evaluation and Result

Here are the comparisons among the original method, proposed version 1, version 2 and version 3.

Table 8. Comparation among methods and proposed versions 1, 2 and 3

	Method	**Proposed version 1**	**Proposed version 2**	**Proposed version 3**
Type	Polar graph	Calendar garph	3D bar chart	3D bar chart
Displayed data size	7-day	1-month	1-month	1-month
Dimensional	2D	2D	3D	3D
Mask	Circle, color shade	Grid, color shade	Bar, grid, color shade	Bar, grid, color, color shade
Accuracy	Low	Medium	High	High
Scalability	Low	High	High	High
Perception	High	Medium	Medium	High

The first polar graph only displays 7-day data whereas all method we proposed can show one-month data. And the previous two methods are 2-dimensional and later two are 3-dimensional. Polar graph consists of circle and color shade to represent the number of accidents, and the calendar graph use different color shade of grid to show accident condition in calendar form. However, the last two method display the number of accidents by height of bar and the top different color shade of grid can represent every day data. Besides, the last method includes 3 colors (red green blue) to indicate different severity.

The polar graph and calendar graph use color shade to represent number of accidents is not really precise, but the later methods we proposed are use height corresponding to different values, which is really precise.

Apart from that, if the data set is huge, the polar graph cannot accommodate all data, but later 3 methods can be applied to large sample size. As for perception, the clock form of the polar graph makes it intuitive to see the time changes within a day, but the changes in different days are not very obvious. Although the calendar map can reflect the accident situation within one month and one day, it is difficult to compare each other.

3-D bar charts can be compared from different dimensions (different days of one-week, different times of one day, data within one month), and then different results are obtained. The amount of information is large, but the color shade is single, resulting in low information acquisition. In the last 3-D bar chart, we used different colors in the y-z axis view to distinguish the severity, and the viewer can obtain information more intuitively.

The significance of this part can be divided into Four parts:

Firstly, in the way of data representation, the regular grid provides a neat data visualization model from different angles (x-y, x-z and y-z), and through the form of a 3d bar chart, different dimensions are integrated into one, which can clearly display and compare data for each dimension by rows or columns.

Secondly, in the data dimension, this model provides more different dimensions of time representations: hour, day and week (the x-axis is the hour, the y-axis is the week, and the z-axis is the day), so it is easy for user to represent and compare the data of the accident occurrence from more different aspects, such as the amount of accidents on different days of the same time period (such as 0:00 am to 2:00 am) and the number of accidents at different time periods in same day. In the existing data visualization model, the time dimension is only 7 days, so it is impossible to provide so many contrast dimensions and information volume.

Thirdly, in the term of the noticeable of the data, through improvement, a strongly contrasting color was used to indicate the number of accidents on different days of a month. Red and green as a pair of contrast colors which were used to represent data in the top view allow the user to intuitively feel which time period or day of the week has a high risk of accidents. Also, the colors with different hue are used to indicate different weeks of a month, which there is a strong discrimination between each week.

Finally, this model also has good scalability. When the amount of data is more, it is easier to find the similarity of data in the same time period, and it is easy to get road congestion information and dangerous information. This model has no spatial limitations and the expression dimension limit of other existing models (like the polar model is limited by the spatial position size and circle size).

Among the users surveyed, all felt that the improved model could provide more information. At the same time, it is generally believed that the visual model can be clearly understood, and the legends provide enough information. However, suggestions have been put forward for the representation of similar data. If two numbers are close, it is difficult for them to distinguish, that is to say, there are not enough categories and grades of colors. Meanwhile, it is suspected that if the number of data increases, it may be difficult to express it in this model.

5. Conclusion

In this study, we first selected the paper Research on the Map-matching and Spatial-temporal Visualization of Expressway Traffic Accident Information and improve the visualization model. Regarding the traffic accident rate of a certain area every two hours, every day and every week, the two-dimensional time-space visualization model of the original paper was first changed to calendar model to display more data. And then the calendar model was changed to three-dimensional model so that each attribute can be clearly compared. Finally, the 3d model was further improved, more colors were used to represent the accident rate and the legend was also improved. In a word, compared with the models in the paper, the final three-dimensional model has the advantages of displaying more data, easier comparison, clearer representation information and easier understanding.

Reference

Fang, A., Peng, X., Zhou, J., & Tang, L. (2018, September). Research on the map-matching and spatial-temporal visualization of expressway traffic accident information. In *2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE)* (pp. 23-27). IEEE.