J Vis (2020) 23:1035–1052
https://doi.org/10.1007/s12650-020-00674-7

REGULAR PAPER

Yong Zhang · He Shi · Feifei Zhou · Yongli Hu · Baocai Yin

# Visual analysis method for abnormal passenger flow on urban metro network

Received: 15 August 2019 / Revised: 2 April 2020 / Accepted: 5 April 2020 / Published online: 10 July 2020

**Keywords** Metro data visualization · Anomaly verification · Visualization analytics

## 1 Introduction

Public transit, especially the metro system, has become one of the main facilities to carry mega passenger flows in big cities like Beijing and New York. However, with the increase in passenger flow, some unexpected situations pose negative effects on the metro system, such as the gatherings of passengers and bad weather. If anomalies cannot be effectively observed and controlled in time, they may spread rapidly with the expansion of negative influences. Therefore, the accurate detection of abnormal events in public transit, as well as the effective reasoning of potential causes can help relevant departments propose effective emergency measures to prevent abnormal events from occurring again.

Several approaches have been developed to overcome the aforementioned issues. However, they mainly focus on detecting abnormal stations without exploring the reasons behind the abnormal passenger flow in subways. In addition, it is difficult to verify their results as well. Only through the complex analysis of the original traffic data may users be able to understand whether the test results are reasonable. These kinds of methods are not suitable for users without the background of computer science, and therefore they cannot satisfy the requirement of handling huge and complex traffic data efficiently. At the same time, the rapid development of social network encourages people to express their views and ideas on the social media platform such as Weibo. The massive social network data may contain clues of abnormal situations in metro systems. Manual queries of relevant social network data are labor-consuming and may lead to heavy workload with low efficiency. Thus, it is necessary to develop an intuitive and interactive visual analytic system to detect abnormal events and reason their causes with an auto-integration of semantic data and smart card data.

In this paper, we try to display the original traffic data and verify the anomalies through 3D visualization-based method. At the same time, we also provide a set of visualization views based on Weibo data to explore the causes of anomalies. The interactive visual analytic system interface we have designed is shown in Fig. 1. The main contributions of the paper are as follows:

Y. Zhang · H. Shi · F. Zhou (✉) · Y. Hu · B. Yin
Beijing Key Laboratory of Multimedia and Intelligent Software Technology, Faculty of Information Technology, Beijing Artificial Intelligence Institute, Beijing University of Technology, Beijing 100124, China
E-mail: 1441650574@qq.com

Y. Zhang
E-mail: zhangyong2010@bjut.edu.cn

- A 3D hierarchical view is proposed to analyze spatio-temporal data of metro systems and each layer shares the same geographic information. It not only displays multiple attributes of metro systems (e.g., load factors, passenger flow), but also demonstrates abnormal stations as well as abnormal section flow.
- Multiple coordinated views are proposed to locate suspicious Weibo data which potentially contain the causes of anomalies. Through these views, users can efficiently find the causes of anomalies from the original Weibo texts.

## 2 Related work

### 2.1 Visualizations of anomaly detection

Anomaly detection combined with visualization technology provides many useful tools for users. For examples, the anomaly detection visualization scheme proposed by Wang et al. (2012) combined GIS, GPS, database and other technologies to create a usable visual analysis tool and offer other researchers new visual project guidance. Riveiro and Falkman (2011) discussed the role of visualization and interactivity in the process of anomaly detection as well as analyzed the challenges on existing anomaly detection and behavior analysis methods for customer use and maintenance. At the same time, anomaly detection visualization is widely used in many specific aspects. In the terms of safeguarding network security, Fan et al. (2019) proposed a novel smart labeling method to detect network anomalies through the iterative labeling process of the users. Their method can achieve a good result on anomaly detection with fewer labeled samples. In the field of marine dynamics, ocean surface current visualization plays a significant role. Zhang et al. (2019b) proposed a visualization approach widely used for mesoscale eddy detection. This method is based on an amended Helmholtz–Hodge decomposition and 2D/3D ocean surface current. In road traffic, there are also many researches about anomaly detection visualization. Riveriro et al. (2017) proposed a visual analytic framework, bridging the gap between computational and human approaches by visual analytics to detecting anomalous behavior in road traffic. Kalamaras et al. (2018) proposed an interactive visualization analysis platform which allows exploration of historical data and the prediction of future traffic through a real-time interactive interface, and this platform can also realize visual detection of abnormal events. Another interactive tool for visual analysis and exploration of urban traffic congestion was proposed by Wang et al. (2013). It extracts information about traffic congestion from the blocked place to the entire city in a multi-level way. In addition to letting users understand unexpected traffic conditions through visualization tools, Cheut et al. (2014) also proposed a system with an interactive platform for users to explore, analyze and visualize hidden traffic patterns.

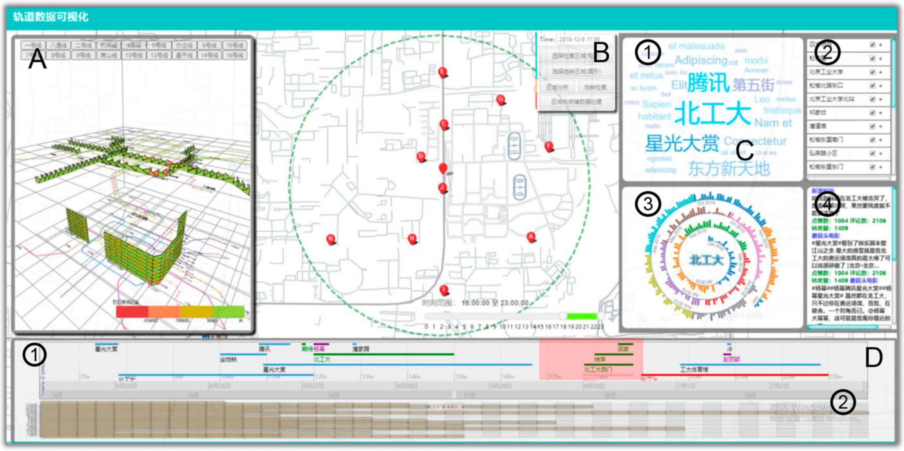


**Fig. 1** System interface: (*A*) 3D hierarchical view, for verifying anomalies. (*B*) Map view, for displaying anomalies and providing interactions. (*C*) The first part is word cloud. The second part is the list of keywords of POIs around the abnormal stations. The third part is the spiral view of topic words and the forth part is a list of Weibo text (*D*). The first part is for showing the changes of topic words over time during the day. The second part is a overview of anomalies

Anomaly detection and visualization are widely used in many fields, but they have been rarely applied in public transportation. In this paper, we provide an interactive visualization way which not only shows the results of anomaly detection, but also provides users with an effective tool for detecting and analyzing the causes of traffic anomalies.

### 2.2 Visualizations of traffic data

The urban rail transit visualization is closely related to the visualization of geographic information system (GIS). There are currently many studies on the application of GIS information system visualization. Yang et al. (2019) designed and realized a forest disease visual analysis system to help researchers and decision makers on issue related to forest diseases. Zhang et al. (2019a) proposed a visual analytics system for studying the evolution and correlations to a number of weather factors of haze. They also developed a comparative visualization to consistently overview trends of scalar variables and wind directions.

For traffic flow visualization, many studies have discussed the 2D visualization methods (Andrienko and Andrienko 2017) and it is common to represent traffic flow by points (Andrienko et al. 2017), lines (Alper et al. 2011) and regions (Collins et al. 2019). In order to display the traffic data better, Andrienko and Andrienko (2008) and Slingsby and Wood (2010) divided map into regular rectangles by multiple heat-maps, and each heat-map represents the time overview of traffic volume in designated areas. Wang et al. (2014) also used heat-maps to visualize in traffic speed of selected road sections over time. In addition to researching traffic data within the area, many researchers also concentrate on values between two points or flows, which is called OD movements. Zeng et al. (2016) proposed waypoint-constrained OD visual analytics which is a new approach for exploring path-related OD patterns in an urban transportation network. As for approaches mentioned above, they mainly studied cell status pattern and the correlation of flow patterns between pairwise cells, while rarely used social network data to explore the cause of the abnormal situation, which is our main research focus in this paper. The TripVista system developed by Guo et al. (2011) provided an interactive method to analyze the trajectory data of pedestrians and different types of vehicles and pedestrians at intersections. TrajectoryLenses (Krüger et al. 2013) allowed users to select specific regions and time to analyze the trajectory data from the origin to the destination. Zeng et al. (2014) proposed a series of visualization techniques to analyze passenger routes and traffic efficiency in public transport system. Chen et al. (2016) proposed an interactive visualization method to discover motion patterns. Wang et al. (2013) put forward a method of traffic congestion analysis based on vehicle trajectory information visualization.

Considering the high-dimensional characteristics of traffic data, some researchers devoted to 3D visualization of traffic data. Tominski et al. (2002) showed the effective use of 3D trajectory bands on visualizing trajectory data. During the process of visualization, attribute data of individual trajectories are visualized as color-coded bands and sets of trajectories are visualized by stacking the bands. Cheng et al. (2013) also utilized 3D staked bands to represent the overview of spatiotemporal changes of attribute data on the road network, but this method is likely to cause visual confusion when displaying multiple data in the same period. Itoh et al. (2016) proposed a limited Ribbon view which could provide a good horizontal contrast of passenger flow among metro lines in a certain period and display multiple metro flow attributes simultaneously. However, its ability to display historical data in the same period was weak.

This paper manages to verify anomalies by displaying spatio-temporal traffic data of metro. To be specific, multi-attributes should be displayed simultaneously in the visualization. Moreover, the visualization should provide not only horizontal comparisons of passenger flow in different metro lines in a certain period, but also vertical comparisons of passenger flow in the certain subway lines and stations at the same time in the historical time.

### 2.3 Visualizations of social network data

With the development of information technology, social network data visualization has become a research hotspot. ThemeRiver visualizes user-selected themes of news stories as horizontally centred stacked graphs (Havre et al. 2000). A recent revisit of the application and adaptation of ThemeRiver approach entertainment datasets suggests new methods for ordering and colouring the streams (Byron and Wattenberg 2008). The main limitations of these stacked graph techniques are the fixed length of time bins, the static selection of themes and the lack of zooming or filtering operations. To solve those limitations, several approaches have been suggested for the visual exploration of blog posts by visualizing tags and comments arranged along a time-axis (Vassileva and Gutwin 2008) or by providing facet visualization widgets for visual query formulation according to time, place and tags (Dörk et al. 2008). However, it's difficult for these methods to explore the causes of subway anomalies through social media data.

Besides the work above, Itoh et al. (2016) proposed TweetBubble view to explore the information related to Japanese metro. They emphasized the interactions but neglected the external factors of metro stations. Moreover, their visualization method failed to fully show the evolution of the topics in each period of a day and could not make a detailed comparative analysis between the topics and the changes of the metro traffic. Compared with TweetBubble view (Itoh et al. 2016), this paper not only considers the number and frequency of topic words, but also pays more attention to the distribution of each topic word in a day. Therefore, the human–computer interaction in our system is prone to be more flexible and comprehensive.

## 3 Data description

### 3.1 Smart card data

The smart card data in this paper are provided by Beijing Traffic Information Center. Each record has 37 attributes including user card number, entry and exit time, entry and exit lines and stations, etc. We use Dijkstra algorithm to estimate passenger flow between stations. For passenger flow at a subway station, because most of the subway anomalies are caused by uncertain abnormal weather, unexpected human events, and large-scale celebrations, the abnormal events are irregular and have low probability, and therefore the sample content of the traffic data set is relatively small. At the same time, according to the regular travel of residents, the traffic flow is able to take year, month, day and hour as the operating cycle, so the traffic flow matrix has a low rank structure (Zhou et al. 2016). However, due to the existence of abnormal conditions, the low rank structure of traffic flow is destroyed. The low-rank representation model RPCA can decompose the original matrix into a normal traffic matrix with low rank and an abnormal traffic matrix, which can be used to detect and extract abnormal passenger flow in the subway. At the same time, considering the time correlation between data, an improved RPCA model (Wang et al. 2018) is used to detect the abnormality of subway passenger flow data.

The subway passenger flow can be represented by two components: the expected flow $X$ and the anomalous part $A$. The anomalous part includes special events in or around the site. Except for some apparent daytime cycle anomalies derived from passenger flow measurements, for the subway passenger flow matrix $\mathbf{D}$, two adjacent rows of the same working day in different weeks are generally approximately equal. The matrix $\mathbf{D}$ is constructed based on the raw riding data, which is organized with the row and column corresponding to the date and the time interval of each day, respectively. This property is significant for the corresponding exception flow $X$, while current RPCA model has no specific description for this important property. The time constraint matrix visually expresses the fact that the nominal passenger flow matrix for the same time interval of the same working day is generally similar. Therefore, item $\|\mathbf{HX}\|_1$ containing the time constraint matrix $H$ can be used to maintain consistency between $X$ lines (Wang et al. 2018).

The improved RPCA model is given by

$$\begin{aligned} \min_{\mathbf{X},\mathbf{A}} \quad & \|\mathbf{X}\|_* + \lambda_1\|\mathbf{A}\|_1 + \lambda_2\|\mathbf{HX}\|_1, \\ \text{s.t.} \quad & \mathbf{D} = \mathbf{X} + \mathbf{A} \end{aligned} \tag{1}$$

where $\mathbf{H} = \text{Toeplitz}(0, 1, -1)$, $X$ is the predicted passenger flow matrix, and $A$ is noise. The first two terms penalize the low rank and sparse properties of $X$ and $A$, respectively, and the third term penalizes the sparse consistency among all rows of $X$.

### 3.2 Social network data

Now more and more people choose to express their views and opinions on social platforms. At the same time, social platforms have also become a gathering place for people to express their feelings. Sina Weibo currently has more than 400 million registered users and allows them to post short messages, the number of words is limited to 140, and there are about 30–60 million active users every day, generating nearly 150 million Weibo data. After massive manual queries, we found that there are a lot of comments and blog posts related to the abnormal passenger flow events on Weibo. Through these data, we can accurately grasp and analyze abnormal passenger flow conditions and causes of the abnormalities.

In order to explore the cause of abnormal passenger flow events in Weibo deeply, we mainly extract the author, timestamp and content of Weibo. The number of comments, retweets and likes are also recorded. The author and timestamp of Weibo's release are easy to obtain, and they play important roles in revealing the activities of participants and the development of social events. When an overactive Weibo blogger is found, his or her information will be deleted, because most of the information is advertising, and it has no reference value for deeper exploration. The main content of the Weibo is highlighted by extracting the text topic to distinguish whether the Weibo is effective for the visual analytics system.

Therefore, the latent Dirichlet allocation (LDA) model is used to extract the topic words from the original Weibo data. LDA is a three-layer Bayesian probability model extended on the probabilistic implicit semantic index. It is a document generation probability model. The model contains three layers, namely documents, topics and terms. The basic idea is to treat documents as a collection of topics, and each topic is represented as a probability distribution of related terms. Finally, the subject vocabularies extracted from all Weibo documents are summarized and sorted through a program. Among all the Weibo vocabularies, some words with a high repetition frequency and their related words often point to the same social event. In addition, the impact of these social events is often proportional to the intensity of the subject words. According to the number of repetitions, the topic words are sorted by the number of repetitions, the index of the topic vocabulary and the original Weibo are established, and all the relationships are stored.

## 4 Overview

To solve the aforementioned problem, several specific visualization tasks should be supported by our visualization systems:

- The anomalies detected by models should be visually displayed, and the abnormal interest area can be interactively selected for analysis according to user's interest.
- Visualization system should not only provide horizontal comparisons of passenger flow among metro lines and stations in a certain period, but also vertical comparisons of passenger flow among the same time in history of certain subway lines and stations, so that users can intuitively obtain all the original traffic data which are needed for the verification of anomalies.
- When exploring the reasons from Weibo data, keywords of all POIs around abnormal stations in a certain range should be obtained to retrieve the original Weibo text, and visual interactions for Weibo topic words should be provided to help users obtain reliable and important information.

In order to accomplish these analysis tasks, we design a visual analytics system. The overview of our system is shown in Fig. 2. It consists of two main parts. One part is the data processing and analysis in back-end, the other part is the visualization system in front-end. Smart card data and Weibo data are stored in the back-end database after preprocessing. Our system is implemented based on a browser-server architecture, where the server hosts a data-analysis phase on exploring abnormal events, verifying them and reasoning the causes, while the browser provides a visual analytical phase by referring to the results from the former phase.

The anomaly detection module is mainly used to display anomalies and provide suggestions of suspected stations. It has spatiotemporal interaction controls, which can help users select regions of interest in the 2D map, so as to study the metro spatiotemporal data of anomalies in more details in 3D views.

The anomaly verification module displays the spatiotemporal data of the stations and routes according to the user's area of interest in a 2D map. The multi-dimensional attributes for anomaly verification include abnormal stations, station intervals and their corresponding passenger flow, passenger load rate, geographical location and time. To show these attributes precisely and accurately, we need more visual coding methods for high-dimensional data. In order to ensure the intuitiveness of the data visualization and avoid the interference between those attributes, we design a 3D layered view. 3D space has the advantage of integrating additional visual information into the presentation (Tominski et al. 2005). The 3D view could sort, classify and display hierarchically important data related to abnormal passenger flow. We develop anomaly verification modules by employing 3D layers to combine the attributives, e.g., into one view to represent spatial dependencies. Besides, we visualize data with temporal dependencies with an approach named 'focus + context' developed by Tominski et al. (2005), in order to hide irrelevant information and

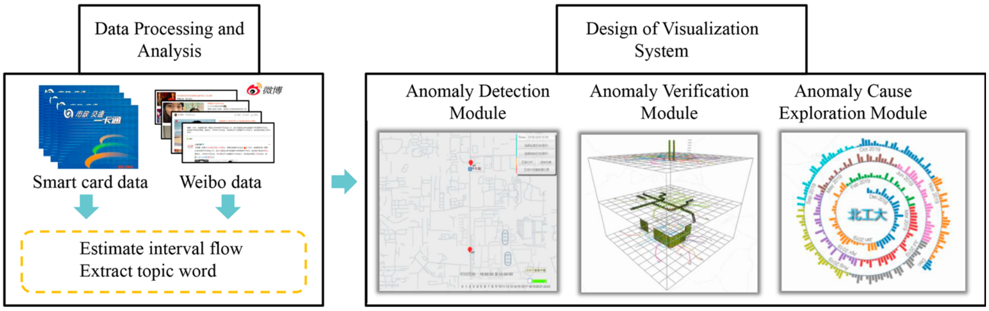

Data Processing and Analysis
Smart card data
Weibo data
Estimate interval flow
Extract topic word
Design of Visualization System
Anomaly Detection Module
Anomaly Verification Module
Anomaly Cause Exploration Module
北工大

**Fig. 2** Overview of our system

concentrate on detected events rather than the whole dataset. Therefore, we can clearly display the information of the user's attention in each layer of the 3D hierarchical view and effectively avoid the interference of information in other areas.

The main process of exploring the causes of anomalies is to extract topics related to the abnormal stations firstly and then visually help users discover topics that may be related to the subway anomalies, and finally find reasons in Weibo texts. Instead of an accurate answer, we provide an interactive visualization tool for users.

## 5 The visual analytical system

As shown in Fig. 2, the visualization part of our system has three modules: anomaly detection module, anomaly verification module and anomaly cause exploration module.

### 5.1 Anomaly detection module

The visualization of this module consists of two components: map view and overview of anomalies. The map view is designed as Fig. 1(B), which has a time selector as Fig. 3(C) and space selectors as Fig. 3(B) to help users select time and regions. In map view, abnormal stations in the selected time are shown as Fig. 3(A). The overview of anomalies is shown in Fig. 4. The abscissa represents stations, and the ordinate represents the metro line. The yellow grids represent normal stations and the red grids represent abnormal stations. When clicking on the grids of the view, they can interact with the map view and locate the abnormal stations in the map.

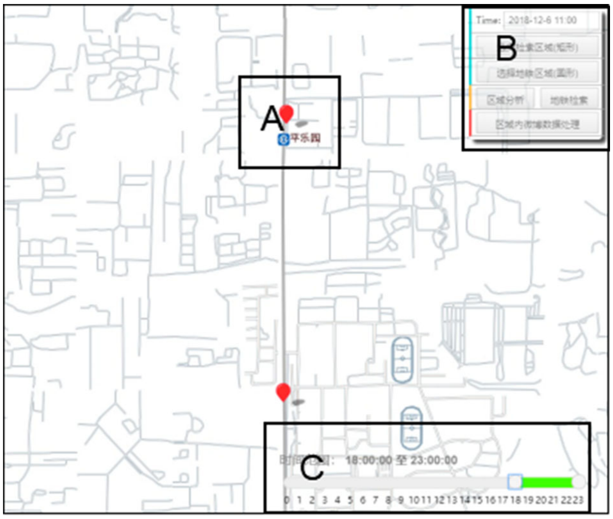


**Fig. 3** Map view of the system: (*A*) The abnormal stations detected by models. (*B*) Operation area which offers space selectors and some operation buttons. (*C*) Time selector which offers the function of selecting time range

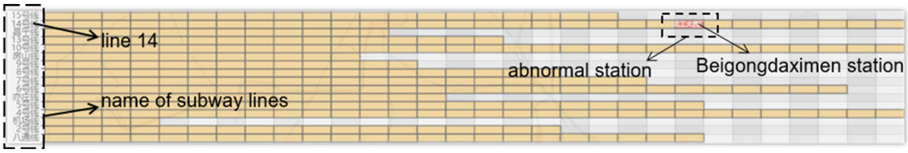


**Fig. 4** The overview of anomalies. This view shows an abnormal overview of the whole metro system at a certain time. The longitudinal coordinates represent the line name. Each grid represents the corresponding stations of the subway line. The normal stations are expressed by yellow. The abnormal stations are expressed by red and displayed name

## 5.2 Anomaly verification module

The 3D layered view we designed is divided into three layers from top to bottom. As shown in Fig. 5, the first and third layers only display the spatiotemporal information of the stations and related lines within the scope of choosing. The second layer shows the information of all lines to provide a horizontal comparison of traffic flow. Each layer of the view is drawn with a Beijing subway map. The horizontal positions of three subway maps are the same, but the vertical positions are different.

In the subway map of each layer, the stations are represented by the circles in the corresponding positions. The subway line is represented by the connection lines between stations. The 3D layered view is designed as a scalable and draggable view. The whole view can be dragged and zoomed in or out by mouse, so that users can observe the details they want. In order to ensure the user's visual effect when analyzing passenger traffic using 3D-layered views, the zoom range of the view is controlled to a certain extent. In addition, we provide hidden and appear functions of subway lines. Users can display subway lines according to their needs and hide unnecessary lines.

As shown in Fig. 6, the first layer mainly shows data of stations in the metro system. Each station is represented by a circle in its corresponding geographical location, and the different color of the circle represents the number of the passengers flow during a period. We use red, pink, yellow and green to represent the decline of the passenger flow in turn. We place a cuboid at the abnormal station, which is made up of several small cuboids. The color of the bottom cuboid represents the passenger flow of the metro station at a selected time. From the second cuboid, each one from bottom to top represents the passenger flow at the same time in the past every week. Through the view, we can intuitively see the difference of passenger flow between the interest stations and other stations at the selected time, and whether the stations have similar passenger flow at the same time in history.

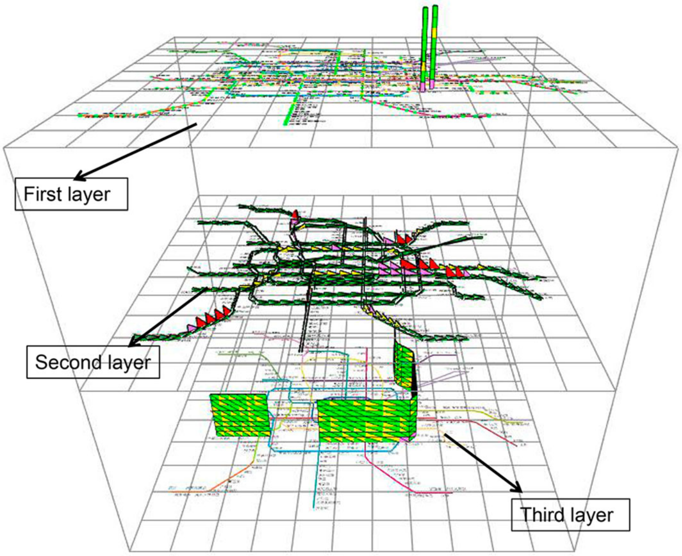


**Fig. 5** The 3D hierarchical view. The first layer is the information about stations. The second layer is the information about sections between stations at the selected time. The third layer shows the historical information of metro sections between stations

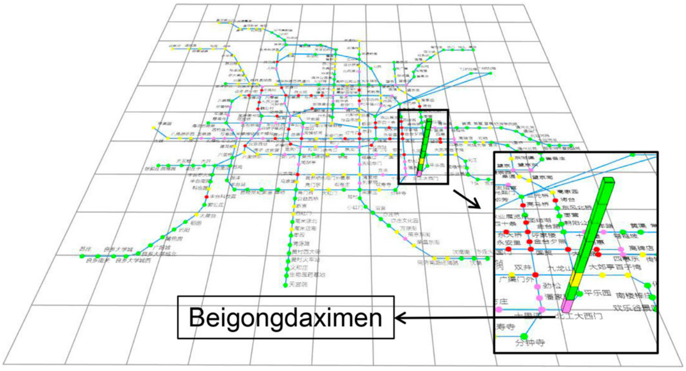


**Fig. 6** The first layer of the 3D hierarchical view, it shows the passenger flow in *Beigongdaximen* at 5 pm on December 3rd, 2017, and historical passenger flow in the same time of each week

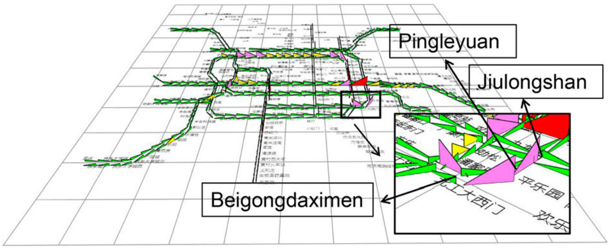


**Fig. 7** The second layer of the 3D hierarchical view, this layer relies on Beijing Metro Map to provide geographical location. We use vertical right triangle to connect two stations to represent the section information. The color of right triangle represents the load factors which more than 120% is expressed by red, 100–120% is expressed by pink, 80–100% is expressed by yellow, below 80% is expressed by green, and different heights represent different traffic flow. The higher the height, the greater the traffic flow. This figure shows the section passenger flow between stations around *Beigongdaximen* station at 5pm on December 3rd, 2017

Figure 7 shows the second layer that mainly illustrates the passenger flow of each line at the selected time. We use right triangles to represent the information of the sections between adjacent stations, where the color represents the load factors with the same setting as the first layer and different heights represent the passenger flow. To distinguish directions of a metro line, the right triangles on a metro line are connected to form a serrated shape. The directions of the serrated shape represent the directions of the metro line, as shown in Fig. 8. This layer can intuitively represent the passenger flow and the load factors of each metro line in the whole metro system at a selected time. The sections related to abnormal stations and other metro sections can be compared horizontally.

The last layer is shown in Fig. 9. A 3D wall view is used to show the passenger flow on a specific line where the abnormal stations are located. The view is divided into many layers just like the cuboids in Fig. 6. The adjacent stations on each layer are connected by rectangles with equal heights. Each rectangle is divided into two right triangles by a diagonal line, representing two directions, respectively, at the same time as shown in Fig. 10. From bottom to top, different colors (which follow the same setting as the first layer) of the right triangles means different load factors at the same time in every past week.

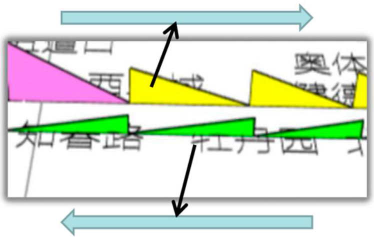


**Fig. 8** The right triangles on a metro line are connected to form a serrated shape, and the direction of the serrated shape represents the direction of the metro line

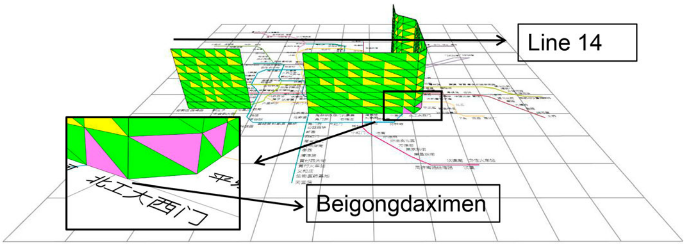


**Fig. 9** The third layer of the 3D hierarchical view, it shows the passenger flow of sections between stations around *Beigongdaximen* station at the same time in history

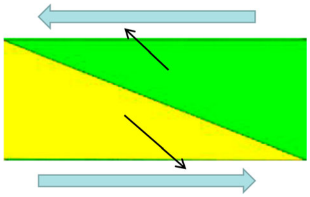


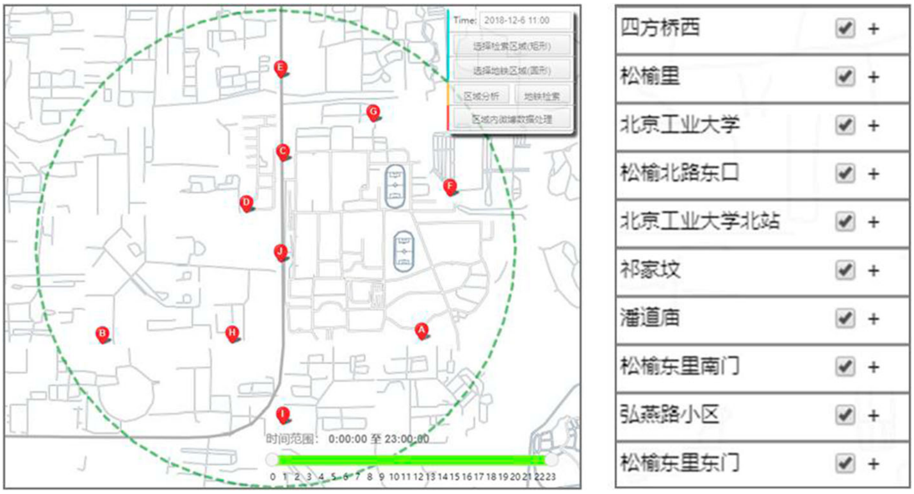
**Fig. 10** Each rectangle is divided into two right triangles by a diagonal line, representing two directions, respectively, in the same time

Through the three layers view, users can not only verify abnormal stations, but also find abnormal sections around the abnormal stations.

## 5.3 Anomaly cause exploration module

We hope to be able to explore the cause of the abnormal situations of subway stations through Weibo data. Therefore, the first step is to detect the original Weibo text related to the abnormal subway station. Instead of directly retrieving the names of stations among all the original Weibo texts, we observe that the stations are affected by not only the stations themselves, but also some activities within a certain range of the stations. Thus, we provide a keyword acquisition tool, which can obtain POIs and station names in a certain range of the stations, aiming to obtain more useful Weibo data and provide more convincing visualization results.

The keyword acquisition tool is based on the map area of Fig. 1(B). Users can take an abnormal station as a center and circles an area to obtain keywords of the stations and POIs. Our system can automatically acquire POI keywords in this area such as the names of schools, shopping malls and bus stops. These keywords can be displayed in the form of a list as shown in Fig. 11. The words that users are not interested



四方桥西 ☑ +
松榆里 ☑ +
北京工业大学 ☑ +
松榆北路东口 ☑ +
北京工业大学北站 ☑ +
祁家坟 ☑ +
潘道庙 ☑ +
松榆东里南门 ☑ +
弘燕路小区 ☑ +
松榆东里东门 ☑ +

**Fig. 11** The acquisition tool of the keywords, the left part is the POIs around *Beigongdaximen* station on the map. The right part is the edit list of keywords

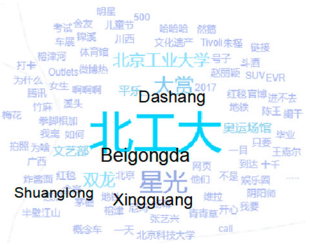


**Fig. 12** Word Cloud of topic words related to *Beigongdaximen* station

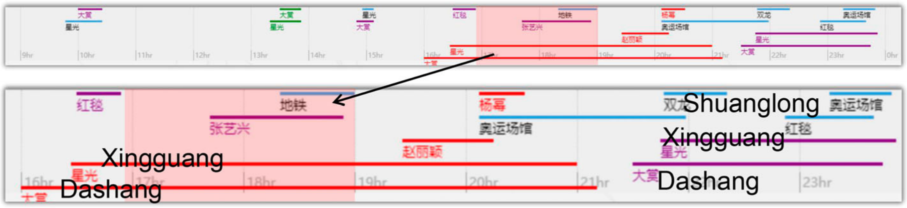


**Fig. 13** This view is used for showing the changes of topics around *Beigongdaximen* station over time during the day

in can be deleted. The keywords that are expected to appear can be added, and keywords that do not conform to idioms can be edited. We provide three views which cooperate together to explore the causes of anomalies.

*Visual discovery of abnormal topic words* As shown in Fig. 12, we extract the topics and calculate the frequency of each topic to generate a word cloud based on an iterative force-directed layout algorithm. High-frequency words gather in the center and display closely, while low-frequency words loosely distributed around. Through the word cloud, we know which topics of micro-blogs appear frequently on a selected day. Therefore, it is reasonable to speculate that some related activities are being held near the subway station in the selected time, which may cause the subway station to be abnormal.

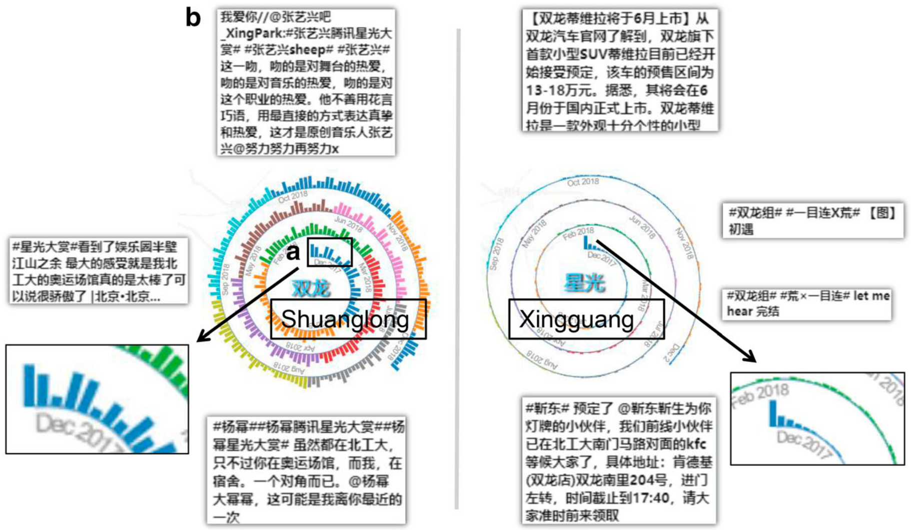
*Temporal correlation between abnormal topics and stations* As shown in Fig. 13, we combine topic words with time axis to show how topics change over time during a day. When a topic word appears, if it appears repeatedly within half an hour, we connect the two time points with a straight line. When that word reappears in the next half an hour, we connect them with a line again. These steps are repeated until all the words have been traversed. Then the average frequencies of the topic words corresponding to each line are calculated in every half an hour. We color the line with green, blue, purple and red to represent the increase in topic frequencies. Red transparent areas on the time axis indicate a metro anomaly occurs at the station during a period of time on the selected day. Therefore, we can clearly notice the topics with numerous and frequent appearance during the anomaly occurrence through the view. Combined with word cloud, we can lock some topic words which are most related to anomalies.

*Exploring the history of abnormal topics* We hope to find out whether some topic words appeared before or exhibited similar frequencies of occurrence in the history. Hence, we display a histogram in a spiral way as shown in Fig. 14. The first bar at the center shows the frequency of an abnormal topic word on a selected day, and the rest bars show the frequencies of the topic word in 365 days before the selected day. The words are displayed in different colors and months. The word at the center of the spiral diagrams represents the suspicious topic word. By clicking on it, users can interact with the list of Weibo. From this view, we can intuitively get the topic frequency of every day in the past year.

Combining the three views together, we can quickly know whether a word is connected with metro anomalies. Clicking the word of the word cloud, the corresponding original Weibo texts can be found in the list as shown in Fig. 1(C)(4). We can find the answers from the original Weibo texts that most likely contain the causes of anomalies.

## 6 Case studies

In order to verify the effectiveness of the system, three factual cases are introduced in detail in this section.



我爱你//@张艺兴吧_XingPark:#张艺兴腾讯星光大赏# #张艺兴sheep# #张艺兴# 这一吻，吻的是对舞台的热爱，吻的是对音乐的热爱，吻的是对这个职业的热爱。他不善用花言巧语，用最直接的方式表达真挚和热爱，这才是原创音乐人张艺兴@努力努力再努力x

#星光大赏#看到了娱乐圈半壁江山之余 最大的感受就是我北工大的奥运场馆真的是太棒了可以说很骄傲了 |北京·北京...

#杨幂##杨幂腾讯星光大赏##杨幂星光大赏# 虽然都在北工大，只不过你在奥运场馆，而我，在宿舍。一个对角而已。@杨幂大幂幂，这可能是我离你最近的一次

【双龙蒂维拉将于6月上市】从双龙汽车官网了解到，双龙旗下首款小型SUV蒂维拉目前已经开始接受预定，该车的预售区间为13-18万元。据悉，其将会在6月份于国内正式上市。双龙蒂维拉是一款外观十分个性的小型

#双龙组# #一目连X荒# 【图】初遇

#双龙组# #荒×一目连# let me hear 完结

#靳东# 预定了 @靳东靳生为你灯牌的小伙伴，我们前线小伙伴已在北工大南门马路对面的kfc等候大家了，具体地址：肯德基(双龙店)双龙南里204号，进门左转，时间截止到17:40，请大家准时前来领取

**Fig. 14** The spiral view of topics, area (*a*) represents the frequencies of a topic word in the past one year. The left spiral view is about the word “Shuanglong”. The right spiral view is about the word “Xingguang”. Dialogs around the spiral views is the micro-blog texts corresponding to topic words which usually display as Fig. 1(C)(4)

### 6.1 The anomalies around *Beigongdaximen* station

Our system detects some traffic anomalies that may occur at the *Beigongdaximen* station around 5:00 p.m. on December 3rd, 2017, so we use our system to analyze the station. Firstly, we verify the anomalies via 3D hierarchical view. Through Fig. 6, which is the first layer, we can see that the traffic flow of the station is not too high compared with other high traffic flow stations such as *Tiantongyuan* and *Xidan*. Nevertheless, compared with its historical data, we can find that in the same period of the past 8 weeks the traffic volume of the station is much lower than the volume on Dec 3rd 2017, so the model test results are basically reasonable. Through the second and the third layers of the 3D hierarchical view, we can find the abnormal section flow between stations. As shown in Fig. 7, the passenger flow from *Jiulongshan* to *Pingleyuan, Pingleyuan* to *Beigongdaximen*, and *Shilihe* to *Beigongdaximen* at 5:00 p.m. are relatively high. According to the 3D wall in Fig. 9, it can be seen that their passenger flow in this period is obviously different from the traffic flow at the same time of history. We can discover that the traffic flow of these sections is abnormal. Moreover, these metro sections are all abnormal in the directions to the *Beigongdaximen* station, so it is very likely that some events which attract passengers occurred near the station.

We continue to explore the causes of anomalies. There are many schools, shopping malls and hotels near the station. We can get all POI keywords in this area through our system, and get micro-blog text for further visualization analysis. First of all, through word cloud in Fig. 12 we can find that the topic of “Beigongda” appears most frequently, which is normal, because the name of the subway station is “Beigongdaximen”. However, there are also some strange words like “Xingguang”, “Dashang” and “Shuanglong”. For their unknown relationship with the station, we mainly focus on these words and make further analysis. In Fig. 13, we find that “Xingguang” and “Dashang” appear frequently from 4:00 p.m. to 7:00 p.m., which coincides with the time of the occurrence of anomalies. In the spiral chart (Fig. 14), we find that the topic of “Xingguang” has not appeared so frequently in the history except the selected day. We suspect that some of its relating events are the causes of anomalies. From the original micro-blog, we find that a very important activity of Tencent company was held in the Olympic Stadium of Beijing University of Technology named “Xingguangdashang” (including “Xingguang” and “Dashang” in Chinese). We check some of the original texts corresponding to other topics such as “Shuanglong”, but do not find other large-scale activities or abnormalities. At last, we speculate that the anomalies are caused by an activity named “Xingguangdashang” held by Tencent company. So at about 5:00 p.m., the abnormal increase in traffic flow occurred in the metro sections around the *Beigongdaximen* station but only happened in the directions toward the station.

## 6.2 The anomalies of multiple metro stations in *Chaoyang* district

We detect some anomalies may occur at several subway stations in *Chaoyang* district around 7:00 p.m. on August 7th, 2015. We analyze several stations with large passenger flow, including *Guomao* station, *Shuangjing* station and *Yonganli* station. Firstly, we validate these stations through 3D hierarchical view and find that the traffic flow of these stations is different from the traffic flow in history. As shown in Fig. 15, the passenger flow at *Jianguomen* station is basically normal, because the corresponding cuboids are all green, indicating that the station has no congestion at the same time in the history, but *Yonganli* station, *Jintaixizhao* station, *Dawanglu* station, *Guomao* station and *Shuangjing* station have obvious traffic anomalies. At 7:00 p.m. on August 7th, their passenger flow color is pink, but at 7:00 p.m. in the history at the same time, their colors are red which means extremely crowded. The passenger flow on August 7th is obviously much less.

As shown in Fig. 16, we find a lot of abnormal passenger flow through the second and third layers. The arrows of line 10 in Fig. 16(3) indicate the sections of *Hujialou–Jintaixizhao*, *Jintaixizhao–Guomao*, *Guomao–Shuangjing*, the arrows of line 1 in Fig. 16(4) indicate *Jianguomen–Yonganli*, *Yonganli–Guomao*, *Guomao–Dawanglu* and *Dawanglu–Sihui*. The flow in the same period of history is obviously much larger. Comparing with the dentate view from 7:00 p.m. (Fig. 16(1)) to 8:00 p.m. (Fig. 16(2)), we can see that the traffic flow of some sections of Line 1 and Line 10 begins to increase around 8:00 p.m. and the evening peak of some lines is obviously delayed. To find the reasons for the large-scale delay of the evening peak in *Chaoyang* District, we inquire the accident records of the metro management department. There were neither accidents in the subway stations on August 7th, nor road closure.

We find that there are some topic words such as "heavy rain", "hail", "waiting" through the word cloud (Fig. 17(1), Fig. 17(3)) In addition, "heavy rain" and "hail" begin to appear frequently at 7:00 p.m. and decrease after 8:00 p.m., which basically coincides with the time of anomalies. As shown in Fig. 17(2), "heavy rain" rarely appears and the "hail" has never appeared in the past year.

Therefore, we argue that the large-scale metro anomalies may be caused by the sudden weather change corresponding to the word like "heavy rain" and "hail". By querying the original texts of Weibo in the text list of our system, we find that there is a sudden heavy rainstorm mixed hail in the evening peak period on August 7th, and *Chaoyang* District is the main rainfall area. It can be concluded that many people may had not immediately come back home because of the sudden rainfall, which led to the delay of the passenger flow peak.

The above events affected by external factors cannot be queried on traffic websites and do not have records in the metro management department. If we understand the causes of these anomalies, we can take effective preventive measures when the similar events occur again. For example, if we know that an anomaly occurred in a certain place is caused by a large event (case 1), we can continue observing the changes of the traffic flow after the anomaly happens and summarize the regular of it. When similar activities are held again in this area, relevant departments can regulate in advance. If the relevant departments understand the influence of bad weather (case 2), for example, the delay of evening peak due to the

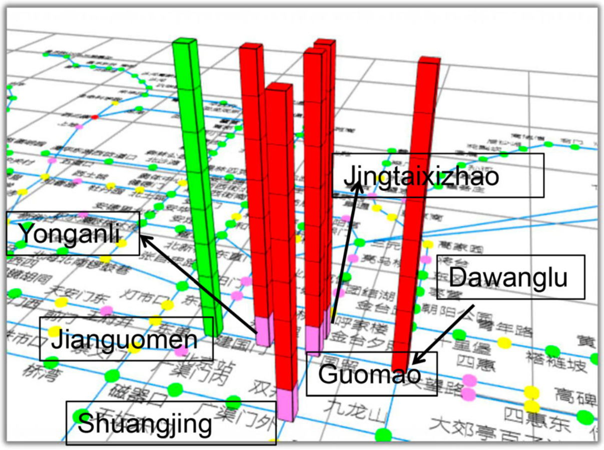


**Fig. 15** Several abnormal stations in *Chaoyang* district

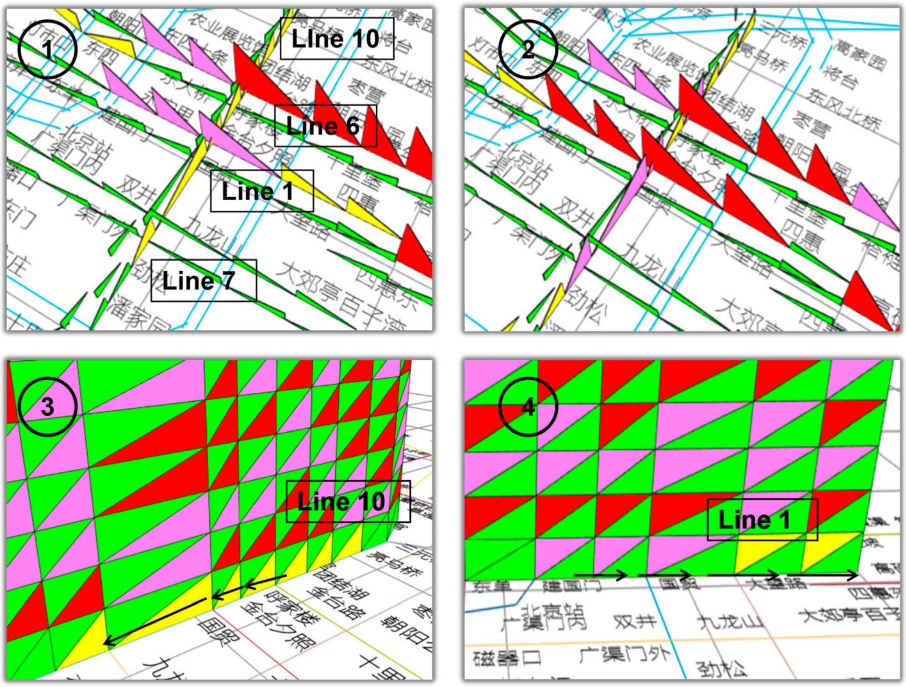


**Fig. 16** (*1*) The traffic flow of several sections in *Chaoyang* district at 7:00 p.m. (*2*) The traffic flow of several sections in *Chaoyang* district at 8:00 p.m. (*3*) The historical traffic flow of several sections of Line 10 in *Chaoyang* district at 7:00 p.m. (*4*) The historical traffic flow of several sections of Line 1 in *Chaoyang* district at 7:00 p.m.

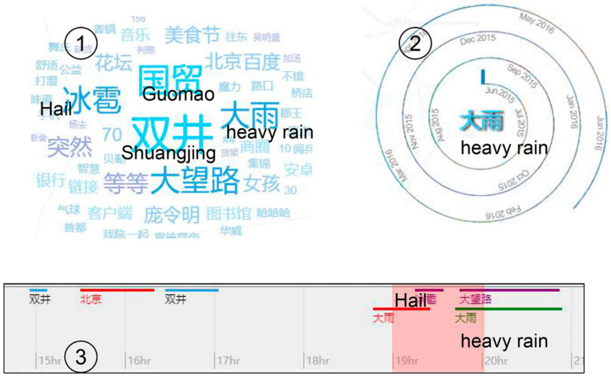


**Fig. 17** The topic analysis of case 2

heavy rain, they can increase transportation capacity at the proper time to cope with time changes of evening peak.

## 6.3 The anomalies of multiple metro stations in *Olympic Forest Park* district

At around 6:00 p.m. on August 23th, 2015, our system detected an abnormal passenger flow event at three stations near *Olympic Forest Park*. Circle these possible anomalies on the map view and obtain passenger flow data to generate a 3D-layered view as shown in Fig. 18.

In the first layer, three stations: *the Olympic Sports Center* station, *Forest Park South Gate* station, and *Olympic Park* station are detected as suspected abnormal stations by the model. The bottom cuboid is red, and the historical traffic of these three stations is basically yellow. Therefore, there is a sudden increase in traffic at 6:00 p.m. on August 23th. Through the second and the third layer, we can basically confirm abnormal section flow, which are *Lincuiqiao-Senli Park South Gate*, *Beishatan-Olympic Park*, *Beitucheng-*

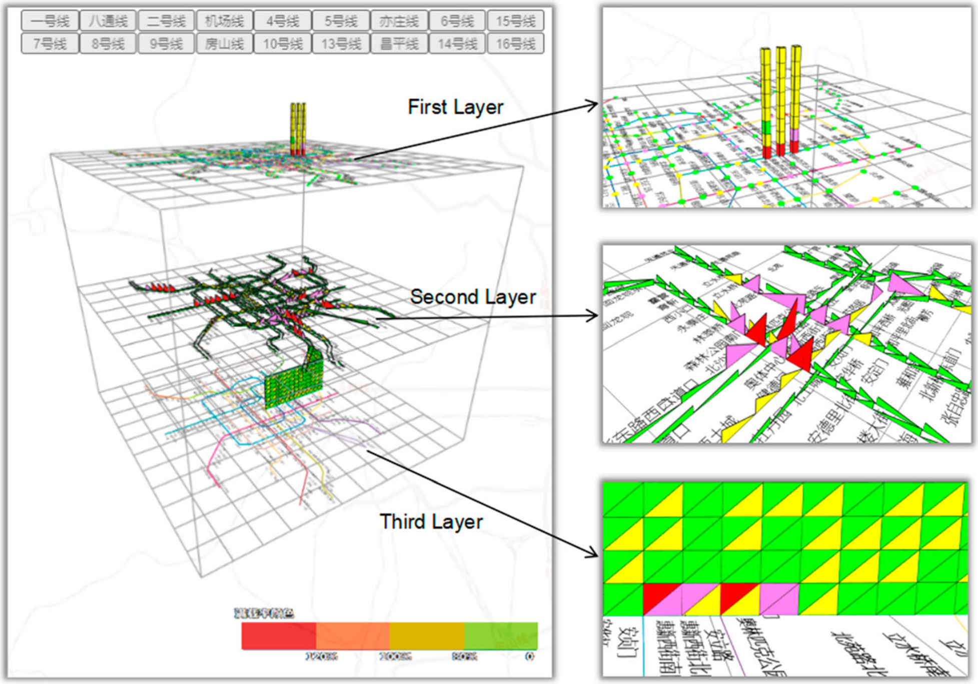


**Fig. 18** The 3D-layered view of case 3: Abnormalities occurred at three stations near the Olympic Sports Center at 19:00 on August 23

*Austrian Sports Center*, *Anli Road-Olympic Park*, *Olympic Sports Center-Olympic Park*. The direction of the anomalous section is all directions to the three anomalous stations. Thus, we can obtain an anomalous area caused by an event to be analyzed.

In order to explore the cause of the event near *the Olympic Sports Center*, the keyword extraction tool is introduced to obtain the keywords and POI of this area. After simple editing of keywords, all Weibo contents related to the three stations and the surrounding POI are obtained. Besides, the topics are extracted to generate a topic word cloud map. As shown in Fig. 19, the subject word cloud indicates that "the Olympics", "opening ceremony" and "tournament" are the keywords that appear on that day.

As shown in Fig. 20, the system generates thematic spiral diagrams of the three keywords of "the Olympics", "the opening ceremony" and "the championship". Check the number of daily occurrences of the three words in the past year. Although the two keywords of "the opening ceremony" and "the championship" have appeared in history, they have appeared in large numbers and intensively, and they appeared on the same day on August 23rd. With the theme time-line diagram corresponding to Fig. 20, it can be found that the occurrence time of the anomaly is about 6 to 7:30 p.m., and the words of "the opening ceremony" and "the tournament" also appear numerous times in the corresponding time interval. The social event corresponding to this theme may be the cause of the abnormal passenger flow at the subway station. The Weibo list generated by the two words demonstrates at 7:00 p.m. on August 23th, *the Bird's Nest Stadium* hosted the opening ceremony of the IAAF World Championships. The opening ceremony attracted a large number of spectators and resulted in short-term increase in the traffic at three stations near the *Olympic Sports Center*.

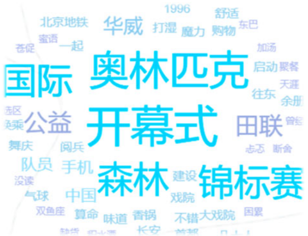


**Fig. 19** The topic word cloud map of case 3: Topic cloud of relevant micro-blog texts in the Olympic Sports Center

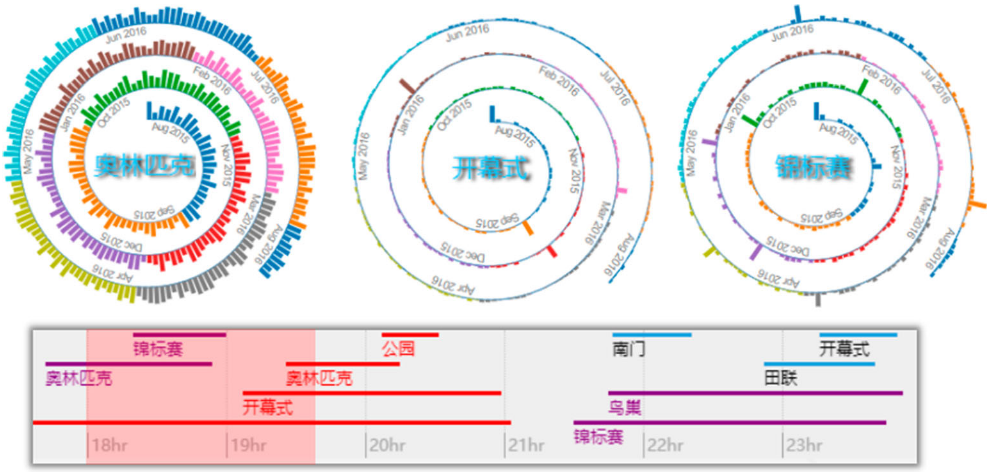


**Fig. 20** Anomaly cause analysis of abnormal traffic flow at stations near Olympic Sports Center: The thematic spiral diagrams of the three keywords of the Olympics, the opening ceremony and the championship.

## 7 System performance analysis

To evaluate the reliability and usability of our system, two more experiments are conducted.

### 7.1 System reliability experiment

The first experiment mainly tests the loading capacity of the system when the amount of data is gradually increasing. The loading capacity of the system is measured by the time it takes for the system to load a 3D hierarchical view. We repeated the experiment six times based on the number of abnormal stations and doubled the amount of data in each group of experiments. The experimental results are shown in Table 1. In Table 1, NAS, NAL, SLT and FPS represent number of abnormal statistics, number of abnormal lines, the system's loading time and frames per second, respectively. From Table 1, we can get the system initialization time is 2.32 s on average, and the frame rate is 60 at system initialization with Vertical Hold turned on. The system load time will inevitably increase with the increase in the amount of data. Obviously, the system load time is not directly proportional to the increase in the amount of data. As the amount of data increases, the system load time increases more slowly. The cases used to prove the performance of the system are based on the 14 lines of Beijing metro transportation system. Under usual circumstances, the number of anomalous sites is about 3–6 and the number of anomalous lines is 1–2. At the same time, considering that the system is actually used for local retrieval, the real-time loading time when the system is really running should be the time to retrieve abnormal sites and lines in the system minus the corresponding system initialization time in the table. Hence, the average time should be 3.58 s. Even in extreme cases, the system load time is around 10 s. In general, the system can complete the analysis task interactively even with a large amount of data.

This experiment is conducted using a computer with Windows 10 Professional 64-bit operating system. The CPU of this computer is Intel (R) Core (TM) i7-7700K CPU @4.20 GHz, the memory is 16.00 GB, and the graphics card is NVDIA GeForce GTX 1080.

**Table 1** The results of system reliability experiment

| | NAS | NAL | SLT | FPS |
|---|---|---|---|---|
| A | 0 | 0 | 0.87 | 60 |
| B | 3 | 1 | 3.26 | 29.5 |
| C | 6 | 2 | 5.63 | 28.1 |
| D | 9 | 3 | 7.68 | 11.4 |
| E | 12 | 4 | 9.86 | 8.0 |
| F | 15 | 5 | 12.49 | 6.8 |

### 7.2 System usability experiment

The second experiment is a system evaluation experiment in which users used the system to fill out questionnaires based on their usage. In this experiment, we contacted a total of 38 users. Before the beginning of the experiment, we introduce the using method of the system to the users, and the users can take 15 min for voluntarily free practice in a certain degree. The evaluation form is designed and summarized according to the SUS system availability scale. The scale has twelve questions. The first two questions are the user's visualization and interactive evaluation of the system. The last 10 questions are users' evaluation of the use experience. The specific problem settings are as follows:

- T4_1: overall visual intuition;
- T4_2: interactive interface a whole;
- T5_1: I will use this system frequently;
- T5_2: This system is complex to use;
- T5_3: This system is easy to use;
- T5_4: Technical training before the experiment helped me to use this system;
- T5_5: This system integrates multiple functions;
- T5_6: This system has functional conflicts;
- T5_7: Other users should soon be able to learn to use this system;
- T5_8: This system is cumbersome to operate;
- T5_9: I am confident when operating this system;
- T5_10: I need to be warmed up before the experiment.

Among the randomly selected volunteers, there were 6 women and 27 men. The evaluation results were the descriptive statistical analysis of SPSS software.

It can be seen from the analysis results that the minimum statistical value of the system evaluation result is 50.0, the maximum statistical value is 77.5, and the average statistical value is 63.092. According to the availability scale evaluation method, the results are in 60 seats. The above shows that the system performance is good, and the experimental results are shown in Table 2, which shows that the system designed in this paper fully meets the standards. In Table 2, the descriptive statistics include ten items. They are N statistics (NS), minimum statistics (MinS), maximum statistics (MaxS), mean statistics (MeanS), standard deviation statistics (SDS), variance statistics (VS), skewness statistics (SS), standard deviation of skewness (SSE), kurtosis statistics (KS) and kurtosis standard error (KSE). According to the univariate ANOVA test based on gender, the evaluation results of male subjects are generally lower than those of female subjects. At the same time, considering the difference in the number of male subjects and female subjects, the system has universal applicability.

Through two experiments, it can be seen that our system has certain reliability and availability and is suitable for local anomaly analysis and has no special requirements for the users.

**Table 2** The descriptive statistics of experiment results

| | NS | MinS | MaxS | MeanS | SDS | VS | SS | SSE | KS | KSE |
|---|---|---|---|---|---|---|---|---|---|---|
| T41 | 38 | 3 | 5 | 4.11 | .689 | .475 | − .139 | .383 | − .795 | .750 |
| T42 | 38 | 3 | 5 | 4.32 | .702 | .492 | − .533 | .383 | − .783 | .750 |
| s1 | 38 | 2 | 4 | 3.42 | .599 | .358 | − .477 | .383 | − .602 | .750 |
| s2 | 38 | 2 | 5 | 2.79 | .811 | .657 | .734 | .383 | − .094 | .750 |
| s3 | 38 | 2 | 5 | 3.84 | .718 | .515 | − .218 | .383 | .018 | .750 |
| s4 | 38 | 2 | 5 | 4.00 | .771 | .595 | − .373 | .383 | − .202 | .750 |
| s5 | 38 | 3 | 5 | 4.39 | .595 | .353 | − .384 | .383 | − .636 | .750 |
| s6 | 38 | 2 | 5 | 2.79 | 0.905 | .819 | .673 | .383 | − .863 | .750 |
| s7 | 38 | 3 | 5 | 4.24 | .542 | .294 | .138 | .383 | − .131 | .750 |
| s8 | 38 | 1 | 4 | 2.47 | .725 | .526 | .320 | .383 | − .076 | .750 |
| s9 | 38 | 2 | 5 | 4.05 | .804 | .646 | − .428 | .383 | − .421 | .750 |
| s10 | 38 | 2 | 5 | 3.68 | .962 | .925 | − .269 | .383 | − .785 | .750 |
| sc | 38 | 50.0 | 77.5 | 63.092 | 6.1357 | 37.647 | .440 | .383 | .373 | .750 |

## 8 Conclusion

In this paper, we propose a novel visualization analytical system for metro network, including three modules: anomaly detection module, anomaly verification module and anomaly cause exploration module. In the anomaly detection module, we apply the improved RPCA model to detect the anomalies. In the anomaly verification module, we propose a 3D hierarchical view to verify the possible abnormal stations detected by the improved RPCA model and further discover the anomalies of the section passenger flow section between stations. In the anomaly cause exploration module, we can quickly locate the original Weibo text that may be related to the causes of the anomalies. Nevertheless, massive data increase the pressure of loading data in the 3D hierarchical view. We will find ways to optimize the system and improve efficiency in the future. Our system is designed as a real-time system, but we only use historical data in this paper due to some reasons not related to technology. In future work, the system only needs to obtain two authorizations of Weibo data and real-time traffic data to realize the function of real-time interaction.

**Acknowledgements** This work was supported in part by the National Natural Science Foundation of China under Grant U1811463, U19B200261, 61632006, 61672066, 61876012, in part by the Natural Science Foundation of Beijing under Grant 4172003, in part by the Beijing Municipal Science and Technology Project under Grant Z171100004417023.

## References

Alper B, Riche N, Ramos G, Czerwinski M (2011) Design study of linesets, a novel set visualization technique. IEEE Trans Vis Comput Graph 17(12):2259–2267

Andrienko G, Andrienko N (2017) Visual analytics of mobility and transportation: state of the art and further research directions. IEEE Trans Intell Transp Syst 18(8):2232–2249

Andrienko G, Andrienko N, Fuchs G, Wood J (2017) Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. IEEE Trans Vis Comput Graph 23(9):2120–2136

Andrienko G, Andrienko N (2008) Spatio-temporal aggregation for visual analytics of movements. In: IEEE symposium on visual analytics science and technology, pp 51–58

Byron L, Wattenberg M (2008) Stacked graphs-geometry and aesthetics. IEEE Trans Vis Comput Graph 14(6):1245–1252

Chen S, Yuan X, Wang Z, Guo C, Liang J, Wang Z, Zhang X, Zhang J (2016) Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data. IEEE Trans Vis Comput Graph 22(1):270–279

Cheng T, Tanaksaranond G, Brunsdon C, Haworth J (2013) Exploratory visualisation of congestion evolutions on urban transport networks. Transp Res Part C Emerg Technol 36:296–306

Collins C, Penn G, Carpendale S (2019) Bubble sets: revealing set relations with isocontours over existing visualizations. IEEE Trans Vis Comput Graph 15(6):1009–1016

Cheut D, Sheets DA, Zhao Y, Wu Y, Yang J, Zheng M, Chen G (2014) Visualizing hidden themes of taxi movement with semantic transformation. In: 2014 IEEE Pacific visualization symposium, pp 137–144

Dörk M, Carpendale S, Collins C, Williamson C (2008) Visgets: coordinated visualizations for web-based information exploration and discovery. IEEE Trans Vis Comput Graph 14(6):1205–1212

Fan X, Li C, Yuan X, Dong X, Liang J (2019) An interactive visual analytics approach for network anomaly detection through smart labeling. J Vis 22:955–971

Guo H, Wang Z, Yu B, Zhao H, Yuan X (2011) Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: IEEE Pacific visualization symposium, pp 163–170

Havre S, Hetzler B, Nowell L (2000) Themeriver: visualizing theme changes over time. In: IEEE symposium on information visualization, pp 115–123

Itoh M, Yokoyama D, Toyoda M, Tomita Y, Kawamura S, Kitsuregawa M (2016) Visual exploration of changes in passenger flows and tweets on mega-city metro network. IEEE Trans Big Data 2(1):85–99

Kalamaras I, Zamichos A, Salamanis A (2018) An interactive visual analytics platform for smart intelligent transportation systems management. IEEE Trans Intell Transp Syst 19(2):487–496

Krüger R, Thom D, Wörner M (2013) TrajectoryLenses: a set-based filtering and exploration technique for long-term trajectory data. In: Computer graphics forum, pp 451–460

Riveiro M, Falkman G (2011) The role of visualization and interaction in maritime anomaly detection. In: Conference on visualization and data analysis

Riveriro M, Lebram M, Elmer M (2017) Anomaly detection for road traffic: a visual analytics framework. IEEE Trans Intell Transp Syst 18(8):2260–2270

Slingsby A, Wood J (2010) Treemap cartography for showing spatial and temporal traffic patterns. J Maps 6(1):135–146

Tominski C, Schumann H, Andrienko G, Andrienko N (2002) Stacking-based visualization of trajectory attribute data. IEEE Trans Vis Comput Graph 18(12):2565–2574

Tominski C, Schulze-Wollgast P, Schumann H (2005) 3D information visualization for time dependent data on maps. In: IEEE symposium on information visualization, pp 175–181

Vassileva J, Gutwin C (2008) Indratmo, Exploring blog archives with interactive visualization. In: Proceedings of the working conference on advanced visual interfaces, pp 39–46

Wang ZX, Chong CS, Goh RSM, Zhou WQ, Peng D, Chin HC (2012) Visualization for anomaly detection and data management by leveraging network, sensor and GIS techniques. In: International conference on parallel and distributed systems—proceedings, pp 907–912

Wang Z, Lu M, Yuan X, Zhang J, Van De Wetering H (2013) Visual traffic jam analysis based on trajectory data. IEEE Trans Vis Comput Graph 19(12):2159–2169

Wang Z, Lu M, Yuan X, Zhang J, Wetering H (2013) Visual trafic jam analysis based on trajectory data. IEEE Trans Vis Comput Graph 19(12):2159–2168

Wang Z, Ye T, Lu M, Yuan X, Qu H, Yuan J (2014) Visual exploration of sparse traffic trajectory data. IEEE Trans Vis Comput Graph 20(12):1813–1822

Wang XH, Zhang Y, Liu H, Wang Y (2018) An improved robust principal component analysis model for anomalies detection of subway passenger flow. J Adv Transp. https://doi.org/10.1155/2018/7191549

Yang B, Cao W, Tian C (2019) Online visual analysis of forest diseases. J Vis 22:197–213

Zeng W, Fu CW, Arisona SM, Erath A, Qu H (2014) Visualizing mobility of public transportation system. IEEE Trans Vis Comput Graph 20(12):1833–1842

Zeng W, W C, S F, Arisona M, Erath A, Qu H (2016) Visualizing waypoints-constrained origin-destination patterns for massive transportation data. Comput Graph Forum 35(8):95–107

Zhang W, Wang Y, Zeng Q, Wang Y, Chen G, Niu T, Tu C, Chen Y (2019a) Visual analysis of haze evolution and correlation in Beijing. J Vis 22:161–176

Zhang C, Wei H, Bi C, Liu Z (2019b) Helmholtz–Hodge decomposition-based 2D and 3D ocean surface current visualization for mesoscale eddy detection. J Vis 22:231–243

Zhou Z, Meerkamp P, Volinsky C (2016) Quantifying urban traffic anomalie