Introduction of ChinaVis Data Challenge 2019

1. ChinaVis and ChinaVis Data Challenge

China Visualization and Visual Analytics Conference (ChinaVis), co-sponsored by scholars and researchers in the visualization and visual analytics field in China, aiming to improve the communication in the visualization and visual analytics communities in China and surrounding regions, discuss the directions and opportunities of this field in the era of big data, promote the development and progress of related research and applications, and facilitate talent training and communications. ChinaVis builds a communication platform for renowned experts, entrepreneurs, and application departments domestically and internationally to further discuss frontier technologies and applications of visualization and visual analytics, so that to form a new ecosystem of industry, academia, research and application. ChinaVis has been held successfully in Beijing (2014), Tianjin (2015), Changsha (2016), Qingdao (2017), Shanghai (2018), Chengdu (2019), and Xi'an (2020), where domestic and international researchers and experts were in attendance.

The ChinaVis Data Challenge, first took place in 2015, has been successfully held for six years. As a key part of ChinaVis conference, the data challenge provides a series of attractive scenarios, detailed data, and specific problems inviting researchers, developers as well as amateurs to employ their visual analysis technologies to fulfill data analysis tasks. Furthermore, the data challenge is primarily a platform for participants to provide self-training and communication with each other, and helping them assess to evaluate the effectiveness of their systems and tools. The rapidly growing Data Challenge over the past six years has made it the leading domestic contest in the field on visualization and visual analytics.

2. ChinaVis Data Challenge 2019

The ChinaVis Data Challenge 2019 offers two tracks. Track one focuses on indoor movement trajectory visual analytics, where the ICMTD-2019 dataset is used by participants to accomplish the following four tasks: (1) inferring the schedule of the conference; (2) analyzing the conference personnel types in the venue and summarize the movement patterns of each type; (3) identifying at least 5 abnormal events; (4) summarizing the deficiencies in conference organization and management.

Track one, was published with data and tasks on March 22, 2019, and the submission deadline was June 10, 2019. A total 75 entries were received from 359 contestants, and all entries went through expert reviews and evaluation. Scores and comments were released for each entry, with an award rate of 33%. All contestants were invited to the Data Challenge Seminar of ChinaVis 2019 on July 24, 2019, that included an overall review of the data challenge, a release of the ground truth, invited presentations, winners' presentations, an award ceremony, and a poster exhibition.

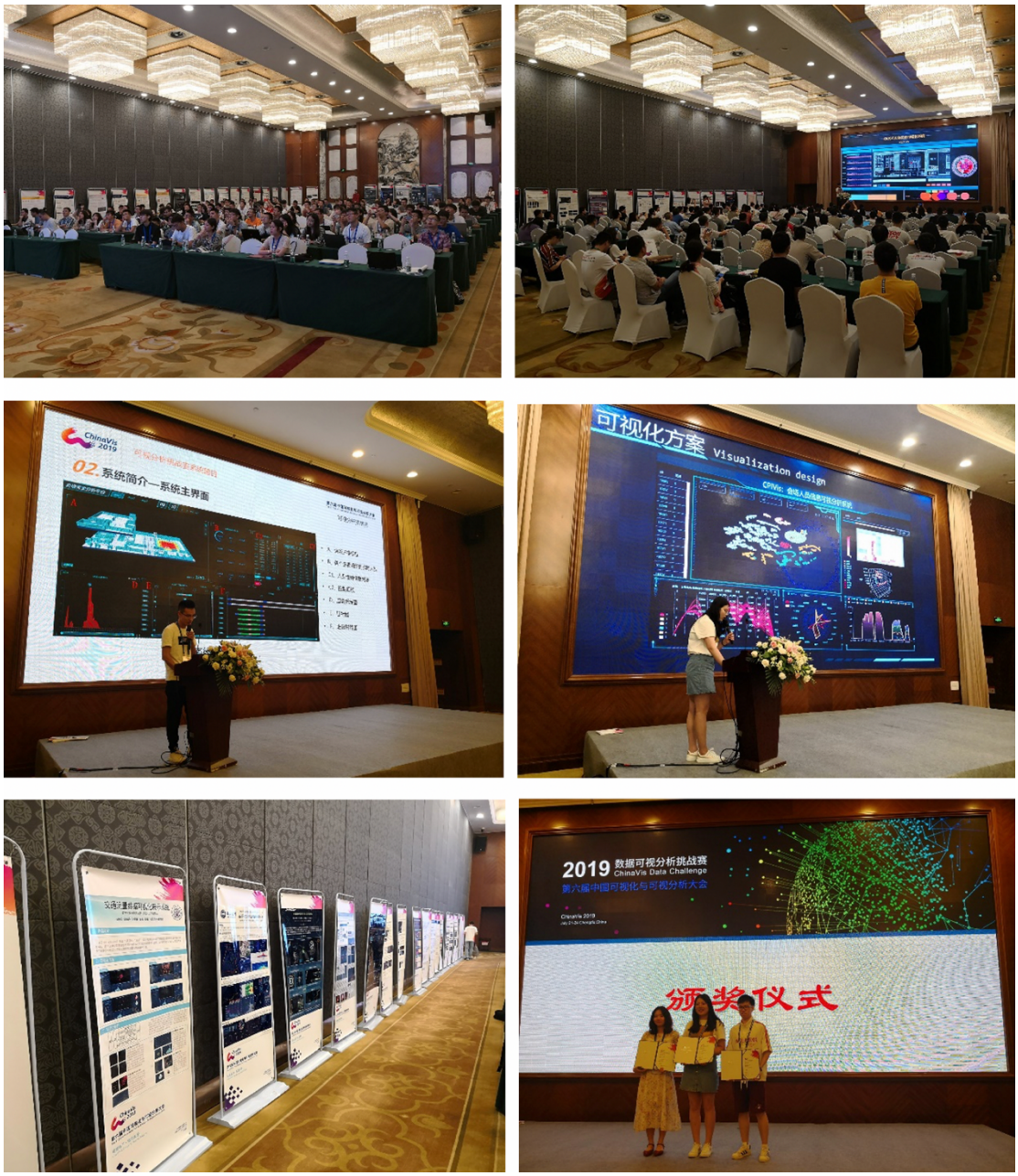

Figure 1. Photos from the seminar of ChinaVis Data Challenge 2019.

3. Excellent Entries

Six excellent entries are briefly introduced as follows.

3.1 Entry 1

Figure 2 shows the entry of Li Zhechuan's team from Chengdu University of Technology. The entry combines rich 2D and 3D information visualization forms and adopts a set of statistical analysis functions. With this work, the accuracy rate of completing data analytics tasks was the highest among all the entries, thus this entry won the 1st Prize.

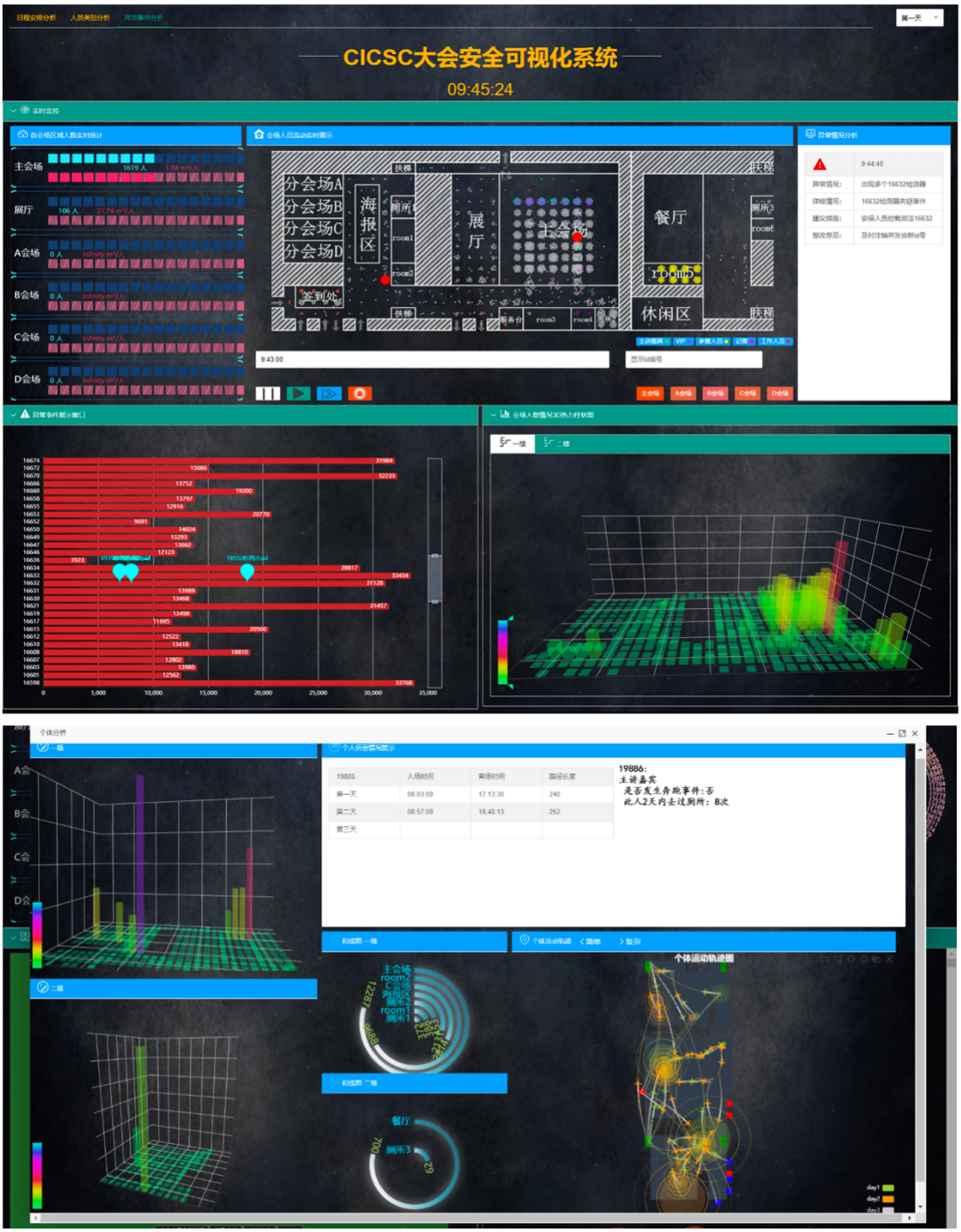

Figure 2. Entry of Li Zhechuan's team from Chengdu University of Technology.

As shown in Figure 3, the movement bubble chart in this work uses bubbles with different colors and sizes to encode areas of the venue and durations of stay respectively. In the bubble chart, X-axis represents time and Y-axis represents person-ID. Figure 3 shows a bubble chart of the movements of VIP guests, with the red bubble representing the main venue and sub venues, the black bubble representing the dining room, and the purple bubble representing the VIP Lounge. As we can see from the bubble chart, VIP guests usually attend the conference in the morning, enter the restaurant for a meal after the conference, and then return to the VIP lounge, and enter the venue at the start of the afternoon conference, finally leave the venue at the end of the conference.

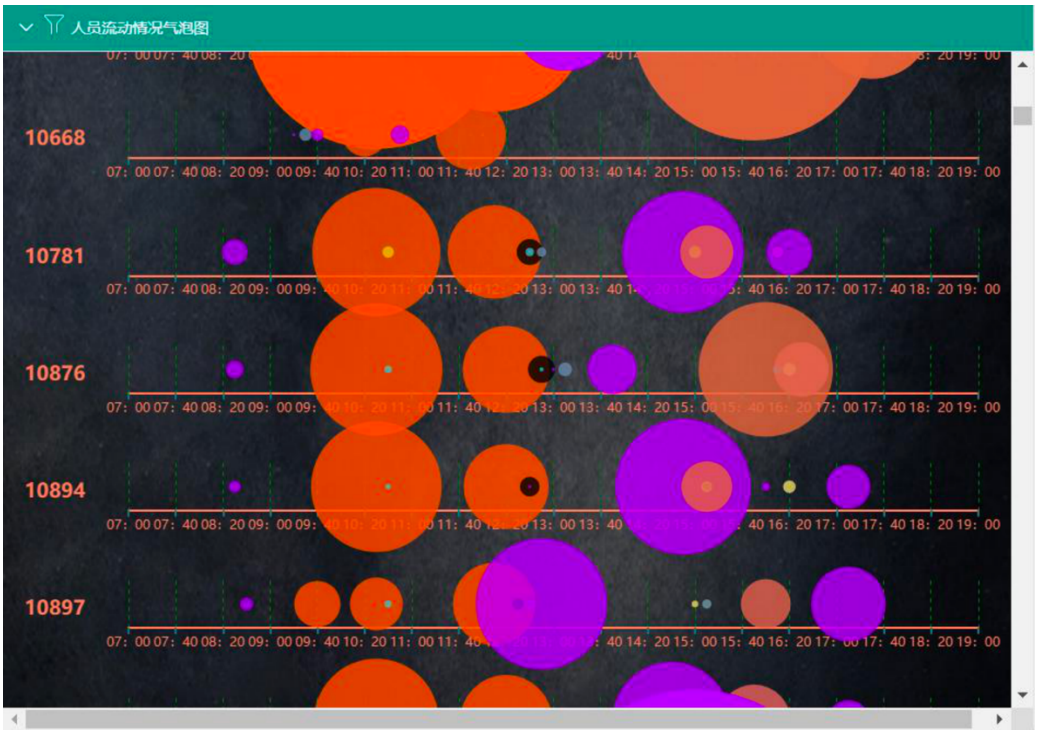

Figure 3. Bubble chart of the movements of VIP guests.

In addition, this work uses 3D bar charts to represent the density of people in various areas of the venue, providing a good monitoring of the distribution of people in the venue. As shown in Figure 4, the upper and lower bar charts represent the distributions of people on the 1st floor at the start of the conference and at the end of the conference respectively. As seen in the graph, the columns in the entrance and exit areas of the venue are significantly higher than those in the other areas at that time.

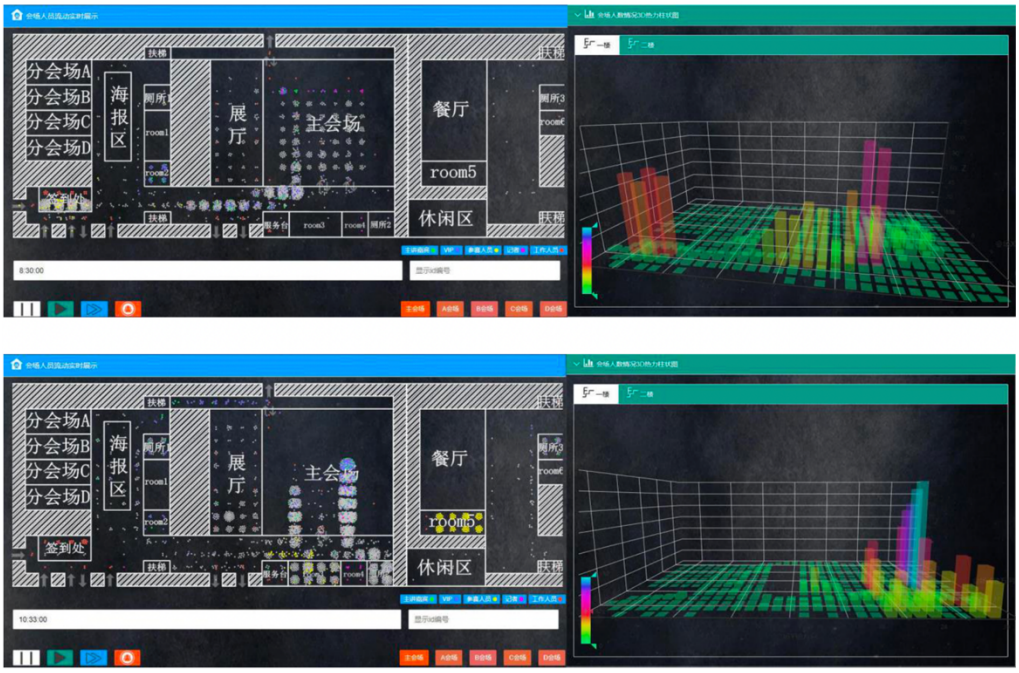

Figure 4. 3D bar charts of people density in the venue.

3.2 Entry 2

Figure 5 shows the entry of Liu Li's team from Chongqing University. Their entry exhibits powerful analytical functionals, outstanding aesthetics, and a set of well-designed interactions through multi-perspective visual analysis. As a result, the entry acquired a great accuracy in fulfilling data analytics tasks and received high scores from reviewers, thus won the 1st Prize.

Figure 5. Entry of Liu Li's team from Chongqing University.

This work analyzes patterns of crowd movement behaviors by combining semantic trajectory visualization and spatio-temporal trajectory visualization. Figure 6 shows a semantic trajectory visualization, where the X-axis represents time, the Y-axis represents person-ID, the color of the bar indicates the different areas, and the length of the bar reflects the time spent in the area. This visualization enables us a semantic analysis of the movement patterns within the venue and makes an easy comparison of the movement patterns of different types of people.

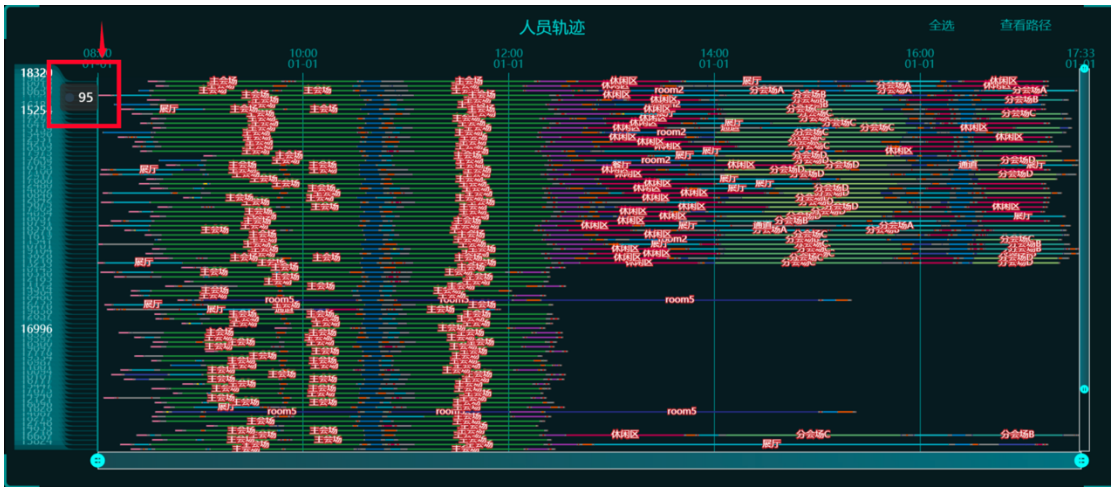

Figure 6. Semantic trajectory visualization of movement patterns.

Figure 7 shows the 3D spatial-temporal trajectory visualization of this work. Plotting the trajectory to a two-story 3D map allows the inspection of the - movement process of multiple or specific people. This figure shows the trajectory of media reporter-13612 interviewing hacking contestant-10409 in the corridors of the hacking contest room on the second floor.

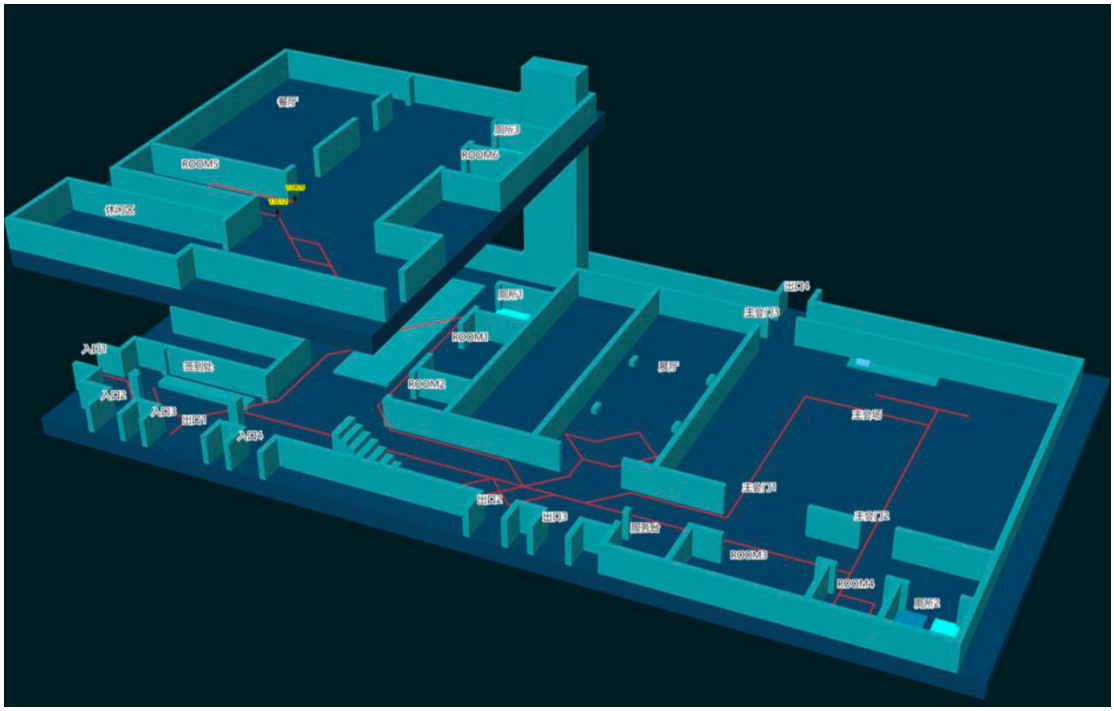

Figure 7. 3D spatial-temporal trajectory visualization of person 13612.

3.3 Entry 3

Figure 8 shows the entry of Li Wentao's team from Northeast Normal University. This entry introduces a series of automatic data analysis techniques, such as the LDA document theme generation model, K-means clustering algorithm, and t-SNE projection algorithm. This is an excellent work combining automatic analysis with interactive data analysis effectively, so it won the 2nd Prize.

Figure 8. Entry of Li Wentao's team from Northeast Normal University.

This work extracts the movement features of people and uses t-SNE to project people into a two-dimensional plane. As shown in Figure 9, a point represents a person and clustered points represent these people with similar movement patterns, thus making it easy to recognize

different types of people quickly. Further, users can interactively select a point or group of points in the projection map to continue exploring their movement patterns in other views.

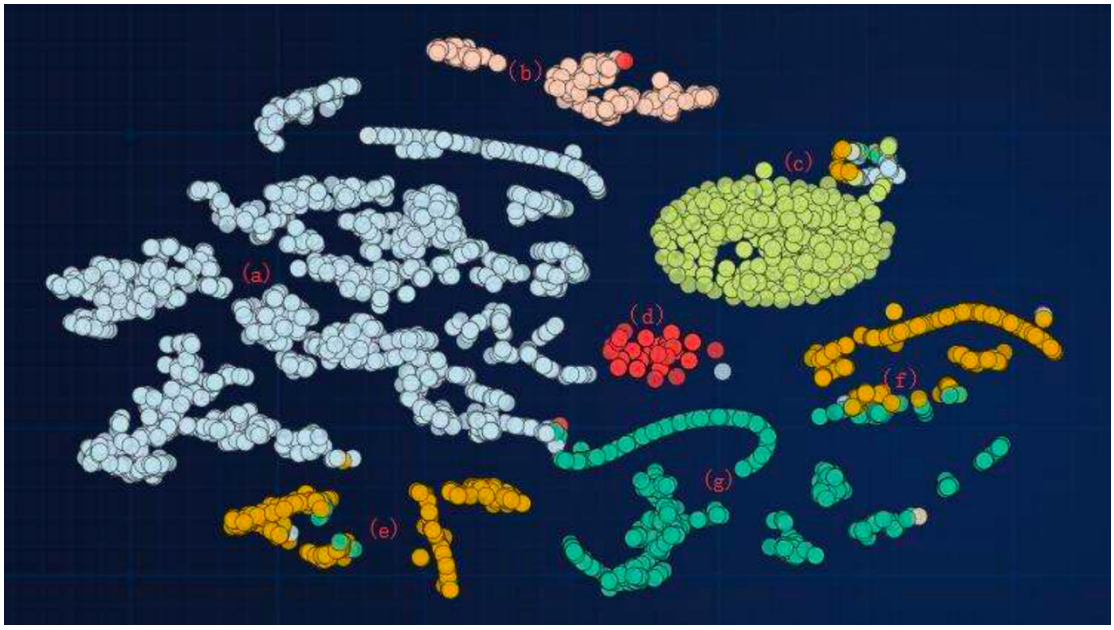

Figure 9. Two-dimensional projection map of the people.

Brushing the persons in group-(a) in the projection map and exploring the spatial characteristics of these persons in the radar map, Figure 10-a reveals that the persons spend the most time in the main venue, followed by the sub-venue and Figure 10-b shows that the participation time of these persons is generally consistent over the three days of the conference. Then, further analysis in the parallel coordinate chart in Figure 10-c shows that the departure times of these persons are obviously divided into two sub-groups, with one sub-group leaving at the end of the main venue conference and the other sub-group leaving at the end of sub-venue conferences. Figure 10-d and Figure 10-e show the spatio-temporal movement trajectories of a typical person in the two sub-groups respectively, where the X- and Y-axes represent the coordinates of the venue layout and the Z-axis encodes the length of time in seconds.

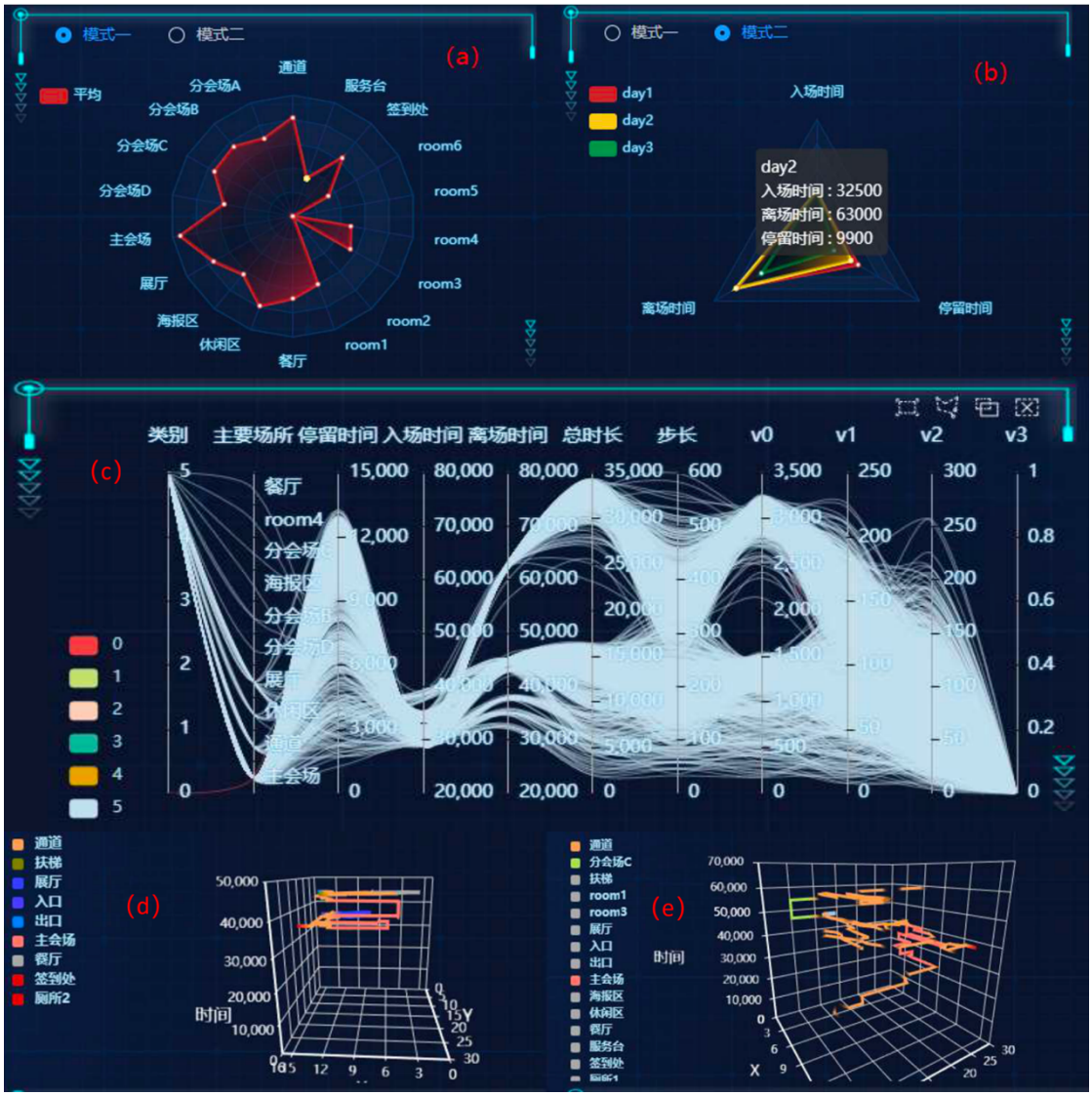

Figure10. Multi-dimensional patterns and trajectory analysis view.

3.4 Entry 4

Figure 11 shows the entry of Han Xiaoyang's team from the Computer Network Information Center of the Chinese Academy of Sciences. The fascinating highlights of this entry were the AR/VR data analysis environment constructed by using iPad and Hololens and the ability in supporting collaborative multiple people analysis, which won the Virtual Reality Application Award.

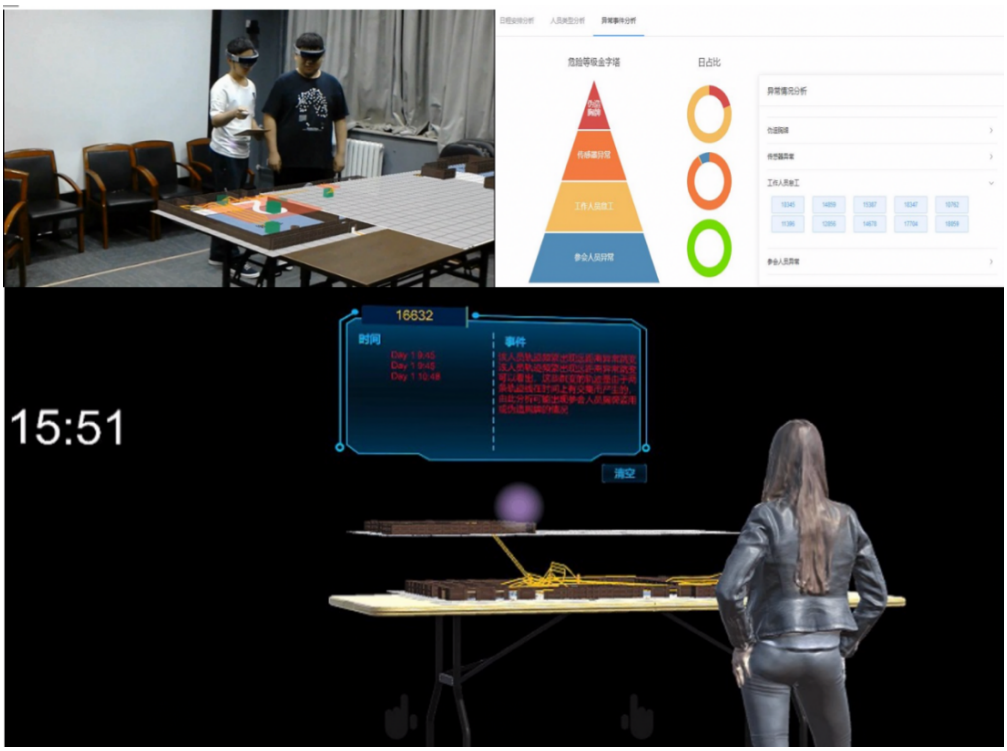

Figure 11. Entry of Han Xiaoyang's team from the Computer Network Information Center of the Chinese Academy of Sciences.

This work is highly immersive in 3D movement trajectory analysis due to the combination of VR/AR. As shown in Figure 12, the 3D trajectory map in the HoloLens terminal shows the typical trajectory of a VIP guest, with the numbers representing the order of passing and the green columns showing the time spent at that location. It is clear that the VIP guest stayed in the VIP lounge and the main conference hall for a longer period of time.

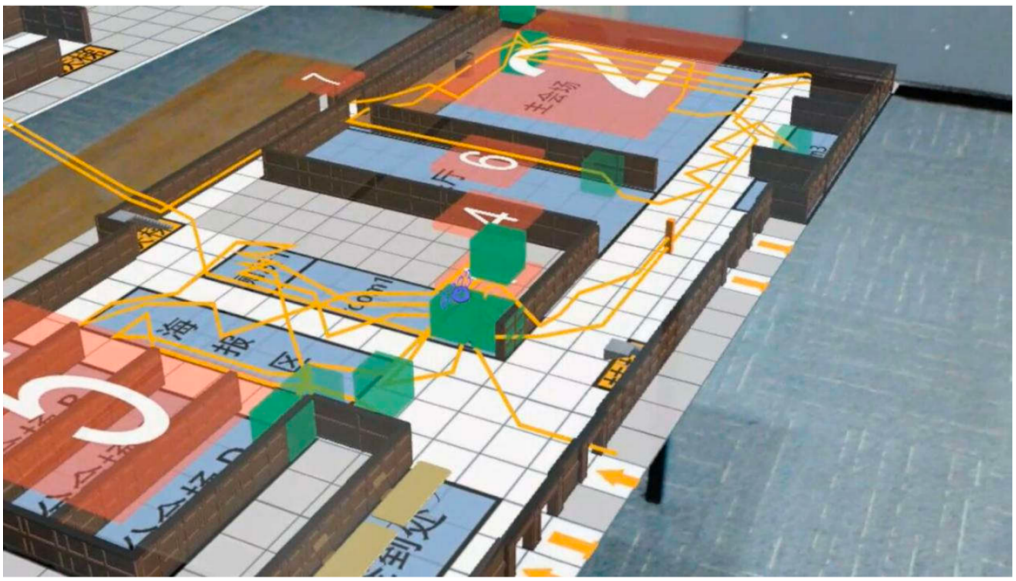

Figure 12. 3D trajectory map of HoloLens terminal.

3.5 Entry 5

Figure 13 shows the entry of Li Quan's team from WeBank. The entry displays an excellent visual analytics system with aesthetically pleasing interfaces and rich interactive features. The

reviewers expressed their approval of some of the novel visualizations in the entry, and as a result the entry won the Visual Analytics Design Award.

Figure 13. Entry of Li Quan's team from Webank.

The work is well designed with a path overview chart as shown in Figure 14-a, where the X-axis represents the venue area with different colors, Y-axis represents time, and an arc indicates the time span of a person staying in a certain area. Therefore, the longer a person stays in an area, the larger its arc. Figure 14-b shows the river chart in which different colors encode different areas, the X-axis represents the time and the river width represents the number of people. The graph shows that on the first morning, the main venue area coded by green has a high flow of people.

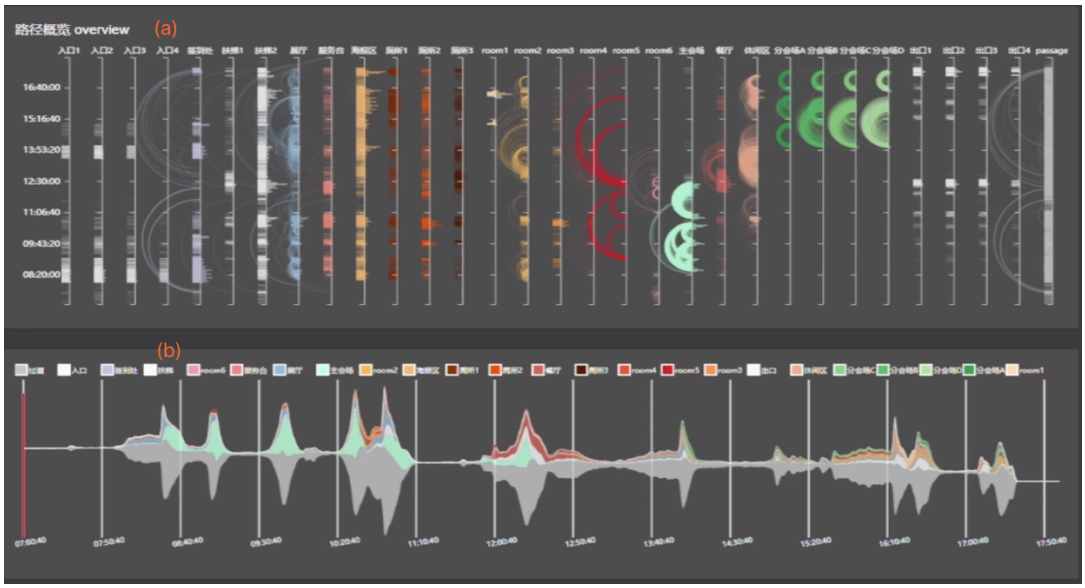

Figure 14. Path overview chart and river chart of people.