10-2017

Interactive Visual Analytics Application for Spatiotemporal Movement Data VAST Challenge 2017 Mini-Challenge 1: Award for Actionable and Detailed Analysis

Yifei GUAN
Singapore Management University, yfguan.2016@mitb.smu.edu.sg

Tin Seong KAM
Singapore Management University, tskam@smu.edu.sg

Interactive Visual Analytics Application for Spatiotemporal Movement Data

VAST Challenge 2017 Mini-Challenge 1: Award for Actionable and Detailed Analysis

Guan Yifei*, Kam Tin Seong

Singapore Management University

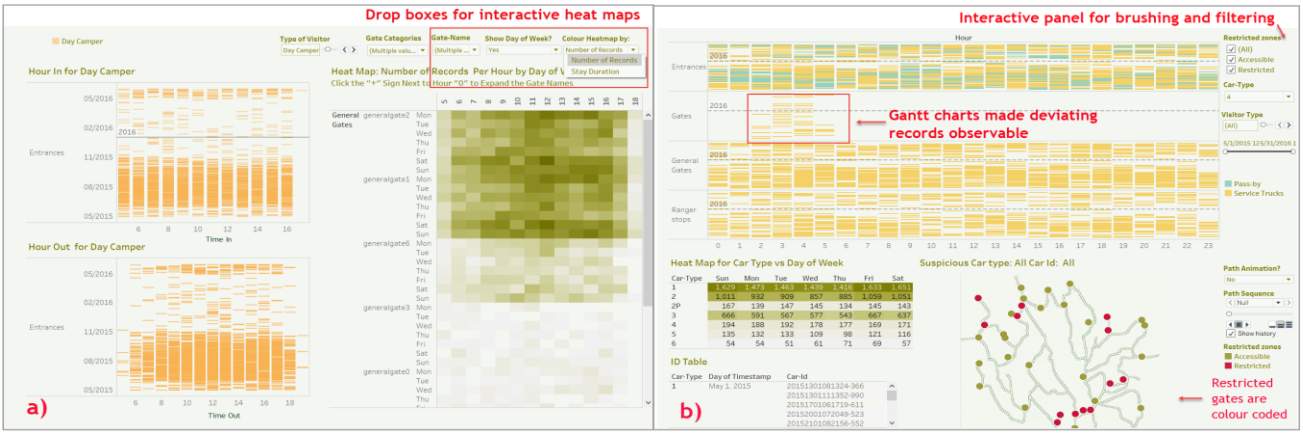

Fig 1 a) Dashboard for pattern of life analysis. b) Dashboard for anomaly detection

ABSTRACT

The Visual Analytics Science and Technology (VAST) Challenge 2017 Mini-Challenge 1 dataset mirrored the challenging scenarios in analysing large spatiotemporal movement tracking datasets. The datasets provided contains a 13-month movement data generated by five types of sensors, for six types of vehicles passing through the Boonsong Lekagul Nature Preserve. We present an application developed with the market leading visualisation software Tableau to provide an interactive visual analysis of the multi-dimensional spatiotemporal datasets. Our interactive application allows the user to perform an interactive analysis to observe movement patterns, study vehicle trajectories and identify movement anomalies while allowing them to customise the preferred visualisation configurations.

Keywords: Spatiotemporal visualisation, Movement Analysis, User interaction, Tableau

1 INTRODUCTION

This paper is in supplement to our submission for the 2017 VAST Challenge (MC1). The dataset provided in this challenge contains 170k movement records in the Boonsong Lekagul Nature Preserve. The challenge involved analysing the movement of traffic through the Preserve to validate the hypothesis that there is some link between the traffic going through the preserve and the decline in the nesting of Rose-crested Blue Pipit.

The challenge was made intriguing by the multidimensional interaction of the attributes in the datasets. There are 40 different gates in the Preserve, categorised into five types: entrances, general-gates, gates, ranger-stops and camping. The accessibility of each type of gate differs for different types of vehicles (all six types) passed through the Preserve. The following paper discusses our rationales and methodologies in handling such large-scale traffic movement data to gain actionable insights with visual analytics.

*yfguan.2016@mitb.smu.edu.sg
tskam@smu.edu.sg

2 METHODOLOGY AND TOOLS

We analysed the complex multidimensional datasets with commercial off-the-shelf software Tableau. [1] The adoption of an available software allowed us to focus more on visual analysis rather than application development. Aside from Tableau's renowned interactive visualisation capabilities, its other advanced functionalities like table calculations, level of details (LOD) calculations and parameters allowed us to easily slice and dice the datasets to gain deeper insights.

2.1 Data Preparation

We used Tableau's "Index" function to derive the order of movements for the vehicles based on the timestamps provided. Further application of Tableau's table calculation allowed the movement sequence calculation to restart from 1 for each unique Car-ID. We also tagged certain vehicles as "day campers" or "extended campers" based on whether the total number of days they stayed in the Preserve is greater than 1.

To make the anomaly analysis more expedient, we identified the "restricted gates" in the Preserve with reference made to the data descriptions and the Preserve map. This identification made the trespassing cars more prominent, as supposedly the restricted gates are only accessible by preserve rangers.

2.2 Segmenting the Visitors

Visitors entering the Preserve were divided into six segments including day campers, extended campers, rangers, service truckers, sightseeing visitors and passer-by based on their choice vehicles and their distinguish activities patterns. This made our patterns of life analysis more scalable as we would be able to analyse the homogeneity and heterogeneity of patterns within and between different segments of visitors.

2.3 Understanding Vehicle Trajectories

In our analysis, a *path* was defined as the series of gate categories visited by vehicles in a chronological order. In this case, a vehicle moves from entrance1 ->camp1->entrance3 will be taken as being travelled along the same *path* as a vehicle which moves from entrance2 ->camp1->entrance4. The path analysis simplified our analysis through data abstraction, where we hide detail of data while maintaining their essential characteristics. [3]
On the other hand, a *sequence* was defined as the series of individual gates visited by vehicles in the chorological order. Taking the same example given above, now we consider the two vehicles are travelling by different *sequences*.

3 Dashboards and Visualisations

We developed four thematic dashboards for a comprehensive storytelling by deploying Tableau dashboard's actions.

Patterns Analysis Dashboard – Refer to Fig 1a), the first section of this dashboard contains a Gantt chart to visualise the enter and exit time by different segments of visitors in one view. Gantt bars with higher saturation symbolise regular movements while rare records symbolise suspicious movements. [3] To have a conjoint analysis of the patterns of life in the Preserve, we also included an interactive heat map for this dashboard, which creates a data matrix to visualise the spatiotemporal interactions in the multi-dimensional datasets and symbolizes intersections of higher density with darker shades. We made the heat map configurable whereby the date interval of the heat map was made interchangeably based on user preferences. We also provided the option for the users to view an aggregated heat map, or a detailed one showing the day of the week at each row. Lastly, the user could choose to visualise the density of stay duration instead of the visitor traffic size on the heat map.

Vehicle Trajectory Analysis Dashboard – This dashboard was designed specifically to implement our methodology of differentiating the *path* and *sequence* of vehicles discussed in section 2.3. We first used colour coded sequential cubes to provide an overview of the various paths adopted by different segments of visitors, with most popular paths displayed on the top. Click on one path, the variations of sequences under it will be displayed, refer to Fig 2. The path-sequence analysis allowed us to draw interesting insights. For instance, unlike other visitors, we noticed the rangers tend to travel by long paths with fixed sequences, indicating they are working on routine shifts.

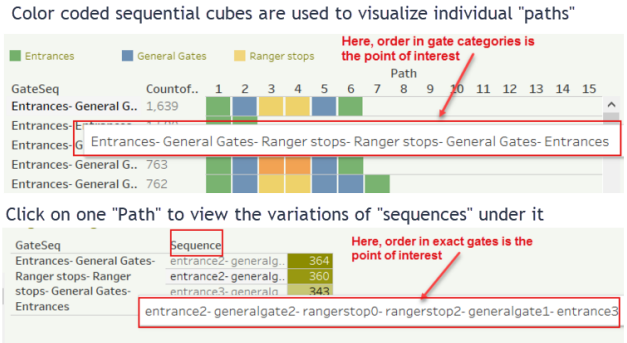

Fig 2 Differences between a path and a sequence

We supplemented the trajectory analysis dashboard by mapping a Preserve map to Tableau to visualise the vehicle trajectories. We allowed the users to make an option to view the entire path at one glance or to view an animated map tracing the movements. The latter would be a better choice if the path consists of repeated stops. We provided several visual encodings to enhance the visualisations where we denoted the gate categories with colours and configured the node sizes in proportion to traffic sizes. With simple "drag and drop" actions, user could further configure the visual encoding of the maps, as shown in Fig 3. The flexible mapping function is one salient feature of Tableau, which is highly aligned with our design concepts.

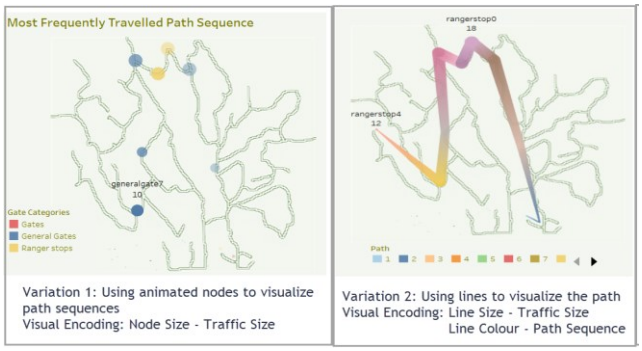

Figure 3 Two variations of mapping visual encoding

Anomaly Detection Dashboards – We have designed two dashboards for anomaly detection. The first one focuses on movement anomaly detection, refer to Fig 1.b). Gantt chart was used again for its effectiveness in the immediate exposure of deviating records. To enhance the effectiveness of anomaly detection, we added several filters to allow the users to focus on their interested records for investigation. For example, user could filter away preserve rangers and keep only the restricted gates to identify trespassing vehicles easily. A heat map showing the distribution of records by day of week and vehicle type, an ID-table showing the individual car-IDs and a map visualising the trajectory of a selecte car-ID are included in this dashboard to provide a comprehensive investigation, brushing on the Gantt chart will filter these complementary charts accordingly.
The second dashboard focuses on duration anomalies by deploying scatter plots to compare stay duration versus number of stops visited by each vehicle, which made outliers immediately prominent. The scatter plots were supplemented with the same heat map, ID-table and interactive map bundle described earlier also to deliver a complete investigation package.

4 Conclusion

Our interactive visual analytics application allowed us to discover significant patterns from the intriguing spatiotemporal datasets, and provided huge rooms for customisation in accordance to the data discovery needs of the users. Applying the anomaly detection features, we have successfully identified the most significant suspicious activities that would be detrimental to the natural habitat of blue pipits, such as illegal dumping. With the analysis on vehicle trajectories, we would like to raise some thoughts to ponder: the overly regular patrolling activity of the rangers might also pose a risk to the wildlife by allowing malicious actors to plan nefarious activities around their fixed schedules.

References

[1] Tableau software. Make your data make an impact. [Online] https://www.tableau.com

[2] Q. Cui, M.O. Ward, E. Rundensteiner, J. Yang, "Measuring Data Abstraction Quality in Multiresolution Visualizations", *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 5, pp. 709-716, Sept./Oct. 2006.

[3] Wood, J. (2016). Visual Analytic Design for Contextualising Sensor Data. Paper presented at the VIS 2016, 23-28 Oct 2016, Baltimore, USA