MC1: A Bespoke Analysis Tool for Spatio-temporal Park Traffic Data

Dimitar Kirilov* Isabel Lindmae* Andrew Burks* Chihua Ma* G.Elisabeta Marai*

Electronic Visualization Laboratory, University of Illinois at Chicago

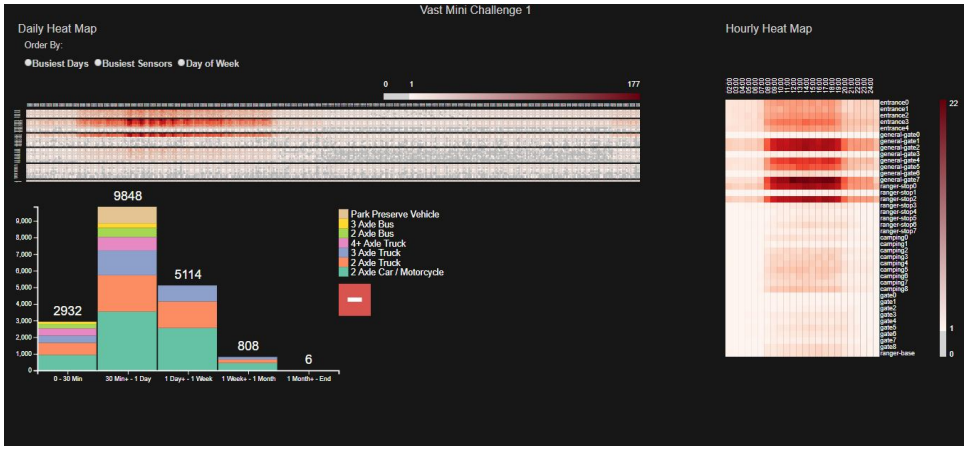

Figure 1: Traffic data analysis tool, with a large overview interactive heat map and a detailed heat map to the side; the maps capture spatio-temporal patterns A stacked histogram allows for temporal analysis; bins can be zoomed in for closer inspection.

Abstract

This paper describes a web-based traffic data analysis tool developed for the VAST 2017 Mini Challenge 1. The tool consists of two linked heat maps which allow for the inspection of daily activity for vehicles, as well as a histogram which allows for the analysis of total time spent by vehicles in the park. Combined, these views allow for the analysis of both spatial and temporal patterns in the park preserve.

1 Introduction

The overall goal for the VAST 2017 Mini Challenge 1 was to find repeating patterns of activities by individuals or groups, patterns that appear over longer periods of time, and patterns of unusual activity that changes from an established pattern, or are just difficult to explain. The provided dataset included a .CSV file of vehicle information, and a road map of the park preserve with labels of the sensor locations of where the vehicles were detected. The .CSV file contained information about each vehicle which included the time stamp, car ID, car type, and the gate name of where the vehicle was located in the park preserve.

2 Approach

For this challenge, we designed a web-based tool which consists of two linked heat maps which show the daily and hourly activities of vehicles, and a stacked histogram which shows the total time spent by a vehicle in the park preserve. The two views allow for filtering and ordering of the data which support the analysis of both spatial and temporal patterns in the park preserve. The tool was implemented using a set of JavaScript libraries which include: D3, Bootstrap, and JQuery.

*{dkiril4,ilindm2,aburks3,cma6,gmarai}@uic.edu

The daily heat map allows for the user to analyze general daily vehicle group activities in the park preserve. Each row in the heat map corresponds to a sensor location, and each column corresponds to a different day. The user can sort the heat map rows by busiest day, month, or busiest overall sensor, as well as filter the data by specific vehicle type. These operations allow the analysis of interesting patterns of different groups of vehicles. The hourly heat map supports a more detailed analysis of a particular day's results. By inspecting a particular column in the overview map, the hourly heat map shows that day's vehicle traffic distributed over the hours of the day. These two views allow the users to see what time of day certain activities happen in the park and where.

The second encoding we use is the stacked histogram in the bottom left of the tool which shows the total time vehicles have spent in the park. The histogram uses stacked bars to distinguish between the different types of vehicles and supports zooming when a bin is clicked to show a detailed view of that time period. The temporal bins are divided as follows: [Less than 30 Min], [Less than 1 Day], [Less than 1 Week], [Less than 1 Month], [Less than 1 Year]. The bins account for temporal subsets, so that, for example the Less than 1 Day bin does not include the same information that the Less Than 30 Min Bin has.

3 Patterns of Groups

The first question of the challenge required us to find and recognize repeating patterns of activities by individuals or groups.

The first pattern we discovered was about overnight campers. Our stacked histogram shows time spent in the park vs. number of vehicles. We can see that only 2 Axle Cars / Motorcycles, 2 Axle Trucks, and 3 Axle Trucks stay more than a day in the park preserve. This indicates that the people who stay overnight in the park are probably families. If large numbers of campers that come by bus stay overnight, then they most likely only move around the park on foot.

The second pattern we found was by looking at the hourly heat map (Fig. 2). From this figure, we can see that the most vehicle activity is during the day between 7:00 am and 4:00 pm, which

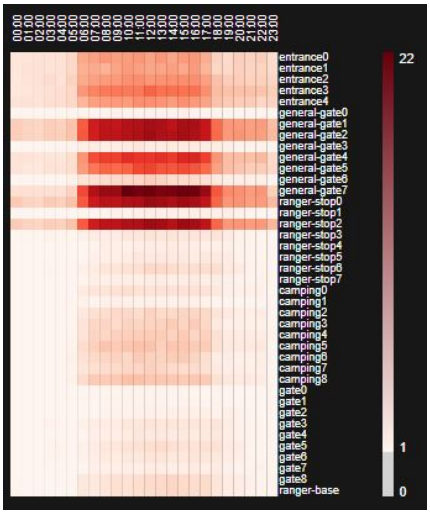

Figure 2: Heatmap showing general traffic data by the hour.

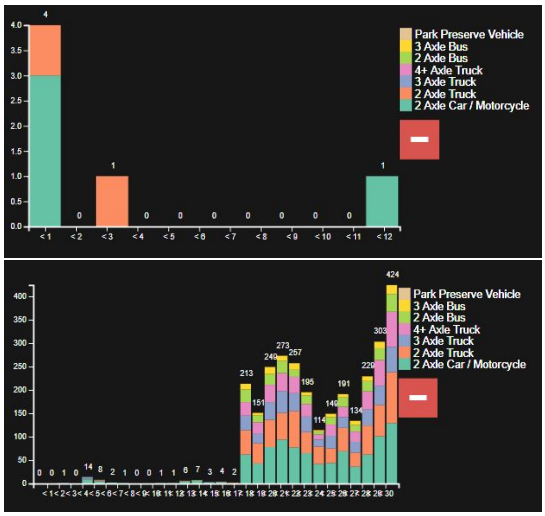

Figure 3: Monthly histogram temporal bins (top) and Detailed by-the-minute histogram temporal bins (bottom).

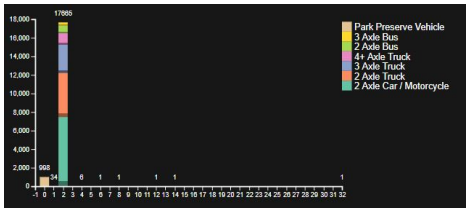

Figure 4: Filtered histogram, which graphs the number of vehicles (y axis) vs the number of entrances they have been detected at (x axis). Vehicles that have only been to 2 entrances over their whole route are highlighted.

makes sense since that's when the general public goes outdoors to the parks.

The third pattern we noticed became apparent by sorting the daily heat map by busiest sensors (clustered in groups). We can see that general gates and ranger stops receive the most activity. Also, it appears that ranger stops 0 and 2 receive significantly more activity than the rest of the sensors.

4 Patterns Over Multiple Days

The second question of the challenge asked us to find and recognize patterns that occur over multiple days and involve vehicles traveling through and within the park.

The first pattern we found for this question was about vehicles that stay over multiple days (perhaps for camping purposes) in the park. By looking at the stacked histogram it is evident that there are many people who stay overnight in the park, and upon closer inspection of the [Less than 1 Week] bin, we see that there is a gradual decline for the amount of days people spent in the park.

The next few patterns come from the daily heat map view and they are all closely related. First, by ordering the heat map by busiest month, we were able to see that the busiest month was July and that the busiest time the park has been is during the summer. By ordering the heat map by specific day of the week, we were also able to notice that Fridays, Saturdays, and Sundays are also the busiest days of the week which makes sense since that's usually when people are off work and can travel. On top of that, there is also increased vehicle activity when a holiday is near (e.g. 4th of July).

5 Unusual Patterns

The third question that we were asked for this challenge was to find unusual patterns, which may be patterns of activity that change from an established pattern, or are just difficult to explain from what we knew of the situation.

The first unusual pattern is one anomalous vehicle that stays in the park for a long time (Fig. 3). The general overview shows that there are 6 vehicles that stay for longer than a month and haven't left the park preserve. Upon clicking the bin, the detailed view shows that the vehicles that have stayed for more than a month include a 2 Axle Truck (nearly 2 months), and a 2 Axle Car / Motorcycle (nearly a year). The most notable outlier is a vehicle that has been in the preserve for a year and has id: "20155705025759-63". At the opposite end of the spectrum, upon inspecting the 0 - 30 min bin, we are able to see that there are around 25 vehicles that have stayed in the park for less than 7 minutes. It is possible these vehicles are using the park as a shortcut to major roads.

The third unusual pattern is shown via filtering the histogram (Fig. 4). Vehicles that have been detected at 2 entrances give us the first unusual pattern. This pattern means that people are cutting across the park to possibly avoid traffic as their whole route consists of entering the park and leaving it.

The same filtering shows the next unusual pattern. This time we highlight vehicles that have been detected at more than 2 entrances. This seems unusual, as normally people are assigned an id upon entering the park and then they are supposed to give it up when they exit the park preserve. From this view it is evident that there are 11 vehicles that don't do that, which indicates that they find a way to bypass this check and avoid paying the fee to re-enter the park.

An additional finding involves a recurring pattern of two 4 Axle Trucks that arrive at certain locations on Mondays and Wednesdays.

6 Conclusion

Overall, we were able to identify specific patterns, as well as unusual behavior regarding traffic in the park preserve involving both individuals and groups of people.

Acknowledgments

We thank the anonymous reviewers for their generous and supportive feedback. We thank everyone at the Electronic Visualization Laboratory for their support. This work was supported in part by NSF awards CNS-1456638 and IIS-0952720.