Probe Vehicle Based Trajectory Data Visualization and Applications

Anna Petrone, M.S. and Mark L. Franz, Ph.D.[*]

Center for Advanced Transportation Technology Laboratory
University of Maryland, College Park
*apetron2@umd.edu, *mfranz1@umd.edu*

Abstract

The proliferation of devices equipped with Global Positioning System (GPS) technologies such as smartphones and in-vehicle navigation systems has given rise to an exciting opportunity to better understand traveler behavior. Devices providing the timestamped location of travelers at frequent intervals provide details on trip origins and destinations (OD), departure times, route/mode choices, and travel times. In this paper, we utilize such a GPS data set to develop visualization and analysis methods that allows transportation professionals to better understand the nature of congestion on specific road segments. This novel data set was obtained from INRIX, which operates a vast network of traffic sensors, and includes all available trips in Maryland for the month of July 2015: a total of 4.8 million trips, each with potentially thousands of waypoints. A real-world application was generated to illustrate the benefits of the visualization and analysis methods in understanding vehicular travel patterns.

Keywords: Vehicle Trajectory Data, Origin-Destination Data, Travel Behavior Patterns

1. Background and Motivation

Understanding highway travel demand patterns is critical for transportation network analysis and planning. Traditionally, analysis of travel demand patterns relies on transportation surveys. While such surveys offer valuable information on general travel behaviors, the need for dynamic, continuous, high resolution traveler data has driven the development of cutting edge data collection methods based on Global Positioning System (GPS) information. The initial advancement along these lines was the utilization of vehicle probe data that uses GPS information to measure the travel times and speeds on pre-defined road segments called traffic message channels (TMCs). While traditional probe data has transformed how transportation professionals monitor and analyze the roadway network, it did not fully utilize the individual trip level spatio-temporal capabilities of GPS data. To meet this need, several transportation data providers and way-finding service providers have developed technologies to passively collect anonymous traveler data based on timestamped location information, which will be referred to in this paper as trajectory data. Upon processing trajectory data, information on trip origins and destinations (OD), departure times, route choices, and travel times can be extracted to understand travel behavior patterns and the related transportation network performance.

As will be described in the literature review section, there has been tremendous progress made on utilizing trajectory data for traffic network monitoring and analysis with cutting edge data visualizations. However, a majority of these applications have been

focused on using trajectory data to understand when and where congestion occurs. Recognizing the potential of trajectory data to reveal critical information about the origins and destinations of trips that contribute to traffic congestion suggests that there are several research paths for using this unique data that remain unexplored.

While there are a multitude of potential uses of trajectory data, this paper focuses on the utilization of trajectory data to assess the origins, destinations, and travel times of trips traversing a target highway section. These results can be used to better understand congestion patterns and to inform infrastructure improvement investments or congestion mitigation strategies such as incentivizing tele-work, flex hours and car-pooling.

To achieve this goal, this research was guided by the following tasks:

Task 1: Review Literature - Examine previous studies on the application and visualization of trajectory data

Task 2: Pre-Process Data - Snap raw trajectory data to the roadway network map and aggregate trip OD's to common transportation analysis geographies including counties, traffic analysis zones, census tracts and zip codes

Task 3: Develop Visualizations - Establish intuitive visualizations for the trajectory based analysis of a congested highway segment

Task 4: Demonstrate Research Contribution - Apply and discuss the analysis and visualization methods using real-world data and describe potential use cases.

Though this researches focuses on the analysis of vehicular trips, a similar task framework can be used to evaluate other modes such as transit, pedestrian, or bike. The above tasks form the structure of this paper, whose remainder is organized as follows. The next section presents the previous work related to trajectory data applications. In Section 3, a description of the data used in this study and the processing effort is presented. Section 4 presents a sample application. The final section summarizes the contributions of this research and defines future research paths.

2. Literature Review

The ubiquitous use of location sensing devices such as mobile phones and GPS devices has enabled researchers to collect and analyze valuable information about human mobility patterns. The topic of human mobility patterns has become a rigorously studied subject for its potential to address a multitude of topics including epidemic spreading, social events and interactions, and urban planning [1-3].

2.1. Trajectory Data Processing

Recognizing that raw trajectory data is recorded as timestamped latitudes and longitudes, most applications require data pre-processing to match the observations to the road network. On this front, the work of Brakatsoulas *et al.*, [4] describes two common map-matching challenges related to gaps in observations and un-reported changes to the transportation network map. To address these challenges, the study developed three map-matching algorithms including one incremental and two global algorithms based on computation efficiency and accuracy of mapping results. The study found that while the incremental method ran faster, the global matching methods provided more accurate mapping results. Next, Punzo *et al.*, [5] investigated the accuracy of vehicle trajectory data and the ability of various smoothing methods to correct for data inaccuracies for vehicle jerk analysis and platooning. The research determined that trajectory data accuracy varies across different traffic conditions and care should be taken in resolving data inconsistencies before an analysis is conducted. The work by G. Andrienko *et al.*, [6]

developed a clustering algorithm tailored for trajectory data analysis based on selection of prototype cluster objects, distance thresholds, and distance functions. The initial clustering result was then manually inspected using cutting edge data visualizations to allow for fine tuning and clustering adjustments based on user knowledge that is not represented in the trajectory data. Once pre-processed, the spatio-temporal nature of trajectory data lends itself to visual data analytics.

2.2. Trajectory Data Visualization

In the realm of transportation studies, two publications lead by the Andrienko team provide comprehensive discussions on studies conducted on the visualization of spatio-temporal data [7] and general movement data [8]. More specifically, N. Andrienko *et al.*, [9] created innovative visualizations for traffic congestion analysis. The related work by N. Andrienko *et al.*, [10] developed several vehicular movement visualizations for analyzing traffic intensity and speed on a spatially abstracted transportation network.

The visualization of vehicular trajectory data has revolutionized traffic network monitoring and planning analysis by providing high-resolution movement patterns of individual vehicles. To utilize this unique data, several studies have been conducted to develop visualizations to communicate network performance and roadway traveler behavior patterns. For example, the study by Tominski *et al.*, [11] developed three-dimensional visualizations of trajectory data by stacking time stamped vehicle positions on a traditional two-dimension network map. In a similar study, Guo et al. [12] introduced multiple visualizations for evaluating vehicular movements at intersections. Next, the work of Liu *et al.*, [13] built the Visual Analytics for Intelligent Transportation (VAIT) system. VAIT was demonstrated on a two month data set of taxi trajectories in China. The data was used to build an analysis tool to query data and visualize high volume taxi regions within an urban network. In a similar study, Ferreira *et al.*, [14] developed a data query and visualization interface for New York City taxi trips. The application of the interface was applied to several case studies including the impact of hurricane on taxi movement. The work of Wang *et al.*, [15] analyzed a non-traditional trajectory data set that was collected as vehicles passed by pre-determined sensors on the roadway network. The authors developed several visualizations including map views, plots of flow and speed versus time at a given location, and link/route performance visualizations. A visual based method for analyzing traffic jams was demonstrated in the related paper by Wang *et al.*, [16]. The visualizations assisted in traffic jam detection and analysis in terms of spatial and temporal evolution of congestion. A comprehensive review of traffic data visualizations, including trajectory data, can be found in Chen *et al.*, [17].

2.3. Trajectory Data Applications

Building on these vehicular trajectory data visualization techniques, several studies have utilized trajectory data for the specific purpose of monitoring, predicting or evaluating roadway performance. In the study by Kong *et al.*, [18], a curve fitting based method was compared to a vehicle tracking based method in terms of estimation accuracy of vehicle speeds and computational efficiency. The vehicle tracking method was more accurate but had longer computation time compared to the curve fitting method. The authors demonstrated that these traffic state estimation methods have the ability to be used for real-time network monitoring. Next, Herrera *et al.*, [19] developed and tested cellular phone GPS based traffic monitoring system called Mobile Century. The system was tested on an interstate in California. The proof of concept successfully demonstrated that the system was able to collect and feed data to an online interface for real-time traffic monitoring with just a 2-3 percent technology penetration rate in the traffic stream. The related research of Hiribarren and Herrera [20] investigated the use of trajectory data to estimate real-time traffic conditions on arterials based on the Lighthill-Whiteman-

Richards traffic flow model. Using ground truth from micro-simulation, the method was able to reasonably estimate and predict traffic states on arterials with a low technology penetration rate of just 1 percent. Lastly, the paper by Tang *et al.*, [21] presented a method to estimate time-space paths and dwell times. The work used unique three-dimensional time-space plots to conceptualize the proposed method. The method was tested for trip estimation and dwell time accuracy using real-world trajectory data from three U.S. cities. The pilot study showed that the method was able to assign the correct path and accurately estimate the link travel times with computation speeds sufficient for real-time applications.

In contrast to trajectory data applications for roadway network performance, few studies have rigorously investigated the use of trajectory data for understanding trip OD's and their impact on network performance. Here, the work lead by G. Andrienko [22] developed methods to visualize general OD patterns using diagram maps rather than flow maps. The methods were applied to two case studies that illustrated the ability of the proposed data visualizations to reduce dimensionality using spatial and temporal abstraction. Recognizing the limited penetration of trajectory data, the work of Yang *et al.*, [23] investigated two methods to estimate true OD matrices using trajectory data and link counts. The study showed that the distribution of trajectory based OD observations in the network affects the performance of the models.

This section has presented a summary of the studies related to processing, visualizing, and applying trajectory data for transportation analysis. While these studies have advanced the use of trajectory data in traffic analysis, most have been geared toward network monitoring and analysis. This paper builds on these studies by presenting a unique application of trajectory data that allows an analyst to conduct an in depth investigation of travel demand patterns on a target segment. Such results can be used for data driven traffic demand management decisions.

3. Data Processing and Visualization Methodology

3.1. Data Description

This novel data set was produced by INRIX, a transportation data service provider. The authors of this paper received the data from the MDSHA, which has a partnership with INRIX. Each trip is given a unique ID that is stripped of any personally identifiable information (PII). As the motorists travel to their destination, the INRIX routing technology updates and records their time-stamped position. The dataset includes all available trips in Maryland for the month of July 2015. Each trip is categorized as being taken by a light, medium, or heavy vehicle. For the purposes of this paper, we assume light vehicles correspond to personal trips, and ignore trips by medium or heavy vehicles. The raw format of this data includes trip ID's and timestamped latitude and longitude coordinates taken along each trip, representing a total of over 4.8 million trips and over 100 million waypoints. A preliminary analysis was conducted to determine the technology penetration rate of travelers equipped with INRIX navigation technologies (separate publication forthcoming). To do so, the number of trajectory observations were compared to volume counts for several Maryland corridors equipped with a permanent automated traffic recorder (ATR). This investigation found that this data-set represents approximately 1 percent of the total traffic stream in Maryland during July 2015.

3.2. Data Processing

Before the trajectory data can be analyzed and visualized, several steps must be taken to link the raw data to useful transportation network elements. These steps include (1) data mapping, (2) trip processing and (3) data querying. The framework of the data processing is presented in Figure 1 and is followed by a detailed description of each component in the procedure.
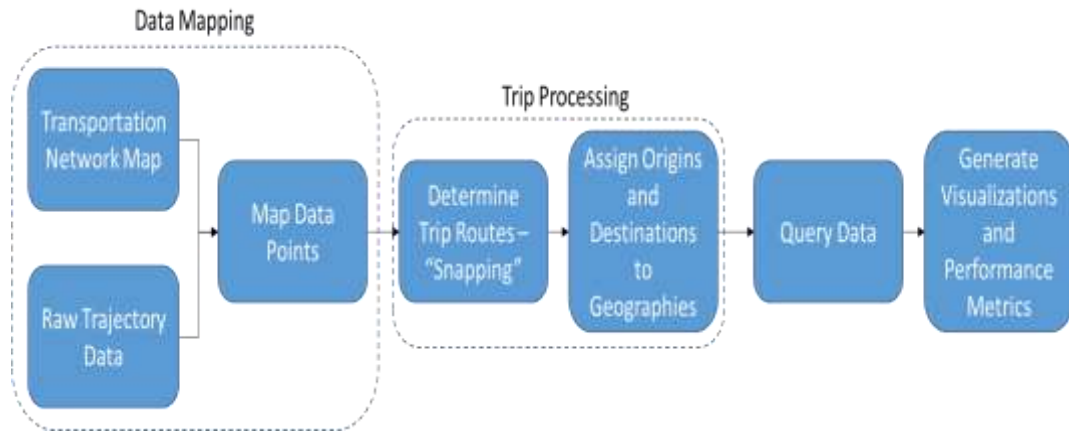
**Figure 1. Trajectory Data Analysis Procedure**

The first step is to match the raw data points to the road network. The term "matching" refers to the procedure of determining the most likely routes for each trip based on the possibly infrequent waypoints. Matching is the most challenging step within the data processing effort. The inherent nature of GPS-based waypoint data creates inconsistent time intervals between consecutive data recordings (*i.e.*, ping intervals). In addition, increased trip path uncertainty is introduced with longer ping intervals. The histogram of ping intervals in the July 2015 data-set is presented in Figure 2. Furthermore, GPS devices are subject to some error, so the recorded location may not lie directly on the road segment that was actually travelled. Map matching addresses these issues.
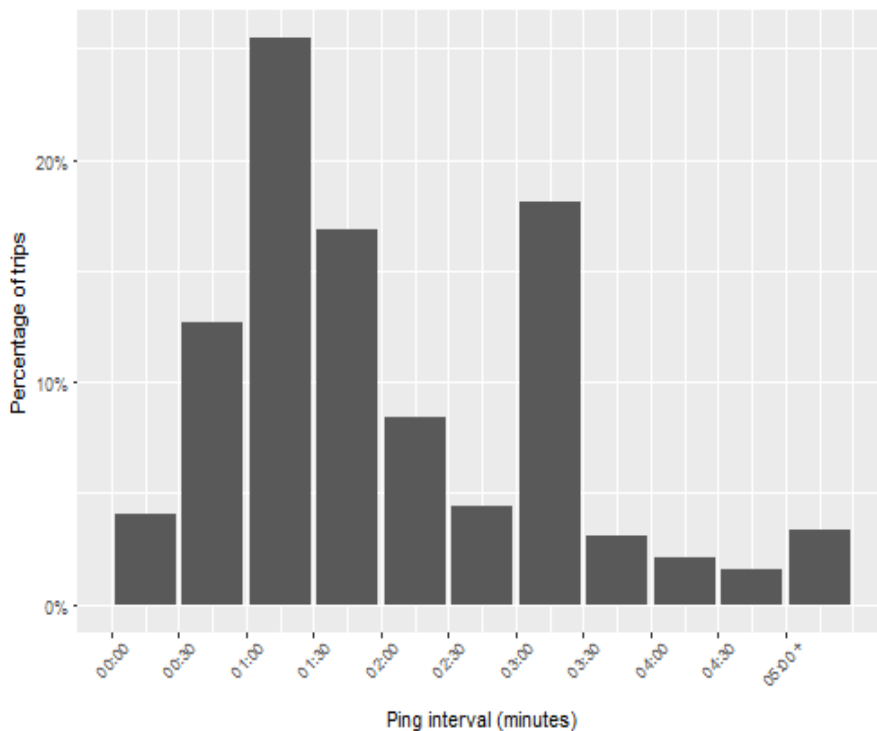
Figure 2. Trajectory Ping Interval Histogram

While the details of the map matching procedure are beyond the scope of this data visualization paper, a brief description is provided. The employed procedure used an application programming interface (API) created by Open Source Routing Machine

(OSRM) [24] that runs a hidden Markov chain model developed by Newson and Krumm [25]. The procedure calculates the probabilities of transitioning from one potential match road segment to another. The route that maximizes the overall transition probability is then selected as the travel path for that trip. This process is repeated for each trip in the dataset. An illustration of the result of the matching process is shown in Figure 3, where the original raw waypoints are shown as circles, and the corresponding matched road segments are highlighted underneath.
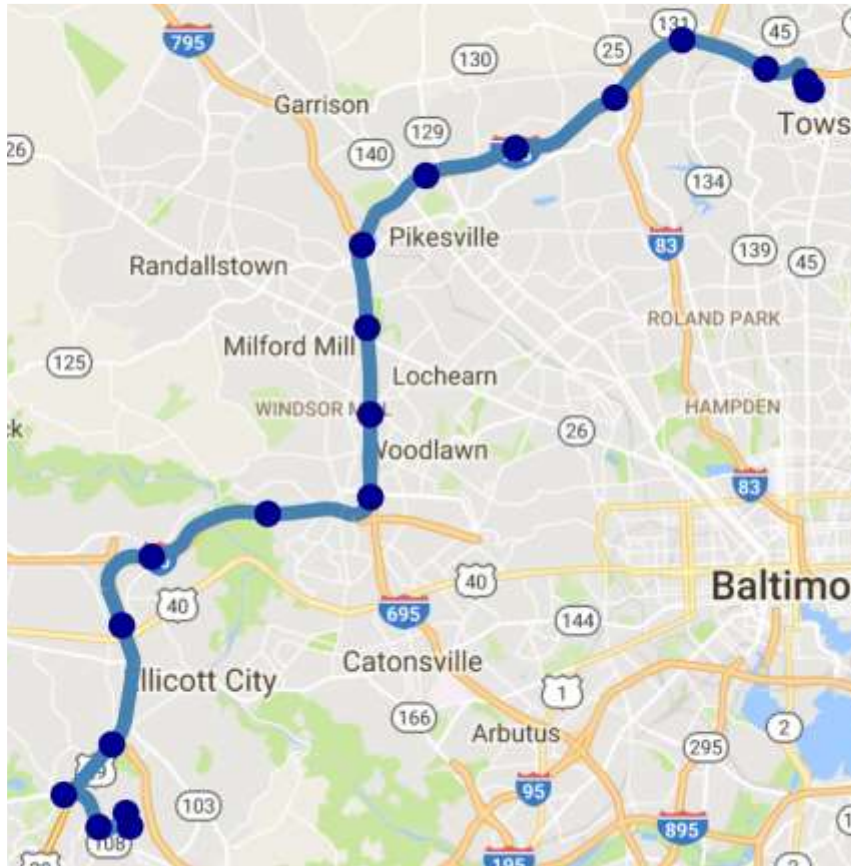
**Figure 3. Map Matching Raw Data to Create Trip Trajectories**

The OSRM service generates map matching results with respect to the open source roadway network OpenStreetMap (OSM) [26]. Within OSM, there are two types of relevant objects: (1) nodes and (2) ways. A node is a single point, and a way (henceforth road segment) is a sequence of nodes. OSRM's matching service returns a sequence of OSM node ID's for each trip. The node ID sequence is then used to retrieve the corresponding road segments for the trip. Next, it is necessary to estimate at what time the vehicles travelled each of the segments, since the time stamps from the raw data could be sparse. A simple space mean speed was used to interpolate the time of travel on any intermediate road segments.

The final step in trip processing is to assign the trip origin and destination locations to analysis geographies. Common geographies in transportation analysis include states, counties, sub-counties, zip code tabulation areas, census blocks, census tracts and traffic analysis zones (TAZ's). These pre-identified geometries were incorporated using basic spatial queries available to assign the origin and destination location of each trip to each of the analysis geographies. At this point the data was ready to be queried, visualized and analyzed.

4. Sample Application of Visualizations

To illustrate the potential use of the data visualizations, a sample application is presented. Here, the visualizations were used to better understand the trips contributing to congestion. These visualization include (1) origin density maps, (2) destination density maps, (3) OD Matrix, (4) route map, and (5) travel time histograms. As mentioned in the background section, the resulting analysis can be used to inform congestion mitigation strategies such as target demand management or infrastructure investments.

The target segment of southbound (SB) I-270 from the interchanges of I-370 to Montrose Road was selected for demonstration, as it was a top ranked site for congestion in the 2016 Maryland State Highway Mobility Report [27]. This 5.3 mile segment of SB I-270 suffers from recurrent congestion during AM peak period (6 – 9 AM) as commuters from suburbs such as Rockville, MD, and Frederick, MD travel to the employment magnets of Washington, D.C., Bethesda, MD and Northern Virginia. The analysis segment is shown as the black arrow in Figure 4 [28].
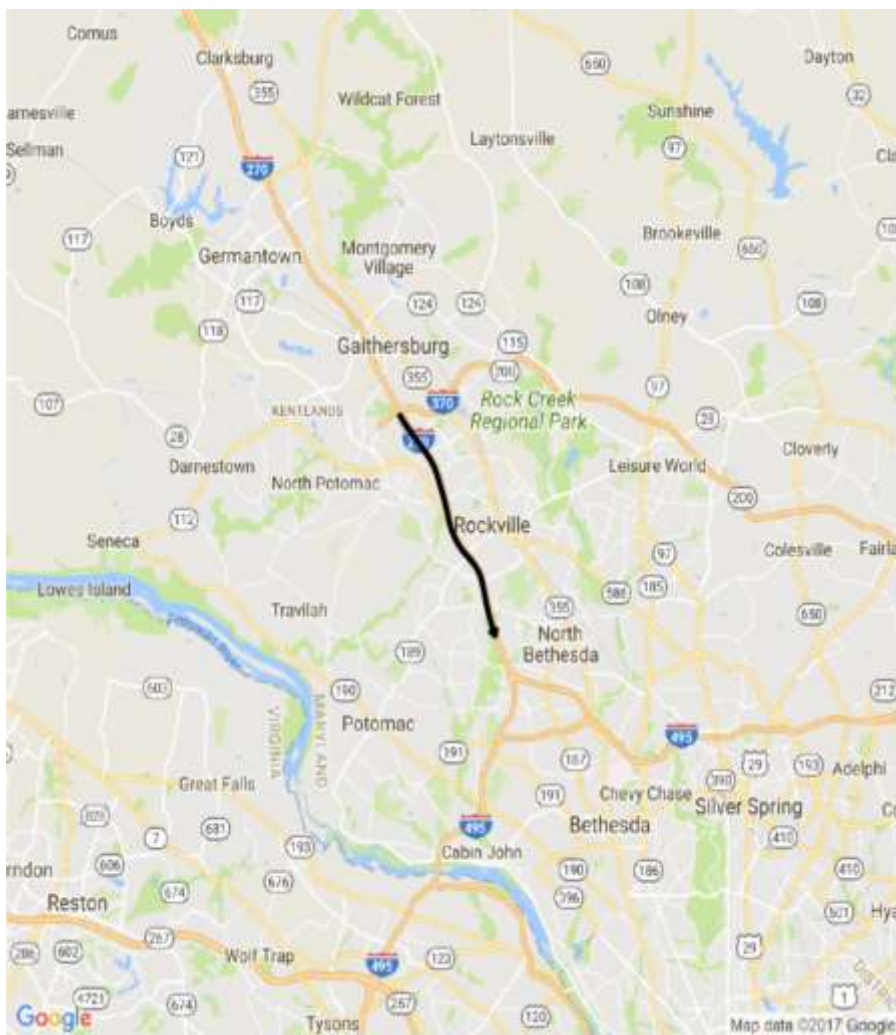
Figure 4. Map of Analysis Segment on SB I-270

The analysis procedure starts by querying the database for all trips taken by light vehicles that traversed the target segment on I-270 between 6 and 9 AM on weekdays in July of 2015. Next, the desired data aggregation level must be defined. For this example, locations are summarized at the sub-county level, as defined by the U.S. Census Bureau

[29]. Analysis on a finer geographic scale is theoretically possible, however caution must be taken with respect to the low sampling rate of the data. The remainder of this section presents the various visualizations and interpretations for assessing travel patterns that contribute to congestion.

4.1. OD Analysis

The first visualization is a sub-county level trip origin density map. Figure 5 shows the basic origin density map with the target segment represented as a black arrow and the sub-county levels colored in various tones of green to represent the number of trip origins. Here, lighter tones represent fewer trip origins while darker tones show higher numbers of trip origins. The density plot shows the main sub-counties that serve as origins for trips that traverse the highlighted segment of I-270. Labels can be overlaid on the map as shown in Figure 5, in which the label displays the number of trips (July total), the average time in which individuals departed from that location to make their trip, and standard deviation of these departure times. The data label in Figure 5 shows that 598 trips that traveled on the target segment originated in the Frederick sub-county within Frederick County, MD and their average departure time was 6:01 AM with a standard deviation of 53 minutes. Figure 5 can be used to identify locations (in this case, sub-counties) with relatively high numbers of origins, and to identify when those trips started. It is important to reiterate that the data used in this study represents approximately 1 percent of all trips in Maryland, so the absolute numbers of these trip origins should not be taken literally, but only as a way to determine larger patterns. However, if given a larger sample, the same visualization technique could be used to display the larger sample size.
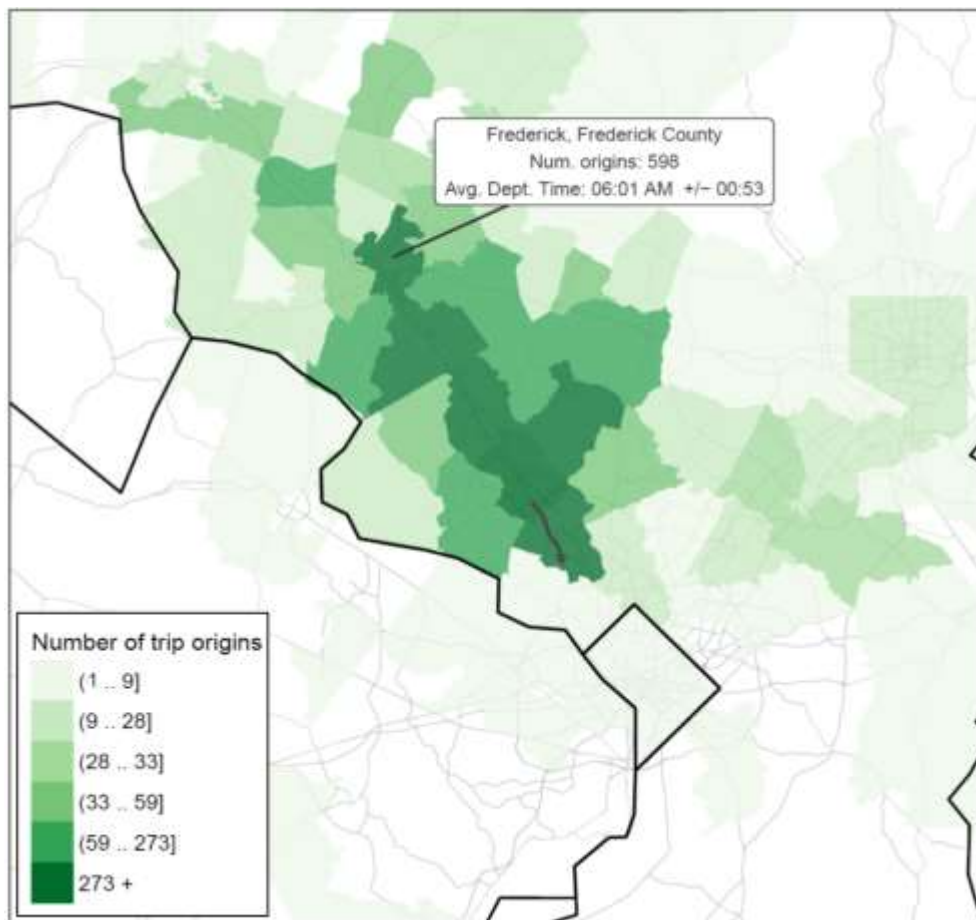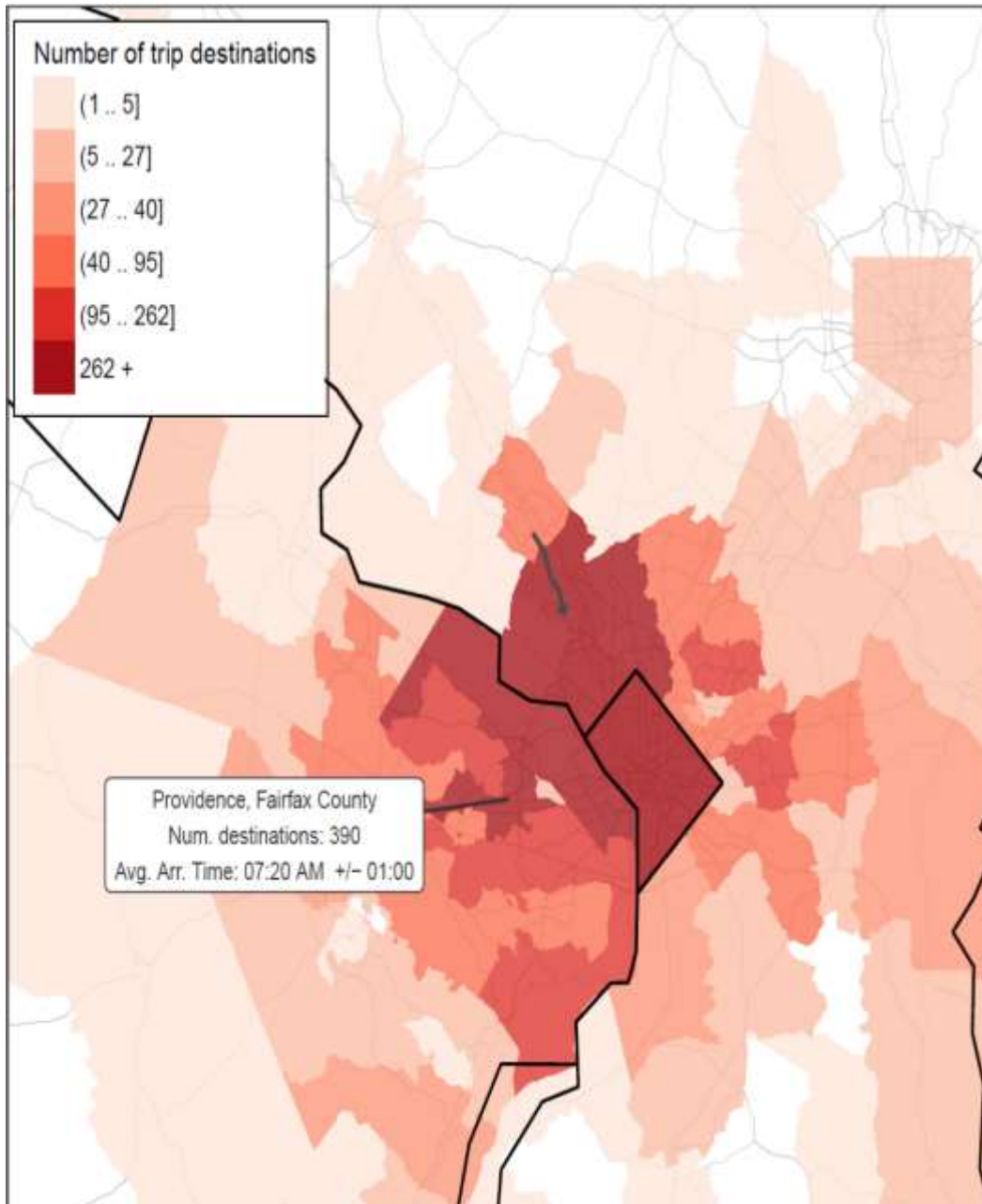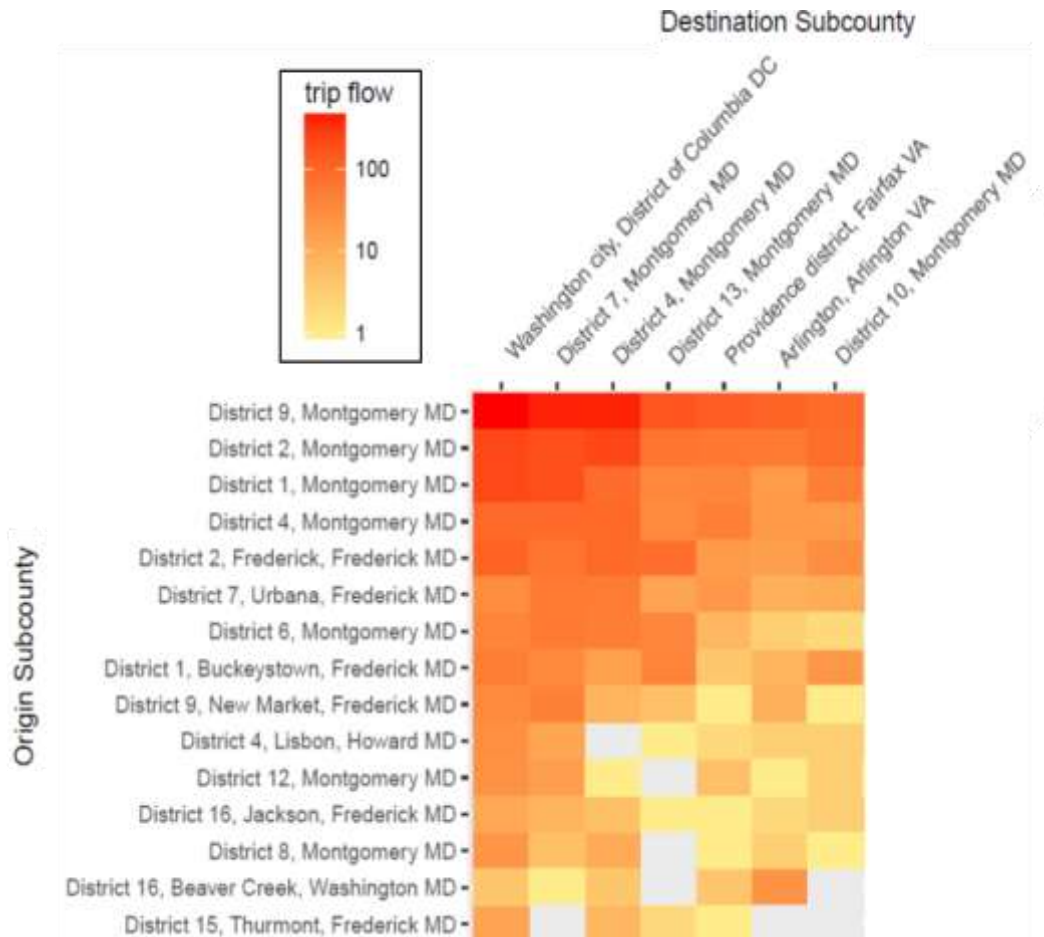

Figure 5. Sub-County Level Origins

Next, the sub-county level trip destination density map is presented in Figure 6. Similar to the origin maps, the destinations are represented by various tones of red, with lighter tones representing fewer trips with destination there, and darker tones symbolizing a greater number. As expected, a majority of the AM peak trips that traversed the target segment of SB I-270 were destined to sub-counties within Montgomery County, MD (the county containing the target segment) Washington, D.C., and Fairfax County, VA. For example, the Providence sub-county within Fairfax County, VA was the destination of 390 trips that had an average arrival time of 7:20 AM and an arrival time standard deviation of 1 hour.

Figure 6. Sub-County Level Destinations

Figure 7 shows a subset of the origin-destination matrix for trips that traveled on the target segment between 6 – 9 AM on weekdays. The full matrix would contain 215 rows

and columns, so for the sake of readability, only the top origin and destination locations are shown. Here, it is clear that the largest OD pairs were for trips that started within sub-counties in Montgomery County, MD and ended at sub-counties in Montgomery County, MD, Washington D.C., and Fairfax, VA. This observation is in line with transportation demand gravity models.


Figure 7. Partial Sub-County Level OD Matrix
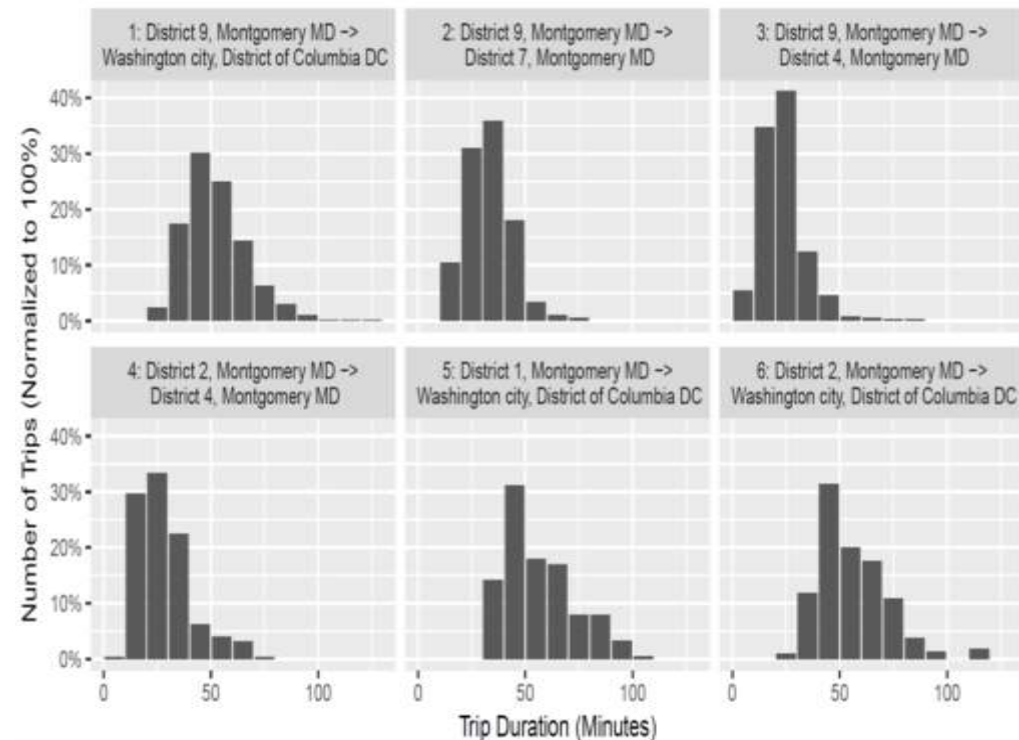
4.2. Travel Time Analysis

Supplementary to the OD observations, trajectory data allows for OD specific travel time analysis. Note that in traditional probe vehicle data analysis, travel times are obtained for each separate TMC segment, but are not linked to a specific trip. While such information is useful in identifying when and where congestion occurs, it lacks the details of trip travel times (the entire travel time that a specific traveler experiences). Trajectory data allows for the travel time measurement of entire routes between OD pairs. Figures 8 and 9 present histograms of travel times between the top six sub-county OD pairs in this sample application. Figure 8 shows the raw count trips on the y-axis and 10 minute travel time bins on the x-axis. Meanwhile, Figure 9 shows the percentage of trips within the OD pair on the y-axis and the same 10 minute travel time bins on the x-axis. These plots can be used to analyze the distribution of travel times, assess travel time reliability, and identify OD pairs with a significant number of outliers in terms of travel time.

As expected, the histograms for OD pairs that are separated by shorter distances tend to have the peaks of the distribution shifted towards shorter travel times. More interestingly, the distribution of travel times between the District 9 and District 2 sub-county in

Montgomery County, MD and Washington, D.C. show a more disperse distribution. This observation may be explained by the fact that a majority of the road network within the Washington, D.C. border consists of signalized arterials, causing a more variable travel time for trips terminating within city boundaries.
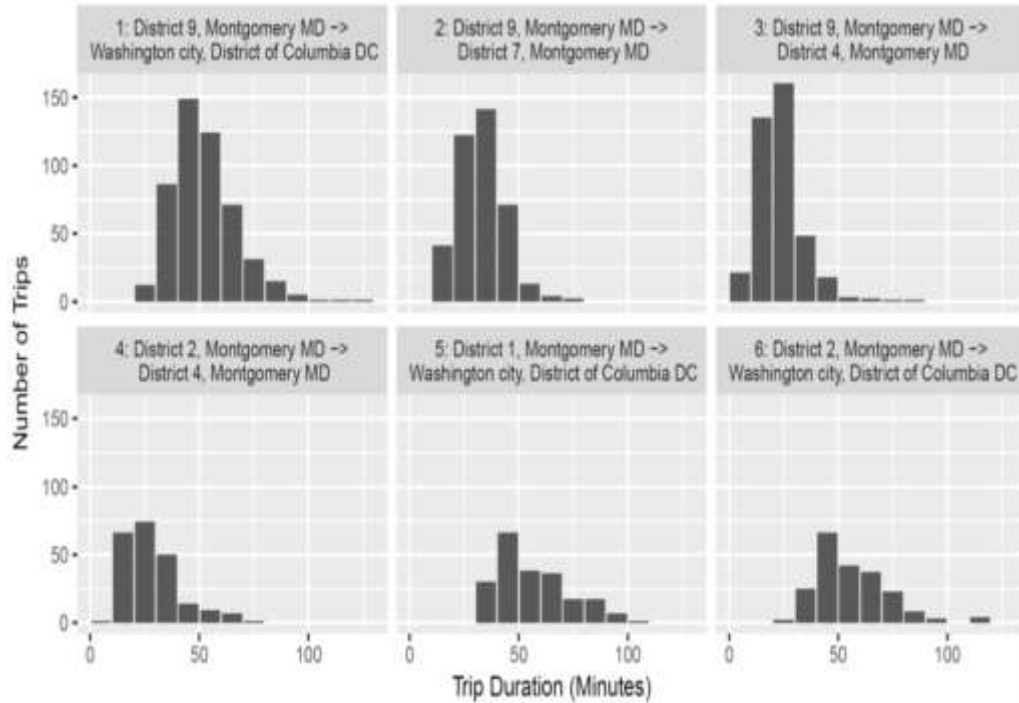

Figure 8. Travel Time Histogram-Raw Count

Figure 9. Travel Time Histogram-Normalized

4.3. Route Usage

Related to understanding OD travel times, trajectory data allows for the analysis of route usage. Figure 10 shows each of the routes that were used in the queried data-set. Thicker lines represent segments with more trips. Note that the ability to scale the line widths in this way is only possible thanks to the map matching process, which snaps the intermittent location pings to the road network, and computes the path taken between successive pings. As evident in Figure 10, many trips travel south on I-270 to gain access to the I-495 Capital Beltway, which is another top ranked congested segment in Maryland. Understanding such information with real-world data – as opposed to travel surveys, simulation, or engineering judgement – allows fewer biases and assumptions in transportation planning and demand analysis applications.
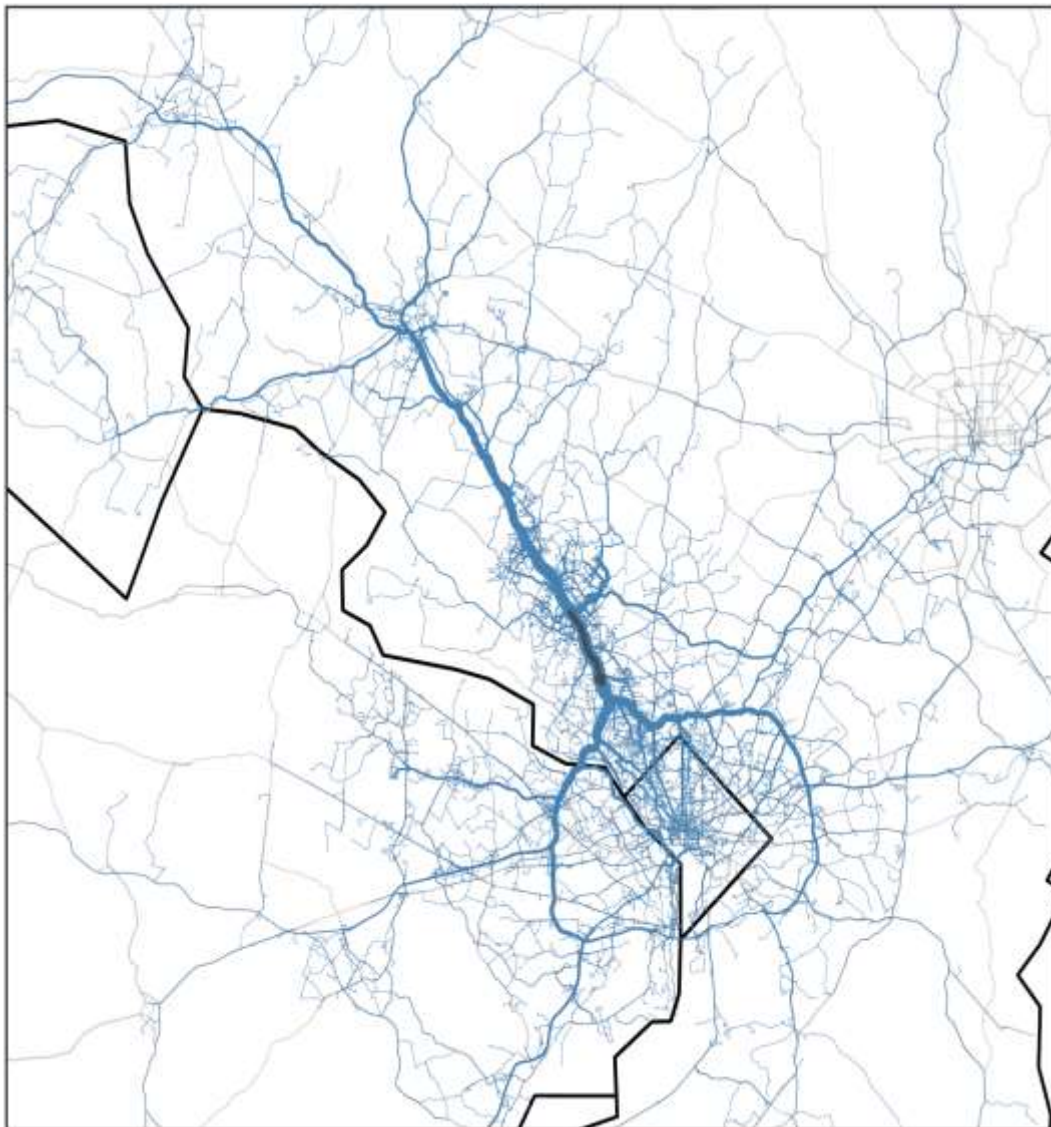

Figure 10. Trajectory Path Usage Map

4.4. Analysis of Temporal Demand Patterns

Another advantage of trajectory data is the ability to investigate temporal travel patterns. Using the same target segment of SB I-270, Figures 11 and 12 present the hourly evolution of sub-county origins and destinations from 6-9 AM, respectively. Such

visualizations may be used to understand how demand patters progress over time. To illustrate this point, consider the cluster of sub-counties to the north and west of the target segment in the hourly origin plots of Figure 11. During the 6 AM hour, many trips from this area traversed the target segment. However, as the AM peak period progressed, the number of trips departing from the north and west of the target segment steadily declined.
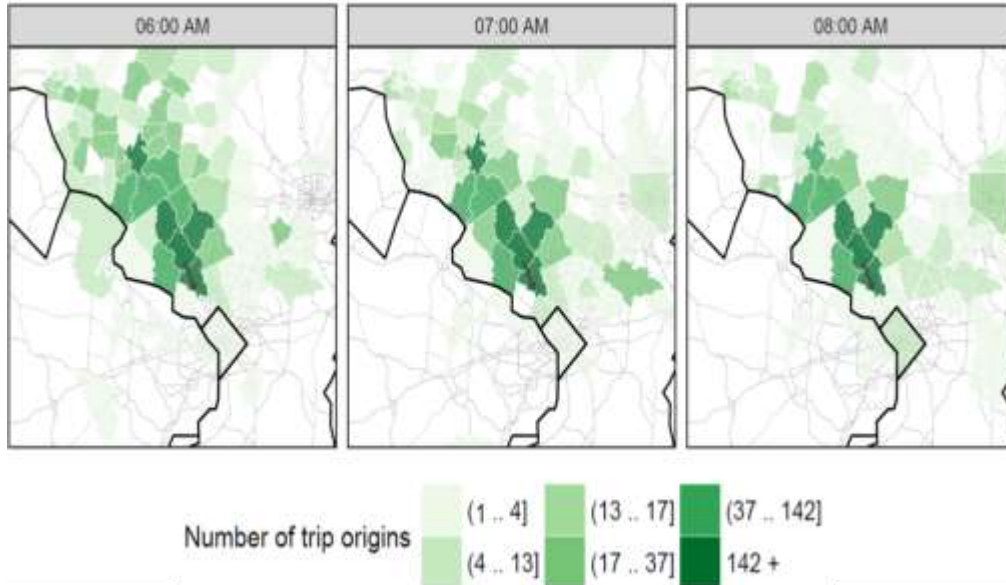
Figure 11. Hourly Evolution of Trip Origin Patterns (6-9 AM)

The same strategy can be applied to understand how destination patterns change over the AM peak period. Figure 12 shows the hourly sub-county destination map for 6-9 AM. Here, the Montgomery County District 13 is shown in a blue border. In comparing the number of trips for each hour within the AM peak period it is evident that within this dataset, arrivals drop off significantly in the 8 AM hour, despite being a major trip attractor for the first two hours. The two analyses described above illustrate how trajectory data supports real-world spatio-temporal demand analysis.
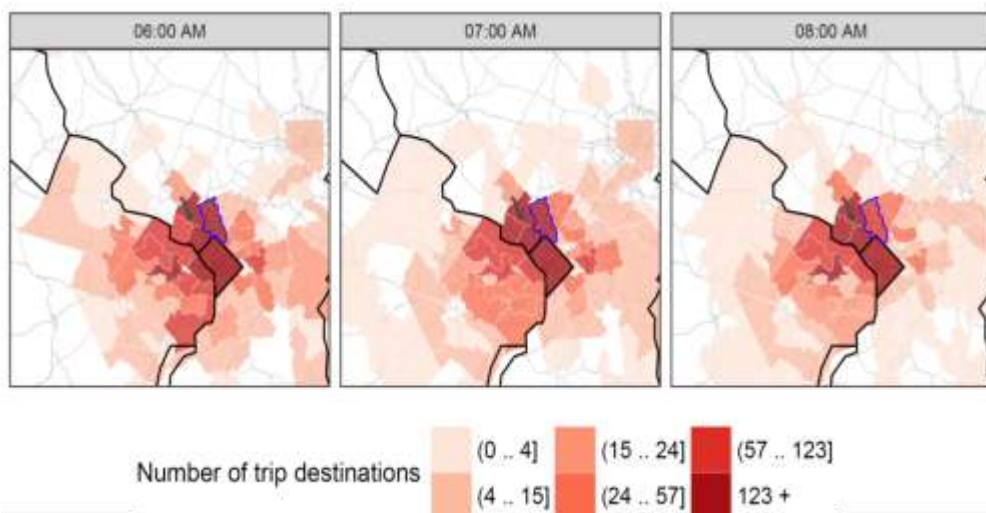
Figure 12. Hourly Evolution of Trip Destination Patterns (6-9 AM)

This section has presented several applications of trajectory data in analyzing real-world travel demand patterns including the development of OD matrices, visualization of routes used, travel time between OD pairs, and segment analysis that filtered for trips using a target segment and assessed the evolution of demand at the sub-county aggregation level. Recognizing that large scale trajectory data analysis is a relatively new path in transportation research, the applications in this section highlight the potential of this data to revolutionize traffic demand analysis.

5. Conclusions and Future Work

This paper presented a procedure to process, visualize and analyze vehicular trajectory data. While this trajectory data represents a very small portion of the total traffic stream, a proof of concept was demonstrated using real-world data on a recurrently congested freeway segment in Maryland. In doing so, several data visualizations were presented and interpreted to illustrate how the developed procedure can be used to analyze traffic congestion. While traditional OD visualization such as origin maps, destination maps and OD matrices are supported by this data, the unique features of trajectory data allow analysts to link OD information with departure time and route choice. This information allows for a more comprehensive understanding of vehicular demand patterns that can ultimately be used to inform real-world data driven congestion mitigation decisions.

Several lesson were learned by analyzing this unique data set. First, the raw data requires significant pre-processing before it can be used for useful travel pattern analysis. Next, matching intermittent GPS waypoints to the road network presents a non-trivial computational challenge, especially when the ping rate is low.

Future research paths include the use of trajectory data to investigate changes in travel patterns, such as route usage under non-recurrent congestion events such as crashes or work zones. Next, this study applied a space-mean speed to determine the travel times between consecutive waypoints. This research team has begun to investigate alternative methods to estimate traffic conditions between ping intervals. Once the procedures are refined, this research team intends to develop a web-based interface for trajectory data analysis.

Acknowledgement

References

[1] K. Siła-nowicka, J. Vandrol, T. Oshan, J. A. Long, A. S. Fotheringham, J. Vandrol, T. Oshan and J. A. Long, "Analysis of human mobility patterns from GPS trajectories and contextual information", International Journal of Geographical Information Science, vol. 30, no. 5, (2016), pp. 881-906.

[2] L. Pappalardo, F. Simini, S. Rinzivillo, D. Pedreschi and F. Giannotti, "Returners and explorers dichotomy in human mobility", Nature Communications, (2015).

[3] S. Gao, "Spatio-Temporal Analytics for Exploring Human Mobility Patterns and Urban Dynamics in the Mobile Age", Spatial Cognition & Computation, vol. 15, no. 2, (2015), pp. 86-114.

[4] S. Brakatsoulas, R. Salas and C. Wenk, "On Map-Matching Vehicle Tracking Data", Proceedings of the 31st VLDB Conference, Trondheim, Norway, (2005), pp. 853-864.

[5] V. Punzo, M. Teresa and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data", Transportation Research Part C, vol. 19, (2011), pp. 1243-1262.

[6] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi and F. Giannotti, "Interactive Visual Clustering of Large Collections of Trajectories", IEEE Symposium on Visual Analytics Science and Technology, Atlantic City, New Jersey, USA, (2009).

[7] N. Andrienko and G. Andrienko, "Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach", Springer Science & Business Media, (2006).

[8] G. Andrienko, N. Andrienko, P. Bak, D. Keim and S. Wrobel, "Visual Analytics of Movement", Springer, New York City, NY, **(2013)**.

[9] N. Andrienko, G. Andrienko and S. Rinzivillo, "Leveraging Spatial Abstraction in Traffic Analysis and Forecasting with Visual Analytics", Information Systems, vol. 57, **(2016)** April, pp. 172-194.

[10] N. Andrienko, G. Andrienko and S. Rinzivillo, Experiences from Supporting Predictive Analytics of Vehicle Traffic, in Proc. IEEE VIS Workshop: Viualization. Predictive Analysis, Paris, France, **(2014)**.

[11] C. Tominski, H. Schumann, G. Andrienko and N. Andrienko, "Stacking-Based Visualization of Trajectory Attribute Data", IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 12, **(2012)** December, pp. 2565-2574.

[12] H. Guo and Z. Wang, "TripVista : Triple Perspective Visual Trajectory Analytics and Its Application on Microscopic Traffic Data at a Road Intersection", IEEE Pacific Visualisation Symposium, **(2011)**, pp 163-170.

[13] S. Liu, J. Pu, Q. Luo, H. Qu, L. M. Ni and R. Krishnan, "VAIT : A Visual Analytics System for Metropolitan Transportation", IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 4, **(2013)**, pp. 1586-1596.

[14] N. Ferreira, J. Poco, H. T. Vo, J. Freire and T. Silva, "Visual Exploration of Big Spatio-Temporal Urban Data : A Study of New York City Taxi Trips", IEEE Transactions on Visualization and Computer Graphics, vol. 19, no. 12, **(2013)**. pp. 2149-2158.

[15] Z. Wang, T. Ye, M. Lu, X. Yuan, H. Qu, J. Yuan and Q. Wu, "Visual Exploration of Sparse Traffic Trajectory Data", IEEE Transactions on Visualization and Computer Graphics, vol. 20, no. 12, **(2014)**, pp. 1813-1822.

[16] Z. Wang, M. Lu, X. Yuan, J. Zhang and H. Van De Wetering, "Visual Traffic Jam Analysis Based on Trajectory Data", IEEE Transactions on Visualization and Computer Graphics, vol. 19, no. 12, **(2013)**, pp. 2159-2168.

[17] W. Chen, F. Guo and F. Wang, "A Survey of Traffic Data Visualization", IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 6, **(2015)**, pp. 2970-2984.

[18] Q. Kong, Q. Zhao, C. Wei and Y. Liu, "Efficient Traffic State Estimation for Large-Scale Urban Road Networks", IEEE Transactions On Intelligent Transportation Systems, vol. 14, no. 1, **(2013)**, pp. 398-407.

[19] J. C. Herrera, D. B. Work, R. Herring, X. Jeff, Q. Jacobson and A. M. Bayen, "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment", Transportation Research Part C, vol. 18, no. 568, **(2010)**, pp. 568-583.

[20] G. Hiribarren and J. C. Herrera, "Real time traffic states estimation on arterials based on trajectory data", Transportation Research Part B, vol. 69, no. 19, **(2014)**, pp. 19-30.

[21] J. Tang, Y. Song, H. J. Miller and X. Zhou, "Estimating the most likely space – time paths, dwell times and path uncertainties from vehicle trajectory data: A time geographic method", Transportation. Research Part C, vol. 66, no. 176, **(2016)**, pp. 176-194.

[22] G. Andrienko, N. Andrienko, G. Fuchs and J. Wood, "Revealing Patterns and Trends of Mass Mobility through Spatial and Temporal Abstraction of Origin-Destination Movement Data", IEEE Transactions on Visualization and Computer Graphics, vol. 2626, **(2016)**.

[23] X. Yang, Y. Lu and W. Hao, "Origin-Destination Estimation Using Probe Vehicle Trajectory and Link Counts Origin-Destination Estimation Using Probe Vehicle Trajectory and Link Counts", Journal of Advanced Transportation, vol. 2017, **(2017)**.

[24] D. Luxen and C. Vetter, "Real-time routing with OpenStreetMap data", In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '11), **(2011)**, pp. 513-516.

[25] P. Newsom and J. Krumm, "Hidden Markov Map Matching Through Noise and Sparseness", 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, **(2009)**.

[26] OpenStreetMap contributors. Planet dump November 25, 2016, (2016), retrieved from https://planet.openstreetmap.org

[27] S. Mahapatra, M. Wolniak, E. Beckett and K. Sadabadi, 2015 Maryland State Highway Mobility Report, Maryland Department of Transportation, **(2016)**.

[28] D. Kahle and H. Wickham, "ggmap: Spatial Visualization with ggplot2", The R Journal, vol. 5, no. 1, **(2013)**, pp. 144-161.

[29] U.S. Census Bureau, Geography: County Subdivision, (2015). Accessed on July 21, 2017. https://www.census.gov/geo/reference/webatlas/cousubs.html

Authors

Mark Franz earned his BA in Physics and BS in Astronomy from the University of Florida, his MSCE from West Virginia University, and his Ph.D. in Civil and Environmental Engineering from the University of Maryland (UMD) with emphasis on transportation engineering. Currently, Mark is a Transportation Analyst at the Center for Advanced Transportation Technology Laboratory (CATT Lab) at UMD where he is developing and improving online transportation analysis tools and visualizations for public and private sector clients.

Anna Petrone earned bachelor's degrees in Mathematics and Economics from Drexel University, and completed a Master's degree in Civil Engineering in the Transportation department at the University of Maryland. Her interests include data visualization, big data, and operations research, with the ultimate goal of increasing the role of transit and shared rides to relieve traffic congestion.