Exploring Google Maps API as a Tool for Measuring Regional and Overall Traffic

Takamoto Kodani
University of Texas at Austin
Austin, TX 78712
kodani@utexas.edu

Sae Saito
University of Texas at Austin
Austin, TX 78712
ssaito@utexas.edu

ABSTRACT

ACM Classification Keywords
Information systems—Location based services

Author Keywords
traffic; commute; Austin; traffic analysis; Google Maps; API; Distance Matrix

INTRODUCTION

On a day-to-day basis, drivers are often confronted with the travails of traffic congestion endemic to urban areas. On a personal level, traffic congestion wastes time, induces stress and frustration, causes delays, and increases money spent on fuel and on fixing the wear and tear that comes from frequent stop-and-go traffic; on a societal level, it leads to increased air pollution, higher chances of road collisions, and heavy traffic spilling over onto side streets.

Austin, Texas is a city that has experienced its fair share of traffic congestion. For its 2016 score report, INRIX has ranked Austin 13th in the U.S. for worst road traffic. The average car commuter spends 47 hours a year in congested traffic, leading to an average of $1,453 a year in extra cost, and the total comes to approximately $810 million per year for the entire city [2]. This is aggravated by the fact that, between 2010 and 2016, Austin was the second fastest growing metropolitan area [1], which consequently put a higher number of vehicles on the road.

With traffic congestion only worsening year after year in urban areas like Austin, there have been many attempts at countermeasures, and one has been the use of intelligent transportation systems (ITS). These are systems with a broad scope of applications that seek to understand traffic patterns for the purpose of developing smarter and more efficient transport networks. They usually take advantage of sensor technologies to detect traffic flow and automate traffic management [9].

The data generated from these sensors are typically beyond the access of most people, but there are now GPS-based navigation applications like Google Maps, Waze, and Apple Maps that allow anyone to see traffic status and to receive optimal route-mapping for their travels. While individuals does not have the ability to change traffic lights or control the number of other vehicles on the road, they can nonetheless make their travels easier by being aware of current and future traffic conditions and by molding their commutes around them. With such massive amount of traffic data at one's fingertips, there is a large potential for data analysis, for understanding traffic patterns without having to rely on sensors, which are expensive and therefore limited in terms of how many you can place alongside the many, many roads in Austin. With a service like Google Maps, you can learn traffic conditions at any time, for any street, and at no cost.

For data to be analyzed, it first needs to be quantified. However, the user interface for applications like Google Maps and Waze is graphical, as well as somewhat non-static, in nature, rendering it difficult to quantify the maps. To combat this, we propose a solution involving the use of Google Maps API to convert the status of various traffic routes into numbers, which are then aggregated to generate scores for different neighborhood regions of Austin, which can then be combined again to culminate in an average traffic severity score for the city of Austin as a whole.

By quantifying traffic maps in such a way, we can more easily discover traffic patterns. This is useful not just for larger entities like the Austin Transportation Department but also for individual drivers, as will be discussed.

CONTRIBUTIONS

We conducted Austin's regional traffic severity analysis using Google Map API which is publicly available and enables light-weighted and inexpensive traffic data acquisition. The fact that this study pertains to Austin traffic makes it relevant for Austin commuters in particular.

We were able to build a prototype of an application that utilizes Google Map API to acquire traffic data and provides a quick glance on the traffic conditions on a given date and time. The regional traffic severity scores can possibly provide users an idea of traffic volume without having to search particular locations and routes. This function is useful for personal uses such as when users look for stores and restaurants in regions with less traffic; when users search for a place to move in to with less traffic time on their commute; when users would

ISBN 123-4567-24-567/08/06.

DOI: http://dx.doi.org/10.475/123_4

like to have an idea of traffic severity in a quick and simple manner.

PRIOR WORK

Many researches have been conducted in interest of traffic volume analysis and short-term traffic forecasting. Short-time traffic forecasting considers prediction of traffic made over the time period of few seconds to few hours into the future based on current and historic traffic information [10].

Most of the studies focused on developing new models and algorithms to conduct statistical analysis of traffic data. Statistics is the "mathematics of collecting, organizing and interpreting numerical data, particularly when these data concern the analysis of population characteristics by inference from sampling" [5]. Hence, statistical analysis is useful to interpret the data and causalities. Recently, trends are moving towards the analysis based on data driven modeling utilizing Computational Intelligence and Neural Network approaches [5]. Neural Network has been mainly favored as data analytics methods for traffic research because of its generalization ability, adaptability, and good predictability. It can also deal with nonlinear models while statistical (regression) approach often cannot deal effectively with nonlinearity [5] . Therefore, Neural Network is often regarded as more useful compared to statistical models when modeling complex datasets with nonlinearities. It is also effective when obtaining good predictions is important, and explanation and interpretation of results are not on focus.

The data collection for traffic analysis has been classically done by putting cameras on streets and counting the vehicles that have passed by. Another way to analyze traffic volumes is to use the data from loop vehicle detectors installed under the roads [3]. There were also some traffic studies and travel time prediction done using simulations [4, 8].

The effect of various factors such as weather, heavy vehicles, speeds, and type of day on traffic volume has been studied in some literatures. Konstantopoulos et al. [6] indicated that rainfall can affect the drivers' visibility and therefore increase the risks of accidents. Rainfall can also make driving more dangerous due to decreased friction on a wet road. These increased risks affect driving behaviors and hence the traffic volume [7]. Li and Chen [7] has found that including the scheme of the days of the week in the traffic analysis can improve the accuracy of travel time prediction compared to when including only scheme of weekdays and weekends. Using the time as an input variable in the traffic analysis can also improve the traffic prediction model [7].

Few studies have been done to identify traffic patterns using Google Maps. Juntunen et al. [3] presented a lightweight system utilizing Google Maps and Charts Application Programming Interfaces (API) to provide traffic engineers a way to interact and explore the traffic data more easily and without the data mining. This system consists of web user interface, a server-side PHP API for database access that handles queries and calculations of traffic volume information, and a MySQL database containing information about detector loops, intersection structure information, and traffic volume data. In this system, Google Map is used on web user interface as a visualization tool, while real-time traffic data is collected from detector loops.

METHOD

We used Google Maps for the purpose of this research, due in part to its free and accessible platform, but also due to the web APIs that it offers. As of this writing, there are seventeen different types of Google Maps APIs - they range in kind from allowing a map to be embedded on a web page to giving directions based on two locations provided, to giving elevation based on a set of coordinates.

For this study, we focused on the Distance Matrix API, which outputs the distance and travel time between an origin point and a destination point. In particular, we were interested in the travel time, as we hypothesized it can serve as an indicator of how heavy or light traffic is along a route. By calling the Distance Matrix API with a script repeatedly, it becomes possible to capture and record the travel time of a route at different times of the day, every day, for as long as Google offers this service.

In total, we collected the travel times for 44 different routes, which were in turn categorized into seven regions as well as a Cross-regional category (I-35 and MoPac) within the city of Austin (see Figure 1). We selected highways and major streets that usually sees heavy traffic during rush hour. Data for all routes were collected automatically with a Python script every fifteen minutes or so for 6 to 12 hours a day for a period of two weeks. Because the Distance Matrix API has a daily usage limit on the number of times a request can be made (2,500 elements per day), we decided not to include any additional routes to avoid unexpectedly reaching the daily quota.

The data we gathered were all then converted into traffic severity scores. This was done by finding the shortest travel time and the longest travel time for each route, and then calculating each data point's place within this range - for example, if the shortest time recorded was 10 minutes and the longest time 30 minutes for route A, a travel time of 20 minutes would yield a score of 50 on a scale of 0 to 100. We averaged the travel times across regions and as a whole for each moment when traffic was captured. Overall average of Austin traffic was weighted to place greater emphasis on highways.

In addition to collecting traffic data, we gathered personal car commute data, which involved manually recording departure time, arrival time, route taken, and subjective traffic severity rating.

DATA ANALYSIS

The severity score for each region organized by the day of the week is shown in Figure 2. (The color of the graph lines corresponds to the color-codes used on a map of Austin shown in Figure 1) As expected, the traffic is more severe on weekdays with Thursday being the most severe for overall Austin. Cross-regional has the highest peak traffic among all the regions. Southeast region has the lowest peak traffic. However, if we take a median of traffic severity score in each region, South Central region is shown to have highest median severity

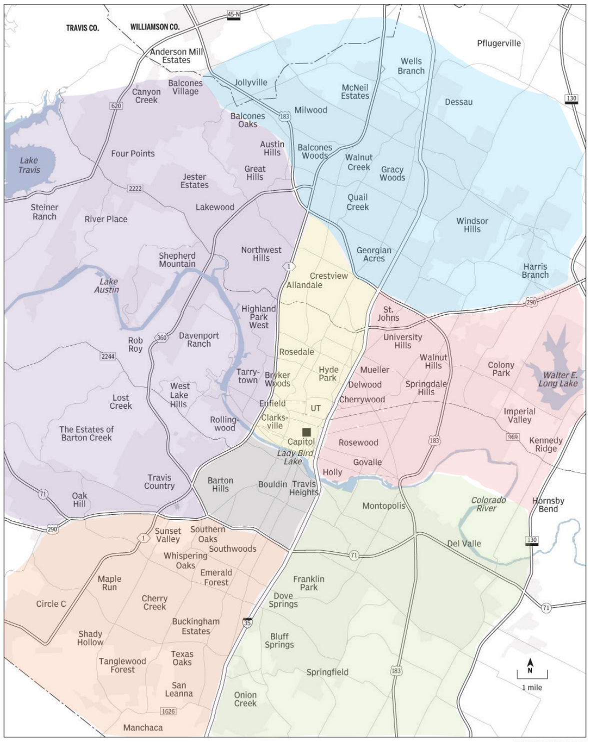

Figure 1. Map of Austin divided by region

score and North region has the lowest score (Figure 3). These data indicate that South Central region keeps relatively high traffic volume throughout a week. South Central region is also unique in a way that it has peak traffic on Saturday (Figure 2). These phenomena could be due to the South Central region having event centers, parks, and restaurants that are popular on weekends.

We also conducted analysis of the severity score at different time of a day for each region (Figure 4). There are two high peaks at 7AM and 4PM which correspond to the commuting hours. Cross-regional has the highest peaks again as expected, while Southeast region has lowest peaks. There is also smaller peak at 1PM for all regions. This peak likely corresponds to the lunch hour traffic. Traffic calms down at around 7PM for most regions. However, South Central is unique again that traffic severity retains higher value until midnight compared to other regions.

DISCUSSION

By converting travel times into scores and then averaging them by region we were able to discover unique traffic patterns associated with each region. Based on personal experience, the findings seem not too far removed from reality. For example, highways see significantly increased traffic during rush hour; Tuesday through Thursday traffic is heavier than all other days of the week; and traffic in South Central remains more or less consistent throughout the week, having the greatest severity scores on Friday through Sunday. If we had separated

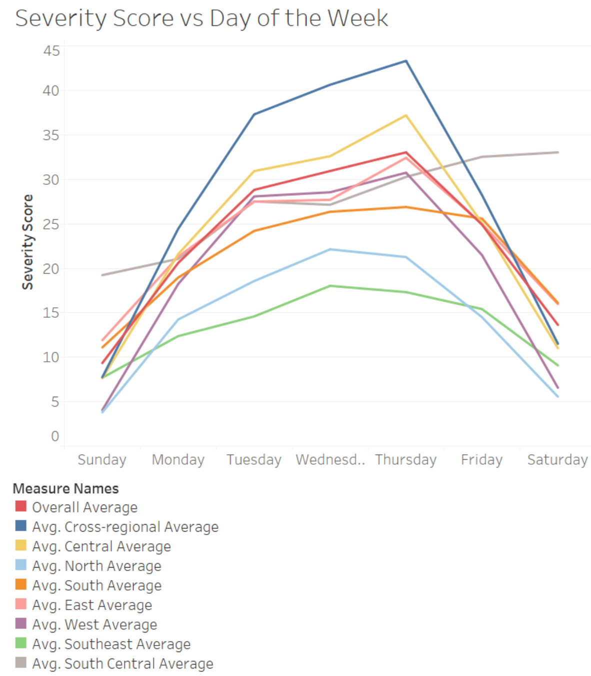

Figure 2. Severity score for each region organized by day of the week

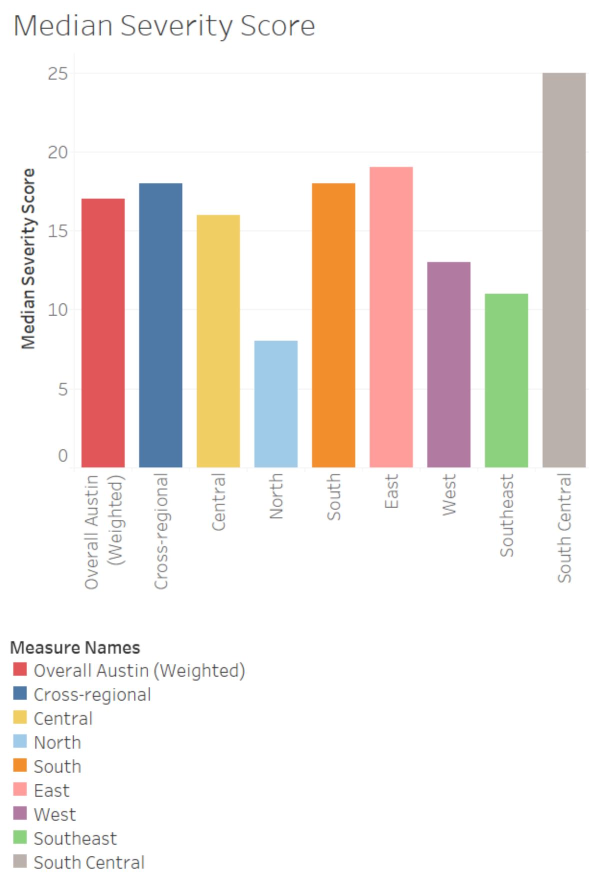

Figure 3. Median severity score for each region organized

downtown from Central as its own region, we expect we would have seen a similar behavior.

We realize checking our results with personal experience is a poor way of validating our data. In the future, steps should be taken to confirm that the data is valid and that the scores accurately reflect the severity of traffic in reality. One way this can be done is by making comparisons with traffic sensor data from the city. If it can be determined that there is a strong relationship between results produced from our Google Maps method and those from more traditional methods, the former method should be given greater consideration in the future, as it is inexpensive, accessible to anyone, scalable to other regions and cities, and customizable to any route.

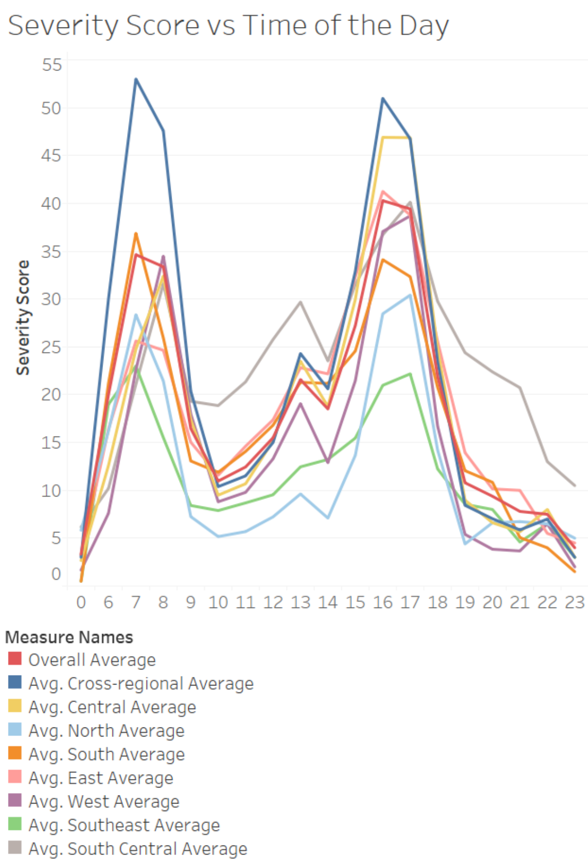

Figure 4. Severity score for each region organized by time

Limitations

One of the biggest issues in our study is that data was collected over a very short time period of two weeks within a single month (April 2017). Traffic patterns vary by season, by weather, by the presence of road construction and of traffic accidents, among many other variables. Continuing to collect traffic data for an additional number of years (not just weeks or months) will likely bear more meaningful results. Austin, for instance, is a college town that probably sees more traffic during the fall and spring semesters than during the summer - April traffic could very well be different from May traffic.

Another limitation with our research method is related to how regions and routes were selected. The decision on how Austin should be divided was somewhat of an arbitrary process; any and all of the demarcations we ultimately settled on could very well be contested. At the same time, the point of the study is not so much about how the regions are culturally different from one another, as it is about exploring the potential for extracting data within each specific region, and we believe we have been able to do that.

The crux of being able to produce an accurate traffic severity score for a region using Google Maps Distance Matrix API hinges on proper route selection. The important question that needs to be addressed is, do the routes selected actually represent the traffic severity of the entire region that they belong to? If they do not, then the score is meaningless. We have tried to choose routes that were a) composed of major roads and highways because they see greater traffic and b) traversed the entire expanse of a region, both north to south and east to west, at the very least. Nonetheless, it is a concern that should be looked more closely into in future studies.

Before writing a script leveraging Google Maps API, we wrote a script that analyzed the traffic map of Google Maps by counting the number of pixels based on the four colors (green, orange, red, and maroon) that it showed to indicate traffic severity. Although this idea was abandoned due to it being inconvenient, more involved, and unable to be easily localized within certain regions, it is something that should be re-explored, as it is able to take into account a far greater number of routes, unlike the method we ultimately we decided to go with.

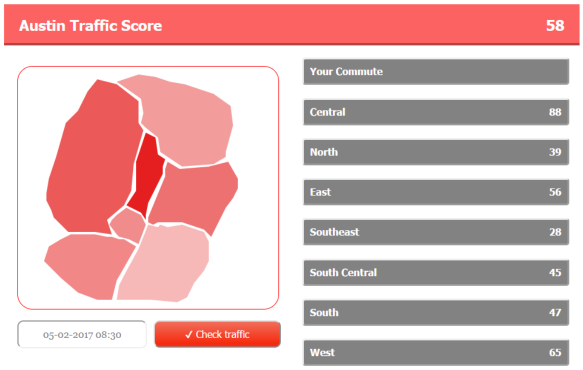

Figure 5. Simple app developed to demonstrate use of traffic scores

APPLICATIONS

Although aggregating traffic severity scores can find their use among city officials for understanding and improving the transport network, we present potential uses for them on a personal level - as a way for individuals to keep tabs on traffic status. As a demonstration, we have built a very simple, proof-of-concept app where we put these scores into action.

The app (Figure 5) pulls actual route travel time data using the same Google Maps Distance Matrix API we used for the study. Then, using the same calculation method we used for converting that data, scores are generated for traffic status of Austin overall and of the seven regions we initially identified. The app serves as a one-stop dashboard to quickly learn traffic conditions, either now or sometime in the future, as predicted by Google Maps. While not yet possible, we envision it being able to add one's own personal routes, which will eventually show the traffic severity score for each route when enough data is gathered. Another feature that could be added in the future is one where the user can draw a route on the map directly, and the app would provide an estimated time of travel, instead of having to enter an address one by one.

While mobile apps like Google Maps and Waze are already popular and very useful for navigating traffic and managing one's commutes, they are heavyweight applications that can be too much for some purposes. By relying on traffic severity scores, use cases for our app include being able to check traffic condition of one's personal commute on the fly and being able to see a rough estimate on traffic severity within multiple routes and multiple regions at different times of the day and at different days of the week with very little effort.

Although not yet possible with current system, the app would in concept be useful when one has the option of going to a store or restaurant within various regions; when one is moving and wants to know his or her commute estimates based on location; and when one has the leisure of going to work, home, or school at any time but want to know when traffic would likely be tolerable enough. Of course, these are all things that can be done with Google Maps and Waze but are more time-consuming and can be rather inconvenient to do.

FUTURE WORK

One of the limitation of this study was the small sample size of the collected traffic data. In our future work, we need to include the data from longer time period. The number of routes used in this study was also small due to the limited quota of the Google Map API. The selection of routes to include in the regional traffic severity analysis needs further study.

For better traffic severity predictability, we need to consider other possible variables other than time and day of a week, such as weather, accidents, and events.

Our application require further development with accuracy in prediction of traffic severity, better user interface, and other features such as interactability with the map.

While we relied entirely on data provided by the Google Maps API, there are also other web services, such as Microsoft, MapQuest, and HERE, which offer their own traffic APIs; rather than depending on a single service, using a wide range of services can help augment the size and reliability of data. In addition, it may be beneficial to take advantage of weather APIs (from Weather Underground, for example) and incident report APIs to automatically gather secondary data on top of traffic data.

CONCLUSION

In this research we have proposed a novel, cost-effective, and efficient way of quantifying traffic flow. Although the data we collected was not as comprehensive as it could have been, early results are promising. We expect fine-tuning our route selection, region mapping, and score calculation formula, as well as expanding our data set by collecting traffic data for a greater length of time, will yield better and more accurate scores, which can then be used to quantitatively analyze traffic patterns and build personal commute-management apps.

DIVISION OF WORK

Takamoto Kodani: Logged traffic; wrote scripts for capturing data from Google Maps API and for calculating traffic severity scores; built the app using HTML, CSS, JS, and Python; and wrote part of the article. **Sae Saito**: Logged traffic; cleaned, aggregated, and analyzed the data; created data visualizations; and wrote part of the article.

ACKNOWLEDGMENTS

We thank Professor Thomaz for his help and valuable comments on this project.

REFERENCES

1. Data Access and Dissemination Systems (DADS). 2010. Your Geography Selections. (Oct 2010). `https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk`

2. Inrix. 2017. INRIX 2016 Traffic Scorecard - U.S. (2017). `http://inrix.com/resources/inrix-2016-traffic-scorecard-us/`

3. Tomi Juntunen, Vassilis Kostakos, Mikko Perttunen, and Denzil Ferreira. 2012. Web Tool for Traffic Engineers: Direct Manipulation and Visualization of Vehicular

Traffic Using Google Maps. In *Proceeding of the 16th International Academic MindTrek Conference (MindTrek '12)*. ACM, New York, NY, USA, 209–210. DOI: http://dx.doi.org/10.1145/2393132.2393173

4. Natalia Juri, Avinash Unnikrishnan, and S. Waller. 2007. Integrated Traffic Simulation-Statistical Analysis Framework for Online Prediction of Freeway Travel Time. *Transportation Research Record: Journal of the Transportation Research Board* 2039 (Dec. 2007), 24–31. DOI:http://dx.doi.org/10.3141/2039-03
5. M. G. Karlaftis and E. I. Vlahogianni. 2011. Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies* 19, 3 (June 2011), 387–399. DOI: http://dx.doi.org/10.1016/j.trc.2010.10.004
6. Panos Konstantopoulos, Peter Chapman, and David Crundall. 2010. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention* 42, 3 (May 2010), 827–834. DOI: http://dx.doi.org/10.1016/j.aap.2009.09.022
7. Chi-Sen Li and Mu-Chen Chen. 2013. Identifying important variables for predicting travel time of freeway with non-recurrent congestion with neural networks. *Neural Computing and Applications* 23, 6 (Nov. 2013), 1611–1629. DOI: http://dx.doi.org/10.1007/s00521-012-1114-z
8. Y. Ma, M. Chowdhury, A. Sadek, and M. Jeihani. 2012. Integrated Traffic and Communication Performance Evaluation of an Intelligent Vehicle Infrastructure Integration (VII) System for Online Travel-Time Prediction. *IEEE Transactions on Intelligent Transportation Systems* 13, 3 (Sept. 2012), 1369–1382. DOI:http://dx.doi.org/10.1109/TITS.2012.2198644
9. Sumit Mallik. 2014. Intelligent Transportation System. *International Journal of Civil Engineering Research* 5, 4 (2014), 367–372.
10. Eleni I. Vlahogianni, Matthew G. Karlaftis, and John C. Golias. 2014. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies* 43 (June 2014), 3–19. DOI: http://dx.doi.org/10.1016/j.trc.2014.01.005