FIGHTING TRAFFIC WITH EMERGING TECHNOLOGIES: AN ANALYSIS OF AUSTIN TRAFFIC USING DYNAMIC TIME WARPING AND BLUETOOTH DATA

HONORS THESIS

Presented to the Honors College of
Texas State University
in Partial Fulfillment
of the Requirements

for Graduation in the Honors College

by

Patrick Peter Ray Sullivan

San Marcos, Texas
May 2019

FIGHTING TRAFFIC WITH EMERGING TECHNOLOGIES: AN ANALYSIS OF AUSTIN TRAFFIC USING DYNAMIC TIME WARPING AND BLUETOOTH DATA

by

Patrick Peter Ray Sullivan

Thesis Supervisor:

____
Yihong, Yuan, Ph.D.
Department of Geography

Approved:

____
Heather C. Galloway, Ph.D.
Dean, Honors College

Table of Contents

1. Abstract

There is a growing field of research surrounding geography and big data. With the explosion of the internet of things, large collections of data hide unforeseen patterns waiting for future research to uncover. As technology continues to grow, new applications in geography and city planning will continue to develop. One technology currently in development for urban planning applications is Bluetooth technology. In the past decade, institutions began to analyze trends in Bluetooth data for traffic analysis. An example of Bluetooth traffic analysis is tracking how a city's population moves over time. This research paper extracts data from a City of Austin file and performs a dynamic time warping analysis using Python. To be able to calculate the dynamic time warping analysis this research paper extracted a frequency count from a normal Saturday, a normal weekday, and Memorial Day in 2016. The data was made to compare traffic disparities between the busy Memorial Day and the other two "normal" days in Austin. Dynamic time warping is a statistical analysis which compares the similarity between two different data sets. The results of the dynamic time warping demonstrated that the center of Austin had a greater disparity when compared to the locations on the outskirts of town. After conducting the research, the Python program illustrated a larger volume of traffic during Memorial Day when compared to normal traffic conditions. Cities across the United States will be able to predict traffic volume and urban population growth with this technique. More research will bring cities into the 21st century as urban centers begin to create smart cities which could provide emergency services, alternative traffic routes, and create alerts for severe weather conditions.

2. Introduction

Technology is progressing at a rapid rate. As the internet of things grows and expands, there will be unforeseen innovations in the next decade. The internet of things is a general term used to describe how more electronic devices are added to different networks and how those devices communicate with each other. Volumes of data are already starting pile up, and programmers are beginning to uncover uses for this data. One such use is Bluetooth traffic data. Bluetooth devices are becoming common in automobiles. With more cars acquiring technology, city governments are taking advantage of this opportunity. These new advancements are allowing city planners to make observations in anthropomorphic migrations amongst highly urbanized regions. A Python script was created to study the difference in the volume of traffic on three separate occasions in the city of Austin for this report to explore such new technological methods.

We will explore how transportation planners can alleviate congestion in the streets of the Austin Metroplex with a dynamic time warping analysis with Bluetooth data. More specifically, the analysis in this project will highlight areas with extreme traffic fluctuations throughout the day. Dynamic time warping is a mathematical algorithm used to measure the average difference between two time series. The mathematical algorithm will allow urban planners to create new roads were necessary to help end the many traffic problems presently plaguing Austin.

After the dynamic time warping analysis, ArcMap Pro will generate two maps. The maps show the difference in traffic when comparing Memorial Day in 2016 which has a large volume of traffic when compared to a normal weekday and a regular Saturday. The dynamic time warping was able to highlight differences between a normal day and a busier than average day.

As more devices become connected, it becomes easier to research the humanistic behaviors of a population. However, many individuals do not know how Bluetooth towers work or what their implementations are (Yao et al, 2019). Such behaviors are important to consider when mapping and processing Bluetooth data. The lack of knowledge might cause some hindrances in future studies because of the confirmation biases many individuals possess. Also, Bluetooth networks have plenty of faults on their design such as a slow connection and a lack of network support which exist outside the realm of human error. Therefore, when using Bluetooth data in the future, it is important to keep in mind problems for statistical studies for this project and other geographical studies.

3. Literature Review

3.1 Analyzing Urban Land Use and Dynamics from Big Geodata

Big data is an emerging field which is exploding at a rapid pace. As the area of study grows, it will spread into different academic disciplines. For instance, geographical information science will benefit from big geodata.

One application of big geodata is to find correlations in social media platforms. By using data collected from social media such as Twitter, we can gauge what a population of a region knows and find relationships within the information. A study was conducted to assess an average person's knowledge of cities within the United States (Han et al, 2015). Tweets were analyzed to predict a pedestrian's understanding of towns in a local and regional basis. To analyze the data, the researchers used Twitter's API to filter through the data and collect the names of cities relative to their location. Researchers used the global awareness index to determine how much Twitter users comprehended cities relative to their location.

After calculating and normalizing the global awareness index, the researchers found that Twitter users' understanding of geographic locations of towns followed the First Law of Geography. The First Law of Geography created by Waldo Tobler states everything is related to everything else, but near things are more related than distant things. Results from the study demonstrated the phenomenon by illustrating Twitter users knew more about the local towns than cities farther away. As research continues, more applications for the First Law of Geography will grow. Additionally, the results of the Bluetooth traffic analysis followed the First Law of Geography.

Mapping population density with cell phone data is another application of big geodata. With cell phone usage and GPS data, it is possible to track spatiotemporal trends from a regional population (Deville et al, 2014). In their research, they learned cell phone data is better at calculating dynamic population density than conventional methods. A regression line was created to find the mean of the areas studied. The research team was able to accurately calculate population trends in portions of France and Portugal which is crucial to climate change adaptation, public health, and food security. In the future, the City of Austin might use Bluetooth data collected from cell phones instead of cars to receive a more accurate representation of the study area.

Big data is used to create a statistical analysis over a given study area. On example is by studying housing prices and economical price influencers. The data for finding the causation can be created by mining social media data. In China, a research team created a heat map which illustrated where housing prices fluctuated relative to characteristics in the provinces (Wu et al, 2016). Heat maps in the study concluded green spaces, the construction materials of a house, and the quality of the education systems surrounding the neighborhood affected housing prices. However, there are many other applications of spatial statistics by studying Twitter responses to natural disasters, marketing tactics based on big data, and linguistics which can change based on geographical location. All these factors could be used to see how housing prices can affect traffic in the area.

There are large amounts of data produced every day just waiting to be analyzed. Air pollution monitoring stations have a large collection of data which could be analyzed. In this research study, patterns in wavelength clusters of remote sensing data identified the region where the air pollution stations were located (VoPham et al, 2018.). The study predicted the amount of air pollution that would occur in each area because of the historical data and environmental factors. Data collection for traffic analysis is like the air pollution stations because the data comes from several locations in a region. Although the towers collected the data in separate ways, illustrating the results are similar because they come from point data.

As the global network of devices grows, cities will need to integrate new and secure networks to adapt to changing technology. In *Proposal and Application of Bluetooth Mesh Profile for Smart Cities' Services*, the author describes how a Bluetooth Mesh Profile (BMP) could be created to securely integrate more Bluetooth devices to the urban network (Veiga et al, 2019). BMP is a protocol Bluetooth data follows to avoid overloading the system. The steps of the protocol are the model layer, the foundational model layer, the access layer, the upper transport layer, the lower transport layer, the network layer, and the bearer layer. Essentially, the modal and foundation model layers define what the purpose of the data is for, the access and the upper transport layers encrypt and secure the data, and the lower transport, network, and bearer layers segment and transport the data packets to there location. This technology can illustrate geographical concepts for transportation, arbitrary data collection for actions performed on the network for big data analysis, and for critical infrastructure performance such as a Bluetooth network amongst water plants.

3.2 Applying Bluetooth Data to Urban Planning and Transportation

As the internet continues to grow, there will be more devices added to localized area networks across cities all around the globe (Krco, 2019). In the creation of more localized area networks, it will be possible to create intelligent cities that could predict traffic, send traffic assistance when needed, and provide advanced data analytics for future spatiotemporal pattern recognition. In China, provinces have begun to track cell phone Bluetooth signals from cell phones to predict where traffic patterns will occur in the future. In *Understanding Road Performance Using Online Traffic Condition Data,* the researchers used a variety of techniques with a self-stimulating behaviors dataset (Chen et al, 2019). One technique called tidal harmonic analysis was able to accurately know where Chinese citizens would be located depending on what part of the day it was. It used Bluetooth data and traffic patterns to accurately predict when traffic would start and stop in the concept of how waves and tides function. As technology continues to grow, more applications will allow urban planners to create smarter cities which can adapt to traffic issues.

Following the motif of Bluetooth Data and traffic analysis, there are several other methods which can work with Bluetooth networks. For instance, if the Bluetooth devices communicated with each other, it would be possible to send assistance when necessary. With a cooperative intelligent transport system, it would be able to use machine learning to allow traffic systems to know when accidents are most likely to occur and to alert emergency respondent teams when necessary (Javed et al, 2019). Furthermore, civil engineers are starting to use Bluetooth vector analysis to predict traffic congestion by using spatial-temporal trends (Laharotte et al, 2015). The technique uses the original position of a Bluetooth device and the location of the device over a period of time to predict the speed of the device. The spatiotemporal analysis creates a model showing where traffic speeds fluctuate around the city (Yang et al, 2015). Along with Bluetooth data, machine learning techniques can be applied to big Bluetooth data to create traffic forecasts.

Data generated from smart cities can be manipulated to produce models for emergency simulations and mass urban migration patterns. Using a variety of inputs such as weather conditions, time of the day, static features such as speed limits, and human factors such as population density it is possible to create a supervised machine learning program which can accurately foresee where traffic accidents occur. Such processes are explored more in *Using Machine Learning to Predict Car Accident Risk* where ESRI creates a machine learning script using a Python module called XGBoost to create traffic simulations with decision trees with the input data listed above (Wilson, 2018). The result was a shapefile which displayed where traffic accidents are likely to occur. Machine learning techniques will help urban planners to create software in the future which will be able to predict where traffic accidents will occur, and how to alleviate traffic jams as more data becomes readily available. Data collected in this project could be used to create machine learning modules in the future.

There are several other methods which can create analyzes for urban planners to use outside the realm a machine learning. One type of statistical method is game theory. Game theory is the mathematical model of interactions between rational decision makers. Behavioral relationships include statistical models such as kin selection, direct reciprocity, and indirect reciprocity (Cortés-Berrueco et al, 2016). Such experiments were able to predict traffic data without accessing Bluetooth data. Cameras are also used to collect information for traffic.

Though Bluetooth data may have its perks, the collection of it can be tricky because of the limitations of the number of cars compatible with the technology. Furthermore, environmental factors might cause interference with the Bluetooth wireless connection. Cameras are placed around cities to observe traffic and calculate traffic speed could fix these issues (Shafie et al, 2019). With the camera counting the physical number of cars, the results would not skew the number of cars not compatible with Bluetooth technology. Perhaps in the future, the best solution would be to use a hybrid method of Bluetooth data and cameras. The cameras and Bluetooth sensors can be used to detect differences in traffic amounts caused by auto and horse races, university move-in days, festivals, major concerts, and seasonal shopping (U.S. Department of Transportation, 2018). By using a combination of methods, urban planners and geographers can use the data available from Bluetooth data and other means to collect information and make predictions for new traffic models.

3.3 Time Series Analysis and Its Applications in Urban Studies

For this research project, the main analysis will be done using dynamic time warping. The purpose of dynamic time warping is to measure the average difference between two-time series and adjust for nuances in the data to help create the best possible alignment between two-time series sets. This process is better than the Euclidian distance method because the Euclidian distance does not account for time adjustments. Euclidian distance is the difference between two-time series without the distance matrix.

Dynamic time warp follows a set of protocols which adjusts for data abnormalities. To calculate dynamic time warping, one must get a two-time series and create a distance matrix. In this context, the two-time series represent the x and y-axis for the distance matrix. For this study, the x-axis is the larger frequency count, and the y-axis is the smaller frequency data count. Data used from the x-axis was used to compare the normal Saturday and weekday. Once the distance matrix comes into completion, the values for both time series will fill the matrix. The following formula calculates the distance matrix.

$$\left| Ai - Bj \right| + \min(D[i-1, j-1], D[i-1, j], D[i, j-1])$$

This formula subtracts the x value and y values and then adds the smallest value from the left matrix value, bottom left matrix value or the bottom matrix value from the current matrix coordinate. The process must follow through for every unit in the matrix. Once finalized, the graphical representation will follow the smallest values from the topmost right corner until it reaches the bottom of the distance matrix (Elsworth, 2016). The follow images show how dynamic time warping is calculated. The first image demonstrates how to fill in the distance matrix and the second image illustrates how to select the correct numbers for the results.

		Dynamic Time Warping Example			
5	8				
7	8				
6	6	9	5		
8	5	8	4		
3	2	3	7	8	8
	5	2	7	4	3
Step 1	7 - 6 = 1				
Step 2	4 is the small number in the series 9,8 and 4				
Step 3	1 + 4 = 5				

Fig. 1. Steps to calculate the distance matrix

		Dynamic Time Warping Example			
5	8	11	7	6	8
7	8	11	5	8	10
6	6	9	5	6	9
8	5	8	4	8	13
3	2	3	7	8	8
	5	2	7	4	3

Fig. 2. How to select the appropriate values for dynamic time warping.

Think of dynamic time warping as a football team trying to score a touchdown. Making the distance matrix is like planning the plays. To create the distance matrix, the football coach produces a series of winning plays to score a field goal or a touchdown. Each slot in the distance matrix represents a play. To generate the distance matrix, start by subtracting the numbers on the x and y-axis, and then adding the smallest value from the bottom the left, bottom left, or bottom to the location currently analyzed on the matrix. Upon completion, the football team knows how to move through the opponent's defensive line. In terms of dynamic time warping, the numbers selected for the algorithm start with the top right most value, then in the next iteration selects the smallest value in the left, bottom left, or bottom value and the process continues until the algorithm reaches the bottom of the matrix. In football terms, the selected values in the matrix show the plays the football team will do to score a touchdown in the endzone and hopefully win the game.

A variety of academic fields and institutions use dynamic time warping. Newly created technologies are rapidly evolving by using dynamic time warping. Some of the technologies presently being created are for traffic applications. Traffic is a phenomenon created by fluctuating car speeds which sends ripple effects through the entire highway resulting in stopped cars. One way to study the behavior of traffic is by observing the speed at which individualized groups of cars move. With derivative dynamic time warping, scientists can record how long cars remained stopped for in a localized area using cell phone GPS data (Chandrasekaran et al, 2011). By following derivative dynamic time warping, the group of researchers was able to predict when bottlenecks would occur by observing the start and stop time of individual vehicles with their GPS data. Plus, scientists were able to create personalized routes for avoiding traffic with a user's GPS data history (Nawaz et al, 2014).

Furthermore, in a similar study, researchers concluded dynamic time warping has applications to tracking cell phone data. It was much more accurate when compared to traditional methods such as the Bayesian method (Kittipong et al, 2010). The Bayesian method is a field of statistics that studies continuous random actions. Dynamic time warping is becoming a go-to for the source of traffic information as the technology becomes more readily available. Plus, dynamic time warping is better than the Bayesian method because it uses observations physically made and not predictions. These studies are related to this research project because they are alternative miscellaneous methods to quantify traffic and which could be used in combination with the Bluetooth data analysis in the future.

Python is the language of choice when it comes to data science. Data scientists from all fields of study need an easy computer language to master and perform the tools necessary for their analyses. The data science tool NumPy was used in this research to create data sets for dynamic time warping modeling.

4. Research Design

The data used in this research study was collected from the City of Austin's website. Austin has a network of Bluetooth travel sensors which compile the frequency of cars passing by the towers in a given hour. Each detection records a randomly generated record ID, the host read time, the field device read time, the reader identifier, and the device Media Access Control (MAC) address. The Bluetooth towers are able to determine if the devices in the network is a car by looking at the device's MAC address and the car's Controller Area Network (Brown, 2019) The following figure is a segment of the data captured from the city of Austin's website to show the original data (City of Austin Transportation Department, 2019).

65dc72ef2b5802900576727bd...	01/01/2016 12:00:00 PM	01/01/2016 11:55:51 AM	south_1st_stassney	ae:c9:45:28:5f
ff02e661588edc2384746d6144...	01/01/2016 12:00:00 PM	01/01/2016 11:55:51 AM	anderson_mill_us183	20:e0:db:94:fd
76534ff4dcf445e5cf89be73dcb...	01/01/2016 12:00:00 PM	01/01/2016 11:55:51 AM	congress_slaughter	19:6b:33:b0:08
8012fdb4503cc287e99af27e69...	01/01/2016 12:00:01 PM	01/01/2016 11:55:52 AM	south_1st_wm_cannon	e4:62:6e:f0:45
452d487687d37a4328ba61cb3...	01/01/2016 12:00:02 PM	01/01/2016 11:55:53 AM	ih_35_riverside	80:c9:65:f8:9c
8fc90112f968529f837c6fe821f...	01/01/2016 12:00:02 PM	01/01/2016 11:55:54 AM	lavaca_6th	29:3e:ef:0b:ce
469efa8973fbce822408d697f1...	01/01/2016 12:00:03 PM	01/01/2016 11:55:54 AM	guadalupe_21st	bf:cc:02:eb:c3
860910897d30d5d76f25636ce...	01/01/2016 12:00:03 PM	01/01/2016 10:16:49 AM	south_1st_oltorf	14:50:b9:b8:e5
1c1c0ec122eea8a595b56a2852...	01/01/2016 12:00:04 PM	01/01/2016 11:55:55 AM	mlk_lavaca	c0:0e:43:31:5c
41372e5ea1f27f513f87f4d3092...	01/01/2016 12:00:04 PM	01/01/2016 11:55:55 AM	south_1st_wm_cannon	9f:68:7a:91:2d
744979f92011bf16e776969b78...	01/01/2016 12:00:04 PM	01/01/2016 11:55:55 AM	burnet_duval_loop1	92:ef:04:5b:48
a810e7de979021606acedf5b3...	01/01/2016 12:00:05 PM	01/01/2016 11:55:55 AM	lamar_mlk	96:64:bc:86:48
ee3f71ec9937f228c68cacc25ae...	01/01/2016 12:00:05 PM	01/01/2016 11:55:56 AM	congress_5th	ca:29:56:3c:31
328e862abbb72a53d9817b8da...	01/01/2016 12:00:06 PM	01/01/2016 11:55:57 AM	lavaca_6th	34:10:90:17:7e

Fig. 3. Original data from the City of Austin's website.

There are multiple data processing methods used to illustrate different patterns. One purpose would be to demonstrate short term trends in the traffic data. Furthermore, given the MAC address of certain devices, we can track the distance recorded for individual MAC addresses and speculate patterns within the data.

5. Methodology

The Python script written for this project will be able to compare two data sets and adjust for abnormalities in the data by using dynamic time warping (DTW). The following two flow charts were made to understand the algorithm behind the program that extracted the data from the City of Austin's Bluetooth Dataset (CABD) and the Dynamic Time Warping (DTW) program.

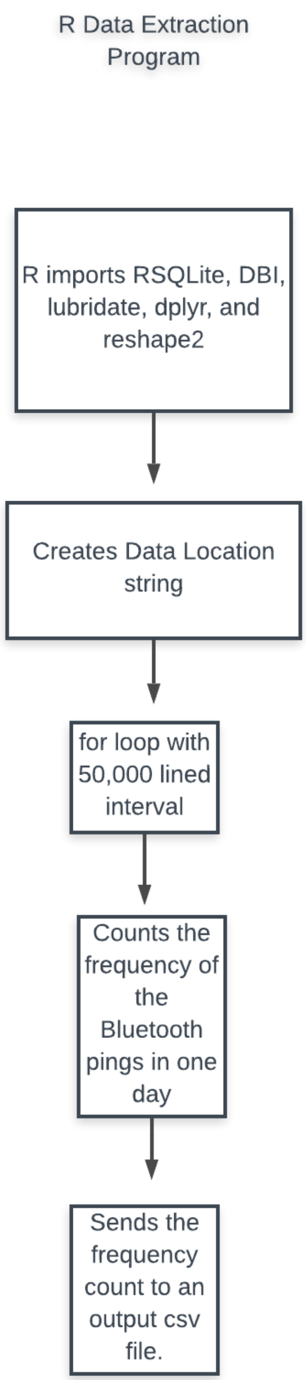

Fig. 4. Algorithm for the R code that extracted data from the City of Austin's dataset.

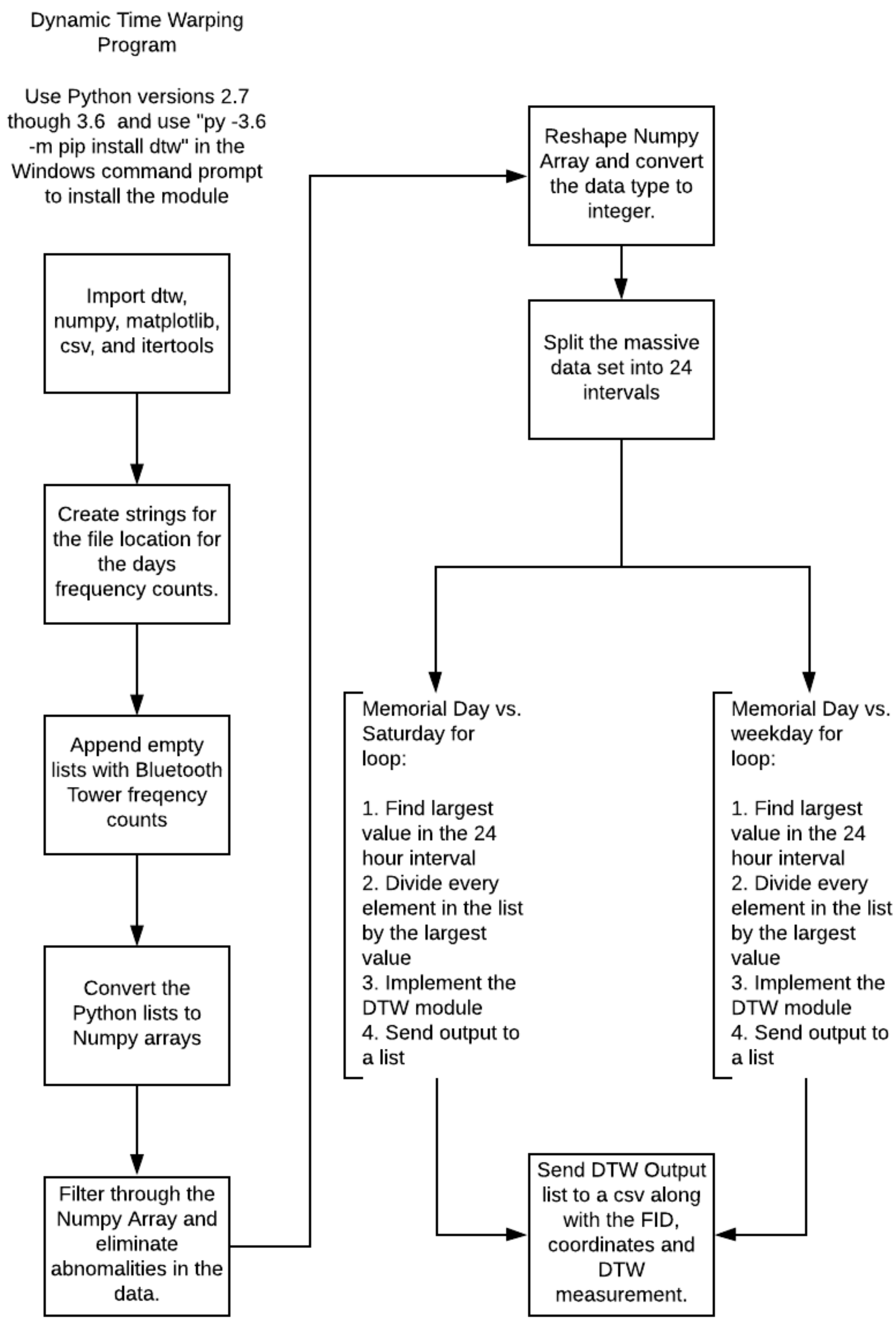

Fig. 5. Algorithm for the Python program that calculated dynamic time warping distance.

5.1 Data Extraction

The algorithm for the data extraction program is straight forward. It works by using tools provided by the R software suit to calculate the frequency of compatible cars able to connect to the Bluetooth Towers. Then it filters through the massive 8 gigabytes CABD file with a 50,000 lined interval. While the 50,000 lined intervals for loop goes through the CABD file, it counts the number of cars that pass by each Bluetooth tower during a designated amount of time. The result is a CSV file containing the FID, the street name where the Bluetooth Tower is, the count for each hour in the day, the normalized version of the same day count, the latitude, and the longitude. A CSV file stands for comma-separated value file. CSV files are used for databases and can translate into an excel program grid.

5.2 DTW Calculator Program

The following is the algorithm created for the program was developed to calculate the dynamic time warping data series for this project. Before executing or troubleshooting the program, make sure to use a version of Python between 2.7 and 3.6. To download the DTW module, use the command “py -3.6 -m pip install DTW” in the Windows command prompt. Without having the DTW module and the correct Python version, the program will not work.

To simplify the flow chart above, the DTW Calculator Program accomplishes three processes: data formatting, normalization, and DTW computation. Originally the data is not formatted in the correct arrangement for the DTW module. Therefore, NumPy is used to edit the data to where the parameters of the DTW module will accept the array. After the data processing, the hour frequency counts become normalized.

Normalizing the data is a statistical method used in this research. Normalization requires the largest number in a data set to be divided into every number including itself. The result is a set of relatable decimals that are easier to compare the quantifiable size to another data set rather than a set of random numbers. The resulting decimals are in more of a spectrum meaning they are more similar when they are closer to 0.

Once the data is normalized, the DTW module calculates the average DTW distance in the time series. A CSV file containing the results is used to generate the following maps.

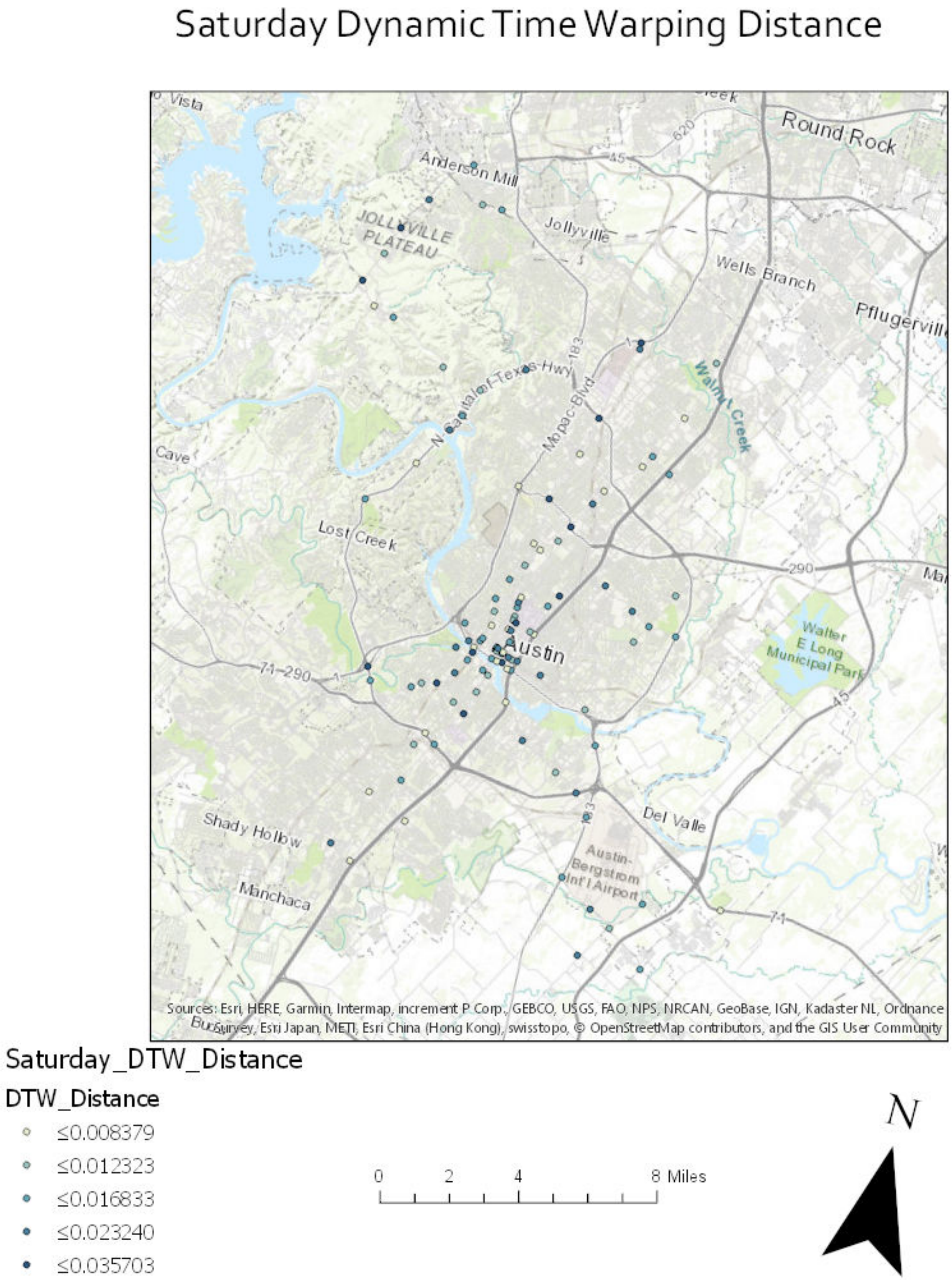

Fig. 6. DTW results for the normal Saturday and busy Memorial Day comparison.

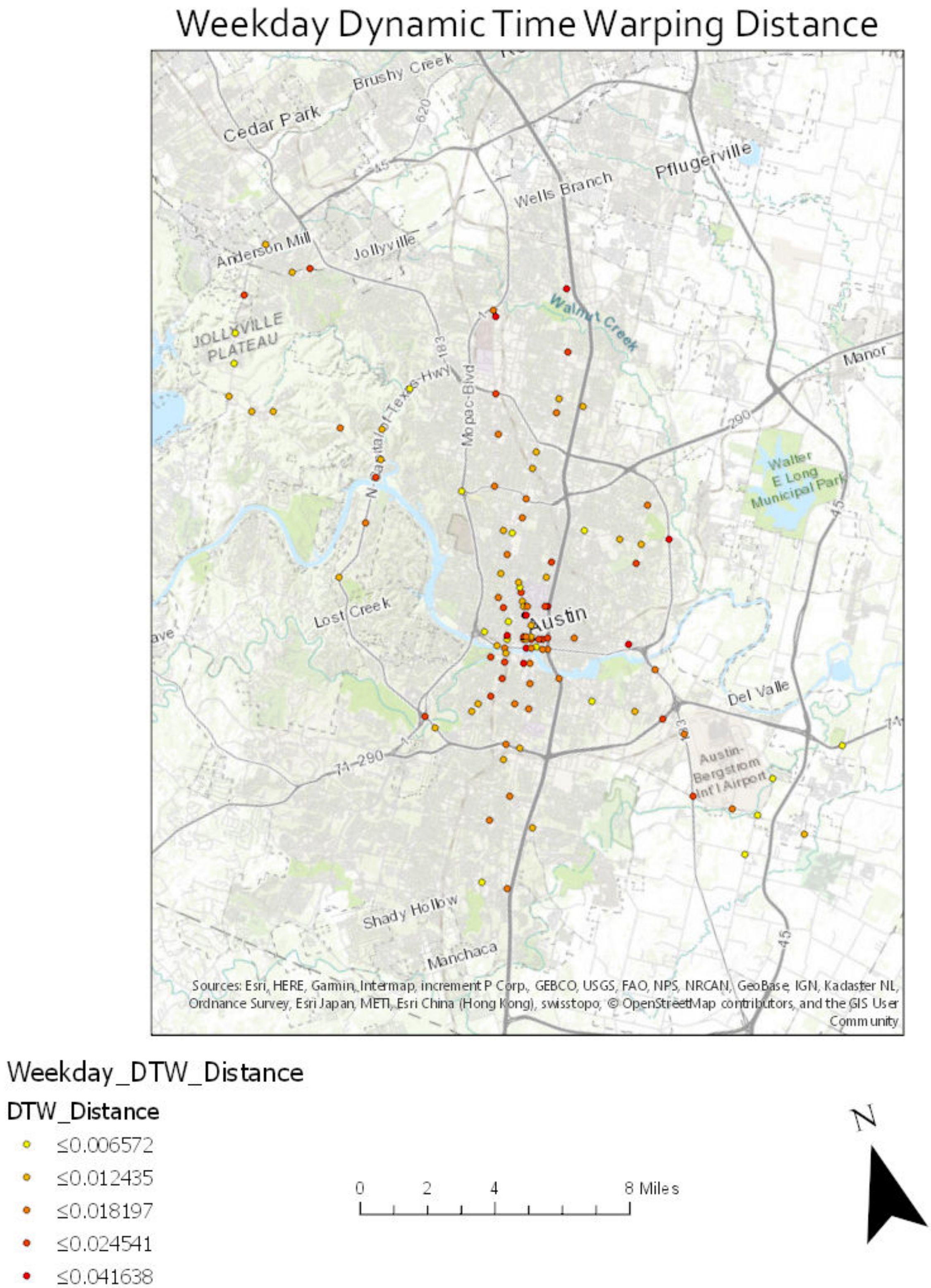

Fig. 7. DTW results for the normal weekday and busy Memorial Day comparison.

6. Results

The DTW distance measurements were underwhelming similar. The normal weekday and Saturday were like the busy traffic day. The average DTW distance for the Memorial Day and a normal weekday was 0.014901, and the average DTW for Memorial Day and the normal Saturday was 0.014719. Small differences in the measurements mean both data sets were more like each other. However, they are still slightly different.

Although the DTW is small, there is still a difference in the data. With the data normalized, the average DTW is on a scale to which the reader can understand more effectively. Combining the nature of the thousands of cars passing through Austin's highways every day and the difference in normalization; the small difference could be the result of an extra few hundred cars on the highways which causes more congestion. Plus, there is a pattern illustrating different parts of town received more traffic than other parts of town.

Figures 6 and 7 both show spatial trends in the data. The areas subject to the most change in traffic were in the center of downtown Austin, along with the highways, and around the outskirts of town close to the freeways. All these areas have a high volume of traffic, and the cause of change is the result of the mass migration to and from the city. The demand for work and entertainment drove the Austinites out of their homes and caused a slight uptick in traffic for Memorial Day weekend. Furthermore, there was an influx of visitors to visit family and to enjoy the entertainment centers like Rainy Street in Austin. Although the numbers are small, the average distance which is 0.01481 can result in hundreds of more cars on the road during Memorial Day weekend which would slow the entire city down especially if one of those newly added cars were to crash.

An example of an anomaly in the data is where the smallest dynamic time warping distance is measured. The smallest dynamic time warping distance was collected on 2nd street. However, the adjacent Bluetooth data towers recorded a large change in the volume of traffic in the surrounding area because this specific region is in the heart of downtown Austin. There are several different scenarios for this to occur. The first scenario is this area receives a constant stream of traffic throughout the day and night which would cause the DTW measurement to be low. Secondly, there might have been some construction in the area surrounding the data retrieval site. Construction on 2nd street for the latest set of high-rises could have caused traffic flow to be none existent because the construction site would have blocked off the street. Thirdly, the Bluetooth tower itself could have had a technical malfunction which would have prevented it from collecting all the data from cars which passed by.

Another pattern in figures 6 and 7 is although some parts of the outskirts of town receive great change, many Bluetooth sensors outside of the city center did not change very drastically. On the map, these Bluetooth sensors are unlikely to change because they do not exist in high traffic areas along highways. For instance, the largest dynamic time warping distance recorded close to Little Walnut Creek which was caused by the large influx of traffic the highway from rush hour traffic from suburbia. A possible study in the future might need to analyze these areas specifically within their group because the larger frequency changes in the other data sets might have drowned out the statistical reverence in outlying areas.

7. Discussion

There are several issues from the collected Bluetooth data. One major issue is some cars are not Bluetooth compatible. The average car age is 11.6 years old (Walsworth, 2016) and Bluetooth did not become standard in some car brands until 2010 (Hamel,2019). Therefore, cars made before 2010 will not have Bluetooth compatibility. Older cars might skew the data results to where the study represents a smaller number of cars than what is actually represented.

There are other methods which might be more accurate than Bluetooth monitoring. As discussed in the literature review, video monitoring is another method traffic engineers use to measure traffic flow. Video monitoring is where a camera measures the speed of cars passing by the camera. Not only does the camera capture the speed, but it can also give a more accurate count for the frequency of cars pass by each camera station. The camera count is more accurate because it is counting the physical number of cars in a given area rather than cars with compatible technology with the Bluetooth towers. The frequency yield will be much more accurate because it will be counting all the cars passing by the camera instead of cars which are Bluetooth compatible.

Bluetooth reception is not the most accurate. There are environmental factors which cause inference with the Bluetooth sensors. Some environmental factors that can block network signaling include physical objects, radio frequency interference, electrical interference, and weather conditions. Physical objects such as buildings, hills, and trees can interfere with the Bluetooth connection between the car and the Bluetooth tower. Wireless frequency interference can occur if the Bluetooth wireless signal falls within the range of 2.4GHz and 5 GHz. Networking issues are problematic because most WiFi and Bluetooth devices fall within this range. Electrical interference can occur in the car depending on the number of devices in each car. Lastly, weather conditions such as rain, fog or lightning can disrupt Bluetooth signals because the medium at which the Bluetooth signals pass through is distorted and might not reach the Bluetooth tower. The data lost through any of the events described above might cause large data inaccuracy in the future because there will be missing frequency counts.

8. Conclusion

After implementing the DTW Calculator Program, the Python script showed a slight increase in traffic on Memorial Day weekend in the city of Austin in 2016. Using the dynamic time warping module in Python, there was a quantifiable increase of 0.01481 of traffic between the two days examined and the Memorial Day Saturday. After discussing the reliability of the data, it determined the time series data was accurate and worth examining. As more cars become Bluetooth compatible, the number of devices connected to Bluetooth networks will increase which will represent a larger portion of the population rather than only the Bluetooth compatible cars. This research was one application of how urban planners can use smart cities to illuminate new insights on behavioral patterns which will make cities of the future safer and more efficient than ever before.

9. Bibliography

Brown, David Scott. 2019. “Your Car, Your Computer: ECUs and the Controller Area Network.” *Techopedia.com*, Techopedia, 2 Jan. 2017, www.techopedia.com/your-car-your-computer-ecus-and-the-controller-area-network/2/32218.

City of Austin Transportation Department. 2019. "Bluetooth Travel Sensors - Individual Address Files (IAFs) | Open Data | City of Austin Texas." *Data.austintexas.gov*, The City of Austin , 5 Oct. 2017, data.austintexas.gov/dataset/Bluetooth-Travel-Sensors-Individual-Address-Files-/qnpj-zrb9.

Chandrasekaran, Gayathri, Tam Vu, Alexander Varshavsky, Marco Gruteser, Richard P. Martin, Jie Yang, and Yingying Chen. 2011. "Tracking Vehicular Speed Variations by Warping Mobile Phone Signal Strengths." In *2011 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 213–21. Seattle, WA, USA: IEEE. https://doi.org/10.1109/PERCOM.2011.5767589.

Chen, Song, Xiaoyan Wei, Nan Xia, Zhaojin Yan, Yi Yuan, H. Michael Zhang, Manchun Li, and Liang Cheng. 2019. "Understanding Road Performance Using Online Traffic Condition Data." *Journal of Transport Geography* 74 (January): 382–94. https://doi.org/10.1016/j.jtrangeo.2018.12.004.

Cortés-Berrueco, Luis E., Carlos Gershenson, and Christopher R. Stephens. 2016. "Traffic Games: Modeling Freeway Traffic with Game Theory." Edited by Natalia L. Komarova. *PLOS ONE* 11 (11): e0165381. https://doi.org/10.1371/journal.pone.0165381.

Deville, Pierre, Catherine Linard, Samuel Martin, Marius Gilbert, Forrest R. Stevens, Andrea E. Gaughan, Vincent D. Blondel, and Andrew J. Tatem. 2014. "Dynamic Population Mapping Using Mobile Phone Data." *Proceedings of the National Academy of Sciences* 111 (45): 15888–93. https://doi.org/10.1073/pnas.1408439111.

Elsworth, S. 2016. "Friday 29th September, 2017," 45.

Hamel, James. "What Cars Have Bluetooth Capabilities?" *What Cars Have Bluetooth Capabilities?*, It Still Runs, 10 Jan. 2019, itstillruns.com/cars-bluetooth-capabilities-5669426.html.

Han, Su Yeon, Ming-Hsiang Tsou, and Keith C. Clarke. 2015. "Do Global Cities Enable Global Views? Using Twitter to Quantify the Level of Geographical Awareness of U.S. Cities." Edited by Alejandro Raul Hernandez Montoya. *PLOS ONE* 10 (7): e0132464. https://doi.org/10.1371/journal.pone.0132464.

Javed, Muhammad Awais, Sherali Zeadally, and Elyes Ben Hamida. 2019. "Data Analytics for Cooperative Intelligent Transport Systems." *Vehicular Communications* 15 (January): 63–72. https://doi.org/10.1016/j.vehcom.2018.10.004.

Krco, Srdjan. 2019. "Bluetooth Based Wireless Sensor Networks –Implementation Issues and Solutions," 7.

Laharotte, Pierre-Antoine, Romain Billot, Etienne Come, Latifa Oukhellou, Alfredo Nantes, and Nour-Eddin El Faouzi. 2015. "Spatiotemporal Analysis of Bluetooth Data: Application to a Large Urban Network." *IEEE Transactions on Intelligent Transportation Systems* 16 (3): 1439–48. https://doi.org/10.1109/TITS.2014.2367165.

Nawaz, Sarfraz, and Cecilia Mascolo. 2014. "Mining Users' Significant Driving Routes with Low-Power Sensors." *ArXiv:1409.2264 [Cs]*, September. http://arxiv.org/abs/1409.2264.

Kittipong HIRI-O-TAPPA, Setha PAN-NGUM , Wasan PATTARA-ATIKOM , and Sorawit NARUPIT . 2010. "(PDF) A Novel Approach of Dynamic Time Warping for Traffic Congestion Detection and Short-Term Prediction." n.d. Accessed March 4, 2019. https://www.researchgate.net/publication/279951318_A_Novel_Approach_of_Dynamic_Time_Warping_for_Traffic_Congestion_Detection_and_Short-Term_Prediction.

Shafie, A.A. & Ali, Md & Zaman, Fadhlan & M. ALI, ROSLIZAR. (2011). "(PDF) Smart Video Surveillance System for Vehicle Detection and Traffic Flow Control." n.d. Accessed March 4, 2019. https://www.researchgate.net/publication/289016817_Smart_video_surveillance_system_for_vehicle_detection_and_traffic_flow_control.

U.S. Department of Transportation. 2018. "Traffic Info.Pdf." n.d.

Wilson, Daniel. 2018. "Using Machine Learning to Predict Car Accident Risk." n.d., 19.

Veiga, Adonay A., and Claudia J. B. Abbas. 2019. "Proposal and Application of Bluetooth Mesh Profile for Smart Cities' Services." *Smart Cities* 2 (1): 1–19. https://doi.org/10.3390/smartcities2010001.

VoPham, Trang, Jaime E. Hart, Francine Laden, and Yao-Yi Chiang. 2018. "Emerging Trends in Geospatial Artificial Intelligence (GeoAI): Potential Applications for Environmental Epidemiology." *Environmental Health* 17 (1). https://doi.org/10.1186/s12940-018-0386-x.

Walsworth, Jack. “Average Age of Vehicles on Road Hits 11.6 Years.” *Average Age of Vehicles on Road Hits 11.6 Years*, Automotive News, 22 Nov. 2016, www.autonews.com/article/20161122/RETAIL05/161129973/average-age-of-vehicles-on-road-hits-11-6-years.

Wu, Chao, Xinyue Ye, Fu Ren, You Wan, Pengfei Ning, and Qingyun Du. 2016. “Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China.” Edited by Kim-Kwang Raymond Choo. *PLOS ONE* 11 (10): e0164553. https://doi.org/10.1371/journal.pone.0164553.

Yang, Su, Shixiong Shi, Xiaobing Hu, and Minjie Wang. 2015. “Spatiotemporal Context Awareness for Urban Traffic Modeling and Prediction: Sparse Representation Based Variable Selection.” Edited by Wen-Bo Du. *PLOS ONE* 10 (10): e0141223. https://doi.org/10.1371/journal.pone.0141223.

Yao, Yaxing. *Unpacking People's Understandings of Bluetooth Beacon Systems - A Location-Based IoT Technology*, ScholarSpace, 1 Jan. 1993, dblp.uni-trier.de/pers/y/Yao:Yaxing.