2021 NOCoE Transportation Technology Tournament

Interactive Data Platform for Non-recurring Congestion Detection and Analysis

NOCoE National Operations Center of Excellence

Mizzou University of Missouri

SAINT LOUIS UNIVERSITY

UMSL

MoDOT

1	Report Number	TTT2021-06
2	Designed Agency	Missouri Department of Transportation
3	Sponsoring Agency	Transportation Technology Tournament, sponsored by the National Operations Center of Excellence (NOCoE) and the U.S. DOT ITS JPO PCB program
4	Type of Report	Final - Concept of Operation
5	Project Title	Interactive Data Platform for Non-recurring Congestion Detection and Analysis
6	Performing Organization Name and Address	University of Missouri - Columbia (Mizzou) | Columbia, MO 65211 Saint Louis University (SLU) | 1 N Grand Blvd, St. Louis, MO 63103 University of Missouri - St. Louis (UMSL) | 1 University Blvd, St. Louis, MO 63121
7	Primary Contact	Alex Wassman, MoDOT, alexander.wassman@modot.mo.gov Eric Kopinski, MoDOT, eric.kopinski@modot.mo.gov
8	Faculty Advisors	Yaw Adu-Gyamfi, Mizzou, adugyamfiy@missouri.edu Jalil Kianfar, SLU, jalil.kianfar@slu.edu
9	Additional Contribution	Linlin Zhang, Mizzou, lz5f2@mail.missouri.edu Mustafa Attallah, SLU, mustafa.attallah@slu.edu Xiang Yu, Mizzou, xytm4@mail.missouri.edu Esther J. Robinson, UMSL, erobinson@cbbtraffic.com
10	Submission Date	June 13, 2021
11	Abstract	The rapid detection of non-recurring congestion is necessary since it will effectively mitigate and eliminate some potential risks to traffic flows and roadways. In Missouri, since Interstate 270 (I-270) is one of the most heavily traveled interstates, has higher severe crash rates than average, and is majorly impacted by non-recurring congestion, the I-270 northbound traveling direction is selected as the mainly research object. The research team will focus on developing a neural network model to detect, identify, and predict the non-recurring congestion along the I-270 northbound traveling direction using the historical probe and Waze incident data. The research team also creates an interactive data platform to provide real-time non-recurring congestion locations, traffic flow information, roadway geometry, and CCTV videos to users.
12	Number of Pages	Concept of Operations: 15 Pages | Additional Resources: 1 Pages

Table of Contents

List of Figures

List of Tables

1 Overview

1.1 Background and Motivation

Safe and reliable transportation networks are important in everyday life. Everyone is affected by the function of the transportation system, especially frequently used roads and highways. Congestion, recurring and non-recurring, can have a major impact on the safety and efficiency of a roadway. According to the American Association of State Highway and Transportation Officials (AASHTO), non-recurring traffic incidents are about a fourth of the traffic congestion in the United States. When non-recurring congestion occurs, it can make managing the network more challenging and cause safety hazards. Non-recurring congestion can be caused by accidents, construction, special events, weather conditions, etc. When this happens, the event must be noticed and action must be taken to make drivers aware of the incident and to provide a solution to minimize impacts to the system. When drivers are not expecting congestion, they may not be prepared to reduce their speed or stop, especially at higher speeds. The ability to detect where there is non-recurring congestion and notify drivers quickly and effectively is crucial to the safety of all users on the road. This would require a system that can communicate and respond promptly and appropriately.

One example of a frequently used roadway that is majorly impacted by non-recurring congestion is Interstate 270 (I-270). In Missouri, I-270 is one of the most heavily traveled interstates. About 140,000 vehicles use it daily. Both directions of I-270 have reported having rates of severe crashes that are higher than average. The study area of I-270 is shown in Figure 1.

Figure 1. Project Study Area I-270 Northbound Direction

1.2 Project Objectives

Due to the irregularity of non-recurring congestion and the severe impact of the transportation system it brings, MoDOT shows interest in it. This project aims to use historical and real-time traffic conditions data to detect, analyze and provide visualization of the non-recurring congestion on I-270 northbound. Thus, the objectives of the project will include:

(1). Integrate or fuse different data sources including Roadway Geometry, Probe, CCTV, and Incident data such as crashes, work zones, or breakdowns.

(2). Develop a real-time database for storing data.

(3). Develop a procedure for detecting non-recurring congestion from given data.

(4). Build an interactive web user interface for visualizing data and analyze non-recurring congestion events.

1.3 Approach

The rapid detection of non-recurring congestion is necessary since it will effectively mitigate and eliminate some potential risks to traffic flows and roadways. To address the objectives, two general approaches are proposed as below:

(1). In the real world, non-recurring congestion is normally caused by unexpected events and conditions, including road construction, adverse weather conditions, vehicular crashes, and so on. In this project, the historical probe and Waze incident data, with the incident types listed above, are used to build a neural network model to predict and identify the non-recurring congestion.

(2). An interactive web platform is developed to provide real-time roadway geometry, speed, and real-time CCTV. To perform a clear visualization, the detailed information of each road segment can be shown by zooming and clicking on the map. In addition, the predicted non-recurring congestion segment is shown on the map. Moreover, a real-time database server, firebase, is utilized to store the dataset during the whole procedure.

1.4 Stakeholders

The stakeholders in this project include:

(1). Designer: The designer agency in this case is the University of Missouri and Saint Louis University team. The team is the leader of this project by identifying the problem, establishing an efficient solution, and generalizing the solution to be used by the users of I-270.

(2). MoDOT: The main goal of the project is to improve the performance of the I-270 which is maintained by MoDOT. MoDOT plans and conducts maintenance activities of this corridor, has a fleet of roadway incident response trucks for this corridor, and utilizes CCTV cameras and traffic sensors to monitor traffic on this corridor.

(3). Municipalities on I-270: I-270 NB is crossing two counties on its way to the north. Saint Louis County and Saint Louis City are benefiting from I-270 before it reaches the

Mississippi River and crosses Illinois. The I-270 frontage roads are impacted by non-recurring congestion on I-270 and are maintained by local municipalities along this corridor.

(4). Missouri State Highway Patrol and Local Police Departments: The police law enforcement agencies are responsible for the instant action taken whenever an incident happens on I-270. Non-recurring congestion might lead to crashes that require a quick response or preparation.

(5). Users: The drivers are the targeted stakeholders and for them, the interstate is being monitored, studied, and improved. Efficient communication with drivers is very important in order to notify them about the current and future state of the road. Good communication with drivers improves the performance of the road and thus making the project reach the supreme objectives. Providing information about non-recurring congestion may improve roadway safety by preventing some secondary crashes.

2 Interactive Data Platform for Non-recurring Congestion Detection and Analysis

2.1 Problem Definition

Compared to recurring congestion, non-recurring congestion events are unpredictable and could dramatically decrease the reliability of the transportation system. Therefore, rapid detection, visualization, and analysis of non-recurring traffic congestion strategies are needed to reduce its impact on the entire transportation system. The goal of the project is to develop a web framework to quickly detect, analyze and communicate non-recurring congestion by fusing data from multiple sources. The web user interface assists Transportation Management Center (TMC) staff with implementing emergency response and congestion mitigation strategies and could be integrated with transit agencies and first responders dispatch and routing centers.

2.2 Description of Solution

2.2.1 Description of Data

As described in one of the objectives, the project aims to integrate data from multiple sources, including the probe vehicle data, Waze incident data, and real-time CCTV data.

The attributes of the probe data, which is provided by HERE, a commercial vendor, involves the vehicle speed, average speed, reference speed, travel time, and confidence of the specific road segments aggregating every five minutes from 2020 to 2021. To specify the potential congestion and identify the non-recurring congestion, a real-time transportation database, Waze incident, is used in this project. The Waze incident provides the real-time incident data in 2021 that contains the traffic information- confidence and reliability, incident information- incident type, incident subtype- and explicit geographical information- country, city, street, direction, latitude, longitude, and so on. In order to give the website platform users real-time traffic conditions along

I-270, MoDOT provides several camera locations and links that could show the live-view real-time CCTV.

2.2.2 Modeling

This project is aiming to detect non-recurring congestion in a timely manner, and inform stakeholders about roadway events. That is helpful in taking proactive actions by all the involved agencies in order to improve the performance of the road which is leading to efficient transport service. The northbound of I-270 is covered by 25 TMC stations on a length of 30 miles. These TMC segments collect speed and travel time at five-minute intervals. In order to get the aimed goal of detecting the non-recurring congestion, a model is built at each location. The model predicts the speed five minutes ahead by using the current and previous speed at all 25 locations. Specifying the locations that influence the speed at a specific location requires further analysis and that requires more time. Therefore, the 25 locations are considered to define the next five-minute speed at a specific location.

The probe data from 2020 used to train and evaluate the models. To capture all the trends in the data, a random split of 60% and 40% training and testing datasets are used. In this case, the probability of having a biased model is very low and that helps to make more reliable predictions. The entire 2021 datasets are used to identify the non-recurring congestion by using the 2020 built models. K-fold cross-validation is utilized to optimize the efficiency of the model by splitting the training dataset into 10 folds. In k-folds, the training dataset is divided into 10 equal parts. Figure 2 below shows how the k-fold method works.

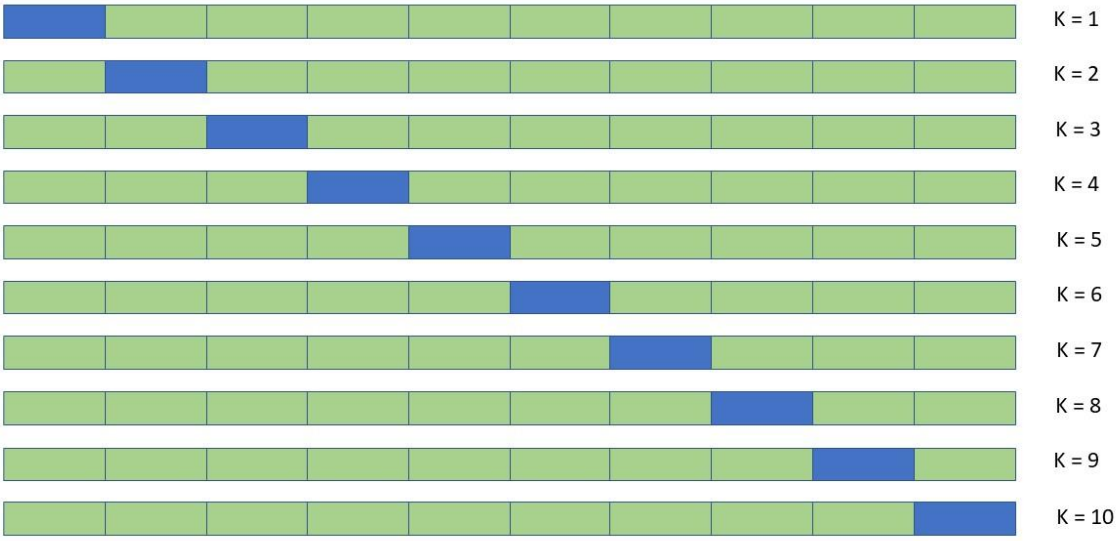

Figure 2. 10-fold Cross-validation

The first line represents the first k test which means all the green parts are used to train and the blue part of the training data is used to evaluate. In the second line, the blue part is left out for

evaluation and all the other parts are used to train the model. In the end, the line that shows the minimum error is utilized.

AvNNet neural networks are used to build the models. This algorithm is efficient and could dramatically minimize the errors by using the model averaging. Figure 3 shows the max and minimum error measures as well as the average error measures for the entire segment.

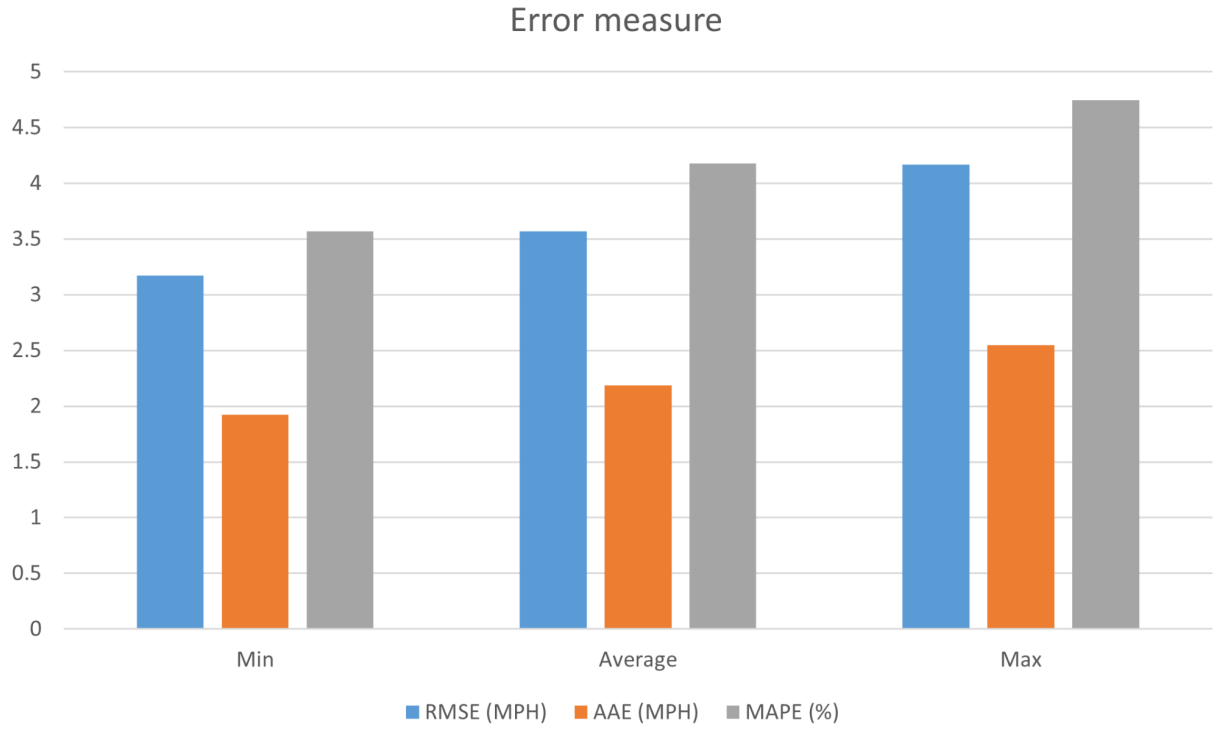

Figure 3. Error measures of the speed prediction model

The general modeling process is a two-step process, first, model the speed and minimize the errors to predict the next five-minute speed on a specific location. Second, the predicted speed is compared to the speed limit at that location to identify the type of congestion as explained in section 2.2.3.

2.2.3 Identify Non-recurring Congestion

After predicting the vehicle speed by the avNNet model of each segment in I-270 in 2021, the column of the speed difference, which represents the difference between predicted speed and the reference speed is created. Since the reference speed is defined as the free-flow speed, which indicates the speed with no vehicle on the road, the big gap of speed difference specifies the slow traffic. For the vehicle whose speed is greater than the reference speed, the speed difference is set as zero. By matching the time of predicted speed difference in 2021 and the Waze incident dataset with selected incident type, which could lead to non-recurring congestion, the threshold of 10 mph in speed difference is used to detect the non-recurring congestion. Since the Waze

incident data we have are from Jan 1st to Feb 21st in 2021, after matching by period, the total incident records are 1275. The incident types- hazard on road construction, hazard on road car stopped, jam caused by heavy traffic, jam caused by standstill traffic, major accident, minor accident are considered to lead to non-recurring congestion. Then, 209 of the 1275 incident records are selected. After filtering the speed difference by 10 mph, 54 of them, which have a big speed difference, are considered as the non-recurring congestion.

Thus, in this project, there are 54 incident records, which represent 52 individual geometrical points that distribute in 16 TMC segments in I-270, are considered as non-recurring congestions during the Jan 1st to Feb 21st in 2021, by using the avNNet model.

Figure 4 and figure 5 indicate the distribution of congestion type, and hour of all the non-recurring congestions from Jan 1st to Feb 21st, 2021 respectively.

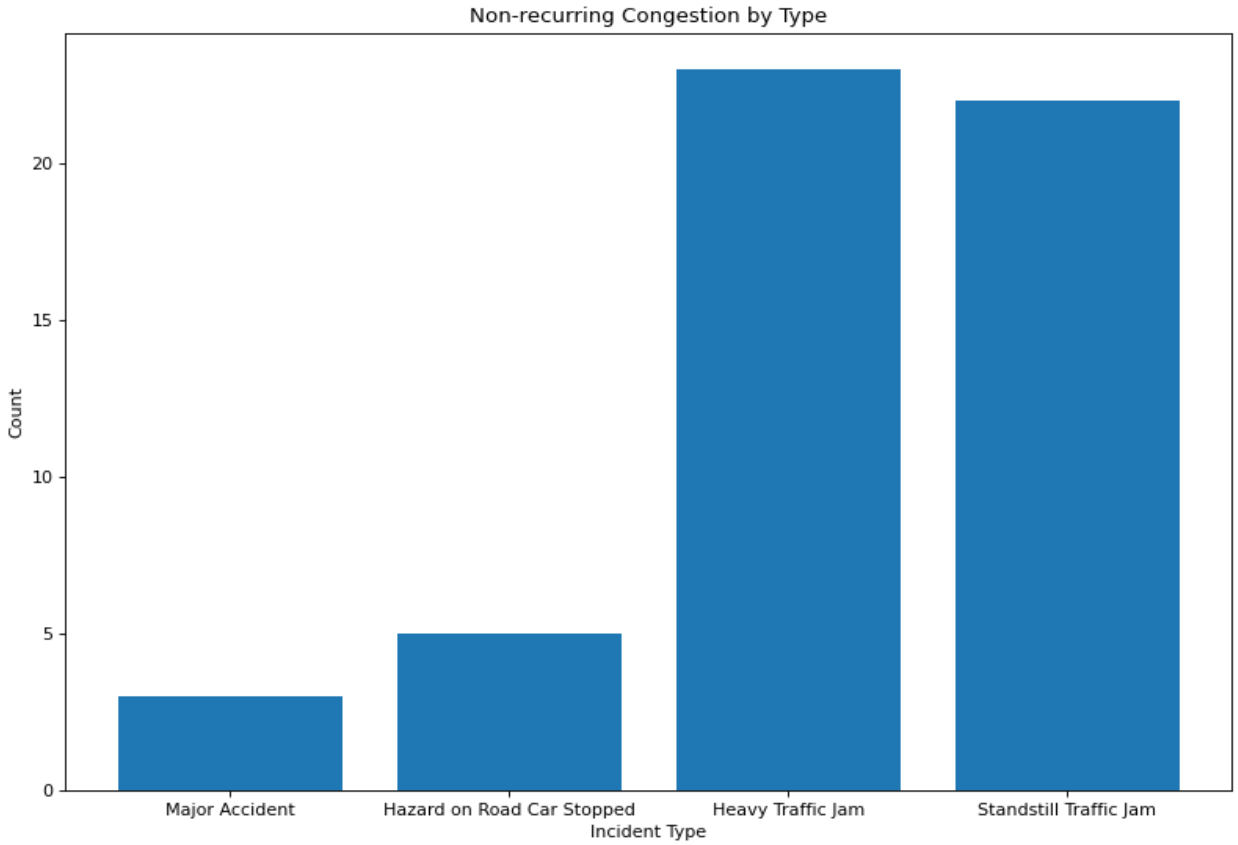

Figure 4. Non-recurring Congestion Distribution by Type

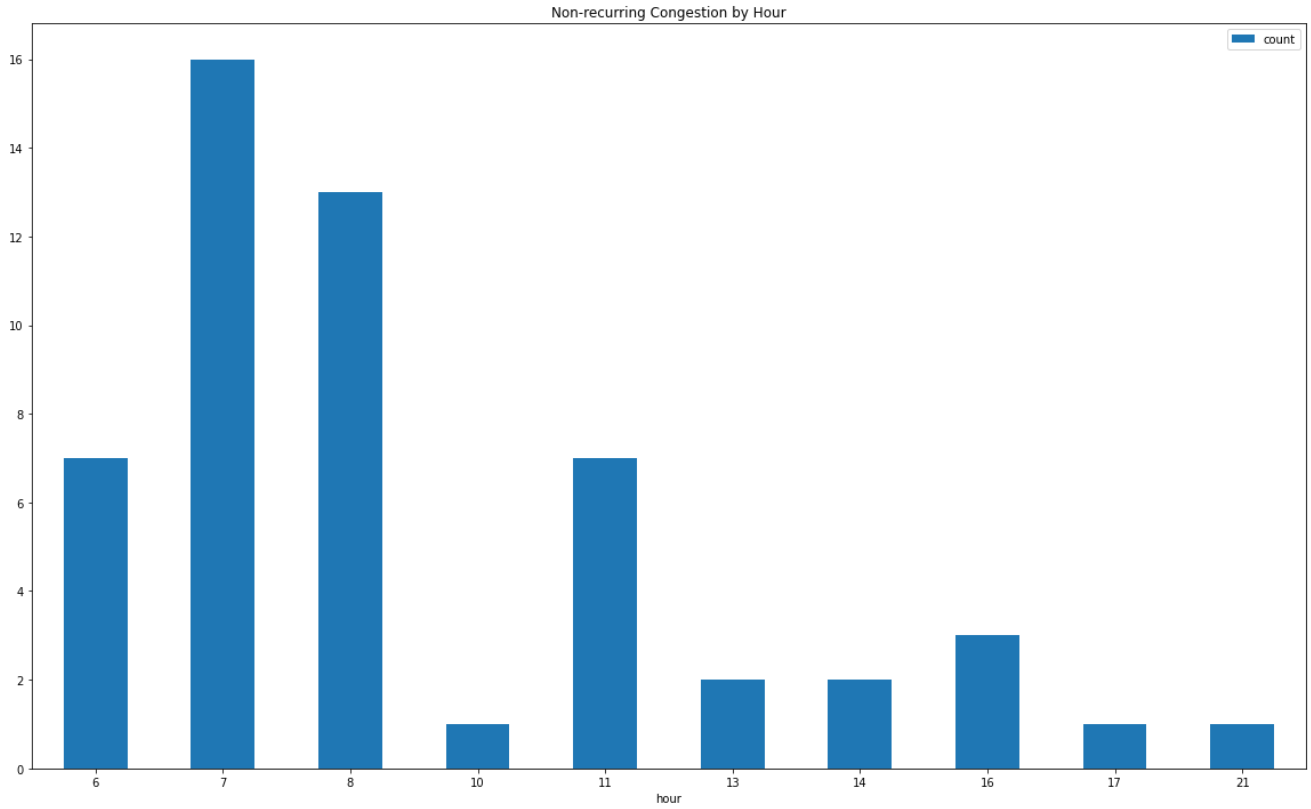

Figure 5. Non-recurring Congestion by Hour

From figure 4, the main congestion types that caused the non-recurring congestion are heavy traffic jams and standstill traffic jams, while only a few cases caused by major accidents, and hazard on road cars stopped. In figure 5, most non-recurring congestion occurred in the morning rush hour from 6 am to 8 am. In addition, a lot of cases occurred during the lunch break time around 11 am.

2.2.4 Website Development

One of the objectives of this project is to create a web framework to quickly detect, analyze, and communicate non-recurring congestion by integrating data from multiple sources including roadway geometry, probe, CCTV, and incident data. Since it is an interactive real-time platform, a database is needed for storing real-time data. In this project, the Firebase is used.

For the interactive web development, several layers were designed with real-time data from the Firebase, including the CCTV camera layer, speed layer, and congestion condition layer. The CCTV camera locations (latitude & longitude) and API information were provided by MoDOT. On the website, the camera icons' locations represent the real-world cameras locations and a total of 10 locations were marked along I-270. The users could click the camera icons to know the real-time traffic conditions.

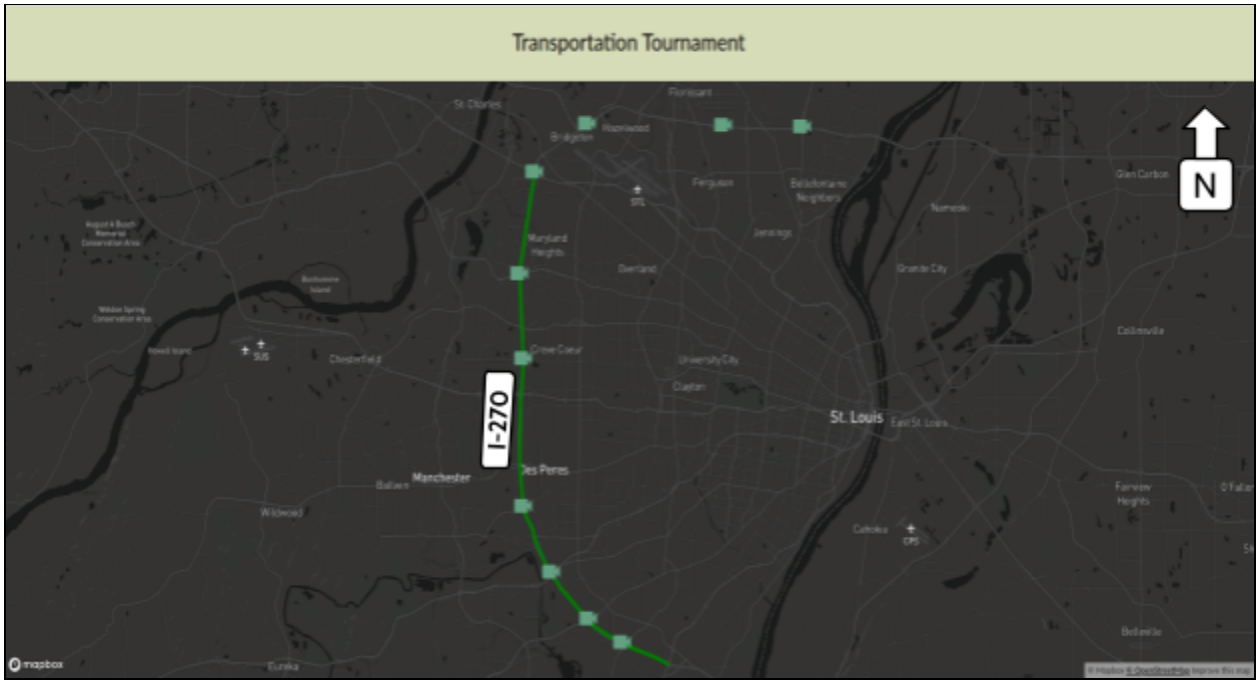

Figure 6. Website Interface

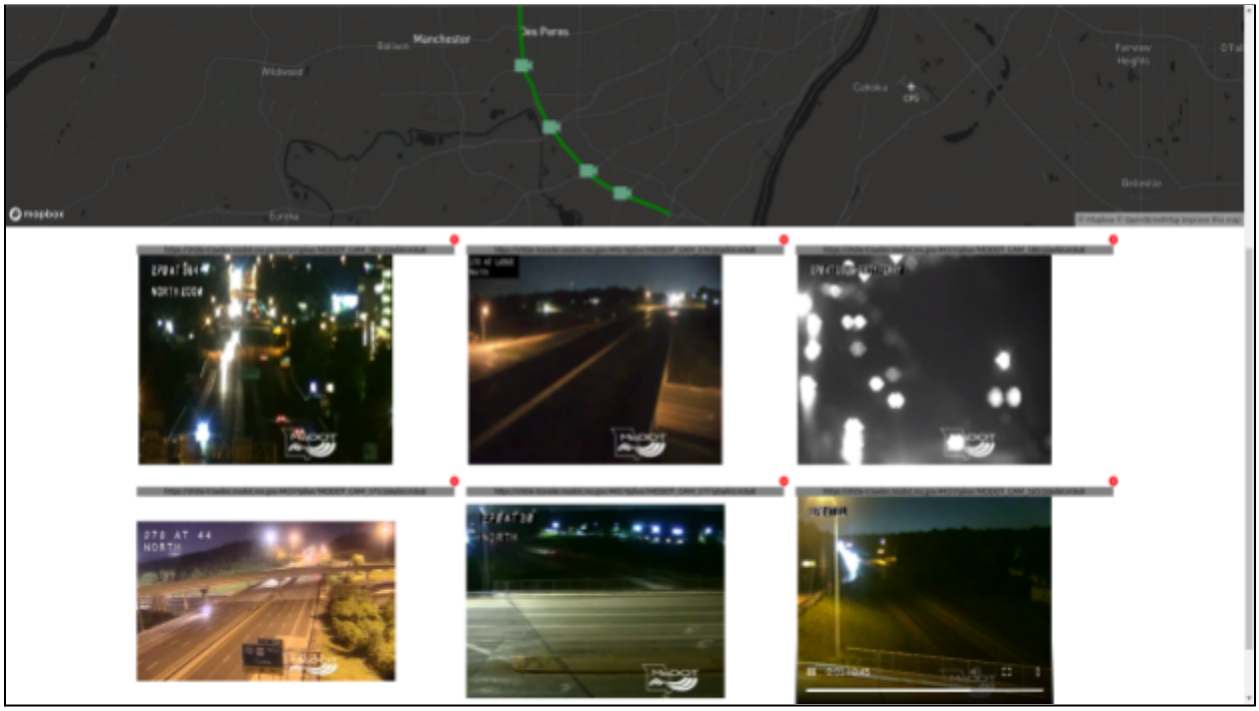

Figure 7. Real-time CCTV

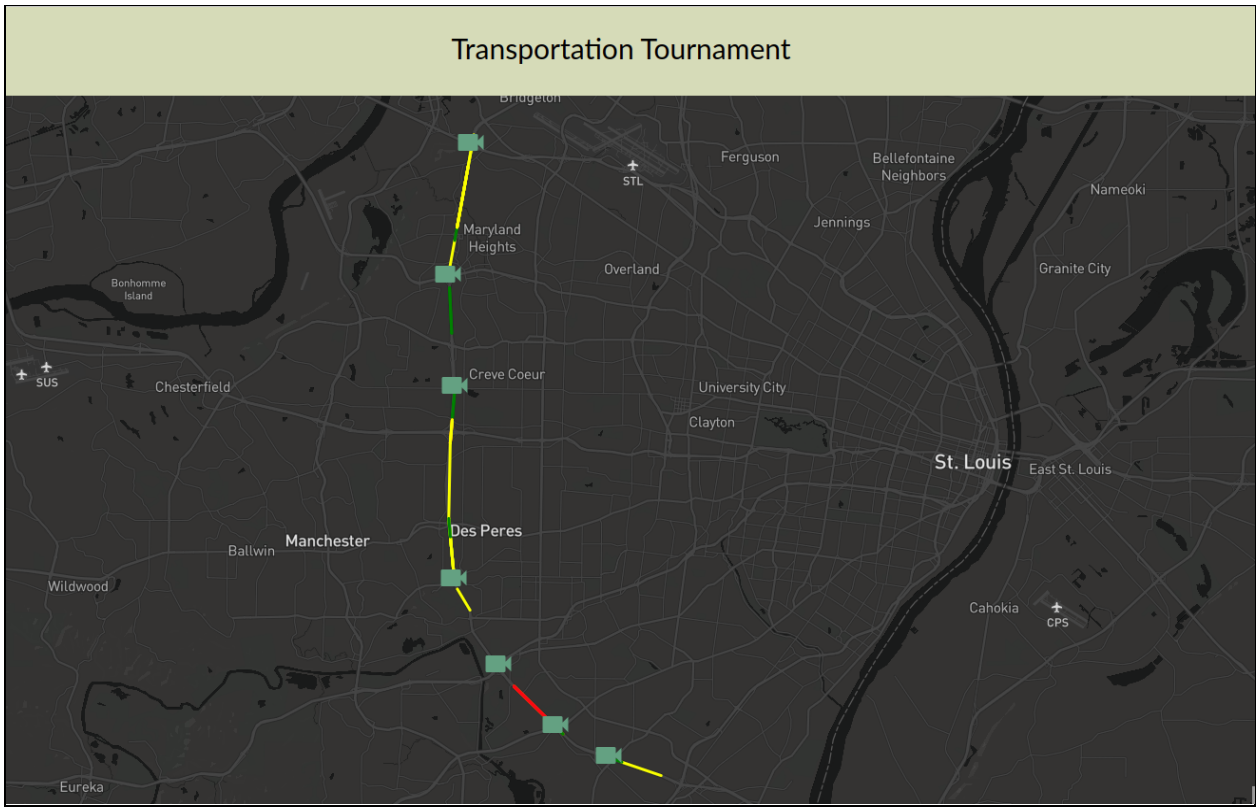

Figure 8. Non-recurring Congestion in Jan and Feb 2021

The green line in figure 6 indicates this project's research area. Figure 7 shows the real-time camera videos that fused from the firebase. As described in 2.2.3, a total of 16 TMC segments that are identified as non-recurring congestion from Jan 1st to Feb 21st in 2021 are colored in figure 8. Based on the speed difference, the congestion level is defined and colored on the map, with the thresholds of 15 mph and 20 mph, and corresponding colors of green, yellow, and red.

2.3 Functional Architecture

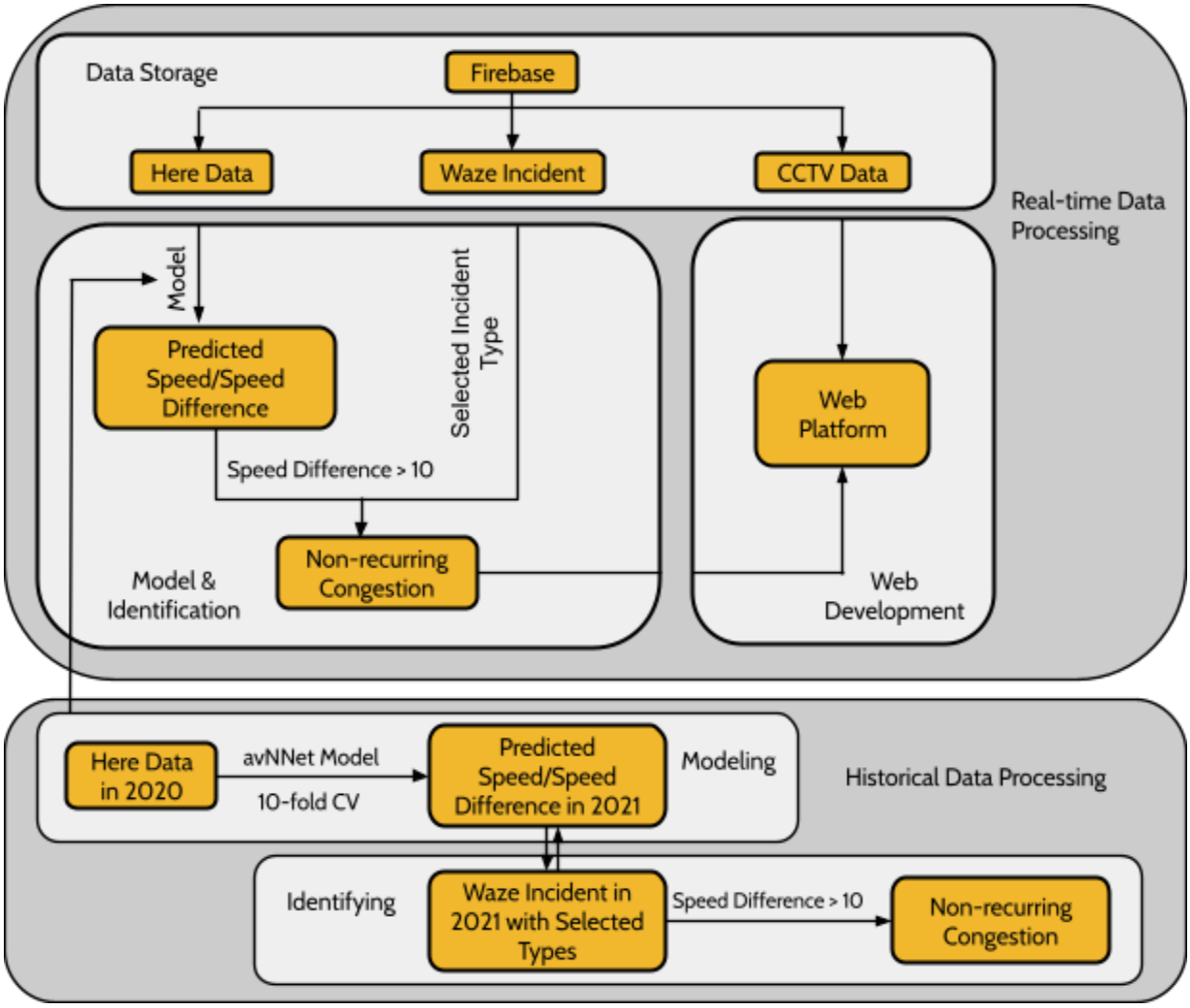

Figure 9. High-level Functional Architecture

The functional architecture, as shown in Figure 9, illustrates how the system operates. The whole system can be divided into two sections- the historical data processing section, and the real-time data processing section.

In the real-time data processing section, to integrate data from multiple sources, all the real-time data, HERE, Waze incident, and CCTV data, are stored and fused from firebase. Based on the fixed avNNet model, which is conducted by using historical data in 2020 and 2021, the predicted speed and speed difference, the difference of predicted speed and reference speed, of the next five-min interval are obtained. Since the large speed difference indicates slow traffic, the predicted speed difference that is greater than 10 mph is selected in this project to represent the slow traffic that may cause congestion. At the same time, the Waze incident data with selected types, which could lead to non-recurring congestion, are also fused from the firebase. As a result,

in the same five-min interval, the records that meet both two conditions are defined as non-recurring congestion. Then, these non-recurring congestion points are conveyed and colored on the real-time map of the Web platform with the real-time CCTV camera data fused from firebase.

The historical data processing section demonstrates how the model is built and the non-recurring congestion is identified by using historical HERE data in 2020, and Waze data in 2021. The HERE data in 2020 is used to create the model and predict the speed and speed difference of the next five-min interval of each segment in 1-270 in 2021. By matching the Waze incident in 2021 with selected incident types in the same period, the records with speed difference that greater than 10 mph are selected as the non-recurring congestion. In this paper, all the results of non-recurring congestion shown in the following chapter occurred from Jan 1st to Feb 21st in 2021.

2.4 Physical Architecture

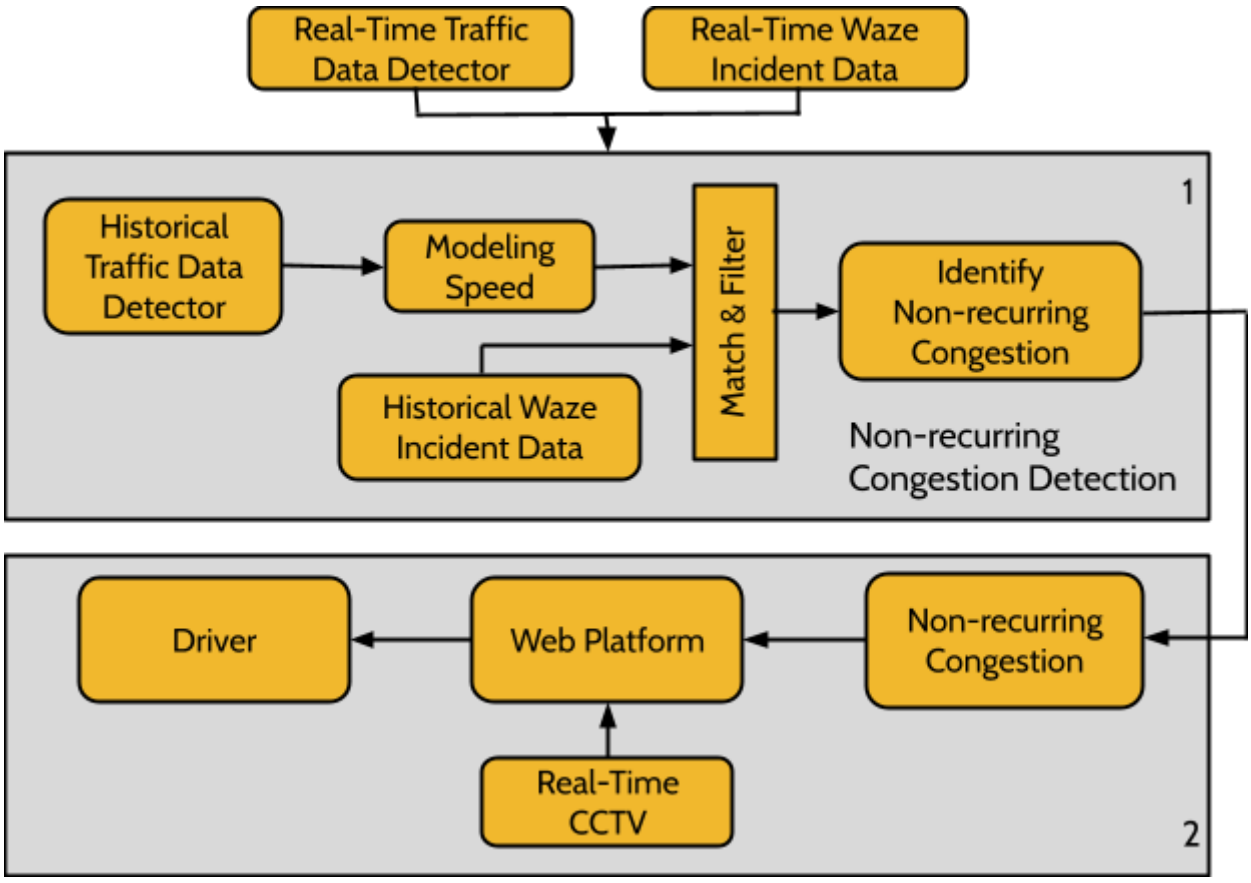

Figure 10. Physical Architecture

Figure 10. above shows the physical architecture of the project and it is majorly divided into two sections:

(1). Historic data collection, data modeling, and real-time data processing.

(2). Real-time road data sharing: These steps start when the real-time data starts to be processed to identify the non-recurring congestion. The second step is to share the result on a website that the users can use.

2.5 Enterprise Architecture

Figure 11. shows the Enterprise Architecture of this project. There are a total of four layers, business layer, data layer, application layer, and technology layer. In each layer, all used tools details are listed.

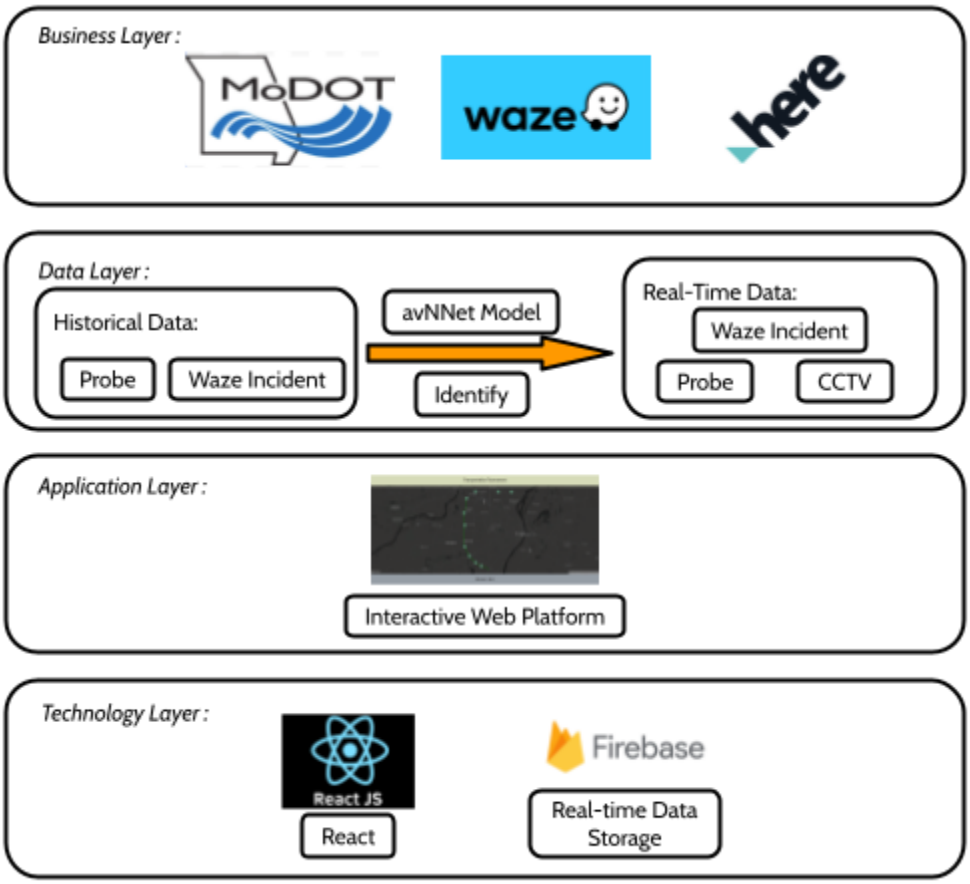

Figure 11. Enterprise Architecture

2.6 Cost and Timeline

The cost and timeline estimates for the development of the system are listed in Table 1 and Table 2.

Table 1. Timeline Estimate

Timeline-2021																				
Tasks	Start Date	End Date	Duration	7-Feb	14-Feb	21-Feb	28-Feb	7-Mar	14-Mar	21-Mar	28-Mar	4-Apr	11-Apr	18-Apr	25-Apr	2-May	9-May	16-May	23-May	30-May	June-6
Identify & Meet Stakeholders	2/7/2021	2/14/2021	1 weeks																		
Define Problem	2/7/2021	2/14/2021	1 weeks																		
Assess Needs	2/14/2021	3/28/2021	6 weeks																		
Training	3/28/2021	5/16/2021	7 weeks																		
Modeling	5/28/2021	6/3/2021	1 week																		
Website Development	6/1/2021	6/13/2021	2 weeks																		
Fuse Data from Firebase	6/1/2021	6/13/2021	2 weeks																		
Identify Non-Recurring Congestion	6/3/2021	6/10/2021	1 week																		

Table 2. Cost Estimate

Task	**Cost Estimate**
Identify & Meet Stakeholders	Personnel Cost
Define Problem	Personnel Cost
Access Needs	Personnel Cost Cost for getting access and downing the HERE, Waze incident, and CCTV data
Training	Personnel Cost
Modeling	Personnel Cost (Data Scientist) Cloud server cost (about $2 per hour for 2 weeks) GIS License (about $3,000 per year)
Identify Non-recurring Congestion	Personnel Cost Python is available for free
Website Development	Personnel or Developer Cost Cloud Server (about $120 per month) React is available for free
Fuse Data from Firebase	Personnel Cost Firebase is available for free

2.7 Benefits and Limitation

Creating an interactive web user interface will have several benefits. One of the most important benefits is having quick and effective communication to drivers, which will greatly improve safety. When a driver is unaware of an occurring incident, they may be driving too fast for the upcoming queue, especially at interstate speeds. Giving drivers ample warning will prepare them for the congestion ahead and will give them the necessary time to slow down or stop when needed. Unaware drivers who are traveling at higher speeds also can create safety hazards for law enforcement or roadside assistance who are working to resolve the incident. If there is a clear report of the situation, drivers can be more cautious when traveling through that area. Additionally, when drivers do not expect a delay, they may drive more aggressively or dangerously. Being able to decrease delays will help prevent aggressive or dangerous driver behavior.

From an operational perspective, the faster the appropriate action is taken, the less delay will occur. This will not only help the vehicles on the congested roadway but surrounding roads too. The reason for this is when there is non-recurring congestion on a frequently used road like I-270, people use surrounding routes that usually cannot accommodate the additional volume. Additionally, mobility is impacted in the same way. If I-270 was closed for a crash, then people will have fewer options to get to their designations. The travel delay can have other harmful impacts like increased emissions and increased vehicle maintenance. The economy is affected by an increase of delay as well. People can be late for work and the shipping of goods can be slowed down. Some delays can not be avoided when a crash or unexpected incident occurs but being able to resolve the issue efficiently can reduce these delays and negative impacts.

2.8 Future Work

In terms of the reasons that cause the non-recurring congestion, other data sources, for example, the historical and real-time weather data, scheduled event data, road construction data could be involved in future research, in making better detection for the non-recurring congestion by improving the speed models and identification process.

3 Conclusion

This document presents a high-level concept of operation for detecting, analyzing, and visualizing the non-recurring congestion in I-270 northbound. In this document, a two-step approach is proposed. According to this approach, the non-recurring congestion detection, analyzing, and visualizing is divided into two processes- using historical data and using real-time data.

By using historical probe data in 2020, the avNNet model is built with the acceptable error of speed prediction in 2021 on the studied segments. In addition, cross-validation with 10-folds strengthens the model by avoiding overfitting problems. Then, the non-recurring congestion is

identified with a big speed difference and selected incident type by using the historical Waze incident data in the same period, and is labeled and colored in the web platform.

Based on the same solution method, the web platform user will get a quick, accurate, dynamic, and up-to-the-minute predicted result of non-recurring congestion in I-270 northbound using the real-time probe, Waze incident, and CCTV camera data which are fused from the real-time data storage server- firebase.

Reference

Dakota, N. (2010). American Association of State Highway and Transportation Officials.

Reducing Non-Recurring Congestion. FHWA Operations - Reducing Non-Recurring Congestion. https://ops.fhwa.dot.gov/program_areas/reduce-non-cong.htm.

ArcGIS Desktop Pricing and License Levels: Buy Now. and License Levels | Buy Now. (2021). https://www.esri.com/en-us/arcgis/products/arcgis-desktop/buy.