DECISION SUPPORT TOOL FOR PATTERN RECOGNITION TO SUPPORT SIGNAL CONTROL

2023 TRANSPORTATION TECHNOLOGY TOURNAMENT

NATIONAL OPERATIONS CENTER OF EXCELLENCE

Title Decision Support Tool for Pattern Recognition to Support Signal Control		**Report Date** May 26, 2023
Team Name	Intelligent Transport Titans	
Team Members	Mahmud Hasan Mamun	Ph.D. Candidate, Florida International University
	Atika Jabin	
	Kamar Amine	
	Samia Akter	
	Daniel Frolich	
	Jasmine Hadi	Undergraduate Student, Florida Atlantic University
Academic Advisor	Mohammed Hadi, *Ph.D., P.E.*	Professor and Director, Lehman Center for Transportation Research, Florida International University.
Industry Advisors	Naresh Machavarapu, *PE, RSP$_{2I}$*	City Traffic Engineer, City of Boca Raton, Florida
	Stephen Kulisek	Traffic Signal Supervisor, City of Boca Raton, Florida
	Vince Picardi	Traffic Analyst, City of Boca Raton, Florida
Sponsoring Agency	2023 Transportation Technology Tournament, National Operations Center of Excellence (NOCoE)	

Abstract

Effective traffic signal control is crucial for enhancing traffic flow performance on signalized arterials, reducing travel time, increasing capacity, improving safety, and lowering emission and fuel consumption. However, current settings of signal timing plans rely on limited data collected for one day and fail to account for yearly traffic pattern variations throughout the year. Most traffic signals use a single timing plan for an entire year, disregarding daily variations. Continuous data from multiple sources, including traffic sensors and probe vehicles, offer opportunities to estimate advanced automated traffic signal performance measures in high resolution. However, agencies underutilize these measures, leading to missed operational and safety improvements. Improving signal timing plans can enhance system performance cost-effectively. Supervised and unsupervised will play a major role in supporting this improvement.

This project aims to create a tool that identifies traffic patterns throughout the year using clustering analysis and generates optimized signal timing plans for the identified clusters, considering recurring and non-recurring congestion. This proactive tool aims to increase signalized arterials' mobility and reliability by accounting for the different traffic patterns throughout the year. To accomplish this goal, unsupervised machine learning based on clustering analysis will be conducted first using data from multiple sources to identify different clusters of days and time intervals within the days that exhibit significant differences in traffic patterns. For each cluster, the team will develop a signal timing plan that will be developed and optimized, and then the different plans will be compared and evaluated. The result from the tool can also be used as a basis for next-generation traffic-responsive control systems that predict traffic patterns in real time.

ACKNOWLEDGEMENT

The team would like to express their sincere gratitude to the Transportation Management Center of the City of Boca Raton, Florida, for their invaluable support in this endeavor. Furthermore, the team extends their heartfelt acknowledgment to the City of Boca Raton, Florida, officials, namely Naresh Machavarapu, PE, RSP-2I, City Traffic Engineer; Stephen Kulisek, Traffic Signal Supervisor; and Vince Picardi, Traffic Analyst, for their dedicated efforts in producing this document. The team is indebted to their esteemed Academic Advisor, Dr. Mohammed Hadi, P.E., for his unwavering guidance, support, and insightful suggestions throughout the entire process. Additionally, the team would like to express their gratitude to the National Operations Center of Excellence (NOCoE) for organizing such an exceptional tournament.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

Traffic pattern recognition refers to the process of analyzing and identifying recurring patterns in the parameters and performance of traffic flow. This process involves obtaining information from traffic data, such as vehicle counts, speeds, and occupancy, to understand the characteristic traffic patterns that emerge over time and space. Traffic pattern recognition aims at gaining insights into traffic dynamics, identifying trends, and supporting informed decisions regarding traffic management and control. By recognizing patterns, transportation professionals can develop effective strategies to optimize signal control, improve traffic flow mobility and reliability, enhance safety, reduce congestion, and reduce emission and fuel consumption.

Traffic pattern recognition involves employing various techniques, including statistical analysis, data mining, machine learning, and pattern-matching algorithms. These methods allow for identifying patterns in traffic data that may not be immediately apparent to human observers. Common patterns that are often analyzed include daily and weekly variations, seasonal trends, and special event impacts. Traffic pattern recognition can be applied in various domains, such as transportation system management and operations (TSMO), intelligent transportation systems, and transportation planning and design. The procedure helps transportation agencies and engineers understand traffic behavior, predicts future traffic conditions, and develops proactive measures to mitigate congestion and improve overall transportation efficiency.

With the advancement of technology and the availability of large-scale traffic data from sources like sensors, cameras, and Global Positioning System (GPS) devices, traffic pattern recognition continues to evolve. Sophisticated algorithms and artificial intelligence techniques are being developed to handle the complexity and volume of traffic data, enabling more accurate and real-time recognition of traffic patterns. Overall, traffic pattern recognition plays a vital role in understanding traffic dynamics, optimizing transportation systems, and ultimately improving the mobility and efficiency of our road networks.

1.2 Current Practice by Agency

Agencies employ various methods for traffic signal timing to optimize traffic flow. The process typically involves collecting traffic data using traffic studies such as manual turn movement counts, and segment tube counts. The agencies use additional studies or data sources d to estimate travel times. Transportation agencies then analyze this data to understand traffic patterns, observe the traffic for a limited period in the field, and use signal timing optimization software to determine appropriate signal timing parameters. These parameters are later fine-tuned based on field observations and include green and red signal durations, cycle length, and offset timings between adjacent signals. Agencies may use traffic simulation models to evaluate different timing scenarios and assess their impact on traffic performance.

Additionally, agencies consider factors such as pedestrian movements, transit priority, and emergency vehicle response times when developing signal timing plans. Once the optimal signal timing parameters are determined, agencies implement the plans by reprogramming traffic signal controllers at each intersection. Periodic evaluations and adjustments are made based on feedback from field observations and ongoing data collection. Advanced technologies, such as adaptive signal control systems, are also being adopted by some agencies.

The current practice followed by transportation agencies for signal planning involves using traffic data collected for a single day to create signal timing plans. However, this approach fails to consider the significant variations in traffic patterns that occur throughout the year. Most traffic signals in the United States operate on a fixed timing plan that remains unchanged for an extended period, typically three to five years or longer, disregarding the changing traffic patterns.

2 DESCRIPTION OF PROBLEM

The identification of effective traffic signal control strategy and timing planning is essential for improving traffic flow performance on signalized arterials, benefiting road users by reducing travel time, increasing capacity, improving safety, and reducing fuel consumption and emission. However, when creating signal timing plans, transportation analysts tend to rely on traffic data for one day that does not represent the traffic patterns that occur throughout the year. Most traffic signals in the United States operate using a single signal timing plan for a given period of the day for the whole year with no consideration of the variation in the traffic patterns between the days of the year. In addition, these plans stay in effect for three to five years, or even longer, despite the changing traffic patterns. Detailed continuous data are becoming available from multiple sources detectors that cover all days of the year, including traffic sensors, high-resolution controller data, and third-party crowdsource data based on probe vehicles. These data sources allow the estimation of advanced automated traffic signal performance measures (ATSPM) with high resolution in time and space.

However, most agencies are still not using these measures effectively to develop signal timing plans that better consider the variations in traffic patterns for a whole year. The lack of a methodology to process the available detector from different sources to support signal control and timing efforts can lead to missed opportunities for operational improvements, resulting in avoidable delays and increased traffic congestion. Adaptive Signal Control Technologies (ASCT) can adapt to the traffic fluctuations that occur from cycle to cycle, allowing systems to operate more efficiently. However, agencies face additional costs of installation and maintenance for deploying ASCT. Thus, improving signal control based on the time of day has the potential to improve system performance in a cost-effective manner.

3 DESCRIPTION OF SOLUTION

3.1 Proposed Solution for the Problem

The goal of this effort is to create a tool that identifies the different traffic patterns throughout the year using data analytics and generates optimized signal timing plans for the identified patterns. This proactive tool aims to increase the mobility and reliability of signalized arterials by accounting for the different traffic patterns anticipated throughout the year. To accomplish this goal, a clustering analysis will be conducted first using data from multiple sources to identify different clusters of days and time intervals within the days that exhibit significant differences in traffic patterns. For each cluster, a signal timing plan will be developed and optimized, and then the different plans will be compared and evaluated. The derived signal plans can also be used as part of next-generation traffic-responsive systems that use predictive machine learning algorithms in real-time operations to predict the traffic pattern in the next 15-30 minutes and implement the corresponding timing plan to the pattern. Overall, this tool is intended to improve traffic operations for the majority of days in a year using existing data, thereby benefiting transportation agencies seeking to enhance their traffic management strategies.

3.2 Requirements for the Solution

The requirements for solutions to address the mentioned problem of traffic signal control and timing planning are as follows:

- Enhanced Data Collection: There is a need to gather detailed and continuous traffic data from various sources, including traffic sensors, high-resolution controller data, and third-party crowdsource data based on probe vehicles. This data should cover all days of the year to capture the variations in traffic patterns. This study uses data from different sources, including volume, travel time, and signal timing.
- Advanced Analytics and Clustering: Advanced analytical techniques and modeling tools are essential to process the available traffic data. This analysis includes developing methodologies to analyze and interpret the data from different sources, allowing for a better understanding of traffic patterns and performance measures. This study utilizes the k-means clustering method to recognize the traffic patterns from the collected data.
- Improved Signal Timing Plans: The development of signal timing plans should go beyond relying on data from a single day and consider the variations in traffic patterns throughout the year. The utilization of ATSPM can enable agencies to create more effective signal timing plans that optimize traffic flow and consider the specific characteristics of different days. As a part of the proposed solution, specific signal timing plans are developed for each cluster or pattern from the clustering and pattern recognition.
- Implementation of Next Generation Traffic responsive plan selection (TRPS): Several issues have been identified with traditional TRPS (please note that TRPS is a different control strategy than ASCT). The availability of data from multiple sources combined with supervised (neural network) and unsupervised learning (clustering) allow the development of a new generation of TRPS that will allow the application of proactive traffic control strategies that are based on predicted traffic conditions in real-time operations. Such strategies will take advantage of ATSPM measures that were not available for traditional TRPS. Please note that this capability in the developed system is optional and only applicable to agencies that want to implement the TRPS. If this is not the case, then the agency can implement the plans for traffic patterns in time-of-day settings.
- Cost-Effectiveness: Solutions should consider cost-effectiveness in terms of installation and maintenance compared to current practices of signal control, outlined earlier. While ASCT may offer benefits, the associated costs may restrict their deployments if agencies do not have the required resources. Therefore, alternative approaches that improve signal control based on the time of day should be explored to achieve system performance improvements in a cost-effective manner.

By addressing these requirements, transportation agencies can develop more efficient and effective signal control strategies that improve overall transportation system performance.

3.3 Stakeholders

Multiple stakeholders have been identified for this project, with critical roles assigned to various entities, including the following.

- The City of Boca Raton, Florida (the City): The City is a vital stakeholder as they will be the agency adopting the use of the methodology and tool and will have to collect and archive the data on which the model and application are built.
- Florida Department of Transportation (FDOT) District 4: The City of Boca Raton is in Palm Beach County. The state roads in the county are under the jurisdiction of FDOT District 4. The City continuously interacts with FDOT District 4 regarding their management. Thus, the tool and method will be shared with FDOT District 4 for their input and possibly shared with other agencies in their jurisdiction.
- Other FDOT districts and signal control agencies in Florida and the Nation: The findings from this effort will be shared with other agencies through webinars and conference presentations.
- Traffic Management Center (TMC) operators: The TMC of the City operators will utilize the developed tools. Thus, their inputs are collected and incorporated into the developed tool. In addition, the operators will be trained in utilizing the developed tools and using their results.
- Private sector data providers: The City has agreements with private sector providers that can be used as inputs to the tool.
- Other stakeholders: Other stakeholders that may be involved in selecting and implementing the plans may include the maintenance and construction management department, the emergency management center (EMC), and facility users.

4 CON-OPS FOR SOLUTION:

4.1 High-level functional architecture

The high-level functional architecture for this tool is comprised of five functions.

1. Data Input and Preprocessing: This function receives the traffic data as input from the user and performs the following sub-functions:
    a. Data Input: This sub-function retrieves data from various sources linked to the tool. The sources include traffic sensors, probe vehicles, crowdsourcing tools, or online data archiving tools. The tool retrieves the data for a specified period and location of analysis and checks for quality and format to make sure the requirements of the tool are met.
    b. Data Preprocessing: This sub-function performs the necessary data preprocessing steps that are required to ensure a successful clustering exercise. Preprocessing includes data cleaning by removing outliers and errors and filtering the data for the

relevant attributes based on the goal of the performed analysis. Additionally, data aggregation over specific time intervals and locations is performed in this sub-function.

2. Clustering and Traffic Pattern Identification:
    a. K-means Clustering: This sub-function applies the K-means clustering algorithm, an unsupervised learning method proven effective with traffic data. This method groups similar data points together into clusters based on their similarities. The number of clusters can be identified using optimization techniques.
    b. Principal Component Analysis: Depending on the aggregation of the data, the number of input parameters might be greater than the number of days (observations). In this case, this sub-function performs principal component analysis on the data to reduce its dimensionality and retain the essential features. This sub-function helps identify the key patterns in the data by transforming it into a lower-dimensional space while preserving the important information. It helps in understanding the significant factors that are influencing traffic patterns.
    c. Pattern Identification: This sub-function labels the obtained clusters and identifies the traffic patterns based on the characteristics of each cluster. Patterns might include free-flow conditions or congested conditions.
3. Signal Timing Optimization:
    a. Initial Signal Timing: This sub-function retrieves the initial signal timing plans for the study area and time interval from the TMC.
    b. Evaluation: This sub-function evaluates the existing signal timing plans and cross-checks the timings with the clustered data as a way of evaluating the performance of the plans.
    c. Optimization: This sub-function employs optimization algorithms to refine and update the time-of-day signal timing plans. Based on the obtained clusters and the identified traffic patterns, the objectives of the optimization process will be set accordingly.
    d. Next-Generation Traffic Responsive: This is an optional function and involves using a supervised machine learning algorithm (neural network) to identify the traffic pattern in the next 15-30 minutes in real-time operations and implement the corresponding signal timing plan for the pattern. This function is optional and is only used by agencies and locations that elect to use traffic responsive system. In other cases, the agencies can implement the plans for different patterns using time-of-day control.
4. Output and Visualization:
    a. Results Presentation and Visualization: This sub-function will generate corresponding heat maps of the grouped clusters. The heat map will show the relevant parameters, such as demand, capacity, travel time, and green-to-cycle length ratio. Heat maps before and after signal timing optimization will be generated to compare and evaluate the optimization process. This sub-function will support decision-makers and will facilitate communication with the involved stakeholders. Figure 1 and Figure 2 show an example of a heat map based on clustered data before signal optimization. The data were clustered into four clusters for the generated sample heat map. The data were aggregated for 15-minute

intervals over two hours in the mid-day period, resulting in eight-time intervals in total. Each row in the heat map represents a time interval of the corresponding parameter. The parameters are the upstream volumes (demands), capacity, travel time, and green/cycle ratio. This visualization will help understand the underlying issues with the signal timing and the causes of congestion for specific patterns such as increased demand, insufficient green time due to higher times for the cross street, and/or spillbacks from downstream intersections.

5. Deployment: This function integrates the tool for deployment in its intended environment. Installation, configuration, and setup of relevant software and servers, and cloud platforms will be performed in this function.

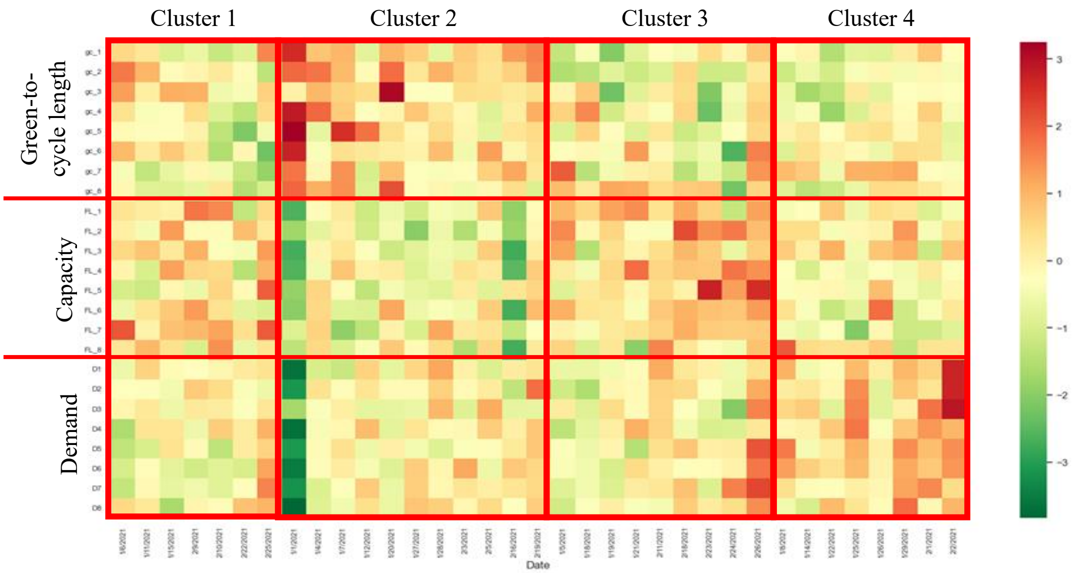

Figure 1 Sample Heat Map- Green/Cycle Ratio, Capacity, and Demand

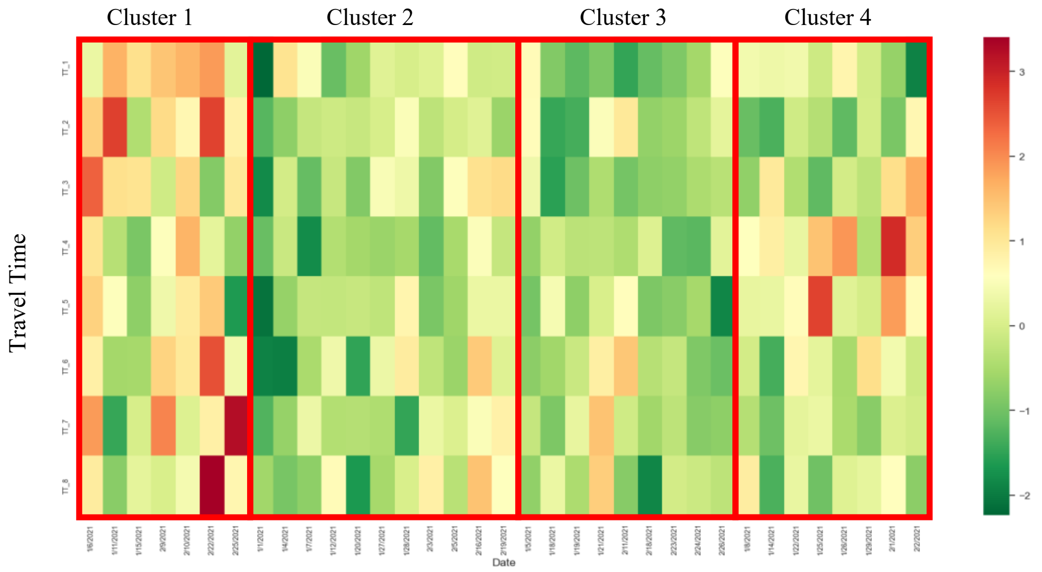

Figure 2 Sample Heat Map - Travel Time

4.2 High-level physical architecture

The physical architecture will be composed of the following:

1. Traffic collection devices and sources
    a. Traffic sensors: These devices could be existing loop detectors or non-intrusive detection technologies such as video image detectors and/or microwave radars to measure traffic volume, occupancy, and speed.
    b. Bluetooth readers: The readers provide data to estimate travel times for use as input to the tool, or this data can be collected from third-party vendors.
    c. Signal controller data: At a minimum, this data should include the signal timing provided by the actuated controller for each corridor for each period in all analyzed days. Detection event data, if available, are also very useful as they allow assessing other measures such as green occupancy ratios, split failures, and arrivals on green.
2. Data Processing and User Interface Infrastructure:
    a. Existing Computer: An existing computer will be necessary to handle the traffic data and perform data processing tasks efficiently.
    b. Web Server: The web server hosts the user interface of the application, allowing traffic engineers and operators to interact with the tool through a web browser. It handles user requests, serves the application's interface, and facilitates data exchange between the user and the tool.
3. Deployment infrastructure:
    a. Traffic control devices: These devices could include existing central systems, and field traffic signal controllers used to control traffic flow.
    b. Communications: Communications are required to exchange commands and information between the field and the center and monitor operations.

4.3 High-level enterprise architecture

1. City of Boca Raton, Florida:
    a. Provides necessary support for system deployment
    b. Facilitates infrastructure deployment and maintenance for data collection
    c. Implements system changes, including new signal timing plans based on clustering
    d. Installs the Next Generation TRPS
2. Traffic Management Center (TMC) of the City of Boca Raton, Florida:
    a. Uses the system to monitor traffic patterns and manage signal timings
    b. Maintains the traffic signals
    c. Collaborates with the City of Boca Raton in implementing the plan
1. Third-party Data Providers:
    a. Supplies additional data such as video image processing data, high-resolution controller data, or third-party crowdsource data
2. Maintenance and Construction Management Department:
    a. Conducts physical modifications to the traffic infrastructure, if needed
    b. Installs new sensors or upgrades existing ones, if needed
    c. Maintains the infrastructure

3. Emergency Management Center (EMC):
    a. Monitors traffic conditions for emergency response planning
    b. Provides input on system accommodation for emergency vehicles
4. System Operators:
    a. Operates the system on a day-to-day basis
    b. Monitors traffic conditions and adjusts signal timings
    c. Responds to system alerts or issues
5. Facility Users:
    a. Provides feedback on road usage
    b. Contributes to the evaluation of system performance

4.4 Work Needed to Develop and Implement the Solution

4.4.1 Cost breakdown

This proposed solution is entirely based on the existing facilities, including high-resolution controller data from the existing traffic sensors and traffic signal controller, third-party crowdsource data, and existing computational devices. Therefore, this solution will not require additional costs for infrastructure except for implementing the next-generation TRPS, including installation, maintenance, and labor costs. The following section will describe a cost breakdown of implementing the Next Generation TRPS.

- Additional Sensors for TRPS: This includes additional sensors at strategic locations to measure traffic patterns in real time.
- Software modifications: This includes modifications to the central software to select signal timing plans based on the predicted pattern based on real-time operations.
- Installation and maintenance: The costs associated with the additional hardware and software will need to be added.

Table 1 Item Wise Cost Breakdown of the Proposed Solution

Items	Estimated Costs
Additional Sensors	$5,000.00
Central Software Modification	$2,000.00
Hardware Installation	$3,000.00
Project Cost	**$10,000.00**
Maintenance (30% of project cost)	$3,000.00
Overhead Cost (10% of project cost)	$1,000.00
Total Cost	**$14,000.00**

4.4.2 Timeline

Overview of Steps:

- Data Collection and Processing: These include collecting detailed and continuous traffic data from the high-resolution controller and third-party crowdsourcing based on probe

vehicles. This data should cover all days of the year to effectively capture the variations in traffic patterns.
- Clustering Traffic Pattern: After collecting and processing the data, the clustering algorithm will be developed to identify the traffic pattern, which may take up to 30 days.
- Signal Timing Plan Development: This involves developing specific signal timing plans based on time-of-day and Next Generation TRSP for each cluster or pattern from the clustering and pattern recognition.
- Unit Test: The unit testing will be performed before implementing Next Generation TRSP.
- Implementation: Developed signal timing plans for each cluster or traffic pattern will be implemented.
- Evaluation: This evaluation assesses the effectiveness of the implemented plans and makes necessary adjustments.
- Operation and Maintenance: This solution is expected to be operated and maintained for at least five years.

Table 2 Item Wise Cost Breakdown of the Proposed Solution

Items	Duration	2023						
		Jun	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Data Collection and Processing	60 days	🟦	🟦					
Clustering Traffic Pattern	30 days			🟦				
Signal Timing Plan Development based on Next Generation TRSP	30 days				🟦			
Unit Test	30 days					🟦		
Implementation	30 days						🟦	
Evaluation	30 days							🟦
Operation and Maintenance	5 years							

5 POTENTIAL BENEFITS

This application aims to optimize traffic operations by leveraging existing data, leading to a more efficient and proactive response to traffic conditions. By utilizing this tool, transportation agencies can significantly enhance their traffic management strategies, resulting in improved mobility, reliability, safety, fuel consumption, and environmental impacts. Its primary objective is to provide better traffic management for most days of the year, revolutionizing how transportation agencies handle and address traffic challenges.

5.1 Operational benefits

This application optimizes traffic operations by leveraging existing data, resulting in a more efficient and proactive response to traffic conditions. It revolutionizes how agencies handle traffic challenges by empowering them with comprehensive data-driven insights.

5.2 Safety benefits

The application's adaptability to various traffic patterns reduces congestion for the majority of days of the year, resulting in a safer transportation environment. This proactive approach to traffic management decreases the likelihood of primary and secondary crashes, ensuring enhanced safety for road users. The tool can be extended to explicitly consider safety in clustering and the predictive models, and this will be done in a future version of the tool.

5.3 Mobility benefits

This tool generates new signal plans tailored to different traffic patterns throughout the year. These plans enhance mobility on signalized arterials by reducing average delays, mitigating congestion, and preventing queue spillbacks. The proactive application improves traffic flow and reduces congestion, benefiting commuters and businesses with shorter travel times, improved accessibility, and greater overall transportation network efficiency. Travel time reliability is also expected to improve. A recent evaluation of the concept by the team indicates a potential reduction of about 7% in travel time compared to the existing practice of driving time-of-day plans.

5.4 Environmental benefits

Its proactive traffic management approach also fosters eco-friendly practices, promoting a cleaner and greener environment. The application effectively reduces fuel consumption and vehicle emissions by minimizing congestion and decreasing total and average delays. The tool can also be extended to explicitly consider environmental impacts in clustering and the predictive models, and this will be done in a future version of the tool.

5.5 Other benefits and risks

The application offers numerous other benefits for traffic management. It enhances overall transportation efficiency, improving public perception and satisfaction among road users. With its data-driven insights, the application enables informed decision-making, evidence-based strategies, and more efficient resource allocation. It is scalable and adaptable, applicable to transportation networks of various sizes, making it valuable for agencies of all scales. The real-time monitoring and adjustment capabilities allow for proactive management of traffic incidents, congestion, and disruptions, enhancing overall traffic flow and emergency response.

However, the implementation of the application is not without constraints. Data accuracy and reliability are crucial, as suboptimal data can compromise decision-making and performance—the system integration and compatibility present technical challenges, requiring careful planning and additional resources. Stakeholder acceptance and adoption are also vital for successful implementation, as resistance to change and the need for training and education can pose challenges. Addressing these risks through robust data validation, thorough testing, and stakeholder engagement is essential to ensure the application's effective utilization and maximize its benefits while minimizing potential challenges.