Solution Definition and High-Level Con-Ops

Improving the Communication of Work Zone Data

Prepared by
Iowa State University
Center for Transportation Research and Education
2711 S. Loop Drive, Suite 4700
Ames, IA 50010-8664

Authors
Smrithi Ajit, Sayedomidreza Davami, Tongge Huang, Skylar Knickerbocker, Varsha Ravichandra-Mouli, and Archana Venkatachalapathy

Advisor:
Anuj Sharma, Associate Professor, Iowa State University
Sinclair Stolle, Traffic Management Systems Engineer, Iowa DOT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

OVERVIEW

Purpose

The National Highway Traffic Safety Administration (NHTSA) defines a highway work zone "as an area where roadwork takes place [which] may involve lane closures, detours, and moving equipment." (1) Work zones are crucial to building a safe and healthy road system. It involves regular construction and maintenance of roadworks such as pavements and bridges to prevent degradation of traffic conditions. Significant emphasis has been placed on work zones at this time with the Federal Highway Administration (FHWA) leading the Work Zone Data Initiative (WZDI), which is trying to provide an architecture of the when, where and how work zones are deployed, and the U.S. DOT leading the Work Zone Data Exchange (WZDx) which is developing a standard way of communicating work zone information between infrastructure owners and operators and third-party data consumers. Both of these efforts at this time appear to focus on static lane closures with little focus on moving operations that can be difficult to track in real-time. This document will describe a high-level concept of operation that uses automatic vehicle location data to locate the beginning and ending location of slow-moving maintenance operations.

Background

Work zones often disturb the driving environment. When a work zone is in place, the existing traffic system is severely impacted in terms of mobility and safety. Closures and detours at work zones reduce the road capacity leading to an increase in congestion resulting in travel delays. A 2017 Federal Highway Administration (FHWA) report described seven sources of traffic congestion and highlighted work zones second on the list of important causes. The report estimated that work zones cause 10% of congestion and 24% of unexpected freeway delays (2). In the past 5-year, FHWA estimates 4,400 fatalities and, 200,000 injuries have occurred in work zones. Rear-end type crashes are the most common type, with distractions, slow or stopped traffic and improper following distance being the most important causes for crash. (3)(4)

Monitoring and evaluating a work zones performance is vital to ensuring smooth traffic flow on roadways, which can be performed with the data collected from devices or sensors deployed in the field. Most relevant data related to work zones include traffic volume and speed. With advancements in technology, there has been significant improvement in data collection devices. Devices built now-a-days collect more volume and detailed information related to traffic events.

Many agencies are beginning to focus on improving this data through the use of connected temporary traffic control devices such as smart arrow boards or traffic cones. Significant improvement has been made over the last several years to begin incorporating these devices into an agency to improve work zone location information including Minnesota DOT who developed a concept of operations (5) and the Iowa DOT with their smart arrow board deployment plan (6). Information on maintenance operations is an area of focus which must be improved because of the short duration and the dynamic nature of the work. Currently, a manual notification of work status is made to notify the TMC of operations underway. In some cases, when the work is stopped midway or completed early, the TMC may not be notified. The accurate information of ongoing maintenance operations can be essential for travelers, which is currently not available with the TMC for dissemination due to the non-availability of a reliable data collection system.

To improve this data, agencies have equipped their maintenance vehicles with Automated Vehicle Location (AVL) systems. These vehicle systems are primarily used for snowplow operations, but the same vehicles are used for painting, pothole repairs and other maintenance operations. To minimize the work by the maintenance staff and improve the quality of work zone data, the AVL data points can be used to classify a vehicle in maintenance mode then cluster all of the surrounding vehicles to get the extents of the maintenance operation. As the vehicles move, the data will be updated in real-time to accurately communicate the extents of the maintenance operation through the agency's WZDx or ATMS/ATIS system. This project will develop an automated process that identifies maintenance activities and clusters AVL data to communicate actual maintenance operations in real-time through the WZDx.(7).

Stakeholders

Stakeholders of this data include internal departments of State Transportation agency (Maintenance Field Staff, TMC Operators and DOT staff), the public, third party navigation companies and connected vehicles.

State Transportation Agency – Maintenance Field Staff

The maintenance field staff are the primary producers of this maintenance activity location currently. This is a manual process that requires them to call into the Traffic Management Center (TMC) multiple times. The initial activity is a notification of their work activity and the approximate location of that activity, which can cover multiple miles. This area likely will encompass the area they will be working throughout the day to minimize the number of calls needed to the TMC. It is not expected that this activity will change as the estimated location of activities will still be needed in this concept of operations.

In addition to the initial call, multiple additional calls can be made, which include notification of work starting, notification when changing directions or roadways, and notification of ending. All of these notifications are less likely to occur by field staff since they are focused on the ongoing work that is being completed, and the calls to the TMC can be forgotten. The most common activity that can be forgotten is the notification of work completion. It was stated by field staff that there are multiple occasions when weather sets in and they cancel their activity but forget to notify the TMC since they are needing to get started on other activities.

The system described in this concept of operations will be automated which will eliminate the need for the multiple phone calls to the TMC notifying them when conditions change. This system will require only the initial phone call of the approximate locations of the work activity and duration. The maintenance trucks, as shown in Figure 1, are equipped with AVL which reports their location approximately every six seconds that can then be associated with the planned work activity submitted. The remaining activities that were called into the TMC can be eliminated as the system will be continuously updating the location of the work activity based on the location of the maintenance vehicles.

State Transportation Agency – TMC Operators

The TMC operators are currently the users entering this information into the ATMS or ATIS system that reports the maintenance activities to the public. The operators are the ones taking the calls from field staff and entering this information for the public to see in addition to their other activities including identifying and responding to crashes or other impacts along the transportation system . If time is available, they will also monitor the activity on cameras and post message on dynamic message signs if available. In Iowa, there are typically 3-5 operators on staff 24/7 in the Traffic Management Center in Ankeny, IA, as shown in Figure 1.

The system should minimize the work activity needed by the operators since they will just need to input the initial location of the work zone activity reported by the field staff. The AVL information will update the activity based on the presence of the maintenance vehicles within the area. Begin and end times will be automated since the system will be able to identify the initial activity within the area as well as when the maintenance activity ends.

Figure 1. Images of Iowa DOT Maintenance Vehicles with AVL and the Iowa DOT Traffic Management Center.

State Transportation Agency – DOT Staff

State Transportation agencies are most interested in this information, which gives them vital insight to improve the work zone configurations implemented for better traffic management. However, the accuracy of the work zone locations and times are difficult for many agencies to collect and maintain. This can make it difficult to monitor the performance of a work zone or identifying ways to improve work zone safety and mobility. Providing an automated system that associates AVL data with work zone activity will allow for this information to be archived and used to improve the research needed to improve work zone safety and mobility.

Public/Third Party Navigation/Connected Vehicles

These three groups of users are external users of the data. They are grouped together since they are all end users of the data and the goal is to provide the data to them through the US DOT led Work Zone Data Exchange (WZDx) which provides a standard protocol to push agency data to third party users. The end users are critical users to the system as this information attempts to provide accurate work zone activity data to those users to improve the safety and mobility. The current system provides little accuracy of the actual location of work activity which can leave users to ignore the notifications of work zones. If the information is accurate and users know the exact extents of the activity, they may be more likely to slow down which should improve the safety and mobility through the area.

The third parties are critical to this system since the 511 systems that are currently used to communicate this information are used minimally by the public compared to Waze, Google, HERE, TomTom, etc. Being able to seamlessly provide this information to third parties should allow this information to be received by more of the public and eventually to connected vehicles. This is the goal of the WZDx that will be leveraged in the proposed concept of operations as the end product of the system.

Solution Requirements

Based on the discussion with stakeholders and the focus by the FHWA and US DOT in work zone activity data, the following solution requirements were developed.

- Minimize the input of work zone activity information by maintenance staff and TMC operators. This will allow them to focus on other activities and improve the data from the system

- Provide an automated way to identify slow-moving maintenance activity with the existing AVL system deployed on snowplow/maintenance vehicles.
- Provide an automated way of associating maintenance operations with planned information entered into the ATMS/ATIS system
- Provide the data in a format that can be consumed by third-party systems that are more likely to get information to the traveling public to improve the safety and mobility in work zones.

DESCRIPTION OF SOLUTION

To improve this accuracy of slow-moving maintenance operation, many agencies currently equip their vehicles with AVL systems, to track their snowplows, which can also be used to track maintenance operations. The snowplows, in most cases, use the same vehicles for maintenance operations such as painting, pot-hole repairs, etc. The next step is to identify what kind of analysis can be done on the data gathered, which would be essential for state DOTs to improve work zone management.

To achieve the solution, 3 steps are required.

- Conflating data to linear referencing system
- Clustering the process to obtain the length of the maintenance operation and track them.
- Associating them with the planned work activity

Linear referencing is used as the basis for clustering and associating data along the same roadway. Linear referencing is the method of identifying geographic location by using relative positions along a measured linear feature and has traditionally been used in asset management. This process attempts to operationalize the linear referencing system (LRS) by leveraging the system for real-time operations. The linear referencing system provides complete information on the entire roadway system for the State of Iowa and is managed through a Geographic Information System (GIS) interface. Coordinates can be passed to the LRS through a REST service that will conflate the data to the LRS and return the RouteID and measure that are fundamental to locating the data spatially. The 'RouteID' and 'Measure' values that are obtained by conflating can be used through database functions to cluster and relate other spatial data without the computational expense for similar GIS related tasks.

The input AVL contains a variety of different variables that are relevant for winter operations but those values are stripped with the exception of the coordinates, timestamp and velocity. Table 1. Sample AVL Data after LRS Conflation.Table 1 shows a sample of the data after it has been conflated to the LRS. The last two columns "RouteID' and 'Measures' represent the values returned by from the LRS. A similar process is completed on the ATMS/ATIS maintenance events with the beginning and ending coordinates.

PingID	Longitude	Latitude	Date	Time	TruckID	Velocity	RouteID	Measure
48054217	-93.35623	43.14007	12102019	10:18:20	A33661	15	S001910035S	13.025
48054219	-93.97150	42.75416	12102019	10:18:21	A34665	21	S001930017N	13.436
48054214	-92.66557	43.04446	12102019	10:18:20	A30121	0	-99999	0
48093288	-91.55803	40.62576	12102019	10:55:50	A34665	30	S001920218S	10.661

Table 1. Sample AVL Data after LRS Conflation.

Once the AVL data has been conflated to the LRS it must be clustered and tracked. Cluster analysis is also known as a classification analysis which is an approach for grouping similar observations in such a way that all similar observation get grouped together. This process is used to convert multiple AVL data points into a single line as shown in Figure 2. In the Figure 2 (on the right), multiple points or pings for a 1 minute interval are shown. The two colors represent two different trucks. Since the pings from the two trucks fall into the same 1 minute interval, they are clustered or grouped together to represent a single maintenance activity as shown on the left. The single line represents the extent of the maintenance operation for that minute.

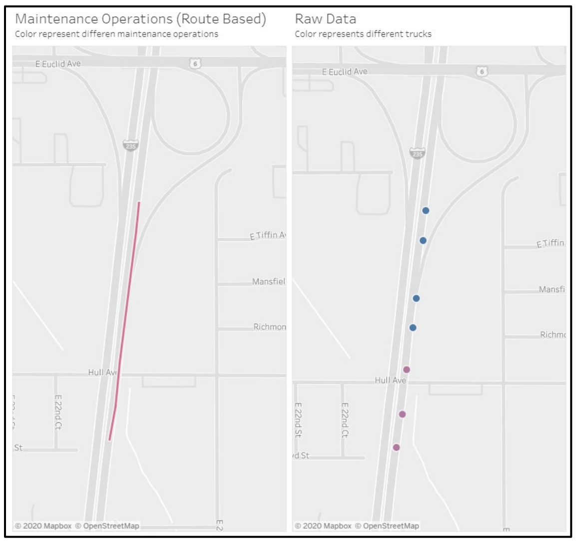

Figure 2. Example of clustering AVL Data pings for 1 minute.

In order to cluster trucks into unique operations, a buffer distance must be selected to cluster trucks. The goal is to determine if there is another cluster operation happening in the same area in the same time period and include all of them under the same unique maintenance operation. A buffer distance can be defined as a distance limit (in miles) under which action needs to be done or captured. This distance can be tested to refine results in operation but it should be less than 1 mile. Since the data has been conflated to the LRS which provides a measure in miles this operation can be easily completed without any additional GIS processes and has a higher accuracy since it will only relate pings that are along the same roadway.

Maintenance operations often extend across a stretch of road or multiple roads over a time period. Because of this, each of the clusters must be tracked over time to know that they are related to the same maintenance activity. To do this, post processing of the clustering is done to traceback the clusters and provide all cluster the same unique identifier comparing the location extents and times of each unique cluster.

The final step in the process is the integration of the AVL data clusters with the ATMS/ATIS maintenance events using the LRS values that were obtained from each. The reason this step is necessary is because the AVL data alone does not contain all of the necessary information needed for the WZDx data feed that can be ingested by third party data providers. The ATMS/ATIS event can provide additional details such as the estimated end time, the number of lanes closed, any restrictions due to the maintenance operations, etc. Figure 3 shows an example of how the two data sources are integrated using the LRS values for multiple maintenance operations along the roadway. For example, planned maintenance #1 would be called into the traffic management center by the field staff with estimated locations where they will work, in this case from mile markers 110 to 113.5. The AVL data in Figure 3 is processed and clustered every minute and shows the extent of the work to be from 110.11 to 113.47 falling within the extents of the planned work reported by the field staff. Since both AVL cluster and ATMS/ATIS operation are on the same route and there is overlap in their extents they are related and the AVL extents can be used to update the existing extents of the ATMS/ATIS event to be included in the WZDx data feed with verified instead of estimated locations. A similar second cluster is also shown from 98.76 to 99.82 that also falls within the extents of what was planned by the field staff. In addition to these, a third cluster is identified based on the AVL data from 107.67 to 108.26. Since no data on pre-planned work in that small section is available, it may be believed that the work in that zone was not reported by the field workers, an added advantage of the system that has been developed.

A majority of the system will now be automated which will allow for accurate maintenance activity extents to be reported to the public and used internally to improve the maintenance activities impacts to mobility and safety. The biggest player in the solution part is the AVL data provided by the DOT fleet since the solution is data-driven. The field staff and operators will still need to communicate to submit planned maintenance activity which will contain details needed for the WZDx that can not be automated through the AVL data including the duration/end time, the lanes closed and any vehicle restrictions.

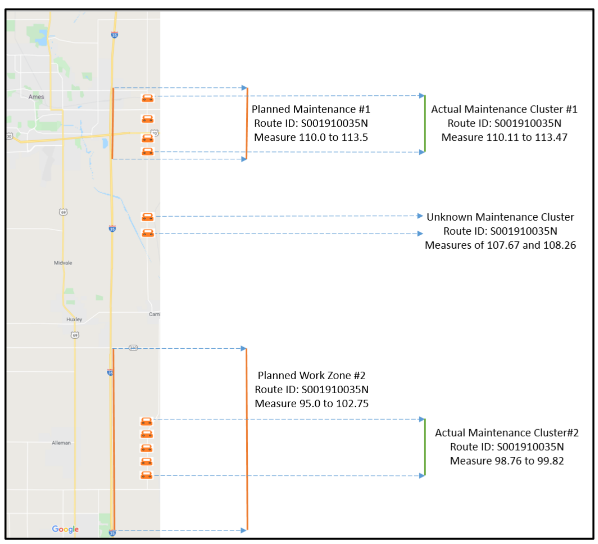

Figure 3. Example integration of AVL data clusters and ATMS/ATIS maintenance events.

CONCEPT OF OPERATIONS

Physical Architecture

The High Level Physical architecture may be thought of as having three layers namely, the data provider layer, the backend and data storage layer, and the front end layer as shown in Figure 4. The physical architecture leverages the existing AVL data that the Iowa DOT is currently collecting from each of the over 800 maintenance vehicles statewide through a separate vendor contract. The AVL data signals (pings) from the physical GPS sensor in the truck are sent to Iowa DOT where the data is stored and made available through an external API. The data in the API includes all of the pings from the previous minute that were received. The system pulls the data from the external API provided by Iowa DOT into cloud hosted storage. The data is pulled into the Hadoop cluster using a python process. The data is passed to the PostgreSQL database in order to retrieve the LRS elements corresponding to the AVL data collected. PostgreSQL is being used instead of the ESRI Roads and Highways REST services since it can scale to the increase in data that need to be conflated to the Iowa DOT's LRS. The processed data with LRS elements is then archived in the Hadoop cluster and made available to the end user either through a visualization tool or a feed that is compliant with WZDx. The archived data allows users to interact with the data from the Hadoop cluster through a Tableau server.

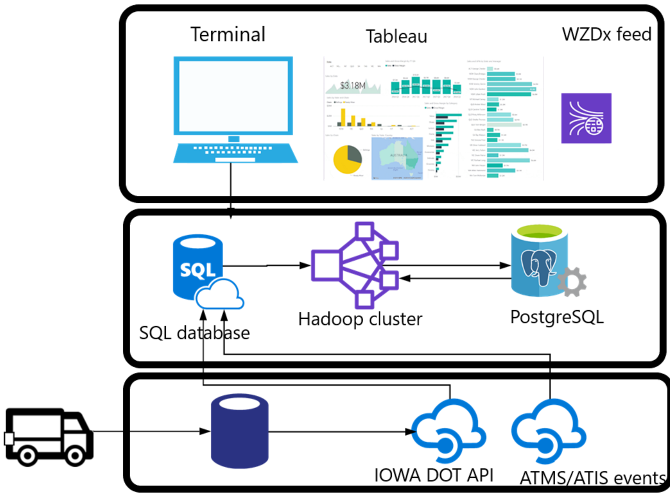

Figure 4. Physical Architecture

Enterprise Architecture

The High Level Enterprise architecture shows the types of architectures used, the concepts, principles, resources, and models that are part of the underlying architectures as depicted in Figure 5. Enterprise architecture encompasses the business architecture, application architecture, system architecture and information architecture. The subcomponents of the business architecture include the Iowa DOT, ATMS/ATIS and TMC, the major stakeholders in the process. The application architecture includes the modules for work zone association, identifying maintenance operation and time clustering and tracking, the working of which are explained in the functional architecture description. The system architecture includes the Linear Referencing System (LRS), the data storage module, and the WZDx modules. Finally, the information architecture includes the relevant data being ingested and processed through the system and includes snowplow AVL data, CCTV camera data and snowplow camera data.

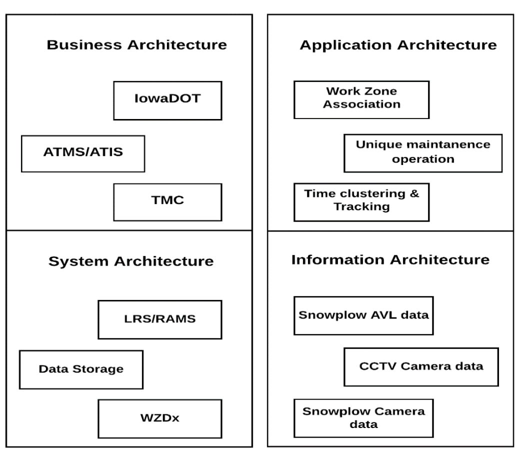

Figure 5. Enterprise Architecture.

Functional Architecture

The overall functional architecture is shown in Figure 6. Below are detailed descriptions of the functions within each module.

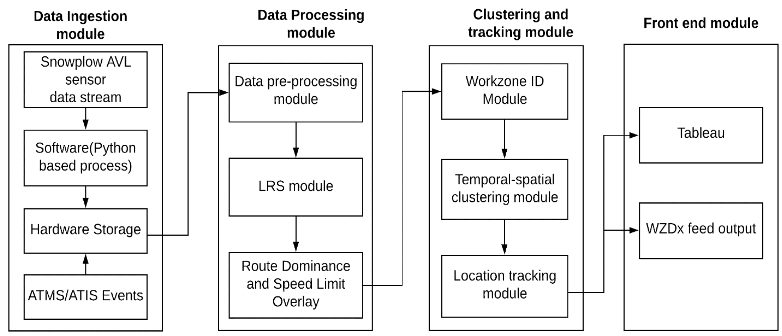

Figure 6.Functional Architecture.

Data Ingestion

The data from the maintenance AVL and the ATMS/ATIS is ingested with the help of a python process that connects to an external API provided by Iowa DOT. The AVL data is downloaded every minute while the ATMS/ATIS data is downloaded every 15 minutes. The ATMS/ATIS events are filtered to only include construction/maintenance events and the AVL data is filtered to only include pings where the speeds are less than 25 mph (classified as in maintenance mode). After the data is downloaded it is stored in a SQL database in the cloud. The data can then be pulled into the Hadoop cluster to handle the processing of the AVL data being received.

Data Processing

The data processing step not only involves addressing data inconsistencies and non-conformities like missing values, date-time issues but also additional steps of sending the data through the Iowa DOT's linear referencing system. The LRS provides a RouteID and measure value that may be used to obtain any attributes that are also associated with the LRS. The Iowa DOT provides an ArcGIS REST service that allows coordinates to be provided and RouteID and measures values will be returned for those coordinates but the system is not capable of handling the amount of data provided from the maintenance vehicles in a timely fashion. Instead, a PostgreSQL database is being used to process the data locally. Initial testing showed that the ArcGIS REST service can process 14 pings/second while the PostgreSQL using PostGIS with geometry could process 361 pings/second. Both the ATMS/ATIS and maintenance AVL pings are processed through the PostgreSQL database to return LRS route and measure values.

Temporal-spatial clustering module

This module involves the use of temporal clustering by grouping maintenance AVL pings into 1-minute aggregates assigning a unique identifier to each cluster. The pings may be from different trucks but as long as they fall within the 1 minute time period they are grouped together. Spatial clustering is ensured through grouping sensors by their closeness or proximity, clearly within a tolerance band. This process is made easier by the LRS values which allows pings to be clustered based on their RouteID and the proximity of the measure values.

During the clustering module, each unique identifier is then associated with an ATMS/ATIS. This allows the maintenance operation to be automatically associated with the maintenance event that was called into from the field staff. The unique clusters and ATMS/ATIS events are associated based on overlaps in the LRS values as described in the description of the solution.

Location tracking module

This module helps track the movement of the trucks spatially progressing through the clusters with the help of the start and end locations. Operations or activities occurring within a pre-defined buffer distance are grouped together as one operation and given the same unique identifier. This ensures that a moving operation that spans multiple miles of roadway is provided the same unique identifier for every minute instead of providing a different unique identifier each minute the data is processed.

Front end module

The front end module provides an external feed that is WZDx compliant and capable of being consumed by third-party data consumers that can then be used in navigation systems such as Google, Waze, HERE, and TomTom. In addition to this, the database is accessible to a tableau server which helps to visualize the progression of each work zone activity as clustered pings spatially progressing with time.

LEVEL OF EFFORT TO IMPLEMENT

The table below provides a cost breakdown and estimate duration for each of the activities for the proposed solution. The solution is leveraging most of the services already available within the DOT and adding this enhancement on top of the existing systems. The system currently leverages open source software that is available and will require a server as well as development time to implement. Cost is based on data from the US DOT ITS deployment cost database. The estimated cost for three years of service is $97,500.

Task	Duration	Cost Items	Estimated Cost
Database Setup/Design	3 months	1 Server ($20k) and Personnel Cost ($10k)	$35,000
System Integration	6 months	Personnel Cost ($50k)	$50,000
System Testing	1 month	Personnel Cost ($10k	$10,000
System Maintenance	Ongoing	Personnel Cost (20 hours/year)	$1,250/year

Table 2. Cost Breakdown.

ANTICIPATED IMPACTS

Operational benefits

The WZDx acts as a primer or a standard for communication and provides a significant operational benefit of improved communication among stakeholders and various organizations regarding the work zone. Dynamic tracking of work zones and updating in a centrally accessible point is reducing operational complexity. When a work zone becomes active or when an operation is complete, and it becomes inactive, state DOT operators, traffic management centers (TMCs), and the public are provided with timely information. Since the work zone management task will be automated, it will reduce the workforce and ensure enhanced work accuracy.

The system will also be able to identify unauthorized or unknown maintenance activities if there are no planned maintenance events in the ATMs/ATIS systems. This should prevent any unexpected maintenance activity from occurring or allow improved coordination if it does occur.

Safety benefits
This maintenance system will leverage ITS data and allow for the dynamic tracking of work zones by the public, leading to improved traffic flow and safety. Providing the data in a WZDx format will allow more third party consumers to provide this information once it is fully adopted and increase the change that a driver, or connected vehicle, will receive this data. The data will provide also provide a verified location of the work activity that can better inform the drivers and increase the trust with the public that the data is accurate and not to be ignored. Since the process reduces the amount of work being asked of the maintenance workers, it will allow them to focus on the maintenance operation, increasing their safety. The accurate work zone information will also facilitate on-going work zone safety analysis by enabling the synchronization of work-zone and incident data.

Mobility benefits
Knowing the accurate start and finish time/location of each work zone makes it easier to plan travel and hence helps the regulation of traffic flow near work zones. This will help allow usage of parallel corridors or alternate routes and reduce bottlenecks. By increasing visibility of work-zone information and delay times, we reduce crashes and hence avoid further road or lane blockage. DOT staff can also use this information to understand, in real-time, the impacts the maintenance is having on mobility and request that the operation end if an excess delay is being caused. This cannot be accomplished currently since the location of the work activity is not precisely known to trigger alerts.

Environmental benefits
Improved work zone management and its communication between agencies reduce unstable traffic caused by work-zones. Since the accuracy of the work zone is improved, it reduces emissions from vehicles due to queuing and traffic congestions, frequent acceleration and deceleration and due to reduced speed in work-zones. It limits stop-and-go traffic, thus reducing overall fuel consumption. The reduction of congestion in work-zones also facilitates a decrease in ozone and other pollutant levels.

Other benefits and risks
Further enhancements can be made to validate work zone accuracy by utilizing video feeds of cameras on roads as well as the usage of cameras present in AVL trucks. This can eventually be used to confirm work activity or potentially identify the lanes that are closed. Knowing the location of work zones can also allow for better performance monitoring to understand the impacts different maintenance activities have on the public and allow for the agency to better plan the time and type of activity to improve mobility and safety in the future.

The team did not identify any current risk with the system identified since it is an additional enhancement on existing systems. But as the public and connected vehicles may begin to rely on the more accurate maintenance activity information, the system may be critical for operations outside of the DOT.

REFERENCES

1. Council, N. S. Work _ Construction Zones Work _ Construction Zones. 2008, pp. 1–2.
2. FHWA Work Zone Facts and Statistics - FHWA Office of Operations.
3. Federal Highway Administration. Work Zone Safety for Drivers. 2003.
4. J. Paracha. Work Zone Safety for Drivers - Safety | Federal Highway Administration. https://safety.fhwa.dot.gov/wz/resources/fhwasa03012/.
5. Minnesota DOT, Real-Time Integration of Arrow Board Message into Traveler Information Systems. https://www.dot.state.mn.us/its/projects/2016-2020/arrowboard/conops.pdf
6. Iowa DOT, Smart Arrow Board Deployment Plan. https://iowadot.gov/workzonereferencelibrary/docs/Smart-Arrow-Board-Deployment-Plan.pdf
7. Varsha R.M. 2020, Using Automatic Vehicle Location (AVL) for real-time maintenance identification and tracking