



**Vehicle Infrastructure Integration (VII)
Data Use Analysis and Processing
PROJECT SUMMARY REPORT**

Final
March 2012

Prepared By:



Mixon/Hill, Inc.

12980 Metcalf Avenue, Suite 470

Overland Park, Kansas 66213

(913) 239-8400

All documentation, software, and data related to this project are proprietary and copyrighted. Use is governed by the contract requirements as defined in the Michigan Department of Transportation contract 2007-0371. Unauthorized use of this documentation is a violation of law except as provided for in said contract.

Copyright © 2012 Mixon/Hill, Inc. All rights reserved.

Technical Report Documentation Page

1. Report No. RC-1568	2. Government Accession No. N/A	3. MDOT Project Manager Steven J. Cook, P.E.	
4. Title and Subtitle Vehicle Infrastructure Integration (VII) Data Use Analysis and Processing – Project Summary Report		5. Report Date March 2012	
		6. Performing Organization Code N/A	
7. Author(s) Lee T. Mixon, J. Kyle Garrett, Bryan C. Krueger		8. Performing Org. Report No. N/A	
9. Performing Organization Name and Address Mixon/Hill of Michigan, Inc. 12980 Metcalf Avenue, Suite 470 Overland Park, KS 66213		10. Work Unit No. (TRAIS) N/A	
		11. Contract No. 2007-0371	
		11(a). Authorization No. N/A	
12. Sponsoring Agency Name and Address Michigan Department of Transportation Research Administration 425 West Ottawa Street Lansing MI 48933		13. Type of Report & Period Covered Final Report 2/15/2007 to 8/31/2011	
		14. Sponsoring Agency Code N/A	
15. Supplementary Notes N/A			
16. Abstract <p>The purpose of the Data Use Analysis and Processing (DUAP) project is to support the Michigan Department of Transportation (MDOT) and its partners in evaluating uses and benefits of connected vehicle data in transportation agency management and operations. The project complements efforts throughout the transportation community to design and deploy connected vehicle infrastructure, vehicle equipment, and initial applications and investigates how data from connected vehicles may benefit the ways MDOT and other transportation agencies do business. DUAP research has been constrained by the relative unavailability of connected vehicle data, but has successfully demonstrated: the capability to collect, aggregate, process and provide data from connected vehicles; pragmatically acquiring diverse data from a variety of sources; and applications that may enhance traffic monitoring, pavement defect and condition assessment, and origin-destination studies for transportation planning. The project concludes that there is substantial potential for the use of connected vehicle data in transportation management and operations. It is recommended that the next phase of research focus on development of reliable data sources and specific applications for implementation in MDOT.</p>			
17. Key Words Connected vehicle; VII; applications		18. Distribution Statement No restrictions. This document is available to the public through the Michigan Department of Transportation.	
19. Security Classification - report Unclassified	20. Security Classification - page Unclassified	21. No. of Pages 58	22. Price N/A

Table of Contents

EXECUTIVE SUMMARY	IV
ACTION PLAN FOR RESEARCH.....	V
1 INTRODUCTION.....	1
1.1 Objective.....	1
1.2 Scope	1
2 METHODOLOGY	3
2.1 Systems Engineering.....	3
2.2 Data Interfaces	4
2.3 Applications Analysis.....	5
3 RESULTS AND ACCOMPLISHMENTS.....	7
3.1 DUAP System	7
3.1.1 Data Collection.....	7
3.1.2 Storage.....	8
3.1.3 User Interface	9
3.2 Data Sources.....	17
3.2.1 MITS Center.....	17
3.2.2 Chrysler Fast FeedBack Fleet Input	19
3.2.3 MDOT Fleet Vehicles	21
3.2.4 Android Applications	23
3.3 Applications Analysis.....	23
3.3.1 Traffic Condition Monitoring	24
3.3.1.1 Objective.....	24
3.3.1.2 Concept	25
3.3.1.3 Research Findings	26
3.3.2 Pavement Condition Monitoring	27
3.3.2.1 Objective.....	27
3.3.2.2 Concept	28
3.3.2.3 Research Findings	29
3.3.3 Origin-Destination Studies	34
3.3.3.1 Objective.....	34
3.3.3.2 Concept	35
3.3.3.3 Research Findings	36
4 CONCLUSIONS.....	38
4.1 DUAP System Capabilities.....	38

4.2	Data Collection and Standards	38
4.3	Applications	39
5	RECOMMENDATIONS FOR FURTHER RESEARCH	41
6	RECOMMENDATIONS FOR IMPLEMENTATION.....	42
	APPENDIX A - ACRONYMS AND ABBREVIATIONS	44
	APPENDIX B - REFERENCES	46
	APPENDIX C - UNITS CONVERSION TABLE.....	47
	APPENDIX D - MDOT FLEET DATA SPECIFICATION	48
	In-Vehicle Device Configuration.....	48
	Data File Specification(s)	48
	File Identification and Transport.....	50
	APPENDIX E - DUAP INPUT DATA SPECIFICATION.....	51
	Data File Specification	51
	File Identification and Transport.....	52

List of Figures

FIGURE 1 – DUAP DATA FLOWS	2
FIGURE 2 – MAIN USER INTERFACE	10
FIGURE 3 – OBSERVATION COMPARISON.....	12
FIGURE 4 – AVERAGE SEGMENT SPEED.....	14
FIGURE 5 – DATA QUERY	15
FIGURE 6 – EXAMPLE DATA QUERY RESULTS	15
FIGURE 7 – ACCELEROMETRY TRACKING	16
FIGURE 8 – MITS CENTER VEHICLE DETECTION STATION LOCATIONS (AS OF 2007)	18
FIGURE 9 – SAMPLE MITS CENTER DATA (2007)	19
FIGURE 10 – SAMPLE CHRYSLER FAST FEEDBACK FLEET VEHICLE LOCATIONS	20
FIGURE 11 – SAMPLE CHRYSLER FAST FEEDBACK FLEET DATA.....	21
FIGURE 12 – MDOT PROBE VEHICLE FLEET COMMUNICATIONS	22
FIGURE 13 – EXAMPLE VEHICLE SPEED AGGREGATION	26
FIGURE 14 – PAVEMENT CONDITION DEMONSTRATION CONFIGURATION	30
FIGURE 15 – EXAMPLE PAVEMENT CONDITION DEMONSTRATION ROUTE.....	30
FIGURE 16 – MDOT PROFILOMETRY VAN - FRONT VIEW	31
FIGURE 17 – MDOT PROFILOMETRY VAN - REAR VIEW.....	31
FIGURE 18 – PAVEMENT DEFECT RESPONSE 1	32

FIGURE 19 – PAVEMENT DEFECT RESPONSE 232
FIGURE 20 – PAVEMENT DEFECT RESPONSE 333

Revision History

Revision	Issue Date	Status	Authority	Comments
01.00	2012.03.05	Draft	LTM	No comment.
02.00	2012.08.31	Final	LTM	No comment.

Electronic File

Saved As: 06022-rp901rpt0100_Final_DUAP_Project_Summ_Rpt

EXECUTIVE SUMMARY

The purpose of the Data Use Analysis and Processing (DUAP) project is to support the Michigan Department of Transportation (MDOT) and its partners in evaluating uses and benefits of connected vehicle data in transportation agency management and operations. As such, the project complements parallel efforts of MDOT, the U.S. Department of Transportation (USDOT), the Vehicle Infrastructure Integration Consortium (VIIC), and others to design and deploy the connected vehicle infrastructure, vehicle equipment, and initial applications. The DUAP project builds on that foundational work to investigate how the availability of data from connected vehicles throughout the road network may impact the ways MDOT and other transportation agencies do business.

DUAP research has, however, been constrained by the relative unavailability of connected vehicle data. The research was initiated and developed on a presumption of data being available from the efforts of other projects such as the Vehicle Infrastructure Integration (VII) Proof of Concept, many of which were not able to meet those expectations. Access to other potential data sources pursued by the project was constrained by data privacy and sharing policies, or simply by the cost of acquisition. In the end, the best data sources proved to be those developed by MDOT for deployment on its own fleet of vehicles.

The project has nonetheless successfully worked through a disciplined systems engineering process to develop and demonstrate three connected vehicle system capabilities.

- The DUAP *system* itself collects, aggregates, processes, and provides interactive views of the connected vehicle data. The system design is flexible and can accommodate data of varying types, dimensions, and resolutions.
- The DUAP data source *interfaces* demonstrate the ability to pragmatically acquire connected vehicle data from whatever sources may be available. Interfaces were developed for several specific data sources along with a standard DUAP interface.
- DUAP *applications* demonstrate the potential for enhancing DOT operations with connected vehicle data. Applications evaluated with data made available to the project include traffic monitoring, pavement defect and condition assessment, and origin-destination studies for planning.

From these, the project concludes that there is substantial potential for the use of connected vehicle data in transportation management and operations. The next phase of research will focus on development of reliable data sources and specific applications for implementation in MDOT.

ACTION PLAN FOR RESEARCH

The potential for vehicles to pass information about their operating state and environmental conditions to other vehicles and to the infrastructure has been a topic of transportation interest and research for many years. Such a capability could greatly facilitate safety, mobility, and environmental enhancements in transportation. The USDOT established formal initiatives to evaluate Vehicle-Infrastructure Integration (VII) in 2005.

Michigan was an early proponent of and partner in VII initiatives and in 2006-2007 developed the *Vehicle-Infrastructure Integration Strategic and Business Plan*.¹ That plan identified several areas of needed VII research including safety, traffic management, and asset management. The DUAP program was then established to facilitate and implement research that would evaluate and demonstrate the potential benefits of VII to agency operations.

¹ Michigan Department of Transportation. *Vehicle-Infrastructure Integration Strategic and Business Plan*. Lansing, Michigan. January 9, 2007.

1 INTRODUCTION

This document summarizes the results, accomplishments, and lessons learned from the Michigan Department of Transportation's (MDOT's) Vehicle Infrastructure Integration (VII)² Data Use Analysis and Processing (DUAP) project. It also includes recommendations for further research and development and provides an interface specification for potential data providers.

1.1 Objective

The *Vehicle-Infrastructure Integration Strategic and Business Plan* laid out a plan that focused on partnering, developing, and deploying a connected vehicle infrastructure and test beds; increasing safety and mobility; improving asset management; developing outreach programs to better expose others to connected vehicle concepts in Michigan; justifying the need for connected vehicle research and systems; and determining creative investment funding venues for connected vehicle activities.

Within this context, the DUAP project objectives were to investigate and evaluate the utility of connected vehicle data and its integration with other transportation agency sources in enhancing safety, increasing mobility, and improving asset management. Tasks within DUAP identified uses for the data, developed algorithms to use and process the data, developed prototype applications and data management software, and evaluated the utility of the processed data for MDOT and its partners. Data processing required acquisition from a variety of sources, standardization and integration, storage, synthesis for particular applications, and dissemination. Phase 1 of this work focused on collection and dissemination of information, and later phases will deal with development of applications that use the data collected by the system.

1.2 Scope

The purpose of the DUAP project is to support MDOT and its partners in evaluating uses and benefits of connected vehicle data in transportation agency management and operations. As such, the project complements parallel efforts of MDOT, the U.S. Department of Transportation (USDOT), the Vehicle Infrastructure Integration Consortium (VIIC), and others to design and deploy the connected vehicle infrastructure, vehicle equipment, and initial applications. The DUAP project builds on that foundational work to investigate how the availability of data from connected vehicles throughout the road network may impact the ways transportation agencies do business. The project focuses specifically on data uses and benefits in responding to safety concerns, managing traffic, and managing MDOT's transportation assets. The work also supports other connected

² The set of programs, technologies, and services referred to as "Vehicle Infrastructure Integration" at the DUAP program's inception has been programmatically repackaged several times. References in this and any supporting documents to "VII" and "IntelliDriveSM" can be understood to be synonymous with "connected vehicle" for technical purposes. "Connected vehicle" will be the preferred use throughout this report, except when referring to specific programs, projects, and documents.

vehicle projects, technology development for MDOT, and economic growth for the state.

The key tasks within the DUAP project are:

- to identify uses for the connected vehicle data;
- to develop algorithms to use and process the connected vehicle data;
- to develop prototype applications and data management software; and
- to evaluate how well the data and the algorithms function in a department of transportation.

As illustrated in Figure 1, the DUAP system is intended to draw data from existing MDOT data sources and other relevant data sources to be integrated with connected vehicle data. The integrated system output would be returned to the existing MDOT applications as an enriched data stream or could be used in new applications for MDOT. Other MDOT projects may have influence on or facilitate the data integration. Applications outside MDOT could get data through a new MDOT gateway application.

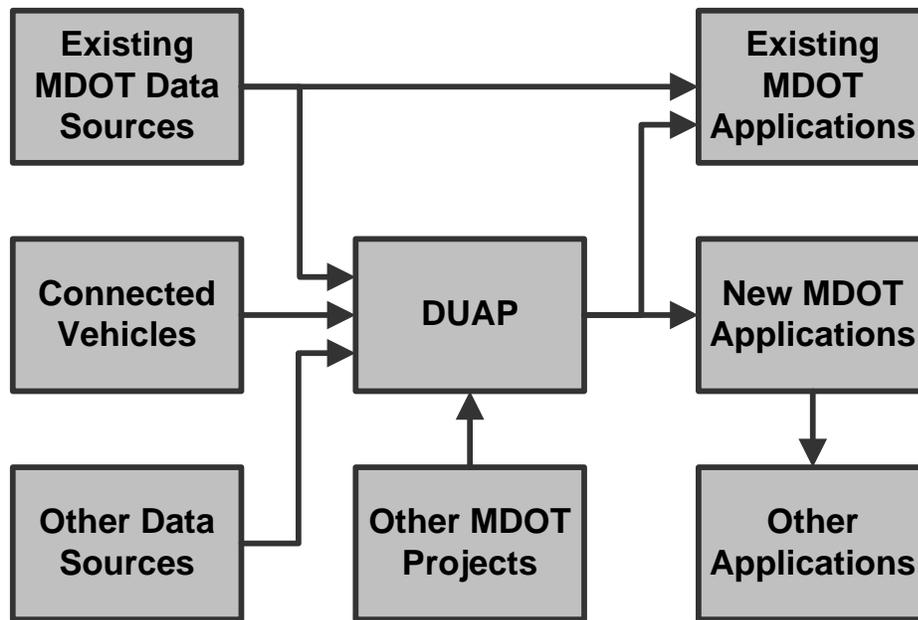


Figure 1 – DUAP Data Flows

2 METHODOLOGY

2.1 *Systems Engineering*

It has become widely accepted that system development projects should follow a disciplined systems engineering process from planning through operations. Indeed, the use of a standard systems engineering process is a requirement of Title 23 of Code of Federal Regulations (CFR) Section 940.11, which defines eligibility for Federal Intelligent Transportation System (ITS) funding. The systems engineering approach used in DUAP development follows the USDOT's definition of systems engineering and consists of developing a concept of operations, generating requirements, synthesizing a system architecture, designing the system and components, testing, deployment and ongoing operations. Results of this process are captured in a series of systems engineering documents.

The DUAP systems engineering process started with determining the scope and objectives of the system through extensive stakeholder meetings. The initial meetings included representatives from across the MDOT organization, partnered Michigan agencies, USDOT, academia, transportation consulting firms, and technology vendors. These gatherings identified opportunities and challenges in transportation operations and emerging connected vehicle technologies, with the goal of finding a set of high-value applications for DUAP demonstrations. This process also identified MDOT strategies and operational policies that might enhance or constrain the scope of DUAP development. Follow-up meetings and research identified existing MDOT systems and data for which DUAP interfaces would need to be created.

The Concept of Operations (ConOps) captures the results of the stakeholder interactions and analyses to determine what the system should do. It describes the existing systems relevant to transportation operations data; establishes the case for change; describes the capabilities and features of the proposed system; and discusses scenarios illustrating future system operation. The document is used to facilitate communications on connected vehicle and transportation operations with agency stakeholders and partners and as a basis for generating more detailed user needs and expectations for connected vehicle data, use, and processing.

The System Architecture Description (SAD) identifies the system's components and describes its internal and external interfaces. It expands on the initial system concept as described in the ConOps to provide more detailed descriptions of the system interfaces and internal structures. The description is provided in the form of several views of the architecture: an operational view based on user and system interactions; an information view focused on the flow and storage of data through and within the system; a decomposition view describing the functions to be performed by the system components; a hardware view of the system and communications interfaces; a technology view focused on standards and specifications; and a view relative to the National ITS Architecture.

The System Requirements Specification (SRS) is generated from user needs described in the ConOps and analysis of the system interfaces as described in the

System Architecture Description. The SRS contains descriptions of required system functions for each of the system's major components; design constraints presented by policy and standards; quality characteristics; and any external factors that may impact the system design. The document also provides traceability to the underlying source of the requirements, generally to user needs in the ConOps.

The System Design Description (SDD) documents the intended system implementation. In the process applied for DUAP, the documentation takes the form of a set of Unified Modeling Language (UML) diagrams for object-oriented designs. The process and documentation address each component and its interfaces in detail.

System testing provides assurance that the system as described in the SDD meets its requirements as described in the SRS. Unit testing is performed on individual system components as they are developed. Integration testing addresses the interfaces between components and is performed as the system is assembled. Full system testing and user acceptance testing demonstrate the completed system.

2.2 *Data Interfaces*

The DUAP project was based on the presumption of data being available from the efforts of other connected vehicle projects. Activities pursued as part of DUAP were intended to define and develop applications of that data within MDOT and other transportation agencies.

DUAP research and systems development paralleled the development of the data sources by USDOT, the VIIC and others. The means of collecting and presenting data were coordinated with those developments and implemented on an accelerated schedule. Unfortunately, the data sources originally envisioned for DUAP did not become available as expected.

The IntelliDrive Proof of Concept (POC) schedule was focused on an IntelliDrive deployment decision to have been made by USDOT in 2008. Its original program milestones were met only up to the point of prototype implementations. As the schedule slipped, applications were reprioritized. Eventually all actual POC prototype applications that might have been used by MDOT as data sources for DUAP were dropped by USDOT from the formal demonstration due to schedule overruns. Probe data for a few days of operations was made available to DUAP only after the completion of the POC tests.

In parallel with the POC, MDOT pursued procuring data from other sources.

- Fixed vehicle detection station observations of speed, volume and lane occupancy were made available from the Michigan ITS (MITS) Center for comparison with any probe vehicle data that might become available.
- Discussions were held with several automakers to assess opportunities for demonstrating probe data gathering. While discussions were informative and all parties expressed interest, potential agreements were complicated by data security concerns and undercut by the economy.

- Arrangements were eventually made with Chrysler and MTS Technologies to obtain data from Chrysler's field diagnostic fleet of employee-operated vehicles. Data has been intermittently available from Chrysler vehicles and loaded into DUAP since 2008. Its usefulness has been constrained by the limited data types and complete anonymity of the data.³
- Discussions were also held with Orbital to obtain data already being collected for the Southeast Michigan Snow and Ice Management (SEMSIM) system. The system did not have an existing interface for exporting data and would not have provided enough data to justify the needed enhancements, so it was decided to pursue other sources.

In light of these obstacles, it was determined that the only recourse for demonstrating applications for MDOT would be to gather MDOT's own data. Provisions were made for collecting data from MDOT's own fleet vehicles across the state. Two alternative systems were used to collect the data from MDOT vehicles: a custom on-board unit using Wi-Fi hotspots at MDOT facilities for data download; and an Android application running on a commercial 3G smartphone. The MDOT Infrastructure Monitoring Data Management project evaluates the ability of sensors to deliver infrastructure data from the Cut River Bridge in a remote location of the Michigan upper peninsula to the south tower of the Mackinac Bridge. Infrastructure remote sensing and monitoring systems will support DUAP by providing data that may support applications defined by DOT user needs. These prototype systems have enabled the DUAP program to demonstrate the ability to collect data and provide applications specifically related to the improvement of DOT operations.

2.3 *Applications Analysis*

The objective of the application analyses, as a subset of DUAP's larger goals, is to investigate and, if possible, demonstrate enhanced transportation agency operations based on connected vehicle systems and data. Within the frame of that objective, each application area has particular objectives for safety, mobility, environmental affect, cost, and timeliness. These objectives are likely to be driven by established agency performance measures for that application area. Objectives may also relate to the existing agency systems—providing a particular kind or format of data to fulfill regulatory, policy, or procedural goals.

The concept for each application area starts and ends with the needs of the particular application, but is built around the DUAP system itself and a common set of connected vehicle capabilities and systems. The needs of a particular application derive largely from the objective served by that application, its associated performance measures, and the physical and operational nature of the application. At a conceptually high level, some applications are focused on a vehicle's behavior in the transportation system and data pertaining to vehicle

³ Some applications—compiling origin-destination pairs for planning studies, for example—are not possible with completely anonymized data.

location and movement may be all that is needed; other applications may need data on the roadway environment itself, and it may be necessary to derive or infer conditions from observations of the vehicle's interfaces with the roadway.

The application research consists of three phases, each of which serves as a gate for further development and testing. The initial system *development* testing asks "Can the system collect usable quality data?" A successful development test then confirms accurate data collection and synthesis into usable metrics. *Acceptance* testing asks "Are the data reflective of actual conditions?" Successful acceptance testing demonstrates correlation of metrics with pavement conditions and traditional performance measures over known test segments. *Application* testing asks "Can the data collection be used to enhance DOT operations?" Successful application testing finally demonstrates the ability to consistently and accurately identify field pavement conditions of interest from probe data.

Development testing is used to conform and qualify the data gathering process. As noted earlier, the form, scope and quality of data to be gathered from vehicles is part of the research. Once data are available and have been successfully collected by DUAP, the data are analyzed for usefulness in particular applications. Key issues for being able to use the data may include its temporal resolution, precision, accuracy, consistency, and reliability. The potential issues will be assessed for individual devices, vehicles, and for the fleet. It may be necessary to modify device installation and data gathering configurations on the vehicle to facilitate particular applications.

Acceptance testing is used to correlate probe data with known conditions and standard measurements. Once the base data sets are available, data are collected with the new techniques over the same areas and conditions as the base sets. Testing varies the collection parameters—for example, vehicle and vehicle type, speed, suspension tuning and tire condition. Parametric studies are then used to create algorithms for generating performance metrics based on probe data for correlation with standard metrics. Data collection on other test areas with known conditions would then be used to validate algorithms and correlations.

Application testing is used to demonstrate the ability to generate useful data to facilitate DOT operations. The application would be deployed throughout the DOT's areas of interest to gather data for monitoring condition trends and anomalies. Data would be processed by methods determined in acceptance testing. When anomalies are found, DOT staff would be notified to confirm the conditions detected by the probe data. For ongoing condition monitoring, metrics derived from probe data would be compared to those obtained in normal monitoring procedures.

The analysis of each application is completed with conclusions on the relative success in meeting the objectives for that application and suggestions for next steps in research and implementation.

3 RESULTS AND ACCOMPLISHMENTS

The efforts of gathering requirements from stakeholders, performing detailed analysis, and gathering and evaluating the practical application of gathered data resulted in a software suite that is the DUAP system. The DUAP system collects data of interest from participating data sources and processes those data into useful information that can be visually presented to decision makers.

3.1 DUAP System

As was shown in Figure 1, the DUAP system collects data from many different sources, stores those data, and applies algorithms to the available data—new and previous—in order to support existing and new MDOT applications. The output from the DUAP system can be used as data for yet another system to process or as an application in a geographic context supported directly by the user interface described in this section.

The DUAP system is built upon a foundation of modular computing blocks. This modular approach provides flexibility to the system in that all possible data sources, analysis algorithms, and applications are indeterminate. Constructing the DUAP system with a singular purpose from what was known at the time would have potentially limited its future application and increased the costs associated with updating it.

3.1.1 Data Collection

For the DUAP system to be able to perform useful work, it must have data. The data sources currently used by the DUAP system are described in Section 3.2.

The design of the DUAP system necessarily accommodates various data sources through its modular architecture. No single data representation mechanism exists that will accurately convey information in every possible instance. The best approach is to create collection software that can deal with source data in its native format and copy what it defined as relevant to the DUAP system.

Evidence strongly suggests that the most popular data formats for information exchange are text-based—humanly-readable letters and numbers. Both character-delimited text and XML are text-based data exchange mechanisms. The former typically defines columns by a header, with data separated by a special character with comma, semicolon, tab, and space being popular choices. The latter uses tags formatted to the XML standard specification to identify data elements.

Even with common data exchange formats consisting of human-readable characters, data content arrangement can vary considerably. One delimited text format might contain a single unique identifier column with its associated data located in several adjacent columns within the same record. Contrast this with another delimited format that identifies all data in three columns where the first two columns represent the unique identifier and data type and the third column contains the data value. One might refer to these two approaches as short-and-wide versus tall-and-narrow.

XML data sources can easily be classified as anything goes. As long as the structure conforms to the published standard, elements may be labeled anything, located anywhere, and use any defined units necessary to describe data.

Even though text-based data exchange formats are popular, many data sources convey their information in native computer data structures for compactness and processing speed. These data formats are arguably as complex and diverse as their XML counterparts with the added complexity of not being easily read and understood by people.

The DUAP system, through its modular data collection approach, is capable of handling the heterogeneous nature of existing and potential data sources:

- MITS center loop detector data arrive in a delimited format that identifies speed, volume, and occupancy at a particular time for each road loop.
- Chrysler test-fleet data arrive in the tall-and-narrow style of delimited text file.
- MDOT test fleet accelerometry data arrive in multiple file formats. One delimited text file for CAN bus data, a second delimited text file for GPS location, and a third binary file containing 100 Hz 3-axis acceleration data.
- Android accelerometry data are similar to MDOT test fleet accelerometry data, but arrive in a single delimited text file.

The purpose of the DUAP system data collection modules is two-fold: to gather data of interest from participating sources and to extract those data for homogenous storage and retrieval by application modules to process.

The collector modules can retrieve data using different network protocols. HTTP and FTP are the most common. When data are collected, they are stored in a file and archived with a timestamp to maintain the source data in its original format. As data are being copied to the archive file, they are also being processed in their native format and parsed into the DUAP system storage structure. The storage architecture is described in the next section.

3.1.2 Storage

Information can be considered complete when it answers six questions: who, what, how, where, when, and why. The DUAP system has adopted a data model that fulfills the answers to these questions. The fundamental unit of data within the DUAP system is called an observation. An observation can briefly be described as a direct association of a specified measurement retrieved from a source for a location at some time with a magnitude. An observation directly addresses five of the six questions for information completeness:

- Who – each observation keeps track of the source of information. The source is usually called a contributor and can be an organization or data system operated and maintained by an organization.
- What and How – each observation is associated with a type, and the type defines what and how. For example, a temperature type could be defined

that indicates the measurement was made using a Pt1000 sensor and the units are Celsius.

- Where – this could be defined as any location within the context of the DUAP system. In this case the location is defined by geo-coordinates that include latitude and longitude recorded in microdegrees and elevation to the nearest decimeter.
- When – this is the time a measurement was taken. It is measured in milliseconds and is relative to Coordinated Universal Time (UTC). It is possible to use future times to represent predicted states.

Answering the why question is the purpose of the DUAP system. Applications use the stored observations to derive new observations and provide the context in which the observations can be compared to infer the reasons for any given transportation state. If a set of vehicles are presented as having a speed of zero for a few minutes on an arterial road segment, an application could infer that the reason is that the traffic signal was presenting the stop condition. If the queue persists for many minutes, an application might infer there is a collision or a stalled vehicle as the reason.

The DUAP system data collectors read and process data from each source in its native format. The data are transformed into observations based on the definition in this section. An archival process also stores the collected data in its native format in a file that is tagged with the source and collection time for future retrieval and verification needs.

The DUAP system applications provide measurement type, and geographic and time context. To this end, data are indexed by type, geographic location, and time as they are converted to system observations. These indexes greatly improve processing speeds so that there is very little delay between new data being received and an application being able to inform a user of any resulting transportation system states.

3.1.3 User Interface

The user interface is accessible using a common Internet browser that can interpret and display HTML 4 and JavaScript1.8.5/ECMAScript 5. The interface provides the tools needed to visualize DUAP application data in a geographic and time-related context. There are four directly supported applications: side-by-side comparison, segment average speed, sensed accelerometry (ride roughness), and potential pothole location.

The current URL for the DUAP system is <http://duap.mixonhill.com>. However, the system can be hosted on a variety of web servers and deployed with varying domain names as needed by the particular deployment.

The main user interface shown in Figure 2 is the default presentation when accessing the system from the deployed URL. Interface inputs and controls allow the user to display system information in a geographic context for a particular time range. Consequently, the input controls allow the selection of time zone, start time, end time, observation type, and application layer.

The geographic context is provided by the large map area. The Google Map application programming interface (API) is used to provide the map functions. Map controls allow panning and zooming in and out, and to display a plain map, satellite image or combination background. In some instances the satellite background provides more clarity for viewing application information.

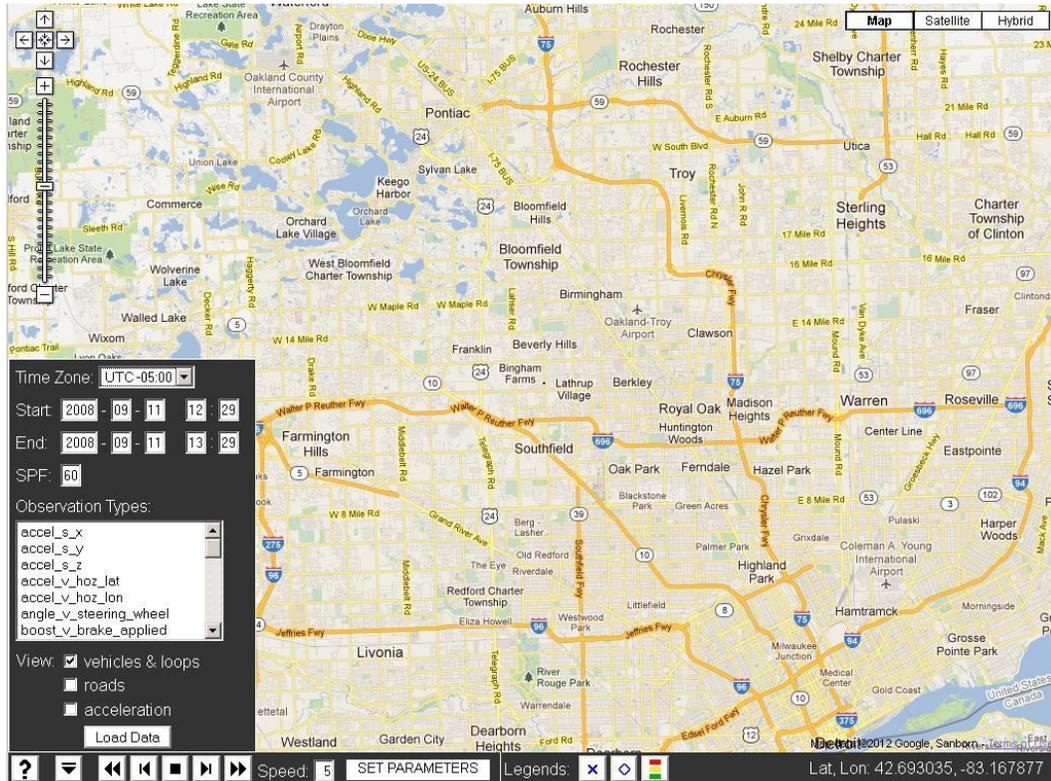


Figure 2 – Main User Interface

In the lower-left corner is the query selection dialog. Just below it in the bottom border is a question button and down arrow button. Clicking on the question button displays a pop-up window with additional instructions on how to use the interface. The down arrow button toggles the query dialog on and off to maximize the map viewing area after data have been selected.

The top third of the query dialog deals with time parameters: time zone, start time, end time, and seconds per frame.

Associating time with data can be complex, but users generally work to their current local time. The time zone drop down list attempts to accommodate some of the complexity arising from the need to display data that may have been recorded in a different time zone or while a different daylight saving policy was in effect. The control value defaults to the time zone defined by the browser, but it also contains offsets for Eastern and Central Standard Time, and allows for Daylight Saving Time.

The start time and end time controls are straight forward. The default values set the time window to one hour ending 24 hours earlier than the current local time as

determined by the browser. Both sets of controls allow for the selection of year, month, day, hour, and minutes. Hours are reported in 24-hour format where midnight is 00, noon is 12, and 11 p.m. is 23.

The SPF control is a seconds-per-frame setting with a minimum value of 1 and a maximum value of 60. This correlates to the concept of a time window that can be as narrow as one second or as wide as one minute. This construct enables data with different time resolutions to be displayed in a common presentation. The DUAP system collects observations that do not necessarily occur at exactly the same instant. Without the SPF setting, only a few data values would be presented together on the map. The data values would appear to pop around the map and not provide a very useful comparison.

Generally a 60 second time window works well for observation type comparisons. Consider one vehicle reporting an air temperature value at 12:01:22 and another vehicle reporting an air temperature value at 12:01:29. A one-minute time window enables both values to be displayed for the 12:01 time interval.

A one-second time window setting is better suited for higher frequency data collection such as vehicle accelerometry. The accelerometry and pothole application is described later in this section.

The middle third of the query dialog holds the observation type multiple selection control. The observation types currently understood by the DUAP system are listed here. Up to six different observation types can be selected at one time. Single selections are made by clicking the mouse on an individual entry. Multiple selections can then be made by holding down the control key and making other individual selections, or by holding down the shift key and making an adjacent range selection.

The bottom third of the query dialog contains check boxes for choosing which application to activate and a button to retrieve data applying the user-configured parameters. For the observation comparison application, only the vehicles and loops check box should be checked. Pressing the load data button executes the query. If successful, icons with their associated values are displayed on the map as illustrated in Figure 3. Otherwise, a no data message is displayed in the status box in the bottom control bar.

The bottom control bar is complementary to the query dialog. The query dialog accepts data filtering parameters and the control bar enables the interaction with the retrieved data filtered by the query.

The control bar consists of five sections. The two left-most buttons correspond to displaying the help dialog, and toggling the query dialog between its hidden and shown state with the purpose of maximizing the viewable map area.

To the right of the status box are the legend controls. The “X” button corresponds to mobile data, such as vehicles, and the “O” button corresponds to stationary data, such as road loops. The rainbow button is for road segment speed definition.

Each legend button toggles the hidden or shown state of a legend dialog. Once opened, a legend dialog can be closed by clicking the button in the dialog or on the control bar. Legend dialogs automatically arrange themselves from left to right at the bottom of the screen relative to their activation order.

In the case of executing a query for comparing observations, the legend dialogs serve an additional purpose. When more than one observation type is selected, up to a maximum of six, observation values surround each displayed icon starting at the 2 o’clock position and proceeding counter-clockwise. When the corresponding legend dialog is active, it also presents the observation type that occupies each of the value positions around the icons. When observation types are selected that possess similar numeric ranges, the legend dialog provides additional context.

The bottom right control displays the current latitude and longitude coordinates of the cursor on the map. This information is provided so the user can get better sense of relative scale and distances to known landmarks.

Segment average speed is the next application to be described. Road segments are defined in the DUAP system using the Michigan Geographic Framework. Segment average speeds are calculated at five-minute intervals and are color coded red, yellow, and green. Using this application only requires that at least one observation type be selected and a time range be specified, along with clicking the box next to the roads label in the query dialog.

Segment average speeds have more visual contrast if the map is placed into satellite mode. Yellow road segments are easier to discern from the gray and dark green background. It is also recommended, although not necessary, to select vehicle speed—“speed_v”—as at least one of the observation types for the query and also select the vehicle and loops application. While other observation values besides vehicle speed can be displayed in relation to segment average speed, not all of them necessarily make useful sense. Figure 4 provides an illustration of this display.

Since the seconds-per-frame setting has a maximum of 60 seconds and segment average speeds are five-minute intervals, iterating through data sets displays any given group of segment average speeds for five steps before they potentially change color. Selecting the vehicles and loops application and vehicle speed displays vehicle speeds on the color-coded segments.

Some road segments are not colored. At least two data points must be available on any given segment to calculate a five-minute average speed. Uncolored segments have less than the minimum number of observations required for the calculation. Some segments may be colored but not appear to have any vehicle associated with the color. This is another result of a five-minute average speed overlaid with one-minute vehicle data, i.e. only one data point may be coincident giving the appearance that a vehicle is not present.

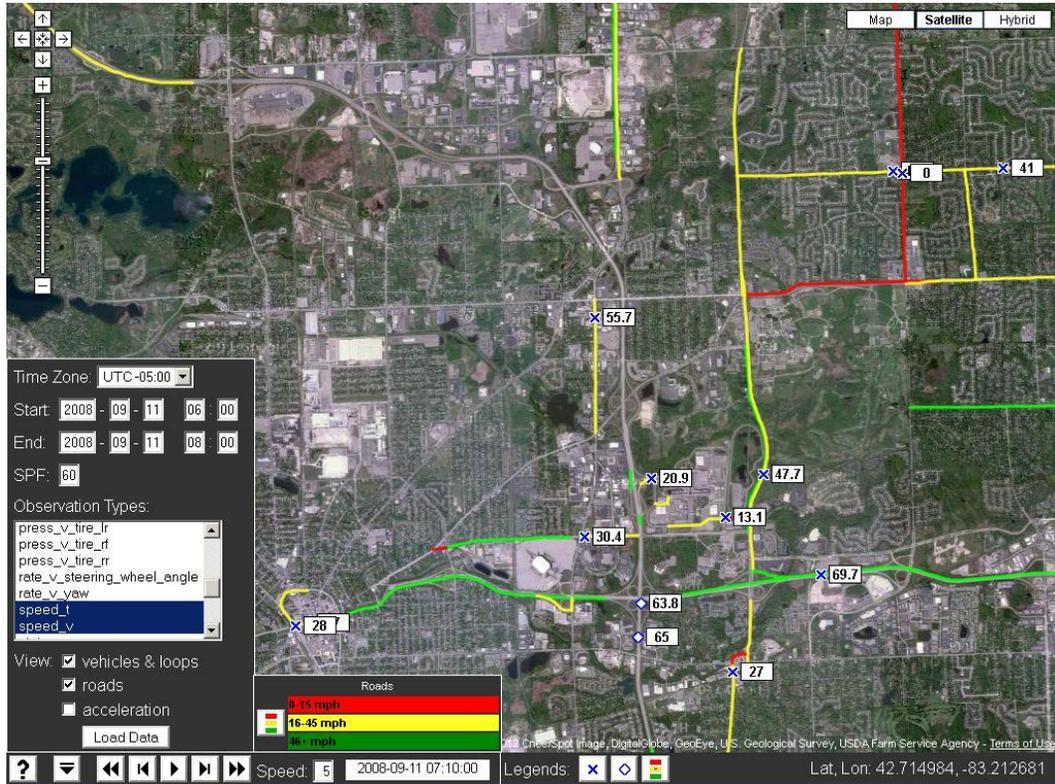


Figure 4 – Average Segment Speed

In addition to the map presentation, the DUAP system user interface also has a direct data query capability. Where the main user interface is reached using the system URL, the data query user interface is reached through the system URL plus the resource name, “/DUAP_Data_Dump.html”. The name is case-sensitive. The current data query location is

http://duap.mixonhill.com/DUAP_Data_Dump.html .

The data query interface is intentionally simple, as shown in Figure 5. It consists of a query dialog similar to the default map interface, but without the map. The parameter inputs function exactly as they do for the default map interface described earlier in this section. Without the map to provide geographic context, however, a latitude and longitude bounding region must be defined using the latitude and longitude range inputs.

Once the query parameters are set, clicking the request data button executes the query against available data and presents a link to download the results of the query. Clicking the link should initiate the browser prompting the user to download a file with a default name to a default location or a name and location chosen by the user.

Figure 5 – Data Query

The downloaded file is a standard compressed zip file of a single comma separated value file that contains the results of the query. The CSV file may be extracted from the compressed archive and then opened by common text editors or spreadsheet software such as Microsoft Excel as shown in Figure 6. The data may also be loaded into other applications for further processing as desired.

Observation Type	Time	Latitude	Longitude	Value
speed_t	9/11/2008 19:30	42.488699	-83.491821	67.1607
speed_t	9/11/2008 19:30	42.22887	-83.383677	8
speed_t	9/11/2008 19:30	42.249585	-83.308524	63.5
speed_t	9/11/2008 19:30	42.377717	-83.229115	61.2927
speed_t	9/11/2008 19:30	42.430449	-83.219248	60.6889
speed_t	9/11/2008 19:30	42.378891	-83.198355	57
speed_t	9/11/2008 19:30	42.31508	-83.169886	68.6129
speed_t	9/11/2008 19:30	42.416326	-83.086488	64.717
speed_t	9/11/2008 19:30	42.494262	-82.953798	67.7907
speed_t	9/11/2008 19:30	42.48397	-83.445853	65
speed_t	9/11/2008 19:30	42.382851	-83.353026	73.5476
speed_t	9/11/2008 19:30	42.385644	-83.257416	68.4793
speed_t	9/11/2008 19:30	42.378204	-83.229067	58
speed_t	9/11/2008 19:30	42.276348	-83.203231	79.5319
speed_t	9/11/2008 19:30	42.379401	-83.198364	57.56
speed_t	9/11/2008 19:30	42.38542	-83.15931	66.0667
speed_t	9/11/2008 19:30	42.491523	-83.031649	60.0638

Figure 6 – Example Data Query Results

The last application to be described is the accelerometry tracking application shown in Figure 7. This application presents vehicle icons, a second-by-second vehicle path, and three-axis accelerometer data alongside potential pothole locations.

In this application it is best to select only the `accel_s_x`, `accel_s_y`, and `accel_s_z` observation types. Only the acceleration application should be checked. The time range can be chosen normally. Accelerometer data arrives at a minimum of 100 Hz, so a 1 second per frame (SPF) setting is also necessary. These recommendations are made as the accelerometry represents a considerable volume of data in a short time period and can quickly and easily overwhelm network connection and web browser capabilities.

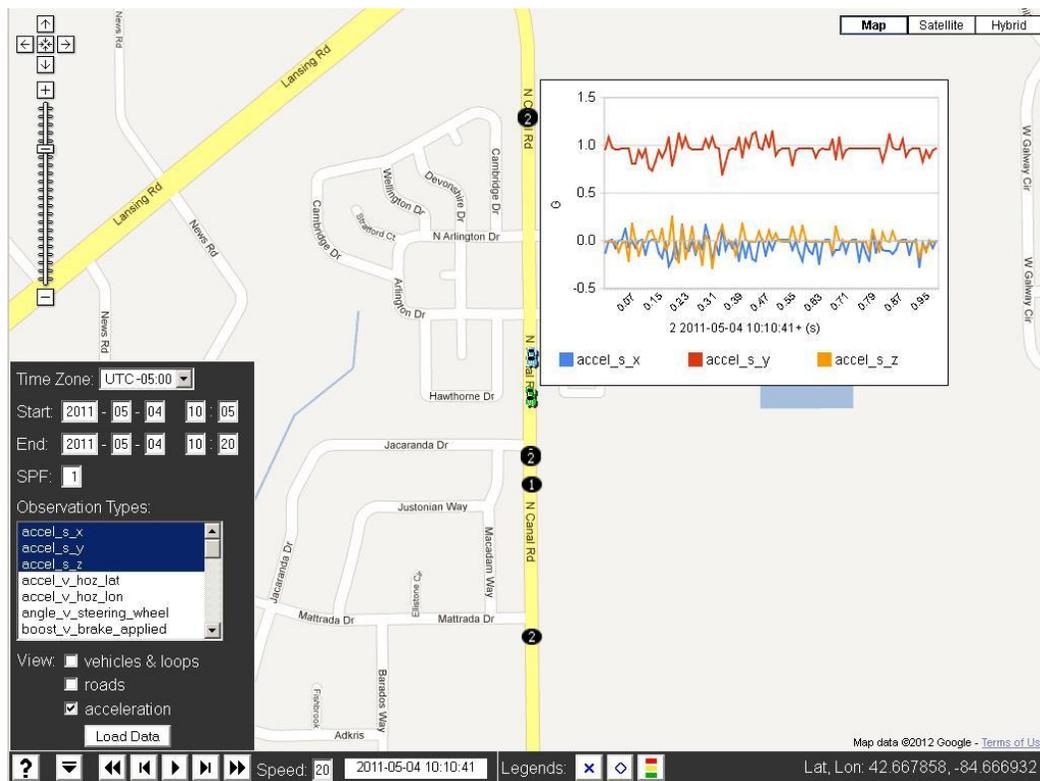


Figure 7 – Accelerometry Tracking

If acceleration data are available for the given time range, some vehicle icons and pothole icons may appear. If no vehicle icons appear immediately, it is likely that the start of their path is slightly later than the beginning of the specified time range. Playing through the available data presents a vehicle icon at some point. Pothole icons may not appear if the time range is short and the road segments were relatively smooth for the same period.

Up to nine different vehicles can be displayed with potential potholes numbered for each vehicle path. More than nine vehicles and pothole sets are possible, but the pothole numbering is limited to a single digit and will start over, making it difficult to determine if two potholes with the same label are from different vehicle paths.

It is possible to zoom out far enough on the map to view all of the vehicles with a good portion of their path at the expense of detail. A typical, and better, approach is to select a specific vehicle and zoom in closer for more detail. Selecting one particular vehicle icon changes its color from blue to green. A vehicle icon can be unselected by clicking the same vehicle icon again, or selecting another vehicle.

Selecting a specific vehicle icon centers the map display on that vehicle icon for easy tracking and displays a one-second graph of that vehicle's acceleration data. If the map is zoomed in far enough, a red line centered underneath the vehicle icon that represents the path of the vehicle for that one-second interval is visible. Stepping through the data in this mode is quite informative, allowing the user to see which acceleration axes were affected and to what extent when a potential pothole is encountered.

3.2 Data Sources

The purpose and general design of DUAP's data collection was described in the previous section. This section describes the specific DUAP data sources, the geographic extents to which they apply, the collection mechanism, and file formats.

3.2.1 MITS Center

Data provided by the Michigan Intelligent Transportation Systems (MITS) Center from its vehicle detection stations are brought into DUAP as known background data for correlation with vehicle probe data. Speed, volume, and lane occupancy are provided from all vehicle detection stations every minute and stored in the system's observation cache. Locations of the loops, shown in Figure 8 below on a map of the Detroit metro area, are stored as metadata and saved with the data in the cache.

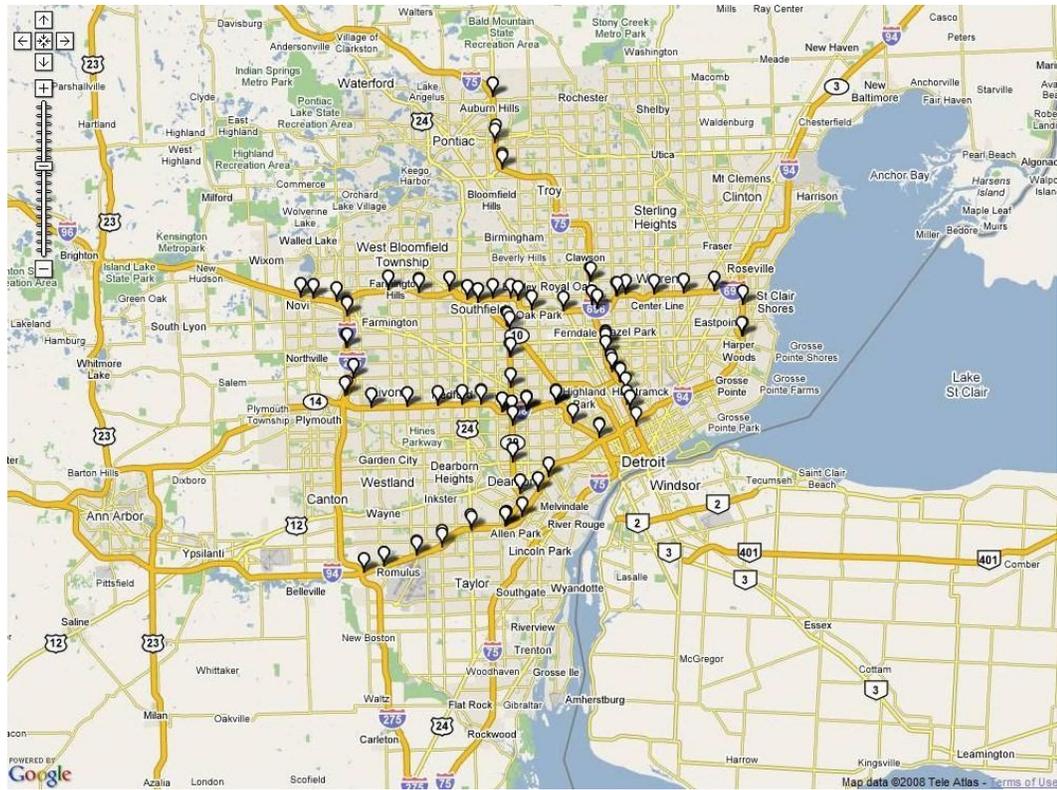


Figure 8 – MITS Center Vehicle Detection Station Locations (as of 2007)

MITS center data are retrieved from the MITS Center server by a DUAP file transport protocol (FTP) request once every minute. As illustrated in Figure 10, data consist of records of date and time, detector station identifier, detector type, volume (vehicles per minute), occupancy (percent), and speed (miles/hour).

2008-03-08	17:11	721156	77	38.0	7.0	64.9737	
2008-03-08	17:11	1114312	77	66.0	8.7	68.1061	
2008-03-08	17:11	852058	77	42.0	2.7	89.2381	
2008-03-08	17:11	786632	77	56.0	10.8	54.875	
2008-03-08	17:11	458812	77	34.0	2.7	68.3529	
2008-03-08	17:11	524528	77	61.0	2.6	67.7541	
2008-03-08	17:11	1441836	77	28.0	3.3	61.3214	
2008-03-08	17:11	721216	77	35.0	4.3	63.0286	
2008-03-08	17:11	524608	77	70.0	4.2	69.9143	
2008-03-08	17:11	721176	77	14.0	5.0	69	
2008-03-08	17:11	131528	77	35.0	6.0	65.1714	
2008-03-08	17:11	131348	77	14.0	2.0	74.7143	
2008-03-08	17:11	851998	77	65.0	5.3	75.3077	
2008-03-08	17:11	524368	77	30.0	3.0	54.2667	
2008-03-08	17:11	459072	77	75.0	3.2	55.92	
2008-03-08	17:11	786652	77	4.0	2.0	52	
2008-03-08	17:12	66056	77	32.0	6.0	67.75	
2008-03-08	17:11	66056	77	19.0	3.0	77.7368	
2008-03-08	17:11	131488	77	35.0	5.5	97.4	
2008-03-08	17:11	1179968	77	50.0	4.3	72.16	
2008-03-08	17:11	1114232	77	34.0	6.0	68.2941	
2008-03-08	17:11	918624	77	37.0	3.3	74.7297	
2008-03-08	17:11	721116	77	76.0	10.0	57.7105	
2008-03-08	17:12	131478	77	12.0	5.0	60	
2008-03-08	17:11	852008	77	24.0	5.0	71	
2008-03-08	17:11	65996	77	4.0	1.0	63	
2008-03-08	17:11	66016	77	10.0	4.0	83	
2008-03-08	17:11	131478	77	40.0	6.5	66	
2008-03-08	17:11	131463	77	36.0	5.0	66.1111	
2008-03-08	17:11	1179928	77	34.0	6.0	66.8824	
2008-03-08	17:11	524348	77	22.0	3.0	49.4545	
2008-03-08	17:11	1376336	77	25.0	2.0	68.08	
2008-03-08	17:12	66016	77	29.0	5.0	79.2069	
2008-03-08	17:11	918604	77	16.0	4.0	67	
2008-03-08	17:13	131348	77	20.0	3.0	71.85	
2008-03-08	17:13	852048	77	43.0	4.5	73.093	
2008-03-08	17:13	1441856	77	23.0	1.7	62.3478	
2008-03-08	17:13	524608	77	68.0	4.0	68.3824	
2008-03-08	17:13	131338	77	11.0	4.0	66	

Figure 9 – Sample MITS Center Data (2007)

3.2.2 Chrysler Fast Feedback Fleet Input

The Chrysler data being read into DUAP originate with the Chrysler Fast Feedback Fleet. Data from this fleet is collected for diagnostic purposes from Chrysler LLC vehicles driven by Chrysler employees and is aggregated from the vehicles at access points around Chrysler facilities in southeast Michigan. MTS Technologies provides the on-board units that extract the data from the vehicle’s controller network bus and provides the network server(s) that publish the data to both Chrysler and DUAP.

The data received from the Chrysler fleet is anonymized⁴—that is, it does not contain any association of a particular vehicle with any of the data points. Figure 10 illustrates locations of probe data points for a single 60-second interval. It further does not directly relate observations of various types as being from a single vehicle, or relate a time sequence of observations of the same type as coming from a single vehicle.

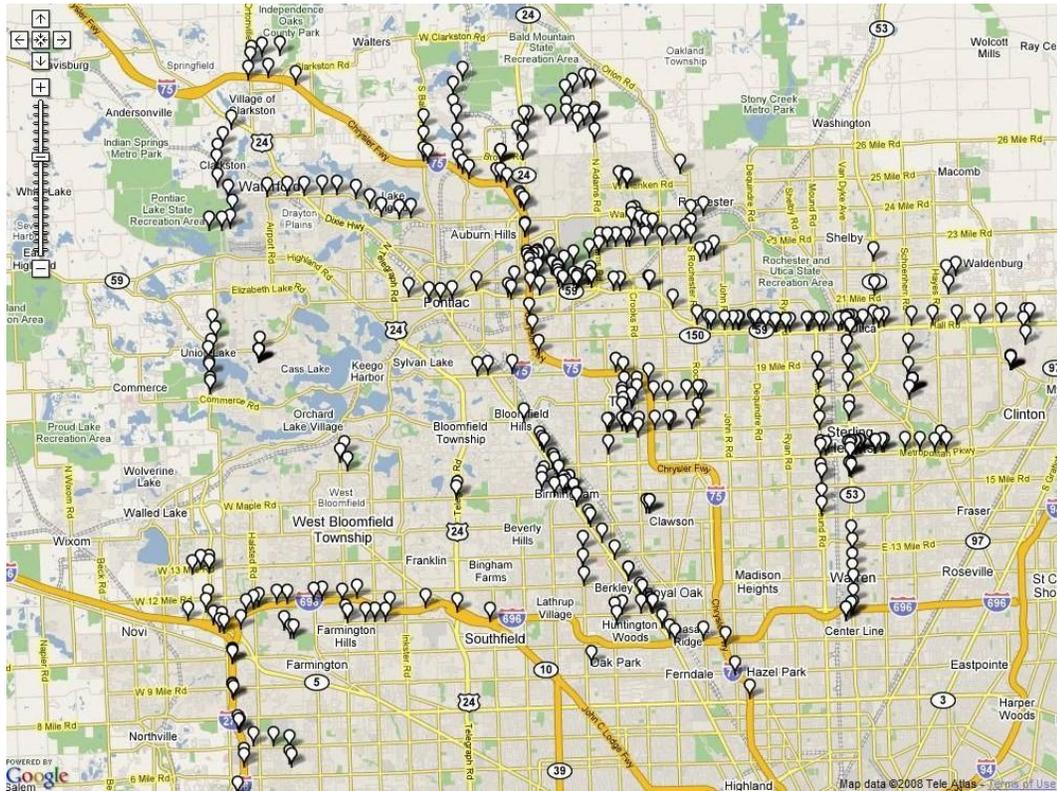


Figure 10 – Sample Chrysler Fast FeedBack Fleet Vehicle Locations

Data provided by Chrysler vehicles includes:

- Latitude
- Longitude
- Date/time
- Vehicle speed
- Ambient air temperature
- Barometric pressure
- Front windshield wiper status

Individual data records consist of latitude, longitude, date/time, and value for one of the observation types. Each type of observation can have multiple variations within the file, depending on the metadata associated with a particular vehicle's data acquisition application. In the Figure 11 example, vehicle speed records are labeled "Wheel-Based Vehicle Speed" or "Vehicle Speed – FDCM4". Units for the observation types likewise vary, but are consistent for a particular label.

⁴ As of 2008.01.25.

2008-03-09	08:00:08.0	Wheel-Based Vehicle Speed,76.109375,+42313295,-083069411
2008-03-09	08:00:08.0	Wheel-Based Vehicle Speed,76.109375,+42313295,-083069411
2008-03-09	08:00:08.0	Barometric Pressure,28.792,+42313295,-083069411
2008-03-09	08:00:08.0	Ambient Air Temperature,17.6,+42313295,-083069411
2008-03-09	08:00:08.0	Front Windshield Wiper Status,0,+42313295,-083069411
2008-03-09	08:00:12.0	Wheel-Based Vehicle Speed,71.796875,+42001779,-084583032
2008-03-09	08:00:12.0	Barometric Pressure,29.028,+42001779,-084583032
2008-03-09	08:00:12.0	Ambient Air Temperature,21.2,+42001779,-084583032
2008-03-09	08:00:12.0	Front Windshield Wiper Status,0,+42001779,-084583032
2008-03-09	08:00:31.0	Wheel-Based Vehicle Speed,31.125,+42440625,-083209968
2008-03-09	08:00:31.0	Barometric Pressure,28.969,+42440625,-083209968
2008-03-09	08:00:31.0	Ambient Air Temperature,14,+42440625,-083209968
2008-03-09	08:00:31.0	Front Windshield Wiper Status,0,+42440625,-083209968
2008-03-09	08:00:35.0	Wheel-Based Vehicle Speed,85.65625,+42410616,-083146447
2008-03-09	08:00:35.0	Barometric Pressure,28.674,+42410616,-083146447
2008-03-09	08:00:35.0	Ambient Air Temperature,15.8,+42410616,-083146447
2008-03-09	08:00:35.0	Front Windshield Wiper Status,0,+42410616,-083146447
2008-03-09	08:00:37.0	Ambient Air Temperature - FDCM4,-9.5,+42238112,-083201695
2008-03-09	08:00:37.0	Vehicle Speed - FDCM4,0,+42238112,-083201695
2008-03-09	08:00:57.0	Wheel-Based Vehicle Speed,72.515625,+42374782,-083034933
2008-03-09	08:00:57.0	Barometric Pressure,29.441,+42374782,-083034933
2008-03-09	08:00:57.0	Ambient Air Temperature,15.8,+42374782,-083034933
2008-03-09	08:00:57.0	Front Windshield Wiper Status,0,+42374782,-083034933
2008-03-09	08:00:59.0	Wheel-Based Vehicle Speed,28.359375,+42323309,-082573079
2008-03-09	08:00:59.0	Barometric Pressure,29.205,+42323309,-082573079
2008-03-09	08:00:59.0	Ambient Air Temperature,15.8,+42323309,-082573079
2008-03-09	08:00:59.0	Front Windshield Wiper Status,0,+42323309,-082573079
2008-03-09	08:01:08.0	Wheel-Based Vehicle Speed,75.5625,+42302464,-083068992
2008-03-09	08:01:08.0	Barometric Pressure,28.792,+42302464,-083068992
2008-03-09	08:01:08.0	Ambient Air Temperature,17.6,+42302464,-083068992
2008-03-09	08:01:08.0	Front Windshield Wiper Status,0,+42302464,-083068992
2008-03-09	08:01:12.0	Wheel-Based Vehicle Speed,73.234375,+42012273,-084585066
2008-03-09	08:01:12.0	Barometric Pressure,29.028,+42012273,-084585066
2008-03-09	08:01:12.0	Ambient Air Temperature,19.4,+42012273,-084585066
2008-03-09	08:01:12.0	Front Windshield Wiper Status,0,+42012273,-084585066
2008-03-09	08:01:31.0	Wheel-Based Vehicle Speed,40.765625,+42433559,-083209644
2008-03-09	08:01:31.0	Barometric Pressure,28.969,+42433559,-083209644
2008-03-09	08:01:31.0	Ambient Air Temperature,14,+42433559,-083209644
2008-03-09	08:01:31.0	Front Windshield Wiper Status,0,+42433559,-083209644
2008-03-09	08:01:35.0	Wheel-Based Vehicle Speed,85.53125,+42417601,-083158325
2008-03-09	08:01:35.0	Barometric Pressure,28.615,+42417601,-083158325
2008-03-09	08:01:35.0	Ambient Air Temperature,15.8,+42417601,-083158325
2008-03-09	08:01:35.0	Front Windshield Wiper Status,0,+42417601,-083158325
2008-03-09	08:01:57.0	Wheel-Based Vehicle Speed,54.65625,+42374957,-083022082
2008-03-09	08:01:57.0	Barometric Pressure,29.441,+42374957,-083022082
2008-03-09	08:01:57.0	Ambient Air Temperature,15.8,+42374957,-083022082
2008-03-09	08:01:57.0	Front Windshield Wiper Status,0,+42374957,-083022082
2008-03-09	08:02:08.0	Wheel-Based Vehicle Speed,75.3125,+42291950,-083069065
2008-03-09	08:02:08.0	Barometric Pressure,28.851,+42291950,-083069065
2008-03-09	08:02:08.0	Ambient Air Temperature,17.6,+42291950,-083069065
2008-03-09	08:02:08.0	Front Windshield Wiper Status,0,+42291950,-083069065

Figure 11 – Sample Chrysler Fast Feedback Fleet Data

3.2.3 MDOT Fleet Vehicles

MDOT has developed a capability of deploying data collection devices to fleet vehicles from the state pool to provide data for DUAP applications. The original concept for this capability would equip about ten vehicles in each of MDOT’s seven regions around the state. The system would also deploy communications infrastructure to provide connectivity from the fleet vehicles to the DUAP servers as shown in Figure 12 below. Each region would also have one or more wireless access points (WAPs) at MDOT’s regional transportation service centers (TSCs) for downloading data from the vehicle on-board devices. Internet connectivity from the WAPs would provide access to the DUAP servers.

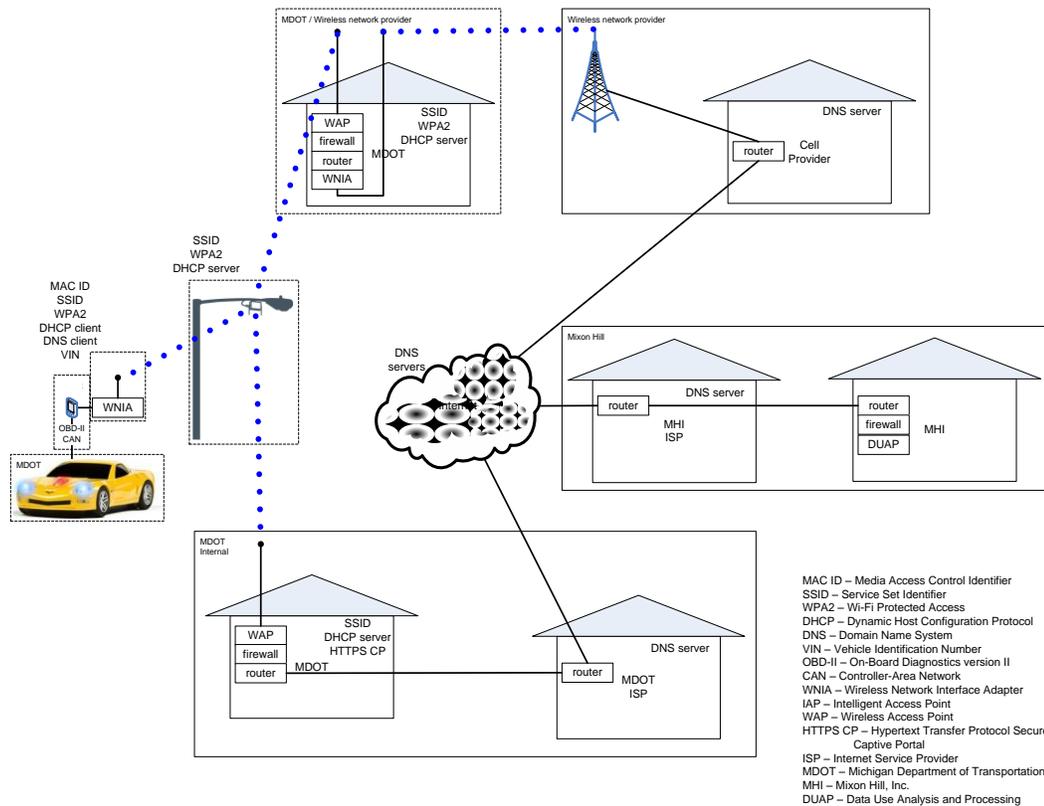


Figure 12 – MDOT Probe Vehicle Fleet Communications

The on-board collection devices consist of data logger hardware and software, an interface to the vehicle’s On-Board Diagnostics II (OBD-II) port, Global Positioning System (GPS) receiver and antenna, a three-axis accelerometer to supplement the vehicle sensors, and Wi-Fi wireless communications. A six-foot data cable connects power and data between the connector-mounted OBD-II unit and the dash-mounted GPS unit. The accelerometer is mounted on the board within the base OBD-II module. Data are stored locally on a MicroSD card that can be removed for manual data retrieval.

The MDOT fleet data specification is described in Appendix D. The availability of particular data types and the configuration of output are specific to vehicle make, model and year of manufacture. The particular data fields being provided by a given vehicle/device are described in the data file headers. Data is uniquely associated with vehicle identification number (VIN) to assure traceability for research purposes.

3.2.4 Android Applications

For those cases where it is not necessary to get data from the vehicle control bus, location and acceleration data can be obtained from many commercially-available consumer electronic devices. In particular, mobile telephones running Android and similar operating systems, generally referred to as “smartphones”, provide both sensors and native communications capability. These features provide an attractive base for development of aftermarket probe data acquisition.

The application used for demonstration of these capabilities for DUAP was developed independently of the project to specifications provided to the developers and shown in Appendix E. The Android application uses the smartphone’s GPS capabilities with its on-board three-axis accelerometer to build data files which are sent over the device’s available data network connection(s). Vehicle speed is calculated from the GPS location and time data.

Use of commercially available devices and supporting infrastructure enables the application to be installed on large numbers of devices at low cost to the agency. The prototype application does, however, have some technical limitations.

- The prototype identifies the data’s origin by the mobile equipment identifier (MEID) associated with the device. There is no explicit link to the vehicle in which the device is being used, or assurance that it is always used in the same vehicle. This could limit the usefulness of the data within some application analyses.
- The prototype does not have any interface to the vehicle data bus. This limits the range of agency applications to those concerned directly with location and acceleration of the device.
- The device’s location and orientation within the vehicle are not necessarily or physically constrained by a particular mounting or interface. As such, the acceleration data from the device may not have the same orientation from one data set to the next. Although it might be possible to detect and account for the orientation, it is a significant complication in the application analyses. MDOT is currently developing solutions for this in their Slippery Road Detection and Evaluation Project.

3.3 Applications Analysis

The selection of the most appropriate applications to develop as prototypes occurred initially in conjunction with the development of functional requirements, and in close collaboration with MDOT and other members of the Steering Committee. The intent, however, was to develop a set of prototype applications that addressed the management and operations needs of a broader group of agencies, users, and time frames—from real-time operations to archived data uses.

The selection of applications was also conditioned by how data latency affects the distribution of data processing across the network of connected vehicle services. For example, a safety application such as intersection collision avoidance with very short latency needs requires the application to run at the roadside.

Conversely, an asset management application where data about pavement conditions is aggregated over a period of months or even years can use data that is aggregated and analyzed on remote servers. The DUAP focus on applications running in more central facilities limited the project's ability to address any real-time applications and suggested a focus on near-term operational and long-term archival and planning applications

No matter where the data processing takes place for a particular application, it was also important to design the DUAP system and its prototype applications with an understanding that an individual data element will typically be used in multiple applications. For example, roadway slip as measured by vehicles' electronic traction control systems could be a data element processed locally as part of an adaptive traffic signal control application (perhaps to extend green times); used in a traveler information application to send messages to dynamic message signs, 511[®] systems, and web sites in near real-time; and part of a winter maintenance application that uses decision-support tools to determine treatment strategies and dispatch plows.

These application considerations have been fundamental to the DUAP design. It has been designed specifically to collect connected vehicle data into a common repository for multiple applications. The data structures enable display of both collected data and synthesized performance measures. Applications built on the data can provide historical trends, identify locations experiencing particular sets of conditions, and return aggregated performance measures.

With this overview, each of the selected application areas will be addressed on its own terms by discussion of the application's objectives, concept, research and results, and conclusions.

3.3.1 Traffic Condition Monitoring

3.3.1.1 Objective

Traditional and current traffic monitoring solutions are almost exclusively based on using spatially-fixed vehicle detection stations to detect, count, and characterize the vehicles passing each station. Independent of the particular technologies used to implement that detection and characterization—induction loops, radar, video, and so forth—the detection is limited to those particular locations at which the detector is stationed. Solutions based on obtaining the location of particular vehicles—for example, automatic vehicle location (AVL) systems—generally have been for fleet operations, or for supplementary probe data.

A broad deployment of connected vehicle systems would dramatically change this situation. If all vehicles can report their locations, fixed detection and counting becomes a means of confirming performance measures that are aggregated from the individual vehicle data. Traffic monitoring would no longer be limited to the number of stations that could be deployed by the transportation agency. For example, the agency can use connected vehicle data for generating accurate and reliable travel times to be utilized by the agency and the traveling public.

The objective of connected vehicle traffic condition monitoring, then, is to enable the measurement of traffic conditions across the transportation network using data provided by the vehicles themselves. For this solution to be successful, the connected vehicle systems need to be able to:

- Collect traffic data throughout all parts of the roadway network to be monitored
- Collect traffic data at finer spatial and temporal resolutions than currently available with fixed detection stations
- Collect traffic data in near real time, consistent with current generation fixed detection stations and networks
- Create actionable information from the data, consistent with current generation advanced traffic management systems (ATMS)
- Publish the information with as little time delay as possible.

3.3.1.2 Concept

A vehicle's relationship to the traffic flow is provided by data describing vehicle motion. These data could include, at a given time, the vehicle:

- Location (latitude and longitude from GPS)
- Speed
- Heading
- Brake status and ABS actuation
- Steering status, yaw rate, and stability control actuation
- Longitudinal and lateral acceleration

Traffic measures would then be derived from the aggregation of data from individual vehicles. Traffic speed, for example, could be represented by the mean and standard deviation of speeds of the vehicles along a given roadway segment within a certain time interval. Incident detections might be synthesized from correlation of vehicle speeds with braking, acceleration, and steering status. Travel time estimates could be derived from analysis of vehicle positions and speeds.

Low data latencies (on the order of seconds rather than minutes) will be required for vehicle probe data to be useful in traffic monitoring and management. Current ATMS sensors and networks operate in near real time, and any significant increase in data latency would compromise both operations and public trust.

The number of vehicles needed to generate useful traffic data may be a constraint on traffic monitoring applications, especially in the near term when connected vehicles would represent a small fraction of overall vehicle populations. A 2007

study⁵ found that probe vehicle populations on the order of 10% of vehicles are needed for accurate and reliable traffic travel time estimates, with slightly more vehicles needed to estimate mean traffic speed. Increasing the number of vehicles would provide both higher confidence and more accurate statistics. Traffic control applications depending on vehicle counts (e.g., signal control) would require even higher population percentages, approaching connected vehicle saturation, to be effective.

3.3.1.3 Research Findings

As described earlier in the discussion of DUAP system interfaces, mean speed calculations have been developed to demonstrate generation of aggregate traffic performance measures from connected vehicle data. The mean value of all vehicle speeds along any roadway segment within a given five-minute window is calculated and assigned to that segment as the “traffic speed”. No calculation is provided for segments with insufficient data for calculating a mean.

Unfortunately, instances of sufficient probe vehicle data density to drive the traffic speed calculation have been sparse. Figure 13 below demonstrates the capability for one such interval when sufficient data were available.

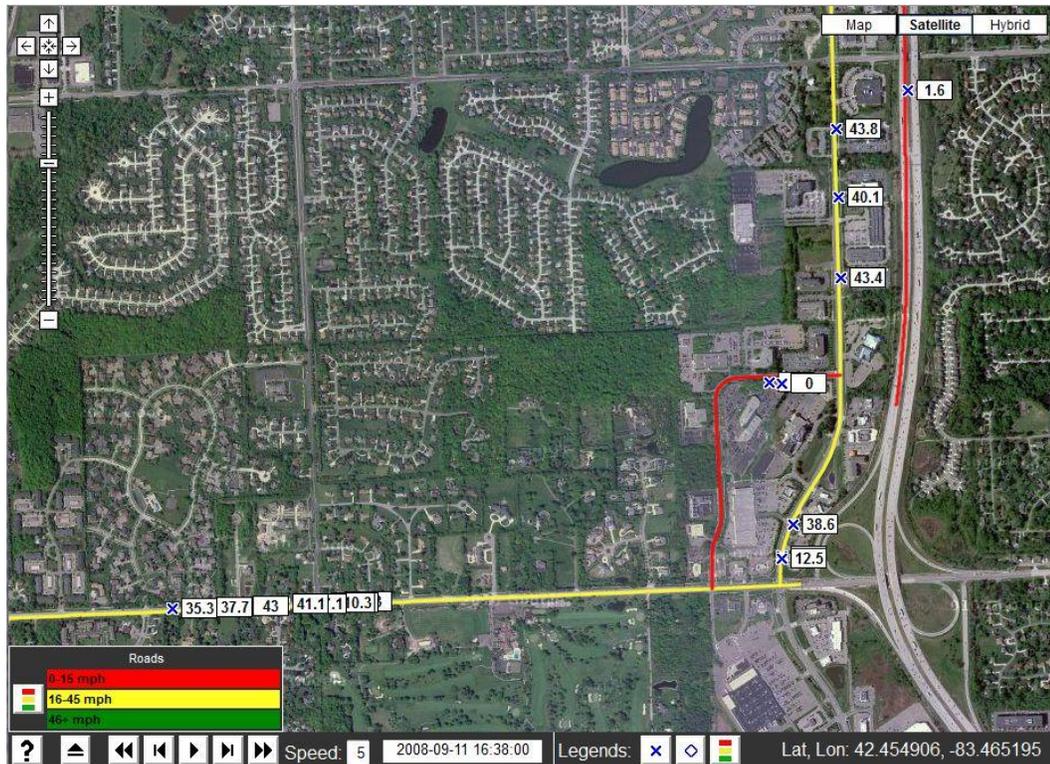


Figure 13 – Example Vehicle Speed Aggregation

⁵ Noblis Inc. *Vehicle-Infrastructure Integration (VII) Probe Data Characteristics: Analysis of Three Field Data Sets*. Falls Church, Virginia. June 2007.

3.3.2 Pavement Condition Monitoring

3.3.2.1 Objective

Connected vehicle systems provide a potential opportunity to improve pavement condition monitoring by capturing observations from more widely-distributed lower-cost sensors than are currently available to transportation agencies. Knowledge of pavement conditions is important to a transportation agency's planning processes and State DOTs are required to provide these data to the U. S. DOT's Highway Performance Monitoring System (HPMS). Unfortunately, pavement condition measurements have been time-consuming and expensive to obtain. Advances in vehicle systems and communications are creating new capabilities that may change the situation. In particular, it may be possible to obtain useful measures of road network pavement conditions from probe vehicle data.

Pavement is one of a transportation agency's most valuable infrastructural assets. As such, the need to preserve that asset will compel the agency to monitor the pavement condition throughout its lifetime. Measurements of newly constructed pavement provide both indicators of construction quality and a baseline for monitoring pavement wear. Periodic measurements while in service are key indicators of the need for maintenance and refurbishment.

Assessment of pavement condition takes several forms. In the broadest sense, drivers make subjective judgments of pavement condition as they use the roadways. Visual inspection may be used routinely by a transportation agency to monitor gross defects like potholes and cracks. Profilographs are often used to measure the vertical deviations in pavement as part of acceptance testing for new construction. Response-type road roughness measuring systems (RTRRMS) record the vertical motion of a vehicle as it travels along the pavement. Road roughness profiling devices measure the distance between the pavement surface and a vehicle's inertial reference point using non-contact laser or ultrasound sensors.⁶

Measures of pavement condition are diverse, reflecting both the equipment used to obtain the measure and the intended application of the measurements. The two most prominent measures are the Present Serviceability Rating (PSR), which represents individual human assessment of ride quality, and the International Roughness Index (IRI), which is an objective measure of pavement surface deviations. Correlations have been developed between these and other measures, and the IRI is generally accepted as the best common measure of pavement condition, required by the Federal Highway Administration as the standard for ride quality index by the state DOTs.

Current MDOT pavement monitoring practices allow standard assessments of pavement statewide about every two years.

⁶ The FHWA's web site for pavement information includes an excellent (if dated) synopsis of pavement measurement at <http://www.fhwa.dot.gov/pavement/smoothness/rough.cfm> (accessed 2010 Jan 25).

Monitoring pavement condition, as for HPMS input, requires consistent data sets over long periods of time—several years. The data have to be indexed precisely to the pavement segments being monitored in order to facilitate correlation to materials, use, and maintenance history. Data also need to be normalized to a consistent set of measures based on techniques and figures of merit, and seasonal biases and anomalies may need to be noted or removed.

In contrast, detecting gross pavement defects (e.g. potholes, blow-ups) requires accurate spatial discrimination. Analysis of the data is looking for particular events at particular locations, and identification with a particular pavement segment is secondary to the actual location. Data need to be normalized to a consistent set of measures (size or severity, for example) and be cleansed of seasonal biases.

3.3.2.2 Concept

Consistent with other connected vehicle applications, the concept for pavement condition monitoring is to use instrumentation commonly available on light vehicles to obtain data that may be indicative of pavement conditions. Indications obtained in this manner could then be used to notify maintenance managers of conditions that may require more precise assessment.

The nature of data from cars and light trucks presents several challenges and opportunities. In general, data available from light vehicle systems relate only to vehicle operations, and not particularly to the environment in which the vehicle operates. Information about the roadway (or, similarly, weather conditions) can only be inferred from the vehicle data. Data available from a vehicle varies among manufacturers and models. Even the vehicle data interface standards typically describe the structures and interfaces to the data, but do not necessarily prescribe what data should be made available. While this may create some limitations, aftermarket products may provide data from additional sensors to supplement the vehicle data stores.

Data obtained from vehicles would be correlated with other observations to synthesize relative measures of pavement conditions and performance. Since vehicle designs can vary, data from each vehicle observation platform would be calibrated to a known set of pavement conditions. Calibration is likely to depend on the observation type, sensor, sensor installation and configuration, and vehicle configuration. Data from all vehicle observation platforms could then be transformed into a standard set of normalized pseudo-observations. Analysis of the aggregated pseudo-observations would be used to identify likely sites for more precise condition assessment.

Since light vehicles do not provide any means of obtaining actual pavement surface profiles, other data are analyzed for potential correlation with profilometric standards. Data acquisition and characterization for data received from the MDOT fleet most closely resemble the response-type road roughness measurements used prior to more detailed profilometric techniques. Parameters of specific interest to pavement condition monitoring include: acceleration on three

axes, vehicle speed, tire pressure, and traction control actuation (if available). Other vehicle data may also be evaluated for correlation with road conditions.

Data collection parameters can generally be inferred from prior pavement condition monitoring practices. For example, longitudinal profilometry is typically recorded at one-inch intervals. Profiles are then typically blanked (small variations, less than 0.2 inches, are ignored) and notch filtered (removing both high and low frequency variation). Profiles can then be aggregated or averaged over various path lengths to get figures of merit for roughness. Comparable accelerometry data at 60 miles/hour (88 feet/sec) would be recorded at approximately 1 millisecond intervals, although this could provide substantially more data than would be needed to generate the relevant statistics.

Generating actual vertical acceleration data may be problematic. The three-axis accelerometers are embedded in the on-board devices, and the mounting of the devices to the vehicles is highly variable. The orientation of the device relative to the acceleration components to be measured can also have a large impact on measurement accuracy⁷. Depending on the accelerometer package implementation, the raw acceleration components may include the gravity component. In that case, even the zero-motion state may show non-zero acceleration values. Even if the mounting of the accelerometer within the vehicle were controlled, the zero-motion state would include non-zero components of the vehicle's position (parked on a hill on a severely crowned street, or on the slope of a drainage ditch).

3.3.2.3 Research Findings

The first survey and analysis of the potential for pavement condition analysis applications used the MDOT fleet described in Section 3.2.3 to collect accelerometry data. MDOT's pavement profiling van was used to collect baseline profilometry for a fixed route along which accelerometry would be collected. Both the fleet vehicles and the profilometry van were equipped with the OBD-II device, GPS receiver, and three-axis accelerometer. Data was collected through a wireless access point at the originating MDOT facility for uploading to the DUAP system. Figure 14 illustrates the demonstration configuration.

⁷ http://www.pcb.com/techsupport/tech_accel.php

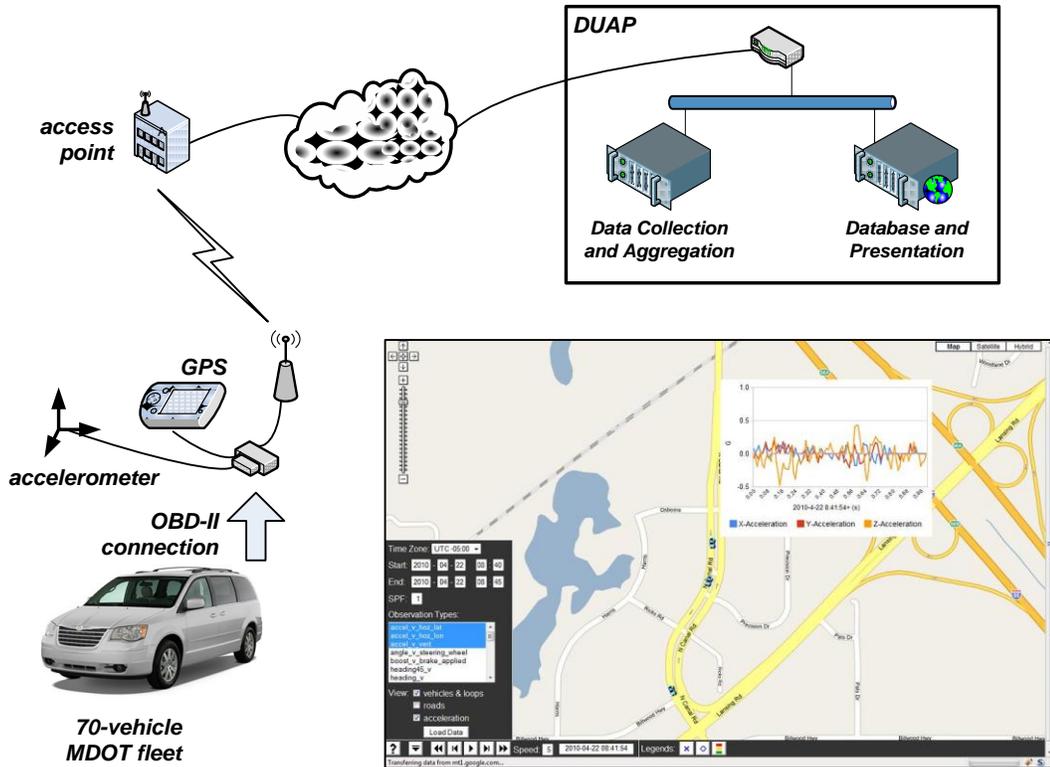


Figure 14 – Pavement Condition Demonstration Configuration

Test routes were selected to provide convenient and effective data collection. Routes started and ended at a MDOT facility with a network access point for data uploads; included a variety of roadway segment types; and showed a variety of pavement conditions and defects. Figure 15 shows one such route. Multiple passes were planned for each route.

The MDOT profilometry van (Figure 16 and Figure 17) is built with single-point laser instruments in the front and band-type profilometers in the rear of the vehicle. For these tests, the van was also equipped with the MDOT fleet devices. Use of both sets of equipment on a single vehicle would provide the most reliable and consistent sets of profilometry and accelerometry.

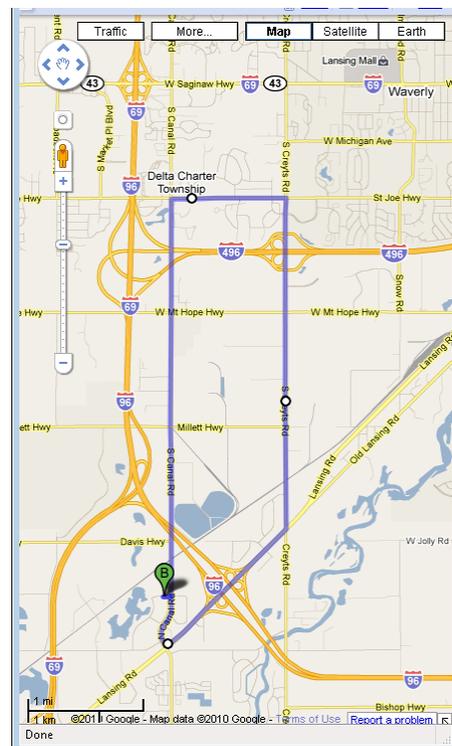


Figure 15 – Example Pavement Condition Demonstration Route



Figure 16 – MDOT Profilometry Van - Front View



Figure 17 – MDOT Profilometry Van - Rear View

Accelerometry, vehicle position, and vehicle data from the OBD-II port were synchronized for data logging. The three axes of acceleration data were recorded every 10 milliseconds. The GPS data of latitude, longitude, elevation, speed, heading, and satellite state were reported and logged every one second. The OBD-II device provided a record of vehicle speed independent of the GPS-derived value. Data logging was stopped and data was uploaded after each trip over a test route was completed.

Typical accelerometry results are shown in Figures 18, 19 and 20 below for three locations with pavement defects along the routes selected for the demonstration. The small graph within each figure shows the three axes of acceleration as the vehicle (the green icon on the satellite photo) moves over the pictured pavement defect along the red line on the satellite photo. These results provided confidence that on-board aftermarket accelerometers can capture distinguishable pavement defect signatures.

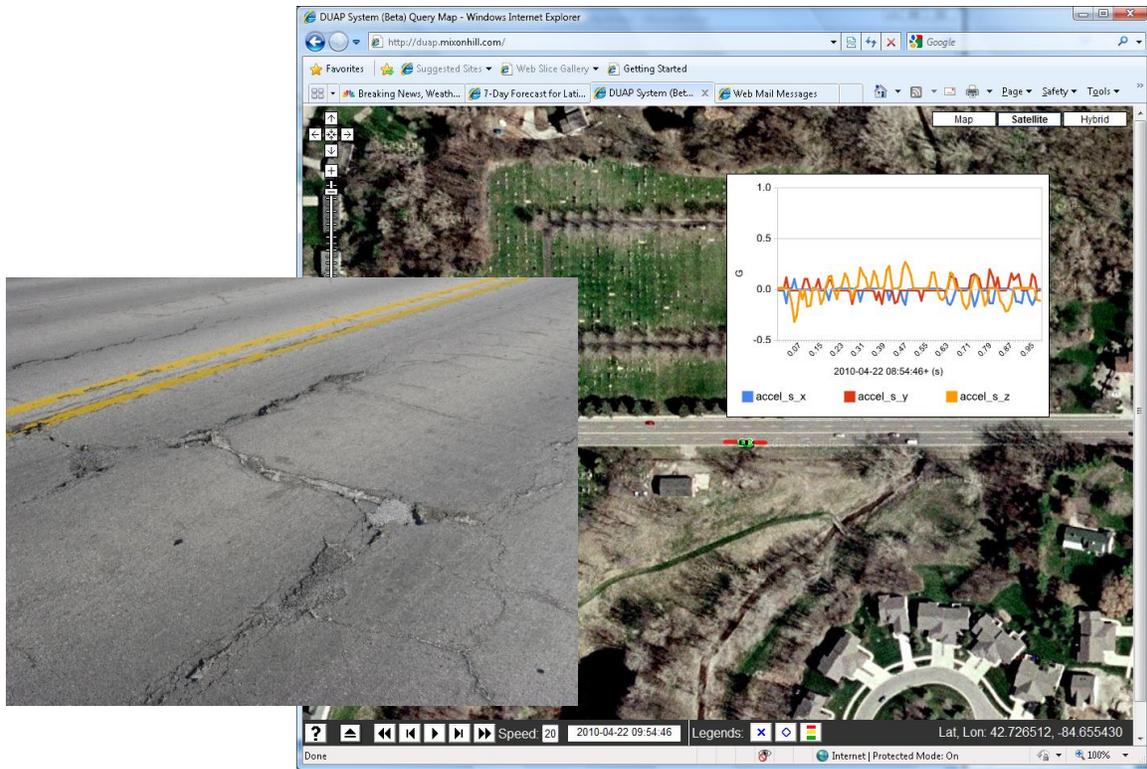


Figure 18 – Pavement Defect Response 1

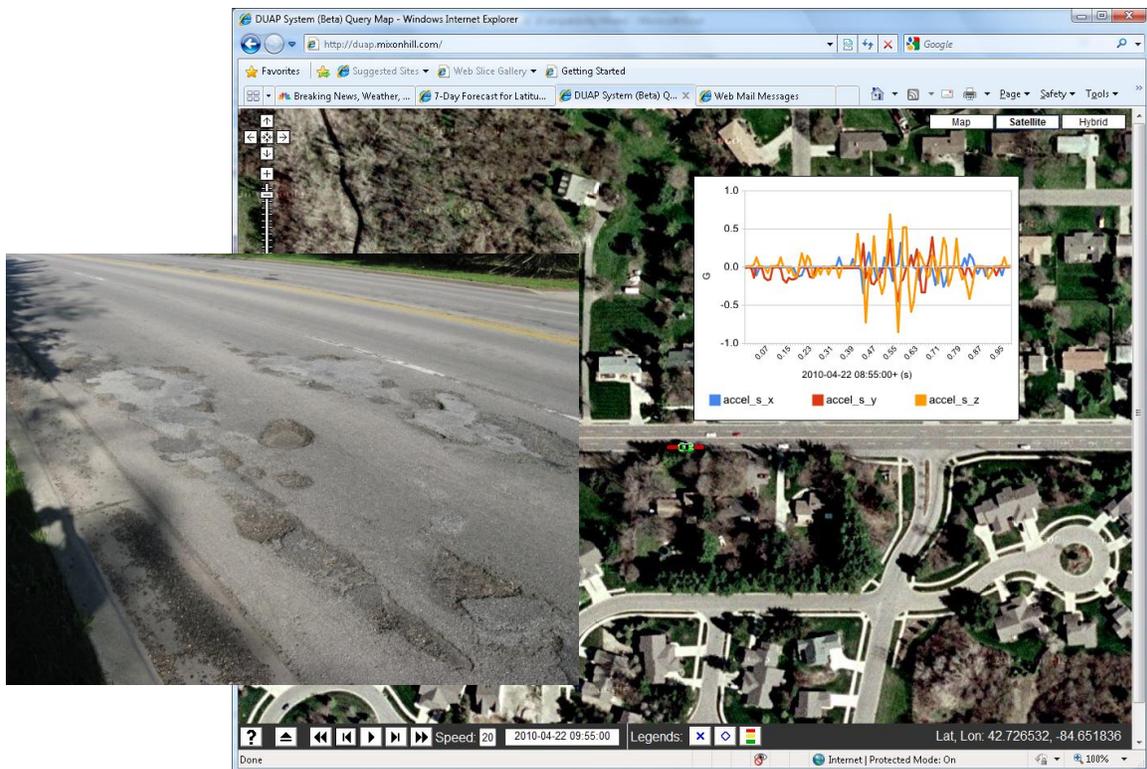


Figure 19 – Pavement Defect Response 2



Figure 20 – Pavement Defect Response 3

Correlation of the profilometry with accelerometry in these runs, however, was not so conclusive. The geographic calibration of the profilometry measurements was limited by lack of a GPS receiver for the profilometer, so correlation with the accelerometry was limited to those data sets with known start/stop conditions. Reliability of the on-board equipment was also a problem, and accelerometry was collected only for a subset of all the test runs.

The promise shown in this first round of field testing was sufficient, however, to motivate an additional set of tests with technical and operational improvements.

- Known issues with the on-board equipment would be resolved to improve reliability.
- Accelerometry would also be collected using a second platform with alternative data communications.
- Routes would be formalized and inspected for PASER rating as well as profilometry/IRI and acceleration.
- Profilometry would be georeferenced at start/stop points.

Another round of field testing was therefore held to address these issues. Similar to the first round, these studies ran a platoon of instrumented vehicles over a set of test segments with varying pavement conditions to gather a base set of data for correlation studies.

Conditions varied from new to severely degraded. Standard measurements included laser profilometry and PASER condition rating studies. Two accelerometry studies—using an Android phone application and the specialized telemetry units from the first round of field testing—were performed concurrently in multiple passes over the same segments.

Measurements were loaded into the DUAP system for comparison, analysis and correlation. DUAP puts all the data feeds into consistent formats and time/space context.

These results are representative of all vehicles, sensing devices, and segments surveyed.

- Accelerometry is, as expected, highly variable between sensing devices and vehicles. It is, however, relatively consistent between passes by the same device and vehicle over a given segment.
- Sensed accelerations provide an adequate base for approximately locating pavement defects. Locations are relatively consistent across devices and vehicles. Damped responses seem to provide more reliable indications, although ground surveys were not detailed enough to be conclusive.
- Statistics derived from sensed accelerations can provide relative indications of vehicle ride response from segment to segment for a given vehicle. It *may* be possible with sufficiently large data sets to correlate a given vehicle and device's characteristic responses to other pavement condition measures like IRI or PASER rating.

3.3.3 Origin-Destination Studies

3.3.3.1 Objective

Planning for the maintenance and growth of a multi-modal transportation network requires a tremendous amount of operational data. The movement of people, goods and services across the network varies throughout the weeks and seasons, but also shifts over time with local economic conditions and new development. Decisions as to where to allocate remediation and construction funds should be based on the best available network condition and usage data. All of the agency's performance measurement and data archiving systems are driven largely by this need for data to support these decision processes.

Traditional infrastructure-based measures of system usage can report conditions at particular locations on the network, but are notoriously lacking in their ability to see the movement of people, goods and services across the network. This capability, generally captured in the category of origin-destination (O-D) studies, could be significantly enhanced by connected vehicle systems.

The objective of this application analysis is therefore to assess the potential for DUAP to support collection of O-D and trip data for DOT planning purposes. To do so, DUAP needs to get, process, and provide observations of vehicle movement on Michigan's road network. This capability would then further

support development and deployment of other connected vehicle and agency planning applications.

3.3.3.2 Concept

Measuring traffic movement on the road network is both important and tediously difficult. Historically, measurement has focused on counts of traffic volume at particular locations as a means to generate statistics on average daily traffic for the network link on which the count is made. These methods result in characterizing roadway use through a set of link volumes. Counts can vary widely in scope and complexity, from human counting over a particular study period to permanent automated counters, or from directional flows past a particular point to lane-by-lane counting of movement through an intersection.

Trip data analysis attempts to characterize the demand for traffic capacity from node to node within the network, capturing the origins and destinations of representative trips and profiling turning movements along travel paths. This is inherently a different type of data gathering than that used in volume counts since it depends on the movement of a particular vehicle from one node to another. Nonetheless, both counts and trip data need large volumes of data to accurately characterize traffic through the network.

The original VII “Day 1” applications for trip data were described in the “Corridor Management Planning Assistance” use cases⁸:

- Analyze Archived Probe Data and Generate Input Data for Planning Process
- Archive Opt-In Information
- Archive Probe Data
- Collect Vehicle Probe Data
- Collect Vehicle Origins and Destinations
- Provide Raw Probe Data
- Provide Vehicle Origin and Destination Information to Planning Entity
- Record Performance Measures

Descriptions of the use cases, however, were not provided in the Day 1 applications document. Applications supporting these use cases were not included in the eventual POC demonstrations, either. As such, recent connected vehicle experience has not developed any particular body of knowledge on these applications.

From the DOT planner’s perspective, almost any information about vehicle travel paths might be useful. From a practical standpoint, planners need data for models

⁸ VII National Working Group, *PD-22.c-1 – Traffic Management: Corridor Management Planning Assistance (Day 1)*, Version 3, 2006 April 06.

used to analyze the impact of changes to the road network on road usage. Traditionally, planning data has focused on

- Traffic counts for links in the network,
- Turning movements at nodes within the network, and
- Origin-destination pairs, at all scales from specific local and regional corridors to national freight movements.

These data, with the metadata from the network maps, form the analytical basis for transportation simulation studies in support of planning.

Existing methods for collection of trip data are expensive and based on small samples of vehicle and network data. The Travel Survey Manual⁹ describes seven “common types of travel survey used to discern more about (and to model) the behavior of users of highway and transit facilities”.

- Household travel/activity
- Vehicle intercept
- Transit onboard
- Commercial vehicle
- Workplace and establishment
- Hotel/visitor
- Parking

Data collection is therefore tuned to the application models. Data collected for planning (region-to-region and corridor models) is generally highly aggregated, whereas data collected for traffic and signal studies (network loading analysis) is more detailed, but of limited geographic scope.

Acquiring vehicle-specific trip data in a connected vehicle environment provides both opportunities and challenges. For example, data on an individual vehicle’s origin, path, time history, and destination(s) is inherent to probe data collection, subject to limitations for the owner/operator privacy. These data could replace or augment data collected in traditional surveys. Connected vehicle data could also supplement current practice with data that are otherwise unavailable, unreliable, or too expensive to gather. On the other hand, the data are provided from the perspective of the vehicles, and only those vehicles able to provide data are represented in the data collection. This may present issues with data scaling that are not inherent in current trip data collection methods.

3.3.3.3 Research Findings

The ability to generate vehicle trip paths with DUAP is more an issue of the limitations in the data provided to DUAP than of the system’s capabilities.

⁹ Available for download at <http://ntl.bts.gov/lib/4000/4500/4529/1392.pdf> as of 2010.02.02.

For example, the Chrysler data discussed earlier in this report was completely anonymized—each data value was uniquely specified, and not associated to a particular vehicle, or even to another value of the same type. It would be theoretically possible to reconstruct probable vehicle paths from the isolated data, but both the effort to provide the data and the uncertainty in the results would increase quickly with larger data sets.

On the other hand, data provided by the MDOT fleet vehicles and the Android application was specifically associated with the data source, either by MEID or vehicle identifier. Generating the vehicle path is implicit to the data.

This variability in the utility of connected vehicle data highlights the need for consideration of data privacy and sharing policies as part of application analysis and development.

4 CONCLUSIONS

4.1 *DUAP System Capabilities*

DUAP has been successful in fulfilling its original purpose—to evaluate uses and benefits of data from probe vehicles in agency management and operations. The program demonstrated, for example:

- Acquisition of data from multiple probe data sources;
- Deployment of aftermarket on-board sensors and data acquisition units for measurements of specific operational interest to DOTs (in this case, accelerometers);
- Sorting and aggregation of multiple data types from each of the sources;
- Synthesis of performance measures for specific DOT applications (for example, segment average speeds, or relative pavement conditions) from the probe data; and
- Presentation of raw probe and processed data in consistent, flexible map-based operator interfaces.

DUAP's success has, however, highlighted the limitations of the prior institutional arrangements and approach. The presumption of large volumes of data from diverse systems led to a research plan exploring what might be done with the data. The current reality is that data will continue to become available only as it is needed and can be obtained for particular applications. The focus of research needs to shift from “what can we do with everything that’s available?” to “what do we need, and how do we get it?”

4.2 *Data Collection and Standards*

DUAP experience with data collection has demonstrated its ability to get and aggregate data from multiple sources and formats into an integrated repository. The system philosophy and architecture of isolating the data collecting components from the data repository enables the system to add new collectors as needed to accommodate interfaces with varying data specifications, timing, and network protocols. If a data source already provides an interface, a new collector component for DUAP will be the most effective means of getting the data from that particular source. DUAP performance can, however, generally be improved by having the data provider conform their data interface to the DUAP standard shown in Appendix E.

Given the capability to collect data from a wide variety of interfaces, the data repository also needs to be able to accommodate data from different sensors and sources at differing time and spatial resolutions. Each data source may provide values for a given parameter with its own resolution, which may or may not match the resolutions needed by particular applications of that data. For example, “speed” could be provided by multiple sensors on a vehicle (for each wheel, for the average wheel, based on GPS coordinates) and by roadway sensors (which

generally provide only time averaged speeds over multiple passing vehicles). Each of these cases needs to be specifically identified within the repository with metadata (for example, source) that may be needed in the application of that data.

4.3 Applications

Viability of any particular application of connected vehicle data depends directly on data characteristics including:

- Availability of the data types relevant to that application
- Availability of sufficient data volumes
- Data spatial and temporal resolutions consistent with the needs of that application
- Metadata linking the data to the sensors and loggers that provided it so as to enable appropriate data quality checks

Given these data conditions and the prototypical nature of connected vehicle systems, applications of value to DOTs are likely to depend in the near term on data from DOT-controlled vehicles. Test bed and demonstration deployments to date have been consistently unable to provide sufficient data to enable DOT application development. As described earlier, DUAP application development proceeded most effectively when data originated with vehicles and system directly under MDOT's control.

Connected vehicle technologies may eventually be valuable for traffic monitoring applications, but only after they are sufficiently distributed within vehicle fleets. This conclusion reflects experience not just within a formal "connected vehicle" context, but results seen in other AVL deployments and commercial traffic data service providers. It is unclear whether a DOT's fleets would ever be large enough to provide statistically useful traffic probe data without other correlating traffic data (e.g., from roadway sensors). In the near term, connected vehicle technologies are likely to be most useful as experimental and confirmatory traffic probes.

Pavement defect detection has been demonstrated in this project and should be achievable across the DOT with currently available technologies. Equipment deployed on a relatively small fraction of the DOT vehicle fleet around the state would gather data as part of the normal operations of the fleet vehicles. This application would enable the DOT to reduce the time and resources needed to identify pavement defects like potholes and blow-ups.

Pavement condition monitoring may be achievable with currently available technologies. Correlation of vehicle accelerometry with known pavement conditions is possible—but will need additional research to confirm and reliably demonstrate.

Connected vehicle technologies could be used for origin-destination and planning studies, subject to protection of individual privacy concerns. While wide scale deployment of connected vehicle systems for safety applications may require complete anonymity, focused studies and deployment allowing vehicle tracking

with driver anonymity are technically feasible. Such a deployment would provide results similar to traditional O-D studies using travel logs, but with significant improvements in data quality and detail.

5 RECOMMENDATIONS FOR FURTHER RESEARCH

DUAP's purpose, simply stated, was to assess the use of data from connected vehicles (originally, vehicle-infrastructure integration) to improve transportation agency operations. DUAP research was therefore based and developed on a presumption of data being available from other connected vehicle projects. Activities undertaken as part of DUAP were directed at applications of that data within MDOT and other transportation agencies.

The DUAP research was forced to take a different approach when the original connected vehicle Proof-of-Concept demonstration did not produce the intended volume and variety of connected vehicle data. Other sources of vehicle probe data outside MDOT had similar issues of scale and diversity, and were additionally found to be cost-prohibitive. DUAP research therefore turned toward identifying and developing sustainable sources of data within MDOT to support transportation agency applications. Provisions were made for collecting data from MDOT's own fleet vehicles across the state. The prototype system used to collect the data has enabled the DUAP program to demonstrate the ability to collect data and provide applications specifically related to the improvement of DOT operations.

Based on this experience, it is recommended that further research using the prototype system demonstrated in DUAP (hereafter called DUAP-1) be developed in the DUAP-2 program. DUAP-2 will identify, develop solutions for, and fulfill the data needs of MDOT as they contribute to improving the cost efficiency and enhancing the effectiveness of its operations, with the emphasis on integrating mobile data gathered as part of MDOT's ongoing operations with other data sources from across the agency. Applications to be considered for continuing research should include, but not necessarily be limited to those addressed in the DUAP-1 research.

6 RECOMMENDATIONS FOR IMPLEMENTATION

The next phase of Data Use Analysis and Processing, DUAP-2, should directly address the research recommendations of the prior section and the goals of MDOT's Strategic Plan and MDOT's Connected Vehicle Strategic and Business Plan. For example, DUAP-2 should identify, develop solutions for, and fulfill the data needs of MDOT as they contribute to *improving the cost efficiency and enhancing the effectiveness of its operations*. The techniques and lessons learned in DUAP-1 can be applied in DUAP-2 to operations throughout MDOT. DUAP-2 will provide a common repository and set of user interfaces for telematics data and data from other ongoing MDOT research and development including the Vehicle-Based Information and Data Acquisition System (VIDAS), the Cut River Bridge data system, Slippery Road Multi-path Signal Phase and Timing (SPAT), and available smart phone/aftermarket devices with applications that support DUAP 2 data needs.

A robust system engineering approach should be used to assure that the needs and solutions are aligned at each step of the development process. Documents and software developed in DUAP-1 should be updated, enhanced, and reincorporated into the DUAP-2 deliverables.

- The concept(s) of operations should be developed to address the needs of MDOT as a whole, looking especially for opportunities to leverage data sources across the organization.
- User requirements should be gathered from practitioners and baselined against existing systems such as the Transportation Management System (TMS), the Michigan Intelligent Transportation System (ITS) Center and the Regional Transportation Management Centers (TMCs). Companies and agencies working in data partnerships with MDOT should be included in the requirements gathering process. Interface requirements for providing new data to existing MDOT systems should be specified.
- The system architecture and design should trace directly to the requirements and use industry standards as much as possible, but should not be constrained to existing ways of generating and collecting data—for example, through fixed vehicle detection stations or road weather information systems (RWIS).
- New data sources and acquisition system development may use new hardware and communications as well as software.
- System implementation and testing should be iterative, assuring that components and assemblies at each level of integration comply with the corresponding requirements.
- The final data acquisition and processing systems should be field operationally tested for acceptance by the practitioners and data partners.

Whereas DUAP-1 was originally based on a limited set of use cases and generic data to be provided without prior consideration of its applications, DUAP-2

should more precisely assess MDOT's data needs in light of the intended applications. The process should not be constrained to particular sources of data (as was the original presumption for POC data in DUAP-1), and DUAP-2 can proceed from any facet of MDOT's operations to whatever data sources may be needed to fulfill the operational objectives.

DUAP-2 should begin as soon as possible to leverage the DUAP-1 accomplishments and fit with emerging connected vehicle opportunities such as the Safety Pilot Test Conductor Model Deployment Project in Ann Arbor, Michigan, to utilize the data collected by the 3000 vehicle test fleet.

APPENDIX A - ACRONYMS AND ABBREVIATIONS

The following table provides definitions of terms, acronyms, and abbreviations used in this document.

Term	Definition
3G	Third generation, as applied to commercial wireless telecommunication services
511 [®]	Telephony-based traveler information services
API	Application programming interface
ATMS	Advanced traffic management system
AVL	Automatic vehicle location
CFR	Code of Federal Regulations
ConOps	Concept of operations
CSV	Comma-separated variable
DOT	Department of Transportation
DUAP	Data Use Analysis and Processing
ECMA	Originally, the European Computer Manufacturer's Association; now not considered an acronym, but still used as part of the names for some standards specifications
FTP	File transfer protocol
g	Unit gravitational acceleration; 1 g equals 32.2 feet per second-squared
GPS	Global Positioning System
HPMS	Highway Performance Monitoring System
HTTP	Hypertext Transfer Protocol
Hz	Hertz; 1 Hz equals one cycle per second
IEEE	Institute of Electrical and Electronic Engineers
IRI	International Roughness Index
ITS	Intelligent Transportation Systems
JAR	Java Archive
MDOT	Michigan Department of Transportation
MEID	Mobile equipment identifier
MicroSD	Micro- Secure Digital; memory card format
MITS	Michigan Intelligent Transportation Systems

OBD-II CAN	On-Board Diagnostics II Controller Area Network
O-D	Origin-Destination
PASER	MDOT subjective pavement condition measure
POC	Proof of Concept
PSR	Present Serviceability Rating
RMS	Root mean square
RTRRMS	Response-type road roughness measuring system
RWIS	Road Weather Information Systems
SAD	System Architecture Description
SDD	System Design Description
SEMSIM	Southeast Michigan Snow and Ice Management
SPaT	Signal Phase and Timing
SRS	System Requirement Specification
TMC	Traffic Management Center
TSC	Transportation Service Center
UML	Unified Modeling Language
URL	Uniform Resource Locator
USDOT	U.S. Department of Transportation
UTC	Coordinated Universal Time
UTF	Unicode Transformation Format
VIDAS	Vehicle-based Information and Data Acquisition System
VII	Vehicle Infrastructure Integration
VIIC	VII Consortium
VIN	Vehicle identification number
WAP	Wireless access point
Wi-Fi	A trademark of the Wi-Fi Alliance and a brand name for IEEE 802.11 wireless networking services

APPENDIX B - REFERENCES

The following documents contain additional information pertaining to this project and the requirements for the system:

- 1 Michigan Department of Transportation. (January 9, 2007). *Vehicle-Infrastructure Integration Strategic and Business Plan*. Lansing, MI: Michigan Department of Transportation.
- 2 Michigan Department of Transportation. *VII Data Use Analysis and Processing: Concept of Operations*. Lansing, MI: Michigan Department of Transportation, 2007.
- 3 Michigan Department of Transportation. *VII Data Use Analysis and Processing: Requirements Specification*. Lansing, MI: Michigan Department of Transportation, 2007.
- 4 Michigan Department of Transportation. *VII Data Use Analysis and Processing: System Architecture Description*. Lansing, MI: Michigan Department of Transportation, 2007.
- 5 Michigan Department of Transportation. *VII Data Use Analysis and Processing: System Design Description*. Lansing, MI: Michigan Department of Transportation, 2008.
- 6 Noblis Inc. *Vehicle-Infrastructure Integration (VII) Probe Data Characteristics: Analysis of Three Field Data Sets*. Falls Church, Virginia. June 2007.
- 7 VII National Working Group. *PD-22.c-1 – Traffic Management: Corridor Management Planning Assistance (Day 1)*. Version 3. 2006 April 06.
- 8 Cambridge Systematics, Inc. *Travel Survey Manual*. Washington, DC: U.S. Department of Transportation and U.S. Environmental Protection Agency. June 1996
- 9 Michigan Department of Transportation. *Slippery Road Detection and Evaluation, Project Final Report*. May 31, 2012

APPENDIX C -UNITS CONVERSION TABLE

Measurement	U.S. standard to SI	SI to U.S. standard
Linear Measure	1 in=2.540000 cm 1 ft=0.3048000 m 1 mile=1609.3 m=1.6093 km	1 cm=0.3937008 in 1 m=3.280840 ft 1 m=6.2137x10 ⁻⁴ mile 1 km=0.62137 mile
Mass	1 lb=0.4535924 kg 1 oz=28.34952 g	1 kg=2.204622 lb 1 g=0.03527397 oz
Pressure	1 psi=6.894757 kPa 1 psi=0.06895 Bar 1 mm Hg=0.133 kPa	1 kPa=0.1450377 psi 1 Bar=100 kPa=14.504 psi 1 kPa=7.5 mm Hg
Volume (liquid)	1 gal=3.785412 l 1 ft ³ =0.02831685 m ³	1 l=0.2641720 gal 1 m ³ =35.31466 ft ³
Density	1 lb/ft ³ =16.01846 kg/m ³	1 kg/m ³ =0.06242797 lb/ft ³
Force	1 pound-force=4.4482 N	1 Newton = 0.2248 pound-force

Abbreviations for units of measure are as follows:

Unit of measure and abbreviation:

(SI): millimeter, mm; centimeter, cm; meter, m; gram, g; kilogram, kg; kiloPascal, kPa; liter, l; milliliter, ml; cubic meter, m³; Newton, N

(U.S.): Inch, in; foot, ft; ounce, oz; pound, lb; psig, psi; gallon, gal; cubic feet, ft³

APPENDIX D -MDOT FLEET DATA SPECIFICATION

In-Vehicle Device Configuration

The in-vehicle equipment used to obtain probe data from MDOT fleet vehicles consists of

- a data logger unit with a three-axis accelerometer mounted to the circuit board and a connector to the vehicle's OBD-II port
- a GPS receiver with serial cable communications to the data logger unit
- a Wi-Fi antenna for communication with the wireless access point

The data logger obtains data from the vehicle Controller Area Network (CAN) at the maximum speed the vehicle is able to return data through the OBD-II port, typically about 30 parameters per second. It uses its own on-board clock, and can be synchronized to the GPS time, a network time server, or be set manually.

The GPS chipset is a SiRFstarIII GSC3LT. It is sensitive down to -159dBm, with a two-dimensional RMS precision of 16 feet (5 meters) and a data rate of 1 Hz.

The accelerometer has a data range of ± 2 or ± 8 g and a sampling frequency of 100 or 400 Hz, depending on the device settings. For the DUAP accelerometry data collection, it is set to ± 2 g and 100 Hz.

Data from the logger are stored and delivered in three data files, each with its own file specification as described below: a probe vehicle data file (.IOS), a positions file (.GPS), and an accelerometry file (.ACC). New files are created when "key on" data is received from the vehicle through the OBD-II port and closed when the vehicle stops responding for more than 20 seconds (i.e., "key off").

Data File Specification(s)

The probe vehicle data file consists of a first-line parameter list, some header lines describing the file format version, start time for the file, and codes for the vehicle data corresponding to the file content. As shown in the example below, the time-dependent data follow the header lines and is in an ASCII-encoded hexadecimal format, except for the first parameter and any error parameters (e.g., "NODATA"). The first value (T) in each line is the elapsed time since the start of recording and the second value (dT2) refers to the time elapsed in acquiring all of the data in that row; both are in units of 10 msec. Parameters that may be in the data file (though not shown in the example) include:

- B0000010D Vehicle speed
- B00000110 MAF air flow rate
- B0000010C Engine RPM
- B00000105 Engine coolant temperature
- B0000010F Intake air temperature
- B00000133 Barometric pressure
- B0000011F Run time since engine start

- B00000111 Throttle position

```
T,dT2,B0000010D,B00000110
VERSION: 330
TIME: 04/22/2010 09:47:15
PROT: ISO15765-4 (CAN11/500) (6)
VIN: 1GDHG316991139713
Serial: 135
SUP 00: 983B8017
SUP 20: 00000001
SUP 40: 40000000
SUP 60: 00000000
SUP 80: 00000000
SUP A0: 00000000
SUP C0: 00000000
SUP E0: 00000000
335,335,00,04F3
920,585,00,04EB
1057,137,00,04F6
1197,140,00,0507
[...]
74003,56,13,23B1
74059,56,14,24CA
74123,64,14,25AE
74179,56,14,25AE
74236,57,14,267C
74300,64,14,267C
74364,64,15,272B
74420,56,15,272B
74476,56,15,2797
74537,61,15,27E5
74602,65,15,27E5
74656,54,16,2852
[...]
```

The second of the three files contains geo-location information. It is also a text file with a header row, VIN, and device serial number. This file differs from the IOS file in that there is no initial timestamp. Collectors processing these files must have all three files to form a complete set and are dependent on the IOS file for the initial time of recording.

The contents of the GPS file are a subset of the information recorded from a GPS receiver. Columns of interest to the DUAP system are the T (time), latitude, longitude, altitude, and NumSats.

```
T, dT2, Latitude, Longitude, Altitude, Velocity, Heading, Date, Time, FixType,  
NumSats  
VIN: 1D4GP24E17B151460  
Serial: 118  
19856, 1214, 42.675464, -84.663307, 271.1, 0.72, 353.32, 220410, 133434, 3, 5  
20673, 817, 42.675464, -84.663307, 271.2, 0.20, 334.47, 220410, 133435, 3, 5  
21507, 834, 42.675464, -84.663307, 271.3, 0.42, 331.14, 220410, 133436, 3, 5  
22492, 985, 42.675460, -84.663307, 271.2, 0.00, 0.00, 220410, 133437, 3, 5  
23475, 983, 42.675460, -84.663307, 271.2, 0.00, 0.00, 220410, 133438, 3, 5  
24497, 1022, 42.675460, -84.663307, 271.2, 0.00, 0.00, 220410, 133439, 3, 5  
25490, 993, 42.675460, -84.663307, 271.2, 0.00, 0.00, 220410, 133440, 3, 5  
[...]
```

The time column represents the number of milliseconds that have elapsed relative to the initial recording time. Latitude, longitude, and altitude (measured in meters), are self-explanatory. The number of satellites column is used to determine an accuracy of location. Fewer than four satellites usually results in the location information, and hence an entire record, not being useable.

The last of the three files is the ACC file. It contains the three axis acceleration data. The file is mostly binary data with a few text strings intermixed. No excerpt of the file is given in this text as the binary data would not be intelligible or very useful. The text content of the file includes the VIN and device serial number in the header and millisecond time tags interspersed at approximately one-second intervals among the binary data. <1519> is an example of a millisecond time tag.

The binary data are arranged in sets of approximately one hundred 4-byte structures. The first byte is ignored as it deals directly with the embedded accelerometer hardware configuration. The subsequent three bytes represent the X, Y, Z axes respectively, with one byte allocated to each. The ± 2.3 g range is mapped to the byte values 0 to 255 resulting in a minimum step measurement of approximately 0.018 g.

File Identification and Transport

File name formats are encoded as <vin>-yMddhhmm.ext, where “<vin>” is the vehicle identification number, “y” is the last digit of the year; “M” is the alpha-sequenced month (e.g., “A” is January); “dd” is the day of the month; “hh” is the hour of the 24-hour day; “mm” is the minutes past the hour; and “ext” is “IOS”, “GPS”, or “ACC”, depending on the content type. The VIN can be read from the vehicle through the OBD-II port for most vehicles, or defaulted to a value in the configuration file.

Files are transmitted whenever a Wi-Fi connection at a wireless access point is determined to be available. In the current DUAP configuration, the data files are accepted through the HTTP POST protocol at:

<http://duapmdot.net/tf01/<filename.ext>>

Files are deleted from the device cache where the file transfer is determined to be complete and is acknowledged. Transfers that are interrupted or time out are incomplete, so the files are retained and transfer is retried at the next available connection.

APPENDIX E - DUAP INPUT DATA SPECIFICATION

Data File Specification

The DUAP input data specification is intended to provide a standard for exchanging connected vehicle data that does not put explicit requirements on the data types to be transferred. The concepts behind this specification are that it be flexible, easily readable, as terse as practical, and transportable over many different communication media. To these ends it borrows heavily from successful technology standards such as web log files and hypertext transfer protocol (HTTP).

The file encoding is UTF-8. The file format itself consists of a header section containing key-value pairs, a comma-separated-value list, and a footer section also containing key-value pairs. Every line must have a line terminator. Header and footer sections are separated from the main data content by double line terminators. A carriage return character in the line terminator is optional and, in the interest of brevity, the lone newline character is preferred. The time window each file represents is flexible, but file sizes should be kept small enough to ensure reliable transmission over relevant media.

The first and last records in the body represent the exact recorded values and are critical to an accurate representation of the intervening data. The intervening data records are delta values from the initial record, except for the accelerations, which must each represent the measured value. The delta time column is always the milliseconds after the timestamp specified in the header.

The fields key in the header description is a list that specifies the order of the data columns and the labels for each data column with their respective units separated from the label by a single hyphen. The units were chosen to save some space by eliminating the need for a decimal point—using micro-degrees for geo-coordinates, for example.

It should be emphasized that the delta values are relative to the first record and not each subsequent record. This method minimizes any potential cumulative error. The comma delimited format allows representation of null values and effective mixing of data types with different sampling frequencies—in the example, GPS data and acceleration—into a single file with minimal overhead. The text in brackets is informative for this example only and is not a commenting mechanism.

```
type: android
version: 1.0
source-id: <UUID> or <MAC>
timestamp: yyyy-MM-ddTHH:mm:ss.SSSZ [2001-07-04T12:08:56.235-0700, for example]
fields: dt-ms,ax-mm/s2,ay-mm/s2,az-mm/s2,lat-u°,lon-u°,alt-dm,gps_sats,gps_est_spd-cm/s
0,9800,-200,400,42558104,-106845336,1953,5,0
10,9810,-190,420,,,,,
.
.
99,4900,-110,250,152,-21,84,4,376
100,4817,984,2646,,,,,
.
.
5921, 8820, -981, 1960, 42559205, -106842991, 1523, -1, -100
record-count: 5922 [should be 6000 for 60 seconds at 100 Hz]
```

Data may also be accepted in Java archive (JAR) files that typically achieve better than 3 to 1 compression. The system automatically extracts the data when it is submitted. This option may be important as smaller files have a better chance of being successfully transmitted over occasionally unreliable networks, more data can be stored locally in the event a transmission fails and a later retry is needed, and costs may be lower if there is a per-byte charge by the telecom carrier. Each JAR file should contain only one file using the <source-id>-<timestamp>.csv naming convention.

File Identification and Transport

In the current DUAP configuration, the comma-separated variable (CSV) data files are accepted through the HTTP POST protocol. Two variations of URL forms are accepted:

<http://duapmdot.net/tf02/<source-id>/<timestamp>.csv> or

<http://duapmdot.net/tf02/<source-id>-<timestamp>.csv>

The <source-id> in the URL is the same as the source-id within the file. Any text string recognizable in the URL is acceptable. DUAP itself does not attempt to make any association between this <source-id> and any other identifier of the source of the data.

The timestamp format within the URL should be yyyyMMddTHH:mm:ssZ, where the “T” delimits the time from the date, and the “Z” defines the UTC (zenith) offset for the time zone of data origination. A timestamp represented within the data file as “2010-07-20T19:10:45.503-0500” (where “-0500 represents Eastern Standard Time in the US and Canada) would be presented in the destination filename as “20100720T191045-0500”.

Similar to the CSV files, the JAR files may be submitted through HTTP POST using either one of two different URL forms:

<http://duapmdot.net/tf02/<source-id>/<timestamp>.jar> or

<http://duapmdot.net/tf02/<source-id>-<timestamp>.jar>

The source-id and timestamp should likewise conform to the convention discussed for the CSV files.