OVERVIEW OF THE VEHICLE DATA TRANSLATOR

S. Drobot, M. Chapman, E. Schuler, G. Wiener, and W. Mahoney III
National Center for Atmospheric Research, Boulder, CO, USA
drobot@ucar.edu, mchapman@ucar.edu, eschuler@ucar.edu, gerry@ucar.edu,
mahoney@ucar.edu

P. Pisano
USDOT/Federal Highway Administration, Washington, D.C., USA
paul.pisano@dot.gov

B. McKeever
USDOT/Research and Innovative Technology Administration, Washington, DC, USA
ben.mckeever@dot.gov

ABSTRACT

In a typical year, there are 1.5 million weather-related vehicle crashes in the U.S., leading to 673,000 injuries and nearly 7,400 fatalities. Adverse weather and the associated poor roadway conditions are also responsible for 554 million vehicle-hours of delay per year in the U.S., with associated economic costs reaching into the billions of dollars. With funding and support from the U.S. Department of Transportation's (USDOT) Research and Innovative Technology Administration (RITA) IntelliDrive[SM] initiative and direction from the Federal Highway Administration's (FHWA) Road Weather Management Program, the National Center for Atmospheric Research (NCAR) is conducting research to develop a Vehicle Data Translator (VDT) that incorporates vehicle-based measurements of the road and surrounding atmosphere with other, more traditional weather data sources, and creates road and atmospheric hazard products for a variety of users. This is one possible solution for mitigating the adverse impacts of weather on the transportation system. This paper presents an overview of the current VDT development and upcoming modifications.

1. INTRODUCTION

The transportation community is well on its way toward the development of wireless vehicle capabilities whereby vehicles communicate with other vehicles and the road infrastructure to improve safety and mobility and to reduce environmental impacts [1–4]. In the near future, millions of vehicles will anonymously collect direct (e.g., temperature) and indirect (e.g., ABS) measurements of the road and atmospheric conditions in their immediate surroundings. This will greatly expand the current weather observation network, particularly with respect to the roadway environment.

The volume, anonymity, and mobility characteristics of vehicle-based observations pose several challenges related to data quality and consistency. These must be addressed before the data will be broadly usable and acceptable. In contrast, weather observations collected from standard stationary platforms have a constant, known location. One solution to address these challenges is to develop a "Vehicle Data Translator" (VDT). The main function of the VDT is to quality-check (QCh) individual vehicle probe data elements, such as temperature and pressure, and then combine them into "derived observations" that are valid along a given length of roadway over a given time period.

Under contract with the USDOT, research and development are being conducted by the National Center for Atmospheric Research (NCAR) to design, develop, and test the VDT concept [5–6]. Data from the Detroit Developmental Test Environment (DTE) Proof of Concept (POC) test, DTE 2009 Winter Demonstration, and Michigan Data Use Analysis and Processing (DUAP) Project are being used to evaluate and test the prototype VDT.

This paper outlines recent progress on development of a VDT.

2. VEHICLE PROBE DATA - WEATHER DATA PROCESSING

Vehicle data can be complex and they pose a significant analytical challenge, particularly for measuring or deriving weather and road condition data. For example, some obvious issues include dealing with the large data volume, timeliness of the data, data quality and representativeness, and data format(s). The uncertainties and complexity associated with raw vehicle data (i.e., unprocessed data from vehicles) may deter many end users from using the raw data. In our opinion, most end users will either not be able to—or not want to—handle the immense volume of vehicle data, let alone deal with data quality questions. We anticipate that many, if not most, users of vehicle weather data will require *processed* data. In this context, *processed* data means vehicle data that are extracted from the network of vehicle probe data, quality checked, and disseminated to active data subscribers in near real time. It is also likely that, due to the volume of data, many users will prefer statistically-derived data representing specific geographical areas (such as road segments) or times. Accordingly, we developed the VDT with several primary processing components as shown in Figure 1. Data from IntelliDrive^{SM}-equipped vehicles (e.g., personal and fleet vehicles) are communicated to the Road Side Units (RSEs) when the vehicles are within range of the receivers. The RSEs are connected to the IntelliDrive^{SM} communications network where most of the data will flow. Individual processing components of the VDT are described in the following sections.

The prototype VDT includes a data parser function that extracts relevant weather and road condition vehicle probe fields from the vehicle data network. The data elements selected for extraction will be determined by research results and feedback from stakeholders in both the atmospheric and surface transportation communities. Data elements can also be added or subtracted as needs vary. The data flowing out of the data parser is still considered raw as it has not been processed in any way.

Data filtering algorithms are then applied to chosen data elements to remove data that are not likely to be representative of the true conditions. For example, outside air temperature measurements may not be representative of the true ambient conditions if the vehicle speed is less than 25 mph [7]. Filters can also be applied to data collected at particular locations that are known to generate errors (e.g., data measurements from inside tunnels). The process of deciding when and how to filter data will require considerable more research and will need to be done with great care, as one would not want to remove data that may have some value.

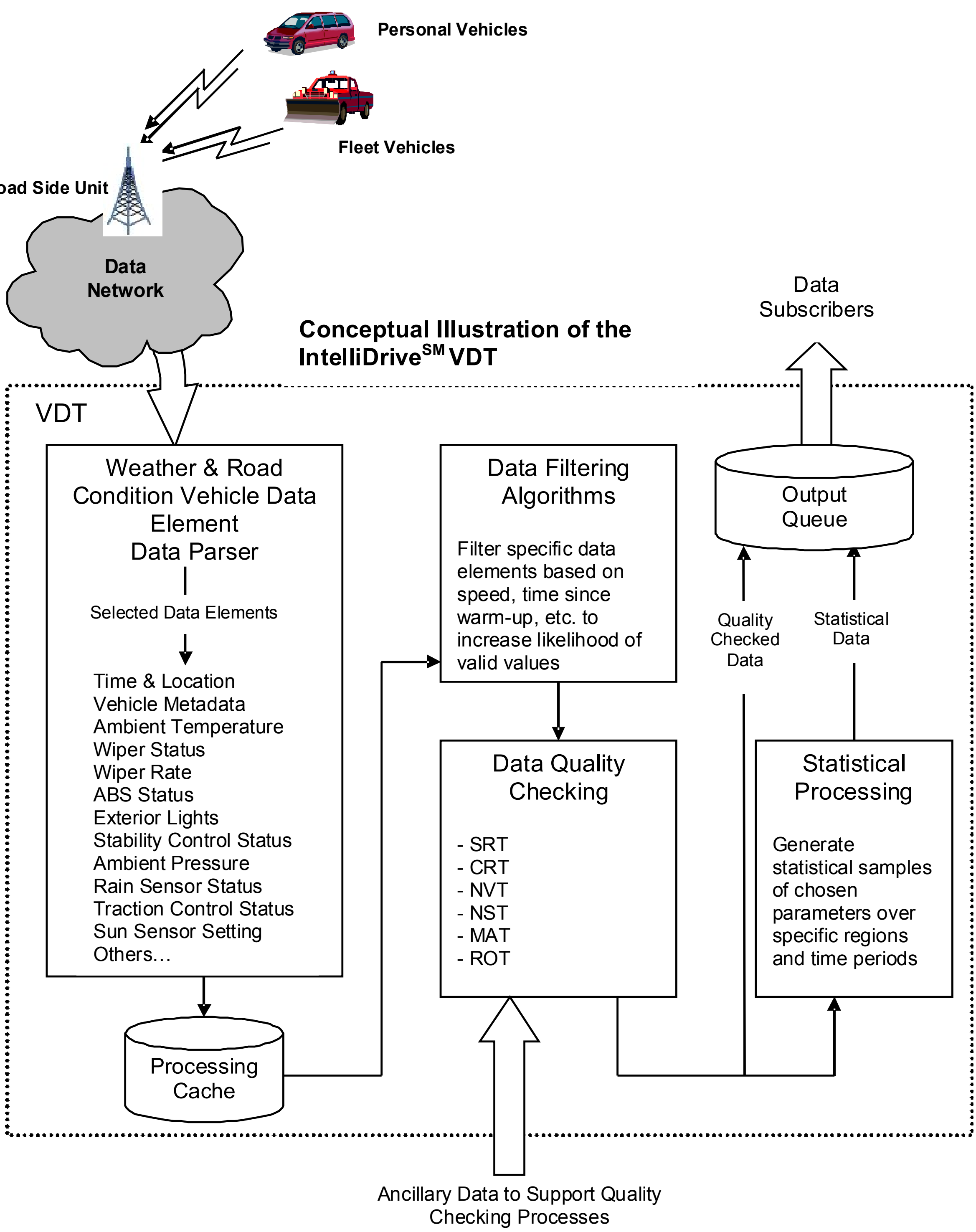

Figure 1 – Vehicle Data Translator (VDT) Conceptual Illustration

A benefit of removing data that are considered unrepresentative in the filtering procedure is a reduction in the data volume that will need to undergo more complex and computationally intensive quality checking (QC) procedures. The QC tests include many of the common tests that are applied to surface weather data in the *Clarus* system [8], as well as more complex tests to handle road condition data. In both cases, ancillary data, including as surface weather observations, satellite, radar, climatological data, and model output statistics are required to conduct many of the QCh tests. The quality checking methods used in the VDT include:

Sensor Range Test (SRT): Identifies observations that fall outside the range of the known sensor hardware specifications. Unlike the *Clarus* system, data processing algorithms included in the VDT will not have the advantage of "knowing" what type of sensor (e.g., make, model) produced each measurement. In order to develop an effective SRT, it will likely be necessary to conduct research on the automotive sensors that are available in the market place and make an educated assessment of the reasonable bounds for this test.

Climatological Range Test (CRT): Flags observations that fall outside of location-specific climatological (or historical) ranges. This is a more complex task than the SRT because of the variability of the climatological range values over times, dates, locations, and seasons. Once a climatological range is assigned for a specific road segment or geographic region (e.g., based on weekly values), the test will simply allow for observations within that range and flag observations that fall outside of the range.

Neighboring Vehicle Test (NVT): Compares the observation to neighboring vehicles in the road segment. If the observation value falls outside of a dynamic threshold, then the observation will fail the test. The threshold likely will be determined by the number of neighboring observations that are available for comparison. In other words, the higher the quantity of vehicles in a road segment, the tighter the threshold range.

Neighboring Surface Station Test (NST): Compares neighboring surface station observations (e.g., Road Weather Information System (RWIS), Automated Surface Observing Systems (ASOS), Automated Weather Observing Systems (AWOS)) to each vehicle observation. This test will obviously depend on the availability of a surface observation within a suitable distance of the road segment. Currently, the VDT uses a 50-km radius from the vehicle location.

Model Analysis Test (MAT): Compares observations with a model surface analysis, such as the Real-Time Mesoscale Analysis (RTMA) or the Rapid Update Cycle (RUC) Surface Assimilation Systems (RSAS). This test will compare the observed value to either a range of grid values along a road segment or within a predetermined threshold of the grid value closest to the location of the vehicle.

Remote Observation Test (ROT): Compares vehicle observations to remotely sensed data from either satellite and/or radar. Both the radar and satellite can be used as either a complimentary step with the other QC tests or as the final option if no neighboring vehicles or surface stations are available. As an example, pavement temperature, rain and sun sensor observations are good candidates for this test.

In the QC process, data quality flags will be applied to the raw vehicle data so that data subscribers will have the flexibility of utilizing the raw data or taking advantage of the quality checking flags. After the QC procedure, some data will flow directly to the output queue to minimize data latency (Figure 1).

Some of the data will be cached and processed to generate statistical values for a given location (e.g., grid cell or point) and time period. The statistical processing will create two separate data streams: "processed data" and "derived data". To generate processed data the application of a statistical technique (e.g., mean, median) to a set of observations over a known grid segment is performed, and it should result in a more robust sample and reduce the overall data load for users that either cannot or do not want individual vehicle data. To generate derived data, a more computationally intensive procedure is necessary that focuses on deriving new or enhanced road and atmospheric variables of interest to the surface transportation community [9]. The derived variables will be a blend of weather and non-weather related IntelliDrive℠ data elements in conjunction with ancillary data sets. Table 1 lists possible derived observations that are currently being considered during the incipient phases of IntelliDrive[SM], the ancillary data required to construct each derived observation, and the relevant vehicle observations (observed and input).

Table 1 – Sample Matrix for VDT Derived Weather-Related Observations

Derived Observation	Ancillary Data	Vehicle Observation
Ambient Air Temperature	RSAS, Surface Stations	Vehicle Speed, Temperature, Hours of Operation
Barometric Pressure	RSAS, Surface Stations	Vehicle Speed, Barometric pressure, Hours of Operation
Precipitation Occurrence	Radar, Surface Stations, Satellite	Wiper Status, Vehicle Speed, ABS, Stability Control, Traction Control, Headlight Status
Precipitation Type	Radar, Surface Stations, RSAS, Sounding	Wiper Status, Vehicle Speed, Temperature, Elevation, ABS, Stability Control, Traction Control
Precipitation Rate	Radar, Surface Stations	Wiper Status, Rain Sensor, Vehicle Speed, ABS, Stability Control, Traction Control, Headlight Status
Fog	Radar, Surface Station Visibility, Satellite Cloud Classification	Wiper Status, Rain Sensor, Vehicle Speed
Pavement Conditions (wet, dry, icy)	Radar, Surface Stations	Vehicle Speed, ABS, Stability Control, Traction Control, Temperature, Wiper Status
Boundary Layer Water Vapor	Radar, Soundings, Surface Stations	Temperature, Wiper Status
Pavement Temperature	Surface Stations	Temperature, Vehicle Speed
Smoke	Satellite Smoke Algorithm, Surface Station Visibility,	Vehicle Speed, Headlight Status

It should also be noted that the merging of different data sources (vehicle observations and ancillary data) will require the use of expert systems (e.g., fuzzy logic, neural networks) and/or decision tree algorithms in order to produce robust observations. It is likely that a combination of techniques will be required for some of the products.

3. VDT DISPLAY

A key component of the VDT concept includes the capability to display the data. A prototype display has been developed to support VDT research and to convey the utility of vehicle-based data. The VDT display is Java-based and it will be initially configured to display vehicle probe data sets covering the Detroit DTE region. The display can be re-configured to support other locations using an XML configuration file. The display will be capable of rendering the following data types:

- Vehicle probe messages
- Standard surface weather observations
- Gridded radar data
- RSAS data
- Satellite cloud mask data
- Road segment statistics (data range, variance, means, modes, etc.)

- Derived road segment values

The display is divided into 4 major areas, a toolbar on top, a map area in the center, a time slider window below the map and a data layer section on the bottom (Figure 2). The data layer section displays the available datasets using a list with checkboxes. The user can select a dataset by checking the corresponding data layer and selecting a time of interest in the time slider section. Point observations, such as ASOS reports and probe messages, are displayed as points on the display. The user can click on a display point to bring up underlying data values in text form. In order to display road segment statistics and weather, the user can click on or mouse over the appropriate road segments.

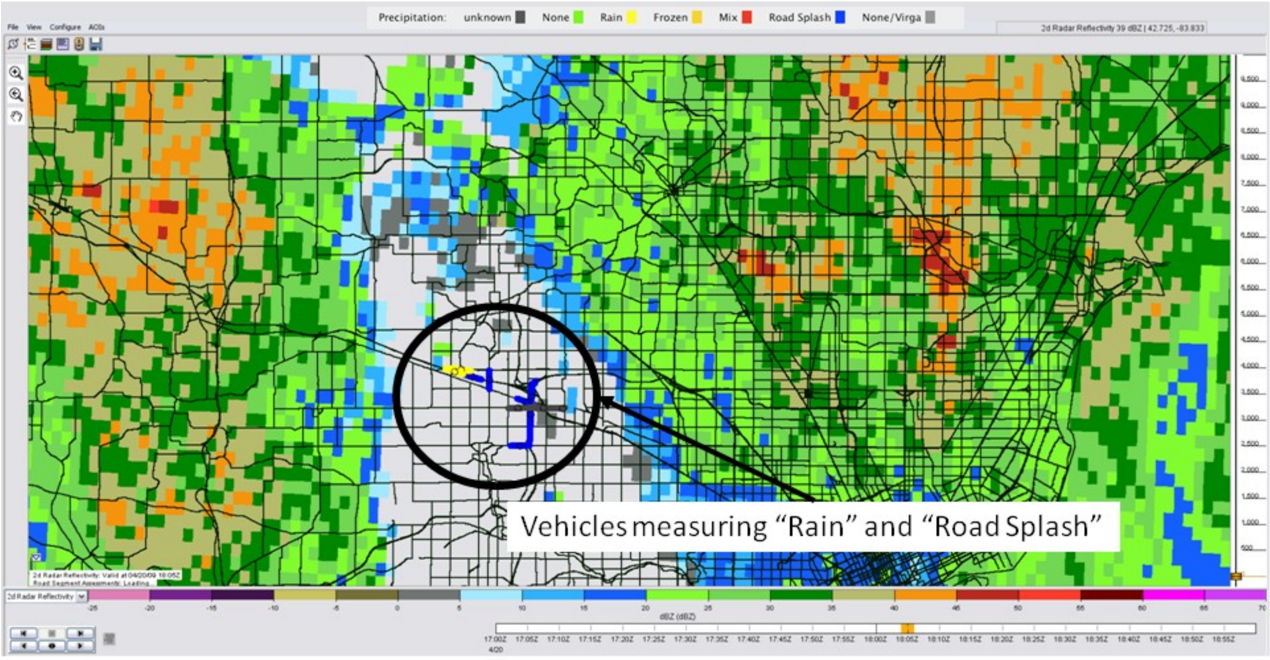

Figure 2 – Example of prototype VDT display.

4. CONCLUSIONS

From a weather perspective, the overarching goal of the IntelliDrive^{SM} initiative is for the public and private organizations that collect, process, and generate weather and transportation products to utilize vehicle data to improve weather and road condition products, and then to provide those products to transportation system decision-makers, including travelers. Nonetheless, the utilization of data from vehicles poses significant technical challenges, particularly with respect to data quality and quantity. The amount of data potentially flowing through a vehicle-based data network, such as IntelliDrive^{SM}, could be immense and it is likely that many prospective users will not be capable of handling the vast quantities of data that are expected.

The VDT discussed in this paper is one approach to preprocess weather-related vehicle data before they are distributed to data subscribers. The function of the VDT is to extract data elements needed to derive weather and road condition information from vehicle probe data, filter the data to remove samples that are likely to be unrepresentative, quality check the data utilizing other local surface observations and ancillary datasets, generate statistical output for specific road segments, and disseminate the quality-checked and statistically processed data to subscribers, which may include other data processing and dissemination systems such as the U.S. Department of Transportation's (USDOT) *Clarus* System.

5. ACKNOWLEDGEMENTS

Research conducted during this project was funded by FHWA Contract DTFH61-08-C-00012.

6. REFERENCES

(1) Srinivasan, K.K., and P. P. Jovanis. Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1537, Transportation Research Board of the National Academies, Washington, D.C., 1996, pp. 15–22.

(2) Cheu, R.L., H. Qi, and D.-H. Lee. Mobile Sensor and Sample-Based Algorithm for Freeway Incident Detection. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1811, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 12–20.

(3) Shladover, S.E., and T.M. Kuhn. Traffic Probe Data Processing for Full-Scale Deployment of Vehicle–Infrastructure Integration. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 2086, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 115–123. DOI: 10.3141/2086-14

(4) Department of Transportation IntelliDrive^{SM} Website. http://www.intellidriveusa.org/. Accessed July 27, 2009.

(5) Drobot, S.D., M. Chapman, T. Jensen, G. Weiler, and E. Schuler (2009). Creation of VII-based Derived Observations Recommendation Report. National Center for Atmospheric Research, Boulder, Colorado. 37 pp.

(6) Drobot, S.D., M. Chapman, T. Jensen, G. Weiler, E. Schuler, and C. Johansen (2009). Post-Demonstration Analysis Report. National Center for Atmospheric Research, Boulder, Colorado. 63 pp.

(7) Stern, A.D., P. Pisano, P. Kennedy, K. Petty, W. Mahoney (2008). A next generation of observations based on passenger vehicles. American Meteorological Society Annual Meeting, New Orleans, Louisiana, January 24, 2008, and published in the 88th Annual Meeting of the American Meteorological Society Proceedings.

(8) United States Department of Transportation (2006). Clarus Weather System Design: System Design Description Version 1.1, USDOT, ITS Joint Program Office. Washington, D.C. 2006.

(9) Petty, K.R. and W.P Mahoney (2007). Weather Applications and Products Enabled Through Vehicle Infrastructure Integration (VII): Feasibility and Concept Development Study. National Center for Atmospheric Research, Boulder, Colorado. 2007. 116 pp.