SUPPLEMENT A

MONITORING SYSTEM ARCHITECTURE

OVERVIEW

This supplement supports the discussion presented in Chapter 2 and provides more detailed examples of general types of tables for storing and managing data. As discussed in Chapter 2, five general types of tables are commonly used in travel time reliability monitoring systems:

1. Configuration information.
2. Raw data.
3. Travel times.
4. Travel time density functions for the segment and route regimes.
5. Reliability summaries.

Database design usually takes the form of a schema, which formally describes the database structure, including the tables, their relationships, and constraints on data value types and lengths. Rather than defining an implementable schema, this chapter presents example tables that can store information generated during all steps of the reliability monitoring computation process, from the raw data to the travel time density functions and reliability metrics. The exact tables, fields, and relationships are flexible to the needs of the agency, the data available, and the desired reporting capabilities.

CONFIGURATION INFORMATION

The configuration information stored by a reliability monitoring system must include certain details about the freeways and arterials that underlie the routes being monitored. The most important information associates sensors with specific locations on a facility (usually marked by post miles), so that the physical distance between sensors can be computed. In addition to the geographic location, it is important to know which lane(s) a sensor monitors, and if the lane is a mainline or managed facility lane. After information has been established for key freeways and arterials, routes must be designed in the configuration tables. In this context, a route is composed of a number of contiguous segments of facilities, and can include both freeway segments and arterial segments. The travel times for these contiguous segments are ultimately aggregated to produce a total route travel time. Exhibit A-1 shows an example of the database table for route configuration information, and Exhibit A-2 shows an example of the definition of route segments.

Exhibit A-1
Route Configuration Table

Column #	Field	Description
1	Route ID	Unique identifier for the route.
2	Route Name	Name of the route.
2	Length	Length of the route.
3	Region ID	Unique identifier of the route's region.

Exhibit A-2
Route Segment Configuration Table

Column #	Field	Description
1	Route ID	Unique identifier for the route.
2	Route Segment ID	Unique identifier for the route segment. This is the relative position of the segment down a route.
3	Facility ID	Unique identifier for the facility and direction that the route is on. Links to tables with detector/sensor configuration information for the facility.
4	Managed Facility?	Whether the route segment has managed facility lanes.
5	Start Point	Description of the route segment's start point.
6	Start Point Postmiles	Absolute postmiles of the route segment's start point.
7	End Point	Description of the route segment's endpoint.
8	End Point Postmiles	Absolute postmiles of the route segment's end point.

RAW DATA

The form of the raw data, and thus the structure of the raw data tables in the database, depends on whether the data is from an infrastructure-based detector, an AVI sensor, or an AVL technology. Infrastructure-based detectors usually transmit data every 30-seconds, and data consists of some combination of flow, occupancy, and speed values. In the raw database table, each record represents the 30-second data summary. AVI sensors transmit data every time a vehicle equipped for sampling passes by. In the raw database table, each record represents a vehicle passing a sensor. AVL systems provide similar data if virtual monuments are defined that function the same as AVI tag readers. The raw AVL data are message packets containing the latitude, longitude, speed, and heading of the vehicle at some sampling rate, often every few seconds. Since the data from all three technology types is different, each will require its own table in the database.

Exhibit A-3, Exhibit A-4, and Exhibit A-5 show sample tables for raw infrastructure-based data, raw AVI data, and raw AVL data, respectively.

Exhibit A-3
Raw Infrastructure-based Data Table

Column #	Field	Description
1	Time ID	Timestamp for the 30-second period.
2	Detector ID	Unique identifier for the reporting detector. Links to information about the detector's facility-direction, location, and lane number.
3	30-sec flow	The vehicles counted in the 30-second period.
4	30-sec occupancy	The average occupancy over the 30-second period.
5*	30-sec speed*	The average speed over the 30-second period.*
		* where observed

Exhibit A-4
Raw AVI and AVL Data Table – Timestamps at Specific Locations

Column #	Field	Description
1	Time ID	Timestamp of the vehicle's arrival at the sensor or monument location.
2	Sensor ID	Unique identifier for the sensor or monument. Links to information about the sensor or monument's facility- direction, location, and lane number.
3	Vehicle ID	Unique identifier for the reported vehicle. For example, the tag ID, the vehicle ID, or a Bluetooth address (with changes to protect privacy).

Exhibit A-5
Raw AVL Data Table

Column #	Field	Description
1	Time ID	Timestamp of the vehicle polling.
2	Vehicle ID	Unique identifier for the vehicle being polled.
3	Longitude	The vehicle's longitude.
4	Latitude	The vehicle's latitude.
5	Speed	The vehicle's speed.
6	Bearing	The vehicle's bearing.

TRAVEL TIME INFORMATION

The information in the raw data tables needs to be processed so that travel time information can be developed for each segment, route, and time period. For infrastructure-based sensors, estimating travel times requires imputing missing data values, computing speeds from volume and occupancy values, and extrapolating point speeds over spatial segments to derive travel time information. For AVI systems, it requires computing travel times for matched vehicles, filtering out bad travel times that are likely representative of longer trip times, and calculating travel time information from the good samples. For AVL technologies, it requires matching the raw data to segment or route end points, and calculating travel time information from the observed values. Where agencies blend these technologies to develop travel time information for a given segment or route, data fusion can provide more accurate average travel time information. From a database standpoint, data fusion requires the travel time information derived from each individual technology type to be aggregated up to a common temporal and spatial level so that fusing of the data can occur. The examples in this section assume that travel time information is aggregated at the 5-minute level and spatially aggregated to the segment and route level. This section illustrates sample database tables for each technology type. The flow of tables reflects the computational process used to turn raw data into final travel time information. Finally, it presents data tables that can store fused 5-minute travel time information and hourly travel time information summaries, from which reliability statistics can be computed.

Infrastructure-Based Detectors

This section shows four sample database tables for infrastructure-based detectors, which store data produced by the steps used to compute route-level, 5-minute travel time information. Exhibit A-6 stores the results following imputation of missing or bad 30-second flow and occupancy raw data samples. The Ensuring Data Quality section of Chapter 11 describes the methods for performing this imputation. Exhibit A-7 shows the results of aggregating all data to the 5-minute level and computing speeds from the 5-minute flow and occupancy values. The methods for performing this speed calculation are described in the Calculating Travel Times section of Chapter 11. It also includes a field for storing the percent observed, a measure of the data validity. Exhibit A-8 stores the results of extrapolating the 5-minute detector speed data over a defined segment, to compute segment-level travel time information. In this table, the travel time information represents travel times for vehicles ending their travel on that segment during the 5-minute period. Finally, Exhibit A-9 stores the final infrastructure-based travel time information for each defined route. The route-level travel time data are computed by “walking the travel time field” along each segment in the route, a process described in the previous chapter. Chapter 14 and Appendix M describe the data management of inductive loop and radar-based sensor data for the Northern Virginia Case Study.

Exhibit A-6
30-second Infrastructure-based Data with Imputation

Column #	Field	Description
1	Time ID	Timestamp for the 30-second period.
2	Detector ID	Unique identifier for the reporting detector. Links to information about the detector’s facility-direction, location, and lane number.

3	30-second Flow	The 30-second volume count.
4	30-second Occupancy	The 30-second average occupancy.
5	Imputed?	Whether the data values are observed or imputed.

Exhibit A-7
5-minute Infrastructure-based Data with Speeds

Column #	Field	Description
1	Time ID	Timestamp for the 5-minute period.
2	Detector ID	Unique identifier for the reporting detector. Links to information about the detector's facility-direction, location, and lane number.
3	5-min Flow	The vehicles counted in the 5-minute period.
4	5-min Occupancy	The average occupancy over the 5-minute period.
5	5-min Speed	The average speed over the 5-minute period. Computed from the flow, occupancy, and stored g-factor for that detector, time of day, and day of week.
6	% Observed	The percentage of data points directly observed from the detector, as opposed to imputed.

Exhibit A-8
5-minute Infrastructure-Based Segment Travel Times

Column #	Field	Description
1	Time ID	Timestamp for the 5-minute period.
2	Segment ID	Unique identifier for the segment.
4	Travel Time	Average travel time for the route segment.
5	Operative Regime	Regime that was operative during the 5-minute period.
6	% Observed	The percentage of data points directly observed on the route segment, as opposed to imputed.

Exhibit A-9
5-minute Infrastructure-Based Route Travel Times

Column #	Field	Description
1	Time ID	Timestamp for the 5-minute period.
2	Route ID	Unique identifier for the route.
3	Travel Time	Average travel time for the route.
4	Operative Regime	Regime that was operative during the 5-minute period.
5	% Observed	The percentage of data points directly observed on the route, as opposed to imputed.

AVI Data

This section shows three sample database tables designed to store AVI information derived during the processing of trip time information. The database tables would contain travel times that have been extracted from raw trip times collected by AVI sensors. Exhibit A-10 stores travel times for all matched vehicles between specific sensor pairs. In some instances the starting (Sensor ID 1) and ending (Sensor ID 2) sensor represent a single network segment. In other instances, they represent a segment sequence or an entire route. Exhibit A-11 filters the data in Exhibit A-10 to describe travel times for routes that have been defined in the system. It also stores whether the travel time is considered valid following the filtering process.

Exhibit A-12 stores the final results of the AVI-based travel time computations. It stores 5-minute travel time information for each segment and route. The Data Management section of Appendix N describes the primary data management software system for the Lake Tahoe case study, which examined travel time reliability based on Bluetooth data.

Exhibit A-10
AVI Vehicle Travel Times for Segments and Sensor-to-Sensor Pairs

Column #	Field	Description
1	Time ID	Timestamp of the vehicle's arrival at the first sensor.
2	Sensor ID 1	Unique identifier for the first sensor.
3	Sensor ID 2	Unique identifier for the second sensor at which the vehicle is matched.
4	Vehicle ID	Unique identifier for the reported vehicle. For example, the tag ID or Bluetooth address (with changes to protect privacy).
5	Travel Time	The travel time between the two sensors.
6	Valid?	Whether the travel time is considered valid after filtering.

Exhibit A-11
AVI Vehicle Route Travel Times

Column #	Field	Description
1	Time ID	Timestamp of the vehicle's arrival at the first sensor on the route.
2	Route ID	Unique identifier for the route.
3	Vehicle ID	Unique identifier for the reported vehicle. For example, the tag ID or Bluetooth address (with changes to protect privacy).
4	Travel Time	The travel time between the first and last sensor on the route.

5	Valid?	Whether the travel time is considered valid after filtering.

Exhibit A-12
5-minute AVI-based Representative Segment or Route Travel Times

Column #	Field	Description
1	Time ID	Timestamp of the 5-minute period.
2	Segment or Route ID	Unique identifier for the segment or route.
3	Number of Valid Samples	The number of valid vehicle travel times observed.
4	Average Travel Time	The average travel time measured from the valid AVI matches for the segment.
5	Operative Regime	The regime that was operative on the segment or route during the 5-minute period.

AVL Technologies

This section shows two database tables used to store AVL-based travel time information. Exhibit A-13 stores the results of matching the raw data to a route in the system and calculating route travel times for individual vehicles. It also contains a field indicating whether each vehicle's travel time is considered valid following the filtering process.

Exhibit A-14 stores the final results of the computation process for AVL data: a representative travel time for the 5-minute period, computed from all valid vehicle samples collected in the time period. Similar to the final AVI travel time table, this table includes a field to store the travel time variability among individual travelers, as well as the error inherent in the representative travel time estimate. The data management system for AVL data in the San Diego case study is described in Chapter 13 and Appendix L.

Exhibit A-13
AVL Vehicle Travel Times

Column #	Field	Description
1	Time ID	Timestamp of the vehicle's arrival at the beginning of the segment or route.
2	Segment or Route ID	Unique identifier for the segment or route.

3	Vehicle ID	Unique identifier for the vehicle being observed.
4	Travel Time	Route travel time for the vehicle being observed.
5	Valid?	Whether the travel time is considered valid after filtering.

Exhibit A-14
5-minute AVL-based Segment or Route Travel Times

Column #	Field	Description
1	Time ID	Timestamp of the 5-minute period.
2	Segment or Route ID	Unique identifier for the segment or route.
3	Number of Valid Samples	The number of valid vehicle travel times observed.
4	Average Travel Time	The average travel time on the segment or route measured from the valid AVL samples.
5	Operative Regime	The regime that was operative on the segment or route during the 5-minute period.

Data Fusion

For routes that are being monitored by more than one technology type, data fusion can be used to improve the accuracy of the travel time information derived from a single technology. Data fusion requires travel time data from the individual sensor types to be aggregated to the same temporal and spatial level. These technology-specific data can be combined using weighting factors to produce enhanced information.

Exhibit A-15 shows a sample database table that contains the final, fused travel time information. It includes fields indicating the time of day and day of week, to facilitate specific query requests. It also includes fields indicating which technologies contributed data to the final information. Exhibit A-16 shows the same data aggregated up to the hourly level, to facilitate more high-level travel time and reliability analysis.

Exhibit A-15
Fused 5-minute Segment or Route Travel Times

Column #	Field	Description
1	Time ID	Timestamp of the 5-minute period.
2	Time of Day	5-minute period (without the day).
3	Day of Week	Day of the week.
4	Segment or Route ID	Unique identifier for the segment or route.
5	Average Travel Time	Average segment or route travel time during the 5-minute period.
6	Includes Infrastructure-based Estimate?	Whether infrastructure-based data are included in the average travel time.
7	Includes AVI-based Estimate?	Whether AVI-based data are included in the average travel time.
8	Includes AVL-based Estimate	Whether AVL-based data are included in the average travel time.
9	Operative Regime	The regime that was operative on the segment

		or route during the 5-minute period.
10	Error	An estimate of error based on the % observed (infrastructure-based) and sample sizes (AVI/AVL).

Exhibit A-16
Fused Hourly Segment or Route Travel Time Summaries

Column #	Field	Description
1	Time ID	Timestamp of the hourly period.
2	Time of Day	Hourly period (without the day)
3	Day of Week	Day of the week
4	Segment or Route ID	Unique identifier for the segment or route.
5	Average Travel Time	Average segment or route travel time during the hour.
6	Includes Infrastructure-based data?	Whether infrastructure-based data are included in the summary.
7	Includes AVI-based data?	Whether AVI-based data are included in the summary.
8	Includes AVL-based data?	Whether AVL-based data are included in the summary.
9	Operative Regime	The regime that was operative on the segment or route during the hour.
10	Error	An estimate of error based on the % observed (infrastructure-based) and sample sizes (AVI/AVL).

OPERATING REGIMES

Determination of the travel time reliability regime for each segment or route and 5-minute time period is based on a matching process in which real-time observations of segment or route travel times are compared against non-parametric probability density functions that represent the regimes in which the segment or route has been found to operate.

This section describes the data tables in which the non-parametric probability density functions are stored that describe the various regimes. The methodology described in Chapter 3 is based on the assumption that these regimes can be described by the percentiles of the distribution. The table shown in Exhibit A-17 stores the percentile values.

Exhibit A-17
Non-Parametric Density Function Summary

Column #	Field	Description
1	Segment or Route ID	Unique identifier for the segment or route.
2	Regime	Unique identifier for the specific regime to which the information in the record pertains
3-103	Percentile Value	Percentiles of the non-parametric density function from 0% to 100%, one value for each percentile

RELIABILITY SUMMARIES

There are some summary database tables that make sense for storing highly aggregated reliability measures. In terms of temporally aggregated measures, these are tables that store, for a given route, reliability information for a single calendar month, a quarter, or a year. This aggregation is useful for users who want to examine route-specific reliability trends over time for selected days of the week or times of the day.

Exhibit A-18 illustrates an example database table that summarizes reliability information for each route by month. To create this table, travel time distributions are reviewed for each time of day (either 5-minute period or hourly) and day of the week, and reliability measures (e.g., the operative regime) are reviewed and tabulated. The same table can also be generated for quarterly and yearly time periods. These types of tables support queries letting users investigate other measures such as the buffer time for all weekdays over a month, or the planning time on Sundays at 5:00 PM over a year.

Reliability information can also be aggregated in the spatial dimension. An example would be the storage of quarterly reliability summaries for each region within the network. This allows high-level comparisons of performance between time periods and across regions. Exhibit A-19 shows an example of the storage of quarterly reliability information by region.

Exhibit A-18
Monthly Segment or Route-level Reliability Summary Table

Column #	Field	Description
1	Timestamp	The month and year.
2	Segment or Route ID	Unique identifier for the route.
3	Day of Week	The day of the week.
4	Time of Day	The time of day (can be either a 5-minute period or higher aggregations, such as an hour).
5	Regime	The regime to which the record pertains.
6	Percent Time	Percentage of time during which the regime specified in column #5 was operative during the time and segment or route to which this record corresponds.

Exhibit A-19
Quarterly Regional-level Reliability Summary Table

Column #	Field	Description
1	Timestamp	The quarter and year.
2	Region ID	Unique identifier for the region.
3	Weekday?	Whether the reliability summary is for weekdays or weekends.
4	Time of Day	The time of day (can be either a 5-minute period or higher aggregations, such as an hour).
5	Average Buffer Time Index	The median travel time over the quarter for that time of day and day of the week.
6	Average Planning Time Index	The 95th percentile travel time over the quarter for that time of day and day of the week.

SUMMARY

The final database design for a reliability monitoring system must reflect the technologies used to collect data and the processes used to derive travel time estimates from the raw data. In general, four types of tables are needed to fully describe the monitoring system and its outputs. Configuration tables are needed to define the routes and route segments for which travel times are to be computed, including their starting and ending point, where their detectors are located, and what type of detection they have. Raw data tables are needed to store the unaltered inputs from the various detection types. Travel time tables are needed to store the representative travel times calculated for each route segment, time period, and detection type, as well as the intermediary information generated during the computation. Finally, reliability tables can be used to store high-level monthly, quarterly, and yearly summaries of reliability statistics for individual routes or higher spatial aggregations.

APPENDIX A

HIGH LEVEL FUNCTIONAL REQUIREMENTS LIST

This appendix is intended to serve as a summary of the functional requirements described in the document. For agency staff, this list represents an overview of the types of capabilities that should be required for a travel time reliability monitoring system. These items are organized into the same categories used in the main document to describe the components of a travel time reliability monitoring system.

Data Collection

- A defined plan for collecting traffic data, including detector types and locations, detector spacings, and the frequency of data collection.
- Communication hardware in place to collect data and transmit it to a central hub.
- A defined plan for collecting incident information, including the data source, the types of data needed, how they will be obtained, and the frequency of data collection.
- A defined plan for collecting weather information, including the data source, the types of data needed, how they will be obtained, and the frequency of data collection.
- A defined plan for collecting work zone information, including the data source, the types of data needed, how they will be obtained, and the frequency of data collection.
- A defined plan for collecting special event information, including the data source, the types of data needed, how they will be obtained, and the frequency of data collection.
- A defined plan for collecting traffic control information, including the data source, the types of data needed, how they will be obtained, and the frequency of data collection.
- A defined plan for measuring/estimating demand and demand fluctuations.
- A defined plan for measuring capacity and determining if it is inadequate.
- A defined plan for collecting data on exogenous events (optional)
- A defined plan for collecting transit-specific data from AVL-equipped vehicles. (optional)
- Interagency agreements in place for sharing data in real-time. (optional)
- Agreements with private data distributors to obtain needed data in real-time. (optional)

Data Management

- Use of an industry standard database.
- A defined data warehouse model.
- A database architecture that supports the storage of both traffic and non-traffic data.
- The capability to store all raw sensor data where privacy issues do not disallow it.
- The intention to store data for every sensor at the lowest level of granularity possible, where it cannot be stored in raw form.
- The intention to store both raw data and imputed data in parallel and never replace raw data with imputed data.
- Clearly specified filtering methods and algorithms for removing bad data from malfunctioning infrastructure-based sensors.

- Clearly specified filtering methods and algorithms for removing unrepresentative travel time data from AVI sensors.
- Clearly specified algorithms for map-matching vehicle-based data to specific routes.
- Defined thresholds for the vehicle sampling rates required for valid data.
- Clearly specified methods and algorithms for imputing missing or damaged data.
- Clearly specified methods for tracking imputed data points.
- Clearly specified methods for tracking the imputation measure used on a data point.
- Clearly specified methods for storing metadata on what percentage of data points have been imputed and how they have been imputed to evaluate the statistical validity of reliability estimates.
- Clearly specified methods for imputing reliability measures in locations that lack detection technologies. (optional)

Computation Engine

- Defined methodologies for calculating speeds and travel times from infrastructure-based sensors.
- Calculations for deriving travel times from vehicle-based sensors and calculating the standard error of estimates.
- Methods for fusing travel times from different data sources that account for validity differences between different detection technologies and travel time estimation techniques.
- Performance metrics that report the statistical validity of reliability estimates to account for the uncertainties inherent in data aggregation and fusion.
- Clearly defined reliability metrics and corresponding equations.
- Defined spatial and temporal aggregations that the system will perform and the corresponding methodologies for performing them.
- Spatial aggregation capabilities to compute and store reliability measures for links, routes, subareas, and regions.
- Temporal aggregation capabilities to compute and store reliability measures for minute-to-minute, hourly, daily, weekly, monthly, quarterly, and yearly time periods.
- Defined algorithms for predicting travel times from historical and current data.
- Defined algorithms for linking travel time variability with each of the seven sources of unreliability.

Report Generation

- A fully defined list of reports that the system can deliver, including the reliability measures that they will convey and the format they will take.
- For each report, a description of the user it is intended to serve, the steps the user will need to take to create the report, and what the report will look like.
- A defined maximum amount of time that it should take to generate a report.
- User flexibility to choose the spatial and temporal aggregation levels in reports.

Systems Interactions

- A list of applicable ITS standards to which the system must adhere, with clearly specified methods for adhering to these standards. (optional)
- A clearly defined plan for adhering to the applicable regional architecture.
- A method of transferring data between regional systems using TMDD. (optional)
- A method of transferring reliability measures to other user interfaces in real-time using the Data Dictionary for ATIS. (optional)

APPENDIX B

SUMMARY OF APPLICABLE ITS STANDARDS AND GUIDES

Standard	Author Organization/ Location/Contact	Summary
National ITS Architecture	USDOT http://www.iteris.com/itsarch/index.htm	A flexible framework for planning, defining, and integrating ITS systems. Defines the logical (processes and information flows) and physical (transportation agencies and communications) architectures that govern ITS systems.
Real-time System Management Information Program	FHWA http://ops.fhwa.dot.gov/	Focuses on providing traveler information in real-time to decrease congestion; includes an *Information Sharing Specifications and Data Exchange Formats* guide that standardizes the communications interface for exchanging traffic data and event information.
Travel Time Reliability: Making it There On Time, All the Time	FHWA http://ops.fhwa.dot.gov/ Rich Taylor, FHWA Office of Operations Rich.taylor@fhwa.dot.gov (202) 366-1327	A general informational guide to measuring and distributing travel time reliability information.
Traffic Management Data Dictionary (TMDD) and Message Sets for External Traffic Management Center Communications	ITE/AASHTO http://www.ite.org/standards/TMDD/	Defines data elements for roadway links, incidents, traffic-disruptive events, traffic control, ramp metering, traffic modeling, video camera control traffic, parking management, weather forecasting, detectors, actuated signal controllers, vehicle probes, and CMSs. Defines message sets for communications between TMCs and other ITS centers.
TMDD Guide	ITE/AASHTO	Describes how to use the TMDD and provides context for the system engineering process.
Data Dictionary for Advanced Traveler Information Systems (ATIS)	SAE	Provides a set of core data elements needed by information service providers for ATIS. Data Dictionary provides the foundation for ATIS message sets for all stages of travel (pre-trip and en route), all types of travelers, all categories of info, and all platforms for delivery of information (in-vehicle, portable devices, kiosks, ets).

APPENDIX C

ADDITIONAL ITS STANDARDS

Standard	Author Organization/ Location	Summary[1]
Standard Specifications for Archiving ITS-Generated Traffic Monitoring Data	ASTM Available for purchase from ASTM website	Specifies a data dictionary for archiving traffic data; including conventional traffic monitoring data, data collected directly from ITS systems, and travel-time data from probe vehicles
Standard Practice for Metadata to Support Archived Data Management Systems	ASTM Available for purchase from ASTM website	Describes a hierarchical outline of sections and elements to be used in developing metadata to support archived data management systems.
NTCIP 1201: Global Object Definitions	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	Defines those pieces of data that are likely to be used in multiple device types such as actuated signal controllers and dynamic message signs. Examples of this data include time, report generation, scheduling concepts, etc.
NTCIP 1206: Object Definitions for Data Collection and Monitoring (DCM) devices	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	Specifies object definitions that may be supported by data collection and monitoring devices, such as roadway loop detectors.
NTCIP 1209: Data Element Definitions for Transportation Sensor Systems (TSS)	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	Object definitions that guide the data exchange content between advanced sensors and other devices in an NTCIP network; including video-based detection sensors, inductive loop detectors, sonic detectors, infrared detectors, and microwave/radar detectors
NTCIP 1204: Object Definitions for Environmental Sensor Stations	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	Defines objects specific to Environmental Sensor Stations
NTCIP 2202: Internet (TCP/IP and UDP/IP) Transport Profile	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	A set of transport and network layer protocols to provide connectionless and connection-oriented transport services.
NTCIP 8003: Profile Framework	ITE/AASHTO/NEMA http://www.ntcip.org/library/documents/	A framework and classification scheme for developing combinations and/or sets of protocols related to communication in an ITS environment.
Standard for	IEEE	Enables consistent standardized communications

1 http://www.standards.its.dot.gov/StdsSummary.asp?ID=372

Standard	Author Organization/ Location	Summary[1]
Traffic Incident Management Message Sets for Use by Emergency Management Centers	Available for purchase from IEEE website	among Incident Management centers, fleet and freight management centers, information service providers, emergency management centers, planning subsystems, traffic management centers and transit management centers
Standard for Common Incident Management Sets for Use by Emergency Management Centers	IEEE Available for purchase from IEEE website	Standards describing the form and content of the incident management messages sets for emergency management systems (EMS) to traffic management systems (TMS) and from emergency management systems to the emergency telephone system (ETS) or (E911).
Standard for Message Sets for Vehicle/Roadside Communications	IEEE Available for purchase from IEEE website	Standard messages for commercial vehicle, electronic toll, and traffic management applications.
ISP-Vehicle Location Referencing Standard	SAE Available for purchase from SAE website	For the communication of spatial data references between central sites and mobile vehicles on roads. References can be communicated from central sites to vehicles, or from vehicles to central sites. May be used where appropriate by other ITS applications requiring location references between data sets.

APPENDIX D

ITS STANDARDS GLOSSARY

Center: A connection point on a network (ranging in size from a single-laptop to a complex multi-computer environment) that is capable of exchanging messages with other centers.

Data Dictionary: An organized listing of dialogs, messages, data frames, data elements, and their properties that are required so that both the user and the system developer have a common understanding of input, output, components of storage, and intermediate calculations.

Data Element: A syntactically formal representation of some single unit or information of interest (such as a fact, proposition, observation, etc.) with a singular instance value at any point in time, about some entity of interest (e.g., a person, place, process, property, object, concept, association, state, event). A data element is considered indivisible in a certain context.

Data Type: A classification of the collection of letters, digits, and/or symbols used to encode values of a data element based upon the operations that can be performed on the data element.

Dialogue: A sequence of messages.

Event: Broadly defined to include any set of travel circumstances an agency may wish to report. This includes incidents, descriptions of road and traffic conditions, weather conditions, construction, and special events. Events can be current or forecasted.

Interchangeability: Reflects the capability to exchange devices of the same type on the same communications channel and have those devices interact with other devices of the same type using standards-based functions.

Interoperability: Allows system components from different vendors to communicate with each other to provide system functions and work together as a whole system.

Message: Groupings of data elements that include information about how the data elements are combined and used to convey information among ITS centers and systems. Messages are abstract descriptions using a message set template, not specific instances of transmissions

APPENDIX E

APPLICABLE ITS MARKET PACKAGES

Market Package	Description	Applicability to Travel Time Reliability
Network Surveillance	Includes traffic detectors and other surveillance equipment that transmit data back to the Traffic Management Subsystem through fixed-point to fixed-point communications	Freeway and arterial traffic data collection from infrastructure-based sensors, incident detection
Traffic Probe Surveillance	Supports wireless communications between the vehicle and the center or dedicated short range communications between passing vehicles and the roadside for traffic data collection	Traffic data collection from AVI and AVL technologies
Weather Information Processing and Distribution	Processes and distributes the environmental information collected from the Road Weather Data Collection market package	Correlating travel time reliability with weather conditions
Maintenance and Construction Activity Coordination	Supports the dissemination of maintenance and construction activity to centers	Correlating travel time reliability with lane closures and construction activities
ISP Based Trip Planning and Route Guidance	Offers the user trip planning and en-route guidance services	User output of travel time reliability monitoring process
Broadcast Traveler Information	Collects traffic conditions and other information and broadcasts the information to travelers using technologies such as FM subcarrier, satellite radio, cellular data broadcasts, and Internet web casts	User output of travel time reliability monitoring process
Interactive Traveler Information	Provides tailored information in response to a traveler request	User output of travel time reliability monitoring process
Dynamic Route Guidance	Offers advanced route planning and guidance that is responsive to current condition	User output of travel time reliability monitoring process
ITS Data Warehouse	Includes all the capabilities outlined in ITS Data Mart and adds the functionality and interface definitions that allow the collection of data from multiple agencies and data sources spanning across modal and jurisdictional boundaries	Archiving reliability data
ITS Virtual Data Warehouse	Provides the same broad access to multimodal, multidimensional data from varied data sources as in the ITS Data Warehouse Market Package, but provides this access using increased interoperability between physically distributed ITS archives that are each locally managed.	Archiving reliability data

APPENDIX F

APPLICABLE TMDD SECTIONS

<u>Volume 1 Concept of Operations:</u>

2.3.4 Need to Share Event Information
2.3.4.1. Need for An Index of Events
2.3.4.2 Need to Correlate an Event with Another Event
2.3.4.3 Need to Provide Free Form Event Descriptions
2.3.4.4 Need to Provide Free Form Event Names
2.3.4.6 Need for Current Event Information
2.3.4.7 Need for Planned Event Information
2.3.4.8 Need for Forecast Event Information
2.3.4.9 Need to Share the Log of a Current Event
2.3.4.10 Need to Reference a URL
2.3.4.11 Need to Filter Events
2.3.4.11.1 Need to Filter Event Recaps
2.3.4.11.2 Need to Filter Event Updates

2.3.5 Need to Provide Roadway Network Data
2.3.5.1 Need for Roadway Network Inventory
2.3.5.1.1 Need for Node Inventory
2.3.5.1.1 Need for Link Inventory
2.3.5.1.3 Need for Route Inventory
2.3.5.2 Need to Share Node, Link, and Route Status
2.3.5.2.1 Need to Share Node State
2.3.5.2.2 Need to Share Link State
2.3.4.2.3 Need to Share Route State
2.3.5.3 Need to Share Link Data
2.3.5.4 Need to Share Route Data

2.3.6 Need to Provide Control of Devices
2.3.6.1. Need to Share Detector Inventory
2.3.6.1.2 Need Updated Detector Inventory
2.3.6.1.3 Need to Share Detector Status
2.3.6.1.4 Need for Detector Metadata
2.3.6.1.5 Need for Detector Data Correlation
2.3.6.1.6 Need for Detector Data Sharing
2.3.6.1.7 Need for Detector History
2.3.6.5 Need to Share Environment Sensor Data
2.3.6.5.1 Need to Share ESS Inventory
2.3.6.5.2 Need to Share Updated ESS Inventory
2.3.6.5.3 Need to Share ESS Device Status
2.3.6.5.4 Need to Share ESS Environmental Observations
2.3.6.5.5 Need to Share ESS Environmental Observation Metadata

2.3.6.5.6 Need to Receive a Qualified ESS Report
2.3.6.5.7 Need to Share ESS Organizational Metadata
2.3.6.6 Need to Share Lane Closure Gate Control
2.3.6.6.1 Need to Share Gate Inventory
2.3.6.6.2 Need to Share Updated Gate Inventory
2.3.6.6.3 Need to Share Gate Status
2.3.6.6.7 Need to Share Gate Control Schedule
2.3.6.8 Need to Share Lane Control and Status
2.3.6.8.1 Need to Share Controllable Lanes Inventory
2.3.6.8.7 Need to Share Controllable Lanes Schedule
2.3.6.9 Need to Share Ramp Meter Status and Control
2.3.6.9.1 Need to Share Ramp Meter Inventory
2.3.6.9.2 Need to Share Updated Ramp Meter Inventory
2.3.6.9.3 Need to Share Ramp Meter Status
2.3.6.9.8 Need to Share Ramp Metering Schedule
2.3.6.9.9 Need to Share Ramp Metering Plans
2.3.6.10 Need to Share Traffic Signal Control and Status
2.3.6.10.1 Need to Share Signal System Inventory
2.3.6.10.2 Need to Share Updated Signal System Inventory
2.3.6.10.3 Need to Share Intersection Status
2.3.6.10.8 Need to Share Controller Timing Patterns
2.3.6.10.9 Need to Filter Controller Timing Patterns
2.3.6.10.10 Need to Share Controller Schedule
2.3.6.10.11 Need to Share Turning Movement and Intersection Data

2.3.7 Need to Share Data for Archiving
2.3.7.1 Need for Traffic Monitoring Data

Volume 1: Requirements

3.3.4 Events Information Sharing
3.3.4.3 Subscribe to Event Information
3.3.4.4 Contents of Event Information Request
3.3.4.6 Required Event Information Content
3.3.4.7 Optional Event Information Content
3.3.4.8 Action Logs
3.3.4.9 Event Index

3.3.5 Provide Roadway Network Data
3.3.5.1 Share Traffic Network Information
3.3.5.2 Share Node Information
3.3.5.3 Share Link Information
3.3.5.4 Share Route Information

3.3.6 Provide Device Inventory, Status and Control
3.3.6.1 Generic Devices

3.3.6.2 Traffic Detectors
3.3.6.6 Environment Sensors
3.3.6.7 Lane Closure Gates
3.3.6.9 Lane Control Signals
3.3.6.10 Ramp Meter
3.3.6.11 Traffic Signal Controllers

3.3.7 Share Archive Data
3.3.7.1 Share Traffic Monitoring Data for Data Archiving
3.3.7.2 Share Processing Documentation Metadata

Volume 2: Design Content

3.0 TMDD ISO 14817 ASN.1 and XML Data Concept Definitions
3.1 Dialogs
3.1.1 ArchivedData Class Dialogs
3.1.3 ConnectionManagement Class Dialogs
3.1.4 Detector Class Dialogs
3.1.5 Device Class Dialogs
3.1.7 ESS Class Dialogs
3.1.8 Event Class Dialogs
3.1.9 Gate Class Dialogs
3.1.11 IntersectionSignal Class Dialogs
3.1.12 Lane Control Status Dialogs
3.1.13 Link Class Dialogs
3.1.14 Node Class Dialogs
3.1.16 RampMeter Class Dialogs
3.1.17 Route Class Dialogs
3.1.18 Section Class Dialogs
3.1.19 TransportationNetwork Class Dialogs

3.2 Messages
3.2.1 ArchivedData Class Messages
3.2.3 ConnectionManagement Class Messages
3.2.4 Detector Class Messages
3.2.5 Device Class Messages
3.2.7 ESS Class Messages
3.2.8 Event Class Messages
3.2.9 Gate Class Messages
3.2.11 IntersectionSignal Class Messages
3.2.12 LCS Class Messages
3.2.13 Link Class Messages
3.2.14 Node Class Messages
3.2.16 RampMeter Class Messages
3.2.17 Route Class Messages
3.2.18 Section Class Messages

3.3 Data Frames

3.4 Data Elements

3.5 Object Classes

APPENDIX G

APPLICABLE SECTIONS OF THE DATA DICTIONARY FOR ADVANCED TRAVELER INFORMATION SYSTEMS

6.43 Information Request, linkTravelTime(28)

6.99: Estimate of travel time returned to the traveler based upon route.

6.100: Estimate of travel time between way points or from/to origin/destination and way point.

APPENDIX H

DETECTOR DIAGNOSTIC ALGORITHM

Diagnostic Algorithms:

$$S_1(i,d) = \sum_{a \le t \le b} (k_i(d,t)) = 0$$

$$S_2(i,d) = \sum_{a \le t \le b} (k_i(d,t) > 0) \cdot (q_i(d,t) = 0)$$

$$S_3(i,d) = \sum_{a \le t \le b} [(k_i(d,t)) > k^{*}], k^{*} = 0.35$$

$$S_4(i,d) = (-1) \sum_{x:p(x)>0} \hat{p}(x) \log(\hat{p}(x))$$

$$\hat{p}(x) = \frac{\sum_{a \le t \le b} 1(k_i(d,t) = x)}{\sum_{a \le t \le b} 1}$$

Where:
d =day index
t =30 second sample number
k =density
q =flow

$$S^{*}_{1} = 1200$$

$$S^{*}_{2} = 50$$

$$S^{*}_{3} = 200$$

$$S^{*}_{4} = 4$$

Malfunction if: $$\left\{ \begin{array}{l} S_1(i,d) > S^{*}_{1} \\ S_2(i,d) > S^{*}_{2} \\ S_3(i,d) > S^{*}_{3} \\ S_4(i,d) > S^{*}_{4} \end{array} \right\}$$

APPENDIX I

PEMS CALCULATIONS

PeMS Speed Calculation

PeMS uses a g-factor, which represents the effective length of a vehicle, to calculate speed from flow and occupancy detector outputs. The g-factor is a combination of the average length of the vehicles in the traffic stream and the tuning of the loop detector. Traditionally, a constant value for the g-factor is used, but this leads to inaccurate speed estimates because the g-factor varies by lane, time-of-day, and loop sensitivity.

PeMS estimates a g-factor for each loop every 5 minutes over an average week to provide accurate speed estimates. The algorithm implemented in PeMS is adapted from the paper "Statistical methods for estimating speed using single-loop detectors" by van Zwet, Chen, Jia and Kwon. The steps for estimation of speed from that paper are as follows:

1. Assume that speed on the freeway at free-flow conditions is known and constant.
    - Free-flow is defined by having occupancy less than a certain threshold
    - The free-flow speed is only a function of the type of freeway (meaning the total number of lanes) and the particular lane that the detector is in.
2. Using this assumption for each loop, work backwards and compute the g-factor for a number of points during a number of days.
3. Smooth this using a robust adaptive regression method to obtain a g-factor for each loop in the system over a typical week.
4. Use the g-factor to compute the initial estimate of speed for each loop in real-time.
5. Pass the initial estimate through an exponential filter with weights that vary as a function of flow. When the flow at the loop is low the smoothing is severe and when the flow is high, there is little smoothing. This allows the estimate to quickly adapt to periods of congestion as well as to have stable speeds when there is very little data (such as in the middle of the night).
6. The resulting speed is the speed estimate.

The free-flow speeds assumed in the g-factor calculation were taken from double-loop detector data in the Bay Area and are as follows:

Type	# of Lanes	Lane 1	Lane 2	Lane 3	Lane 4	Lane 5	Lane 6	Lane 7
HOV	1	65.0						
HOV	2	65.0	65.0					
ML	1	65.0						
ML	2	71.2	65.1					
ML	3	71.9	69.7	62.7				
ML	4	74.8	70.9	67.4	62.8			
ML	5	76.5	74.0	72.0	69.2	64.5		
ML	6	76.5	74.0	72.0	69.2	64.5	64.5	
ML	7	76.5	74.0	72.0	69.2	64.5	64.5	64.5

PeMS Linkage between Delay and Events

PeMS uses a Congestion Pie Algorithm to assign delay on a freeway to one of three categories:

1. **Collisions**: Based on the set of all accidents that take place on the freeway system.
2. **Bottlenecks**: Based on anything caught by the PeMS bottleneck identification algorithm. The cause of a bottleneck on any one day is not determined.
3. **Miscellaneous**: All of the delay that cannot be assigned to either of the two previous categories.

Delay is assigned a cause on a quarterly basis. The steps used to assign delay to its cause area as follows:

1. Compute Total Delay, D_{tot} :
   - Calculate with respect to 60mph for each county-freeway-direction in the quarter
2. Compute Delay Due to Collisions, D_{col}:
   - Extract the number of collisions per day from the incident data set provided by the agency.
   - Compute a straight line linear regression relating the delay on each day to the number of collisions. The intercept, alpha, is the average daily delay for a collision-free day.
   - Compute the delay due to collisions, $D_{col}=D_{tot}$-alpha (limited by zero)
3. Compute Delay Due to Bottlenecks, D_{bn}
   - Take the recurrent bottleneck locations that are active more than 20% of the days in the quarter.
   - For these bottlenecks, estimate the average daily delay due to the bottleneck, D_{bn}, from the results of the bottleneck identification algorithm
   - Limit D_{bn} such that $D_{bn} + D_{col} <= D_{tot}$
4. Compute the Miscellaneous Delay, D_{misc}
   - Any delay that cannot be assigned to either bottlenecks or collisions
5. Subdivide D_{bn} into potential delay savings, D_{pot}, and excess delay, D_{excess}
   - For each corridor, if there are any bottlenecks, compute the potential savings that result from running an ideal ramp metering algorithm at each bottleneck. An ideal ramp metering algorithm is a ramp metering strategy that restricts the rate at which vehicles enter the freeway such that the traffic at this location operates at capacity. Capacity is defined as the maximum observed 15-minute flow at each location.
   - The result of restricting the demand at the bottleneck is that the freeway operates at free-flow conditions and the delay on the freeway is reduced to zero.
   - The side effect of this is that the vehicles that have to wait at the ramps incur delay. This is excess delay, D_{excess}.

The limitations of this strategy are:

- No delay attributed to weather, lane closures, or special events.
- Incident data supplied by local agency can be incomplete or incorrect.

- There are sections of freeways that are not covered by detectors but where incidents are reported. This leads to a mismatch between the regime covered by fixed measurement devices and the regime covered by the incident data collection.
- The Ideal Ramp Metering Algorithm relies on many ideal assumptions which include:
    - The ramps at each location have enough storage capacity to hold all metered vehicles.
    - It is politically feasible to run this algorithm
    - Drivers do not take any detours or diversions

Thus, the potential delay savings result needs to be interpreted with care as an estimate of the maximum savings possible (rather than the realizable savings).

Travel Time Predictions

PeMS graphically displays a prediction of the travel time for a selected route from the time selected until the rest of the day. The travel time prediction is done by examining the collection of historical travel times for the route and choosing the days with the three closest travel time profiles. A weighted vector over the last few samples is used to measure distance between travel time profiles. The prediction is formed by taking the median of those three closest travel time profiles and plotting that for the rest of the day. The user plot shows the measured travel time up until that point and then the prediction for the rest of the day.

APPENDIX J

EXISTING AVI FILTERING ALGORITHMS

Transguide: San Antonio, TX

Summary: Transguide filters AVI travel time estimates by defining a set of valid recorded travel times during each 2 minute evaluation period based on those that are within 20% of the estimated travel times between the same two points for the previous 2 minute time period.

Algorithm (developed by the Southwest Research Institute):

$$Stt_{AB_i} = \left\{ t_{Bi} - t_{Ai} \mid t - t_w \le t_{Bi} \le \mathrm{t}, tt'_{AB_i}(1 - l_{th}) \le t_{Bi} - t_{Ai} \le tt'_{AB_i}(1 + l_{th}) \right\}$$

$$tt_{AB_i} = \frac{\sum_{i=1}^{|Stt_{AB_i}|} (t_{Bi} - t_{Ai})}{|Stt_{AB_i}|}$$

Where:

Stt_{ABi} is the set of valid recorded travel times that is used at each evaluation time to estimate the current average travel time between two AVI readers A and B;

t_{Ai} and t_{Bi} are the detection times of vehicle I at readers A and B;

t is the time at which estimation takes place;

$tt`_{ABi}$ is the previously estimated travel time from reader A to reader B;

t_w is the rolling average window, which determines the period of time that should be considered when estimating the current average travel time (Transguide uses 2 minutes);

l_{th} is the link threshold time, which is used to identify and remove outlier observations (Transguide uses 0.20); and

tt_{ABi} is the estimated average travel time for the time period.

TranStar: Houston, TX

Summary: TransStar filters AVI travel time estimates by defining a set of valid recorded travel times during each 30 second evaluation period based on those that are within 20% of the estimated travel times between the same two points for the previous 30 second time period.

Algorithm (developed by the Southwest Research Institute):
TranStar's algorithm is the same as that used by TransGuide, but travel times are updated each time new travel time information is obtained from a vehicle instead of being updated at fixed intervals. TranStar also uses a link threshold parameter of 0.2, but uses a shorter rolling average window of 30 seconds.

TRANSMIT: New York City Metropolitan Area

Summary: TRANSMIT filters AVI travel times by averaging link travel times over a 15-minute observation interval. An equation is then used to smooth the estimated travel time against historical data from the same 15-minute interval in the same day of the previous week to obtain an updated historical average travel time that is used as the current estimate.

Algorithm:

$$tt_{AB_k} = \frac{\sum_{i=1}^{n_k} (t_{Bi} - t_{Ai})}{n_k}$$

$$tth"_{AB_k} = a \cdot tth_{AB_k} + (1-a) \cdot tth"_{AB_{k-1}}$$

Where:

tt_{ABk} is the estimated current average travel time for the interval;

n_k is the number of link travel times collected for each interval k, up to a maximum of 200 observations;

tth_{ABk} is the historical smoothed travel time for the k^{th} sampling interval;

tth''_{ABk} and tth''_{ABk-1} are updated historical smoothed travel times for the current (k) and previous ($k-1$) sampling intervals; and

α is the smoothing factor (set at 0% when an incident is detected and 10% otherwise).

Dion and Rakha (2006)

"Estimating Dynamic Roadway Travel Times using Automatic Vehicle Identification Data for Low Sampling Rates". Francois Dion and Hesham Rakha, 2006.

Summary: This paper presents an AVI filtering algorithm that is designed to handle stable and unstable traffic conditions, provide accurate travel time estimates in areas where there is low market penetration of AVI sensors, and work for both freeway and signalized arterial roadways. It does this by applying the following filters:

1. Expected average trip time and trip time variability in future time interval
2. Number of consecutive intervals without any readings since the last recorded trip time
3. Number of consecutive data points either below or above the validity range
4. Variability in travel times within an analysis interval

Algorithm:

The first part of the filtering algorithm calculates the expected smoothed average travel time and smoothed travel time variance between a pair or readers A and B for a given sampling interval through a smoothing low-pass filter.

The algorithm then utilizes a robust data-filtering process that identifies valid data within a dynamically varying validity window. The size of the validity window varies as a function of the number of observations within the current sampling interval, the number of observations in the previous intervals, and the number of consecutive observations outside the validity window.