

13 Planning of Traffic and Movement

PORTAL: An On-Line Regional Transportation Data Archive with Transportation System Management Applications

Steven Hansen
Andrew Byrd
Andy Delcambre
Andy Rodriguez
Spicer Matthews
Robert L. Bertini

Department of Civil & Environmental Engineering
Nohad A. Toulon School of Urban Studies and Planning
Portland State University
P.O. Box 751
Portland, Oregon 97207-0751
Phone: 503-725-4249
Fax: 503-725-5950
Email: bertini@pdx.edu

2005 CITE Quad/WCTA Regional Conference
Vancouver, BC, Canada
April 7-9, 2005

PORTAL: An On-Line Regional Transportation Data Archive with Transportation System Management Applications

Steven Hansen
Andrew Byrd
Andy Delcambre
Andy Rodriguez
Spicer Matthews
Robert L. Bertini

Department of Civil & Environmental Engineering
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
P.O. Box 751
Portland, Oregon 97207-0751
Phone: 503-725-4249
Fax: 503-725-5950
Email: bertini@pdx.edu

ABSTRACT

The Portland Oregon Regional Transportation Archive Listing (PORTAL) is an online transportation data archive that has been developed and is being expanded in cooperation with the Oregon Department of Transportation (ODOT) and other regional transportation agency partners. Since July 2004, PORTAL has been archiving a live data stream consisting of 20-second data from the 436 inductive loop detectors comprising the Portland area's Advanced Traffic Management System (ATMS). PORTAL, developed and maintained at Portland State University and run by students and faculty, is designated as the Portland region's official transportation data archive entity, as part of the region's ITS Architecture. The objective of this paper is to describe the development of PORTAL, to explain the steps taken to implement the Portland region's first functional ITS data archive, and to describe our plans for future development and expansion. The paper includes a discussion of PORTAL's structure, the techniques used for data storage, the data processing capabilities and the web-based user interface. We also present detailed descriptions of automated reports and performance measures that can be generated by users with access to the archive. In the future, the data archive will expand to include a statewide perspective, regional transit data, freeway incident data, city traffic signal data, and truck weigh-in-motion data.

INTRODUCTION

In the context of intelligent transportation systems (ITS), data archiving refers to the systematic retention and re-use of data generated for operational purposes. In most cases, the original and primary purpose of generating these data is for real-time management functions of transportation systems. With the right insights and careful planning, archiving these otherwise-discarded data can offer a rich source of information for various entities and agencies with a heightening need to evaluate system performance and characteristics on a continuing basis (1). Once retained in their most raw form, the vast amount of ITS-generated data can be used in transportation management, planning, administration and research by key stakeholders including metropolitan planning organizations (MPOs), transportation planners, traffic management operators, transit operators and transportation researchers (2).

Archived data may be used to augment or as a substitute for conventional data sources that are both time consuming and costly to implement, and the high level of spatial and temporal detail found in ITS-generated data lends itself to new and useful methods of analysis (2). Further, the cost effectiveness

of existing data collection infrastructure can be improved and sampling bias can be reduced due to the continuous data stream. In addition, a better understanding of variability in system performance and response can be achieved (2). For example, transportation agencies are required to spend large amounts for manual traffic count programs that collect data at small number of sites on only a few days per year. In contrast, often ITS data are available at multiple locations on a 24-hour basis, 365 days a year. This can improve the calculation of seasonality factors. The biggest challenge is to document and retain the ITS data so that future users can understand its limitations and capabilities. Use of archived data can also assist in maintenance planning for ITS components. Creating the meta-data that define the ITS data is another big challenge.

To encourage the retention and use of ITS-generated data, the United States Department of Transportation (USDOT) revised the National ITS Architecture in 1999 to include the Archived Data User Service (ADUS). ADUS is designed to promote “the unambiguous interchange and reuse of data and information throughout all functional areas.” ADUS requires that data from ITS systems be collected and archived for historical, secondary and non real-time uses, and that these data be made readily available to users (3).

In cooperation with the Oregon Department of Transportation (ODOT) and other regional partners, the Portland Regional Transportation Archive Listing (PORTAL, located at portal.its.pdx.edu) was inaugurated via a direct fiber optic connection between ODOT and Portland State University (PSU). In July 2004, the data archive “went live,” receiving 20-second data from the 436 inductive loop detectors comprising the Portland area’s Advanced Traffic Management System (ATMS). This paper describes the steps taken to successfully implement the first functional ITS data archive in the Portland region and future expansion plans. We include a discussion of the archive structure, data storage, and data processing. Next, we describe the user interface, followed by a description of data retrieval features currently available. The paper concludes by discussing additional features, including other data sources that are currently under development.

DATA PROCESSING AND ARCHIVING

ODOT’s Region 1 transportation management operations center (TMOC) gathers data from 436 inductive loop detectors which comprise the Portland region’s ATMS. These detectors have been implemented as part of the comprehensive ramp metering system, so dual mainline loops (speed traps) are located just upstream of on-ramp locations, and the on-ramps themselves are also instrumented. Recently, additional detection has been installed at several locations where there are long segments without ramps. These data (count, occupancy and speed) are used in the day to day management of the transportation system to identify congested areas and incidents and to dispatch incident response or emergency vehicles to the appropriate locations. Drivers can be alerted to problems and rerouted using variable message signs and television and Internet-based traveler information systems (www.tripcheck.com). As ODOT upgrades their current pre-timed ramp metering system to an adaptive/dynamic system, the quality of these data is increasingly important for real-time ramp operations.

At 20 second intervals, each loop detector records vehicle count, the average speed of these vehicles, and occupancy, or percentage of the sample period when a vehicle was over the detector. Prior to implementation of the Portland data archive, the raw 20-second data were discarded. Now, permitted users can access and query the archived data through a web interface. The overall structure of the Portland ADUS is shown in Figure 1. The data are sent from ODOT TMOC to PSU in XML format. A script on the PSU server parses the XML data as it is sent every 20 seconds. The data are stored and accessed using a PostgreSQL relational database management system (RDBMS) running on a Linux platform.

The RDBMS stores data physically on a redundant array of independent disks (RAID) providing both high-speed access and increased reliability through redundancy in the event of hardware failure. The web-based user interface is provided by the Apache web server, with database connectivity and additional

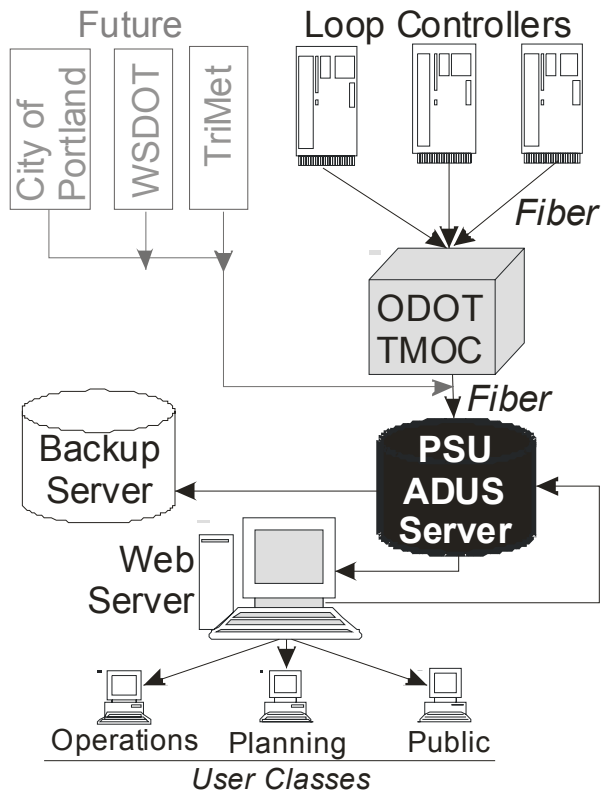


Figure 1: Portland Metropolitan Area ADUS Data Flows

redundancy, and the ability to rebuild missing data upon device failures. Finally, hardware maintenance and security updates are provided for all computer systems by experienced systems administration personnel.

National ITS Architecture

The development of the Portland ADUS is adhering to the recommendations of the National ITS Architecture. The archive fulfills five distinct processing functions, in accordance the National ITS Architecture ADUS addendum to the ITS Program Plan (3):

1. Store data in the same format as received from ITS subsystems
2. Accommodate levels of aggregation and reduction of the data flows, depending on the type of data represented
3. Sample raw data flows for permanent storage in accordance with user specifications. Permanent storage of the sampled data should be either online, offline, or both.
4. Apply quality control procedures to the data, including the flagging of suspect data and the editing of data.
5. Distinguish between the following data types: unprocessed (raw), edited, aggregated, and transformed.

processing by the PHP hypertext preprocessor module. Detector locations within the database have been geo-coded for interoperability with geographic information systems (GIS).

This archive implements a data warehousing strategy by retaining large amounts of raw operational data for analysis and decision making processes. These data are stored independently of their operational sources, allowing the execution of time-consuming queries with no impact on critical operations uses.

Data security is a key component of the data archive. In this spirit, the working copy of the database maintained on the primary server is replicated in a compressed format at a remote site. These daily backups ensure that if one copy of the database is lost, the archive service can be rapidly returned to operation with no significant loss of data. Both the primary database server and the backup storage are located in climate controlled machine rooms with uninterruptible power supplies (UPS) and generator backup power, preventing data loss or gaps in data availability due to power outages. The working copy of the database is stored on a RAID device, providing error detection,

Raw Data Storage

As the 20-second loop detector data is transferred from the ODOT TMOC to the PSU relational database, it is archived with no loss of resolution. Volume, speed, and occupancy data are archived indefinitely at 20 second intervals. Prior to this, ODOT had discarded the 20 second data.

Data Aggregation

At this time, the archive accommodates levels of aggregation appropriate to the type of data being recorded. Each night at 3 a.m., the 20 second data are aggregated to 5 minute data. These data are appended to a table for the relevant month. The data are also aggregated to 1 hour and appended to a table which contains data for an entire year. Future plans call for other user-defined levels of aggregation as the need is demonstrated.

In addition to total volume, average occupancy and average speed, the 5 minute and 1 hour tables also include calculated values of vehicle miles traveled (VMT), vehicle hours traveled (VHT), delay, and travel time. These calculations depend on the highway segment length associated with each loop detector. These segments, or loop detector influence areas, are defined as the segment of the freeway extending halfway to the next upstream station and the segment extending halfway to the next downstream station. Where no next station existed, the detectors influence extended to the end of the highway. PORTAL is currently using 60 mph to test the delay calculation. In the future we envision that the user could specify a desired value of target speed so that results could be compared.

Sampling Data

Users have the ability to sample raw data flows and store data permanently in accordance to their specifications, including aggregation of the data at desired resolutions. These data can either be output in tabular form on the screen, plotted on a relevant graph, or downloaded in comma separated value (CSV) format for permanent offline storage and further manipulation.

Quality Control

As with all ITS data, we experience data quality issues. When dealing with 180 megabytes of data per day, this is to be expected. Fortunately, PSU has been the first organization to carefully examine and validate the ODOT 20-second loop detector data. As a result, ODOT has been implementing improvements to their system to remove several systematic errors that were discovered.

Quality control procedures are applied to identify and mark suspect or erroneous data. For example, as one quality control measure, we have implemented an algorithm similar to the Daily Statistics Algorithm (DSA) used in the California PeMS database (4). DSA flags four error types:

- Type 1: Occupancy and flow are mostly zero
- Type 2: Non-zero occupancy and zero flow
- Type 3: Very high occupancy
- Type 4: Constant occupancy and flow

PORTAL uses thresholds to flag each error type. For example, a loop is given a type 1 error if more than a particular number of observations have occupancy and flow values of zero. An error type 2 is triggered if more than a small number of show a non-zero occupancy with zero count. A type 3 error is triggered if more than a certain percentage of observations have occupancies greater than 0.35. The fourth diagnostic measures randomness of occupancy and flow. If a loop records constant occupancy values, this diagnostic will flag a type 4 error.

Distinguish Between Data Types

The final data processing requirement outlined in the National ITS Architecture ADUS addendum is that distinctions be maintained between raw, edited, aggregated and transformed data. The ADUS user can query the data in three distinct areas of the web interface – raw data, time series, and grouped data. These areas are discussed further in the next section. In each case the user knows whether the query being performed is using raw or aggregated data.

RETRIEVING THE DATA

Figure 2 shows the current PORTAL user interface. The map shows an example of how ODOT currently uses loop detector data in real time. The highways are color coded to indicate the current traffic speed on each road segment. The map, maintained by ODOT, is updated every 20 seconds when the loop detectors record average speed for the time period. This map is also available on ODOT's website www.tripcheck.com along with detailed information on road conditions.

The welcome screen gives the user the following options for querying archived data, as indicated by the menu tabs near the top of the screen:

- Time Series
- Grouped Data
- Data Fidelity
- Raw Data
- Weather

Time Series

The time series tab allows the user to perform the queries on the 5 minute aggregated data to extract volume, speed, occupancy, vehicle miles traveled, vehicle hours traveled, travel time, and delay. The user can specify a specific loop detector station as well as a full highway segment. Additionally, the user can choose the time period and specific travel lanes. The user has the option of either plotting the data,

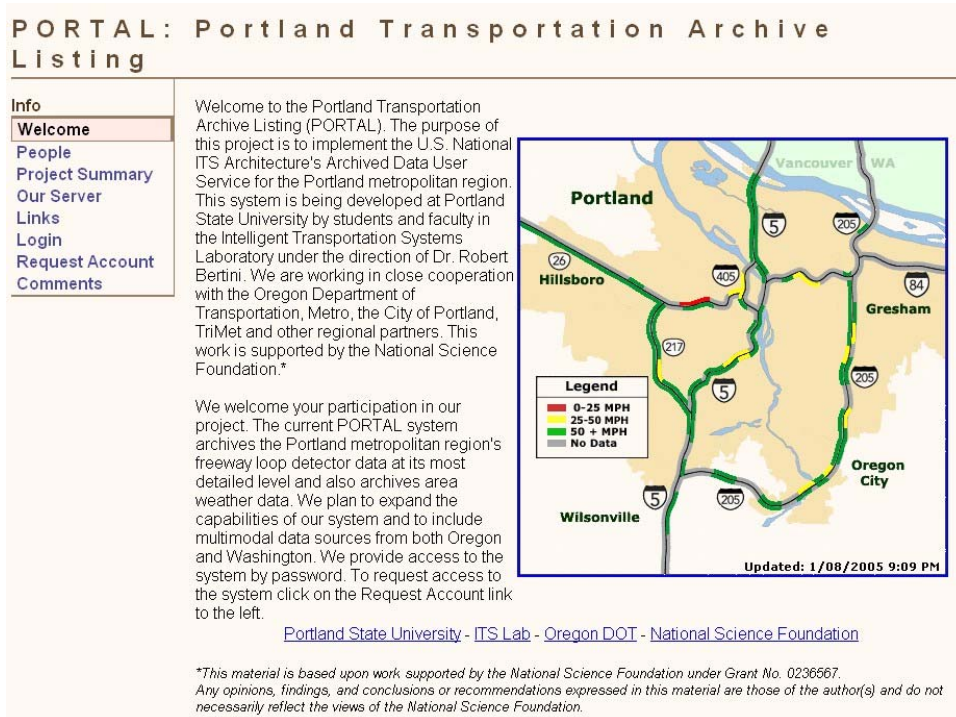


Figure 2: PORTAL Homepage

viewing the numbers in a table on the screen, or exporting the data as a comma separated value file. Figure 3 shows a time series plot of volume for one station along I-5 North for October 28, 2004. Figure 4 shows a speed contour plot for I-5 North on October 28, 2004. For convenience, the contour plot also includes a schematic of the freeway on- and off-ramp configuration through the corridor.

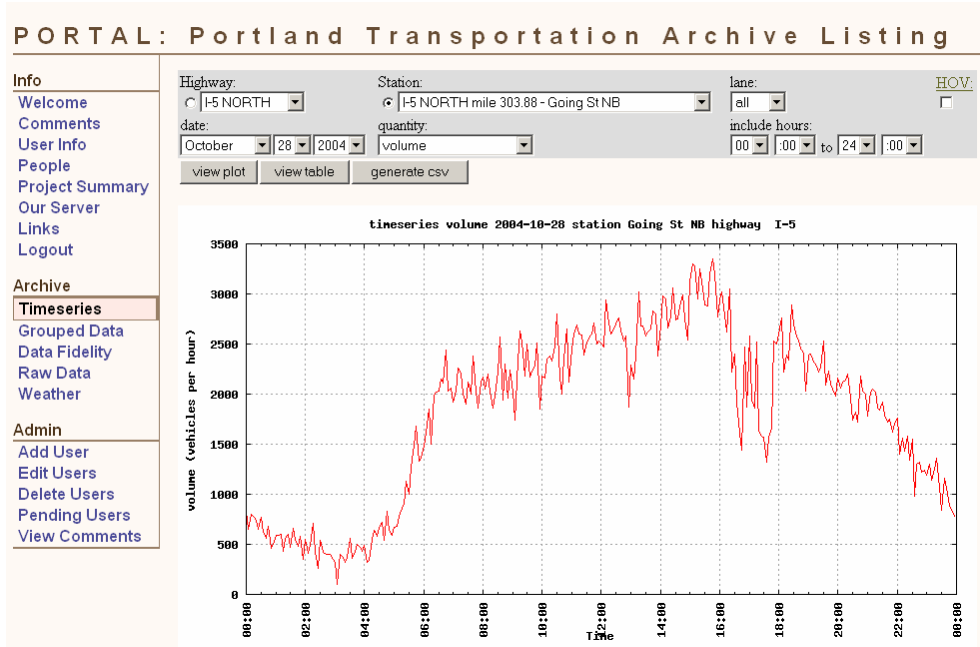


Figure 3: Sample Volume Plot for I-5 North Going Street Loop Detector Station, October 28, 2004



Figure 4: Sample Speed Contour Plot for I-5 North, October 28, 2004

Grouped Data

This section allows the user to display summaries of data from multiple days. The data can be grouped by hour of day, day of week, or week of year. The user can also determine the statistics they wish to see, including, mean, minimum value, maximum value, and standard deviation. As in the time series section, the user has the choice of plotting the data, displaying a table, or downloading a CSV data file. As an example, Figure 5 shows a plot of the mean plus or minus one standard deviation of speeds measured at one location (Portland Blvd.) on I-5 North during September and October 2004.

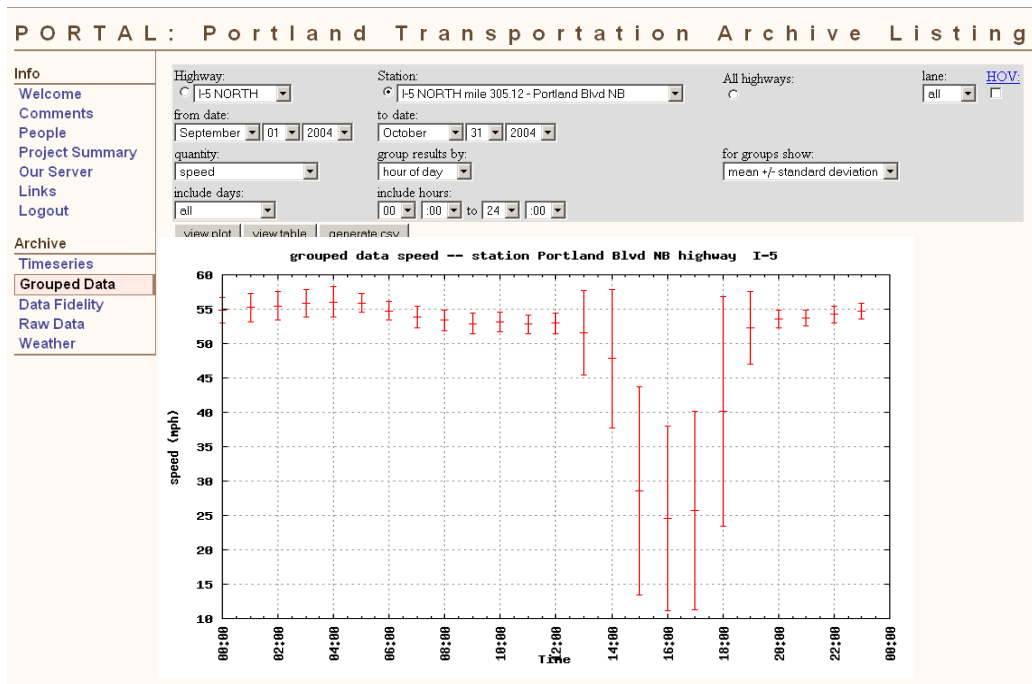


Figure 5: Grouped Data Plot of Mean +/- One Standard Deviation of Speed

Data Fidelity

This section allows the user to determine whether the loops queried were determined to be good according to the data quality algorithms described above. The user can either get information on the health of a specific detector, or, can get the percentage of healthy loops on a specified highway section. Figure 6 shows an example of the detector health information available. This has been useful for assisting ODOT with the upgrade of some of their traffic management system software and in updating some of their ramp meter controller hardware. We plan to continue to refine this component by soliciting input from TMOC staff.

Raw Data

The user currently has the option of downloading raw 20 second data in CSV format for specific detector stations or highway segments.

Weather

When conducting traffic studies in the future, it will be important to know what weather conditions were like on the days analyzed. For example, previous studies in Oregon have found substantial differences in the numbers of crashes on rainy days as compared with dry days (5). In order to provide at least a minimal amount of weather data to future analysts, PORTAL collects and stores METAR weather

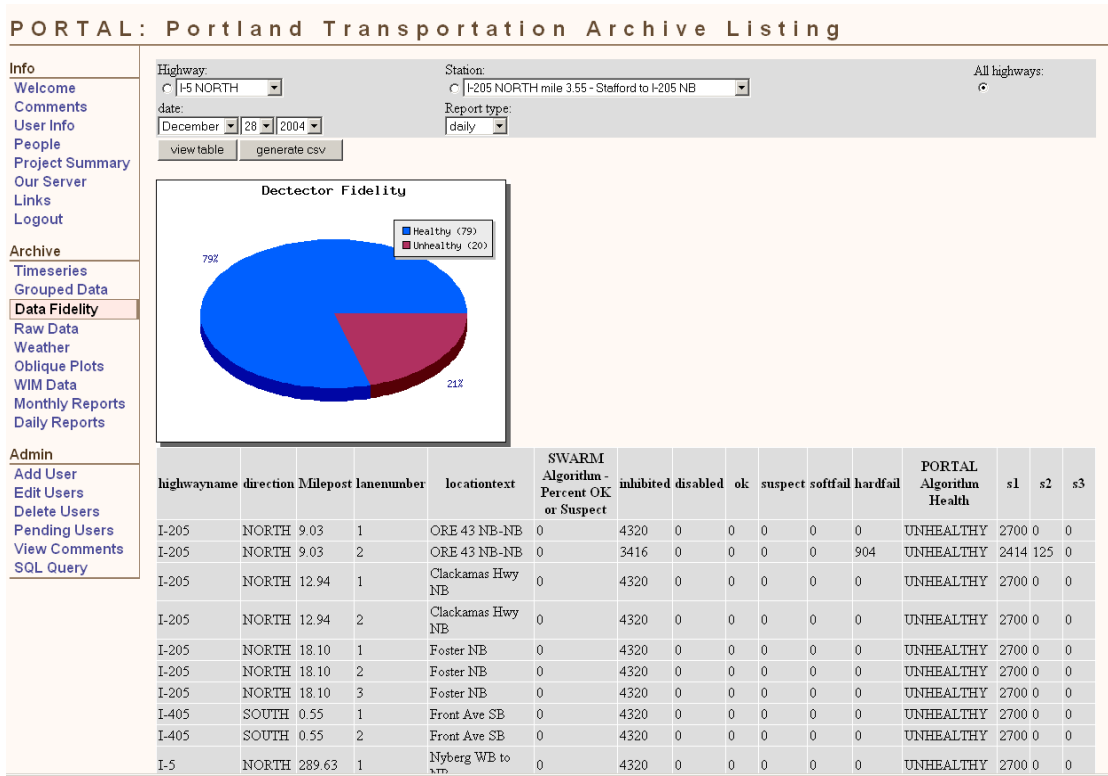


Figure 6: Data Fidelity Output

observations from National Oceanic and Atmospheric Administration (NOAA). Weather observations are collected from three points in the Portland region: Portland International Airport as well as at the airports located in suburban of Hillsboro and Troutdale, located west and east, respectively, of the city of Portland. PORTAL contains a script that collects and processes the weather data, providing easy access to NOAA’s weather reports (6). Users can query the database to get the weather conditions for a specified time period. There is a table in the database which indicates which weather station is geographically closest to each loop detector. Users have the option of extracting data from all three weather stations, or the station that is nearest to a particular loop detector. Figure 7 shows an example of a weather pop-up that is designed to display hourly temperature (left hand axis) and hourly rainfall quantities (right hand axis). Road Weather Information Systems (RWIS) data are also collected from several locations in the region. In the future these data will also be archived in PORTAL.

Monthly Congestion Reports

The FHWA has embarked on a program to develop monthly Urban Congestion Reports for approximately ten cities across the U.S. Figure 8 shows a sample presentation of these data. In anticipation of the desire to present monthly congestion data for the Portland region, PORTAL can automatically generate the congestion data contained in Figure 8 in a tabular form for easy incorporation. This is just one example how PORTAL can be customized to suit a variety of reporting schemes.

FUTURE EXPANSIONS

Research across industries indicates that effective data archiving systems start as small prototypes with a single data source. Studies consistently indicate that detector data is the most useful archived data (1), and for this reason we have chosen to first archive these data. The Portland region has a wealth of

additional data sources that will be archived at the appropriate time. Two data sources, the computer aided dispatch data and TriMet bus data, are discussed below.

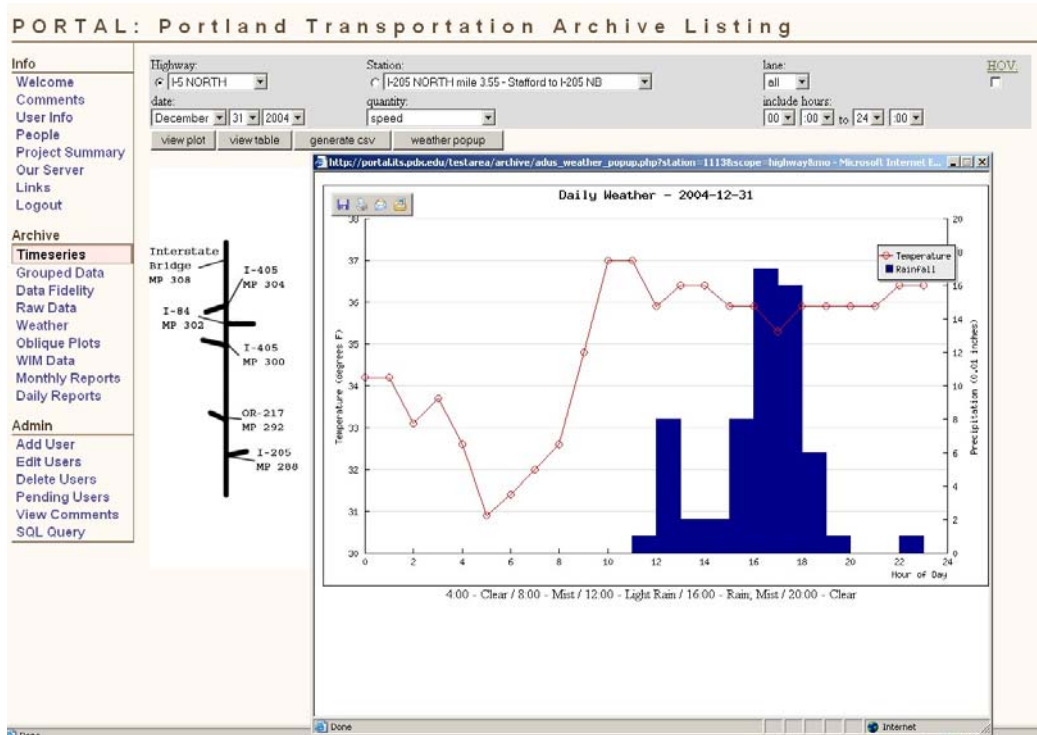


Figure 7: Weather Pop-up Example

	UCR NATIONAL COMPOSITE INDICATORS								
	Month vs. Last Month			Month vs. Last Year*		Quarter vs. Last Year*			
	SEP	OCT	Change	2003	2004	Change	2003	2004	Change
% Cong. Travel	49.4%	47.1%	↓ 2.3%	38.7%	44.5%	↑ 5.8%	36.5%	44.0%	↑ 7.5%
Travel Time Index	1.658	1.649	↓ 0.009	1.619	1.636	↑ 0.017	1.612	1.638	↑ 0.026
Buffer Index	2.132	2.156	↑ 0.024	2.056	2.148	↑ 0.092	2.064	2.136	↑ 0.072

* excludes Los Angeles

URBAN CONGESTION REPORT October 2004



EXECUTIVE SUMMARY

CITY	PERCENT CONGESTED TRAVEL			TRAVEL TIME INDEX			BUFFER INDEX			Max Usable Days: 21	CONTRIBUTING FACTORS						
	New 12-mo. LOs	New 12-mo. HIs	0	New 12-mo. LOs	New 12-mo. HIs	2	New 12-mo. LOs	New 12-mo. HIs	3		Monthly Precip. vs. Norm	Rain > 0.5"	Rain > 1.0"	Frozen Precip.	Fog	Work Zone Count	Monthly Incidents Hours vs. Norm
	0%	50%	100%	1.45	1.70	1.95	1.65	2.30	2.85		(NUMBER OF DAYS)	(NUMBER OF DAYS)	(NUMBER OF DAYS)	(NUMBER OF DAYS)	(NUMBER OF DAYS)	(NUMBER OF DAYS)	(NUMBER OF DAYS)
Atlanta	24	44		1.56	1.70	1.85	1.95	2.20	2.52	16	70%	0	0	0	0	N/A	48%
Boston	39	46		1.64	1.65		2.04	2.32		20	50%	1	0	0	2	N/A	78%
Chicago	52		84	1.60	1.71		2.11	2.29		17	105%	2	0	0	1	N/A	91%
Houston	26	41		1.61	1.70		2.04	2.21		18	52%	1	0	0	3	N/A	63%
Los Angeles	55		65	1.56	1.80		1.87	2.21		20	1232%	4	1	0	3	N/A	105%
Louisville	12	21		1.51	1.63		1.84	2.10		17	313%	1	1	0	0	N/A	168%
Miami	24	41		1.51	1.63		1.93	2.14		20	89%	3	1	0	0	N/A	58%
Philadelphia	14	34		1.52	1.64		1.87	2.16		19	81%	1	1	0	0	N/A	23%
Pittsburgh	9	27		1.50	1.59		1.83	2.04		18	149%	1	0	0	3	N/A	38%
San Antonio	36	53		1.21	1.31		1.94	2.37		21	245%	0	0	0	0	N/A	108%
Seattle	17	36								20	89%	0	0	0	6	N/A	80%
Portland										21							

LEGEND
 12-mo. LO: 0% (white), 50% (light gray), 100% (dark gray)
 12-mo. HI: current month (red), New 12-mo. HI (green), New 12-mo. LO (blue)

Figure 8: Monthly Congestion Report Example

Computer Aided Dispatch Data

Portland's ATMS includes a comprehensive incident management system (5), which in turn generates a large computer aided dispatch (CAD) database which then contains information about all recorded incidents on Portland's freeway system. This information includes the type of incident, which lanes were blocked as a result of the incident, and the start and end time of the incident. The TMOC can record the X-Y coordinates of an incident, but it is only required if the incident severe enough to cause significant delay. In 2001, coordinates were recorded for 10% of all incidents (5). We plan to archive the CAD data so that it can be further used to assess the performance of the incident management program and identify times and locations when incidents affected traffic flow.

TriMet Bus Data

TriMet is the transit provider for the Portland, Oregon metropolitan region. As part of their overall operations management system, TriMet has developed a "smart" bus that enables their automated bus dispatch system (BDS). The BDS includes Automatic Vehicle Location (AVL) using a satellite-based global positioning system (GPS) on all buses as well as Automatic Passenger Counters (APCs) installed on a majority of the existing fleet and all new bus acquisitions. As part of the real-time component of the BDS, each vehicle transmits its location to the transit management center (along with several other data points) periodically (roughly every 1-3 minutes). These real-time data are currently archived by TriMet.

Each vehicle in the fleet includes an on-board memory card that is pre-programmed with that bus' schedule information. Therefore the bus "knows" where it is located at all times, and links that with the schedule, enabling a schedule status (late/early) report to the operator. Further, as part of the archived data component of the BDS, a data record is written to the memory card at each bus stop. These records contain the route number, direction, trip number, date, vehicle number, operator ID, bus stop ID, stop arrival time, stop departure time, boardings, alightings, and passenger load. Also recorded is whether the doors of the bus were opened, whether the lift was used, the dwell time, maximum speed between the previous and current stop, and GPS coordinates. The memory card data are downloaded at the end of each day and compiled into a huge database. This information is used by TriMet to measure system performance on an ongoing basis and to aid in the bus route scheduling effort that occurs each quarter.

One of the major benefits of TriMet's real time and archived BDS data is that it can be used to extract performance characteristics from the region's arterial system. We plan to begin testing the use of both data sources as part of the Portland archive in the near future for this purpose.

CONCLUSIONS

In cooperation with the Oregon Department of Transportation (ODOT), TriMet and the City of Portland, Portland State University (PSU) is a major user of regional ITS data for transportation research. Other recent uses of ITS data by ADUS stakeholders such as Metro, the Portland regional planning organization, have been encouraged and facilitated by PSU researchers already familiar with manipulating this type of data.

As a result, PSU is in an ideal position to host transportation data, and its Intelligent Transportation Systems Laboratory (www.its.pdx.edu) has been designated as the regional archiving site for ITS data from Portland and adjacent areas of southwestern Washington State. The design and implementation of this archive are being carried out in accordance with the functional requirements for ADUS set forth in the National ITS Architecture, while also taking into account experiences and suggestions from other systems that are being developed in larger, more populous states and regions (1,2,3,13).

This paper has sought to document the experience thus far in implementing PORTAL (see portal.its.pdx.edu), so that other regions around the country can benefit from our experience. By openly sharing algorithms and experiences, in the spirit of the PeMS system, all regions can benefit from past experiences. In Portland, we are emphasizing the development of a multimodal resource that is benefiting

all transportation agencies via a unique collaborative environment established by the TransPort committee.

ACKNOWLEDGEMENTS

The authors thank ODOT, TriMet and the City of Portland for their ongoing support and guidance in the development of the Portland ADUS. In particular Dennis Mitchell, Galen McGill and Jack Marchant of ODOT; Ken Turner and Steve Callas of TriMet; and Bill Kloos, Rich Johnson and Willie Rotich of the City of Portland have helped making this project possible. Dick Walker, Kyle Hauger and Dr. Gerry Uba of Metro have also encouraged these efforts. This work is supported by the National Science Foundation under Grant No. 0236567, the Portland State University Department of Civil and Environmental Engineering and the Oregon Engineering and Technology Industry Council (ETIC). The Portland metropolitan area's TransPort ITS committee has also encouraged development of this project.

REFERENCES

1. Turner, S. *Guidelines for Developing ITS Data Archiving Systems*. Report 2127-3. FHWA, U.S. Department of Transportation, Texas Department of Transportation, and Texas Transportation Institute, 2001.
2. ADUS Program. *ITS Data Archiving: Five-Year Program Description*. USDOT, March 2000.
3. ADUS Program. *Archived Data User Service (ADUS): An Addendum to the ITS Program Plan*. USDOT, September 1998.
4. Chen, C., Kwon, J., Rice, J., Skabardonis, A. and Varaiya, P. (2003) "Detecting errors and imputing missing data for single loop surveillance systems," *Transportation Research Record*, No. 1855 (2003), p. 160-167.
5. Bertini, R.L., Rose, M. and El-Geneidy, A. *Using Archived Data to Measure Operational Benefits of ITS Investments: Region I Incident Response Program*. Portland State University, Center for Transportation Studies, Research Report, June 2004.
6. <http://www.schwarzvogel.de/software-pymetar.shtml>, accessed July 30, 2004.
7. Bureau of Public Roads. *Traffic Assignment Manual*. U.S. Dept. of Commerce, Urban Planning Division, Washington D.C., 1964.
8. Federal Highway Administration, *Traffic Monitoring Guide*. <http://www.fhwa.dot.gov/ohim/tmguide/index.htm>, accessed 11/13/2004.
9. Email Correspondence with Kyle Hauger of Metro Travel Forecasting, 10/29/2004.
10. *Highway Capacity Manual 2000*. Transportation Research Board, Special Report 209, Washington, D.C., 2000.
11. *State Highway Inventory Report*, Oregon Department of Transportation, 2004.
12. Cassidy, M.J. and R.L. Bertini. Some Traffic Features at Freeway Bottlenecks. *Transportation Research Part B*, Vol. 33, 1999, pp. 25-42.
13. PeMS System, California Department of Transportation and University of California at Berkeley. <http://pems.eecs.berkeley.edu/Public/>. Accessed July 30, 2004.