Transit Performance Measurement and Arterial Travel Time Estimation Using Archived AVL Data

Mathew Berkow, John Chee, Robert L. Bertini, and Christopher Monsere

Abstract—**Portland's local transit provider (TriMet) began using an automated bus dispatch system to manage and collect data about the performance of its fleet of buses in the late 1990s. These data provide TriMet with an abundance of useful information that it has used to successfully improve the performance and efficiency of its transit system. Each day, TriMet buses travel Portland's city and suburban streets on more than 90 different bus routes, collecting data at each scheduled and unscheduled stop. This rich source of data also has the potential to aid traffic engineers in evaluating arterial performance using the bus fleet as probes. This paper explains the type of data being collected by TriMet buses, gives examples of how it is used to evaluate and improve transit service, and demonstrates how these data might be used in the future, perhaps in conjunction with other technologies such as central signal system data, to aid traffic engineers in monitoring and improving arterial performance.**

I. Introduction

It is generally held in the field of transportation that we cannot build our way out of congestion. As a result, much attention has been focused on improving the operation of the system including the expanded use of Intelligent Transportation Systems (ITS) technologies. In the area of public transportation, many agencies have implemented systems that collect a wealth of operating information that can be used to evaluate and improve the operating efficiency of the transit fleet.

One such agency that makes extensive use of these technologies is the transit operator in the Portland, Oregon metropolitan region, TriMet. TriMet has a long history and experience with advanced public transportation operations. The Bus Dispatch System (BDS) includes automatic vehicle location technology, automated passenger counters, traffic signal and radio communication capabilities, and computer aided dispatching. At the start of each vehicle day, a data storage card programmed with schedule information for the assigned route is inserted into the vehicle control head (VCH) which is the operator's interface with the system. Since the bus "knows" its location at each point in time through its global positioning system (GPS), the onboard computer manages the front route display, announces stops, requests priority at traffic signals equipped with transit priority control, and reports to the central dispatch its location for management and customer information purposes. As the bus traverses the route, operational measures are stored on the data card. At the end of the day, the card is removed from the bus and uploaded to the BDS data warehouse. TriMet's scheduled regular weekday service contains nearly 10,000 trips and on a typical weekday the BDS records about 500,000 entries [1]. Only a portion of information is currently available in real time due to communication limitations [2].

While TriMet makes extensive use of the data for internal operations, the BDS data also have strong potential for other uses such as measuring arterial performance. Commonly used measures of arterial performance include volume, speed, occupancy, vehicle miles traveled, vehicle hours traveled, travel time and delay [3]. Depending on their end use, these data may be reported at various temporal and spatial aggregations. Because buses are essentially acting as probes, if a sufficient number of buses are deployed, several of these measures can be calculated from the bus data. One challenge in using the data, of course, is that a bus is not a car. Buses operate on a schedule and stop to pick up and drop off passengers but are still subject to congestion and other road conditions. Other researchers have explored the uses of buses as probes on arterials by identifying and separating out activity that is unique to buses, such as dwell times, acceleration and deceleration [4,5,6,7].

This paper begins by describing the BDS archived data and demonstrating its usefulness for transit operations by recreating a number of common transit performance measures for a sample route. The paper then describes the preliminary work being done to use the data to develop measures of arterial performance, including graphical representations of arterial congestion levels.

II. Data Description and Filtering

TriMet has made several weeks of stop level data available for inclusion in PORTAL (Portland Oregon Regional Transportation Archive Listing), a sample of

Manuscript received March 31, 2007. This work was supported in part by the Oregon Department of Transportation, the Federal Highway Administration, and the National Science Foundation.

M. Berkow is with the Toulan School of Urban Studies and Planning, Portland State University, Portland, OR 97201 (e-mail: mathewb@pdx.edu). J. Chee is with the Department of Computer Science, Portland State University, Portland, OR 97201 (e-mail: chee@cecs.pdx.edu). R.L. Bertini is with the Department of Civil and Environmental Engineering, Portland State University, Portland, OR 97201 (e-mail: bertini@pdx.edu). C. Monsere is with the Department of Civil and Environmental Engineering, Portland State University, Portland, OR 97201 (e-mail: monsere@pdx.edu).

TABLE 1
SAMPLE TRIMET BUS DISPATCH SYSTEM DATA

Service Date	Leave Time	Badge	Route No	Direction	Service Key	Trip No	Stop Time	Arrive Time	Dwell	Location ID	Door	On	Off	Maximum Speed	Train Mileage
2006-09-21	23534	720	72	1	W	1080	23280	22002	1484	115	1	2	0	44	26.4
2006-09-21	23536	720	72	1	W	1080	23299	23534	0	8681	0	0	0	36	26.4
2006-09-21	23610	720	72	1	W	1080	23336	23556	0	9399	0	0	0	37	26.7
2006-09-21	23660	720	72	1	W	1080	23414	23648	0	2195	0	0	0	37	27.1
2006-09-21	23684	720	72	1	W	1080	23452	23676	0	2201	0	0	0	37	27.3
2006-09-21	23704	720	72	1	W	1080	23481	23700	0	2226	0	0	0	32	27.4
2006-09-21	23738	720	72	1	W	1080	23518	23730	0	9403	0	0	0	35	27.7
2006-09-21	23766	720	72	1	W	1080	23539	23746	7	3160	1	2	0	31	27.8
2006-09-21	23784	720	72	1	W	1080	23562	23780	0	3179	0	0	0	27	27.9
2006-09-21	23818	720	72	1	W	1080	23586	23794	5	3162	1	1	0	28	28.0

which is presented in Table 1. PORTAL is the region's archived data user service (ADUS) and is a research effort in the Intelligent Transportation Systems laboratory at Portland State University. The PORTAL project (portal.its.pdx.edu) is creating an extensive archive of transportation data and has nearly three years of freeway data archived [8]. The example in Table 1 and all remaining examples in this paper were generated using data for the outbound runs of TriMet Route 72 Killingsworth/82nd Ave on September 21, 2006. A schematic of the route is shown in Fig 1. The outbound direction follows the route from the northwest corner (Swan Island) to the southeast corner (Clackamas Town Center). The route is approximately 18 miles in length.

In order to identify an individual bus trip, it is necessary to obtain a unique combination of data elements from Table 1 (*Service Date*, *Trip Number*, *Route Number, Service Key* and *Direction*). *Service Date* refers to the date of the trip. *Trip Number* refers to the route and may appear in the data on multiple days (each weekday, for example) and is attached to both the inbound and outbound segment of a given run. *Service Key* refers to either weekday, Saturday or Sunday service while *Direction* indicates whether it is an inbound or outbound run.

A number of location-related data elements are captured. Each individual bus stop in the TriMet system is

Fig 1. Map of TriMet Route 72 Killingsworth/82nd Ave.

assigned a unique *Location ID* (this location identification is also used for real-time traveler information). The BDS captures two measures of distance. *Pattern Distance,* which is recorded in feet, is an estimation of the linear distance of the bus from the beginning of the route. *Train Mileage* is the cumulative actual distance traveled by the bus and is recorded in miles. The bus also records the *Maximum Speed* that it achieves at an instant between two stops.

There are four time variables, all of which are recorded in seconds past midnight. *Arrive Time* is the time that the bus either opens its door or comes within 30 meters of a scheduled stop. *Leave Time* is the time that the bus leaves the vicinity of the stop location. *Stop Time* is the time that the bus was scheduled to arrive at that particular stop, which enables TriMet to evaluate the on-time status of a bus and correlate on-time performance with other variables. The schedule information changes over time so the BDS data warehouse also captures the schedule information. *Dwell* time indicates the number of seconds that the door is open at a given stop. Passenger boarding, alighting, and dwell time can be evaluated to identify the amount of time that a bus spends picking up or unloading passengers.

Finally, the buses are equipped with automatic passenger counters (APC) that record the number of passengers boarding (*On*) and alighting (*Off*) the bus. These two fields are used to estimate the passenger load for each stop record.

Much of this research has been done working with the data exactly as it was received from TriMet. As with any "real" data source there are missing values, errors, and data nuances that must be fully understood prior to using the data to develop other measures. PORTAL and other archived data sources are concerned with ways to measure data quality and present this quality metric to the user so that only good data are used for analysis. A script has been created which flags bad records as follows:

- Arrive time of the current record is less than the depart time of the previous record

TABLE 2
ROUTE LEVEL PERFORMANCE MEASURES

Metric	Value
Scheduled number of trips	108
Actual number of trips	104
Scheduled hours of service	126.2
Actual hours of service	124.6
Scheduled number of miles	1,796
Actual number of miles operated	1,750
Schedule average running speed	14.2
Actual average running speed	14.0
Actual average 'Max' speed	29.8
Number of passengers carried	7,680
Total boardings and alightings	15,360
Average passenger load during each trip	17.2
Number of passengers per mile	4.4
Number of operators	62

- Train mileage of the current record is less than the train mileage of the previous record
- Dwell = 0 and (door > 0 or ons > 0 or offs > 0)
- Dwell <>= 0 and (door = 0 and ons = 0 and offs = 0)
- Max speed > 80
- Pattern distance = 0

In the following sections, the usefulness of the BDS for transit performance and arterial conditions is demonstrated.

III. TRANSIT PERFORMANCE MEASURES

The BDS data can be used to generate a wide array of standard route-level performance measures. Samples of these performance measures, suggested in [9], are presented in Table 2 for Route 72, outbound on September 19, 2006. These measures provide a quick snapshot of how the route performed on a particular day. Clearly, in addition to the point estimates in Table 2, with proper analysis, these performance metrics can be measured over time, and along with their variance can provide a more robust picture of performance. Two examples of the more robust metrics are presented in the following sections. Note that there is a multitude of metrics that can be presented; for additional examples refer to [9,10,11,12,13,14].

A. Variability in Run Time

In constructing accurate bus schedules and analyzing performance, it useful to know how measures vary over the day. The scatter plot in Fig 2 shows the total time it takes for a bus to complete its route as a function of start time. As one would expect, the time required for a bus to complete its route varies with the time of day. Longer travel times result when the bus encounters congestion and increased passenger activity during the peak periods (as seen at around 7:00 AM and 4:00 PM). For this outbound route, peak passenger times occur in the evening. While this plot shows one day of data, additional days could easily be added and the average as well as the variance of the run time for a particular time period could be estimated. This information is used to modify schedules to more accurately represent conditions.

B. Headway

Headway, which refers to the time spacing of buses at a particular point, represents one of the most important transit performance measures. Headway and how it varies is important because passengers are extremely sensitive to wait times when determining which transportation mode they will use. In fact, economic studies indicate that commuters value a reduction in walking and waiting time roughly three times as much as they value a reduction in time spent on the bus.

Fig 3 shows a bar plot of actual minus scheduled headway for Route 72 over the day at a particular stop. The figure highlights some fundamentals of transit operations. For each arrival at the stop, the difference between the actual and measured headways oscillates

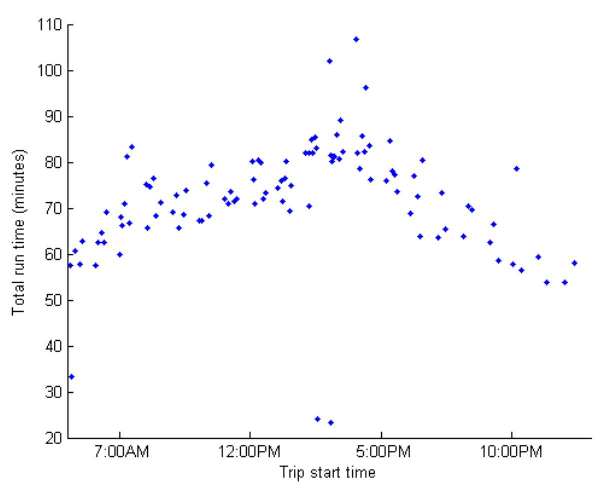

Fig 2. Scatter plot of total trip time for Route 72, outbound, September 21, 2006

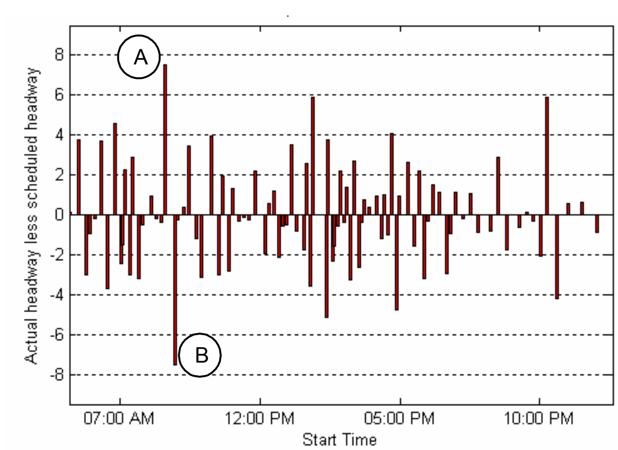

Fig 3. Bar plot of actual minus scheduled headway for Route 72, outbound, September 21, 2006 at stop 115.

between being positive and negative. The arrival shown at time A was approximately 7 min behind the scheduled headway. The headway between the next bus, B, is 7 minus less than scheduled headway. Minimizing headway variance allows for more efficient bus operations. If a bus reaches a given stop late, there are likely to be more than the normal number of passengers waiting to board. Passenger boarding takes time, as many people pay their fare with cash rather than with a bus pass. As a result, the buildup of passengers at the given stop tends to result in the bus falling further behind schedule, a situation that compounds on itself as the bus progresses along its route. Meanwhile, the following bus, if it is on time, arrives only a few minutes after the previous bus. Because there are few, if any, passengers at each stop, the second bus steadily gains on the increasingly late bus. The result is one bus that is overly full, one that is mostly empty and poor service for passengers (due to overcrowded buses and inconsistent wait times). This is a common operational challenge known as bus bunching. From an operator's perspective, bus bunching results in an inefficient use of equipment and the BDS data allows this performance to be explicitly monitored.

Fig 4 offers a graphical representation of bus bunching. The figure shows a time-space diagram of sequential bus trip trajectories with time on the *x*-axis and distance on the *y*-axis. The trajectory is annotated with dots to show locations where passengers boarded the bus. The horizontal distance between any consecutive bus trajectories at any distance, *x*, represents the headway at that point. The slope of the trajectory at any point is the speed of the bus. An ideal figure would show trajectories spaced at the scheduled headways with essentially similar slopes over the time period. In the figure, bus A which began its route from $x = 0$ at 7:27 AM, begins to fall behind schedule (slope of the line decreases relative to previous trajectory). The next bus which departs $x = 0$ at 7:40 AM begins to catch up and as it does so, the horizontal (headway) and vertical distance (spacing) between it and the following bus gets progressively smaller until at approximately 8:10 AM the buses are essentially together. The dots on the plot indicate segments where the buses picked up passengers. It is clear that in this instance, the late bus is picking up many passengers while the early next bus is picking up progressively fewer. Transit operators can evaluate bus routes using similar graphs to decide if schedule adjustments or additional equipment needs to be added.

This section has described the sample performance measures for transit operation that can be gleaned from the TriMet BDS system. The following section describes how the BDS might be used for real-time arterial performance measurement.

IV. Arterial Performance Using Bus Probes

Traditional measures of collecting data with which to evaluate arterial performance, such as sending out test probe vehicles equipped with GPS units, is costly. With so many buses operating on the arterial network at any given time, it is potentially feasible to consider the buses as probes for measurement of arterial performance [5]. Measuring real-time conditions of arterials has a number of potential advantages. Travelers could gain information about the current traffic conditions on all routes so that potential alternate routes (particularly for short trips) could be evaluated and chosen. Managers and operators of the system could use the data to evaluate changes in arterial performance over time.

Many metropolitan areas have extensive detection infrastructure on freeways and use these data for communicating real-time traffic conditions to the public and other operating needs. Estimating current traffic conditions on arterials is more complicated primarily due to the presence of traffic signals. Arterials typically are instrumented less comprehensively and if they are, the

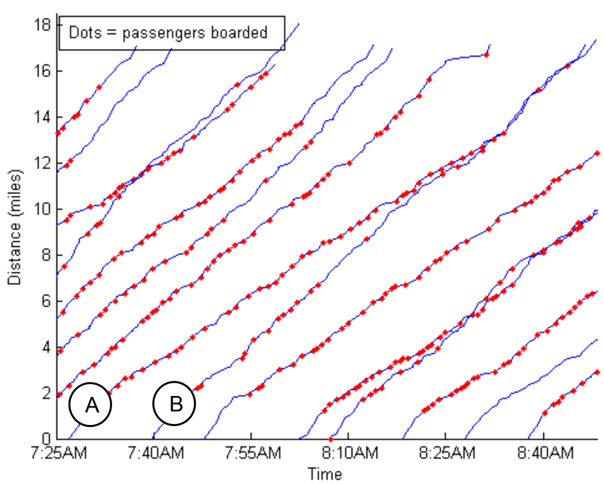

Fig 4. Time-space diagram of bus trajectories annotated with dots showing passenger boarding locations.

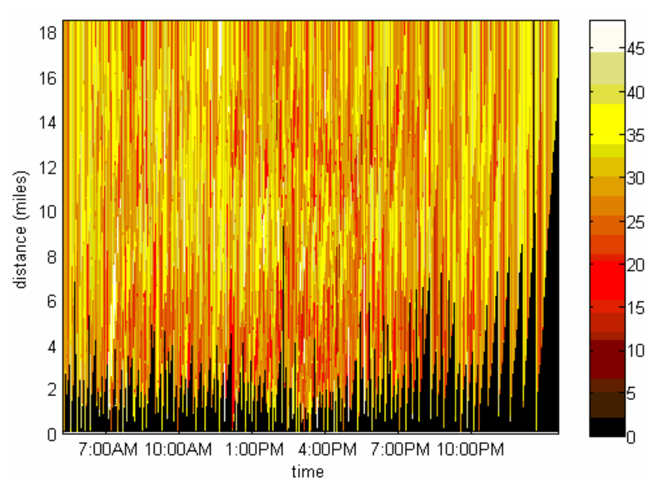

Fig 5. Time-space contour plot of speed, for Route 72, outbound, September 21, 2006

detection is primarily designed to operate the signal system rather than collect traffic information. In many cases, central communication is also limited. These challenges are the primary reasons why using buses as probes is being considered.

There are number of challenges with using bus vehicles to measure arterial performance in real-time. First, as discussed in the introduction a bus does not exactly behave the same as other vehicles in the traffic stream. However, the BDS data is configured in such a way that methods to estimate the performance of other vehicles are possible. Second, while there are a substantial number of buses on the network, each given link only will be traversed by a bus at a time interval equal to the scheduled headway. The most frequent service in the off-peak period on TriMet routes is 15 minutes (some higher volume routes have shorter headways in the peaks). Also, along a particular bus route there will only be a limited number of buses on the route at one time. For example, in Fig 4 at 8:10 AM there are 8 buses on the route. In summary, arterial performance using bus probes can only be measured at the resolution dictated by bus operations. On some dynamic arterials, this limitation may be significant. In the following subsections, two methods to represent arterial conditions are presented.

A. Speed Contour Plot

Fig 5 shows a *t*-*x* plane for one complete day (September 21, 2006) of Route 72 Killingsworth with time on the *x*-axis and distance on the *y*-axis. The maximum speed (in mph) is expressed as a color with darker colors indicating slower speeds. The speed estimates of the arterial are created by assigning the measured maximum speed to a portion of the actual route at that time. Referring to Fig 4, at each vertical time slice there are a number of buses on the route. The speeds at each time slice are plotted in Fig 5. If there is data missing, interpolation was performed to create a more complete plot. In this case, over 12,000 data points were used.

The red colors on the map indicate lower speeds and potential congestion. In Fig 4, mile four seems to be a consistently slower section. The morning commute (7:00-8:00 AM), early afternoon (1:00-2:00 AM) and the evening commute (4:00-5:00 AM) can all be identified as vertical sections on the map that are predominantly red.

B. Condition Maps

Fig 6 and 7 demonstrate another way that the bus data can be used to represent an arterial condition map (similar to many freeway condition maps). PORTAL has developed a number of map-based displays using Google Maps as the underlying map display platform. PORTAL provides an interface for drawing markers and lines on the Google maps. In Fig 6 and 7, this software platform is also used. In Fig 6, a marker is placed at each stop. The

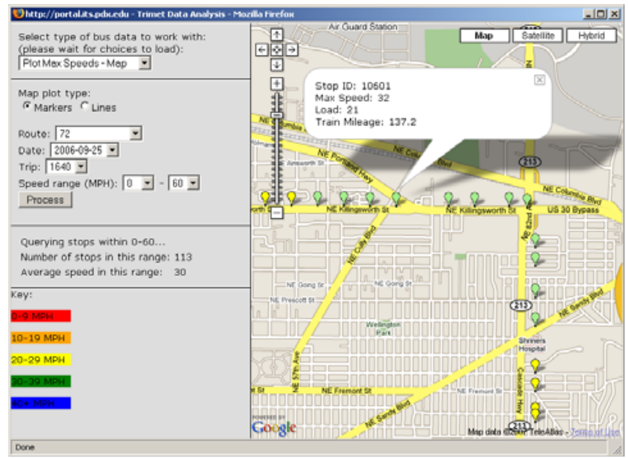

Fig 6. Condition map with speed between stops plotted as a point representation in PORTAL.

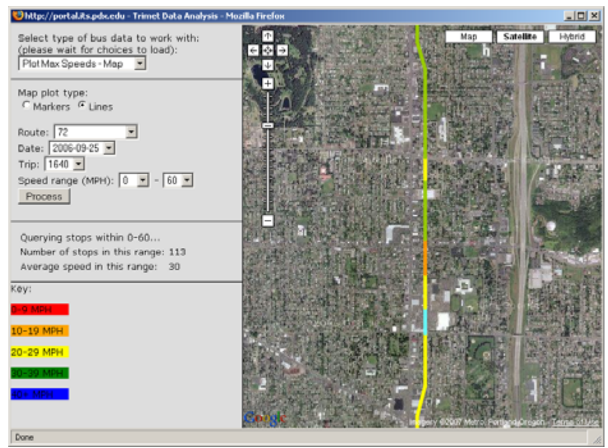

Fig 7. Condition map with speed between stops plotted as a line representation in PORTAL.

color of the marker is determined by the *Maximum Speed* value for that record. The PORTAL archive has a web-based interface to the data. Users can click on a marker to see more information from the data, such as the *Location ID, Maximum Speed*, *Estimated Load*, and *Train Mileage* at that particular stop. Fig 7 is similar to Fig 6, except that line segments, instead of markers, are drawn between each stop. The color coding is again based upon the *Maximum Speed*.

These figures show an example of one particular trip. The problem is that this trip takes over an hour, so the map does not present a "snapshot" of the roadway at a particular time. However, it is possible to create an interface which will enable users to view the bus route, or even the entire road system at ten minute intervals. For example, from 7:00 to 7:10AM, the user would see color coded speed lines for all buses currently on the road system (essentially the contour plot for a time slice displayed on a map).

V. DISCUSSION

TriMet has designed the BDS so that it would be most useful for the types of analysis that they consistently perform. When making use of these data to evaluate arterial performance, there are some facets of the database structure that make it a bit cumbersome and could challenge real-time application. These are for the most part minor problems that could be easily addressed with cooperation from TriMet.

For example, buses do not all begin offering service, and thus recording data, at the same location along the route. This creates a problem when attempting to compare different trips, especially in a time-space plot such as Fig 3. The two fields that record distance traveled indicate either the distance traveled from the start of the particular trip (*Pattern Distance*) or from the start of the bus's recorded service for the day (*Train Mileage*), rather than from some consistent location such as the bus terminal. If a bus does not pull up to a certain location at the end of a trip, the AVL unit does not recognize that the trip is over. When the bus driver then opens the door and alights the bus for its layover, the layover time is recorded in the *Dwell* field. An example of this phenomenon can be seen in the first record of Table 1. While not a major problem, it does add a level of complexity to computing different types of dwell measures. Finally, the *Maximum Speed* field, which records the maximum speed attained by the bus between two recorded stops, sometimes records speeds in excess of 80 miles per hour. This reflects the fact that GPS signals can be lost particularly in the downtown area.

VI. Conclusions

This paper provides a sample of the types of analysis that can be done using archived transit data. Given the large amount of data that buses collect, there are indeed many more ways in which the data can be explored. Much of this research up until now has focused on understanding the data by computing transit performance measures. In the future, this research will extend this process into the creation of measures of arterial performance.

One challenge with working with bus data to measure arterial performance is separating out delay caused by congestion from delay caused by things such as passenger movement. One possibility for addressing this problem is to use the bus data in conjunction with other traffic sensor data such as that from a central signal system.

Much of the analysis has focused on viewing the data one day at a time. Other applications can be imagined that would enable the user to view trends over time. Average trip times, for example, can be analyzed week to week, month to month or even year to year. Since it is possible to account for the amount of passenger movement and dwell times, changes in average trip times can be used to by traffic professionals to track changes in congestion levels.

The bus route that was used for this analysis completes over 100 trips each day, in each direction. Each one-way trip records over 100 rows of data. As a result, one day's worth of data in one direction can comprise over 12,000 rows. Multiply this by the more than 90 different bus lines that TriMet operates each day and it becomes apparent that the BDS is a rather large data source. The task is to filter through this data in order to find its usefulness for measuring arterial performance, both in the moment and over time.

Finally, transit database structures have typically been designed with the evaluation of transit performance in mind. By studying this data, it is likely that improvements to the database structure (i.e. additional fields of data) can be recommended that would render the data more useful for measuring arterial performance.

Acknowledgement

Funding and support for this project is provided by the National Science Foundation, Oregon Department of Transportation, Federal Highway Administration, City of Portland, and TriMet. The authors particularly thank David Crout and Steve Callas of TriMet.

References

[1] TriMet, "Data resources derived from BDS card data," unpublished.

[2] D. Crout, "Accuracy and precision of TriMet's transit tracker system," unpublished.

[3] R. L. Bertini, S. Hansen, S., A. Byrd, and T. Yin, "PORTAL: Experience implementing the ITS archived data user service in Portland, Oregon." *Transportation Research Record 1917*, TRB, Washington, D.C., pp. 90-99, 2005.

[4] J. Nee and M. Hallenbeck, "Surveillance options for monitoring arterial traffic conditions," Research Project T1803, Task 14. Washington State Transportation Center (TRAC), 2001.

[5] S. Tantiyanugulchai and R.L. Bertini, "Arterial performance measurement using transit buses as probe vehicles," in *Proceedings of the IEEE 6th Annual Conference on Intelligent Transportation Systems*, Shanghai, China, October 2003.

[6] P. Chakroborty and S. Kikuchi, "Using bus travel time data to estimate travel times on urban corridors," *Transportation Research Record*, No. 1970, TRB, Washington, D.C., pp. 18-25, 2004.

[7] R. Hall and N. Vyas, "Buses as a traffic probe," *Transportation Research Record 1731*, TRB, Washington, D.C., pp. 96-103, 2000.

[8] R.L. Bertini, S. Hansen, A. Byrd, and T. Yin, "PORTAL: Experience implementing the ITS archived data user service in Portland, Oregon." *Transportation Research Record 1917*, TRB, Washington, D.C., pp. 90-99, 2005.

[9] R.L. Bertini and A. El-Geneidy, "Generating transit performance measures with archived data," *Transportation Research Record 1841*, TRB, Washington, D.C., pp. 109-119, 2003.

[10] *Web Document No. 6: Transit Capacity and Quality of Service Manual*. TCRP, TRB, Washington, D.C., 1999.

[11] G.J. Fielding, *Managing Public Transit Strategically,* Jossey-Bass, San Francisco, CA, 1987.

[12] G.J. Fielding, "Transit performance evaluation in the U.S.A.," *Transportation Research*, Vol. 26A, No. 6, 1992, pp. 483–491.

[13] J. Strathman, T. Kimpel, K. Dueker, R. Gerhart, K. Turner, D. Griffin, and S. Callas. "Bus transit operation control: review and an experiment involving TriMet's automated bus dispatching system," *Journal of Public Transportation*, 4(1) 1, 2001, pp. 1–26.

[14] H.W. Levinson, *Synthesis of Transit Practice 15: Supervision Strategies for Improved Reliability of Bus Routes*, TRB, National Research Council, Washington, D.C., 1991.