Perspectives on Transit
Potential Benefits of Visualizing Transit Data

Colin Stewart, Ehab Diab, Robert Bertini, and Ahmed El-Geneidy

Advancements in information and communication technologies have enabled transit agencies around the world to generate streams of data on a high-frequency basis. Increasingly, these agencies are interested in new methods of visualizing these data to communicate the results of their planning efforts, operational investments, and transit performance to decision makers and stakeholders. Most agencies today collect and provide numerous kinds of data, including Google’s general transit feed specification schedule data, automatic vehicle location data, and automatic passenger count data. This paper aims to demonstrate the untapped potential of these data sources; specifically, the paper uses transit data from Montreal, Quebec, Canada, to generate performance measures that are of interest to both transit planners and marketing professionals. Some of these measures can also help in communicating the positive attributes of public transportation to the community. Performance measures are generated at different scales, including transit system, neighborhood, route, and stop levels. This paper expands on previous research on transit performance research and visualization by adopting currently available resources for so-called big data.

In the past century, collecting data about public transit systems was difficult and expensive. It typically required manual counting of transit metrics (e.g., the number of passengers boarding and alighting) (1). However, with the advent of digital transit monitoring systems [GPS, automatic vehicle location (AVL), and automatic passenger counters (APC)], as well as open data format, such as general transit feed specification (GTFS), there is currently an abundance of transit-related data available for analysis. AVL and APC data represent one of the main sources of so-called big data used for transit planning (2).

At the same time, concerns over global warming, urban sprawl, and quality of life, plus generational and demographic changes, are making transit an increasingly important part of the public conversation. Thus transit frequently dominates the headlines (3–5). In addition, because major transit projects can run into serious opposition, the need for informed public debate, and for honest and positive images of transit, is more critical than ever. This paper aims to show how the wealth of data gathered by transit agencies can be used to demonstrate transit performance and benefits, either internally in a transit agency or externally to policy makers and the general public, through the use of visualization techniques.

Much of the previous research on transit data visualization and performance has been targeted internally toward transit agency management with a focus on cost efficiency and service effectiveness (6–8). For example, Levinson primarily addresses transit service reliability, discussing the organizational structures and technology that can be used to ameliorate reliability issues (9). However, a limited number of studies focus on visualizing general aspects of transit that can be used to illustrate planning efforts and to help planners on negotiating service modification projects with stakeholders. Bertini and El-Geneidy mostly concentrate on generating performance measures at the route and stop levels, using such metrics as route spacing, fluctuations in passenger volumes, average route speed, and schedule adherence (1). Berkow et al. extend this work, providing additional performance measures and some visualizations using automatically collected AVL and APC data (10).

By contrast, transit performance measures that are targeted at customers, decision makers, or higher-level administration tend to concentrate on convenience, comfort, and safety (11). Similarly, most transit information systems convey only the mechanics of how a transit system works, largely through maps and schedules (12). However, many transit agencies also invest in broader marketing that takes a view of the bigger picture, such as advertising that aims to create a positive image of transit (13). This paper, in part, takes its lead from this kind of marketing practice, but also seeks to go beyond it by making greater use of the vast and increasing quantities of available transit data.

The overall objective of this paper is to demonstrate the untapped potential of using commonly available data for transit agencies to generate performance measures that are of interest to transit planners and professionals, which could be used to communicate transit aspects and improvement efforts to the public and decision makers. It focuses on generating visualization techniques rather than evaluating the performance of transit agencies, which has been extensively done in the literature (14–17). Thus the visualizations will focus on more general issues and aspects, rather than one aspect of transit, giving planners a wide variety of methods. While keeping that in mind, the generated visualizations will help planners to negotiate and support the implementation of bus stop consolidation and all-door-boarding plans, to illustrate the current network features and changes, and, finally, to offer more initiatives to enrich the transit data visualization literature.

DATA AND RESEARCH METHODS

Montreal, Quebec, is the second most populous metropolitan area in Canada with 3.7 million inhabitants. The Société de transport de Montréal (STM), the city’s transit service provider, serves 1.8 million
people daily. STM operates 69 km (43 mi) of subway lines and 222 bus routes. Many of the visualizations in this paper were generated with STM’s archived AVL and APC data. STM has equipped a total of 306 buses with this technology, which are assigned to varying routes to obtain a sample of the entire bus network’s performance. AVL and APC data are recorded at the stop level, including bus arrival and departure times and number of passengers boarding and alighting.

Several visualizations also use STM’s GTFS data. The GTFS format is an increasingly popular open data format (18) used by Google Maps for its transit trip applications. GTFS data are frequently updated by hundreds of transit agencies around the world (19) and include transit supply information. The data are represented as a series of text files with comma-separated values; the main files used in this study are stop_times.txt and stops.txt, which include the scheduled arrival and departure times by stop, trip, and route ID, and stop unique name, sequence, and coordinates.

Other data used in this paper come from a variety of sources. These sources include Bixi bicycle-share data, DMTI Spatial, and the 2003 Montreal origin–destination survey. These data sets were processed with Microsoft Excel and Access and then visualized with ArcGIS and ArcScene 10.1 and Adobe Creative Suite 6 (primarily Illustrator).

This research generates performance visualizations at four hierarchical spatial levels. At the system level, the visualizations deal with transit coverage and spread, improvement in system speed, and waiting time at stops. At the neighborhood level, the interaction between Montreal’s night bus service and the Bixi bicycle-share system is explored, as is the relationship between transit service and population density. At the route level, various passenger performance indicators are analyzed, such as bus type, load factor, route seating capacity, and stop spacing. Finally, at the stop level, bus stop demand and headways are examined. For each visualization, the type of data, methods, and the visualization purpose are discussed.

**SYSTEM-LEVEL ANALYSIS**

**System Service Span**

Recently, several transit agencies have equipped their transit stops with LCD screens to show various media and transit real-time information, cutting down passengers’ negative perceptions of waiting time (20–22). STM, for example, has equipped the majority of its metro stations with screens that show real-time information about the transit system, as well as different types of entertainment media. To date, these LCD screens have not been used to show how Montreal’s transit network functions. With that in mind, a video (screenshot shown in Figure 1 (23)) was created. It shows the flow of all buses and metro trains on the island of Montreal over one typical weekday. With the use of GTFS data from winter 2013, it was possible to identify the scheduled location of buses and metro trains at each minute during the day. The ArcGIS linear referencing tool was used to approximate where each bus and metro was at 30-s intervals; a series of Python scripts were written to automate this process for all the bus routes and metro lines. The final video, which was assembled by using Adobe Premiere, Illustrator, and AfterEffects, runs for 96 s, with each hour of the day spanning 4 s in the video.

Different modes of transport are assigned different colors, and time is indicated by the clock at the top right. White points indicate a regular bus service, purple points show the frequent bus service (10-min headways or less), while the other colors show the metro service. The video shows that metro service starts at about 5:30 a.m. and ends around 1:00 a.m., while the buses run for 24 h. The video also highlights the complexity of the transit network, the length of service provision over the day, and the daily activity peaks. This short movie can be displayed on the screens at STM’s metro stations and bus stops, which will help to increase public awareness.
about the quality of the service, which is strongly linked to positive increases in passengers’ perceptions and overall satisfaction (24).

System Coverage

Catchment areas around transit stops are useful for understanding service coverage. Overlap between catchment areas indicates that an area is served by multiple buses or metros; thus the level of transit service can be measured by the number of overlapping catchment areas. Figure 2a shows the Island of Montreal’s overlapping catchment areas. Each catchment area was based on the 85th percentile of the walking distance estimates (524 m for bus service and 873 m for the metro) during the morning peak period (06:30–09:30). This distance is calculated by using the 2003 Montreal origin–destination survey and the network analyst extension in ArcGIS.

The map was generated by intersecting 100-by-100-m grid cells with the network distance service areas (using network analyst in ArcGIS) for each scheduled stop; the color of each grid cell corresponds with the number of catchment areas that overlap it. The intersection is done with the Intersect function in the Overlay toolset/ArcGIS Toolbox. Network analyst was used, as recommend in the literature (25), to provide an accurate estimation of the actual shape of the service area according to the pedestrian and street networks. Here, scheduled stop refers to the stops of each bus route rather than the physical stops; thus if a physical stop serves multiple bus routes, multiple catchment areas were generated. For each route, only one direction was randomly considered for illustration to avoid double counting.

Red denotes areas with high levels of supplied service with more than 136 scheduled stops, indicating that a high level of transit service is within walking distance. The shades from orange to yellow denote areas with medium levels of supplied service ranging between 14 and 136 scheduled stops per cell. The green color represents fewer than 14 scheduled stops per cell, indicating a lower level of transit supply. This figure can serve two purposes. The first purpose is providing a visual indicator for passengers to understand where concentrations of transit service are present, within a walking distance. Transit availability has long been seen as an important factor in residential and firm location decisions (26). The second purpose is to highlight the transit agency’s efforts to provide high levels of service in some neighborhoods.

Density of Buses

Figure 2b shows the density of buses in the Montreal system during the morning peak period (06:30–09:30). Darker green indicates fewer buses while brighter red indicates more buses. This figure was generated by using winter 2013 GTFS data. As with Figure 2a, the city was divided up into 100-by-100-m grid cells; then, the number of buses entering each cell was calculated, and a kriging function was used to create a smooth interpolated surface.

This figure shows how buses are distributed throughout the city. In the west, the greatest concentrations of buses are around single points, primarily malls and commuter rail stations, whereas in the middle of the island, buses are increasingly concentrated on entire streets rather than at single points. This distribution reveals a higher degree of bus resources directed in these neighborhoods. This knowledge could be used, among other system-level visualizations, to identify the location of required improvement strategies (e.g., reserved lanes or articulated buses) or high-capacity transit, such as bus rapid transit or light rail, along these corridors.

Speed Improvement

Transit agencies strive to improve their service and, accordingly, to increase ridership and revenue. With the use of GTFS data for the winters of 2009 and 2013, it was possible to identify the improvements in bus route speed between the two periods and the speeds of the new routes according to schedule in the winter of 2013. This information may be reported while considering the percentage of on-time buses. According to STM, around 83% of buses were on time (1 min early to 3 min late) over the past few years, with minor improvements. This finding indicates that the improvement in speed did not affect the buses’ on-time performance.

Figure 2c shows these changes during the morning peak period (06:30–09:30). The speed improvement between the winters of 2009 and 2013 is represented by red lines. Thicker lines indicate higher segment travel speed and thus greater time savings for passengers. The speeds for each route segment were calculated by finding overlaps between route segments. This method was helpful for analyzing the impacts of a reserved bus lane or a transportation system plan along a segment shared by more than one route. The maximum segment speed improvement was approximately 3.2 km/h, a 10% speed increase from 2009 to 2013. This type of map would help planners to visually identify changes in speed and closely investigate the reasons behind the changes. In addition, it could be used by a transit agency to highlight its efforts and the use of improvement strategies (e.g., bus lanes), as has been done by STM between 2009 and 2013 (27).

The figure also shows the speeds of the new routes added between 2009 and 2013. These speeds ranged between 10 and 35 km/h. These numbers can help passengers realize their travel speeds on the transit network and may affect commuters’ distorted perception regarding transit travel time. Correcting this distortion and increasing the awareness of travel mode alternatives and their qualities may help increase transit use (28–30).

Waiting Time at Stops

Waiting time is one of the most tedious aspects of transit use (31–33). Research has found that the time spent waiting is perceived to be two to three times more valuable than that of in-vehicle travel time (34); it has also been reported that transit users prefer to walk farther distances than to wait longer at stops (35).

Figure 2d shows the average waiting time at the system level for Montreal during the morning peak period (06:30–09:30). Mean bus wait times were derived from winter 2013 GTFS data and were calculated by halving the headway for frequent routes, and by assigning a value of 8 min for those with headways exceeding 15 min. The logic here is that on routes with sub-15-min frequencies, passengers just arrive at their stop when it suits them, and then wait on average half the headway; by contrast, on less frequent routes, passengers typically consult schedules instead (36–38). For the figure, bus stops that served multiple routes were duplicated so that waiting time values were included for each route. After all the waiting times were calculated, a surface was interpolated using the stops’ scheduled headway values, creating a continuous, smooth surface. The kriging function in ArcGIS was used to create the surface, in which the surrounding measured values were weighted (based on distance between the measured points) to derive a predicted value for an unmeasured location.

This map is useful, from a passenger’s point of view, for spatially linking the expected waiting time to where he or she lives or works. It should be noted that this figure, as well as Figure 2a and c, represents...
FIGURE 2  Maps of (a) overlapping service area, (b) density of buses during the a.m. peak period, (c) positive changes in speed and new added bus routes, and (d) wait time at bus stops in winter 2013. (Source: Data, STM, DMI, GTFS; projection, NAD 1983 MTM 8.)
the three main components of transit trip from the users’ point of view (i.e., walking, in-vehicle, and waiting time).

**NEIGHBORHOOD LEVEL**

**Bus and Bicycle-Share Systems Overlap**

Sometimes it is important to show how different transport options are working together at smaller scales than at the systemwide level. This point is true when one considers the Bixi bicycle-share network, which only covers an area directly surrounding the downtown core of Montreal. Unlike the metro network, Bixi does not have a schedule; it is available to riders all day and night. Thus Bixi can be an effective option for transit riders after the metro closes (1:00 a.m. on weekdays). The other main option for transit riders in this time period is the night bus network. Depending on the situation, Bixi may be a better alternative than night buses or vice versa.

Figure 3 shows the accessibility to both of these services when the metro is no longer operating in the evening. Using accessibility measure is common on the literature to generate maps (39). The figure was derived with data from STM GTFS and bixi.com and was created by measuring the level of accessibility of each postal code area (typically equivalent to a block face) to night buses and to Bixi bicycles. For each postal code, the distance (dist) from its centroid to each night bus stop ($i$) and Bixi station ($j$) within walking distance [400 m (¼ mi) along the pedestrian network] was determined by using ArcGIS. Then, the following formulas were used to calculate two accessibility scores for each postal code.

\[
\text{accessibility}_{\text{bus}} = \sum_i \frac{\text{night buses per hour}}{\text{dist}_{\text{bus}}^i}
\]

\[
\text{accessibility}_{\text{Bixi}} = \sum_j \frac{\text{mean available Bixi bicycles at night}}{\text{dist}_{\text{Bixi}}^j}
\]

The accessibility maps for each of the above two measures were visually overlaid to show where night bus accessibility (blue) overlaps night Bixi accessibility (yellow); the areas of overlap are shown in green. What the figure ultimately shows is that between the night bus network and the Bixi network, a large proportion of Montreal near downtown is completely covered. This figure is also of use to people who need to get home after the metro closes and who need to decide whether bus or Bixi is the most suitable mode. This method of analysis could be extended to other cities or modes. Planners may use this figure to identify neighborhoods that are underserved at night.

**Passenger Activity in Different Boroughs**

Understanding passengers’ activity changes over a day is important for transit agencies to provide better service for a reasonable cost. Different boroughs in the city may have completely different tran-
sit patterns over a day. Figure 4 shows the difference in passenger activity between three boroughs of Montreal: Pointe-Claire, a suburban borough in the west of the island of Montreal; Ville-Marie, the downtown borough of Montreal; and Plateau–Mont-Royal, a mixed-use borough directly northeast of downtown. From each of these boroughs, a representative bus stop was chosen by using the following method: the total daily passenger activity (boardings and alightings) of each bus stop in each of these three boroughs was calculated, and the bus stop in each borough with the median level of activity was selected. The hourly activity of these three stops was then plotted over a day. This figure was generated by using STM’s AVL and APC data collected throughout the entire year of 2013 (around 600,000 records).

As the figure shows, different patterns can be identified over the day. The figure shows that the downtown borough, Ville-Marie, despite being the core of the city, has significantly lower bus usage than does Plateau–Mont-Royal at all periods of the day other than the morning peak; also, activity in Plateau–Mont-Royal continues late into the night. Passenger activity downtown (Ville-Marie) peaks sharply during the morning hours, then has two smaller peaks in the afternoon before dropping off completely in the evening. In the suburban borough (Pointe-Claire), there are two waves of morning commuters at 6 a.m. and 8 a.m., but after that, activity is fairly low throughout the day. The figure can help planners visually identify variation over a day. It also shows the population density per square kilometer for each borough. This density can be linked to the variation of demand to justify increasing the offered service during a particular period.

ROUTE AND SEGMENT LEVEL

Route Seating Capacity

Viewing transit data at the route level can help planners better understand the characteristics and demand changes along a specific route and provide passengers with a better idea about how their most commonly used routes work. In this section, the majority of visualizations were generated by using STM’s AVL and APC data.

Figure 5 presents a three-dimensional illustration of the seating capacity of Route 67 northbound. Three-dimensional visualizations are useful analysis and planning support tools for transportation planning (40). Route 67 runs for 9.4 km (5.8 mi) along Boulevard Saint-Michel and connects two metro stations east of downtown Montreal. The average bus load factor (i.e., the percentage of full seats on the bus) is represented by translucent vertical bars, and the average daily boarding and alighting are shown with green and blue columns, respectively. The figure was generated by using STM’s AVL and APC data collected between January and May of 2013 (around 50,000 records); bus load factors were calculated by determining the number of passengers boarding and alighting at each stop, in conjunction with the number of seats installed on the specific buses used. The data were displayed using ArcScene 10.1.

It is clear from the figure that the highest levels of boardings and alightings are around the Saint-Michel metro station. The key policy element, the load factor, shows an increase from about 65% near the Joliette metro station to a peak of 76% near the Saint-Michel metro
station (shown in red); it then drops off as the route continues north. Transit planners can use images like this to suggest and support short-turn route policies to decrease crowding and increase passenger comfort. This information can also be displayed to passengers in real time to help them decide which stop or which trip they should take to increase their chance of getting a seat.

**Bus Stop Consolidation**

Bus stop consolidation can be one of the most effective strategies for increasing system efficiency and decreasing running times (41). Consolidation can be defined as decreasing the number of stops being served along a route while maintaining customer accessibility to the route. However, the decision of whether to consolidate bus stops is more complex from the passengers’ point of view; they must make a trade-off between the potential increase in their walking distances with the potential decrease in both their travel and waiting times. These decreases are linked to increases in ridership and rider satisfaction levels (42–45).

Figure 6 presents a conceptual geographic information system planning approach for bus stop consolidation along Route 57. A 400-m (¼ mi) buffer using network distances was defined around each bus stop; this buffer is a standard-size catchment area for buses (46–48). Next, these buffers were intersected by 20-m (66-ft) grid cells. For each grid cell, the number of intersected service areas was calculated, representing the number of stops serving that cell. As seen in Figure 6a, substantial redundancy and overlaps in the bus stop service areas exist on the southern end of the route.

Figure 6b proposes an altered route with reduced redundancy. To create this figure, AVL and APC data for the route from 2010 were used: 30,000 stop-level records were analyzed, and the stops with the lowest passenger activity (boarding and alighting) and not at critical locations (transfer or terminal points) were identified as candidates for consolidation. Ultimately, nine of the 31 stops were recommended for removal, and the greatest number of overlaps per cell decreased from nine to five.

Such consolidation would decrease the bus running time and passengers’ waiting time, while having minimal impact on passengers’ accessibility to the route. More specifically, each skipped bus stop decreases bus running time by 10 s (49–52) and its variation by 0.1% on average (50). Running time savings along the route can reach 90 s (1.5 min) on average, which represents 8% of the route total trip running time (18 min). This savings means that STM would be able to add a new trip after every 13 trips, increasing bus frequency and decreasing passengers’ waiting time, while reducing the variation in service (i.e., recovery time at layovers) by 0.9% on average. Figure 6, along with these results, can help transit planners to communicate and demonstrate the benefits of bus stop consolidation measures. In addition, instead of using mainly ridership to determine the stops that can be consolidated, more complex stop-consolidation strategies could be tested, visualized, and compared to the initial situation by using the same technique discussed above, as many times as needed.
Boarding, Alighting, and Route Direction

All-door bus boarding, allowing passengers to board buses through any door on the vehicle, is becoming an increasingly widespread practice. There are many ways of implementing all-door boarding; it can be allowed on every bus in a transit system (system level), on specific bus routes (route level), or at specific bus stops (stop level). Different patterns of passenger boarding make different implementations more appropriate than others. The ideal implementation strikes a balance between serving as many high-boarding stops and as few low-boarding stops as possible, as well as being simple and consistent for the public to understand. To facilitate acceptance of a new all-door boarding policy, it is necessary to communicate to the public and decision makers why particular implementations are chosen.

Figure 7 shows the daily boardings and alightings at each stop along Route 165, one of the busier routes in Montreal. The figure employs AVL and APC data obtained in 2013 (about 40,000 records), and was created manually in Adobe Illustrator. Figure 7 shows a schematic diagram of the stops along the route, allowing for a more immediate understanding of the passenger activity figures than a standard graph. The northbound direction of the route, on the top half of the figure, has high boarding levels at just two stops, while having low boarding...
levels at its other stops. By contrast, the southbound direction, on the bottom half of the figure, has no individual stops with such high boarding levels, but instead has moderate levels of boarding spread across about a quarter of its stops. Over the whole day, the two directions see similar levels of boarding (10,729 northbound; 9,133 southbound). The boarding pattern seen in the northbound direction could justify all-door boarding at its two main stops only (a stop-level implementation). Conversely, this figure could make the case for a route-level implementation of all-door boarding, either in the southbound direction only or on the full route.

**BUS STOP LEVEL**

**Demand at Bus Stop**

Finally, this paper discusses stop-level visualizations. Figure 8 presents a diagram and a map that illustrate the variation of demand and dwell times at three bus stops along Montreal’s northbound Route 67 throughout the day. By using AVL and APC data (mentioned in Figure 5) for the route, the average dwell time per hour and the average total passenger activity per trip at each stop were calculated. Three consecutive bus stops were selected with an average spacing of 170 m (550 ft). Dwell times are represented with lines, while the average total passenger activity per trip is shown with vertical bars. This figure shows that passenger activity at the Holt stop (the middle stop) is consistently much lower than at the other two stops; as such, Holt is an excellent candidate for bus stop consolidation. Removing this stop would eliminate the portion of dwell time required by the bus for accelerating and decelerating at this stop, thus decreasing the route’s overall running time. This figure can be used in conjunction with Figure 6 to support bus stop consolidation projects and to show the overall benefits, helping planners to visualize their service enhancement plans.

**Combined Benefits**

At stops that are served by more than one route in the same direction (e.g., local and express service), combined schedules can play a role in showing the real benefits that passengers experience. Figure 9, which was generated with 2013 GTFS data, shows the combined frequencies of Routes 67 and 467 at the Saint-Joseph Boulevard northbound stop. The x-axis represents time, and the y-axis represents buses per half-hour. The Route 467 bus, which is a frequent express service running from 7:00 a.m. to 7:00 p.m., is shown in purple, while
the Route 67 bus, a parallel local all-day service, is shown in orange. The green area represents the combined headway, which in turn represents passengers’ actual experience of bus frequency. According to recent surveys conducted in June 2011, 2012, and 2013 along the corridor, about 18% of the corridor users indicated that they used both routes equally (49). As such, knowing the combined frequency by time period is more useful than knowing the schedules for each individual route. Planners may use such a figure to demonstrate and support transit projects (e.g., implementing new limited stop service).

CONCLUSION

Data visualization is a powerful tool for communicating with planners, decision makers, and the public and can take advantage of big data in the transit industry. The main objective of this paper was to move beyond the generation of internally focused performance measures that relate to transit agency operations (e.g., number of passengers per mile) and to instead introduce new performance measure visualizations that demonstrate general aspects of transit, helping planners to negotiate and support service-change projects. This research used a variety of common data sources that most transit agencies collect and provide. Nevertheless, planners could use whatever data are available by a particular agency, such as real-time AVL data instead of historical AVL data and GTFS data, to generate the same visualizations.

Performance measures were generated at the system, neighborhood, route, and stop levels and followed a hierarchical spatial approach. The system-level visualizations dealt with transit coverage and service spread, system improvement in speed, and waiting time at stops. The neighborhood-level visualizations covered the integration between night bus service and bicycle sharing, and the fluctuations of transit demand over the day. The route-level visualizations focused on route seating capacity, passenger activity by direction, and communicating the potential impacts of implementing stop consolidation. Finally, the stop-level visualizations looked at variation of bus stop demand and dwell time and at combined headways on parallel stops.
routes. Several of these figures can be used together to help planners visualize their plans at different spatial levels.

One important use of these big data tools is to demonstrate and test different ways of conveying and communicating positive messages about changes or potential changes in service to stakeholders and decision makers. Transit data visualization techniques are capable of displaying multiple transit performance measures and data sets simultaneously to illustrate important transit concepts to more general audiences with relative ease. Furthermore, the required data to create these images are readily available, as the data are already being automatically collected and archived. A good example of using a big data source (AVL and APC data) is demonstrated in the paper. Indeed, the value of an open framework for streaming big data being generated by transit agencies or other entities cannot be overstated. An extension of this study could be done to understand the potential of other data sources, such as smart card data, social media, and cellular data, in conveying positive messages about the transit service. Finally, efforts presented in this paper to incorporate new data visualization techniques will help facilitate communication between planners and increase public awareness of transit service attributes and improvements, thus helping transit planners to justify their decisions. This increased public awareness may in turn help to increase public involvement, satisfaction, and loyalty toward transit agencies.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Michel Tremblay, Anna Guinzo, and Sébastien Gagné of STM for providing the AVL and APC data. This research was funded by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES
