

# Improving Travel Information Systems with the Assistance of GPS

Udai Hassein, Kaarman Raahemifar, and Lisa Kadoury

**Abstract**—Travel delays and the ever increasing traffic queues negatively affect the everyday lives of commuters. Users of public transit systems, such as the TTC, are left unaware of the length of the delay they might experience on their journey. One of the most important factors in an Intelligent Transportation System (ITS) is the ability to predict the bus's arrival time accurately which not only increases the current customers' satisfaction, but also attracts new customers for the system. The objectives of this paper are (1) to predict the bus travel times using GPS data, and (2) to develop a parametric optimization model to improve bus arrival times based on the speed distribution. The bus travel time optimization model proposed in this research explicitly includes arrival time, dwell times, schedule adherence, and traffic congestion, all critical factors for accurate bus arrival times. In this paper, a test bus route was used in downtown Toronto, Canada. The optimization model performed significantly better than the historical data-based models. The proposed model identified the relationship between travel times and the independent variables, leading to superior results.

**Index Terms**—Optimization model, bus arrival time, road safety, public transportation.

## I. INTRODUCTION

The Advanced Traveler Information System (ATIS) is an important component of Intelligent Transportation Systems (ITS) [1]-[5] that provides accurate real-time information to the commuters. The system helps reduce the city traffic by encouraging the travelers to use public transit. To obtain the real-time information, an Automatic Vehicle Location (AVL) unit is used to track and monitor remote vehicles. All transit vehicles have an embedded AVL unit that enables the Global Positioning Systems (GPS) to monitor their current location, speed and direction.

The Toronto Transit Commission (TTC) operates just over 1800 buses that are equipped with a dead-reckoning (DR) system to locate buses in real-time. To locate a bus at any time, the departure location must be known. Knowing the departure location, the TTC uses gyroscopes and odometers to determine the bus's current location. The TTC currently uses DR systems and signposts to locate buses. However, these technologies are limited due to the drift caused by the DR system and the inability of the signpost system to continuously locate vehicles. Three main elements affect the

accuracy of the predicted bus travel time: bus arrival time, bus dwelling time and schedule adherence. Transit use is increasing and customer satisfaction can be maintained and improved by providing real-time traveler information [1]-[5].

This paper presents an optimization model, in which the bus travel time, speed, schedule adherence and arrival time at bus stops are considered as decision variables, for determining the minimum travel time based on the speed distribution. The following section describes the data collection effort for the predicted arrival times using GPS technologies.

## II. DATA COLLECTION AND ANALYSIS FOR BUS TRAVEL TIME

This section presents the experimental work, data collection, and the predicted bus arrival time. The input data used for this experiment are the following: arrival time, dwell time and schedule adherence for each bus stop. The methodology of data analysis is similar to the one used by Jeong (2004) for the best resolution [6]. For this experiment, bus route number 81N located within downtown Toronto was used. The bus departs Pape subway station and travels north. This bus route is comprised of six main stops. The bus schedule for this route is available online. The total distance of the route is 4.9 km, which was calculated using Google Maps, as illustrated in Table I and Fig. 1. The GPS data provides the following information: location, speed, time and direction.

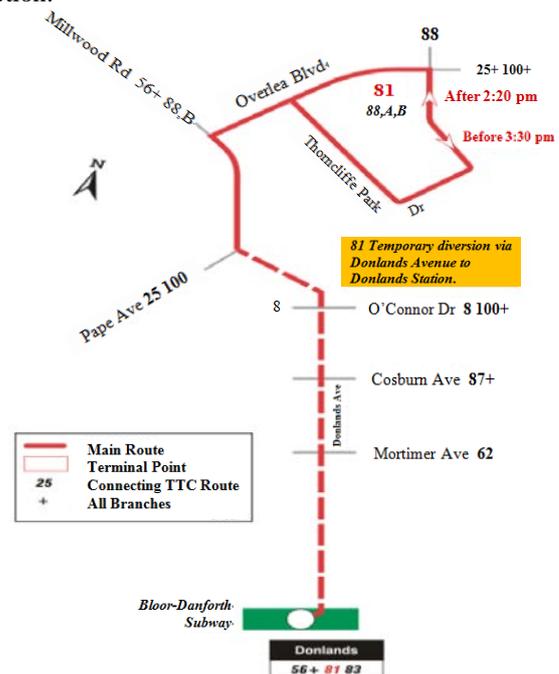


Fig. 1. Map of Route 81N (from TTC map) [7].

Manuscript received June 26, 2015; revised August 26, 2015.

Udai Hassein and Lisa Kadoury are with the Department of Civil Engineering Ryerson University, Toronto, Ontario, Canada (e-mail: uhassein@gmail.com).

Kaarman Raahemifar is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, Ontario, Canada (e-mail: kraahemi@ee.ryerson.ca).

### III. DATA COLLECTION PROCESS

Data collected was bus arrival and dwell times at the six main stops. The procedure was repeated 12 times. GPS information was calculated by using the GPS data logger. Bus schedule and route information were collected from the TTC website. The time period in which the different bus schedules run was collected. The distances between every two consecutive stops were calculated using Google Maps.

#### A. Arrival Time

The arrival time is the time at which a bus arrives at a bus stop. Furthermore, the arrival time at a bus stop is defined by the GPS time of the nearest location data from each bus stop, hence the shortest distance from the bus stop. Arrival time can be calculated as follows:

$$t_i = \frac{d}{v_{avg}} \quad (1)$$

where,  $t_i$  = travel time to the next bus stop  $i$ ;  $d$  = the distance from the bus's current location to the next bus stop;  $v_{avg}$  = average speed of the bus. Note that here it cannot use the current bus speed obtained from the GPS since the current bus speed could be zero due to many reasons such as the bus arriving at a bus stop or stopping at traffic signals. If the bus speed is zero, i.e.,  $v_{GPS} = 0$ , then the travel time would be infinity. To overcome this problem, the average speed of the bus is considered. The average speed is calculated using the historical data during similar periods and the current bus speed.

$$v_{avg} = \frac{v_{GPS} + v_a}{2} \quad (2)$$

where,  $v_{avg}$  = average bus speed;  $v_{GPS}$  = current bus speed obtained from the GPS;  $v_a$  = average bus speed from the historical data. The arrival time to the destination stop  $k$  is calculated using the equation below:

$$AT_i = t_i + \sum_{j=i+1}^k t_j \quad (3)$$

where,  $AT_i$  = arrival time at bus stop  $j$ ;  $t_j$  = link travel time. The time of the data  $i$  is the arrival time at bus stop  $j$ , and the time of the nearest data before arriving at the bus stop is considered the arrival time at stop  $j$ .

#### B. Dwell Time

Dwell time is defined as the total time that a bus remains stationary at a bus stop. This will vary depending on many factors, such as the number and types of passengers getting on and off the bus, the traffic lights, schedule adherence and the driver's overall behavior. It can calculate the dwell time at each bus stop by taking the difference between the arrival time and the departure time:

$$WT_i = DT_i - AT_i \quad (4)$$

where,  $WT_i$  = dwell time at stop  $j$ ;  $DT_i$  = departure time from stop  $j$ ;  $AT_i$  = arrival time at stop  $j$ . Departure time is defined as the time when a bus starts moving from a stationary position and therefore, its speed exceeds from zero. When a bus passes a bus stop without stopping, the dwell time is zero. Fig. 2 shows the dwell time at each bus stop. There are many bus stops in downtown Toronto which means the distance between the two consecutive ones is relatively short. When

the bus stop is very close to an intersection and the traffic light at an intersection is red, the bus usually remains at the bus stop. The traffic light data for route 81N was not readily available; therefore, it could not distinguish the dwell times due to the red lights at the intersections. Therefore, in this study the duration of time a bus spent at each bus stop was considered as the dwell time, regardless of whether it was due to encounter with a red light or not.

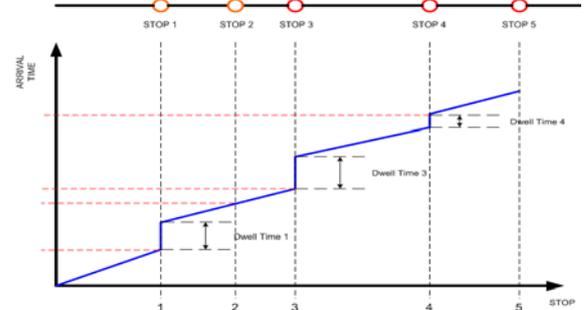


Fig. 2. Calculating arrival time/dwell time at bus stop.

#### C. Schedule Adherence

Schedule adherence refers to the difference in the scheduled and actual arrival time:

$$ST_i = AT_i - PT_i \quad (5)$$

where,  $ST_i$  = schedule adherence at bus stop  $j$ ;  $AT_i$  = actual arrival time at bus stop  $j$ ;  $PT_i$  = scheduled (TTC) arrival time at bus stop  $j$ . A negative difference indicates that the bus arrived early, whereas a positive difference indicates late arrival. It is necessary to add schedule adherence as an input variable to the prediction model [6].

### IV. DATA OBSERVATION

Twelve sets of data were collected for this experiment. Fig. 3(a) shows the arrival times for each stop on route 81N. As explained before, the variability in the data is due to the traffic congestion, traffic signals, and passenger demand. Furthermore, the data was not collected at the same time of the day, which further contributes to its variability. The traffic is heavier at certain times of the day further delaying the arrival of the bus. Likewise, there are more passengers at certain times causing further delays. Fig. 3(b) shows the dwell time variability for each of the twelve sets of data. According to the graph, dwell time is higher for bus stop 1, the initial bus stop located at Pape subway station. This delay is caused by passenger demand, bus driver behavior, and traffic congestion at the subway station.

Bus stop 4 also shows higher variability in dwell time than the other stops. This is because bus stop 4 is a control stop at which the bus driver must pace himself. If ahead of schedule, the bus must remain at the control stop (bus stop 4) for as long as needed. On the other hand, if the bus is behind its schedule, the driver can pass by the control stop without actually stopping, unless there are passengers getting off/on the bus. Fig. 3(c) displays the schedule adherence for the same twelve sets of data. As previously mentioned, a negative value in schedule adherence means that the bus arrived early at the specified stop, whereas a positive value means the bus arrived late.

Looking at the graph, for most of the data sets, bus stop 1 has a negative value meaning that most of the times the bus arrived at this stop earlier than it was scheduled. Bus stop 2, on the other hand portrays a positive value for schedule adherence for almost all data sets, meaning that the bus usually arrived late at this stop. Furthermore, the relatively

large negative and positive schedule adherence values for bus stop 6 suggest that practically no bus arrived at this stop at its scheduled time: buses either arrived later or earlier than predicted. This clearly shows that an improvement in predicting and scheduling bus arrival time is required to further satisfy transit users, operators, and bus drivers.

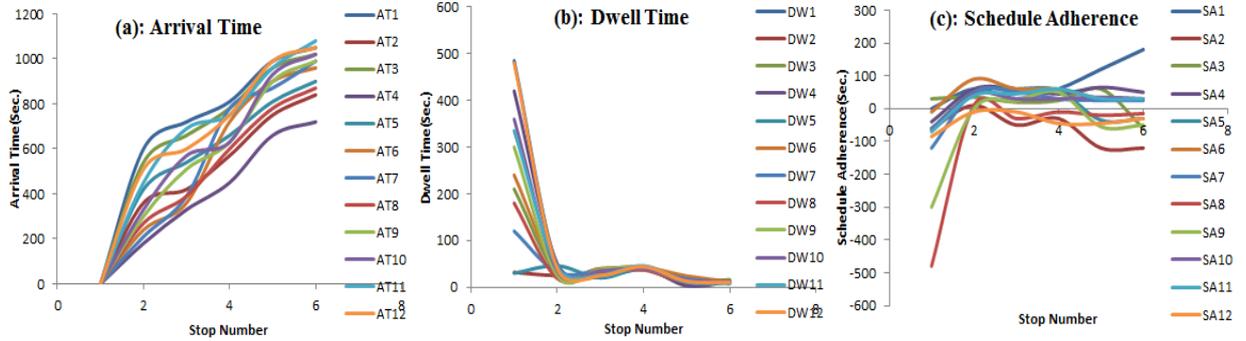


Fig. 3. Data observation (a) Arrival times; (b) Dwell time; (c) Schedule adherence.

## V. DEVELOPMENT OF NEW MODEL FOR SPEED DISTRIBUTION

This section presents the development of a new parametric optimization model for speed distribution. The optimization model for the speed of the bus is a mathematical procedure consisting of three parts: objective function, decision variables, and constraints. This is accomplished using the speed of a bus obtained by GPS, which minimizes travel time. Pertaining to the parametric model, Lingo is one of the programming systems utilized to resolve decision variables subjected to assured constraints for a specified objective function. The objective function minimizes travel time based on the bus's speed obtained by GPS data, which affects safety as well as performance in terms of speed distribution. The decision variables are represented with four variables and a subscript  $i$  ( $t_i$ ,  $V_i$ ,  $AT_i$ , and  $TS_i$ ) to satisfy the constraint.

## VI. PARAMETRIC OPTIMIZATION MODEL

The Parametric Optimization Model of the speed variation is to improve the bus arrival time. Given that the variance within the decision variables is linear, it is possible to include it within the objective function. As an alternative, minimizing the bus travel time could minimize the variance. Thus, the objective function can be displayed as:

$$\text{Minimize } Z = t \quad (6)$$

where,  $Z$  = objective function and  $t$  = minimum travel time. The travel time ( $t_i$ ), average speed ( $V_{avg_i}$ ), arrival time ( $AT_i$ ), dwell time ( $WT_i$ ), and schedule adherence ( $ST_i$ ) are calculated by:

$$t_i = \left( \frac{d_i}{\frac{1000}{V_{avg_i}}} \right), \forall_i \quad (7)$$

$$V_{avg_i} = \frac{V_i + V_{GPS_i}}{2}, \forall_i \quad (8)$$

$$AT_i = t_i + TS_i, \forall_i \quad (9)$$

$$WT_i = DT_i - AT_i, \forall_i \quad (10)$$

$$ST_i = AT_i - PT_i, \forall_i \quad (11)$$

The constraints related within the objective function are:

$$t_i \leq t_{max_i}, \forall_i \quad (12)$$

$$t_{max_i} \geq 0; \forall_i \quad (13)$$

$$AT_i \geq 0; \forall_i \quad (14)$$

$$V_i \geq 30; \forall_i \quad (15)$$

$$V_i \leq 50, \forall_i \quad (16)$$

$$V_i - V_{avg_i} \leq 10, \forall_i \quad (17)$$

where  $t_i$  = travel time for bus stop group  $i$  calculated by Eq. (7);  $TS_i$  = travel time between two consecutive bus stops;  $t_{max_i}$  = the maximum travel time allowed for bus stop group  $i$ ; and  $V_i$  = speed for bus stop group  $i$ . Eq. (15) and (16) show the lower and upper speed limits for bus stop group  $i$ . The objective function in Eq. (6) and the constraints in Eq. (12)-(17) symbolize a linear optimization model. The model contains 81 constraints and 4 decision variables categorized by the following four variables:  $t_i$ ,  $V_i$ ,  $AT_i$ , and  $TS_i$ . The linear optimization model can be solved using the existing optimization software like Lingo.

## VII. APPLICATIONS AND RESULTS

The purpose of this section is to apply the collected data to the proposed model. When designing suitable bus travel times, the distribution of speed rates is critical. When evaluating the travel time of a bus, the GPS speed of the bus is applied to determine the best speed distribution design as judged by the arrival time accuracy.

### Input Data

Bus schedules and route information were collected from the TTC website as shown in Table I.

TABLE I: INPUT DATA FOR THE OPTIMIZATION MODEL

Bus Stop	d (m)	V_GPS (km/hr)	WT (min)	PT (min)	t_max (min)	Q
1	0	13.07	8.033	0	3	1
2	730	21.74	0.75	9	12	1
3	440	20.60	0.417	11.5	14.5	1
4	450	14.79	0.633	12.5	15.5	1
5	2150	17.52	0.25	14.5	17.5	1
6	1220	16.14	0.133	15.5	18.5	1

The 2<sup>nd</sup> column shows the distances ( $d$ ) between the two consecutive bus stops. The 3<sup>rd</sup> column shows the GPS speeds ( $V\_GPS$ ) (taken from the GPS data logger). Column four displays the experimentally calculated dwell times ( $WT$ ). The fifth column displays the predicted time ( $PT$ ) taken from the TTC schedule. Column six shows the maximum time ( $t\_max$ ) a bus can be delayed at each stop. The  $t\_max$  is assumed to be at a maximum of 3 minutes for each bus stop. Column 7 shows the quantity of buses ( $q$ ). In this example, it only assumed one bus for quantity.

#### A. Numerical Example (Parametric Optimization Model)

Results of applying the optimization model to the input data in Table I are shown in Table II. Using the Lingo software, based on the local optimal solution, the objective function value was obtained as 0, and the parameters of the optimization model determined that  $t = 0, 1.577, 0.863, 1.089, 4.688, 2.800$  min,  $TS = 0, 7.423, 10.637, 11.411, 9.812, 12.700$  min and  $AT = 0, 9, 11.5, 12.5, 14.5, 15.5$  min. For stop number 5, the  $t$  value was 4.688 min. This is due to the fact that the distance between stops 4 and 5 is the longest (2150 m). The  $t$  value for stop 6 is also large ( $t = 2.800$  min) due to the large distance of 1220 m between stops 5 and 6.

Additionally, for stop number 3, the high  $V$  value was

40.601 km/hr due to the low traffic congestion between stop 2 and stop 3. For stop number 5, the  $V$  value was 37.517 km/hr due to the large travel time between stops 4 and 5. Stop 6 has a  $V$  value at 36.143 km/hr. This is also due to the fact that there is a large distance of 1220 m between stop 5 and stop 6. The range of the speed of the bus is from 30 to 50 km/h. The optimization model presented another solution that is better than the traditional method. The speed  $V_i$  and travel time  $t_i$  for the Parametric optimization model is illustrated in Table II. The optimization model minimizes the bus arrival time for a specific segment.

#### B. Sensitivity Analysis

Arrival time is the combination of  $t$  (link travel time) and  $TS$  (travel time between two consecutive bus stops) (Fig. 4.a). The difference between the arrival and departure times for each bus stop is very small since the link travel time is very small based on the model improvement (Fig. 4.b) meaning that the bus follows the schedule accurately. Fig. 4.c demonstrates schedule adherence, i.e., the difference between the predicted and actual arrival times. The figure shows that the schedule adherence values are equal to zero, which means there were no delays since the optimization model improved the arrival time. The model improved the arrival time by using speed distribution between each two consecutive bus stops. The model makes the TTC bus schedules more reliable which leads to happier commuters.

This study provides a parametric optimization model which improves the bus travel time using GPS data. The results demonstrated here appear in relation to the linear distribution. Table II shows the parametric model, where the model improves  $t$  travel time that is dependent upon the  $V$  distribution. The parametric model is a better method to calculate  $AT$  arrival time (see Table II), and it is demonstrated in Fig. 4(a), Fig. 4(b), and Fig. 4(c).

TABLE II: RESULT OF THE PARAMETRIC OPTIMIZATION MODEL

Bus Stop	V	V_GPS	V_avg	T	TS	AT	WT	DT	PT	ST
1	30.000	13.07	21.534	0.000	0.000	0.000	8.033	8.033	0.000	0.000
2	33.807	21.74	27.774	1.577	7.423	9.000	0.750	9.750	9.000	0.000
3	40.601	20.60	30.601	0.863	10.637	11.500	0.417	11.917	11.500	0.000
4	34.794	14.79	24.794	1.089	11.411	12.500	0.633	13.133	12.500	0.000
5	37.517	17.52	27.517	4.688	9.812	14.500	0.250	14.750	14.500	0.000
6	36.143	16.14	26.143	2.800	12.700	15.500	0.133	15.633	15.500	0.000

$V\_GPS, V\_avg = \text{km/hr}$   
 $t, TS, AT, WT, DT, PT, ST = \text{min}$

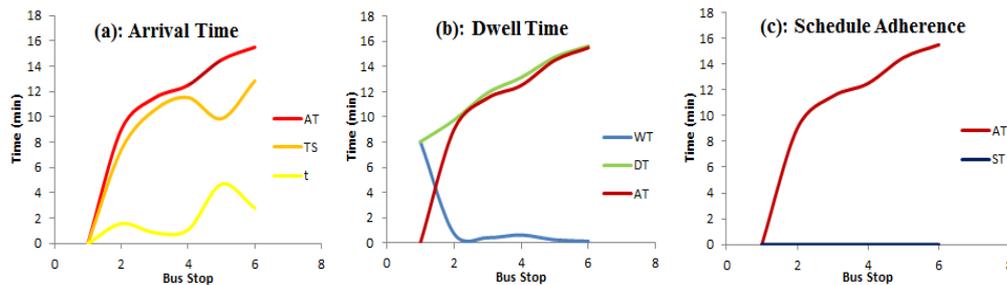


Fig. 4. Optimization model: (a) Arrival time; (b) Dwell time; (c) Schedule adherence.

### VIII. CONCLUSIONS

This paper has presented the arrival time, dwelling time and schedule adherence of TTC buses in downtown Toronto by collecting data on site. Schedule adherence data showed that none of the buses arrived at their scheduled time. From this, an optimization model was developed to improve bus travel time based on the bus speed distribution. The proposed model finds the best bus speed distribution for a specified GPS speed data by utilizing the complete speed design area. The developed parametric optimization model determines the best distribution of the bus's speed, where such an optimization model has an objective function subject to constraints. The parametric optimization model equations for speed distribution are simple to apply and provide speed rates that are relatively comparable to the speed limited.

In summary, the parametric optimization model provides advantages such as determining the speed rate that minimizes the objective function from one bus stop to another, and increasing its function in travel time design. Optimization is clearly a powerful and beneficial tool, and the method presented in this paper can help improve bus travel times. Recommendations for future analysis and numerical research are (1) to improve operational reliability conditions that can be analyzed using a reliability index as travel time; and (2) to develop the mathematical example utilizing more data.

### ACKNOWLEDGMENT

The authors are grateful to Bassim Ibrahim for his assistance in the thorough data collection and have encouraged ideas that appear in the paper. Special thanks to Salma Taqwfiq for her professionalism, wisdom, experience, suggestions, guidance, encouragement for the research.

### REFERENCES

- [1] BRT (Bus Rapid Transit). (2003). BRT Reference Guide. *Marketing, Information, and AVL, Passenger Information*. [Online]. Available: <http://www.fta.dot.gov/brt/guide/info.html>
- [2] R. F. Casey, L. N. Labell, E. J. Carpenter, J. A. LoVecchio, and Lo. Moniz, *Advanced Public Transportation Systems*, Publication FTA-MA-26-7007-98-1, Federal Transit Administration, U.S. Department of Transportation, 1998.
- [3] R. F. Casey, L. N. Labell, L. Moniz, J. W. Royal, and M. Sheehan, *Advanced Public Transportation Systems*, Publication FTA-MA-26-7007-00-1, Federal Transit Administration, U.S. Department of Transportation, 2000.
- [4] J. J. Fruin, *Passenger Information Systems for Transit, Transfer Facilities, Synthesis of Transit Practice 7*, Transportation Research Board, National Research Council, Washington, D.C. Institute for Transportation Research and Education, 1985.
- [5] T. B. Reed, "Reduction in the burden of waiting for public transit due to real-time schedule information: A conjoint analysis study," in *Proc. the 6th International Vehicle Navigation & Information Systems*, Seattle, Washington, 1995.
- [6] R. H. Jeong, "The prediction of bus arrival time using automatic vehicle location Systems data," Doctoral dissertation, Texas A&M University, 2004.
- [7] TTC website. [Online]. Available: [http://www3.ttc.ca/About\\_the\\_TTC/Operating\\_Statistics/2010.jsp](http://www3.ttc.ca/About_the_TTC/Operating_Statistics/2010.jsp)

**Udai Hassein** is a Ph.D. Candidate at the Department of Civil Engineering, Ryerson University, Toronto, Ontario, He earned his M.A.Sc. from Ryerson University and has published many journal & conference papers since the start of his doctoral studies in 2011. His interest lies in the area of highway geometric design and road safety.

**Kaarman Raahemifar** is a professor in the Dept. of Electrical and Computer Engineering, Ryerson University, Toronto, Ontario. He earned his PhD, in Electrical and Computer Engineering from Windsor University, in 1999. He earned his MASc, in electrical and computer engineering from Waterloo University in 1993 and has published many journal & conference papers.

**Lisa Kadoury** earned her M.A.Sc. in civil engineering from Ryerson University in 2014. Her interest lies in the area of transportation.