

Quantitatively Understanding Transit Behavior from the Rider's Point of View

Colin Bick

Abstract

Arrival time uncertainty is a major source of negative perception by riders, yet how this uncertainty manifests in the rider's experience is not well-studied. While operators constantly make efforts to improve reliability, and real-time arrival predictions reduce uncertainty for riders in transit, it is also possible to lessen frustration by better informing riders of system behavior beforehand. This work introduces a new method for understanding transit behavior through an analysis of historical arrival time data from San Francisco. The results identify impacts of timeliness on rider experience, such as that average wait time is minimized by showing up five minutes early, or that a five-minute transfer window will be successful 80 percent of the time. Categories of rider experience also are discovered, such as between daytime and evening users. More importantly, it is demonstrated how operators and trip planners can make use of this method to improve rider experience.

Introduction

Of the many obstacles faced when using public transportation, one of the largest is the information barrier. Transit schedule and route information add several dimensions to what is already contained in a standard road map. The agencies responsible for distributing the information rarely are successful at communicating what one

might call the intent of their system; in other words, it often is difficult to learn how a system is meant to be used.

Paper schedules that display route details and timetables are one of the most fundamental formats of transit information, yet using them to plan a trip is far from a simple task. Cain (2007), for example, found that 47.5 percent of sampled riders were unable to correctly plan a trip using only a system map and paper schedules. Some studies examine ways to improve presentation of the information, but rarely are there efforts that find new information to present. Even after significant improvements in presentation, one study (Sollohub et al. 2006) was unable to determine if improvements were effective since the riders generally were unable to correctly use them.

An online trip planner is a tool that improves on a schedule by providing a rider with step-by-step directions from one point to another for a given day and time. The trip planner tries to construct optimal answers using the same information provided in a paper schedule and decides on a priori costs for transfers, waits, walks, etc., before a well-defined optimization problem is possible, not to mention useful (Modesti et al. 1998, Sherali et al. 2006). An especially important parameter is that of minimum transfer time, defining a safety margin of time between vehicles to ensure the feasibility of a transfer.

What these tools do not take into account is that transit behavior is inherently random. The arrival times provided in a schedule are estimates only, making their use in trip planning much less straightforward. A rider must therefore draw heavily on experience and make guesses when creating and evaluating trip plans, weighing perceived risks against convenience. Online trip planners, similarly, must choose a minimum transfer time as a heuristic meant to compensate for nondeterministic behavior, and, consequently, the possibility of making shorter transfers is ignored.

Riders, on the other hand, are all too aware of the uncertainty involved in taking the bus. The uncertainty of arrival time is a major source of frustration (Caulfield et al. 2009), and the perception of time spent waiting or riding increases disproportionately whenever the actual time spent is longer than expected (Li 2003). Even the uncertainty of arrival time itself increases the perceived amount of time passing (Mishalani et al. 2006), and without any sources of information beyond schedules or schedule-based tools, the rider does not know what to expect. While these effects on perception and frustration are understood, precisely how the events in using transit are impacted by timeliness (or the lack thereof) is not well-studied.

The next bus predictor NextBus was developed to decrease this uncertainty. Approached in many manners, this tool uses real-time data to predict the time of arrival of the next bus of a given route, in a given direction, at a given stop. To make these predictions with reasonable accuracy, machine learning techniques are trained against historical data (see, for example, Chien et al. 2002; Jeong 2004; Shalaby et al. 2004; Wall et al. 1999). The value and utility of this tool has been investigated several times, due in large part to the costs involved; it is understandable that not every agency has invested in it.

While the next bus predictor reduces uncertainty during a trip (i.e., while at the bus stop), it does not help the rider anticipate behavior beforehand. This uncertainty—the uncertainty of transit behavior in general—has yet to be addressed. Riders and trip planners constantly are forced to make guesses that try to remediate the disagreement between scheduled and actual arrival times. Any treatment of transit as a discrete or deterministic process is bound to result in errors and frustration.

If transit behavior is approached mathematically and as a random process, on the other hand, not only can the guesswork in trip planning be removed, but the entire process of using public transportation can be made as exact and predictable as possible. The tradeoff is the added complexity of treating many aspects of riding the bus as probability distributions, including time spent waiting, actual arrival times, trip durations, and vehicle transfers. These distributions naturally all reduce to some combination of actual arrival times. Thus, by viewing actual arrival times in relation to the schedule—that is, how late the bus is—the true nature of transit behavior can be understood, and therefore communicated, from the rider's point of view.

This research introduces a new method for understanding transit behavior by modeling vehicle lateness as a random variable. An analysis is performed on data from four weeks of system-wide arrivals from San Francisco. The data were calculated by matching archived GPS data provided through NextBus with the corresponding schedule provided in GTFS (General Transit Feed Specification) format.

A similar work was performed by Berkow et al. (2009), but from the operator's perspective. They demonstrated an approach that provides far greater insight into a system's performance than the generation of performance measures, applying statistical as well as visual tools to a year's worth of data recorded by Portland's TriMet. Properties of performance-related random behaviors such as those exhibited by passenger boardings, lift use, overall ridership, vehicle headway, and lateness were investigated at several resolutions. Importantly, they noted that the large size

of the dataset allowed the entire analysis to be done without making estimates or assumptions.

This work focuses on the rider's perspective. It identifies the quantitative impacts of lateness on rider experience, including waiting for the bus, making transfers, and overall trip time. It then discovers and explores categories of rider experience, showing how different groups of users are impacted by different behavior. Finally, it is demonstrated how the method and results can be used by operators and tools such as trip planners to provide the rider with better information.

Model and Data

Description of Model

Conceptually, the idea of lateness as a random variable is fairly straightforward, if one imagines a rider waiting for the bus with a schedule in hand. Addressing lateness as a property of the entire system requires a more careful definition. This study models lateness as measuring

$$\text{lateness} = t_{\text{actual}} - t_{\text{scheduled}}$$

where t_{actual} and $t_{\text{scheduled}}$ are the actual and scheduled arrival (or departure) times for a scheduled stop chosen at random, with uniform probability, from the population of all scheduled stops during the period of data collection. To make the results representative of a weekly schedule, a period of data collection was chosen such that its duration is precisely four weeks, without holidays.

Additionally, the study makes a distinction between lateness computed using arrival times and departure times. Each computation below states whether it is referring to arrivals or departures. When using arrival times, the first stop of each run is excluded from the population; similarly, when using departure times, the last stop of each run is excluded. This is done to make the model more meaningful.

Description of Data

San Francisco, by releasing both schedule and real-time data to the public, offers a valuable opportunity to explore this approach in a dense and heterogeneous transit system. San Francisco Municipal Rail (Muni) transports 200 million passengers per year inside an area of 47 square miles, employing bus, light rail, cable car, and a historic street car (San Francisco 2011). GPS tracking data collected every minute from Muni for a period of four weeks starting in March 2009 was matched to

schedules described in a GTFS (General Transit Feed Specification) format for that month (see references for data sources). This match-up describes the actual arrival and departure times of each vehicle at each stop along its assigned route, as well as the scheduled arrival time. Much information about the routes is available from the GTFS data, including stop location, route shape, and vehicle type. The software written to perform the match-up has been made available online at <http://cbick.github.com/gps2gtfs>.

Of the approximately 14 million rows of raw GPS data, only around 40 percent were qualified to survive preprocessing. While equipment errors causing this may have been uniform, any human influence (for example, new drivers or incorrect use of tracking equipment) in creating bad data cannot be assumed uniformly distributed. For this reason, the data collected in some areas may be much less populated than in others. Additionally, the match-up of GPS data may be distorted by geographical (or geological) influences. For example, a visualization of the data shows that GPS signals neighboring bodies of water are an order of magnitude more erroneous than those between buildings. For this reason, data may be less accurate for certain stops than for others.

The GTFS data are far from perfect as well. Stops in close proximity often are given the same scheduled time, decreasing specificity in the meaning of the schedule in general. In other instances, detours can cause a vehicle to shift its schedule or miss designated stops entirely.

Treatment of Data

Computation of Lateness

To compute the arrival and departure time estimations, the GPS data points were projected onto the corresponding route's path and the times were interpolated. The arrival and departure of a vehicle at a stop were defined to be the points at which the vehicle entered and exited a 25-meter radius around the stop.

Outliers and Errors

Processing raw GPS data and matching it up to schedule data introduce many opportunities for error. Examining the data shows that 98 percent of the weighted distribution falls between 8 minutes early and 20 minutes late, but the remaining 2 percent ranges from 35 minutes early to 90 minutes late. While these outer bounds of lateness are suspiciously large, the following steps were taken to eliminate errors:

- Only vehicles whose signal existed for and matched a significant portion (>50 percent) of their designated route paths were considered admissible.
- All vehicles were put through a secondary processing to ensure there were no better matches in the schedule.

These steps establish confidence that there is an insignificant level of error in the matching process. Any errors in the GTFS data are impractical to eliminate, so its correctness must be assumed. Errors resulting as artifacts from abnormal events or incorrect GPS data cannot be eliminated using any consistent method. Therefore, any apparent outliers cannot be discarded and should be included in the distribution.

Sample Bias and Correction

From the nature of the data collection, it is clear that the data set is not representative of the population in the model. The observations, therefore, have been weighted and normalized to provide representative estimates of the population as defined by the schedule. In particular, for any particular scheduled stop, its weight w is defined as

$$w = \frac{f_{\text{scheduled}}}{f_{\text{observed}}}$$

where $f_{\text{scheduled}}$ is the number of occurrences of the stop according to the model's population, and f_{observed} is the number of actual observations that were made for that stop.

Independence

As the model is constructed, observations in the data are not independent. Specifically, arrival times within the same run are highly correlated. While it may be reasonable to assume that lateness samples from different vehicle trips are (sufficiently) independent, several of the confidence interval widths shown below are potentially underestimated. Since calculating more accurate intervals would add computational complexity without contributing to the discussion, such efforts have not been made.

It is necessary, however, to at least provide a reasonable estimate of the true error bounds. For this, the reader is directed to examine the confidence intervals in Figure 10. These represent a meaningful upper bound on width, as each population in that figure consists of independent samples.

Analysis

The processing described above resulted in a little over 2 million recorded observations of lateness. With a large dataset such as this, rather than try to fit some analytical model, it is much more straightforward to simply compute an empirical cumulative distribution function (ECDF). From an ECDF, one can quickly compute quantiles and their confidence intervals, as well as take computationally simpler approaches to simulation. The ECDFs and their confidence intervals in this work were computed using Horvitz-Thompson estimates as described in, e.g., Diaz-Ramos et al. (1996).

This section proceeds by examining first some overall properties of lateness as a random variable, then a series of comparisons of hand-picked partitionings. In each case, it is identified how the extracted information is useful to a bus rider.

Overall Behavior of Lateness

Overall ECDF of Lateness

The overall (weighted) ECDFs of arrival lateness and departure lateness are shown in Figure 1. They are nearly identical. The 95 percent confidence interval for each ECDF is too small to discern in the figure, each having a width of only about 0.25 percent. The estimated means of lateness, the positions estimated by the schedule (i.e., that of lateness = 0), and the 5 percent quantiles have also been marked.

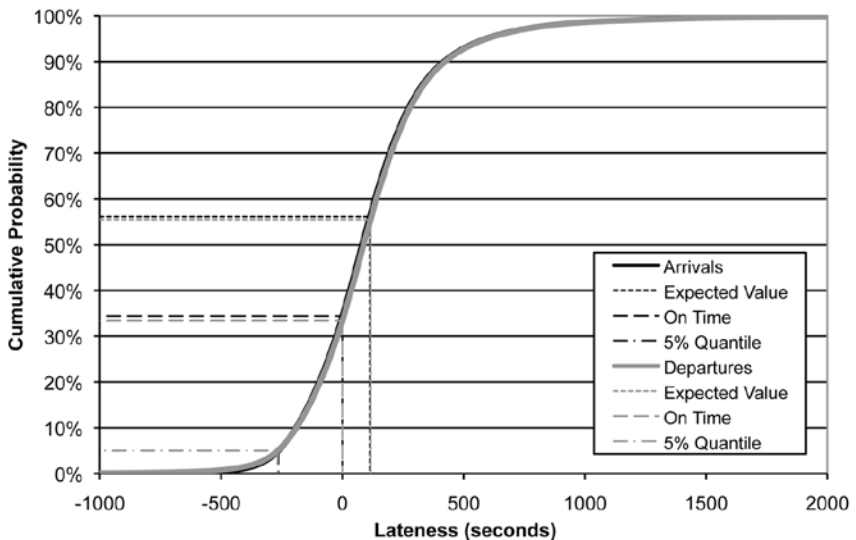


Figure 1. ECDF of lateness

From this plot, it is seen that the schedule is approximately equal to the 33 percent quantile for lateness of departure—in other words, a rider showing up (spontaneously) at a bus stop at precisely the scheduled time has only a 67 percent chance of catching that bus. To have a 95 percent chance of catching the bus, the rider must arrive 267 seconds, or about 5 minutes, early.

The 95 percent confidence interval for the average lateness of arrivals is between 113 and 114 seconds, while that for departures is between 116 and 117 seconds; that is, the bus is about 2 minutes late on average. The ECDFs also show that the rider can expect to arrive (or depart) 5 minutes or more past the scheduled arrival around 18 percent of the time; 10 minutes or more 5 percent; and 15 minutes or more 2 percent. Such knowledge can help the rider not only to better plan a trip, but also to experience less frustration since the behavior is now in a sense predictable.

Average Waiting Time

An interesting metric to consider at this point is that of expected waiting time, or how long a rider can expect to wait on average for a vehicle given the time he or she arrives at a stop relative to the schedule. To compute this, a simulation was constructed with scheduled arrivals occurring at a constant frequency of arrival (headway). The actual arrivals were simulated by sampling from the empirical distribution of departure lateness, implicitly assuming that the distribution does not change significantly for different headway values. The resulting values for expected wait time are depicted in Figure 2, which plots the average wait time as a function of the passenger's arrival time at the stop for various headways. The figure shows that the minimum average wait time is achieved by the rider arriving about five minutes early, with slightly earlier minimums for larger headways since the cost of missing the bus is higher.

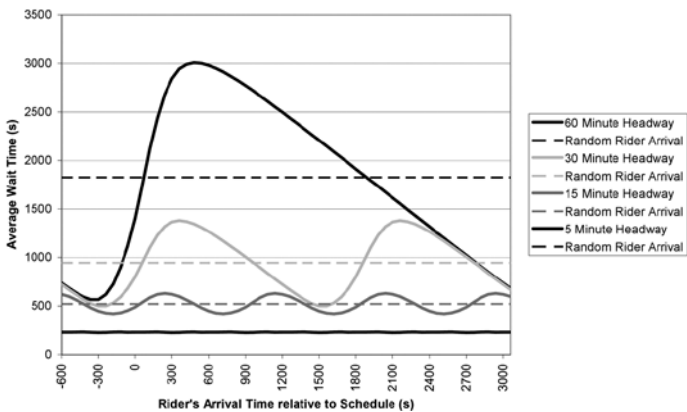


Figure 2. Average wait time

Figure 2 also shows a dashed line for each headway, marking the average wait time for a rider who arrives at the stop randomly without consulting the schedule. It is interesting to see that a rider arriving “on time” has approximately the same expected wait time as one arriving randomly. For the shorter headways, there is almost no practical difference between arriving randomly and consulting the schedule; it is clear from this why the behavior observed by Balcombe and Vance (1998, cited by Cain 2007) might occur, where riders arrive randomly for headways of 10 minutes or better, but consult the schedule for those of 15 minutes or more.

Making a Transfer

One of the critical determinations that must be made when planning a trip on public transportation is the feasibility of a transfer between two vehicles. Riders have only their experience to draw on, and software trip planners simply use a threshold minimum transfer time. Using the lateness data, however, it is simple to construct a simulation of a rider transferring from one vehicle (sampling from the arrival distribution) to another (sampling from the departure distribution), producing the informative plot displayed in Figure 3. Here, the likelihood of making a transfer—that is, the likelihood that the second vehicle will depart from the stop after the first vehicle arrives—is plotted against different transfer window sizes. The 90 percent mark is not reached until there is a 7-minute window between scheduled arrivals, and to be 95 percent sure of making the transfer requires a window of 10 minutes. This chart enables even a new rider to make an informed decision about the acceptability of a particular trip plan.



Figure 3. Overall probability of making a transfer

An element that is missing from this model is the concept of a designated transfer point, where transfers can be coordinated between vehicles (meaning that the second vehicle will wait for the first), or the transfer window is elongated by scheduling the departure of the bus to be some minutes later than its arrival. This element is missing, too, from the SF Muni GTFS data. We can hypothesize that the probability of making a transfer increases at these transfer points, but no empirical observations can be made.

Hour of Weekday Comparisons

This first set of comparisons examines the behavior of lateness exclusively on weekdays, partitioned according to which hour of the day an arrival is scheduled. This is determined strictly according to the clock hour of the scheduled arrival, meaning, for example, that a scheduled arrival of 7:59 a.m. belongs to the 7 a.m. category.

This kind of comparison is useful because it reveals categories of user experience. Riders using transit mainly to commute during rush hour will observe a different class of behavior than those who use transit in the evening or at night. As different categories of riders have different intents in their use of transit, it is important to understand the behavior of transit from each perspective.

Hour of Weekday ECDFs

The lateness ECDFs for the 8 a.m., 5 p.m., 8 p.m. and 1 a.m. blocks were selected for visual comparison as displayed in Figure 4. The overall ECDF is included as well. In all cases only the arrival lateness data was used.

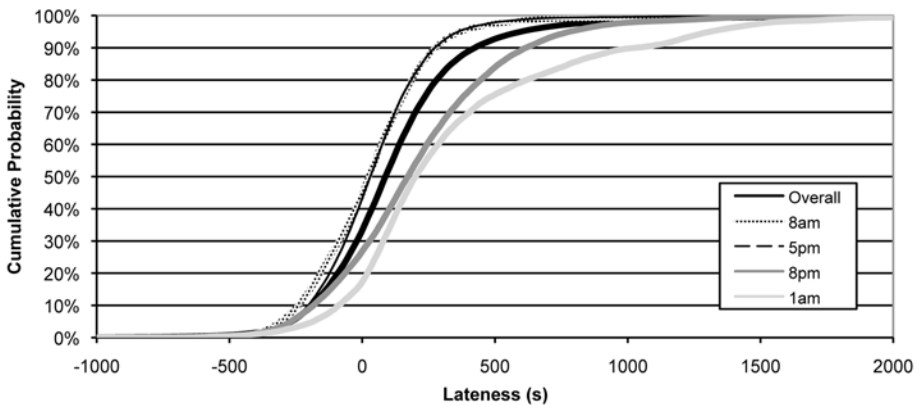


Figure 4. Hour of weekday ECDF comparisons

From these curves it can be seen that the 8 a.m. and 5 p.m. hours tend to be earlier than overall and have a lighter tail. The 8 p.m. hour shows the opposite, being considerably later than the overall and having a much heavier tail. The 1 a.m. behavior is the most distinct of the group, being quite late and wearing a long tail. It also has a wider 95 percent confidence interval (not shown), ± 1.4 percent compared to ± 0.6 - 0.8 percent for the others. Approximately 20 percent of the 1 a.m. data is more than 10 minutes late.

These differences in behavior during different phases of the day are quite important to the rider. Commuters are interested specifically in the rush hour behaviors. Late-night users should know that the behavior is significantly later and less reliable than during the daytime. The quantiles for late night behavior are especially useful when taking the last bus home. By making further analysis, it can be seen how the differences in behavior should influence the rider's use of the system.

Hour of Weekday Average Waiting Time

As in the overall case shown in Figure 2, Figure 5 plots average wait time as a function of the rider's arrival time at the stop. Only the departure data was used. For ease of comparison, the plot shows curves for the 5 p.m. and 1 a.m. hours only, alongside the overall lateness, and considers only a 30-minute headway. The 5 p.m. data show a lower minimum average wait than the overall for a rider arriving five minutes early and has a higher maximum average wait as well. These are due to the increased timeliness of that hour. The 1 a.m. data's minimum average wait, in contrast, are considerably higher than the other curves and occur where the rider arrives only two minutes early, as expected from the increased lateness of the hour (this assumes another bus is coming). It is interesting that the heavier tail in the 1 a.m. distribution actually causes the maximum average wait to be lower, since a rider arriving late is less likely to have missed the bus; but, of course, the variance of the wait time is much higher, though this is not shown in the figure.

There are several important messages for the rider in this plot. For one, a rider can expect to wait a long time for the bus at 1 a.m., which may be uncomfortable and possibly even dangerous. Rush hour riders, on the other hand, can expect short and pleasant waits—provided they arrive five minutes early.

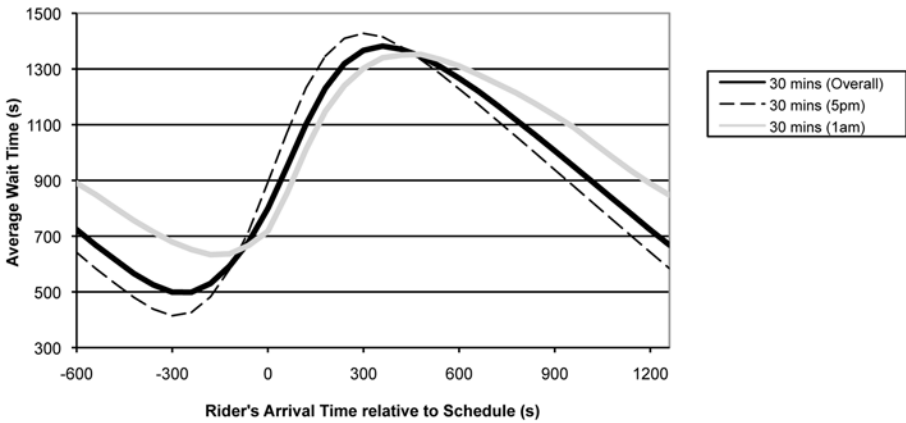


Figure 5. Average wait time comparison for hour of weekday

Hour of Weekday Probability of Transfer

We now inspect the differences the time of day on a weekday can make in the feasibility of a transfer. Figure 6 compares the probability of making a transfer for selected hours of arrival along with the overall probability. Note that the curves intersect near the 50 percent point where the transfer window is 0. This is an effect of using near-identical distributions for each transfer (e.g., the 8 a.m. line represents transferring from an 8 a.m. arrival to an 8 a.m. departure). As expected, the increased timeliness of the 8 a.m. and 5 p.m. data results in lower probability than the overall for negative transfer windows and in higher probability for positive windows. The opposite applies to the 8 p.m. and 1 a.m. hours. In fact, the 90 percent point is reached with windows as small as six minutes for the 5 p.m. data and as large as 15 minutes for the 1 a.m. data.

It is helpful to the commuter to know that transfers are easier to make during the rush hours. It is even more important, however, for the timeliness-concerned rider to note the probabilistic behavior: given (for example) a commute with a scheduled transfer window of five minutes during the 8 a.m. hour, the rider will make that transfer approximately 85 percent of the time—that is, the transfer will be missed 3 out of 20 times. By making this behavior predictable, the rider can expect and plan for its occurrence, removing much of the frustration associated with transfers.

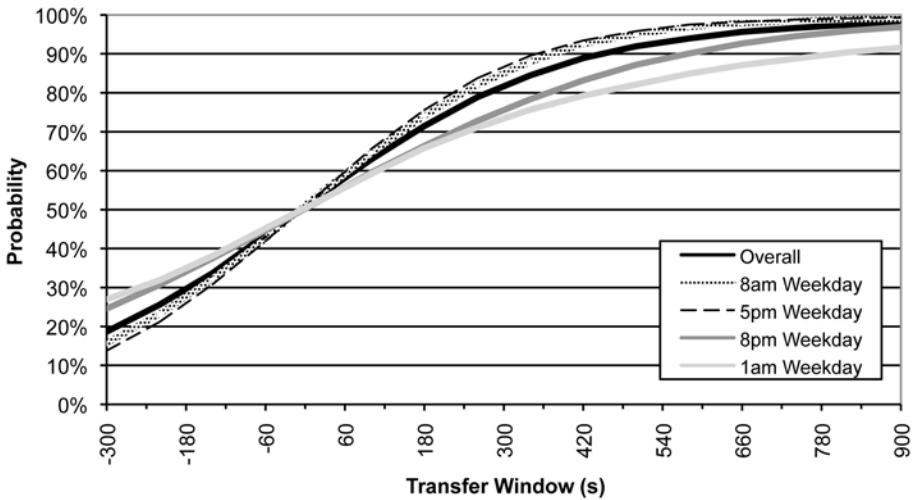


Figure 6. Probability of making a transfer for hour of weekday

Hourly Trend

It is instructive to look at information comparing all hours of the day, instead of just the four selected in the last few plots. Figure 7 compares the quartiles (25, 50, and 75 percent quantiles) and means, along with 95 percent confidence intervals, for each hour of arrival on a weekday. These values were computed using the arrival lateness ECDFs from each hour.

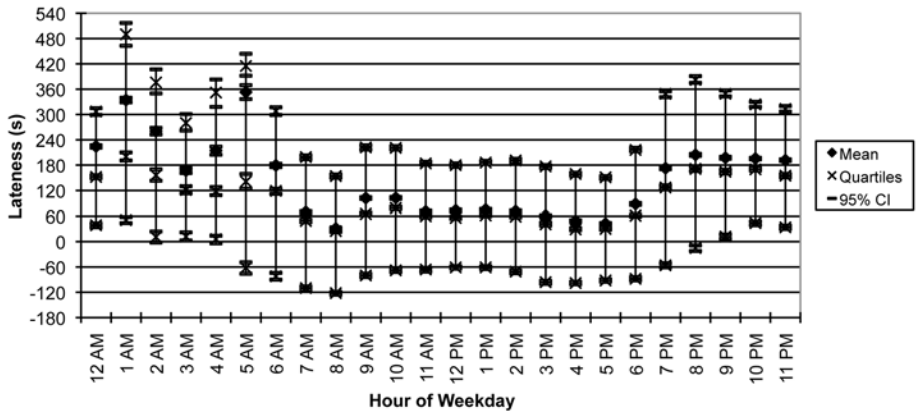


Figure 7. Quartiles and means: Hour of weekday comparison

The information here is both interesting and valuable. If “reliable” is defined as having a small distance between the first and third quantiles, then the most reliable times of bus operation are at 12:00 p.m. and 5 p.m. The level of reliability is, in fact, fairly good and does not vary much throughout the course of the workday. Then, the reliability deteriorates sharply from 5 p.m. to 8 p.m., and the skew of the distribution (as indicated by the distance between the mean and the median) increases. Between 9 p.m. and 12:00 a.m., there is a surprising increase in reliability, though the mean lateness stays relatively high. Finally, there is inexplicable behavior, as all reliability is lost between 12:00 a.m. and 1 a.m., rendering bus behavior almost unpredictable.

It should be noted that it is quite possible that the erratic behavior observed in the early hours is caused by erroneous GTFS data, leading to faulty treatment of the GPS data. The outlandishly large skew on the 5 a.m. data, in particular, is quite suspect. It is difficult to make any conclusions here, since routes run much less frequently at these hours and so there is much less data to draw upon.

It is clear, however, that a rider should expect distinctly different behavior when using public transportation during business hours and in the evening. Again, this is especially relevant to the commuter, who is typically highly concerned with timeliness and reliability.

Progress of Route Comparisons

This next set of comparisons splits the data according to how far along its route a transit vehicle has progressed at each stop. This can be done in a number of ways, and two were selected here: first is the stop number, which enumerates the stops along a route in sequence, and second is route portion, defined as the stop number divided by the total number of stops for the trip.

Route Portion ECDFs

Figure 8 shows the arrival lateness ECDFs for the start of the route (those stops having route portion between 0 and 0.25), middle of the route (0.25 to 0.75), and end of the route (0.75 to 1.0), as well as for overall arrival lateness. These curves are relatively similar, so it is difficult to make interesting conclusions offhand. They do demonstrate a trend of increasing lateness and a heavier tail as a route progresses, but whether this makes a perceptible effect is unclear.

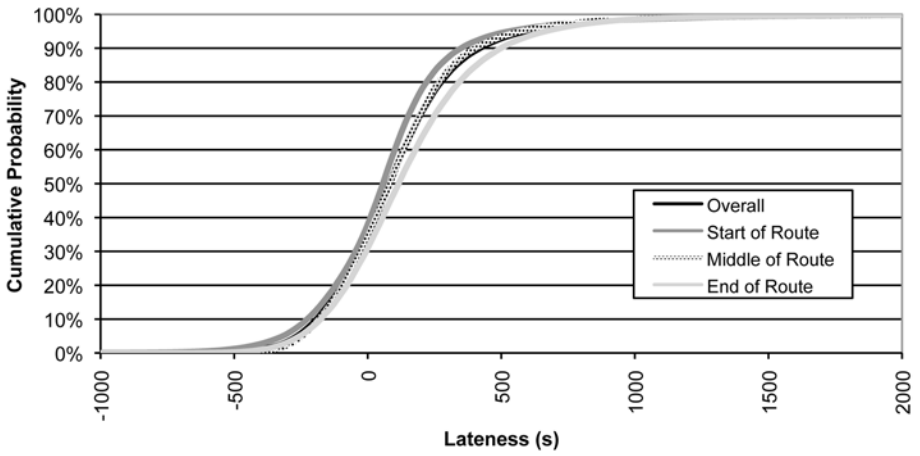


Figure 8. Route portion ECDF comparisons

Probability of Transfer between Route Portions

To better understand the effect of route progress on lateness behavior, the probability of making a transfer according to transfer window size was once more calculated. In this case, four scenarios were selected, representing transfers between different portions of the route, using the same definitions as the ECDFs plotted previously. The results are shown in Figure 9. The cases of making transfers from the beginning of a route or to the end of the route have been omitted, as they are ostensibly less useful.

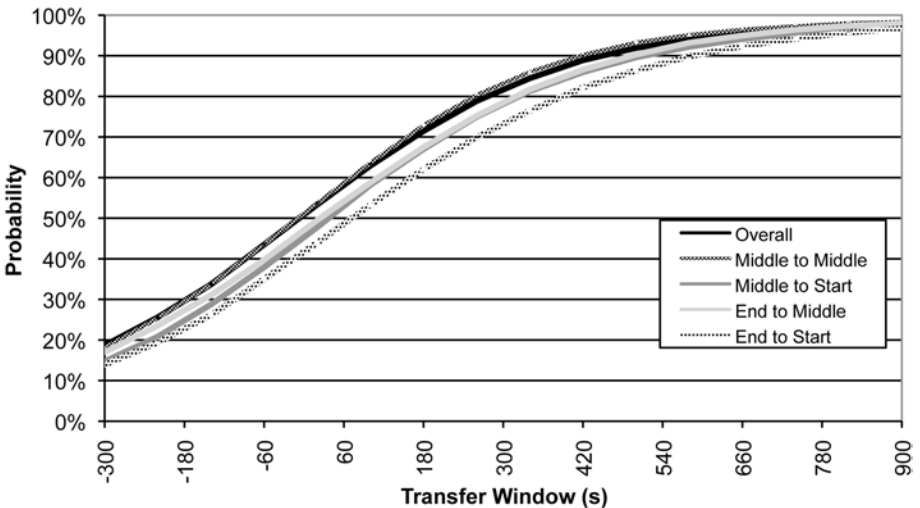


Figure 9. Probability of making a transfer for route portion

Observe that the curves no longer all intersect the 50 percent mark where the transfer window is zero, since the transfers are no longer symmetric. All but the middle-to-middle curve appear to have shifted below the overall curve in probability, suggesting that the previous results were optimistic. For some windows, there is as large as a 10 percent decrease in probability. Fortunately, the shift is not so large in the higher probabilities. To achieve 90 percent probability in the worst case (transferring from the end of one route to the start of another), the transfer window need only be 1 minute larger than in the overall case.

Trend along Route

We also can clarify the effect of route progress by looking more closely for a trend. Using the same type of plot that demonstrated an hourly trend, Figure 10 shows the quartiles and means of the arrival lateness distribution according to stop number. It is plain to see that lateness increases as a route progresses, and that reliability (again using the distance between the 1st and 3rd quartiles as an indicator) decreases. The lack of data in the higher stop numbers causes much larger confidence intervals, but the trend is still apparent, and a trip planner should account for it.

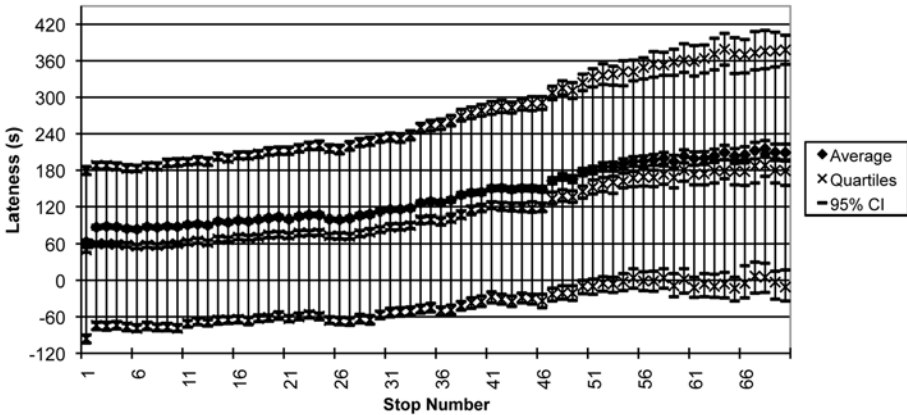


Figure 10. Quartiles and means: Stop number comparison

Trip Plan Evaluation

Taking a step further in our observations, this final analysis makes a quantitative evaluation of a simple trip plan. Suppose a commuter is evaluating the option of taking the bus to work and wishes to arrive at 8 a.m. Looking at the system map, the commuter finds that he must make one transfer. The schedule shows that the first bus is scheduled to arrive at the transfer point at 7:00 a.m., 7:30 a.m., and 8:00 a.m.

The second bus is scheduled to depart from the transfer point every 20 minutes beginning at 7:10 a.m. and to arrive at the final destination 20 minutes later.

The naïve conclusion is that there is no chance to make the transfer from the 7:30 a.m. arrival to the 7:30 a.m. departure, and so he must wait 20 minutes for the next bus, arriving at work at 8:15 a.m. (15 minutes late). To arrive on time, therefore, he must take the bus which arrives at the transfer point at 7:00 a.m., which is highly inconvenient.

Now, instead, it is possible to make a probabilistic evaluation using the knowledge of behavior derived thus far. Figure 11 shows the ECDF of the commuter's arrival time at his destination given the schedule data above, where he takes the first bus scheduled to arrive at the transfer point at 7:30 a.m. and then transfers to the next bus that he sees. This ECDF was generated through simulation on the 7 a.m. and 8 a.m. hour of arrival datasets, using the arrival and departure data as appropriate. The bimodal characteristic of this curve represents the two main outcomes of the trip plan: that half the time, the 7:30 a.m. transfer is successful, and half the time it is not. The commuter's arrival time, on average, is only 154 seconds, or about 3 minutes, past 8 a.m. There are occasions where he will be more than 20 minutes early, and occasions where he will be more than 20 minutes late; each of these represents less than 2 percent of occurrences, or about 1 day every 11 weeks. This method not only has discovered the reality of the situation to be better than initially thought, but it also has formed the correct expectations for the rider.

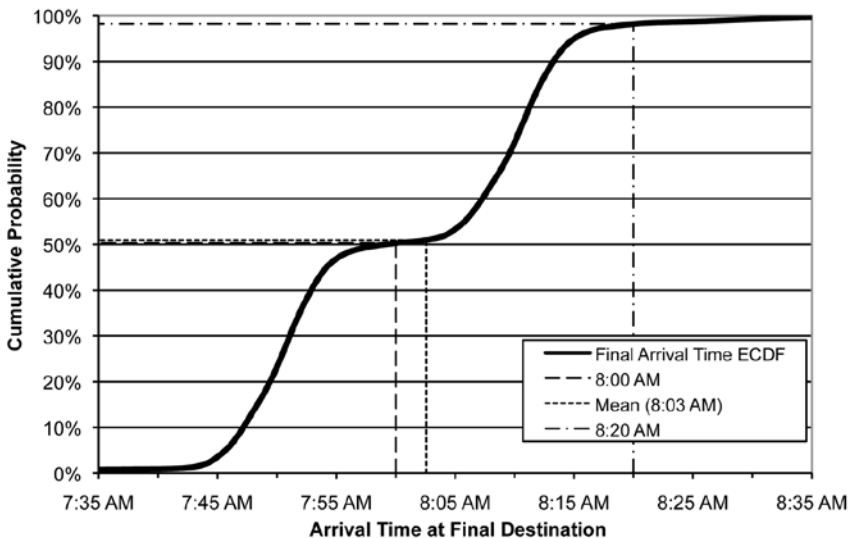


Figure 11. ECDF of trip plan for daily commute

Conclusions and Further Work

This work has demonstrated a new approach to the analysis of arrival data that creates an understanding of transit behavior from the rider's point of view. Using this approach, operators can understand precisely how on-time performance impacts their users and identify sources of frustration. By releasing timeliness or tracking data to the public, agencies can allow the community to use these methods to provide better information and to improve the rider's experience, in the same way that releasing schedule and arrival prediction data has seen growing success.

Online trip planners, or new tools entirely, can use the model from this research to provide better information to the rider. Incorporating it directly into the trip planner would make it possible to operate on explicit probabilities of transfer, instead of guessing a minimum transfer time. This easily could be extended to consider all outcomes and their probabilities, giving every trip plan an evaluation like that shown in Figure 11. From there, the optimization problem can be redefined in a very flexible and meaningful manner.

It also is desirable to communicate this knowledge in a more direct form, so that a rider is always equipped to form the correct expectations ahead of time. Of course, as is evident in the case of paper schedules, presenting information in a format that is easy to use and understand is a difficult problem. A small step is to provide simple rules of thumb that will guide expectations in a way that reduces frustration: show up five minutes early, give small transfer windows a (small) chance, expect poor behavior late at night. A more complete and utile form of communicating this information is worth pursuit.

Acknowledgments

The author would like to thank the reviewers for their effort in improving this paper. The author would also like to thank Robert Damphousse for his help in making sense of the initial GPS data. Finally, the author thanks the San Francisco Municipal Transit Agency for releasing their tracking and schedule data to the public.

References

Balcombe, R. J., and C. E. Vance. 1998. Information for bus passengers: A study of needs and priorities. *Transportation Research Laboratory—Report 330.*

- Berkow, M., A. M. El-Geneidy, R. L. Bertini and D. Crout. 2009. Beyond generating transit performance measures: Visualizations and statistical analysis with historical data. *Transportation Research Record* 2111: 158-168.
- Cain, A. 2007. Are printed transit information materials a significant barrier to transit use? *Journal of Public Transportation* 10(2): 33-52.
- Caulfield, B., and M. O'Mahony. 2009. A Stated preference of real-time public transit stop information. *Journal of Public Transportation* 12(3): 1-20.
- Chien, S. I., Y. Ding, and C. Wei. 2002. Dynamic bus arrival prediction with artificial neural networks. *Journal of Transportation Engineering* 128(5): 429-438.
- Diaz-Ramos, S., D.L. Stevens, Jr., and A.R. Olsen. 1996. EMAP Statistics Methods Manual. EPA/620/R-96/XXX. Corvallis, OR: U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory.
- Google. 2009. General Transit Feed Specification (previously Google Transit Feed Specification). http://code.google.com/transit/spec/transit_feed_specification.html [accessed March 2009].
- Jeong, R.H. 2004. The prediction of bus arrival time using automatic vehicle location systems data. PhD Dissertation, Texas A&M University.
- Li, Y. 2003. Evaluating the urban commute experience: a time perception approach. *Journal of Public Transportation* 6(4): 41-66.
- Massart, P. 1990. The tight constant in the Dvoretzky-Kiefer-Wolfowitz Inequality. *The Annals of Probability* 18(3): 1269-1283.
- Mishalani, R. G., M. M. McCord and J. Wirtz. 2006. Passenger wait time perceptions at bus stops: Empirical results and impact on evaluating real-time bus arrival information. *Journal of Public Transportation* 9(2): 89-106.
- Modesti, P., and A. Sciomachen. 1998. A utility measure for finding multiobjective shortest paths in urban multimodal transportation networks. *European Journal of Operational Research* 111: 495-508.
- San Francisco Municipal Transit Agency. 2009. SFMTA GTFS transit data license agreement and download. <http://www.sfmta.com/cms/asite/transitdata.htm> [accessed March 2009].

- San Francisco Municipal Transit Agency. 2010. SFMTA GPS vehicle location data. <http://www.sfmta.com/cms/asite/nextmunidata.htm> [accessed January 2010].
- San Francisco Municipal Transit Agency. 2011. <http://www.sfmta.com/cms/ahome/indxabmu.htm> [accessed February 2011].
- Shalaby, A., and A. Farhan. 2004. Prediction Model of bus arrival and departure times using AVL and APC data. *Journal of Public Transportation* 7(1): 41-61.
- Sherali, H. D., C. Jeenanunta, and A. G. Hobeika. 2006. The approach-dependent, time-dependent, label-constrained shortest path problem. *Networks* 48(2): 57-67.
- Sollohub, D., and A. Tharanathan. 2006. A multidisciplinary approach toward improving bus schedule readability. *Journal of Public Transportation* 9(4): 61-86.
- Wall, Z. and D. J. Dailey. 1999. An Algorithm for predicting the arrival time of mass transit vehicles using automatic vehicle location data. 78th Annual Meeting of the Transportation Research Board, January 1999.

About the Author

COLIN BICK (colin.bick@gmail.com) has used public transit almost exclusively for several years. Professionally, he is a software developer and holds an M.S. in applied mathematics from the Colorado School of Mines.