An Investigation of Bus Headway Regularity and Service Performance in Chicago Bus Transit System

Minyan Ruan¹, Jie Lin²

ABSTRACT In frequently served bus routes, passengers are more concerned about bus headway regularity than actual punctuality of bus arrival according to the schedule. Buses arriving within very small (bus bunching) or very large headway influence bus service performance. In this paper, a stop-level probability based headway regularity metric is proposed to measure bus service reliability. The metric is formulated as a function of bus dwell time, passenger activities and expectation. The usefulness of the metric is demonstrated through a case study on one of the bus routes operated by the Chicago Transit Authority (CTA), using automatic vehicle location (AVL) data. The empirical findings from the data are consistent with what is found in the mathematical model. The study findings are useful in bus planning and operations (especially concerning bus dispatching and holding) to improve bus service performance and increase ridership.

INTRODUCTION

One of the main concerns of transit agencies on competing with auto transportation is the performance of service. Bus service may vary due to many factors such as traffic conditions, bus operation and passenger demand, all of which have been studied in a large body of literature (Polus, 1978; Nakanishi, 1997). In particular, bus headway irregularity discourages commuters’ use of public transit (Polus, 1978; Fu and Yang, 2002). Aware of the effects of those factors, transit planners and engineers work to improve transit services. In the process, effective service
performance metrics will be useful tools for measuring, diagnosing, and pinpointing problem areas.

The most commonly used metric is the average passenger waiting time proposed by Osuna and Newell (1972). Assuming uniform passenger arrival, average passenger waiting time is further derived as the sum of one half of the average headway and the ratio of headway variance to twice the average headway, i.e., \( E(Waiting) = 0.5E(Headway) + V(Headway)/2E(Headway) \). Beside the observed platform waiting time, Furth and Muller (2005) suggested that there be potential waiting time included in the measurement of service reliability. This potential waiting time, \( W_{potential} = W_{0.95} - E[W] \), together with the mean waiting cost, creates the total waiting time, \( W_{equivalent} = 0.5(E[W] + W_{0.95}) \). Both of these two studies concerned the headway variance, indicating that the average waiting time would be less if the headway variance was small when holding the average headway equal. In Polus’s study (1978), the observed travel time on links along a certain route for several days was found to follow Beta distribution, whose parameters were estimated and found linearly related with each other and also with the link length in the off peak hours. Link reliability was then given as the deviation of travel time. The concept of utilizing fitted distribution and estimating parameters on system characteristics is plausible. However, the link based travel time deviation, while useful to transit agencies to improve service regularity, hardly gives passengers the direct knowledge of when the next bus arrives or service reliability. Even when the predicted arrival time of the next bus is available, passengers waiting at a bus stop would want to know the probability of the bus arriving within a tolerable range of that predicted arrival time.

Other studies explored schemes to improve service reliability. The zone scheduling strategy proposed by Jordan and Turnquist (1979), aimed at improving the reliability defined as the sum of variance in passenger trip time over all passengers that used the route. Under this scheduling scheme, the average and variability of in-motion and dwell time could be reduced due to the decrease of service stops. Furthermore, Turnquist (1982) differentiated the passenger’s viewpoint of bus service reliability into schedule adherence and headway regularity. That is, with infrequent service routes passengers tend to learn the schedule and coordinate their arrivals at the bus stops with the scheduled bus arrival time so as to minimize wait time. This means that schedule adherence is a good measure of service quality provided to the passengers and thus induces the schedule-based holding strategy. If demand is large and service is frequent enough, passengers will be arriving randomly at the bus stops without referring to the bus timetable. Under this situation, the measure of bus service can be formulated as average waiting time introduced earlier and the headway-holding strategy is appropriate to reduce the variance of headway.

As advance technologies, e.g., automatic vehicle locator and passenger counter (AVL/APC), become available, transit agencies (Hammerle, et al, 2005; Strathman, et al, 2002; Bertini and EI-Geneidy, 2003) have been actively pursuing novel ideas of exploiting the rich AVL/APC data to inform and enhance their service performance. Bertini and EI-Geneidy (2003) demonstrated using AVL/APC data to derive performance measures of the entire fleet, a route, a segment and a point. The performance indicators derived were the hours of service, number of trips, number of miles operated, number of passengers carried, number of passengers per mile, average speed,
number of operators, trip time and dwell time. It is believed that there are more potential and novel applications of AVL data and more research is needed.

Focusing on a frequently served urban bus route, we propose a probability based service reliability metric that is meaningful to both passengers and management. In this paper, we first mathematically formulate the probability based headway regularity metric and then demonstrate the use of the proposed metric through a case study of a CTA bus route recorded with AVL/APC. Finally, we summarize the contributions and limitations of the work.

FORMULATION OF PROBABILIT BASED HEADWAY REGULARITY

As indicated in the literature review, service reliability varies between different bus service conditions. In urban areas where bus service is frequent, passengers concern mostly about bus headway regularity. Extending that idea, we present in the following a probability based headway regularity metric, which integrates headway, number of stops, bus dwell time, passenger activities at stops and passenger tolerance level into one single measure. This metric gives the probability of a bus operating with regular headway as expected.

First, we defined bus service reliability at a given stop of a route.

**Definition:** bus service reliability at a given stop of a route is defined as the probability of a bus arriving at the stop within the maximum passenger anticipated headway. That is, at a given stop \( k \), service reliability is defined as,

\[
p = P\{\text{headway}_k \leq H_b\}
\]

(1)

where \( H_b \) is the maximum passenger anticipated headway. Note that Eq. (1) does not constrain the lower bound of the headway. That is because from the passengers’ point of view the more frequent bus service the better. From the agency’s point of view, frequent bus service means higher operating costs, which is a much less desirable situation. Such consideration is outside the scope of this paper. In addition, this formulation can include the scenarios where there are bus bunching, which is a constant service phenomenon that plagues transit agencies.

Now consider any two consecutive stops \( k-1 \) and \( k \) on the route. The link defined between the two stops is denoted as link \( k \).

\[
\begin{array}{c}
\bullet \quad h_{k-1}, d_{k-1} \quad t_k \quad h_k, d_k \\
\bullet \quad h^p_{k-1}, d^p_{k-1} \quad t^p_k \quad h^p_k, d^p_k
\end{array}
\]

We introduce the following notation:

- \( h_k \) : Headway between two consecutive buses at stop \( k \).
- \( t^p_k \) : Previous bus travel time on link \( k \).
- \( t_k \) : Current bus travel time on link \( k \).
- \( d^p_k \) : Previous bus dwell time at stop \( k \).
- \( d_k \) : Current bus dwell time at stop \( k-1 \).
Thus, we have the following headway definition at stop $k$:

\[
    h_k = h_{k-1} + d_{k-1} + t_k - d_{k-1}^p - t_k^p \\
    = h_{k-1} + d_{k-1} - d_{k-1}^p + \Delta t_k
\]

where $\Delta t_k$ is the difference in travel time between the current and previous buses on link $k$.

If we further assume uniform passenger arrival and boarding (alighting), the dwell time $d_{k-1}$ or $d_{k-1}^p$ in Eq. (2) can be written as the following.

\[
    d_{k-1} = \lambda_{k-1} (h_{k-1} - d_{k-1}^p) / \mu_{k-1}^b \\
    = \rho_{k-1} (h_{k-1} - d_{k-1}^p)
\]

where:

$\lambda_{k-1}$ : Passenger arrival rate at stop k-1.

$\mu_{k-1}^b$ : Passenger boarding process rate at stop k-1.

The term $(h_{k-1} - d_{k-1}^p)$ is the difference between the headway and dwell time of the preceding bus at stop k-1, i.e., the maximum passenger waiting time between two consecutive buses, which is equal to the difference between the time difference between the arrival time of the current bus and the departure time of the last bus. $\rho_{k-1} = \lambda_{k-1} / \mu_{k-1}^b$ is the service intensity at stop k-1 defined as the ratio between the passenger arrival rate and boarding rate at the stop. Here we also assume that the capacity of the bus is sufficient to accommodate all passengers that are waiting to board the bus at the stop. $\rho$ is larger if service rate is smaller or arrival rate is higher and is no greater than 1. Another assumption is that the alighting time can be ignored.

In Eq. (2), if we assume the travel time on link $k$ experienced by the two consecutive buses (when they are not too far apart in time) remains constant, meaning there is no difference in traffic condition, signal delay or drivers’ driving habit, $\Delta t_k = 0$. This is a reasonable assumption for experienced bus drivers who are familiar with the route. Substituting $d_{k-1}$ in Eq. (3) for that in Eq. (2), we have the following definition of headway.

\[
    h_k = h_{k-1} + \rho_{k-1} (h_{k-1} - d_{k-1}^p) - d_{k-1}^p \\
    = (1 + \rho_{k-1}) (h_{k-1} - d_{k-1}^p) \\
    h_{k-1} = (1 + \rho_{k-2}) (h_{k-2} - d_{k-2}^p) \\
    h_{k-2} = (1 + \rho_{k-3}) (h_{k-3} - d_{k-3}^p) \\
    \vdots \\
    h_2 = (1 + \rho_1) (h_1 - d_1^p)
\]

So $h_k$ equals
Finally, substituting Eq. (5) into Eq. (1), we have:

\[
p = P\{h_k \leq H_b\} = P\left\{ \prod_{j=2}^{k}(1 + \rho_{j-1})h_1 - \sum_{i=2}^{k}\sum_{j=i}^{k}(1 + \rho_{j-1})d_{i-1}^p \leq H_b \right\}
\]

Eq. (6) defines the probability based headway regularity metric, which is jointly determined by service intensities (\(\rho_{j-1}\)), dwell time of the preceding bus (\(d_{i-1}^p\)) at all stops into the trip till stop \(k\), headway at the first stop (\(h_1\)) and the maximum anticipated headway (\(H_b\)). This definition is operationable for transit agencies to improve headway regularity because those variables can be measured and controlled by the agency, whereas traffic related variables, which no doubt influence bus headway regularity, are generally outside of transit agencies’ control.

Eq. (6) can be simplified if we further assume a constant service intensity and dwell time of the preceding bus at those stops, i.e., \(\rho_{k-1} = \Lambda \Lambda = \rho_3 = \rho_2 = \rho_1 \Lambda \Lambda\), and \(d_{k-1}^p = \Lambda \Lambda = d_3^p = d_2^p = d_1^p = d^p\). Then Eq. (6) can be rewritten as the following:

\[
p = P\{h_k \leq H_b\} = P\left\{ (1 + \rho)^{k-1}h_1 - \sum_{i=2}^{k}(1 + \rho)^{k-i+1}d_{i-1}^p \leq H_b \right\}
\]

\[
= P\left\{ (1 + \rho)^{k-1}h_1 - d_p \sum_{i=2}^{k}(1 + \rho)^{k-i+1} \leq H_b \right\}
\]

\[
= P\left\{ (1 + \rho)^{k-1}h_1 - d_p \left[ (1 + \rho)^{k-1} \right] - \frac{1}{\rho} \right\} \leq H_b \right\}
\]

\[
= P\left\{ h_1 \leq \frac{d^p (1 + \rho) - (1 + \rho) + H_b}{\rho (1 + \rho)^{k-1}} \right\}
\]

\[
= P\left\{ h_1 \leq \frac{d_p (1 + \rho) - 1}{\rho(1 + \rho)^{k-2}} + \frac{H_b}{(1 + \rho)^{k-1}} \right\}
\]
Hence, the headway regularity probability at stop \( k \) is expressed as the probability of the headway at the first stop less than a threshold. Eq. (7) is the dispatching headway probability with a scaling threshold. That is, the headway regularity at any stop \( k \) into the trip is related to the dispatching strategy. If the dispatching probability is known, the headway regularity probability can be determined mathematically. Eqs. (6) and (7) provide useful insight to bus headway regularity and mathematical formulation of the influence of service parameters on headway regularity.

It can be observed from Eqs. (7) that when \( H_b > \frac{1 + \rho}{\rho} d^\rho \), the probability of headway within the maximum anticipated headway will be decreasing as the number of stops increases. This suggests that there exists a threshold for \( H_b \) under which the headway regularity probability increases for later stops.

In the next section, we demonstrate the use of the proposed headway regularity measure in Eq. (7) to evaluate a CTA bus route in downtown Chicago.

**CASE STUDY**

Chicago Transit Authority (CTA) serves the City of Chicago and 40 surrounding suburbs with over 150 routes (see Figure 1) and more than 2000 buses and 2,273 route miles\(^3\). CTA buses provide about 1 million passenger trips a day and serve more than 12,000 posted bus stops. Downtown Chicago, located in the middle east of the city, generates the largest demand for bus service.

---

\(^3\) Statistics are from the CTA official web page: [http://www.transitchicago.com](http://www.transitchicago.com)
CTA began installing AVL and APC devices on select buses in 2002 (Hammerle, 2004), as part of the Automated Vehicle Annunciation System (AVAS), which is an enormous effort by CTA to provide real-time broadcasting of bus arrival time and other route information at bus stops. Currently all CTA buses are AVL equipped and over 40% are APC equipped. These AVL systems collect the so-called archived AVL data, which means the recorded data is only uploaded to a CTA server after the bus’ daily operation. Currently only a small portion of the CTA bus routes are capable of transmitting AVL data in real time.

AVL data is event-based, in which events are serviced stops, un-serviced stops and time points. Serviced stop events are recorded when the bus stops at the pre-specified stop, while the un-serviced stop events are recorded when the AVL sensor detects the bus is passing the pre-specified stop without stopping. Time points are planned geographic locations along the routes with the scheduled arrival times for CTA bus operators. In other words, CTA buses are “timed” at time points rather than at stops. Therefore, time point rather than bus stop is chosen as the analysis unit in this study.

The study route is a 14-mile north-south bound CTA key route, which goes through one of the main roads located in the west loop of downtown Chicago. Only southbound service in the morning peak hours (7:00AM-9:00AM) is considered in this paper. It has relatively high demand.
due to the intersection with many other bus routes and train lines connecting to downtown area. The published timetable for the route reveals a maximum headway of about 15 minutes during morning peak hours. According to the CTA planning guidelines, there should be 5 or 6 stops within a mile of a bus route, resulting in 120 stops in total in one direction, many of which are not serviced since no passenger boarding or alighting occurs at those stops. There are thirteen time points in the southbound, including two very closely spaced ones and one each at the two terminals.

**DATA PREPARATION AND EXPLORATION**

We obtained from CTA one-month original AVL records of September 2006 and the time point based headway data. Respectively, they have 41 and 32 fields, including all kinds of bus running information, such as time, distance and locations. The time point based data is a derived set from the original AVL data by CTA personnel. We extracted the weekday morning peak hour (7:00AM-9:00AM) records from both datasets on the study route southbound. Approximately 15% of the bus trips within the study period contain APC records, giving extra information of the number of passengers alighting and boarding at the stops. There are 2023 time point level records for the 7 time points, while the stop level APC has 1159 observations.

In this study, the thirteen time points are aggregated to seven roughly evenly spaced time points (see Figure 2) so that each of the six segments contains equal number of stops. In doing so, the seven time points act as pseudo stops, which contains the aggregated passengers arrival and bus dwell time from all the stops between the current and the preceding time point. Data processing and analytical burden are thus reduced. In addition, it is found that the average service ratio, defined as the fraction of bus trips that serviced a specified point (e.g., a stop or time point), at each of the seven time points is roughly the same at about 0.45. With an equal service ratio, it is further assumed in the study a uniform demand and equal dwell time at each time point, which is discussed later in the paper.
Figure 2. Study route with stops and time points

Bus Trajectory

The bus trips trajectory on Sep 1st, 2007 during 5:00 AM and 9:00 AM below in Figure 3 shows that there are much more bus services during rush hours from 7:00 AM to 9:00 AM. The dots on lines marked the arrivals at each of the planned 13 time points along the trip. As we can see, many of the trips bunched together and some even overtook the other during some segments of the route during peak hours. When a bus bunches with the preceding bus, it is usually followed by a larger headway, i.e., longer waiting time for passengers, till the next bus comes. Bus bunching is considered inefficient use of resources because the lead bus usually carries a lot more passengers than the following bus, resulting an almost wasted bus trip in the following bus. As an interesting observation point, the bus trajectory plots seem to suggest that bunching is more likely to happen when the dispatching headway (i.e., headway at the first stop) is irregular. This observation is consistent with our theoretical derivation shown in the previous section.
Headway distribution

CTA maintains a schedule-based operation system, which schedules bus arrival time at time points. This management strategy determines the dispatching decisions. That is, buses are dispatched in such a way to meet the scheduled arrival time at time points. The dispatching headway distribution of the first time point is shown in Figure 4, in which the highest probability headways range from 6 to 9 minutes. Special cases like extreme short or long headway may be owing to the garage supervisor’s judgment of the traffic condition, late or early returning of buses or personnel shortage, etc.

Figure 5 depicts the headway distributions for the rest of the time points (B – G) respectively. Generally speaking, higher percentage of larger headways is observed when higher percentage of bunching headways is present, as expected. Compared with the first time point (A) in Figure 4, the headway distributions at F and G appear to be less centralized and more skewed to larger headways, other than an extreme large proportion of short headways less than 1 minute. At the 15-minute headway, the probability of headways to the left is generally decreasing from B to G.

This is consistent with what we have found in Eq. (7) when $H_n > \frac{(1 + \rho) \rho^d}{\rho}$.
Figure 4. Headway (min) distribution at the first time point A

Figure 5. Headway (min) distribution at the time point B-G
FITTED DISPATCHING HEADWAY DISTRIBUTION

Using the time point headway data for weekday morning peak hours of Sep. 2006 at time point A, we fitted different distributions of the observed headways with the analysis software Crystal Ball. Table 1 shows the fitting results, as well as the goodness of fit for four distributions, gamma, Weibull, logistic, and lognormal. Although gamma distribution ranks the second best on Kolmogorov-Smirnov, it outperformed all the other on Anderson-Darling and Chi-square rankings. Collectively, gamma distribution provides the better fit.

Table 1: Fitting distribution comparison

<table>
<thead>
<tr>
<th>Distribution</th>
<th>A-D</th>
<th>Chi-Square</th>
<th>K-S</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>3.16</td>
<td>43.0743</td>
<td>.0824</td>
<td>Location=-1.98, Scale=2.41, Shape=4.60</td>
</tr>
<tr>
<td>Weibull</td>
<td>5.17</td>
<td>65.6997</td>
<td>.0991</td>
<td>Location=0.29, Scale=10.64, Shape=1.81</td>
</tr>
<tr>
<td>Logistic</td>
<td>5.28</td>
<td>53.8854</td>
<td>.0819</td>
<td>Mean=8.49, Scale=2.82</td>
</tr>
<tr>
<td>Lognormal</td>
<td>13.54</td>
<td>97.4644</td>
<td>.1652</td>
<td>Mean=10.17, Std. Dev.=9.73</td>
</tr>
</tbody>
</table>

The general formula for the probability density function of gamma distribution is:

\[ f(x) = \frac{(x-\mu)^{\gamma-1} \exp(-\frac{x-\mu}{\beta})}{\beta^\gamma \Gamma(\gamma)} \quad x \geq \mu; \gamma, \beta > 0 \]  

where \( \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \) and there are three parameters, location \( \mu \), scale \( \beta \), and shape \( \gamma \). In the fitted curve in Figure 6, their values are: location \( \mu = -1.98 \), scale \( \beta = 2.41 \) and shape \( \gamma = 4.60 \).

CALCULATED HEADWAY REGULARITY PROBABILITIES

Using the fitted gamma distribution of the first time point headway and Eq. (7), we calculated the headway regularity probabilities. The following assumptions were made to facilitate the calculation. The average dwell time was set at 2.5 minutes while serving approximately 15 passengers, estimated from the sample APC data. Moreover, the average passenger arrival rate was estimated to be 0.03 person/min at time points, based on the mean passenger boarding 15 and mean headways around 9 minutes. Also, it was assumed an average of 10 seconds passenger boarding time at each time point because the average dwell time is 2.5 minutes and passengers boarding are 15 on the timepoints. As a result, service intensity rate would be 0.3. The maximum anticipated headway was set to be equal to the maximum timetable headway of 15 minutes.

The calculated headway regularity probabilities at the seven time points are shown in Table 2. In addition to the actual scenario where \( \rho = 0.3 \) and \( H_b = 15 \). We considered three other scenarios: ( \( \rho = 0.3, H_b = 10 \)), ( \( \rho = 0.5, H_b = 15 \)) and ( \( \rho = 0.5, H_b = 10 \)). All the scenarios are with \( H_b > \frac{(1+\rho)d^o}{\rho} \) except for the last one ( \( \rho = 0.3, H_b = 10 \)).
Table 2: Service reliability in aggregated stops with different estimated parameters

<table>
<thead>
<tr>
<th>Aggregated Stop</th>
<th>Probability ($\rho=0.3 \ H_b=15$)</th>
<th>Probability ($\rho=0.3 \ H_b=10$)</th>
<th>Probability ($\rho=0.5 \ H_b=15$)</th>
<th>Probability ($\rho=0.5 \ H_b=10$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.8381</td>
<td>0.6388</td>
<td>0.7727</td>
<td>0.5654</td>
</tr>
<tr>
<td>3</td>
<td>0.8089</td>
<td>0.6487</td>
<td>0.6805</td>
<td>0.5225</td>
</tr>
<tr>
<td>4</td>
<td>0.7836</td>
<td>0.6563</td>
<td>0.6062</td>
<td>0.4929</td>
</tr>
<tr>
<td>5</td>
<td>0.7625</td>
<td>0.6619</td>
<td>0.5514</td>
<td>0.4728</td>
</tr>
<tr>
<td>6</td>
<td>0.7451</td>
<td>0.6663</td>
<td>0.5127</td>
<td>0.4592</td>
</tr>
<tr>
<td>7</td>
<td>0.7311</td>
<td>0.6696</td>
<td>0.4862</td>
<td>0.4501</td>
</tr>
</tbody>
</table>

In the first scenario, the service reliabilities are decreasing along consecutive time points. These results are consistent with the headway distributions trends in Figure 5. This is expected as headway irregularity gets carried over and the service is further exacerbated. The headway regularity probabilities with the maximum anticipated headway of 15 minutes are generally higher than those with 10 minutes. Maximum anticipated headway represents passengers’ tolerance of the bus service. Passengers anticipating 15 minutes are more tolerant than those anticipating 10 minutes, so the bus service reliability is perceived as better when the tolerance level is 15 minutes even under the same bus service performance.

When the service intensity is 0.3, the service reliability is higher than those where the service intensity is 0.5. Service intensity is a measure of passenger activity level at a stop or time point. High service intensity means high passenger activity level at the time point and tends to negatively associate with worse service reliability.

In the second scenario where $H_b > \frac{(1+\rho)d^\rho}{\rho}$, the service reliabilities are increasing. The fact that headway probability within 15 is decreasing while those of 10 is increasing from the first stop to the last one, suggests that headway probability between 10 and 15 minutes is largely decreasing. This result can be supported by the headway distributions plot in Figure 5, the headway stretches to the two extremes instead of having peaks in the medium headways.

In summary, we consider the service reliabilities at the time points of the study segment during weekday peak hours are generally tolerable. When passengers’ expectation and service intensity are both high, the worst service reliability is 0.4501, meaning a 45% possibility that the next bus will be arriving within 10 minutes after the last bus arrives.

CONCLUSION

This paper proposed a probability based headway regularity metric for service reliability. The proposed metric was demonstrated through a case study on one of the Chicago southbound key routes on weekday morning peak hours. For frequent bus service, headway regularity is more concerned than schedule adherence. The probability based headway regularity not only gives the service reliability measure but it associates the measure with important transit operational factors. Thus, transit agencies can pinpoint the problems easily with one single metric. AVL data provides large amount of bus running information and thus greatly reduced the uncertainty of study results. Finally, this metric is easy for the public and passengers to understand. In the end, both passengers and transit agencies are benefit from the improved services.
One limitation of the study is the assumptions of constant passenger arrival rate, bus dwell time and link running time. However, those assumptions are not necessary once more passenger data, e.g., APC data and passenger demand, are available.

ACKNOWLEDGEMENT

We would like to thank CTA for providing the data. The views expressed in the paper are those of the authors solely.

REFERENCES

Chicago Transit Authority website, www.transitchicago.com, last accessed on August 20, 2007
Rossetti, M.D., Turitto T. (forthcoming) “Comparing Static and Dynamic Threshold Based Control Strategies”, Transportation Research Part B.