Malaysian Journal of Mathematical Sciences 15(2): 313–322 (2021)

Malaysian Journal of Mathematical Sciences

Journal homepage: https://einspem.upm.edu.my/journal

Travel Time Reliability Modelling with Burr Distribution

Khoo, W. C.*, Lim, K. S., and Cheong, H. T.

School of Mathematical Sciences, Sunway University, Malaysia

E-mail: wooichenk@sunway.edu.my
**Corresponding author*

Received: 22 October 2020
Accepted: 27 April 2021

Abstract

Travel time reliability is the key factor for the travelers to make decision in choosing the public transport, such as bus. However, current methodology may not be sufficient to explain the travel time and its reliability. This paper aims to specify a new approach in the study of travel time modelling for bus route with Burr (which is also known as Burr XII) distribution. Two real day-to-day travel time bus routes data over 6 months in the Klang Valley area, Malaysia has been investigated for illustration purpose. The travel time data often demonstrates strong positive skewed, which could be explained explicitly using the Burr distribution. Mathematical expression and the statistical properties of the distribution will also be presented. The parameter estimation is done by Maximum Likelihood Estimation (MLE). Comparison has been done with some continuous distribution counterparts such as gamma, Weibull, log-normal and log-logistics. The results show that Burr distribution outperformed and appeared as a viable distribution to characterize the bus travel time. Some dynamic insight of the bus route reliability will also be examined by buffer time index and skew-width methods.

Keywords: Travel time; reliability; bus route; positive skewed; Burr distribution.

1 Introduction

1.1 Models

Travel time is defined as the accumulated time for the movement from a point to another point. Reliability is explained by the operation consistency of a particular route. The travelers rely solely on the travel time reliability in the journey planning, therefore, the reliability is worthwhile to be explored in the Klang Valley region in order to provide more useful information. In Malaysia, traffic condition always shows uncertainty pattern for diverse reasons, i.e. bottlenecks are the notorious factors would occur in the major routes and causing massive traffic congestion. This paper aims to diagnose the travel time pattern analysis by using probability density distribution (PDF). Particularly, the travel time data from two main bus routes in Klang Valley, Malaysia will be analyzed. It is expected that the analysis results may help to provide some significant information for the travelers in transport and logistics planning. Due to the growing demand of public transport and the concern of the eco-environment in this new era, applying statistical analysis on the travel time distributions is inevitable to enhance the understanding in travelling pattern.

Travel time data usually exhibit strong positive skew and long upper tails [14] and this pattern of the data is usually catered well by the Burr distribution as it holds such characteristic. The introduction of the Burr distribution is innovated by [2], where the distribution is named after the inventor. The Burr distribution is popular in the reliability modelling as the cumulative and reliability function has a closed form. Refer to [18, 12] and [15] for some extensive studies relevant to the model characteristics, statistical and probabilistic properties. The effectiveness of Burr in failures modelling has also been discussed. Recently, [14] analyzed the travel time variability on urban roads with Burr distribution. In [9], collected the bus routes data with automatic vehicle. Comparison of the Burr distribution and the Gaussian mixture models has been evaluated. Evidence shows that the Burr distribution may be the appropriate distribution for travel time modeling analysis. Hence, it motivates us to analyze the bus routes travel time pattern in Klang Valley using Burr distribution. Some continuous distribution counterparts have been considered in this paper for comparison purpose. For instance, log-logistic distribution is a well-known for failure times modelling, and it is a special case of Burr distribution. Weibull distribution is also often suggested in the reliability analysis. Burr distribution has advantage against Weibull distribution as it gains a better approximation. This paper takes other potential candidates into the consideration for comparison, such as log-normal and gamma distributions. Comparison will be carried out in terms of the distribution fitting and the results will be shown in the later part.

1.2 Motivating Examples

It is motivated to specify an appropriate distribution to explain the travel time pattern for bus transport from terminal to terminal in Klang Valley. In this paper, the real secondary data were provided by the Rapid KL. Rapid KL is a public transportation system built by Prasarana Malaysia and its operation covering Kuala Lumpur (KL) and Klang Valley area. Two bus routes with different characteristics has been considered. The chosen bus routes are No. 421 and No. 708. The bus routes both cover urban and sub-urban areas, including central business district (CBD) area, residential area, arterial road and highway. The data cover a 6-month period for the duration of 1 June 2014 to 31 December 2014 with the operating time from 6 o'clock in the morning until 12 o'clock midnight.

The data is stored in a spreadsheet with some detailed information of the bus routes. For instances, the route identification (ID), trip ID, stop ID and name, street name, name of bus driver and etc. The data is filtered for every 5-minute interval, and the travel time is taken from 0 - 240 minutes. Figure 1 and 2 show the Google map and the histogram for the route ID No. 421 and No. 708, respectively. It is noticeable from the histogram that the data exhibits long right skewed, which may be appropriate to be fitted by the Burr distribution and the proposed counterparts. Route No. 421 departs from the terminal Hub Taman Dagang, passing by the KL downtown, and arrives at the Wisma Olympic. The journey consists 32 bus stops and the total distance of 10.875km. Route No. 708 starts from Sunway Pyramid, which is one of the main city in Petaling Jaya (PJ) and arrives at Klang Utara, covering a total distance of 26.15 km and 34 bus stops. This paper considers intermediate and long distances coverage, where route No. 421 is the intermediate distance while route No. 708 is a long distance coverage. Table 1 shows the summary of the chosen bus routes for the real analysis in this study.

Table 1: Summary of the bus routes.

Route No.	Number of Bus Stop	Departure Stop	Arrival Stop	Distance of Route(km)	Sample Size
421	32	Hub Taman Dagang	Wisma Olympic	10.875	4683
708	34	Sunway Pyramid	Klang Utara	26.150	3494

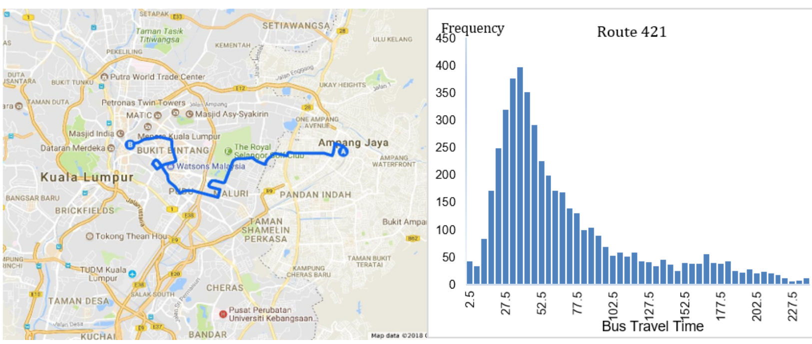

Figure 1: Map showing and histogram of the bus route No. 421.

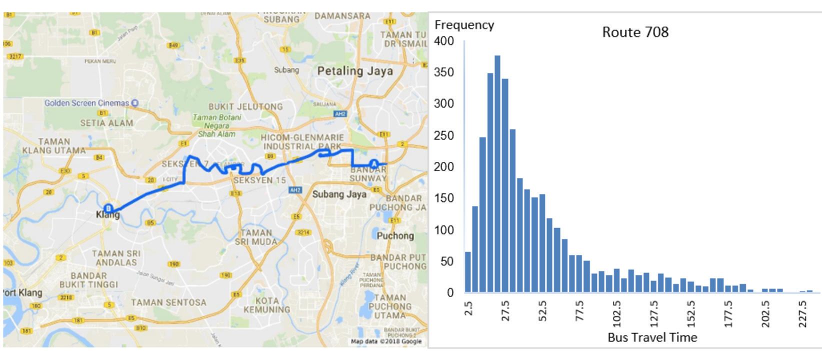

Figure 2: Map showing and histogram of the bus route No. 421.

1.3 Structure of Paper

In this article, Section 2 will discuss some existing works on travel time reliability modelling. The drawback of the existing approaches is the motivation for this paper. The model for travel time modelling will be studied in Section 3. Particularly, a 3-parameter Burr distribution will be presented for the bus route travel time modelling. The statistical properties such as cumulative distribution function (CDF) is also studied for the travel time reliability analysis as the reliability analysis based upon the percentile rely solely on the CDF computation. Here, three reliability measurements have been used for the analysis, namely buffer time index, skew-method and width-method. In Section 4, the parameter estimation is done by maximum likelihood estimation (MLE). The numerical computation is run by the built-in function of Matlab R2017a. Section 5 presented the data analysis results and discussion. For illustration purpose, two bus routes in Klang Valley area are chosen for the distribution fitting. Comparison with other potential distributions will be carried out and examined by Akaike Information Criterion (AIC). Section 6 concludes.

2 Background

In transportation, the measurement of travel time variability determines the level of reliability. The reliability analysis has been done since a few decades ago. In 1970s, [8] measure the mean travel time with standard deviation. Then, the analysis improved the bus service reliability. In [10], the author modelled the reliability of the arterial routes by the inverse of the standard deviation. Regression and statistical models are considered as the techniques in predicting the reliability. Recent discovery shows that [6] measured the reliability with fundamental statistical properties. Mean and standard deviation is applied to explain the travel duration. Subsequently, [5] investigated the travel time reliability on an urban road with distribution.

The estimation of the travel time to enhance the reliability can be somehow complicated. There are some researchers who did the research in the area of estimating the travel time. In 1959, [4] estimated the freeway travel time reliability by using real-life dual loop data. After four decades, [1] developed a new methodology of degraded link capacities to estimate the reliability in the transportation system. Other than that, [11] presented a study of travel time reliability for private car commuter journey conducted in Melbourne, Australia. [7] analyze the reliability with simulation. Later in 2008, [3] estimated the reliability of travel time for freight corridors and [13] estimated the journey time and analyzed the reliability using stochastic model. The algorithms and case studies are presented in the paper.

The reviews show that although some study has been developed in travel time reliability analysis, it is still lack of study in the development on methodology. Extensive study shown also the travel time reliability has been done by conventional variation coefficient in which the method considers solely on the fundamental statistical properties, i.e. mean and standard deviation. This is not sufficient to describe the travel time reliability. Furthermore, not many research has been engaged in bus route travel time modelling. Therefore, our study is to propose a suitable statistical distribution to explain the bus route travel time modelling for Klang Valley area, and the reliability measurement other than mean and standard deviation, i.e. the skew-width methods.

3 The Burr Distribution

3.1 The Mathematical Expression

We present the Burr distribution to analyze the travel time reliability. The fundamental statistical properties have been given here. The CDF is given by

$$F(x) = 1 - \left[1 + \left(\frac{x}{\alpha}\right)^{c}\right]^{-k}, \tag{1}$$

where $x > 0$, $c > 0$ and $k > 0$. The PDF is obtained by taking the first derivative of the cdf, which is given by

$$f(x) = \frac{kc}{\alpha}\left(\frac{x}{\alpha}\right)^{c-1}\left[1 + \left(\frac{x}{\alpha}\right)^{c}\right]^{-(k+1)}. \tag{2}$$

3.2 Travel Time Reliability Metrics

Skew-width methods are applied to analyze the reliability. Refer to [17, Figure 2] for more explanation about the skew-width method framework. The conventional variation coefficients which using standard deviation and mean provide not much insight as these properties tend to obscure important characteristics of the distribution under certain condition. Here, we adopted Buffer Index (BI) and skew-width methods to capture the reliability element. Taylor [16] defines BI as

$$BI = \frac{t_{95} - \bar{t}}{\bar{t}}. \tag{3}$$

The skew method is presented in the form of

$$\lambda^{skew} = \frac{t_{90} - t_{50}}{t_{50} - t_{10}}, \tag{4}$$

and the width method is given by

$$\lambda^{var} = \frac{t_{90} - t_{10}}{t_{50}}, \tag{5}$$

where t_{95} is the 95 percentile, which it can be obtained by Equation (1).

The skew-width methods are the robust measurement for reliability. The analysis is based on percentile and median, which is able to capture the information for long-tailed distribution. The skew method is considered as the ratio of the distance between the 90^{th} and 50^{th} percentile and the distance between the 50^{th} and 10^{th} percentile. Generally, small λ^{skew} defines the distribution is highly right skewed, and large λ^{skew} shows that the distribution is strongly left skewed. The larger λ^{skew} yields the more unreliability. Besides that, the width method is based on the ratio of the range of travel times. Similarly, larger value of λ^{var} indicates larger range relative to the median, which consider lower reliability.

4 Parameter Estimation

4.1 Maximum Likelihood Estimation

Let X be the n sample size of Burr distribution with the PDF $f(x)$ with parameters k, c, α. The log-likelihood function is given by

$$\log L(k,c,\alpha) = n\log\left(\frac{kc}{\alpha}\right) + \sum_{i=1}^{n}(c-1)\log\left(\frac{x_i}{\alpha}\right) - (k+1)\log\left[1+\left(\frac{x_i}{\alpha}\right)^{c}\right]. \tag{6}$$

Taking derivatives of Equation (6) and let it be equal to zero will give the result of the parameter estimation.

5 Data Analysis and Results

5.1 Distribution Fitting

This section discusses the distribution fitting for both routes. Maximum likelihood estimation in Section 4 has been applied for parameter estimation of all distributions in this paper. The estimation is run by MATLAB R2017b with the built-in function and the results are tabulated in Table 3. The histograms in Figure 1 and Figure 2 show that Burr distribution is a potential candidate for the real data. Other distributions such as gamma, log-logistic, log-normal and Weibull distribution are also taken into consideration for comparison purpose. For readers' convenience, the PDF of the other continuous distribution is given in Table 2.

Table 2: Summary of the continuous distribution.

Distribution	Probability Density Function (pdf)	Log-likelihood Function
Gamma	$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}$, $x > 0$	$\log L(\alpha,\beta) = (\alpha-1)\sum_{i=1}^{n}\log(x_i) - n\log[\Gamma(\alpha)] - \alpha n\log(\beta)$, $i = 1, 2, \ldots, n.$
Log-logistics	$f(x) = \frac{1}{\sigma}\frac{1}{x}(\frac{x}{e^{\mu}})^{\frac{1}{\sigma}}[1+(\frac{x}{e^{\mu}})^{\frac{1}{\sigma}}]^{-2}$, $x > 0$	$\log L(\mu,\sigma) = \sum_{i=1}^{n}\frac{1}{\sigma}\log\left(\frac{x_i}{e^{\mu}}\right) - 2\log\left[1+\left(\frac{x_i}{e^{\mu}}\right)^{\frac{1}{\sigma}}\right] - \sum_{i=1}^{n}\log x_i - n\log\sigma$, $i = 1, 2, \ldots, n.$
Lognormal	$f(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}$, $x > 0$	$\log L(\mu,\sigma) = -\sum_{i=1}^{n}\log(x_i) - \frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[\ln(x_i)-\mu]^2$, $i = 1, 2, \ldots, n.$
Weibull	$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(\frac{x}{\lambda})^{k}}$, $x \geq 0$	$\log L(k,\lambda) = n\log(\frac{k}{\lambda}) + (k-1)\sum_{i=1}^{n}\log(\frac{x_i}{\lambda}) - \sum_{i=1}^{n}(\frac{x_i}{\lambda})^{k}$, $i = 1, 2, \ldots, n.$

Figure 3 shows all distribution fittings for Route 421 and 708 respectively. It is noticed that the Burr distribution and its counterparts are appropriately used to explain the pattern of the data. Next, the comparison is carried out by Akaike Information Criterion (AIC). The AIC value is computed by log-likelihood function, which is given by $AIC = 2k - 2\ln(L)$, where k is the number of parameters of the distribution function and L represents the likelihood function. Table 4 presents the goodness-of-fit for the respective distributions. The results show that the Burr distribution outperformed the other distributions. Distribution fittings for both routes are shown in Figure 3.

Table 3: Parameter estimation for the Burr and other continuous distributions

Route	Distribution	Parameters
421	Burr	$c = 2.2826$, $k = 1.1471$, $\alpha = 58.0535$
	Gamma	$\alpha = 2.0928$, $\beta = 32.7212$
	Weibull	$a = 76.1357$, $b = 1.4636$
	Log-logistic	$\mu = 3.9712$, $\sigma = 0.4176$
	Lognormal	$\mu = 3.9690$, $\sigma = 0.7506$
708	Burr	$c = 1.77369$, $k = 1.49861$, $\alpha = 49.6178$
	Gamma	$\alpha = 1.52863$, $\beta = 33.0439$
	Weibull	$a = 1.2318$, $b = 54.3285$
	Log-logistic	$\mu = 3.5805$, $\sigma = 0.4972$
	Lognormal	$\mu = 3.5608$, $\sigma = 0.9226$

Table 4: Goodness-of-fit test for the distributions

Route	Distribution	Log-likelihood	AIC
421	Burr	−23832.1	47670.2
	Gamma	−23870.2	47744.4
	Weibull	−23833.9	47671.8
	Log-logistic	−23888.3	47780.6
	Lognormal	−23977.7	47959.4
708	Burr	−16980.4	33966.8
	Gamma	−17029.2	34062.4
	Weibull	−16992.6	33989.2
	Log-logistic	−17117.4	34238.8
	Lognormal	−17076.0	34156.0

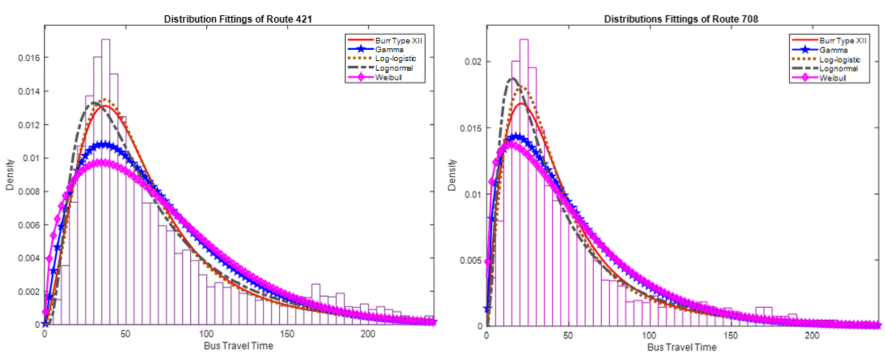

Figure 3: Distribution fittings of Route 421 and Route 708.

5.2 Travel Time Reliability Metrics

This section evaluated the travel time reliability. The Burr distribution has been fitted with the estimated parameters to compute the travel time reliability metrics explained in Section 3.2. We compare two routes in terms of its reliability by BI, skew-width methods and width of travel time. It is reported that the very small value of λ^{skew} indicates the distribution is highly right skewed. Comparatively, the value of λ^{skew} for route 421 is 2.3805 which presents higher right skewness rather than route 708, which means route 421 has better reliability. Besides that, route 708 has higher width of travel time, which is 2.4902 comparable to route 421 which has 2.0653, which means route 708 has lower reliability. This paper serves a good preliminary study for travel time reliability modelling. For more in-depth discussion, it will be carried out elsewhere.

Table 5: Travel time reliability metrics for route 421 and route 708

Route	Buffer Index (BI)	Skew Method, λ^{skew}	Width Method, λ^{var}
421	1.5742	2.3805	2.0653
708	1.7930	2.5740	2.4902

6 Concluding Remarks

Burr distribution is well-known in reliability analysis due to its flexibility in modelling failure times. This paper provide some useful analysis for the bus route travel time reliability, in Klang Valley region, with Burr distribution. The Burr distribution is presented, particularly, the cumulative distribution function is derived to explain the reliability metrics. Existing methodology suggested using some potential continuous distributions, such as gamma, Weibull, log-normal and log-logistic for travel time modelling. Here, the comparison has been carried out between the Burr distribution and the nominated continuous distributions. The comparison results show that the Burr distribution outperformed the other distributions.

The conventional variation coefficient which considering the ratio of variance and mean may not be able to provide dynamic description for the data distribution. Instead, in this paper, three reliability metrics, namely buffer time index, skew-width methods have been introduced to handle the reliability measures. Comparison in terms of the reliability performance to both bus routes has been done. The results show that route 421 has better reliability.

This paper provides two important insights in the travel time reliability study. First, the Burr distribution which handles well in illustrating the distribution of the data, secondly, it is more reliable to use the skew-width methods in explaining the bus route reliability rather than the conventional coefficient of variation. Some extension of the study is expected in the area of travel time modeling in the near future. It is noticeable that bimodal is observed in travel time data. Therefore, mixture distribution could be considered to tackle such problem.

Acknowledgement The authors would like to acknowledge support from Sunway University Internal Grant INT-2018-SMS-01.

Conflicts of Interest The authors declare there is no conflict of interest in this article.

References

[1] H. Al-Deek & E. B. Emam (2006). New methodology for estimating reliability in transportation networks with degraded link capacities. *Journal of Intelligent Transportation System*, *10*(3), 117–129.

[2] I. W. Burr (1942). Cumulative frequency functions. *The Annals of Mathematical Statistics*, *13*(2), 215–232.

[3] H. C. Chu (2010). *Estimating Travel Time Reliability on Freight Corridors*. The National Academies of Sciences Engineering Medicine, Washington D. C. Retrieved from https://trid.trb.org/view/909907.

[4] E. B. Emam & H. Al-Deek (2006). Using real-life dual-loop detector data to develop new methodology for estimating freeway travel time reliability. *Transportation Research Record*, *1959*(1), 140–150.

[5] M. Fosgereau & D. Fukuda (2012). Valuing travel time variability: characteristics of the travel time distribution on an urban road. *Transportation Research Part C-Emerging Technologies*, *24*, 83–101.

[6] M. Fosgereau & A. Karlstrom (2010). The value of reliability. *Transportation Research B*, *44*(1), 38–49.

[7] Y. Hollander & R. Liu (2008). Estimation of the distribution of travel times by repeated simulation. *Transportation Research Part C-Emerging Technologies*, *16*(2), 212–231.

[8] W. C. Jordan & M. A. Turnquist (1979). Zone scheduling of bus routes to improve service reliability. *Transportation Science*, *13*(3), 242–268.

[9] Z. Ma, L. Ferreira, M. Mesbah & S. Zhu (2016). Modeling distributions of travel time variability for bus operations. *Journal of Advanced Transportation*, *50*, 6–24.

[10] A. Polus (1979). A study of travel time and reliability on arterial routes. *Transportation*, *8*(2), 141–151.

[11] A. Richardson & M. Taylor (1978). Travel time variability on commuter journeys. *High Speed Ground Transportation Journal*, *12*(1), 77–79.

[12] R. N. Rodriguez (1977). A guide to the Burr type XII distributions. *Biometrika*, *64*, 129–134.

[13] A. Sumalee, T. Pan, R. Zhong, N. Uno & N. Indra-Payoong (2013). Dynamic stochastic journey time estimation and reliability analysis using stochastic cell transmission model: algorithm and case studies. *Transportation Research Part C: Emerging Technologies*, *35*, 263–285.

[14] S. Susilawati, M. A. P. Taylor & S. V. C. Somenahalli (2013). Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, *47*(8), 720–736.

[15] P. R. Tadikamalla (1980). A look at the Burr and related distributions. *International Statistical Review*, *48*(3), 337–344.

[16] R. Taylor (2017). Travel time reliability: making it there on time, all the time. *U. S. Department of Transportation Federal Highway Administration*. Retrieved from https://ops.fhwa.dot.gov/publications/tt_reliability/ttr_report.htm.

[17] J. W. C. Van Lint & H. J. Van Zuylen (2005). Monitoring and predicting freeway travel time reliability: using width and skew of day-to-day travel time distribution. *Transportation Research Record 1917*, *1917*(1), 54–62. https://doi.org/10.1177/0361198105191700107.

[18] W. J. Zimmer, J. B. Keats & F. K. Wang (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, *30*(4), 386–394.