Kei HIROI, Hitomi IMAI, Nobuo KAWAGUCHI

Dynamic Arrival Time Estimation Model and Visualization Method for Bus Traffic

August 31, 2018

Springer

Acknowledgements

We wish to thank Meitetsu Bus Company Limited for insightful suggestions and provision of bus traffic data. This research and development work was supported by the JST OPERA and the MIC/SCOPE #172106102.

Contents

Chapter 1
Dynamic Arrival Time Estimation Model and Visualization Method for Bus Traffic

Abstract Bus transportation service is more strongly influenced than other public transport modalities by various factors such as traffic congestion, weather conditions, number of passengers, and traffic signals. These factors often cause delays, and users may feel inconvenienced when waiting at a bus stop. Few studies have analyzed the relationship between operational situations and multiple different factors by visualization. Thus, we propose an arrival time estimation method and a visualization model. The arrival time estimation model dynamically updates the accuracy via an estimation method using a combination of a multiple-regression model and a Kalman filter. The visualization model analyzes relationships between delays and various factors. The goal of this study is to realize a society where people can use buses more comfortably.

1.1 Introduction

Many people use public transport in the form of bus service. According to a survey by the Ministry of Land, Infrastructure, Transport and Tourism[13], in Japan, approximately 12 million people use this service every day. Recently, traffic data have begun to be collected by various systems[8], for example, bus arrival information systems. Such systems obtain bus information using GPS: arrival or departure time, traveling locations (latitude and longitude), etc. In addition, the number of passengers and the behavior of the driver are also recorded. On the other hand, bus operation situations are more strongly influenced than other public transport modalities by traffic congestion[12], weather condition[11], number of passengers[6][7], traffic signals[9], etc. These factors[10] are related to delays, and the motion of buses changes in a complicated manner[4][5]. Many services inform the users of departures from a bus stop[14][15], but few services provide specific estimated arrival or delay times. Few studies have analyzed the relationships between operational situations and multiple factors by visualization. When a bus is delayed, passengers may feel inconvenienced when waiting at the bus stop. Thus, we propose an arrival

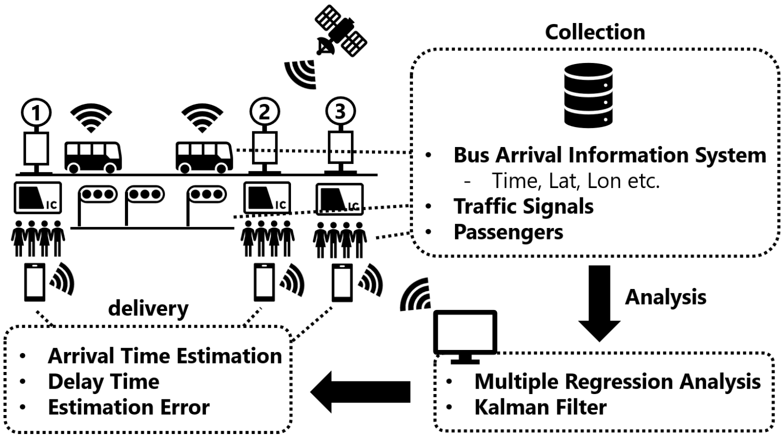


Fig. 1.1 Outline of the Proposed System.

time estimation method and a visualization model: "EMRF (Extended Multiple Regression Filter) model". The arrival time estimation model dynamically updates the accuracy via an estimation method using a combination of a multiple-regression model and a Kalman filter. The multiple-regression model estimates the trend in advance, and the Kalman filter updates the estimation to the optimal state based on the trend in Figure 1.1. As a feature of this method, the closer the bus is to the terminal station, the better the accuracy. "Bus Tapestry" is the visualization model, which analyzes relationships between delays and factors. This model creates a heat map of operational situations (delays or premature arrival) and adds bus stop positions, signal positions, and bus traffic data. We can thereby visually find factors related to delays. The goal of this study is to realize a society whereby people can use buses more comfortably.

1.2 Literature Review

1.2.1 Multiple-Regression Analysis

Multiple-regression analysis is a linear model and derives the dependent variable Y using multiple independent variables $X_i (i = 0, 1, \cdots)$ by Equation 1.1 as follows:

$$Y = a_0 + a_1 X_1 + \cdots + a_i X_i \tag{1.1}$$

where a_i is a coefficient calculated for each independent variable. In the study by Patnaik et al[1], the dependent variable was the estimated time taken between bus stops. Independent variables, such as the time required for the timetable, the distance between bus stops, the number of passengers, and the time to open and close the door, were factors that influenced delays. The estimation under this model was highly accurate. However, multiple-regression analysis is a static estimation based on the past data and does not consider increased passengers due to rainy weather or events being held near the bus stop. Therefore, it is difficult to respond to such a real-time changing environment and present the estimated arrival time to users.

1.2.2 Kalman Filter

The Kalman filter is a powerful mathematical tool for estimating the future states of variables even without knowing of the precise nature of the system modeled. In the study by Chen et al[3], the time required in the next interval was dynamically estimated based on the time required for the timetable and information accumulated from the starting station. Although the Kalman filter can process information including errors and perform estimation dynamically, accurate estimation is difficult when a bus stop interval is characteristic.

1.3 Model Development

1.3.1 Arrival Time Estimation

The EMRF model consists of a multiple-regression model and Kalman filter. Before departure, the multiple-regression model estimates changes in inputs, and after departure, the Kalman filter performs estimation dynamically from the difference between the measured value and the estimated value based on the preliminary estimation.

First, we explain the multiple-regression analysis in this study. The dependent variable is the estimated time taken between bus stops[1]. The independent variables are factors that influence delays[2] such as bus stop sections, delays ahead of n stations, the time zone, the day of the week, the time required for the timetable, and the number of passengers. The delay is defined as the difference between the required time for the timetable and the actual required time. For the time zone, Early Morning is defined as until 7:00, Late Morning is defined from 8:00 to 10:00, Early Noon is defined from 10:00 to 13:00, Late Noon is defined as from 13:00 to 17:00, Evening is defined from 17:00 to 19:00, and Night is defined after 19:00. For the number of passengers, the number of people getting onto the bus is compared with the number getting off, and the higher number is recorded.

Second, we describe our Kalman filter in Figure 1.2. Based on the estimation of the multiple-regression model, the Kalman filter estimates the required time for each bus stop interval. The end point is defined as the bus stop N. At a given bus stop k, the EMRF model estimates the times required for bus stop intervals $k-(k+1)$, $k-$

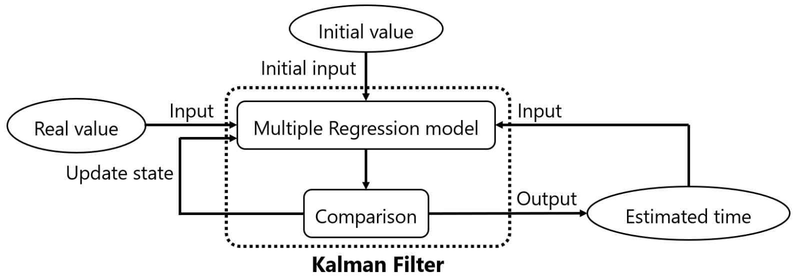


Fig. 1.2 Outline of the Proposed Method.

$(k+2), \cdots, ,k-N$. We input values estimated by a multiple regression model as the initial state of the system. Specifically, for $k=1$, an estimated value is calculated using past data. For $k>1$, the model calculates the real time differences (delay or premature arrival) using the estimated value for the bus stop interval $(k-1)-N$, together with the timetable used in the calculation, and it inputs the results that are re-estimated using the multiple-regression model. After the bus leaves the starting station, the model updates the system status and estimated value each time it arrives at the bus stop and repeats this motion until it reaches the end point. By repeating the update, it is possible to correct the value even if the estimated and actual measured values are different from each other. Additionally, as the bus approaches the end point, the accuracy of the estimation can be improved.

In general, the Kalman filter estimates the state of the system at time $(k+1)$ using the state equation based on the previous state by Equation 1.2 as follows:

$$x_{k+1,j} = \Phi_{k+1} x_{k,j} + u_k + Wk, j \tag{1.2}$$

where $x_{k+1,j}$ is the state of the system at time $(k+1)$, Φ_{k+1} is the state-transition model, u_k is the state vector, and Wk, j is noise. We use a multiple-regression model (Equation 1.1) instead of the state vector u_k in Equation 1.2. The relationship between an observation value z_k and a state variable $x_{k,j}$ is expressed by the observation equation of Equation 1.3.

$$z_k = H_k x_{k,j} + v_{k,j} \tag{1.3}$$

where H_k is the observation model and $v_{k,j}$ is noise. We define the state variable $x_{k,j}$ as the estimated time $E_{k,j}$ and the real time required R_k in Equation 1.4 as follows:

$$x_{k,j} = (E_{k,j}, R_k) \tag{1.4}$$

where $E_{k,j}$ is the total value from an arbitrary bus stop k to the bus stop j and R_k is the total value from the starting station to the bus stop k.

1.3.2 Visualization

Bus Tapestry creates a heat map of operational situations (delays or premature arrival) and adds bus stop positions, signal positions, and bus traffic data. This visualization method attempts to determine and find a tendency of bus delays and their reasons to use large-scale bus traffic data as a scatter plot. The vertical axis is the time zone, the horizontal axis is the distance from the starting to the ending bus stop, and each position is the accumulated distance.

First, we explain how to process the data. Our visualization expresses operational situations by the difference between the actual time required and the required time. The value is positive when the bus is later than the timeline suggests, and it is negative when the bus is earlier than the timeline suggests. The route distance is the total value based on the bus stop and the location information of the signal (latitude and longitude) in Figure 1.3.2. Our method calculates the interval distances d_{n+1} from L_n to L_{n+1} using the information at each position. The total distance L_{n+1} is the sum of their values in Equation 1.5 as follows:

$$L_{n+1} = L_n + d_{n+1} \tag{1.5}$$

In this study, our method does not consider the curvature of the road, for example. We did not process data from the bus arrival information system because these data might include errors. Similarly, the travel distance of the bus is also the sum of the distances between the locations traveled in Figure 1.3.2. For the travel distance, the location of the bus stop is the point at which the departure information was recorded.

Second, we describe how to visualize the data. Our method visualizes operational situations and the signal position using our created data. The heat map represents operational situations (delays or premature arrivals), where the vertical axis is the time zone (hour) and the horizontal axis is the distance (km). The horizontal axis expresses the distance between bus stops by putting the same data every 0.01 km within each bus stop interval. We add the signal position data (mileage) and make the number of traffic signals between routes visible. Consequently, we can analyze the influence of the signals between bus stops and operational situations of each time zone. Furthermore, we can evaluate the operational situations in greater detail by adding the location information of each bus to the visualized data.

1.4 Data Collection

In this study, data were collected from the bus arrival information system. The recorded area is located in the Aichi Prefecture and includes information on position, time, route, bus stop, etc. These data were provided by the Transportation Bureau City of Nagoya[14] and the Meitetsu Bus Company Limited[15] through the Location Information Service Research Agency (Lisra)[16]. The above data were recorded when arriving and departing a bus stop and during communications at 30

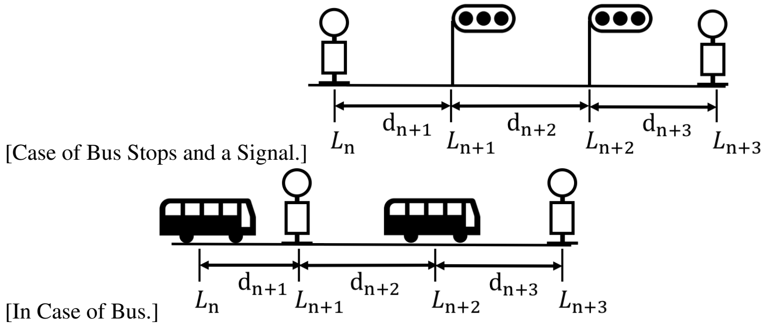


Fig. 1.3 Definition of interval and total distance.

second intervals. The range for the data collection was from December 13-22, 2014. This dataset includes 1030 buses, 3784 bus stops and 664 routes. These data were recorded only when departing the bus stop. The range for data collection was for July 1-15, 2016 and from January through October 2017. This dataset includes 710 buses, 1539 bus stops and 523 routes. The number of passengers was provided by the Meitetsu Bus Company Limited. In addition, we indicate each position of traffic signals on the target bus routes.

1.5 Analysis of Results

1.5.1 Independent Variables and Coefficients of Multiple-Regression Model

We analyze whether the independent variables assumed in 1.3.1 are necessary and sufficient manner. First, we add "The amount of precipitation", "Temperature", "The number of signals in the bus stop sections" and "Interval distance" as the independent variables of our model. Table 1.1 shows a result of the multiple-regression analysis on the data for 1 month in March 2017 and 10 months from January to October 2017 In the multiple-regression analysis with the independent variables in 1.3.1, the coefficient of determination was 0.77 for the one month of data and 0.79 for the 10 months of data. In the multiple-regression analysis using 1.3.1's variables, the amount of precipitation was 0.77 and 0.79. Although the coefficients slightly improved due to the increase in the data volume, we cannot confirm large changes due to the addition of independent variables.

Table 1.1 Independent Variables and Coefficients.

Independent Variables	Coefficients (1 month)	Coefficients (10 months)
-	0.7715	0.7934
The amount of precipitation	0.7716	0.7935
Temperature	0.7715	0.7934
The number of signals	0.7715	0.7934
Interval distance	0.7715	0.7934

1.5.2 Influence of Data Volume on p Value

Figure 1.4 shows the comparison results for the p value calculated by the multiple-regression analysis using the independent values of 1.3.1 and 1.5.1. The independent variables in Figure 1.4 are arranged in descending order of p value for the data for 10 months. The coefficients of "bus stops interval 1-2", "Early Morning" and "Monday" are 0, and the significance level (p value ¡0.05) is represented by a red dotted line. According to Figure 1.4, there are variables with significantly different p values for the data for 1 month and 10 months. Specifically, with regard to "The amount of precipitation", "Temperature", some "bus stop interval", "Time zone", "Day of week", and "Delay in front of 2 stations", the p value using the data for 1 month did not satisfy the baseline value. These independent variables resulted in less influence on the required time than the other factors; however, the p values changed significantly for the 10 months of data, and the p values that met the significance level increased. We think that this is because although these independent variables have minimal influence on the required time in the case using the data for one month, the data for 10 months include the day on which these independent variables work well. However, some independent variables using data for 10 months exceeded the level of significance, with the result that "Tuesday", "Wednesday", "Thursday", and "Friday" had minimal impact. It seems that this was because of the similar operating conditions on weekdays and weekends. In addition, it was found that this situation considered more data, and the effectiveness of the independent variable increases. On the other hand, when using 10 months of data for "The amount of precipitation", "temperature", "number of signals" and "interval distance", no p value remarkably increased compared to other p values. Therefore, these independent variables are thought to affect the required time.

1.5.3 Correlation between the Independent Variables

The added variables of "The amount of precipitation", "temperature", "number of signals" and "interval distance" were found to affect the required time; thus, we investigate the correlation between the independent variables. We calculate the correlation coefficient R for all independent variables assumed in this study, and we

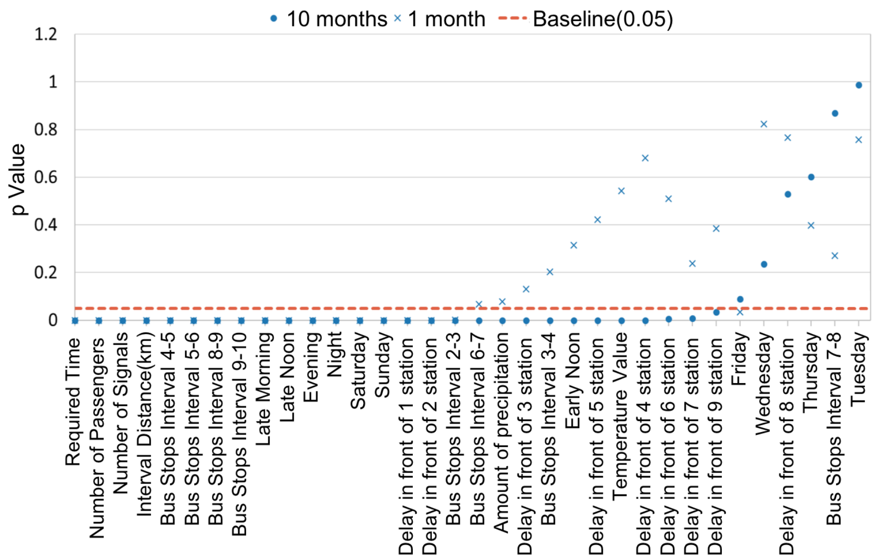


Fig. 1.4 p Value of Each Independent Variable.

show the variables with $|R| > 0.4$ in Figure 1.5. "The delay in front of n stations" showed a strong correlation overall. This seems to be because the delay at the previous bus stop influences the delay of the next bus stop directly. There was also a strong correlation between "The required time" and "Number of signals". This is why no significant change was observed in the decision coefficient even when "number of signals" was added as an independent variable; thus, "the number of signals" could be one of the factors in setting the required time. In addition, the "number of signals" also shows a strong correlation with "interval distance", and this tends to increase the number of signals as the distance increases. "Bus stop interval" sometimes showed strong correlation with other independent variables; however, "Bus stops interval" was a dummy variable and occurred because the values indicated by the factors were biased. However, "Early Noon" and "Late Noon" are dummy variables set under the same condition, and they must be independent variables. Since this time zones were arbitrarily categorized, we thought that the correlation could be weakened by rearranging it according to the bus location data of each area. Therefore, we classified the time zones into 6 classes based on the median, maximum and minimum of the time zone lag using k-means clustering. K-means is a typical non-hierarchical classification method that divides a set of data into l clusters. First, an arbitrary centroid $\mu_i(i = 1,,l)$ is defined in the cluster as an initial value. Then, each point of the data is assigned to the cluster c_i having the closest centroid, and the centroid is updated to the average point of the data included in the cluster. The cluster assignment and the update of the centroid are repeated until the dispersion within the cluster is minimized to calculate the optimum classification result. Figure

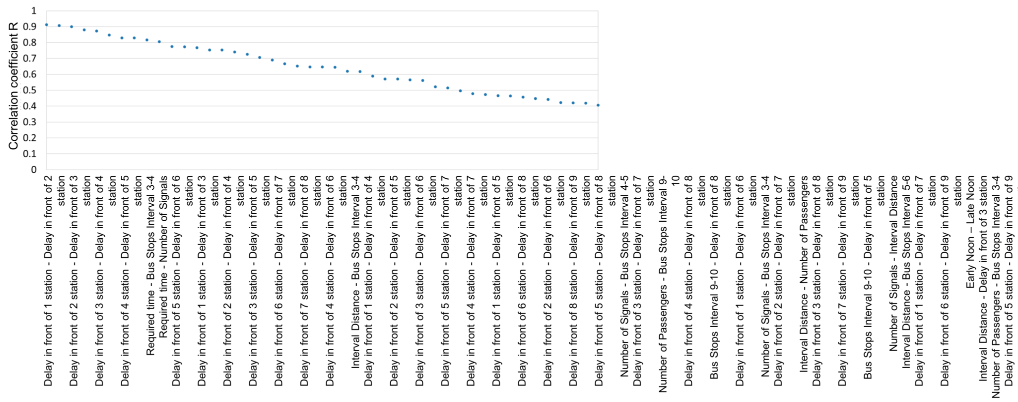


Fig. 1.5 Independent Variables whereby $|R| > 0.4$.

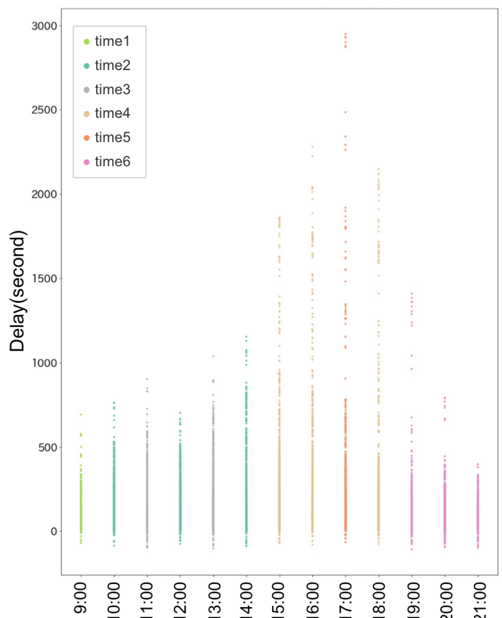


Fig. 1.6 Clustering Result for Time Zone.

1.6 shows the result of the classification of time zones in a route. Using this classification result, the correlations between time zones all achieve $|R| > 0.4$. Then, we removed “Number of signals” from the independent variables and improved the multiple-regression model for “time zone” tailored to each route.

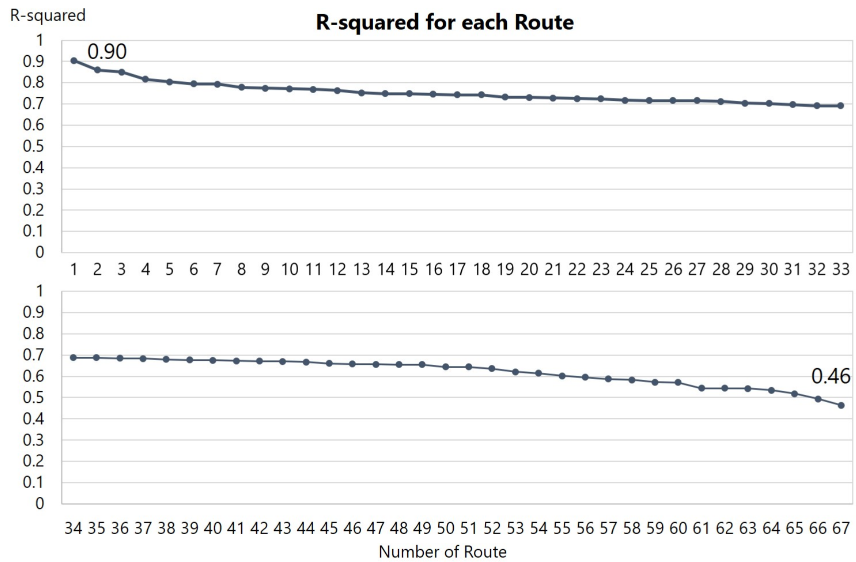


Fig. 1.7 R-squared for Meitetsu Bus.

1.5.4 R-squared by Multiple-Regression Analysis

The results of a multiple-regression analysis are shown in Figure 1.7. The data are from the Meitetsu Bus Company Limited, having a range of March 1-31, 2017. For comparative purposes, the R-squared for the Nagoya City Bus data is collected in Table 1.7. The range of the utilized data is for December 13-19, 2014. The R-squared value indicates how well an independent variable accounts for the variability of another, dependent variable. The value of R-squared ranges from zero to one, with values closer to one indicating a lower degree of relative error. The highest R-squared value was 0.90. However, since the coefficients had abnormally large values, such as 5.86×10^{11}, multiple-regression analysis could not be performed properly. This is because the route contains 25 bus stops, and as such, there are too many independent variables. On the other hand, the smallest R-squared value is 0.46. This seems to be caused by irregular congestion in a bus stop interval. The average value for Meitetsu Bus was 0.69, which was close to the average value of the Nagoya City Bus (0.76). However, the value of Meitetsu Bus was slightly lower than that of the Nagoya City Bus because there was less data on Meitetsu Bus than on Nagoya City Bus. Data of Nagoya City Bus were recorded when arriving and departing the bus stop and when communicating every 30 seconds, but data of Meitetsu Bus were only recorded when departing the bus stop. The relationship between the R-squared and the number of bus stops is shown in Figure 1.8. The R-squared is 0.0053, and there was no correlation between the multiple-regression analysis and the number of bus stops in Figure 1.8.

Table 1.2 R-squared for Nagoya City Bus.

RouteID	R-squared
8415	0.79
8471	0.76
8784	0.69
8921	0.58
8939	0.80
8990	0.80
9014	0.80
9015	0.88
Average	0.76

Table 1.3 Variation of R-squared for Route 9.

The type of data	R-squared
14 days	0.34
101 days	0.44
excluding	0.55

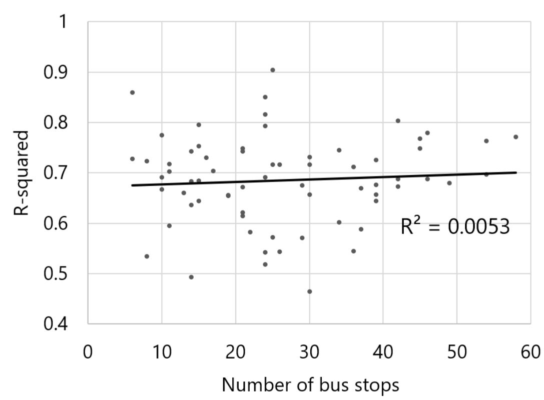


Fig. 1.8 The Relationship between the R-squared and the Number of Bus Stops.

1.5.5 Accuracy Verification by Changing the Amount of Data

We verified how a change in the amount of data affects the estimation accuracy using the multiple-regression model using the data from Meitetsu Bus Company Limited. The compared data were data for 14 days (July 1-14, 2016), 101 days (July 1-14, 2016 and from January to March 2017), and 101 days excluding the abnormal values. The estimated date is July 15, 2016. We removed the abnormal values using the interquartile range. We calculated the R-squared by comparing the estimated and actual values for route 9 (Figure 1.7) in Table 1.3. Table 1.3 shows that the R-squared increased as the amount of data increased. Additionally, excluding the abnormal values further improved the estimation accuracy.

1.5.6 Visualization

We used the data from the Transportation Bureau City of Nagoya on December 16 and 21, 2014. The result of our visualization is presented in Figure 1.10. The black

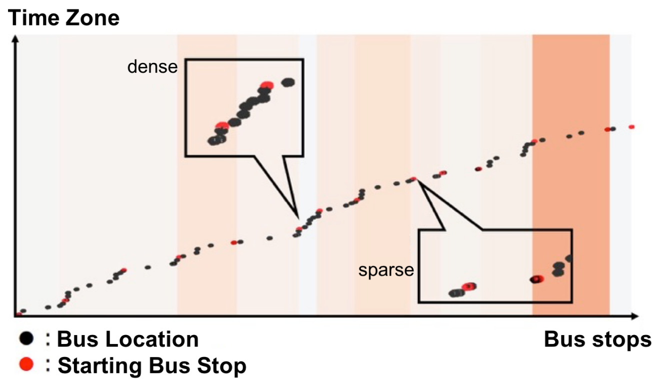
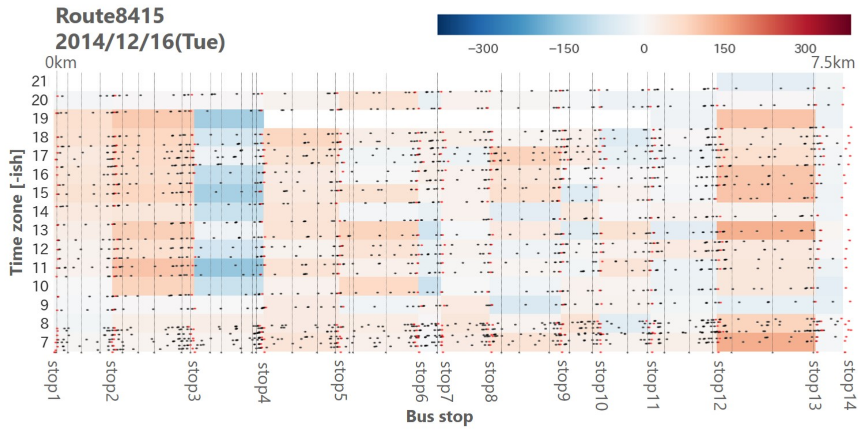


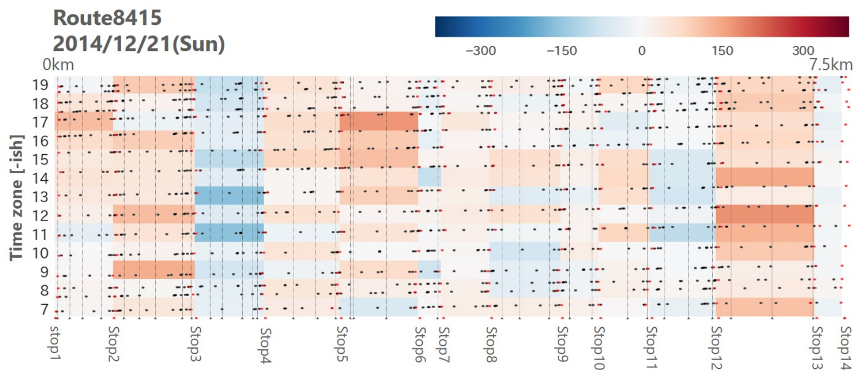
Fig. 1.9 Bus Tapestry Sample.

dots are the running positions of the bus, the red dots are the bus stop positions, and the dotted lines are the signal positions. The white areas in the heat map are the time zones during which the bus was not running. For example, **??** is an enlarged view of the initial time in Figure 1.5.6, where the bus travels along the y axis as time elapses. The points representing the positions of the bus are divided into sparse and dense points, where the bus does not move much when the points are dense but does move when the points are not dense.

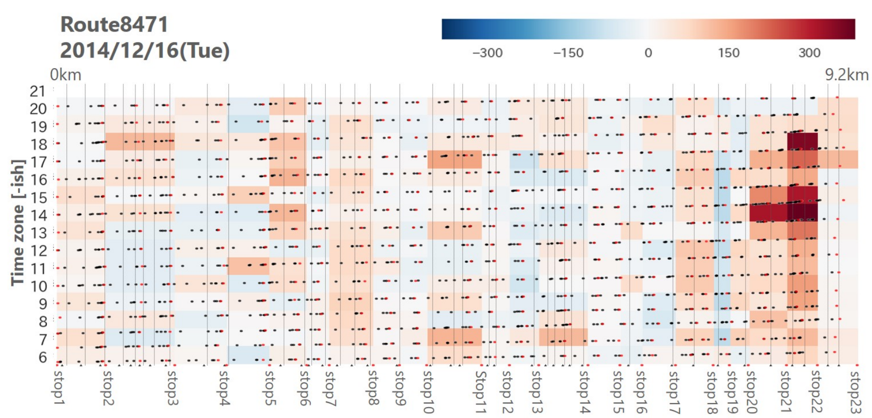
Figure 1.5.6 and Figure 1.5.6 are other days of the same route and show a similar delay condition overall. However, in the range of bus stop 5 to bus stop 6, we find that the delay on Sunday is greater than those on Tuesdays from 16:00 to 17:00. Figure 1.5.6 and Figure 1.5.6 are the other route on the same day. They show that route 8471 has a large delay near three stations before the end point compared to route 8415. Moreover, in Figure 1.5.6, the signal between bus stop 2 and bus stop 3 does not significantly affect the delay because there are few points before it. On the other hand, the signal between stop 11 and stop 12 is likely to affect the delay because there are many points before it. Thus, we can visually identify the relationship between delay and factors using our visualization method.



[Visualization for Route 8415 (Tue).]



[Visualization for Route 8415 (Sun).]



[Visualization for Route 8471 (Tue).]

Fig. 1.10 Result obtained by Bus Tapestry.

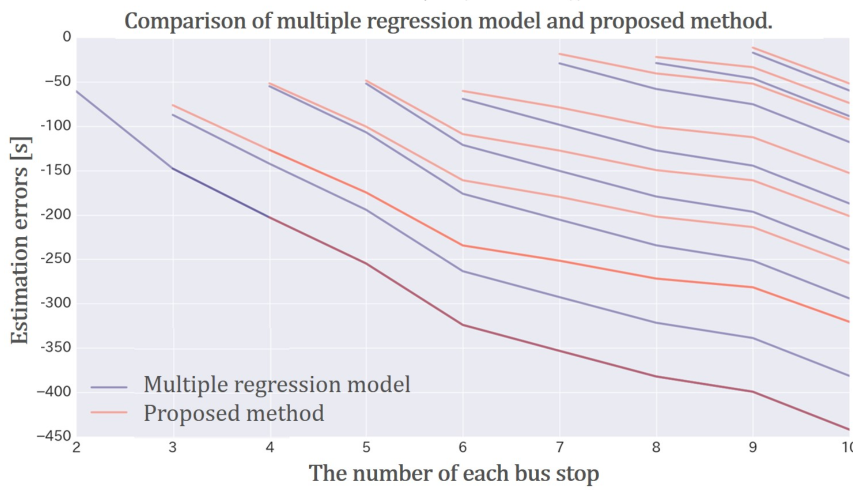


Fig. 1.11 Estimation Error in Route 9, Schedule: 10430.

1.6 Evaluation of the Model

We used the data from Meitetsu Bus Company Limited for March 1-31, 2017 for Route 9. The estimated date is Tuesday, January 31, 2017. The model was created for 30 days, excluding the estimated date. For the estimated date, the number of passengers and delay in front of *n* stations were the average of 30 days.

1.6.1 Comparison by Estimation Errors

For Schedule 10430, the estimation errors by the multiple-regression model and our model are presented in Figure 1.11. Schedule 10430 is a bus running from 20 to 21 hours. The estimation errors are the difference between the estimated value and the actual value. The error is positive when the EMRF model estimates are longer than the actual value and negative when the model estimates are shorter than the actual value. Figure 1.11 shows that the estimation errors are smaller than those of the multiple-regression model, and the EMRF model corrects the estimation.

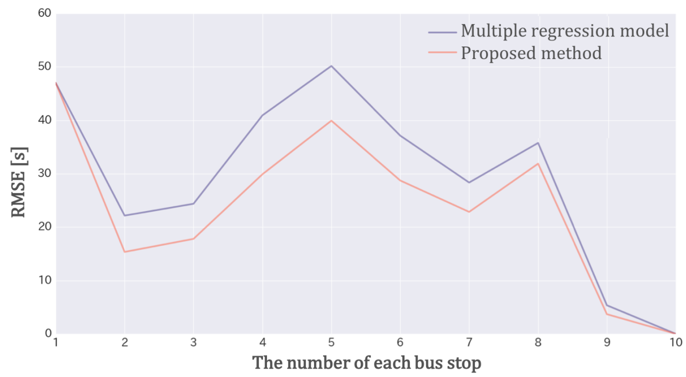
1.6.2 Comparison by RMSE

We evaluate the models using the RMSE (Root Mean Squared Error) in Equation 1.6 as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad [s] \tag{1.6}$$

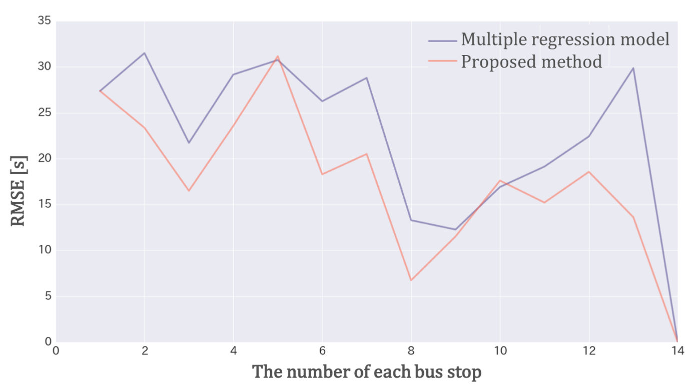
where N is the number of bus stop intervals, y_i is the actual value of the i-th bus stop interval, and $\hat{y}_i$ is the estimated value of the i-th bus stop interval. The RMSE is an evaluation method that quantifies the difference between the estimated value and the actual value. An RMSE closer to 0 indicates a more accurate estimation. 1.12 presents the the RMSE obtained by the multiple-regression model and our model. Schedule 10430 is a bus running from 10 to 11 hours. The estimated date varied from March 1 to 31. The horizontal axis is the estimated bus stop, and the vertical axis is the RMSE. Figures 1.6.2 and 1.6.2, presenting March 31 using data of other dates, show that the RMSE by our model is smaller than that obtained solely by using the multiple regression model. Especially in Figure 1.6.2, the estimation was well corrected at bus stop 2. In Figure 1.6.2, showing March 13, the results are approximately equal to those under the multiple-regression model. Then, we estimated the other route in Figure 1.6.2. Most of the RMSE values under our model showed a higher accuracy than the multiple-regression model except for bus stops 5 and 10. Similarly, for all schedules, the average RMSE is presented in Figure 1.13. This figure shows that the RMSE was smaller in our model even in the case of using the average value for the data of one month. Therefore, it is assumed that our model can improve the estimation accuracy.

[Route 9, Schedule: 10400 (3/31).]



[Route 9, Schedule: 10425 (3/31).]

[Route 9, Schedule: 10400 (3/13).]



[Route 10, Schedule: 11209 (3/31).]

Fig. 1.12 The RMSE for Route 9 and Route 10.

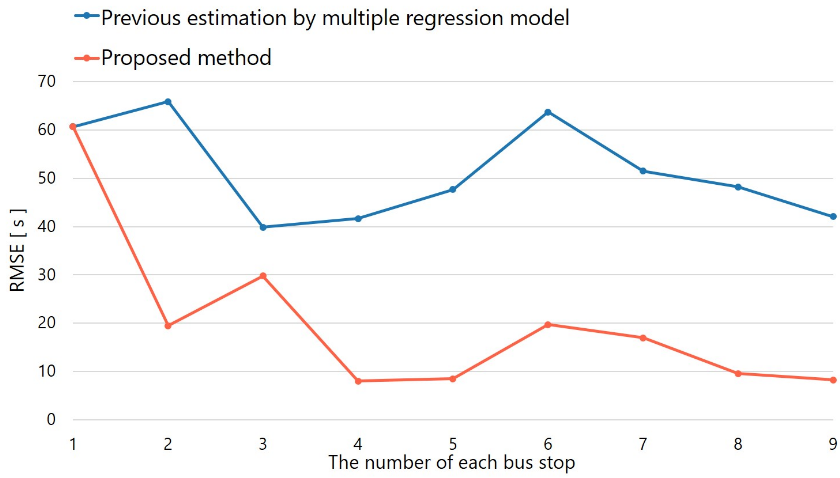


Fig. 1.13 The Average RMSE.

1.7 Examination of Presentation Method

We propose a method for presenting the arrival time including estimation errors. Using the standard deviation, the estimated required time is calculated with some leeway and presented to users with an accuracy of approximately 95% using Equation 1.7 as follows:

$$E' = E \pm 2SD(E - R) \tag{1.7}$$

where E' is the required time including estimation errors, E is the estimated required time, and R is actual required time. Showing users the earliest arrival time allows them to broaden their choice of actions, as in Figure 1.14. For example, users might think "If this time is the earliest possible, let's go to a convenience store" or "Since there is no need to hurry, let's walk slowly". Presenting the latest arrival time has the effect of alleviating the anxiety of "How long will I have to wait at the bus stop?" Our method can also show the estimated arrival time at the destination stop and inform the users of it because our model can be applied to all bus stop intervals. Presenting the specific estimated arrival times in this way gives users a more accurate idea of operational situations, making it easier to act upon such data.

Moreover, to investigate the viewpoints of the users on the presentation method of the estimated arrival time, we performed an investigation using questionnaires. This period was for January 29-30, 2018. We obtained responses from 184 people using SNS, where valid responses from 169 people were obtained. Respondents were asked to evaluate how they viewed estimation errors with 5 responses: "Never", "Hardly ever", "Neutral", "Some of the time", or "All of the time". When the estimation errors are less than 1 minute, "All of the time" accounted for approximately 90 % of all responses. When the estimation errors are within 1-5 minutes,

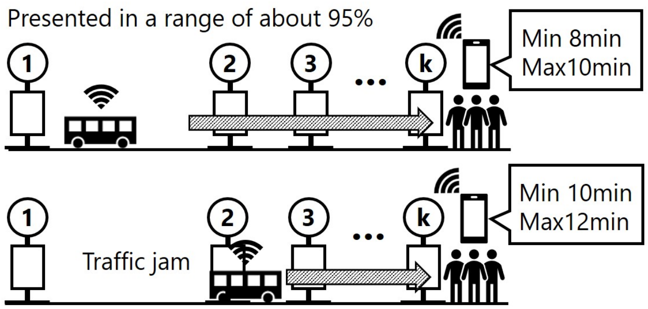


Fig. 1.14 How to Present the Results to Users.

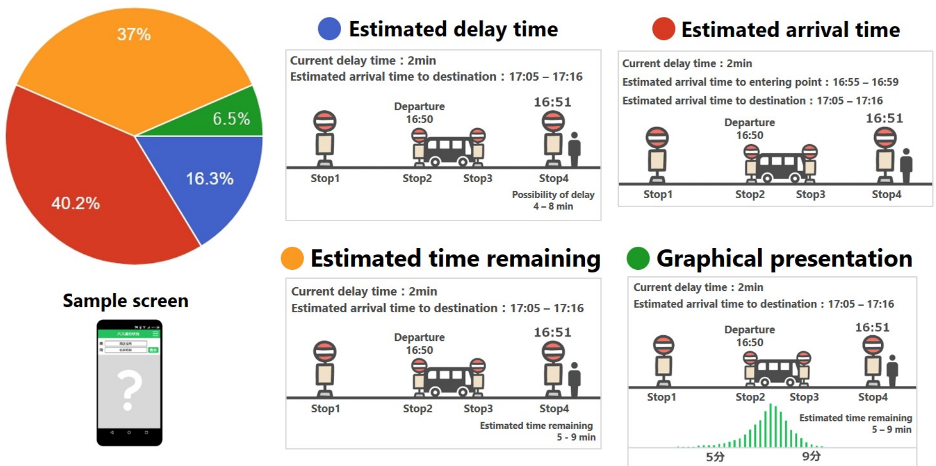


Fig. 1.15 The Result of the Questionnaire about the Presentation Method.

"All of the time", "Some of the time" and "Neutral" accounted for approximately 90 % of all responses. Therefore, it is assumed that the standard for estimation errors is less than 5 minutes. Figure 1.15 shows the results of the questionnaire on the presentation method. There were 4 types of sample types: "Estimated delay time", "Estimated arrival time", "Estimated time remaining", and "Graphical presentation". "Estimated arrival time" and "Estimated time remaining" each accounted for approximately 40 % of all responses. Thus, we found that the users prefer to display the arrival time over the delay time. It is assumed that the users could comfortably use an application in which they could select the presentation method because the answers were divided.

1.8 Conclusion

In this study, we proposed the EMRF model and Bus Tapestry. The EMRF model is a dynamic model for arrival time estimation combining the multiple-regression model and the Kalman filter. We verified the accuracy of the estimation using the R-squared and evaluated the EMRF model by the RMSE. The results showed that the average estimation error improved from 186 seconds to 17 seconds. We also presented the estimated arrival time including estimation errors. We performed an investigation using the questionnaires and obtained 184 answers concerning the presentation method. The results showed that the standard estimation error was less than 5 minutes, and the users preferred to display the arrival time rather than the delay time.

Bus Tapestry is a visualization method for analysis of operational situations. We can visually compare the operational situations of other days or routes and potentially find different features. Additionally, we can see the relationship between delays and number of signals in greater detail. In the future, it may be possible to estimate abnormal values and use machine learning. Furthermore, to start an estimation service, it is necessary to conduct a demonstration experiment and collect the opinions of users.

Acknowledgement

We wish to thank Meitetsu Bus Company Limited for insightful suggestions and provision of bus traffic data. This research and development work was supported by the JST OPERA and the MIC/SCOPE #172106102.

References

1. Jayakrishna, Patnaik Steven Chien, Athanassios Bladikas: Estimation of Bus Arrival Times Using APC Data, Journal of Public Transportation, Vol. 7, No.1, pp.1-20 (2004)
2. Alejandro Tirachini: Estimation of Travel Time and the Benefits of Upgrading the Fare Payment Technology in Urban Bus Services, Transportation Research Part C: Emerging Technologies, Vol.30,pp.239-256 (2013)
3. Mei Chen, Xiaobo Liu, Jingxin Xia: Dynamic Prediction Method with Schedule Recovery Impact for Bus Arrival Time, Transportation Research Record Journal of the Transportation Research Board, pp.208 - 217 (2005)
4. Takashi Nagatani: Dynamical Transitions to Chaostic and Periodic Motions of Two Shuttle Buses, Physica A:Statistical Mechanics and its Applications,Vol.319,No.1, pp.568 - 578 (2003)
5. Takashi Nagatani: Chaos Control and Schedule of Shuttle Buses, Physica A:Statistical Mechanics and its Applications, Vol.371,No.2,pp.683 - 691 (2006)
6. Chen Zhang, Jing Teng: Bus Dwell Tme Estimation and Prediction : A Study Case in Shanghai-China, Procedia - Social and Behavioral Sciences,Vol.96,pp.1329 - 1340 (2013)

7. Amer Shalaby, Ali Farhan: Prediction Model of Bus Arrival and Departure Times Using AVL and APC data, Public Transportation,Vol.7,No.1,pp.41-61 (2004)
8. Bo Yu, Jing Lu, Bin Yu, Zhongzhen Yang: An Adaptive Bus Arrival Time Prediction Model, the Easten Asia Society for Transportation Studies", Vol.8,pp.1126 - 1136(2010)
9. Chin-Woo Tan, Sungsu Park, Hongchao Liu, Qing Xu, Peter Lau: Prediction of Transit Vehicle Arrival Time for Signal Priority Control:Algorithm and Performance, IEEE Transactions on Intelligent Transportation System, Vol.9,No.4,pp.688 - 696 (2008)
10. Andrei Iu, Adrian Friday: Statistical Modelling and Analysis of Spare Bus Probe Data in Urban Areas, International IEEE Annual Conference on Intelligent Transportation Systems Madeira Island, pp.1256-1263 (2010)
11. Victor W. Stover, Edward D. McCormack: The Impact of Weather on Bus Ridership in Pierce County, Washington, Journal of Public Transportation,Vol.15,No.1,pp.95-110 (2012)
12. Michael L. Anderson: Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion, American Economic Review,vol.104,No.9, pp.2763-2796(2014)
13. Ministry of Land, Infranstructure, Transport and Tourism, http://www.mlit.go.jp/en/index.html. Cited January 2018
14. Toransportation Bureau City of Nagoya, http://www.kotsu.city.nagoya.jp/en/pc/OTHER/TRP0001448.htm. Cited January 2018
15. Meitetsu Bus Company Limited, http://www.meitetsu-bus.co.jp/english/index.html. Cited January 2018
16. Location Information Service Research Agency, http://lisra.jp/en. Cited January 2018