International Journal Of Engineering Research And Development
e-ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 13, Issue 11 (November 2017), PP.47-59

Determination of The optimal Aggregation Interval Size for Each Time Period of Individual Vehicle Travel Time Collected Using DSRC in The Interrupted Traffic Flow Section

*Hyun Suk Park
Korea Institute of Civil Engineering and Building, 283 Goyangdae-Ro Goyang-Si, Korea.
Corresponding Author: Hyun Suk Park

ABSTRACT:- The National Highway ITS utilizes the daily fixed 5-minute interval, which is the result of the study of the optimal aggregation interval size of data for the daytime hours in sections in which individual vehicle travel time shows a normal distribution. The characteristics of the individual vehicle travel times in the interrupted traffic flow section of national highway differ from those of the continuous traffic flow section, which include that in the interrupted traffic flow section of national highway, the individual vehicle travel times often show a non-normal distribution in the case of non-congestion, and that the number of collections of DSRC decreases significantly at night. Thus, it is necessary to identify the optimal aggregation interval size that fits the situation.

In this study, I determined the optimal aggregation interval size at which the error is minimized, by varying the aggregation interval size of individual vehicle travel time and estimating the mean square error (MSE), in order to determine the optimal aggregation interval size of data for a representative interrupted traffic flow section of national highway, which shows a bimodal asymmetric distribution of individual vehicle travel times during non-congestion. The maximum estimation error equation of t-distribution, which can be used even in asymmetric distribution, was used as the bias estimation equation for MSE calculation.

I calculated the optimal aggregation interval size at which the MSE is minimized, for a representative interrupted traffic flow section, using the maximum sampling error equation of t-distribution (r = 68%). The analysis results showed that the optimal aggregation interval size was 1 minute for morning and afternoon hours (7:00 to 22:00) and 5 to 20 minutes for transition hours (6:00 to 7:00, 22:00 to 24:00). The optimal aggregation interval size increased to 25 to 30 minutes at dawn hours (0:00 to 6:00) with a decrease in the number of collections. The optimal aggregation interval size for morning and afternoon hours (7:00 to 22:00) changed to 3 to 5 minutes when excluding the time interval size (1 to 2 minutes) at which data is always missing due to a stop signal. It was analyzed that in order to minimize the calculation error of the representative travel time value and maintain the reliability of the information, the current, daily fixed 5–minute aggregation interval size needs to be extended up to 30 minutes for dawn hours.

Keywords:- DSRC, estimation of the travel time, interrupted traffic flow section, ITS, optimal aggregation interval size of data

Date of Submission: 22 -11-2017 Date of acceptance: 30-11-2017

I. INTRODUCTION

The national highway ITS (Intelligent Transport Systems) of the metropolitan area was established in 1997 in order to reduce traffic congestion, starting the pilot project for a 30 km section between Seongnam and Icheon in National Route 3. It is currently in operation in the road of 827.1km, which takes up about 65% of the national highway in the metropolitan area [8]. 5.8GHz DSRC (dedicated short range communication), which was first introduced for the automatic fare collection of expressways, was introduced to the national highway ITS in 2007. As the number of vehicles with OBU (on-board unit) rapidly increased, a pilot project of the national highway ITS using DSRC was started in 2009 for a 27.6km section between Goyangsigye and Jayu IC in National Route 77 [4]. DSRC has proven to be more stable in terms of equipment operation and is cheaper in cost than previous AVI section detectors. Consequently, since 2010, its use has been expanded, including substitution of DSRC for older collection equipment. As of 2017, it is installed in a 425km section of National Route 1 and 10 other routes at intervals of about 1 to 5km and is in operation as a main collection system for ITS traffic information generation [5]. The representative value of the individual vehicle data collected by DSRC should be calculated in order to provide the data to traffic information users. Here, it is important to determine the aggregation interval size at which the error is minimized because the reliability of the representative value varies depending on the aggregation interval size of data. The individual vehicle travel times at the interrupted traffic flow section in national highway mostly follow a non-normal distribution due to waiting for signal. However, the travel time aggregation interval size for the calculation of the representative value thereof is fixed at 5-minute interval, which is the optimal aggregation interval size of the individual vehicle travel time of the continuous traffic flow section, which follows a normal distribution, because no study has been conducted to determine the optimal aggregation interval size applicable to the interrupted traffic flow section. This reduces the reliability of the traffic information. The interrupted traffic flow section differs from the continuous traffic flow section in the travel time arrival pattern. The difference includes a repeated pattern during non-congestion in which a vehicle does not arrive due to a stop signal and the disappearance of the repeated pattern of data loss during congestion. Thus, the optimal aggregation interval size of the interrupted traffic flow section should differ from the continuous traffic flow section. Further, the number of collections of DSRC significantly decreases at night, so that the travel time arrival pattern of individual vehicles in the interrupted traffic flow section at night are much more different from the normal distribution of individual vehicles in the continuous traffic flow section. I seek to improve the reliability of the representative value of the travel times in the interrupted traffic flow section collected using DSRC, by proposing the optimal aggregation interval size of the interrupted traffic flow section, which shows a bimodal non-normal travel time distribution, according to congestion status.

II. REVIEW OF PREVIOUS STUDIES

A. Concept of Error according to Aggregation Interval Size [3]

If the parameter is defined as θ and the estimator of θ is defined as $\hat{\theta}$, when the expectation of the estimator is taken, an estimator identical to the parameter is referred to as an unbiased estimator or an unbiased estimate, which satisfies the relation of Equation (1):

$$E(\hat{\theta}) = \theta \qquad (1)$$

In general, the estimator ($\hat{\theta}$) of field DSRC collection data may be similar to the population mean (θ). However, generally, it is a biased estimator because it does not satisfy Equation (1) due to sampling errors and residual outliers. If I have two unbiased estimators $\hat{\theta}_1$, $\hat{\theta}_2$ for θ, I can consider the variance as a measure for estimating the more effective estimator among the two estimators. The smaller the variance of $\hat{\theta}$ is, the better the estimator is, because the values are densely distributed around θ. The unbiased estimator with the minimum variance is called the best unbiased estimator or the minimum variance unbiased estimator. If bias is allowed as in the estimators of DSRC collection data, it is not meaningful to compare the variances only, and the effectiveness should be evaluated by using the mean square error (MSE), which is the sum of the variance and the bias2, as in Equation (2). That is, the aggregation interval size with the minimum MSE can be determined as the optimal aggregation interval size, after calculating and comparing the MSEs of the estimators of DSRC data for each aggregation interval size.

$$MSE[\hat{\theta}] = E[\hat{\theta} - \theta]^2 = E[\{\hat{\theta} - E(\hat{\theta})\}^2]+\{E(\hat{\theta})- \theta\}^2 = V[\hat{\theta}] + [B(\hat{\theta})]^2 \qquad (2)$$

If an unbiased estimator is applied to the MSE equation, the bias becomes 0 and the equation becomes an equation for calculating an unbiased estimator with the minimum variance.

B. Study of Determination of Optimal Travel Time Aggregation Interval Size

Park (2000) defined the aggregation interval size at which the mean square error (MSE) is the minimum with the sum of variance and square of bias in the method of setting the optimal aggregation interval size for travel time estimation and prediction. Using the AVI data collected from the US Houston Urban Expressway (US-290) and the Gaussian kernel method, the MSE of each aggregation interval size was calculated. According to the results, the optimal aggregation interval size was 5 minutes. The study suggested that the optimal aggregation interval size may change as the degree of congestion increases [7]. Lim (2005) proposed an optimal aggregation interval size using the AVI section data for the sections in rural highway showing a normal distribution of travel time. The method of moment, the maximum likelihood method, and the F-test was used to calculate the MSE for each aggregation interval size. As a result, the method of moment and the maximum likelihood method yielded the optimal aggregation interval size of 5 minutes. The F-test yielded 1 minute, 5 minutes, and 30 minutes depending on the time period, showing a significant variation according to the time period. Thus, the study analyzed that the optimal aggregation interval size obtained by the method of moment and the maximum likelihood method has more explanatory power [6].

Choi (2012) calculated the daily fixed aggregation interval size at which the MSE is minimized, for the section of the Gangwon province national highway ITS which shows a normal distribution of travel time, using the daily data of AVI section detectors and the method of moment. The study analyzed that as the daily traffic volume is smaller, rural highway has a relatively longer aggregation interval size than urban highway. However, sections with the traffic volume of 5,000 vehicles or less per day are excluded in the analysis due to difficulties in application of statistical techniques [1]. In addition, studies on the optimal aggregation interval size for sections with spot detectors, not section detectors, were conducted for expressways using the CVMSE method and the F-test. In these studies, the aggregation interval size at which the MSE is minimized was determined as the optimal aggregation interval size [2][11].

Park (2017) estimated the optimal aggregation interval size at which the MSE is minimized in the interrupted traffic flow section of national highway during the time period between 8:00 and 21:00 except for night time and dawn hours to be 3 to 5 minutes. The study analyzed that during congestion or transition hours, at which the average travel time slightly increases, the optimal aggregation time decreases to 3 minutes. The study suggested that considering the efficiency of operation and the improvement of the reliability of the calculation of the representative travel time value, the aggregation interval size should be basically maintained at 5 minutes as usual, and reduced to 3 minutes in the case of congestion [9].

C. Limits of Previous Studies

Most of the previous studies have determined the optimal aggregation interval size of data only for the sections where travel time shows a normal distribution. It cannot be used as the aggregation interval size of the interrupted traffic flow section, which shows a bimodal asymmetric travel time arrival distribution due to a stop signal. The study [9] conducted on the section where travel time shows a non-normal distribution, which excluded night time and dawn hours, also failed to suggest the optimal aggregation interval size for the total operation hours. The aggregation interval size of data of national highway ITS center is fixed at 5 minutes per day, which is the same as the results of previous studies [6][7]. When considering the characteristic of the interrupted traffic flow section as discussed in the previous chapter, which is that the travel time distribution differs between off-peak hours and peak hours, the aggregation interval size of data for reliable calculation of the representative travel time value needs to be different depending on the traffic condition.

III. DETERMINATION OF OPTIMAL TRAVEL TIME AGGREGATION INTERVAL SIZE

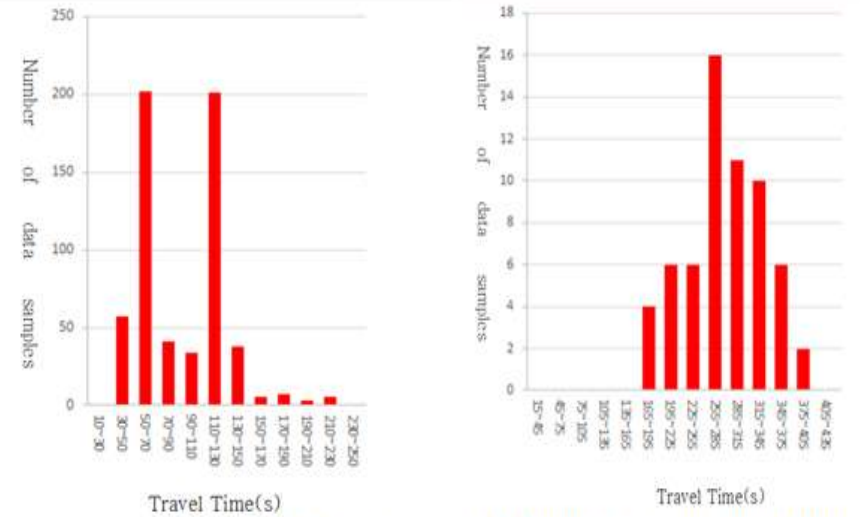
As shown in Table 1, among the sections of the general national highway ITS of the metropolitan area, the section at which DSRC is used to collect traffic information covers a total of 425km of 10 national routes in addition to National Route 1. The section has 219 DSRCs and 437 intersections. The average DSRC section length is 2 km and ranges from about 1 to 6 km when calculated for each route. If I exclude National Route 46 and National Route 77, which are continuous traffic flow sections, from the routes, the average DSRC section length is reduced to about 1 to 2 km. Analysis of the number of intersections in the DSRC section showed that the average number of intersections is two, but that the average number of intersections in the DSRC section of National Route 1 and National Route 3, which are main routes of the metropolitan area and where congestion occurs most frequently, is four [9].

Table 1: DSRC Section of National Highway in Metropolitan Area

Line	Section Length (km)	Number of DSRC	Number of intersections	DSRC Section Length (km)	Number of intersections in DSRC Section
1	35	24	81	1.5	4
3	62	30	102	2.1	4
17	30	14	13	2.3	1
37	2	2	5	2.0	5
38	96	53	106	1.9	2
39	48	23	22	2.2	1
42	51	27	54	2.0	2
43	35	23	35	1.6	2
45	15	9	19	1.9	2
46	21	8	0	3.0	0
77	29	6	0	5.8	0
Sum or Average	425	219	437	2.0	2

Section T030070 of National Route 3, which has a similar DSRC section length and average number of intersections within section to the averages of general national highway in the metropolitan area, and where congestion occurs frequently, was selected as the representative section for determining the optimal aggregation interval size of individual vehicle travel times to be collected in the interrupted traffic flow section of national highway. As shown in Fig. 1, section T030070 ranges from 12003DSE00402 (Sadong-ri, Icheon city) to 12003DSE00502 (Ami-ri, Icheon city) and has an extension of about 1km. As shown in Fig. 2, it is a typical, interrupted traffic flow section of general national highway, which exhibits a clear bimodal asymmetric travel time distribution in the case of non-congestion due to a stop signal.

Fig. 1: Status of Target Section



(a) During a off-peak period in T030070 (14:30-16:30 on June 15, 2015)

(b) During a peak period in T030070 (18:40-18:50 on June 15, 2015)

Fig. 2: Variation of the Distribution of Individual Vehicle Travel Time Depending on the Presence of Congestion

A. Determination of Mean square Error Bias Estimation Equation

The effectiveness of the optimal travel time aggregation interval size is evaluated by using the mean square error (MSE), which is the sum of the variance and the bias2, as in Equation 2 in the previous chapter. For the bias estimation equation for MSE calculation, the maximum estimation error equation of t-distribution, which Park (2017) applied because it can be used for sections whose travel time shows a non-normal distribution, rather than the method of moment, which is applicable only to sections whose travel time exhibits a normal distribution. The maximum estimation error of t-distribution is used because travel times whose population variance is unknown and which exhibit a non-normal distribution have an approximate t-distribution [10]. Because the value depends on the confidence level, the bias point estimation equation is completed by comparing with the MSE true values obtained by the parametric survey and determining a confidence level which has the same MSE ranking at each aggregation interval size and has a high correlation.

During the off-peak hours (13:50 to 15:50 on July 24, 2015) of T030070, the representative interrupted traffic flow section, the vehicle license plates that passed through the start point/end point of DSRC were surveyed. The MSE true value for each aggregation interval size was determined using the travel time matching each individual vehicle. Using the DSRC collection data as the input data, the confidence level (68%) with standard error ($\sigma/\sqrt{n}$), which is the standard deviation of the estimator, and $2\sigma/\sqrt{n}$ confidence level (95%) were applied to the maximum sampling error equation of t-distribution, and the results were compared as shown in Fig. 3. For the comparison of MSEs under the same conditions, I used and compared the data of the intervals of 2 minutes or more, which does not have missing data caused by the decrease of the number of collections.

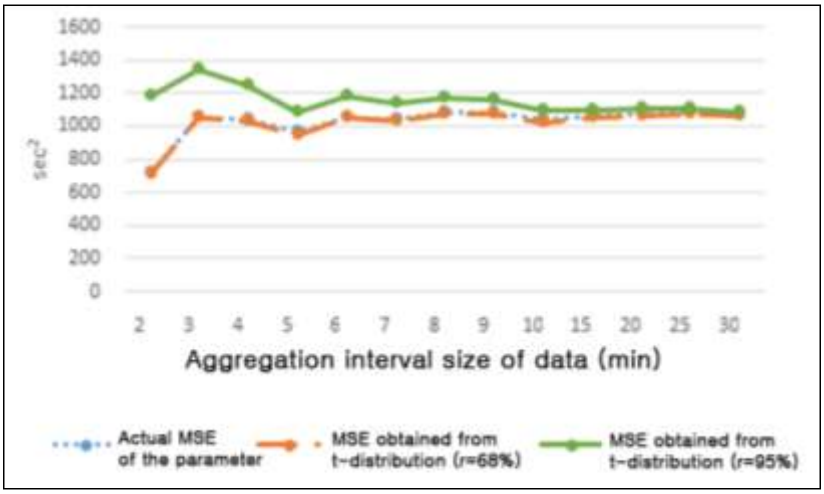


Fig. 3: MSEs of True Value vs. t-distribution (r = 68%, 95%) for Each Aggregation Interval Size

When the confidence level was 68%, the increase and decrease pattern of MSE for each aggregation interval size was considerably similar to the true value, and the correlation coefficient (R^2) with the true value was 0.99, which is considerably high, as shown in Fig. 4. Thus, the maximum sampling error equation of t-distribution with 68% confidence level as shown in Equation (3) was determined as the sampling error equation of the bimodal asymmetric travel time distribution.

$$\text{Sampling error } (\mu - \overline{X}\,) = \left| \, t_{0.32}(n-1)\frac{s}{\sqrt{n}} \right| \tag{3}$$

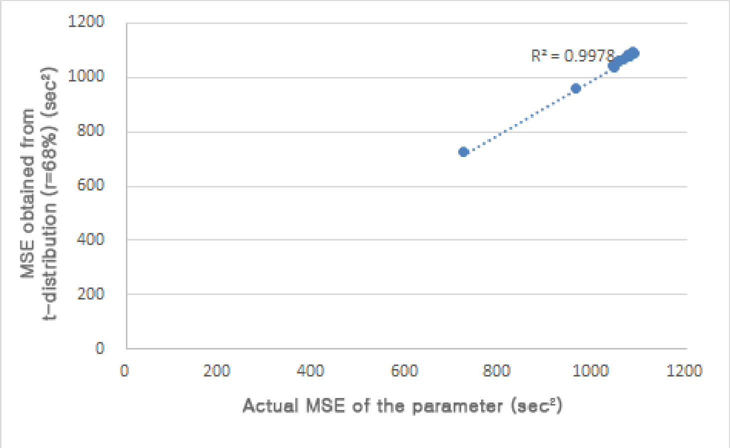


Fig. 4: Correlation Coefficient (R^2) of MSEs of True Value vs. t-distribution (r = 68%)

The MSE estimation equation is based on the deviation between the mean value and individual data, and thus allows MSE estimation if the number of collections is three or more. Thus, as shown in Table 2, MSE of 1-minute aggregation interval size with the number of collections of 4.7 was also compared. According to the result, it is analyzed that although there is a slight deviation from the actual MSE of the parameter in comparison with other aggregation interval sizes without missing data as shown in Fig. 5 and Fig. 6, it can be used for the MSE ranking comparison for each aggregation interval size.

Table 2: The Average Number of Collections and the Status of Missing Data for Each Aggregation Interval Size (average number of collections/aggregation interval size)

1 minute	2 minutes	3 minutes	4 minutes	5 minutes	6 minutes	7 minutes
4.7	9.42	14.1	18.8	23.5	28.3	33
8 minutes	9 minutes	10 minutes	15 minutes	20 minutes	25 minutes	30 minutes
37.7	42.4	47.1	70.6	94.17	117.7	141.3

※ The shaded area denotes the aggregation interval sizes where data is missing at least once

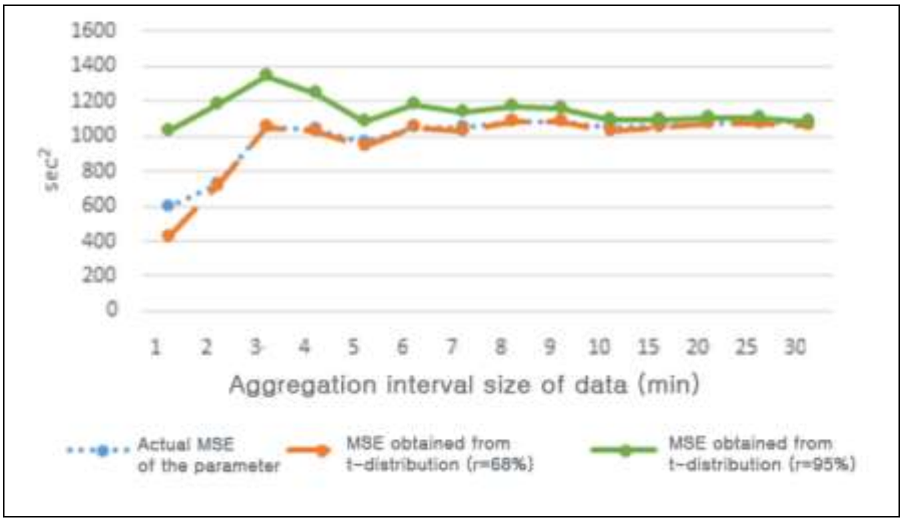


Fig. 5: MSEs of True Value vs. t-distribution (r = 68%, 95%) for Each Aggregation Interval Size (the number of collections of three or more)

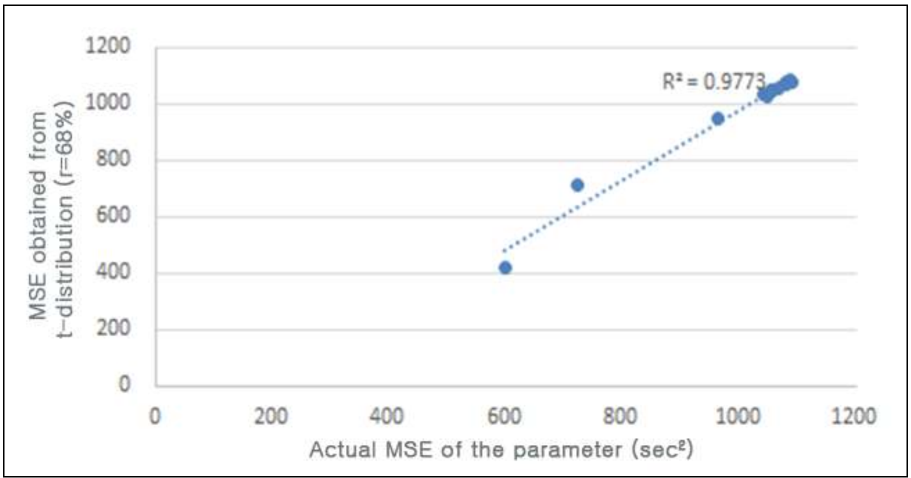


Fig. 6: Correlation Coefficients (R^2) of True Value vs. t-distribution (r = 68%) for Each Aggregation Interval Size (the number of collections of three or more)

B. Determination of the Optimal Aggregation Interval Size of Data

1) Status of Missing Data and Congestion for Each Time Period Considering the reliability of the analysis result, the analysis of the optimal aggregation interval size of section T030070 was performed on the data of the two weeks before the parameter survey date (Friday) of July 24. However, data of July 3 was used instead of that of July 10 because on July 10, data were missing during daytime hours due to system abnormality, not a decrease in the number of collections.

As shown in Table 3, the aggregation interval size with the number of collections of three or more was 1 to 2 minutes or more at 06:00 to 23:00, which is daytime and night time hours, when congestion may occur due to an increase of traffic volume, and the aggregation interval size increased up to 3 to 20 minutes at other hours. The analysis of the status of congestion for each time period showed that congestion in which the average travel time increased up to 450 seconds or more occurred at 17:00 to 20:00, and that the average travel time intermittently increased to 150 seconds or more at 13:00 to 15:00, as shown in Fig. 7.

Table 3: Aggregation time interval for each time period with the number of collections of three or more

Time	00-01 o'clock	01-02 o'clock	02-03 o'clock	03-04 o'clock	04-05 o'clock	05-06 o'clock
Aggregation interval size	more than 6minutes	more than 15minutes	more than 15minutes	more than 20minutes	more than 20minutes	more than 6minutes
Time	06-07 o'clock	07-08 o'clock	08-09 o'clock	09-10 o'clock	10-11 o'clock	11-12 o'clock
Aggregation interval size	more than 2minutes	more than 1minute	more than 1minute	more than 1minute	more than 1minute	more than 1minute
Time	12-13 o'clock	13-14 o'clock	14-15 o'clock	15-16 o'clock	16-17 o'clock	17-18 o'clock
Aggregation interval size	more than 1minute	more than 1minute	more than 1minute	more than 1minute	more than 1minute	more than 1minute
Time	18-19 o'clock	19-20 o'clock	20-21 o'clock	21-22 o'clock	22-23 o'clock	23-24 o'clock
Aggregation interval size	more than 1minute	more than 1minute	more than 1minute	more than 1minute	more than 2minutes	more than 3minutes

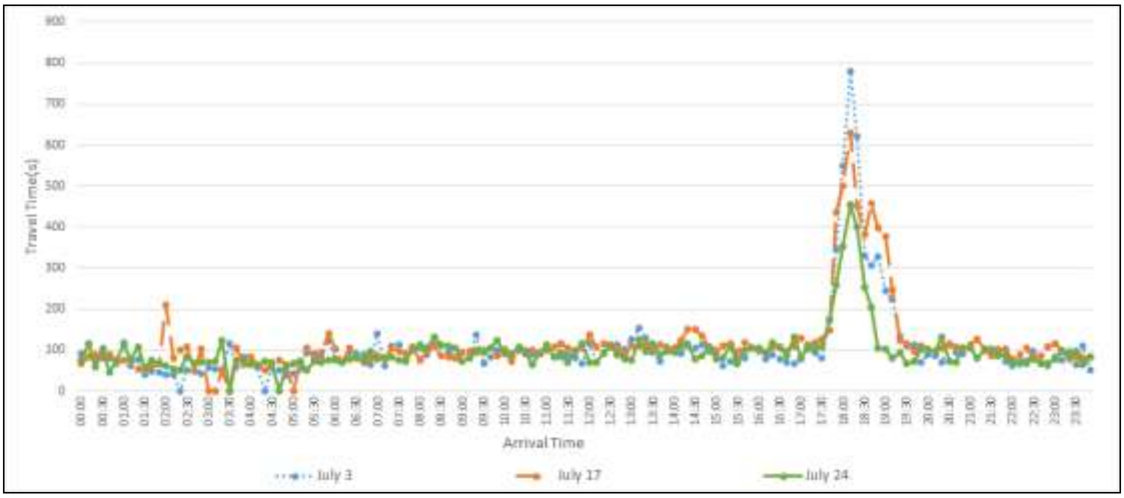


Fig. 7: Change of Travel Time

2) Determination of the Optimal Aggregation Interval Size for of Data Each Time Period

The MSE of the aggregation interval size with the number of collections of three or more was calculated for each time period. The result showed that the optimal aggregation interval size at which the MSE is minimized was 25 to 30 minutes at the dawn hours of 0:00 to 6:00, and that the optimal aggregation interval size started to decrease to 20 minutes from 6:00, at which the number of collections began to increase, as shown in Fig. 8 and Table 4. From 7:00 to 22:00, the minimum aggregation interval size of 1 minute was the optimal aggregation interval size, and from 22:00 till midnight, when the number of collections decreased again, the aggregation interval size increased to 5 to 10 minutes.

However, the dawn hours showed a pattern in which the MSE of the short aggregation interval size increases sharply from 22:00, when the number of collections decreases, and the MSE of the short aggregation interval size decreases again from 6:00, when the number of collections increases, although some of the short aggregation interval sizes are excluded in the analysis. Thus, it is analyzed that the optimal aggregation interval size at dawn hours increases, unlike daytime.

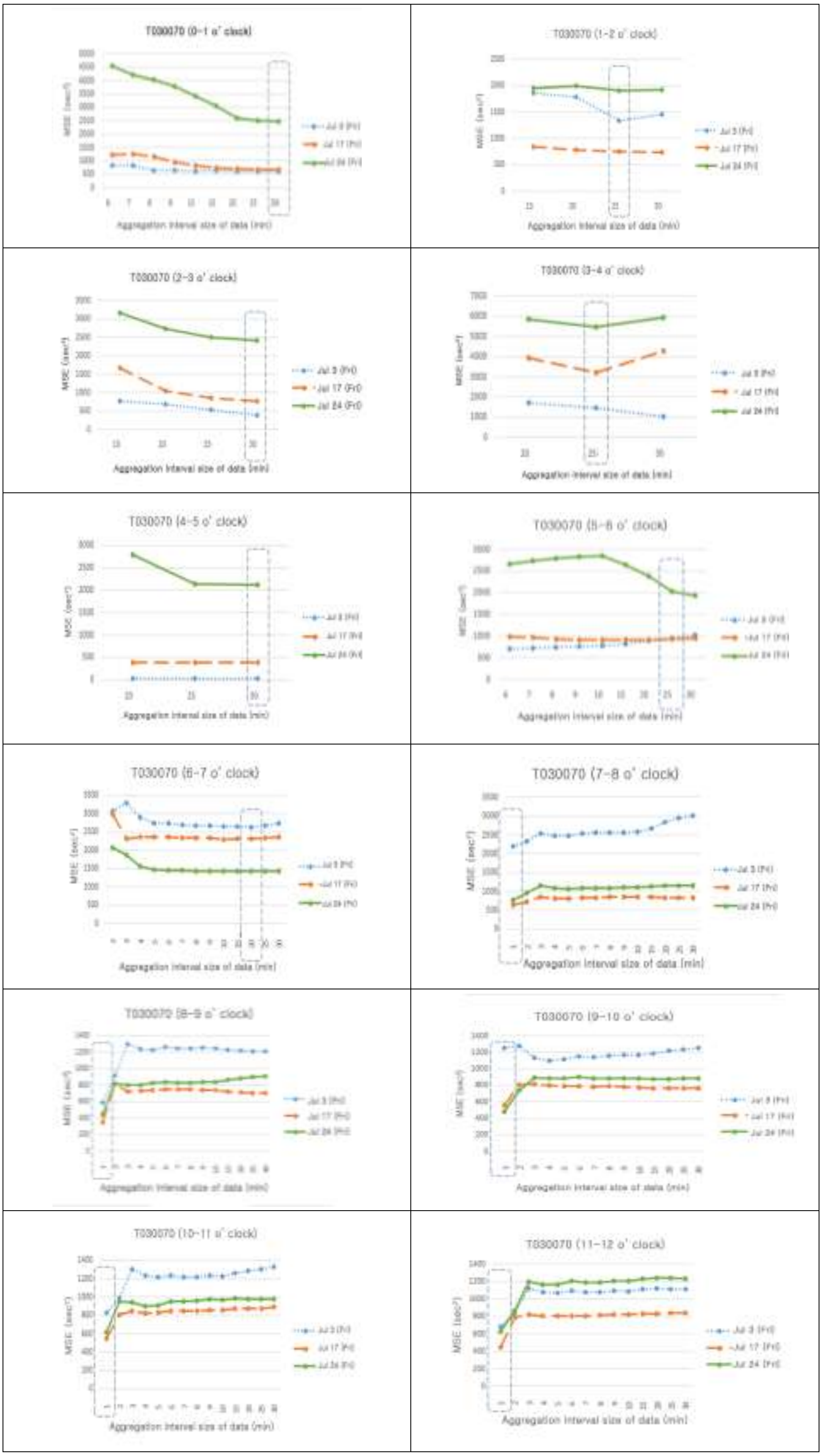


Fig. 8: MSE of Aggregation Interval Sizes for Each Time Period

Fig. 8: MSE of Aggregation Interval Sizes for Each Time Period (continued from above)

Table 4: Optimal Aggregation Interval Size of Data for Each Time Period

Time	00-01 o'clock	01-02 o'clock	02-03 o'clock	03-04 o'clock	04-05 o'clock	05-06 o'clock
Aggregation interval size	30 minutes	25 minutes	30 minutes	25 minutes	30 minutes	25 minutes
Time	06-07 o'clock	07-08 o'clock	08-09 o'clock	09-10 o'clock	10-11 o'clock	11-12 o'clock
Aggregation interval size	20 minutes	1 minute	1 minute	1 minute	1 minute	1 minute
Time	12-13 o'clock	13-14 o'clock	14-15 o'clock	15-16 o'clock	16-17 o'clock	17-18 o'clock
Aggregation interval size	1 minute	1 minute	1 minute	1 minute	1 minute	1 minute
Time	18-19 o'clock	19-20 o'clock	20-21 o'clock	21-22 o'clock	22-23 o'clock	23-24 o'clock
Aggregation interval size	1 minute	1 minute	1 minute	1 minute	5 minutes	10 minutes

Since the optimal aggregation interval size for daytime hours was different from the current 5 minutes, I checked the distribution of arrival times of individual vehicles from 15:00 to 15:30 of July 24 as a sample, as shown in Fig. 9. The analysis result showed that the section of T030070 showed a repeated pattern in which the vehicle group with a high arrival travel time which waited at a red light at the downstream intersection arrives, gradually followed by the vehicle group with a low arrival travel time, and then a vehicle does not arrive for about 100 seconds due to a stop signal.

Thus, the error from the population mean will be minimized when the data of the individual vehicles which waited for signal for the longest time and that of the individual vehicles which were least affected by a stop signal are not collected at the same time. Therefore, it is analyzed that it is reasonable to determine 1 minute, which is the aggregation interval size less than 100 seconds, during which a vehicle does not arrive, as the optimal aggregation interval size.

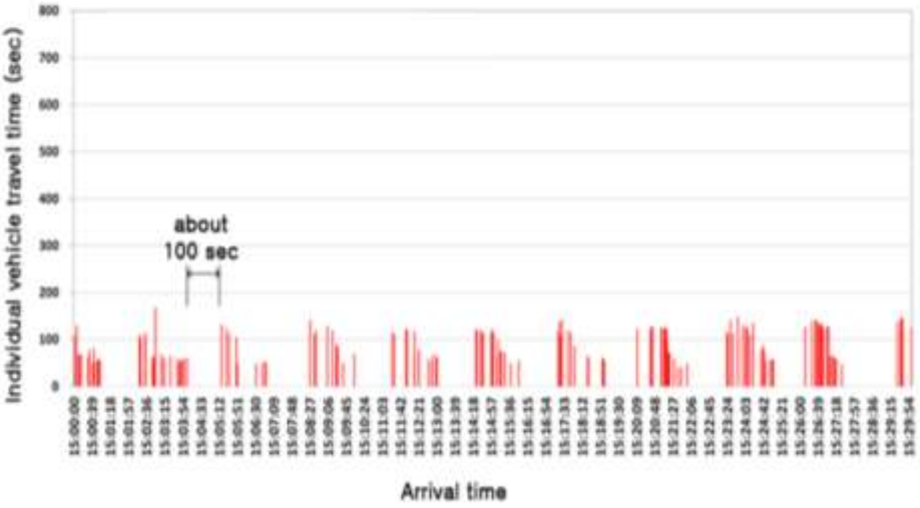


Fig. 9: Distribution chart of arrival times of individual vehicles during daytime hours (July 24, 15:00 to 15:30)

Also, it is analyzed that as shown in Fig. 10, at dawn hours, the optimal aggregation interval size becomes long due to a sudden drop in the amount of aggregated data because the arrival times of individual vehicles are collected without a repeated pattern of the change in the travel times.

It is analyzed that at transition hours, the optimal aggregation interval size becomes shorter than the previous time period because the arrival times of individual vehicles start to form a travel time pattern where the vehicles are classified by whether they waited for signal, due to the increase of incoming traffic volume, as shown in Fig. 11.

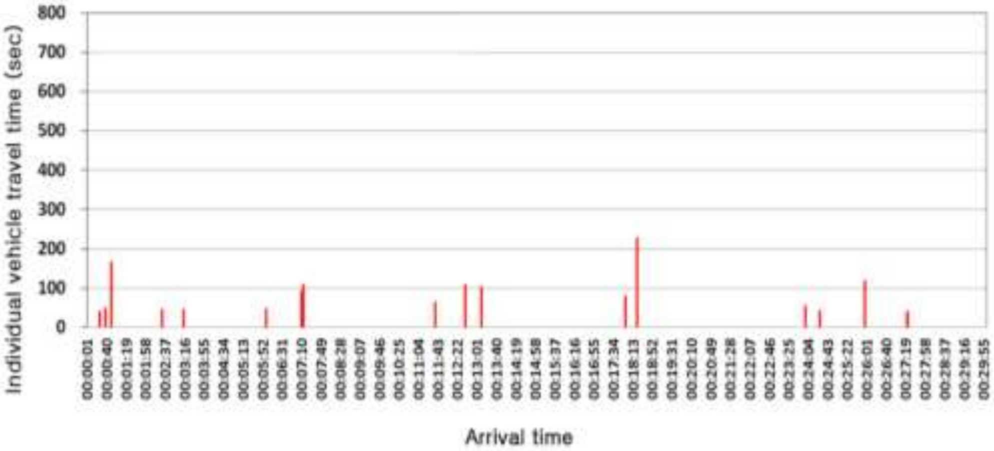


Fig. 10: Distribution Chart of Arrival Times of Individual Vehicles at Dawn Hours (July 24, 0:00～0:30)

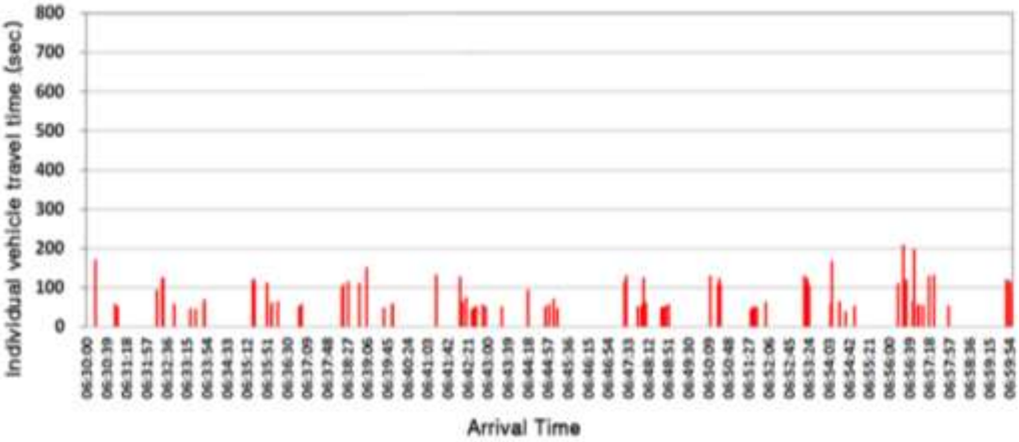


Fig. 11: Distribution Chart of Arrival Times of Individual Vehicles at Transition Hours (July 24, 6:30 to 7:00)

Further, it is analyzed that at the daytime zone of July 24, the aggregation interval size increased from 1 minute to 2 minutes because the interval of 100 seconds at which a vehicle does not arrive due to a stop signal disappeared with the rapid increase of travel time, as shown in Fig. 12, and because the travel time distribution changed from a bimodal asymmetric distribution to a unimodal distribution as shown in Fig. 13.

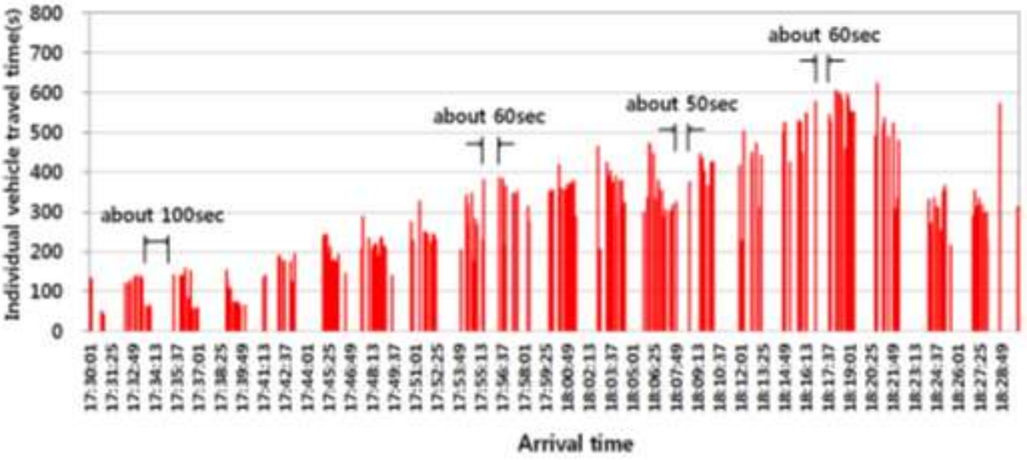
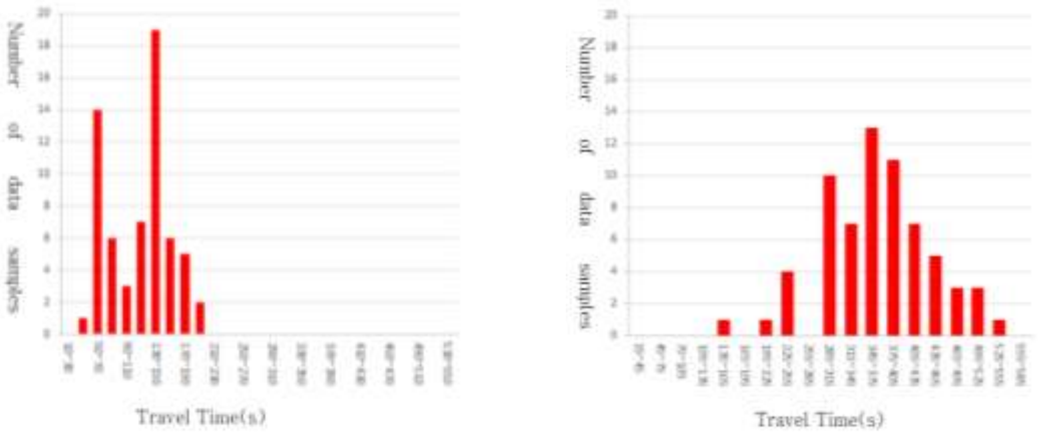


Fig. 12: Distribution Chart of Arrival Times of Individual Vehicles at Congestion Hours (July 24, 17:30 to 18:30)



(a) Travel Time Distribution (17:35-17:45)

(b) Travel Time Distribution (18:00-18:15)

Fig. 13: Change of Distribution of Travel Times of Individual Vehicles during Congestion Transition (July 24)

3) Determination of Optimal Aggregation Interval Size Considering Normal Data Loss

If the aggregation interval size is set to 1 minute, which is less than the time interval at which a vehicle does not arrive due to a stop signal, the representative value is not repeatedly generated even under normal operating conditions, so that an additional calibration is required and another error may be caused during the process. In this section, I will consider the aggregation interval size of data that avoids the repetitive representative value loss resulting from traffic conditions and short aggregation interval size. Given that the typical stop signal time (R) of national highway is less than 2 minutes, I will calculate the minimum MSE of the aggregation interval sizes of 3 minutes or more only, to determine the optimal aggregation interval size.

I calculated the MSE of the aggregation interval sizes of 3 minutes or more with the number of collections three or more for each time period. The result showed that the optimal aggregation interval size at which the MSE is minimized was 25 to 30 minutes at dawn hours of 0:00 to 6:00 and that the optimal aggregation interval size began to decrease to 20 minutes from 6:00, when the number of collections began to increase. The optimal aggregation interval size was 3 to 5 minutes from 7:00 to 22:00, and the aggregation interval size started to increase again to 5 to 10 minutes from 22:00, when the number of collections began to decrease again.

Table 5: Optimal Aggregation Interval Size of Data for Each Time Period
(3 minutes or more, the number of collections of three or more)

Time	00-01 o'clock	01-02 o'clock	02-03 o'clock	03-04 o'clock	04-05 o'clock	05-06 o'clock
Aggregation interval size	30 minutes	25 minutes	30 minutes	25 minutes	30 minutes	25 minutes
Time	06-07 o'clock	07-08 o'clock	08-09 o'clock	09-10 o'clock	10-11 o'clock	11-12 o'clock
Aggregation interval size	20 minutes	5 minutes	4 minutes	5 minutes	4 minutes	5 minutes
Time	12-13 o'clock	13-14 o'clock	14-15 o'clock	15-16 o'clock	16-17 o'clock	17-18 o'clock
Aggregation interval size	4 minutes	3 minutes	3 minutes	4 minutes	4 minutes	3 minutes
Time	18-19 o'clock	19-20 o'clock	20-21 o'clock	21-22 o'clock	22-23 o'clock	23-24 o'clock
Aggregation interval size	3 minutes	3 minutes	4 minutes	4 minutes	5 minutes	10 minutes

IV. CONCLUSION AND FUTURE STUDY DIRECTION

In the case of calculating the representative value of individual vehicle data collected by DSRC and providing it to traffic information users, the reliability of the representative value may vary depending on how the aggregation interval size of data is calculated. In this study, I determined the optimal aggregation interval size at which the MSE of the travel time collected with DRSC is minimized, for a representative interrupted traffic flow section which shows a bimodal asymmetric travel time distribution. In order to calculate the MSE for each aggregation interval size, I used the maximum estimation error equation of t-distribution, applicable to the section whose travel time shows a non-normal distribution, as the bias estimation equation, applied 68% confidence level and calculated the optimal aggregation interval size at which the MSE error is minimized.

The calculation result showed a repeated pattern in which the optimal aggregation interval size at which the MSE is minimized was 1 minute at morning and afternoon hours (7:00 to 22:00) and 5 to 10 minutes at transition hours (22:00 to 24:00), when the number of collections started to decrease, and the optimal aggregation interval size increased up to 25 to 30 minutes at dawn hours (0:00 to 6:00) and then decreased again at morning hours with the increase of the number of collections. However, if the aggregation interval size is set to be less than the time interval (1 to 2 minutes) during which a vehicle does not arrive due to a stop signal, the representative value will not be repeatedly generated even under normal operating conditions, which requires a separate calibration and thus may cause another error. Thus, I recalculated the optimal aggregation interval size using the aggregation interval sizes of 3 minutes or more. According to the result, the optimal aggregation interval size that minimizes the error of the calculation of the representative value was 3 to 5 minutes for morning and afternoon hours (7:00 to 22:00), 5 to 20 minutes for transition hours (6:00 to 7:00, 22:00 to 24:00), and 25 to 30 minutes for dawn hours (0:00 to 6:00). Therefore, it is analyzed that when the representative value is set at the fixed aggregation interval size of 5 minutes per day, the reliability of the representative value decreases when the number of collections decreases. Thus, it is analyzed that the aggregation interval size needs to be extended up to 25 to 30 minutes for dawn hours in order to maintain the reliability.

The expected effects of this study are as follows: Although the ITS centers of the regional construction and management offices have collected and provided traffic information of national highway, there was no sufficient basis for the setting of the aggregation interval size of data. However, this study provides the logical basis therefor. Also, if the ITS centers of the regional construction and management offices revise the aggregation interval size so that it changes according to the traffic conditions, based on the result of this study, it will contribute to the improvement of the reliability of the traffic information provided to the public. In this study, I found that, based on the MSE analysis of aggregation interval size for each time period, the optimal aggregation interval size for dawn hours needs to be set to be longer than that of daytime hours in order to increase the reliability of the representative travel time value. In future studies, it is necessary to develop an algorithm that minimizes the representative value estimation error by determining the optimal aggregation interval size in real time with the change of the number of collections, etc. as variables.

REFERENCES

[1]. Choi D. W. (2012), Determine Optimal aggregation interval size for Travel Time Estimation on Rural Interrupted Traffic Flow Considering Traffic Flow Characteristics, University of Science and Technology Master’s Thesis.

[2]. Gajewski B.J., Turner S.M., Eisele W.L., Spiegelman O.H. (2001), ITS Data Archving: Statistical Technique for Inductance Loop Detector Speed Data, Transportation Research Record 1719.

[3]. Kim H. J., Kim J. S. (2013), Lecture Statistics, Myung Jin.

[4]. Korea Institute of Civil Engineering and Building Technology (2009), Final Report of Design Service for the National Highway ITS of Seoul Regional Construction Management Office 2009(2nd), Seoul Regional Construction Management Office.

[5]. Korea Institute of Civil Engineering and Building Technology (2016), Final Report of ITS Operation Management of Seoul National Highway in 2016, Seoul Regional Construction Management Office.

[6]. Lim H. S. (2005), A Study on the Optimal aggregation interval size for Travel Time Estimation on the Rule Arterial Interrupted Traffic Flow, Ajou University Master’s Thesis.

[7]. Park D. J. (2000), Determining Optimal aggregation interval size for Travel Time Estimation and Forecasting With Statistical Models, J. Korean Soc. Transp., 18(3), Korean Society of Transportation, 55-76.

[8]. Park H. S., Kim Y. C. (2014), A Study on the Setting RSE Considering the Reliability of Traffic Information, Conference of Korean Society of ITS, Korean Society of ITS, 257-261.

[9]. Park H. S., Kim Y. C. (2017), Determination of the Optimal aggregation interval size of Individual Vehicle Travel Times Collected by DSRC in Interrupted Traffic Flow Section of National Highway, Journal of Korean Society of Transportation, Korean Society of Transportation, 63-78.

[10]. Yoo J. S., Oh C. S. (1999), Modern Statistics, Park Young Sa.

[11]. Yoo S. Y., Rho J. H., Park D. J. (2004), Investigating Optimal aggregation interval size of Loop Detector Data for Travel Time Estimation and Prediction, J. Korean Soc. Transp., 22(6), Korean Society of Transportation, 109-119.

*Hyun Suk Park. “Determination of The optimal Aggregation Interval Size for Each Time Period of Individual Vehicle Travel Time Collected Using DSRC in The Interrupted Traffic Flow Section.” International Journal Of Engineering Research And Development, vol. 13, no. 11, 2017, pp. 47–59.