

A BENCHMARK DATA SET FOR VALIDITY EVALUATION OF ROAD NETWORK SIMULATION MODELS

Ryota HORIGUCHI ¹⁾
Toshio YOSHII ²⁾
Hirokazu AKAHANE ³⁾
Masao KUWAHARA ⁴⁾
Masahiko KATAKURA ⁵⁾
Haruo OZAKI ⁶⁾
Takashi OGUCHI ⁷⁾

- 1) Kumagaigumi Co., Ltd., 2-1 Tsukudo-cho, Shinjuku-ku, Tokyo, Japan
Phone: +81-3-52615526, Fax.: +81-3-52619350, e-mail: rhoriguc@ku.kumagaigumi.co.jp
- 2) Institute of Industrial Science, Univ. of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo, Japan
Phone: +81-3-34026231, Fax: +81-3-34016286, e-mail: yoshii@nishi.iis.u-tokyo.ac.jp
- 3) Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba, Japan
Phone: +81-474-780444, Fax.: +81-474-780474, e-mail:
- 4) Institute of Industrial Science, Univ. of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo, Japan
Phone: +81-3-34026231, Fax: +81-3-34016286, e-mail: kuwahara@nishi.iis.u-tokyo.ac.jp
- 5) Tokyo Metropolitan Univ., 1-1 Minamiosawa, Hachioji, Tokyo, Japan
Phone: +81-426-772781, Fax.: +81-426-772772, e-mail: katakura-masahiko@c.metro-u.ac.jp
- 6) Toyo Univ., 2100 Kujirai, Kawagoe, Saitama, Japan
Phone: +81-492-391393, Fax.: +81-492-314482, e-mail: ozaki@krc.eng.toyo.ac.jp
- 7) Tokyo Metropolitan Univ., 1-1 Minamiosawa, Hachioji, Tokyo, Japan
Phone: +81-426-772781, Fax.: +81-426-772772, e-mail: oguchi-takashi@c.metro-u.ac.jp

ABSTRACT

We have been developing and validating a traffic simulator AVENUE which deals with urban road networks. Through the experience of the model validation, we have spent substantial effort to collect precise data set including time-varying O-D demand and individual vehicle trajectories. In this paper, the outline of the survey in Kichioji-Mitaka area is described at first. In order to collect precise data, the survey was made with recording number plates for 'every' vehicles traveling the area. Then, the extraction of individual vehicle trajectories from the survey data is briefly explained. The extraction process interpolates misreading data by observers with heuristics rules. The extracted vehicle trajectories are subsequently aggregated to time-varying O-D demand. Finally, the validation of AVENUE is presented as an example of the usage of this benchmark data set. The recommended model validation process and the future topics conclude this paper.

INTRODUCTION

This research aims to construct a precise data set for validity evaluations of road network simulation models. Needs to estimate the effects of planning and/or

improvements on traffic management has encouraged activities of developing traffic simulation models. Especially, dozens of simulation models to deal with the traffic on road networks have been developed during the last two decade.

However, they often have a problem in the model validation stage using real world data to evaluate their ability to reproduce traffic conditions, nevertheless the most of developers and users consider that the validity check is important. The major difficulty to prevent developers from model validations is that such simulation models generally require the data which is hard to directly observe, such as time-varying origin-to-destination (O-D) demand and individual vehicle trajectories.

For the last five years, we have been developing and validating a traffic simulator 'AVENUE' which deals with urban road networks^{[1][2][3]}. Through the experience of the model validation, we have spent substantial effort to collect precise data set including time-varying O-D demand and individual vehicle trajectories.

In this paper, the outline of the survey in Kichijoji-Mitaka area is described at first. In order to collect precise data, the survey was made with recording number plates for 'every' vehicles traveling the area. Then, the extraction of individual vehicle trajectories from the survey data is briefly explained. The extraction process interpolates misreading data by observers with heuristics rules. The extracted vehicle trajectories are subsequently aggregated to time-varying O-D demand. Finally, the validation of AVENUE is presented as an example of the usage of this benchmark data set. The recommended model validation process and the future topics conclude this paper.

NUMBER PLATE SURVEY IN KICHIOJOJI-MITAKA AREA

Figure 1 indicates Kichijoji-Mitaka area, western part of central Tokyo, where the benchmark data set was collected. The area spreads about 2 km from east to west and 1 km from north to south. The road network in the area consists of four major north-south streets and two major east-west streets. We selected this area by two reasons; there were plausible alternative routes for several O-D pairs and traffic condition changed during morning peak period.

The survey was carried out on 30 Oct. 1996, during morning peak period from 7:00 a.m. to 10:00 a.m. by more than 200 observers. There were 70 roadside points where the observers were located within the area. Every 'in' and 'out' approaches for the major intersections have a roadside point near the intersections.

Observers recorded vehicle types and the four large digits in the number plate of all passing vehicles. The passing time were simultaneously stamped in minutes. In order to increase the reliability of the data, two observers for one lane were assigned to each roadside point. One of the observers recorded number plates with an audio tape recorder and another one wrote down the numbers. Through this survey, totally 71,837 data elements were collected during valid time period, 7:50 - 10:00.

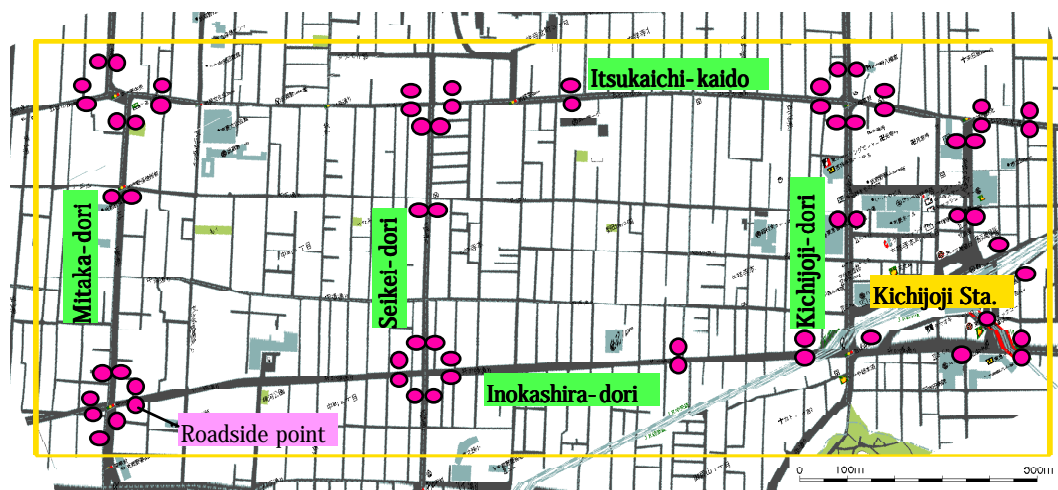


Figure 1: Kichijoji-Mitaka area and roadside points.

DATA PROCESSING WITH NUMBER PLATE MATCHING

Vehicle trajectories can be extracted from the result of the number plate survey by number plate matching between two adjacent roadside points and by chaining the matches. Travel times can be calculated by subtraction of the passing times. However, long trajectories are easily disconnected by misreadings.

We have proposed the interpolative method to extract vehicle trajectories considering observers' misreadings^[4]. The extraction process has six steps; 1) resolving inconsistencies in the data, 2) number plate matching between two roadside points, 3) pruning irregular match, 4) disentangling twisted vehicle trajectories, 5) retouching irregular trajectories, and 6) reuse of unused data elements.

Resolving Inconsistencies

The major data inconsistency is double counts of number plates at the roadside points beside two-lanes links. As the observers are assigned to each lane, they sometimes read same vehicle at the same time when the vehicle shows untidy movement over two lanes. Thus, one of the data elements which have the same number plate, the same vehicle type and the same time stamp is thought to be eliminated from the data of a roadside point. In this step, 838 double counts were detected and the total number of the data elements became 70,999.

Number Plate Matching

Between the data of two adjacent roadside points, two data elements of which the number plates and the vehicle types are same are 'matched' and may be a part of a vehicle trajectory. If the time difference between the matched data elements are too longer than the ordinary link travel time, the match should be out of consideration. In this case, the link travel time is limited within five minutes.

When the number plate of a vehicle was misread at one roadside point, a vehicle trajectory would be extracted as 'broken' two trajectories. In the preliminary analysis, since the misreading occurrence was estimated as much as 3.2 percent for normal observers, it is appropriate to assume that the number plate of a vehicle is not misread at two close roadside points. Therefore, the broken trajectories can be interpolated by skip matching, that is, a data element which have no match in the data of adjacent roadside point may match with the data element of same number plate at farther adjacent roadside point.

Pruning Irregular Match

A chain of matches forms the trajectory of a vehicle. However, the connectivity of roadside points sometimes affect to form irregular vehicle trajectories such as 'loops' and 'shortcuts'. These kinds of irregular matches easily happen when a vehicle travels within short time through the area where the roadside points sparsely located, but a loop can be detected by depth-first-search from the start point of the trajectory, and a shortcut can be done by width-first-search.

Disentangling Twisted Vehicle Trajectories

When two vehicles of which the number plate and the vehicle type are same as each other closely travel in the area, their trajectories will be entangled with unnecessary matches. In order to disentangle those trajectories, we give preferences to ambiguous matches using following heuristic rules and take higher ones; i.e.

- Rule-1:* To give higher preference to the matches that the disentangled trajectories are as possible as long when those matches are canceled.
- Rule-2:* To give higher preference to the matches that the disentangled trajectories are as possible as straight when those matches are canceled.
- Rule-3:* To give higher preference to the matches that the disentangled trajectories keep first-in-first-out order as much as possible when those matches are canceled.

Figure 2 illustrates example cases that these rules are to be applied. The thick arrows in the figure mean the matches with higher preferences.

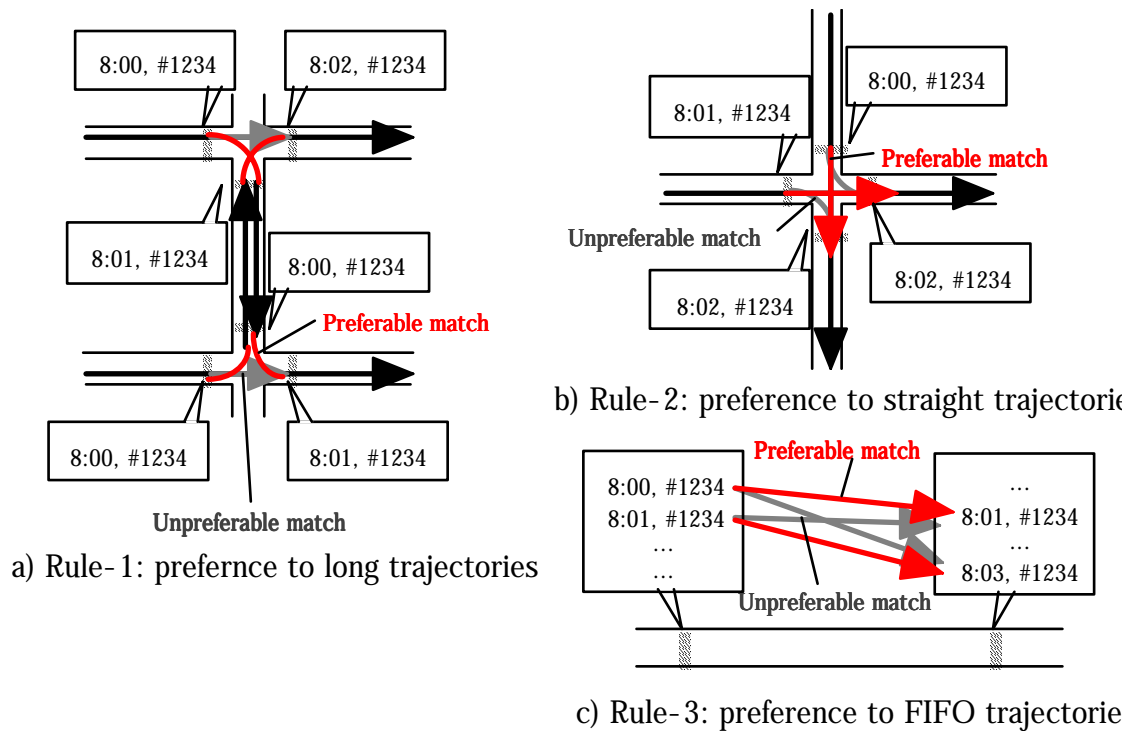


Figure 2: Heuristic rules to disentangle trajectories.

Retouching Irregular Trajectories

There are irregular trajectories that have either end inside an intersection. Such trajectories should be retouched as to move their irregular ends to the outside of the intersections. For that purpose, those irregular ends are connected to 'similar' data elements at neighbor roadside points with the following rules.

Rule-4: If four digits are same but in different order, those two data elements are similar.

Rule-5: If three digits out of four are same, those two data elements are similar.

Rule-6: If one of two data elements has no information for number plate, those two are similar.

If there are no similar data elements, new data elements will be imposed within the data of the nearest roadside point.

Reuse of Unused Data Elements

There are still remaining data elements which are 'unused' for extracted trajectories. According to the previous rules, detected similar pair of unused elements between two adjacent roadside points across an intersection are connected to be a trajectory. If there are no similar data elements to be connected, new data elements will be imposed within the data of the nearest roadside point and create a trajectory with them.

Result of Extraction

Following list shows the number of trajectories and the number of used data elements (percentage out of 70,999) at each step of the extraction process. As the step proceeds, the number of trajectories and the number of used data increases. The total number of the data elements increases up to 71,265 at the last step because of newly imposed data. Finally, the number of O-D pairs become 747 and the number of routes used by vehicles are 930.

Step	# of traj.	# of used	(%)
1. Resolving Inconsistencies	-	-	
2. Number Plate Matching	13,936	62,713	(88.3)
3. Pruning Irregular Matches	14,208	64,038	(90.2)
4. Disentangling Twisted Trajectories	14,953	68,334	(96.2)
5. Retouching Irregular Trajectories	14,953	69,123	(97.4)
6. Reuse of Unused Data Elements	16,043	71,265	(100.4)

MODEL VALIDATION OF AVENUE USING BENCHMARK DATA SET

AVENUE was, so far, validated with the real world data observed in Kinshicho area of Tokyo^[1] and Kanazawa-city^[2], where the both network has no alternative route for each O-D pair. In that case, AVENUE reproduced traffic congestion with satisfactory accuracy as well as non-congested traffic conditions. Further validation using the network of Toyoda-city^[3] that has room for drivers' route choice faced difficulty to obtain accurate O-D demand and traffic conditions to be compared with simulation results.

In this area, the drivers' route choice behavior has not be analysed yet, therefore we have applied AVENUE to Kichijoji-Mitaka area as an example to show the usage of the benchmark data set. Since AVENUE can assign each vehicle to a specific route, this is equivalent to incorporate the route choice behaviors perfectly into the simulation model. Figure 3 illustrates the road network used for the simulation. As there exists traffic volume that appears and/or disappears within links, O-D demand is assigned not only network-end nodes but also zone centroids.

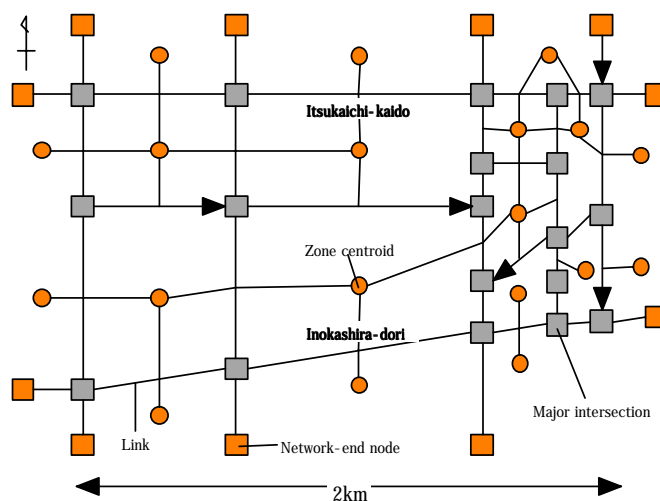


Figure 3: Road network used for the validation of AVENUE

Figure 4 shows the comparison of the simulation result and the survey in terms of throughput traffic volume in 10 minutes at each roadside point. The correlation of them indicates so high value as to conclude that AVENUE well reproduced link flow. While Figure 5 is the comparison of the average travel speeds on links for every 10 minutes and the correlation is quite low. This might be meaning there is still room to pay effort for parameter fitting. However, several problem can be pointed out of the benchmark data set; i.e.

- Link travel times obtained from survey involve error from the discreteness of passing times, which are recorded in minutes.
- As the observers wrote down the time in every minutes when the headway of coming vehicles was enough long, passing times of some vehicles may be different in 1 minute.

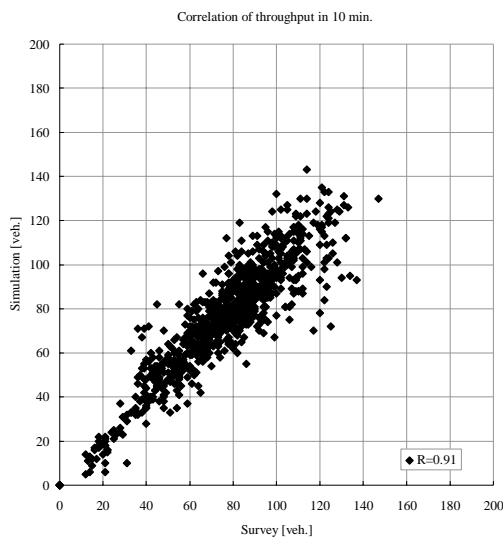


Figure 4: Correlation of throughput volumes.

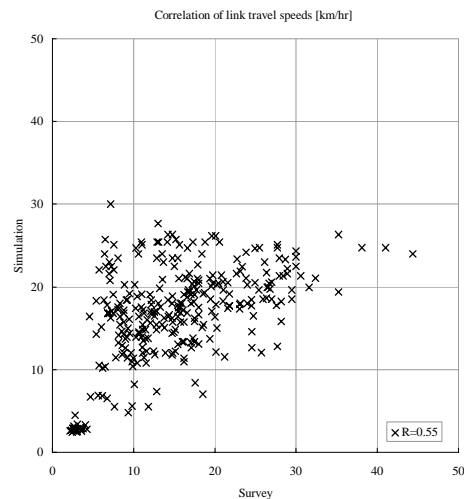


Figure 5: Correlation of link travel speed.

In order to cancel out such errors, it is recommended that travel times should be compared within long sections. Figure 6 shows the average travel speeds of vehicles traveling along Itsukaichi-kaido from east to west. The simulation result well follows the survey values.

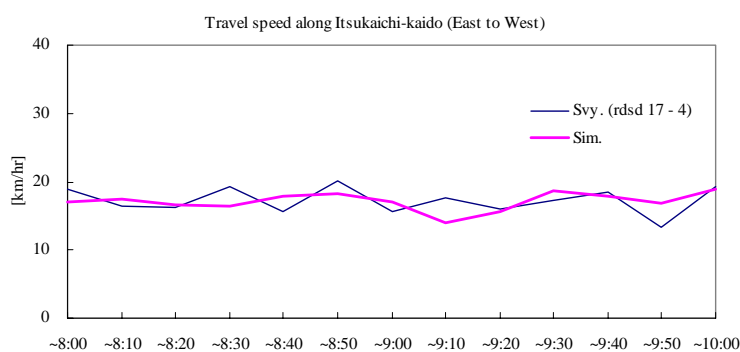


Figure 6: Comparison of average travel speed along Itsukaichi-kaido (East to West)

CONCLUSION

In this paper, the benchmark data set for the validation of road network simulation models was presented. To accurately estimate time-varying O-D traffic demand, the trajectories of almost all vehicles were extracted. As an example of the model validation using this benchmark data set, we applied AVENUE to the data set. The evaluation of the simulation result was made with the comparisons of throughput volumes and travel speeds. Especially travel speeds should be compared within the sections that have long distance in order to cancel out the errors in the survey values.

This benchmark data set will be distributed through the internet. We are now constructing WWW site providing the contents of this data set. Further analysis for drivers' route choice behavior and validations of various simulation models will be included in the WWW site as well.

REFERENCES

- [1] R. Horiguchi, M. Katakura, H. Akahane and M. Kuwahara : A development of a traffic simulator for urban road networks: AVENUE, Proceedings of Vehicle Navigation & Information Systems, Yokohama, pp.245-250, 1994
- [2] R. Horiguchi, M. Kuwahara and I. Nishikawa : The model validation of traffic simulation system for urban road networks: 'AVENUE', Proceedings of the Second World Congress on Intelligent Transport Systems '95, Yokohama, Vol. IV, pp.1977-1982, 1995
- [3] R. Horiguchi, M. Kuwahara, M. Katakura, H. Akahane and H. Ozaki : A network simulation model for impact studies of traffic management 'AVENUE Ver. 2', Proceedings of the Third Annual World Congress on Intelligent Transport Systems, Orlando, CD-ROM, 1996
- [4] R. Horiguchi, T. Yoshii, H. Akahane, M. Kuwahara, M. Katakura, H. Ozaki and T. Oguchi : Constructing a benchmark data set for validity evaluation of road network simulation models, (will be published in) Proceedings of the 8th World Conference on Transport Research, Antwerpen, 1998