FEASIBLE USE OF CROWDSOURCED TRAFFIC DATA IN TRANSPORTATION ENGINEERING APPLICATIONS

Sajith UDAYANGA[1], Teppei KATO[2], Aya KOJIMA[3] and Hisashi KUBOTA[4]

[1]Associate Member of JSCE, Student, Dept. of Civil & Env. Eng., Saitama University
(255 Shimo-okubo, Sakura-ku, Saitama 338-8570, Japan)
E-mail: pathiraja.a.u.s.760@ms.saitama-u.ac.jp
[2]Member of JSCE, Dept. of Civil & Env. Eng., Saitama University
E-mail: tkato@mail.saitama-u.ac.jp
[3]Member of JSCE, Dept. of Civil & Env. Eng., Saitama University
E-mail: akojima@mail.saitama-u.ac.jp
[4]Member of JSCE, Dept. of Civil & Env. Eng., Saitama University
E-mail: hisashi@mail.saitama-u.ac.jp

Insight into traffic engineering data sources has appeared in raising new challenges and opportunities when adopting cloud-based novel data mining techniques. This paper proposes the use of crowdsourced travel time data as possible mechanism to incorporate with traffic engineering. Travel time data obtained from the Google Distance Matrix API and review its usefulness in transport planning and evaluation. Applications have been evaluated while assessing the limitations such as network-level analysis for congestion zone identification, performance evaluation of the traffic calming measures and average queue length estimations at the controlled intersections in the Japanese context. Limitations were identified about the data characteristics. Successful validation of methodology creates a tool to improve traffic management and transport decision-making in regional transport planning.

Key Words : *crowdsourced data, google api, spatio-temporal graphs, cell transmission model*

1. INTRODUCTION

Accessibility and convenience of human movements using road transportation became the most prominent against the other transportation modes. There are many problems and issues arising with the development of transportation infrastructure and increasing demand for transportation. Cities, towns and major urban centers has been tackled with improving many problems of mobility and the safety. The significant influence of road transportation in the economy and personal life illustrate the broad impact of research in this area.

Transport planning and management is a critical contributor which can provide solutions for major problems. The transport planning ranges to a highly diversified area in which many types of research were conducted. It required to provide quality and reliable transport service. Congestion is a major problem faced by many transport systems which need to be managed and maintained at optimum levels to ensure an excellent service for users. Therefore, the traffic congestion management, crowd handling and implementing new policies are frequent tasks involved by the transportation planner and managers. But the availability of data for transport planning was not enough. Therefore, many developments were initiated over the past several decades to obtain data required for transport planning. There are many parameters to identify in controlling traffic. Traffic flow, speed, capacity, safety, compatibility, accessibility are some parameters required in transport planning to ensure efficient planning and design[1)].

(1) Crowdsourcing

The participation of a large and varied group of people in the planning process has long been encouraged to increase the effectiveness and acceptability of plans. It had been alternately defined as the outsourcing of a job to a large undefined group[2)].

Crowdsourcing is a distributed problem-solving model in which undefined size of people engaged in achieving a common goal. When looked at the generation of crowdsourced data, it could be categorized as an active crowdsourcing and passive crowdsourcing. In active crowdsourcing, people who get involved in crowdsourcing are aware of the objectives, and they work towards the project and involve actively as agents. In situations such as surveys, polls, contests, data-entry and protests people involve directly and actively towards achieving the objective of the project. In passive crowdsourcing, the activity of people is being traced as data and people are not actively taking part in the project.

Crowdsourced data analysis could be an optimum approach in such a complex situation to gather information from the undefined amount of people who are actively using the traffic network in real-time. With crowdsourced data gathering, it enables to gather information instantaneously on both temporal and spatial[3)]. Cloud-based big data approach which has the potential of collecting and procession large scale traffic network data has been reviewed in the given paper while proposing detailed data validation method.

(2) Google distance matrix API

Google distance matrix application programming interface (API) the service provided by Google Inc under Google Maps platform. By accessing this API, it is possible to collect trip information such as distance and travel time about a route between given origin and given destination. By using Google Maps, a user can find the distance and estimated time to arrival (ETA). The user must give the origin and destination to Google map if he/she wants to find travel information on traveling between two locations. Then Google provides information such as distance, directions, travel time-based on different modes of traveling.

(3) Travel time data

Travel time is the most understandable parameter when compared to other parameters used to evaluate traffic on roads such as traffic flow speed etc. Being the most fundamental source of information in planning and designing, it's enabled researchers to form solid research base. For an instance, a classic application of travel time-based policy is developed by the Boston Region Metropolitan Planning Organization. In the identification of performance need studies and assessments, travel time is widely used. The travel time index is being used on several occasions to rank and prioritize transportation improvement projects for funding.

(4) Congestion monitoring

Traffic congestion can create social and economic dilemma which can affect the harmony of the people in roads. On the other hand, it might be a cause for environmental problems in terms of emissions. Traffic jams become infestation of modern life specially in a megacity. It is important to keep alert for sub-urban areas neighboring big cities which can be subjected to urban sprawl.

Saitama prefecture is the forth (4th) lower congested travel speed record in Japan. Identification and monitoring of urban area formation and major traffic jams play important role in planning and design of traffic jam countermeasures. **Fig.1** illustrates the mostly jammed areas namely, Tokorozawa, Kawagoe, Saitama central and Saitama eastern area.

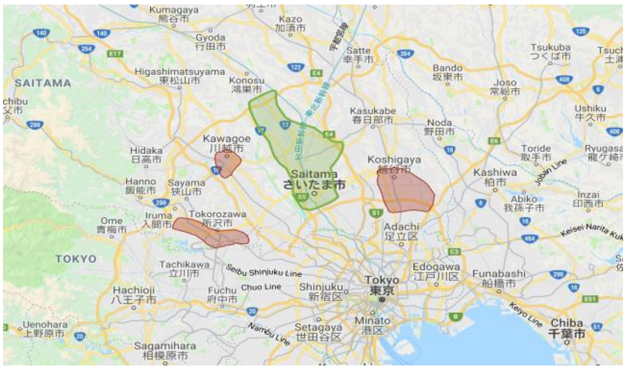

Fig.1 Congested areas in Saitama

(5) Preliminary Studies

Novelty of the proposed data collection method encouraged authors to conduct few pilot studies with the intention of identifying the limitations and the data characteristics in Japanese context. Investigation of crowdsourced approach to identify the speeding hots-spots of the low volume local roads which can be incorporated in decision making process of location traffic calming measures such as speed humps was conducted.

a) Congestion zone identification

Preliminary study was carried out to understand the applicability of the given method identifying jammed area of the network level. Link speed data of Kawagoe road network was obtained to identify the patterns and the data characteristics. **Fig.2** shows the speed variation of the road link in the central Kawagoe which explain the temporal variations. It was noted that, useful information can be gathered in given scenario.

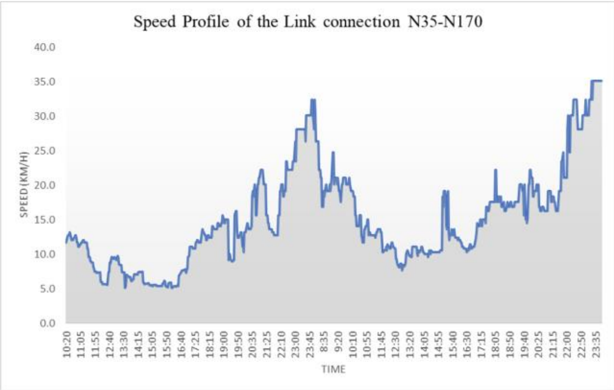

Fig.2 Speeding hot-spot identification

b) Performance evaluation of speed hump

Pilot study was conducted to understand the effectiveness of permanently installed speed hump located at Omiya-Sengencho. This is a low volume local road with large number of connecting roads also catering pedestrians and cyclists. Road stretch was divided into segments. Travel time data was collected using license plate recognition (LPR) system and proposed crowdsourced (API) approach along with the manual data entering. Due to several limitations identified, given technique might not be fit to the purpose. Facts will be elaborated in detail in the discussions. Further works to be carried out for more detailed justification of the accuracy level obtained for the given scenario.

Number of users take part in data feeding process of the crowdsourced database can be identified as a limitation to the data sensitivity. With low volume traffic with many pedestrian movements lead to generate more outliers to the database.

c) Speeding hot-spot identification

Preliminary study on identification of speeding hot-spots was carried out at Nakaya-machiku-Kyujitsukykanshinryojo, Yokohama with the special care was taken to define 19 road segments. Field reconnaissance survey was carried out prior to define the segments. **Fig.3** illustrates the information about the focused road stretch along with the output speed profile. Accuracy of the output may not be enough for such micro-level analysis, since lesser crowd is taking part in the data feeding to the data base. In this particular case, theoretical difference in space mean speed and time mean speed cases accuracy reduction in addition to the low user density. Compared speed profiles were obtained from the spot speeds recorded from the speed gun (Store-Car) and mean speeds from Google API which indicated in the bold line in the **Fig.3**. Although it is having lesser sensitivity, two peaks can be spotted corresponding to higher speed values.

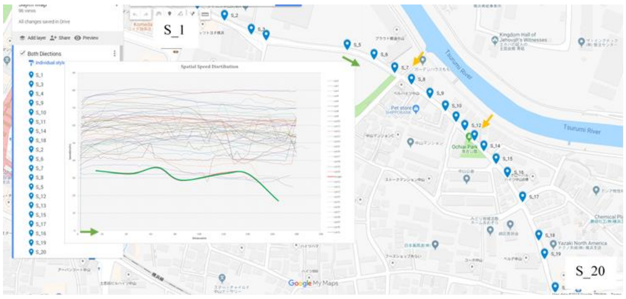

Fig.3 Speeding hot-spot identification

2. LITERATURE REVIEW

Development of information and communication technologies enabled researchers to enter an era of big data and many sources of novel data, including mobile phone data have been applied for transportation research. Big data is believed to have the potential to make our cities smarter by facilitating the discovery and explanation of urban development and its dynamics, which will enable city managers and citizens to make more informed decisions and to enjoy better city life[4]. Similarly, in the transportation field, big data has promoted intelligent transportation systems by providing a better understanding of where, when, and how people travel around the globe.

(1) Crowdsourced traffic data

As a concept, social computing focuses mainly on building a network of collaborators and facilitating online communication between groups. This has eventually given rise to open source platforms and forums where people with similar motivation and outlook can come together to solve issues and find answers to problems that affect their community.

Crowdsourcing is an example in which an organizer or an o organization can use the network of collaborators to solve a problem that would otherwise be cost- or labor-intensive, or in which within a defined organization the expertise is unavailable or insufficient. Most of the crowdsourcing systems use devices and technologies that are readily available and low cost; often crowdsourcing is based on devices that are owned by individuals (as in cycling data collection in CycleTracks and Cycle Atlanta), involving no major financial investment on the part of the system[2].

The widespread availability of internet and mobile devices has made crowdsourced reports a considerable source of information in many domains in the world. Traffic managers and planners, among others, have started using crowdsourced traffic incident reports (CSTIRs) to complement their existing sources of traffic monitoring. One of the prominent providers of CSTIRs is Waze. Quantitative analysis was conducted to evaluate Waze data in comparison to the existing sources of Iowa Department of Transportation. The potential added coverage that Waze can provide was also estimated[5)]. These findings were offered an open-source software package which can implement the clustering in near real-time.

Comprehensive review of travel behavior and mobility pattern studies were done by Mario et al. which used the mobile phone data. Many studies have explored the potential of passively collected data to supplement traditional surveys. Moving forward, it will be possible to tap into the full potential of these data sources and supplement them with minor surveys. This approach has the potential to decrease respondent burden and cost while improving data quality and prediction accuracy[6)].

a) Use of Google travel time data

O-D travel time matrix was estimated using Google Maps API by Wang at el. while developing desktop tool to perform their expected task. Multiple calls can be requested from the API with this tool. By doing so, authors would be able to tap into the dynamically updated transportation network data and the routing rules maintained by Google and obtain a reliable estimate of O–D travel time matrix. **Fig.4** illustrated below shows the estimated travel time from the city center by ArcGIS and Google API. This study uses a study area in East Baton Rouge Parish of Louisiana. The estimated travel time by either method correlates well with the (Euclidean) distance from the city center (with a R^2 = 0.91 for both methods).

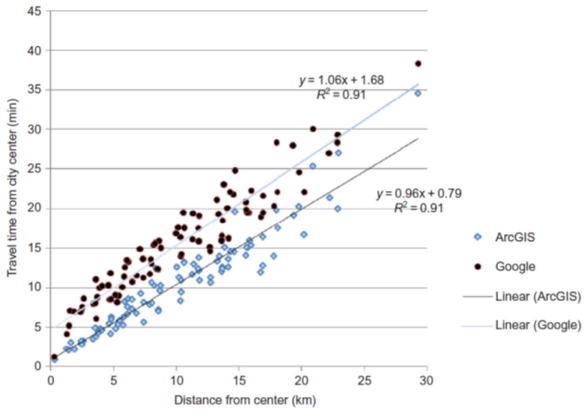

Fig.4 Speeding hot-spot identification[7)]

However, the regression model of travel time against corresponding distances by Google has a significant intercept of 4.68 minutes (a negligible 0.79-minute intercept in the model of travel time by ArcGIS). The 4.68-minute intercept by Google probably reflects the elements of starting and ending time on a trip (getting on and off a street on the road network). This is consistent with our daily travel experience and empirical data.

Case study was conducted by applying estimated travel time in assessment of hospital accessibility. In summary, Google API can generate the real-time travel time by accessing the most updated road network data in Google[7)].

Travel duration and distances to and from airports, under varying conditions/parameters, have been gathered using the Google Maps Distance Matrix API. airport access and egress times have been analyzed for 22 European airports against the background of the European Commission's Flightpath 2050 goal of enabling four hours door-to-door intra-European travel, with the intention to provide further insight for policymakers when deciding on the implementation of measures to reduce overall door-to-door travel time and enhance the seamlessness of the transport system. For this purpose, travel times and speeds have been gathered via Google Maps for airport access and egress via private and public transport. For the analyses, extensive use has been made of the origin, destination, transport mode, and departure time parameters. Ferries have been set to be avoided and the best guess traffic model has been chosen for all following Google Maps Distance Matrix API requests. The presented results originate from repeatedly sending requests to the Google Maps servers with alterations in the described parameters.

Both, the generation of requests as well as their querying, have been performed automatically by two self-developed Python scripts. Google Maps only returns valid travel options and, with that, valid travel durations and distances when the Google Maps services could identify, either, a drivable road connection (in the case of driving) or a suitable public transport service (in the case of transit). Obtained data was successfully used to analyze variations in airport and egress speeds between European airpors[8)].

Car travel time data collected from Google Distance matrix API while employing more than 1.6 million API requests. Analysis firstly resulted in the graph of relative travel times. Later, they investigated travel time variability, also produced the mean travel time matrix which might be used as a skim matrix to validate the macro model of Kaunas city. Finally, we concluded with the eval action of accessibility of all 206 centroids. e[9)].

b) Identification of road bottlenecks

Mitigation of the traffic congestion in urban expressways, much previous studies has focused on the analysis of basic traffic parameters, while, congestion is a random event and drivers pay more attentions to the probability of congestion occur and the time it will last.

A traffic bottleneck is a localized disruption of vehicular traffic on a road segment in which separates upstream queued traffic and free-flowing downstream traffic. When compared to a traffic jam, a bottleneck is a result of a specific physical condition, often caused by merging and diverging traffic, lane drops, grade changes or badly timed traffic lights, intersections. Identification of time and duration of congestion in bottlenecks can enable investigations and provide mitigatory actions[10].

Study on identifying the static road bottlenecks and dynamic road bottlenecks based on the dynamic traffic data gathered from Google distance matrix API was concluded for arterial road in Colombo metropolitan area in Sri Lanka. Authors validated the data using license plate matching survey on 3.06 km long speed bypass road in Colombo. **Fig.5** shows the obtained temporal travel time variations for different vehicle types.

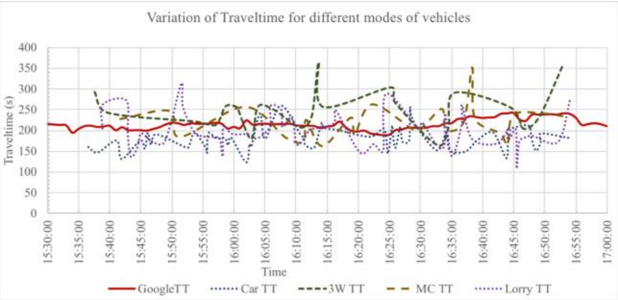

Fig.5 Temporal variation of travel time[11]

Identification of road bottlenecks with a spatiotemporal variation of speed data was graphically visualized o traffic state matrix and impact over time is evaluated. The reliability and significance of road bottlenecks were identified using three evaluator measures, travel time index, bottleneck influence factor and speed variation form overall network. Therefore, has concluded that the methodology is very advantageous in planning and forecasting traffic management activates which require higher accuracy and low cost of implimentation[11].

Hence this is a promising and economical method to identify congestion patterns in large urban networks in developing countries. Also, for academics in developed countries where travel time data is not accessible.

c) Queue length estimation

Quantifying congestion was charged with developing ways to reassess and enhance congestion measurements. Growth of traffic congestion is a major concern to community at large[12].

Average maximum queue length was estimated by using probe vehicle GPS data with long reporting intervals by Xin et al. while summarizing the queue length estimation studies done by other researchers. Most of the authors used conventional techniques rather crowdsourced approach[13].

The analysis carried out to develop a cell transmission model to identify average space mean speed contour which is responsible for the location of the traffic queue at a given time for controlled and uncontrolled intersections which face heterogeneous traffic in Sri Lanka[14]. Minimum queue length and maximum queue length of a controlled intersection were estimated using the cumulative travel time graphs. This methodology has been verified using actual queue length data collected from CCTV video footage and manual transcription. With the comprehensive validation of this technique can replace the data obtaining from traffic monitoring sensor networks[15].

d) Space mean speed and time mean speed

The mean of a certain mix of vehicles with different speeds in a given length of road is the space mean speed. Consider a cohort of vehicles classified by speed with values V1, V2...Vn. The times which they take to cover a certain length L, are T1, T2…Tn. Space mean speed is derived from the average of the times taken by the vehicles as shown in the equation (1a). Where, V_s is space mean speed, L is distance traveled, n is number of vehicles, T_{i} is travel time.

$$V_s = L/((\sum_{i=1}^{n} T_i)/n) \tag{1a}$$

Time mean speed: the same distribution of vehicles will produce a different mix while passing a fixed point over a certain period. The average of these speeds is the time-mean speed as shown in the equation (1b). Where, V_t is time mean speed, n is number of vehicles. V_i is given vehicle speed.

$$V_t = (\sum_{i=1}^{n} V_i)/n \tag{1b}$$

There will be a greater representation of faster vehicles. Radar (Store Car), speed camera & loop detectors can be used to determine time mean speed[16].

3. PROPOSED METHODOLOGY

Proposed methodology was developed to evaluate traffic congestion point and jammed road segment identification framework in Japanese context. **Fig.6** shows the overall methodology flow chart to be employed in the study.

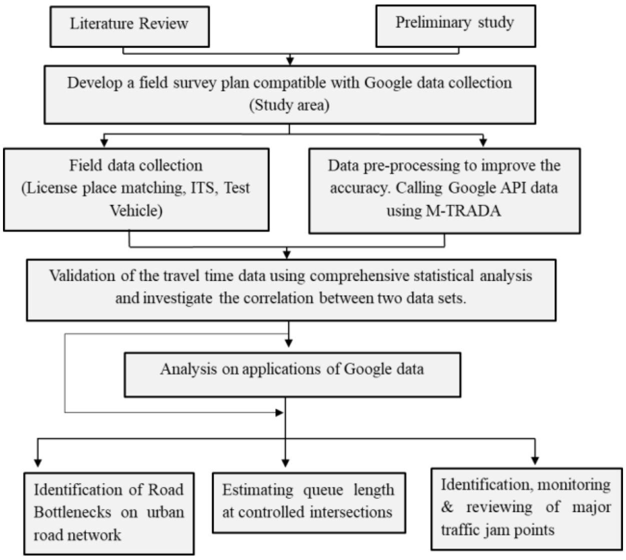

Fig.6 Methodology flow chart

(1) Data collection

This research uses a desktop tool (M-TRDA) for implementing the task by calling the Google Maps Application Programming Interface (API) as shown in the **Fig.7** below. Detailed information about the data mining will be presented at the oral presentations.

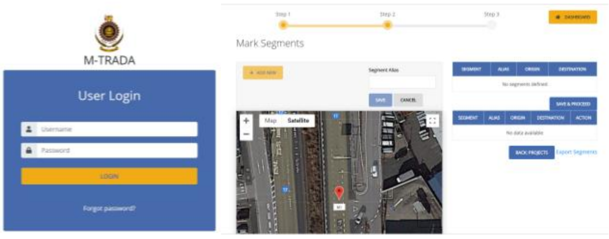

Fig.7 M-TRDA data collection tool

License plate recognition system to be employed for travel time data collection for validation. **Fig.8** shows the detection capacity of the software "movie-N~ CLP-MV01S. Ver. 2.10" with common license plate types in Japan.

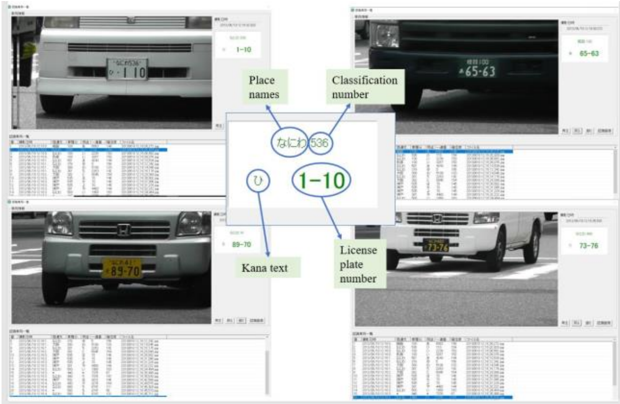

Fig.8 License plate recognition system

(2) Study area

Road links in the Central Saitama area was selected for data study including Shin-Omiya bypass road. Estimations of road bottlenecks in southbound Shin-Omiya bypass road has been evaluated in this paper which subjected to expand at the time of oral presentation with the data validation. Study area has been illustrated along with sample data collected from Google API in the **Fig.9** bellow.

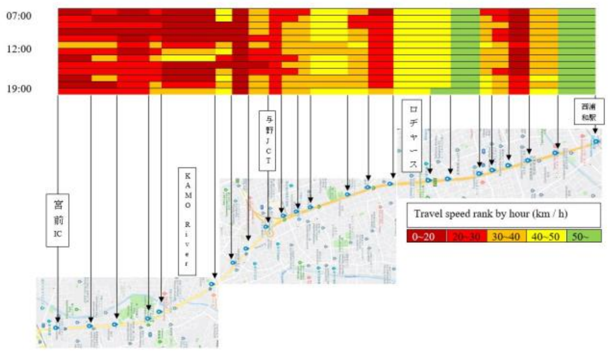

Fig.9 Study area of Shin-Omiya bypass road

Warm colors were used to indicate average segment space mean speed lesser than 21 km/h[17]. In terms of utilizing obtained data

Travel Time Index (*TTI*) and percentage effect must be calculate using equation (2a) and (2b) respectively as show below. Where, $TT_{bottleneck}$ is travel time at the bottleneck and $TT_{freeflow}$ is the travel time under free flow condition. Travel time at the bottleneck will be obtained after sorting the low average space mean speed values based on the 21 km/h. Travel time at the free flow condition will be obtained as per the highest space mean speed recorded during the off-peak time. Similarly, speed values will be obtained from the given scenario.

$$TTI = (TT_{bottleneck})/(TT_{freeflow}) \tag{2a}$$

$$\%Effect = (V_{Net} - V_{Bot})/(V_{Net}) \tag{2b}$$

Evaluated congested locations will be subjected to review under process illustrated in the **Fig.10** which might be used as a framework for monitoring, reviewing and traffic jam countermeasure practicing by transport planners in Japan.

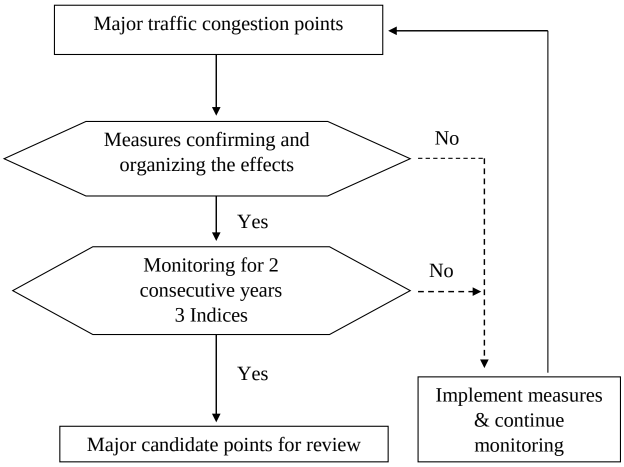

Fig.10 Traffic jam monitoring framework

3 indices were defined as **Table 1** below for evaluation of jam point as indicated in the **Fig.10**.

Table 1 Index definition for monitoring framework.

	Index Definition
1	12 h weekday Avg. speed < 20 km/h
2	Weekday peak time Avg. speed < 10km/h
3	Holiday 12 h dynamic peak < 10 km/h

(3) Queue length estimation

Cell transmission model will be derived from that concept the space mean speed calculated from travel time data obtained from Google Distance Matrix API for the intersections as shown in the **Fig.11** below.

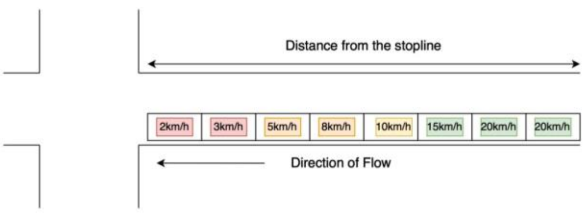

Fig.11 Cell transmission model

When the cell transmission model is developed, threshold speed values will be identified using video footage and manual data collection. Objective function of queue length estimation by cell transmission model will be illustrated in equation (3a), (3b), (3c) as shown below, where V_s is the limiting threshold speed value, *a* is the lower bound of the speed, *b* is the upper bound of the speed. Vehicles having speed a km/h will be classified as a queue and speed more than *b km/h* will be classified as a free moving vehicle without any interruptions.

$$V_s < a(km/h) – Queue \tag{3a}$$

$$a < V_s < b – Transition \tag{3b}$$

$$V_s > b(km/h) – Free\_Flow \tag{3c}$$

Speed contour graphs will be developed with observed queue length variations for lower bound and upper identification as shown in the **Fig.12** below.

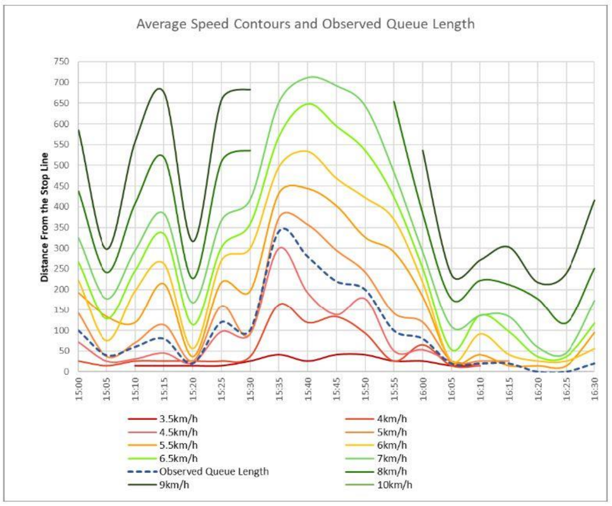

Fig.12 Speed contour graph

Considering the mean standard error, contour line will be selected which gives a better estimate. Cumulative travel time graphs will be plotted to identify the maximum and minimum queue lengths as shown in the **Fig.13** below.

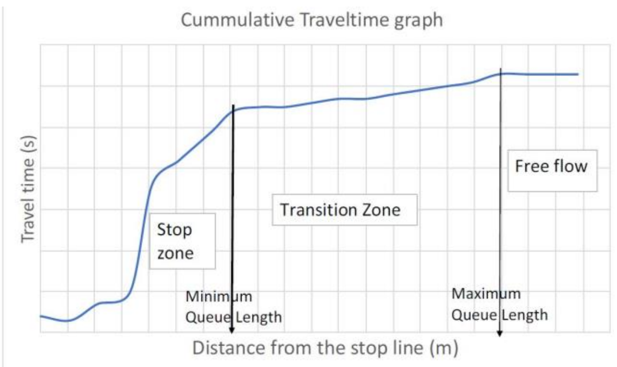

Fig.13 Cumulative travel time graph

Observed queue lengths and estimated queue lengths are tested with paired sample T test with the hypothesis test. Estimated values will be verified with statistical analysis aforementioned

Finally, identified bottleneck points will be classified under 3 indices given in the **Table 1**. Estimated queue lengths will be employed to demarcate the congested road sections.

4. ANALYSIS AND RESULTS

Given road segment was analyzed using the data obtained from the Google API with 5-minute interval for 12 hours (07:00-17:00) for road segment specified in the **Fig.10** above. Hourly averaged space mean speeds were illustrated in the **Fig.12** for visualization.

Time	1--2	2--3	3--4	4--5	5--6	6--7	7--8	8--9	9--10	10--11	11--12
7:00	13.1	30.5	16.9	23.1	9.8	40.2	17.8	31.9	27.3	36.9	28.3
8:00	22.3	26.8	13.5	18.1	9.4	40.9	17.3	30.8	26.8	35.1	29.2
9:00	29.1	33.8	20.4	26.1	12.5	15.8	11.8	31.8	27.7	37.3	32.2
10:00	19.6	17.1	11.7	18.5	9.6	12.3	11.6	30.3	28.4	35.6	32.5
11:00	25.1	27.6	17.4	22.3	11.9	15.1	11.2	28.2	26.7	34.8	31.8
12:00	31.5	37.1	24.4	31.9	28.4	38.3	16.9	30.5	27.6	36.9	32.1
13:00	22.0	24.2	14.3	23.6	33.6	42.3	17.8	29.5	29.9	35.2	31.2
14:00	18.5	30.4	20.7	30.9	17.3	27.9	13.9	20.6	27.1	35.8	29.3
15:00	25.8	33.3	20.2	25.6	11.7	12.8	10.5	15.8	26.7	33.8	31.3
16:00	27.2	35.7	21.6	25.9	10.9	30.2	15.2	27.6	27.6	34.3	29.9
17:00	19.4	34.3	22.4	29.4	15.0	40.7	17.6	30.9	26.2	35.6	31.0
18:00	10.6	36.5	24.0	32.7	21.4	36.8	16.1	31.2	26.6	34.8	31.8
19:00	29.6	39.7	25.7	34.3	35.2	45.5	19.0	32.5	28.6	36.8	31.5
12h avg	22.6	31.3	19.5	26.3	17.4	30.7	15.1	28.6	27.5	35.6	30.9
peak	15.5	28.6	18.3	26.3	14.5	29.3	14.9	28.5	27.3	35.8	30.5

Time	12--13	13--14	14--15	15--16	16--17	17--18	18--19	19--20	20--21	21--22	22--23
7:00	28.4	18.2	44.6	47.8	54.8	27.7	22.9	14.8	38.1	53.2	43.7
8:00	36.8	22.9	45.1	47.7	55.1	31.8	24.7	14.7	37.7	52.2	52.0
9:00	39.0	26.3	46.6	48.8	55.8	38.5	30.3	17.6	38.7	54.7	54.9
10:00	37.6	24.7	45.4	48.2	55.8	37.6	26.3	14.8	37.1	52.3	52.3
11:00	37.5	22.2	44.5	47.5	55.3	38.4	27.7	16.4	38.5	52.2	52.0
12:00	40.2	27.2	47.2	49.7	57.8	40.6	34.7	20.3	38.6	55.4	55.0
13:00	37.0	22.0	44.9	48.2	55.8	38.2	30.0	18.2	37.7	55.9	55.1
14:00	35.7	24.0	44.7	46.8	54.7	39.9	31.7	19.1	38.2	54.1	55.2
15:00	37.3	25.8	46.2	48.0	55.4	38.1	30.4	18.5	38.2	54.4	54.4
16:00	36.2	24.2	45.0	48.8	56.5	38.0	28.8	18.4	39.4	54.6	53.4
17:00	37.6	23.1	44.7	48.1	56.5	41.8	36.4	21.2	40.2	57.8	57.7
18:00	37.7	23.9	46.6	49.5	58.3	41.9	34.5	20.6	39.5	55.6	56.2
19:00	41.2	31.8	48.1	50.9	59.1	45.2	38.8	24.5	40.1	55.5	55.7
12h avg	37.1	24.3	45.7	48.5	56.2	38.3	30.5	18.4	38.6	54.4	53.7
peak	34.8	22.7	45.3	48.1	55.9	36.8	28.8	17.3	38.2	53.8	51.9

Fig.12 Spatio-temporal graph of bottleneck formation

3 major bottlenecks can be identified namely segment 3-4, 5-6, 7-8 and19-20. Segment identity as per the **Fig.9** left most point is segment 1. Segment number is increasing towards south bound direction of the Shin-Omiya bypass road.

Fig.13 shows the spatial variation of the travel time index along the road segments. Percentage effect also indicated higher values on segments. This decomposed an idea on how worse the performance of the bottleneck to the overall road traffic. Similarly, indices 1 and 2 defined in the **Table 1** also indicated lowest values for corresponding segments as identified as bottlenecks.

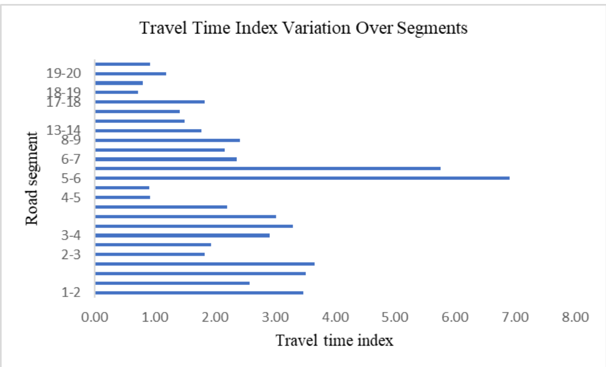

Fig.13 Travel time index distribution

5. DISCUSSION AND FUTURE WORK

This paper reviewed the feasible use of Google distance matrix API travel time data for transport engineering application in general, while presenting results of the preliminary works done. Also captured the data characteristic limitations of the proposed method which have been identified through the background studies. Network-level full analysis results will be presented at the oral presentation sessions to be held.

Importantly, data characteristics were identified with their limitations. Google Maps API approach is not free of concerns. When it applied for low volume local roads, data sensitiveness and representativeness is low. Since we are dealing here with big data, user density which the system gathering raw data required to be reasonably large. But it has a proven track record with published literature of meaningful use in macro-level analysis with acceptable accuracy and sensitivity.

ACKNOWLEDGMENT: The support by the Asian Development Bank Japanese Scholarship Program for graduate studies at International Graduate Program on Civil and Environmental Engineering, Saitama University is greatly acknowledged. Also, teachers of planning and design laboratory, Department of Civil and Environmental Engineering, Saitama University support extended for additional funding for software and tools is highly acknowledged.

REFERENCES

1) Samarajiva, R., S. Lokanathan, K. Madhawa, G. Kreindler and D. Maldeniya : Big data to improve urban planning, *Economic & Political Weekly* 50(22): 43,2015.
2) Misra, A., A. Gooze, K. Watkins, M. Asad and C. A. Le Dantec : Crowdsourcing and its application to transportation data collection and management, *Transportation Research Record* 2414(1): 1-8, 2014.
3) Chatzimilioudis, G., A. Konstantinidis, C. Laoudias and D. Zeinalipour-Yazti : Crowdsourcing with smartphones, *IEEE Internet Computing* 16(5): 36-44, 2012.
4) Steenbruggen, J., E. Tranos and P. Nijkamp : Data from mobile phone operators: A tool for smarter cities?, *Telecommunications Policy* 39(3-4): 335-346, 2015.
5) Amin-Naseri, M. : Adopting and incorporating crowdsourced traffic data in advanced transportation management systems, 2018.
6) Sun, Y. and M. Li : Investigation of travel and activity patterns using location-based social network data: A case study of active mobile social media users. *ISPRS International Journal of Geo-Information* 4(3): 1512-1529, 2015.
7) Wang, F. and Y. Xu : Estimating O–D travel time matrix by Google Maps API: implementation, advantages, and implications, *Annals of GIS* 17(4): 199-209, 2011.
8) Rothfeld, R., A. Straubinger, A. Paul and C. Antoniou : Analysis of European airports' access and egress travel

times using Google Maps, *Transport Policy*, 2019.
9) Dumbliauskas, V., V. Grigonis and A. Barauskas : Application of Google-based data for travel time analysis: Kaunas city case study, *Promet-Traffic&Transportation* 29(6): 613-621, 2017.
10) Zhang, L. and D. Levinson : Some properties of flows at freeway bottlenecks, *Transportation Research Record* 1883(1): 122-131,2004.
11) Kumarage, S. P. and J. Dimantha De Silva : Identification of Road Bottlenecks on Urban Road Networks Using Crowdsourced Traffic Data, 2018.
12) Lomax, T., S. Turner, G. Shunk, H. Levinson, R. Pratt, P. Bay and G. Douglas : NCHRP report 398: quantifying congestion, *TRB, National Research Council*, Washington, DC,1997.
13) Qi, X. and H. Jiang : Estimating average maximum queue length at an intersection using GPS data with long reporting intervals, 2017
14) Daganzo, C. F.: The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory, *Transportation Research Part B: Methodological* 28(4): 269-287, 1994
15) Kumarage, S. P., J. Dimantha De Silva and J.M.S.J Bandara, Average Queue Length Estimation Using Crowdsourced Travel Time Data Under Heterogeneous Traffic Conditions : *World Conference on Transport Research*, 2018.
16) Knoop, V., S. P. Hoogendoorn and H. van Zuylen : Empirical differences between time mean speed and space mean speed. *Traffic and Granular Flow'07, Springer*: 351-356, 2009.
17) Manual, H. C. : *Transportation Research Board of the National Academies*, Washington, DC, 2010.