Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Exploring the Propagation Pattern of Traffic Congestion through Analyzing and Visualizing Vehicle Detector Data

Chia-Wei Hsu[a], Yu-Ting Hsu[a]*

[a]*Department of Civil Engineering, Transportation Division, NTU, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan (R.O.C.)*

Abstract

The monitoring of roadway traffic conditions is critical for traffic management, where the detection of traffic congestion is one of major concerns. Traffic congestion may have various causes, including the increase of traffic volume due to higher private vehicle usage, inappropriate design or lack of capacity of road network and layout changes on the road segment owing to non-recurrent incidents such as traffic accidents or construction work. Hence, further understanding of how traffic congestion is formed, propagated and dissipates, and identifying possible bottlenecks are critical for overall traffic management, which may enable drivers and relevant agencies to more actively prevent traffic congestion and thereby improve the quality of traffic management strategies. Based on high-resolution VD data, this study integrates the consideration of data processing, pattern recognition and visualization to develop a data analysis framework for better understanding of traffic congestion in an urban network. Based on different thresholds of congestion detection, the spatio-temporal locations of congestion occurrences are recorded. The network structure is constructed based on the actual coordinate of VDs, map information and the concept of the adjacent matrix. An adjusted kernel density estimation approach is proposed and applied to case studies, in order to investigate the effects of congestion propagation on road segments with different characteristics in terms of connection type and adjacency. Finally, a general principle describing the propagation pattern of traffic congestion is concluded and presented through data visualization.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: congestion propagation pattern; *kernel density estimation*; *traffic state detection*; *vehicle detectors*; *data visualization*

1. Introduction

Traffic congestion is one of the main focuses of traffic management. It is a state when traffic demand exceeds roadway capacity. The characteristics of traffic congestion occurring within urban road networks can be considerably

* Corresponding author. Tel.: +886-2-33664237.
E-mail address: yutinghsu@ntu.edu.tw

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

different from those taking place on freeways due to the effects of traffic signals, intersections and the complexity of road networks. Traffic congestion can be further categorized into recurrent one which usually occurs during peak hours and non-recurrent one which results from a variety of incidents, such as traffic accidents, road construction as well as the activities of large events.

Researchers are interested in several related topics, including the formation of traffic congestion, the estimation of negative effects caused by traffic congestion, the bottlenecks in the road network and the strategies to prevent/ease congestion. To address the issues mentioned above, congestion incidents need to be identified from traffic data first. How to detect traffic congestion through a systematic approach of data collection and analysis has been the fundamental task for traffic management. In previous studies, traffic data are often extracted from loop detectors. However, there are some obvious shortcomings, such as the difficulties in facility maintenance, high malfunctioning and misdetection rate. Hence, other types facilities for detection, for example, electronic toll collection (ETC) sensors, monitors and microwave vehicle detectors (VD) are installed. ETC system has been operating on the freeways to enhance the performance of the freeway system. In addition to improving the service level of the freeway system, it also contributes to the collection of a large amount of traffic data. These data can be used for traffic management and opened to both academia and individuals for extended applications. In some urban areas, vehicle detectors are widely installed within the urban road network, and high-resolution traffic data are collected. They provide abundant records of traffic conditions including point travel speed, traffic volume and occupancy. The daily VD data are provided on the governmental open data platform. Through the investigation of these data, the characteristics of traffic flows can be observed, and a baseline traffic condition can be determined. By comparing the traffic data of a set of target VDs within a Region Of Interest (ROI) during a certain time period with the baseline traffic condition, congestion incidents can be detected. Traffic congestion may be manifested as a chain reaction, forming a shockwave across a certain scope of a roadway network (Li et al., 2013). Some studies on traffic congestion forecasting have been conducted by employing pheromone communication models (Kurihara et al., 2009), density wave models (Nagatani, 2002). To understand how a congestion incident may propagate throughout a network and dissipate based on the exploration of real data can be the research direction to further enhance urban traffic management.

Whether there are some differences in terms of the traffic flow characteristics and congestion propagation pattern between arterials under construction and the others in urban areas is worth investigating. The traffic data collected during the bike lane construction in Taipei is accounted for the case study section. To provide pedestrians and cyclists a safer environment, Taipei City government has been implementing a bike lane network plan since 2014. Considering the the time needed to eliminate the queue at traffic signals, three north-south arterials and three east-west arterials are selected. Each of them has a width of at least 40 meters, and metro routes were built along four of them. For those with wider sidewalks, for example, Jen-Ai road and Zhong-Shan N. road, marking lines for bike lanes are painted on the original sidewalks. For the other four routes, sidewalks are first broadened, and then marking lines are drawn. Residents had been reporting the congestion and inconvenience during the bike lane construction on Fu-Xing S. Road and Xin-Sheng S. Road from March to September in 2016. According to the travel speed collected from vehicle detectors, during the construction, travel speed slightly decreased by 6.49% to 7.91% and the service level had been degraded (Taipei City Traffic Engineering Office, 2016). However, the service level had almost recovered after the construction work was completed. Moreover, more detailed understanding of the relationships among neighboring road segments may also provide traffic management agencies and individuals valuable information for evaluating the influences of construction decisions, determining traffic management strategies and providing enhanced navigation service. Hence, high-resolution VD data during the construction in an ROI covering the arterials under construction are extracted for further analysis. Characteristics of the congestion propagation pattern including the conditional probability that a congestion may occur given the occurrence of another congestion, the potential relationship between adjacent road segments and how traffic congestion contribute to different road segments can be observed.

In this study, we seek to obtain better understanding of the pattern of how traffic congestion propagates and influences a roadway network. The traffic control center of Taipei City has provided a system for real time traffic status inquiry by plotting the road performance information on Google Map. Information that directly express traffic flow characteristics can be extracted based on the collected traffic data (Chen et al., 2015), while the cascading traffic pattern may further suggest driver behavior of diverting to circumvent congested road segments. Hence, the main purpose of this study is to go deeper to investigate the effects of congestion afterwards. To monitor where and when traffic congestion occurs, we take point vehicular speed for primary consideration. Based on the traffic data collected

from VDs, we cluster these data by capturing the spatiotemporal variation of vehicular speed over the network so as to identify congestion incidents. Based on the congestion incidents identified, the affected road segments can also be further determined. Ultimately, this research seeks to investigate the propagation of congestion incidents within an urban road network. By visualizing the bottle necks and shockwave after congestion occurs, some general principles in regard to the propagation can be observed so that a precautionary traffic management strategy may be taken.

In this study, we expect to have further understanding about the cascading pattern of traffic congestion based on the high resolution VD data, which may be a reference for the determination of traffic management strategies. The system for real time traffic status inquiry provided by the traffic control center of Taipei City and Google Map visualize instant traffic status in terms of vehicular speeds over the roadway network via a web-based inquiry interface, but we are more interested in the probability of congestion expanding to neighboring areas. To be more specific, the research objectives are summarized as below:

- Propose an alternated probability density estimation approach to properly compute the conditional probability that a congestion may occur on a certain road segment given the occurrence of another congestion.
- Determine the potential relationship between traffic of adjacent road segments based on the degree of adjacency and turning (straight, left turn or right turn) pattern.
- Visualize the density estimation result and discuss how congestion on a road segment may contribute to adjacent road segments and affect neighboring areas.

The remainder of the paper is organized as follows: Section 2 introduces the proposed methodological framework applying kernel density estimation (KDE) and explains the data processing and analysis procedure step by step. Next, case studies using the road network of Taipei City are performed, and results are visualized in Section 3. Finally, the conclusions of research findings and recommendations for future research are summarized in Section 4.

2. Methodology

This section proposes a methodological framework which applies an adjusted network KDE approach to determine the congestion cascading pattern in terms of the conditional probability of congestion incidents and the potential relationship between traffic of adjacent road segments within a portion of urban road network. To clarify how the road segments are connected, the conception of adjacency matrix is applied considering the coordinates of VD locations and using Open Street Map. The procedures to extract network information, preprocess VD data and perform KDE estimation by employing the proposed approach are also explained in the following sections.

2.1. Data Description

The two main components of our data are road network structure and pre-processed VD data. The conception of the adjacency matrix is introduced to form the road network structure of the selected ROI. In most networks in previous research, binary variables are used to indicate the presence of the links within, the edges between nodes are either existing or not (Newman, 2004). A network with such an attribute can be represented by an $n \times n$ adjacency matrix A with elements

$$\mathrm{A}_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise} \end{cases}$$

However, since all the VDs are located on the road segments, the associated adjacency matrix need to be edge based. Furthermore, most of the arterials in the selected road network are bidirectional, and thereby the direction of traffic is also considered. The connection of road segments within the selected ROI can be properly described by an $n \times n$ adjacent matrix with elements A_{du} where d is the entrance of a downstream road segment with respect to the exit of an upstream road segment u with elements

$$A_{du} = \begin{cases} 1 & \text{if } d \text{ and } u \text{ are connected,} \\ 0 & \text{otherwise} \end{cases}$$

Fig. 1 shows a sample road network, while Table 1 represents its 1st order adjacency matrix. When $A_{du} = 1$, a 1st order adjacency occurs. Even for those road segments with no VDs installed, their connection with other road segments can still be described by the 1st order adjacent matrix. The other matrix containing the turning information is shown in Table 2, the attributes are tagged as S (straight), L (left turn) and R (right turn). For those connection that cannot be identified easily due to non-orthogonal intersection, coordinate information in terms of longitude and latitude is utilized. This study requires both the 1st and 2nd order adjacency relationship and turning information. The 2nd order adjacency matrix can be obtained by performing a dot product of the 1st order adjacency matrix itself. In the 2nd order adjacency matrix, the elements with value 1, is named 2nd order adjacency. A 2nd order adjacency relationship indicates that two road segments are connected through another in-between road segment.

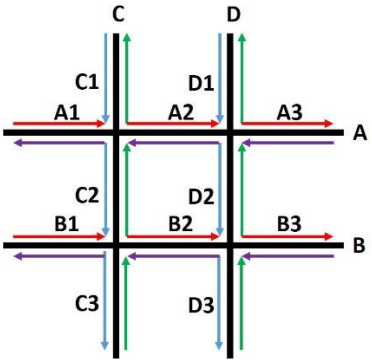


Fig. 1. Sample road network

Table 1. Sample adjacency matrix

		Upstream	*Arterial*	*A*	*B*	*C*	*D*	...
			ID	A2E	B2E	C2N	D2N	…
Downstream			*Direction*	East	East	North	North	…
Arterial	*ID*	*Direction*	*Segment*	CD	CD	AB	AB	…
A	A2E	East	CD	1	0	1	0	
B	B2E	East	CD	0	1	0	0	
C	C2N	North	AB	0	0	1	0	
D	D2N	North	AB	0	1	0	1	
…	…	…	…					…

Table 2. Sample turning matrix

		Upstream	*Arterial*	*A*	*B*	*C*	*D*	...
			ID	A2E	B2E	C2N	D2N	…
Downstream			*Direction*	East	East	North	North	…
Arterial	*ID*	*Direction*	*Segment*	CD	CD	AB	AB	…
A	A2E	East	CD	self	0	R	0	
B	B2E	East	CD	0	self	0	0	
C	C2N	North	AB	0	0	self	0	
D	D2N	North	AB	0	L	0	self	
…	…	…	…					…

The original dataset contains high resolution VD data (recorded every 5 minutes) in Taipei City from January, 2015 to March, 2017, provided by the Traffic Control Center of Taipei City Traffic Engineering Office. The raw data contain information including device ID, time and date, lane number, volume and travel speed of large vehicles and regular passenger cars, lane occupancy, and average interval between vehicles. Table 3 illustrates the important components in the VD data, which are useful for this study. Average travel speed is the major indicator for traffic congestion.The average travel speed of large vehicles and normal passenger cars are weighted respectively to account for the overall average travel speed.. Travel speeds on different lanes of a road segment are averaged. Data are then filtered by ROI and the time interval of interest (different peak periods of weekdays).

Table 3. Contents of columns

Columns	*Contents*
DeviceID	Name of vehicle detectors
DateTime2	Tag of date and time
LaneOrder	Number of lane
BigVolume	Volume of large vehicles
BigSpeed	Speed of large vehicles
CarVolume	Volume of regular passenger cars
CarSpeed	Speed of regular passenger cars
LGID	Identifier of the direction of traffic

Level of Service (LOS) C and average speed are the two criteria chosen in this study. The data pre-processing procedure is described as follows. First, data of the set of VDs within the ROI and target time intervals are filtered by string matching techniques. 96 road segments and 66 VDs are included in the considered ROI. Weekday data and weekend data are then separated. Second, obviously unreasonable values are viewed as erroneous data and removed. Third, incident charts for the two different criteria can be constructed, respectively. For LOS C criterion, according to section 19.6 in 2011 Taiwan Highway Capacity Manual (Transportation Planning Division, 2011), we consider level C as our threshold, which is often taken as the standard of light congestion by transportation management agencies. The definition of LOS is shown in Table 4. Under LOS C, a congestion is recorded if the travel speed is lower than 30 km/hr. The differences between the actual travel speed and the LOS C threshold are also calculated. By the criterion of average speed, the normal traffic condition is defined first so that we construct a baseline for reference. The baseline is set based on the weekly average of travel speed within the targeted time interval. Those lower than 80% of this baseline value are recorded; the difference between the actual travel speed and the baseline value is calculated as well. Finally, as a preparation step for further processing, data recorded from the previous step are transformed to a binary data structure. For negative values of the difference between the actual travel speed and the threshold, 1 is assigned

for them, while others are assigned 0. The value 1 indicates that a VD detects possible congestion or incident during the associated time interval, while 0 indicates an acceptable level of service.

Table 4. LOS criteria for urban road network with 50km/hr speed limit

Average Travel Speed V (km/hr)	*LOS*
V≥35	A
30≤V<35	B
25≤V<30	C
20≤V<25	D
15≤V<20	E
V<15	F

2.2. Adjusted Network Kernel Density Estimation (Adj. Network KDE)

In order to interpret the spatio-temporal characteristics of congestion propagation within an urban road network, an adjusted network kernel density estimation approach is applied. As a nonparametric probability density estimation approach (Rosenblatt 1956; Whittle 1958; Parzen 1962), kernel density estimation does not require any assumption for the distribution of data points. This provides more flexibility and allows researchers to discover more characteristics beneath the data set such as the actual distributions of relevant variables.

Rosenblatt (1956) and Parzen (1962) developed current form of kernel density estimation. Assuming $(x_1, x_2, ..., x_n)$ is a univariate independent and identically distributed sample extracted from some distribution with an unknown density f, the standard form of kernel density estimator can be written as:

$$f_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}) \tag{1}$$

K is the kernel function which is a non-negative symmetric function and satisfies $\int K(u)du = 1$. Since K is a probability density function, f also possesses the characteristics of the probability density function. In practice, several kinds of probability density functions are commonly selected as K, including uniform, triangular, parabolic, quartic and Gaussian. h is a positive number named bandwidth or smoothing parameter which controls the smoothness and preciseness of kernel density estimation.

In order to perform density estimation of various spatial related issues, the standard kernel density estimation concept is then extended to 2-D planes. The general form of the planar kernel density estimator in a 2-D space can be written as:

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{\pi r^2} k(\frac{d_{is}}{r}) \tag{2}$$

$\lambda(s)$ is the density at location s, d_{is} is the distance from point i to location s, and r is the bandwidth in the planar KDE. k is the kernel, modelled as a function of $\frac{d_{is}}{r}$ ratio. Instead of giving an equal weight to all points within bandwidth r, a distance decay effect is taken into account. That is, as the distance between a point and location s increases, that point is weighted less while calculating the overall density. Some commonly applied kernel functions are used to account for the distance decay effect, including Uniform, Triangular, Epanechnikov, Quatic and Gaussian function (Gibin et al., 2007; Levine, 2004).

To perform density estimation of point events with network constraints, network KDE is proposed (Xie et al., 2008).

Network space is used in the point event context, and the kernel function is developed based on network distance instead of Euclidean distance. Hence, it performs better on density estimation while a planar KDE may over-detect clustered patterns. The general form of the network KDE can be expressed as:

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{r} k(\frac{d_{is}}{r}) \tag{3}$$

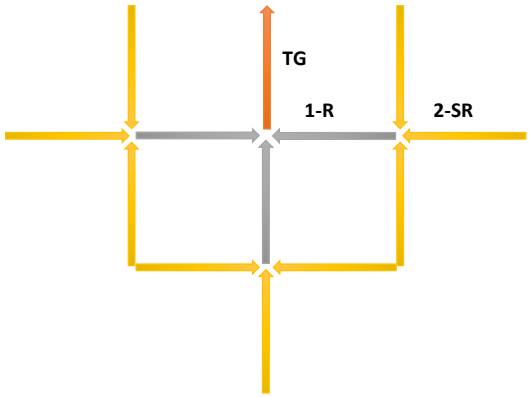
Instead of using network distance, this research employs the conception of "degree of adjacency" based on the structure of the road network and adjacency matrix. The locations of VDs do not follow a specific rule, for example, they can be at the front, middle or the end of road segments. Hence, there are difficulties measuring precise network distance. Furthermore, the conditional probability that congestion occurs on the upstream road segment given the occurrence of another congestion on the downstream road segment is also considered. The adjusted form of the network KDE can be written as:

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{r} p_{is} k(\frac{adj_{is}}{r}) \tag{4}$$

adj_{is} is the degree of adjacency of upstream road segment i with respect to downstream road segment s. p_{is} is the conditional probability that congestion occurs on i given another congestion occurring on s. To be more specific, s and i are both locations of VDs. In addition, each s can also be viewed as the center of several neighboring road segments including itself, which causes the effects of congestion on adjacent i s.

2.3. Analysis Procedure

Base on the road network structure and pre-processed data, the 1st order and 2nd order adjacency relationships of the road segments and binary incident charts of the VDs within the selected ROI can be obtained. The analysis procedure will be explained as follows. First, possible incidents are detected. A cell of the binary incident chart with value 1 indicates that possible congestion or incident takes place. For a single VD, if there is a sequence of value 1 that lasts for at least 4 time intervals (20 minutes), we define it as a possible congestion incident. Second, the conditional probability that incidents occur on neighboring road segments p_{is} is calculated. Duration of each congestion incident is recorded in the previous step. During the congestion incident on a certain road segment, the numbers of consecutive time intervals identified as congested on neighboring road segments are also recorded. We define the ratio of the latter (upstream adjacent road segment) and the former (downstream road segment) as the conditional probability that neighboring road segments are affected by the congested road segment. Third, the kernel density of each road segment is calculated. The kernel density can be calculated through Equation (4). A simple example is provided for illustration as follows. For the road network shown in Fig. 2(a), congestion occurs on the target road segment TG, road segment 1-R having the 1st order right-turn relationship with respect to TG, and another road segment 2-SR in the 2nd order straight-right-turn relationship with respect to TG. Two congestion incidents were detected on TG; one started from 6:45 AM and ended at 7:15 AM, while the other started from 8:20 PM and ended at 8:50 PM. Both lasted for six time intervals (30 minutes). Fig. 2(b) and Fig. 2(c) show the calculation of p_{is} for these two congestion incidents, respectively. 4 and 3 congestion intervals were detected on 1-R during the two congestion incidents on TG, respectively. On 2-SR, 3 congestion intervals are detected during the both congestion incidents on TG, respectively. Therefore, the kernel density calculation is $\lambda(\text{1-R}) = \frac{1}{3}\frac{4}{6}k(\frac{1}{3}) + \frac{1}{3}\frac{3}{6}k(\frac{1}{3})$ for 1-R and $\lambda(\text{2-SR}) = \frac{1}{3}\frac{3}{6}k(\frac{2}{3}) + \frac{1}{3}\frac{3}{6}k(\frac{2}{3})$ for 2-SR if $r = 3$ is chosen as the search bandwidth.



(a)

Segment \ Interval	~6:50	~6:55	~7:00	~7:05	~7:10	~7:15	p_{is}
TG							
1,R							4/6
2,SR							3/6

(b)

Segment \ Interval	~8:25	~8:30	~8:35	~8:40	~8:45	~8:50	p_{is}
TG							
1,R							3/6
2,SR							3/6

(c)

Fig. 2. (a) Road network sample for KDE; (b) p_{is} calculation example for the first congested incident; (c) p_{is} calculation example for the second congested incident

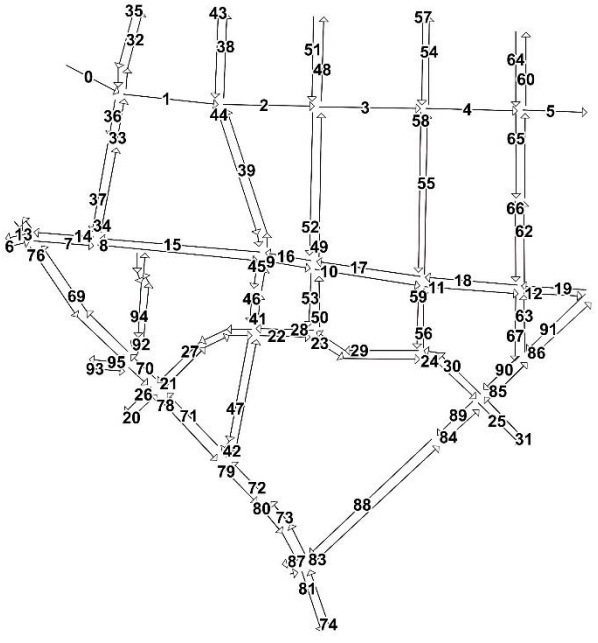
3. Case Study and Results

A case study is performed using the urban road network of Taipei City with the proposed approach applied. The dataset includes a real arterial network in part of the Daan district, Taipei City, Taiwan. The VD data from January, 2015 to March, 2017 are provided by Traffic Control Center of Taipei City. The point location of each VD is paired with a road segment. The network of the selected ROI is analyzed in terms of the kernel density of congestion. Different scenarios, including a day with a special event and a week during the construction of bike lanes are investigated. The analysis for each scenario is organized as overview, segment-wise perspective and summary. The overview perspective shows the plain view of the KDE results within the whole ROI, while the segment-wise perspective shows KDE results of the segments with relative high density and their neighboring segments. The criteria of both LOS C and average travel speed are adopted.

3.1. Descriptions of the Case Study

The selected ROI is defined by the boundaries constructed by 6 arterials within Taipei City. The ROI and locations of VDs installed are shown in Fig. 3(a), where road segments are represented by thicker lines, while the square dots represent VDs. Road segment within the ROI are numbered and expressed in Fig. 3(b). The ROI contains 96 road segments with different traffic directions separated. Totally 66 VDs are installed within this road network. Traffic data of peak hours during weekdays are extracted for analysis.

(a)



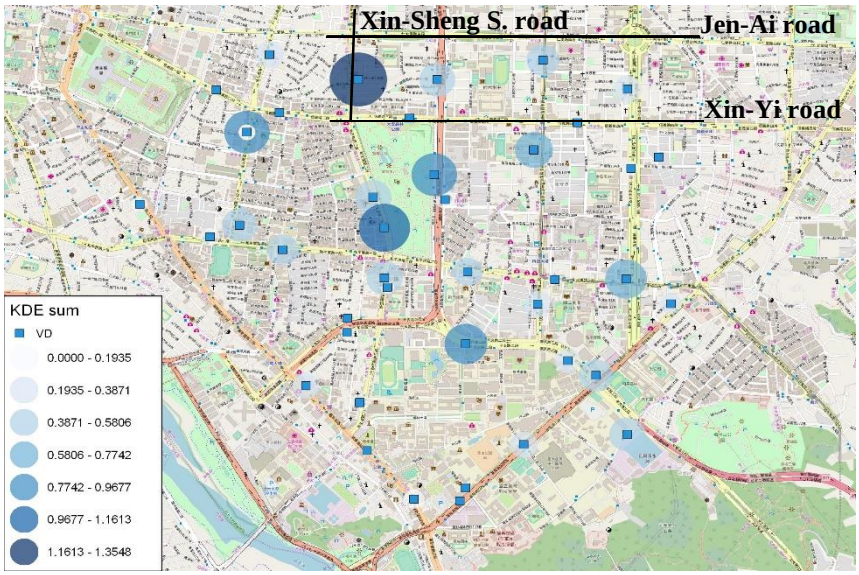
(b)

Fig. 3. (a) Road network of the ROI and location of VDs; (b) Numbered road segments of the ROI

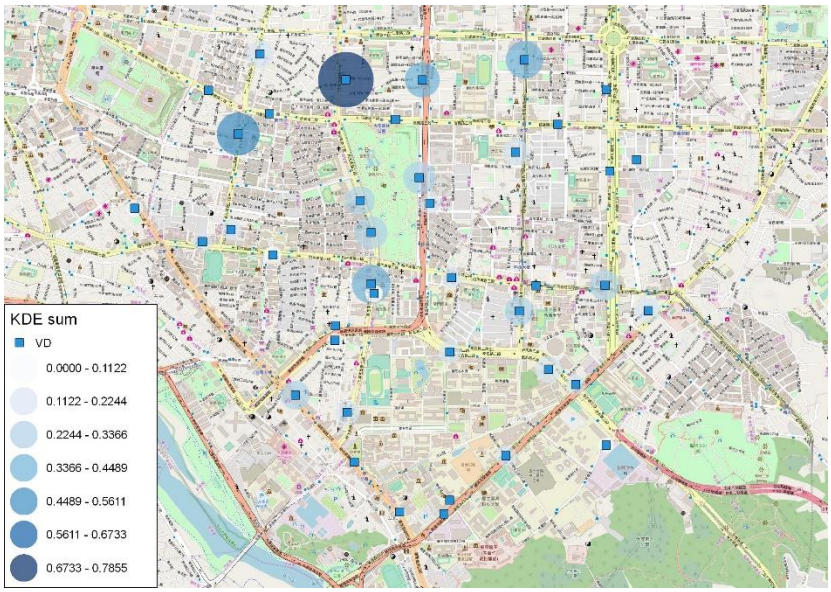
3.2. Result Analysis

The following result analysis is based on the kernel density estimation results over the selected ROI. The kernel density on each adjacent road segments of each day within the interested time periods are calculated respectively and then summed. The time period of scenario 1 is delimited from Dec. 28, 2015 to Dec. 31, 2015. The purpose of scenario 1 is to investigate the congestion propagation pattern during normal weekdays and a day with a special event, which is the New Year's Eve Celebration in this case study. During this time period, the bike lanes on Fu-Xing S. road and Xin-Sheng S. road were still under construction. The VD data from Dec. 28 to Dec. 31 in 2015 are extracted. The kernel density estimation of congestion is calculated for each road segment in the selected ROI.

For the LOS C criterion (30km/hr), the visualization of a KDE plain view is shown in Fig. 4(a) and 4(b) from an overview perspective. Larger circle and darker color represent relatively higher density. The VDs with relatively high density are located on Xin-Sheng S. road, especially for the segment between the two largest arterials of Taipei City: Jen-Ai road (north bound of the ROI) and Xin-Yi road (road segment 0~5). Other road segments can generally maintain LOS C within peak hours. However, comparing Fig. 4(a) with Fig. 4(b) (for Dec. 31), only the color saturation at the locations of hot spots on the heat map becomes slightly higher, indicating higher probability of the occurrence of congestion. For the road segments with relatively high density, further investigation and observation are needed, since they can be the potential sources of congestion propagation. Road segment 43 appears to have the highest density among all segments. However, since segment 43 is located at the north edge of the selected ROI, none of its upstream segments are accounted in this study. The possible congestion propagation to the upstream segments from the congestion originating at road segments 33, 43 and 44 is visualized in Fig. 5(a), 5(b) and 5(c), respectively. The thickness and darkness of the color mark represent the degree of influence. Segments without a color mark can be either providing minor contributions or having no VDs installed. The red color represents the road segments as congestion sources, while the upstream road segments are shown in gray scale. Darker color and thicker line segment indicates larger impact. The arrows indicate the travel directions. For road segment 33, the effect of congestion propagation to the upstream road segments are too minor to be observed. For the analysis of road segment 44, road segments of the 1st order adjacency are more likely to be affected, while road segments of the 2nd order adjacency are influenced less. Among all road segments of the 2nd order adjacency, the one that enters road segment 44 by a left turn may receive more contribution from the congestion source.

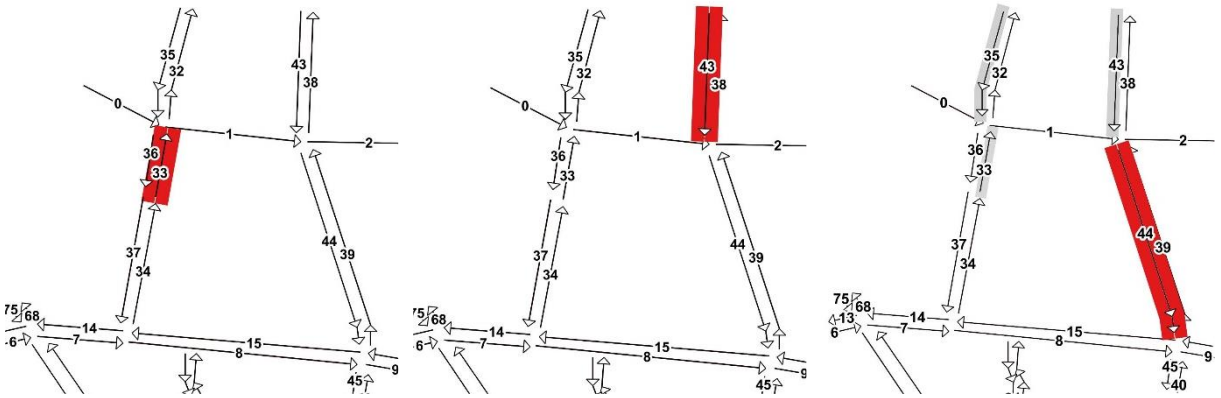


(a)



(b)

Fig. 4. (a) KDE of Scenario 1 with LOS C (2015/12/28~30); (b) KDE of Scenario 1 with LOS C (2015/12/31)

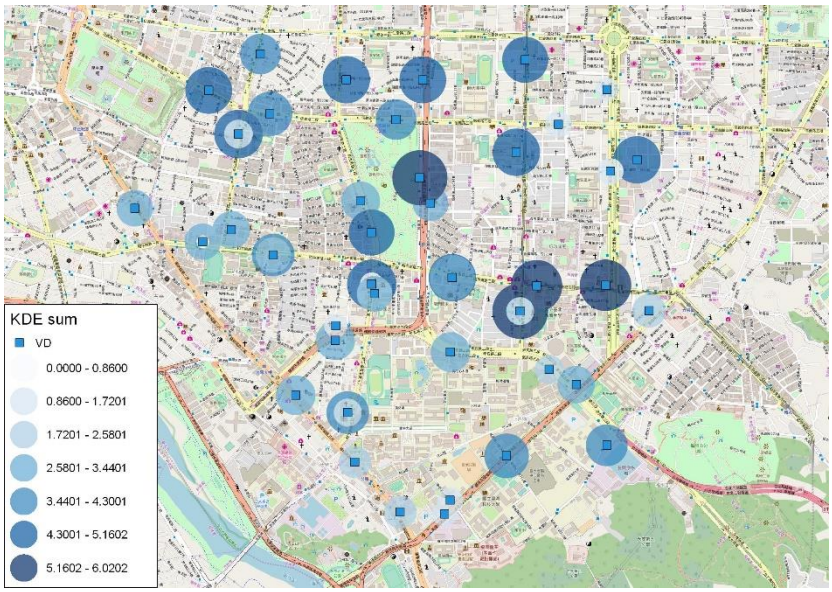


(a) (b) (c)

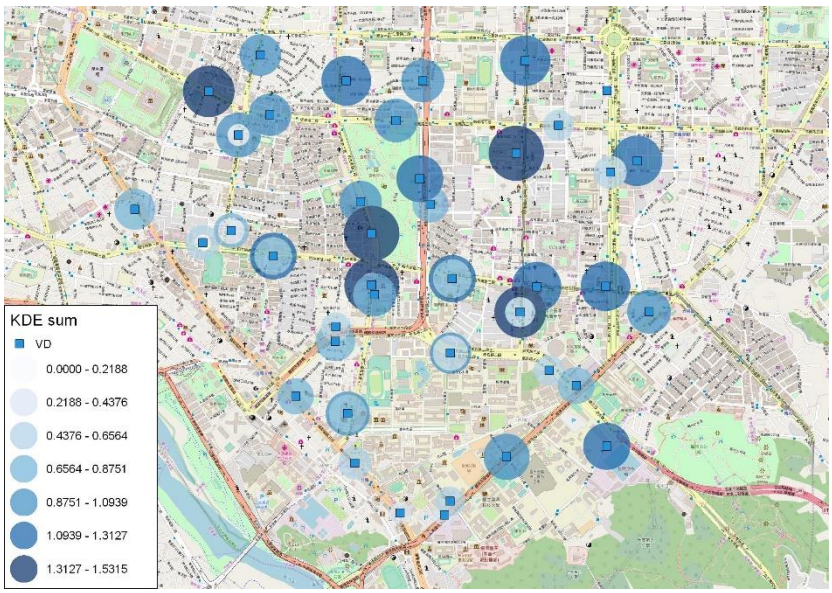
Fig. 5. (a) Upstream Influence from The Congestion of Segment 33 (S1_C); (b) Upstream Influence from The Congestion of Segment 43 (S1_C); (c) Upstream Influence from The Congestion of Segment 44 (S1_C)

For the criterion based on average speed, the visualization result presenting kernel density within the whole week from Dec. 28, 2015 to Dec. 31, 2015 is shown in Fig. 6(a). Another figure specifically presenting the kernel density on Dec. 31, 2015 is shown in Fig. 6(b). The locations with larger circles and darker colors are the road segments with higher kernel density. Similar locations of congestion hot spots can be observed through Fig. 6(a) and Fig. 6(b) compared with Fig. 4(a) and 4(b). We can observe that the road segments on Xin-Sheng S. road, Jian-Guo S. road and Fu-Xing S. road near He-Ping E. road have relatively higher density than other road segments, indicating higher probability of the occurrence of congestion. Road segments 40, 44 and 56 have the highest density among all segments; segment-wise analysis is performed with respect to these segments. The possible propagation to the upstream segments from road segment 40, 44 and 56 is visualized in Fig 7(a), 7(b) and 7(c), respectively. For the analysis of road segment 40, the road segments of the 1st order adjacency are more likely to be influenced. Comparatively, the road segments of the 2nd order adjacency are not affected as much as the road segment of the 1st order adjacency. Among those adjacent upstream road segments, the ones connected without turning have higher density. For the analysis of road segment 44, traffic on the road segments of the 1st order adjacency are more likely to be influenced as well. Since Xin-

Yi road only allows one-way traffic, there are no road segments entering road segment 44 by left turning. Among the two road segments of the 1st order adjacency, the effects are almost the same. The road segments of the 2nd order adjacency are not significantly affected. Among the three upstream road segments of the 2nd order adjacency, the effects are also nearly the same. For the analysis of road segment 56, the road segments of the 1st order adjacency are more likely to be affected. Among the road segments of the 1st order adjacency, the one that enters by left turning has higher density than the one that enters by right turning. The road segments of the 2nd order adjacency following the road segment of the 1st order adjacency by right turning are influenced less. For all the road segments of the 2nd order adjacency, the ones that are not connected with left turning may receive more effects on traffic from the road segment 56 (source of congestion) than the other.



(a)



(b)

Fig. 6. (a) KDE of Scenario 1 with Average Travel Speed (Dec. 28-30, 2015); (b) KDE of Scenario 1 with Average Travel Speed (Dec. 31, 2015)

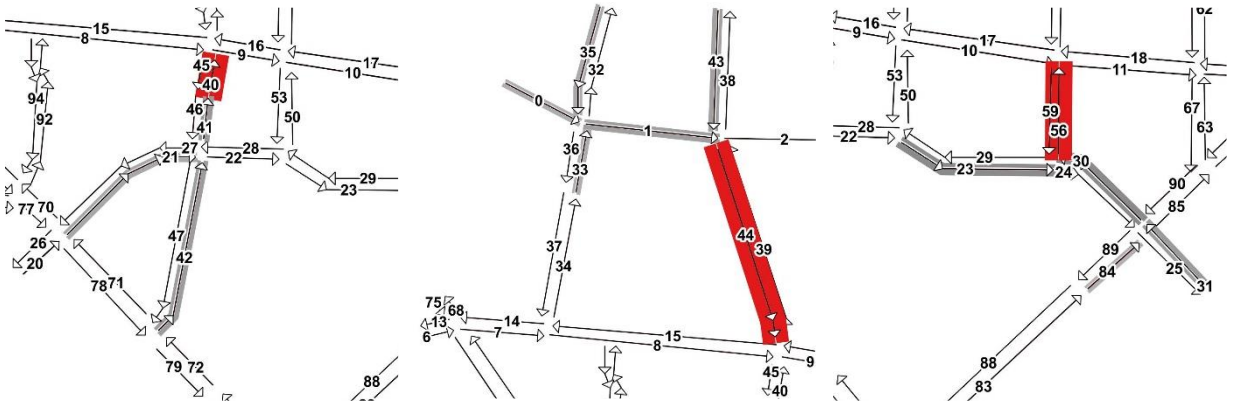


Fig. 7. (a) Upstream Influence from The Congestion of Segment 40 (S1_ avg); (b) Upstream Influence from The Congestion of Segment 44 (S1_ avg); (c) Upstream Influence from The Congestion of Segment 56 (S1_ avg)

The construction of the bike lanes had caused occupation of lanes originally used by motorized vehicles and changed the layout of road segments. To investigate the congestion propagation pattern on weekdays during the construction of bike lanes, a week (Apr. 18-22, 2016) around the end of the construction is chosen for Scenario 2. Kernel density estimation is performed on each road segments in the selected ROI.

From the overview analysis, the results based on the LOS C criterion is shown in Fig. 8. The area with high density primarily lies on Xin-Sheng S. road. However, the service level of some road segments on the parallel arterials including Jin-Shan S. road and Jian-Guo S. road may be degraded as well. The impact on the road segments seems to be relatively local, indicating that most of the congestion incidents do not cause wide propagation. Similar to what we have observed from the overview perspective, even for the road segments with higher density, congestion does not spread extensively. The possible congestion propagation to the upstream segments from the source on road segments 33, 39 and 84 is visualized in Fig. 9(a), 9(b) and 9(c), respectively. Road segments 43 and 51 have higher density than road segment 84; however, their upstream road segments are not included in ROI. By contrast, in the analysis of the road segments for this period, the effect of congestion propagation to upstream road segments are too minor to be observed.

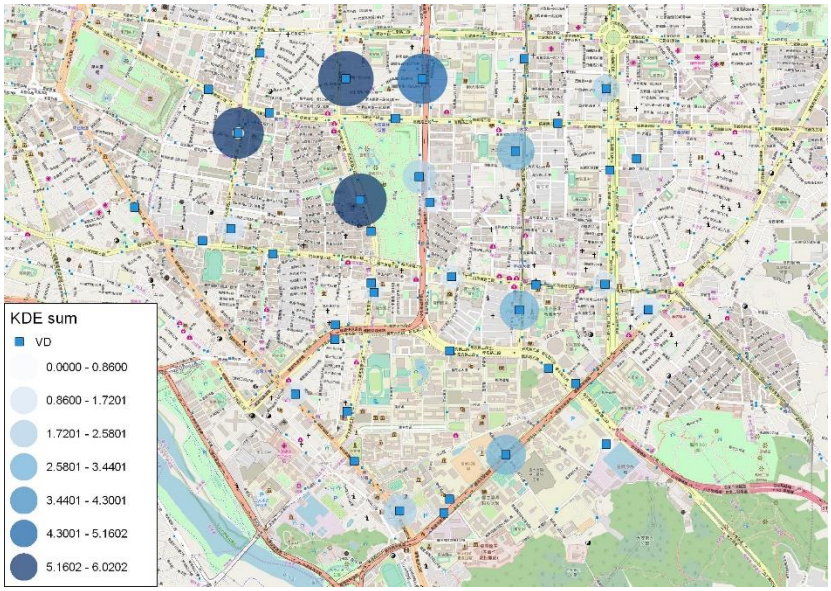


Fig. 8. KDE of Scenario 2 with Level C of LOS (2016/4/18~22)

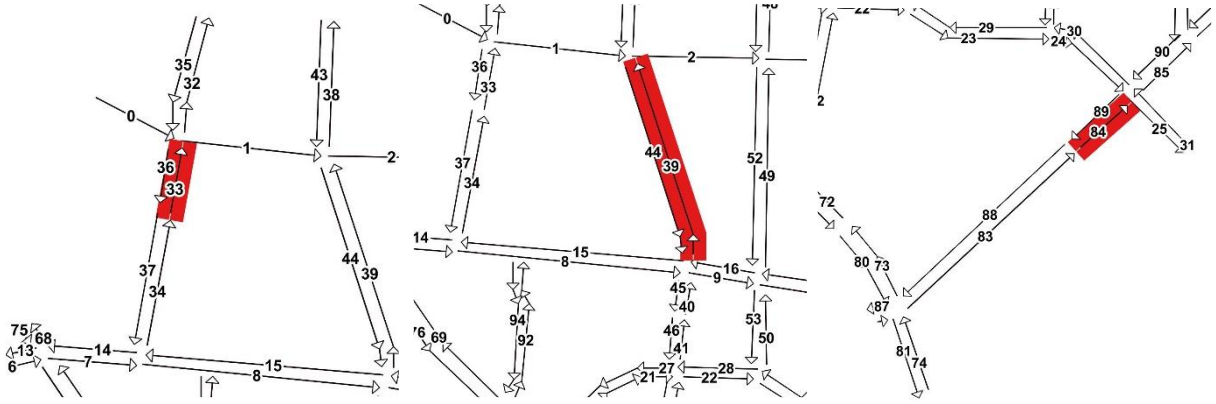


Fig. 9. (a) Upstream Influence from The Congestion of Segment 33 (S2_C); (b) Upstream Influence from The Congestion of Segment 39 (S2_C); (c) Upstream Influence from The Congestion of Segment 84 (S2_C)

For the average speed criterion, the overview visualization result is shown in Fig. 10. The distribution of higher density areas is roughly unchanged, however, with lower color saturation. This may indicate that the impact from the construction had eased to some extent. Road segments 40, 46 and 51 are associated with the highest density among all road segments. However, the adjacent road segments of road segment 51 is not covered within ROI. Hence, the effects on road segment 58 with the 4th highest density is investigated, and the pattern on road segment 56 is observed once again to be compared with Scenario 1. The possible propagation to the upstream segments from the congestion sources on road segments 40, 46, 58 and 56 is visualized in Fig. 11(a), 11(b), 11(c) and 11(d), respectively. For the analysis of road segment 40, there are one 1st order and two 2nd order adjacent upstream road segments. For the road segment of the 2nd order adjacency, the one that enters the road segment of the 1st order adjacency by left turning has smaller density than the other one connected without turning and the road segment of the 1st order adjacency. However, the road segments of the 2nd order adjacency and connected without turning has higher density than the 1st order adjacent one. For the analysis of road segments 46 and 58, the effect of congestion propagation to upstream road segments are too minor to be observed. For the analysis of road segment 56, similar patterns can be observed as they are in Scenario 1. The road segments of the 1st order adjacency are more likely to be affected. Herein, the road segment that are connected with left turning has higher density than the one connected with right turning. However, the angle of right turning here is around 135 degrees, which may be defined between a right turn or no turn. The road segments of the 2nd order adjacency entering the road segment of the 1st order adjacency by right turning are influenced less. Among all road segments of the 2nd order adjacency, there are no significant difference between them in terms of the possible contribution to congestion received from the source occurring on road segment 56.

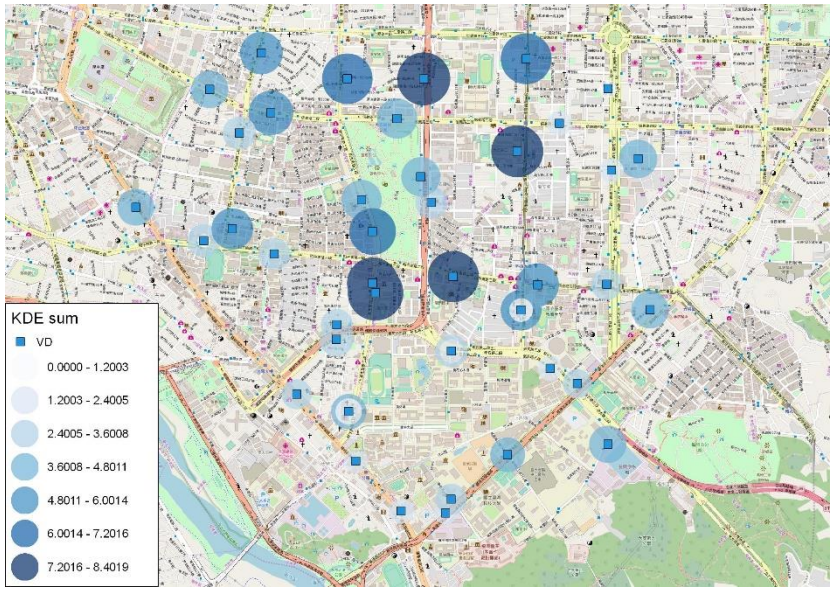


Fig. 10. KDE of Scenario 2 with Average Travel Speed (Apr. 18-22, 2016)

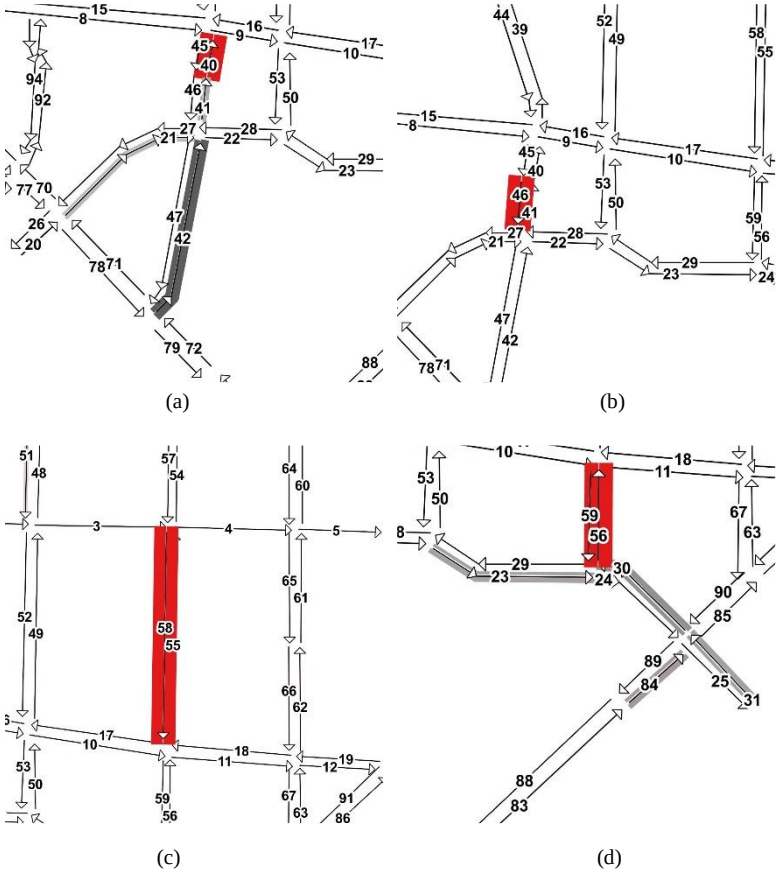


Fig. 11. (a) Upstream Influence from The Congestion of Segment 40 (S2_avg); (b) Upstream Influence from The Congestion of Segment 46 (S2_avg); (c) Upstream Influence from The Congestion of Segment 58 (S2_avg); (d) Upstream Influence from The Congestion of Segment 56 (S2_avg)

3.3. Summary of Insights from Case Study

During the construction of bike lanes, the layouts of road segments were changed. The probability of the occurrence of the congestion had changed as well, compared with the situation before the construction, and was higher on the arterials with bike lane construction than elsewhere. The analysis results from Scenario 1 show that for a day with special event, the size of the impacted area may increase, indicating that a single congestion incident is likely to spread wider. Based on threshold of LOS C, most of the road segments performed well even during the bike lane construction. Nevertheless, if the criterion of average speed is applied, which represents the daily traffic baseline, the result shows that the fluctuations of travel speed and congestion incidents actually happen from time to time. Hence, a varying value of average travel speed may be a better standard for more active traffic management.

From the segment-wise perspective, the road segments of the 1st order adjacency usually have higher density than the road segments of the 2nd order adjacency, which is consistent with the common knowledge of traffic management that spatially closer locations have stronger connection to each other. For the connection between congestion source and its 1st order adjacent upstream road segments, we may conclude that generally upstream flows which go straight to the congested road segment are affected most by the source. The segments connected with left turning come second, and the segments with right turning receive the least influence. For the connection between the road segments of the 1st order adjacency and 2nd order adjacency, similar characteristics can be observed. Each road segment in a grid network can have at most three 1st order adjacent upstream road segments and nine 2nd order adjacent upstream road segments. However, not every road segment within the road network is orthogonally connected with each other. Additionally, not all of them have VDs installed. Hence, there may be some simplifications in this study. In general, the proposed approach can be sufficient for the research objectives and flexible for extended applications.

4. Conclusion and Outlook

Due to the advanced sensor technology, high-resolution vehicle detection data are accessible and can provide abundant information for traffic management. However, in the existing literature, these data have not been fully explored and utilized. This research develops a framework by using an adjusted KDE approach to estimate the effects of congestion propagation in an urban roadway network, which includes the procedure of data preprocessing, analysis and visualization. Based on the VD data in Taipei City, this research presents a VD data analysis framework composed of congestion (incidents) detection, KDE, visualization of congestion hot spots and propagation patterns. The bike lane network construction since 2014 is used for the case study to investigate the propagation pattern during network layout changes.

Extended from the research background, literature review, construction of a KDE based spatial analysis framework and the case study, the research insights of this study can be concluded as follows. This research utilizes the adjusted kernel density estimation approach to compute the effects of congestion on road segments. By constructing the network structure, we not only record the location and adjacency of neighboring road segments, but also identify how they are connected in terms of traffic flow dynamics. This non parametric approach allows us to better understand the spatial characteristics of traffic flow evolution over a network. This study displays a complete framework for analyzing VD data. We apply the criteria suggested from the latest 2011 version Highway Capacity Manual and the average travel speed calculated from the data itself. The case study is conducted to test the feasibility of applying them as the congestion thresholds and provide the visualized results, which can help identify the characteristics of congestion propagation patterns under different event and network layout changes. The relationship between neighboring road segments and the influence contributed by the congestion source segment are clarified. A pattern of congestion propagation can be found which is consistent with the general knowledge about traffic management. However, that some part of road network that is not a typical grid network has slightly different outcomes while most part of the network is typical and follows the general pattern. The proposed framework can still provide the visualization of propagation pattern for each road segment. Instead of plotting data on the time line to observe the fluctuation. The propagation of congestion may be visualized to some extent. However, this study contributes on several different perspectives. Our proposed methodology can not only visualize the propagation itself but also extract its characteristics. Furthermore, it has good expandability to compare with historical data and the ability to predict future traffic state with urban road network, which is valuable reference for traffic management.

To further enhance the analysis of urban VD data and its applications for traffic management, there are several considerations and suggestions for the future work, which may expand the use of the data analysis framework and provide referential information for better quality decision-making. The relevant aspects are listed as follows. To make the results more reliable, the malfunctioning rate and the amount of missing data need to be decreased. On the other hand, it may be addressed by either installing more VDs to fill the vacancy spot or developing proper data imputation approaches. The reliability of different data imputation approach need to be further tested as well. New data can be further included to form a larger data set. More generalized base line traffic conditions can be determined. Also, by arbitrarily choosing certain part of data set for more case studies, the congestion propagation pattern under different circumstances can be identified. The outcome can be provided as a network evaluation reference for transportation engineering and management agencies. For the adjusted network KDE approach proposed in this study, the spatial relationship of road segments and the conditional probability p_{is} are considered in the equation. However, the closeness between congestion incidents in terms of time dimension is not directly used in calculating p_{is}. Hence, elements that can properly represent the time dimension can be further investigated. There are some difficulties in extracting road network information. For example, road networks may not always be typical grid network and the map information is not well organized or reliable. Problems for constructing small networks may be fixed manually. However, to implement this approach to a larger network, automatic network extracting technique need to be further developed. More studies can be conducted by focusing on how to simplify different network structures while extracting their commonality. We may be able to make some basic explanation about some phenomenon slightly different from the general findings we come up with. However, what kind of simplification is allowed and more types of road segment connection patterns still need further investigation.

References

Anbaroglu, B., Heydecker, B., & Cheng, T. (2014). Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. Transportation Research Part C: Emerging Technologies, 48, 47-65.

Chang, C. T.(2012). Integrating Data Mining and Kernel Density Estimation to Assess Common Physiological Indicators of Multiple Diseases. Doctoral Dissertation of Department of Industrial Engineering and Management, Yuan Ze University.

Chen, W., Guo, F., & Wang, F. Y. (2015). A survey of traffic data visualization. IEEE Transactions on Intelligent Transportation Systems, 16(6), 2970-2984.

Chen, Z., Liu, X. C., & Zhang, G. (2016). Non-recurrent congestion analysis using data-driven spatiotemporal approach for information construction. Transportation Research Part C: Emerging Technologies, 71, 19-31.

Coifman, B. (2002). Estimating travel times and vehicle trajectories on freeways using dual loop detectors. Transportation Research Part A: Policy and Practice, 36(4), 351-364.

Coifman, B. (2003). Identifying the onset of congestion rapidly with existing traffic detectors. Transportation Research Part A: Policy and Practice, 37(3), 277-291.

Daqing, L., Yinan, J., Rui, K., & Havlin, S. (2014). Spatial correlation analysis of cascading failures: congestions and blackouts. Scientific reports, 4, 5381.

Hu, Y. H. (2012). Using GPS Logging Data to Estimate Saptial Distribution of visitors-A Case Study of Yehliu Geopark. Master's Thesis of Department of Geography, National Taiwan University.

Ji, Y., Luo, J., & Geroliminis, N. (2014). Empirical observations of congestion propagation and dynamic partitioning with probe data for large-scale systems. Transportation Research Record: Journal of the Transportation Research Board, (2422), 1-11.

Jiang, Y., Kang, R., Li, D., Guo, S., & Havlin, S. (2017). Spatio-temporal propagation of traffic jams in urban traffic networks. arXiv preprint arXiv:1705.08269.

Kerner, B. S., Demir, C., Herrtwich, R. G., Klenov, S. L., Rehborn, H., Aleksic, M., & Haug, A. (2005, September). Traffic state detection with floating car data in road networks. In Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE (pp. 44-49). IEEE.

Kurihara, S., Tamaki, H., Numao, M., Yano, J., Kagawa, K., & Morita, T. (2009, May). Traffic congestion forecasting based on pheromone communication model for intelligent transport systems. In Evolutionary Computation, 2009. CEC'09. IEEE Congress on (pp. 2879-2884). IEEE.

Li, X., She, Y., Luo, D., & Yu, Z. (2013). A traffic state detection tool for freeway video surveillance system. Procedia-Social and Behavioral Sciences, 96, 2453-2461.

Lin, F. B., Tseng, P. Y., Lin, K. S., Su, C. W., Chan, C. W., Cheng, C. Y., Lu, Y. C., Liu, K. C., Chen, C. Y., Wang, Y. F. (2011). 2011 Highway Capacity Manual in Taiwan，Institute of Transportation, Ministry of Transportation and Communication.

Long, J., Gao, Z., Ren, H., & Lian, A. (2008). Urban traffic congestion propagation and bottleneck identification. Science in China Series F: Information Sciences, 51(7), 948.

Meng, L. M., Han, L. S., Peng, H., Zhang, B., & Du, K. L. (2014, August). A machine learning approach to urban traffic state detection. In Intelligent Control and Information Processing (ICICIP), 2014 Fifth International Conference on (pp. 163-168). IEEE.

Mirge, V., Verma, K., & Gupta, S. (2017). Dense traffic flow patterns mining in bi-directional road networks using density based trajectory clustering. Advances in Data Analysis and Classification, 11(3), 547-561.

Nagatani, T. (2002). The physics of traffic jams. Reports on progress in physics, 65(9), 1331.

Ni, D., & Wang, H. (2008). Trajectory reconstruction for travel time estimation. Journal of Intelligent Transportation Systems, 12(3), 113-125.

Parzen, E. (1962). On estimation of a probability density function and mode. The annals of mathematical statistics, 33(3), 1065-1076.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics, 832-837.

Sun, Z., & Ban, X. J. (2013). Vehicle trajectory reconstruction for signalized intersections using mobile traffic sensors. Transportation Research Part C: Emerging Technologies, 36, 268-283.

Treiber, M., & Helbing, D. (2002). Reconstructing the spatio-temporal traffic dynamics from stationary detector data. Cooperative Transportation Dynamics, 1(3), 3-1.

Wang, Z., Lu, M., Yuan, X., Zhang, J., & Van De Wetering, H. (2013). Visual traffic jam analysis based on trajectory data. IEEE Transactions on Visualization and Computer Graphics, 19(12), 2159-2168.

Wu, J. J., Sun, H. J., & Gao, Z. Y. (2007). Cascading failures on weighted urban traffic equilibrium networks. Physica A: Statistical Mechanics and its Applications, 386(1), 407-413.

Wu, Z. X., Peng, G., Wang, W. X., Chan, S., & Wong, E. W. M. (2008). Cascading failure spreading on weighted heterogeneous networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(05), P05013.

Xie, Z., & Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. Computers, environment and urban systems, 32(5), 396-406.

Yeh, Z. H. (2016). The Impact of Bike Lane Network Construction Press Release. Taipei City Traffic Engineering Office.

Yu, M. C. (2009). A Study on Kernel Density Estimation. Master's Thesis of Department of Mathematics, National Tsing Hua University.

Zhu, W., & Barth, M. (2006, September). Vehicle trajectory-based road type and congestion recognition using wavelet analysis. In Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE (pp. 879-884). IEEE.