東京大学大学院新領域創成科学研究科
社会文化環境学専攻

2020 年度
修　士　論　文

Data-based Travel Time Sensitive Analysis and Bottleneck Detection in Urban Road Network

データによる旅行時間の感度分析および都市道路網におけるボトルネック検出

2020 年 7 月 10 日提出

指導教員　柴崎 亮介　教授

喩　騰瑶
Yu Tengyao

Acknowledgements

I would like to express my acknowledgements to all the people who provide me help to finish this master thesis.

First I want to acknowledge my supervisor, Prof. Shibasaki. He gave me many valuable advices on my master thesis, about the directions, the objectives and the methodologies. During writing this thesis, Prof. Shibasaki has spent much time discussing with me and pushing me to think carefully about the details. In addition, the most important thing I learn from Prof. Shibasaki in this three year is how to find a meaningful question and to solve it. Instead of learning a new method and try to apply it, I think the difference on study motivation will help me to do more useful and successful project in the future.

I would also like to give my sincere gratitude to my seniors in Shibasaki Lab, Dr. Zhang, Dr. Chen and Dr. Huang. Dr. Zhang gave me many suggestions about the methodologies to achieve my objective. Dr. Chen spent much time discussing with me about the framework of my work and this thesis. Due to their suggestion, I can complete this master thesis. As my tutor, Dr. Huang provided a lot of help on daily life and study abroad. The same I want to give my acknowledgement to all the members in Shibasaki Lab, who gave me help in my study and often discuss to push new idea.

Then I want to appreciate all the teachers in CSIS, University of Tokyo, who has helped me on my study and this thesis. The teachers in CSIS collect and process the data, providing more potential for us to do analysis.

Finally I would like to express my gratitude to my parents. They are not rich but they work hard to support my dream to study abroad. Their behavior and words always give me power to finish the master thesis.

Abstract

Traffic is a key topic for both scholars and governments as an important part of people's daily life. Even though there exist many successful studies now, due to the rapid change of the road network and information system, there is still a lot of potential that can be explored by analyzing the road network. In recent years, as the popularization of data-based Intelligent Traffic System (ITS), more and more dynamic methods are presented to optimize the traffic network, which require us to know about traffic condition and traffic change more dynamically. Especially, it attracts how traffic flow influences travel time when traffic flow and changed by strategies like dynamic traffic signals and road congestion tolling schemes. In this thesis, I present a framework centered on travel time estimation and sensitive measurement to evaluate the sensitivity of a road network to indicate the degree of travel time changes according to the small change of traffic volume. The estimation approach for link travel time can support other dynamic traffic models, which requires the estimation under stochastic or changing traffic demand. And the sensitive measurement provides an approach to evaluate the robustness of the whole road network. In detail, based on GPS data, the change of travel time is first estimated on each link according to the traffic volume on the individual links. Then a transformation matrix is constructed to connect the traffic volume on each link and the travel time for each route. By analyzing the matrix, the sensitive measurement is calculated and analyzed during one day. Finally, based on the sensitive measurement, a bottleneck for road network is defined and detected as the links which influence the sensitivity of the whole road network most. From the analysis of the results, the sensitivity measurement can indeed reflect the traffic condition in the network and it is influenced by the traffic volume and traffic volume distribution on links.

Key words: Traffic Network, Travel Time Sensitivity, Traffic Bottleneck

Contents

List of Table

List of Figures

Chapter 1 Introduction

1.1 Background

1.1.1 Road Traffic Network

In recent years, as the development of traffic, people move frequently every day, by walk, by car, by bus, by subway, or even by train and airplane. Among all of the travel mode, we rely most on cars. A national travel survey in England in 2016 shows that, most of trips are taken by car, around 62 percent. And the travel distance by car takes up 78%. Similarly, in most countries, the car is the main travel mode when people go out to shop, to commute and to relax. There is no doubt that, road traffic takes a very important role and indeed influences our daily life now. For individuals, they need road information to make route choice and time schedule. For commercial companies, they can utilize the road traffic information to provide better service and make more profit. Map apps like Google maps are hoped to recommend the best path with the shortest time for drivers. It requires them to predict travel time accurately according to real-time traffic condition data. Other companies related to public transportation can also make use of the traffic information and innovate new kinds of service. For example, car sharing service grows rapidly in the recent years. It allows people to rent a car conveniently for a short period of time (usually a few hours) and pay for it. Companies have to decide where to build car stations and how many cars should be put in the station to meet most trip demand (Fig. 1). Another example is hitch hikes. With the analysis of a large number of daily trips, it is found that quite a bit of trips has similar trajectory. Companies like DiDi (in China) set up a system for passengers to hitch a ride and drivers can earn money from the hitch. And finally, for traffic administrators, traffic management is surely an important part of city management. In order to let traffic efficient, the department of traffic management must have adequate knowledge of traffic condition. Based on this knowledge, they can improve traffic by building new roads, adjusting traffic signals or charge a road congestion tolling. In conclusion, traffic is so important for us to study and make it better.

Table 1 Travel Mode Statistics in England (2016)

TRAVEL MODE	TRIPS	DISTANCE
Car	62%	78%
Bus	5%	4%
Subway	3%	10%
Walk	25%	3%
Bicycle	2%	1%
Others	2%	4%

Fig. 1 Car Sharing Station of Zipcar in the U.S.A

The researches on traffic started along with the invention of automobiles. In the history of automobile, when the Ford Model T created by the Ford Motor Company in 1908, it became possible to mass-produce automobiles. After that, as many companies optimized their production to reduce production time and cost, more and more working families could support to buy an automobile. Meanwhile, traffic congestion became a serious problem. Government and researchers began to study and optimize the traffic condition. They built new roads, set different lanes, set traffic signs and signals in intersections and so on. As the develop of the road building in these tens of years, urban road network has become a large and complex network for us to study.

To keep the efficiency and safety of urban traffic network, scholars made a lot of effort on the study of traffic network. And there are mainly four kinds of researches on traffic topic: 1) traffic simulation, 2) traffic prediction, 3) traffic analysis and 3) traffic optimizations.

The first one is about traffic simulation, which aims to simulate real driving behavior and traffic network accurately. For example, Daganzo (1992) presented a cell transmission model to simulate macroscopic traffic flow and Gipps (1981) used a microscopic view to simulate car-following behavior. Engineers also built many traffic simulation tools based on these simulation methods, e.g. Sumo, MatSim and AnyLogic, which are commonly used in traffic related projects or academic studies. These traffic simulation researches are the basis of most researches in the field of traffic.

The second one is working on traffic prediction. Ben-Akiva (1998) presented a framework to predict or estimate traffic on the traffic simulation system DynaMIT. Shu et al. (1999) presented a FARIMA model to predict time-series traffic flow data dynamically. In recent years, there are many advanced methods that can predict more accurately and quickly, for example, Lv et al. (2014) achieved the prediction via deep learning approach. These prediction studies of traffic flow are the foundation for the building of smart transportation system and applications.

The third type focus on the traffic analysis, which evaluate the traffic condition or find the issues in the network. For example, Crovella & Kolaczyk (2003) did a spatial network analysis via graph wavelets and Tian et al. (2016) built a topology network to evaluate the efficiency of the roads. These studies set up measurement and tell the traffic managers how much the traffic condition is good or bad in a specific area. They also provide directions where it should be improved.

And the final kind of studies aim to improve and optimize traffic network. For example, many studies presented their approaches to get a optimal traffic signal schemes (Sanchez et al., 2004; Sun et al., 2006; Christofa & Skabardonis, 2011). And in recent years, traffic tolling becomes a hot topic, which is hoped to control traffic flow and reduce traffic congestion. Some scholars present their tolling schemes to maximize the social welfare (Bergendorff et al., 1997; Yin & Lou, 2009; Sharon et al., 2017). This kind of studies can directly give out suggestions to the department of traffic management.

Until now, traditional traffic flow theory or traffic simulation tools (e.g. sumo, MatSim, AnyLogic) got difficult to solve the large network problem. Fortunately, new methods and data collection technologies provide more potential and new chances for traffic researches. The department of traffic management are also pushing the cooperation between academic researches, hoping that researches can provide them evidence to understand traffic features more deeply and to lead better policies. Many cities are now pushing the building and spread of an Intelligent Transportation System (ITS), an intelligent application to monitor real-time traffic condition, control traffic flow and provide other innovative service. This system and the data collection technology allow us to excavate more potential in data.

1.1.2 Travel Time Estimation

In the studies about traffic network, it is no doubt that travel time is the most important measurement that every scholar focuses on. Individuals hope that the travel time can be short and government rely on the travel time to evaluate traffic condition or the traffic optimization schemes. On the other hand, when a company presents a new service, travel time is certainly a key word that they must consider about and can attract users. As the building of ITS, both drivers and traffic managers are supposed to get known of real-time traffic condition and make better decision based on it. Then, it becomes a new requirement that travel time must be estimated under different traffic condition in a quick way so that it could be applied into other dynamic or real-time traffic optimization model, such as route choice model with congestion and road charging.

In past researches, there are three main approaches to estimate travel time. The first one bases on traditional traffic flow theory. For example, cell transmission models (CTM) (Daganzo, 1992) is commonly used in papers of traffic fields. After many scholars' improvement and adjustment, it is supposed to be accurate and near to real traffic. However, it requires complex calculation when calculating the model. Therefore, it becomes difficult when it comes to solve the question in a large scale.

The second approach is to utilize multi-agent simulation like sumo and MatSim. Those who aim at traffic optimization models in the field of operation researches usually prefer this approach because these tools allow them to focus on their optimization model with less knowledge of traffic. Garcia-Nieto, Olivera & Alba (2013) presented a traffic signal optimization framework and they estimated travel time under different traffic signal schemes by simulation tool sumo. But the tools also limit their research by the function they can provide. For example, if a traffic tolling scheme is supposed to set in a certain area, the tolling way is limited by the tolling scheme that the simulation tools provide. In addition, it usually needs to be calculated iterations by iterations when searching for the optimal solution, which means the traffic simulation has to be run so many times. As it costs too much time, it is difficult to be applied in a dynamic situation.

The mentioned two traditional approaches are similar that they can estimate travel time accurately to a certain degree but may cost a lot of time when the network is large and the calculation process must be repeated. In recent studies, the progress of intelligent algorithms and machine learning bring new potential to the study of estimating time. With many hidden layers, the estimation result more accurate than the previous two method and calculation speed is also very fast (e.g. Jenelius & Koutsopoulos2013).

However, both simulation results and intelligent algorithms let the estimation process is like a "black-box". They can't reflect the travel time trend or the relationship between other traffic parameters. In fact, in some special cases, the estimation of travel time can be a little rougher when we care more about the trend instead of how much the time is. For example, Çolak et al. (2016) gave a mathematical model to simulate how drivers choose route according to the change of travel time. When the travel time of different route is estimated, it is represented as the function of the trip number which passes this link. By this approach, it can simulate how the drivers compete with other ones, how the drivers response to the traffic condition and how travel time will change according to their choice. As the formulation is easy in implement, it can be efficient in this kind of dynamic traffic models. Similarly, Clark &Watling (2005) presented link travel time as a polynomial function of traffic flow. Because it focused on the travel time under stochastic demand, the estimation of travel time must be simplified.

In conclusion, there are many successful studies providing good approaches to estimate travel time. When the travel time result is more important, the existing researches has provided many advanced and accurate methods. On the other hand, in the case that travel time is changed along with drivers' choice and the change of traffic volume, the estimation of travel time can be simplified as long as it is meaningful.

1.1.3 Bottleneck in a Network

A bottleneck in traffic indicates a local area that limits the traffic flow on the link or on the whole network. It is often where congestion starts or congested the most seriously. The bottleneck, sometimes, can be caused by a physical design defect, such as a sharply narrowing down as Fig. 2 shows. The capacity of this road is limited by the narrower segment, so the vehicles over the capacity will be congested in the bottleneck. This is a kind of stationary bottleneck that can be found easily.

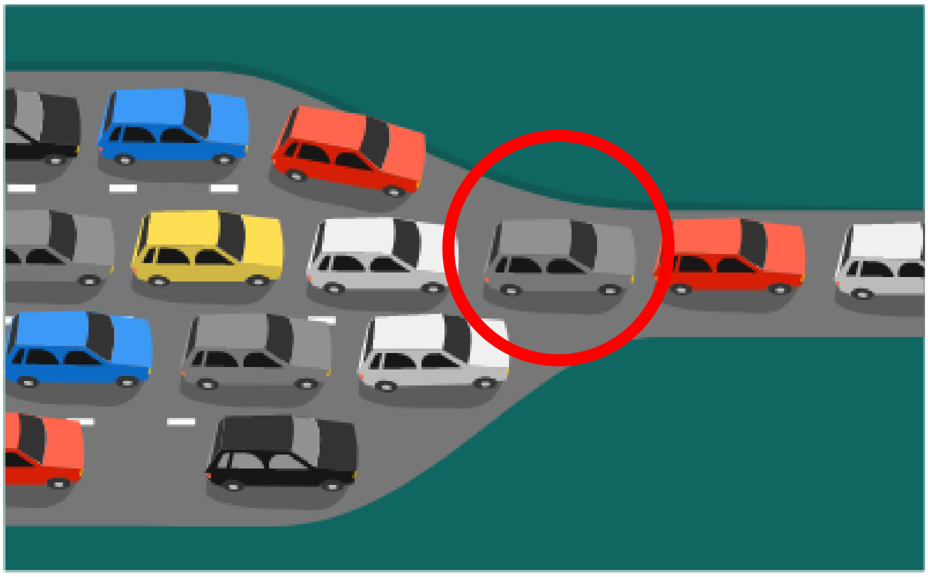

Fig. 2 A physical bottleneck in traffic flow

Then scholars hope to find the moving bottleneck from traffic flow. Newell (1998) presented his approach to detect the bottleneck in traffic flow by analyzing travel time, speed and density. And then Daganzo, Carlos and Jorge (2005) utilized a kinematic wave model to analyze the moving bottleneck. In these studies, the bottleneck is defined as the segments where vehicles pass at a lower speed. And the bottleneck point is reflected by traffic density and speed. These studies effectively gave definition and detection of the bottleneck in traffic flow. But as the real road traffic network is very complex, these models are limited in faraway single road or freeways. For traffic managers in urban city, the detection of bottleneck in the road network is still a significant task, so that then can improve traffic efficiency more directly.

However, when it comes to the whole network, it is hard to define and detect such a bottleneck in traffic flow. As an example, Long et. al (2008) gave their standard to detect traffic bottleneck according to link average journey velocity. They defined the bottleneck as the links with low average velocity and calculate the velocity through a CTM approach. An example network with 20 nodes and 60 links are built to explain the approach. Table 2 gave their standard for congestion.

Table 2 Long et. al standard to detect bottleneck

Condition	Velocity
Freely	$v \geq 30\ km/h$
Light Congestion	$20\ km/h \leq v \leq 30\ km/h$
Congestion	$10\ km/h \leq v \leq 20\ km/h$
Serious Congestion	$v \leq 10\ km/h$

In Long's study, traffic bottlenecks are detected by link velocity and it is found that the bottleneck is related to road construction and traffic demand. Indeed, Long's study is just an application of the CTM traffic flow theory. Until now, there are many technologies to estimate velocity and to count traffic volume real-time, there is no necessary to estimate a historical velocity through CTM. ITS can even monitor and visualize the vector of links real-time for a network. Under this background, it requires more understanding about the bottleneck where the travel time or velocity are more easy to change with the stochatic traffic demand may be adjusted by some traffic policies or strategies.

1.1.4 Data Source

As for the data, travel time and traffic volume data in urban road network is necessary for this study. According to the past researches in the traffic field, here are usually three kinds of data used in traffic problems: count data, CDR data and GPS data.

The first one, count data, is usually recorded by camera and sensors. In the case of camera, the raw data will be videos, which are a little bit difficult to process. A more common kind of count data is collected by single or double loop detectors. It can record passing vehicles and the time it passes the recorder. Thus, it can record all the passing vehicles and get the real traffic demand in the recorder position. However, it is limited spatially by the location of sensors. So, it is more common in freeway problems (e.g. Muñoz, Sun, Horowitz & Alvarez, 2003).

The second kind of data is CDR data, which is common in recent studies. CDR data, which means call details records, are recorded by cell phones when the user makes a call or other telecommunications transaction. The time of the call is recorded by the cell phone and the position relies on the location of the nearest Base Transceiver Station (BTS). Obviously, this kind of data bases on the telecommunications transaction so it can not collect all the traffic demand but a sampled one. In addition, as the recording frequent is limited, it is more suitable to an O/D problem, in which just origin point and destination point are the target (e.g. Iqbal et. al, 2014). Another feature of CDR data is that the position is not accurate information but the location of the BTS. If it is applied in a study of road traffic field, it must be processed to map-match into road network (e.g. Çolak et. Al, 2015). So it is more common to be used for the human mobility analysis.

The final one is GPS (Global Positioning System) data, similarly recorded by cell phones but at a certain frequent. GPS is a satellite-based navigation system owned by the United States government which records the geolocation and time information. Nowadays, there are many mature methods to transform the raw GPS data into latitude and longitude information, or even map-match it into the road network. Although it is also sampled by cell phone use and Internet connection, GPS data is still the best choice for most of the studies in traffic road fields due to the low cost, high resolution and wide range.

In summary, among the three kind of data, count data can collect the real traffic volume but limited in spatial. So it is more suitable for a freeway problem. The CDR data can be utilized for human mobility analysis or traffic analysis. But for the latter one, the position information should be processed to match into road network. Finally the GPS data is the most common data utilized in the studies on traffic road network. Even though the GPS data is sampled by cell phone use and Internet connection, for which we can just get a sampled traffic volume in stead of the real one, it is the best choice for most traffic related researches as it can cover a large area with a low cost.

1.2 Research Objective

In this master thesis, I focus on the travel time sensitivity analysis from historical trajectory data and hope to find some key roads, named bottleneck in this thesis, in the network, where the increase of traffic volume is easier to cause congestion in the whole network.

The first goal is to find the relationship between travel time and traffic volume on each individual link through the analysis of trajectory data. According to the relationship, the travel time under different traffic volume will be estimated quickly. This part of result is supposed to support other researches which need travel time estimation result under different traffic demand, like route guidance problem and traffic tolling problem.

Second, according to the relationship between link travel time and link traffic volume, I want to find a sensitive measurement of the whole network. The sensitive measurement can reflect when traffic volume increases in one or several certain links, whether the whole network will change seriously or not. By this sensitive measurement, we can conclude how much the whole network is easy to be congested and slow down.

The final goal is to detect some “bottleneck” links through the sensitive analysis of the network. These links are more sensitive to traffic volume increasing and when traffic volume increase a little, the whole network condition will be influenced more seriously. By detecting these bottlenecks, it can tell some key roads in the complex urban road network for traffic managers to prevent congestion in a certain area or give priority to these roads when they take some improve strategies.

1.3 Thesis Outline and Contributions

This section briefly introduces the structure of this thesis. I will summarize each subsequent chapter by its target, methodologies, expected results and contributions.

In Chapter 2, "Travel Time Estimation", aims to find out the relationship between travel time and traffic volume on each individual link. First, the data used in this thesis is introduced and summarized. Then, the GPS trajectory data is processed to get travel time and traffic volume for each half an hour. A linear matrix equation approach is used to calculate the mean travel time in half an hour for each link. And then, a linear regression approach is used to find the travel time function of traffic volume. This part of results can reflect the trend of travel time when traffic volume changes, which is meaningful in similar cases when the change of travel time is important. For example, it can estimate the time influence by user equilibrium in a route choice model.

Chapter 3, with the title "Sensitive Analysis", aims to give a measurement to evaluate how much the whole network is sensitive to some small change of traffic volume on certain road segments at a certain time point, given route information and traffic volume information in recent half an hour. In this part, the travel time function learned from chapter 2 is seemed as the attributes of each link and the estimated parameters are used to construct a linear matrix according to the route information. And then I analyze the matrix by introducing a mathematical conception, condition number, which reflect how much stable a linear matrix solution is. By this mathematical analysis of the matrix, I can get the index of the network to tell how much the network is easy to be congested or slow down. From the analysis, it is found that during daytime, travel time sensitivity has positive correlation with traffic volume and traffic volume difference. In addition, there are negative correlation between travel time sensitivity and mean speed in the network. According to the results, it is concluded that this sensitive measurement is effective to evaluate how much the network gets easy to be slow down.

Chapter 4, with the title "Bottleneck Detection", defines the bottleneck road segments in the road traffic network as the ones which contributes more to the sensitive measurement. When traffic volume increases on these bottlenecks, the whole network is more sensitive and easier to be congested. In this part, a Monte Carlo method is applied to simulate small increase of traffic volume on different road segments. And then, it produces new matrixes, based on which I do the matrix analysis again and get the sensitive measurement of each case. By this way, those which contributes more to the sensitive measurement are seemed as the bottleneck of the network at that time point. This part does a new define to bottleneck in the traffic network instead of a bottleneck point in traditional traffic flow theory. As the sensitive measurement is related to real-time traffic volume and route information, it can be a method to find the key roads detect where are more easily congested for the department of traffic.

Chapter 2 Travel Time Estimation

In order to analyze the network travel time sensitivity, the first step to estimate the relationship between travel time and traffic volume on each individual link. As mentioned in the introduction part, there are many approaches presented in the past studies that can estimate travel time very well, such as traffic flow theory, traffic simulation tools and intelligent algorithm like deep learning. While these approaches focus on the accuracy of the estimation result, it cannot reflect the relationship between travel time and other traffic parameters. So when the target is to find out the travel time trend, the estimation procedure can be simplified to be reasonable.

In this part, I will first introduce the data I used in the thesis and give a simple data description. Then I will explain the methodology to process the raw data and get the travel time and traffic volume information. According to the information, I build linear regression model for each link to estimate the travel time trend to the traffic volume.

2.1 Data Description

Road Network Data. The road network data is Digital Road Map Database extended version 2016 (DRM), provided by CSIS (Center for Spatial Information Science) of the University of Tokyo and the data is organized by the Sumitomo Electronic Group. The database includes the whole road network in Japan and provides their location, connection relationship, length and class information.

Travel Data. Travel information is from an interpolation GPS data, provided and processed also by CSIS, the University of Tokyo. The trajectories have been map-matched into the road network and the trajectory points are interpolated into each road intersection.

This interpolation GPS data includes user ID, trip ID, travel mode, latitude, longitude, time, link ID and mesh ID in Tokyo for one month. The fields in GPS data are explained in Table 3.

Table 3 Fields in GPS data

Uid	User id of the record
Subid	Trip id for the user
Pointid	The pointe order in the trip
Mode	Estimated travel mode, including car, bike, walk and stay
Time	Timestamp of the recorder
Latitude	The position information of latitude of the point
Longitude	The position information of longitude of the point
Node_id	The interpolated node id in DRM
Link_id	The interpolated point locates on which link end
Mesh_id	The interpolated point locates on which mesh

The study area in the thesis is a 3-level mesh near Ikebukuro Station in Tokyo, Japan, as the Fig. 2 shows. In Japan, the map is divided into small grids by the Standard Grid Square system and the grids are coded by Standard Grid Square Code. A 3-level mesh is a rectangle area of which the length of side is near to one kilometer. In this area, there are 555 roads and 365 intersections, which match 555 links and 365 nodes in the network. Ikebukuro is one of the business centers in Tokyo and traffic is busy near Ikebukuro Station. In the data record, there are over 137 thousand trips during the month.

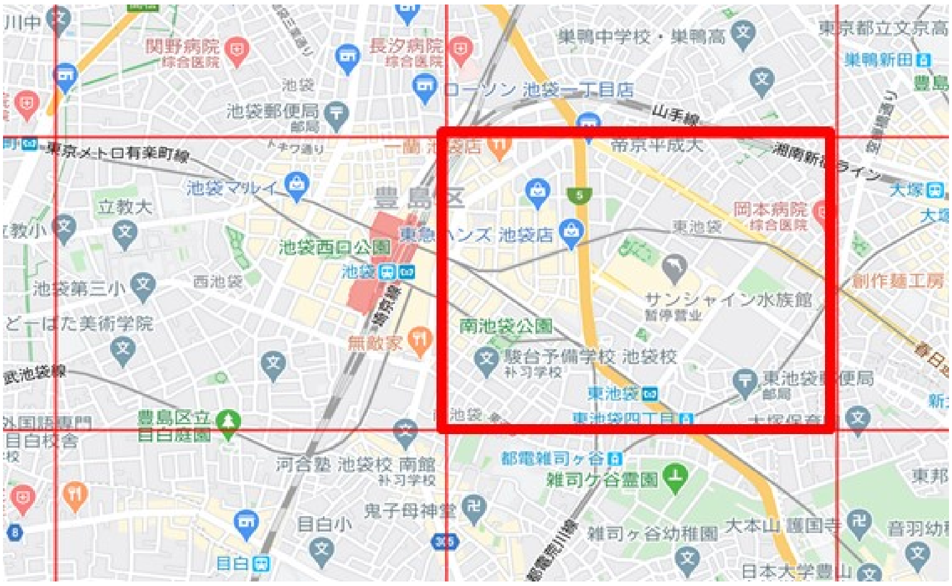

Fig. 3 The study area mesh near Ikebukuro Station

The general traffic condition in this area is as shown in Fig. 3 and Fig. 4 . In one day, there are about 5,000 trips in the network and Fig. 3 draws the traffic volume tendency in each hour. In Fig. 2.1.2b, it draws the tendency of mean speed in each hour instead of travel time. Although travel time is the object of this study, it is hard to summary and visualize to reflect traffic condition because it differs according to the route length. On the contrary, mean speed can replace travel time to show how much the traffic network is congested for data description. For each trip, a lower speed will match a longer travel time.

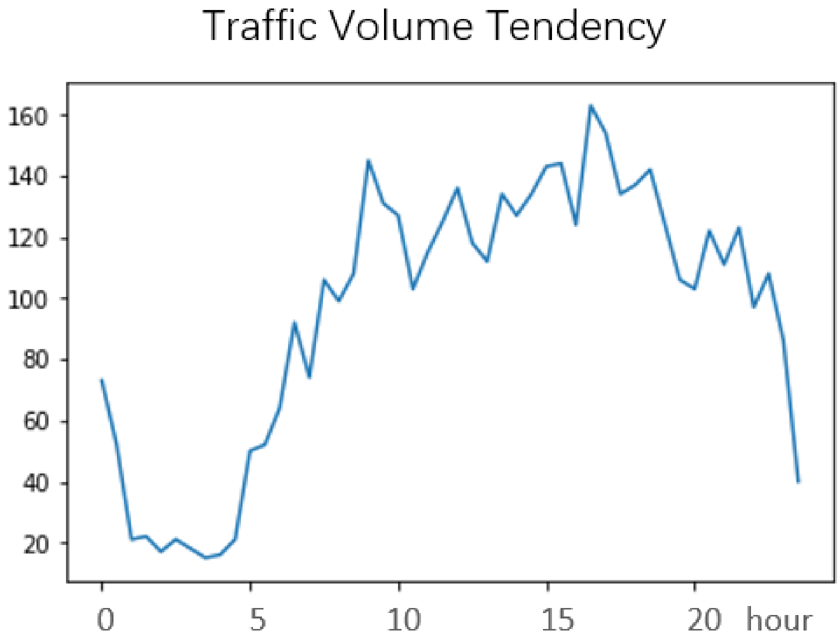

Fig. 4 Traffic volume tendency during one day

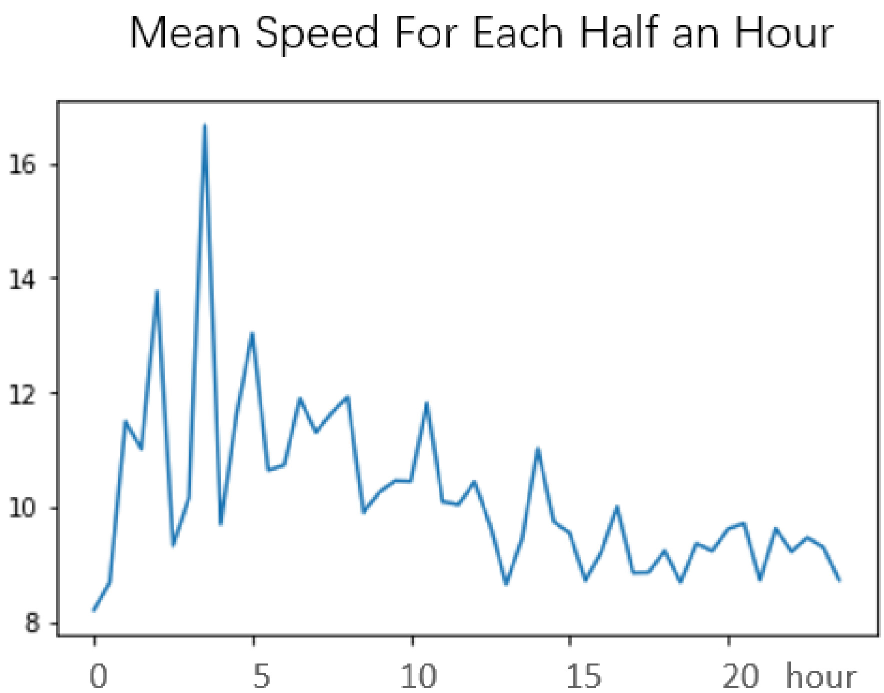

Fig. 5 Average speed tendency during one day in the whole network

From the tendency line, the traffic volume keeps a low level during night time, from 0 am to 6 am. At 6am, when some people have to go to work or school, traffic volume gets increasing until 9 am, when morning peak hour ends. During the day time from 8 am to 18 pm, the traffic volume keeps at a high level. After 8 pm, there is another evening peak and lasting to around 10 pm, when those who finish work or have fun outside are going to go home.

And from Fig. 5, the mean travel speed can also tell some information about the traffic condition in this area. During the night time, when traffic volume is low, the travel speed is higher than daytime. And as the traffic increase from morning peak, the mean speed decreases gradually. In special, at 10 am, when traffic volume gets a trough level, travel speed arrives a peak. After that, as the traffic volume keeps a high level, travel speed continues to reduce until 6 pm. From 6 pm to 8 pm, travel speed rises again while traffic volume reduces. And then, along with the evening peak at 8pm, travel speed slow down again.

Combining the two lines of traffic volume and travel speed, it is no doubt that travel speed is significantly influenced by traffic volume. When traffic volume is lower, travel speed is relatively higher. And when traffic volume increases and keeps at a high level, travel speed will slow down. From the value of travel speed and traffic volume, I define a peak hour in this area, from 5 pm to 7 pm, with both high traffic volume and low travel speed.

During this peak time, I count the trip number of each link and draw the distribution in Fig. 6. The horizontal axis means the trip number and the vertical axis indicates the total link number with this demand. From the histogram, most of roads (over 300 roads in total 555) are hardly used, where the trip number is smaller than 25. With the trip number between 25 and 100, these roads can be considered as medium-frequently used in the network while the other ones are most busy, with over 100 trips during the peak hour.

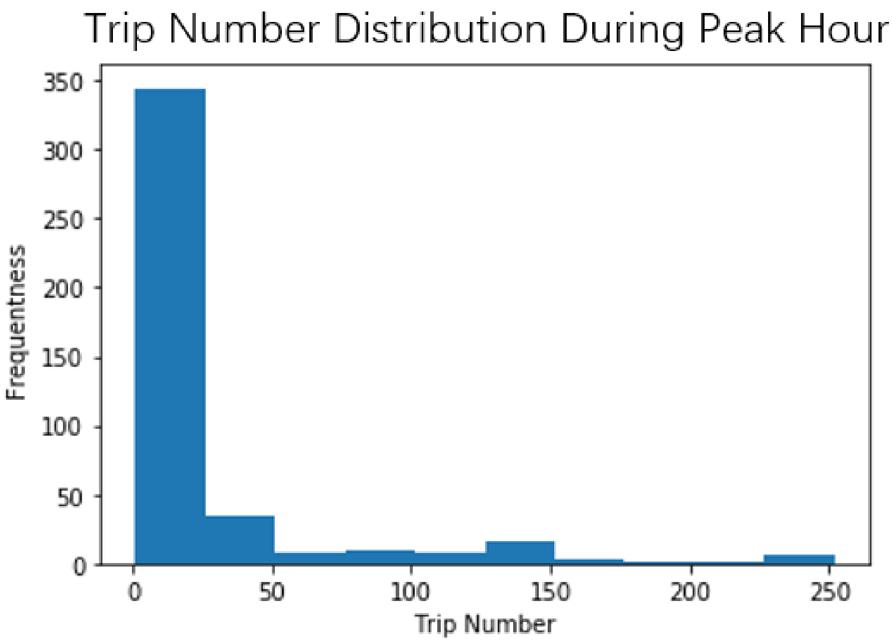

Fig. 6 Trip Number Distribution

In order to inspect the traffic condition of these links, then I draw a scatter diagram to observe whether the mean travel speed of links is related to its trip number. As shown in Fig. 7, each point indicates one link. The horizontal axis indicates the trip number of links during peak hour and the vertical axis indicates the mean speed. I divided the links into three group as the red circles show. The first group, circle 1, are the links where the trip number is high and the travel speed is medium compared to other groups. The second group, circle 2, are the links where the trip number is medium but travel speed is high. The third group, circle 3, indicates those links which are hardly used, but the travel speed is also low.

Combined to the Fig. 6, most of the links are hardly used. Meanwhile, when a vehicle passes these less-used links, the travel speed may be low. On the other hand, among the frequently used links, which take up around one-third of the total links, the more the trip number is, the lower the travel speed may be. But in the same busy group (circle 3), the travel speed is quite stable. As the trip number increases, the travel speed doesn't change at all, which indicates the congestion in this area is not seriously.

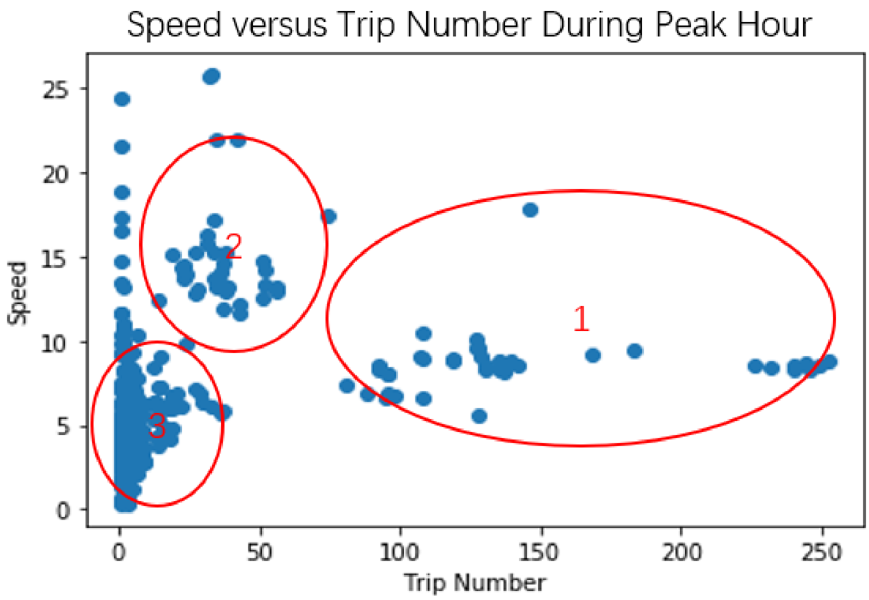

Fig. 7 Mean Speed of Links Versus Trip Number

After the basic understand of the traffic network, the next step is to find a suitable time interval to count traffic volume and calculate the average travel time for each link. The time interval should not be too short, otherwise the trip number will be too little for regression. But it also should not be too long in view of real-time application. Fig. 8 is an example link, which is a busy road and I draw the mean travel time value versus trip number by 4 different time intervals. It is found that when the time interval is very short (5 min and 10 min), the trip number is too little and it is difficult to conclude the travel time tendency. When the time interval is longer (20 min and 30 min), trip number is more and it is more significant that travel time increases as the traffic volume increases on the same link. Considering that it is a busy link, for most links in the work, a 30 minutes interval is better to find out the regression function of travel time.

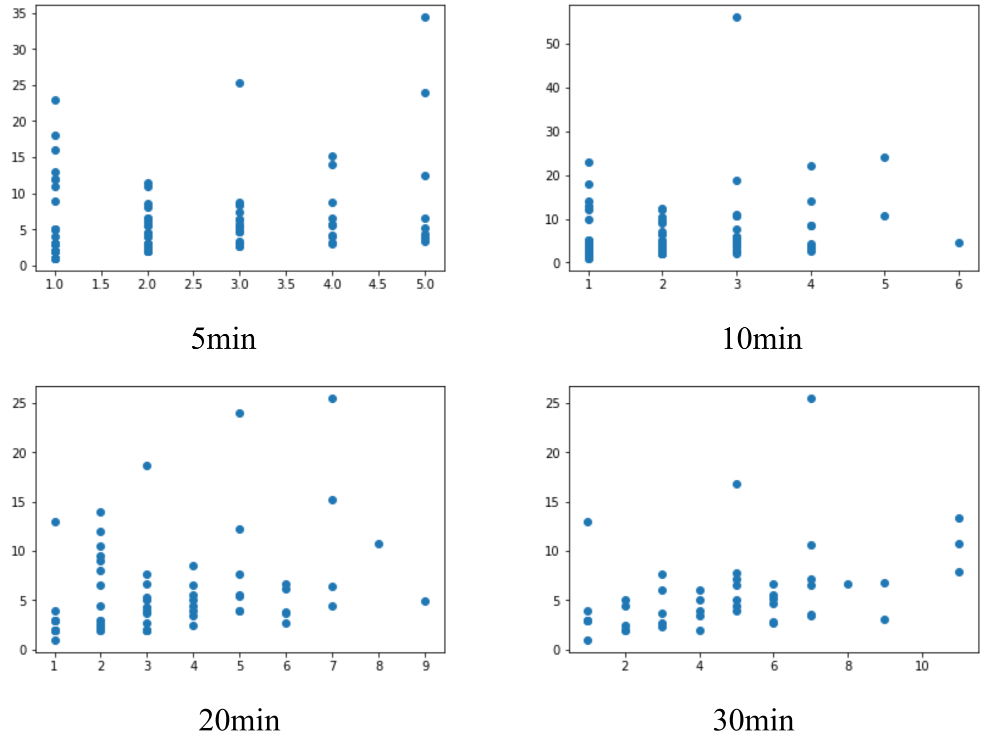

Fig. 8 Mean Travel Time versus Trip number by different time intervals

2.2 Linear Matrix Solution

After deciding 30 minutes as the time scale, I count traffic volume for each half an hour and use a linear matrix to solve the mean travel time according to the interpolation data. In this section, I explain the method by which I get the travel time for each link.

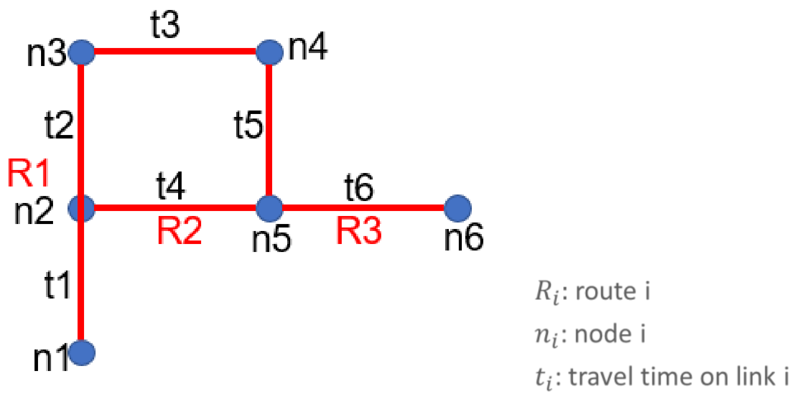

Fig. 9 An example network

In Fig. 9, I draw a small network as the example to explain how to construct the linear matrix. In this network, there are 6 nodes and 6 links. Assume that the mean travel time on link i is t_i and the total travel time for route i is $T(R_i)$. If we collect three route information in this work:

Route 1: driving through link1, link 2 and link3;

Route 2: driving through link1, link 4 and link 6;

Route 3: driving through link 5 and link6;

Then I can write three equations, with the assumption that the waiting time in intersections is part of the travel time on links:

$$T(R_1) = t_1 + t_2 + t_3$$

$$T(R_2) = t_1 + t_4 + t_6$$

$$T(R_3) = t_5 + t_6$$

In the background of big data, the route number can be very large and I can write many such equations, which make it possible to give enough constraints to mean travel time on links. Compared to use the interpolation link travel time and naively calculate the mean value, the total route travel time is more accurate to get. By this way, the error made from interpolation process will be reduced. After getting the equations, they construct a linear matrix equation and the mean travel time of each link can be solved by a least square solution.

2.3 Travel Time Estimation

With the estimation result of mean travel time on each link, linear regression models are built for each link. There are some basic assumptions for the regression.

$$Travel\ Time = a \times traffic\ volume + b$$

Assumption 1: $a \geq 0$ for all the links. Travel time is supposed to be longer as traffic volume increases. The situation that travel time reduces when traffic volume is less is not reasonable, which will cause a mistake to evaluate and optimize a road network.

Assumption 2: When the point that used for linear regression is less than ten, a is equal to zero. The spark point for linear regression means the road is hardly used. Usually, these roads are and travel time on these links are little influenced by traffic volume. To avoid over-fitting, the gradient of linear regression model on these links are set to be zero.

Before the regression model is built, I calculate the correlation coefficient of the links which have enough points to do regression. Correlation coefficient is a mathematical parameter to measure if there are linear relationship between two variables. Theoretically, when the correlation coefficient is near to 0, there are nearly no linear relationship between the two variables. When the correlation coefficient is near to 1, it means the two variables are positively linear relative. On the contrary, when the correlation coefficient is near to -1, the two variables are negatively linear relative.

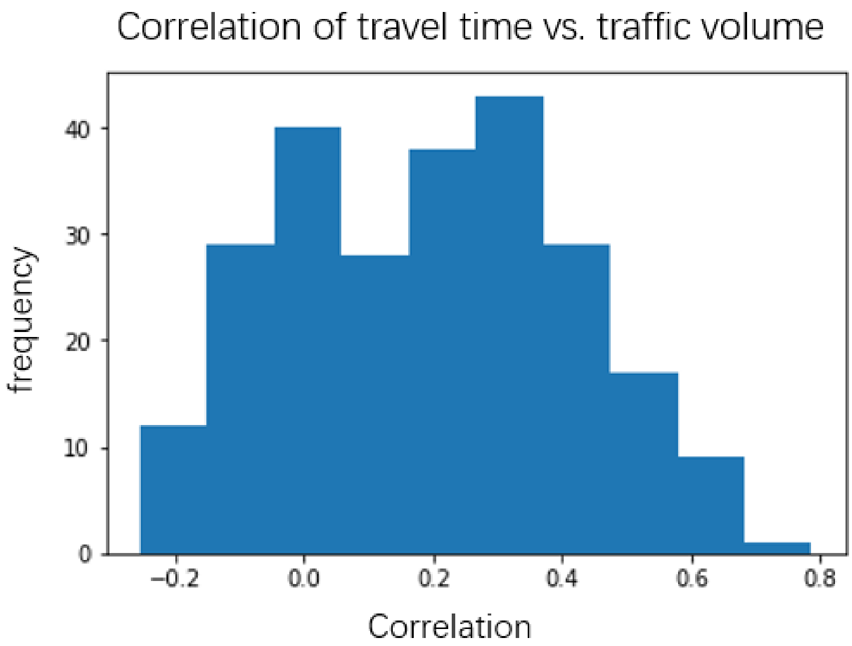

Fig. 10 The correlation coefficient of travel time versus traffic volume

Fig. 10 shows the correlation coefficient histogram of the links. Those with the correlation between -0.2 and 0.2 are considered as no linear relationship between traffic volume and travel time. On the links with the correlation between 0.2 and 0.4, there are weakly linear relationship between travel time and traffic volume (around 100 links). And on the links with the correlation bigger than 0.4, it is seemed that the travel time is strongly linear relative to traffic volume (around 50 links). Fig. 11 and Fig. 12 give two examples of the regression result. The link shown in Fig. 11 is one where travel time is weakly linear relative to traffic volume. The fitting result shows there is indeed an increasing tendency of travel time as traffic volume increases, but the points are scattered. The link in Fig. 12 is a link where travel time is strongly relative to traffic volume. The points are near to the regression line.

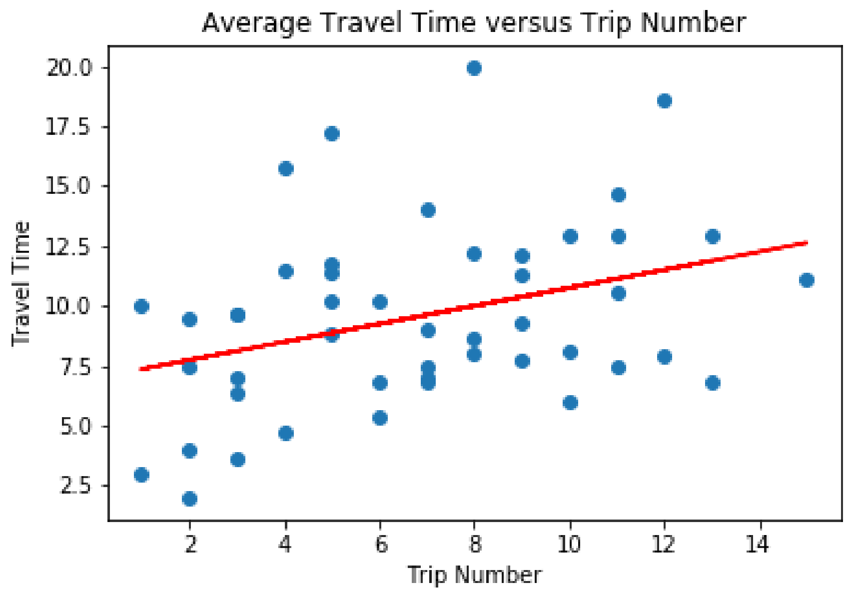

Fig. 11 A weakly relative link example

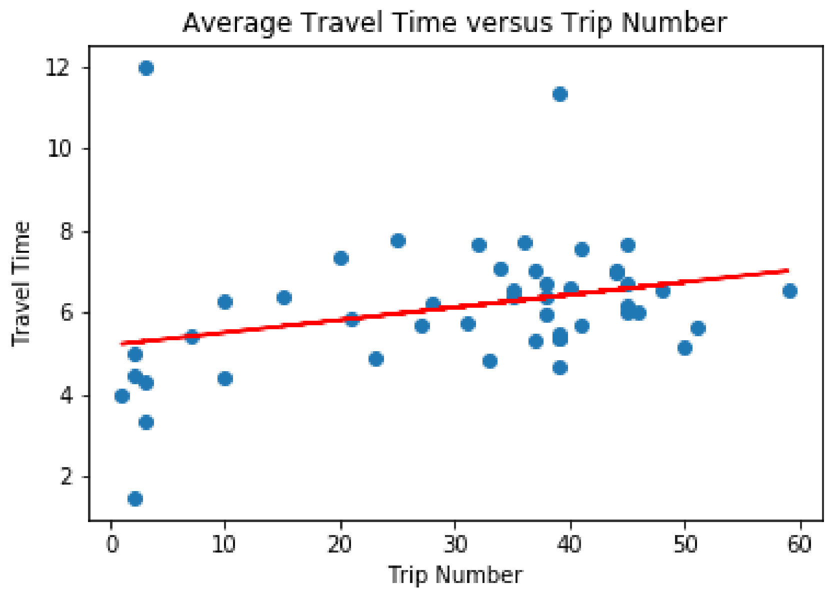

Fig. 12 A strongly relative link example

In a summary of this part, mean travel time is solved by a linear matrix equation, and then the relationship between travel time and traffic volume on each link is estimated by linear regression models. In around one-third of the total links, the travel time are significantly influenced by traffic volume. In other two-thirds of the links, the travel time is more stable. The gradient of the linear function for these links are zero and the travel time estimation result is just the mean travel time.

Chapter 3 Sensitive Analysis

Given the relationship between travel time and traffic volume of each link, in this part, I want to find out a measurement to evaluate the relationship between travel time and traffic volume in the whole network. I explain how I construct a linear matrix according to the linear functions in Chapter 2 and get the measurement. Finally, I give a simple analysis of this measurement.

3.1 Matrix Construction

Considering the same example in Chapter 2, showed by Fig. 2.2.1, I have got the Equations 1. In these equations, the link travel time can be replaced by the linear function estimated in Chapter 2 and equations are transformed into Equation 2. By reorganizing Equation 2, I get Equation 3, which can be abbreviated to linear matrix Equation 4.

$$\begin{cases} T(R_1) = t_1 + t_2 + t_3 \\ T(R_2) = t_1 + t_4 + t_6 \\ T(R_3) = t_5 + t_6 \end{cases} \quad \ldots\ Equation\ 1$$

$$\begin{cases} T(R_1) = (a_1x_1 + b_1) + (a_2x_2 + b_2) + (a_3x_3 + b_3) \\ T(R_2) = (a_1x_1 + b_1) + (a_4x_4 + b_4) + (a_6x_6 + b_6) \\ T(R_3) = (a_5x_5 + b_5) + (a_6x_6 + b_6) \end{cases} \quad \ldots\ Equation\ 2$$

x_i indicates the traffic volume of link *i*.

$$\begin{cases} T(R_1) = a_1x_1 + a_2x_2 + a_3x_3 + 0x_4 + 0x_5 + 0x_6 + (b_1 + b_2 + b_3) \\ T(R_2) = a_1x_1 + 0x_2 + 0x_3 + a_4x_4 + 0x_5 + a_6x_6 + (b_1 + b_4 + b_6) \\ \quad T(R_3) = 0x_1 + 0x_2 + 0x_3 + 0x_4 + a_5x_5 + a_6x_6 + (b_5 + b_6) \end{cases} \ \ldots \ Equation\ 3$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1(n+1)} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{m(n+1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots\cdots \\ x_n \\ 1 \end{bmatrix} = \begin{bmatrix} T(R_1) \\ T(R_2) \\ \cdots\cdots \\ T(R_m) \end{bmatrix} \quad \ldots Euqation\ 4$$

For Equation 4, it can be written like $Ax = T$. Matrix A indicates the relationship matrix between travel time and traffic volume, with m rows and (n+1) columns. The elements of the matrix consist of the estimated parameters of the linear regression model in Chapter 2. Vector x indicates the traffic volume on each link plus a constant term. And the vector T indicates the travel time for each route.

By this reorganization process, the relationship between travel time and traffic volume has been expended into the whole network and represented as a matrix A. When traffic volume changes, the multiply result, which is the route travel time vector, will also change. In this study, I focus on how travel time will be changed according to the change of traffic volume, which is just the sensitivity of the travel time. By mathematical analysis introduced in the following section, I provide the measurement according to the analysis of matrix A.

3.2 Condition Number Measurement

For a linear matrix equation like $Ax = T$, vector x and vector T are connected by the matrix A. If vector x or T is given, we can calculate another one from the equation. Thus, the relationship between vector x and vector T can be reflected by the matrix A. However, other than the value of vector x or T, many studies also care about sensitivity how much vector T will change when vector x changes a little, which can also be reflected by Matrix A.

In the field of mathematical, a concept of condition number is usually used to measure how sensitive the output is to the small change of input. The idea of condition number was first invented by Turing (1948) and credited by Wilkinson (1963). The definition of condition number for matrix A in linear equation $Ax = T$ is given as Equation 5. $\kappa(A)$ indicates the condition number of matrix A and $||\cdot||$ indicates the norm of a matrix. Cleve Moler (2017) explained why condition number can measure the sensitivity of matrix A.

$$\kappa(A) = ||A||||A^{-1}|| \quad \dots \; Equation\ 5$$

Considering the equation $Ax = T$, with a small change δx of x, the equation will be:

$$A(x + \delta x) = T + \delta T \quad \dots \quad Equation\ 6$$

As δx is a very small change, we can assume that:

$$A\delta x = \delta T \quad \dots \quad Equation\ 7$$

From the definition of the norm for matrix A, which reflects the stretching and shrinking ability, two variables M and m are introduced as follows (in these equations, x indicates any vectors):

$$M = max\frac{||Ax||}{||x||} = ||A|| \quad \dots \quad Equation\ 8$$

$$m = min\frac{||Ax||}{||x||} = min\frac{||T||}{||A^{-1}T||} = max\frac{||A^{-1}T||}{||T||} = \frac{1}{||A^{-1}||} \quad \dots \quad Equation\ 9$$

Then, we can get Equation 10 by combing Equation 7 and 8, and get Equation 11 by combing Equation 7 and 9.

$$\|b\| \leq M\|x\| \quad \dots \quad Equation\ 10$$

$$\|\delta T\| \geq m\|\delta x\| \quad \dots \quad Equation\ 11$$

Finally by combining Equation 10 and 11, it gets:

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A)\frac{\|\delta T\|}{\|T\|} \quad \dots \quad Equation\ 12$$

Equation 12 tells that when a small change happens on x, the change rate is limited by condition number of transformation matrix A. The bigger condition number is, the more sensitive it means in this situation.

In past researches, condition number is commonly used to evaluate the sensitivity or robustness of a system (e.g. Christiansen & Hansen, 1994; Skogestad & Havre, 1996; Kenney et. al, 2003). It is a good method to evaluate the sensitivity via condition number between two groups of data. With the transformation matrix, the change of several variables can be estimated in one parameter, by which the calculation is fast and easy in implement.

3.3 Results and Analysis

In this section, I give the results of condition number and compare it with other traffic parameters including traffic volume, traffic volume differences and mean travel speed.

3.3.1 Condition Number Tendency

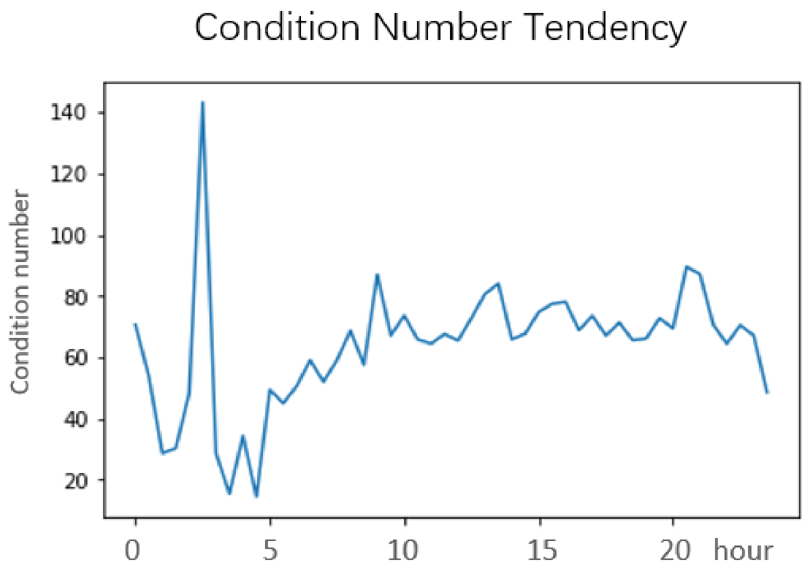

Fig. 13 Condition Number Tendency of the whole Network During One Day

Fig. 13 shows the condition number tendency of the whole network during one day. From the tendency line, the peak of the condition number is around 3 am during night. There are two possible reasons for the night peak. The first reason is on account of a large error. The traffic volume during night is lower than daytime and as GPS data is a sampled data, the collecting records are spark, for which the error may be large. The second reason is maybe the sensitive does change a lot at that time. As there are little trips during night, every vehicle is supposed to pass the links freely. If there is any slow-down behavior, it will influence the statistic result seriously.

Then, it turns to the daytime. Around 5 am the condition number keeps the lowest value in one day and after 5 am, the sensitivity keeps increasing until 10 am. During 9 am and 10 pm, when it is called daytime, the condition number line generally keeps a high level. During daytime, there are three peaks, happening at 9:30, 14:00 and 21:00. After daytime, the condition number reduces to a low level.

3.3.2 Condition Number vs. Traffic Volume

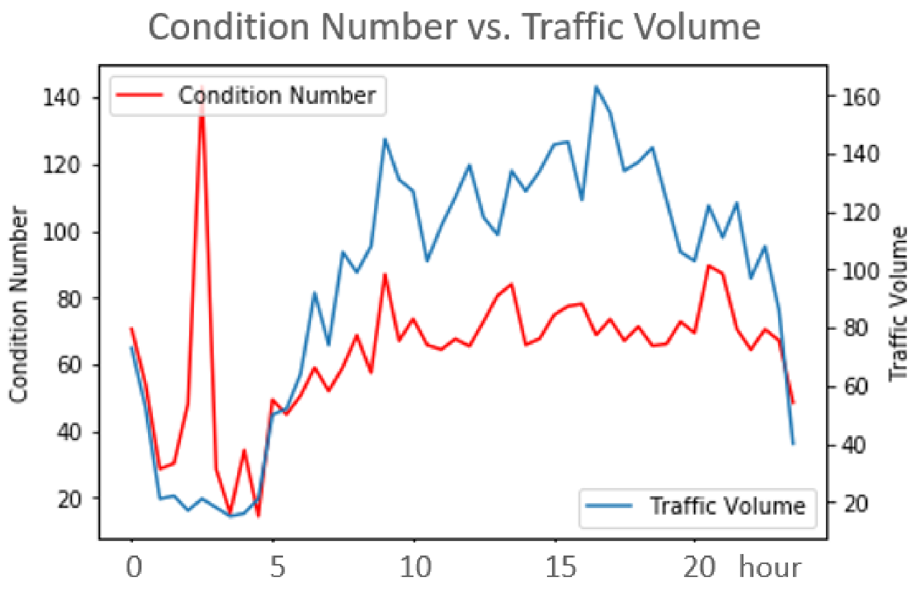

Fig. 14 Condition Number Tendency versus Traffic volume tendency of the whole Network During One Day

Fig. 14 compares the tendency of condition number and trip volume in the whole network. The horizonal axis indicates the time period From this picture, I hope to find out how traffic volume will influence the condition number. First, at the condition number peak during night, the traffic volume hardly changes and keeps at a very low level, about 20 trips for each half an hour. During the daytime, the tendency of condition number and traffic volume are similar in some time period. For exaple, after 5 am when traffic volume increases, the condition number also increases at the same time and when traffic volume reduce between 9 am and 10 am, the condition number also dereases. The coordinated change also happens during 14:00 ~ 15:00 and 20:00 ~ 02:00. However, at the peak hour starts at 17 o'clock, when traffic volume increases rapidly, the condition number is not significantly changed by this demand increase and it almost keep the same value during the peak hour between 15 o'clock from 17 o'clock.

3.3.3 Condition Number vs. Traffic Volume Difference

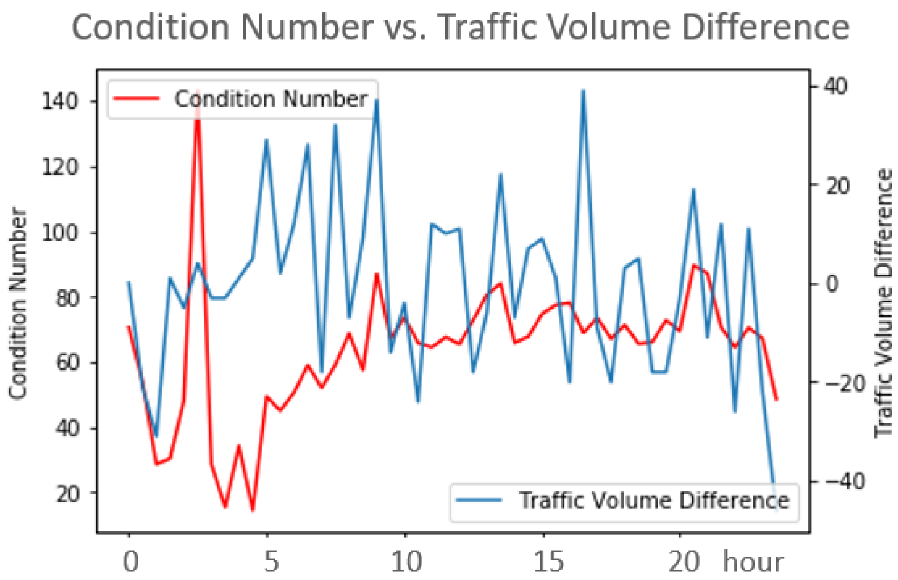

Fig. 15 Condition Number Tendency versus Traffic volume difference tendency of the whole Network During One Day

According to the analysis in section 3.3.2 for the Fig. 14, I hope to furtherly observe how condition number changes according to the traffic volume variation, which is the traffic volume difference compared to the last half an hour. Generally, the change between condition number and traffic volume difference is coordinated. For the three condition number peaks during daytime, at 9 o'clock, 13 o'clock and 20 o'clock, the traffic variation is also at a high level. The rapid decrease of demand difference also brings a decrease to condition number, such as 9:30. Another regulation is that, during 5 am and 10 am, when traffic volume almost keep increasing, as the variation of traffic volume difference doesn't keep increasing, the condition number changes along with the tendency of traffic volume variation. In addition, a difference larger than 20 or smaller than -20 is easier to be followed by a significant of network condition number.

In special, at the traffic volume peak at 16:00, when traffic volume increased a lot, the condition number is not high. It tells that the condition number is not only influenced by the total traffic demand of the whole network, but also the distribution of the traffic volume on each link.

3.3.4 Condition Number vs. Mean Speed

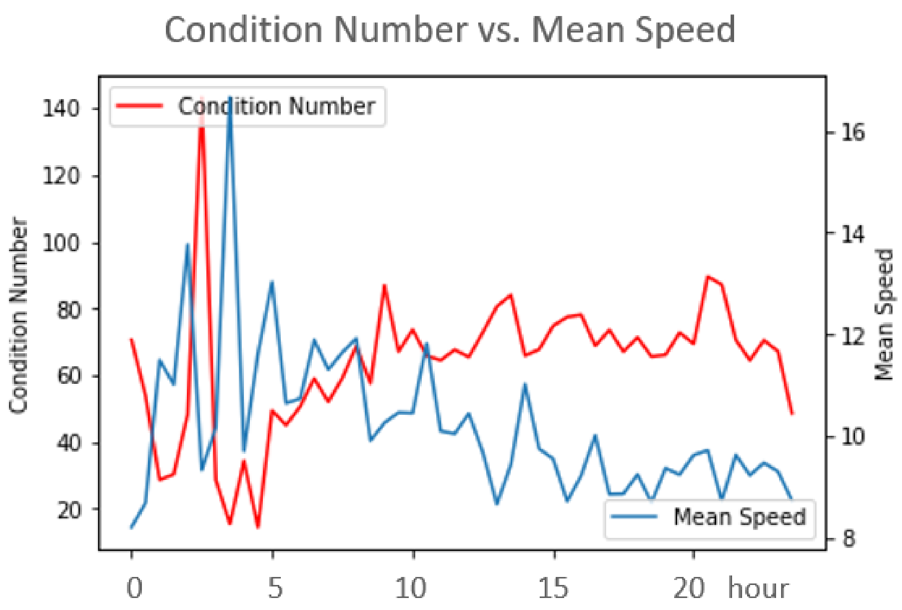

Fig. 16 Condition Number Tendency versus mean travel speed of the whole Network During One Day

Another traffic variables I care about is mean travel speed, which is very direct to reflect traffic condition. A high speed means a shorter time to finish the same route and when speed decreases, the traffic gets congested. From Fig. 16, there is a sudden and rapid reduce of travel speed during night time around 3 o'clock. Correspondingly, it is the time when condition number increases rapidly. And during the daytime, it is a general regulation that the increases condition number is accompanies with the decrease of the travel speed. The three peaks of condition number match the troughs of travel speed. On the other hand, the peak of average speed at around 10:30, 14:00 and 16:00 also match a low-level condition number.

Then it turns to the traffic volume peak at 16:00. According to the traffic volume and condition number, the sensitive doesn't change a lot along with the rapid increase of traffic volume. From the speed tendency, similar with the condition number, it doesn't decrease a lot. It is also an evidence that condition number can estimate the change of travel time.

3.3.5 Summary

Table 4 Covariance between condition number and other variables

	Traffic volume	**Traffic volume Variation**	**Mean speed**
Covariance between Condition number &	509.75	54.66	-16.88

As a summary of the analysis between condition number and other traffic variables, Table 4 shows a numerical analysis between condition number and traffic volume, traffic volume variation and mean speed. The covariance parameter estimates whether there are positive or negative relationship between two variables. When the value of covariance is positive, it means the two variables are positively related. In this case, the greater values of one variable mainly correspond with the greater values of the other variable. On the contrary, when the value is covariance is negative, the two variables are negatively related. This relationship is not limited to a linear relationship but a tendency similarity. In special, when the value of the covariance is equal to zero, it is seemed that the two variables are not related.

So, first according to the tendency and covariance between condition number and travel speed, condition number is negatively related to travel speed. It shows that, condition number is indeed a parameter which is able to reflect the traffic condition in the whole road network.

Second, from the result shown in Table 4, the covariance of condition number between both traffic volume and traffic volume variation are positive, which indicate the sensitivity is positively related to these two variables. So, it is concluded that the sensitivity of the road network is influenced by the traffic demand and traffic demand variation in the whole network. However, considering the exceptions when traffic volume is high and increases a lot but condition number is not increasing, it gives an evidence that condition number is also influenced by the traffic volume distribution but not only the total value.

In a summary, the sensitive measurement given by condition number can be an effective measurement to evaluate the travel time sensitivity to the change of traffic volume on each link. When the sensitive measurement is large, it means the traffic network is easier to slow down or to be congested.

Chapter 4 Bottleneck Detection

In chapter 3, it is found that the sensitivity is influenced by the distribution of traffic demand and traffic volume variation on each link. Then it becomes a question that for a certain situation, at which link the traffic volume increases will cause a larger condition number change. In this chapter, the bottleneck is defined as the road where traffic volume increases a little, the condition number will increase a lot. By this definition, the bottleneck will point out the road segment where traffic congestion starts or easily happens.

By a Monte Carlo method, the influence of each link to the condition number is estimated and those with the largest influence are detected as the bottlenecks. Detecting these bottlenecks can help traffic managers predicting and preventing traffic congestion in a local area.

4.1 Monte Carlo Methods

As the change of condition number for a matrix is difficult to calculate directly by mathematical methods, a Monte Carlo method is introduced to estimate the influence of each link to the whole network.

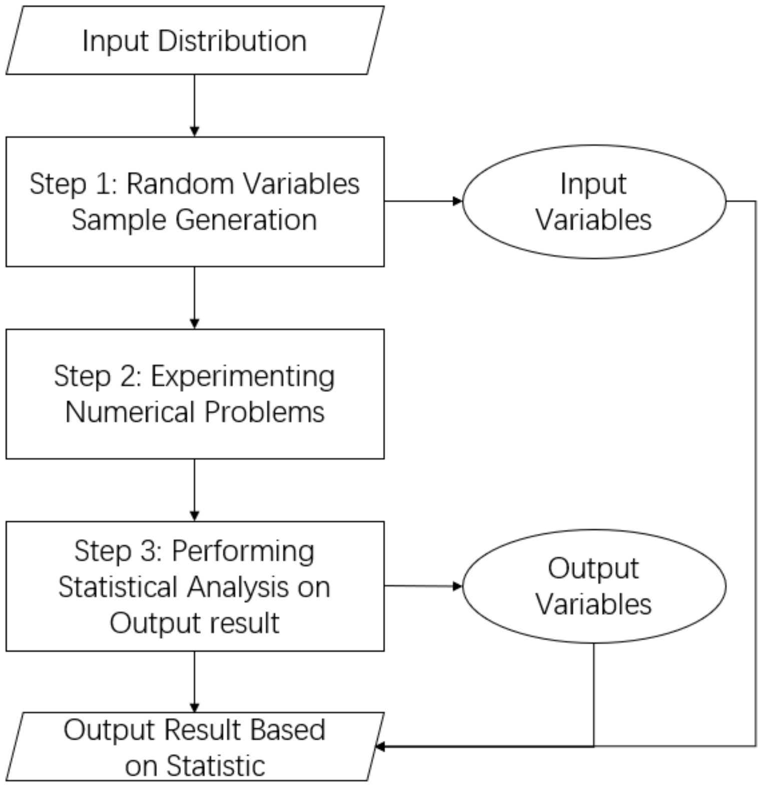

Fig. 17 The Procedure of Monte Carlo Method

Monte Carlo method is a random simulation method which gets the estimation result by repeating random sampling and calculating the average value. According to the law of large numbers, the experimental average value will be near to the expected value of a random variable if it is repeated many times. The procedure of Monte Carlo Method is as Fig. 17 gives. Before simulation, it must make sure the distribution of input variables. And then, a sampling data should be generated according to the distribution. This sampling dataset, as the series of input, should be experimented as the real system works and give a series of output. By statistical analysis of the input and output variables, the result will be given.

Monte Carlo Method can solve both probabilistic and deterministic problems (Hammersley, 2013). If the output is probabilistic, the Monte Carlo can give out the distribution of the output by many experiments. And when it comes to a deterministic output, the experiments can be designed to get the expected value. Fig. 18 show an example to calculate the value of π. Points are generated in the square with the width of side $2r$ with a uniform distribution. And then, the distance between each point and the center point will be calculated. If the distance is smaller than r, this point in the circle of radius r. Then, the ratio between the are of square and circle is equal to the ratio between the total points number and the points in the circle. So we have:

$$\pi = 4 \times \frac{\textit{the points number in the circle}}{\textit{the total point}}$$

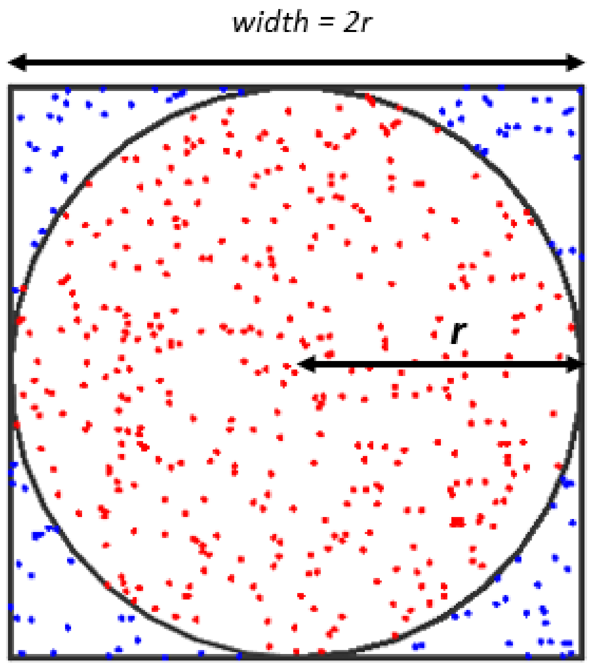

Fig. 18 An example to calculate pi value by Monte Carlo Method

Because it is easy in implement, Monte Carlo Method is usually to be applied to simulate systems or process with many coupled degrees of freedom. In this thesis, in order to analyze the influence of traffic volume change on each link, Monte Carlo is a good method to simulate with different input for such a traffic network problems. Then, it follows the process I simulate the change of traffic volume and estimate the change of condition number via Monte Carlo Method.

Step 1. Generating 1 ~ 5 trips from the real trip data.

Step 2. Adding the generating trips into the real data.

Step 3. Reorganizing the transformation matrix A and calculate the condition number of it.

Step 4. Repeat the procedure from step 1 to step 3 for 500 times

Step 5. The impact of the small change on one link is estimated as the total increase of condition number if this road is passed by the adding trips/ the number it is generated.

Especially, it is possible that some links are never passed by in the generating links. In this case, the influence of these links is ignored because they are assumed to be hardly used.

4.2 Results and Analysis

By a Monte Carlo method, the bottlenecks are detected and I compare the traffic volume and linear regression coefficient of the bottlenecks with the other links.

Table 5 Bottleneck Traffic Volume During Peak Hour

BOTTLENECK	17:00-17:30	17:30-18:00	18:00-18:30	18:30-19:00
1	12	11	12	9
2	8	12	9	8
3	7	8	6	7

Table 5 gives the traffic volume on the traffic volume on the bottleneck links for each half an hour during the peak hours from 17:00 to 19:00. When busiest roads have around 40 trips for each half an hour, the traffic volume on these bottlenecks are not as much as assumed.

Table 6 Bottleneck Linear Regression Coefficient During Peak Hour

BOTTLENECK	17:00-17:30	17:30-18:00	18:00-18:30	18:30-19:00
1	2.25	2.43	2.38	2.05
2	1.65	2.05	2.27	2.25
3	2.11	2.68	2.68	2.11

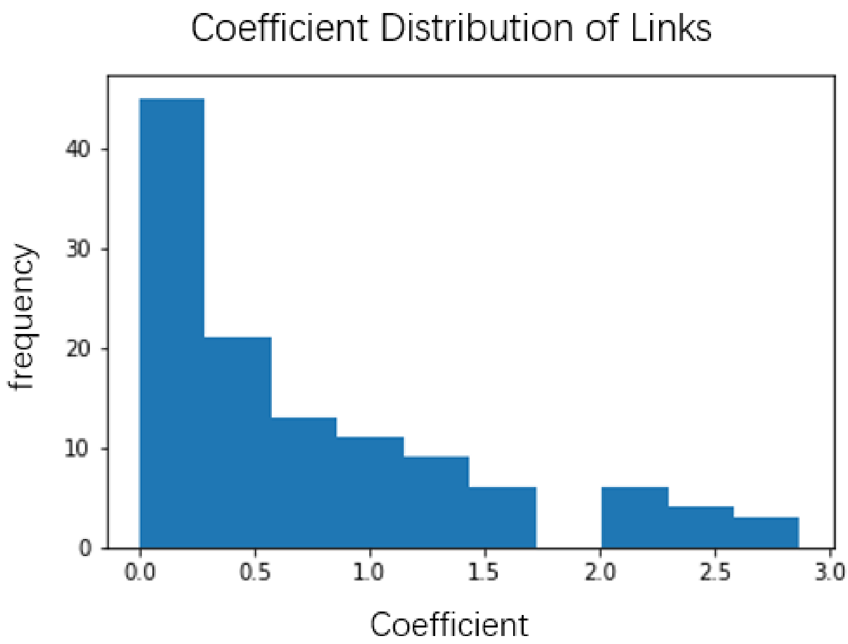

Fig. 19 Coefficient Distribution of the Links

As for the coefficient of the links estimated by the linear regression models, all of the bottlenecks have a high coefficient level compared to other links. It means that the travel time on this links are more sensitive to the traffic volume originally and then the influence are expanded to the whole network.

Chapter 5 Conclusion and Discussion

5.1 Summary

In this thesis, based on the analysis of the GPS data, a sensitive measurement of travel time is defined and estimated to evaluate how much travel time is sensitive to a small change of traffic volume. With a higher sensitive value, the road network is easily to be congested. The core idea of this method is to construct a transform matrix to connect two vectors, traffic volume on each link and travel time for each route. With the matrix, it is possible to transform the change between two vectors into one value, which can directly estimate traffic condition and lead traffic optimizations. On the other hand, according to the sensitive measurement, we can go back to the traffic volume vector to find out which roads impact the sensitivity the most, which are defined as the bottleneck of the network. By pointing out the bottlenecks, we can detect where is easier to be congested.

And from the analysis of the study area, it confirms that the sensitive measurement I present can indeed reflect the traffic congestion. A higher sensitivity is often accompanied by the decrease of travel speed, which indicates a more serious congestion. In addition, the increase of traffic volume and traffic variation in the road network usually causes the increase of the sensitivity but it differs where the increase happens.

The first contribution of this thesis is the method I present can be applied in other traffic models, in which travel time has to be estimated under a changing traffic demand. Secondly, the sensitivity measurement can be an effective approach to estimate traffic change and condition. And finally, by this measurement I give the method to find bottlenecks in the network, which can predict where traffic congestion starts and lead some traffic optimization strategies.

5.2 Limitation

As this thesis is based on data analysis, the estimation result is somehow limited by the historical data. For example, if there is an event in the study area and many cars pass through the links which used to be less-passed or medium-frequently passed, the travel time tendency according to traffic volume is likely to be completely different from the estimation I have got from historical data. Thus, the estimation result from common day can just estimate a general situation. In addition, I hope to observe how condition number changes when traffic volume and traffic volume difference continuously increases, but unfortunately, this continuous increase doesn't happen in the historical data. So I can't conclude the sensitivity under this situation.

In addition, when the travel time on each link is estimated by a linear regression model, the parameter varies with many other factors like weather. It is convinced that when it rains or snows, the travel time will be longer. But in this thesis, all the other factors are ignored. Travel time is assumed to be influenced just by the traffic volume on the link.

5.3 Future Direction

In the future, the similar analysis can be done in more study areas, with different traffic features. As the data limitation, there is no serious congestion in Tokyo but it indeed exists in many cities all over the world. It is hoped to find out how the sensitivity will change along as traffic volume and travel speed.

It can be also extended to a larger area suitable for the time scale of half an hour. The best range for this time scale is a district, with which a car can drive through almost from one side to another. If the study area is too large, the links even in one network is weakly connected with another one.

Finally, the sensitivity measurement and bottleneck I get can be used to create a global or local optimization scheme. For example, a traffic signal scheme or a road tolling scheme can adjust traffic volume on each link, which can cause the transformation matrix. And a local optimization can be a new road to share some traffic flow or widen the existing roads, which changes the estimated parameter of the linear regression models.

Reference

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, *15*(2), 105-111.

Daganzo, C. (1992). The cell transmission model. Part I: A simple dynamic representation of highway traffic.

Daganzo, C. F. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79-93.

Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. (2014). Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, *16*(2), 865-873.

Min, W., & Wynter, L. (2011). Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, *19*(4), 606-616.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998, February). DynaMIT: a simulation-based system for traffic prediction. In *DACCORD Short Term Forecasting Workshop* (pp. 1-12). Delft The Netherlands.

Shu, Y., Jin, Z., Zhang, L., Wang, L., & Yang, O. W. (1999, June). Traffic prediction using FARIMA models. In *1999 IEEE International Conference on Communications (Cat. No. 99CH36311)* (Vol. 2, pp. 891-895). IEEE.

Raymond, J. F. (2001). Traffic analysis: Protocols, attacks, design issues, and open problems. In *Designing Privacy Enhancing Technologies* (pp. 10-29). Springer, Berlin, Heidelberg.

Crovella, M., & Kolaczyk, E. (2003, March). Graph wavelets for spatial traffic analysis. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE*

Computer and Communications Societies (IEEE Cat. No. 03CH37428) (Vol. 3, pp. 1848-1857). IEEE.

Raymond, J. F. (2001). Traffic analysis: Protocols, attacks, design issues, and open problems. In *Designing Privacy Enhancing Technologies* (pp. 10-29). Springer, Berlin, Heidelberg.

Tian, Z., Jia, L., Dong, H., Su, F., & Zhang, Z. (2016). Analysis of urban road traffic network based on complex network. *Procedia engineering*, *137*, 537-546.

Sun, D., Benekohal, R. F., & Waller, S. T. (2006). Bi-level programming formulation and heuristic solution approach for dynamic traffic signal optimization. *Computer-Aided Civil and Infrastructure Engineering*, *21*(5), 321-333.

Christofa, E., & Skabardonis, A. (2011). Traffic signal optimization with application of transit signal priority to an isolated intersection. *Transportation Research Record*, *2259*(1), 192-201.

Sanchez, J. J., Galan, M., & Rubio, E. (2004, June). Genetic algorithms and cellular automata: A new architecture for traffic light cycles optimization. In *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753)* (Vol. 2, pp. 1668-1674). IEEE.

Sharon, G., Hanna, J. P., Rambha, T., Levin, M. W., Albert, M., Boyles, S. D., & Stone, P. (2017, May). Real-time adaptive tolling scheme for optimized social welfare in traffic networks. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2017)*.

Bergendorff, P., Hearn, D. W., & Ramana, M. V. (1997). Congestion toll pricing of traffic networks. In *Network Optimization* (pp. 51-71). Springer, Berlin, Heidelberg.

Yin, Y., & Lou, Y. (2009). Dynamic tolling strategies for managed lanes. *Journal of Transportation Engineering*, *135*(2), 45-52.

Çolak, S., Lima, A., & González, M. C. (2016). Understanding congested travel in urban areas. *Nature communications*, *7*(1), 1-8.

Clark, S., & Watling, D. (2005). Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological*, *39*(2), 119-140.

Lepanjuuri, K., Cornick, P., Byron, C., Templeton, I., & Hurn, J. (2017). *National Travel Survey 2016*. Technical Report Prepared for the Department for Transport.

Muñoz, L., Sun, X., Horowitz, R., & Alvarez, L. (2003, June). Traffic density estimation with the cell transmission model. In *Proceedings of the 2003 American Control Conference, 2003.* (Vol. 5, pp. 3750-3755). IEEE.

Çolak, S., Alexander, L. P., Alvim, B. G., Mehndiratta, S. R., & González, M. C. (2015). Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. *Transportation Research Record*, *2526*(1), 126-135.

Iqbal, M. S., Choudhury, C. F., Wang, P., & González, M. C. (2014). Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies*, *40*, 63-74.

Garcia-Nieto, J., Olivera, A. C., & Alba, E. (2013). Optimal cycle program of traffic lights with particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, *17*(6), 823-839.

Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological*, *53*, 64-81.

Turing, A. M. (1948). Rounding-off errors in matrix processes. *The Quarterly Journal of Mechanics and Applied Mathematics*, *1*(1), 287-308.

Wilkinson, J. H. (1963). *Rounding errors in algebraic processes*. Courier Corporation.

Cleve Moler, (2017, July 17). What is the Condition Number of a Matrix? [Blog post]. Retrieved from https://blogs.mathworks.com/cleve/2017/07/17/what-is-the-condition-number-of-a-matrix/

Skogestad, S., & Havre, K. (1996). The use of RGA and condition number as robustness measures. *Computers & chemical engineering*, *20*, S1005-S1010.

Kenney, C. S., Manjunath, B. S., Zuliani, M., Hewer, G. A., & Van Nevel, A. (2003). A condition number for point matching with application to registration and postregistration error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(11), 1437-1454.

Christiansen, S., & Hansen, P. C. (1994). The effective condition number applied to error analysis of certain boundary collocation methods. *Journal of Computational and Applied Mathematics*, *54*(1), 15-36.

Newell, G. F. (1998). A moving bottleneck. *Transportation Research Part B: Methodological*, *32*(8), 531-537.

Daganzo, Carlos F., and Jorge A. Laval. (2005) "Moving bottlenecks: A numerical method that converges in flows." *Transportation Research Part B: Methodological* 39.9 (2005), 855-863.

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.

Hammersley, J. (2013). *Monte carlo methods*. Springer Science & Business Media.

２０２０年度　修士論文　データによる旅行時間の感度分析および都市道路網におけるボトルネック検出

喩　騰瑶