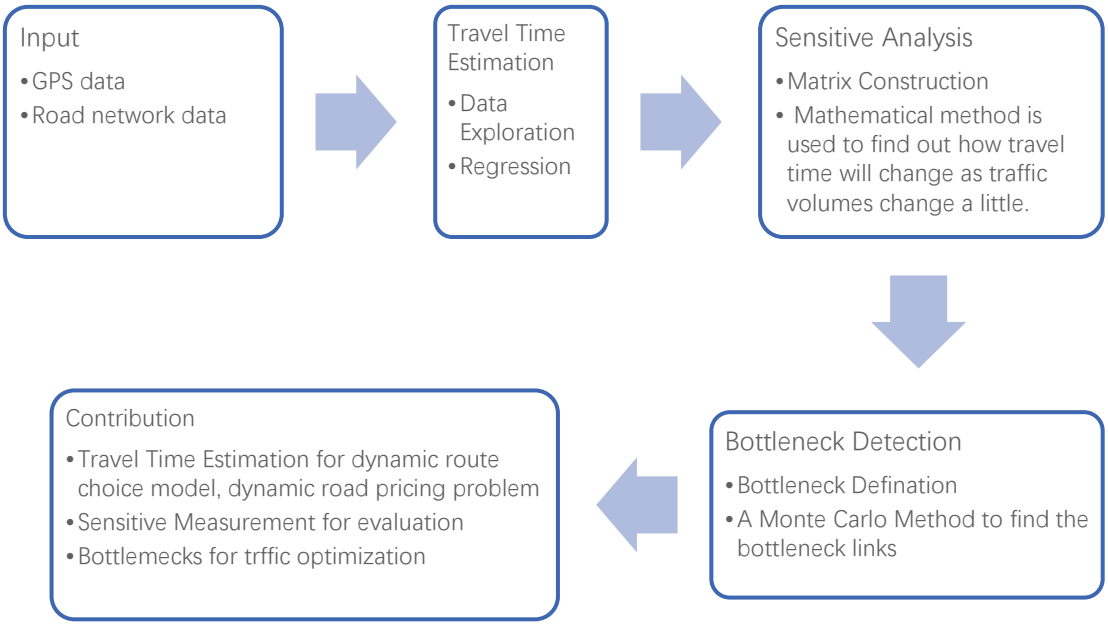
Data-based travel time estimation and sensitive analysis in urban traffic network

学籍番号　47186837
氏　名　喻　腾瑶（Yu Tengyao）
指導教員　柴崎 亮介 教授

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 requires 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. Until now, 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. However, few studies have pointed out the travel time according to the traffic volume.

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.

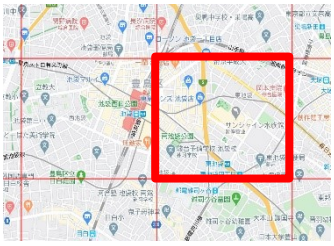
In detail, based on GPS data, the change of travel time is first estimated on each link according to the traffic volume on the individual links. Then a transformation matrix is constructed to connect the traffic volume on each link and the travel time for each route. By analyzing the matrix, the sensitive 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.



Pic. 1 framework of research

Travel Time Estimation

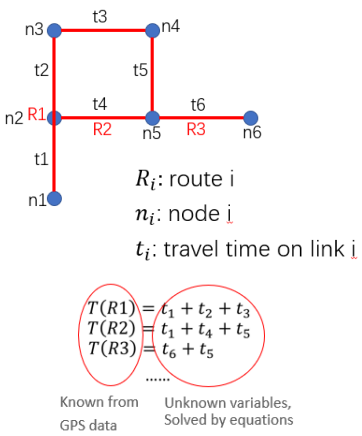
In this part, mean travel time is estimated on each link for half an hour. So, traffic volume is counted half an hour for each link and travel time is also calculated by this time scale. Then linear regression is used to fit the function between these two variables. There are some assumptions in my work: 1) mean travel time is just influenced by its own traffic volume but not influenced by other links. 2) waiting time in intersections is merged into link travel time.



Pic. 2 the selected mesh

The GPS data records the trajectory information in Tokyo for one month, including travel mode, position information and time. A mesh near Ikebukuro (around 1 km², as picture b shows) is selected to do some exploration. In this mesh, there are 555 road segments and 365 nodes. And the GPS data records over 137 thousand trips during the month in this mesh.

Then I tried to utilize the route information to solve the mean travel time on individual links by linear equations. The idea of this method is that if big data provides enough route information, there are enough constraints to solve the travel time on links. For example, in picture 4 there is an easy road network. We have route 1 - passing link 1, 2 and 3; route 2 – passing link 1, 4 and 5; route 3 – passing link 5 and 6. The route travel time can be known form GPS data, and link travel time we need can be solved by the linear equations if we have many such equations.



Pic. 4 An easy example

With the travel time and traffic volume on each link for each half an hour, I built a linear regression model:

$$Mean\ travel\ time = a \times traffic\ volume + b$$

In this formula, a is supposed to bigger than zero because travel time should not decrease as traffic volume increase. And when regression point is smaller than 10, which means this link is less used, a is also set equal to zero. According to the result, it is found that travel time indeed increase as the increase of traffic volume but the point is scattered. The R-square is also too small and MSE is large.

Sensitive Analysis

In this part, I aim to find how travel time is sensitive to the change of traffic volume. The parameters from regression is used to construct a matrix. By mathematical method, sensitive analysis in done to find out how much travel time changes as traffic volume changes a little according to the matrix.

In the same example of Pic. 4, link travel time is replaced by the formula $t_i = a_i * x_i + b_i$. After transforming, it becomes a linear matrix equation like Ax=T, or $A^{-1} \cdot T = x$.

T(R1) = (a1*x1+b1) + (a2*x2+b2) + (a3*x3+b3)+···
T(R2) = (a1*x1+b1) + (a4*x4+b4) + (a5*x5+b5)+···
T(R3) = (a6*x6+b6) + (a5*x5+b5)+···

T(R1) = a1*x1+ a2*x2+ a3*x3+ ···+(b1+b2+b3+···)
T(R2) = a1*x1+ a4*x4+ a5*x5+···+(b1+b4+b5+···)
T(R3) = a6*x6+ a5*x5+···+(b5+b6+···)
.....

$$A = \begin{bmatrix} \square & \cdots & \square \\ \vdots & \ddots & \vdots \\ \square & \cdots & \square \end{bmatrix}_{(m+1)\times n} \quad x = \begin{matrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_m \\ 1 \end{matrix} \quad T = \begin{matrix} T(R1) \\ T(R2) \\ T(R3) \\ \cdots \\ T(Rn) \end{matrix}$$

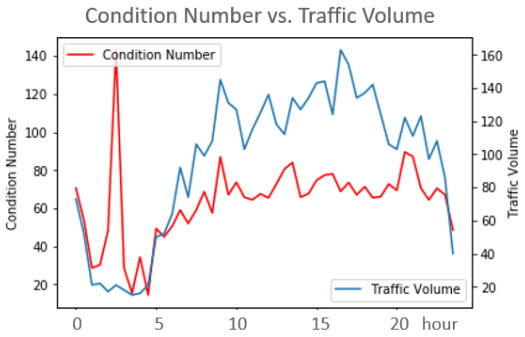
Consisting of Regression Coefficient — Traffic Volume — Travel Time

$$A^{-1} \cdot T = x$$

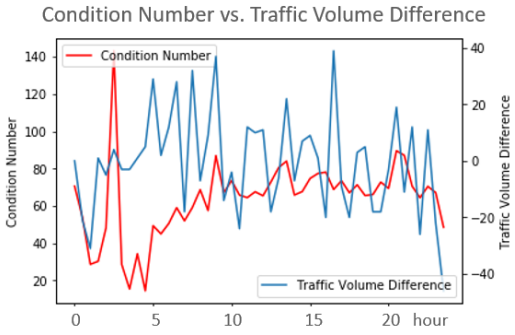
Pic. 5 Matrix Construction

The linear matrix equation $Ax = T$ means that the route travel time vector and traffic volume vector are connected by a transformation matrix A. Then, in matrix theory, a condition number of matrix A, calculated by formula (1), can evaluate how much vector T will change when x changes a little.

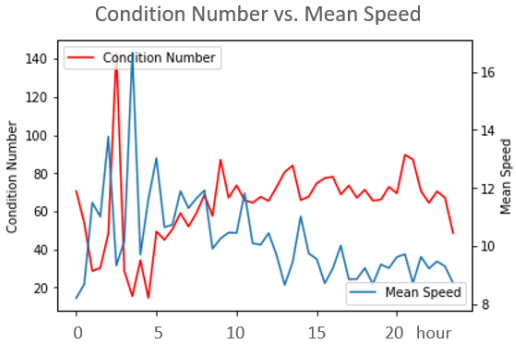
$$\kappa(A) = \|A\|\|A^{-1}\| \quad \cdots \text{formula (1)}$$



Pic. 6a



Pic. 6b



Pic. 6c

Picture 6 shows the result of condition number tendency and the comparison to traffic volume,

traffic volume difference and mean speed. 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.

Bottleneck Detection

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.

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.

By statistics to the traffic volume and linear regression coefficients of the bottleneck links, it is found that when busiest roads have around 40 trips for each half an hour, the traffic volume on these bottlenecks are not as much as assumed. 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 is expanded to the whole network.

Summary

In the master thesis, Chapter 1 will talk about the background, explain why I choose this topic and the research objectives. Chapter 2 will give the description of the data I used and explain how I get travel time and traffic volume from the data and then estimate the travel time relationship function of traffic volume for each link via linear regression models. Next, in Chapter 3, the regression result is used to construct a transformation matrix between route travel time and link traffic volumes. With the matrix analysis, a condition number, which can reflect how much travel time will change as the change of traffic volume, will be calculated to respect the sensitive of the network. In Chapter 4, a bottleneck link is defined as the links where traffic volume increase influences the network sensitivity the most. Monte Carlo Method is introduced to find the bottlenecks then. Finally, In Chapter 5, the thesis is summarized and the limitation and future work are discussed.