A two-factor evaluation of bus delays based on GIS-T database and simulation

Li Zhang
Xi Ren

May 2010

Master’s Thesis in Geomatics

Master Programme in Geomatics
Examiner: Prof. Bin Jiang, Prof. Anders Östman
Supervisor: Dr. Julia Åhlén
Abstract

During the urbanization process, vehicles quantity increase with expansion in population. Under this situation, bus transportation system also suffers from bus delay. Bus delay could be caused by a series of factors, for instance, overload passengers, traffic jam, traffic accident and other unpredictable situations. Therefore, choosing crucial elements to efficiently evaluate bus delay is a complex problem in bus delay researches and operation management. The thesis propose an approach to evaluate and explain bus delay by two elements: traffic congestion and passengers’ waiting time. Those two elements would represent the action of external and internal factors on bus operation. This approach could be adaptive to explain the reasons for bus delays, thus to help the optimization of bus lines and give useful information for decision making of transportation company. To achieve the research aim, a GIS-T database was created by combining the GIS database and TIS database. Spatial data as well as attribute data are combined in the database to represent the crucial information for bus delay. Based on GIS-T the database, the impact of traffic congestion and passengers’ waiting time was calculated using the bus line simulation. By implementing the above steps, the main cause of bus delay was studied. A case study application of this method is narrated; focusing on optimize the bus system of Guiyang city, South China. Different methods are used to find out the problem of system and the reason for delay. Moreover, optimization suggestion is proposed according to result. Compared with other methods, the two-factor method has the advantage of locating the reason of delay for each station. The time performance is not superior to other methods. By comparing the situation of adjacent station, the proportion of traffic congestion and overload passenger in bus delay was determined. The two-factor method is applicable for other transit system in different cities which has similar structure as Guiyang. However, for cities with other structure, a feasibility should be made to select an appropriate model.

Keywords: Geographic information system (GIS); Transportation Information System (TIS); Geographic information system for Transportation (GIS-T); Simulation; Traffic congestion index; Passengers’ waiting time, Two-factor evaluation
Acknowledgements

The authors would like to thank Professor Anders Östman, Professor Bin Jiang and Dr. Julia Ahlen for their personal help, critical and constructive instruction as well as information they provide. Furthermore, thank Guiyang Public Traffic Company and Guizhou University for providing the data and opportunity for us to optimize the transit lines of Guiyang city. We would like to express our heartfelt thanks to PTV AG Company for giving us free version of software and technical support which are significant help for our thesis.
Content

Abstract .............................................................................................................................................. 3
Acknowledgements ......................................................................................................................... 4

1. Introduction ........................................................................................................................................ 6

2. Basic theoretical concepts ............................................................................................................. 7
   2.1 GIS-T .............................................................................................................................................. 7
       2.1.1 Data model .......................................................................................................................... 10
       2.1.2 Microscopic model ............................................................................................................. 11
   2.2 Delay factors .............................................................................................................................. 12
       2.2.1 Traffic congestion .............................................................................................................. 12
       2.2.2 Passengers’ waiting time .................................................................................................. 13
       2.2.3 The bus line interaction research ..................................................................................... 13

3. Methodology ...................................................................................................................................... 14
   3.1 Study area ................................................................................................................................... 14
   3.2 Data .............................................................................................................................................. 15
   3.3 Process ......................................................................................................................................... 15
   3.4 Data preparation ........................................................................................................................... 16
   3.5 Simulation .................................................................................................................................... 17
   3.6 Optimization ............................................................................................................................... 20
       3.6.1 Traffic congestion calculation ............................................................................................. 21
       3.6.2 Passengers waiting time calculation .................................................................................... 23
       3.6.3 Optimisation ......................................................................................................................... 25

4. Applying the method on Guiyang City ......................................................................................... 25
   4.1 Data preparation ........................................................................................................................... 25
   4.2 Data transformation ..................................................................................................................... 26
   4.3 GIS-T database ............................................................................................................................ 27
   4.4 Simulation ................................................................................................................................... 30
       4.4.1 Simulation design (Traffic flow model) ................................................................................. 30

5. Results and analysis ....................................................................................................................... 33
   5.1 Calculation and analysis .............................................................................................................. 36
       5.1.1 Traffic congestion ................................................................................................................. 36
       5.1.2 Passenger waiting time ....................................................................................................... 39
       5.1.3 Two-factor delay assessment ............................................................................................ 41

6. Discussion .......................................................................................................................................... 46
   6.1 Simulation method ....................................................................................................................... 46
   6.2 impact of traffic congestion and overload of passengers ............................................................. 47
   6.3 applicability and algorithm complexity ...................................................................................... 48

7. Conclusion ......................................................................................................................................... 48

8. Further research ............................................................................................................................. 50
1. Introduction

The urban expansion and rapid development has increased the demand for education, housing, transportation and other aspect. As one of them, transportation demand has increased steadily, as the world’s population continually grows. However, increasing population and automobile dependency has also brought delay problem to bus transportation systems. Bus delay could be caused by a series of factors, for instance, over load passengers, traffic jam, traffic accident and other unpredictable situations.

Then, the evaluating methods for urban bus system can be divided into two categories: First, from general perspective, calculating the route cost or efficiency score for bus system to locate the problem, as Mandi (1980), Lao (2009) and Meignan (2007); Second, from detail perspective, by using traffic congestion, as Taylor (2000); or passengers’ waiting time (Goverde, 1998) to evaluating the bus system. The first method uses the financial or social cost of route to find the problem in system. And most of them consider the social economy in it to locate problem and propose the solution. However, few of them had mentioned the method to research the exact cause for bus delay. Just as mentioned in previous part of the paragraph, it could be resulted from various factors. Among the study of them, the traffic congestion (Taylor 1999) and passengers’ waiting time (Goverde 1998) are widely used to represent the complicated delay factors from perspective of traffic and passengers. They also mentioned how to evaluate the delay with several factors, like congestion index or acceleration noise. Nevertheless, these methods only consider the reason of delay unilateral. Only the traffic factors or the passenger influence is considered in the evaluation. Therefore, there are no valid research methods to analyze and validate the impact between traffic congestion and passengers’ waiting time for bus delay. Either for transit planner or researchers aims at evaluating the short distance bus system in small city, such as transit system in Guiyang city, China, a sufficient method is needed and should be proposed to solve this kind of problem.

Therefore, the research problems of this paper would be defined as follows:

- How would critical factors for bus delay be represented in a database?
- How to evaluate and validate the impact of traffic congestion and passengers’ waiting time?
- How is the effectiveness of this method as compared to other methods?

To solve these problems, this research should be based on specific GIS technology. A new evaluation method which integrate the traffic congestion and passenger waiting time will be proposed and validate in thesis. Based on this method, a case study application of Guiyang city, China will be studied. And the aim of application is to locate the reason for bus delay and propose the solution. In the case study, different methods will be used to calculate to bus delay from different aspects. Then the new method will be proposed and applied for case study. By evaluating the different bus lines for bus system in Guiyang city, the effectiveness of new method which we proposed will be validated.
2. Basic theoretical concepts

GIS is a computer-based system, which is fit for capture, storage, manipulation, visualization and analysis of geographic information (Thill, 2000). In geographic information science view, each object has two unique geographic attributes. The GIS could manage the objects and their information with their spatial characteristics (Miller, 2001). They are “spatial dependency” (the tendency for things closer in geographical space to be more related) and “spatial heterogeneity” (the tendency of each location in geographical space to exhibit some degree of uniqueness) (Anselin, 1989). In previous research, Geographic Information System is known as a powerful tool applied to solve transportation problems, especially on manipulation, visualization and animation of urban traffic data (Claramunt 2000). It claims that GIS have special ability on spatial representation and modeling to describe a segment of the surface of the earth (Frank 1992). In the transportation research area, three classes of GIS models are relevant (Goodchild 1992).

- Field model: this model is often responsible for representing continuous variation of a phenomenon over space, such as terrain elevation.
- Discrete model: this model describes discrete entities space, such as highway rest areas, urban areas, and etc.
- Network model: this model is represented by topologically connected linear entities, such as roads, railway lines, etc.

The discrete model and network model will be adopted in this research. Discrete model could regard urban features as background information, and network models describe roads and bus line in the database.

Based on early pioneers’ researches of GIS, it is concluded that GIS is a special database compared to normal database management engine, due to its geo-visualization capability, analytical capability and features management capability. Generally, there are plenty of methods that could be used to manage transportation system and solve transportation problems, therefore GIS is not the only one. This research will analyze and simulate bus line based on perspective of GIS since it is better than traditional methods to visualize and locate important transportation factors, such as bus stations.

2.1 GIS-T

Transportation is not only regarded as a major component of the daily life and further development direction (Wiggins et al. 2000) but also as one of the most basic application aspects in GIS. However, GIS has some problem to share data with other database. For instance, transportation information system data are modeled by network data model, represented by a planar graph, such as structure of bus network in the application of this paper. If the network data model in GIS, stored by location referencing, and two-dimensional graph presented by X and Y coordinates or
longitude and latitude, the network data could be directly shared between TIS and GIS. Considering this situation, GIS – T (geographical information systems for transportation) is created by combining the advantages of both GISs and TIS for geographic information to store, visualize, plan, and analyze transportation information. Fletcher (2000) defines that GIS-T embraces hardware, software and data as well as methods to process the information related to the geographical regions and transportation system. Comparing to traditional GIS, GIS-T has dominance in data sharing with the transportation management system and in professional capacity as a toolkit for the sake of managing transportation problems. Therefore GIS-T is conceptualized since the traditional GIS meet several limitations on handling the transportation data and communicating with the existing systems. With the development of the computer techniques and recent emergence of a “second GIS-T renaissance” (Fletcher 2000; Wiggins et al. 2000), the GIS-T application covers much of the broad scope of transportation. GIS tools are not only praised by transportation analysts and decision makers on design and management, environment assessment and public transit planning subjects, but also are developed with new framework to handle the transportation problems in emergent ways.

GIS – T has reliance on the network data model but excludes any other data model (Thill 2000). As pointed by Goodchild in 1998, despite “the basic network data model is already a domain-specific departure from conventional GIS data modeling, it does not suffice to handle the complexity embedded in transportation network data.” Thus, the extensions, such as turning tables and links and other transportation factors which comes from conventional TIS, are needed to handle particular structures.

Actually, GIS – T is an extension from GISs and TIS. The final report of NCHRP (National Cooperative Highway Research Program Report) explained the GIS – T as the production of cross-fertilization of an enhanced GIS and enhanced TIS (Vonderohe et al. 1993). The structure of GIS – T is showed as Figure 1.
Nowadays, there is variety of choices among data models employed in transportation applications, such as Geographic Data File Standard, NCHRP 20-27 (Vonderohe et al., 1998), enterprise GIS-data model (Dueker and Butler, 1998) and TIGER (Topologically Integrated geographic Encoding and Referencing system). In this paper, the node-arc representation (NAR) (Miller 2001) data model is adopted to create a new database. The detailed explanation will be illustrated in the following chapters. By following Bulter’s and Dueker’s (1998) research, an important data sharing principles be referenced in this dissertation, which is: “Transportation features must be uniquely identified to facilitate sharing of data among participants. Participants need to identify common features in sharing data”.

The GIS-T database design stage could be regarded as a crucial stage in system design. The database should be designed as the GIS-T database integrating the spatial information as well as the attribute information. As Thong et al., (1997) conclude the GIS-T design should follow four major steps: first of all, the collection of the spatial and attribute data should be done previously, then the transformation is needed to translate the raw data for system input, thirdly, the computer techniques are necessary for visualization and graphical representation; finally, the multimedia such as the images, video and sound should be connected to the GIS-T database to support the application. Also the design should follow the rules of Codd (1970) in database structure design that defines types of the value where the column is contained and
how to keep the second normal form to maintain the dependency of elements on the primary key. The database design could also be organized according to the system implementation phases. Wang and Stauffer (1995) built up the database on basis of different phases of study. For the purposes of study and analysis, they stored the needed data in different phases. There is no doubt that an abundant database could help the traffic relief research, whereas the structure seems a little disorder without a clear plan at first. Moreover, there might lead to data duplicate and disaccord problems. It is undoubtedly saying that a well-arranged database would benefit the system. The logic Entity-relation model for application in Guiyang City will be introduced in Application chapter.

GIS-T database plays an essential role in this paper. It is an information platform that stores transportation data and background data. The background information is stored in the personal geo-database which contains geographic reference data, feature of urban areas and roads. Bus Corporation provides TIS data which includes velocity data, GPS data, structure of bus system data, and etc. GIS-T database will be shaped by combining personal Geo-database with TIS database of Bus Corporation. The rules mentioned above will be strictly followed in the design.

2.1.1 Data model

Data model and analysis relating to database are the key components in the system. Despite there exist various kinds of data models, the network representation and linear reference system are two classic representations of transportation system (Miller, 2001). Traditional planar network model system referred as node-arc representation (NAR) divides a transportation system into separate sub networks, and unit them with the “transfer arcs”. Each separate network could be treated as a specific model. When applied to public transit network, a node stands for bus stop while an arc stands for a road connection. Combined with traffic analysis zones, the computation could be simplified (Miller, 2001). If the node-arc model is adopted, in accordance with Goodchild’ s opinion on the data model for NAR, the Arc table including the arc-ID and its direction information, the turn tables including the different direction cost related with intersection and the reference address, are all necessary in the data implementation (Goodchild, 1998).

The network can be represented by the node-arc method or linear referencing method. In this paper, the node-arc representation method will characterize bus network even though it may be changed from node-based to point-based. Generally, node-based transit network illustrates location at physical highway intersections, while actually, stops are normally pictured as points on arcs rather than nodes. Therefore, point-based method is better than node-based method in describing network (Choi and Jang 1997). In thesis research, stations are represented as nodes while the road segments expressed by arcs. At the simulation part, the transit network would be constructed with this representation method.
2.1.2 Microscopic model

The microscopic model is an abstraction of bus network structure. Compared with the crude structure proposed by Schulze in 1993, the bus-network structure created by Meignan et al. in 2007 (Figure 2) seems more operational. The itinerary, line, bus stop and bus station comprise the network. In transportation information system of Guiyang Public Traffic Company, this structure is also adopted nowadays. As figure 2 shown, Structure A represents the interaction between two bus lines, where they share a bus station. Then the station shared will be unique station instead of two duplicate ones in network design as well as simulation. Since the same bus line is operated by several bus tours, they arrive at a certain station at different times (Figure 2b).

![Figure 2: The Meignan bus-network structure](image)

Different behavior models depend on different opinions regarding to agent behaviors. Meignan created the behavior model of a traveler, shown in Figure 3, and a bus agent which is quite similar to common bus behavior. The behavior of bus component in passenger boarding simulation follows the behavior model mentioned. Due to the difficulty in tracing the mobility of a traveler, the traveler behavior is modified. The model focuses on time consumed in waiting on a bus station and getting off at a destination but ignores the steps of proceeding to bus station and departing from a bus stop.
From the above description, it is easy to find out that the microscopic model concentrates more on moving vehicles. And the necessary elements for microscopic model concerning on single moving vehicle are fulfilled. Therefore, in model design of this thesis, microscopic model is the major part of simulation, and elements like vehicle density, mean speed and quantity vehicles statistics for characterizing the road situation are also been used as complements.

2.2 Delay factors

2.2.1 Traffic congestion

“Traffic congestion is the phenomena of increased disruption of traffic movement on an element of the transport system ……occurrence of congestion is the presence of queues.” (Taylor 1999). As quoted above, traffic congestion is an existing part of transport system, though its specific definition and identification is not immediately obvious (Taylor 1992). This phenomenon is caused by common factors and special factors. Common factors could be traffic flow, traffic signals and etc. And special factors are the incidents, such as road construction, traffic accident and etc. Actually, traffic congestions are phenomena caused by common factors. In reality, several traffic congestion indexes provide different basic information for traffic situation of network. There are two methods to collect such information. One is collecting information from fixed sites, e.g. monitors equipped on the crossroad. The other is collecting information from moving observers, for instance the onboard GPS instruments. Traffic congestion may occur in any part of transport system, whereas the level of congestion is recognized by some threshold value (Taylor 2000), exampled as traffic delay, congestion index and etc. In this paper, traffic congestion is of great importance parameter to investigate the reasons of bus delay.
2.2.2 Passengers’ waiting time

The bus delay will cost passenger extra time to wait for the bus. Higher passenger waiting time can imply large bus delay. Generally, it could be combination result of bus delay and accumulation of delay passengers.

The passenger waiting time could be defined as two categories: primary waiting time and secondary waiting time. As Goverde mentioned in 1998, the primary waiting time was defined as summing all involved waiting times at the transfer stations while the secondary waiting time refers to the sum of all waiting times at subsequent stations. The secondary waiting time is associated with the road congestion and traffic situation which belongs to the road section part, while the primary waiting time is more related with the bus stop part.

It is not difficult to figure out ways to decrease passenger waiting time. Meignant et al. mentioned in 2007 that the passengers’ waiting time could be decreased through adding more operating buses. They concluded a relation between the total number of buses in service and average waiting time, as Figure 4 shown.

![Figure 4: The simulation result for passenger waiting time by Meignant](image)

It is obvious that more buses operating in the network will increase operating cost and traffic congestion. As a result, the tradeoff between the number of buses in service and the passengers’ waiting time is a critical factor in the optimization.

But aim of simulation in application is not to minimize the passenger waiting time, but use it as a crucial factor to judge the traffic system operation efficiency instead.

2.2.3 The bus line interaction research

The bus system is comprised of several bus lines. It is undeniable that there exists interaction between bus lines. According to the spatial relationship, bus lines might be divided into two types, one is bus lines who share the same part of road lane and bus lines that do not. In order to clarify the mainstream of research problem, there exists the need to discuss the interaction between the bus lines. Concluded with two basic components, the interaction is partitioned by the impact of traffic and passenger. The method mentioned above using two crucial parameters to measure and judge the reasons of decreasing the service efficiency are traffic congestion and passenger
waiting time. In order to validate the method, it is necessary to analyze the two different types of bus line.

Regarding to the bus lines sharing the same part of road lane, the impact of traffic is addressed with the traffic congestion. Since the bus lines share the same road lane, they share nearly the same road situation at same time. If a road is jammed, not only vehicles of certain bus line is blocked, but also the other buses of different bus lines sharing the same road lane are delayed. Thereby the traffic congestion that represents the road situation of certain road segment is a parameter who might express the interaction of busses between different bus lines as well as impact of other vehicles that share the same road lane with bus studied.

Regarding to bus lines share no route, the interaction still exists. As the traffic jam affects the road situation of certain road segment, the route choice might be changed by drivers when they approach the blocked segment, in other words, the parallel path or called alternate route could be chosen in this situation. The traffic situation in alternate route could change due to this external traffic input. It is obvious that not only buses in these alternate routes, but also buses in jammed road segment will be affected by this action. Still, the changes would be pictured with the traffic congestion.

Analyzing the interaction from the perspective of passengers, the interaction between two bus lines could be established the passengers transfer between two bus lines. Nevertheless, the passenger mobility is hard to trace without the pre-store card system. With the approach mentioned in thesis, the interaction of passengers could also been taken into account. One of the most fundamental reasons is despite how difficult the passenger mobility is hard to trace, the passengers will finally exit the bus line or reenter the bus line as the waiting passenger at certain stations. The leave passenger is unnecessary to study for subsequent part of bus line, meanwhile, the passenger reenter the certain bus line at certain station could be evaluated with passenger waiting time. Thus, the interaction between the bus lines, no matter it is caused by impact of traffic or passenger, could be evaluated with the aforementioned approach in thesis. Other factors affect bus lines like the census or traffic accidents could also be evaluated with the approach.

3. Methodology

3.1 Study area

The study area is in Guiyang city which is the capital city of Guizhou province, China. Guiyang city locates at 106°27′ E ~ 107°03′ E, 26°11′ N ~ 26°55′ N, covers 2403 square kilometer and the average altitude is 1071m. The province consists of seven regions which are Yunyan region, Namning region, Xiaohe region, Jinyang region, Baiyun region, Wudang region and Huaxi region, and its population is 2,300,000. The main districts of the city are Yunyan and Namning
region, they are the oldest region in Guiyang city herewith the traffic problems often appear in these regions.

Guiyang Public Traffic Company is the unique transportation company in Guiyang city running the whole transportation system of this city comprising of 87 bus lines. Among them, 16 bus lines pass through the oldest region. The customers’ satisfactory still need be improved, since there still exists jam situation and bus delay. At the same time Guiyang Public Traffic Company would benefit from the optimization of bus lines.

In this paper, the research will be focused on those bus lines which pass through Yunyan and Nanming region.

3.2 Data

The existing data come from the Guiyang Public Traffic Company and Traffic control center of Guiyang, and consists of:
- The number of passengers who have boarded the bus, at each station.
- The number of people who leave the bus, at each station.
- Amount of passengers onboard at each station.
- Average bus velocity between stations. (They are calculated from on-board GPS data.)
- Driving time between bus stations.
- The bus real arrival time on each station.
- The traffic on each intersection in main area of city (they are classified into three groups.).
- The velocity of vehicle on main road in the main area of city.
- The location and characteristic of all Pedestrian Bridges in main area of city.
- The distribution and characteristic of traffic light in main area of city.
- Time tables for No. 4, 5, 6, 9, 10, 12, 13, 18, 19, 20, 27, 30, 51, 54, 69, 202 bus lines. 2007-2008.

3.3 Process

The process of research is depicted by Figure 5.

Data preparation: the original data and secondary data should be prepared before data transformation. One reason is that the original data and secondary data are not all included in the GIS-T database, for instance fuel consumption and engine revolution.

Data transformation: not all types of data are qualified for the GIS-T database, e.g. population density data which may to be modeled by a field model for future analysis. Herein, data sharing is of importance when establish GIS-T database.
GIS-T: GIS-T database is the key part of this research. It produces integrated and detailed data to support simulation, and provide a platform for visualization.
Simulation: simulation is a method to present original bus transportation system and optimized bus transportation system. Comparing improvement between pre-optimal and post-optimal will be on basis of the results of simulation.

Results: this part will clarify improvement after optimization.

![Diagram](image)

**Figure 5:** The process of research

### 3.4 Data preparation

The original data and secondary data sources contain mass of information. In the data preparation phase, only relevant information is selected.

As mentioned before, there are three types of traffic simulation models, i.e. microscopic models, macroscopic models and mesoscopic models (Lieberman and Rathi, 1997; Hoogendoorn and Bovy, 2001). Among them the microscopic model concerns single moving vehicles while the macroscopic model concerns the external factors like traffic density. In this design, the microscopic model acts as the major part of simulation and the traffic density was added as constraints on the behavior of the agency.

Considering the current situation, there are several differences as compared to previous studies.

1. The limit of data availability. A few data like the environmental factors and economic factors (e.g. census data) are hard to receive.
2. Each vehicle, as a component of bus-network, would exhibit its own operational features, which reflect the vehicle density and road situation where macroscopic simulation is responsible for, and it is believed that optimization to each vehicle operation could affect the situation of road segment.

3. The GPS data on vehicles make it possible to calculate the Mean speed of vehicle and provide more accurate estimation of road situation.

4. The mission given by Guiyang Public Traffic Company Company is to optimize the existing bus lines. Meanwhile, a relatively successful microscopic optimization would be evaluating an easier and quicker plan other than come forth huge arrangement as adding a new road or other complicated actions.

Based on all the listed consideration, it is decided to choose microscopic model as simulation model and utilize GPS data on vehicles to compute the mean speed and other factors as congestion index to express the road situation which could be the outcome of macroscopic model. Combining both of them (microscopic model results and those factors like mean speed), the simulation could be closer to reality.

### 3.5 Simulation

Generally, the simulation begins with designing a simulation model. The advantages of simulation as compare to real experimentation are cost saving and faster. Analysis target transportation system and Setting a simulation object and expected result are the first stage of simulation. The basic processes could be addressed by twelve steps (Banks 1996). It is worthy of clarifying that, taking cost of simulation and characteristics of bus transportation into account, the simulation in this research will not follow these steps exactly. The cost of simulation will be too expensive and the time of simulation will be too long if all steps above are followed to simulate bus transportation system. However, simulation time could be shorted by simulation software when computer is used to simulation. Eventually, based on simulation theory, the steps of simulation will be changed according to the requirements of research. The new steps are depicted as figure 6.

The type of transportation system is one of the significant measurements for simulation. Different types of systems might be simulated via different approaches. Therefore, there should define the types of system before simulation is conducted. It can be divided into two parts of the bus transportation process. One is on traffic flow part, the other one is on bus station part. The on traffic flow part describes the traffic flow system regarding as a dynamic system (Pal and Bose 2009). The bus is moving in this system, and the change of velocity depends on traffic congestion. The bus station part defines single service operations (Fitzsimmons 1998), which is similar with a checkout system in the supermarket. The structure of single service operation system is described in figure 7.
Figure 6: New steps of simulation, modified after Banks (1996)

In the following, more information of each element of the system is outlined:

- **Arriving passengers**: The arriving passengers part specify people who come to bus station and start to wait for buses. The passenger arriving time is an essential factor for simulation.
- **Queue of passengers**: The length of queue describes the quantity of passengers during waiting hours. In real world, the quantity of passengers waiting for a bus
might sometimes beyond capacity of a bus. Under this situation, more buses should be added to absorb excessive passengers.

- Cashier’s station: In most cases, the cashier is the bus driver. The service time of cashier will affect dwell time of the bus. Therefore, improving service time of cashier will be helpful to reduce dwell time of a bus, though it is not the main approach.
- Onboard passenger: The number of onboard passengers should be maximized to the maximum capacity of one bus.

![Figure 7: Bus station operation system](image)

The queue of passengers is related with passenger arriving time and bus transportation system delay. In general, the passenger arriving time is affected by population density of neighborhoods and passengers’ travel demands. And the bus transportation system delay is measured by traffic congestion. The relationship between on traffic flow part and on bus station part is illustrated as figure 8:

![Figure 8: The relationship between two systems](image)

Traffic congestions are delaying bus transportations, causing the queue of passengers to be larger than expectation value. However, if delay time of system is in the expectation interval, the queue of passenger is unrelated with traffic congestion. Therefore, the traffic congestion should be measured before discussing dwell time of bus on station. The simulation process should be focus on those two parts.
In order to fit the requirements of simulation, the software Vissim is adopted to simulate mobility of passengers and the road traffic model. Just as Federal Highway Administration claimed, the VISSIM is a microscopic simulation model using several detailed information, especially some information manifests multiple settings, e.g. the friendly merging behavior, and the priority rule that the following vehicles would slow down to let vehicles merge into traffic flow. It is those features and settings that polish the fidelity of the simulation.

Several ascendancies that the Vissim embedded would bring it great adaptability for traffic simulation. The most important advantage is VISSIM can not only simulate alighting, boarding and waiting passengers in each station, but also simulate the road traffic situation. Hence, the road network should be modeled similar to real road network; the length, the speed distribution and other parameter like passenger quantity are similar to the real situation. The advantages of this software are as follows:

1. Simple and convenient interface.
2. Direct representation. Network is constructed on supported background pictures, such as JPG and BMP format as well as aerial photos. Combined with the three dimensional models used for visualization, it provides a clear and direct presentation of simulation result.
3. Advanced Signal Control Mechanism, which makes the simulation of crossroad intersection closer to reality.
4. Customized results. The output results can be customized. The data collection speed can be controlled by user through adjusting the sampling time. Some output results exemplified as drive delay are quite available to repress the situation.
5. Seamless interface. The network constructed in VISUM would be imported into VISSIM to simulate meanwhile the simulated VISSIM network ought to be transferred into VISUM to get a detailed graphical representation like cobweb graphs.
6. Innovative pedestrian simulation improves the urban traffic simulation accuracy since it involves the effect of pedestrian into system.

Despite the model applied in the thesis is only a simple part of software functions, the optimization deserves necessary help and benefits from it. Due to limitation in the data, the full potential software potentialities have not been well exploited. The further researches will certainly benefits from the full use of the software.

3.6 Optimization

To fulfill the requirements of the Guiyang Bus Corporation, two ways of optimalization are formed. One is to reduce the waiting time of passengers as to satisfy customer requirements. The other way is trying to avoid getting more buses on road segments which experience high traffic congestion. These simulation methods
avoid traffic congestion and satisfy customers’ requirements through resetting bus schedule, bus line and adding more buses during busy period. The traffic congestion is calculated according to the GPS data, and measured by traffic delay, congestion index (CI), proportion stopped time (PST), acceleration noise (AN) and mean velocity gradient (MVG). And the waiting time of passengers are calculated according to the queue length, and measured by amount of customer in waiting queue.

3.6.1 Traffic congestion calculation

Delay

Traffic delay measures the extended travel time (Taylor 1992). It is defined as the actual travel time minus the minimum travel time (equation 1.)

\[ DTR = T_a - T_0 \] (1)

Where, DTR is traffic delay, \( T_a \) is actual travel time and \( T_0 \) is the free flow travel time, traffic delay should be a nonnegative number.

The minimum travel time (free flow travel time) is an assumed value which is estimated by traveler expectation. Therefore, if the expected travel time is not known, traffic delay is not usually used when studying traffic congestion. Generally, in bus system, the value of \( T_0 \) equals to scheduled time interval at each two station according to time tables. In this research, this indicator can be adopted to analyze bus transportation delays as long as the bus corporation specify an expected travel time. Traffic delay is the basic concept for congestion index.

Congestion index

The Congestion index measures the level of delay as compared to the estimated travel time (free flow travel time) (Richardson and Taylor 1978). The index formula may be specified as:

\[ CONI = (T_a - T_0)/T_0 \] (2)

Where, CONI is congestion index value, \( T_a \) is actual travel time and \( T_0 \) is the free flow travel time, The Coni index should be a nonnegative number.

If the CONI value is zero it means that the actual travel time is equal to the expected travel time. Then the transportation system meets traveler’s requirement. If the CONI value is one it means that the actual travel time is twice as long as the expected travel time. The system delay makes up 50% of total travel time. Generally, speaking, high CONI value means low user’s satisfaction in transportation system. The congestion index is often used to measure the utilization rate of transportation system (Quality of transportation system).
Proportion stopped time

The Proportion of stopped time (PST) on a trip segment measures journey quality. The index can be used to compare different routes (Taylor 2000). The formula for this indicator is:

\[ PST = \frac{T_s}{T_t} \quad (3) \]

Where, PST is the proportion stopped time, Ts is the stopped time in the segment of road, and \( T_t \) is the total travel time.

The stopped time defines the time when the vehicle is stationary or when it is moving at less than a threshold speed value (2km/h) (Zito et al. 1995). Based on stopped time, the running time (Tr) can be calculated by total travel time minus the stopped time.

\[ T_r = T_t - T_s \quad (4) \]

The running time is a crucial parameter to calculate acceleration noise index. In this thesis the proportion stopped time will be applied to indicate the quality of each bus route.

Acceleration noise

The acceleration noise measures how the speed varies when a vehicle moves along a route Underwood (1968). The acceleration noise index is used to measure the quality of traffic flow and the level of congestion and can be written as.

\[ AN^2 = \frac{1}{Tr} \sum_{i=1}^{n} \frac{\Delta V_i^2}{\Delta T_i} \quad (5) \]

Where, AN is the acceleration noise, Tr is the running time of the whole journey, \( \Delta T_i \) is the time interval taken for a speed change \( \Delta V_i \).

If the travel process is at a steady speed, the acceleration noise will present a low value, which means the quality of journey is good. If there are a lot of stop-start driving during the travel process, the acceleration noise will be large, which means the quality of journey is bad. In this paper, the acceleration noise is used to measure the quality of bus travel process. Actually, bus traveling is different to other types of traveling, since buses always stop at the bus station and when approaching red traffic signals. Therefore, acceleration noise will be a secondary parameter to describe the level of congestion in this research. However it is a powerful complementarity for CONI and PST.

The mean velocity gradient

The mean velocity gradient (MVG) is based on the AN index. It is also a parameter to describe the quality of travel process. The difference between acceleration noise and mean velocity gradient is that the acceleration noise focus on the frequency of speed changes to describe traffic congestion, while the mean velocity gradient is focus on the speed to describe traffic congestion. The mean velocity gradient is specified as:
\[ \text{MVG} = AN/V_a \]  \hspace{1cm} (6)

Where, \( \text{MVG} \) is the mean velocity gradient, \( V_a \) is the average of speed in the whole travel process, calculated as

\[ V_a = \frac{1}{T_t} \int_0^{T_t} V dt \]  \hspace{1cm} (7)

Where, the \( V_a \) is the average of speed in the whole travel, \( T_t \) is the total travel time. In this paper, the mean velocity gradient will be used together with acceleration noise to measure the quality of journey.

### 3.6.2 Passengers waiting time calculation

Both the primary waiting time and the secondary waiting time, was consists of two main parts namely the waiting time in bus and waiting time on platform. As Van der Waard and Van Goeverden indicated in 1989 and 1990 respectively, waiting time on the platform is perceived the major part of the total waiting time. The relation between them can be expressed by Formula 8.

\[ W_T = W_B + 3W_P \]  \hspace{1cm} (8)

Where \( W_T \) represents the total waiting time while the \( W_B \) express the waiting time in bus, \( W_P \) represent the waiting time on platform.

As passengers are classified into four main classes, i.e. originating passengers, through passengers, transferring passengers and terminating passengers, the waiting time of passengers ought to be evaluated for each group of passengers respectively. The passenger types are defined by Goverde (1998) as:

- **Originating passengers**: passengers who just starts travel at the station
- **Through passenger**: passengers inside the bus stopping at the station,
- **Transferring passenger**: passengers who change bus at the station
- **Terminating passenger**: refers to passengers who end the trip at the station.

However, the transferring passengers and terminating passengers should be excluded in the waiting time computation, since the transferring passengers who leave the operating bus line might wait for another bus. It is here believed that they have acted as the originating passengers on another line. Meanwhile, terminating passenger have already leaved the concerning line and won’t need bus service any more. Their waiting time at current station is zero since they leave the bus, while in previous stations; they had already been incorporated in the classes of through passengers or originating passengers. Thus, those two types of passengers should not be included in passengers’ waiting time estimation.

Originating passengers’ waiting time is expressed by equation 9, which is a modification of Goverde’s formula. For the bus service \( J \), the waiting time \( W \) equals the time between arrival of customer and scheduled bus arrival times, plus the bus delay (D).
\[ W_j^o = S_j + D_j \] (9)

Through passengers, waiting time is expressed by equation 10, which is also modification of Goverde’s formula. For bus service J, the waiting time W equals the waiting time at boarding station (S) plus the excess stop time of bus which is over minimum stopping time and the previous waiting times at early stations.

\[ W_j^T = S_j + E_j + \sum_{n=j-1}^{1} W_n^T + W_0^o \] (10)

Nevertheless, the passengers’ mobility is hard to trace, although there exists data on boarding passengers and traffic flow data. The calculation of individual waiting time seems impractical to implement. Thus, the formula 10 is simplified:

\[ W_j^T = S_j + E_j \] (11)

For the current study, the formulas 9 and 10 cannot be applied properly due to the difficult of tracing the passenger mobility. However, for the cities whose public transit system installed pre-store value card system combine with on-off card validation facility, a tracing of the passenger mobility seems practical. In that case, the original formula with accumulated waiting time can be applied to estimate the customers’ waiting time more accurately. Take a passenger service process as an example. The passenger onboard a bus checks in its IC card, and the on board validation facility records the passenger’ IC card and start station. When passenger gets off the bus at destination, use IC card again at debarkation facility, the destination station will be recorded. In some public bus systems, such as in Beijing, China, this method has been widely used to calculate traveling cost and the distance of individual travels for years. It appears that this method could also get the passenger routing information. There is no hard to believe that with the existence of this information, the tracing of passenger mobility and further application for customers’ waiting time is possible and practical. The cost for storage of such daily information is one crucial factors which should be considered by Traffic Company. But a periodic usage of this method for statistic need may be adopted for further research.

The waiting time of originating passengers express the waiting time on platform, while the waiting time of trough passengers express the waiting time in bus. According to formula 8, the waiting time on platform is perceived three times longer than the waiting time in the bus. The total perceived waiting time, is expressed as:

\[ W = 3 \sum W_j^o + \sum W_j^T \] (12)

Furthermore, relying on the Mohring et al.’s conclusion, waiting time at boarding station halve the scheduled interval time between buses. Assume the R is the interval time between two buses according to schedule. Therefore the \( S_j \)can be specified as the Equation 13 below.

\[ S_j = 0.5R \] (13)

24
The passenger waiting time will be an important issue to evaluate the reason of bus delay.

3.6.3 Optimisation

The optimisation consists of two aspects. The first one is optimizing new routes of buses; the second one is optimizing bus schedules. The new routes of buses are optimized through avoiding high traffic congestion segment of road. For instance, a bus line consists of line segments between the bus stations. Each segment has its own traffic congestion index. To optimize routes of buses, segment with high traffic congestion index will be re-linked to low congestion segment to create new routes. This method could reduce system delay of each bus transportation line and improve the continuity of bus transportation system. Then, optimizing bus schedule should be based on optimized bus route. The optimized bus route can give the bus transportation system more stability and continuity. Therefore, the customers’ extra waiting time due to bus delay will be decreased. If passenger capacity of a bus is exceeded, more buses should be added to service. However this excess do not relate to the entire bus line. In general, this need for additional buses only concerns some busy bus stations. Thus, extra bus will be added only to those bus stations.

4. Applying the method on Guiyang City

Bus lines No. 4, 5, 6, 9, 10, 12, 13, 18, 19, 20, 27, 30, 51, 54, 69, 202 are the major lines of Guiyang downtown area. They pass through Yunyan and Nanning region and enhance the traffic jam and bus delay every week. Therefore, this research will concentrate on those bus lines to find out underlying problems and to propose solutions. All bus lines comprise two reverse routes. Working time is from 6:00 to 22:00.

4.1 Data preparation

To follow the process of research defined earlier in this paper, the primary data must be prepared before data transformation. Therefore, the primary data were divided into three parts in accordance with characteristics of the data. Data related to the quantity of passengers are integrated as the first type of data called Passengers data. It contains:
- The number of passengers who have boarded the bus at each station.
- The number of people who leave the bus at each station.
- The quantity of passengers onboard when depart each station at each station.

Bus velocity data are calculated from on-board GPS data which will be used to calculate average bus velocity between stations. Hence, velocity data and driving time data are defined as the second type of data named On-road data. It comprises of:
- Average bus velocity between stations.
Driving time between bus stations inclusive stop times at each station.
Time tables for No. 4, 5, 6, 9, 10, 12, 13, 18, 19, 20, 27, 30, 51, 54, 69, 202 bus lines. 2007-2008.
The third data are called Traffic environment data. It is used to record traffic information which affects working of bus transportation system. It comprises of:
The traffic volume on each intersection in main area of city (they are divided by three groups.).
The velocity of vehicle on main road in the main area of city.
The location and characteristic of all Pedestrian Bridges in main area of city.
The distribution and characteristic of traffic light in main area of city.

4.2 Data transformation

Data transformation is a key step for creating the GIS-T database. A large amount of useful data can be picked from primary data. In this study, the primary bus transportation information is stored in a TIS database. This TIS database consists of the bus transportation structure and related information. The whole bus transportation network structure is based on a microscopic model. Each related factor is stored in an independent excel file. As mentioned in the introduction part, some data in TIS must be transformed to a common format compatible with GIS database. Therefore, the location of bus stations and bus lines should be geo-coded. After geo-coding, the bus transportation structure in the TIS database will be transformed to a digital map using WGS84 (world geographic system 1984) coordinate system. For the related factors data only a subset is of interest and transformed to the GIS-T database. The picked related factors data are:
The number of passengers who have boarded the bus, at each station.
The number of people who leave the bus, at each station.
Amount of passengers onboard when departure each station.
Average bus velocity between stations. (They are calculated from on-board GPS data.)
The bus real arrival time on each station.
The traffic volume on each intersection in main area of city (they are divided by three groups.).
The velocity of vehicle on main road in the main area of city.
Time tables for No. 4, 5, 6, 9, 10, 12, 13, 18, 19, 20, 27, 30, 51, 54, 69, 202 bus lines. 2007-2008

To follow the requirements of Bus Corporation, the waiting time of passengers should be reduced in order to satisfy customers in addition to this more buses on road segments where high traffic congestion occurs must be avoided. The passengers’ data are transformed to time series in excel sheets, while the on-road data are also transformed to excel sheets to calculate average bus velocity and traffic congestion during different time. The statistics of passengers’ data are focused on the number of
passengers boarding at each station. An example of such data is presented in figure 9:

![Figure 9: The Number of passengers (On and off) data, No. 4 Bus line, Bus code 2253.](image)

Generally, many passengers will appear during the rush hour. In the meantime, traffic jam also occurs. Therefore, Passengers waiting time maybe the most likely to increase during rush hour. To locate the problem using traffic simulation, the data of rush hour could be used for evaluation the weakness of the bus system. The bus capacity was here limited to 80. In this study, the number of passengers in result can be used as an indication to find the delay stations and bus line in bus system.

### 4.3 GIS-T database

The GIS-T database, which is created in the ArcGIS environment, provides detailed information to support simulations. In order to complete the database, a digital map based on the WGS84 coordinate system is needed as a background map. The digital map is produced from a GIS database at Guizhou University. Due to national security, the DEM model and basic environment data are deleted from the original digital map. Therefore, the background map only contains information on block, road, railway, water area and green area. All bus lines and bus stations are located on this digital map with WGS84 coordinates. The whole GIS-T database consists of Shapefiles and Personal Geodatabase. The Shapefiles store the location and background information, such as roads, bus stations, while the Personal Geodatabase stores relation factors information, such as passengers’ number and bus velocity. The GIS-T database consists of 47 Shapefiles and 20 Personal Geodatabases. The structure of the database is illustrated in figure 10:
Figure 10: the logical E-R model

The personal geodatabases and shapefiles are created in ArcCatalog following the structure of database. The chosen relation factors data are imported into geodatabases using corresponding field names. The 47 shapefiles consists of 19 point shapefiles, 25 polyline shapefiles and 3 polygon shapefiles. The point shapefiles show Bus stations and traffic volume recording points. The bus stations are presented by colored cycles. The traffic volume recording points are presented by colored triangles. The color varies according to different bus lines. Polyline shapefiles show bus lines and other objects, such as road, green area, water area and block area are presented by polygon shapefiles. The final visualization is presented by figure 11:
The next step is to form the relationship between the shapefiles and geodatabases. By following the structure of GIS-T database, all personal geodatabases are linked to the attribute_table of the shapefiles and geodatabase. A segment of bus stations and traffic volume information is tabled as figure 12 and figure 13 (Detailed attribute tables will be presented in the Appendix):

<table>
<thead>
<tr>
<th>OBJECTID</th>
<th>Name</th>
<th>Direction</th>
<th>Number</th>
<th>Time</th>
<th>On</th>
<th>Off</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2253</td>
<td>0</td>
<td>1</td>
<td>7:38</td>
<td>9:10</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>2253</td>
<td>1</td>
<td>2</td>
<td>9:10</td>
<td>11:21</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>2253</td>
<td>2</td>
<td>3</td>
<td>11:21</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>2253</td>
<td>3</td>
<td>4</td>
<td>13:11</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>2253</td>
<td>4</td>
<td>5</td>
<td>15:22</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>2253</td>
<td>5</td>
<td>6</td>
<td>17:13</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>2253</td>
<td>6</td>
<td>7</td>
<td>19:20</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>2253</td>
<td>7</td>
<td>8</td>
<td>21:22</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>2253</td>
<td>8</td>
<td>9</td>
<td>23:47</td>
<td>1:00</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

Figure 12: Attributes of Bus-station
3.4 Simulation

The computer simulation is implemented using Vissim software and based on data export from the database. The simulation system consists of two parts, the bus station and road network. The data used for the bus station part are the passengers’ data. For the road network part, traffic volume, base map, average speed and scheduled bus are used in the simulation.

4.4.1 Simulation design (Traffic flow model)

The road traffic situation consists of several parts such as the scaled road network, the bus line, the passengers and stops, the traffic factors of each road segments and the traffic composition which has interaction with the bus operation states, as figure 14 shown.

![Simulation diagram](image)

**Figure 14:** Composition of the on road simulation

To improve the accuracy, the road network is adjusted to conform to the real length of each road segment. Boarding Passenger, leaving passengers and waiting passenger at each station is set according to the database. The traffic volume and average speed for each type of vehicle are also set according to the database. The same applies for the road district on certain road.
Figure 15 gives an image of the simulation network on top of the background map. When resolved into basic components, the network is composed of 260 bus stops, 10700 road segments with certain speed restriction, 16 bus line operated according to the time schedule and passengers input at each bus station. The input data were processed from the original data of Guiyang public traffic Company. Among different periods of a day, the peak hours (mainly from 18:00 to 20:00) are selected for the simulations, since the worst traffic congestion and overloading of customers occurs at that time. For comparison, Data between 16:00 to 17:00 and 21:00 to 22:00 are also included in the simulations. The passengers’ quantity, the average speed of each road segments and other data were processed by the method mentioned in previous sections. With the help of formulas 1, 2 and 3, the congestion of delay was evaluated. On the other hand, a user could have a direct visualization and cognition about the whole system operation during the simulation. It is convenient to study the performance of a single bus line in detail. However, under normal circumstance, searching among huge data sets and calculation of the congested bus lines is quite a tough work.

In Figure 16, the segment inside the black rectangle is a bus station. The grey lines joining the station are road segments as white rectangle shown. The cyan block marked with red arrow is the transportation bus studied in the system, while the small white, blue and black blocks marked with purple arrows, are other small vehicles components such as sedan cars in traffic system. The red block marked with green arrow is the large vehicle like truck in the system. The road network is designed according to map file exported from ESRI Arc map with 1:38500 map scale.

For each of the 10700 road segments, the speed restrictions are set together with the real velocity data collected by inspectors. Since the speed data of each road
segment reflects the general traffic situation, setting the speed restriction on each road segments is important for a reliable simulation. Each station’s passenger flows volume per hour was also entered. When two different bus lines share the same bus stop, the individual passenger input data and alighting rate were arranged for each, just as figure 17 shown. The traffic flow input is designed according to the real data, comprised of taxi, bus, truck, coach. Moreover, the bus service rate was set based on the real bus schedule. Thus, the whole simulation network has been constructed completely with real data input.

The network was tested for two simulation hours, and vehicle information was collected for output. During the simulation, the time resolution for collecting data is set to 1 time steps, which equals to approximately 10s in reality. For output, the numbers of waiting passengers, the occupancy of vehicle, the average waiting time for passenger at each station were collected.

During the simulation, the congest road were extracted from the result. By analyzing in latter part, the performance of system will be evaluated. The total delay is the main result of the simulation. It indicates the overall performance of the public transit network.
5. Results and analysis

The output of the simulation runs is formatted as a text file. The file contains several crucial factors, such as waiting passengers, occupancy, alighting passengers, boarding passenger and passenger waiting time at current stop. The data were collected for every time step of simulation time, which equals to approximately 10s in reality.

The simulation result is a large text file with status information for each time step. The objective of the simulation is to emphasize the weakness of the transit system and to identify improvements. One task is to identify the week bus stations. Figure 18 shows the bus station with at least 40 waiting passengers. The horizontal axis indicates the bus station number, while the vertical axis means the passenger’s amount. The occupancy means passengers in the bus when it approaches the station. Besides, waiting passengers stand for passengers waiting at the stop when the bus stops at the station.

Even if the vehicle is highly occupied and there are many passengers on station, the vehicle won’t necessarily be overload. Only if the leaving passengers are as many as possible to empty the space on vehicle for waiting passengers who intend to get on vehicle. Thus the alighting passenger and boarding passengers’ records, which is shown as figure 19, are also necessary for evaluation of the result.

![Waiting passenger&Occupancy](image)

**Figure 18:** The waiting passenger and occupancy at concerning stops

The relationship between Alighting and Boarding passengers can be found out in figure 19. As in previous figures, the horizontal axis represents the bus stop number and corresponding vehicle number, while the vertical axis is the passenger’s quantity.
The calculation of the delay is a time consuming job when the transportation system consists of a large number of stations and complex bus lines distribution. The simulation has been done first to locate which bus line of bus system has the delay problem as it has necessity of simplifying the computation procedure.

As discussed before, the bus delay is caused by traffic congestion and overload passengers. Reviewing simulation result, it is not difficult to figure out that overload problem exists at several bus stops. Figure 19 indicates that the bus station No.151, No.100, No.80, No.53, No.234, No.101, No.22, No.83, No.41 and No.80 all have over 40 passengers waiting at a bus station, while at least 60 passengers are already on the buses. Since the capacity of bus is 80 passengers and the spare seats on vehicle are less than waiting passengers on stations, bus overload problem could happen. For the delay calculations in the next step, it is necessary to identify which bus lines are suffering the overload problem. According to check list table 1, these bus stops belong to bus line: No.9, No.10, No.13, No.18, No.19, and No.69 respectively.

![Figure 19: The Alighting and Boarding passengers at concerning stops](image-url)

<table>
<thead>
<tr>
<th>Simulation time</th>
<th>Veh No.</th>
<th>Stop No.</th>
<th>Bus Line No.</th>
<th>Bus Line Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1834</td>
<td>8</td>
<td>151</td>
<td>9</td>
<td>No.19</td>
</tr>
<tr>
<td>2001</td>
<td>12</td>
<td>100</td>
<td>5</td>
<td>No.10</td>
</tr>
<tr>
<td>2027</td>
<td>13</td>
<td>80</td>
<td>4</td>
<td>No.9</td>
</tr>
<tr>
<td>2223</td>
<td>14</td>
<td>53</td>
<td>7</td>
<td>No.13</td>
</tr>
<tr>
<td>2361</td>
<td>13</td>
<td>234</td>
<td>4</td>
<td>No.9</td>
</tr>
<tr>
<td>2435</td>
<td>12</td>
<td>101</td>
<td>5</td>
<td>No.10</td>
</tr>
<tr>
<td>2500</td>
<td>16</td>
<td>16</td>
<td>8</td>
<td>No.18</td>
</tr>
<tr>
<td>2520</td>
<td>8</td>
<td>221</td>
<td>9</td>
<td>No.19</td>
</tr>
</tbody>
</table>
Table 2 The data for figure 19

<table>
<thead>
<tr>
<th>Vehicle No.</th>
<th>Station No.</th>
<th>Occupancy</th>
<th>Boarding passengers</th>
<th>Leaving passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>151</td>
<td>70</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>100</td>
<td>79</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>80</td>
<td>80</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>53</td>
<td>68</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>234</td>
<td>80</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>101</td>
<td>67</td>
<td>55</td>
<td>67</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>61</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>221</td>
<td>80</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>13</td>
<td>83</td>
<td>80</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>80</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>1179</td>
<td>80</td>
<td>80</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>

If leaving passengers is equal to or larger than waiting passengers at a station, the bus will have enough seats and could function without overloading problem. In this study, such a correction has also been made.

A quantitative analysis of the relations between alighting passengers, boarding passenger and occupancy is required, Figure 19. Although these vehicles face overload problem, the situation is not exactly same for all of them. These eleven vehicles can be divided into two types: the potentially unstable and overload. In case when alighting passengers are more than boarding passengers, the bus still has places for downstream passengers. The operation situation of the vehicle is then defined as potentially unstable. The number of waiting passengers on downstream stations could then decide whether empty seats on vehicles are available or not. When the alighting passengers are not enough to empty the bus for the waiting passengers at the current station, the vehicle is defined as overload. The bus cannot fulfill the need of passengers on current station. In figure 19 and table 2, the vehicle 12 of No.10 stop at station 101 belongs to first category, potentially unstable, while the others belong to second. Although vehicle No.12 had 67 passengers when it approached the station, all of them had alighted at the current stop. Furthermore, although 55 boarding passengers are higher than other vehicles, such as vehicle 1179 and vehicle 8, its status is more stable than those mentioned ones. There still exist 25 empty seats on bus, whether it is overload or not depending on passenger quantity boarding at next station or succeed stations. While on other vehicles, the alighting passengers are not enough to accommodate the waiting passengers, the delay of those bus lines should be calculated.
The waiting time collected in simulation is also a criterion for system performance (figure 20).

The result corresponds to the same vehicles and stops as in figure 19-20. As illustrated, the highest average waiting time appears at No.83 bus Stop while the minimum average waiting time among these selected results is at No.100 bus stop. However, the waiting time at downstream stations can be affected by the earlier stations. Once the bus was delayed at a certain station, the delay might be reduced by fast driving. Otherwise, the delay will be passed on to downstream stations and accumulated causing an increased average waiting time there. In order to find the central tendency, the mean waiting time is calculated without the max value and minimum value involved. The result for the entire bus transportation system is 1155 second. It is close to the average waiting time of No.10 bus line. Consequently, No.10 bus line was chosen as an example in the following analysis.

The simulation software produces useful data, eg. leaving passengers, boarding passengers and waiting passengers. Problematic bus stops and delay vehicles could be filtered out of the original results. However, the simulation result only contains the information such as which vehicle are delays and how many passengers are detained on stations. The reasons for delays, whether it is traffic congestion or poor bus arrangement, cannot be clarified without further analysis. Thus the analysis in the following chapters is of great importance for evaluating the bus system and analyzing the reason for the delays.

5.1 Calculation and analysis

5.1.1 Traffic congestion

Traffic congestion is an essential parameter for bus transportation system
optimization. Each traffic congestion index has its own benefit to describe the characteristics of traffic delay. Therefore, the traffic congestion should be calculated and analyzed. To follow optimization requirements and the simulation results, the traffic congestion index calculated for No. 9, 10, 13, 18, 19, 69 bus lines’ data. The results of No. 10 bus line are presented by figure 21 to figure 23, see equation (1)-(7).

The longest delay, from Fu Yuanbei Road station to Qianlin Park station, appears at 17:50 (see figure 24). The delay is 1657 seconds, running by No. 2361 bus. For the other buses of No.10 bus line, 79% of their delay values are negative number. It is because the actual running time is lower than travelers’ expecting travel time. Indeed, the travelers’ expecting travel time could be identified as standard travel time in bus schedule, which is about 35 minutes. In some situations, the totally running time of a bus maybe lower than 35 minutes, for instance for No. 2155 bus which starts at...
6:48:54 am with only 13 passengers getting on bus and the stop time is quite fewer. Besides, since the on-road vehicles are also quite few, before the rush hour, the total running time is only 28 minutes.

The average bus velocity might be calculated from the GPS data. The GPS data record location information by latitude, longitude and time. The Instantaneous Velocity could be calculated by the differences of GPS location divided by the difference in time. In this research the instantaneous velocity is a crucial factor to compute acceleration noise and the mean velocity gradient. The results are presented by figure 24 and figure 25.

Because of more vehicles in the daytime than nighttime, low average velocities usually occurs during 10:00 to 17:30. This also appears at rush hour when buses are waiting. When the amounts of vehicles have exceeded the capacity of a road, traffic
jam happens at some segments of the road. The highest acceleration noise value is almost close to 0.3 m/s², and the lower value is only 0.011256 m/s². In general, acceleration noise during the rush hour reveals higher than those when the traffic flow is low.

The above phenomena are caused by frequent velocity changes during a segment of travel time. At rush hour, the passengers travel demand increase rapidly, the quantity of vehicle also increase with travel demand. A majority of buses change speed frequently between stopping and acceleration. On the other hand, compare with rush hour, with less traffic flow, the acceleration noise should be lower. However, all buses might stop at any station with passengers waiting. It makes the bus accelerated from a station to another.

![Figure 25: AN, MVG during of No. 10 Bus line](image)

Generally, it is called good quality of journey when acceleration noise (AN) values is lower than 0.1 or proportion of stopped time (PST) is smaller than 5%. But in this case, the PST values are higher than that for all bus lines. If these criteria express the quality of bus travel in a city, the quality of the bus transportation system in Guiyang needs improvement. However, those assessment indexes are not made for short-distance bus transportation system. The AN, PST and MVG assessment indices are good for private travel, but when they come to short-distance bus transportation system, these buses almost stopped at the same time and the same location. These frequent stops is a defect for AN, PST and MVG. These indicators are mainly useful as criteria of quality of journey between two contiguous stations. Therefore, to assess quality of bus travel in short-distance, this study does not only focus on the stopping and acceleration process but also on the passengers waiting time.

### 5.1.2 Passenger waiting time

Being an important criterion the bus system delay evaluation and optimization, the waiting time will be applied to evaluate the bus system performance. Figure 20
shows the average waiting time collected from the simulation result. The waiting time of No.10 bus line will here be analyzed in more detail. A similar analysis may also be applied to the other bus lines.

**Figure 26**: The passenger average waiting time in seconds of No.10 bus line as function of time series and bus number

The passenger waiting time according to formula 11 is shown in figure 26. All buses of No.10 bus line has been evaluated with passenger waiting time, from first bus No.2150 starting at 6:43 till the last one No.2143 starting at 22:15. It can be observed that in figure 26, the maximum passenger waiting time appears at No.2361 bus, approximately 9663 seconds. That is the accumulate sum of all passenger waiting time on every station in No.10 bus line when No.2361 bus operated.

No.2361 bus, which has maximum average waiting time, was dispatched at 17:50. Besides the maximum value, the second largest value is No.2148 bus dispatched at 7:31. Coincidentally, 7:31 am and 17:50pm are rush hour not only for office staff but also students and other people. The large number of waiting passengers could be an interaction of congested traffic and too many passengers on their way to work or school.
In order to identify the reasons for the delays, the passenger waiting time on each station of No.10 bus line is calculated for the No.2361 bus, as figure 27 shown. Accidentally, the maximum waiting time appears at the last station, Qianling Park, while the second high value is provincial hospital. After checking the passenger data, it is found that there is no passenger boarding at the last station. And very few of them are boarding at penultimate station. The average value is only 1 passenger for a whole day, just 1.39% of the boarding passenger in the whole operation process. However, there is a large number of passengers alighting at the last station and penultimate station. At penultimate station, it is about 13% of total alighting passengers, the second highest alighting rate of the whole line. The alighting rate at the last station is 8.15%. Due to fewer boarding passengers and high alighting rate, this is the reason to believe that the passenger waiting time at the last station takes large portion of responsibility for high passenger waiting time. The traffic congestion might be the main cause for delay. However, traffic congestion calculated in previous subsection and passenger waiting time mentioned here are still two separate criteria. At the moment, in order to combine them and complement each other, a new method combining both is necessary.

5.1.3 Two-factor delay assessment

The bus transportation system delay time is easy to count by comparing schedule time and real running time. However, the reason of delay is not clear. As mentioned above, the whole bus transportation process consists of a traffic flow part and on a bus station stopping part. The transportation delay is also caused by obstacles in those two parts. Based on this idea, the calculation model for short-distance bus transportation system delay can be described that total delay time is equal to traffic flow delay time plus bus station stop delay. This calculation model is called the two-factor assessment equation:
Figure 28: The diagram of bus line structure

Figure 28 shows a bus line where circles represent bus stations and solid lines represent the bus route. There are two aspects to explain bus delay according to the analysis. One is the driving delay and the other is the waiting delay at each bus station. Driving delay results in late arrival time and it would cause more passengers waiting at a bus stop. Moreover, more passengers would bring extra boarding time of a bus, which is around 5s per person based on the testing results. As a result, the extra time accumulated in a bus stop could arouse more waiting delay to next stops which is like chain effect. Sometimes no traffic jam could decrease the time delay and in an ideal situation, it could entirely compensate for the delay. Therefore, the calculation model can be defined as follows equations:

\[
T_{ai} = \frac{D_i}{V_i} \quad \text{(14)}
\]

\[
DTR_i = T_{ai} - T_i \quad \text{(15)}
\]

\[
PND_i = \frac{DTR_i}{P_{pi}} \quad \text{(16)}
\]

\[
ST_i = PND_i \times T_{as} \quad \text{(17)}
\]

\[
DTT = ST_{i-1} + DTR_i \quad \text{(18)}
\]

Where,
- \(T_{ai}\): Standard travel time between (i-1) and i, unit s. (Time i)
- \(D_i\): Distance between (i-1) and i, unit m. (Distance i)
- \(V_i\): Average velocity between (i-1) and i, unit m/s.
- \(DTR_i\): Driving delay between (i-1) and i, unit s. (Delay time of running i)
- \(T_i\): Actual travel time between (i-1) and i, unit s. (Actual travel time)
- \(PND_i\): Excess passengers caused by bus delay from (i-1) to i. (passenger number of delay)
TPi: Passenger inter-arrival time at station i. (passenger inter-arrival time i)
STi: Extra stop time caused by traffic delay from station (i-1) to station i. (stop time i)
Tas: Average check-in system service time when passengers are boarding.
DTT: Total delay time.

Those equations can be integrated as follow equation:

\[
F(i)=\sum_{n=1}^{i} ST_n = \frac{(D_1 - T_1) \times Tas}{TP_1 \times TP_{i-1} \times TP_{i-2} \times \ldots \times TP_1} + \frac{(D_2 - T_2) \times Tas^{i-1}}{TP_1 \times TP_{i-1} \times TP_{i-2} \times \ldots \times TP_2} + \frac{(D_3 - T_3) \times Tas^{i-2}}{TP_1 \times TP_{i-1} \times TP_{i-2} \times \ldots \times TP_3} + \ldots + \frac{(D_j - T_j) \times Tas}{TP_i} \tag{19}
\]

It can be fixed as:

\[
F(i)=\sum_{j=1}^{i} \frac{(D_j - T_j) \times Tas^{i+1-j}}{\prod_{i=j}^{i} TP_i} \tag{20}
\]

Therefore, the total delay time can be defined as follow:

\[
DTT = \sum_{j=1}^{i-1} \frac{(D_j - T_j) \times Tas^{i-j}}{\prod_{i=j}^{i} TP_i} + DTR_i \tag{21}
\]

On the basis of simulation result and two factors formula introduced above, the delay of bus line 9, 10, 13, 18, 19, 69 are recalculated. The results are shown in figure 29.
Figure 29: bus line delay of No.9, No.10, No.13, No.18, No.1, and No.69.

The vertical axis represents the delay in seconds while the horizontal axis not only stands for vehicle number, but also been reordered according to time sequence. It is evident that, among these six lines, most of them have peak values in the latter half of the route. When checking with the bus schedule, the peak values’ occurs is around 18:00. What’s more, some of the bus lines has another delay clusters in the first half, around the 7:30 in the morning. Upon that, a tendency can be observed that the delays cluster around the rush hour in the morning as well as rush hour in the afternoon. Because of the clustering of delays around the rush hour, it is probably the traffic congestion that is the main reason for traffic delay.
The No.10 bus line was chosen again as an example for analysis of delays at the station. As table 3 shown, the maximum delay of No.10 bus line, approximately 2476 seconds, appears at No2361 bus, starting at 17:50; while second is No.5306, 1207 seconds, which starts at 7:16am. For comparison, the delay of No.2148 bus whose delay is 1070 is the third largest one and is also added in the figure.

The result of using formula 21 and 13 are similar. The last station, Qianling Park, is accused for high delay of the No.2361 bus. By checking the passenger data, it is shown that there is no passenger boarding at the last station. And very few are boarding at penultimate station, only 1 passenger in average for a whole day. However, there is a large number of passengers alighting at the last station and penultimate station. At penultimate station, it is approximately 13% of the total of passengers, while the second highest alighting rate in whole line is at last station (8.15%). It is undeniable that traffic congestion has greater impact than overload passengers for the delay of this bus. In order to validate this assumption, the service time of No.10 buses was imported, graphed as figure 31. Upon that, it is distinct that, the No.2361 bus starting at 17:50 has the highest service time, 3733 seconds. this means that the bus was delayed by awful traffic congestion, since over 90% of others are less than 3000 seconds. Due to fewer boarding passengers, high alighting rate and long service time, traffic congestions is the main reason for long delay. This also affects the passenger destined to the last station. On the other hand, the peak value of No.5306 bus and peak value of No.2148 bus appears at Yanwu Street and Guizhou Daily Agency. These two bus lines serve at the morning peak hour, and delayed stations are located at the downstream direction of central city. These two
buses might be affected by traffic congestion in central city area. The congestion postpone the service time in stations around the central city area, and as a result, prolong the total service time.

One last important principle needs being noted is that despite the No.2361 bus is different from others, the top seven buses who has high delay all started their service at rush hour in the morning and the afternoon. Traffic congestion is still the main factors for delay.

![Service Time of all No.10 bus](image)

**Figure 31**: Service time of each No.10 bus (Blue bars represent service time)

It is necessary to clarify that the No.10 bus is not the worst delayed bus line. Instead it was here chosen because it is considered as a representative bus line.

However, it is not hard to find out, common characteristics of bus lines with long delays, such as:
1. Urban lines.
2. Bus line crossing the central area of city.
3. The delay station locates at the downstream of the city central area.

## 6. Discussion

### 6.1 Simulation method

If the proposed recommendations are implemented, an improvement on traffic situation is expected.

As a method which aims at evaluating the traffic delay, the feasibility of method should be discussed.

The traffic congestion is considered as an important problem of transportation system, especially when comparing the quality of routes. Therefore, some indicators are used to represent the level of congestion, and in each segment of road and at different time
interval. However, all calculations of congestions need detailed traffic data. If such data are not available, the results of calculation would produce errors. Meanwhile, traffic congestion is related to the entire traffic net. To compare congestion during different time periods, it is better to study only one segment of the road. In this paper, all traffic data are collected from buses in Guiyang. Hence, the traffic congestion indicators cannot be applied to all vehicles, like taxi or truck. Those indicators are only able to research bus transportation of Guiyang, and transit system of other cities. For taxi dispatch system or other types of traffic system, the applicability of the data should be considered before use.

As Underwood (1968) mentioned, the acceleration noise is quite helpful to compare the quality of car travel. However it is not appropriate to use for public transportation system, such as bus and taxi transportation system. The reason is that velocity and stop time of public transport vehicles are decided by amount of passengers and direction of passengers rather than vehicles on-road situation. Therefore, the acceleration noise is not suitable as a indicator of travel comfort.

The optimization of passenger waiting time would sufficiently improve the customers’ satisfactory. However, due to the limitation of data, there still exist gaps between the reality and the simulation system. Some conditions in practice have not been involved in simulation system, such as pedestrians and the traffic light at cross roads. Obviously, the inclusion of such variables is believed to increase the accuracy. However, data for these variables is hard to collect. The Goverde’s waiting time formula was firstly applied to calculate the railway passenger waiting time. Besides, it can also be utilized to trace the passenger mobility in railway system. Due to system’s characteristic difference between the railway and bus system, such data are hard to collect in bus system. Passenger waiting time is still a fundamental element to evaluate bus delay. In the analysis, the waiting time is weighted more than traffic congestion in bus delay. That does not mean the waiting time is more significant than traffic congestions. Even though the weight varies with the different application, the approach proposed in the thesis is adaptive for various weightings. It could state that the two-factor evaluation method mentioned in thesis is applicable for other transit system in different cities which has similar structure. However, the feasibility on certain case should be evaluated according to real situation.

6.2 impact of traffic congestion and overload of passengers

Both the traffic congestion and overload of passengers contribute to the delay of the bus. When causing bus delay, these two factors are acting and interacting with each at the same time. For instance, traffic congestions on certain road may cause the late arrival of the bus, waiting passengers to be accumulated. Hence the more passengers accumulated, the more time will be taken for boarding. This extra time of boarding, plus with the former delay time, will continuously aggregate to the following stations. These factors of the traffic congestion and overload passengers are interacting. Nevertheless, they are not playing an equal role in causing the delay time. In that sense, the simulations may clarify if the delay situation is caused mainly by the
traffic congestion or by the overload passengers.

Consider figure 28 as an example. Assume that the bus suffers delay when traveling from station 1 to station 2 (DDT_2). Assume the result of station 3 is DDT_3.

If DDT_3 − DDT_2 ≤ 0, it implies that the traffic condition of the road linking station2 and station 3 is in ideal situation, it compensates the delay inherited from preceding station. Thus the delay reason in station 3 is mainly because of overload passengers.

If DDT_3 − DDT_2 > 0. And DTR_3/DDT_3>ST_3/DDT_3. Then the traffic congestion is the main reason for the delay.

Consequently, the impact of traffic congestion and overload passenger can be evaluated using the method proposed in this thesis.

6.3 applicability and algorithm complexity

Applying traffic congestion or passenger waiting time, the delay at stations and on bus lines can be calculated. It is not difficult to find out that neither traffic congestion nor passenger waiting time can specify the reason of delay, which this two-factor method can. For other transit system in different cities which has similar structure, two-factor method could be applied for evaluating bus delay. It could even be applied for railway transportation system. However, for other kind of transportation system which has random time-table and routine, such as taxi transportation system, the two-factor method is not suitable for evaluation.

The two-factor method has the advantage of finding the reason of delay. But concerning the algorithm complexity, two-factor method has some weakness compare to passenger waiting time methods.

Asymptotic Notation (also called Big O notation) may be used to compare algorithms. For example, the algorithm complexity of the passenger waiting time method isO(N^2). The two-factor method’s is alsoO(N^2). Consequently, the two-factor method does not have advantage in computation. For different application, which does not require specifying the reason of delay, the choice between two-factor methods or other methods such as passenger waiting time, should be evaluated according to real situation.

7. Conclusion

This paper describes an application based on GIS-T database and bus line simulation. GPS data play significant role in this application. They are not only supplementary data for GIS-T database but also essential sources of traffic congestion evaluation. The optimization of bus lines in Guiyang city may be based on simulation tools, operated in two parts, the passenger simulation and the road traffic simulation. Two crucial factors were applied to evaluate the system operation situation. The passenger waiting time has been used for evaluating service efficiency of buses. Combined with data on traffic congestion and road condition, it is also used for optimization and
decision making.

This thesis proposes an approach to evaluate and estimate the bus delays by two elements: delay due to traffic congestions and delay because of overload passenger. Those two elements represent the interaction of external and internal factors on bus operation. Even for the bus system with multiple bus lines, the interaction between bus lines could also be included into the evaluation. The weight of the two factors could be changed for different applications. The result of approach could be useful for bus operation quality evaluation and bus delay problem solution for bus transportation companies. And the data employed in the method are common data for any bus transportation company. Further, these data are easy to get. The evaluation on bus delay with this approach is low cost and time saving.

The following conclusions were made after analysis.

- Firstly, in database, the representation of crucial factors for bus delay is implemented through two types of data: attribute data and spatial data. Both geographical location and other characteristics of bus delay can be represented.
- Furthermore, both traffic congestion and passenger waiting data can describe the bus delays. However, the two-factor method may also specify the reason of delay by comparing the result of adjacent station. The proportion of traffic congestion and overload passenger in bus delay could be examined.
- Finally, the two-factor evaluation method discussed in thesis is applicable for other transit system in different cities which has similar structure. It could even apply for railway transportation system. However, for other kind of transportation system which has random time-table and routine, such as taxi transportation system, two-factor method is not suitable for evaluation. Besides, concerning the algorithm complexity, two-factor method passenger waiting time method does not have advantage among three methods. The feasibility of method should be evaluated according to real situation.

Based on the simulation results, the following recommendations are given:

1. Rearrange the bus lines which are passing the city central area. In order to save the cost, a simple way is to divide the bus lines into two parts at the city central area. For example, the bus line 10 may be divided into two parts at the Guizhou Daily Agency, which is the nearest station after city center. Thus the bus delay accumulated at first half and central area will not pass to the latter half.

2. Set restraints on heavy trucks and light trucks at suburb area during the peak hour. The traffic congestion in city central area is mainly caused by heavy traffic flow. To ease the traffic congestion in city central area, the traffic flow should be controlled from suburb area to decrease the whole traffic flow in city.

3. Establish other restraint such as allows only vehicles with even-number license plates on the streets on alternating days with odd-numbered vehicles can be enforced if other methods cannot optimize the situation. (after consult to Traffic control rules during the Beijing Olympic Games )
8. Further research

For current data and situation, the formula for passenger waiting time is simplified due to the difficulty of tracing the passenger mobility. However, in cities where public transit system have installed a value card system combined with on-off card validation facility, the trace of the passenger mobility seems practical. Information of passenger mobility is not only useful for calculating the waiting time, but also helpful to design the travel analysis zone and constructing forecasting model which are useful for multi-line optimization system. The simulation model described in this thesis is a little far from reality; whereas, data like pedestrian and the signal control at cross road makes the simulation system more complex, it may improve the quality of the simulation system. With more bus line data and the help of the forecasting model as well as choice predicting model, the optimization of whole system may be achieve. Besides, the census data, the traffic flow statistics, the travel demand data are not only useful for multi-line system, but necessary for analysis the exact delay reason as well. For example, if the census data and travel demand statistics are available, the passengers’ mobility can be analyzed, and then the optimization can take the travel demand as one of the criterion for consideration. If data about the accommodation area is available, the reasons for the delays may be identified.
References


Richardson, A.J. and Taylor, M.A.P., (1978). Travel time variability on commuter


