Chapter 7 Transportation Analysis

Transportation analysis is the last part of the four planning analytical methods covered in this book. Various economic activities for a given population in an area occur at different locations, which are associated with different land uses. The interaction of those activities requires a network to connect places for moving people and goods. The function for such a system is the focus of transportation analysis. Transportation analysis provides the basis for transportation planning. Transportation planning is a process of finding feasible alternatives and components of a transportation system to support human activities in a community. A transportation system consists of many different subsystems to accommodate different modes of transportation. Transportation is a broad category which includes air, water, and land transportation systems. The land-based system includes motor vehicles, pedestrians, bicycles, and rail and public transits. In addition to the modes of transportation, a transportation system consists of networks such as roads, and supporting facilities, such as traffic lights.

A comprehensive plan can give indication of the future land uses in a community. Transportation planners estimate the amount of traffic associated with the planned land use allocation, the options of travel modes, the alternative routes, and the required roadway features to support the estimated traffic volume. During the process of evaluating the various transportation alternatives required to meet the future demand, planners must consider community characteristics, available funds, environmental impacts, and other factors. Transportation engineers design appropriate transportation systems to support a community's desired mobility. For example, after traffic volume increases and congestion occurs in a community, a transportation engineer can design additional roadway lanes needed to accommodate the additional traffic volume while maintaining the original travel speed.

The most critical challenge for transportation engineers and planners is the dynamic feature of a transportation system, both spatially and over time. Transportation facility and traffic volume are constantly affecting each other. Unlike housing development, where a house or an apartment is built for one family and the demand for housing can be met by building more houses, roads are built and shared among travelers. On any particular road, bad traffic indicates that more people are traveling on the road than has capacity to handle. One common practice to improve traffic conditions is to increase the road capacity, such as adding a new lane to an existing road. However, this increased capacity can effectively relieve the traffic congestion only for a short period. More people start to enjoy the easy travel from the road expansion and soon the traffic

congestion occurs again. The theory of induced traffic explains this phenomenon as the increases in the carrying capacity of a road attracts more vehicle traffic to the road (Norland, 2001).

People's travel choices can be summarized as the time of departure, travel mode, and route. Many people travel to work within a short time period, called morning rush hours. Because many work places are close to each other in areas such as Central Business Districts (CBD), traffic volume tends to increase on roads that lead to the work places. People may choose to drive private cars or take public transit to work. People's decisions regarding the three choices are often based on the comfort level, convenience, flexibility, privacy, and travel time. In the United States, the majority of the travelers drive private vehicles for those reasons. The trend in China shows similar pattern as automobiles become affordable to more and more people. Although effort has been made to have different work hours, the nature of business determines the vast majority of businesses will have similar working hours. Once on the road, travelers normally want to reach their destination quickly. This leads to the route choice as the only major factor affecting the traffic volume. Travelers can easily switch routes on the way. Radio stations in many metropolitan areas report road traffic conditions and suggest alternative routes.

Assume a CBD is connected by a highway and a local street, more people would choose the highway in order to avoid traffic lights and to able to travel at higher speeds. As more people get on the highway, the vehicle moving speed will decrease. Eventually, there will be no difference in travel time between traveling on a highway and the local streets, which indicates that traffic has reached equilibrium.

Now assume a new lane is added to the highway to solve the highway congestion problem, the immediate outcome is that travelers on the highway will be able to move faster. However, people who travel on the local streets realize that they may travel faster on highway and, consequently, switch their route to the highway. As a result of this switch, the travel speed on the highway will decrease. If this is the only consequence we may still expect that the new lane has relieved the traffic congestion to some degree. In reality, those who use public transit may realize the highway improvement switch to driving private vehicles. And those who leave to work earlier or later than their preferred time to avoid congestion may switch back to their normal time. This switch of routes, modes, and time is called "triple convergence" (Downs, 2004). According to the triple convergence principle, increasing the roadway capacity does not alleviate traffic congestion during the rush hours unless the roadway capacity is increased to the level that can accommodate all the traffic, which is spatially and financially impossible for many metropolitan areas. In short, the net affect of roadway improvement is that the improved travel condition induces more trips during the rush hours. This dynamic phenomenon presents a big challenge to

transportation planners.

Detailed discussion of transportation engineering and transportation planning goes beyond the scope of this book. Students who are interested in further study should look into transportation courses offered in planning and transportation engineering programs. The goal of this chapter is to introduce the fundamental concepts and calculations in transportation analysis. You will be able to understand the travel demand modeling and its applications in planning.

7.1 Basic Concepts in Transportation Analysis

Let us begin by reviewing some terms commonly used in transportation analyses, using Fig. 7.1 as an example. A transportation study analyzes the traffic conditions on road networks. A road network is a special application of the "network flow problems", which is part of linear programming theory. The mathematical base of network analysis is to determine a static maximal flow from one point to another in a network, subject to capacity limitations of the network (Ford and Fulkerson, 1962). A road consists of multiple segments and one road connects to other roads at intersections. The collective features of each road segment represent the overall traffic conditions in a region.

Figure 7.1 Illustration of a street network

A **street network** refers to all the surface roads that are connected to each other and to different places of human activities. A street network consists of segments and nodes. A network may be a real geometric representation of roadways or straight lines connecting the nodes.

A **node** is an intersection where two or more streets are connected or the end of a street. For example, in Fig. 7.1, four street intersections are labeled as nodes 1 through 4.

A **segment** is a line connecting two nodes. The traffic on a segment remains same and may only change from one segment to another. Figure 7.1 shows four complete segments (labeled as 920,1070,1019, and 1131).

A **link** is a segment associated with direction. If we use letter *L* to represent a link, the link from node *i* to nod *j* is often represented as L_j . A one-way street is represented as a link with only one direction. A link representing a two-way

street will have traffic data for two opposite directions. Therefore, a street network segment may be represented as L_{ij} and L_{ji} . A link is the smallest unit of

analysis in transportation studies. Many of the basic features of a street network are associated to links. Variables normally describe link-based traffic include design capacity, design speed, number of lanes, traffic volume, and actual travel speed.

A **chain** is a series of connected links directed the same way. The travel from a node *i* to a node *j* may go through a link or a chain of links.

A study area is divided into areas, instead of points. Those areas are called **Traffic Analysis Zones (TAZs)**. A traffic analysis zone is delineated as the smallest area of the study region. Although there may be numerous residential locations in a TAZ, to include each location in the model would be rather cumbersome in practice. Therefore, travelers within a TAZ are treated in the same way as an aggregated group (Oppenheim, 1995). Although people living in the same TAZ may access to the street network at many different nodes, transportation studies treat all traffic from a **centroid** of the zone. Traffic generated from a TAZ or end at a TAZ is connected to street networks through one or more **connectors**. One end of a connector is a node on the street network. The other end is the centroid of a TAZ. Connectors may not be real roads. Figure 7.2 illustrates TAZs and their connectors to the street network.

A **trip** is normally the focus of a transportation analysis. It represents the path people make from one place to another, for instance, from home to office. One type of trip is a **vehicle trip***—*the number of automobile trips traveling in a transportation system. Another type of trips is a **person trip***—*the number of people traveling through the transportation system. When there is more than one passenger in a vehicle it becomes necessary to distinguish the two. In this case, an estimation of number of people per vehicle is required to convert person trips to vehicle trips.

The two places connecting a trip are called **trip ends**. Trip ends can be further divided into two categories when trip direction is considered. The trip end at the beginning of a trip is called **origin** and the trip end at the end of a trip is called **destination**.

Figure 7.2 Traffic analysis zones and a traffic network

Travel Time Index measures the additional time for a peak hour trip when it is compared with the same trip during non-peak hours. It is expressed as the ratio of the peak hour trip time and the non-peak hour trip time. For example, the trip from my house to office takes 40 minutes during the peak hours and 25 minutes during non-peak hours. The travel time index is $40/25 = 1.6$.

Traffic flows between TAZs are normally expressed in an **origin-destination (O-D) matrix.** Table 7.1 illustrates an O-D matrix for a study area of four TAZs. Reading horizontally, the matrix shows that the traffic volume generated from Zone 1 is 11,774, among which 4,340 trips go to other parts of the same zone. Three other numbers represent the trips from Zone 1 to Zones 2, 3 and 4. For example, 3,180 trips go from Zone 1 to Zone 2. Reading vertically, the matrix shows the number of trips ending in each zone. For instance, 8,220 trips originate in Zone 2 and end in Zone 1. Numbers in the last column are trips that start from each origin zone. Numbers in the last row represent trips ending in each destination zone. The matrix shows that Zone 1 generates fewer trips than the other three zones. Zone 2 receives more trips than other zones. This could be an indication that Zone 2 may be dominated by industries or shopping centers that make Zone 2 the major employment center.

To From		$\overline{2}$	3	4	Total Trip Origins
	4,340	3,180	1,769	2,485	11,774
2	8,220	16,493	5,804	10,525	41,042
3	4,954	6.287	6,828	5,179	23,247
$\overline{4}$	8.989	14,727	6,691	20,347	50,754
Total Trips	26,502	40,687	21,092	38,536	126,817

Table 7.1 An O-D matrix for a hypothetical study area of four TAZs

Design capacity is the maximum number of vehicles that can pass the end of a link within a given time period without causing traffic delay. It is measured as number of vehicles per hour, such as 1,000 vehicles per hour.

Design speed is the maximum travel speed for a given link when there is no $delay^{\circ}$. A design speed reflects the function of a road and is normally limited by physical, social, economic, and aesthetic conditions. For example, the design speed of a link on steep slope is usually lower than the design speed of a road on flat land due to safety concerns. The design speed is also related to road functions. A local road passing through residential areas is likely to have a lower design speed than a highway. In Fig. 7.1, link 1070 has larger design capacity and faster design speed than link 1019. This is an indication that link 1070 represents a major road while link 1019 represents a local road.

Number of lanes represents the lanes available for travel. The number of lanes can be the total lanes for both direction or be counted by the travel direction.

Volume is the actual number of vehicles going through the link within a given time period. The volume can be measured as daily volume (24-hours) or one-hour volume. The one-hour volume is usually used to represent the traffic during peak hours.

Average Daily Traffic (ADT) represents the typical daily traffic volume for a link. In transportation planning analysis, traffic volume data are normally collected during a long period in order to calculate ADT. The daily traffic can further be divided into weekday and weekend volume.

Average Peak Hour Traffic (PHV) can be calculated when traffic data are collected only during peak hours on multiple days and the average is calculated from the data.

Vehicle occupancy is the number of people traveling together in one mobile vehicle. It is normally calculated as the number of travelers divided by the number of traveling vehicles in a geographic area such as a traffic analysis zone, or a region.

 \overline{a}

ķ A Policy on Geometric Design of Highways and Streets, American Association of State Highway and Transportation Officials, Washington, DC 1994.

In transportation analysis trips are normally classified by **trip purpose** based on the location of the origin and destination.

(1) If one of the trip ends is home and the other trip end is a workplace, the trip purpose is defined as **home based work** (HBW). A typical HBW trip is the trip from home to work in the morning or going home from work in the evening.

(2) If one of the trip ends is home and the other trip end is not a workplace, the trip purpose is **home-based-non-work** (HBNW). Sometimes this trip purpose is called **home based other** (HBO). When a family goes to a restaurant for dinner, the trip purpose is qualified as HBO.

(3) If neither of the two trip ends is home, the trip is defined as **Non-home based** (NHB) purpose. A good example of a NHB trip is for a person to go shopping at lunch break. One trip end is the office and the other trip end is the store.

(4) Two trips between two trip ends (back and forth) are called a **round trip**. For example, a person goes to work from home in the morning and returns home in the evening will count as one round trip, consisting of two trips.

(5) The above discussed trips can be called simple trips. As the trip making behavior becomes more complex, people may make intermittent stops on a trip. **Trip chaining** is the succession of trip segments (Hensher and Reyes, 2000).

Figure 7.3 illustrates the different trip terms. It shows three round trips. The round trip on the left consists of two HBN trips. The round trip in the middle has two HBW trips. The trip from WORK to BANK then HOME is an example of trip chain. The round trip on the right is made up of two NHB trips.

Figure 7.3 Illustration of trips

7.2 Overview of Transportation Analysis

In general, transportation planning consists of two tasks. The first task is to estimate the traffic flow based on population and economic activities. Human activities are normally reflected in land use composition. The second task is to evaluate the social, economic, and environmental impacts of transportation projects.

Traffic demand analysis (freight and passenger) aims to derive the traffic volumes of a traffic network. The volumes are calculated from a set of origin-destination matrices related to a proper traffic analysis zone subdivision. Computer models are often developed to simulate traffic demands. Most of these types of models work best on a corridor or regional scale for planning purposes. They are not designed for estimating traffic volume on a particular street.

The purpose of a traffic impact analysis is to identify the traffic-related consequences of a proposed development, such as a new commercial center. Normally, the outcome of the analysis is recommendation for minimizing undesirable impacts. In order to do so, the analysis will determine specific traffic volumes associated with a proposed development, the capacity of existing transportation systems to absorb additional traffic volumes, the significant traffic impacts, and possible mitigation measures for those impacts. The traffic impact study results are used to identify and assist in the design of specific transportation improvements required for a project. These improvement requirements are normally incorporated into the conditions of the project approval.

While transportation project is intended to improve mobility, it also incurs implementation costs. The comparison of costs and benefits of different alternatives is often used as the basis for selecting a desirable course of action. A commonly used measure is the **ratio of benefits to costs** (B/C ratio). If the ratio of a project is greater than one, the project is expected to have greater benefits than costs. Everything else being equal, the project with the higher ratio would most likely be the preferable choice.

The challenge of a cost-benefit analysis is the measurement of cost and benefit. Costs may be monetary or non-monetary. The costs to build or maintain a street are an example of monetary costs. The costs incurred from automobile accidents and congestion are examples of non-monetary costs. In addition, a transportation project may affect the land use pattern, air quality, and quality of life in the surrounding areas, which may also be non-monetary costs and/or benefits. Examples of benefits include increased efficiency of the transportation system, positive economic development impacts, and increased real estate values. In general, a Benefit-Cost Analysis (BCA) identifies alternative projects that have positive net social benefits and then selects one from the list as the preferable choice. The final selection could be solely based upon the net benefit or other additional considerations. What can complicate the Benefit-Cost Analysis is that the costs and benefits may be short-term or occur over time.

7.3 Street Classification

Streets have two functions—providing access and facilitating movement. Both functions are critical for an effective transportation system. **Access** refers to connecting the points of interest to the street network. **Movement** means that one

can travel fast along a street. Streets with good access are easy to get on and off. Such a street does not allow high-speed travel. Streets with good movement have limited number of access points; so, traffic can move uninterrupted. Most of the time, a street is normally designed to have one primary function. Some streets are designed to move vehicles quickly and efficiently from one point to another; others are connecting to as many places as possible. At one extreme, a street used for fast moving vehicles has limited access, such as express highways. Local streets are at the other extreme, which connect to individual buildings and have low travel speed limits. This allows the transportation system to provide connections to places of human activities with a safe and smooth traffic flow.

Street classification is a method that reflects the various street functions. A local municipality may use a variation of this general street classification system to accommodate the specific local circumstance. In general, a street classification consists of the following categories:

Expressways or freeways (movement \gg access): The most important function of expressways or freeways is to provide rapid vehicular mobility between cities and major attractions. The access of these roads is limited to major regional destinations, such as airports, large shopping malls, or hospitals. The limited number of access points allows automobile travel at high speeds without much interruption between origin and designation. An expressway traveling through a city only has a few exits for connecting the city and the transportation system and with minimal delay for the through traffic.

Arterials (movement > access): The primary function of arterial streets is still to provide a high degree of vehicular mobility. However, this class of roads can connect more areas to a transportation system within an area, such as a city or a town. Once a vehicle gets on an arterial, the purpose is to enter to an express highway within a short distance or travel to a place that is not too far away. Because of its emphasis on mobility, these streets should be designed to maintain high traffic capacity.

Collectors (movement = access): Access and mobility are equally important for collector streets. The access refers to linking the interior of an area to the transportation system by providing a short travel to the nearest arterial streets. One example of the consideration of movement is the left turn lane and restricted turning movements.

Local streets (movement < access): Local streets primarily provide a high degree of access. Vehicles are constantly merging or leaving traffic along streets, as well as containing pedestrian crossings. Another feature of local streets is onstreet parking. Easy access to local streets is much more important than fast vehicle movement. In fact, most local streets impose low travel speed limits.

The U.S. Census Bureau developed a street classification system as part of its Census Feature Class Code (CFCC), based on the U.S. Geological Survey (USGS) classification code in the DLG-3 file. The CFCC is a hierarchical three-character system for linear features. The first character, A, is used for roads.

The second character is a number representing the major street category. The third character, also a number, describes sub-categories for each major category.

Table 7.2 displays the major CFCC Road categories. For example, A1 represents a "Primary highway with limited access". In this category, A11 is a "Primary road with limited access or interstate highway, unseparated"; and A12 is the "Primary road with limited access or interstate highway, unseparated, in tunnel." Most of the street classifications were field verified by census staff during field operations or through the use of aerial photography or imagery.

CFCC Category	Description				
$A0 - Road With$	Source materials do not allow determination of the road				
Category Unknown	category				
A1-Primary Highway	Interstate highways and some toll highways, which are				
With Limited Access	accessed by way of ramps and have multiple lanes of traffic.				
	The opposing traffic lanes are divided by a median strip				
A11	Primary road with limited access or interstate highway,				
	unseparated				
A12	Primary road with limited access or interstate highway,				
	unseparated, in tunnel				
A13	Primary road with limited access or interstate highway,				
	unseparated, underpassing				
A14	Primary road with limited access or interstate highway,				
	unseparated, with rail line in center				
A15	Primary road with limited access or interstate highway,				
	separated				
A16	Primary road with limited access or interstate highway,				
	separated, in tunnel				
A17	Primary road with limited access or interstate highway,				
	separated, underpassing				
A18	Primary road with limited access or interstate highway,				
	separated, with rail line in center				
A2-Primary Road	Nationally and regionally important highways that do not have				
Without Limited	limited access as required by category A1. It consists of				
Access	highways that connect cities and larger towns. A road in this				
	category must be hard-surface (concrete or asphalt). It has intersections with other roads, may be divided or undivided,				
A3-Secondary and	and have multi-lane or single-lane characteristics Highways that connect smaller towns, subdivisions, and				
Connecting Road	neighborhoods. The roads in this category are generally				
	smaller than roads in Category A2, must be hard surface				
	(concrete or asphalt), and are usually undivided with single-				
	lane characteristics. These roads usually have a local name				
	along with a route number and intersect with many other roads				
	and driveways				

Table 7.2 Major Census Feature Class Code road categories

Continued

Source: U.S. Census Bureau. 2004.

The street classification reflects the variations of road functions. The physical characteristics of streets in different categories vary significantly. The major function of roads in the A1 class is movement. Therefore the roads are wide and have limited access, such as the expressways. The length of streets between exits is long. This is consistent with people that use the A1 class streets for long distance travel, such as between cities. Travelers do not need to make frequent stops and speed is the major concern. Compared to the A1 class, roads in the A4 category are used more provide access than movement. Travel speed is less of a concern for people traveling on roads in the A4 category. The primary purpose of using these roads is to get to the point of interest. The connection to residential, commercial and work places is provided through local streets. Those streets are the most widely dispersed and span to every place of human activities. Each lower category street feeds traffic to the higher category streets. The need for higher category streets is less since there are fewer areas to be connected.

7.3.1 Level of Service

The travel quality of a road is normally expressed with the measurement of level of service (LOS). In the United States there are six LOS categories represented by the letters "A" to "F", where A is for the best traffic condition and F, the worst. In general, the LOS system reflects a user's actual travel experience in relation to desired travel condition. It can be used to measure different modes of transportation, such as automobiles, bicycles, pedestrians, and public transit. Although the scale system may be similar for all transportation modes, the measurement varies. Even for the automobile traffic alone, different measurements are normally adopted for highways and urban streets.

It should be noticed that LOS measures the quality of traffic flow, or the levels of congestion. Although traffic can be by automobile, bicycle, or public transit, the discussion below is specific to automobile traffic. In designing a street, traffic engineers use LOS as the base for selecting design parameters.

7.3.1.1 Highway Level of Service

Highway LOS is calculated according to the volume/capacity ratio as shown in Table 7.3.

LOS	Volume/Capacity Ratio	Description (TRB, 2000)
A	Less Than 60%	Free-flow operation
B	60% to Less Than 70%	Reasonably free-flow
C	70% to Less Than 80%	Flow at or near free-flow speed
D	80% to Less Than 90%	Borderline unstable
Е	90% to Less Than 100%	Operation at capacity
F	100% or Greater	Breakdown

Table 7.3 Highway level of service classification

Level-of-service A represents free-flow operations. A traveler can travel at the designed speed almost completely unimpeded. The distance between vehicles is large enough for the motorist to feel comfortable. The effect of minor incidents can be easily absorbed.

Level-of-service B describes a condition in which a motorist still can experience reasonably free flow at free-flow speeds. Although slightly restricted, the ability to maneuver and level of comfort are still high and minor incidents do not cause much delay.

Level-of-service C is the lowest level in which a motorist can move at, or close to, the free-flow speed. The ability to maneuver is noticeably restricted. Drivers may feel tense while driving due to the additional vigilance required for safe operation. Although minor incidents can still be absorbed, local deterioration in service will be substantial.

Level-of-service D depicts a condition in which motorists can experience reduced travel speeds. The ability to maneuver is severely limited. Drivers may start feeling uncomfortable physically or psychologically. The network has little capacity to absorb minor incidents.

Level-of-service E is a condition in which traffic is at its design capacity. There is no room for any disruption of traffic flow. A simple lane change may significantly affect the traffic. The level of maneuverability is extremely limited and driving is no longer a comfortable experience. Serious breakdown with extensive queuing can result from minor incidents.

Level-of-service F describes a condition of breakdowns in the traffic flow. Queuing and congestion are the norm for LOS F. The traffic flow exceeds the design capacity. There is almost no ability to maneuver and no one will feel comfortable driving.

In planning and designing a transportation network, transportation authorities set up LOS standards for each roadway segment, which specify acceptable LOS. It is critical to set up adequate LOS standards. Achieving a better level of service normally costs more for construction and maintenance. It may also require more land to be committed to transportation. In order to maintain the level of movement that matches the development level of a community within the affordable budget, planners and decision makers must carefully select a proper LOS for new transportation, as well as improvement of the existing transportation system. For example, the designation of LOS can be the basis for establishing a traffic impact mitigation fee system to provide "fair share" funding for transportation improvements. The level of service can also be used as environmental impact review criteria that provide a basis for accepting, modifying, or denying a proposed development.

7.4 Travel Demand Modeling

A travel demand model is used to forecast the transportation arrangement. Because of the difference of various conditions, such as those discussed in the land suitability analysis, human activities in an area are not evenly distributed. Certain areas may not be suitable for all types of development. For those areas that can support development, some may be more suitable for industrial development, while other areas are more suitable for residential development. In addition to the natural condition, the layout of existing land uses also determines the potential for future development. For example, it would normally be considered inappropriate to build a factory in a residential area. The outcome of the suitability and compatibility considerations is the uneven distribution of human activities. Certain areas are predominately for residential uses, some for commercial uses, and some for industrial uses. Such distribution makes it necessary for people to travel among different areas using the transportation system. In addition to the connections, the demands for carrying capacity and other facilities, such as parking, vary spatially and temporarily. The trips to and from a factory may have morning and afternoon peaks while such patterns may not exist for trips to a shopping mall.

In order to estimate the traffic associated with different human activities, a region is divided into TAZs and the TAZs are connected to the transportation network. TAZs are the smallest unit of analysis in travel demand modeling. Two general rules are normally used in delineating TAZs:

(1) A zone should be bounded by the transportation network or natural boundaries, such as rivers.

(2) The zone boundaries enclose a relatively homogeneous area in terms of land use characteristics and traffic conditions and separate the areas that are different. TAZ boundaries may follow community or neighborhood boundaries.

The study area where a travel demand model can be developed must be small enough so that people are likely to travel between all zones within the area. At the same time, the study area must be large enough so that the trips crossing the study area boundary can be ignored. A metropolitan region is normally used as a study area.

Although there are different travel demand models using different variables, the application of travel demand modeling in general contains the following six components, normally arranged in a sequential order:

(1) Specify the regional population and economic activities for the study area. This component focuses on what is expected to occur in the study area. The population and economic analysis methods discussed in Chapters $3 - 5$ are used to determine the characteristics of human activities.

(2) Allocate these population and economic activities to each TAZ based on land uses. All human activities concerned in this context require the use of land. The methods discussed in Chapter 6 are used to identify and allocate land for human activities specified in the first component. Traffic Analysis Zones are the smallest areas used to summarize human activities.

(3) Choose a proper model structure and relevant variables to be included in the travel demand model. This process is called **model specification**. Modelers analyze the trend and special features of the study area and construct a model that describes the connection between human activities and the traffic demand. Such a model incorporates the most important variables in establishing the connection. In addition, model specification also determines how those variables are used to quantify the connection.

(4) Use the travel demand model to calculate traffic flows between TAZs. In this component, the travel demand model is used to calculate the traffic from one TAZ to another. The total traffic can be summarized by the modes of transportation and different routes that connect the TAZs.

(5) Collect actual traffic flow data and calibrate the travel demand model for the study area. The purpose of calibration is to adjust model structure and/or parameters in order to match model outputs with observed data. A model is only useful if its prediction matches the observed data. In this component, the travel demand model is fine tuned with real world data.

(6) Use the calibrated travel demand model to predict traffic flows for different growth scenarios. After the model calibration, the travel demand model is believed to be capable of predicting traffic for a given human activity scenario. Additional predictions may be derived from the traffic forecast, such as travel time and travel costs, street alignment and construction costs, and other social,

economic, and environmental impacts. The modeling results are used to support decisions about different alternatives.

Traffic planners and engineers spend a considerable time and resources to design a model and to analyze issues related to the structure, parameter and application of travel demand models. The rest of the Chapter introduces the general components of travel demand models. Although not all planners need to develop or operate a travel demand model, it is important for planners to understand how a travel demand model establishes the connection between human activities and traffic.

This travel demand modeling covers major travel behaviors that affect travelers' decisions on choice of traveling, destinations, transportation modes, and travel paths. The modeling process consists of four individual parts, commonly referred to as the four-step travel demand forecast modeling process.

(1) Trip Generation: Forecast the number of trips originated from and attracted to each TAZ.

(2) Trip Distribution: Allocate the trips within and between the TAZs.

(3) Mode Choice: Divide trips among different modes of travel.

(4) Trip Assignment: Assign the trips to different routes connecting the TAZs.

These four steps were developed in the 1950s and 1960s. Since then, although the four components are kept intact, many significant modifications have been made to the models in response to the advancement of understanding travel behavior by modelers (Chang and Meyers, 1999).

7.4.1 Trip Generation

As we discussed earlier, a trip is defined as a connection between an origin (O) and a destination (D). Consequently, trip generation is a process to determine the number of trips from and to a particular site or area.

Trip generation establishes the connection of transportation analysis to demographic analysis and economic analysis (Gazis, 2002). The subject of transportation analysis is how people travel. The number of trips generated in a zone depends on the zone's population. In general, the more people, the more trips expected. In addition, people with different characteristics travel differently. One observation from various studies shows that people with higher income levels tend to travel more than those with lower income. A young aged person who is not permitted to drive will have to ride with someone else in a private vehicle or use transportation modes other than the private vehicle. The demographic analysis discussed in Chapter 3 gives the base for estimating trip generation.

Some people may travel for the purposes other than to reach a destination.

For example, one may want to get in his/her car to be alone for awhile. However, the majority of trips have an origin and a destination. The destination is closely related to the trip purposes, such as to go to work, to shop, to dine, or to entertain. All those human activities are closely related to the economics of a region. The availability of employment opportunities can determine the number of people who travel to work. The type and size of retail stores can affect the number of people who travel to shop. Understanding the economic activities in a TAZ can help estimate the number of trips that may end in the zone.

As we discussed in Chapter 6, most human activities require the use of land. A region is divided into different pieces of land that are associated with different human activities. Majority people do not live and work in the same place, although the number of people who do so may increase. With advances in technology, such as high speed internet connection, people may work at home. However, majority of the jobs will still require face-to-face interaction in a traditional work-place. The inventory of land uses, therefore, provides the base for estimating trip generation. For example, the Institute of Transportation Engineers publishes trip generation rates for different land use types. The trip generation rate can be calculated as daily trips or peak hour trips. Trip generation rate are presented as vehicle trips or person trips. A vehicle occupancy variable can be used to convert the two rates:

$$
TG_v = TG_p / VO
$$
 (7.1)

where,

 TG_v — vehicle trip generation rate;

 TG_p — person trip generation rate;

VO — vehicle occupancy rate.

As shown in Table 7.4, trip generation rate can be calculated for a particular site. Depending on the type of land use, trip generation rates may be expressed as number of trips per employee, number of trips per unit land area (i.e., trips per acre), or number of trips per occupied dwelling unit. The ITR report also separates the trips by direction—entering or exiting the site. These trip generation rates are normally derived from observed data, using regression analysis.

The origin and destination are normally represented as Traffic Analysis Zones. After a study area is divided into TAZs, the amounts of different land use types in each zone can be determined. There are two components of estimating trip generation and both are closely related to TAZ-level land uses. Trip production refers to the number of trips that originate from a TAZ. Trip attraction reflects the number of trips that end in a TAZ. The combination of trip production and trip attraction is the outcome of trip generation analysis. People's travel behaviors vary for different trip purposes. To improve the accuracy of trip estimation, the trip generation analysis is usually done separately for different trip purposes.

Table 7.4 Selected trip generation rates

343

Chapter 7 Transportation Analysis

The TAZ-based trip generation is calculated in two steps. In the first step, TAZ-based trip productions are calculated. Then TAZ-based trip attraction rates are estimated. The attraction rates reflect the relative attractiveness of a TAZ in relation to other TAZs in the study area.

The difference between trip production/trip attraction and trip origin/trip destination is worth noticing. An individual trip has two ends, one end is the trip origin and the other end is the trip destination. Trip production and trip attraction refer to aggregated trips associated with traffic analysis zones, rather than individual trips. This distinction is important and becomes the basis of trip generation studies. Furthermore, trip production is only related to residential land in a zone. That means, only the TAZs having residential land can produce trips. Trips can be attracted by both residential and non-residential land uses.

Figure 7.4 illustrates the difference of trip origin-destination and trip production-attraction. The graphic represents a three-zone area. Zone A is residential only and Zones B and C only have non-residential land uses. Assume a person who lives in Zone A goes to Zone B to work. After work the person goes shopping in the same zone (Zone B) and then to a take-out restaurant in Zone C before going back home in Zone A. There are total of four trips. Zone A and Zone C each has one trip origin and one trip destination and Zone B has two trip origins and two trip destinations. For Trip 1, which starts from Zone A and ends at Zone B, Zone A is the origin and Zone B is the destination. For Trip 2, which starts from Zone B and ends in Zone B, Zone B is both the origin and the destination. The origin for Trip 3 is Zone B and destination is Zone C. For Trip 4, Zone C is the origin and Zone A is the destination.

In a trip generation study, trip direction is ignored and all trip productions are only associated with residential land use. Therefore, in the three-zone example in Fig. 7.4, only Zone A, the zone with residential land use, can be associated with trip production. Zone A can also attract trips. The other two zones are only associated with trip attraction, not trip production. In the simplified example of four trips shown in Fig. 7.4, all four trips are treated as being produced in Zone A. Zone B and Zone C attract two trips, respectively. No trips are generated in either Zone B or Zone C.

This certainly introduces errors. For example, Trip 2 does not start from Zone A, nor does it end in Zone A. However, there is no residential land use in Zone B, Trip 2 is still treated as if it is produced in Zone A. In addition, the model assumes that there are no trips being generated in Zone B and attracted to Zone C. This limitation is attributed to the practical operation of a travel demand model. Only with such assumption is it feasible to simulate the traffic. Another reason for such model design is that trips generated from workplaces are much smaller than the trip production from residential land. In addition, most TAZs contain residential and non-residential land, which helps to hide the problem.

Figure 7.4 Difference of trip origin-destination and trip production-attraction

7.4.1.1 Trip Production

A common practice of estimating trip production from a TAZ is based on variables related to population in the zone. The most commonly used variable is number of households. It should be emphasized that the trip production is more complicated than simply house counting. For example, households with automobiles are more likely generating more vehicle trips than households without an automobile. Households of higher incomes generally make more non-work related trips (for example, shopping) than lower income households. The following list contains some of the variables commonly used in calculating trip production. Not all of them need to be considered in a single model. In fact, many of these variables are correlated and should not be included in the same model.

Workers per household. Workers are the most likely people who travel. The more workers in a household the more trips will be made.

Number of households. Trips are generated by residents. The more households in a zone the more trips will be made.

Family income. Costs are associated with travels. The higher the family. income, the more trips will be made.

Number of automobiles available. People with a car likely travel more than those without a car.

Education level. People's travel behaviors may vary depending on the level of education.

Family size. It is likely that large families travel more than smaller sized families.

Family's age distribution. People at different ages levels travel differently. For example, school aged children and retirees are not likely to travel to work.

Number of occupied dwelling units. The dwelling unit occupancy rate is an indicator of the number of residences. With the same dwelling units in a zone, the more occupied dwelling units, the more trips that can be expected from the zone.

Dwelling unit type. Home ownership is normally an indicator of family income.

Residential density. The residential density on one hand reflects the number of residents in a zone. The more people living in a zone the more trips are expected to be generated in the zone. On the other hand, residential density may reflect people's living style and travel pattern. For example, a high density area has more public transit services than a low density area.

Two methods are used most often in estimating trip production from traffic analysis zones.

The first method uses **aggregated zonal characteristics**. The trip production of a zone is the dependent variable and independent variables are those in the list above or zonal summary statistics derived from them. The following formula illustrates an example of the functional relationship between the number of trips produced by zone and three zonal independent variables.

Trip Production *f* (median family income, residential density, mean number of automobile per household) (7.2)

Regression analysis is a common approach to specifying the formula. After zonal variable data for the entire study area are collected, linear regression analysis is applied to derive the coefficients for the prediction model. A major concern of this type of trip production model is that the overall zonal characteristics may not accurately reflect the driving forces of trip production. In other words, the number of trips generated from a zone is not necessarily related to the aggregated zonal variable values. For example, Fig. 7.5 shows the family income distribution for two hypothetical zones.

There are 387 and 349 families in each zone, respectively. The median family income values for the two zones are quite close. If all other independent variables also are similar, an aggregated model using median family income as an independent variable would predict that the two zones produce similar number of trips. However, the family income distribution as shown in Fig. 7.5 reveals the difference between the two zones. Family income in Zone 1 spreads a much wider range than that in Zone 2. If the rationale behind the model is that family income affects people's travel behavior, the trips produced in the two zones are very likely different. However, the aggregated model is unable to reflect the difference.

The second method, **cross-classification method**, addresses this deficiency by classifying zonal households into categories based upon socio-economic characteristics. Trip generation rates are then developed for each individual category. The rationale of this method is that households with similar characteristics are likely to have similar travel patterns. Therefore the trips generated for each individual household category should be calculated separately. The trips from different household categories are added together to provide a more accurate estimate of zonal trip generation.

Figure 7.5 Family income distribution in two hypothetical traffic analysis zones

Normally, no more than three to four variables such as family size, automobile ownership, or household income, are used in household classification. Each variable has a few discrete categories. The cross-classification method requires much more data than the zonal method. For example, if four variables are used in household classification and each has three categories, there would be $3⁴ = 81$ household categories in a zone.

The following formula is an example for calculating trip generation for families of different household income categories:

$$
T_{\rm p} = \sum_{i=1,n} (r_i \cdot H_i)
$$
 (7.3)

where,

 T_{n} — trips generated;

i — household category;

 n — number of household categories;

 r_i — trip production rate for household category *i*;

Hi — number of households in category *i*.

Data used in trip generation can be obtained from government agencies or field surveys. In the previous chapters, we have discussed land use data and census data. Those data are often used for estimating trip generation. It may take considerable effort to collect data for the study area. Once relevant data are collected, regression analysis is a common tool used to derive the trip production rate for each category. Trip production rates from similar cities or regions can be used if no local data are available.

The following example illustrates that the amount of person trips per household is based on the number of households classified by household income. In addition, the distribution of trip purpose varies by household income level. As shown in Table 7.5 households in the lowest income groups make fewer trips than the higher income group.

Income Level: \$1,000					
Personal Trips Per Day		15 $15 - 25$ $25 - 35$ $35 - 45$			> 45
Per Household by Trip Purpose					
Home based work (HBW)	0.443	0.873	1.260	1.867	2.766
Home based non-work (HBNW)	2.283	2.593	2.643	2.826	3.022
Non-home based (NHB)	1.423	1.567	2.092	2.604	3.241

Table 7.5 Number of person trips per day by trip purpose and by household income

The person trips by household income category are calculated as:

$$
PT = N_{hh} \cdot PTPH \tag{7.4}
$$

where,

PT— person trips;

 N_{hh} — number of households;

PTPH— person trips per household.

For example, Table 7.6 shows that there are 352 households in the first household income category ($, $15,000$), the daily HBW person trips produced by those households are estimated as:

$$
352 \times 0.443 = 156
$$

	< 15	$15 - 25$		$25 - 35$ $35 - 45$	> 45	Total		
Number of Households	352	478	491	892	543	2,756		
							Automobile Vehicle	
			Person Trips				Occupancy	Trips
HBW	156	417	619	1.665	1,502	4,359	1.37	3,182
Person Trips								
HBNW	804	1,239	1.298	2,521	1.641	7,502	1.81	4,145
Person Trips								
NHB	501	749	1,027	2.323	1,760	6,360	1.43	4,447
Person Trips								
All Purposes						18.222		11,774

Table 7.6 Trip production by household income and trip purpose

The number of daily HBW person trips produced by the 543 households in the last household income category ($> $45,000$) is calculated as:

$$
543 \times 2.766 = 1,502
$$

Similar calculations can be made to the other household categories. Using the number of households by category in Table 7.5, the total HBW person trips is calculated as the sum of trips in all household categories:

$$
T_p = \sum_{i=1,n} (r_i \cdot H_i)
$$

=352 × 0.443 + 478 × 0.873 + 491 × 1.260 + 892 × 1.867 + 543 × 2.766
= 156 + 417 + 619 + 1,665 + 1,502
= 4,359

The result shows that 4,359 daily home-based-work person trips are expected from this zone. The same calculation can be made for other two trip purposes. The results shown in Table 7.6 are a total of 7,502 person trips for the homebased-non-work purpose and the trips that belong to the non-home-based trip purpose are 6,360.

Another variable—vehicle occupancy—is needed to convert person trips into vehicle trips. This information is normally collected at the local level. In general, the vehicle occupancy for work-related trips is lower than other trips because it is more likely for people to travel alone to work. Table 7.6 includes illustrative vehicle occupancy values for different trip purposes. Using the vehicle occupancy, we can then calculate the vehicle trip production for the zone as follows:

> HBW vehicle trips = $4,359 / 1.37 = 3,182$ HBNW vehicle trips = $7,502 / 1.81 = 4,145$ NHB vehicle trips = $6,360 / 1.43 = 4,447$

From the result, we can see the impact of vehicle occupancy on the vehicles on the road. Because people are more likely to travel together for trips from home to non-work places than the trips not starting from homes, the HBNW vehicle occupancy is greater than that of NHB. Although more people travel for the HBNW purpose (7,502) than the NHB purpose (6,360), there are fewer HBNW vehicle trips (4,145) than the NHB vehicle trips (4,447).

7.4.1.2 Trip Attraction

Trip attraction predicts the number of trips to be attracted to (end in) each zone. The attractiveness of a zone is related to the size and type of land uses that are the destination of trips. The majority of such land uses are non-residential land, such as stores, offices, libraries, etc. The trip attraction is normally expressed as the number of vehicle trips per household or per unit area of non-residential land use. In addition to the zonal characteristics, the number of trips attracted to a zone is related to the attractiveness of other zones in the region. All trip attracting zones in a region compete for the number of trips produced in the region. Zones that are more attractive will attract more trips than the zones that are less attractive.

For example, Table 7.7 lists trip attraction rates for residential and some non-residential land uses. For the residential land use, the trip attraction rate is expressed as number of trips per household. The value of "0.079" in Table 7.7 means that each household can attract 0.079 HBW vehicle trips per day. This can be explained as on average, each 1,000 households can attract 79 trips from home to work. An example of HBW trip to a residence is a nurse who leaves home to take care of a patient at the patient's home.

		Trip Purpose				
Type of Activity	HBW	HBNW	NHB			
Households (unit)	0.079	0.518	0.302			
Retail $(sq \cdot km)$	155.7	560.0	467.6			
Basic $(sq \cdot km)$	131.9	84.3	93.8			
Service $(sq \cdot km)$	112.3	64.3	126.6			

Table 7.7 Daily vehicle trip attraction rates

The vehicle trip attraction rates associated with non-residential land uses are functions of the land use type and the size of land uses. The values in Table 7.7 are the per unit area trip attractions by different land use types. The table shows that retail trade sector is more attractive than basic or service sectors. Similar to the residential land, the trips attracted by different nonresidential land can be divided into different trip purposes. For the retail land use, the rate of homebased work trips (155.7, those who work in retail stores) is lower than that of home

based non-work trips (560.0, those who come to stores from home to shop).

Similar to trip generation, data required for calculating the trip attraction rates can be obtained from government agencies or field survey. The types and amount of different land uses by traffic analysis zone can be derived by overlaying the TAZ and land use data (as discussed in Chapter 6). The number of households is usually part of the census data regularly collected. Once the data are collected, regression analysis is a common tool used derive the trip attraction rate for each land use category. When local data are not available, trip attraction rates from similar cities or regions can be used.

The HBW vehicle trips attracted to a zone is then calculated as:

$$
TA_{HBW_H} = N_{hh} \cdot TAR_{-R}
$$
 (7.5)

where,

TA_{HBW}_R— home-based work vehicle trip attractiveness of the zone by households;

*N*hh— number of household in the zone;

TAR $_{R}$ — trip attraction rate by households.

The HBW trips attracted by retail are calculated from the size of retail land use and the retail trip attraction rates.

$$
TA_{\text{HBW_NR}} = A_{\text{NR}} \cdot \text{TAR}_{\text{NR}} \tag{7.6}
$$

where,

TA_{HBW NR}— home-based work vehicle trip attractiveness of the zone;

 A_{NR} — non-residential land use size in the zone;

TAR $_{\text{NR}}$ — trip attraction rate of the non-residential land use.

Table 7.8 lists an example of the types of land use that attract trips for a zone. The unit for the residential land is the number of households and the units for other land uses are square kilometers. The HBW vehicle trips attracted to the zone by the 2,756 households in the zone are calculated as:

HBW trip attractiveness = number of households \times residential trip attraction rate

$$
= 2,756 \times 0.079
$$

= 218

For example, the HBW trip attractiveness of the zone is:

HBW trip attractiveness $=$ size of retail land use \times retail trip attraction rate $= 14.53 \times 155.7$ $= 2,263$

The trips attracted by different land uses can be calculated in the same fashion. The results are included in Table 7.8. The last row shows the total vehicle trip attractiveness of the zone.

Land Use	Size	Trip Attractiveness			
		HBW	HBNW	NHB	Total
Households (unit)	2,756	218	1,428	832	2.478
Retail $(sq \cdot km)$	14.53	2,263	8,138	6.796	17,197
Basic $(sq \cdot km)$	5.51	727	464	517	1,708
Service $(sq \cdot km)$	1.92	216	123	243	582
Total		3.424	10,153	8,388	21,965

Table 7.8 Trip attractiveness by trip purpose

After the same procedure is applied to all the zones in the study area, a trip production and attraction table is constructed. The following tables are the results for a hypothetical study area of four TAZs. From the number of households by income level in Table 7.9 and trip production rates in Table 7.5, we can calculate the trip productions by income level, as shown in Table 7.10.

Summarizing the trip generations by income level in Table 7.10 we can derive the person-trips by purpose, which is displayed in the left portion of Table 7.11. The vehicle occupancy ratios are then used to convert the person-trips to vehicle-trips. The bottom row of Table 7.11 shows the total trips produced in each zone. From the result, we can see that Zone 4 produces many more trips (50,754) than other zones.

The trip attraction by zone is calculated using the trip attraction rates in Table 7.7 and the land use by zone for the study area, in Table 7.9. The results are displayed in Table 7.12. Table 7.12 is an expansion of Table 7.8. The total trip attraction summarized by trip purpose and zone is included in Table 7.13.

	Number of Households by Zone and Income Level								
Zones									
Income Level $(1,000$ dollars)	1	\overline{c}	3	4					
< 15	352	141	70	70					
$15 - 30$	478	239	598	956					
$30 - 45$	491	4,419	2,946	7,365					
$45 - 60$	892	1,427	714	1,784					
>60	543	2,715	1,086	1,810					
Total	2,756	8,941	5,414	11,985					
		Land Use by Zone							
Zones Land Use	1	$\overline{2}$	3	4					
Households (unit)	2,756	8,941	5,414	11,985					
Retail $(sq \cdot km)$	14.53	12.10	8.62	10.19					
Basic $(sq \cdot km)$	5.51	14.20	7.38	14.19					
Service $(sq \cdot km)$	1.92	2.73	1.63	2.57					

Table 7.9 Trip production data by traffic analysis zone

Number of Vehicle-Trips Produced by Trip Purpose and Zone							
Zones Trip Purpose				4	Total		
HBW	3,423	4,771	5,853	4,695	18,741		
HBNW	10,154	12,783	16,719	13,278	52,934		
NHB	8,388	10,038	13,130	10,043	41,599		
Total	21.964	27.591	35,703	28,016	113,274		

Table 7.13 Trip attraction by trip purpose

For trip generation modeling, the number of trips produced must be compared with the observed trips to select independent variables (predictor variables) and the parameters that connect them to the trip generation.

7.4.2 Trip Distribution

The purpose of the trip distribution process is to allocate the trip productions and trip attractions. "To allocate" means to specify where the trips generated from a particular TAZ will go. From Tables 7.11 and 7.13, we can see that Zone 1 produces 11,774 trips and attracts 21,964 trips. However, we don't know where the 11,774 trips produced at zone 1 end and where the 21,964 trips attracted by zone 1 originate. Trip distribution analysis addresses this question. The rationale behind the trip distribution is quite simple. The trips between any trip production zone, *i*, and trip attraction zone, *j*, are a function of the number of trip production, P_i , the trip attraction, A_i , and the effort associated with the travel between the two zones, W_{ij} *w_{ij}* is normally expressed as the costs or time taken to travel from the production zone to the attraction zone. The equation below is a formula for calculating the trip distribution between the two zones, T_{ij} . Such an equation is called gravity model.

$$
T_{ij} = C \frac{P_i \cdot A_j}{W_{ij}} \tag{7.7}
$$

where,

 i — the zone of trip production;

- j the zone of trip attraction;
- T_{ii} number of trips produced in Zone *i* and attracted to Zone *j*;

$$
P_i
$$
—number of trips produced in Zone *i*;

 A_i — number of trips attracted to Zone *j*;

 W_{ii} — the impedance between Zones *i* and *j*;

C — constant.

The formula is named after Newton's law of gravity, which states that the attractive force between any two bodies is directly related to the masses of the bodies and inversely related to the distance between them. In the case of a trip distribution, the attractive force is the number of trips between two zones, the masses are the trip production and attraction, and the distance is the impedance factor between the two zones. Accordingly, large numbers of trip productions or trip attractions increase the number of trips any two zones, while higher impedance factors reduce the trip volume. Conceptually, the gravity model is straightforward. The challenge in applying the model is to assign proper values to the impedance, *W*, and the constant, *C*.

As we discussed before, the travel demand model assumes there is a balance between the total trip production and attraction. This means that trips produced in Zone *i* must equal to the total of all the zones that receive trips that originated in Zone *i*. That is,

$$
P_i = \sum_j T_{ij} \tag{7.8}
$$

The implication is that all zones compete for the trip production, P_i , based on the zonal attractiveness and impedance. For a study area of *n* zones, the balance can be expressed as:

$$
P_i = \sum_{j=1,n} T_{ij} = \sum_{j=1,n} C \frac{P_i \cdot A_j}{W_{ij}}
$$
(7.9)

Therefore, the constant *C* can be derived as:

$$
C = \frac{1}{\sum_{j=1,n} \frac{A_j}{W_{ij}}} \tag{7.10}
$$

Replacing the constant *C*, in the trip distribution formula Eq. (7.6), we get:

$$
T_{ij} = P_i \left[\frac{A_j}{\sum_{j=1,n} \left(\frac{A_j}{W_{ij}} \right)} \right]
$$
(7.11)

The relative attractiveness of Zone *j* regarding the trips produced in Zone *i* is expressed in the bracketed term, which is related to the attractiveness and impedance from all other zones.

The impedance, W_{ij} , reflects the level of difficulty when traveling between the two zones. Normally, it is related to the physical condition of roadway network, distance, cost of travel, or time of travel. The model developer must

make a decision as to what factors to use for deriving the impedance. As an example, W_{ij} can be a function of travel time between Zone *i* and Zone *j*. The impedance increases as the travel time increases:

$$
W_{ij} = t_{ij}^a \tag{7.12}
$$

where,

 t — travel time: *W* — impedance;

a — a constant.

Assume the constant, *a*, for the four-zone study area is 0.5. If we know the travel time as shown in Table 7.14, we can calculate the impedance, W_{ij} . For example, the travel time within Zone 1 is 5 minutes, the impedance for trips inside Zone 1 is:

$$
W_{11} = t_{11}^a = 5^{0.5} = 2.24
$$

The travel time from Zone 1 to Zone 4 is 25 minutes, the impedance from Zone 1 to Zone 4 is calculated as:

$$
W_{14} = t_{14}^a = 25^{0.5} = 5
$$

Table 7.14 lists the complete calculation of impedance for all possible travel options. From the table we can see that the travel time between Zones 3 and 4 is the longest of all ($t_{34} = 30$ minutes). Therefore, the impedance between the two zones has the highest value ($W_{34} = 5.48$).

Travel Time, t_{ii}							
To From		$\overline{2}$	3	4			
	5	15	20	25			
2	15	6	20	15			
3	20	20	7	30			
4	25	15	30	8			
Impedance, $W_{ij} = t_{ij}^a (a = 0.5)$							
To From		\overline{c}	3	$\overline{4}$			
	2.24	3.87	4.47	5.00			
\overline{c}	3.87	2.45	4.47	3.87			
3	4.47	4.47	2.65	5.48			
4	5.00	3.87	5.48	2.83			

Table 7.14 Travel time and impedance matrix

You probably have noticed that we have changed the question from finding W_{ij} to finding t_{ij} and *a*. The travel time, t_{ij} , can be obtained from historical record. The constant, *a*, is an empirical parameter and its value needs to be estimated and adjusted against observed trip data.

Normally, a friction factor, F_{ij} , which is the inverse of the impedance, W_{ij} , is used in the trip distribution formula to reflect people's wiliness to travel between zones

$$
F_{ij} = 1/W_{ij} \tag{7.13}
$$

Continued

Considerable effort is spent in transportation modeling finding an appropriate function for friction factor. Equations (7.12) and (7.13) are included only as an example. The friction factor values may be specifically chosen to take into consideration such things as a major barrier between zones or, a toll to cross a bridge. Regardless the form of equations, a friction factor represents the likelihood of trips between any two zones. For example, two zones that are close to each other and connected by an express-way will have higher friction factor value than another pair of zones that are far apart. If two zones are so far apart that no one is willing to travel between the two, the friction factor value will be zero. Placing F_{ij} in Eq. (7.13) into Eq. (7.12) we can derive the gravity model for zonal traffic:

$$
T_{ij} = \frac{P_i \cdot A_j \cdot F_{ij}}{\sum_{i=1,n} (A_j \cdot F_{ij})}
$$
(7.14)

This model shows that the amount of traffic between two zones is proportion to the number of trip produce in Zone *i*, the trip attraction of Zone *j* and all other zones, and the friction factor between all possible pair of zones.

In reality, travel decisions more complete than the trip production, trip attraction, and impedances. Studies have shown that many other factors affect people's travel behavior, such as age, income, gender, vehicle ownership, or availability and quality of public transit services (Hensher and Reyes, 2000;

Taplin and Min, 1997; Turner and Grieco, 2000). As a common practice, trip forecast model developers use a set of inter-zonal socioeconomic adjustment factors, K_{ij} , and include them in the gravity model. The U.S. Department of Transportation has summarized the following rationales for the necessity of including K_{ij} in the gravity model (USDOT, 1985).

First, the gravity model assumes that the trip purpose determines travel pattern. Consequently, the largest proportion of HBW tips will be allocated to the closest zones (small friction) with largest employment establishments (large trip attraction). However, different jobs require different skills and employ certain members of work force.

In a similar manner, some zones are more likely to have jobs and housing for certain income levels. For example, people who work at grocery stores may have quite different incomes than those who work in corporation headquarters in central business districts. In the United States, they are not likely to live in the same neighborhoods.

Last, the friction factor in the gravity model is developed for the entire study area. For example, it implies that travel time and the cost of travel have the same affect to people's travel behavior. This assumption does not consider the different responses to the impedances. For example, the travel cost may affect people differently, depending on their income level. Assume a city is implementing a congestion fee on rush hour highway travels. Low income people may be unable to allocate their limited resources to pay for the congestion fee, and consequently, unable to use the highway. Those who can, and are willing to pay for the fee, will be able to travel on less a congested highway.

In practice, it is too difficult to collect accurate data to allow further stratification of employment opportunities and residents. However, the model may not be valid without considering these factors. With the limited knowledge of these factors, they are included in one adjustment factor— K_{ij} . The gravity model Eq. (7.14) is then revised as:

$$
T_{ij} = K_{ij} \frac{P_i \cdot A_j \cdot F_{ij}}{\sum_{i=1,n} (A_j \cdot F_{ij})}
$$
(7.15)

The *K* factors can be added during model calibration to incorporate effects that are not previously captured. Those effects can be interpreted as the extent to which the trips can be increased or decreased because of these unaccounted factors. Because of the complexity of those factors, it is difficult to estimate their values. A common practice is to derive *K* factors in the model development process. The process of developing travel models is also called **calibration**, during which the model estimations is compared with observed data. Various parameter values are adjusted until the model output satisfactorily matches the

observation data.

This Eq. (7.11) shows that the trips between a production Zone, *i*, and an attraction Zone, *j*, increase as the trip production or trip attraction increases. When the friction factor increases (for example, as a result of an improved road condition the travel time is shortened), the trips are expected to increase between the two zones. If the friction factor decreases as a result of congestion that leads to longer travel time, the trips between the two zones will decrease. If the socioeconomic factors can lead to more trips between two zones, the K_{ij} value will be greater than 1; otherwise, the K_{ij} value will be less than 1.

To illustrate the use of the trip distribution model, let us ignore the effect of all other socioeconomic factors. That is, $K_{ij} = 1$. The equation becomes the formula:

$$
T_{ij} = \frac{P_i \cdot A_j \cdot F_{ij}}{\sum_{i=1,n} (A_j \cdot F_{ij})}
$$

The computation of trip distribution can be completed in three steps as shown in the following tables. From the zonal attractiveness A_i of Zone j and friction factor between the Zone i and Zone j , F_{ij} , we can calculate the adjusted attractiveness of Zone *i* to Zone *j*, $A_j \cdot F_{ij}$. As shown in Table 7.15, the attractiveness of Zone 1 to Zone 1 is calculated as:

$$
21,946 \times 0.45 = 9,823
$$

The attractiveness of Zone 1 to Zone 2 is:

$$
27,591 \times 0.26 = 7,124
$$

Similarly, we can calculate the attractiveness of Zone 1 to Zone 3 and Zone 4:

Zone 1 to Zone 3:
$$
17,851 \times 0.22 = 3,992
$$

\nZone 1 to Zone 4: $28,016 \times 0.2 = 5,603$

The last column in Table 7.15 represents the total of attractiveness of a Zone *i* to all zones in the region. For Zone 1, the value is calculated as:

$$
\sum_{j} (A_j \cdot F_{1j}) = 9,823 + 7,124 + 3,992 + 5,603 = 26,542
$$

Similarly, the trip attractiveness of all zones regarding Zone 2 is calculated as:

$$
\sum_{j} (A_j \cdot F_{2j}) = 5{,}671 + 11{,}264 + 3{,}992 + 7{,}234 = 28{,}160
$$

To From		$\overline{2}$	3	4	$\sum (A_i \cdot F_{ij})$
A_i	21,964	27,591	17,851	28,016	
	9,823	7,124	3,992	5,603	26,542
2	5,671	11,264	3,992	7,234	28,160
3	4,911	6,170	6,747	5,115	22,943
4	4,393	7,124	3,259	9,905	24,681

Table 7.15 Product of trip attraction and friction factor, $A_i \cdot F_{ii}$

The values in Table 7.15 reflect the impact of the two factors on the zonal attraction. The product $A_i \cdot F_{ii}$ means that more attractive zones (with higher A_i 's) are likely to attract more trips. Meanwhile, as travel from the zone of production to the zone of attraction becames easier (or more convenient), (higher F_{ij}), more trips will be attracted to the zone of destination.

Two more steps are required to calculate the actual number of zonal trips. In the first step, the values in Table 7.15, $A_i \cdot F_i$, were multiplied by the vehicle trip production from the production zone P_i (the last row in Table 7.11 and included as a column in Table 7.16). The result is saved in Table 7.16. For example, the value for Zone 2 to Zone 3 ($i = 2$ and $j = 3$) is calculated as:

$$
P_2 \cdot A_3 \cdot F_{23} = 41,042 \times 3,992 = 163,825,554
$$

When the value for a trip destination Zone *j* is divided by the total attractiveness from this zone and all other zones in terms of a trip production Zone *i*, $\sum (A_i \cdot F_{ij})$, the result is the actual trip distribution from Zone *i* to Zone *j*. Table 7.17 displays the trip distribution results for the four-zone illustration.

Comparing the sums of trips ending in Zone $j - \sum T_{ij}$ Table 7.17 — with the *i*

original trip constant productions— P_i , Table 7.16—we can see that the row totals, i.e., the trip end productions, remain constant. This shows that the number of trips produced is not related to the travel condition.

To From		2	3	$\overline{4}$	$\sum T^{}_{ij}$	
1	4,358	3,160	1,771	2,486	11,774	
2	8,265	16,416	5,818	10,542	41,042	
3	4,977	6,251	6,837	5,183	23,247	
4	9,034	14,650	6,702	20,369	50,754	
$\sum T_{ij}$	26,634	40,477	21,128	38,580	126,817	
$T_{ij} = P_i \cdot A_j \cdot F_{ij} / \sum (A_j \cdot F_{ij})$						

Table 7.17 Trip distribution

The advantage of the gravity model is its simplistic form. The calculation is straightforward. For a trip generation zone, the higher the trip production, the more trips will originate from the zone. A zone that produces more trips is expected to have more trips to other zones than another zone that produces fewer trips. Similarly, a zone that has higher trip attractiveness will accept more trips than another zone that attracts fewer trips. The larger the friction factor between any two zones will lead to more trips traveling between the two zones. The challenge of the model in practice is that it only uses two parameters, F_{ij} and K_{ij} to

represent travel choices. The lack of behavioral basis to explain how individuals or households decide their travel destinations is the drawback of the method.

7.4.3 Mode Choice

The third step of the travel demand model is to estimate the proportion of travelers using different modes of transportation. There are many alternative modes available for an individual to travel from one place to another, such as driving alone or with someone else, walking, taking the train, bus, taxi, riding a bicycle, etc. Many variables may affect an individual's mode choice. If you take a few minutes to list the reasons you used for choosing particular travel modes for different activities last month, you may have a long list. Of course, you may also find your choices were quite limited. There are so many places in the United States, especially in suburban areas, where there is no public transit. It is quite common to see many streets without sidewalks. People in those areas are forced to drive to travel, even to get the Sunday morning paper. Do you remember how desperate you were last time your car broke down?

The variables affecting mode choice can be organized into three categories traveler, trip, and transportation system. Traveler characteristics include variables such as automobile ownership, income, number of workers per household, and the place of living and place of work. The University of Cincinnati is in uptown Cincinnati. Many students who live on or near campus simply walk to school. As a commuter campus, there are also many students who commute everyday. They take buses, drive private vehicles, or ride bicycles to school.

The second category of factors affecting mode choice refers to the journey characteristics, such as the trip purpose, length of trip, place of origin and destination, or time of day the trip is taken. If we do not count those who just want to take a ride for the fun of travel, people travel with a purpose. Different trip purposes may determine how to travel. For example, although I can ride my bicycle to work, I would not be able to do so to take two small children to the zoo.

The third category of variables affecting mode choice is related to the characteristics of the transportation system. Those variables may include travel costs, time taken for the travel, comfort level of travel, convenience, reliability, and security of different modes. An individual makes his/her decision on travel mode after comparing the characteristics of different travel modes. For example, a raise in the bus fare may induce people who ride buses to switch to driving private vehicles. After an increase in the parking fee, some people who currently drive to work may switch to other modes.

Even though mode choice is individually based, the mode choice model estimates the aggregated number of trips associated with each of the possible transportation modes. The outcome of mode choice model is the percentage of travelers using each available travel modes. There are many different ways of calculating the mode choice. One approach is to use a diversion curve, which illustrates the split of two modes. Figure 7.6 illustrates a hypothetical mode choice diversion curve. It compares public transit with private automobiles using travel time as the variable. The horizon axis is the ratio of transit-to-auto travel time ratio. A value of 1 represents that there is no difference of travel time for transit or auto travel. When the transit is faster than driving, the ratio would be less than 1. The ratio would be greater than 1 if it is faster to drive than using the transit. The vertical axis represents the proportion of transit trips. According to the diversion curve in Fig. 7.6, about 47% of travelers would use transit if the travel time is the same for transit and driving private vehicles (the solid line). If the transit travel time is 3/4 of the automobile travel time, the number of travelers who use transit equals to those driving private vehicles.

The diversion curve approach is simple to use. However, one curve is unable to reflect the vast number of variables that may affect travelers' mode choices. One approach to improve the method is to stratify trips using some of the important variables, such as trip purpose, income level, or cost of travel. One diversion curve is developed for each stratified group. For example, 160 diversion curves were used in Washington, D.C., USA (Wright, et al., 1997).

Another option is to apply a utility function for each possible mode of transportation. A utility function measures the degree of satisfaction (or cost if

Figure 7.6 An illustrative mode choice diversion curve

negative) that is associated with each mode choice. The market share of all mode choices is then calculated based on the utilities. This approach is based on the assumption that travelers make rational choices between available modes. That is, a traveler selects the mode with highest utility value if he/she has access to perfect information about each travel mode.

To develop a utility function, we must first decide the relevant independent variables to include in the model. Those variables should reflect the three categories previously discussed: characteristics of travelers, journeys, and transportation system. One example of the traveler characteristics is age. The utility of private vehicles by elderly or underage travelers is limited since they will have to use public transit or travel with others who can drive. Trip purpose and trip destination are examples of the journey characteristics. Easy and mostly free parking of large shopping malls in the outskirts of American cities attract many customers away from downtown areas. The transportation system characteristics are the most common variables to be included in the utility functions. The modes associated with shorter travel times, lower costs, or more convenience are likely have high utility. Sometimes an independent variable may be derived from other variables. As we discussed before, travel costs may have different effects depending on a traveler's income level. A ratio of travel cost to the traveler's income level may be the variable included in the utility function (Papacostas and Prevedouros, 2001). In this example, travel cost is a transportation system variable and income level is a traveler variable.

The utility function is typically expressed as a linear weighted sum of the independent variables (X_1, \dots, X_n) . The effect of variables not specified in the

model is included in the constant item, a_0 . The general form of the utility function with *n* variables is:

$$
U = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n \tag{7.16}
$$

where,

 U — the utility of the transportation mode;

 a_0 — constant;

 a_1 — weight for the first variable;

 a_2 — weight for the second variable;

 a_n — weight for the *n*th variable;

 X_1 — the first independent variable;

 X_2 — the second independent variable;

 X_n — the *n*th independent variable.

For a study area with *k* types of transportation modes, one utility function is established for each mode:

$$
\begin{cases}\nU_1 = a_{01} + a_{11}X_{11} + a_{21}X_{21} + \dots + a_{n1}X_{n1} \\
U_2 = a_{02} + a_{12}X_{12} + a_{22}X_{22} + \dots + a_{n2}X_{n2} \\
\vdots \\
U_k = a_{0k} + a_{1k}X_{1k} + a_{2k}X_{2k} + \dots + a_{nk}X_{nk}\n\end{cases} (7.17)
$$

Any of the factors mentioned before (traveler, trip, and transportation system variables) may be independent variables. A major task in developing the utility functions is the selection of independent variables, which goes beyond this book. Let's assume a set of mode choice utility functions for two modes of three independent variables are developed. The two mode choices are private vehicle and public transit. The three independent variables are cost of travel, travel time, and comfort level.

$$
\begin{cases}\nU_1 = -0.03 - 0.02X_{11} - 0.015X_{21} + 0.04X_{31} \\
U_2 = -0.035 - 0.025X_{12} - 0.02X_{22} + 0.05X_{32}\n\end{cases}
$$
\n(7.18)

where,

 U_k — the utility, $k=1$ for private vehicle travel and $k=2$ for public transit travel;

 X_{1k} — the cost of travel in cents per kilometer, $k=1$ for private vehicle travel and $k = 2$ for public transit travel;

 X_{2k} — the travel time in minutes, $k = 1$ for private vehicle travel and $k = 2$ for public transit travel;

 X_{3k} — the comfort level, $k=1$ for private vehicle travel and $k=2$ for public transit travel.

You may notice that the coefficients for travel cost and travel time are negative and for comfortable level is positive. The sign of the coefficients reflect the change direction of the independent variables. The negative value implies that as the value of the independent variable increases, the utility will decrease. In the case of travel cost and travel time, people will tend to use the travel mode that costs less and uses less travel time. For the third variable, people will give the travel preference to the mode that is more flexible.

Values for three independent variables are given in Table 7.18. The travel cost is measured as cents and the travel time is measured in minutes. Both variable values may be collected from actual data. The comfort level reflects the travelers' opinion and may be derived from a survey. In this example, the comfort level for automobile travel is twice as much as the transit travel.

Travel Mode	Travel Cost (X_1)	Travel Time (X_2)	Comfort Level (X_2)	
	(cent) (min)			
Private Vehicle (1)	100	20		
Public Transit (2)	60	40		

Table 7.18 Hypothetical variable values for a utility function

Inserting the variable values into Eq. (7.16), we can calculate the utility for private vehicle and public transit:

$$
U_1 = -0.03 - 0.02X_{11} - 0.015X_{21} + 0.04X_{31}
$$

= -0.03 - 0.02 × 100 - 0.015 × 20 + 0.04 × 10
= -0.03 - 2 - 0.3 + 0.4
= -1.9

$$
U_2 = -0.035 - 0.025X_{12} - 0.02X_{22} + 0.05X_{32}
$$

= -0.035 - 0.025 × 60 - 0.02 × 40 + 0.25 × 5
= -0.035 - 1.5 - 0.8 + 0.25
= -2.05

It is worthwhile to point out that the utility-based mode choice model calculates the probability of travelers selecting each travel mode. This is due to the fact that travelers cannot be informed perfectly about the travel modes and the decision making process for a traveler to select a travel mode cannot be perfectly modeled. In the mode choice model, there are many ways to establish the relationship between proportion of travelers using each travel mode and its utility. One such model is the Multinomial Logit (MNL) model.

According to the MNL model, the probability that mode choice *k* is made under the assumption of utility maximization is the fraction of the total utility. This probability is expressed as:

$$
Prob(k) = \frac{\exp(U_k)}{\sum_{k=1}^{n} \exp(U_k)}
$$
(7.19)

where,

 $Prob(k)$ — the probability of mode *k* being selected;

 U_k — the utility value for mode *k*;

 n — total number of mode choices.

Using this model, we can calculate the probabilities associated with the two traffic modes in the previous example. We first need to calculate $exp(U_k)$. For the private vehicle travel, $k = 1$,

$$
\exp(U_1) = \exp(-1.9) = 0.150
$$

for public transit, $k = 2$,

$$
\exp(U_2) = \exp(-2.05) = 0.129
$$

then,

$$
\exp(U_1) + \exp(U_2) = 0.150 + 0.129 = 0.279
$$

From Eq. (7.19) we can calculate the probabilities for the two modes as:

For the private vehicle, $Prob(1) = 0.150 / 0.279 = 0.537 = 53.7\%$ For the public transit, $Prob(2) = 0.129 / 0.279 = 0.463 = 46.3\%$

The model result shows that the market share for private vehicle in this case is 53.7% and the pubic transit share is 46.3%. Use of the multinomial logit model ensures that the sum of trips allocated to all modes of transportation always equal to the total trips.

The multinomial logit model demonstrates that the probability of choosing one transportation mode depends on the utility of the mode and other available modes. The effects of a mode on all other modes are equally distributed, which is far from reality. For example, a trip maker is more easily to switch between a bus and an express bus than from riding a bus to driving a vehicle. To address this uneven effect, researchers have developed nested logit models to expand the MNL model (Koppelman and Wen, 1998; Hensher and Greene, 2002). In a nested logit model, travel modes that are similar are grouped together to form a composite mode choice, which is compared with other mode choices. Figure 7.7 gives an example of a three level nested logit structure.

Figure 7.7 A three-level nested logit structure

The probabilities are calculated from the top level then down to the lower level. At each level, the MNL model solution method is used. For the nested logit structure in Fig. 7.7, the MNL model at first level has three choices. The probability for private vehicle is:

$$
Prob(Pv) = \frac{exp(U_{p_v})}{exp(U_{p_v}) + exp(U_{p_t}) + exp(U_{B})}
$$
(7.20)

For public transit,

$$
Prob(Pt) = \frac{exp(U_{p_t})}{exp(U_{p_v}) + exp(U_{p_t}) + exp(U_{B})}
$$
(7.21)

For bicycle,

$$
Prob(B) = \frac{exp(U_{B})}{exp(U_{p_{v}}) + exp(U_{p_{t}}) + exp(U_{B})}
$$
(7.22)

At the second level there are two sets of MNL models. For the private vehicle model the probability of driving alone is:

$$
Prob(Da|Pv) = \frac{exp(U_{Da})}{exp(U_{Da}) + exp(U_{Cp})}
$$
\n(7.23)

$$
Prob(Cp|Pv) = \frac{\exp(U_{Cp})}{\exp(UD_{Da}) + \exp(U_{Cp})}
$$
\n(7.24)

The probabilities in Eqs. (7.23) and (7.24) are called conditional probabilities. That is, the probability of driving alone among those trip makers who use private vehicles. To calculate the unconditional probability of driving alone, we need to use the following formula:

$$
Prob(Da) = Prob(Da|Pv) \cdot Prob(Pv)
$$
 (7.25)

The unconditional probability of car pooling is calculated with the following formula:

$$
Prob(Cp) = Prob(Cp|Pv) \cdot Prob(Pv)
$$
 (7.26)

Similarly, the probabilities for bus and light rail modes are calculated with the following formulas. Bus:

$$
Prob(Bus|Pt) = \frac{\exp(U_{bus})}{\exp(U_{bus}) + \exp(U_{Lr})}
$$
\n(7.27)

Light rail:

$$
Prob(Lr|Pt) = \frac{exp(U_{Lr})}{exp(U_{bus}) + exp(U_{Lr})}
$$
(7.28)

The unconditional probability of bus mode can be derived from the probability for public transit and the conditional probability for bus:

$$
Prob(Bus) = Prob(Bus|Pt) \cdot Prob(Pt)
$$
 (7.29)

and the unconditional probability of light rail mode is calculated from the formula below:

$$
Prob(Lr) = Prob(Lr|Pt) \cdot Prob(Pt)
$$
 (7.30)

At the third level, probabilities for local bus and bus rapid transit can be calculated in the same fashion. The following formulas show the conditional and unconditional probabilities for the two travel modes: Local bus—conditional probability:

$$
Prob(Lb|Bus) = \frac{exp(U_{Lb})}{exp(U_{Lb}) + exp(U_{Brt})}
$$
\n(7.31)

Bus rapid transit—conditional probability:

$$
Prob(Brt|Bus) = \frac{exp(U_{Brt})}{exp(U_{Lb}) + exp(U_{Brt})}
$$
\n(7.32)

Local bus—unconditional probability:

$$
Prob(Lb) = Prob(Lb|Bus) \cdot Prob(Bus)
$$

$$
= Prob(Lb|Bus) \cdot Prob(Bus|Pt) \cdot Prob(Pt)
$$
(7.33)

Bus rapid transit—unconditional probability:

$$
Prob(Brt) = Prob(Brt|Bus) \cdot Prob(Bus)
$$

$$
= Prob(Brt|Bus) \cdot Prob(Bus|Pt) \cdot Prob(Pt)
$$
(7.34)

369

7.4.4 Trip Assignment

The trip assignment model is concerned with trip-makers' choice of route between all zones, after the origin-destination matrix for a network has been constructed. The result is the traffic volume on specific road links that make up the network. There are many possible route choices to travel from one zone to another. The concept of "Impedance" plays an important role in estimating vehicular traffic assignments on a roadway network. Impedance, similar to that discussed in trip distribution, is normally related to travel time and travel cost: the longer the trip or the higher the cost, the larger the impedance for the trip along that path. One criterion for trip assignment is to minimize the impedance by assigning trips to different routes. Most trip assignment models are based on one of two theories of minimizing the impedance (Gazis, 2002).

(1) An individual trip maker chooses the route that has the minimum impedance.

(2) The average journey impedance for all users is minimized.

The first theory indicates that the route choice is to minimize the individual traveler's impedance (user optimal) while the second theory is to minimize the collective impedance for all travelers (system-optimal). Using travel time as an example, a model may assign travelers to different paths in a way that each traveler takes the shortest time to travel from the trip generation zone to the trip attraction zone. Alternatively, a model may assign travelers to different routes in a way that the summation of travel time of all travelers is minimized. Studies have shown that the two optimizations are not satisfied at the same time. That is, a user-optimal solution will not lead to a system-optimal solution (Gazis, 2002).

In reality, the decision of travel route is much more complicated than simple optimization. The travel conditions on the routes are likely different. For example, the distance when traveling on local roads may be shorter than on a highway; however, one may travel faster on the highway. Someone may still be willing to travel on the highway even if it will take a little longer time and distance. On the other hand, someone may try as much as possible to avoid traveling on the highway. The trip assignment procedure identifies relevant variables that affect travelers' path choices and predicts how route decisions are made based on certain assumptions, e.g., shortest path.

We should be able to realize, from our own experience that, it is very hard to predict which route people may choose to travel from one place to another. Many factors may affect choice on any given day. Nevertheless, there are some variables that are most relevant to people's choice. The trip assignment procedure of a travel demand model uses those variables to predict the possibilities of people choosing a particular path. Examples of such variables are travel time, travel distance, traffic lights, the width or number of lanes, travel volume, etc.

The level of service is normally used to estimate the impedance for private automobile travel. To estimate the impedance for transit travel, the travel time is normally estimated for different segments, such as time spent in-vehicle, waiting for a vehicle, walking to the transit stop, or transferring to another vehicle. Different weights may be applied to the time segments to reflect people's tolerance level of the time spent. For example, spending 10 minutes on a bus will have a different affect than walking 10 minutes to a bus stop on a person's choice of which way to travel.

The simplest trip assignment model is the "all-or-nothing" method (Wright, et al., 1997). According to this method, all trips between any two zones are assigned to the route with minimum travel time or minimum cost. If two routes have same travel time, the trips are equally split between the two routes. No matter how many other paths are available, no traffic is assigned to them. The traffic volume on any path is the total of the zonal traffics going through the route. If the objective is to minimize the cost of travel, the solution is called **minimum cost route (MCR) algorithm** (Oppenheim, 1995).

Figure 7.8 uses the four-zone example to illustrate the trip assignment using the "all-or-nothing" method. There are eight links connecting the four zones. One intersection, A, connects all four zones. The numbers in the oval box indicates the link number and travel time. Let's use the trip distribution result shown in Table 7.17. Among different ways of traveling between Zone 1 and Zone 2, the

Figure 7.8 "All-or-Nothing" trip assignment

path of the shortest time connects the two zones through Intersection A. Therefore, all 3,160 trips from Zone 1 to Zone 2 are assigned to the two links (3 and 4) from Zone 1 to Intersection A then to Zone 2. Similarly, the 8,265 trips from Zone 2 to Zone 1 are assigned to the two links (4 and 3) from Zone 2 to Intersection A then to Zone 1. The shortest path between Zone 2 and Zone 3 also goes through Intersection A. This path then carries the 5,818 trips from Zone 2 to Zone 3 (Links 4 and 5) and 6,251 trips from Zone 3 to Zone 2 (Links 5 and 4). The total trips along Link 4, from intersection A to Zone 2 are then calculated as summation of the trips from Zone 1 and trips from Zone 3:

Total trips on Link 4 from Intersection A to Zone $2 = 3,160 + 6,251 +$ $5,271 = 14,682$

Similarly, the trips on Link #6 are a combination of trips from Zone 1, Zone 2 and Zone 3 to Zone 4. Other links that are not on the quickest path, such as Link #1 and Link #8 on the other hand, are not assigned any trips.

Apparently, this method does not consider many other variables affecting traffic, nor does it even consider the capacity of any single road. To address these problems capacity restraint methods have been developed. According to capacity restraint methods, the travel time is a function of the volume and capacity of roadways:

$$
t = f(V, C) \tag{7.35}
$$

where,

 t — travel time on a link;

 V — trip volume;

C — roadway capacity.

The initial impedance such as travel time is estimated at the free flow condition. As trips are assigned to road links, the travel time on some links, where the volume exceeds the design capacity, will increase. During the trip assignment process, the initial travel time on all links is calculated and an arbitrary portion of the trip volume (such as 20%) is assigned to the links with the minimum travel time. Then, the travel time on those links are updated to reflect the impact of the assigned travel volume. Based on the updated travel time, new minimum-traveltime links are identified. Another arbitrary portion of trip volume is assigned to the links with the new minimum travel time. This process is repeated until all traffic volumes are assigned.

Equilibrium assignment is another type of traffic assignment method. This method can simulate the fact that people may choose a different route to avoid traffic congestion. When trips are assigned to a roadway link, the volume associated with the link is therefore increased. As a result, the volume/capacity ratio for the link will increase. Because the trip assignment is based on the ratio, the likelihood of assigning additional trips to this link is reduced. The method seeks the trip assignment that every traveler uses paths that minimize the objective function, such as travel time.

Diversion methods also can be used in trip assignment. Similar to its use in mode choice, a diversion method allocates trips between two possible paths. The proportion of trips assigned to a path is a function of the ratio of the impedance of the alternative path over its impedance:

$$
P_i = f(1/(W_i / W_a))
$$
\n(7.36)

where,

 P_i — proportion of trips assigned to path *i*;

 W_a — impedance of the alternative path;

Wi — impedance of path *i*.

Figure 7.9 illustrates a diversion curve for two alternative paths, using travel time as the impedance. The *x*-axis represents the ratio of the travel time, *R*:

$$
R = T_1 / T_2 \tag{7.37}
$$

where,

 R — the travel time ratio;

 T_1 — travel time for path 1;

 T_2 — travel time for path 2.

Figure 7.9 Diversion curves for trip assignment

The *y*-axis is the proportion of the trip volume on path 1. The curve shows that when the travel time on the two paths is the same, $R = 1$, about forty percent of trips are assigned to path 1. As *R* increases from the equal time point, the travel time on path 1 is greater than that on path 2, the proportion of trips on path 1 decreases. As *R* decreases from the equal time point, the travel time on path 1 is less than that on path 2, the proportion of trips on path 1 therefore increases. The model shows that the more savings of travel time on a particular path, the higher proportion of trips is expected.

The last words about the travel demand modeling are that once the structure of a model is developed, it must be **calibrated**. The calibration process is to compare the output of the travel demand model with surveyed traffic data, using different parameter values (coefficients). Statistical tests are used to find the set of parameters that can best match the model output with empirical data. Before a calibrated model can be used, it must be **validated**. Model validation is a process of comparing the modeling out with a dataset that is independent from the data used for model calibration. The model validation process tests the model structure and the theories behind it by demonstrating its ability to replicate actual traffic patterns (Edwards, 1999). In order to provide meaningful prediction, a model should only be used after it is validated.

7.5 Critique and Limitations

This chapter introduces the basis of transportation analysis. In particular, the four-step travel demand model describes the traffic features of a transportation system. The trip generation step produces zonal trip production and attraction, which determines the amount of traffic expected to occur in the study region. The challenge in this step is that the estimation of trip production and attraction is based on population characteristics and the types of land use. This calculation may introduce errors because it may not represent other important factors affecting people's travel behavior. Two shopping centers of the same size, however, with different stores which carry different merchants and different brands, may attract people with different travel behaviors. In addition, data used in the model may not accurately reflect reality. For example, in the United States, the Bureau of Census is the main source for population data. Census data are summarized at a given area, such as census tract. It becomes a difficult task to know where people live in a census tract. A common practice is to assume people are evenly dispersed within a census tract. We know this is not correct. One improvement is to limit people to residential land. This can be done once the land use data are available. However, it can be complicated if there is residential land with different densities. There is no single standard procedure to allocate a total population to different residential land densities.

Another limitation of the trip generation is the assumption that all trips originate from the place of residence. We already know that there are nonhome-based trips. The traffic demand model cannot assign trip generation to a non-residential zone. For example, it is common to observe heavy traffic during lunch time in a commercial district. These trips will either be ignored or be assigned to one or more residential zones.

The trip distribution model allocates trip production and trip attraction among zones in the study area. One problem, which could be potentially serious, is the assumption that no traffic will go across the study boundary. With the increased travel distance by automobiles it is not uncommon for people to travel long distance, either to go outside the study area or to enter from outside. The model will not be able to catch this portion of traffic.

The third piece of travel demand model is the mode choice or modal split model. It divides trips into different modes of transportation. The current models simulate the choices based on the characteristics of travelers, the trip types, and the service quality of different modes. It could be difficult to assess the performance of a transportation system after introducing a new travel mode since the factors affecting travelers' mode choices are quite localized, which makes it difficult to use study results from other places.

Trip assignment is the last step in the travel demand model. The procedure allocates the trips of different modes to all available roads. The challenge here is that the model must be sensitive to the dynamic changes of the roads. Travelers normally choose the paths based on the level of service of the roads. Their choices will affect the traffic volume on the roads, which in turn affects the level of service. This looping nature often requires that the trip assignment modeling has to be completed in multiple iterations.

The travel demand model applies a linear sequence of the four steps. The trip generation model produces the number of trips from each traffic analysis zone (TAZ). The trip distribution model allocates the trips to all the zones. The mode choice model divides the trips among available travel modes. For each mode, there are normally multiple routes connecting two zones. Therefore, the trip assignment model further allocates the trips between the origin and destination zones in each travel mode to available routes. Hereby, the model assumes that travelers make rational decisions based on the perfect knowledge of traffic analysis zone characteristics, available travel modes, and possible routes. However, this may not reflect reality. For example, travel time is dependent on the travel volumes and traffic congestion can change the friction factors used in the travel demand model. To accommodate the change, a travel demand model has to run under the revised parameter values. However, the impact of congestion on the parameters may not be clear.

A travel demand modeling process does not have to go through these four phases linearly. For example, trip distribution is about where people go, which could very well be affected by the available modes of transportation or the paths connecting the zones. Rosenbloom (1988) summarized the four most common combinations of the four steps. The sequence of trip generation, trip distribution and trip assignment is the same for all four types of combination. The difference is where in the decision process the mode choice is made.

The Type I combination represents a circumstance where a trip maker's decision of making a trip is directly related to the availability of different travel modes. For example, there has been increased traffic in Chinese cities with the

increase of automobiles. Assume that an individual who normally takes the bus to work now drives a car. This individual may decide, on the way home, to drive to a place where she has never been because it is not easily accessible by bus. This represents an additional trip that is a result of the availability of the automobile as a mode of transportation. In a different example, in an American city where public transportation has been significantly improved, residents, especially those who do not drive, may make more trips. Similarly, the decision to make those trips is related to the availability of transit service.

The Type II combination reflects a situation that once an individual decides to make a trip, the choice of trip destination is affected by the availability of transportation modes. For example, trips generated can be split into automobile and transit modes before trip distribution. Assume there are two shopping centers with similar merchants. One is easily accessible by bus with no parking space and the other has numerous parking spaces, however it is not close to any bus stops. Where an individual goes shopping is pretty much determined by his/her access to a car or bus. This type is only applicable in an area where transit takes a substantial portion of traffic.

The Type III combination differs from the first two types in which a trip maker decides where to go in conjunction with considering the available mode choices. The available modes may affect trip distribution. For example, whether or not individuals own a car may influence their decision on where to buy groceries.

According to the Type $\mathbb N$ combination, the modes of transportation do not affect trip distribution. It represents a situation in which a traveler decides on the transportation mode to use after deciding where to go.

With the travel demand model, a planner can assess the effect of policies and programs on travel demand, the performance of a new or proposed transportation facility, and impacts of a proposed development on traffic. Two primary applications of the traffic demand modeling is Transportation Control Measures (TCM), which is designed to reduce vehicular travel, and Congestion Management Program (CMP), which intends to reduce congestion on the highway network by coordinating land use, air quality, and transportation planning. CMP may provide incentives or implement strategies to affect people's behavior and transportation choices. For example, highway congestion can be alleviated by minimizing single occupancy drivers. Incentives could be provided for people to take the bus or carpool. Although the highway capacity is not increased, the existing capacity is better utilized through a planned process of moving more people.

Researchers have significantly improved the travel demand model since its inception. The behavioral potential has been formalized and their operations are supported by powerful techniques of mathematical programming (Oppenheim, 1995). There are many computer software packages performing the four-step process of traffic demand modeling. Although these models require a large amount of data,

including the roadway links, traffic volume, road capacity, and origin-destination tables, the simulation of different scenarios and planning alternatives can be completed quickly.

One caution of using travel demand modeling results is that the model is designed for transportation planning purposes. Therefore it is most suitable for analyzing general traffic patterns in a large area, such as a metropolitan area. The traffic volume on a particular road normally does not reflect the real traffic volume. At least two reasons for the discrepancy can be attributed to the model structure. First, the model simulates an average condition, typically, morning or afternoon peak hours or daily average. The second cause is related to the trip distribution. The model requires that the zonal trip productions and attractions be linked to road network through a limited number of nodes. Therefore, all trips from one traffic analysis zone (TAZ) may be directed to one or two nodes. In reality, people in the same zone may take different roads for travel. If the needs arrive for a more detailed and realistic understanding of traffic on a particular road, readers may refer to other transportation books about the micro level traffic models.

Review Questions

1. Assign trip purpose to each of the trips shown as lines in the diagram. How many trips and how many trip ends have you identified?

2. Describe the purpose of street classification.

3. Describe the concept of level of service. What are the six level of service classes?

4. What is a travel analysis zone (TAZ)? What are the two general rules in delineating TAZs?

5. What are the four steps in the travel demand modeling process? What is the purpose of each step? What is the outcome from each step?

6. What is the fundamental of the Gravity model?

7. What is a friction factor F_{ij} ? How does it affect the trips between Zone *i* and Zone *j* ?

8. Describe the assumption in the trip distribution model that the total amount of trip generation must equal to the amount of trip attraction.

9. What is a Diversion Method? How is it used in mode choice model?

10. What is the "All-or-Nothing" trip assignment?

Exercises

1. Your instructor will give you a map showing the streets in a portion of the place you live. You will follow the Census Feature Class Code (CFCC) road category scheme shown in Table 7.2 to classify the streets. Allow yourself some time to visit the roads. Prepare a map of the street classification.

2. Download the household income data by census tract for a U.S. County of your choice from the U.S. Census Bureau web site (http://www.census.gov/). Then you will download the census tract map files (i.e., in ArcGIS shapefile format) from the same website. You look for the link to TIGER files. TIGER, which stands for Topologically Integrated Geographic Encoding and Referencing system, is a file format the U.S. Census Bureau used for geographic data. Use the data and the trip generation rates in Table 7.5 to calculate the home-based-work person trips. Then you will make a map showing the spatial distribution of the Home-Based Work (HBW) trips. Describe the spatial distribution of the HBW trips.

3. Use the same data in exercise 2 to complete the calculation for Home-Based-Non-Work (HBNW) and Non-Home-Based (NHB) trips for the county. Use the automobile occupancy data in Table 7.6 to calculate the vehicle trips for each trip purpose.

4. The attached spreadsheet file, ex7 4 trip distribution.xls (Fig. 7.10) contains a simplified gravity model for trip distribution of a 3-zone area:

$$
T_{ij} = \frac{P_i \cdot A_j \cdot F_{ij}}{\sum_{i=1,n} (A_j \cdot F_{ij})}
$$
(7.14)

where, F_{ij} is friction factor and is calculated by combing formulas 7.12 and 7.13:

$$
F_{ij} = t_{ij}^{-a} \tag{7.38}
$$

The shaded cells in the spreadsheet represent the data or parameters you may change. They are travel time between a pair of zones, t_{ij} , trip production for each zone, P_i , trip attraction for each zone, A_j , and the constant for calculating friction factor, *a*.

Microsoft Excel - ex7_4_trip_distribution.xls ш								
	图 File Edit	View Insert Format			Tools Data	Window	Help	×
	E Times New Roman			-12			\bullet \equiv \equiv \equiv \mid % $\frac{42}{20}$ \mid \mid \mid \bullet \bullet \bullet	" E
", ataaabisbishigaten with Changes								
	H ₈	f_{X}						
	A		B	Ċ	D	E	F	
1	Trip Distribution Calculation - Gravity Model							
$\overline{2}$	Travel time						Production	
3	t_{ii}		$i =$	1	\overline{c}	3	P_i	
4		$i =$	1	1	\overline{c}	3	200	
5			\overline{c}	8	4	\overline{c}	300	
6			3	6	7	$\overline{\mathbf{3}}$	400	
7	Attraction, A _i			300	200	200		
8								
9	Friction factor			$a =$	0.5			
10	$F_{ij} = t_{ij}$ ^{-a}		$i =$	1	\overline{c}	3		
11		$i =$	1	1.00	0.71	0.58		
12			\overline{c}	0.35	0.50	0.71		
13			3	0.41	0.38	0.58		
14								
	15 Relative attraction							
16	$A_i^*F_{ii}$		$i =$	1	\overline{c}		$3\Sigma A_j^*F_{ij}$	
17		$i =$	1	300	141	115	557	
18			\overline{c}	106	100	141	347	
19			3	122	76	115	314	
20								
21	Trips							
22	$T_{ii} = (A_i^* F_{ii}) / (\Sigma A_i^* F_{ii})$ j =			1	\overline{c}		3 Trips started, T _i	
23		$i =$	1	108	51	41	200	
24			\overline{c}	92	86	122	300	
25			3	156	96	147	400	
	26 Trips ended, T _i			356	234	311	900	
27	H 4 ▶ H Sheet1 / Sheet2 / Sheet3 /				\vert <			\rightarrow
	:Draw・ V AutoShapes・ \ \ □ ○ �� 4 ☆ 図 図 $\mathbf{A} \cdot \mathbf{A} \cdot \mathbf{B}$							
	Ready							

Figure 7.10 A 3-zone trip distribution gravity model

Perform the following tasks:

(1) Open the spreadsheet file and make sure to understand the calculations.

 (2) Copy and paste the entire model, Cells A1:F26 to Cell H1. Now you have two identical models side by side. Change any one or several of the trip production values in Cells F4ĩF6 and compare the new trip distribution in Cells C23:F26 with the original trip distribution in Cells J23:M26. Describe the difference.

 (3) Change any one or several of the trip attraction values in Cells C7 E 7 and compare the new trip distribution with the original trip distribution. Describe the difference.

(4) Change the constant, *a* (Cell F9), for instance, let $a = 2$, and compare the new trip distribution with the original trip distribution. Describe the difference.

(5) Summarize your understanding of the gravity model.

5. Figure 7.11 illustrates a hypothetical street network. The network has 8 nodes (101 to 108) and 9 links (1 to 9). The travel time on each link are listed in Table 7.19. Identify the shortest trip from node 101 to node 107. Assume the trip distribution from node 101 to 107 is 1,000, assign trip volume to each link sing the All-Or-Nothing assignment approach.

Figure 7.11 Example street network

Link No.	Travel time (min)	Link No.	Travel time (min)	
	12			
			10	
			10	

Table 7.19 Travel time and volume

References

- Chang, Cathy L. and Daniel T. Meyers. 1999. Transportation models. In: Edwards, John D. (ed.) *Transportation Planning Handbook*. 2nd ed. Washington, DC: Institute of Transportation Engineers.
- Downs, Anthony. 2004. *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*. Washington, DC: The Brookings Institute.
- Edwards, John D. (ed.) 1999. *Transportation planning handbook,* 2nd ed. Washington, DC: Institute of Transportation Engineers.
- Ford, L.R. and D.R. Fulkerson. 1962. *Flows in Networks*. Princeton, NJ: Princeton University Press.

Gazis, Denos C. 2002. *Traffic Theory*. Boston, MA: Kluwer Academic Publishers.

- Hensher, David A. and April J. Reyes. 2000. Trip chaining as a barrier to the propensity to use public transport. *Transportation* 27: 341 - 361.
- Hensher, David A. and Greene, William H. 2002. Specification and estimation of the nested logit model: alternative normalizations. *Transportation Research Part B: Me thodological*, $36(1): 1 - 17.$
- Institute of Transportation Engineers. 1997. *Trip Generation: Trip Generation Rates, Plots, and Equations*. Washington DC: Institute of Transportation Engineers.
- Koppelman, F.S. and C.H. Wen. 1998. Alternative nested logit models: structure, properties and estimation. *Transportation Research* 32B(5): 289 - 298.
- Norland, Robert B. 2001. Relationships between highway capacity and induced vehicle travel. *Transportation Research*, Part A 35: 47 – 72.
- Oppenheim Norbert, 1995. *Urban Travel Demand Modeling*. New York, NY: John Wiley & Sons, Inc.
- Papacostas, C. S and P.D. Prevedouros. 2001. *Transportation Engineering and Planning*, 3rd ed. Upper Saddle River, NJ: Prentice Hall.
- Rodwin, Lloyd and Bishwapriya Sanyal (eds.) 2000. *The Profession of City Planning: Changes, Images, and Challenges*. New Brunswick: Center for Urban Policy Research, Rutgers, The State University of New Jersey.
- Rosenbloom, Sandra. 1988. Transportation Planning. In: So, Frank S. and Judith Getzels. (eds.) *The Practice of Local Government Planning*, 2nd ed. Washington DC: The International City Management Association.
- Taplin, John H. E. and Min, Qiua. 1997. Car Trip Attraction and Route Choice in Australia. *Annals of Tourism Research, 24(3): 624 – 637.*
- Turner, Jeff and Grieco, Margaret. 2000. Gender and time poverty: the neglected social policy implications of gendered time. *Transport and Travel*, *Time & Society*, 9(1): 129 – 136.
- U.S. Census Bureau. 2004. Census 2003 TIGER/Line®. Files Technical Documentation. Washington, DC: U.S. Census Bureau. Available online at: http://www2.census.gov/ pub/outgoing/geo/marshall/tgr1990/DOCUMENT/. Accessed March 2005
- U.S. Department of Transportation, Federal Highway Administration. 1985. *Development and Application of Trip Generation Rates*. Washington, DC: U.S. Dept. of Transportation, Federal Highway Administration.
- Wright, Paul H., Ashford, Norman J. and Stammer, Jr. Robert J. 1997. *Transportation Engineering: Planning and Design.* 4th ed. New York, NY: John Wiley & Sons. Inc.