**Transportation Planning**

**1. Trip Assignment**

All of the preceding lectures have dealt with the development of trip tables. Assignment is the fourth step in a four-step travel demand model.

Highway assignment is the process by which vehicle trips for each origin-destination interchange included in the vehicle trip tables are allocated to the roadway network.

The allocation process is based on the identification of paths through the network for each origin-destination interchange. The assignment process may be mode-specific with, for example, paths for single occupant vehicles being determined using different criteria than paths for multi occupant vehicles or trucks.

**2. Model Function**

There are a number of methods by which a trip table can be assigned to a network. All of these methods are basically variations of the formula:





While the algorithms and computer code required to efficiently solve the assignment problem, as well as the requirements for storing the probability matrix, do not often lead to the assignment problem being defined in this way, describing the process in this manner does allow for the identification of features that distinguish the various assignment methods.

When the probability matrix is predetermined in some manner that cannot be changed, the method is called a fixed path assignment.

When the probability matrix takes on the value of one when the link is used and zero when the link is not used it is said to be an all or nothing (AON) assignment.

When the cells of the probability matrix are calculated from a stochastic formula that calculates the percentage of trips to be assigned to a set of links contained in reasonable paths, the method is called a stochastic assignment.

When the probability matrix takes on discrete values associated with the percentages of the trip table which are assigned in successive AON assignments, where between iterations the congested time is updated based on a comparison of the assigned volume on a link to its capacity, new AON paths are then calculated, and those percentages are applied to each of the successive AON probabilities (i.e., one or zero), the method is called incremental capacity-restrained assignment.

When the cells of the probability matrix are calculated from the percentage of the trip table assigned to successive applications of AON as in the incremental capacity-restrained assignment, but those percentages are selected through an iterative process that will result in satisfying Wardrop’s first principle, which states that “the journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route” (Wardrop, 1952), the method is said to be a user equilibrium assignment. A variant of this method, called stochastic user equilibrium, uses stochastic assignment rather than AON assignment in successive steps to arrive at equal journeys on used paths, in which case the perceived times are said to be reasonably equal. A common method to determine the allocation of a trip table to successive iterations is the Frank-Wolfe algorithm (Frank and Wolfe, 1956).

An additional consideration in assignment is the number of trip tables that will be assigned and the manner in which the trip tables are assigned. If the trip table is assigned to the network links prior to a user equilibrium assignment, for example by assigning that trip table to fixed or AON paths that do not consider congestion, that trip table is said to be preloaded. Those trip tables (i.e., classified by vehicle and/or purpose) that are assigned jointly in a user equilibrium assignment are said to be a multimodal multiclass assignment.

The first three assignment processes previously described— fixed path, AON, and stochastic—are insensitive to congestion impacts that occur when demand for a network link approaches the capacity of the link. The last two assignment methods— capacity restrained and user equilibrium—explicitly attempt to account for congestion impacts in the traffic assignment process. The last two procedures are typically preferred for future forecasts because they inject a level of realism into the assignment process through reductions of travel speeds as traffic volumes on links increase. In addition, the last two procedures are required if air quality impacts of various alternatives or land use scenarios need to be estimated from traffic assignment results.

While the first three assignment procedures are insensitive to congestion impacts, these can provide important analysis capabilities. For example, AON assignments are useful for determining travel desires in the absence of congestion impacts and are commonly used to preload truck trips and other external through-trip movements in regional models.

Such information can also be useful in targeting transportation improvements. In uncongested networks, stochastic assignment may be the only method available to represent user choices of similar alternative paths.

In all capacity-restrained and user equilibrium assignments, link travel times are adjusted between iterations using a vehicle delay function (sometimes referred to as a “volume-delay,”“link performance,” or “volume-time” function). These functions are based on the principle that as volumes increase relative to capacity, speeds decrease and link travel times increase.

One of the most common of these vehicle-delay functions was developed by the BPR, the predecessor agency of the FHWA. The BPR equation is:





While $t\_{i}$ represents the link i travel time and is expressed in units of time (usually minutes), it may also reflect other costs associated with travel, especially tolls and auto operating

costs such as fuel costs. The value $t\_{i}$ (and $t\_{0i}$) may therefore be represented by something like Equation above:





Parameter K2, therefore, represents the inverse of the value of time. Note that the value of time is also an implied parameter in mode choice.

It is customary to express capacity in vehicles per hour. In models where daily (weekday) highway assignment is used (and therefore the volume variable is expressed in vehicles per day), the hourly capacity estimates must be converted to daily representations.

This conversion is most commonly done using factors that can be applied to convert the hourly capacity to effective daily capacity (or, conversely, to convert daily trips to hourly trips, which is equivalent mathematically). These factors consider that travel is not uniformly distributed throughout the day and that overnight travel demand is low. The conversion factors are therefore often in the range of 8 to 12, as opposed to 24, which would be the theoretical maximum for an hourly-to-daily factor. [These factors are sometimes referred to as “CONFAC,” the variable name in the Urban Transportation Planning System (UTPS) legacy software on which many aspects of modern modeling software are still based.]

These types of conversion factors continue to be needed in models where time periods for assignment greater than 1 hour in length are used. In such cases, the factors convert the hourly capacity to the capacity for the appropriate time period. For example, if a morning peak period is defined as 6:00 to 9:00 a.m., the conversion factor will convert hourly capacity to capacity for the 3-hour period. It is important to consider that travel is not uniformly distributed throughout the 3-hour period, although it is likely to be more evenly distributed over a shorter time period, especially a peak period that is likely to be relatively congested throughout. The theoretical maximum for the factor is the number of hours in the period (three, in this example), and in a period where there is roughly uniform congestion throughout the peak period, the factor could be close to three. Typical factors for a 3-hour peak would range from two to three. The factors for longer off-peak periods would likely be well lower than the theoretical maximum.

Depending on the application, the value of $c\_{i}$ may not represent the true capacity of the link in a traffic operations sense . In the original BPR function, $c\_{i}$represented the limit of the service volume for LOS C, which is often approximately 70 percent of the “ultimate” capacity (at LOS E), although the conversion between these two values is not simple.

 Current best practice is to use the LOS E capacity for the following reasons (Horowitz, 1991):

1. Ultimate capacity has a consistent meaning across all facility types while design capacity does not. For example, it is a relatively simple matter to relate the capacity of an intersection to the capacity of the street approaching that intersection.

2. Ultimate capacity is always easier to compute than design capacity. Finding the design capacity of a signalized intersection is especially difficult.

3. Ultimate capacity can be more easily related to traffic counts than design capacity, which would also require estimates of density, percent time delay, and reserve capacity or stopped delay.

4. Ultimate capacity is the maximum volume that should be assigned to a link by the forecasting model. Design capacity does not give such firm guidance during calibration and forecasting.

**3. Best Practices**

While there is much ongoing research into the use of dynamic assignment and traffic simulation procedures, the state of the practice for regional travel models remains static equilibrium assignment. There has been some recent research into more efficient algorithms to achieve equilibrium than Frank-Wolfe, and some modeling software has implemented these algorithms. Since most urban areas are dependent on the major proprietary software packages for their model applications, static equilibrium procedures will continue to be used for regional modeling for the time being.

There have been some highway assignment implementations that incorporate node delay as a better way of identifying intersections that may cause congestion on multiple links, sometimes referred to as junction modeling. Some modeling software has incorporated methods to consider node delay.

For project planning and design applications to determine link volumes, the use of post-processing techniques such as those discussed in NCHRP Report 255: Highway Traffic Data for Urbanized Area Project Planning and Design (Pedersen and Samdahl, 1982) are recommended rather than reliance on raw model output. Post-processing techniques are recommended because the assigned volumes on individual links can have substantial error, as noted when comparing highway assignment outputs to traffic counts (although count data are often sampled and also have associated error).

**4. Basis for Data Development**

Horowitz (1991) fit the BPR formula (among others) to the speed/volume relationships contained in the Highway Capacity Software, Version 1.5, based on the 1985 Highway Capacity Manual (Transportation Research Board, 1985). These values were also presented in NCHRP Report 365. There is a wealth of literature on volume-delay function form and parameters, including the 2010 Highway Capacity Manual, that the analyst may wish to consult.

**5. Model Parameters**

The BPR formula parameters estimated by Horowitz are presented in Table 1. The speeds shown in this table represent facility design speeds, not model free-flow speeds.

According to the information in the MPO Documentation Database, the BPR formula is the most commonly used volume delay function. MPOs use a variety of values for the a and b parameters, and most use different parameters for freeways and arterials. Table 2 presents BPR function parameters used by 18 MPOs for which data were available from the database.

Figures 1 and 2 graph the ratios of the congested speeds to free-flow speeds on facilities at different volume/capacity.

**Table 1. BPR coefficients estimated using the 1985 *Highway Capacity Manual.***

****



***Figure 1. Freeway congested/free-flow speed ratios based on BPR functions.***



***Figure 2. Arterial congested/free-flow speed ratios based on BPR functions.***

**Table 2. BPR function parameters (morning peak period).**



ratios using the BPR functions from the 18 MPOs. In addition, each graph includes an “average” BPR function based on the curves shown in Figures 1 and 2. The average BPR functions differ from the parameter averages shown in Table 1 in that the functions were derived via linear regressions to match the averages of the congested/free-flow speed ratios for the different volume/capacity ratios.10 The resulting average BPR functions are:

• Freeways:

- Alpha = 0.312.

- Beta = 5.883.

• Arterials:

-Alpha = 0.514.

- Beta = 3.001.

**6. Transit Assignment**

While highway assignment deals with the routing of automobiles over a highway network, transit assignment deals with the routing of linked passenger trips (including walk and auto access and egress) over the available public transportation network. Differences from highway assignment include the following:

• The transit network includes not only links but also routes comprising the links, which represent the different transit services running between stops or stations;

• The flow unit in the trip table which is being assigned is passengers, not vehicles;

• The impedance functions include a larger number of level-of-service variables, including in-vehicle time, wait time, walk access and egress time, auto access and egress time, fare, and transfer activity; and

• Some paths offer more than one parallel service, sometimes with complex associated choices (e.g., express bus versus local bus service).

**7. Model Function**

Transit assignment is closely tied to transit path building. Typically, person trips estimated using a mode choice model are assigned to the transit paths built as input to the mode choice model. The typical transit assignment process is different from traffic assignment processes, where auto paths based on estimated congested travel times are input to a mode choice model and the output vehicle trips are assigned to the roadway network using an equilibrium or other capacity-restrained assignment method. The mode choice-traffic assignment process may require a feedback or iterative process to ensure that estimated roadway speeds used for mode choice (as well as for trip distribution) match the roadway speeds resulting from the traffic assignment process.

Speeds on the transit network may also be affected by the roadway speeds, depending on the software and network coding methodologies. The transit speeds used to develop the transit paths used to construct the travel time and cost skims for input to mode choice and the resulting transit assignment should match.

In the past, transit path-building and assignment were generally performed in production-attraction format with the production zone being defined as the home zone for home-based trips and the attraction zone being defined by the nonhome location. This procedure can be used to determine boardings by line, revenues, and maximum load points.

It has often been performed by time of day with transit paths and assignments being performed for morning peak and mid-day periods. Such an approach accounts for time-of-day differences in transit services with the afternoon peak period being assumed to be symmetrical to the morning peak period (which is an oversimplification). In regions offering night time transit service, the night service may either be modeled as a separate time period or aggregated with the mid-day service for assignment purposes. Finally, some areas provide the same basic levels of transit service throughout the day and, as a result, perform nontime-specific, or daily, transit path-building and assignments.

More recently, some regions have started building transit paths in origin-destination format. This approach has been used to account for directional differences in service by time of day. Service differences may be due to different frequencies of service, different service periods, or different transit speeds due to different levels of traffic congestion. The information is particularly important for tour-based and activity-based modeling procedures, although it can also be used with trip based modeling procedures.

**8. Best Practices**

Table 3 summarizes the time-of-day directional assignment procedures for 23 MPOs. Of the 20 MPOs reporting the use of time-of-day transit paths, 17 indicated the trip purposes assigned to each time-of-day network. Four of the 17 MPOs assigned home-based work trips to the peak period network and the remaining 13 estimated transit trips for each trip purpose by time of day and assigned the trips using time of- day transit paths.

**Table 3. MPOs using transit assignment procedures.**

****

Transit path-builders can be characterized into two basic groups: shortest path and multipath. Shortest path methods find the shortest path through the network, based on a specified linear combination of impedance components including items such as walk or drive access time, wait time, in-vehicle time, transfer time, additional transfer penalties, walk egress time, and fare. The coefficients of the linear combination are usually based on the relative coefficients of these variables in the mode choice model.12 Multipath procedures find multiple “efficient” paths through the transit network based on similar criteria. The multipath methods may include multiple paths for each interchange even if the alternate paths do not minimize total travel impedance. The inclusion or exclusion of alternate paths is based on a specified set of decision rules.

The use of shortest path or multipath methods should be coordinated with the type of mode choice model used. Some mode choice models incorporate path choice in the mode choice structure. For example, in regions with both bus and rail service, the mode choice model might include walk to bus only, walk to rail only, and walk to bus/rail as separate modes.

If the mode choice model is structured to include path choice, the use of a shortest path procedure is reasonable although careful use of a multipath method is also appropriate.

Alternatively, some regions simply model transit use for all combined transit modes in the mode choice model. In these regions, use of a multipath method can be used to determine path choice. Of the 22 MPOs reporting their transit path building procedures, 17 used shortest path for their peak period and off-peak period walk-to-transit paths and five used multipath procedures. For drive access to transit paths, 20 of the 22 MPOs used shortest path for their peak period and off peak period drive-to-transit paths and two used multipath procedures.

FTA has developed a number of guidelines for transit path-building and mode choice for Section 5309 New Starts applications. The FTA guidelines have influenced path-building procedures and parameters and should be reviewed prior to model development, especially if a New Starts application is being considered for a region. Two issues for transit path-building and the transit assignment process are:

• Source of bus speeds—Are bus speeds related to auto speeds in a reasonable manner, and do they reflect observed speeds?

• Consistency with mode choice parameters—Are transit path-building and assignment parameters consistent with the relationships used in the mode choice model?

Table 4 summarizes the sources of bus speeds and the consistency of the path-building parameters with mode choice parameters for the 21 MPOs reporting the information.

Information is reported for only the morning peak and midday networks since all of the MPOs had those two networks.

**Table 4. Transit assignment consistency reported by MPOs.**



**9. Basis for Data Development**

The basis for data development for the model parameters described below is the information obtained from 23 MPO models in the MPO Documentation Database.

**10.Model Parameters**

The main model parameters for transit path-building are the relationships between the components of transit travel impedance. Common parameters, which are usually expressed in terms of their relationship to in-vehicle time, include:

• Monetary cost/fare (value of time) including transfer costs;

• Initial wait time;

• Transfer wait time;

• Transfer penalty time;

• Dwell time;

• Walk time; and

• Auto time.

Typically, the auto time and dwell time parameters are set to 1.0, as both are actually in-vehicle time. While some MPOs consider fares in their transit path-building and assignment procedures, there is little variation in fares in some locations, and so fare is often excluded from the path-building impedance.

Two of the main parameter relationships that affect transit path-building and transit assignment are the ratio of walk time to in-vehicle travel time and ratio of wait time to in-vehicle travel time. Table 5 summarizes the ratios of walk time to in-vehicle travel time, and Table 6 summarizes the ratios of wait time to in-vehicle travel time, from models included in the MPO Documentation Database. As can be seen in the tables, there is little variation in the mean values of ratios, with all of the means falling in the range 2.0 to 3.0. Detailed inspection of the reported ratios shows that most of the ratios are 2.0, 2.5, or 3.0. This result is not surprising since FTA New Starts guidelines ask applicants to “provide compelling

evidence” if the ratio of out-of-vehicle time to in-vehicle time in a mode choice model is outside of the range of 2.0 to 3.0 and the guidelines also encourage consistency between transit path building and mode choice model parameter relationships.

**Table 5. Ratios of walk time to in-vehicle time reported by MPOs.**



**Table 6. Ratios of wait time to in-vehicle time reported by MPOs.**



**Example:**

A simple network shown in Figure below has two way links. Find the shortest path from nodes A,B,C and D to all other nodes and total volume on each link, assuming all or nothing assignment technique. Knowing the trip volume is given Table (7).

**A**

**B**

**C**

**D**

4

5

3

4

3

7

4

12

9

12

10

5

3

4

7

8

Table (7)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** |
| **A** | **-** | **50** | **40** | **20** |
| **B** | **30** | **-** | **80** | **10** |
| **C** | **90** | **80** | **-** | **20** |
| **D** | **60** | **70** | **50** | **-** |





**Example**

Consider a simple transportation network that has one origin and one destination and two links/paths that provide access from the origin to the destination. One link is 7 miles long and has a capacity of 4500 veh/hr and a speed limit of 60 miles/hr. The other link is 5.5 miles long and has a capacity of 2500 veh/hr and speed limit of 45 miles/hr. Assuming that 4000 vehicle wish to make the trip from the origin to the destination. Find the loaded network? ( capacity restrained technique).

Link 1 : $t=t\_{1}\left[1+1.1492 \left(\frac{v\_{1}}{C\_{1}}\right)^{6.8677}\right]$

Link2 : : $t=t\_{2}\left[1+1.03 \left(\frac{v\_{2}}{C\_{2}}\right)^{5.5226}\right]$