# The Transportation Planning Database and Networks

## Introduction

Transportation planner's emphases on the importance of quality of data for understanding the economic and societal context of transportation system and for monitoring the performance condition of this system.
Data have proved necessary for the more technical aspects of calibrating and applying travel forecasting models and for evaluating the effectiveness and impacts of proposed changes.

## The database of Transportation Planning

The travel data were collected for two basic reasons:

1. To determine the origin and destination of highways users entering and urban areas. To solve site specific problems, such as the location of new bridge or highway bypass of a city.
2. To estimate the total number of vehicles using urban highways. Automatic or manual traffic counters were used to determine vehicle volumes on congested highway so that need for highway expansion or highway control could be clearly established.

Today most agencies, state and provincial and local transportation agencies conduct systematic programs of vehicle counting on highway and networks, a legacy of these earlier efforts.
A developed new approach towards data collection interviewing individuals in a randomly selected sample of households. This procedure was found to provide good estimates of the total travel occurring within an urban area.

The information collected in these home interview usually included the following:

- Type of housing structure
- The number of vehicle available
- Number of persons in the household
- The household income category
- Description of the trips (origin, destination, trip purposes, trip time and travel modes)

These above initial survey of household led to important transportation related data collection through the decennial census that today serves as the foundation for much of the socioeconomic and demographic data that used in transportation planning.

In addition to data on travel behavior and demographics, early transportation studies collected a large amount of data on land use and transportation network characteristics.
The data on the transportation network, called an inventory, identified roadway locations, roadway length, pavement width, speeds, parking restrictions and in the case of transit, vehicle headways, numbers seats per vehicle and overall line capacity.

The transportation planning data base will be divided into three major components:

1. The most expensive, is the data collected on transportation system itself, called inventory. Inventory include the physical conditions of transportation facilities and their characteristics of traveler and freight movement.
2. Associated with developing a community vision and a corresponding set of goals and objectives. These data include public attitudes, economic and fiscal trends, and quality of life indicators.
3. Relates to the monitoring of system performance.

The first two data component act as inputs into the process of developing a system plan; the last component provides feedback on how well the system is performing, particularly after the adopted plan is implemented.

## Classification Schemes for Data Collection

In reality, trips may start and end at every address in the world and may use all available networks, streets, services etc. In order to solve a particular transportation problem there is no need to describe the system with the highest possible detail. Instead, in order to gain insight into the problem and its solutions, and to make decisions understandable, it often is better to simplify matters. Also to make calculations feasible, the analysis manageable, and the study costs bearable, and a simplified description of reality is needed. To this end, in each transportation study a system description needs to be designed (a so called systems model) giving answers to the following questions:

- Which travel markets are relevant (only persons, or also goods? only commuting or also education? etc.);
- Which geographical area is relevant?
- Which transport supply networks should be considered?

The answers to such questions highly depend on the problem at hand: the type of policy measures to be studied, the type of assessment criteria that will be used, etc. In studying a national high-speed rail line, the whole country and at least the direct neighboring countries should be considered. The airline network should be included in the network description but most of the road network need not to be considered in the analysis. In deciding on the set up of a local bus line network, the study area need not to be larger than the town in question including its direct surroundings. A detailed description of the networks for car and bicycle travel is needed for the travel analysis.

## Study area

The definition of the study area consists of several steps:

- The delineation of the study area
- The subdivision of the study area into zones
- The definition of zone centroids.

## Study area delineation

The study area is defined as the area within which transport flows will be analyzed and modeled. It is defined by an imaginary curve on a map. All area outside the curve will not be considered as relevant for the problem at hand. It is assumed that there is no travel demand generated outside the study area. Only within the study area there will be a subdivision into zones and there will be a description of traffic networks. The size and position of the study area depends on the problem at hand.
Transportation planners often grouped data by geographic location or type of transportation facilities. For ex. Every metropolitan area is divided into analysis units, often called zones, which form the basis for the analysis of travel movement within, into and out of the region, see Figure 1.

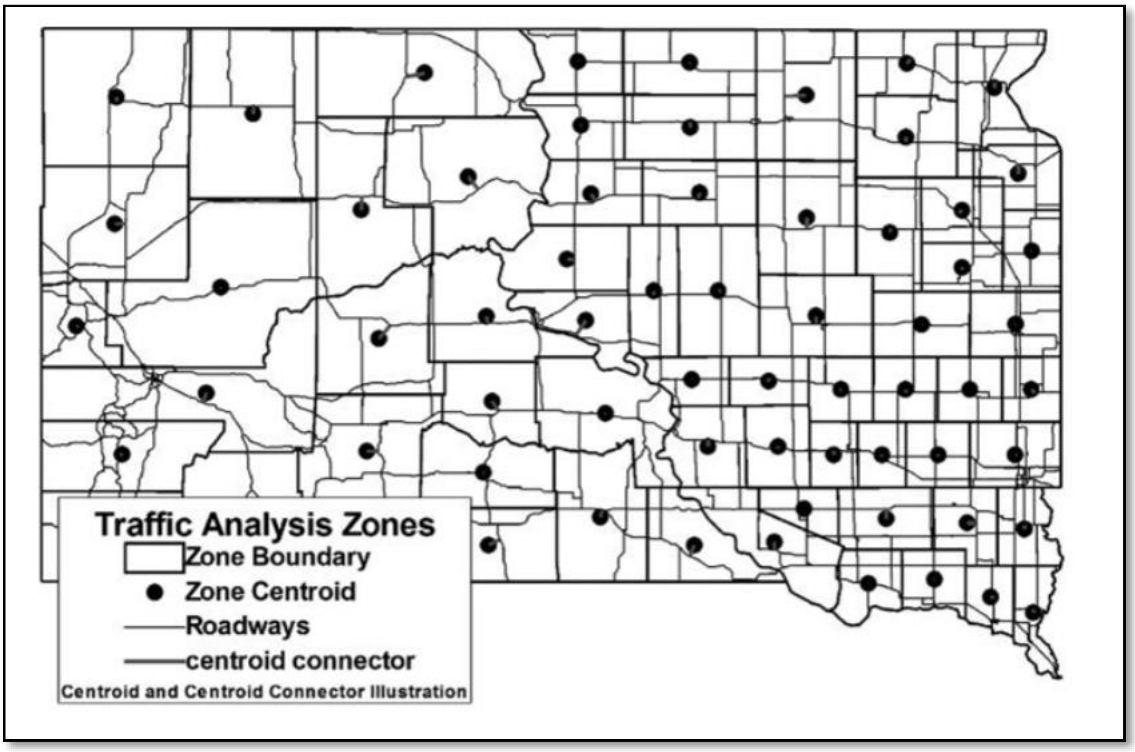


**Fig.1 Traffic Analysis Zones System for Transportation Planning.**

These zones are defined with several criteria:

1. Achieving homogenous socioeconomic characteristics for each zone's population.
2. Minimizing the number of intrazonal trips.
3. Recognizing physical, political, jurisdictional, and historical boundaries.
4. Generating only connected zones and avoiding zones that are completely contained within another zone.
5. Devising a zonal system in which the number of households, populations, area, or trips generated and attracted are nearly equal in each zone.
6. Basing zonal boundaries on census zones.

In general, the finer the subdivision, thus the more zones, the higher the accuracy in model calculations, but also the higher the costs in data collection, and computation. However, there might be an optimum level of detail because the prediction accuracy of zonal data (e.g. population characteristics) declines with the zonal size. If public transport is an important issue, the zonal subdivision needs to be extra fine because the use of public transport strongly depends on the characteristics of the local access and egress transport. Especially modeling the short distance trips will suffer from a coarse zonal system; for long distance trips the error in trip distance or trip time will be limited.

With respect to the delineation of the traffic zones there are a number of existing administrative spatial systems that can be used to define zones. This means that a traffic zone most favorably consists of one or more units of such an existing spatial system. Such systems are:

- municipalities
- census districts
- postal districts
- election districts
- etc.

The characteristics of these spatial systems, consisting of data about the geographical definition of the subareas and about their demographic or socio-economic content are available from various official institutions (Ministry of Transport, Chamber of Commerce, etc.). Adopting such data reduces the costs of data collection and system design. In making municipal transportation plans zone sizes of about 1000 to 2000 inhabitants are advised. For cities of the size of Delft this means about 50 traffic zones. For regional studies a maximum of 500 zones is applied. The Dutch National Model System uses 350 zones.

**Zonal form**

On the one hand, ideally zones should follow available delineations given by official spatial systems in order to save costs and to increase comparability. On the other hand, traffic zones should have a compact convex form in order to minimize errors in trip distances.

**Zonal centroids**

A traffic zone is represented by a single point of the zone, called centroid. It is the geographical representation of the zone. It is assumed that all trips start and end in that point.

The centroid is part of the modeled transport network. It is a fictitious network node that connects the zone to the surrounding networks. It is linked to the network by so-called connectors which fictitious links are representing the underlying local network not included in the network model. The location of the centroid is chosen such that it is indeed the center of gravity of the zone, which means that its location minimizes the distance and time errors in geographically representing the individual trip addresses (see Figure 1). Intrazonal characteristics such as distance or travel times between zones are based on the distances or travel times between the centroids of the zones.

## Zonal hierarchy

In most applications a single zoning system is used for all analysis steps. However, different modeling steps may require different zonal systems. Especially the modal choice analysis may benefit from a more detailed spatial description of trips than the other steps. To this end, many studies apply a hierarchy of zones. In applications of the Dutch National Transportation Model, trip characteristics needed for modal and destination choice analysis are established with a 1200 zone system whereas trip production and traffic assignment work with a condensed 350 zone system.

## Network description

1. **Network types**

   In most planning cases, the final aim of the analysis is to know loads of network elements. For correct choice modeling, travel distances and times in the various networks need to be known. For these purposes a computerized description of the various networks (car, bicycle, public transport) is needed that gives the geographical relations within the network as well as enables calculations of trip characteristics such as speed, travel time etc. These networks are simplified representations of the real networks of which the level of detail depends on the problem at hand. Only the networks within the study area need to be modeled. Such networks consist of nodes and links between nodes. The structure of the networks resembles that of the original real-world network. Zonal centroids are connected to nodes of the modeled network by one or more links. In the case of private travel (pedestrian, bicycle, car) the modeled network may be directly derived from the physical one by selecting network parts or by aggregating sub networks into a single link.

   Nodes and links of the modeled network correspond directly to physical counterparts. In the case of public transport the situation is more complex. Apart from a physical network on which public transport vehicles run, we have a line network that defines services and their characteristics such as service type, frequency, capacity, travel time etc. In order to model choices correctly a specific type of network description is needed called a line description.

   The line network with its stops and transfer points as well as their line characteristics are of prime importance; the underlying physical network is not that important. So, the nodes of a line network description are the stops of the lines, while the links connecting these nodes are the distinct lines available between these stops. Special links are added representing waiting at stops and transferring between lines and stops.

In most analyses today separate analyses are carried out for private and public transport networks. In the near future combined multi-modal networks will become applicable in which these networks are linked using transfer nodes which enable transferring from one type of travel (car, bicycle, bus) to another type (bus, rail). Railway stations are typical multi-modal transfer nodes. Such multi-modal networks enable integrating route and modal choice into one single choice process. A route in a multi-modal network defines automatically the use of the various modes, singly or combined.

2. **Level of network detail**

It makes no sense to include all links and nodes of a real network into the system description. It is too costly and it is not necessary in order to solve the problem efficiently. The question then is: how detailed should the description be? The coarser the modeled network, the less costs for data collection and the quicker analyses will go, but also the less accurate numerical outcomes such as travel times or traffic loads will be. Where the optimum level is depends on the problem at hand: what kind of plans need to be evaluated and what kind of impacts are considered in the assessment criteria? For the usual cases of area-wide planning, the following rules of thumb should be considered. In modeling travel demand about 75% of demand (in terms of kilometers traveled) should be part of the network analysis.

The modeled network should include about 75% of the total network capacity. Because of the hierarchical nature of traffic flow, that means, most travelers try to travel as much as possible on higher order roads, this principle will lead to a sensible reduction in network size. About 20% of the network accounts for about 80% of the traveled kilometers. The modeled network for example need not include residential streets, which as a single road category already forms half of the road network. Groups of residential streets can be represented accurately enough by a single connector link. The selected network part should be a connected network in which flows are possible between all parts of the study area.

In order to set up a modeled network one should make use of a functional classification of the real network. Each transportation network can be divided into a number of sub networks or layers having a distinct transportation function. Each link in the network can be attached a functional class according to the degree it serves a flow or access purpose for the trips on the link. One can distinguish about ten such functional classes.

Examples are:

- motorway (100% flow function)
- urban motorway
- arterial
- collector
- residential street (5% flow, 95% access)
- Residential cul-de sac (0% flow, 100% access).

After having defined the functional classification for the study area network at hand (which mostly is already available) the selection of the modeled network works top down. First, the top functional class is selected completely and the percentage of total capacity selected is calculated. Then, the next level is included and the selected capacity value is determined. The question now is when to stop.

Figure 2 gives an illustration of network selection by showing three modeled descriptions of the same real network (Eindhoven): fine, moderate, and coarse level of detail. The mid level network, having only one-fifth of the number of nodes compared to the fine level, performs best: analysis cost and computing time are only one-fifth but accuracy is nearly the same as with the fine network.

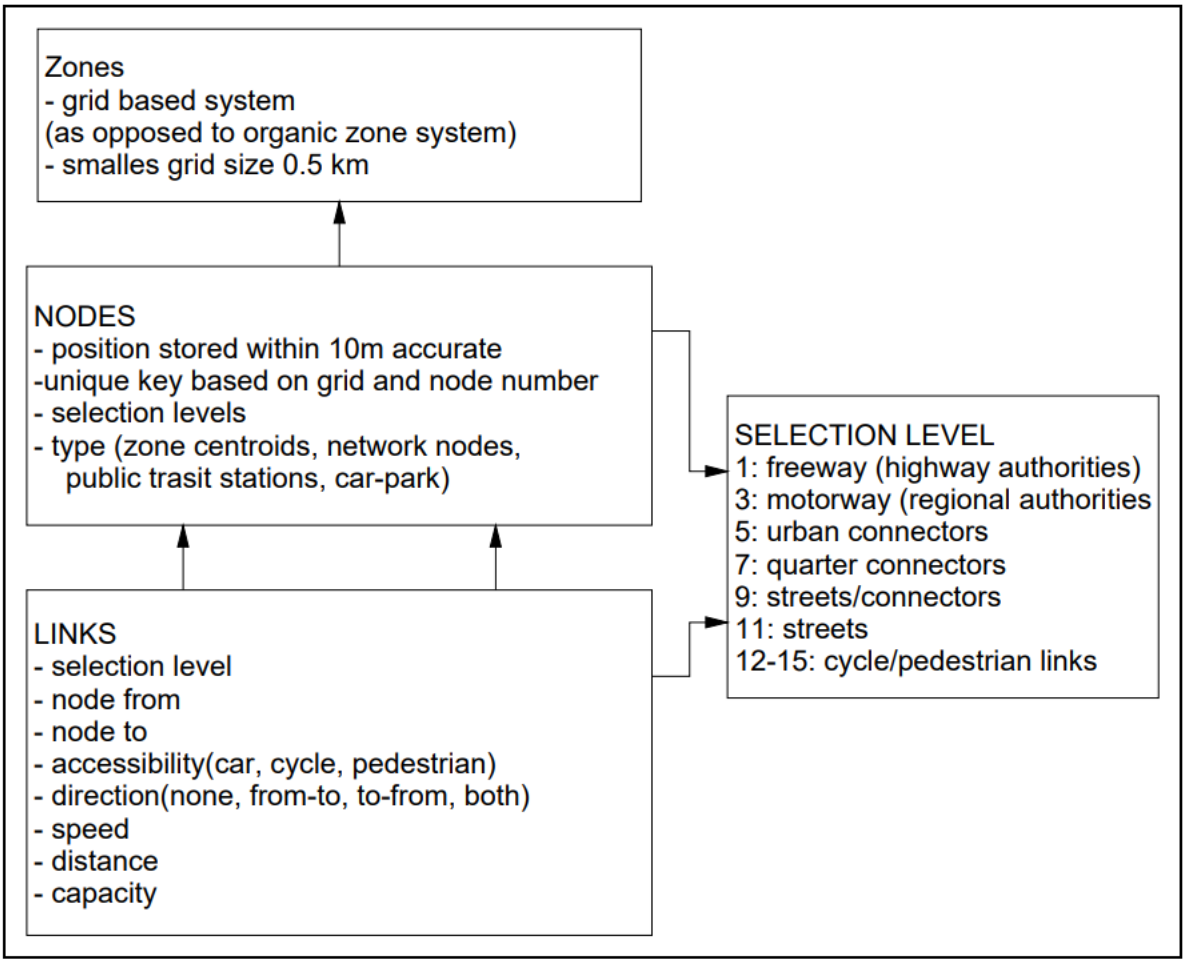


**Fig. 2 Description of a road network at three levels of spatial detail.**

At least all those network parts should be included for which detailed load figures are asked. In order to be accurate enough, the next lower functional layer of the network should be included as well because these offer routing alternatives to the studied links. So, if one is interested in arterials, one should include collector roads in the modeled network as well. If one is interested in the phenomenon of rat-running one is forced to model nearly the

complete network. To be accurate enough for other purposes as well, about 75% of network capacity should be part of the modeled network.
In urban networks, the nodes are the critical elements. The more nodes the more input data are required. In some software packages the number of links joining in a node is limited. This may require the definition of auxiliary nodes in order to represent reality correctly. Of special importance is the way of node coding, this means, whether or not turning movements at nodes are specified explicitly. For an illustration see Figure 3.

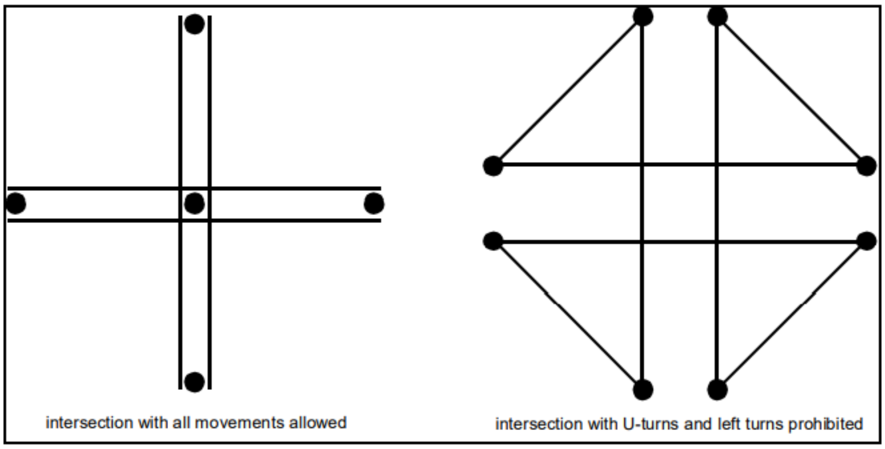


**Fig. 3 Forms of network coding.**

Zonal centroids are part of the modeled network. These are the entry and exit points of the trips. Centroids are connected to the network by connector links. Centroids may be connected to existing nodes or to dedicated auxiliary nodes specially introduced in the links. This way of centroid connecting depends on the software package at hand. In order to achieve sufficient accuracy in modeling, the centroid should have connections in all relevant directions.

3. **Network coding**

The modeled network consists of nodes and links (also called arcs). The nodes represent physical intersections or auxiliary nodes or centroids. Centroids are a special category of nodes where shortest routes start and end. Links represent physical links or auxiliary connections such as connectors. Bi-directional physical links are represented by two unidirectional links.

Zonal centroids are part of the modeled network. These are the entry and exit points of the trips. Centroids are connected to the network by connector links. Centroids may be connected to existing nodes or to dedicated auxiliary nodes specially introduced in the links is way of centroid connecting depends on the software package at hand. In order to achieve sufficient accuracy in modeling, the centroid should have connections in all relevant directions. In larger intersections travel time losses are caused due to waiting at the intersection entries. In addition, the travel time losses at the distinct turning movements may differ significantly. These travel time losses significantly influence the route choice of travelers. To model these travel times accurately enough it may be necessary to introduce in the modeled network description special links that represent these turning movements.

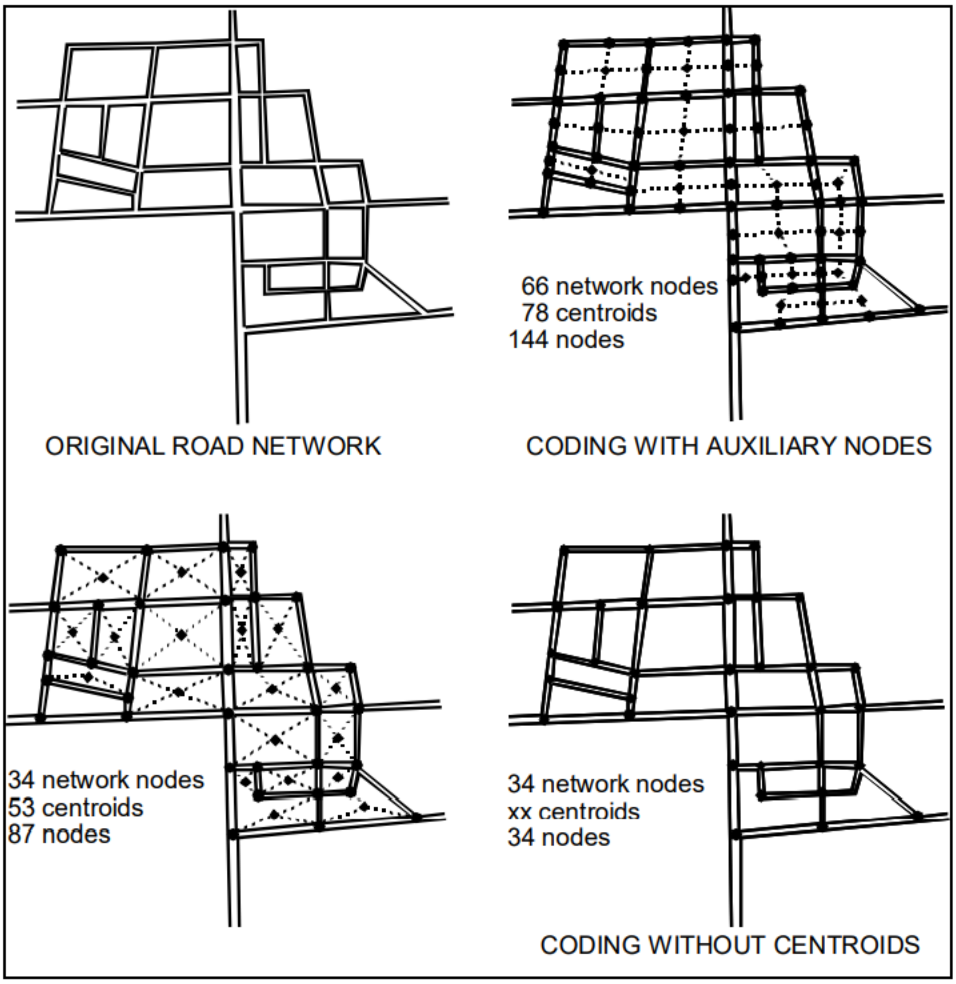


**Fig. 4 Examples of possible specifications of road network structure.**