Title
Performance Study of Dynamic Origin-Destination Estimation for Incremental Expansion of Urban Traffic Network Simulation Models

Author
Vafai, Shayesteh

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,

IRVINE

Performance Study of Dynamic Origin-Destination Estimation for Incremental Expansion of Urban Traffic Network Simulation Models

THESIS

submitted in partial satisfaction of the requirements

for the degree of

MASTER OF SCIENCE

in Civil Engineering

by

Shayesteh Vafai

Thesis Committee:
Professor Jay Jayakrishnan, Chair
Professor Jean-Daniel Saphores
Professor Michael McNally

2014

Table of Contents

List of Figures

List of Figures

Acknowledgement

I would like to express my heartfelt gratitude to my thesis advisor Professor Jay Jayakrishnan for his constant support and guidance. I have learned so much from him. I am especially thankful to Professor Jean-Daniel Saphores, for his help and guidance during my research and analysis of this thesis. Lastly, I would also like to thank Professor Micheal McNally. I learned a lot from his lectures and he was always available for help outside of the classroom.

Several people have helped shape my thoughts through useful insights and enlightening discussions. I would like to thank Ankoor Bhagat and Kevin Wang in particular for their assistance during the research of my thesis. In addition, I thank Dr. Sarah Hernandez and Dr. Qijian Gan for their valuable inputs and suggestions. I would like to thank my dearest friend Sarah Tasnim for her support who has made my experience such a memorable one.

I would like to thank the University of California Transportation Center (UCTC) for funding my studies during my graduate education at UCI.

I would like to express my indebtedness to my mother and sister for their love, encouragement and inspiration, and standing by me all through my life. My mother made a lot of sacrifices for where I am today, and younger sister has been a constant joy and support in my life. They have supported me in ways that were necessary outside of academia. And I reserve my deepest appreciation to my closest friends, for their unconditional support and love.

And finally, I give thanks to God for all my cherished memories, and for giving me the courage and strength to surmount the mountain, both big and small, that lay along the long road to this moment.

Abstract of the Thesis

Performance Study of Dynamic Origin-Destination Estimation for

Incremental Expansion of Urban Traffic Network Simulation Models

By

Shayesteh Vafai

Master of Science in Civil Engineering

University of California, Irvine 2014

Professor Jay Jayakrishnan, Chair

The goal of this research is to suggest a new approach for performing dynamic origin-destination estimation on a controlled simulation network that is expanded from an existing/starting network by adding different sub-areas to the existing network. In micro-simulation analyses, estimating origin-destination (OD) is a very time consuming and data intensive process. This research presents an incremental procedure to develop dynamic OD (origin-destination) matrices for a controlled network with different sub-areas added to the starting network to avoid extensive effort in re-estimating the matrices for an existing area for which matrices are already estimated once, thus reducing the time required to obtain the OD matrices for vehicular micro-simulation. Fractions of traffic from existing external origins and destinations to various other origins and destinations are utilized in the scheme to properly develop the origin and destination demand to/from the new origins and destination in the additional areas.

Chapter 1: Introduction

Microscopic traffic simulators are commonly used in traffic management, traffic operation/control, traffic impact studies, and in evaluating Intelligent Transportation Systems (ITS) strategies by public and private agencies (Lee 2011). Microscopic traffic simulators consist of a series of mathematical traffic flow models, including lane selection models, gap acceptance models, and car-following models.

This thesis defines the framework of the study, as well as the detail of the network and the preparation of demand inputs, including the estimation of time-dependent origin-destination trip matrices. As a starting network (shown in Figure 1), with known demand inputs, was enlarged with additional areas, we developed a new design that does not change the demand inputs for the exiting portion of the network significantly, while the demand tables are estimated for the additional areas to the starting network. With this procedure, the efficiency of simulation model preparation can be improved.

In micro-simulation analyses, estimating origin-destination (OD) demands is a very time consuming and data intensive process (Bhagat 2014). The purpose of this research is to develop an incremental procedure and OD (origin-destination) matrices for the whole network with different sub-areas added to the starting network to avoid extensive effort in re-estimating the matrices for an existing area for which matrices were already estimated once, and reducing the time required to obtain the OD matrices for vehicular micro-simulation.

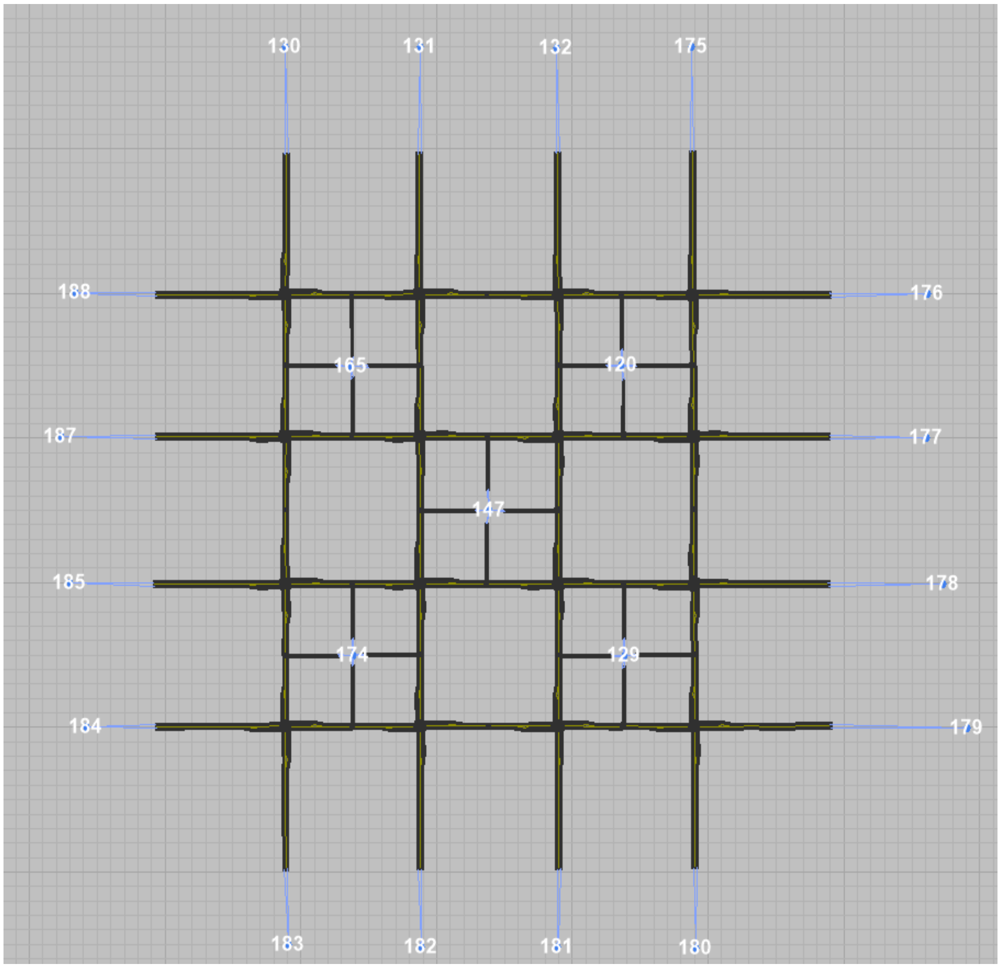


Figure 1: Starting Network

In this thesis, we propose a controlled network system and a procedure to obtain the origin-destination matrices of an extended network using already-estimated origin-destination matrices of a starting network by adding different sub-areas to it. First, larger network was made as shown in Figure 2 and origin-destination matrices were estimated, for the purpose of creating the ground truth data, which represents the real world for the simulation-based study. Since our network is a controlled network, the hypothetical grand truth data is prepared by simulating the larger network and collecting counts as will be explained in the methodology section. The study assumes this network to be the real-world, and then proceeds to the development of a simulation network after OD estimation for a starting network, which is a smaller portion of the larger network. This represents the process that is normally used in practice, when traffic agencies may have data from a larger area, but may develop a simulation model of a smaller area first and then may add additional areas, as per their needs. In our study, the real-world is a simulated network, with an associated static planning model for the network as well, just as is the case in practice. Then further studies proceeds with the estimation of the dynamic OD matrices using the static planning model as the initial static OD pattern. Note that in this study, the static planning model has different OD demand pattern than what is used for the "simulated real world".

The starting network was coded by eliminating some nodes and links from larger network and origin-destination matrices were estimated. Lastly, different sub-areas were added to the starting network. The OD estimation for the extended network is using already-estimated origin-destination matrix of a smaller network.

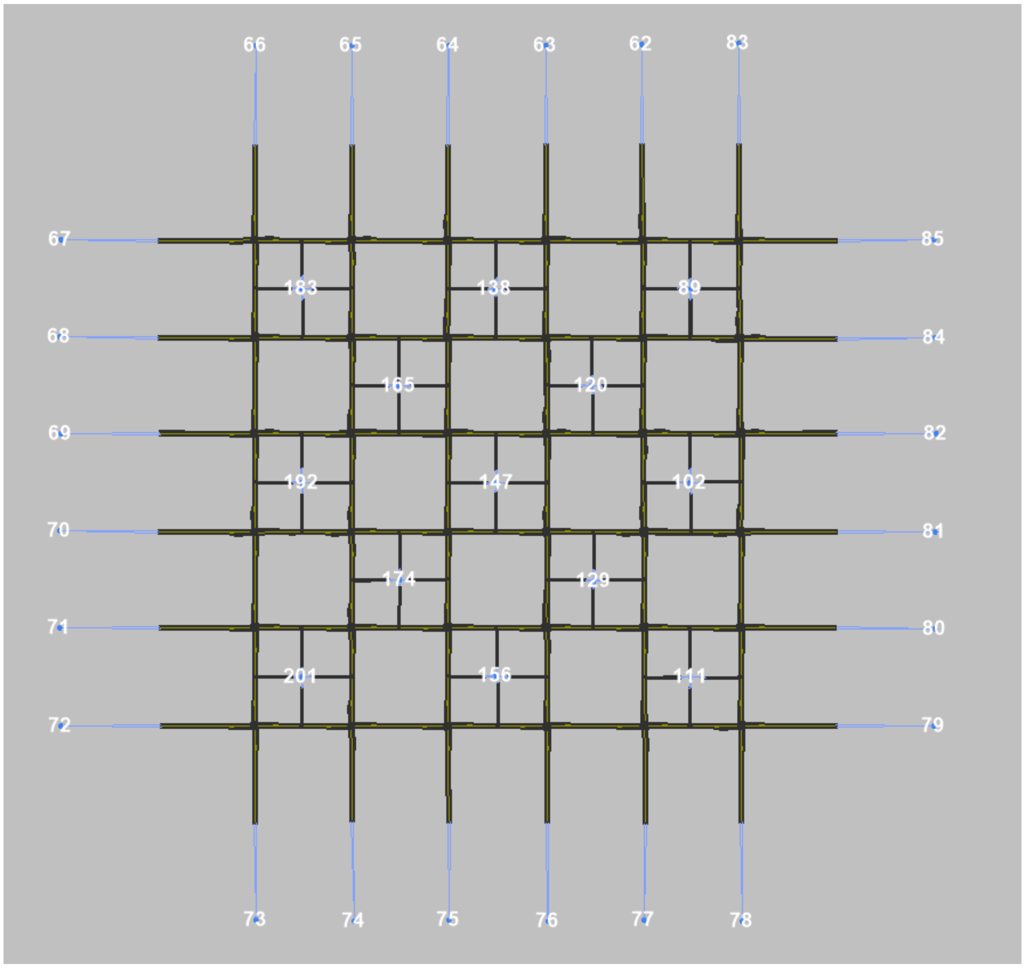
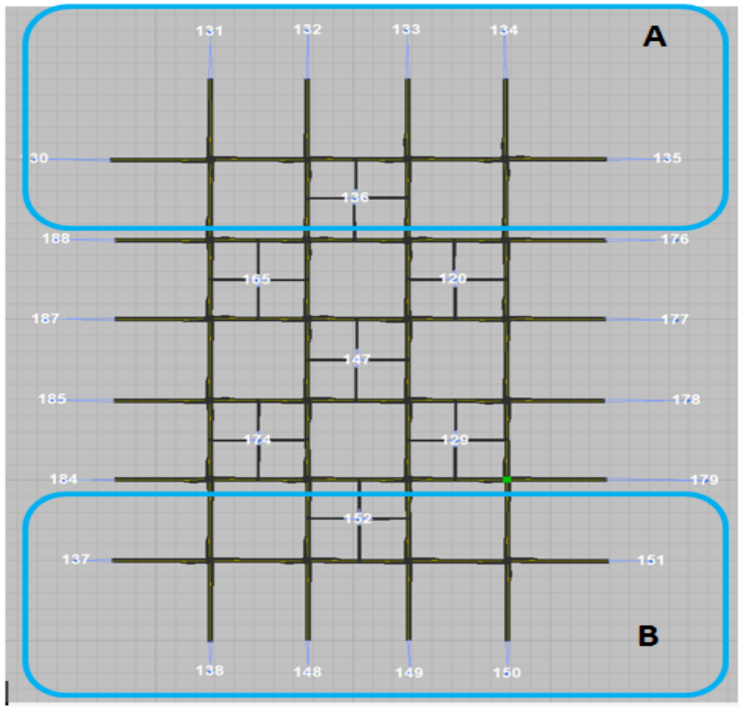
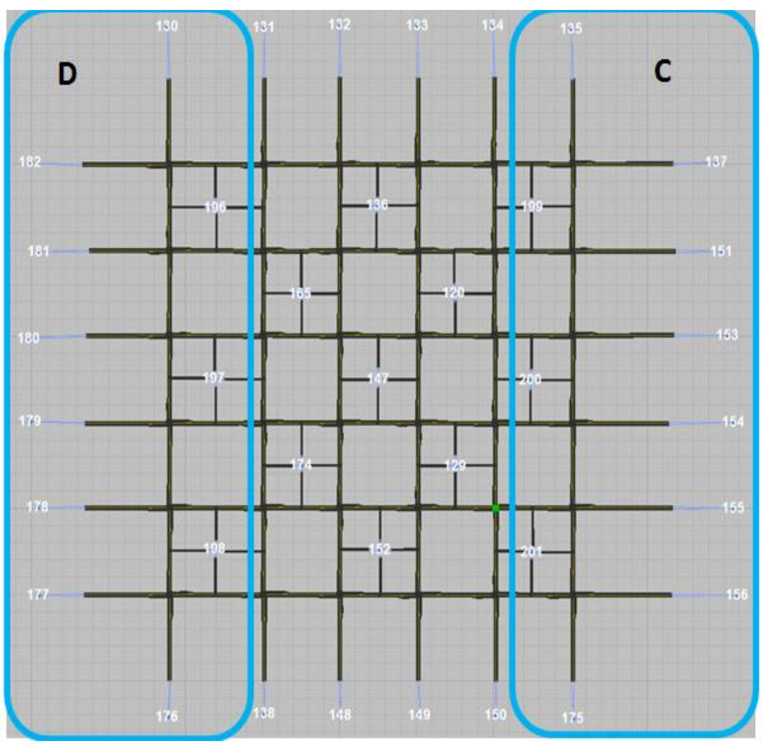


Figure 2: Original Network for Hypothetical Ground Truth Data

We demonstrate our approach on a larger network with 73 nodes which includes 37 centroids to collect the hypothetical grand truth data, and our starting network with 37 nodes which includes 21 centroids to estimate origin-destination matrices, and four different sub-areas. First, sections A and B are added to north-south part of the starting network, and then sections C and D are added to west-east part of starting network as shown in Figure 3. After a brief literature review, details of the network will be explained in the data and methodology section.



(a)



(b)

Figure 3: (a) Extended Network, North-South (b) Extended Network, West-East

Micro-simulation is a time consuming process. In order to conduct micro-simulation analyses, the network geometry must be coded, all traffic control devices must be installed, including intersection signals and ramp meters, and seed OD (origin-destination) matrices must be generated for the OD estimation process. Further, traffic parameters, speed limits, and sensor information are detailed by the micro-simulation system.

Well-estimated OD matrices facilitate the replication of realistic traffic conditions in a simulation. Typically, a license plate matching survey is utilized for predicting OD demand. This process is inefficient with regards to inputs of time, labor, and cost as it entails recording and matching the license plate numbers of all vehicles passing by pre-determined points to estimate the OD demand. Mathematical formulation can be used to estimate OD demand in larger networks. This approach, compared with conducted a license plate matching survey, is efficient and cost effective for static planning model applications. Loop detectors can be used to collect traffic data counts and speed measures. A similar system is suitable for micro-simulation analyses, but with a caveat. Micro-simulation is sensitive to time-dependent demand variation, thus the OD estimation process becomes time inefficient and data intensive due to the necessity of dynamic OD estimation.

Realistic representation of real traffic conditions is obtained by constantly adjusting OD demand based on observed traffic count data. For static networks, the literature covering OD estimation is extensive. However, dynamic OD estimation is not covered extensively in current literature, especially regarding cases of micro-simulation. A path-flow based dynamics OD estimation algorithm is found in previous research on the subject. When applied to a real world example, the algorithm proved to be efficient, as simulated traffic counts and section travel time were relatively close (Choi et al., 2009). The path-flow based model is used for the studies in this thesis.

The following chapter details the fundamental theoretical foundations of the origin-destination estimation process, emphasizing when time-dependent variations in demand over smaller intervals are to be estimated (15-minutes in this study over total time of 2 hours and 30 minutes-morning peak hours).

Chapter 2: OD Estimation

2.1 Background:

Traffic simulation requires traffic origin-destination (OD) demand inputs. Throughout the past 30 years, multiple OD demand estimation methods have been developed for the transportation field.

A license plate matching survey is a traditional method for estimating OD demand. License plate numbers of all vehicles which pass a pre-determined passing point are collected and matched to construct OD demand. A main disadvantage of this method is that, for a large study area, it is not cost efficient due to the labor required to identify all plate numbers. Traffic count data can also be collected from highway loop detectors to estimate OD demand. This method provides fairly accurate traffic counts and speed measures but is incapable of identifying vehicles' origins and destinations. Initial OD demands (or seed demands) are used in mathematical formulations to minimize errors between observed and estimated traffic counts. Although this method generates multiple solutions of OD tables given the traffic counts, it remains the most cost efficient method. Figure 4 shows the procedures for estimating OD demands used in our research (Lee 2011 and Bhagat 2014). These procedures are explained in the following section.

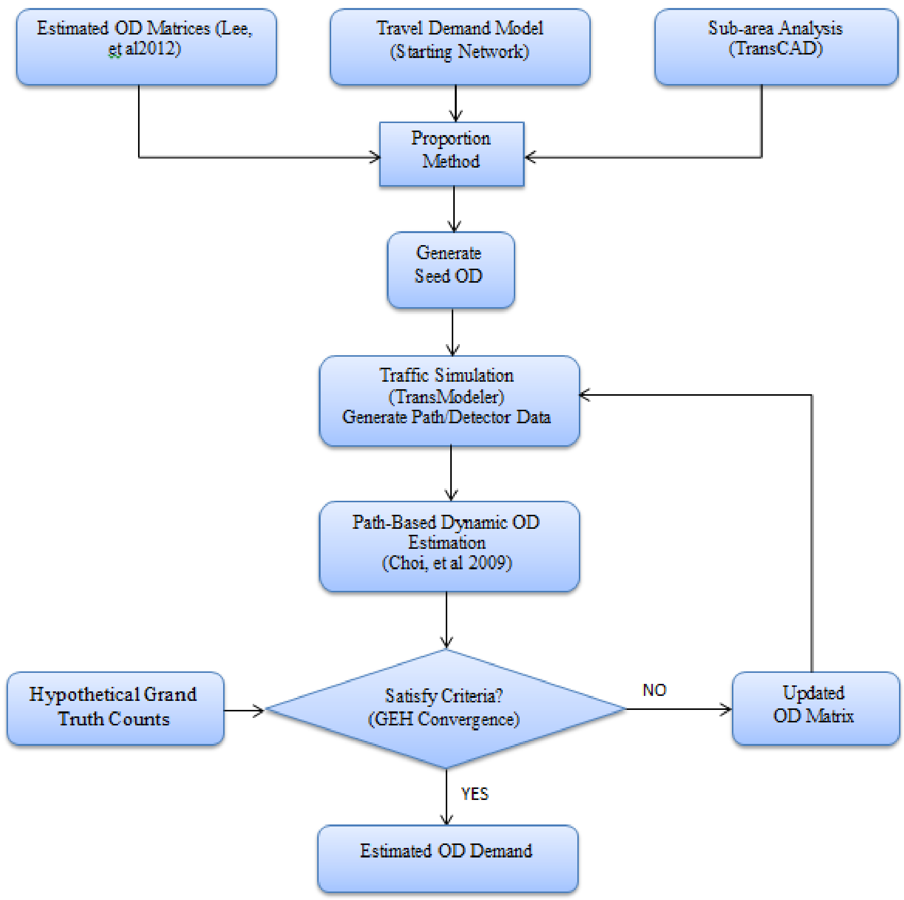


Figure 4: OD Estimation Procedure

2.2 OD Estimation Procedure:

Estimated dynamic origin-destination (OD) demands were used for our simulation. A combination of the hypothetical grand truth counts with a TransCAD sub-area analysis and previously estimated OD matrices were used to derive the travel demand model, which is similar to the real world analysis in Bhagat (2014) which used an existing network with a calibrated OD matrix (Lee et al., 2012). Detector and path data were gathered to create seed OD demands by applying the proportion method as inputs for traffic simulation. Path-based dynamic OD estimation was applied to the traffic simulation data (Choi et al., 2009).

The hypothetical grand truth counts provided traffic count data for the starting network. A comparison between the path-based OD estimation and the count data conducted to reveal whether or not the statistics converged. Path-based OD estimation is used if there is a convergence. If there is no convergence, the process is repeated, beginning with putting the updated OD matrices back into the traffic simulator and the checking for convergence (Bhagat 2014).

Traffic patterns are better portrayed using dynamic OD demands, particularly during peak periods where congestion is found. Five minute count data was aggregated into 15 minute count data for morning peak hour simulation.

The Caltrans Freeway Performance Monitoring System (PeMS) was used by authors of other studies as a source of traffic count data from loop detectors (Bhagat 2014). PeMS generally provides speed and flow data for urban freeways in California but it is necessary to review the data carefully to avoid extraneous data points (Bhagat 2014). Extraneous data points may occur due to the occasional failures of the loop detectors. In our study, the hypothetical grand truth counts from TransModeler are used for traffic counts. Loop detectors are placed symmetrically in our controlled network in order to have more control in the network. TransCAD, TransModeler, and MATLAB were used to aid in the calculation of traffic count data for our controlled network. First, initial OD matrices were adjusted in TransCAD and TransModeler. Then, as explained in the methodology section, initial demands were multiplied by peaking pattern factors and demand factors. Loop detectors in the network then provided the hypothetical grand truth counts.

2.3 Seed OD Generation:

Accuracy of the seed OD table is vital for proper OD estimation. In the earlier real-world study that motivated this work (Bhagat, 2014) the initial static OD demands were obtained from a modified Southern California Association of Governments (SCAG) traffic study. In the SCAG study, existing SCAG travel demand data was combined with stated port truck demand data, taken from a survey conducted for the Port of Long Beach. Sub-area analyses were performed in TransCAD to obtain the initial OD demand for the other simulation study. In this thesis study, since the real-world is itself a simulated real world, the seed OD matrices for our controlled network was as in an associated planning model of the same network, but without the same OD matrix as used in the "real-world" simulation. A larger network was first coded in TransCAD as our SCAG traffic model. Sub-area analysis, typically used to generate seed OD matrices, was assigned to the starting network. An example of the procedure used to make the OD table from the sub-area of the larger network in TransCAD is shown in Figure 5. The flow and V/C ratio on each link is generated by performing a static traffic assignment (multi-class assignment) on the whole network. Then, seed OD matrices are generated by using the sub-area analysis procedure to the selected links in the smaller network.

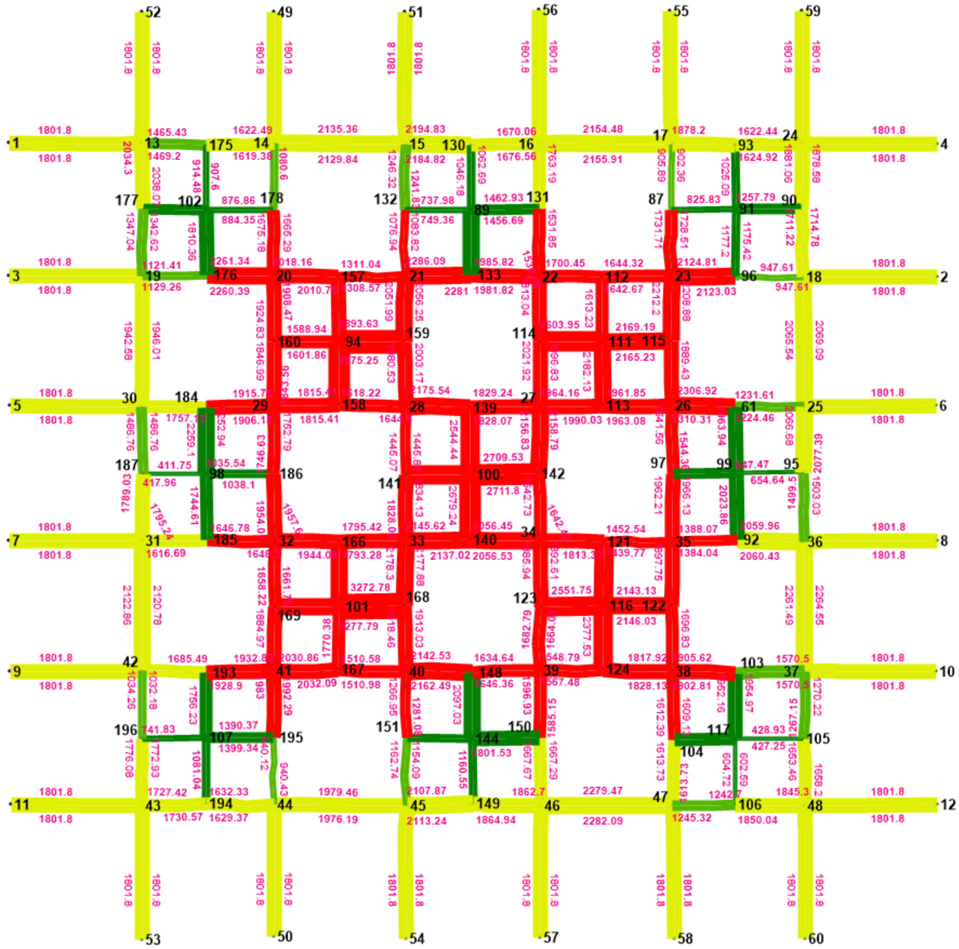


Figure 5: Area in Red Is Sub-area Analysis for Staring Network in TransCAD

As mentioned, estimating origin-destination matrices in micro-simulation analyses is a very time consuming and data intensive process. To avoid extensive effort in re-estimating the matrices for an existing area for which matrices are already estimated once, an incremental procedure is developed in the following section. This procedure for obtaining the origin-destination matrix of an extended network based on the origin-destination matrix on a starting network is demonstrated on an extended network constructed to model traffic in the lager area.

Chapter 3: Data Preparation

3.1 Hypothetical Ground Truth Data:

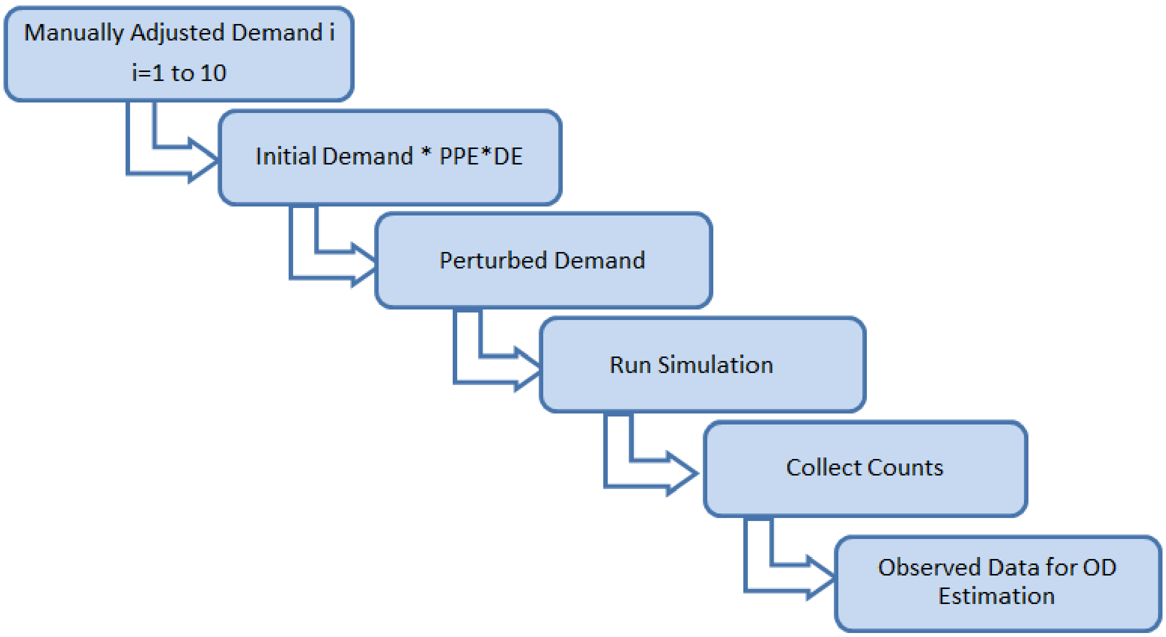


Figure 6: Hypothetical Grand Truth Data Preparation

Hypothetical grand truth data preparation for OD estimation follows the steps shown in Figure 6. The following describes each step in the process in further detail:

- The first step, obtaining initial demands, involves coding the controlled network in TransModeler. The assumed OD matrix for this step was one that was expected not to cause unrealistic congestion on the network, such as only certain areas being unduly congested while other areas did not have enough traffic.
- The initial demand is then multiplied by two factors; peaking pattern factor and demand factor. This step is essential, in that, it captures two primary factors that govern the generation of traffic in the real world. The first factor is the result of travelers departure time decisions which cause the total demand over a longer period of say an hour to have a temporal pattern, which causes the congestion patterns in the network. The second factor

captures the spatial variation across OD pairs. Thus the symmetric patterns that are the starting point for controlled experiment will still need to have spatial variation to not be too unrealistic to be used as a "real world" in the study. This step is then followed by manual adjustment in TransModeler. The initial demand matrix is a 37 x 37 origin-destination matrix with 6 vehicle classes and 10 time periods for each 15 minute time interval. A description of peaking pattern error and demand error follows:

1. **Peaking Pattern Factor (PPF):** An array of random numbers between 0 and 1. Because this is a distribution over 10 time periods, the addition of all the factors should result in a sum of 10. The peaking pattern error is applied in order to bring the data nearer to what may be found in real-world detector sites. Application of the peaking pattern error also results in a dynamic data set rather than the static data taken from TransModeler. Listed in the Table 1 are each 15 minute interval multiplied by the peaking pattern factors for the 2 hour period, with a 30 minute warm up:

Table 1: Peaking Pattern Factors

Time period	Peaking Pattern Factors
6:30AM-6:45AM	0.4
6:45AM-7:00AM	0.5
7:00AM-7:15AM	0.7
7:15AM-7:30AM	1.1
7:30AM-7:45AM	1.4
7:45AM-8:00AM	2
8:00AM-8:15AM	1.6
8:15AM-8:30AM	1.2
8:30AM-8:45AM	0.6
8:45AM-9:00AM	0.5

2. **Demand Factor (DF):** TransModeler produces a smooth initial demand which does not match real world data collection. Thus, the demand error is multiplied by

a factor to render the data collected from the PeMS detectors, resemble a more real world data set. Using uniform distribution with an interval of [0.75 1.25] and with the aid of MATLAB, the demand error was calculated for 6 different vehicle classes.

$$DF = a + (b - a) * rand(n) \quad (1)$$

- The perturbed demand is obtained by multiplying peaking pattern error and demand error by the initial demand matrices. This data set is closer to the data set from PeMS.
- Running a simulation in TransModeler gives us a new demand file:

  demand_i , i=1 to 10 (2)

- Data from sensors and the vehicle counts are collected.
- These data are used as our real world or PeMS data from detectors.

3.2 Planning Data:

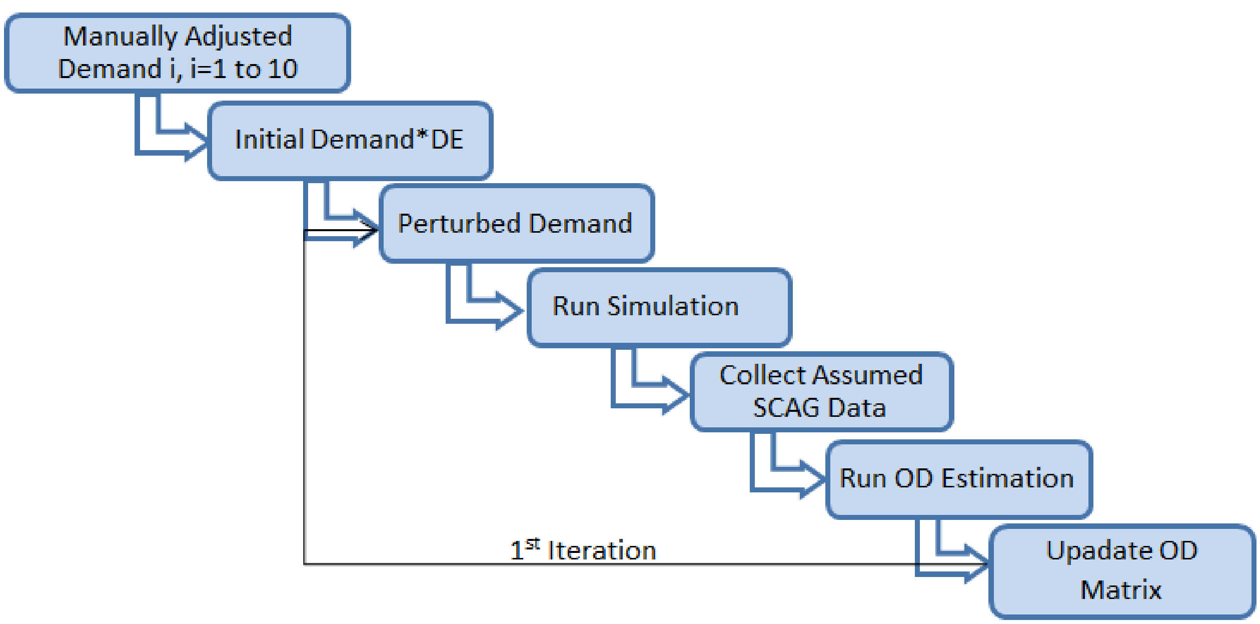


Figure 7: Planning Data Preparation

The steps outlined in Figure 7 show the process for data preparation for the planning matrix for OD estimation. The steps are further detailed below:

- The first step, obtaining initial demands, involves coding the controlled network in TransModeler. No congestion should be found in any link.
- Unlike in the above case of the simulated ground truth "real-world" network, for the planning model, there is no need for a temporal pattern, because planning models involved static OD tables. The only perturbation applied to starting OD table is essentially to capture the error between the total generation in the "real world" over the period for any OD pair, and its corresponding planning model, which is again realistic, in that, all planning models used in practice have estimated origin-destnation flows which have significant differences from actual real-world flow. . As this is an error that we are applying here, it is called a "Demand error" though it is effectively very similar to the "Demand factor" used for the "real-world" above. Factors with a uniform distribution and an interval of [0.75 1.25] are applied to the initial demand matrices. To obtain the demand for each 15 minute interval, that result is divided by 10 for the 10 time periods.
    1. **Demand Error (DE):** TransModeler produces a smooth initial demand which does not match real world data collection. Thus, the demand error is multiplied by a factor to render the data collected from the PeMS detectors, resemble a more real world data set. Using uniform distribution with an interval of [0.75 1.25] and with the aid of MATLAB, the demand error was calculated for 6 different vehicle classes.

$$DE = a + (b - a) * rand(n) \tag{3}$$

- Perturbed demand is calculated by the formula below:

$$perturbed\ demand = (\sum_{i=1}^{10} initial\ demand\ ) * DE \tag{4}$$

- Vehicle trajectories are updated by running the simulation in TransModeler.
- A new trajectories file is generated by running the OD estimation. This process is repeated and iterated in order to obtain GEH values less than 5 or 10 percent. The GEH statistic is a modified Chi-squared statistic that considers differences between observed and simulated traffic counts. It is defined by:

$$GEH = \sqrt{\frac{(O-S)^2}{0.5\ (O+S)}} \tag{5}$$

  - $O =$ Observed traffic flows (vehicles per hour) for each simulated link.
  - $S =$ Simulated traffic flows (vehicles per hour) for each simulated link.

Chapter 4: Methodology

This chapter develops methods to expand extended seed origin-destination (OD) matrices for microscopic traffic simulation by advancing already-estimated OD matrices for a starting network. This technique reduces the time spent on dynamic OD estimation for this extended network.

For real-world simulation studies, the starting seed matrix is derived from a static planning model. However, in using a starting seed matrix for network augmentation schemes, the seed matrix is taken from a larger area than the simulation network. Details of the flow entering the sub-area (the starting network) of the larger network planning model can be utilized effectively for this purpose. For such a sub-area, a tool found in software packages used in regional planned models called “select link analysis”, can be used to show the origins and destinations of traffic (i.e., OD pairs) can be found at the entry points of the sub-area, along with the flow between them, using select-link analysis. The analysis will show how much traffic comes from which OD pairs on a selected link, and thus by selecting a cordon of links around a subarea, an OD matrix can be found that applies to all traffic (including) traffic that is flowing through the sub area and not the flow among the original internal OD pairs in the subarea, which could have been readily found from the larger area planning model.

To generate seed OD matrices for a starting network, the results for the select link analysis on the larger area planning matrix for the links at the boundary of the sub-area are used and the proportion of trips from any group of origins (which could for instance be in a new area being added to the subarea) to any other group of destinations. These proportions are then applied to the known starting OD matrices of the starting network (which already has a calibrated dynamic OD matrix in place), resulting in the seed OD matrices for an extended network. These procedures are further detailed in the following sub-sections. Our goal in this section is to demonstrate the process of obtaining seed OD matrices for an extended network using OD matrices that have already been estimates for different types of network extensions, by retaining the OD matrix within the existing subarea the same at the start as had been calibrated earlier. This is expected to result in faster convergence of the dynamic OD estimation algorithm without substantially changing the estimated matrices for the sub-area. We will first illustrate the design of the selected study network in order to relate it to the different types of network extensions.

In this study, the microscopic traffic simulator used to generate both the ground truth "real world network" and to represent the microsimulation model used in practice for augmented network simulation and OD estimation is TransModeler 3.0. This model offers many advantages over other similar software packages. One of these advantages is its versatility. Various road types can be simulated including local roads and freeways. Traffic flow dynamics can be evaluated while also controlling for traffic signals and ITS operations. It also can be integrated with TransCAD and Geographic Information System (GIS) data with respect to input of network information and result visualization. Vehicle trajectory data can also be generated allowing for link by link processing for estimating emissions of different pollutants (Lee et al., 2012 and Soyoung et al., 2010) as well as for other purposes. In this study, the trajectory data is used to find fractions of any OD's flow across the multiple used paths, which is required in the path-based dynamic OD estimation algorithm that we employed (Choi et al, 2009). In fact the dynamic OD estimation algorithm uses the trajectories to calculate the flows between multiple OD pairs going over every simulated detector, to develop the path-flow fractions for each OD pair, as well as the gradient (rate of increase or decrease needed in each path flow) that is computed in every iteration of the dynamic OD estimation process. Thus the availability of computational modules from earlier research (as in Lee at al., 2012, and Bhagat, 2014) that work with the Transmodeler 3.0 trajectory data output, was deemed a big advantage that this particular software package offered over other microscopic simulation packages.

4.1 Design of Selected Study Network:

The purpose of this research is to develop seed OD (origin-destination) matrices for different sub-areas seed OD matrices for vehicular micro-simulation. This was accomplished by taking some areas out of the larger, 6×6 network (36 internal nodes, with 37 internal and external origin/destination nodes). Thus any subarea of this network has known and already estimated OD matrices, which can be used as the starting network to reduce the time spent on OD estimation when additional areas are added to the simulation model of the subarea and the dynamic OD matrices need to be estimated for the augmented network.

First, a 6×6 network is made, which is considered to be the planning matrix for the larger area, which is similar to the planning matrices that are available to traffic and planning agencies from which subareas are selected for simulation studies. In our study this network of 6×6 is considered to be the "larger" network, while the simulation model is built for a a sub-area of 4×4 is created in TransCAD and the demand from each origin to each destination is being specified OD_{ij}=15. This demand was selected after some preliminary calculations and TransCAD runs so as to generate a reasonable level of traffic flow in the network, so that a ground truth "real-world" network with other details such as intersection signals can be made. This ground truth "real world" in this study is a simulated world of the 6×6 network. The next step is to establish the signal timings, as described below.

4.2 Traffic signal design

Signal timing for the network is designed with the aid of the signal setting software Synchro that is very commonly used in practice. After running the traffic assignment in TransCAD, link traffic flow, an output after traffic assignment, and also the traffic volume for each origin-destination pairs are used as an input to Synchro. The common cycle length selected after some preliminary analyses was 120 seconds for all the intersections. Coordinated actuated traffic signals are used in all the intersections because they have been widely implemented for the past few decades in real world network and they provide better progression along the major corridors through proper coordination. Intersection signal timing are found utilizing the optimization tool in Synchro which yields best green splits (green times for each signal phase as a fraction of the cycle time) for each intersection. These green splits were then used to determine the maximum green time, and the minimum green time (for gap-out) for actuated control at each intersection. These signal timings are then used separately in microscopic simulation with TransModeler, to establish the ground truth data.

A few preliminary trail-error runs were needed with different OD patterns (which are established by multiplying the base number of 15 vehicles for each O-D pair, with factors that depend on the location of the destination in the network. Due to the symmetry of the network, the traffic flow patterns at multiple intersections could be very similar and thus we find there to be six different types of intersections, in terms of the traffic flow patterns on their approach links. The six different traffic signal timings found from Synchro as per the approach volumes and implemented at different intersections in the network as shown in Figure 8, where the color of the intersection shows the signal timings implemented.

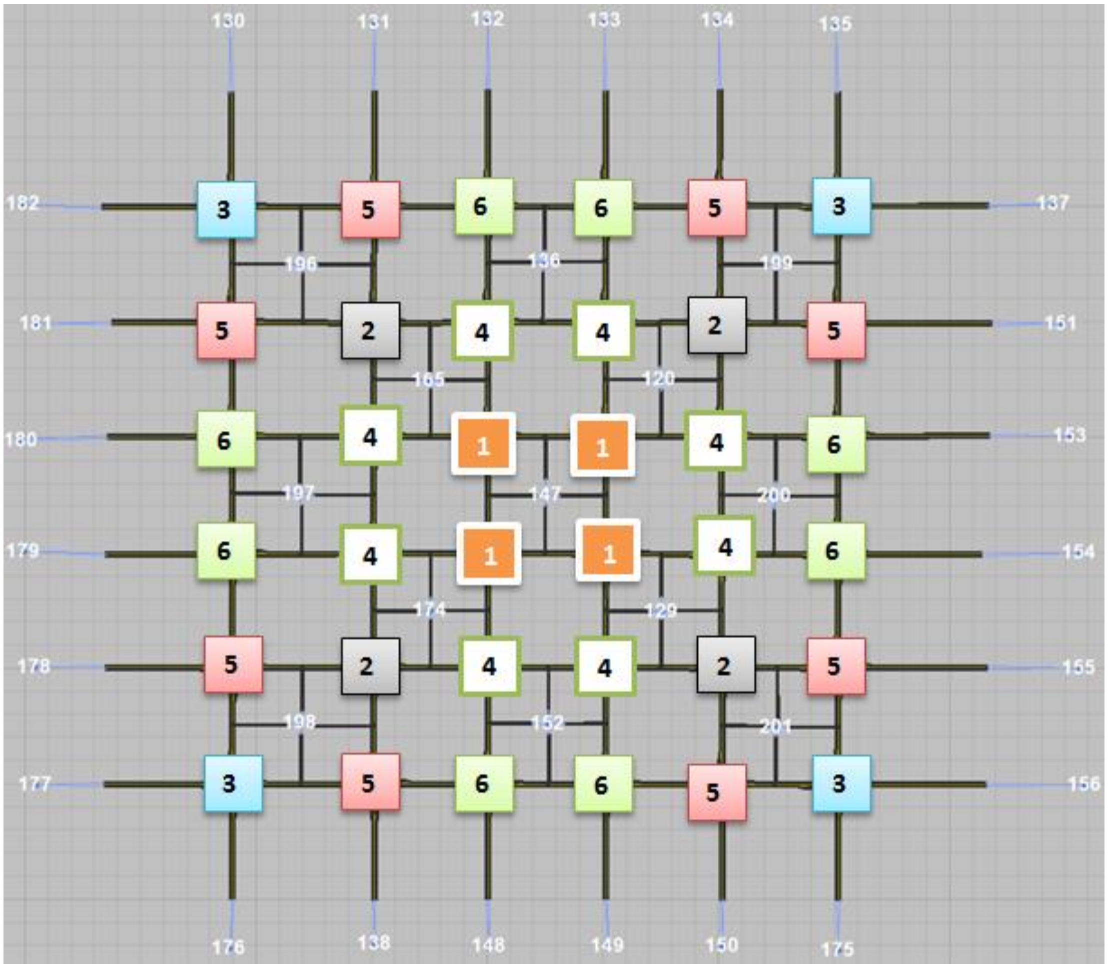


Figure 8: Traffic Signal Coordination

4.3 Iterative process of adjusting OD matrices

TransCAD provides the tools necessary to perform a standard network Traffic Assignment, using algorithms such as the well-known Frank-Wolfe algorithm. The traffic assignment step provides the traffic flow that is coming from each origin and going to each destination. These flows are static and uniform over the whole duration of assignment, as found in planning-model assignments in practice.

The next step is to code this network in the simulation software, TransModeler. The 6×6 network is created in TransModeler, and the same OD matrix as used in the TransCAD assignment run is used as an input to the network in TransModeler, which then simulates the traffic flows on the links and shows time-varying flows on each link during the simulation period. In the initial iterations, the traffic flows will not be show much time variation, as the first simulation uses a static OD matrix for any new area. However, they are compared to the observed data, as yielded by the separate “ground truth observed data” simulation which had both spatial and temporal OD demand variability and thus time-dependent variation in the flows at each simulated detector location as well. The dynamic OD flow adjustment scheme will adjust the static OD pattern to time-varying (dynamic) pattern for every OD pair in an attempt to match the flows in the next simulation run to be closer to the observed data. After the first iteration when the seed matrix used, for all the subsequent iterations the OD matrix for each simulation is the most recent adjusted dynamic OD matrix. During any iteration, TransModeler provides the updated vehicle trajectories and the OD estimation algorithm finds the adusted demand Matrices for multiple vehicle classes (6 classes were used in this study).

The differences between observed and simulated counts for various time steps (we used 15 minute periods in this study) during the analysis period The count data was compared to the path-based dynamic OD estimation to see if the GEH statistics converged. If there is convergence, the path-based OD estimation is used. Otherwise, the updated OD matrices are put back into the traffic simulator and the process is done again to check for convergence (Bhagat 2014).

By reducing the number of nodes and links starting network, 21X21 network, is created. Using sub-area analysis tool in TransCAD package, the flow for the selected sub-area is obtained for the starting network. After performing sub-area analysis, we can demonstrate where the traffic is coming from (origin) and where is it going (destination) and get the maximum flow from each origin to destination. Thereafter, we model the starting network with the same number of nodes and links in TransModeler. By running the simulation, updated trajectories and demands for all 6 vehicle classes are obtained. In the last step GEH convergence is checked, same procedure is applied in this section. If there is convergence, the path-based OD estimation is used. Otherwise, the updated OD matrices are put back into the traffic simulator and the process is done again to check for convergence (Bhagat 2014).

Different sub-areas are added to the starting network to extend the network gradually. First, Sub-area A and B are added to the starting network, get the flow from each origin to destination, and input the flow from TransCAD to TransModeler. Then we run a simulation for the new network with sub-areas added. After iterations are done for the extended network with the additional sub-areas A and B to improve the GEH value, the other sub-areas C and D are added to the network and the same process are repeated to obtain the updated trajectories and demands for all 6 vehicle classes.

4.4 Preparing the Study Networks:

To demonstrate the approach of our network design, we consider the original network with 73 nodes and 37 centroids which all of them are considered as nodes as shown in Figure 2. The 6×6 network is already coded and the OD matrices are already estimated for this network for a two-hour time period and a 30 minute warm up, which is a total of 10 consecutive time intervals of 15 minutes each for 6 vehicle classes. In SCAG planning model there are 10 vehicle classes as shown in Table 2 (Bhagat 2014). However, six classes of vehicles were used in our simulation. To prepare data for the OD estimation procedure, vehicle types 1 and 7 in the planning model were combined and used as PC (passenger car) demand in the TransModeler simulation, and vehicle types 2 and 3 in the planning model were combined and used as PC (HOV – High Occupancy Vehicle) demand. In addition, vehicle type 4 was used for PU (pickup truck) demand, vehicle type 5 for ST (single-unit truck) demand, vehicle type 6 for TT (trailer truck) demand, and combination of vehicle type 8, 9, and 10 were used for TT (Port) demand. This resulted in a total of 6 vehicle type matrices within the seed/initial OD matrices as shown in Table 3.

Table 2: SCAG Planning Model Vehicle Type

SCAG Vehicle Type	Description
1	PC (Passenger Car – Single occupancy)
2	HOV 2+ (High Occupancy PC)
3	HOV 3+ (Very High Occupancy PC)
4	LDT (Light Duty Truck)
5	MDT (Medium Duty Truck)
6	HDT (Heavy Duty Truck)
7	Port PC (Port related Passenger Car)
8	Bobtail: Truck tractor units only without trailer (Port HDT)
9	Chassis: Truck tractor units with trailer, not container units (Port HDT)
10	Container: Tractor trailer with chassis and container units (Port HDT)

Table 3: TransModeler Vehicle Class Description

Class Name	Description
PC	Passenger cars
PU	Pickup trucks, vans, and SUVs
ST	Single-unit trucks
TT	Trailer trucks

Then the sub-area network is prepared by dropping some nodes to create starting network. This yielded a 4×4 network (with37 internal and external nodes and 21 centroids, which are all also nodes), as shown in Figure 1. Subsequently, different additional sub-areas are added to the starting network. Starting network is extended by adding sub-areas A and B as shown earlier in Figure 3(a) in chapter 3.

Instead of starting the origin and destination (OD) estimation process from scratch for this extended network, we suggest a method to create seed OD matrices for the extended network that builds on the already estimated dynamic (time-dependent) OD trip matrices of the starting network. Then the extended network is enlarged by adding sub-areas C and D as shown in Figure 3(b). Therefore, same process applied for the OD estimation for the newly added sub-areas, starting the origin and destination (OD) estimation process on the existing estimated OD trip matrices for each time period for the starting network.

To give more details on how the methodology is done in this research requires differentiating between nodes and centroids in our starting network and also the extended network. Nodes are defined to join two or more links and also considered as a starting and ending points of a link in traffic network. Centroids are defined as a node from which trips begin or terminate.

To demonstrate the approach of our network design, let N_O={20,…,23,26,…, 29, 32,…,35, 38,…, 41, 130,…, 132, 175,…,188} and C_O={130,…, 132,175,…, 188} assign to the nodes and centroids, respectively, in the original network. Additionally, I_O ={20 ,…,23,26,…,29,32,…,35,38,…,41} represent the set of internal nodes in the starting network not including the internal centroids. Centroids and internal nodes are part of N_S ($C_S = N_S \setminus I_S$) (Bhagat 2014) and the number of different origins (and destinations) in the starting network's OD matrix is the number of elements in the set of centroids, which is equal to 21. Since the number of centroids in starting network is 21, the OD matrix is 21×21.

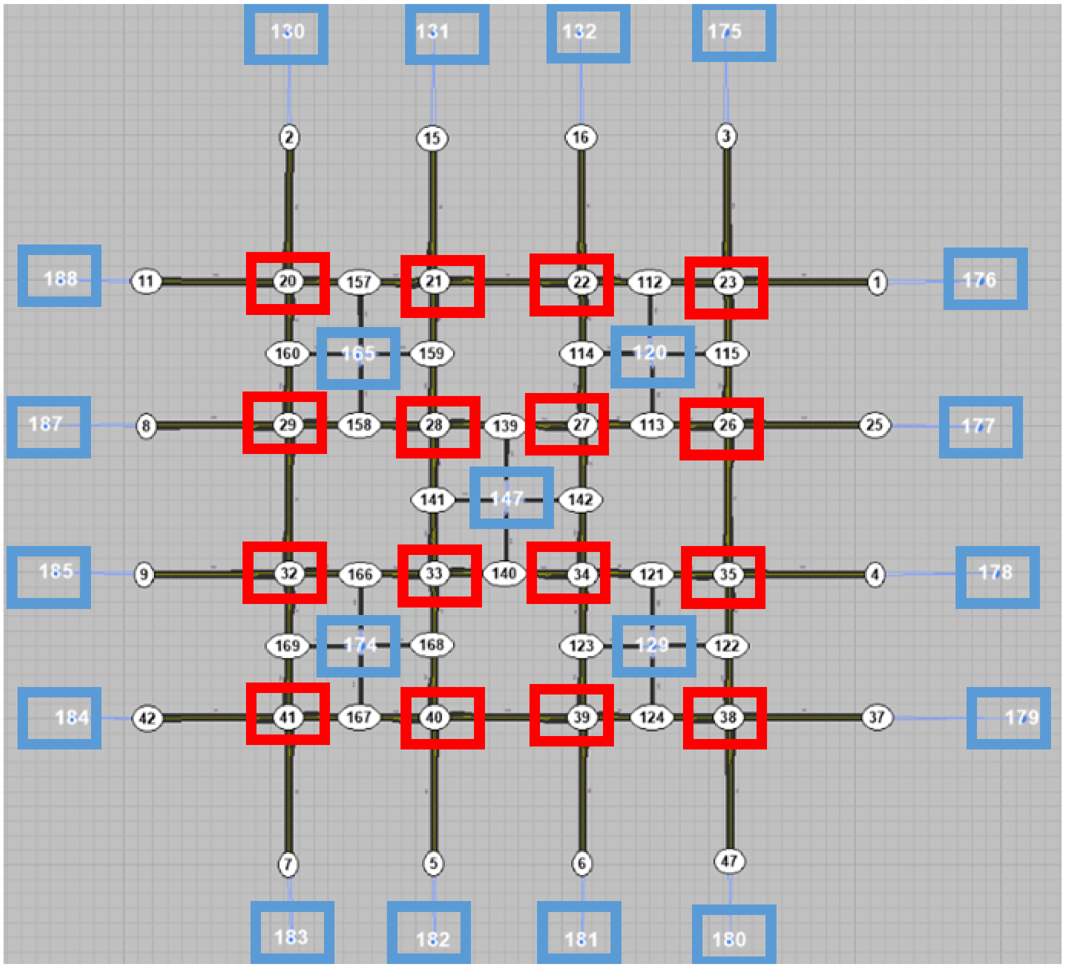


Figure 9: Nodes and Centroids for the Starting Network
(Note: Centroids are blue square boxes and nodes (excluding centroids) are red square boxes)

By extending the starting network, some nodes are added and other nodes, which were a centroid and were located at the end of a link, are deleted. Now sub-areas A and B are added to north and south of the starting network. Then let $N_{E(NS)}$= {3,16,…18,20,…,23,…,26,…,29, 32,…,35,38,…,41,47,52,54,56,120,129,…,138,147,…,152,165,174, 176,…,179,184,…,188} represent the set of all nodes, including removed nodes in our extended network and $R_{E(NS)}$ denote the set of nodes for these added sub-areas which is removed in our extended network. The list of removed nodes is R_E = {2, 3, 5,…, 7,15,16,47}. In the extended network, A_E represents set of all nodes but removed nodes, therefore $N_{E(NS)} = A_{E(NS)} \cup R_{E(NS)}$ and $A_{E(NS)} = N_{E(NS)} \setminus R_{E(NS)}$ (Bhagat 2014). Now let $I_{E(NS)}$ represents the set of all internal in the extended network including all the internal nodes in the starting network. Then let I_E^{New} represent the set of internal nodes just added to our extended network, $I_{E(NS)}^{New}$ ={3,16,…18,20,…,23,…,26,…,29, 32,…, 35,38,…,41,47,52,54,56}. Consequently, since our extended network just adds to our starting network, I_E is the union of the set of internal nodes in the starting network (I_S) and of the set of new internal nodes ($I_{E(NS)}^{New}$): $I_E = I_S \cup I_{E(NS)}^{New}$ (Bhagat 2014). C_E represents the set of centroids in the extended network. $C_{E(NS)}$ ={120,129,…,138,147,…,152,165,174, 176,…,179,184,…,188} is the list which obtained by disregarding the removed centroids in the extended network, so we can write is as $C_E = (N_{E(NS)} \setminus I_{E(NS)}) \setminus R_{E(NS)}$. Now, the size of OD matrix for the extended network is identified by knowing the number of elements in C_E which is 27×27.

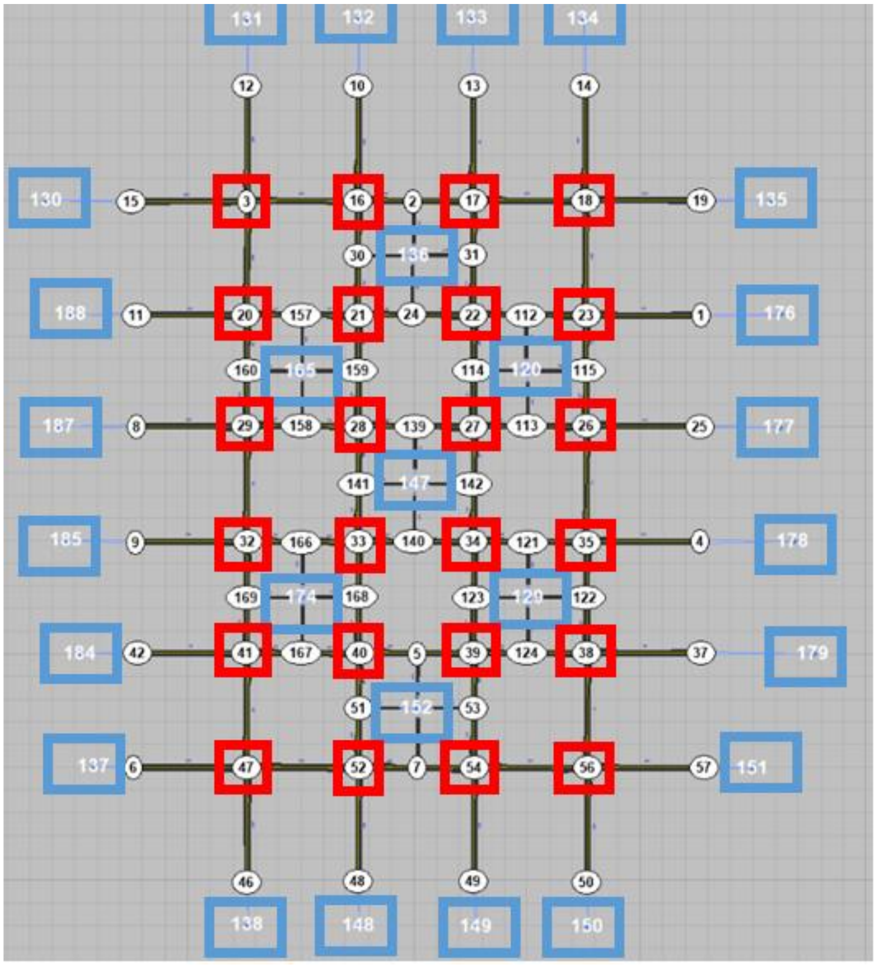


Figure 10: Nodes and Centroids for the Extended Network-North
(Note: Centroids are blue square boxes and nodes (excluding centroids) are red square boxes)

The next step is to add the other two sub-areas C and D to the north and south of the extended network which is used as a new starting network for this section. Then $N_{E(WE)}$ which represent the set of all nodes, including removed nodes in our extended network is N_E = {3,15,…,23,26,…,29,32,…,35,37,…,41,42,47,52,54,57,65,72,74,76,78,80,82,120,129,130,…,135,136,137,138,147,…,156,165,174,…,182,196,…,201}. Knowing $I_{E(WE)}$ represents the set of all internal in the extended network including all the internal nodes in the starting network, $I_{E(WE)}^{New}$, which denoted the set of internal nodes just added to our extended network, is $I_{E(WE)}^{New}$ ={15,19,37,57,65,72,74,76,78,80,82}.Therefore, since the extended network just adds to our new starting network, which includes north and south sections, $I_{E(WE)}$ is the combination of the set of internal nodes in the starting network ($I_{E(NS)}$) and of the set of new internal nodes ($I_{E(WE)}^{New}$): $I_E = I_{E(NS)} \cup I_{E(WE)}^{New}$ (Bhagat 2014). Let $C_{E(WE)}$ represents the set of centroids in the extended network. $C_{E(WE)} = \{120,129,130,\dots,138,147,\dots,156,165,174,\dots,182,196,\dots,201\}$ is the list which obtained by ignoring the removed centroids in the extended network, so it can be written as $C_{E(WE)} = (N_{E(WE)} \setminus I_{E(WE)}) \setminus R_{E(WE)}$. Now, the size of OD matrix for the extended network is identified by knowing the number of elements in $C_{E(WE)}$ which is 37×37.

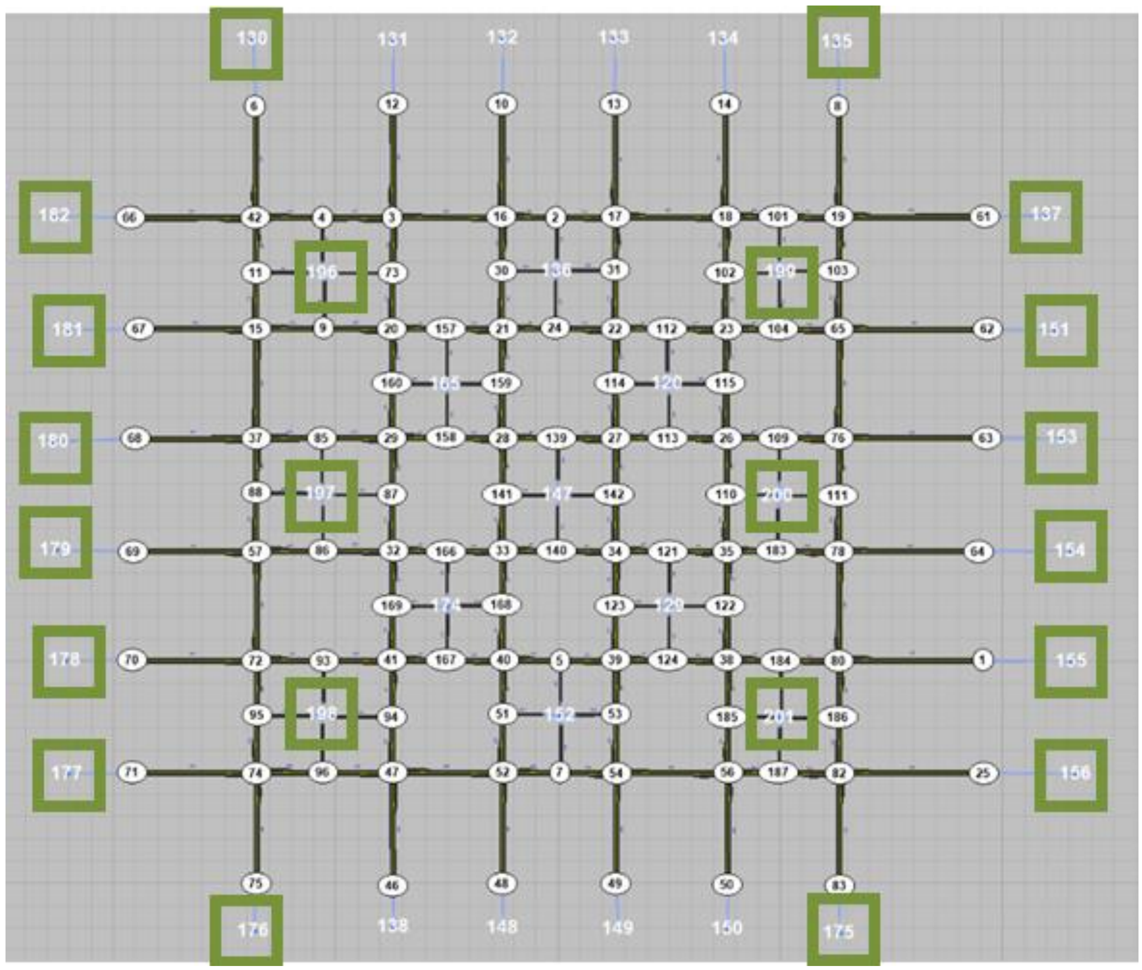


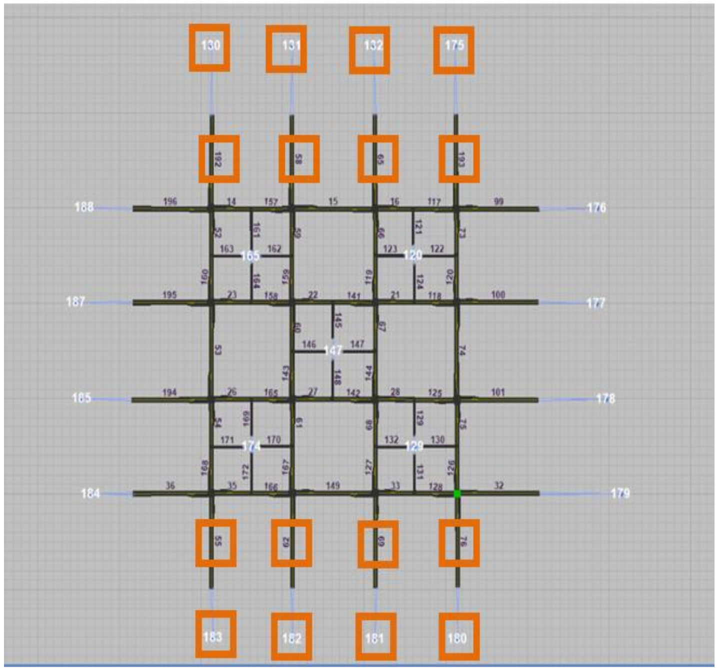
Figure 11: Centroids for the Extended Network-West East

(Note: Green Square Boxes are Newly Added Centroids.)

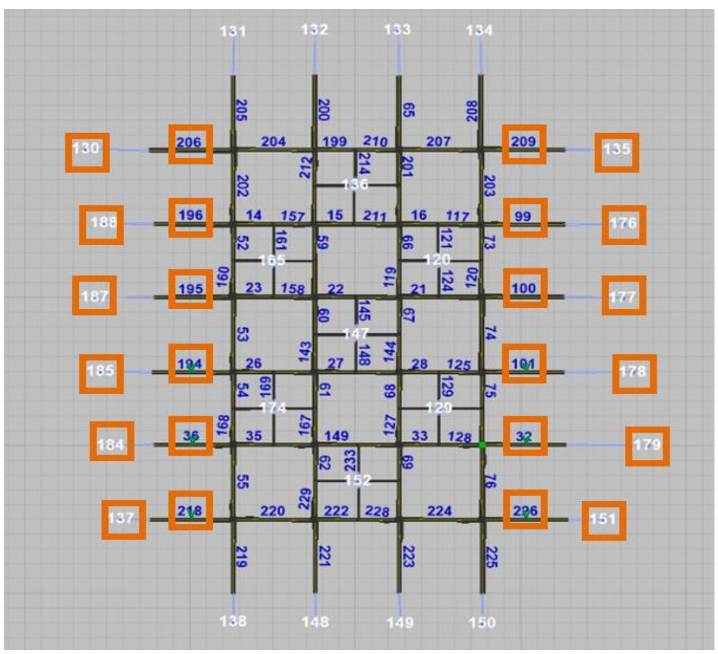
Here sub-area C and D are added to the left and right of the starting network and sub-area B and D are added to upper and lower portion of the starting network which are called "west-east" and "north-south" networks, respectively. In other studies, the network sub-areas can be added to various sides of network with different order, so the representation of the network can be different than what is used in this research. In this study, the various sub-areas are added to the starting network do not have any direct connection. There is no link connecting a node in one added sub-network to a node in the other added sub-network that does not traverse at least one node in the starting network that is retained in the extended network (Bhagat 2014). In this section, the process of adding various sub-areas and the method to add these sub-areas to the starting network are described. To apply the following system to a general network case, the added network areas can be divided in non-contiguous areas and pairs of sub-networks can be considered at a time, just as in the example here so there is no loss of generality in presenting the key concepts in this study.

The new sections are added to the north-south of the starting network is denoted by C^{NS} and is enclosed by rectangle as shown in Figure 3(a). Likewise, the new sections are added to the west-east of the new starting network, which includes north-south sections) is denoted by C^{WE} and is enclosed by rectangle as shown in Figure 3(b). C^{NS}={130,…,138,148,…,152} and C^{WE}={130,135,137,151,153,…,156,175,…,182,196,…,201} are the centroids added to nourth-south and west-east of the starting network, respectively.

The next step is to map link ID's to the removed centroids from the planning model for our extended network. A set of ordered pairs are defined as $(R_{E(NS)}, L_{E(NS)})$ which are the following pairs in the extended network :{(130,192), (131,58), (132,65), (175,193), (183,55), (182,62), (181,69),(180,76)}. As there are no two pairs with identical first component the $(R_{E(NS)}, L_{E(NS)})$ is a function with the domain of $R_{E(NS)}$= {130,..,132, 175, 180,…,183} which is a set of removed centroids and its range of $L_{E(NS)}$={192,58,65,193,55,62,69,76} which is a set of removed links as shown in Figure 12(a). The relation between the removed centroids and set of links is called mapping. Repeating same process, mapping procedure is done for the west- east extension. A set of ordered pairs are defined as $(R_{E(WE)}, L_{E(WE)})$ which are the following pairs in the new extended network: {(130,206), (188,196), (187,195), (185,194), (184,36), (137,218), (135,209), (176,99), (177,100), (178,101), (179,32), (151,226)}. As there are no two pairs with identical first component the $(R_{E(WE)}, L_{E(WE)})$ is a function with the domain of $R_{E(WE)}$= {130, 188, 187,185,184,137,135,176,177,178,179,151} which is a set of removed centroids and its range of $L_{E(WE)}$={206,196,195,194,36,218,209,99,100,101,32,226} which is a set of removed links as shown in Figure 12(b). Following sections will cover the different cases to be considered to generate the seed OD matrix for the extended network.



(a)



(b)

Figure 12: (a) Removed Centroids and Links in North-South Extension (b) Removed Centroids and Links in West-East Extension

4.5 Fraction Method

Different cases need to be carefully reviewed in order to generate an accurate seed OD matrix for our extended network. Figure 13 (Bhagat 2014) is an overview of the five different cases need to be considered and also the changes needed in an extended in an extended OD matrix.

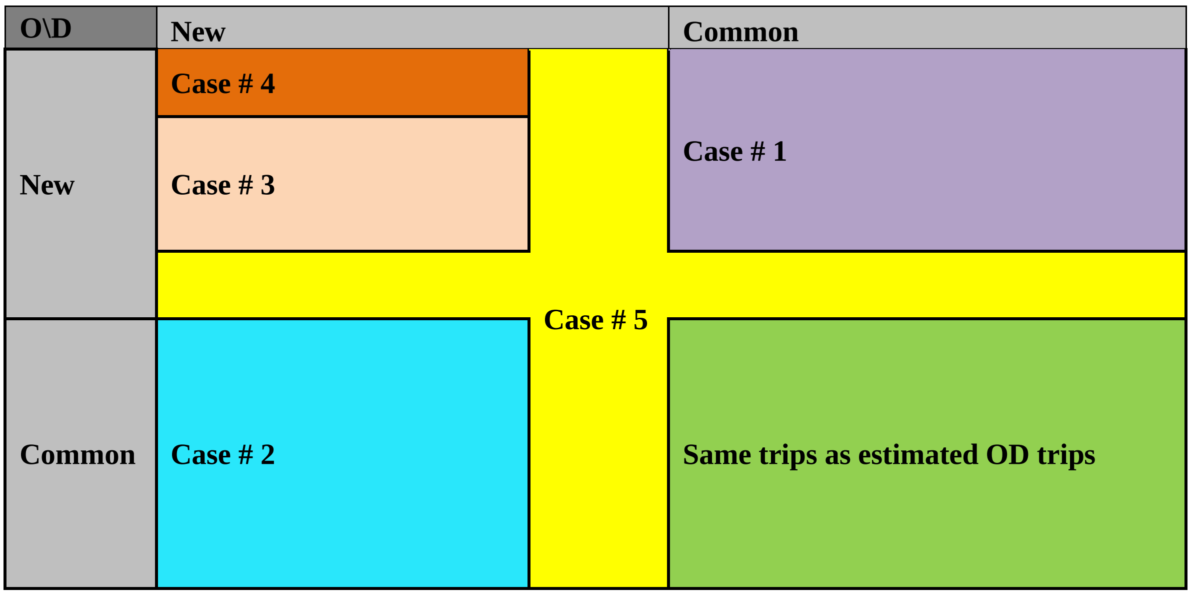


Figure 13: Overview of Changes Needed in the Extended OD Matrices (Bhagat 2014)

4.5.1 Trips from New Centroids to Common Centroids through Links Where Removed Centroids Were Placed with Inbound Direction on the Links:

In this section the trips from new centroids $C_{E(NS)}^{New}$ = {130,...,138,148,…,150,152} to common centroids $C_{S\cap E(NS)}$ = {120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} which pass inbound through links $L_{E(NS)}$ ={192,58,65,193,55,62,69,76}where the removed centroids $R_{E(NS)}$ = {130 ,..,132, 175, 180,…, 183} were located. This method is done by using trips from OD matrix for the starting network which is already estimated. Frist, the "select link analysis" needs to perform in TransCAD for the extended network to obtain the number of trips are inbound from $C_{E(NS)}^{New}$ to $C_{S\cap E(NS)}$ going through the set of link $L_{E(NS)}$ to calculated the following fraction which is a number of trips inbound from $C_{E(NS)}^{New}$ to $C_{S\cap E(NS)}$ passing through $L_{E(NS)}$ over the total flows going to different destinations. Below is the formula to estimate this fraction of trips (Bhagat 2014):

$$f_{ij}^{k} = \varphi_{ij}^{k} \,/\, \left(\sum_i \varphi_{ij}^{k}\right), \tag{6}$$

- f_{ij}^{k}: Fraction of trips from the new centroid i={130,...,138,148,…,150,152} to common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} passing inbound through link k={192,58,65,193,55,62,69,76};
- φ_{ij}^{k}: Select link analysis trips from new centroid i={130,...,138,148,…,150,152} to common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} passing inbound through link k={192,58,65,193,55,62,69,76}.

Next, the starting network estimated trips ψ_{mj}^{t} are divided from removed centroids m={130 ,..,132, 175, 180,…, 183} to common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188}, where t is an assigned time period and $(R_{E(NS)}, L_{E(NS)}) = \{(m,k)\}$. This is an ordered pair mapping set and (m,k) pairs are selected from the ordered set $(R_{E(NS)}, L_{E(NS)})$. Formula below shows how to calculate for the demands (Bhagat 2014):

$$T_{ij}^{t} = \sum_{m,k} \psi_{mj}^{t(R_E,L_E)}, \tag{7}$$

- T_{ij}^{t} : Trips from new centroid i={130,...,138,148,…,150,152} to common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} for time period t;
- $\psi_{mj}^{t(R_E,L_E)}$: Number of estimated trips from removed centroids m={130 ,..,132, 175, 180,…, 183} to common centroids j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188}, where t is an assigned time period and $\left(R_{E(NS)}, L_{E(NS)}\right) = \{(m,k)\}$.

Same procedure is applied to obtain the estimated trips for the west-east extension in the network.

4.5.2 Trips from Common Centroids to New Centroids through Links Where Removed Centroids Were Placed with Outbound Direction on the Links:

This section is an opposite direction of the previous section. In this case the trips from common centroids $C_{S\cap E(NS)}$ = {120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to new centroids $C_{E(NS)}^{New}$ = {130,..., 138, 148,…, 150, 152} which pass inbound through links $L_{E(NS)}$ ={192, 58,65, 193, 55, 62, 69, 76} where the removed centroids $R_{E(NS)}$ were placed. This method is done by using trips from OD matrix for the starting network which is already estimated. Frist, the "select link analysis" needs to perform in TransCAD for the extended network to obtain the number of trips are outbound from $C_{S\cap E(NS)}$ to $C_{E(NS)}^{New}$ going through the set of link $L_{E(NS)}$ to calculate the following fraction passing through $L_{E(NS)}$ over the total flows going to different destinations. Below is the formula to estimate this fraction of trips (Bhagat 2014):

$$f_{ji}^{k} = \varphi_{ji}^{k} / \left(\sum_{i} \varphi_{ji}^{k}\right) \quad (8)$$

- f_{ji}^{k}: Fraction of trips from common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to the new centroid i={130,..., 138, 148,…, 150, 152} passing outbound through link k={192,58,65,193,55,62,69,76};
- φ_{ji}^{k}: Select link analysis trips from common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to new centroid i={130,..., 138, 148,…, 150, 152} passing outbound through link k={192,58,65,193,55,62,69,76}.

Next, the starting network estimated trips ψ_{mj}^{t} are divided from common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to removed centroids m={130 ,..,132, 175, 180,…, 183}, where t is an assigned time period and $(R_{E(NS)}, L_{E(NS)}) = \{(m, k)\}$. Formula below shows how to calculate for the demands (Bhagat 2014):

$$T_{ji}^{t} = \sum_{m,k} \psi_{jm}^{t(R_E,L_E)} * f_{ji}^{k} \tag{9}$$

- T_{ji}^{t} : Trips from common centroid j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to new centroid i={130,..., 138, 148,…, 150, 152} for time period t;
- $\psi_{jm}^{t(R_E,L_E)}$: Number of estimated trips from common centroids j={120, 129, 147, 165,174,176,…, 179, 184, 185,187, 188} to removed centroids m={130 ,..,132, 175, 180,…, 183}, where t is an assigned time period.

Same procedure is applied to obtain the estimated trips for the west-east extension in the network.

4.5.3 Trips from New Centroids to New Centroids through Links Where Removed Centroids Were Placed with Inbound-Outbound Direction on the Links:

This section considered the trips from new added centroids in the C^N , which is the set of centroids added to the north bound of the starting network, to new centroids C^S, which is the set of new centroids added to the south bound of the starting network, passing through in-out link set called io. Likewise, trips from new centroids C^S to new centroids C^N passing via in-out link set called io by utilizing the data from an estimated OD matrix in the starting network.

The set of inbound-links for the inbound-outbound combination of links which started from the north bound of the network (origin) and terminated in the south bound of the network (destination) called io^{NS}. Similarly, the set of inbound-links for the inbound-outbound combination of links which started from south bound of the network (origin) and terminated in the north bound of the network (destination) represented io^{SN}.

To develop the number of trips are traveling from C^N, new centroids added to north bound, to C^S, new centroids added to south bound, which is passing via link set io^{NS}, "select link analysis" is performed in TransCAD for the extended network. Likewise, "select link analysis" is performed to develop the number of trips taken between two new centroids C^S and C^N passing through link set io^{SN}. Then fraction of trips is calculated by formula below:

$$f_{ij}^{NS(in,out)} = \varphi_{ij}^{NS(in,out)} / \left(\sum_i \sum_j \varphi_{ij}^{NS(in,out)} \right) \quad (10)$$

- $f_{ij}^{NS(in,out)}$: Fraction of trips from the new centroids in C^N to the new centroids in C^S passing through inbound-outbound link set io^{NS};
- $\varphi_{ij}^{NS(in,out)}$: Number of trips calculated by select link analysis from new centroids in $C^N$to new centroids in C^S passing through inbound-outbound link set io^{NS};

$$f_{ij}^{SN(in,out)} = \varphi_{ij}^{SN(in,out)} / \left(\sum_i \sum_j \varphi_{ij}^{SN(in,out)} \right) \quad (11)$$

- $f_{ij}^{EW(in,out)}$: Fraction of trips from the new centroids in C^S to the new centroids in C^N passing through inbound-outbound link set io^{SN};
- $\varphi_{ij}^{SN(in,out)}$: Number of trips calculated by select link analysis from new centroids in C^S to new centroids in C^N passing through inbound-outbound link set io^{SN};

Next, the estimated trips ψ_{mj}^{t} are divided between removed centroids $R_{E(NS)}$ = {130 ,..,132, 175, 180,…, 183}, where *r* and *v* span elements of R_E, and *t* denotes as an assigned time period in the hypothetical example, by using calculated fractions (Bhagat 2014).

Now let $r \in \{130,\ldots,132,175\}$ represent the set of removed centroids in the north bound of the network and $v \in \{180,\ldots,183\}$ indicate the set of removed centroids in its south bound, where r denoting to the centroids from where traffic arrives at the north bound entry link, $r\{in\}$ and v referring to centroids where there is traffic from an exit link, $v\{out\}$. Now $(r\{in\}, v\{out\})$ is defined as a one to one mapping between removed centroids. By using formula below, the north to south bound demands can be found for any OD pair in the extended network:

$$T_{i,j}^{t} = \sum_r \sum_v \psi_{r,v}^{t(r\{in\},v\{out\})} * f_{i,j}^{NS(in,out)} \qquad (12)$$

The south to north demands can be similarly found using:

$$T_{i,j}^{t} = \sum_v \sum_r \psi_{v,r}^{t(v\{in\},r\{out\})} * f_{i,j}^{SN(in,out)} \qquad (13)$$

Similarly, same procedure is applied to obtain the estimated trips for the west-east extension in the network.

4.5.4 Trips from New Centroids to New Centroids without Crossing Links Where Removed Centroids Were Placed:

This section considered the trips between new centroids in north bound, C^N, and also trips between new centroids in south bound, C^S, without crossing links where the centroids are removed from the network by using trips from planning model OD matrix. To find these trips, the time-proportion profile is required for all the trips in the starting network utilizing the already estimated OD matrix.

$$f^{t} = \psi_{ij}^{t} / \left(\sum_i \sum_j \sum_t \psi_{ij}^{t}\right) \qquad (15)$$

- ψ_{ij}^{t} : Trips between centroids in the set C_S, starting network centroids,
    - $i = j = 1, 2, 3, \ldots$ is the number of elements in set C_S,
    - $t \in$ Time-Proportion Profile.

Next, time-proportion profile is applied to the planning model trips between new centroids in the north bound, C^N, likewise to the trips between new centroids in the south bound, C^S, in the enclosed areas as shown in Figure 3(a).

The following formula is how to estimate the time-proportion profile to planning model trips:

$$T_{ij}^t = \theta_{ij}^t * f^t \quad (16)$$

- θ_{ij}: Planning model trips:
  - Between centroids in set C^N, $i = j$ =1, 2, 3, …, number of elements in set C^N,
  - Between centroids in set C^E, $i = j$ = 1, 2, 3, …, number of elements in set C^S.

4.5.5 Trips from New Centroids in Common Area (in Starting and Extended Networks) to All Centroids:

This section considered the trips from new centroids in common area C_S^{New} = {136, 152} to all centroids in the extended network except newly added centroids in the common area as shown in Figure 13, i.e. $C_E' = C_E \setminus C_S^{New}$, so that C_E' = {120, 129, 130,…, 135, 137, 138, 147,…, 151, 165,174, 176,…, 179,184,…,188} and similarly from C_E' to C_S^{New} using trips from planning model OD matrix.

$$f^t = \psi_{ij}^t / \left(\sum_i \sum_j \sum_t \psi_{ij}^t\right), \quad (17)$$

- ψ_{ij}^t: Trips between centroids in set C_S, starting network centroids,
  - $i = j$ =1, 2, 3, …, number of elements in set C_S,
  - $t \in$ Time-Proportion Profile.

Following formula is how to estimate the time-proportion profile to planning model trips:

$$T_{ij}^{t} = \theta_{ij}^{t} * f^{t} \tag{18}$$

- θ_{ij} : Planning model trips:
  - From C_S^{New} to C_E';
  - From C_E' to C_S^{New}.

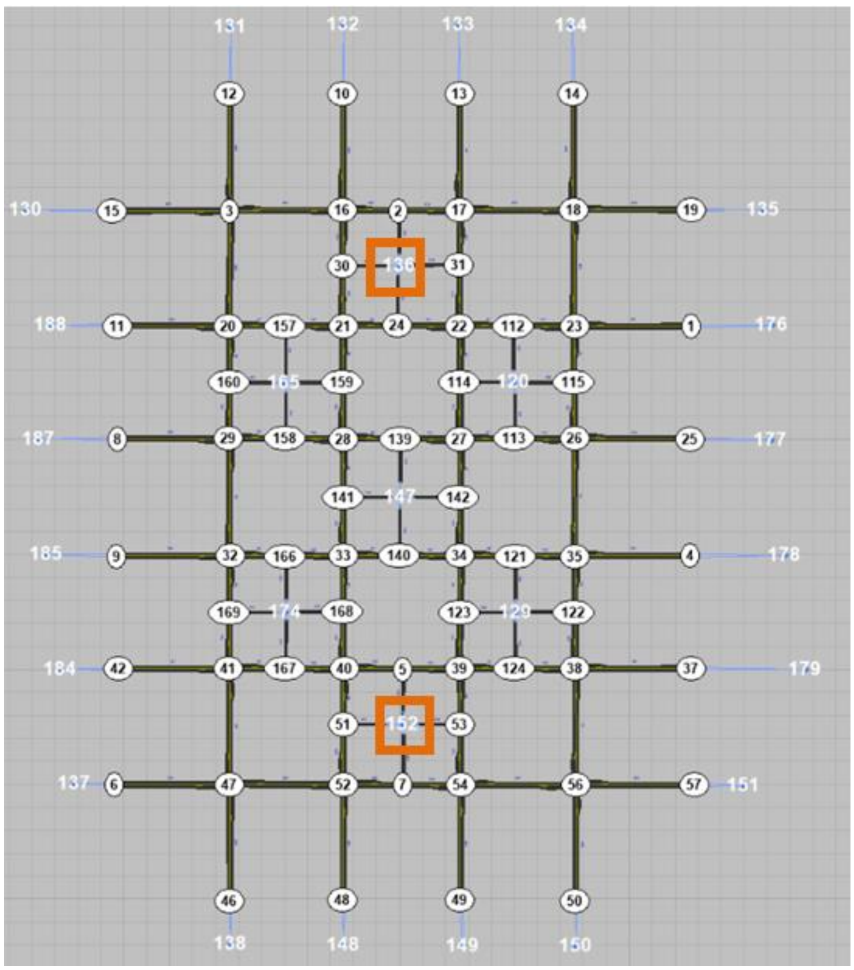


Figure 14: New Centroids in Common Area

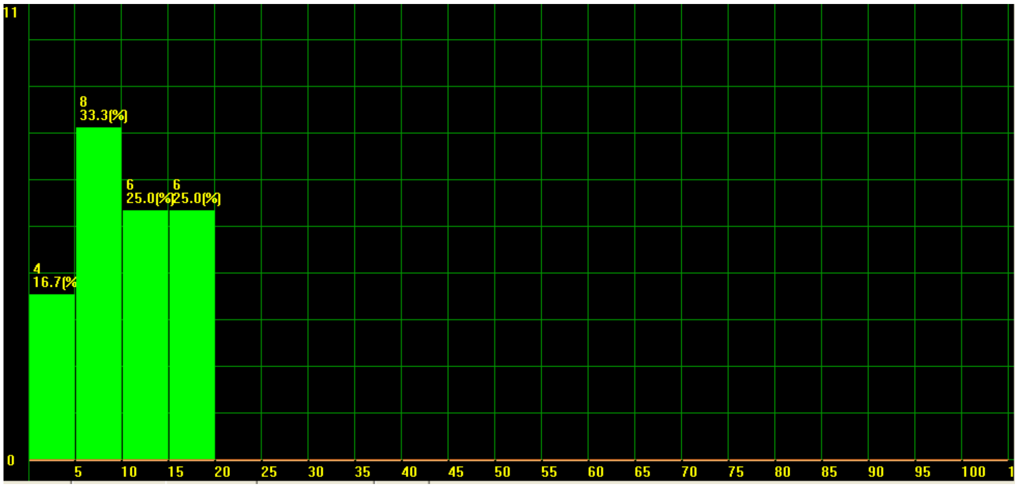
(Note: Orange Square Boxes are Centroids in Common Area.)

4.6. Detector Location:

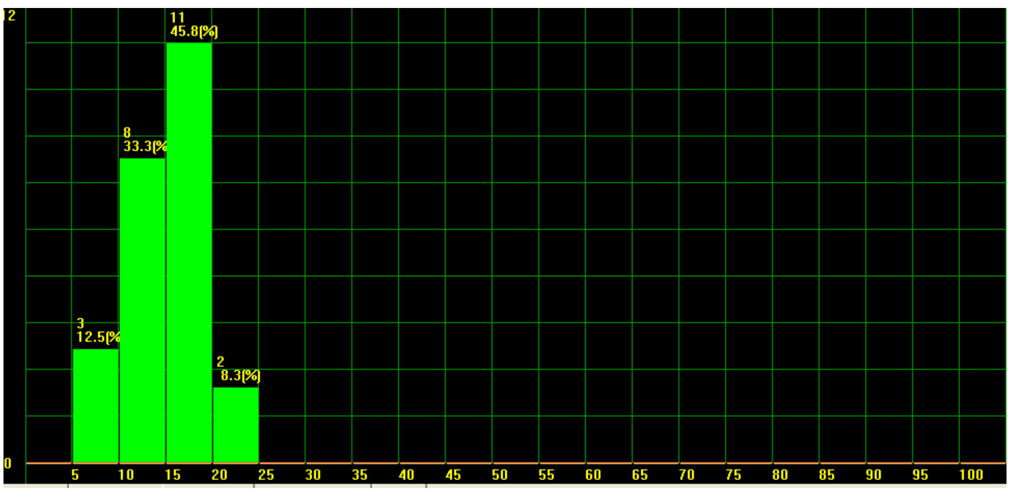
Detectors are partially located in the starting network as well as extended networks. Detectors are all placed symmetrically in the corners of the 4×4 network as well as 6×6 network. In addition, detectors are installed in the beginning of each corridor to count the number of vehicles are entered to the network. For the future studies, network could be fully detected and detectors could be installed in all the corridors and then compare the results with the results in our study. These loop detectors located in this network collect traffic count data by registering the number of vehicles are passing through these loop detectors while simulation is running in TransModeler. Subsequently, collected data from loop detectors are being used for OD estimation process.

Chapter 5: Results

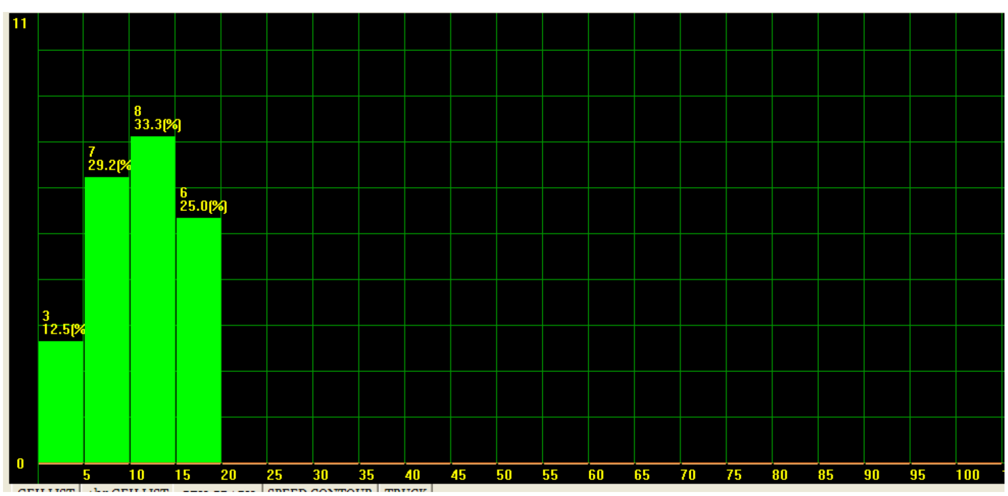
The main results of interest in this study pertain to the convergence characteristics of the incremental OD estimation process described in the previous chapters. This is evaluated on the basis of the match obtained between the observed link traffic flows in the ground truth simulation, and the flows found in the simulations during the iterations using the GEH statistic (Equation 5 in section 3.2). The percentage of detector locations with acceptable GEH statistics is the primary indicator. Shown in Figure 15 are the results of the 1st, 2nd, 4th, 6th, 8th and 10th iterations of the OD estimation algorithm for the starting network. The results indicate that the GEH values were improving over the 10 iterations. Significant improvement in the GEH values was noticed at the 6th iteration. Figure 15 (f) shows that almost 70% of the detectors in the network have a GEH statistic under 10 (vehicles/hour) when network augmentation schemes are used. Over the 10 iterations for the augmented scheme, the rate of improvement seems constant. A reason for this may be due to the limited number of iterations ran (10), without supply side calibrations and other adjustments needed to obtain better GEH statistic results. However, these results do indicate how much better the network augmentation scheme performs and shows the improvement obtained in the first iteration in matching observed traffic flows does not diminish with further iterations. In our study, we ran an augmentation scheme twice to obtain the results for the extended network. First, we ran an augmentation scheme successfully, for the north-south extension. However, the augmentation scheme ran for the west-east extension was unsuccessful due to technical issues. Also, due to a limited time frame, we could not finish this portion of the study, which can, however, be studied further in future research. Results of the GEH statistics are compared with the previous iterations. As shown, there is a significant improvement in GEH values comparing between the 1st and 10th iteration.


11
0
4
16.7(%
8
33.3(%)
6
25.0(%)
6
25.0(%)
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

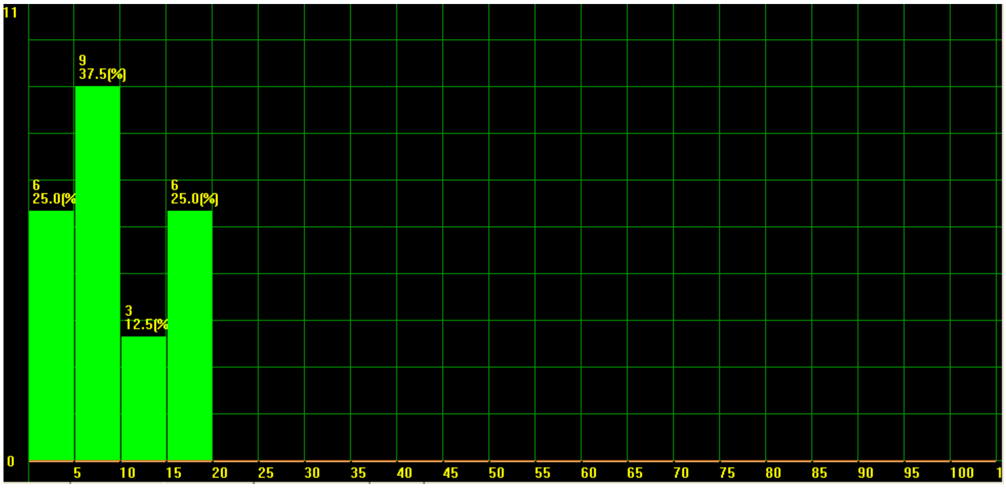
(a)


0
3
12.5(%
8
33.3(%
11
45.8(%)
2
8.3(%)
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

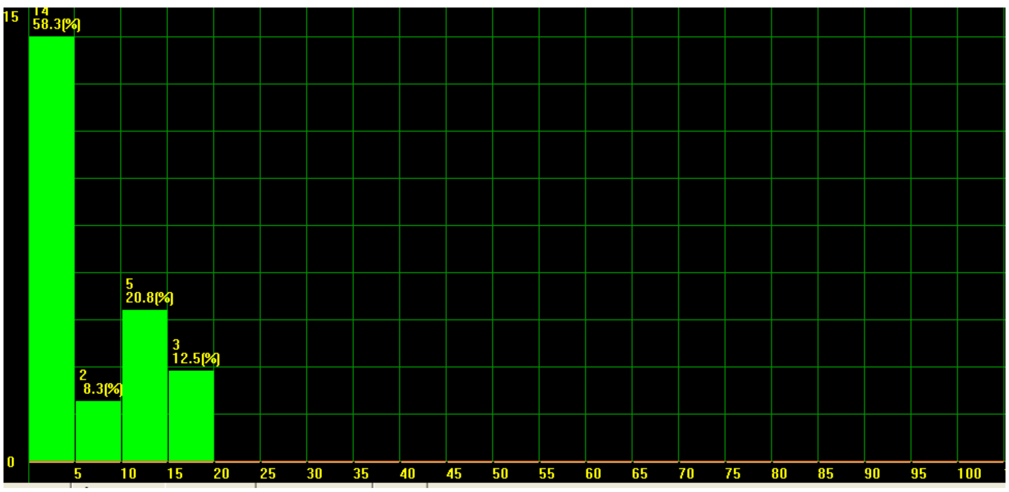
(b)


11
0
3
12.5(%
7
29.2(%
8
33.3(%)
6
25.0(%)
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

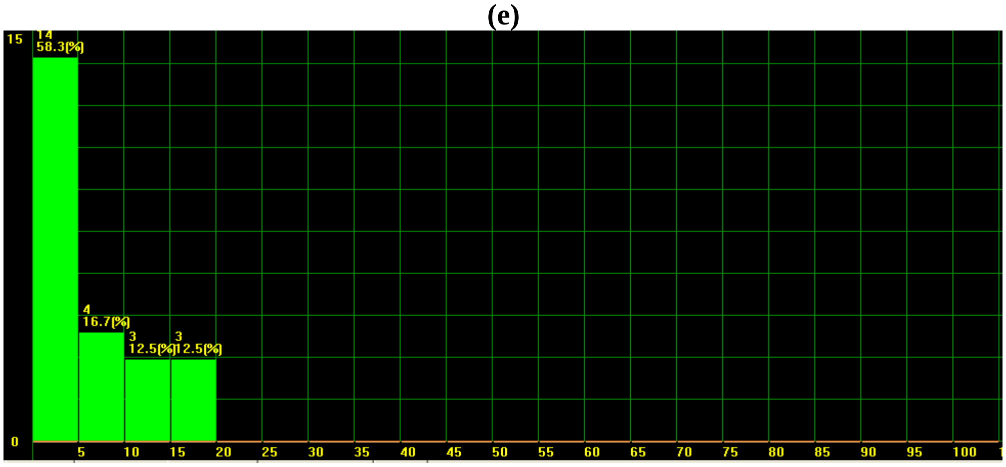
(c)



(d)

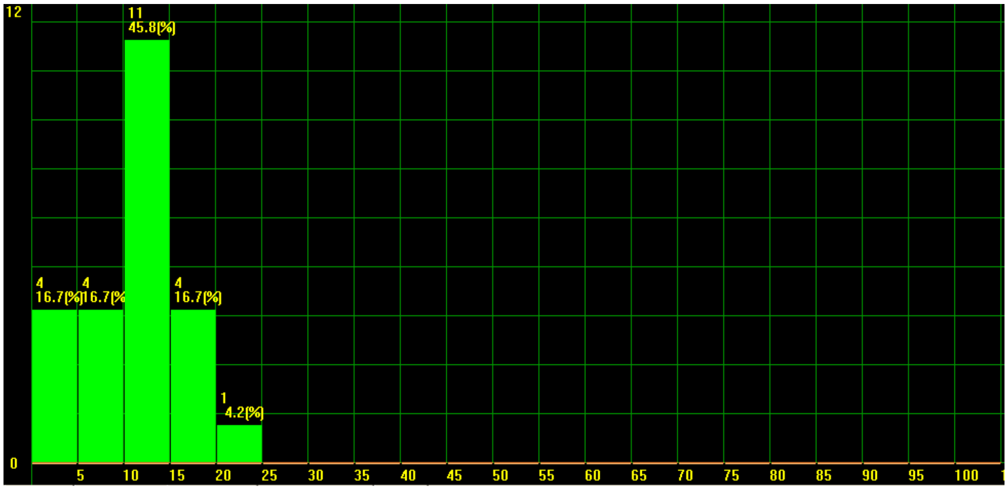


(e)

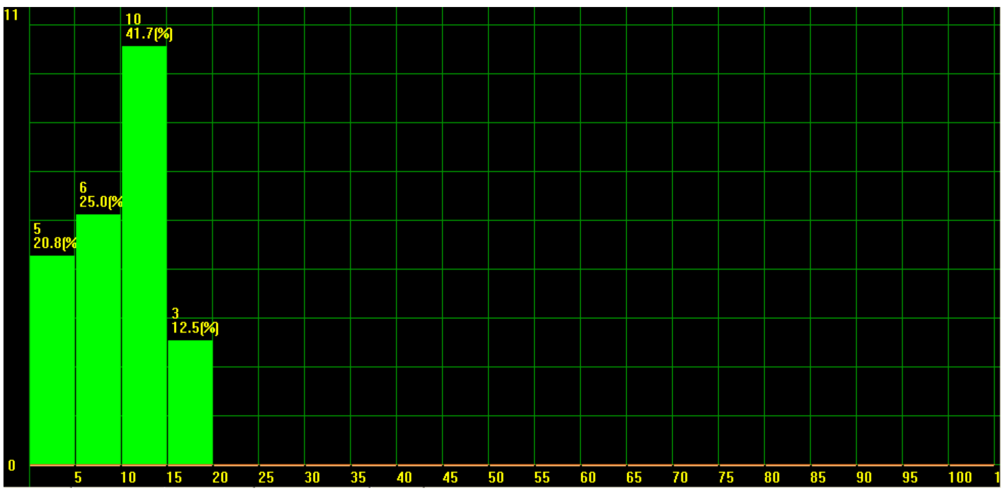


(f)

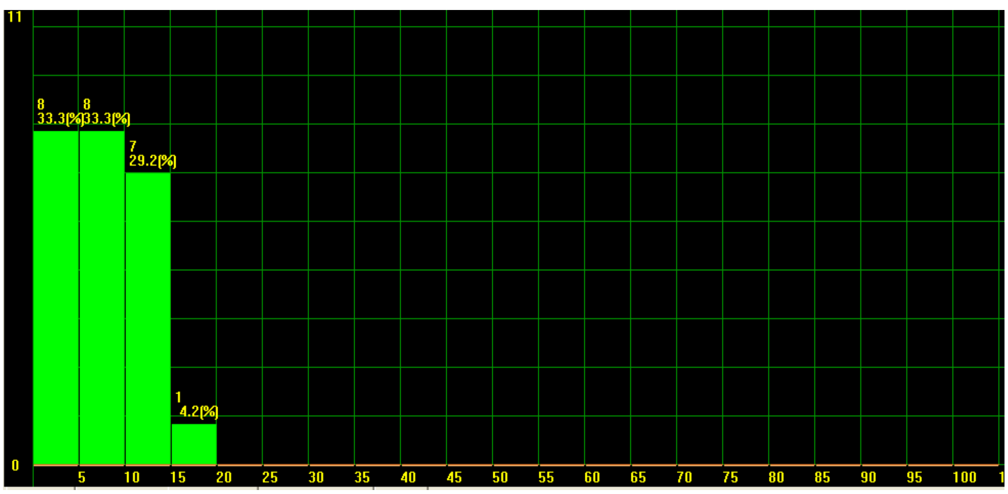
Figure 15: GEH Graphs for Starting Network (a) 1st Iteration, (b) 2nd Iteration, (c) 4th Iteration, (d) 6th Iteration, (e) 8th Iteration, (f) 10th Iteration


12
11
45.8(%)
4
16.7(%)
4
16.7(%)
4
16.7(%)
1
4.2(%)
0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

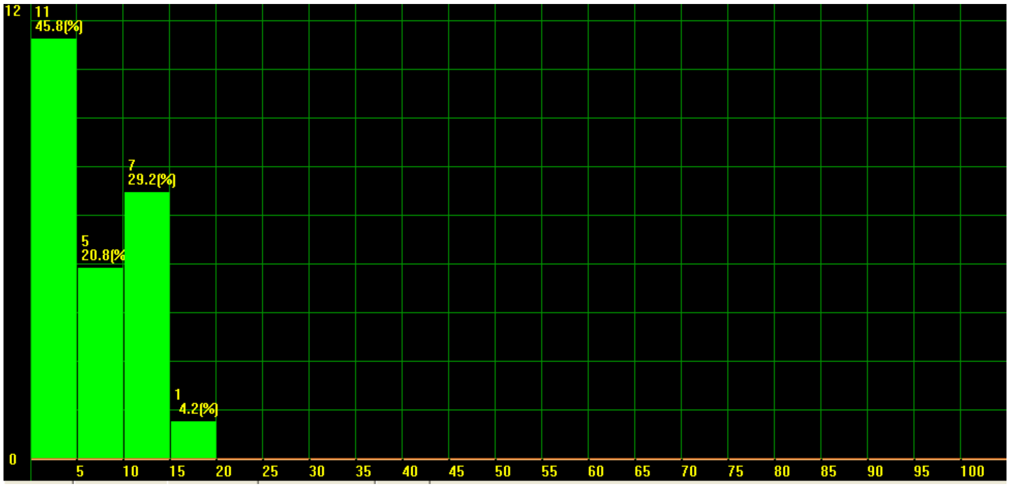
(a)


11
10
41.7(%)
6
25.0(%)
5
20.8(%)
3
12.5(%)
0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

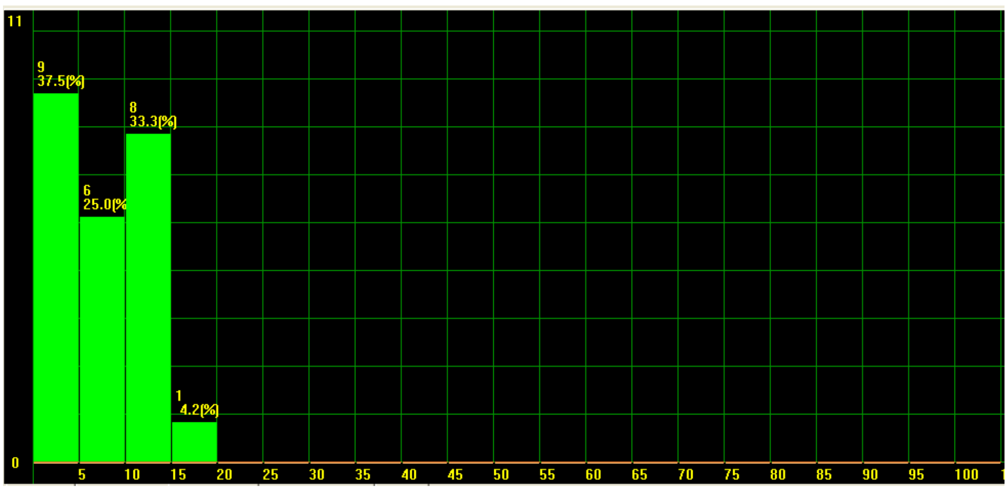
(b)


11
8
33.3(%)
8
33.3(%)
7
29.2(%)
1
4.2(%)
0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

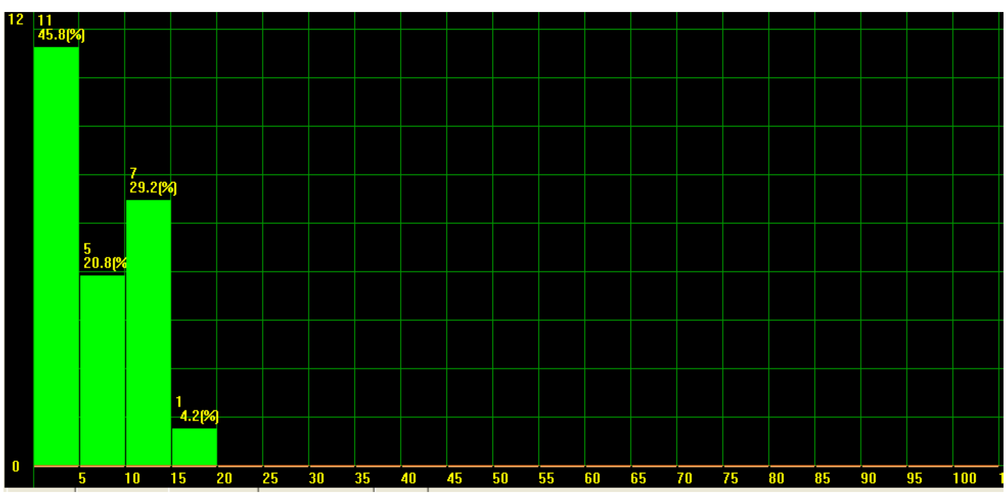
(c)



(d)



(e)



(f)

Figure 16: GEH Graphs for North-South Extended Network (a) 1st Iteration, (b) 2nd Iteration, (c) 4th Iteration, (d) 6th Iteration, (e) 8th Iteration, (f) 10th Iteration

As shown in Figure 16, better performance is obtained in GEH values through the 10 iterations of the OD estimation algorithm in network augmentation schemes. Figure 16 (f) shows that nearly 60% of the detectors in the network have GEH values less than 10 (vehicles/hour) suggesting that most of the simulated link traffic counts sufficiently matched the hypothetical grand truth. As mentioned previously, over the 10 iterations for the augmented scheme method, the rate of the improvement seems constant. A reason for this may be due to the limited number of iterations ran (10), without supply side calibrations and other adjustments needed to obtain better GEH statistic results. The results demonstrate the improved performance of the network augmentation scheme. Comparing the results of the extended network with our starting network, we can conclude that after 10 iterations there is improvement in GEH values.

Chapter 6: Conclusion and Future Research

This thesis aims to suggest a new approach for performing an origin-destination (OD) analysis on a controlled simulation network. This approach entails expanding an existing/starting network by adding different sub-areas to the existing network, allowing us to obtain our controlled simulation network. Because estimating OD is time consuming and data intensive in micro-simulation analyses, an incremental procedure is presented in this research as a way of avoiding excessive effort in estimating. Different sub-areas are added to the starting network in order to forego re-estimating the matrices for an existing area in which matrices have already been estimated. Thus, the time taken to obtain the OD matrices for vehicular micro-simulation is reduced. The network augmentation scheme, which is called the 'Fraction method', shows significant improvement in being able to match the network conditions to hypothetical grand truth data for our controlled network OD estimation iterations. This procedure can substantially reduce time spent on micro-simulation related modeling efforts.

To produce a more technical line of research, our original method developed to generate seed OD matrices could be refined for extended networks based on already estimated OD matrices. Also, practical applications of the method could be explored further. When applied directly to a large scale case study, this method was found to be successful (Bhagat 2014). We applied the same method for this paper and found that when an existing calibrated network is available, using the network augmentation scheme results in new schemes developed for OD estimation. As found, the new scheme shows significant improvement in being able to match the network conditions to hypothetical grand truth data for our controlled network OD estimation iterations. When the fraction scheme that combines a planning model assignment results of the large area are used to roughly preserve the calibrated OD matrices for a starting area, with the OD estimation focusing primarily on the new areas added to the network.

Future studies related to the topic of this thesis may explore better improvements in GEH results by iterating more than 10 iterations. In addition, accuracy in our controlled study could have been improved installing more loop detectors to generate the trajectory matrices for most of the vehicles in our network and then compare the result with this study. Finally, the development of a better seed origin-destination is also important. The role of a seed OD matrix is essential in OD estimation process because the final outcome is very much dependent on this initial value. Heuristic techniques have used without validation in our practical study, but seed OD generation methods needs more improvement for better results.

References:

Bhagat, A. (2014) "Expanding Vehicular Microscopic Traffic Simulation for Policy Analysis An Application to PierPASS in California" *PhD Dissertation,* UC Irvine ITS.

Choi, K., Jayakrishnan, R., Kim, H., Yang, I., Lee, J. (2009) "Dynamic O-D Estimation using Dynamic Traffic Simulation Model in an Urban Arterial Corridor" *Transportation Research Record: Journal of the Transportation Research Board (2133)*, pp. 133–141.

Lee, G. (2011) "Integrated Modeling of Air Quality and Health Impacts of a Freight Transportation Corridor" *PhD Dissertation*. UC Irvine ITS. Source: http://antpac.lib.uci.edu/record=b4751522~S7

Lee, G., You, S., Ritchie, S., Saphores, J., Jayakrishnan, R., Ogunseitan, O. (2012) "Assessing air quality and health benefits of the Clean Truck Program in the Alameda corridor, CA" *Transportation Research Part A (46*), pp. 1177-1193. Source: http://www.sciencedirect.com/science/article/pii/S0965856412000808

Lee, G., You, S., Ritchie, S., Saphores, J., Sangkapichai, M., Jayakrishnan, R. (2009) "Environmental Impacts of a Major Freight Corridor – A Study of I-710 in California" *Transportation Research Record: Journal of the Transportation Research Board* (2123), pp. 119-128. Source: http://trb.metapress.com/index/N45R65Q636275168.pdf

Lee, G., You, S., Sangkapichai, M., Ritchie, S., Saphores, J., Ogunseitan, O., Ayala, R., Jayakrishnan, R., Torres, R. (2010) "Assessing the Environmental and Health Impacts of Port-Related Freight Movement in a Major Urban Transportation Corridor" *University of California Transportation Center (UCTC-2010-08).* Source: http://escholarship.org/uc/item/9zb040jb.pdf

Lee, G., Soyoung (Iris), Y., Ritchie, S.G., Saphores, J.-D., Jayakrishnan, R., Ogunseitan, O. (2012). "Assessing air quality and health benefits of the Clean Truck Program in the Alameda corridor, CA." Transportation Research Part A, 46: 1177–1193