This is to certify that the
thesis entitled

TRAVEL TIME ESTIMATION AND SHORT-TERM PREDICTION IN URBAN ARTERIAL NETWORKS USING CONDITIONAL INDEPENDENCE GRAPHS AND STATE-SPACE NEURAL NETWORKS

presented by

AJAY KUMAR SINGH

has been accepted towards fulfillment
of the requirements for the

Master of Science degree in CIVIL ENGINEERING

Major Professor's Signature

7/6/2006

Date

MSU is an Affirmative Action/Equal Opportunity Institution

TRAVEL TIME ESTIMATION AND SHORT-TERM PREDICTION IN URBAN ARTERIAL NETWORKS USING CONDITIONAL INDEPENDENCE GRAPHS AND STATE-SPACE NEURAL NETWORKS

By

Ajay Kumar Singh

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

MASTER OF SCIENCE

Department of Civil and Environmental Engineering

2006

ABSTRACT

TRAVEL TIME ESTIMATION AND SHORT-TERM PREDICTION IN URBAN ARTERIAL NETWORKS USING CONDITIONAL INDEPENDENCE GRAPHS AND STATE-SPACE NEURAL NETWORKS

By

Ajay Kumar Singh

An important component of Advanced Traveler Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS) is the travel time estimation and short-term prediction on urban arterial networks. This thesis develops robust and efficient average travel time estimation and short-term prediction model for both congested and non-congested conditions appearing throughout a day on a network. A State-Space Neural Network model is proposed. An innovative implementation of Conditional Independence graph is used to identify the independence and interaction between observable traffic parameters that are used to estimate and predict the travel time. This led to the selection of relevant variables from a set of independent variables for travel time prediction. The predictive and computational performance of the Conditional Independence graph coupled with State-Space Neural Network outperformed the traditional State-Space Neural Network model in this study. The travel time estimation and prediction models are developed for links and routes in an arterial network.

Keywords: Travel Time Estimation, Short-term Prediction, Urban Arterials, Networks, Conditional Independence Graphs, State-Space Neural Networks.

Dedicated to my ignorance and endeavors to overcome it

ACKNOWLEDGEMENTS

I am greatly indebted to my advisor Dr. Ghassan Abu-Lebdeh for his continuous guidance, support, and patience throughout my graduate studies. I learnt so much from him in the area of Traffic Operations, Intelligent Traffic Control and modeling and optimization through his courses and interactions during research. I wish to acknowledge Dr. Francois Dion and Dr. R.V. Ramamoorthi for their invaluable comments towards improving this thesis and for the courses they taught to me. I express my sincere regards to Dr. Gilbert Baladi and Dr. Neeraj Buch who accepted me as a Teaching Assistant for their courses. I whole-heartedly thank all the faculty members, former and present graduate students, and the administrative staff of the Department of Civil and Environmental Engineering for making my graduate studies smooth and pleasant.

I wish to acknowledge the love and support of my parents and sisters. I also thank Richa for her love and patience throughout this work and further. I owe so much to my friends Deeps, Devesh, Priya, Tanmay, Shishir, Tithi, and Vinay for their memorable company and encouragement.

Ajay Kumar Singh

Michigan State University

2006

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Traffic congestion is a severe problem on urban arterials and freeways. In 2003, congestion in urban areas caused 3.7 billion hours of travel delay and 2.3 billion gallons of wasted fuel. This is an increase of 79 million hours and 69 million gallons from 2002 and costs more than $63 billion (Schrank and Lomax 2005). Advanced Traveler Information Systems (ATIS) and Advanced Transportation Management Systems (ATMS) are two important components of Intelligent Transportation Systems (ITS). ATIS and ATMS are identified as innovative solutions to traffic congestion problem in urban areas.

Travel time is an easily perceptible measure of traffic conditions on arterials or on freeways and is equally popular among researchers, traffic engineers, and users of the transportation system. Travel time is a key input to the traffic operations of the system as well as transportation planning. It indicates the overall performance of the system, a real-time measure of traffic congestion, and an input to the traffic management strategies (Longfoot 1991).

The main requirements of ATIS are: to help drivers make better and informed choices, to assist drivers in avoiding congestion and unexpected delays, and to reduce the time spent driving. This requires dissemination of reliable travel time information to drivers about the current traffic conditions and also about near future when they actually start their trips. The ATMS must have some sort of predictive capability in order to manage and control the transportation system in proactive manner: "... for these systems (ATIS and ATMS) to be effective, the generated strategies should be proactive (i.e. based on predicted traffic conditions) as opposed to reactive, in order to avoid many undesirable effects such as overreaction, which reflects the situation where many travelers react to a known current traffic condition in a similar fashion resulting in simply transferring the congestion to another location" (Ben-Akiva, et al. 2001).

The traffic management strategies on arterial networks under congested and over-saturated conditions have been proposed by various researchers. The optimal signal control is one of the traffic management strategies applicable on arterial networks. Abu-Lebdeh and Benekohal (2000, 2003) propose traffic management strategies based on dynamic signal control and queue management on arterials. A recent paper by Chen and Abu-Lebdeh (2006) presents dynamic-signal dynamic-speed traffic management strategy for arterials. Wey and Jayakrishnan (2004) present an integer-linear programming formulation for network traffic control scheme optimization. The studies for dynamic traffic signal optimization usually minimize the total delay on the traffic network. This delay is correlated with the travel time on the arterial network.

There is a need of efficient, accurate and reliable travel time estimation and prediction methodologies for urban arterial networks. A significant amount of research for travel time estimation and prediction on freeways and highways has been done. But, the area of urban arterials lacks comparable research. The probable reasons are the vastness of the arterial system, complexity of traffic behavior on arterials due to presence of signal control devices, and lack of extensive traffic surveillance systems as compared to freeways. This implies that travel time prediction on urban arterials should rely only on easily available data that can be obtained in the field using traffic detection technologies like loop detectors or other image-processing based technologies. A methodology is required which can avoid the dependence on extensive traffic surveillance systems for travel time estimation and prediction.

The complexity of traffic dynamics is another aspect of travel time on arterials which should be explicitly taken into account in the modeling approach. The traffic dynamics on arterial networks is more complex than on freeways because of the signal interruptions. These signal interruptions cause interference to the traffic movement which is not present in uninterrupted facilities like freeways. The traffic stream approaching a signalized intersection undergoes a platoon compression, a complete stop during red interval, and then a platoon expansion after the vehicles cross the signalized intersection. The interference that is caused by this chain of compression and expansion within small distances adds to the complexity of process. Moreover, the turning movements and phasing scheme for each movement causes interference to vehicles throughout the length of arterial link.

This thesis is based on the state-space notion of traffic that is suitable for urban arterials. The state-space representation to a complex dynamical system provides insight to model it better. The State-Space Neural Networks are a generic form of Recurrent Neural Networks and are used to model the travel time estimation and prediction in this study. The State-Space Neural Networks or, generally, Artificial Neural Networks (ANN) is an efficient input-output mapping technique. One of the drawbacks of ANN is that it does not provide the understanding of the process and the interaction among variables in the process being modeled. So, Conditional Independence Graphs are used, which are a powerful statistical technique to analyze the independence and interaction among variables involved in a process. The combination of Conditional Independence graphs and State-Space Neural Networks is implemented for the first time in this study. It provides a robust and efficient modeling framework to estimate and predict travel time for short-term in future. This modeling framework relies on data that can be easily available in the field and hence has enormous potential in ATIS and ATMS applications.

Chapter 1

Literature Review

The research on travel time estimation and short-term prediction is similar to short-term traffic forecasting. The short-term traffic forecasting includes forecasting of traffic variables like travel time, average speed, flow rate, occupancy, queue length etc. There has been a lot of research on travel time prediction for freeways. The application area of arterial networks is still lagging as compared to that of freeways. Here, travel time prediction studies have been surveyed for freeways as well as arterial networks. The techniques of Conditional Independence graphs and State-Space Neural Networks are used in this study. A brief survey of literature on the applications of these techniques is presented.

1.1 Travel Time Estimation and Prediction on Freeways

A recent paper by Vlahogianni, et al. (2004) gives a critical discussion on the short-term traffic forecasting techniques for freeways and urban arterials for different types of implementation like Advanced Traveler Information Systems (ATIS) and Advanced Transportation Management Systems (ATMS). The discussion of methodologies for traffic forecasting identifies two broad categories: parametric and non-parametric techniques. This review asserts that the non-parametric modeling techniques like Artificial Neural Networks (ANN) is promising for traffic forecasting problems and gives robust and accurate models. Zwet and Rice (2004) propose a travel time prediction scheme for a freeway section by means of linear regression with time-varying coefficients. These coefficients are subjected to the time of day and time until start of a vehicle on a section. The varying coefficients show the dynamic behavior of the traffic which can also be captured by other advanced parametric and non-parametric techniques. Ishak and Al-Deek (2002) proposed a non-linear time-series approach to make short-term predictions of speed using the most recent speed profile at each loop detector station and then mathematically found travel time from speed. This study is based on freeways data and also presents statistical analysis to identify the parameters like congestion index, rolling horizon, prediction horizon and their interaction terms with congestion index to be significant for model's performance. Wu, et al. (2004) use Support Vector Regression to predict travel time for highways in Taiwan. Chien and Kuchipudi (2003) applied Kalman Filtering to predict the link and path based travel time based on the time-series data collected through wireless technologies on freeways.

There is significant volume of studies related to the application of different type of ANN in short-term travel time prediction on freeways. Mark and Sadek (2004) proposed an ANN model for freeways to predict the experiential travel time under transient traffic conditions, including incidents. They found out that speed appears to be the most influential input variable for travel time prediction on freeways. A special type of Recurrent Neural Networks (RNN) called State-Space Neural Networks (SSNN) is employed by van Lint (2004) and van Lint, et al. (2002) for freeways travel time prediction on the basis of flow and speed. They further show that the analysis of the internal states and weight configurations of SSNN could develop an internal model that is closely linked to the underlying traffic process.

1.2 Travel Time Estimation and Prediction on Urban Arterials

Travel time prediction on urban arterials is still lacking efficient models. Sisiopiku and Rouphail (1994) presented a literature review of methods for travel time estimation on urban arterials. They listed the traffic parameters used by various researchers that influence the travel time on arterials and concluded that most of the existing models are link-specific rather than estimating travel time at a section (path) level. Sisiopiku et al. (1994) proposed the linear regression equations between travel time and percentage occupancy for occupancies in a range of approximately 17 to 60 percent. No model for the congested conditions which results into higher percentage occupancy is reported in this study. Rajaraman (1998) proposed travel time estimation model that has two components, overflow queue estimation and travel time calculation. Stathopoulos and Karlaftis (2003) proposed a multivariate time-series state-space model for traffic flow prediction on arterials and compare it with other time-series techniques like ARIMA modeling. They also assert that different model specifications are appropriate for different time periods of a day. Lin, et al. (2004) decomposed travel time on arterials into two components, free flow travel time and delay. The total delay on an arterial is further decomposed into link delay and intersection delay. The travel time prediction model is based on the flow condition, the proportion of net inflows into the arterial from the cross streets, and the signal coordination level. The inherent limitation in this model is that it relies on existing intersection delay formulas which are not suitable for over-saturated conditions. A recent study reported by Liu, et al. (2006) presents a State-Space Neural Networks and the extended Kalman filter (EKF) hybrid model for travel time prediction. The model is developed for a test bed in the Delft, Netherlands using license plate cameras and loop detectors for data collection of travel time and flow rates. This model is based on a particular section of road and is not taking into account the signal control parameters and the geometrics of the arterial. This limits this model's applicability to a new arterial network where the signal control and geometrics are different. These studies identify the need of a generalized travel time prediction model that explicitly takes into account varying traffic demand, turning movements, signal control, and geometrics on urban arterial networks. The present study overcomes these limitations as the geometrics and signal control are taken as independent variables affecting travel time.

There are some research studies forecasting other traffic variables like average flow rate, queue length, speed etc. for signalized arterials. Since, travel time is dependent on the input variables like average flow rate, queue length, speed etc., the forecasting of these traffic variables can be integrated with travel time estimation and prediction models. Gang and Su (1995) developed a queue length prediction model at signalized intersections based on given flow, occupancy, queue length and signal state from past. Ledoux (1997) also developed models to forecast queue length at signalized intersections from flow and queue length measurements in past time periods. These models when integrated with travel time estimation and forecasting models can result in prediction of more than one traffic parameter for urban arterials.

1.3 Conditional Independence Graphs Applications

The Conditional Independence graphs are a component of Bayesian Belief Networks which have several applications in traffic engineering. Conditional Independence relationships and ideas of Bayesian forecasting are proposed by Whittaker, et al. (1997) for predicting traffic for freeways. A dynamic state-space model is proposed in the study with optimal state estimation is coming from the Kalman filter. Apart from this reference, no other paper is found to explicitly use Conditional Independence graph-based models for short-term traffic forecasting or travel time prediction.

1.4 State-Space Neural Networks Applications

The State-Space Neural Networks have been recently used for short-term travel time prediction on freeways as well as urban arterials. The studies shown in van Lint (2004) and van Lint, et al. (2002) use State-Space Neural Networks for freeways travel time prediction the basis of flow and speed. A hybrid model for urban arterial travel time prediction has been recently introduced by Liu, et al. (2006). This hybrid model is based on State-Space Neural Networks and the extended Kalman Filter. This study also finds the State-Space Neural Networks proposed by these researchers to provide a reliable and generalized travel time estimation and prediction model.

1.5 Conclusions

The literature survey for this study suggests there is a need to develop a travel time estimation and prediction model that relies on traffic parameters that are easily obtained in field. These traffic parameters can be obtained using basic detection technologies like loop detectors and image processing-based technologies. The modeling approach should explicitly take into account the geometrics and signal control parameters. The inclusion of geometrics and signal control in the modeling approach will provide a generalized model applicable to a real arterial network. A State-Space Neural Network (SSNN) model which has been successfully employed by van Lint (2004) and van Lint, et al. (2002) for freeways is found promising here. The SSNN model is expected to capture the non-linear and complex spatio-temporal relationship between traffic parameters and travel time on a section of arterial. The Conditional Independence (CI) Graphs is identified as a statistical tool to understand the interaction and independence among parameters affecting travel time or any other output in a process. The information obtained using CI graphs about the process can be used for SSNN modeling. The combination of CI graphs and SSNN can provide a reliable and efficient travel time prediction model for arterials.

Chapter 2

Problem Statement and Research Objective

The problem of average travel time estimation and short-term prediction is a basic component of Advanced Traveler Information Systems (ATIS) and Advanced Transportation Management Systems (ATMS). With the advancement of traffic surveillance infrastructure on urban arterials and freeways, Intelligent Transportation Systems (ITS) become an integral part of solutions to congestion and transportation system management.

It is important that reliable and efficient methodologies are developed which can make the best use of existing traffic surveillance information and improve traffic operations. This study provides a prediction model that can rely on existing surveillance systems like loop-detectors and other basic information available about the network like geometrics, and existing signal control scheme for use as input information. Thus, it avoids the deployment of more sophisticated and costly surveillance systems which are based on tracking individual vehicles in a traffic network to provide accurate estimates of travel time and prevalent traffic conditions information.

This chapter introduces the urban arterial networks which is the study area for this work. The notion of average travel time and its prediction is illustrated. The difference between travel time estimation and prediction is made clear. The problem formulation of travel time estimation and short-term prediction provides a conceptual base for modeling the travel time on arterial networks. The dependence of average travel time on the traffic variables like flow rate, average speed, queue length, geometrics and signal control variables, which shape the traffic conditions on an arterial network is presented. The last section provides an overview of the different steps in the modeling framework adopted in this research.

2.1 Background and Definitions

The short-term travel time prediction model developed is based on and is applicable to urban arterial networks. This section provides the definition of an arterial link, arterial, and arterial route for which the model is applicable.

An urban arterial network consists of several arterial links aligned in a geo-spatial manner forming a grid. This grid network can be broadly represented by major and minor streets crossing each other. The crossings in this network are signalized intersections. The signalized intersections are main point of operations and control in the network.

First, the definition of an arterial link, arterial, arterial route, and a network is presented below-

Definition 1: An 'Arterial Link' is a two-way length of the road that begins just at the end of a signalized intersection and extends through (and includes) the next downstream signalized intersection.

Definition 2: An 'Arterial' is the set of arterial links which are traversed by a vehicle remaining in through movement only starting from an origin arterial link to the end of its destination arterial link.

Definition 3: An 'Arterial Route' is the set of arterial links which are traversed by a vehicle being in through movement as well as taking left and right turns as required starting from an origin arterial link to the end of its destination arterial link.

Definition 4: An 'Arterial Network' is a system which contains all the arterial links in the area of study. These arterial links are arranged such that the network symbolizes a grid network having major arterials and minor arterials.

In this thesis, unless otherwise stated, a link means an arterial link, and a route means an arterial route as stated in above definitions.

An arterial network is traversed by individual vehicles contained in a traffic stream. The ATIS and ATMS applications are generally dependent on the travel time value of a traffic stream rather than travel time experienced by an individual vehicle traversing an arterial link. This notion of travel time of an individual vehicle and all the vehicles following a traffic stream is stated below.

Definition 5: The 'Individual Vehicle Travel Time' is the travel time obtained by tracking an individual vehicle along an arterial link. Hence, it is the property of an individual vehicle and may not represent the travel time incurred on a particular arterial link or a set of such links.

Definition 6: The 'Average Travel Time' is the average value of the travel time incurred by individual vehicles if each vehicle can be tracked on an arterial link. The average travel time is a property of an arterial link indicating the level of service of link.

Once, a prediction model is developed for the average travel time on a link, it can be easily used for an arterial or an arterial route comprising a set of links. This study aims for a model to be used for ATIS and ATMS applications to improve the level of service and operations on an arterial link or an arterial route in the system.

2.2 Difference between Average Travel Time Estimation and Prediction

The travel time studies done by several researchers deal with the estimation of travel time or its prediction for future unseen traffic conditions. The traffic conditions on an arterial link at any time period can be represented by a set of traffic variables like average flow rate, average speed, queue length etc. The term used for the traffic conditions on an arterial link in this study is the 'state' of the link. The notion of state of the system (arterial network comprising links) is important in this study as it acts as input variable to estimate and predict the average travel time. The definition of the estimated and predicted average travel time is important to state here.

Definition 7: The Estimated Average Travel Time on an arterial link or arterial route is the 'experiential' travel time by the vehicles traversing that link or route. The term 'experiential' denotes the travel time which is already experienced by the vehicles that have departed in the current or past time-periods when the state of the system was known or estimated.

Definition 8: The Predicted Average Travel Time on an arterial link or arterial route is the travel time which will be experienced by the vehicles when their departure is in a future time-period during which the traffic conditions are unseen and unknown.

There is an inherent distinction between the estimation and prediction of average travel time. The travel time estimation is done when the vehicle has already departed and has experienced a travel time in its trip. This travel time is a complex function of current or past state traffic conditions that are signified by independent variables like average flow rate, average speed, queue length etc. These independent variables can in turn be observed in the field directly by using traffic detection technology or can be measured from other known variables, or even can be predicted from the past values of some variables. The average travel time estimation problem requires a modeling technique which is able to efficiently model the static input-output relationships (van Lint 2004).
The average travel time prediction is a different problem from the estimation as it needs a modeling technique to capture the dynamic (time-dependent) relationship among known traffic variables and future traffic conditions or future travel time. The prediction of travel time is done using the current state traffic variable values for the travel time that will be

experienced in the future, when the future traffic conditions are unknown. Thus, travel time prediction is more complex than travel time estimation problem as it requires the consideration of dynamic spatio-temporal relationship that occurs between traffic variables. This thesis provides the methodology for estimating the average travel time as well its prediction for the short-term into the future.

2.3 Formulation for Average Travel Time Estimation and Short-Term Prediction

Let an arterial route on a network consisting of n number of arterial links represented as $X_1, X_2, X_3, \ldots X_n$. Let t_p^* be the current time-period of departure of a vehicle which starts from entry-point (upstream intersection) of link X_1 and ends its trip after clearing the intersection that ends link X_n. This time-period t_p^* is starting from a time-instant t_0^* and ending at time instant t_1^*. So, $t_p^* = [t_0^*, t_1^*]$. The traffic parameters that shape prevailing traffic conditions (like average flow rate, average speed, queue length etc.) on the arterial network in the current departure time-period t_p^* can be obtained in the field through traffic detection system, or can be estimated through other already obtained traffic variables.

It is assumed that the traffic conditions are constant within each time-period if the length of such time-period is taken very small. In many traffic engineering studies and particularly short-term travel time forecasting, it is assumed that traffic conditions are constant within short time-periods of 1 to 15 minutes. The time-period of 1 minute provides high temporal resolution to the data and may be useful if cycle-by-cycle analysis is required. On the other end, 15 minutes is a low temporal resolution for forecasting as the unstable traffic conditions during congestion and over-saturation may get overlooked or averaged due to long time-period. Sisiopiku, et al. (1994) assert in a travel time estimation study that the effects from cycle failures, short-term events, congestion built up downstream of the subject link, and so forth cannot be detected using 15-minutes observation periods. Mark and Sadek (2004) recommend temporal resolution of 5 minutes for travel time forecasting and find that no statistically significant increase in performance was gained by increasing the temporal resolution of the data-set from 5 minutes time averages to 1 minute time averages. In consideration to studies on travel time forecasting on freeways and arterials, here, the length of each time period is taken as 5 minutes which is called the aggregation interval, Φ. Thus, $\Phi = t_1^* - t_0^*$.

Let the known traffic conditions on a link X_1 in current departure time-period t_p^* be $V(X_1, t_p^*)$. The vector $V(X_1, t_p^*)$ which contains the variables depicting traffic conditions on a link X_1 in a time-period t_p^* is called 'State' of an arterial link. The 'State' of an arterial or arterial route is simply a vector which contains the state of each link comprising that arterial or route during a given time-period. This generic term for traffic conditions on a link is very important in the context of travel time estimation and prediction because it is the set of input variables like average flow rate, average speed, queue length etc., which are affecting the travel time. It will be seen from now on that the modeling approach is trying to model the function that exists between the state of a link and travel time on that link.

The average travel time to traverse a link X_1 during the current departure time-period t_p^* is denoted as TT (X_1, t_p^*). The problem of estimation of average travel time can now be defined using the terminology explained in the earlier paragraphs as follows:

Definition 9: The average travel time estimation on arterial link X_1 is the modeling problem to approximate the underlying function between average travel time (TT (X_1, t_p^*)) and the state ($V(X_1, t_p^*)$) on link X_1 during current departure time-period t_p^*.

The average travel time prediction problem is formulated now after stating the average travel time estimation problem. Assume that the average travel time is to be predicted for a future time-period which is just starting when the current departure time-period t_p^* ends (i.e. time instant t_1^*). A term called 'prediction horizon' denoted as Δ is stated, which is a short-interval of time in the future for which the average travel time is to be predicted. The future departure time-period is denoted as $t_{p+1\Delta} = [t_0, t_1]$, where, this time-period begins at t_0 and ends at t_1. This implies that the t_0 value of future time-period is same as t_1^*, the ending time instant of the current time-period. So, $t_0 = t_1^*$. Moreover, the length of this future departure time-period, $t_{p+1\Delta}$ is simply the addition of prediction horizon

(Δ) to the ending time of current time-period t_p^* (t_1^*). So, $t_{p+1\Delta} = t_1^* + \Delta$. Since, the future departure time-period of $t_{p+1\Delta}$ is a single multiple of prediction horizon (Δ), it is called as one-step future time-period and the predicted value of average travel time for this one-step future time-period will be called the one-step predicted average travel time. Conveniently, the prediction horizon (Δ) can be assumed same as aggregation interval (Φ) defined earlier. So, $\Delta = \Phi$ (= 5 minutes, here). This assumption is made so that the temporal scale at which the state of a link and travel time is analyzed can be same for all the time-periods whether it is current, past, or future. Moreover, it makes the task of updating the state of the link and travel time values evolving with time-periods easy.

The predicted average travel time to traverse a link X_1 during the one-step future departure time-period $t_{p+1\Delta}$ is denoted as TT (X_1, $t_{p+1\Delta}$). The problem of one-step average travel time prediction can now be defined using the terminology explained in earlier paragraphs as follows,

Definition 10: The one-step average travel time prediction on an arterial link X_1 is the modeling problem to approximate the underlying function between the one-step predicted average travel time (TT (X_1, $t_{p+1\Delta}$)) during one-step future departure time-period, $t_{p+1\Delta}$, the current state ($V(X_1, t_p^*)$), and the estimated average travel time TT (X_1, t_p^*)) on link X_1 during current departure time-period t_p^*.

It is important to note in the above definition that the state or traffic conditions and the estimated average travel time of a link in the current time period are used as inputs for the one-step predicted average travel time which is expected to occur when unseen and unknown traffic conditions will happen on that link in the one-step future time-period.

The average travel time prediction for any n^{th} step in the future can be developed on the same basis, where $n \geq 1$. Assume that the time-period in n^{th} steps ahead in future is represented as $t_{p+n\Delta}$, when the n^{th} step predicted average travel time is denoted as $TT(X_1, t_{p+n\Delta})$. The known information for average travel time prediction for n^{th} step is still the state of the link, $V(X_1, t_p^*)$ and estimated travel time, TT (X_1, t_p^*) at current time-period t_p^*. Moreover, the predicted average travel time till $n\text{-}1^{th}$ steps are also taken as input variables for n^{th} average travel time prediction. The definition of average travel time prediction for n^{th} step in future is presented as follows:

Definition 11: The n^{th} step average travel time prediction on an arterial link X_1 is the modeling problem to approximate the underlying function between the n^{th} step predicted average travel time (TT $(X_1, t_{p+n\Delta})$) during the n^{th} step future departure time-period, $t_{p+n\Delta}$ and the current state ($V(X_1, t_p^*)$), the estimated average travel time TT (X_1, t_p^*) on link X_1 during the current departure time-period t_p^*, and the predicted average travel time till the $n\text{-}1^{th}$ step.

So, the predicted travel time at the n^{th} future departure time-period = function (current state, estimated travel time at the current departure time-period, and the predicted travel time till $n\text{-}1^{th}$ future time-step).

The average travel time prediction for the n^{th} step in the future is called the Multiple-Step prediction. The Short-term prediction of average travel time which is the ultimate aim of this work can be defined as follows:

Definition 12: The Short-term average travel time prediction is solving the multiple step average travel time prediction problem. The multiple-step average travel time prediction is a chain process of prediction from step =1 to step n as required for the system.

This study focuses on n =1 step average travel time prediction problem. The n ≥1 steps prediction can be done if required using the same methodology adopted in this study for n=1 step. However, it is expected that the accuracy may decrease as the prediction is done farther in n steps in the future. Ishak and Al-Deek (2002) found in a travel time prediction study on freeways that for longer prediction time-steps, less accuracy in predictions are observed due to the dynamic nature of traffic conditions and the increased likelihood of larger deviations from actual conditions. Thus, the short-term average travel time prediction is done in the present work in terms of one-step prediction only.

Figure 2.1 presents a pictorial form of the above formulation in a time-space domain. The space domain shows a single link X_1 and the time scale shows time-periods ranging from the current departure time-period (t_p^*) to the n-step time-period in future ($t_{p+1\Delta}$, $t_{p+2\Delta}$,

$t_{p+3\Delta}$...) for which the prediction is to be made. This time-space domain shows the assumed vehicle trajectory in each time period. The current departure time period, t_p^* is termed as 'Current Horizon' and future departure time-periods are termed as 'Prediction Horizon'. As shown and consistent with the terminology explained in this section, current horizon is starting from time instant t_0^* and ending at time-instant t_1^* forming time period t_p^*(length of t_p^* = aggregation interval, Φ). The state of link X_1 representing prevalent traffic conditions is shown as $V(X_1, t_p^*)$ in the current horizon. The first-step prediction horizon starts from time-instant t_0 (same as t_1^*) and ends at t_1, forming first step future departure time-period, $t_{p+1\Delta}$. Similarly, second-step, and third-step prediction horizons are as shown. Each future departure time-period is of length Δ. As said earlier, Δ = Φ.

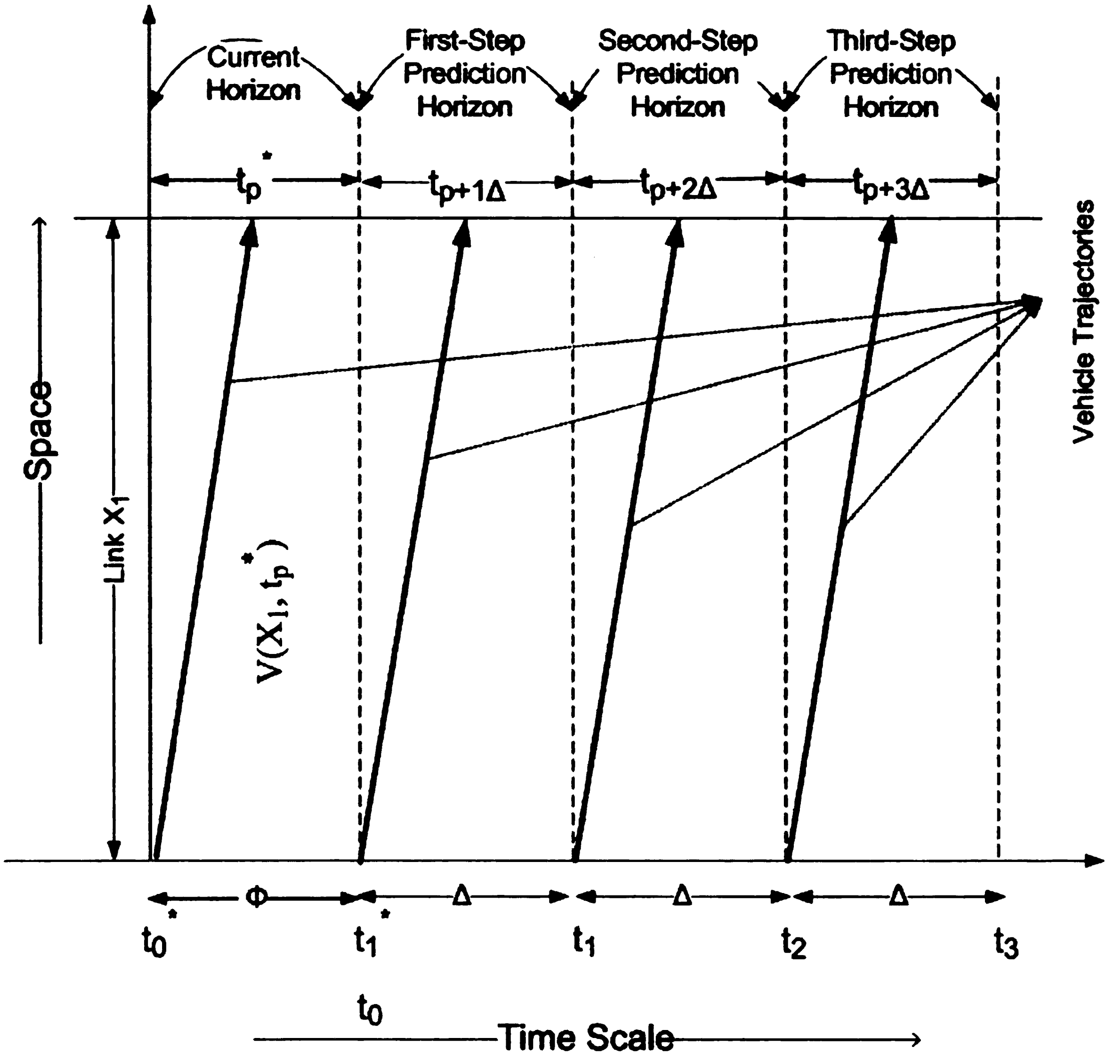

Figure 2.1 – Time – Space Domain of Problem Formulation for Average Travel Time Estimation and Short-Term Prediction.

2.4 Factors Affecting Average Travel Time Estimation and Prediction

The traffic conditions on an arterial link termed as State of the link represent the set of variables which affect average travel time. Here, the traffic variables are identified whose values in a particular time-period can affect the average travel time experienced during that time-period. These variables will form the inputs in the modeling approach for both average travel time estimation as well as prediction.

The average travel time on a link is a complex function of the traffic demand on the link, supply of the link (usually capacity of the intersection, also, capacity of the link if signal spacing is large) and other factors like weather conditions, incidents, driver behavior, vehicle composition, etc. It is obvious that an urban arterial network will have traffic demand fluctuations depending upon the average flow rate on the network during different times of the day. Due to this varying traffic demand it is intuitive that average travel time will be higher for high traffic demand, resulting into low speed and lower for low traffic demand resulting into higher speed. So, average flow rate and average speed are required to be added as input variables affecting travel time.

The varying traffic demand at arterial links in a network interacts with the signal control parameters like cycle length, green splits, offsets and phasing sequence. This interaction of the traffic demand at a link and signal control scheme at the intersection has different outcomes which depend on the traffic demand and signal control scheme itself. If the traffic demand is low (non-congested conditions), the average travel time does not have wide fluctuations because the demand is much lesser than the capacity. In this scenario, the adjacent links in the arterial act independently and have very less interaction among each other. On the other hand, when traffic demand is high (congested and over-saturated conditions), the interaction causes queue build-up and even queue spill-over at an intersection which results in affecting the progression on the upstream links. This suggests including signal control parameters and progression variables like green interval, red interval, and offsets as input variables affecting travel time. Also, presence of queue length at the beginning of green at signalized intersections should be taken into account for through and right-turning lanes and left-turning lane separately.

The other complexity in traffic behavior is due to the lane-changing among vehicles traversing a link. This lane-changing is due to the turning movements on a link and is associated with the driver behavior. The average travel time on a link is a complex function of this lane-changing behavior. The number of lane-changes in a link is not considered as input variable because it is difficult to determine in the field. But, the turning movements which cause the lane-changing are explicitly accounted in this study. The average flow rate and average speed should be accounted for through, left and right-turning movements separately to model average travel time for these three types of movements.

It is also known that average travel time is a basic function of the geometrics of a link, especially link length and posted speed limit. So, speed limit and estimates of the ideal travel time for through, left and right-turning movement which are based on link length, speed limit, and known green intervals for through and left-turning movements should be included in the study.

The present work aims to find estimation and prediction models for average travel time on a link for through, left and right-turning movements separately. The variables that act as independent variables affecting average travel time for through, left and right-turning movement in this study are specified below:

- Traffic demand variables – Average flow rate for through, left and right-turning movements denoted as V_T, V_L, and V_R respectively.

- Average speed for through, left and right-turning movements denoted as S_T, S_L, and S_R respectively
- Signal control variables – Green interval for an approach, denoted as G, and offset between a pair of signalized intersections for an approach, denoted as Off.
- Queue Length variables – Queue length at through and right-turning movement lanes denoted as Q_T. Queue length at left-turning lane denoted as Q_L.
- Speed limit posted for a link, denoted as S.L.
- Geometrics of arterial link variables – Ideal travel time for through and right-turning movement, denoted as ITT_T. ITT_T is the summation of ratio of length of link to speed limit (S.L.) plus the green interval for the left-turning movement. Ideal travel time for the left-turning movement is denoted as ITT_L. ITT_L is the summation of ratio of link length to speed limit (S.L.) plus the green interval for the through movement and red interval for an approach. The phasing scheme for intersections is assumed as exclusive left-turn phase with leading green in this study. Hence, the definition of ITT_T and ITT_L is given according to this assumption of leading green only.

The variables listed above are highly interrelated with each other. For example, average speed is the result of interaction of average flow rate with signal control parameters and queue length at signalized intersection. The fluctuations in average travel time resulting from the interactions discussed above makes it a dynamic and complex spatial-temporal function of the demand and supply of the link under consideration. This thesis aims to understand the interaction among these factors and model the travel time so that it accounts for this interaction.

2.5 Modeling Approach for Travel Time Estimation and Short-Term Prediction

The travel time estimation problem is a first step for prediction of average travel time. The estimation and short-term prediction of travel time has been done separately for through, left and right-turning movement on an arterial link. So, three different models have been developed in this problem, but, the methodology underlying these three models is based on the same techniques and tools.

The literature survey shows that several modeling techniques have been used for traffic variables forecasting and in particular travel time estimation and forecasting. The modeling techniques used in the present work are explained in details in the following chapters. Here, it is worth presenting a brief framework or flowchart as shown in Figure 2.2 of the entire modeling approach to prepare for the upcoming chapters. This figure starts with problem formulation which has already been dealt in this chapter. The conclusion of problem formulation is that travel time should be modeled explicitly for three types of movements. The further chapters will deal with data-collection using TSIS-CORSIM, average speed estimation for each type of movement. This estimation of average speed estimation is needed so as to propose a travel time model that can rely on readily available data in the field. The understanding of interaction between factors affecting travel time is required and is done using a statistical technique called the Conditional Independence (CI) graph. This technique helps particularly in refining the state of a link and in selecting parent input variables. The parent input variables are found sufficient to estimate and predict the travel time. Using these parent input variables, State-Space Neural Networks (SSNN) models are developed for travel time estimation on arterial links. The estimated travel time from the SSNN model, parent input variables from the CI graph, and the average speed estimates are used to develop SSNN prediction models for travel time. Once SSNN models are developed for arterial links, the results are extended for arterial route travel time prediction.

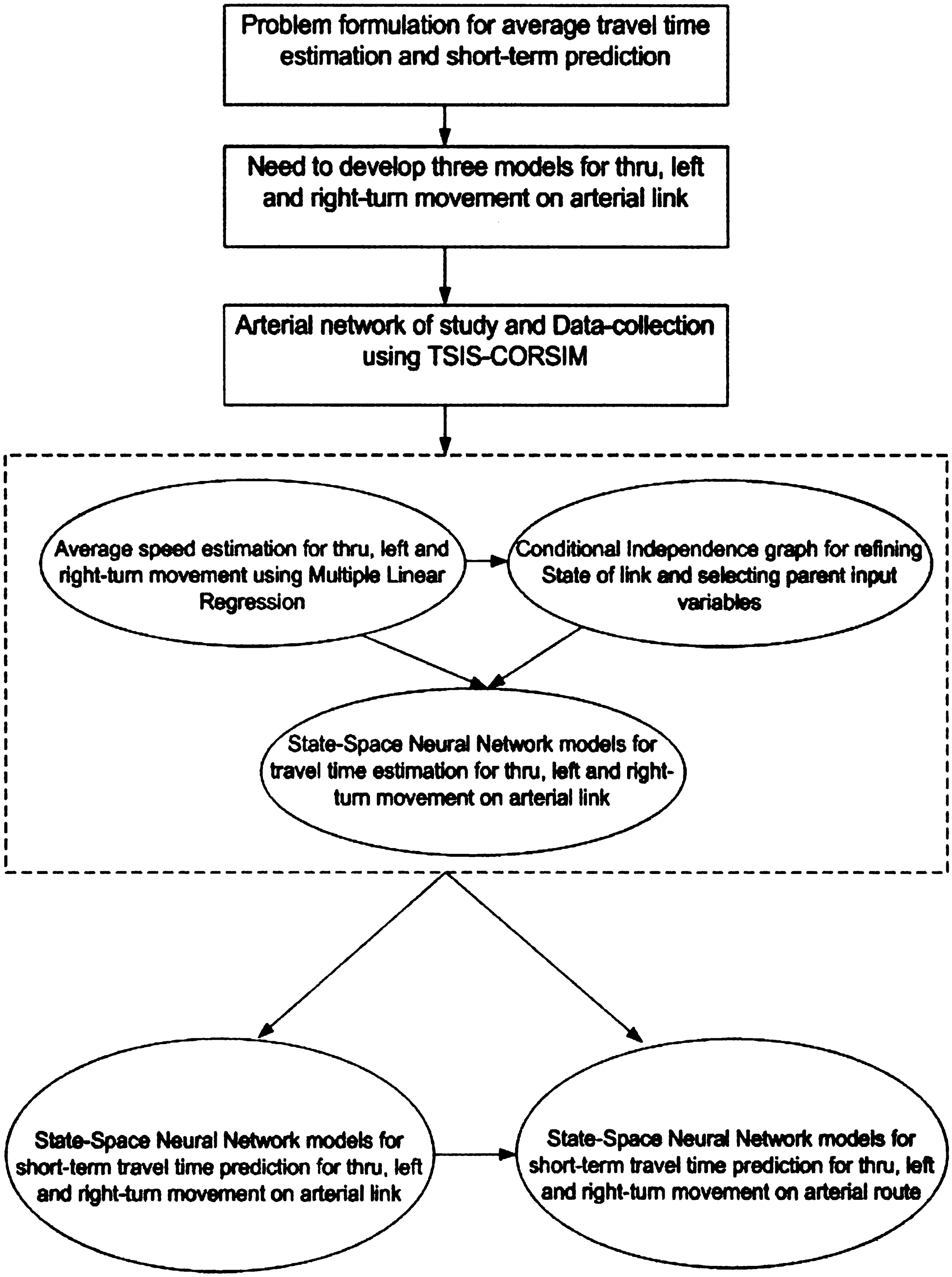

Figure 2.2 – Flowchart of Modeling Approach Adopted for Travel Time Estimation and Short-Term Prediction

Chapter 3

Experimental Set-up

The experimental set-up in this study consists of an urban arterial network for which travel time estimation and prediction models are to be developed. This study uses microscopic traffic simulation software to create a hypothetical arterial network. This arterial network is designed such that it represents a 'real arterial network' as closely as possible. The experimental set-up considers different traffic variables that affect travel time. It is tried through the experimental set-up to simulate a real arterial network by varying geometrics of the arterials as well as traffic demand.

The chapter introduces TSIS-CORSIM traffic simulation software and its suitability for this study. The arterial network as designed and data-collection for travel time study are illustrated. Also, some issues which are related to the use of CORSIM in this study in particular are discussed.

3.1 Overview of TSIS-CORSIM and its Suitability for this Study

An urban arterial network is designed in TSIS-CORSIM, Version 5.1. The TSIS-CORSIM consists of a Graphical-User-Interface (GUI) based traffic network and simulation input editor known as TRAFED. CORSIM is the microscopic traffic simulation module of the TSIS-CORSIM. The microscopic simulation can be viewed in the animation and graphics module of TSIS-CORSIM known as TRAFVU.

The travel time estimation and prediction models for urban arterial networks are still in the developing stage. The travel time estimation and prediction requires analyzing the interaction between a set of traffic variables which are supposed to affect the travel time. So, a large number of experiments should be performed with changing the traffic parameters like geometrics of the network, signal control parameters, and traffic demand, which is difficult to perform in the field and quite expensive. All these constraints require use of traffic simulation software to perform the required experiments. Due to the above mentioned requirements and limitations, TSIS-CORSIM is used for this study.

CORSIM is microscopic traffic simulation software that tracks individual vehicle on a network based on car-following logic and gap-acceptance lane-changing model. Its microscopic approach helps accurately modeling the traffic movement and stops which result due to the control scheme at signalized intersections and varying traffic demand. Almost all possible geometry conditions on the field can be simulated in CORSIM and effect of incidents and other interruptions to the traffic movement can be duly accounted for.

The primary features of CORSIM that made it suitable for the present study are listed as follows:

- CORSIM has an ability to model the time-varying traffic flow throughout a sequence of 'time-periods' as specified by the user. Inside each 'time-period' value, smaller slices of time termed as 'time-intervals' are specified. CORSIM calculates the different traffic parameters like traffic flow, average speed etc in every time-interval. Mostly, the value of time-interval is fixed as one minute. The traffic variables which are calculated at each of the time-interval is aggregated or averaged for each time-period. The signal control parameters can also be varied at each time-period in CORSIM, if some study needs to analyze the effect of variation of signal control scheme for varying traffic demand.

- CORSIM has a unique advantage to model the over-saturated traffic conditions as compared to other empirical and analytical methods (Owen, et al. 2000). The car-following logic inherent in CORSIM simulates the stop and go nature of traffic during heavy traffic demand as well as the queuing delay caused by signal interruptions. The over-saturated conditions resulting in queue spill-over when the queue at a downstream intersection interferes with the traffic movement on the upstream link is also accounted in CORSIM. The present study requires traffic demand to vary from uncongested conditions (low-volume) to the congested conditions (high-volume) and again come back to the uncongested conditions. This variation in traffic flow can be seen throughout the day on an urban arterial. CORSIM captures this variation in traffic flow and provides the traffic parameters values for each time period.
- CORSIM produces an output file having detailed Measures of Effectiveness (MOE's) for a study. This study requires finding out traffic variables which affect the travel time. CORSIM provides extensive report of MOE's which helps to analyze the travel time process.

3.2 Urban Arterial Network Designed in TSIS-CORSIM

An urban arterial network is deigned in TSIS-CORSIM in this study. The network is shown in Figure 5.1 that consists of 25 signalized intersections. These signalized intersections are numbered from 1 to 25 and are shown as smooth-lined circles. The other intersections which are numbered from 26-45 and shown as dashed circles are the entry and exit nodes for this arterial network. There are five arterials running in east-westbound direction. These are numbered as 1-5, 6-10, 11-15, 16-20, and 21-25. These arterials are major arterials as they are designed to serve higher traffic demand than the other cross-arterials. The cross-arterials running along north-southbound direction are numbered as 1-21, 2-22, 3-23, 4-24, and 5-25. These arterials are minor arterials as are designed to serve lesser traffic demand as compared to major arterials. Each arterial consists of four arterial links bounded by signalized intersections on either side. For example, 1-6, 6-11, 11-16, and 16-21 are four links in arterial 1-21. It is to be noted that the notion of major and minor arterials as taken in this study depends on the network in question. The directions of major and minor arterials may change as the network varies. But, it is not critical in this study as the major and minor arterials are considered equally in the modeling approach. Hence, here, it is just a matter of naming the arterials as major and minor based on traffic demand served.

The geometry of links comprising each arterial is varied to account for the variations possible in any real arterial network. Moreover, this will lead to study the variation of travel time with variables that represent geometry of links. The link length varies from 1500 ft to 6000 ft as shown in the Figure 5.1. All links have two lanes in each direction and each intersection approach has a left-turn pocket or bay to accommodate the left-turning traffic. The speed limit posted on arterial links is also varied and is kept as 30 miles per hour for small links (link length of 1500 ft) and 45 miles per hour for the remaining links.

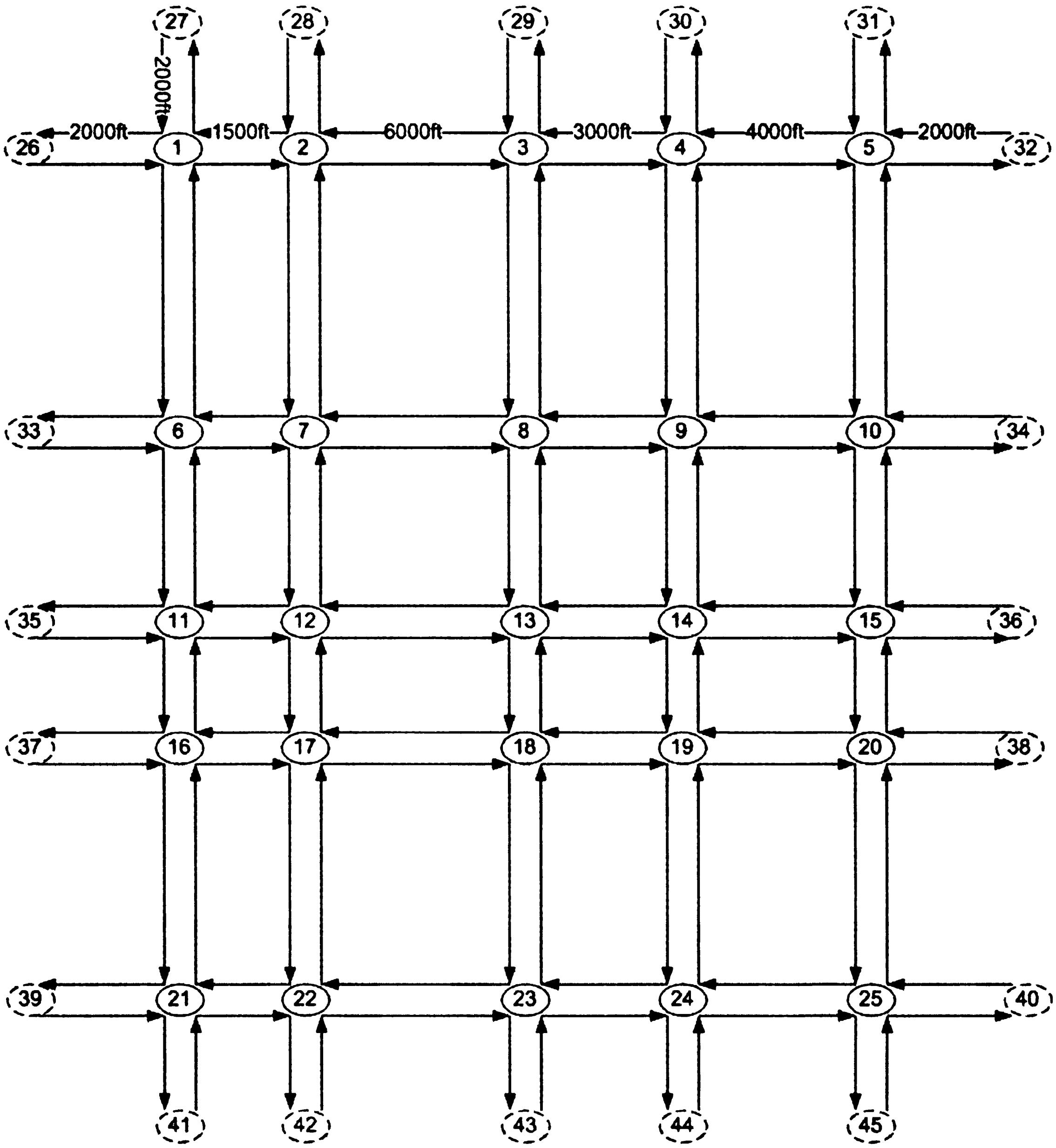

Figure 3.1 - Urban Arterial Network Designed in TSIS-CORSIM for Travel Time Estimation and Prediction.

The provision of time-interval and time-period in CORSIM is used in this study to simulate varying traffic conditions that can be as close to reality as possible. There are total 19 time-periods each of 5 minutes duration. Hence, each simulation run was made for 95 minutes or 5700 seconds (19*5 minutes). As said earlier, CORSIM calculates values of traffic variables at each time-interval which is specified as 1 minute here. So, the values obtained every one minute are aggregated or averaged for a time-period.

The variation of traffic demand in the network is achieved by varying the traffic flow in every time-period at the entry links of the network. The variation of traffic flow in general begins from an uncongested state to a congested state and then gradually back to the uncongested conditions. The traffic flow variation for each arterial is different but is kept the same for both directions in an arterial. For example, the traffic flow variation is different for arterials 6-10 and 16-20 but the same variation is used for eastbound and westbound directions of both arterials. The traffic flow variation in the minor arterials is different from that on the major arterials but is kept the same for all the minor arterials. Figure 5.2 shows the traffic flow variation at five major arterials (arterials 1-5, 6-10, 11-15, 16-20, and 21-25) and one representative minor arterial (arterial 3-23). As can be seen the average flow rate is increasing from as low as 250 vehicles per hour to a maximum of 2300 vehicles per hour. Figure 5.2 clearly shows that enough variability in traffic flow is accounted for which makes travel time estimation and prediction on this network a challenging problem similar to that in field.

This study explicitly considers turning movements at signalized intersections. Each intersection is having average flow rate for through, left and right-turning movement. At each intersection, left-turning traffic is kept as 20% and right-turning as 10% of total traffic flow which is input at the entry node of each arterial.

The signal control in the network is optimized using Synchro which is a macroscopic traffic software. The signal optimization and control phasing scheme are covered in the next section.

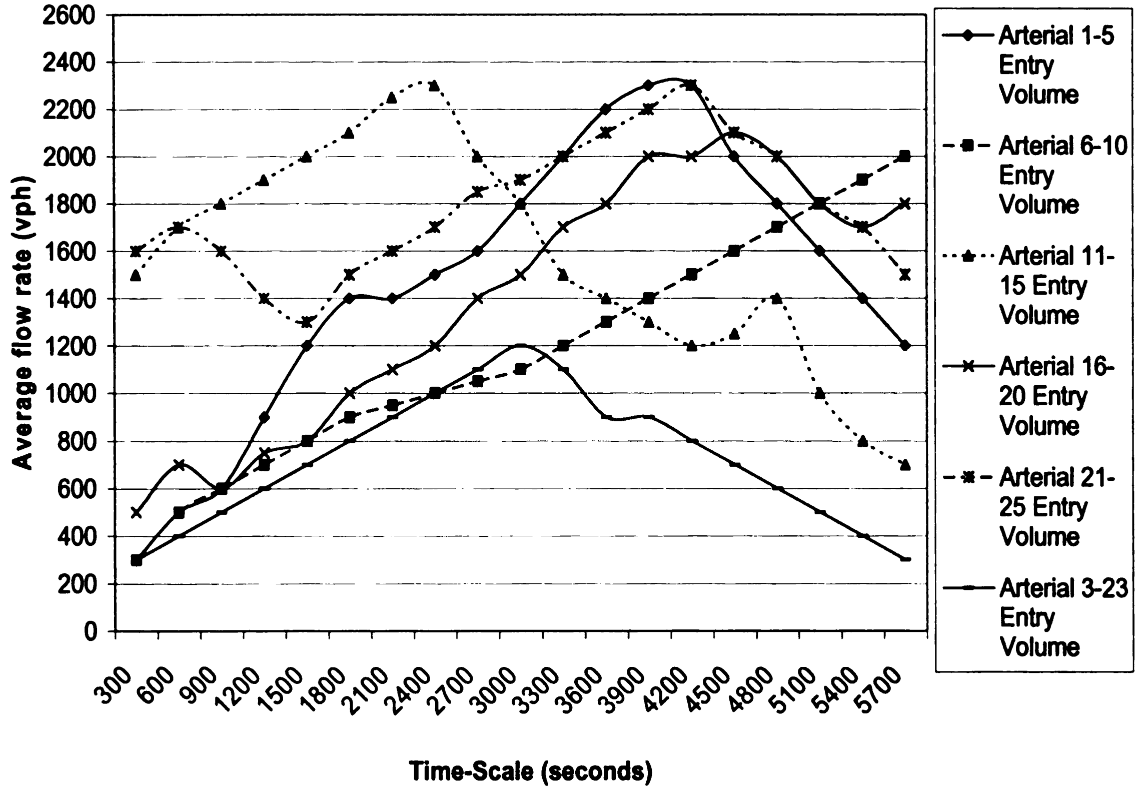

Figure 3.2 – Average Flow Rate Variations at Entry Nodes along Arterials in the Network.

3.3 Signal Control Optimization using Synchro

Synchro is a software application for optimizing traffic signal timing and performing capacity analysis. The software optimizes splits, offsets, and cycle lengths for individual intersections, arterial, or a complete network (http://www.trafficware.com/synchro.htm). Its wide popularity among researchers and practitioners and its efficiency makes it suitable for signal control optimization for this study as well. The optimized signal control scheme is obtained from Synchro for the urban arterial network shown in Figure 5.1. This optimized signal control scheme is then exported to TSIS-CORSIM. The cycle length, green splits, and offsets are optimized at intersection and network level for the peak-hour traffic flow in the network. The pre-timed signals are incorporated in this study which is installed in most of the urban networks. The phasing scheme given by Synchro for a typical intersection (Intersection 9) in the network is shown in figure 5.3. This phasing scheme is an exclusive left-turn phase with leading green which has four phases. Since, the left-turn bays or pockets are provided in the geometrics of network, so, this type of phasing is logical to follow for all the intersections.

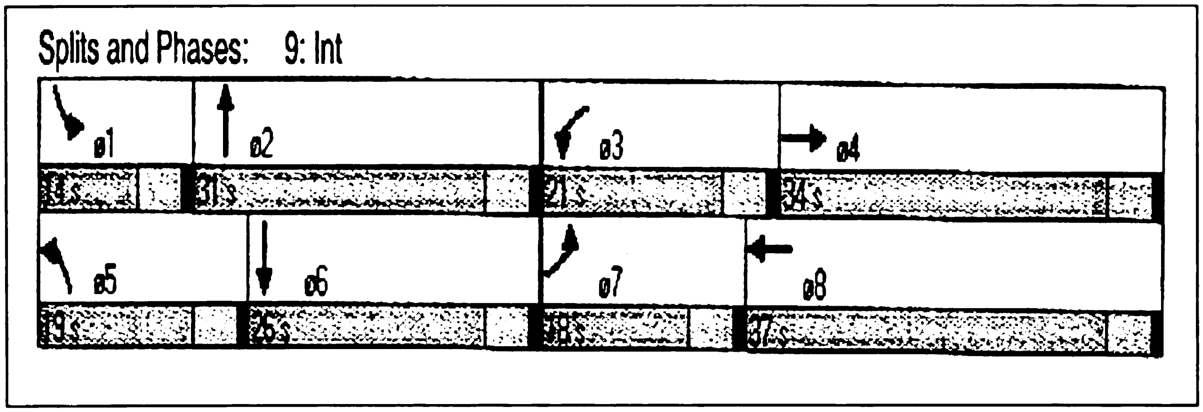

Figure 3.3 – Exclusive Left-Turn Phase with Leading Green Phasing Scheme for a Typical Intersection in the Network of Study.

The signal optimization in Synchro allows green splits, offsets to be taken as explanatory variables affecting travel time in this study. These variables are easily available to a traffic agency responsible for a traffic network and are very easy to use these in the model here.

3.4 Definition of Traffic Variables used in the Study based on TSIS-CORSIM

The TSIS-CORSIM tracks each individual vehicle and aggregates the values of traffic variables for each time-period. So, the values of traffic variables are space-mean values or true values that are representative of traffic conditions on an entire link. The definition of traffic variables involved in the present study as given by CORSIM is stated here (User's Guide, Traffic Software Integrated System (TSIS), Version 5.1):

- *Vehicle Trips* - The number of vehicles that have been discharged from the link since the beginning of the simulation.
- *Average Volume (vehicles per hour)* - Total vehicle trips divided by the simulation time.
- *Vehicle Miles* - The vehicle trips times the length of link.
- *Total Time (vehicle-minute)* - Total time on the link for all vehicles.
- *Average Speed (miles per hour)* - Total vehicle miles divided by the total travel time.
- *Maximum Queue Length by lane (vehicle)* - The maximum queue length that was observed on a lane since the beginning of the simulation (i.e., 100 percentile).
- *Total Time per vehicle (seconds per vehicle)* - The average travel time on a link for each vehicle, calculated by taking the total travel time and dividing it by the number of vehicle trips.

It is to be noted that some variables are given by CORSIM for specific time-periods and others as cumulative for the entire simulation time. All the variables used are finally obtained for specific time-periods by using the above definitions.

3.5 Data Collection from TSIS-CORSIM

The data set having average travel time (seconds per vehicle) and the traffic variables affecting average travel time are obtained from CORSIM. The traffic variables that affect travel time are already mentioned in Chapter 2 (see Section 2.4).

A total of two CORSIM runs were done for data collection with different random seeds. The random seeds are varied in CORSIM runs for headway seed, vehicle seed, and traffic seed. The difference in random seeds provides data-sets which represent different traffic demand patterns for the same network. The data set obtained from the first CORSIM run is termed as 'training set' and the second run data set as 'testing set'. In all modeling steps (i.e. multiple linear regression of average speed, Conditional Independence graphs analysis, and State-Space Neural Networks modeling) presented in Chapters 4-6, models are developed from the training set only, and are tested using the testing set to evaluate the efficiency of the models when new traffic patterns are presented (i.e. model's ability to generalize).

The modeling steps are limited to four arterials selected from the arterial network shown in Figure 5.1. The selected arterials are 6-10, 16-20, 21-25, and 3-23. The entire modeling deals with traffic conditions and travel times on the links of these arterials only. This is done just to reduce the amount of time and effort needed to account for the entire network. The consideration of these four arterials is sufficient to evaluate the efficiency of the models developed as the variability needed in geometrics, traffic demand, and signal control is met with these four arterials.

As stated earlier (Section 3.2), each CORSIM simulation run was made for 95 minutes (5700 seconds), where data is averaged every 5 minutes. So, 19 data-points are available for each arterial link in one run. Here, four arterials mentioned in the above paragraph are selected where each arterial consists of four links. So, a total of 304 data-points are available for training and the same number for testing sets (19 data-points per link * 4 links per arterial * 4 arterials). All 304 data-points are used as the data-set in the modeling steps except for the State-Space Neural Network model that was developed for travel time prediction. The reason is that for first 5 minutes of the simulation on each link, the travel time can not be predicted as there are no values of traffic variables for the past 5 minutes time period earlier than the current time period (because it is the first time-period of simulation). So, 18 data-points are used for each link for travel time prediction, which gives a total 288 data-points for training and a similar number for testing sets (18 data-points per link * 4 links per arterial * 4 arterials).

3.6 Miscellaneous Issues in Data Collection using TSIS-CORSIM

Some issues are identified and discussed here which are encountered during the data collection for the study. These are as follows:

- *Calibration and validation of CORSIM* – An important aspect of all computer simulation programs including CORSIM are calibration and validation whenever the model is applied to field or a new site. This fact has been acknowledged in the discussion of application of simulation programs (May, 1990). Sisiopiku et al. (1994) present calibration of NETSIM (component of CORSIM for surface streets) with field studies. They have found that observed values in the field for variables like occupancy, capacity, flows, and delays are different from those obtained from simulation. But, the differences between simulated and observed values are due primarily to insufficient model calibration rather than actual deficiency in the simulation code to properly model traffic processes. Thus, this study also believes that traffic patterns simulated by CORSIM are approximate depiction of the field situations. It is recommended to calibrate the values of traffic variables especially average travel time for field studies wherever the proposed model is applied.

- *Randomness of traffic patterns and number of random seeds* – This study develops travel time model using one CORSIM run data and tests it on dataset obtained from a second run with different random seeds. Usually, the studies using CORSIM simulation average data from about 2-10 random seed runs. The model in this study does not need to estimate a value of travel time per se which is subjected to specific traffic conditions. The idea is to obtain a dynamic model which can learn the functional relationships between travel time and the varying traffic conditions. A single run in CORSIM is sufficient to fulfill this requirement. Even in a single run, 304 data-points are obtained where each data-point represents the travel time as an outcome of other traffic variables. A total of 16 arterial links are used in training and testing sets to model travel time for any arterial link in question. The testing set considered is a different traffic demand pattern where the model is tested. Thus, sufficient variability has been taken into account in this study to develop travel time estimation and prediction model.
- *Maximum Queue Length in CORSIM* – According to the definition of maximum queue length noted earlier (Section 3.4); CORSIM does not give maximum queue length for each time-period. It identifies the maximum value of queue length observed from the beginning of simulation even for an intermediate time-period. This maximum queue length is not a good estimate of queue length in the situation when simulation observes high traffic demand in the starting time-

periods resulting into a maximum queue length. For later time-periods, even when traffic demand lowers, the maximum queue length stays the same. This does not affect the model much as the inclusion of maximum queue length is done with other traffic variables which act as inputs to the model. Hence, it is required that during application of this model in the field, queue length should be duly considered and accurate value of queue length is obtained. A more accurate value of queue length will increase the efficiency and accuracy of the proposed travel time model in this study.

Concluding this chapter, it can be said that this experimental set-up presents all the complexity and variation which are needed to account for in any travel time study. The following chapters deal with the modeling approach using data collected from this experimental set-up.

Chapter 4

Multiple Linear Regression for Average Speed Estimation

The multiple linear regression is done here to obtain estimated average speed on arterials. The difficulty of obtaining average speed from loop detectors is duly considered in this approach. The proposed model provides estimated speed which is obtained by fitting a multiple linear regression with other traffic variables which are easy to obtain in the field. The true average speed that was obtained from CORSIM simulation is replaced by the estimated average speed obtained here. Using this known relationship in further modeling steps, the travel time can be estimated and predicted by using the estimated average speed instead of the average speed obtained from the field.

4.1 Utility of Average Speed Estimation for Travel Time Modeling

Average speed is an important traffic variable affecting travel time on an arterial link. The travel time prediction models developed for freeways consider speed as one of the input traffic parameters (for example, see vanLint 2004 and Mark and Sadek 2004). Several travel time estimation models developed for urban arterials as reviewed by Sisiopiku and Rouphail (1994) use speed as an input traffic parameter.

The average speed obtained from TSIS-CORSIM in this study can be considered as the true average speed, as it is calculated by tracking individual vehicles in the simulation. The average speed is a traffic parameter which is difficult to observe or estimate in the field. Some of the loop detectors calculate speed on a link. The recommended location of loop detectors for speed estimation in the field for arterial networks is specified in Traffic Detector Handbook (1990). But, this speed is limited to the location of the detector itself. For example, if the detector for estimating speed is located close to the stop line of the signalized intersection, the estimated speed is of lower value. This is because the vehicles slow down and even completely stop when they approach a signalized intersection during red interval. On the other hand, if a detector is installed midway on a link, the estimated speed will be a higher value because of the low interference to the vehicle from the signal control and neighboring vehicles in the traffic stream. The average speed will be a true average speed, if and only if each individual vehicle is tracked and their acceleration and deceleration profiles are obtained. This task is fairly complicated and expensive in the field. Moreover, it is recognized that average speed is different for different types of movement on an arterial link. That is to say that average speed for through, left and right-turning vehicles is different. Recognizing the complexity in finding average speed in the field for the entire traffic stream neglecting their turning movement, it is more complex to find average speed for each traffic movement.

This study aims to provide a modeling approach for travel time estimation and prediction which relies only on easily available data in the field. So, it is required that average speed is calculated from other traffic variables that are easy to obtain from existing traffic detection technologies. The traffic variables other than average speed (like V_T, V_L, V_R, G, Off, Q_T, Q_L, S.L., ITT_T, and ITT_L) affecting travel time can be easily obtained in the field. So, an intermediate modeling step is required to estimate average speed from other known and easily obtainable traffic variables. This estimated average is used in the next modeling steps to model travel time.

4.2 Multiple Linear Regression for Average Speed Estimation

The traffic variables that are easy to obtain in field are V_T, V_L, V_R, *G*, *Off*, Q_T, Q_L, *S.L.*, ITT_T, *and* ITT_L. The function between average speed for each type of movement (i.e. through, left and right-turn) and these traffic variables is to be obtained. The technique used to model or approximate this function is multiple linear regression. The average speed for through, left and right-turn movements (S_T, S_L, and S_R respectively) are termed as response variables in this case, and the known traffic variables like V_T, V_L, V_R, *G*, *Off*, Q_T, Q_L, *S.L.*, ITT_T, and ITT_L are termed as predictor variables. Due to the presence of more than one predictor variable (here 10 predictor variables), multiple linear regression is a promising choice.

Consider a response variable Y and its value Y_i in the i^{th} trial. A set of predictor variables like X_1 and X_2 is identified and X_{i1} and X_{i2} are their respective values in i^{th} trial. The regression model which gives a relationship between Y and X_1 and X_2 is an example of a multiple linear regression model. Consider the model given below-

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \delta_i$$

It is called a first-order multiple linear regression model. The parameters of this model are $\beta_0, \beta_1, \beta_2,$ and the error term is δ_i. The parameters β_1, β_2 are sometimes called as partial regression coefficients because they reflect the partial effect of one predictor variable when the other predictor variable is included in the model and is held constant (Kutner, et al. 2004).

Assuming that the expected value of the error, i.e., $E\{\delta_i\} = 0$, the regression function is

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

The similar equation can be extended to more than two predictor variables and to 10 variables in the speed estimation here. The multiple linear regression models for this study which need to be determined are-

$$S_T = \beta_{0T} + \beta_{1T} V_T + \beta_{2T} V_L + \beta_{3T} V_R + \beta_{4T} G + \beta_{5T} Off + \beta_{6T} Q_T + \beta_{7T} Q_L + \beta_{8T} S.L. + \beta_{9T}\, ITT_T + \beta_{10T}\, ITT_L$$

$$S_L = \beta_{0T} + \beta_{1T} V_T + \beta_{2T} V_L + \beta_{3T} V_R + \beta_{4T} G + \beta_{5T} Off + \beta_{6T} Q_T + \beta_{7T} Q_L + \beta_{8T} S.L. + \beta_{9T}\, ITT_T + \beta_{10T}\, ITT_L$$

$$S_R = \beta_{0T} + \beta_{1T} V_T + \beta_{2T} V_L + \beta_{3T} V_R + \beta_{4T} G + \beta_{5T} Off + \beta_{6T} Q_T + \beta_{7T} Q_L + \beta_{8T} S.L. + \beta_{9T}\, ITT_T + \beta_{10T}\, ITT_L$$

The β_{ij} with appropriate notation of i and j are the parameters of these models which are to be estimated. The value of β_{ij} is found by using the method of least squares.

Let the least square criterion be L for the regression model between Y and X_1 and X_2, as stated earlier. The value of L can be written as –

$$L = \sum_{i=1}^{n}(Yi = \beta 0 + \beta 1\, Xi1 + \beta 2 Xi2)^2$$

According to the method of least squares, the estimators of β_0, β_1, and β_2 are those values that minimize the criterion L for the given sample observations (Kutner, et al. 2004).

The multiple linear regression for average speed estimation was done using SPSS 14.0 for Windows. The stepwise regression method was used which is also called as statistical regression. Stepwise regression is a way of computing ordinary least squares (OLS) in stages. In stage one, the predictor (independent) variable best correlated with the response (dependent) variable is included in the equation. In the second stage, the remaining independent variable with the highest partial correlation coefficient with the dependent variable, controlling for the first independent variable, is entered. This process is repeated, until the addition of a remaining independent variable does not increase square of correlation coefficient (R^2) by a significant amount (or until all independent variables are entered).

The data-set that is used to obtain a multiple linear regression model for average speed is the training set as discussed earlier in Chapter 3. Recalling that training set consists of data collected by first run of CORSIM simulation. The training set consists of 304 data points. The obtained regression model is evaluated by applying it on testing set. The testing set was obtained from second run of CORSIM simulation.

The training and testing sets are transformed by a linear function to scale in a range of 0.1 and 0.9. This transformation is also done with the dataset when it is used for the Conditional Independence graph and State-Space Neural Networks as discussed in later chapters. The scaling of data is required for fast and stable learning in neural networks (Haykin 1999). The scaling of data is done for average speed estimation here just for the sake of uniformity in data set so that obtained speed is in the required range and can be directly fed to next modeling steps.

There are various tests for analysis of appropriateness of regression model and analysis of variance as explained in Kutner, et al. (2004). These tests are provided by SPSS during report generation for a model. The following statistical tests and results can be used to evaluate the efficiency of multiple linear regression models -

- descriptive statistics including mean and standard deviation of set of variables
- model summary
- ANOVA
- Coefficients
- Residual statistics

In the present travel time modeling approach, the travel time estimation and prediction is done on the basis of the training set and is applied on the testing set. Thus, same approach is used for evaluating the efficiency of these regression models also. The reliability of regression model in this study is tested by doing cross-validation i.e. by applying it on a testing set. The performance of regression model on testing set states the efficiency of the model. The statistical tests mentioned above are usually required for multiple linear regression model evaluation. However, the use of testing set for evaluating its performance is assumed enough in the scope of present study.

As noted earlier, three regression models for average speed of through, left and right-turn movement on arterial links were obtained. The speed estimation models were developed for the major and minor arterials separately in the arterial network. The major and minor arterials in the studied arterial network are stated earlier in Chapter 3 (Section 3.2).

4.2.1 Average Speed Estimation for Major Arterials

The average speed estimation for through, left, and right-turn movements on arterial links is discussed here. First, the average speed estimation for through movement vehicles on major arterials is presented. It is noted again that the average speed is scaled by a linear function to lie in a range of 0.1 and 0.9. The training set is used for the regression model development. The following regression models are obtained by using stepwise regression method for average speed estimation on major arterials-

$$S_T = 0.444 - 0.462\, Q_T + 0.317\, S.L. + 0.593\, ITT_T - 0.398\, ITT_L$$

$$S_L = 0.165 - 0.073\, V_T - 0.159\, V_L + 0.261\, G - 0.07\, Off - 0.07\, Q_L + 0.157\, S.L. + 0.403\, ITT_L$$

$$S_R = 0.458 - 0.086\, G - 0.396\, Q_T + 0.336\, S.L. + 0.649\, ITT_T - 0.602\, ITT_L$$

The efficiency of the estimation models is first evaluated for the training set. This is done by plotting best-fit line between scaled estimated and actual average speed. Figures 4.1, 4.2, and 4.3 show the plots for through, left and right –turn movement on major arterials. The R^2 values are as shown in these figures. The R^2 values obtained for the training set are reasonably well and these models are selected for the purpose here. It is to be remembered that average speed estimation is an intermediate step in modeling travel time. There are other variables also which act as input to travel time model. Hence, a reasonable estimate of average speed was required and this purpose is fulfilled here.

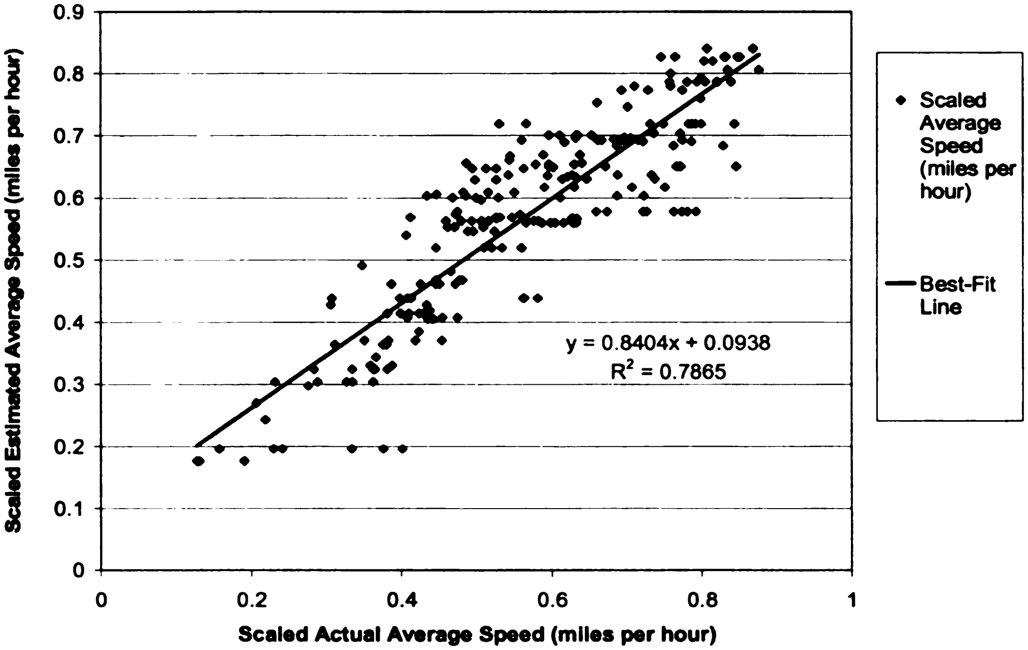

Figure 4.1 – Training Plot between Scaled Estimated and Actual Average Speed for Through Movement on Major Arterials.

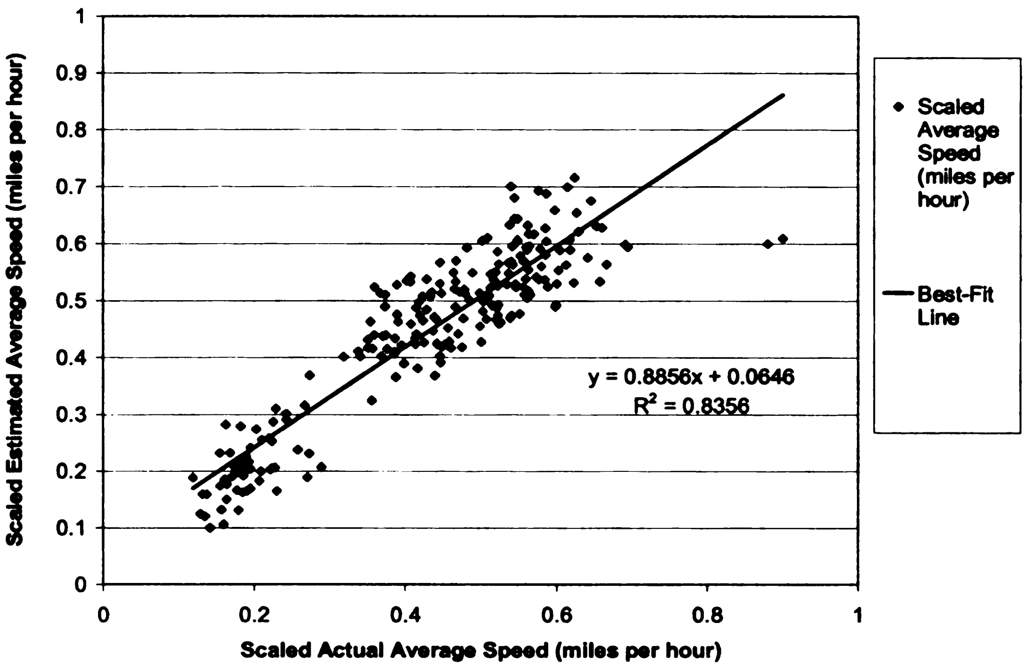

Figure 4.2 – Training Plot between Scaled Estimated and Actual Average Speed for Left-Turn Movement on Major Arterials.

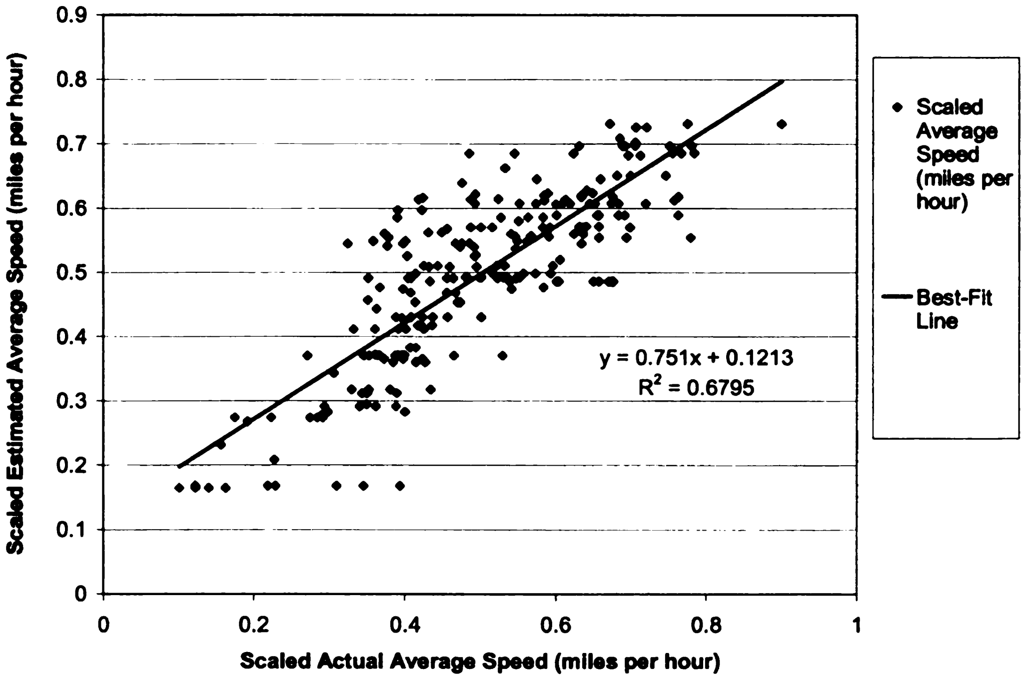

Figure 4.3 – Training Plot between Scaled Estimated and Actual Average Speed for Right-Turn Movement on Major Arterials.

The performance of the average speed estimation models needs to be tested on the testing set. Figures 4.4, 4.5 and 4.6 show the plots between scaled estimated and scaled actual average speed for through, left and right-turn movements on major arterials. It is seen that R^2 values for the testing set are better as compared to that of training set. This observation is just a coincident and may be because the average speed values in the testing set is similar to those average speed values in the training set which were accurately captured by the model. The testing set was not used in the model development and represents different traffic patterns than the training set. This evidence gives support to use these models for average speed estimation. The proposed average speed estimation models are assumed as generalized models on the basis of their comparable and even better efficiency on testing set.

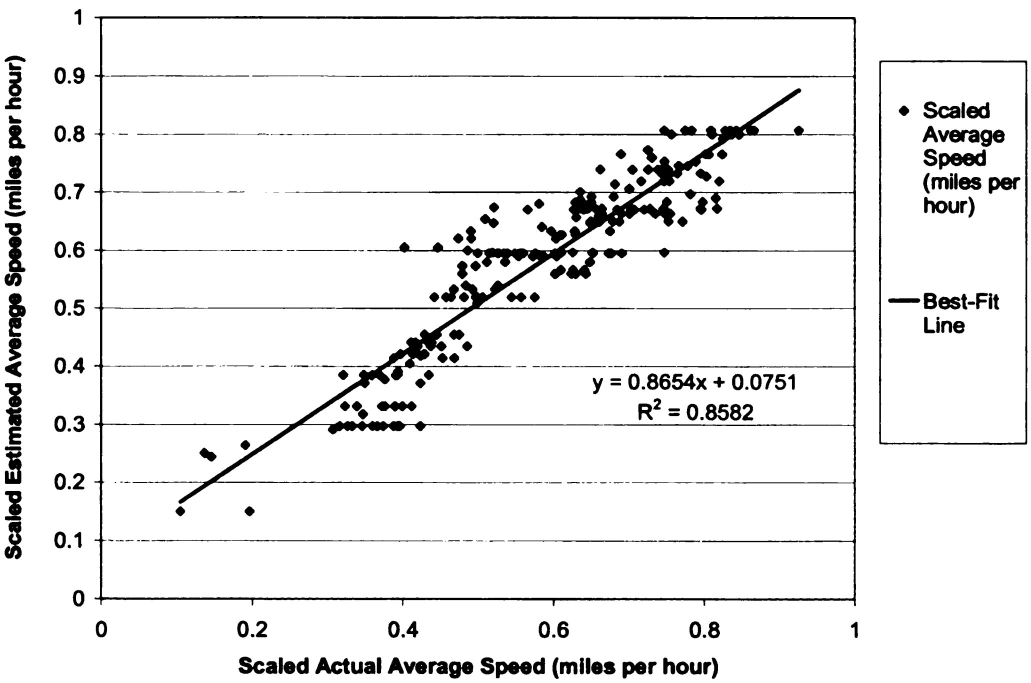

Figure 4.4 – Testing Plot between Scaled Estimated and Actual Average Speed for Through Movement on Major Arterials.

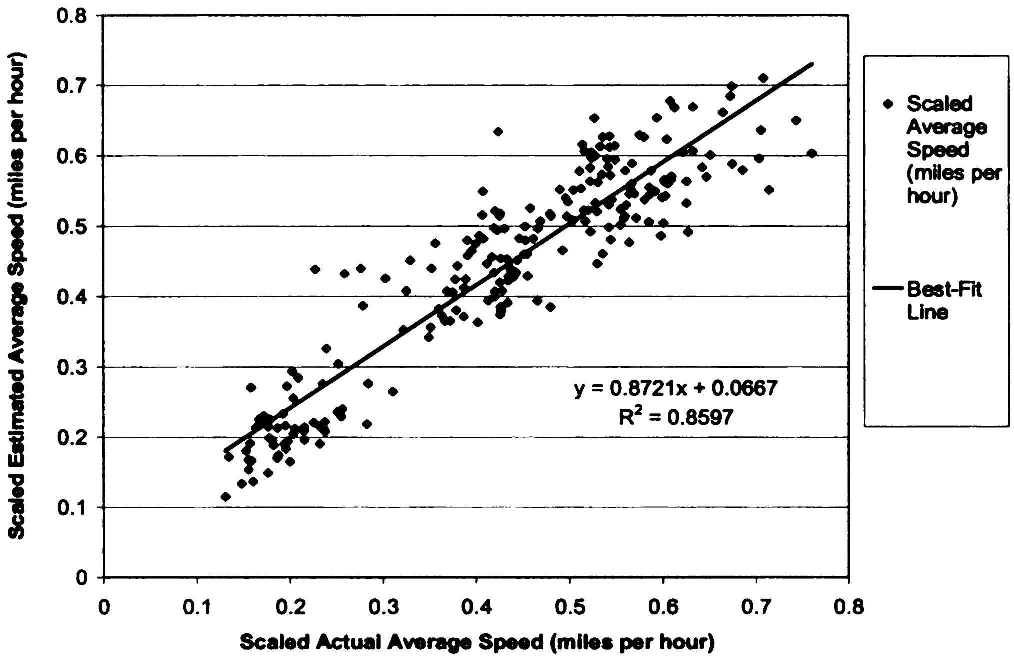

Figure 4.5 – Testing Plot between Scaled Estimated and Actual Average Speed for Left-Turn Movement on Major Arterials.

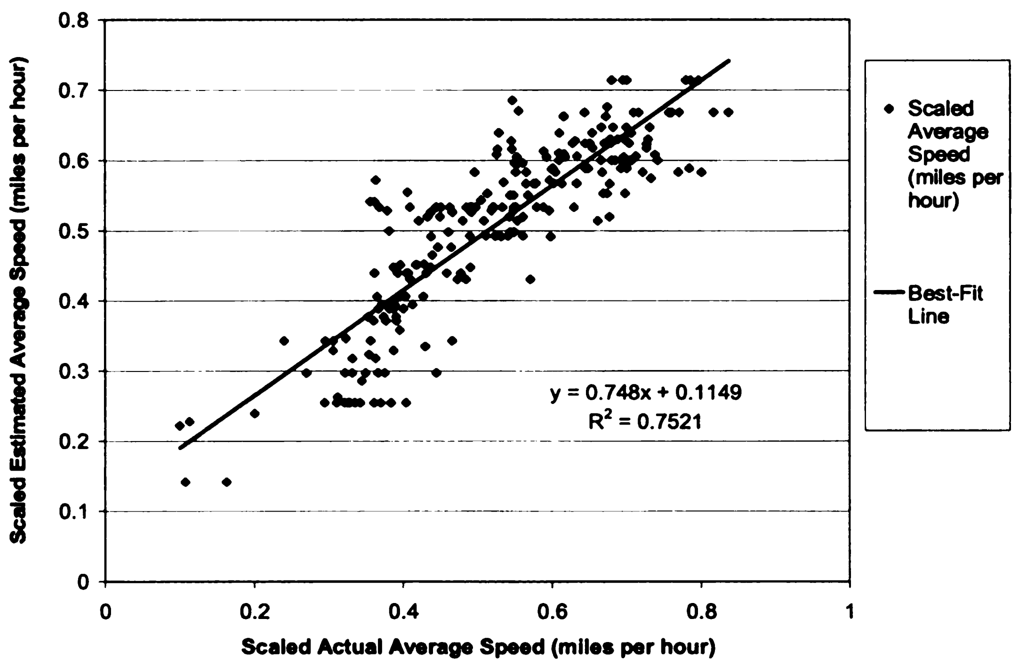

Figure 4.6 – Testing Plot between Scaled Estimated and Actual Average Speed for Right-Turn Movement on Major Arterials.

4.2.2 Average Speed Estimation for Minor Arterials

The average speed estimation for minor arterials is done separately from major arterials. The minor arterials serve lesser traffic demand as compared to the major arterials. The resulting average speed is considered to have different dependence on the input variables. Due to this reason, the minor arterials were considered separately for average speed estimation to improve the accuracy of the models.

The average speed estimation for through movement vehicles on minor arterial links is discussed here. It is noted again that the average speed is scaled by a linear function to lie in the range of 0.1 and 0.9. The training set is used for the regression model development. The following regression models were obtained by using the stepwise regression method for average speed estimation on minor arterials-

$$S_T = 0.564 - 0.286\,V_T - 0.115\,V_R + 0.126\,S.L. + 1.476\,ITT_T - 1.141\,ITT_L$$

$$S_L = 0.466 - 0.179\,V_L - 0.481\,Off + 0.513\,S.L. - 0.104\,ITT_L$$

$$S_R = 0.575 - 0.305\,V_T - 0.126\,V_R + 0.096\,S.L. + 1.04\,ITT_T - 0.829\,ITT_L$$

The efficiency of these estimation models is first evaluated for the training set. This is done by plotting the best-fit line between scaled estimated average speed and scaled actual average speed for each type of movement. Figures 4.7, 4.8, and 4.9 show these plots for through, left and right –turn movements on minor arterials. The R^2 values are as shown in the figures. On the basis of R^2 values, the proposed average speed estimation models are found satisfactory.

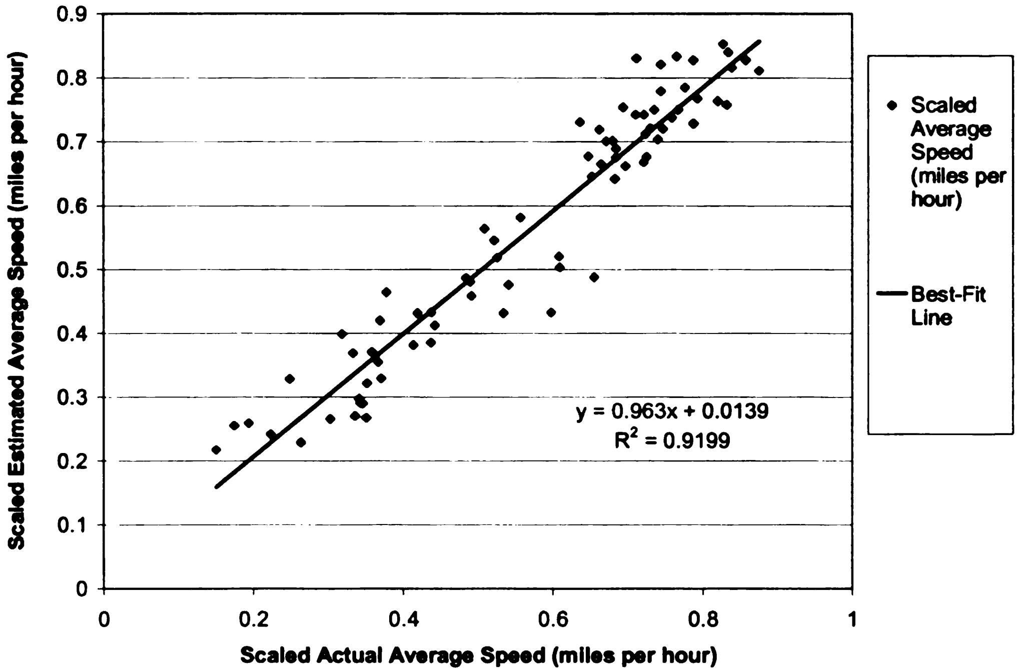

Figure 4.7 – Training Plot between Scaled Estimated and Actual Average Speed for Through Movement on Minor Arterials.

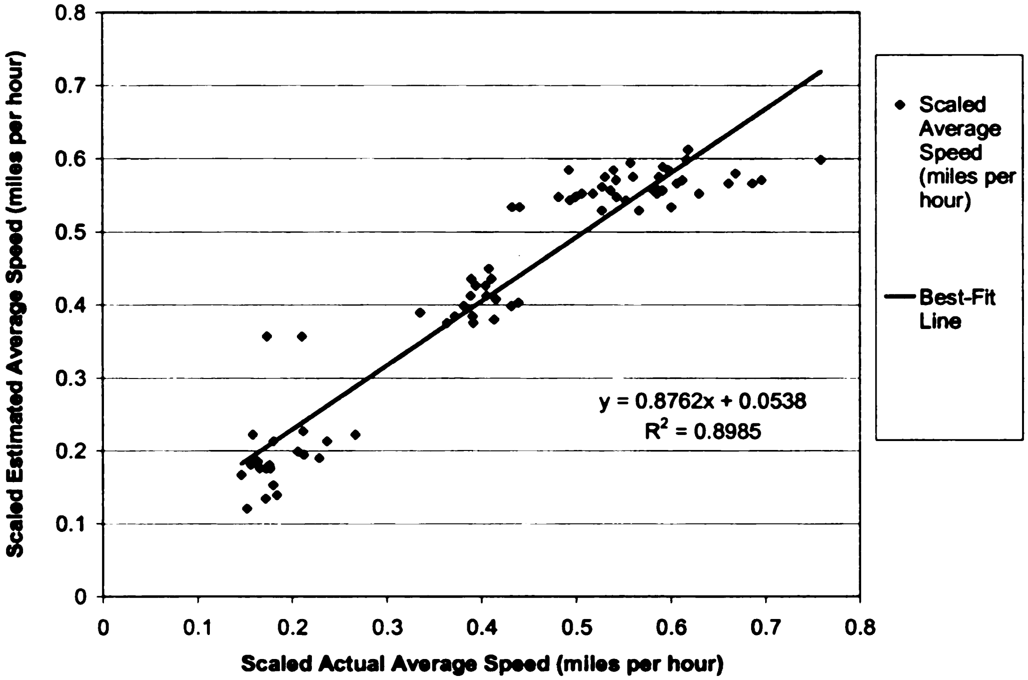

Figure 4.8 – Training Plot between Scaled Estimated and Actual Average Speed for Left-Turn Movement on Minor Arterials.

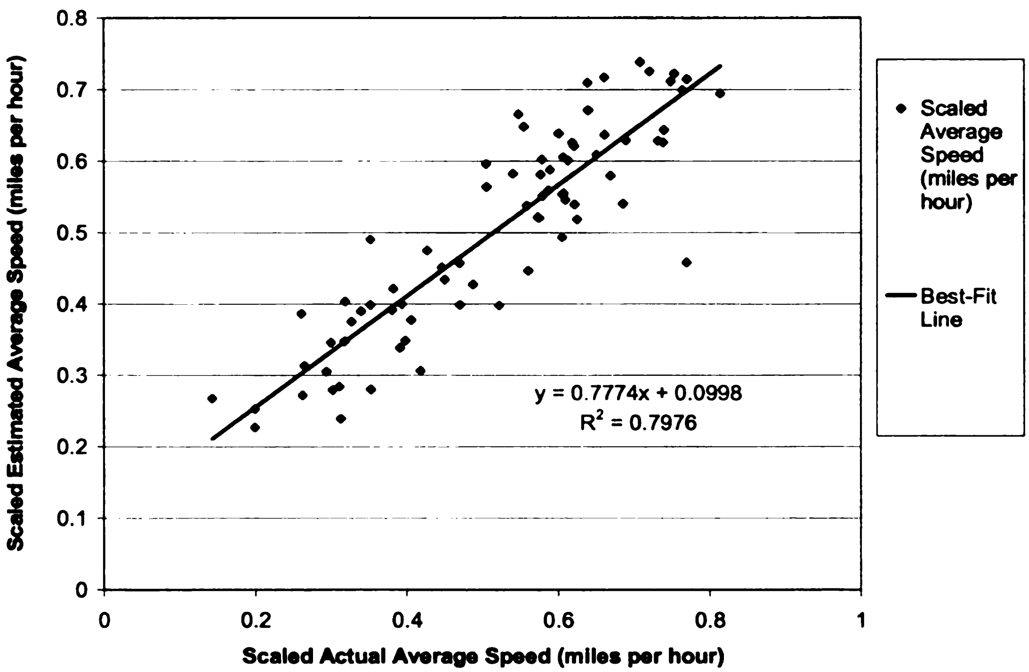

Figure 4.9 – Training Plot between Scaled Estimated and Actual Average Speed for Right-Turn Movement on Minor Arterials.

The performance of the average speed estimation models is further evaluated on the testing set. Figures 4.10, 4.11 and 4.12 show the plots between scaled estimated and scaled actual average speed for through, left, and right-turn movements on minor arterials. It is again observed that the R^2 values for testing set are better as compared to that of the training set. The argument given for this observation for major arterials holds true here also. The testing set was not used in the model development and represents different traffic patterns than the training set. This evidence gives support to use these models for average speed estimation. The proposed average speed estimation models are assumed as generalized models on the basis of their comparable efficiency on the testing set.

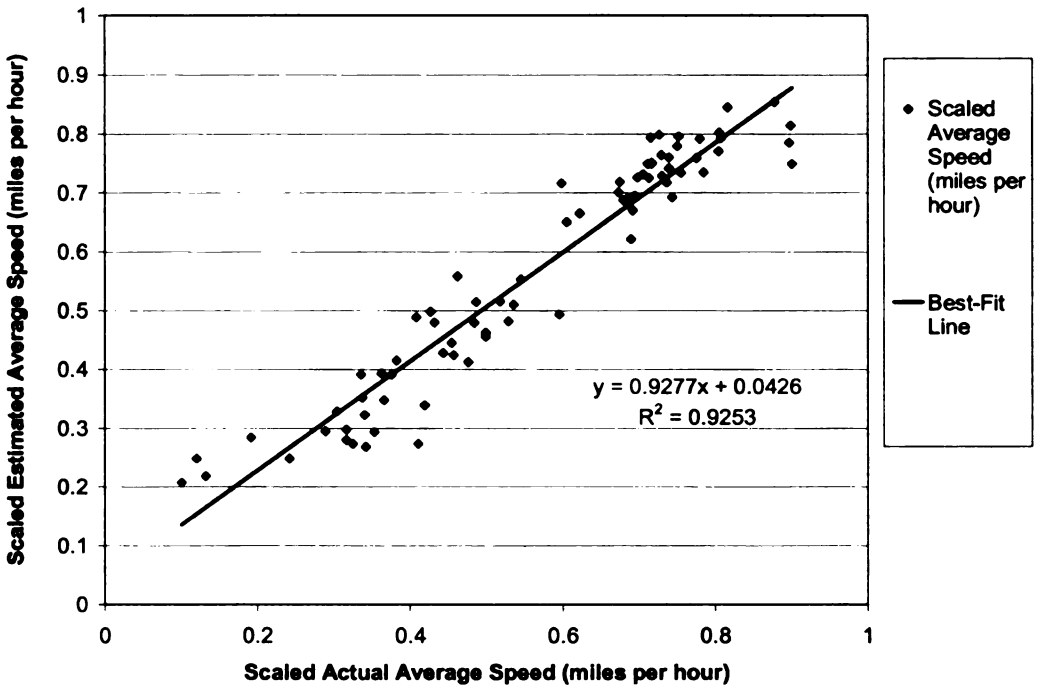

Figure 4.10 – Testing Plot between Scaled Estimated and Actual Average Speed for Through Movement on Minor Arterials.

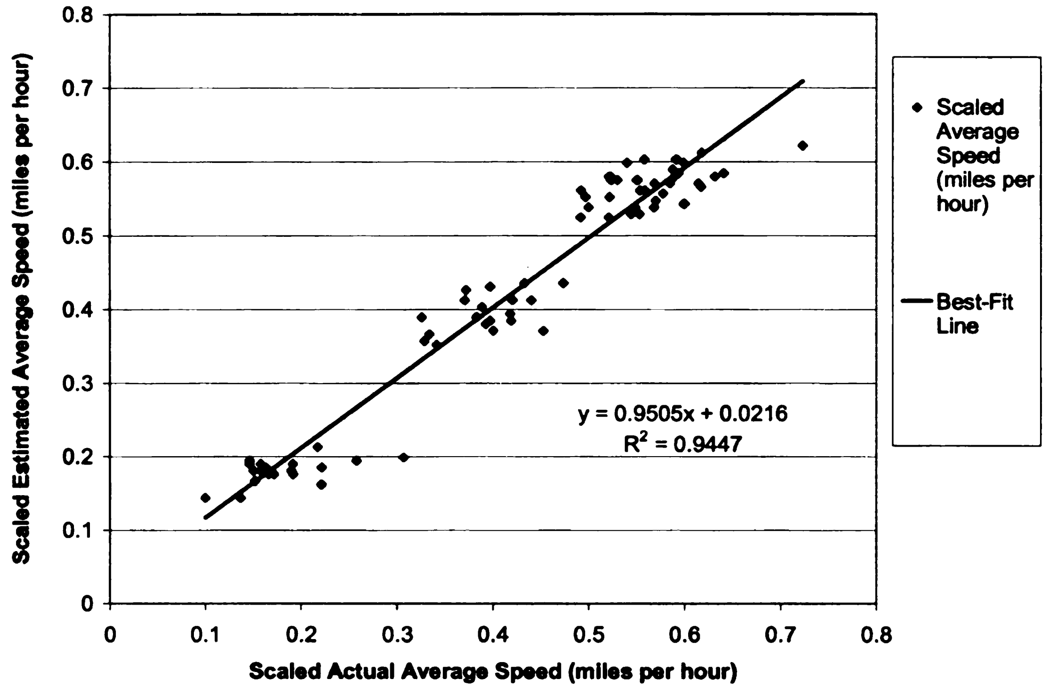

Figure 4.11 – Testing Plot between Scaled Estimated and Actual Average Speed for Left-Turn Movement on Minor Arterials.

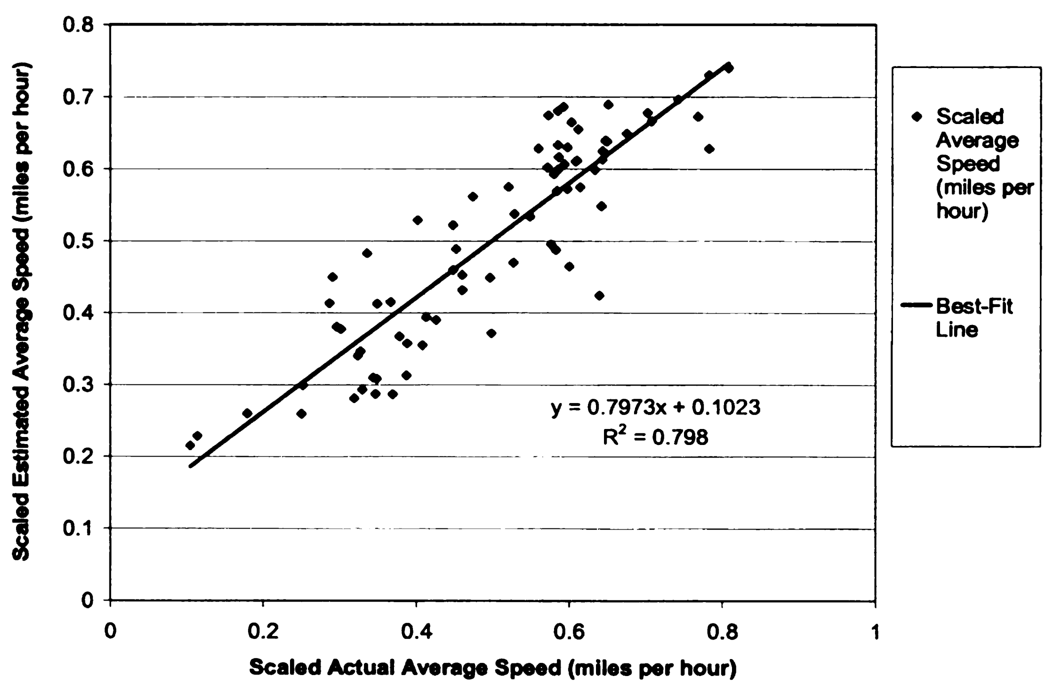

Figure 4.12 – Testing Plot between Scaled Estimated and Actual Average Speed for Right-Turn Movement on Minor Arterials.

The average speed estimation models presented in this chapter provide accurate and reliable estimates of average speed for major and minor arterials. It is also observed that the average speed estimation for right-turning movement is less capable as compared to through and left-turning models in terms of accuracy. The reason for lesser accuracy in right-turning movement models is attributed to the complexity of average speed variation with traffic demand in these movements. The right-turning movement shares the same lane and same signal control state as that of through movement on an arterial link. Hence, none of the input variables except the average flow rate for right-turning movement was distinguishing the average speed for right-turning movement. The lesser accuracy of speed estimation for right-turning movement can impact the travel time estimation and prediction models for right-turning movement. The results of the travel time estimation and prediction for right-turning movement on arterial links is presented in Chapter 7 where this impact is revisited.

Chapter 5

Conditional Independence Graphs Based Analysis of Travel Time Estimation and Prediction

The Conditional Independence (CI) graph is a founding stone of a series of Probabilistic Graphical models like the Markov Networks and the Bayesian Belief Networks. This section uses CI graph as a tool to analyze the interaction and conditional independence between traffic variables that represent the state of a link and hence influence the travel time. The application of the CI graph to this problem serves three-fold purposes: understanding the travel time estimation and prediction process, selecting parent input variables which refine the state of link, and improving the efficiency of State-Space Neural Networks which are used as the ultimate modeling tool for this problem.

This chapter begins with an overview of graphical models and CI graphs and presents the theory behind it. The next sections apply CI graph to analyze average travel time estimation and prediction for through, left, and right-turn movements on an arterial link. Much of the content of this chapter is inspired by the text on 'Graphical Models in Applied Multivariate Statistics' by Whittaker (1990).

5.1 Graphical Models and Conditional Independence Graphs in Applied Multivariate Statistics

5.1.1 Overview of Graphical Models and Conditional Independence Graphs

A Graphical model is a family of probability distributions that incorporates a specific set of conditional independence constraints listed in an independence graph. Graphical models are a statistical technique which provides a graph depicting conditional independence among a set of variables in a process which have multivariate normal distribution. This independence graph is called the Conditional Independence (CI) graph, simply because it is based on the notion of conditional probability among two sets of variables given another variable or set of variables. Thus, CI graph is a powerful way of summarizing the independence and interaction among a set of variables in a process.

The CI graphs are constructed from knowing the pair-wise conditional independencies among variables in a process. The tool that is needed to interpret this graph is the Markovian property. The Markovian property based inferences can be made solely after the construction of CI graph. As can be realized, CI graph provides that information on how different variables in an underlying process are interacting with each other on the basis of the conditional independence. It can be further extended to measure the amount of information contained in one random variable about the values of the other random variable. This information measure is called the Kullback-Leibler Information Divergence.

The construction of Conditional Independence graph can be further used to develop a Graphical Gaussian Model, which assumes a multivariate normal distribution among variables. The graphical Gaussian model can be fitted to the obtained CI graph with Kullback-Leibler Information Divergence to model the joint probability distribution of variables. The techniques of Maximum Likelihood estimation and likelihood ratio tests are used to determine the best fitting parameters in the graphical Gaussian model.

5.1.2 Utility and Scope of Graphical Models and Conditional Independence Graphs in Present Study

In the previous section, it was noted that the CI graphs are usually extended to fit a graphical Gaussian model to obtain the joint probability distribution of variables. The graphical Gaussian models are themselves a modeling technique for a process which has normal distribution for the variables. This thesis uses CI graph and estimation of Kullback-Leibler Information Divergence as a way to understand the process of travel time estimation and prediction. This study does not aim to fit a graphical Gaussian model extending the obtained CI graph. The reason is that State-Space Neural Networks (SSNN) has been found robust and reliable models to approximate complex non-linear processes like the average travel time estimation and prediction. So, no effort has been done to obtain the graphical Gaussian models for average travel time estimation and prediction. Instead, the CI graph and Kullback-Leibler Information Divergence are combined with SSNN for three-fold reasons:

1. *Understanding the travel time estimation and prediction* - The SSNN is capable of modeling a complex non-linear process because of its topology which provides it with associative memory for capturing temporal fluctuations as well as acting as an efficient input-output mapping tool. The associative memory of SSNN captures the state of a link, which is discussed in detail in the next chapter. But, it will be shown there that this associative memory is hard to decipher and is not directly correlated with the effect of different input variables. Thus, it prohibits the modeler from understanding the relationship and interactions among different variables affecting travel time. This deficiency of Artificial Neural Networks in general is called as 'black-box' nature. The CI graph is used to improve the understanding of average travel time estimation and prediction process. The concept of Kullback-Leibler Information Divergence is helpful in evaluating the amount of information contained in traffic variables about travel time, hence increasing the understanding of this process.
2. *Selecting parent input variables* - The SSNN in itself does not distinguish if all variables used are important and needed for modeling the process, or if some variables are redundant. This deficiency does not provide the modeler any information on which variables are significant for modeling the process and which variables are insignificant. The use of the Markovian property on CI graph selects a set of variables termed as parent input variables. These parent input variables give sufficient information about travel time that remaining variables are redundant and can be dropped from the input vector. The selection of parent input variables is covered in the following sections.

3. *Improving efficiency of SSNN* – The efficiency of SSNN may decrease because of the complexity in the topology which in turn appears because of the inclusion of more variables than necessary. So, CI graph has been combined for the first time in this study to reduce the number of variables needed as inputs to model the average travel time. It is found in this study that the efficiency of the SSNN actually increased using the information from CI graph.

5.1.3 Constructing Conditional Independence Graphs and Selection of Parent Input Variables

This section illustrates applying the CI graph in a process and finding the interaction and independence between the variables involved, which have a multivariate normal distribution. First the basic theory of conditional independence and Markovian property is presented which is used later to interpret a CI graph. The construction steps for obtaining a CI graph from a dataset is presented. Later on, Markovian properties are used to interpret CI graph.

Conditional Independence - The theory of independence and conditional independence between events or random variables is cemented into the very foundations of probability and statistics theory.

Consider two random vectors X and Y. The conditional density function of X knowing Y is written as,

$$f_{X|Y}(x;y) = \frac{f_{XY}(x;y)}{f_Y(y)}$$

The random vectors X and Y are independent if and only if the joint probability density function, f_{XY}, satisfies

$$f_{XY}(x,y) = f_X(x)\, f_Y(y)$$

for all values of x and y. The relationship is denoted by $X \perp\!\!\!\perp Y$.

Using the above relation which says that the joint probability density function of X and Y is the product of their marginal density functions, we get,

$$f_{X|Y}(x;y) = f_X(x) \text{ for all } x.$$

Thus, $X \perp\!\!\!\perp Y$ if and only if the conditional and the marginal density functions are identical.

Consider three random vectors X, Y, and Z. In a similar way, the random vectors Y and Z are conditionally independent on X if and only if

$$f_{YZ|X}(y,z;x) = f_{Y|X}(y;x)\, f_{Z|X}(z;x)$$

for all values of y and z and for all x for which $f_X(x)>0$. This is written as $Y \perp\!\!\!\perp Z|X$.

Conditional Independence Graph – After stating the meaning of conditional independence among a set of variables; consider a graph showing these variables and conditional independencies among them.

Let $X = (X_1, X_2, \ldots, X_k)$ denote a vector of random variables. A graph is constructed considering each of the random variable as a node or vertex. These nodes or vertices are either connected by an edge or left disconnected based on a certain rule that corresponds to the conditional independence among two variables knowing remaining all other variables. Let $K = \{1, 2, ..., k\}$ *be* the corresponding set of vertices which denote the random variables. The rule followed for connecting the vertices with an edge or leaving them disconnected in the graph is stated as follows: if a graph is constructed such that there is no edge between two vertices whenever a pair of variables is independent given all the remaining variables, the resulting graph is an independence graph, or more precisely a Conditional Independence graph. Stating a formal definition of conditional independence graph from Whittaker (1990)-

Definition 1: The *conditional independence graph* of X is the undirected graph $G = (K, E)$ where G is the graphical model, $K = \{1, 2, .., k\}$ and (i,j) is *not* in the edge set E if and only if $X_i \perp\!\!\!\perp X_j | X_{K \setminus \{i,j\}}$.

For Example, taking k = 4; i.e. for $X = (X_1, X_2, X_3, X_4)$, assume that the following independence relations are identified among these variables: $X_1 \perp\!\!\!\perp X_3 | \{X_2, X_4\}$, $X_1 \perp\!\!\!\perp X_4 | \{X_2, X_3\}$ *and* $X_2 \perp\!\!\!\perp X_4 | \{X_1, X_3\}$. The CI graph resulting from these relations is shown in the Figure 5.1. It is clear from this figure that the two variables which are conditionally independent like X_1 and X_3 do not have any connecting edge in the graph. The variables X_1 and X_3 are also separated by the rest of variables (X_2, X_4) given which they are independent. The resulting graph gives a picture of the pattern of dependence or association between the variables (Whittaker. 1990).

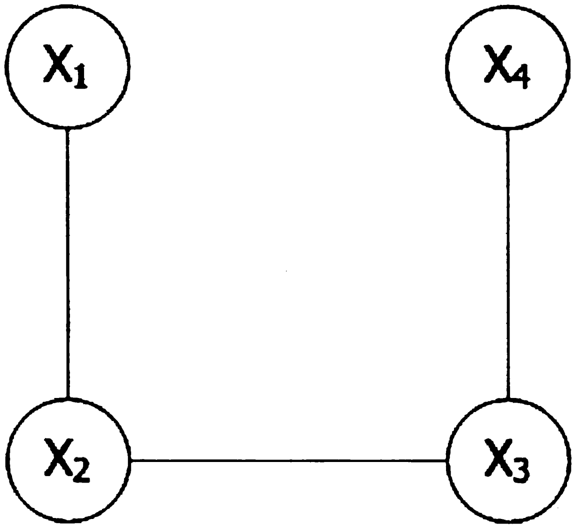

Figure 5.1 - A Simple Conditional Independence (CI) Graph for Conditional Independence Relations of Four Random Variables.

In the present study, a CI graph is obtained following the construction steps as mentioned later. Based on CI graph, the objective is to make some conditional independencies statements which will be similar to the statements made in the earlier example.

Markovian property - The Figure 5.1 is termed as Markov chain. The independence relations as specified for this example are based on three Markovian properties. These properties are listed as below (Whittaker, 1990):

- "the pairwise Markov property: that non-adjacent pairs of variables are independent conditional on remaining variables;
- the local Markov property: that conditional only on the adjacent variables, any variable is independent of all the remaining variables; and
- the global Markov property: that any two subsets of variables separated by a third is independent conditionally only on variables in the third subset."

The details of these Markovian properties are dealt in Whittaker (1990). Here, these are listed and applied to understand the independence statements that can be made for a CI graph.

Constructing conditional independence graph – The construction of the CI graph is presented here given the values of variables involved in the process being modeled. Suppose there are N observations of a k-dimensional random variable $X = (X_1, X_2, \ldots, X_k)$ as noted earlier. The steps for constructing a CI graph for this variable X is adapted from Whittaker (1990) and is presented below:

Step 1 – Estimate the sample correlation matrix $S = \text{corr}_N(X)$.

Step 2 – Compute the inverse of the sample correlation matrix, i.e. S^{-1}. The diagonal elements of this sample inverse correlation matrix are interpreted as being the proportion of variation in the corresponding variable explained by regressing on the remaining or rest of the variables. More explicitly, each diagonal element equals $1/(1-R^2)$, where R is the multiple correlation coefficient between that variable and the rest of the variables.

Step 3- Scale the sample inverse correlation matrix such that the diagonal elements are equal to 1. In this scaled inverse correlation matrix, the off-diagonal elements are the negatives of partial correlation coefficients, i.e. $\text{corr}_N(X_i, X_j \mid \text{rest})$, where 'rest' means rest or remaining of the variables.

Step 4 – Set any sufficiently small element of the scaled inverse correlation matrix to zero.

Step 5 - A graph G = (K, E) is drawn such that there are k nodes representing k dimensions of variable X, and E denotes the set of edges between the nodes. The rule that is followed in drawing this graph is that there is no connecting edge between two variables if their corresponding partial correlation coefficient is zero. The resulting graph is called a CI graph.

Parent input variables – The selection of parent input variables is one of the aims of using CI graph in this work. The parent input variables are a set of variables obtained from the CI graph which have sufficient information contained in them such that the remaining variables do not add any extra information to travel time estimation and prediction. For example, take the independence relation $X_1 \perp\!\!\!\perp X_3 | \{X_2, X_4\}$ used in earlier example in this section. The parent input variables for modeling X_1 are X_2 and X_4, since knowing these two variables X_1 is conditionally independent of X_3.

Kullback-Leibler Information Divergence – Kullback-Leibler Information Divergence, $I\ (f;g)$ is used to evaluate the amount of information available to discriminate between density functions f and g. The calculation of Kullback-Leibler Information Divergence is extended to evaluate the amount of information contained in one random variable about the values of others. In other words, it measures the strength of the specified conditional independencies shown in a CI graph.

Definition 2: The *Kullback-Leibler information divergence* between two probability density functions f and g for the random vector X is

$$I(f;g) = E\log\frac{f(X)}{g(X)}$$

where, E is the expectation taken with respect to the density f.

The measure to evaluate the information in one random vector about another is called information proper, which is directly related to information divergence.

Definition 3: The information in one random vector about another, the *information* proper, is

$$\text{Inf}(X \perp\!\!\!\perp Y) = I(f_{XY}; f_X f_Y)$$

where, I is the Kullback-Leibler information divergence.

It is already mentioned that if and only if $X \perp\!\!\!\perp Y$, $f_{XY} = f_X f_Y$. So, $I(f_{XY}; f_X f_Y)=0$, i.e., $\text{Inf}(X \perp\!\!\!\perp Y) = 0$ in this case. This implies that if $X \perp\!\!\!\perp Y$, no information is contained in X about the values of Y and vice-versa.

The information divergence and the information proper are related to the partial correlation coefficients (ρ) obtained from the scaled inverse correlation matrix. The following equation gives the relation between information proper and ρ,

$$\rho = + \{1\text{-exp}(-2\text{Inf})\}^{1/2}$$

$$\therefore \text{Inf} = -1/2\ \{\log(1-\rho^2)\}$$

So, $\text{Inf}\,(X \perp\!\!\!\perp Y|rest) = -1/2\ \{\log(1-\rho^2)\}$, where $\rho = \text{corr}_N(X, Y \mid rest)$.

5.2 Conditional Independence Graphs for Average Travel Time Estimation and Prediction for Through Movement

This section deals thoroughly with the construction of CI graph for average travel time estimation for through movement in an arterial link. Once the CI graph for travel time estimation is obtained, the inferences for travel time prediction are made. In this study, the CI graphs are developed for the travel time estimation only and the same selected parent input variables are chosen for travel time prediction modeling also. The travel time estimation and prediction modeling have the same output, i.e. the average travel time, with the only difference being whether it is observed in current or future departure time-period. So, the dependence of travel time on the traffic variables is the focus here irrespective of the departure time-period in which it is observed. Hence, the selected parent input variables by constructing the CI graph for travel time estimation can be used for prediction modeling also.

The variables that represent the state of a link as discussed in Chapter 2 are average flow rate for through, left, and right-turning movements (V_T, V_L, and V_R respectively); average speed for through, left, and right turning movements (S_T, S_L, and S_R respectively); green interval for the approach (G); offsets between the two signalized intersections forming the link (Off); speed limit of link ($S.L.$), and the ideal travel time for through and left turn movement (ITT_T and ITT_L, respectively). Chapter 4 deals with the estimation of average speed for each type of movement from basic available data. Since, the modeling approach in this study relies on basic available data only, estimated values of average speed are used for modeling here as well as further in the thesis.

Now, the construction of CI graph for average travel time estimation is presented for the arterial links. The assumption that the distributions of travel time estimation and the variables composing the state of the link are multivariate normal is made. So, all the data can be represented with sample mean vector and sample variance-covariance matrix. The sample correlation matrix is obtained from this variance-covariance matrix because it is easier to interpret the correlation matrix than the variance-covariance matrix.

The sample correlation matrix obtained from the data of average travel time for through movement and the state of the link in the current departure time period is presented in Table 5.1. This sample correlation matrix represents the degree of linear association between travel time and the independent variables depicting the state of the link, but it also shows that the independent variables are correlated among themselves. Hence, the correlation matrix does not provide useful information in the sense that if dependence of travel time is to be analyzed taking into account inter-correlation among independent variables.

The inverse matrix of this sample correlation matrix is found and scaled such that the diagonal elements of the inverse matrix equal to 1. The sample inverse correlation matrix is shown in Table 5.2. The sample inverse correlation matrix gives useful information which is used for constructing a CI graph. The off-diagonal elements of the scaled sample inverse correlation matrix are the negatives of the partial correlation coefficients between the corresponding pair of variables given the remaining variables. If the off-diagonal elements are zero then the corresponding variables are conditionally independent given the remaining variables. So, this scaled inverse correlation matrix is approximated such that the elements which have values close to zero (here, <0.1) are put as zero, and the values ≥ 0.1 are non-zero entities, symbolized as *. This approximated scaled inverse correlation matrix is shown in Table 5.3 gives the information about independence among the variables and can be conveyed by a CI graph.

Table 5.1 – Sample Correlation Matrix for Average Travel Time Estimation for Through Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_T
V_T	1.0													
V_L	0.7	1.0												
V_R	0.6	0.4	1.0											
G	0.3	0.2	0.2	1.0										
Off	-0.0	0.0	-0.0	-0.1	1.0									
S.L.	-0.2	-0.2	-0.2	-0.0	0.2	1.0								
ITT_T	-0.0	-0.1	-0.0	-0.3	-0.1	0.6	1.0							
ITT_L	-0.1	-0.2	-0.1	-0.3	-0.1	0.7	0.9	1.0						
Q_T	0.6	0.4	0.4	-0.0	-0.1	0.1	0.1	0.2	1.0					
Q_L	0.7	0.7	0.5	0.1	0.1	-0.1	-0.0	-0.1	0.6	1.0				
S_T	-0.3	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.3	-0.3	1.0			
S_L	-0.3	-0.4	-0.2	-0.0	-0.1	0.9	0.7	0.8	0.0	-0.3	0.8	1.0		
S_R	-0.4	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.4	-0.3	0.9	0.8	1.0	
TT_T	0.1	-0.0	0.0	-0.3	-0.0	0.6	0.8	0.9	0.5	0.1	0.5	0.7	0.4	1.0

Table 5.2 – Sample Scaled Inverse Correlation Matrix for Average Travel Time Estimation for Through Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_T
V_T	1.0													
V_L	-0.7	1.0												
V_R	-0.3	-0.4	1.0											
G	-0.1	-0.1	-0.2	1.0										
Off	0.0	-0.0	0.0	0.1	1.0									
S.L.	-0.0	0.2	0.2	0.0	-0.2	1.0								
ITT_T	-0.0	-0.0	-0.1	0.3	0.2	-0.6	1.0							
ITT_L	0.0	0.0	0.0	0.0	0.0	-0.0	-0.9	1.0						
Q_T	-0.2	-0.2	-0.4	-0.0	0.1	0.0	0.0	-0.2	1.0					
Q_L	0.0	-0.3	-0.2	-0.1	-0.1	0.0	-0.0	0.2	-0.6	1.0				
S_T	0.0	-0.1	0.0	-0.2	-0.1	-0.2	-0.0	-0.7	0.1	0.3	1.0			
S_L	-0.0	0.2	0.0	-0.0	0.1	-0.1	0.0	-0.1	-0.3	0.0	-0.8	1.0		
S_R	0.0	-0.2	-0.0	0.0	-0.1	0.0	-0.0	0.1	0.3	0.0	-0.3	-0.8	1.0	
TT_T	0.0	0.0	0.0	0.1	-0.0	0.1	0.0	-0.4	-0.4	-0.3	0.2	-0.3	-0.4	1.0

Table 5.3 – Approximated Sample Scaled Inverse Correlation Matrix for Average Travel Time Estimation for Through Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_T
V_T	*													
V_L	*	*												
V_R	*	*	*											
G	*	*	*	*										
Off	0	0	0	*	*									
S.L.	0	*	*	0	*	*								
ITT_T	0	0	*	*	*	*	*							
ITT_L	0	0	0	0	0	0	*	*						
Q_T	*	*	*	0	*	0	0	*	*					
Q_L	0	*	*	*	*	0	0	*	*	*				
S_T	0	*	0	*	*	*	0	*	*	*	*			
S_L	0	*	0	0	*	*	0	*	*	0	*	*		
S_R	0	*	0	0	*	0	0	*	*	0	*	*	*	
TT_T	0	0	0	*	0	*	0	*	*	*	*	*	*	*

Assume that the CI graph $G = (K, E)$ where, K represents the nodes and E represents the edges between these nodes. Each variable in the travel time estimation process is considered as a node resulting in $K = 14$. Now, an edge is put between two variables if there is a non-zero entity (represented by *) in Table 5.3. Also, no edge is put between two variables if there is a zero value between them in this table. The obtained graph following this rule is the CI graph for the average travel time estimation process. The obtained CI graph is shown in Figure 5.2. As seen in the figure, there are two types of edges between variables; directed and undirected. The directed edge is drawn if the general understanding of this process tells that variable at the head of directed edge is caused by variable at the tail of directed edge. For example, TT_T is caused by all other variables as they represent state of link. Similarly, if general understanding does not able to figure out the causation, the edge is undirected edge. For example, whether G and Off cause each other is not clear, so, undirected edge is put in between.

Figure 5.2 shows the pair of variables that are conditionally independent knowing a variable or set of variables. The point of interest in this figure is the average travel time. The travel time of the through movement (TT_T) is shown directly interacting with eight variables namely $Q_T, Q_L, S_T, S_L, S_R, G, S.L.,$ *and* ITT_L. Following the global Markovian property, it can be written as $\{TT_T \perp\!\!\!\perp (V_T, V_L, V_R, Off, ITT_T) \mid (Q_T, Q_L, S_T, S_L, S_R, G, S.L., ITT_L)\}$. The set of variables $Q_T, Q_L, S_T, S_L, S_R, G, S.L.,$ *and* ITT_L is identified as the parent input variables set. It is suggested by definition of conditional independence that the parent input variables are sufficient to estimate the average travel time and that no extra information is available in the remaining variables given the parent input variables.

After identifying the parent input variables, the concept of Kullback-Leibler information divergence and information proper can be used to evaluate the amount of information contained in each of these variables about TT_T. The information proper is calculated using the relation Inf = -1/2 $\{\log(1- \rho^2)\}$, which is stated earlier. The ρ values are obtained from Table 5.2. The information proper is calculated and is shown in Figure 5.2 in parentheses along with the directed edges that are connecting parent input variables to TT_T. This tells that Q_L and S_L have almost the same amount of information about TT_T. The variables like S_R, and ITT_L have almost double information about TT_T as compared to Q_L and S_L. The green interval, G and $S.L.$ have the least information about TT_T. Thus, CI graph not only tells which traffic variables are parent input variables but also their relative importance to estimate travel time.

It becomes obvious choice to use parent input variables alone instead of all the variables which were assumed previously to affect the estimation of average travel time. As discussed earlier in this section, the same parent input variables obtained for the travel time estimation process are used for prediction modeling also. These parent input variables will be used as input variables for the next modeling step, i.e. use of the State-Space Neural Networks (SSNN) for average travel time estimation and prediction of through movement.

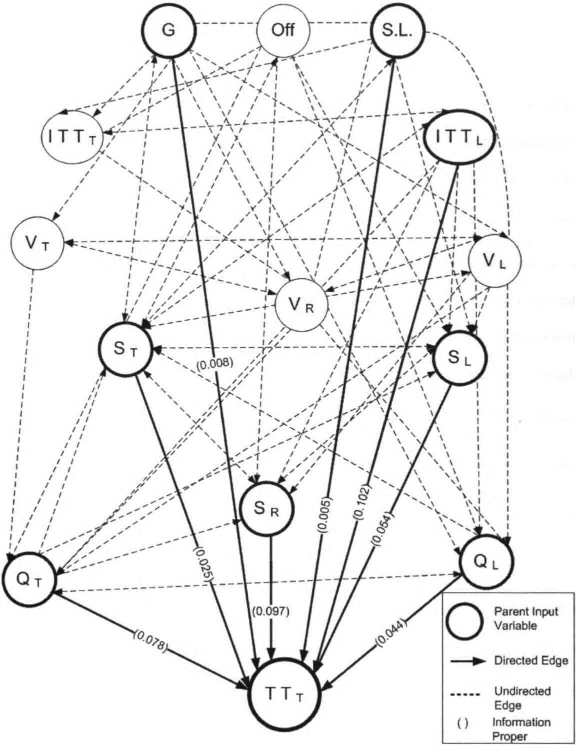

Figure 5.2 – Conditional Independence Graph for Average Travel Time Estimation for Through Movement.

5.3 Conditional Independence Graph for Average Travel Time Estimation and Prediction for Left-Turn Movement

The CI graph for travel time of the left-turn movement is drawn following the same methodology of the through movement presented in Section 5.2. The sample correlation matrix is shown in Table 5.4, which shows that almost all the variables have some correlation with TT_L, the average travel time for left-turn movement. The scaled inverse correlation matrix is shown in Table 5.5. The scaled inverse correlation matrix is approximated such that elements having values <0.1 are assumed to be zero entities and elements having values ≥ 0.1 are non-zero entities, symbolized by *. The approximated scaled inverse correlation matrix is presented in Table 5.6. Using this table, the CI graph is constructed for estimation of TT_L. This CI graph is presented in Figure 5.3, which can be interpreted using the global Markovian property as $\{TT_L \perp\!\!\!\perp (V_T, V_L, V_R, S_T, S_L, Off, ITT_T) \mid (Q_T, Q_L, S_R, G, S.L., ITT_L)\}$. So, the parent input variables to be used for estimation of average travel time, left-turn movement are $Q_T, Q_L, S_R, G, S.L., ITT_L$.

After identifying the parent input variables, the information proper is calculated using the relation Inf = $-1/2\ \{\log(1-\rho^2)\}$, which is stated earlier. The ρ values are obtained from Table 5.5. The information proper is calculated and is shown in parentheses in Figure 5.3 along with the directed edges that are connecting the parent input variables to TT_L. This shows that Q_T, G and $S.L.$ have almost same amount of information about TT_T. The variable Q_L has almost triple information about TT_L as compared to Q_T, G and $S.L.$ The variables S_R *and* ITT_L have almost five times information compared to Q_T, G and $S.L.$, and are most informative about TT_L. Thus, the CI graph not only tells which traffic variables are the parent input variables but also their relative importance to estimate travel time.

As discussed earlier in the Section 5.2, the same parent input variables obtained for the travel time estimation process are used for prediction modeling also. These parent input variables will be used as input variables in the next modeling step, i.e. use of the State-Space Neural Networks (SSNN) for average travel time estimation and prediction of left-turn movement.

Table 5.4 – Sample Correlation Matrix for Average Travel Time Estimation for Left-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_L
V_T	1.0													
V_L	0.7	1.0												
V_R	0.6	0.4	1.0											
G	0.3	0.2	0.2	1.0										
Off	-0.0	0.0	-0.0	-0.1	1.0									
S.L.	-0.2	-0.2	-0.2	-0.0	0.2	1.0								
ITT_T	-0.0	-0.1	-0.0	-0.3	-0.1	0.6	1.0							
ITT_L	-0.1	-0.2	-0.1	-0.3	-0.1	0.7	0.9	1.0						
Q_T	0.5	0.4	0.4	-0.0	-0.1	0.1	0.1	0.2	1.0					
Q_L	0.7	0.7	0.5	0.1	0.1	-0.1	-0.0	-0.1	0.6	1.0				
S_T	-0.3	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.3	-0.3	1.0			
S_L	-0.3	-0.4	-0.2	-0.0	-0.1	0.9	0.7	0.8	0.0	-0.3	0.8	1.0		
S_R	-0.4	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.4	-0.3	0.9	0.8	1.0	
TT_L	0.1	0.1	0.0	-0.4	0.0	0.4	0.8	0.8	0.3	0.2	0.5	0.4	0.4	1.0

Table 5.5 – Sample Scaled Inverse Correlation Matrix for Average Travel Time Estimation for Left-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_L
V_T	1.0													
V_L	-0.7	1.0												
V_R	-0.3	-0.4	1.0											
G	-0.1	-0.1	-0.2	1.0										
Off	0.0	-0.0	0.0	0.1	1.0									
S.L.	-0.0	0.2	0.2	0.0	-0.2	1.0								
ITT_T	-0.0	-0.0	-0.1	0.3	0.2	-0.6	1.0							
ITT_L	0.0	0.0	0.0	0.0	0.0	-0.0	-0.9	1.0						
Q_T	-0.2	-0.2	-0.4	-0.0	0.1	0.0	0.0	-0.2	1.0					
Q_L	0.0	-0.3	-0.2	-0.1	-0.1	0.0	-0.0	0.2	-0.6	1.0				
S_T	0.0	-0.1	0.0	-0.2	-0.1	-0.2	-0.0	-0.7	0.1	0.3	1.0			
S_L	-0.0	0.2	0.0	-0.0	0.1	-0.1	0.0	-0.1	-0.3	0.0	-0.8	1.0		
S_R	0.0	-0.2	-0.0	0.0	-0.1	0.0	-0.0	0.1	0.3	0.0	-0.3	-0.8	1.0	
TT_L	0.0	-0.0	0.0	0.2	-0.0	0.2	0.0	-0.4	-0.2	-0.3	0.0	-0.0	-0.4	1.0

Table 5.6 – Approximated Sample Scaled Inverse Correlation Matrix for Average Travel Time Estimation for Left-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_L
V_T	*													
V_L	*	*												
V_R	*	*	*											
G	*	*	*	*										
Off	0	0	0	*	*									
S.L.	0	*	*	0	*	*								
ITT_T	0	0	*	*	*	*	*							
ITT_L	0	0	0	0	0	0	*	*						
Q_T	*	*	*	0	*	0	0	*	*					
Q_L	0	*	*	*	*	0	0	*	*	*				
S_T	0	*	0	*	*	*	0	*	*	*	*			
S_L	0	*	0	0	*	*	0	*	*	0	*	*		
S_R	0	*	0	0	*	0	0	*	*	0	*	*	*	
TT_L	0	0	0	*	0	*	0	*	*	*	0	0	*	*

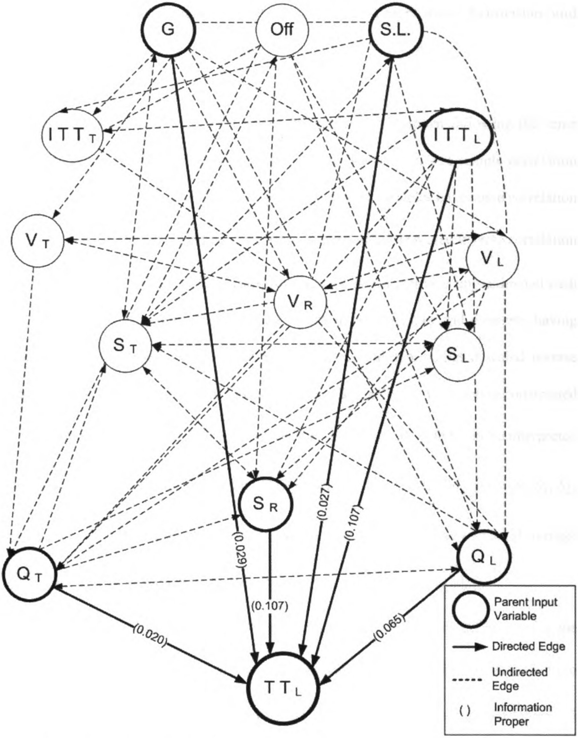

Figure 5.3 – Conditional Independence Graph for Average Travel Time Estimation for Left-Turn Movement.

5.4 Conditional Independence Graph for Average Travel Time Estimation and Prediction for Right-Turn Movement

The CI graph for travel time of the right-turn movement is drawn following the same methodology of the through movement presented in Section 5.2. The sample correlation matrix is shown in Table 5.7, which shows that almost all variables have some correlation with TT_R, the average travel time for right-turn movement. The scaled inverse correlation matrix is shown in Table 5.8. The scaled inverse correlation matrix is approximated such that elements having values <0.1 are assumed to be zero entities and elements having values $\geq$ 0.1 are non-zero entities, symbolized by *. The approximated scaled inverse correlation matrix is presented in Table 5.9. Using this table, the CI graph is constructed for estimation of TT_R. This CI graph is presented in Figure 5.4, which can be interpreted using the global Markovian property as $\{TT_R \perp\!\!\!\perp (V_T, V_L, V_R, Off, ITT_T) \mid (Q_T, Q_L, S_T, S_L, S_R, G, S.L., ITT_L)\}$. So, the parent input variables to be used for estimation of average travel time for right-turn movement are Q_T, Q_L, S_T, S_L, S_R, G, $S.L.$, and ITT_L.

After identifying the parent input variables, the information proper is calculated using the relation Inf = -1/2 {log(1- ρ^2)}, which is stated earlier. The ρ values are obtained from Table 5.8. The information proper is calculated and is shown in parentheses in Figure 5.4 along with the directed edges that are connecting the parent input variables to TT_R. This shows that S_R and ITT_L have the most information about TT_R. The least information is contained in G and $S.L.$ about TT_R. Thus, CI graph not only tells which traffic variables are the parent input variables but also their relative importance to estimate travel time.

As discussed earlier in the Section 5.2, the same parent input variables obtained for the travel time estimation process are used for prediction modeling also. These parent input variables will be used as input variables for the next modeling step i.e. use of the State-Space Neural Networks (SSNN) for average travel time estimation and prediction of right-turn movement.

Table 5.7 – Sample Correlation Matrix for Average Travel Time Estimation for Right-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_R
V_T	1.0													
V_L	0.7	1.0												
V_R	0.6	0.4	1.0											
G	0.3	0.2	0.2	1.0										
Off	-0.0	0.0	-0.0	-0.1	1.0									
S.L.	-0.2	-0.2	-0.2	-0.0	0.2	1.0								
ITT_T	-0.0	-0.1	-0.0	-0.3	-0.1	0.6	1.0							
ITT_L	-0.1	-0.2	-0.1	-0.3	-0.1	0.7	0.9	1.0						
Q_T	0.5	0.4	0.4	-0.0	-0.1	0.1	0.1	0.2	1.0					
Q_L	0.7	0.7	0.5	0.1	0.1	-0.1	-0.0	-0.1	0.6	1.0				
S_T	-0.3	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.3	-0.3	1.0			
S_L	-0.3	-0.4	-0.2	-0.0	-0.1	0.9	0.7	0.8	0.0	-0.3	0.8	1.0		
S_R	-0.4	-0.3	-0.3	-0.1	0.1	0.8	0.7	0.7	-0.4	-0.3	0.9	0.8	1.0	
TT_R	0.1	-0.0	0.1	-0.2	-0.1	0.6	0.8	0.8	0.5	0.1	0.5	0.6	0.4	1.0

Table 5.8 - Sample Scaled Inverse Correlation Matrix for Average Travel Time Estimation for Right-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_R
V_T	1.0													
V_L	-0.7	1.0												
V_R	-0.3	-0.4	1.0											
G	-0.1	-0.1	-0.2	1.0										
Off	0.0	-0.0	0.0	0.1	1.0									
S.L.	-0.0	0.2	0.2	0.0	-0.2	1.0								
ITT_T	-0.0	-0.0	-0.1	0.3	0.2	-0.6	1.0							
ITT_L	0.0	0.0	0.0	0.0	0.0	-0.0	-0.9	1.0						
Q_T	-0.2	-0.2	-0.4	-0.0	0.1	0.0	0.0	-0.2	1.0					
Q_L	0.0	-0.3	-0.2	-0.1	-0.1	0.0	-0.0	0.2	-0.6	1.0				
S_T	0.0	-0.1	0.0	-0.2	-0.1	-0.2	-0.0	-0.7	0.1	0.3	1.0			
S_L	-0.0	0.2	0.0	-0.0	0.1	-0.1	0.0	-0.1	-0.3	0.0	-0.8	1.0		
S_R	0.0	-0.2	-0.0	0.0	-0.1	0.0	-0.0	0.1	0.3	0.0	-0.3	-0.8	1.0	
TT_R	0.0	0.0	0.0	0.1	0.0	0.1	0.0	-0.4	-0.4	-0.3	0.2	-0.3	-0.4	1.0

Table 5.9 – Approximated Sample Scaled Inverse Correlation Matrix for Average Travel Time

Estimation for Right-Turn Movement

	V_T	V_L	V_R	G	Off	S.L.	ITT_T	ITT_L	Q_T	Q_L	S_T	S_L	S_R	TT_R
V_T	*													
V_L	*	*												
V_R	*	*	*											
G	*	*	*	*										
Off	0	0	0	*	*									
S.L.	0	*	*	0	*	*								
ITT_T	0	0	*	*	*	*	*							
ITT_L	0	0	0	0	0	0	*	*						
Q_T	*	*	*	0	*	0	0	*	*					
Q_L	0	*	*	*	*	0	0	*	*	*				
S_T	0	*	0	*	*	*	0	*	*	*	*			
S_L	0	*	0	0	*	*	0	*	*	0	*	*		
S_R	0	*	0	0	*	0	0	*	*	0	*	*	*	
TT_R	0	0	0	*	0	*	0	*	*	*	*	*	*	*

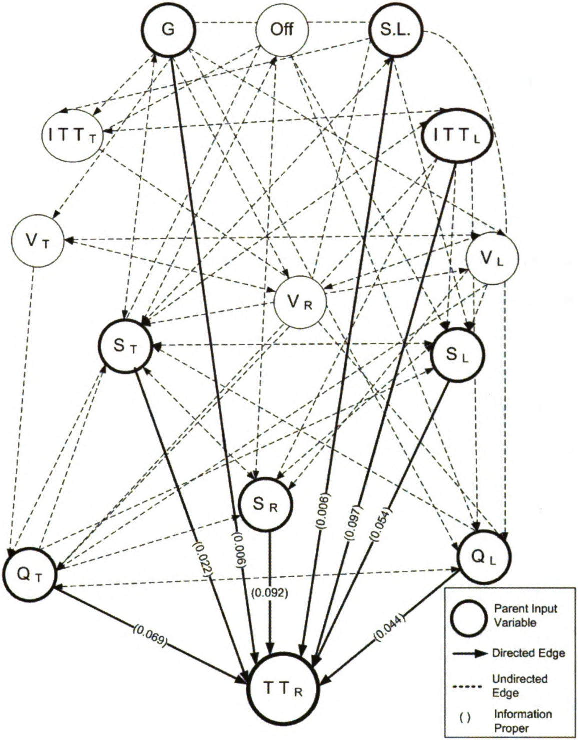

Figure 5.4 – Conditional Independence Graph for Average Travel Time Estimation for Right-Turn Movement.

Chapter 6

State-Space Neural Networks Modeling for Average Travel Time Estimation and Prediction

The Conditional Independence (CI) graphs that were developed in the previous chapter select parent input variables required for travel time estimation and prediction. The next modeling step is to develop the State-Space Neural Network (SSNN) models to use the information from the CI graphs to model the travel time process.

This chapter provides a review of modeling approaches that have been used by researchers for short-term traffic forecasting in general. The SSNN is identified as a promising tool to model the complex nonlinear dynamic process of travel time on urban arterial networks. The SSNN models are developed for the smallest unit of arterial networks, i.e. arterial links. This study develops separate SSNN models based on the same topology for three types of traffic movements possible on each arterial link i.e. through, left, and right-turn movements. The topology of the SSNN models is discussed and found suitable for the study. The training and testing aspects of neural network modeling is illustrated. Finally, the travel time estimation and prediction for an arterial route is developed using SSNN models already developed for its components, i.e. arterial links.

6.1 Modeling Approaches for Travel Time Estimation and Prediction

The problem of short-term travel time prediction is similar in methodology as forecasting traffic conditions on a transportation networks. There has been considerable research on short-term forecasting of average flow rate, average speed, and queue length. Though, travel time prediction is similar to the forecasting of other traffic condition indicators or variables, it is much more complex in formulation. An extensive literature survey has been presented by Vlahogianni, et al. (2004) discussing modeling approaches for short-term traffic forecasting in general.

Here, three broad categories are identified in which modeling approaches for short-term travel time prediction can be classified:

1. Univariate Time-Series Forecasting – This approach is quite popular time-series forecasting approach used in all disciplines including traffic engineering. The travel time is considered a random variable and is forecasted by focusing on the second-order properties only. If *Xt* is a certain random variable representing travel time at any time instant *t*, the expected values of *E(Xt)* and expected products, *E(Xt+h, Xt)*, t = 1, 2,..., and h = 0 ,1, 2,...are specified. The properties of sequence Xt is dependent on these expected values and expected products only, and that is why it is refered to as second-order properties (Brockwell and Davis 2002). Thus, here the dependence of travel time on other variables is ignored and only focused on the second-order properties. The Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) family of time series models are used in this case.

2. Multivariate Time-Series Forecasting – This approach considers that the time series (here, travel time series) is represented as components of some multivariate time series Xt. Here, not only serial dependence within each component series $\{Xti\}$ exists, but also interdependence between the different component series $\{Xti\}$ and $\{Xtj\}$, $i \neq j$ (Brockwell and Davis 2002). The multivariate time-series are modeled using multivariate ARMA processes. The State-Space models and the associated Kalman recursions also belong to the family of multivariate time-series forecasting and are powerful approaches.
3. Non-Parametric Modeling- Unlike the two time-series models mentioned earlier, non-parametric modeling approach does not assume any specific functional form for the dependent and independent variables. These are data-driven approaches and can be again broadly classified into two modeling techniques, namely Non-parametric regression, and Artificial Neural Networks (ANN). Various forms of ANN have been used for the prediction of traffic conditions indicators and travel time.

This thesis uses a variant of ANN called State-Space Neural Networks (SSNN), a generic form of the Recurrent Neural Networks (RNN) to model travel time estimation and prediction.

6.2 State-Space Neural Networks

The State-Space Neural Networks are a generic form of Recurrent Neural Networks. The utility of SSNN lies in its topology which can be represented by state-space equations. The notion of state is considered important in the traffic processes including travel time estimation and prediction. The RNN and SSNN are introduced here and a brief survey of literature is provided in these topics.

6.2.1 Recurrent Neural Networks

Recurrent Neural Networks (RNN) are a powerful class of Artificial Neural Networks (ANN) to model non-linear dynamical systems. There have been vast applications of RNN in system identification, signal processing, and forecasting. The Feed-forward Neural Networks (FNN), which is a basic and general ANN, captures the system dynamics by including present state and past state inputs as separate input nodes in the input layer. However, RNN tends to capture the complex non-linear system dynamics by involving feedback loop from inputs, outputs, or prediction errors. The text by Mandic and Chambers (2001) compares RNN with FNN in terms of non-linear dynamic systems modeling.

The RNN has two functional uses because of its topology (Haykin 1999)-

- Associate memory and
- Input-output mapping network.

Simply speaking, RNN have local feedback loops (recurrent loops) from the hidden layer or outputs and can be trained as global feed-forward networks. The three types of RNN which are most popular in the literature are globally recurrent neural networks (GRNN), output-feedback recurrent neural networks (OFRNN), and fully recurrent neural network (FRNN) as discussed in Medsker and Jain (1999).

Because of the inherent features of RNN like the recurrent loops with time delays and input-output mapping networks, a variant of RNN called State-Space Neural Networks (SSNN) is used in this study.

6.2.2. State-Space Neural Networks

The State-Space Neural Networks (SSNN) is a generic form of RNN which has underlined state-space representation of the process being modeled. The state-space representation of a SSNN is mathematically stated as state-space equations. Let a non-linear time-variant dynamic system to be noise free (Haykin 1999), then,

$$x(t+1) = \varphi\,(W_a x(t) + W_b u(t))$$

$$y(t) = C\,x(t)$$

where,

q, number of unit delays used to feed the output of the hidden layer back to the input layer (order of the model),

m, number of input variables,

p, number of output variables,

t, a discrete time-point,

x(t), a q-by-1 vector denoting the state of a nonlinear discrete -time system,

u(t), a m-by-1 vector u(t) denoting the input applied to the system,

y(t), a p-by-1 vector denoting the corresponding output of the system.

$\mathbf{W_a}$, a q-by-q matrix,

$\mathbf{W_b}$, a q-by-(m+1) matrix,

C, a p-by-q matrix; and

$\boldsymbol{\varphi}$ is a nonlinear mapping function.

These state-space equations suppose that at each discrete time-point *t*, a vector valued observation *y(t)* is related to *x(t)*. The *x(t)* is termed as state of the system at discrete time-point *t*. The second equation representing the relationship between *y(t)* and *x(t)* is called observation equation. The observation is shown as a linear function of state of the system *(C)*, but this relationship can be captured by a non-linear function too. The first equation is termed transition equation or state equation as it models the state of the system at any time-point *t*. The state of the system at any time-point *t+1* depends on the state at past time-points *(t)*, and external inputs *(u(t))* to the system. The function that is modeled between the state of the current time-point and the state in the past time-points and external input vector is a non-linear mapping function *(φ)*. The state of the system is not observed directly; only through a linear or non-linear mapping. Thus, these are the hidden internal states of the system and they stay in the so-called space hence, are called State-Space models (Honkela 2001).

The conditional independence graph representation of these state-space equations is presented by Whittaker, et al. (1997) for better understanding of these equations and relationship between inputs, state, and output. The conditional independence graph proposed by Whittaker, et al. (1997) is adapted and is shown in Figure 6.1. The state of the system at any time-point $t+1$, $x(t+1)$, is connected to input $y(t)$ and the state of the previous time-point, $x(t)$. This dependence is shown by a connecting arrow in the figure. The 'past' represents values of past state (like $x(t-1)$, $x(t-2)$...) and input variables ($u(t-1)$, $u(t-2)$...). There is no edge to $x(t+1)$ or to $y(t)$ from the past, which indicates that enough information is held in the current state $x(t)$ to render past information. This conditional independence graph shows simple state-space models. In other cases, there can be an edge from output ($y(t)$) to the input vector ($u(t)$). Because of its simplicity, this graph is discussed here. The advanced forms of state-space models will be encountered in further discussion.

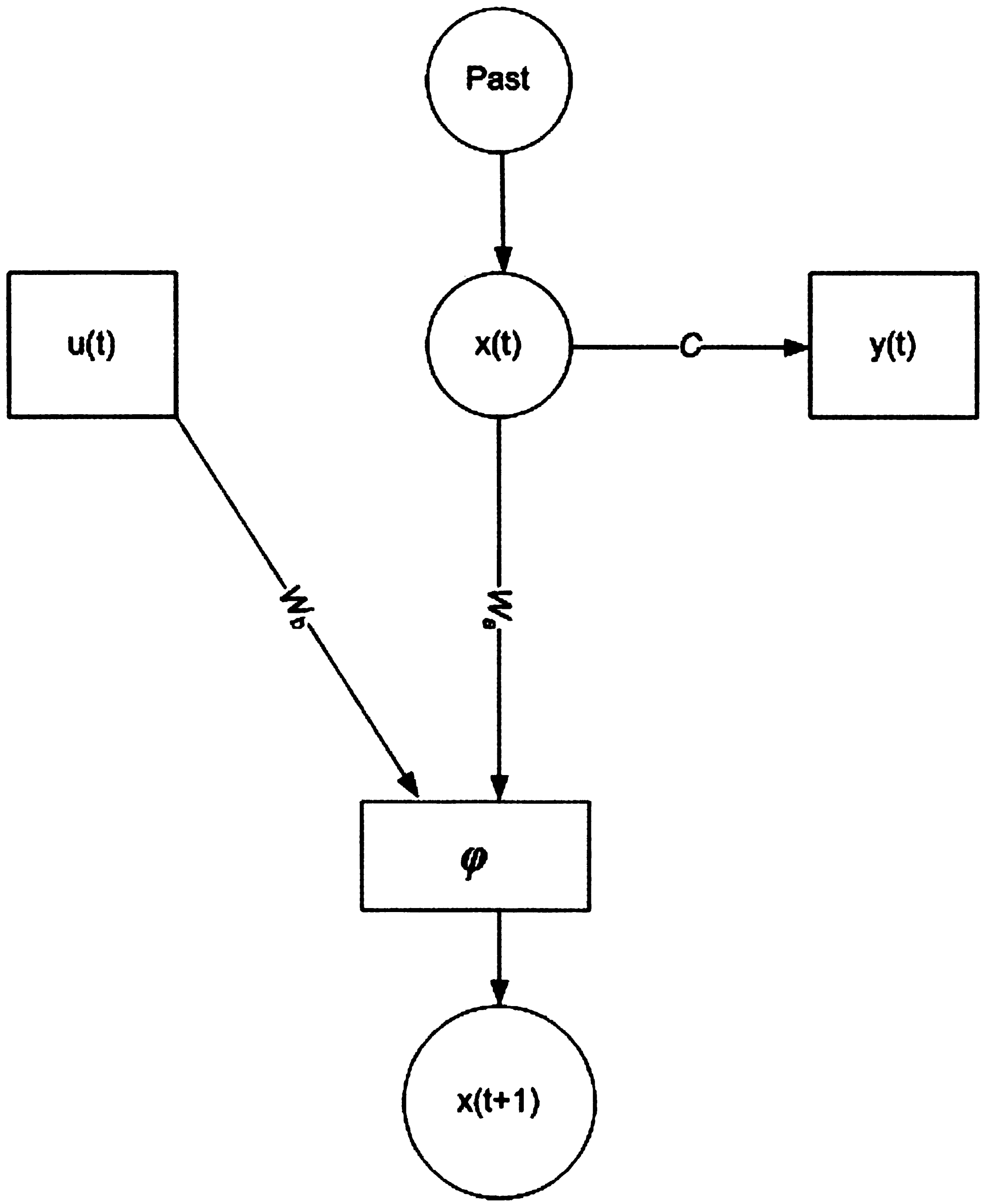

Figure 6.1 – Conditional Independence Graph Representation of a General State-Space Model.

Various types of SSNN models have been presented and studied by researchers. The main difference is based on the nature of feedback loops inside the topology of a SSNN.

Elman (1990) presented a state-space neural network shown in Figure 6.2. In the Elman network, an additional layer called context layer is included which is connected to the hidden layer. The hidden layer values with a time-delay activate the context layer. The dotted lines represent trainable connections whereas the solid arch between hidden and context units represent a non-trainable connection. This non-trainable connection between context and hidden layer has a fixed weight of 1.0. The context layer is also a hidden layer in the sense that it interacts with the nodes that are internal to the system and not with the outside nodes (i.e. the input nodes). The provision of a context layer is useful as it stores the hidden layer values of past time-point. In the state-space notion, hidden and context layers act as the state of the system in current and past time-points respectively. Thus, the context layer provides a short-term memory to the network which is useful in learning dynamic systems.

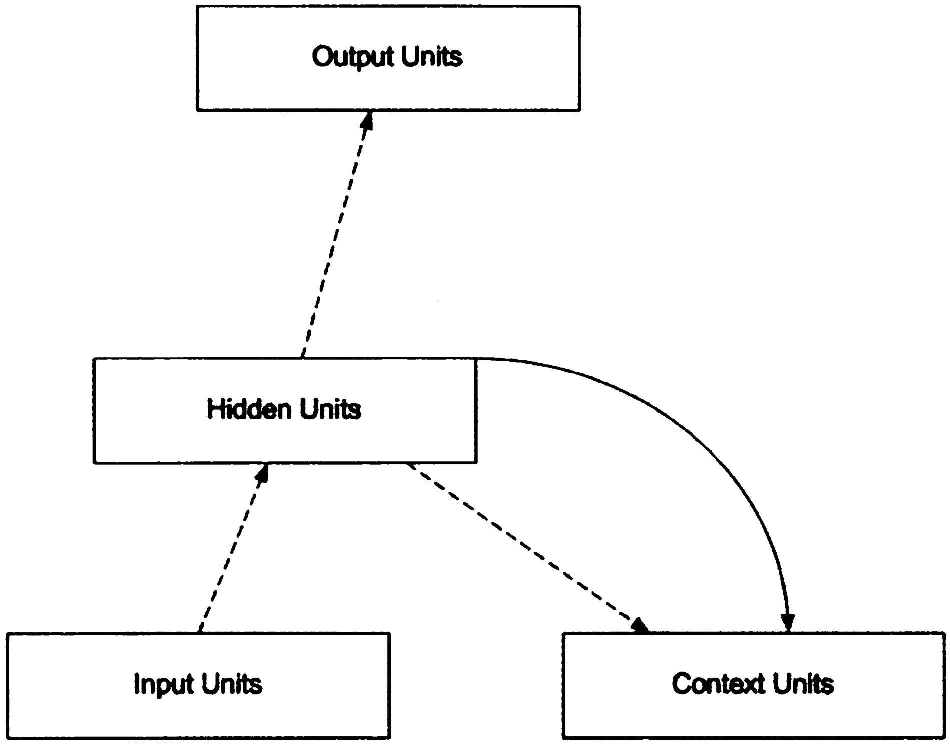

Figure 6.2 – A Simple Recurrent Neural Network Proposed by Elman with Context Layer for Short-Term Memory.

Jordan (1986) also presented a recurrent neural network where a feedback from the output layer is provided to the context layer.

6.3 State-Space Neural Networks Topology for Travel Time Estimation and Prediction

In this research, State-Space Neural Network (SSNN) models are designed for the travel time estimation and prediction problems. The SSNN models for travel time estimation and prediction are developed separately because of the difference in the two problems in terms of time-instants at which inputs, state of system and output are considered in the network.

The smallest unit of an urban arterial network is an arterial link, which is bounded by two signalized intersections one at each end. An arterial link is characterized by the presence of three types of vehicle movements: through, left and right-turning vehicles. In this research, a SSNN topology is developed for an arterial link for each of the three movements.

6.3.1 SSNN Topology for Travel Time Estimation

The SSNN model for travel time estimation is shown in Figure 6.3. The figure shows an input layer whose input vector consists of input variables *($U(t)$)* at the current departure time-period t that affect travel time. This input layer is connected to the hidden layer by a trainable connection, i.e. a connection whose weights are trained by the gradient descent algorithm. The hidden layer in the network represents the state of the system at the time-period t. This state at the time-period t activates the nodes in the context layer which represent the state of system in the past departure time-period $t-1$. The context layer nodes store the hidden nodes values with a time-delay of 1. The context layer has a trainable connection with the hidden layer which provides information about the state at the past time-period to the hidden layer. So, the state of the system (here, an arterial link) at the current and the past time-periods is modeled as internal hidden and context layers respectively.

The hidden layer is then connected to the output layer by a trainable connection. The output layer represents the output of the travel time process, which is the average travel time. Since, the inputs and outputs are both at same time-periods *t*, this process is modeling travel time estimation.

Three similar SSNN models are developed, one for each of the traffic movement. These are termed SSNN-Thru, SSNN-Left, and SSNN-Right models. The output of the SSNN-Thru model for travel time estimation of the through movement on an arterial link is the average travel time at the current departure time-period for through movement. Similarly, the outputs of the SSNN-Left and SSNN-Right models are the average travel time at the current departure time-period for the left and right-turning traffic, respectively. Apart from the difference in the outputs of three SSNN models, the inherent topology is same.

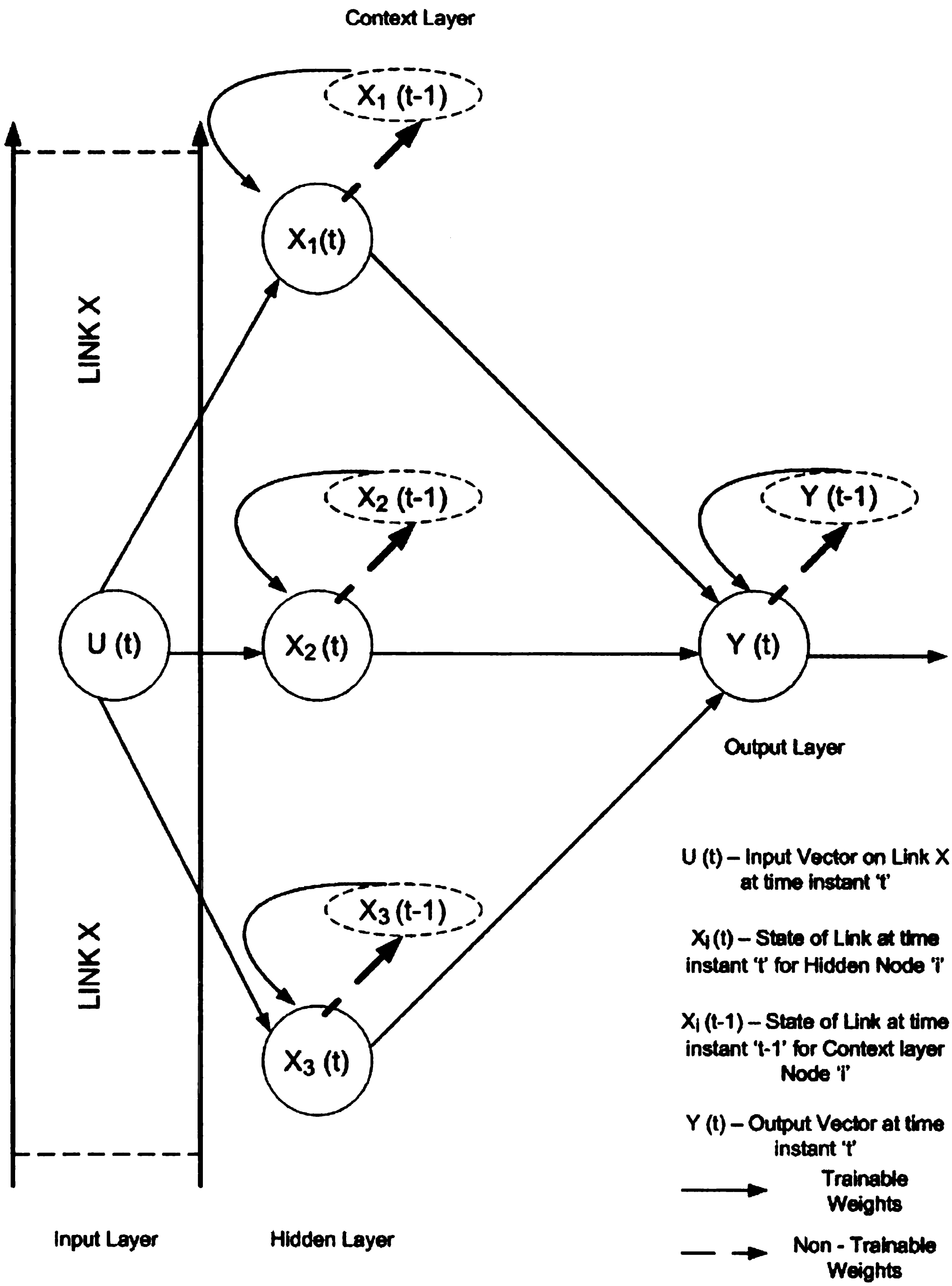

Figure 6.3 – State-Space Neural Network Topology adopted for Travel Time Estimation on an Arterial Link for Specific Through, Left, and Right-Turn Movements.

One important issue about the topology of SSNN models is the number of nodes in the hidden and context layers. Since, the models are developed for each arterial link; the state of a link can be represented by one node only. Hence, there should be a single node in each of the hidden and context layers. This is depicted in the Elman network also which assumes a single non-linear node in the hidden and context layers representing the state of the system. It has been argued by Rivals and Personnaz (1996) that the state variables in a neural network can be computed with as many nonlinear hidden neurons as required. It is also argued that in the Elman network, the nonlinear mapping function between the observed output and the hidden layer is constrained by a single hidden neuron, which leads generally to a non-optimal state-space representation. Following this argument in the present SSNN models, the number of hidden neurons (which is equal to those in the context layer) was selected after testing different numbers ranging from 3 to 10 to obtain a minimal optimal state-space representation. Hence, the SSNN models for each type of traffic movement may have different number of hidden and context neurons in their topology. Figure 6.3 is a generic representation for any type of traffic movement and shows only 3 neurons in the hidden and context layers. The actual number of hidden and context neurons for the SSNN models of the different traffic movements are presented in Table 6.1.

Table 6.1 – Number of Hidden and Context Layer Nodes used in SSNN Travel Time Estimation Models for Through, Left, and Right-Turn Movements

SSNN Travel Time Estimation Models	Number of Hidden and Context Layer Nodes
SSNN-Thru	4
SSNN-Left	10
SSNN-Right	10

6.3.2 SSNN Topology for Travel Time Prediction

The SSNN topology for short-term travel time prediction is similar to that of travel time estimation as discussed in detail in the earlier section. This section deals mainly with the difference in SSNN topology for the travel time prediction process. Figure 6.4 shows the SSNN topology for the average travel time prediction on an arterial link for a specific traffic movement. The input node consists of input vectors which are known at current departure time-period t. The output node is the average travel time which is predicted for one-step ahead future departure time-period $t+1$. The hidden nodes store the state of arterial link for future time-period $t+1$, and the past state values for time-period t are stored in the context layer nodes.

Based on this topology, three SSNN models are developed for travel time prediction for through, left, and right-turning movements on a link. The numbers of nodes in the hidden and context layer are shown to be 3 just for the sake of representation. The actual number of hidden and context nodes vary for the three different SSNN models that were developed for through, left, and right-turning movement. The actual number of hidden and context neurons for the SSNN models for the different traffic movements are presented in Table 6.2.

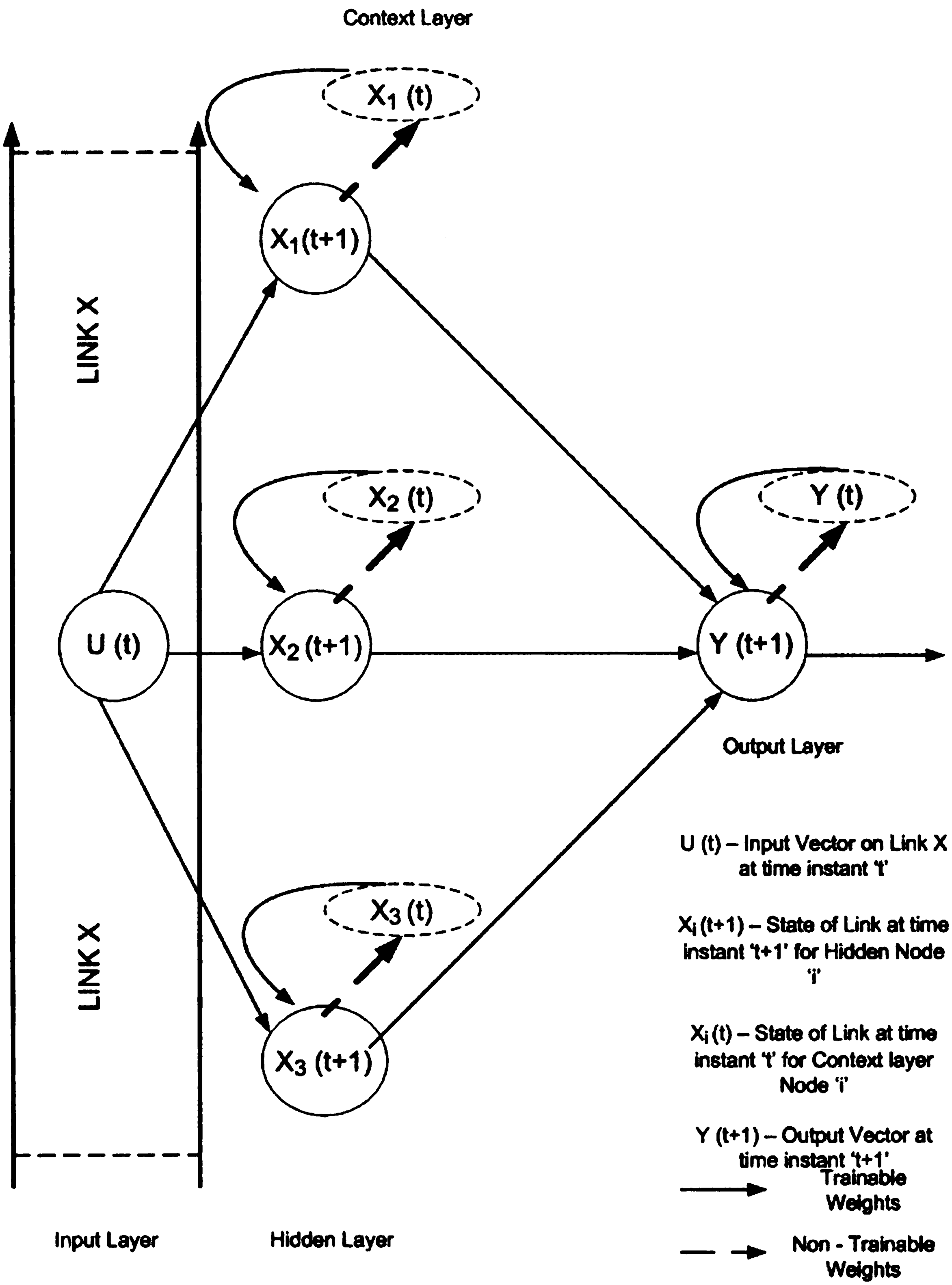

Figure 6.4 – State-Space Neural Network Topology adopted for Travel Time Prediction on an Arterial Link for Specific Through, Left, and Right-Turn Movements.

Table 6.2 – Number of Hidden and Context Layer Nodes used in SSNN Travel Time Prediction Models for Through, Left, and Right-Turn Movements

SSNN Travel Time Prediction Models	Number of Hidden and Context Layer Nodes
SSNN-Thru	4
SSNN-Left	10
SSNN-Right	10

6.4 State-Space Neural Networks based on Conditional Independence Graphs for Travel Time Estimation and Prediction

In the SSNN topology, the input layer consists of one input node in each SSNN model. This input node is fed with the input vector which consists of the set of traffic variables that affect travel time. The Conditional Independence graph application to understand the travel time process has the effect of refining the independent traffic variables and retains only those variables that affect travel time. These input variables were called the parent input variables in Chapter 5. The travel time estimation process is modeled by SSNN models where the inputs to the model are parent input variables only (instead of all the traffic variables). The parent input variables were identified in Chapter 5 for average travel estimation. Table 6.3 lists the parent input variables which act as input vector to the SSNN models along with the corresponding output of the model.

Table 6.3 – Input and Output Variables for SSNN Travel Time Estimation Models for Through, Left, and Right-Turn Movements

SSNN Travel Time Estimation Models	Input Variables (current departure period)	Output Variable (current departure period)
SSNN-Thru	$Q_T, Q_L, S_T, S_L, S_R, G, S.L.,$ *and* ITT_L	TT_T
SSNN-Left	$Q_T, Q_L, S_R, G, S.L., ITT_L$	TT_L
SSNN-Right	$Q_T, Q_L, S_T, S_L, S_R, G, S.L.,$ *and* ITT_L	TT_R

The SSNN models for travel time estimation use input variables as a set of parent input variables which were selected using the conditional independence (CI) graph analysis. The selection of parent input variables is an important component of SSNN models for travel time estimation process. Here, the SSNN models are based on the CI graphs and hence, these models are termed CI-SSNN models. Thus, three CI-SSNN models termed CI-SSNN-Thru, CI-SSNN-Left, and CI-SSNN-Right are developed for travel time estimation on an arterial link for the three types of traffic movements.

The SSNN models for travel time prediction also include the same parent input variables obtained from the CI graphs. Additionally, the input vector contains the estimated travel time of the current departure time-period, which is the output of CI-SSNN models for travel time estimation. Hence, travel time prediction models are also termed CI-SSNN-Thru, CI-SSNN-Left, and CI-SSNN-Right for travel time prediction on an arterial link for the three types of traffic movements. Table 6.4 lists the input and output variables for CI-SSNN models for travel time prediction.

Table 6.4 – Input and Output Variables for SSNN Travel Time Prediction Models for Through, Left, and Right-Turn Movements

SSNN Travel Time Prediction Models	Input Variables (current departure period)	Output Variable (future departure period)
SSNN-Thru	TT_T *(estimated)*, Q_T, Q_L, S_T, S_L, S_R, *G*, *S.L.*, *and* ITT_L	TT_T
SSNN-Left	TT_L *(estimated)*, Q_T, Q_L, S_R, *G*, *S.L.*, ITT_L	TT_L
SSNN-Right	TT_R *(estimated)*, Q_T, Q_L, S_T, S_L, S_R, *G*, *S.L.*, *and* ITT_L	TT_R

6.5 Training and Testing of State-Space Neural Networks for Travel time Estimation and Prediction

The CI-SSNN models as discussed in earlier section are trained and tested to evaluate their performance and efficiency to model travel time. A programming code is written in MATLAB software to create a Recurrent Neural Network based on the topology presented earlier. After creating the RNN, the network is trained using a training set. The input patterns are presented sequentially (retaining the exact sequence of time-periods as it evolves) rather than as concurrent matrix (implying batch-mode of training), which is usually done in simple Feed-forward Neural Networks (FNN). This sequential presentation of the training set allows the network to learn temporal fluctuations in traffic conditions affecting travel time. The code for the SSNN models used in this study is provided in Appendix A.

The training and testing sets are obtained from the data-collection done using CORSIM simulation. The process of obtaining training and testing sets is already covered in Chapter 3. Here, a review is provided of the composition of these sets. The training and testing sets consist of data or values of traffic variables for four arterials namely 6-10, 16-20, 21-25, and 3-23 on the arterial network (refer to Figure 5.1 in Chapter 5). Each of these arterials contains 4 links. The CORSIM simulation was run two times with different random seeds. The first simulation run is taken as training set and the second as testing set.

The features of the training and testing sets for travel time estimation and prediction are presented in Table 6.5. It specifies the CORSIM run and data-points that each set is based on.

It is noted in this study that no validation set is used apart from the training and testing sets. The reason that the validation set is generally used in Artificial Neural Networks is to ensure that the neural network model is not over-trained, so that it performs comparably well on both training and testing sets. The validation set is not used in this study due to the lack of data and moreover, sufficient number of trials have been done to obtain a neural network model that performs comparably well on the training as well as testing sets. This argument holds good as the analysis of the results of neural network models in next chapter show the comparable accuracy on training and testing sets.

Table 6.5 – Training and Testing Sets Composition for SSNN Travel Time Estimation and Prediction Models

CI-SSNN Travel Time Estimation Models	Training Set	CORSIM simulation Run 1	304 data-points
	Testing Set	CORSIM simulation Run 2	304 data-points
CI-SSNN Travel Time Prediction Models	Training Set	CORSIM simulation Run 1	288 data-points
	Testing Set	CORSIM simulation Run 2	288 data-points

There are certain guidelines for training and testing in Artificial Neural Networks as discussed in Haykin (1999). These guidelines were found useful and easy to follow and they provide an efficient learning and testing of a neural network model. The guidelines are usually for selecting learning rate, momentum constant, and number of hidden nodes in the network. These values are obtained using the trial and error approach. Different values within a range are tried and the selected values are those that provide a smooth training graph and meet the goal error requirements. The optimum values of learning rate and momentum constants found for the networks in this study were 0.0001 and 0.9 respectively. The optimum number of hidden nodes obtained for different CI-SSNN models were previously shown in Tables 6.1 and 6.2.

The training graph is plotted by MATLAB between the Mean Square Error (MSE) and the number of iterations the model has run. When the MSE value meets the specified goal error (0.001 in this study), the network is assumed to have converged to an optimum weight configurations. The model is tested using the trained neural network with the obtained optimum weight configuration. A criterion of an efficient training by a neural network is to see the training graph by doing different trials with the learning rate, momentum constant and hidden node values. A training graph which is smooth without many fluctuations represents an efficient learning or training of the network. Figure 6.5 shows a training graph which is representative of an efficient training for the CI-SSNN models in this study. In this figure, the X-axis represents the epochs or iterations and the Y-axis represents the mean square error of training.

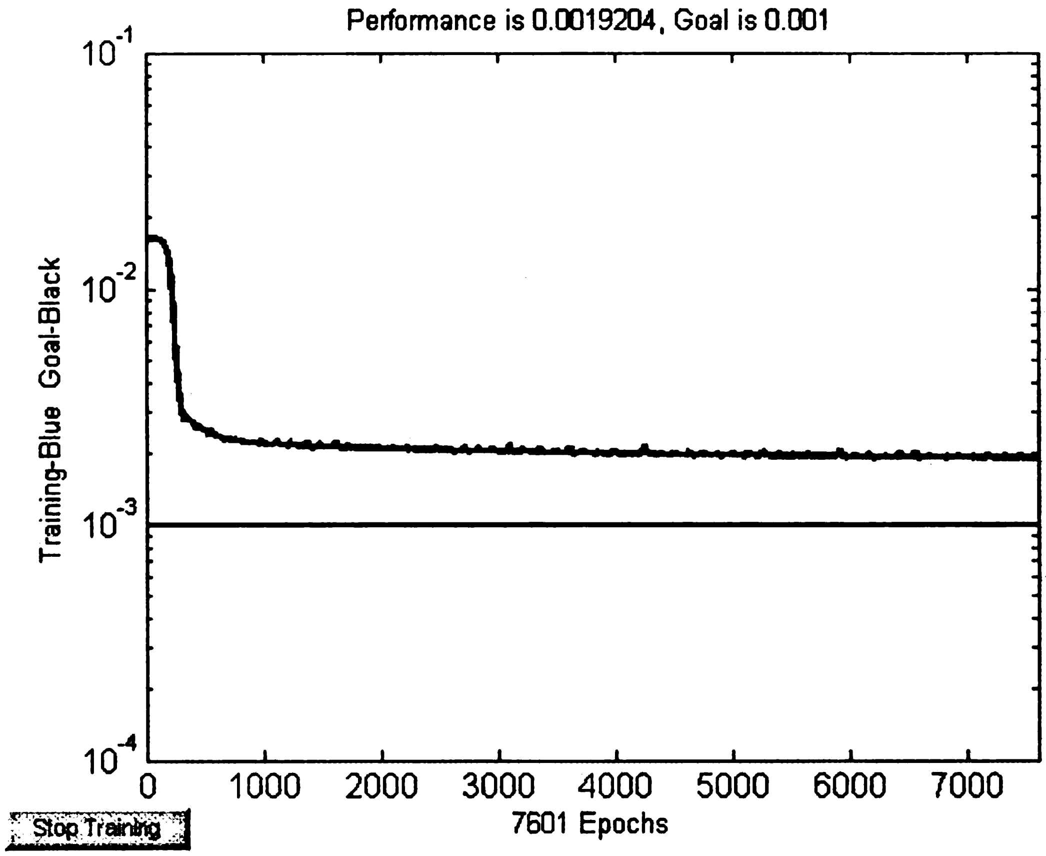

Figure 6.5 – A Sample Training Graph of State-Space Neural Network Model for this Study.

In closing, three CI-SSNN models were developed for the average travel time estimation. Similarly, three CI-SSNN models were developed for average travel time prediction. Hence, each of the three traffic movements has its own CI-SSNN models that provide the travel time value of the specific movement. Once the CI-SSNN models are developed for arterial links, they can be extended to any route that includes the selected arterial links. The results of the CI-SSNN models for arterial links are discussed in the next chapter.

6.6 Conditional Independence Graph based State-Space Neural Networks for Arterial Route

ATIS and ATMS applications are generally interested in travel time along an arterial route rather than an arterial link. After developing CI-SSNN models for travel time estimation and prediction for each type of traffic movement on arterial links, the models can extended to provide results for an arterial route. An arterial route which consists of a finite number of links included in the study can be specified. The CI-SSNN models which are developed for each type of traffic movement on arterial links can be used for travel time estimation and prediction for an arterial route. The travel time on an arterial route is simply the summation of travel time for links comprising that route taking into consideration the movement that is required to traverse that route.

The example arterial route which is used in this study is shown in Figure 6.6 with dark arrows on the urban arterial network. On this route, a vehicle starts its trip at origin intersection 16 and wants to get to destination intersection 25. The route that is taken is marked by turning movements required to traverse the route. The vehicle is considered as a through movement vehicle on arterial link 16-17, right-turning vehicle on arterial link 17-18, left-turning vehicle on arterial link 18-23, through movement vehicle on arterial link 23-24, and through movement vehicle on arterial link 24-25. So, the travel time estimation or prediction problem for this route uses the appropriate CI-SSNN models for the required traffic movements. The travel time on this route is simply the summation of the average travel time for the specified turning movements.

Hence, this study shows that if travel time for all three types of turning movements is modeled by CI-SSNN models, it can be easily implemented for any route in question. The results of implementation of CI-SSNN models for an arterial route are discussed in the next chapter.

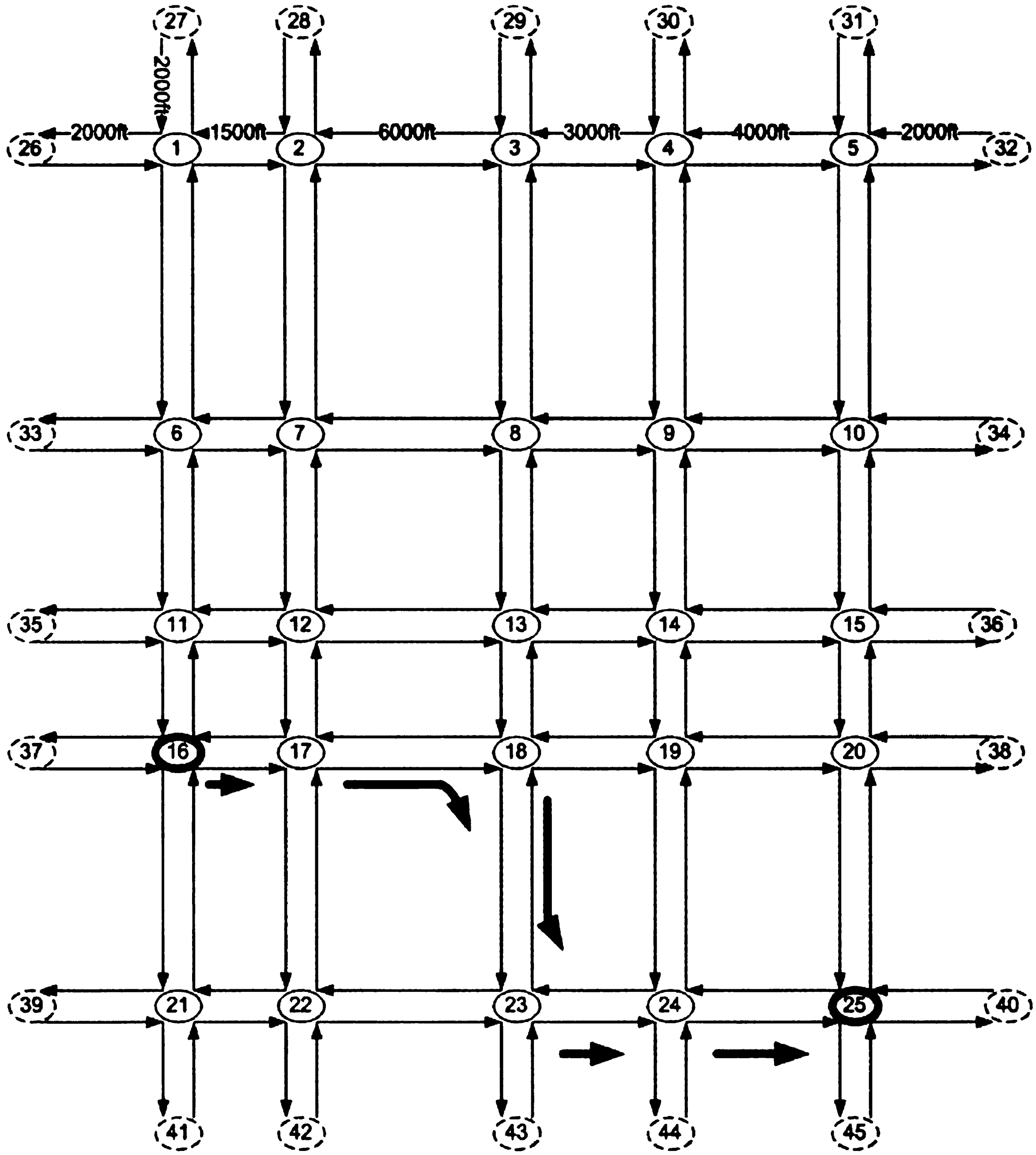

Figure 6.6 – An Example Arterial Route between Intersections 16 and 25 used for Travel Time Estimation and Prediction.

Chapter 7

Results and Discussions

This chapter presents the results of the State-Space Neural Network (SSNN) based on the Conditional Independence (CI) Graph (here, called as CI-SSNN model) for the arterial links and arterial route travel time estimation and prediction. The measures of performance used for analyzing the efficiency of the models are specified. The performance of the CI-SSNN models are presented for through, left, and right turning vehicles on an arterial link for travel time estimation and prediction. A comparison of SSNN model with and without using CI graphs shows the benefit of using the CI graphs in the modeling framework. The last section illustrates the performance of the models for travel time prediction on the example arterial route.

7.1. Measures of Performance

The measures of performance used for evaluating the accuracy and predictive performance of the CI-SSNN models are stated below (Kutner, et al. 2004 and Zhang and Rice 2003):

Square of Correlation Coefficient (R^2)-

A commonly used measure of performance is the coefficient of determination (R^2). The coefficient of determination is interpreted as the proportionate reduction of total variation associated with the use of predictor variable X (here, actual travel time). A larger R^2 value means that more the total variation of the dependent variable Y (here, predicted travel time) is reduced by introducing the predictor variable X.

A measure of linear association between Y and X when both Y and X are random is the coefficient of correlation (r). The coefficient of correlation is the square root of coefficient of determination with positive/negative value attributed to it (i.e. $\pm\sqrt{R^2}$). The positive/negative value determines whether the slope of regression line is positive or negative.

The square of the coefficient of correlation ($r^2 = (\pm\sqrt{R^2})^2 = R^2$) is chosen here as the measure of performance. It is expected that the scatter plot of the predicted travel time versus actual travel time should be a linear line with slope of one and intercept of zero. Thus, R^2 should be close to one in an ideal condition when the predicted travel time from the model is exactly the same as the actual travel time observed.

Mean Absolute Error (MAE) –

Let $\alpha(t, T_{PRED}(t)) = T_{ACTUAL}(t) - T_{PRED}(t)$ be the error of $T_{PRED}(t)$, where $T_{ACTUAL}(t)$ and $T_{PRED}(t)$ are the actual and the predicted average travel time at any time period t respectively. The Mean Absolute Error (MAE) is the expected value of the absolute value of the error, i.e. E $|\alpha(t)|$ over the sample space of the average travel times at all time-periods. Thus, MAE measures the average of the absolute prediction error of average travel time over the entire temporal spectrum on a link. The MAE of the training or testing set implies the MAE for all the links considered in the set. The value of MAE is specific to the magnitude and the range of the average travel time in this study and should be viewed in that context. It is not advisable to compare the MAE values from two test studies with different magnitudes and range of travel times.

Root Mean Square Error (RMSE) –

Let β (t, T_{PRED} (t)) = $(T_{ACTUAL}(t) - T_{PRED}(t))^2$ be the Square Error (SE) of T_{PRED} (t). The Mean Square Error (MSE) is the expected value of the Square Error, i.e. E $|\beta(t)|$ over the sample space of the average travel time at all time-periods. The Root Mean Square Error (RMSE) is the square root of the E$|\beta(t)|$ measuring the root mean square of the prediction error of the average travel time over the entire temporal spectrum on a link. The RMSE of the training or testing set implies the RMSE for all the links considered in the set. The RMSE values are also specific to this study like MAE.

Mean Absolute Percentage Error (MAPE) –

Let γ (t, T_{PRED} (t)) = $\dfrac{T_{ACTUAL}(t) - T_{PRED}(t)}{T_{ACTUAL}(t)} \times 100$ be the Percentage Error (PE) of T_{PRED} (t).

The Mean Absolute Percentage Error (MAPE) is the expected value of the PE, i.e.

$E|\gamma(t)|$ over the sample space of the average travel time at all time-periods. Thus, MAPE measures the absolute value of the percentage prediction error of average travel time over the entire temporal spectrum on a link. The MAPE of training or testing set implies the MAPE for all the links considered in the set. The MAPE is a unit free value unlike MAE and RMSE, and can be compared with the MAPE values obtained in other studies irrespective of the magnitude and the range of the average travel time.

7.2 Results of Travel Time Estimation on Arterial Links

The Conditional Independence (CI) graph analysis of average travel time for through, left, and right-turn movements resulted in selection of parent input variables. These parent input variables are fed as inputs to the State-Space Neural Network (SSNN) model for travel time estimation. Since, the SSNN model is based on parent input variables from CI graph, the resulting model are called the CI-SSNN model. The three different CI-SSNN models are proposed for average travel time estimation for through, left, and right-turn movements on arterial links, termed CI-SSNN-Thru, CI-SSNN-Left, and CI-SSNN-Right, respectively. These models learn the travel time pattern from the traffic conditions in a training set and are applied on different traffic conditions in a testing set.

The output of the CI-SSNN models is the estimated travel time for the current departure time-period for the training and testing sets. The utility of these models is to obtain the estimated travel time values which are then used as an additional input to the travel time prediction models

7.2.1 Results of CI-SSNN-Thru Travel Time Estimation Model for Through Movement

The results of the CI-SSNN-Thru model for travel time estimation of the training and testing sets are presented here. First, the correlation coefficient and goodness of fit between predicted travel time (model-generated estimated travel time) and actual travel time is presented in Figure 7.1 for the training set. The R^2 value is 0.9 which shows efficient learning of the model using the training set.

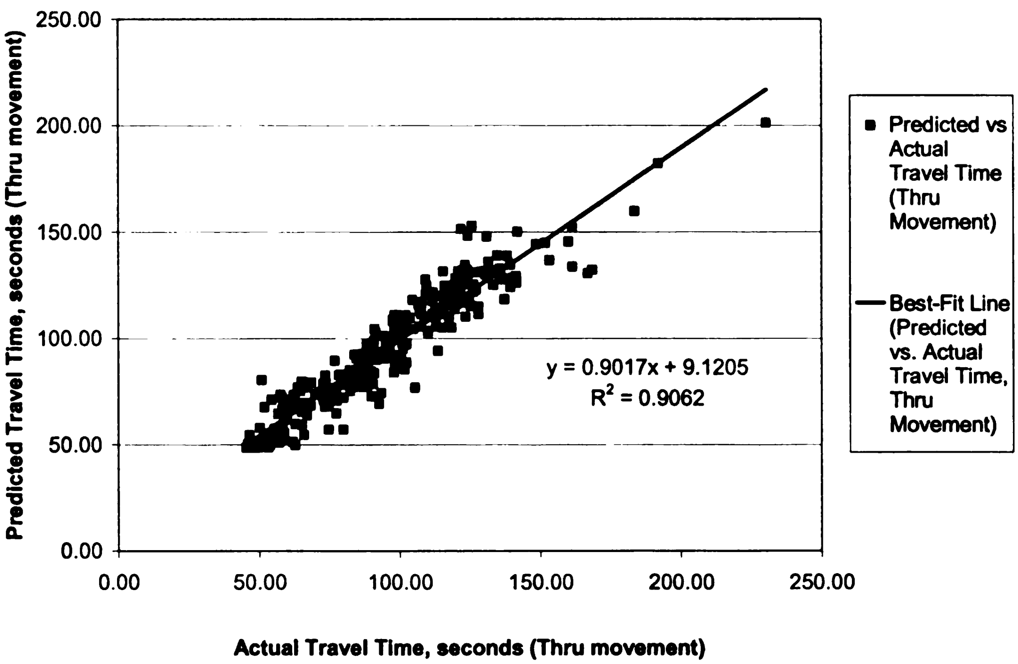

Figure 7.1 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Estimation for Through Movement.

The results in terms of performance measures like MAE, RMSE, and MAPE for the CI-SSNN-Thru model on arterial links comprising the training set is presented in Table 7.1.

The model based on the training set is providing good results with a Mean Absolute Percentage Error value of 7.6%.

Table 7.1 – Performance Measure of CI-SSNN-Thru Model for Travel Time Estimation for Through Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	6.6	8.4	7.6
16-20	6.3	8.4	7.2
21-25	8.1	11.4	7.6
3-23	6.9	9.1	8.1
Total Training Set	6.9	9.4	7.6

The model is then applied to the testing set which includes the same arterials but different traffic demand pattern. The correlation coefficient and scatter plot between predicted travel time (model-generated estimated travel time) and actual travel time is presented in Figure 7.2 for the testing set. The R^2 value of 0.9 is obtained which is the same as that of the training set. This shows that the CI-SSNN-Thru model performed with equal efficiency on both training and testing sets.

The results of the CI-SSNN-Thru model on testing set are presented in terms of performance measures in Table 7.2. The performance measure values for the testing set are comparable to that of the training set. This shows the robustness and efficiency of the CI-SSNN-Thru model for the average travel time estimation for through movement on an arterial link.

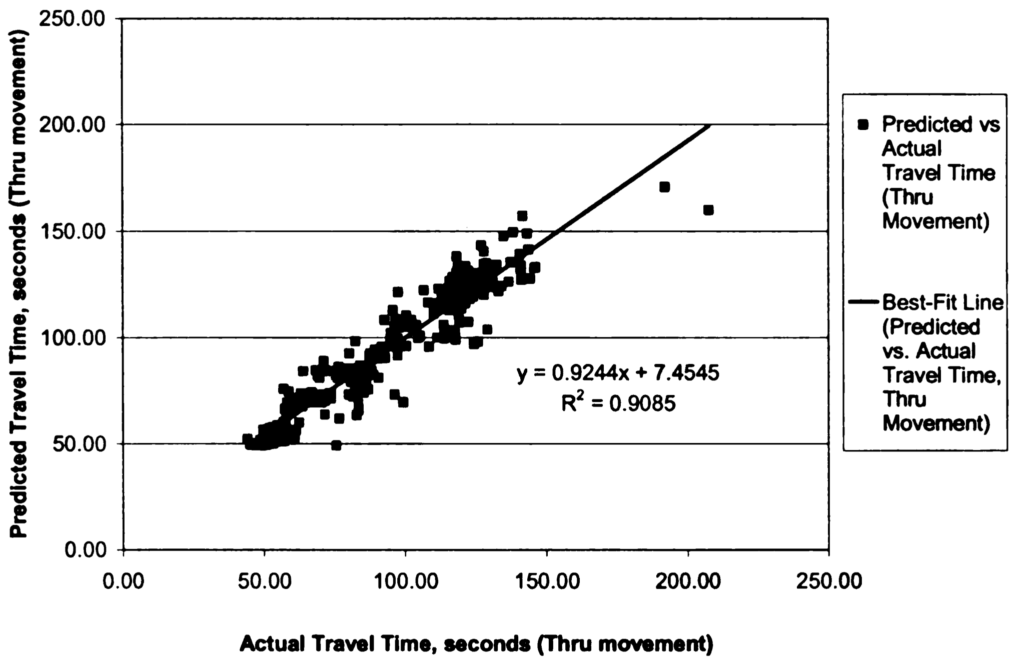

Figure 7.2 – Testing Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Estimation for Through Movement.

Table 7.2 – Performance Measure of CI-SSNN-Thru Model for Travel Time Estimation for Through Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	7.3	10.3	7.7
16-20	5.6	7.0	6.5
21-25	6.2	9.0	6.5
3-23	7.1	9.0	8.5
Total Testing Set	6.6	8.9	7.3

7.2.2 Results of CI-SSNN-Left Travel Time Estimation Model for Left-Turn Movement

The results of the CI-SSNN-Left model for travel time estimation using the training and testing sets are presented. The correlation coefficient and goodness of fit between the predicted travel time and actual travel time is presented in Figure 7.3 for the training set. The R^2 value is 0.87 which shows efficient learning of the model using the training set.

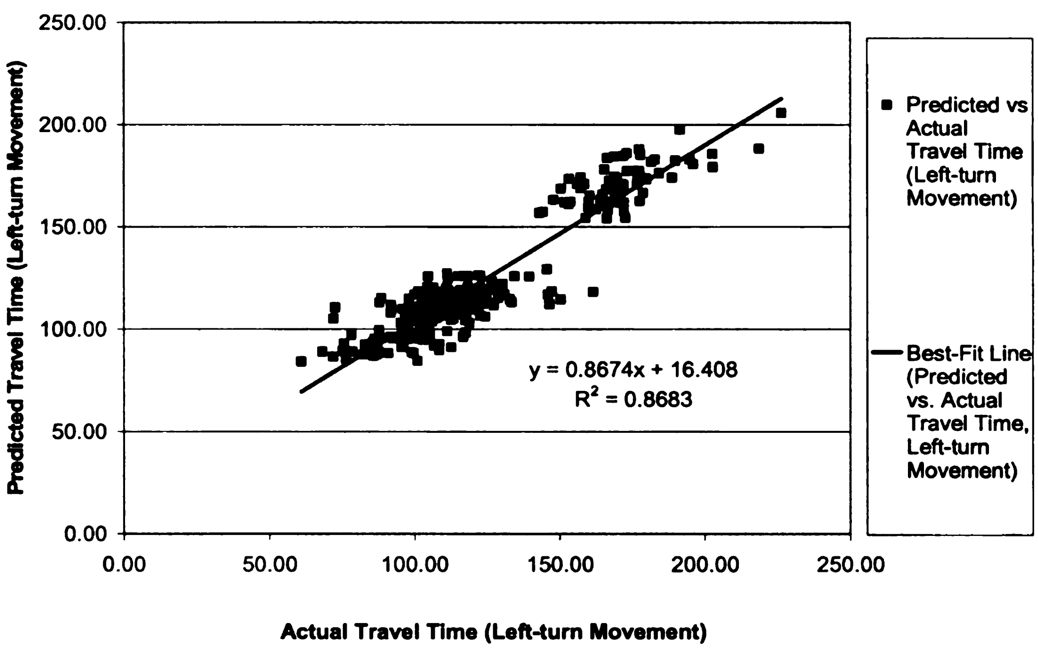

Figure 7.3 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Left-Turn Movement.

It is also noted in Figure 7.3 that the travel time values are falling in two clusters. One cluster has travel time values less than 150 seconds and the second cluster has travel time values greater than 150 seconds. The formation of these clusters pertains to the geometry and signal control scheme at the signalized intersections in the network. It was observed in the data that the travel time values for longer arterial links (link length of 6000 ft and 4000 ft) are generally greater than 150 seconds. The travel time values for the links having smaller length are lesser than 150 seconds. It is also to be noted that this cluster formation is not observed in the scatter plots of through movement (see Figures 7.1 and 7.2). This is because of the exclusive left-turn phasing with lead green that was adopted for the signal control in the network. The traffic conditions as viewed in the CORSIM microscopic simulation showed that majority of the left-turn vehicles could not clear the intersection during the lead left-turn phase in which they arrived at the intersection mostly during congested conditions. They had to wait for the next left-turn green in the next signal cycle. This increases the waiting time (and hence the travel time) for the left-turn traffic movement at signalized intersections in particular for longer links. This explains the higher travel time for left-turn traffic movement especially in longer arterial links.

The proposed CI-SSNN-Left model is expected to capture the different ranges of the travel time that pertains to geometrics and signal control phasing scheme. It will be observed in the later detailed discussion of the results of this model that the travel time was estimated with a similar efficiency for varied link lengths and signal control schemes. This strengthens the claim that the proposed travel time models are generalized and explicitly account for geometrics and signal control state of the arterial links.

The results in terms of performance measures like MAE, RMSE, and MAPE for the CI-SSNN-Left model on arterial links comprising the training set is presented in Table 7.3. The model using the training set provides good results with a Mean Absolute Percentage Error value of 7.5%.

Table 7.3 – Performance Measure of CI-SSNN-Left Model for Travel Time Estimation for Left-Turn Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	9.2	11.7	8.0
16-20	8.9	11.2	7.4
21-25	8.4	11.1	7.3
3-23	8.4	10.9	7.4
Total Training Set	8.7	11.2	7.5

The model developed based on the training set is then applied to the testing set. The correlation coefficient and scatter plot between predicted travel time (model-generated estimated travel time) and actual travel time is presented in Figure 7.4 for the testing set. The R^2 value is 0.72, which is lower a bit than that of the training set. One reason that is evident from the graph is that the model produces poor results when the actual travel time is greater than 250 seconds. Comparing Figures 7.3 and 7.4, it is also clear that none of the travel time value in the training set is more than 250 seconds. So, it is obvious that the CI-SSNN-Left model performed efficiently on both sets within the observed range of data-points in the training set. But, when the model was to extrapolate (because the large travel time values were not present during training), it did not perform well. Neural networks in general and also SSNN are not able to perform well in extrapolation. Thus, during application of this model, it is advisable that the training set should include the extreme boundary values, i.e. maximum and minimum travel time possible so that, the model does not have to extrapolate.

Another observation in Figure 7.4 is the presence of two distinct clusters of travel time values. This observation was also noted in Figure 7.3, where it was discussed. The same explanation holds well here also.

The results of the CI-SSNN-Left model on the testing set are presented in terms of appropriate performance measures in Table 7.4. The performance measure values for the testing set is comparable to that of the training set except arterial 6-10 where the travel time values are found to be more than 250 seconds. This shows the robustness and efficiency of the CI-SSNN-Left model for average travel time estimation for left-turn movements on an arterial link.

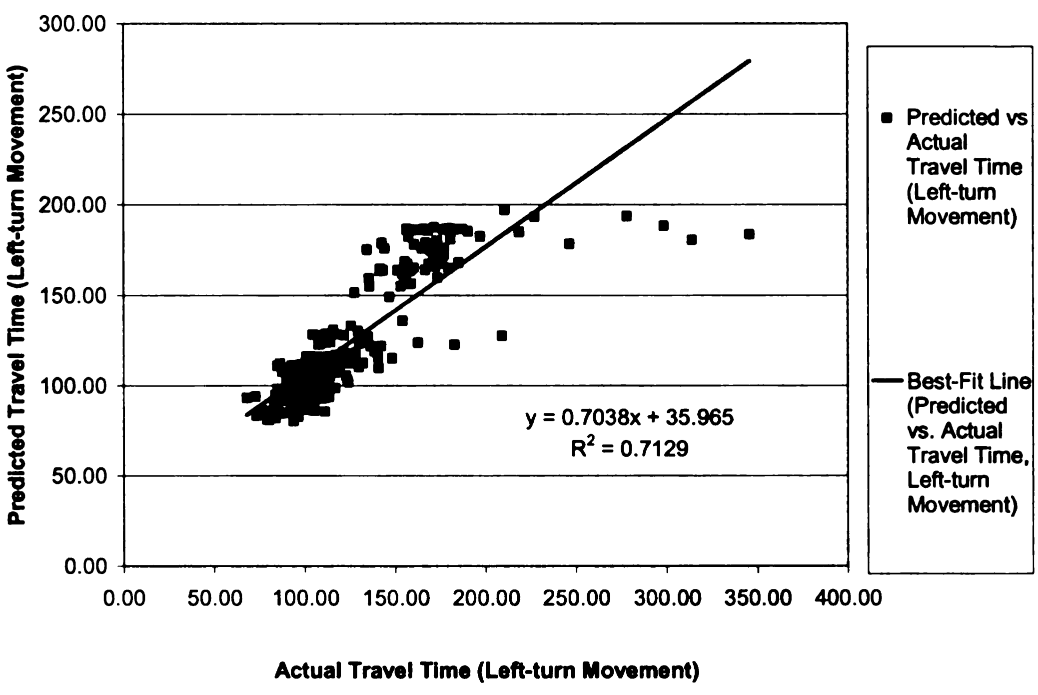

Figure 7.4 – Testing Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Estimation for Left-Turn Movement.

Table 7.4 – Performance Measure of CI-SSNN-Left Model for Travel Time Estimation for Left-Turn Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	16.8	32.4	10.6
16-20	9.0	11.9	7.8
21-25	10.3	13.2	8.2
3-23	9.7	15.8	7.6
Total Testing Set	11.5	20.1	8.6

7.2.3 Results of CI-SSNN-Right Travel Time Estimation Model for Right-Turn Movement

The results of the CI-SSNN-Right model for the travel time estimation using the training and testing sets are presented. The correlation coefficient and goodness of fit between predicted travel time (model-generated estimated travel time) and actual travel time is presented in Figure 7.5 for the training set. The R^2 value is 0.91 which shows efficient learning of the model based on the training set. Moreover, this R^2 value is comparable to that of the CI-SSNN-Thru model which is logical as through and right-turn movements share the same green interval at the intersection unlike left-turn vehicles which have separate green interval as well as a separate left-turn pocket which typically results in different queue lengths.

The results in terms of performance measures like MAE, RMSE, and MAPE for the CI-SSNN-Right model on arterials using the training set is presented in Table 7.5. The model based on the training set gives good results with a Mean Absolute Percentage Error value of 8.4%.

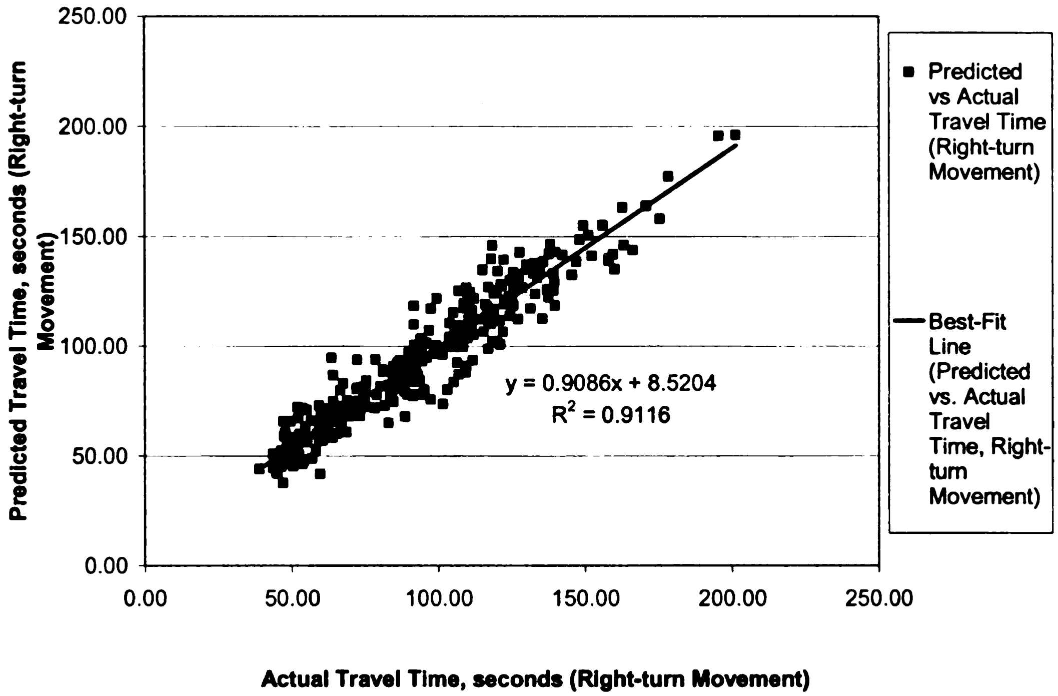

Figure 7.5 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Estimation for Right-Turn Movement.

Table 7.5 – Performance Measure of CI-SSNN-Right Model for Travel Time Estimation for Right-Turn Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	5.7	7.8	6.6
16-20	8.1	10.7	9.5
21-25	8.4	10.7	8.9
3-23	7.1	9.8	8.7
Total Training Set	7.3	9.8	8.4

The model produced using the training set is applied to the testing set. The correlation coefficient and scatter plot between predicted travel time (model-generated estimated travel time) and the actual travel time is presented in Figure 7.6 for the testing set. The R^2 value of 0.83 is obtained which is comparable to that of the testing set. But, this is still a satisfactory outcome taking into consideration that the testing set is a different traffic demand pattern which was not presented to the model while training.

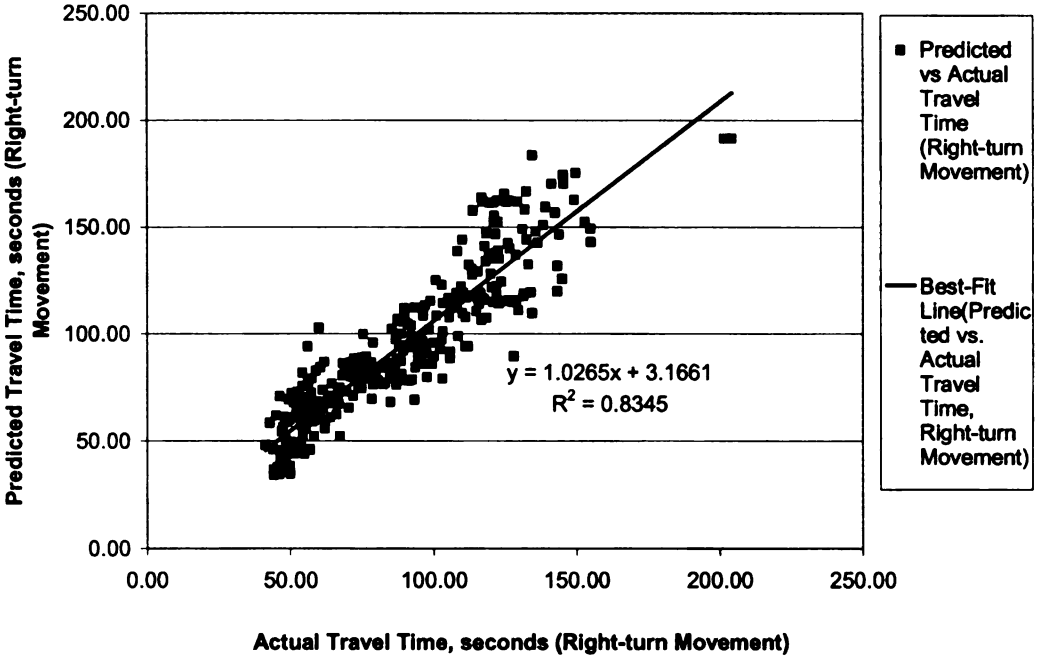

Figure 7.6 – Testing Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Estimation for Right-Turn Movement.

The results of the CI-SSNN-Right model using the testing set are presented in terms of performance measures in Table 7.6. The MAPE values for training and the testing set are 8.4% and 14.0%, respectively. This shows the robustness and efficiency of the CI-SSNN-Right model for average travel time estimation for right-turn movements on an arterial link.

Table 7.6 – Performance Measure of CI-SSNN-Right Model for Travel Time Estimation for Right-Turn Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	11.9	15.5	13.4
16-20	11.3	13.6	14.7
21-25	13.8	18.0	15.0
3-23	10.4	13.8	13.1
Total Testing Set	11.8	15.3	14.0

The above results give satisfactory values of the estimated average travel time for current time-period for through, left, and right-turn movements on an arterial link. The estimated travel time for the current departure time-period is added to the parent input variables used for travel time prediction. The next section deals with the average travel time prediction for through, left, and right-turn movements.

7.3 Comparison of Results for State-Space Neural Networks with and without using Conditional Independence Graphs

The travel time estimation done in the earlier section is based on a CI-SSNN model, i.e. SSNN model combined with CI graph. It is evident that use of CI graph reduces the number of input variables required for modeling of travel time. It is still to be proven that SSNN based on CI graph produces better or comparable results to that of an ordinary SSNN model that is typically used by researchers.

Here, travel time estimation is done by using an ordinary SSNN model without using the CI graph. So, the input vector consists of all traffic variables: V_T, V_L, $V_{R,}$ S_T, S_L, S_R, G, *Off, S.L.*, ITT_T and ITT_L. These input variables are fed to a SSNN-Thru model for estimation of average travel time for through movement. The results are compared with the CI-SSNN-Thru model which was presented in Section 7.2.1. If the results of the CI-SSNN-Thru model are better or comparable to those of the SSNN-Thru model, the use of the CI graph for improving the efficiency of SSNN models is justified and recommended. The results obtained for travel time estimation for through movement using the training and testing sets are presented in Tables 7.7 and 7.8, respectively.

Table 7.7 – Performance Measure of SSNN-Thru Model (State-Space Neural Networks without using Conditional Independence Graphs) for Travel Time Estimation for Through Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	5.7	7.9	6.0
16-20	6.2	8.2	7.1
21-25	7.1	10.2	6.0
3-23	6.6	9.0	7.6
Total Training Set	6.4	8.8	6.8

Table 7.8 – Performance Measure of SSNN-Thru Model (State-Space Neural Networks without using Conditional Independence Graphs) for Travel Time Estimation for Through Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	9.3	11.9	11.0
16-20	6.6	8.7	8.4
21-25	4.7	6.3	5.2
3-23	7.0	9.0	8.5
Total Testing Set	6.9	9.2	8.2

A comparison of Tables 7.7 and 7.8 versus Tables 7.1 and 7.2 does not show much difference between results of the two sets of models. However, the MAPE values for the training set is lower in the SSNN-Thru model, which may suggest that inclusion of all input variables leads to a better learning for the SSNN model. But, the MAPE value of the testing set is higher for the SSNN-Thru model which may suggests that though learning is better, the SSNN-Thru model suffers when tested on a different traffic demand pattern.

The CI-SSNN models use lesser variables than the SSNN models and hence the CI-SSNN models can easily be declared efficient on this practical basis only. The constraint of limited amount of data obtained in any experiment or in the field makes the CI-SSNN models more suitable than SSNN models. Hence, it is recommended that a CI graph analysis be done prior to using artificial neural networks in general in a process that has multivariate normal distribution. The comparison done here serves the purpose for travel time estimation and prediction modeling. The further research in this topic is needed to theoretically and experimentally draw a sound conclusion.

The comparison of the CI-SSNN and SSNN models is shown for through movement only but the conclusion can be extended to modeling for left and right-turn movements also. Hence, travel time prediction on arterial links in the following sections is done using CI-SSNN models only, and it is assumed that the CI-SSNN model will provide better or comparable results to those of the SSNN modeling.

7.4 Results of Travel Time Prediction on Arterial Links

The three different CI-SSNN models are proposed for average travel time prediction for through, left, and right-turn movements (hereby, termed as CI-SSNN-Thru, CI-SSNN-Left, CI-SSNN-Right, respectively) on arterial links. These models for travel time prediction work similar to the travel time estimation models. The only difference is that the output in this case is the average travel time for a future departure time-period instead of the current departure time-period. The parent input variables for average travel time prediction contains the parent input variables used in the travel time estimation models developed earlier. The estimated travel time from the travel time estimation models is also included as an additional input variable for the travel time prediction modeling. The prediction models learn the travel time pattern from the traffic conditions contained in a training set and are then applied to entirely different traffic conditions in a testing set.

7.4.1. Results of CI-SSNN-Thru Travel Time Prediction Model for Through Movement on Training Set

The complex non-linear relationship of average travel time for through movement and the parent input variables is approximated by the CI-SSNN-Thru model. The predicted average travel time (model generated) is plotted against the actual travel time (as obtained from the CORSIM simulation) in Figure 7.7. The R^2 value of 0.87 shows that the proposed CI-SSNN-Thru model is able to learn well from the training set.

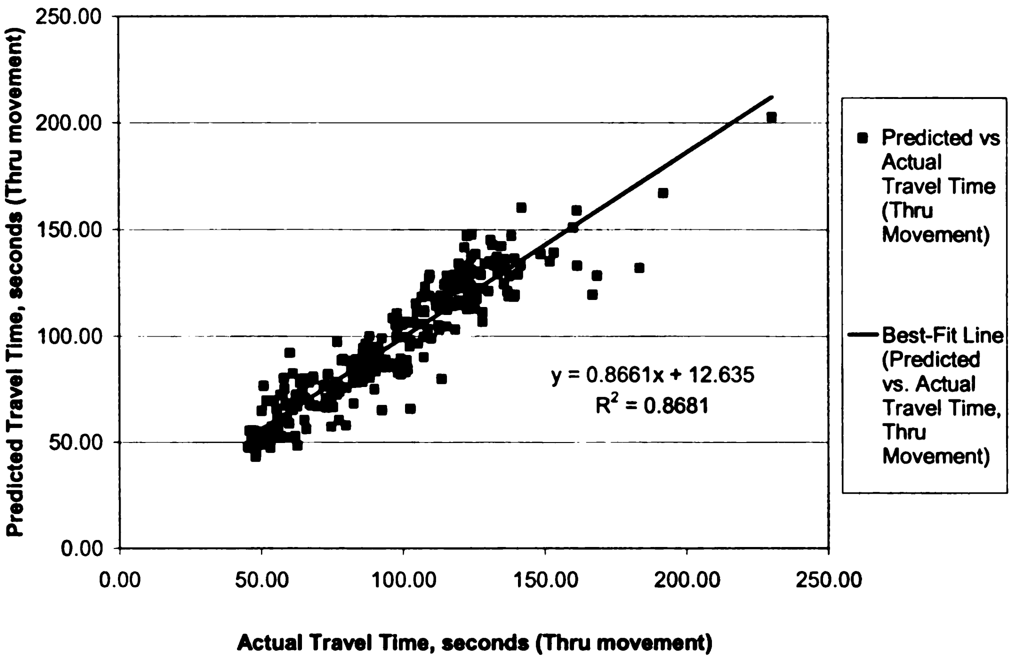

Figure 7.7 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Through Movement.

A detailed analysis of the learning and performance of the CI-SSNN-Thru model is needed for the training set. The detailed analysis enables to analyze the model results with respect to each arterial link contained in the study. The Absolute Percentage Error (APE) is plotted for all of the four arterials whose data was fed into the model. The plot of APE against departure time period (5 minutes to 95 minutes on each link) for each arterial is presented. Figures 7.8, 7.9, 7.10, and 7.11 show these plots for training sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23 respectively. The APE values are mostly lesser than 15% for each individual link in the arterials.

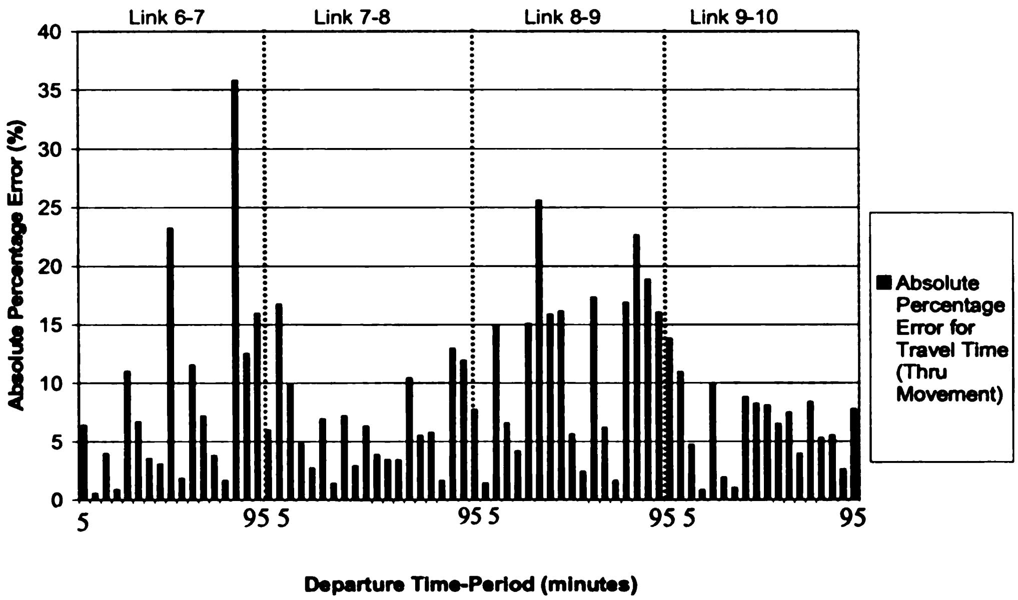

Figure 7.8 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Training Set 1.

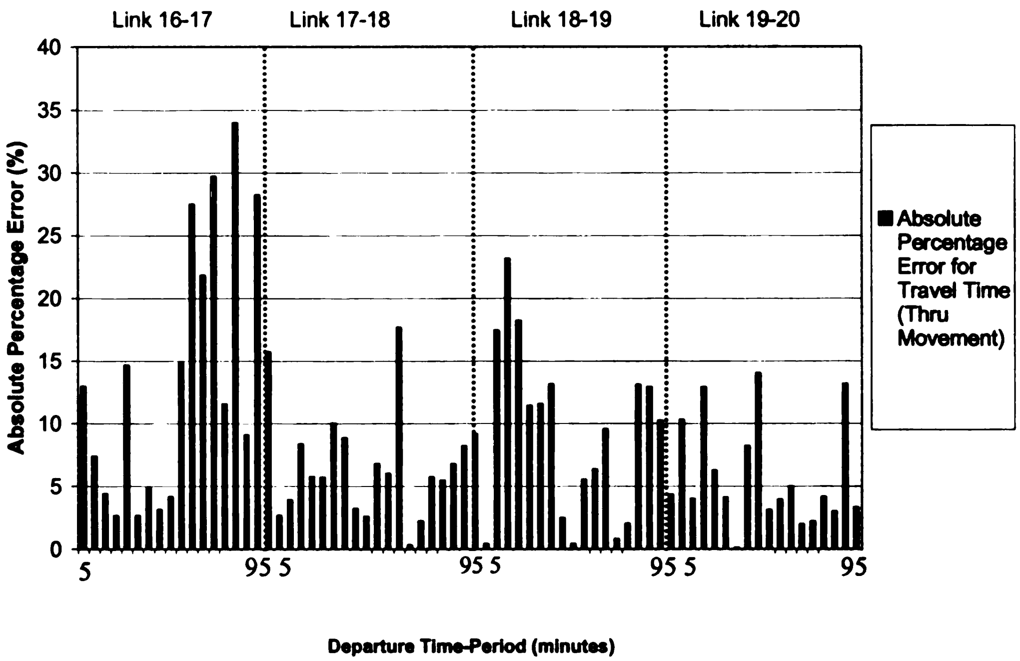

Figure 7.9 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Training Set 2.

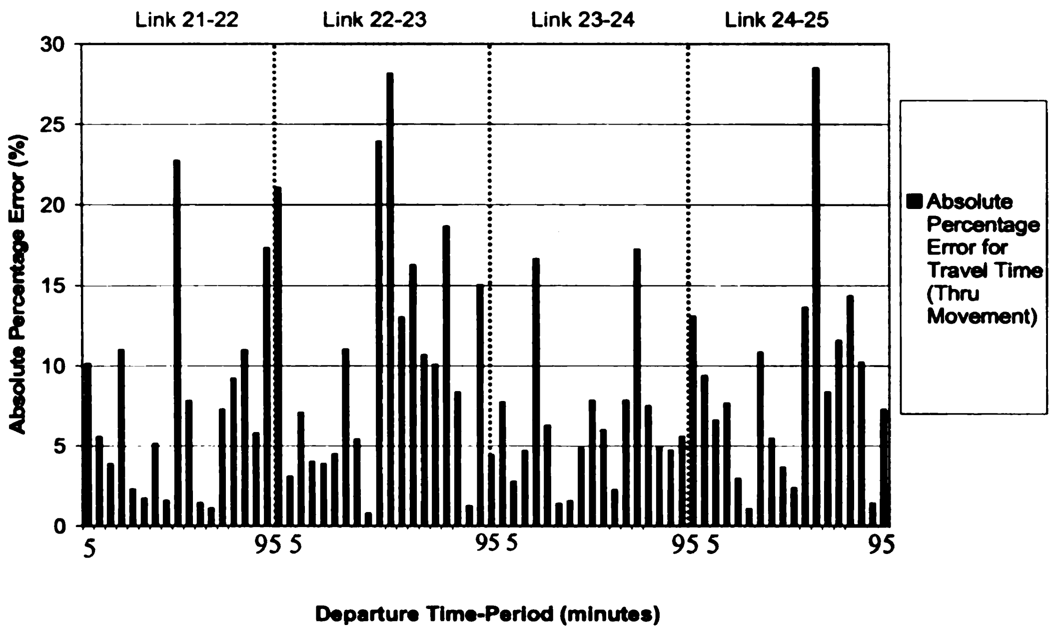

Figure 7.10 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Training Set 3.

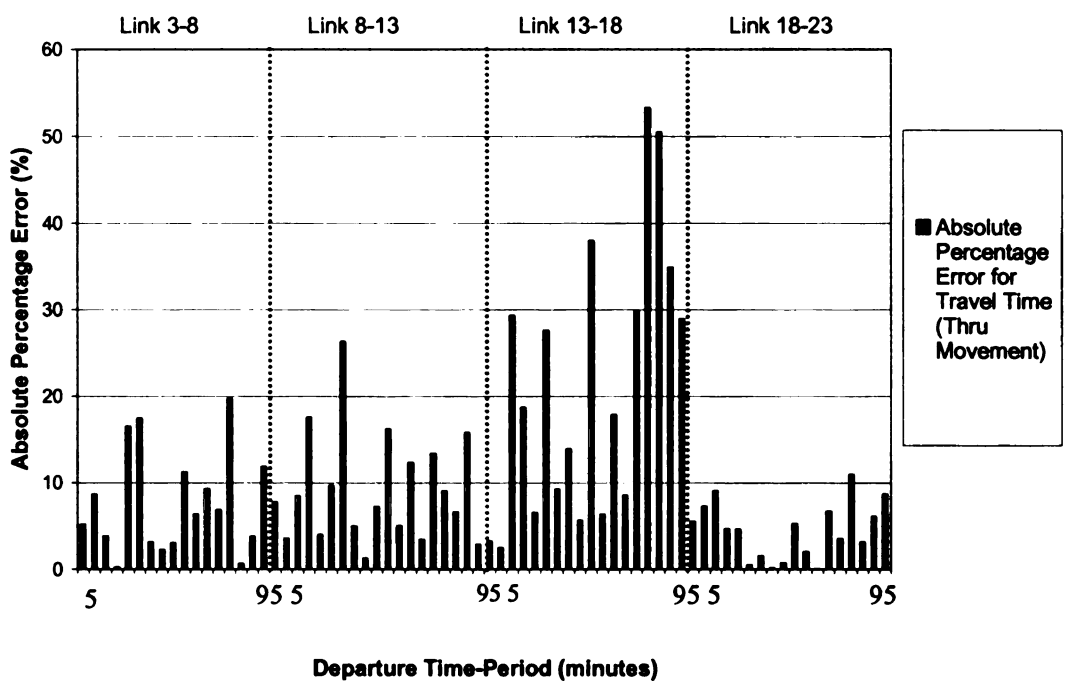

Figure 7.11 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Training Set 4.

It is also noted that the CI-SSNN-Thru model is not able to learn well for some links as depicted from their APE distribution. Links 6-7, 8-9, 16-17, 22-23, 24-25 and 13-18 are remarkable in this aspect. Though, it can not be inferred that the relatively high APE distribution on these links is because these links are at a particular location that they act as a critical link for an arterial. The critical links are those links which have the least capacity in an arterial or a sequence of links and hence may act as bottlenecks. So, the relatively high APE distribution of such links may be due to the inherent stochastic nature of the CORSIM simulation. If there is some unidentified characteristic of a link or traffic stream which is not explicitly accounted in the model formulation, the APE distribution of the testing set for the same links should be relatively high as compared to other links. This argument is revisited in Section 7.4.2., when the testing sets are analyzed.

In conclusion of this section, the measures of performance like MAE, RMSE, and MAPE of the CI-SSNN-Thru model for the training set are presented in Table 7.9. The model gives MAPE value of 9% for the training set. It is expected that this model can generalize and is able to predict the average travel time for through movements when new traffic demand conditions in the testing set are presented.

Table 7.9 – Performance Measure of CI-SSNN-Thru Model for Travel Time Prediction for Through Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	7.6	10.0	8.4
16-20	7.7	9.7	8.8
21-25	8.9	13.2	8.4
3-23	9.0	11.8	10.7
Total Training Set	8.3	11.2	9.0

7.4.2. Results of CI-SSNN-Thru Travel Time Prediction Model for Through Movement on Testing Set

After analyzing the efficiency of the CI-SSNN-Thru model to learn from the training set, the model is applied to the testing set. It is expected that the model should perform with similar efficiency on the traffic pattern of testing set.

The predicted average travel time is plotted against the actual travel time for testing set in Figure 7.12. The R^2 value is obtained as 0.88 which is close to the one obtained for the training set. Thus, the CI-SSNN-Thru model has generalization ability so that when a different traffic demand pattern is presented to the model, it predicts the average travel time with comparable accuracy. This property is a necessary requirement for a reliable prediction model that can be applied in the field.

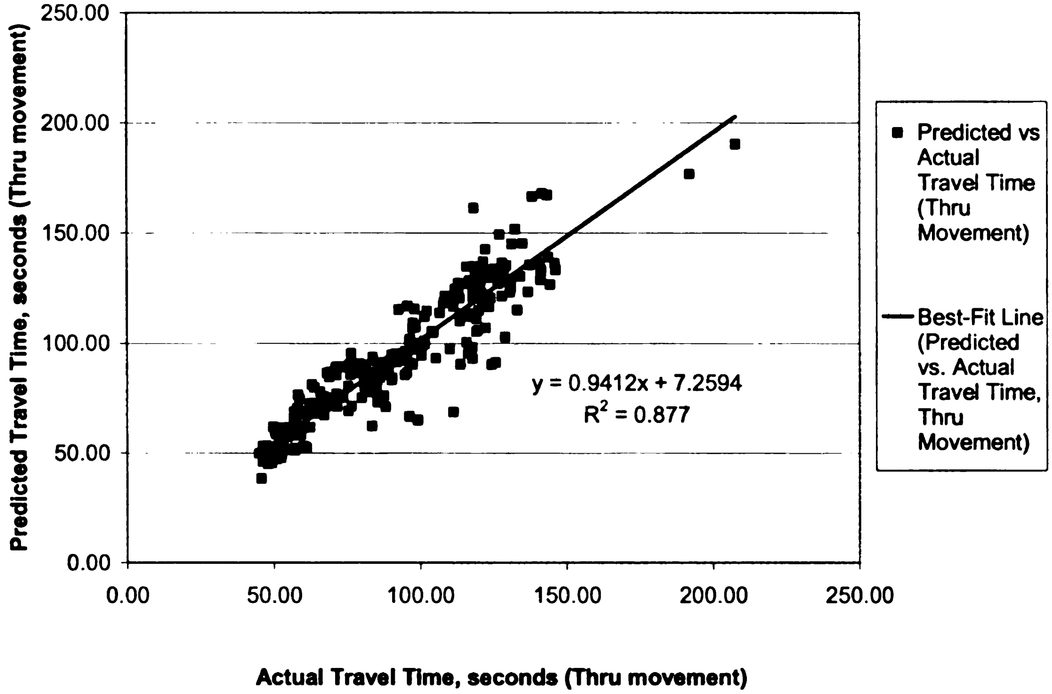

Figure 7.12 – Testing Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Through Movement.

A detailed analysis of the predictive performance of the CI-SSNN-Thru model is needed for the testing set to reinforce the claims made in the earlier paragraph. The Absolute Percentage Error (APE) is plotted for all four arterials that constitute the testing set. The plots of the APE against future departure time-period (5 minutes to 90 minutes on each link) for each arterial is presented. Figures 7.13, 7.14, 7.15, and 7.16 show these plots for training sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23, respectively. These plots are consistent with the training set plots in the sense that the APE values are mostly lesser than 15% for each individual links in an arterial.

The predictive performance of the proposed CI-SSNN-Thru model is relatively weaker for some links where the APE values for some time-periods is as high as 35%. These links are: 6-7, 7-8, 18-19, 23-24, 24-25 and 13-18. It is recalled (as stated in Section 7.4.1) that some links in the training set like 6-7, 8-9, 16-17, 22-23, 24-25 and 13-18 had relatively larger APE values. Though, all these links do not perform poor in the testing set as was expected earlier. So, the large values of APE in training and testing sets can be attributed to the stochastic nature of the CORSIM simulation as stated earlier also in Section 7.4.1.

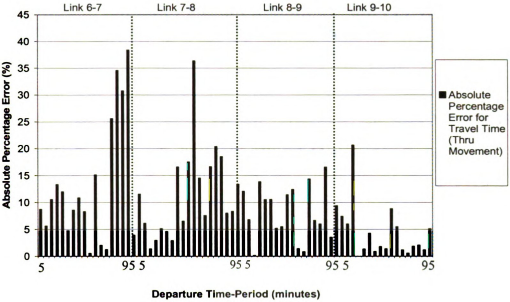

Figure 7.13 – Absolute Percentage Error of Predicted Travel Time vs. Departure Time-Period for Through Movement applied on Testing Set 1.

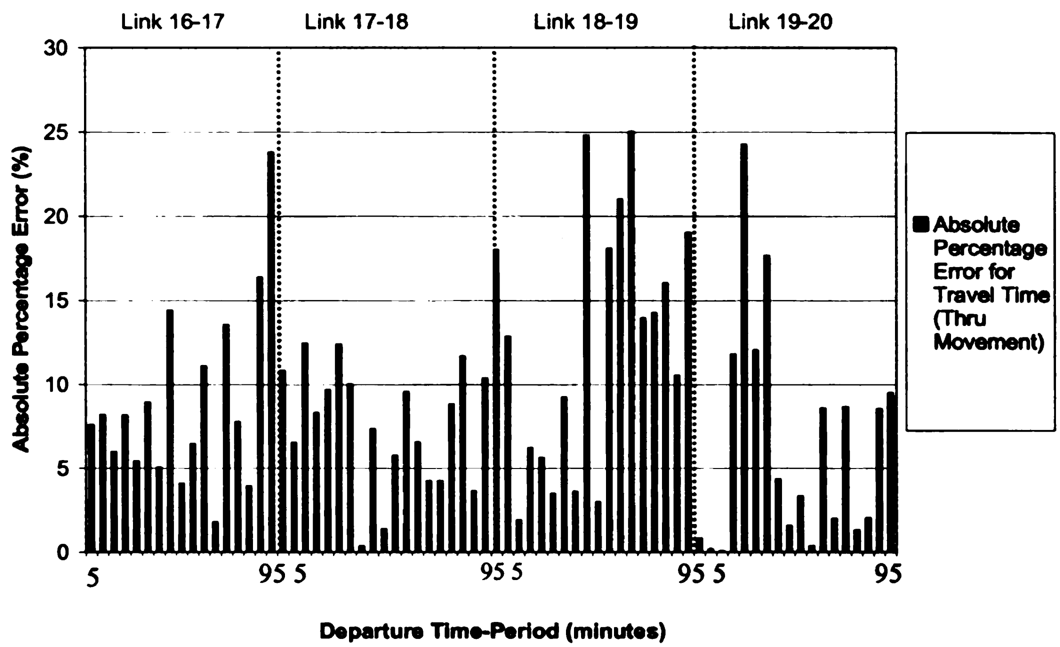

Figure 7.14 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Testing Set 2.

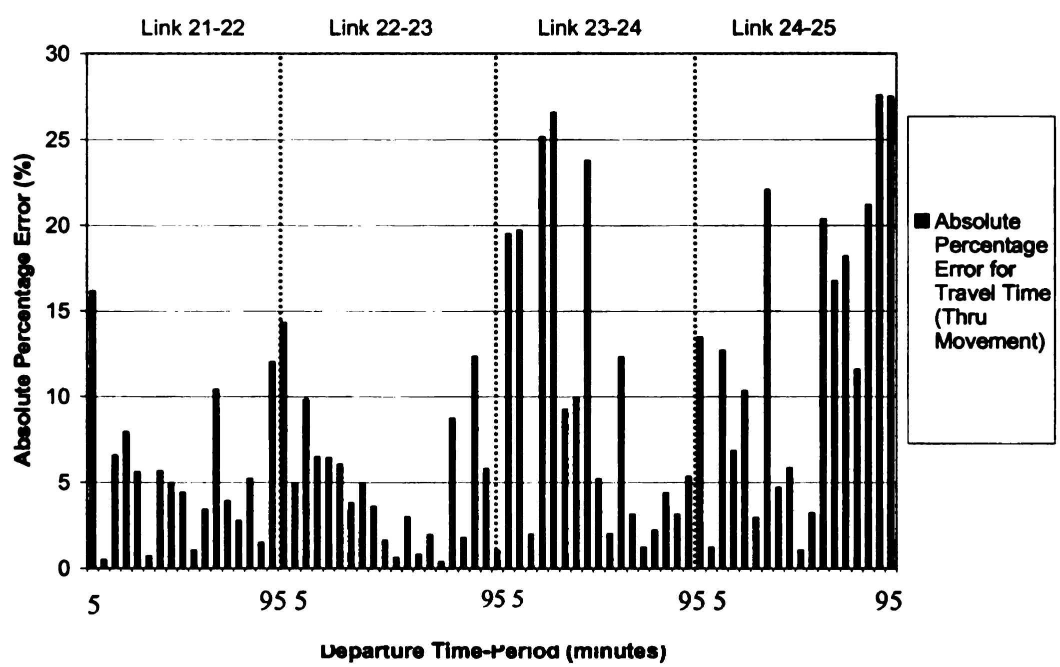

Figure 7.15 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Testing Set 3.

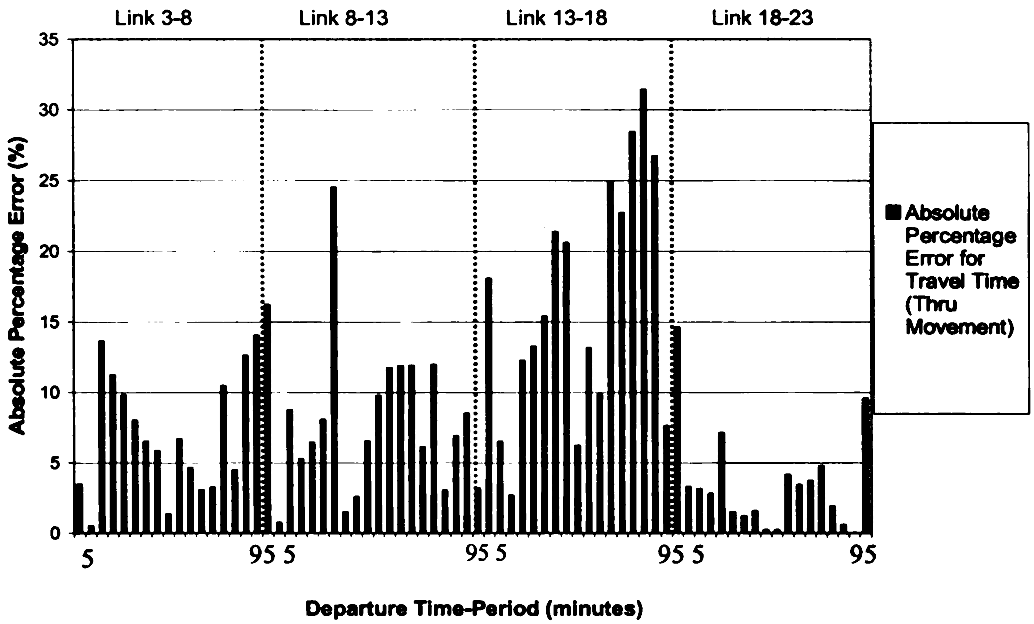

Figure 7.16 – Absolute Percentage Error of Predicted Travel Time versus Departure Time-Period for Travel Time Prediction for Through Movement on Testing Set 4.

The patterns of the actual and predicted average travel time for the testing set are plotted to see how closely the predicted values follow the actual travel time fluctuation. Figures 7.17, 7.18, 7.19, and 7.20 show the actual and predicted travel time patterns where future departure time-period is on the X axis. It is seen that the predicted travel time closely follows the actual travel time for all the arterial links except for link 24-25 shown in Figure 7.19. The reason because of this variation between travel time values is not certain as a similar variation was not observed during the training of the model (see Figure 7.10). It is already mentioned that the links that give poor results during training do not perform poor during testing of the model. Hence, one explanation could be the stochastic nature of the CORSIM simulation.

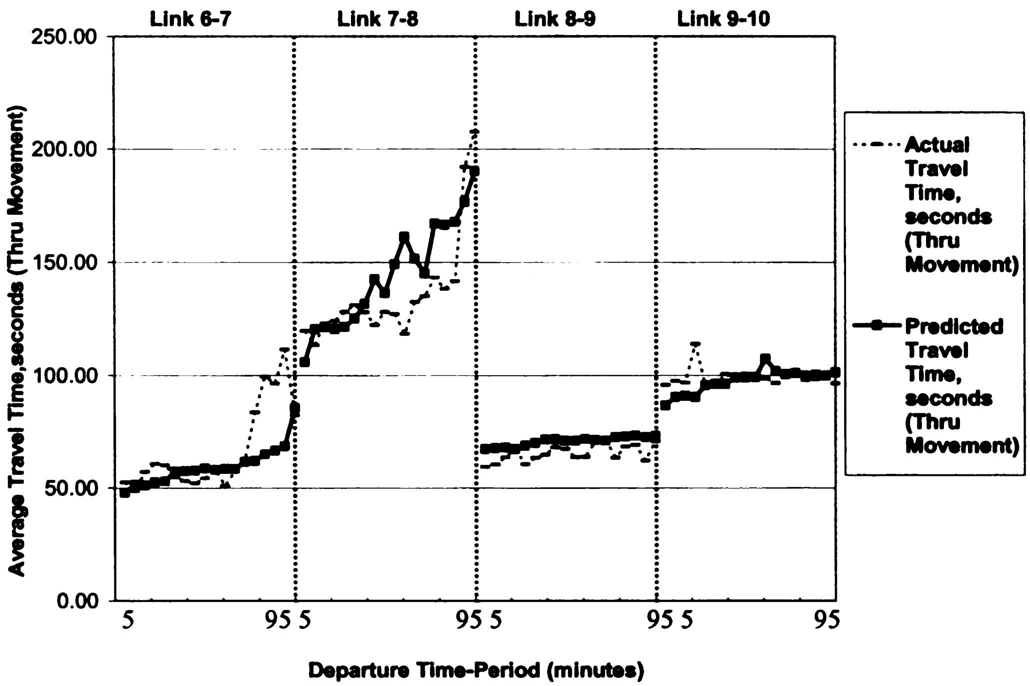

Figure 7.17 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Through Movement on Testing Set 1.

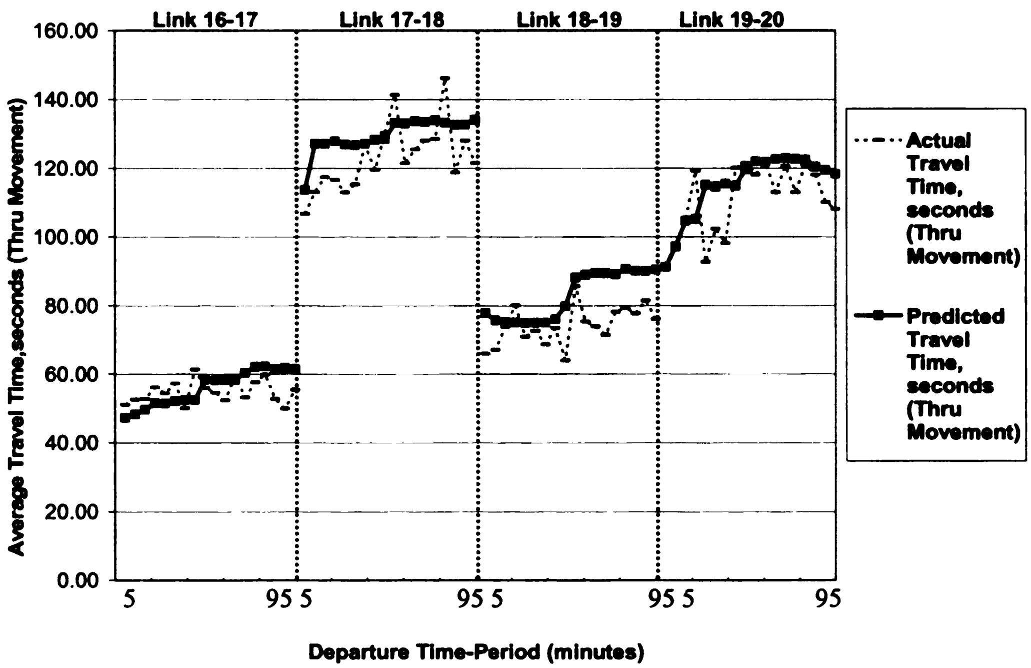

Figure 7.18 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Through Movement on Testing Set 2.

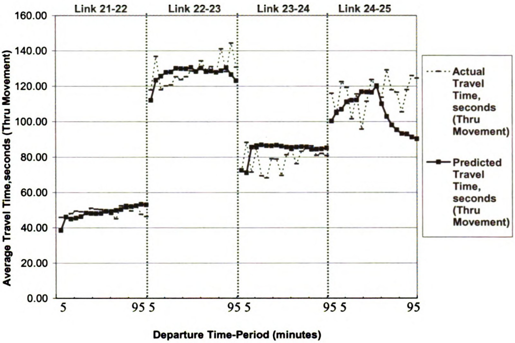

Figure 7.19 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Through Movement on Testing Set 3.

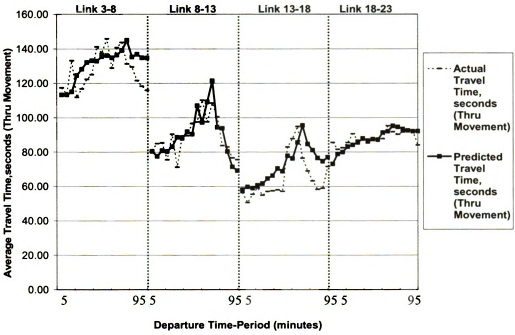

Figure 7.20 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Through Movement on Testing Set 4.

The following two inferences can be made about CI-SSNN-Thru model from these figures:

1. The predicted average travel time lies in the same range as the actual average travel time for all arterial links. This confirms that the CI-SSNN-Thru model is a generalized model which can predict the average travel time within the range that is typical of the given link. The range of average travel time for any link in general is a function of variables like link length, and speed limit.
2. Close inspection of travel time patterns at each arterial link shows that the predicted average travel time follows the actual average travel time pattern. It shows that the CI-SSNN-Thru model captures the traffic dynamics and temporal fluctuations in travel time on each link. This traffic dynamics and temporal nature of average travel time follows from the varying traffic demand on an arterial link throughout the day. The model is able to capture this aspect because of its topology. The hidden and context layers in the SSNN topology provide a short-term associative memory in the network in addition to the input-output mapping feature inherent in artificial neural networks.

In conclusion of this section, the measures of performance like MAE, RMSE, and MAPE of the CI-SSNN-Thru model for testing set are presented in Table 7.10. The MAE, RMSE, and MAPE values in this table are comparable with those obtained for the training set. This proves the generalization ability of the CI-SSNN-Thru model for short-term average travel time prediction for through movements on arterial links.

Table 7.10 – Performance Measure of CI-SSNN-Thru Model for Travel Time Prediction for Through Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	8.9	13.2	9.3
16-20	7.5	9.0	8.8
21-25	7.7	11.1	8.2
3-23	7.5	9.2	8.7
Total Testing Set	7.9	10.7	8.8

7.4.3 Results of CI-SSNN-Left Travel Time Prediction Model for Left-Turn Movement on Training Set

The CI-SSNN-Left model is trained using the training set. The efficiency of the model is analyzed in similar lines with that of the CI-SSNN-Thru model presented in earlier sections. The R^2 value of the predicted versus actual average travel time shown in Figure 7.21 is 0.87. The Figure 7.21 also shows two distinct clusters of travel time values. This observation was also made in Section 7.2.2 where the travel time estimation modeling for left-turn movements on arterial links was presented. The explanation given there holds good here also. The APE distribution for training sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23, is presented in Figures 7.22, 7.23, 7.24, and 7.25, respectively. The APE distribution is mostly within 15% error range. The arterial links that have APE values more than 15% to 45% are 16-17, 19-20, 21-22, 24-25, and 13-18. The detailed analysis of the results is done by calculating the performance measures shown in Table 7.11. The overall MAPE value of 7.3% shows that the CI-SSNN-Left model is able to learn well the temporal fluctuations in predicted values of average travel time.

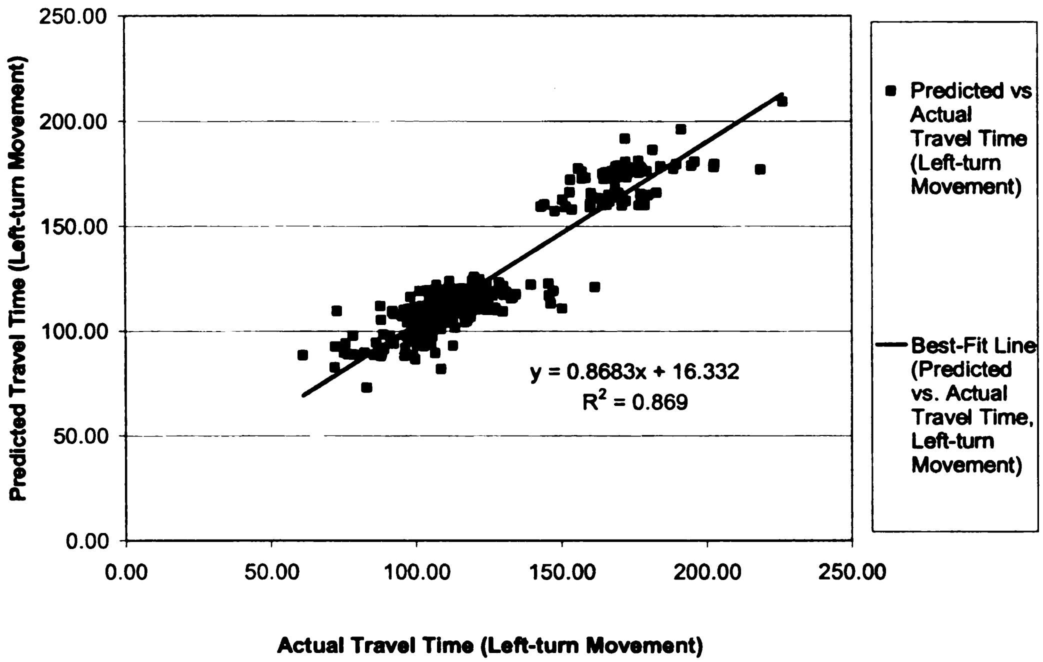

Figure 7.21 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Left-Turn Movement.

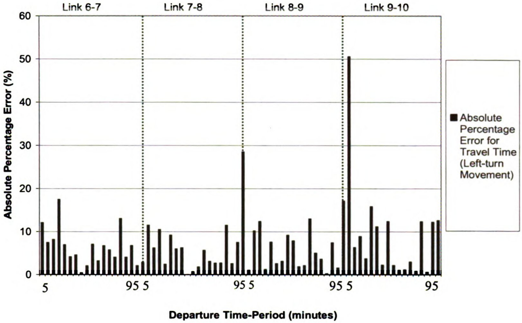

Figure 7.22 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Training Set 1.

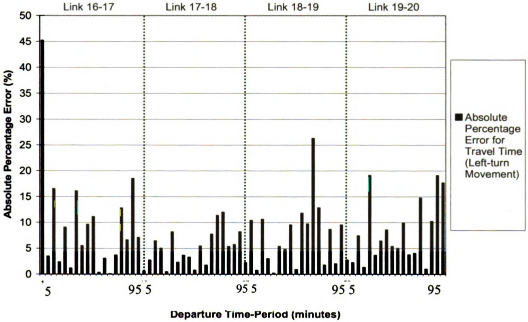

Figure 7.23 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Training Set 2.

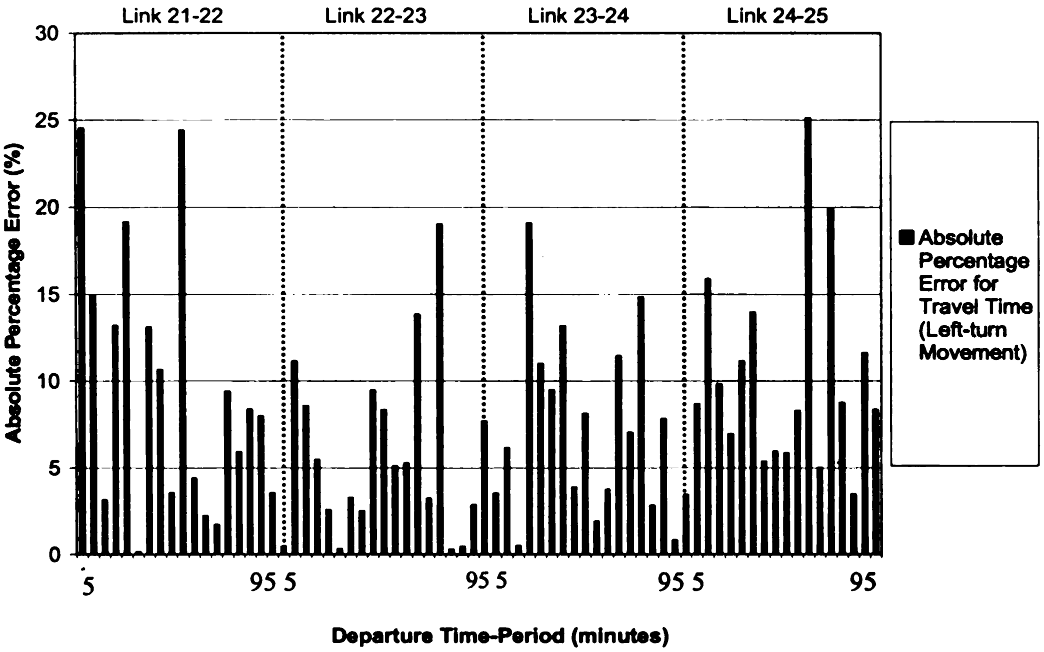

Figure 7.24 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Training Set 3.

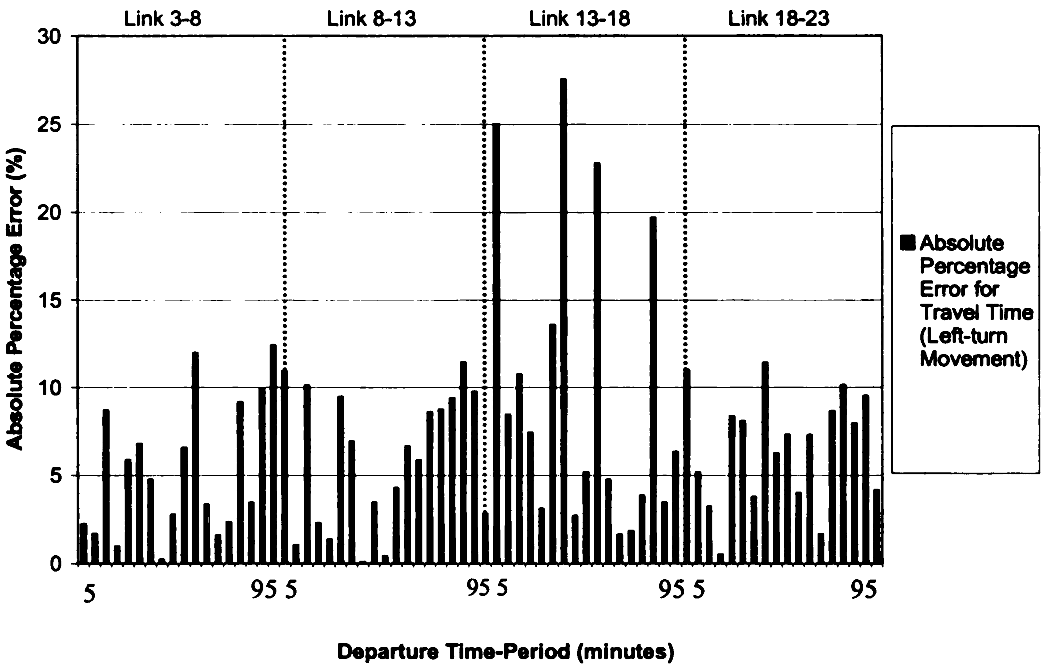

Figure 7.25 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Training Set 4.

Table 7.11 – Performance Measure of CI-SSNN-Left Model for Travel Time Prediction for Left-Turn Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	8.3	10.7	7.0
16-20	8.8	11.6	7.4
21-25	9.5	12.4	8.1
3-23	8.2	10.1	6.9
Total Training Set	8.7	11.2	7.3

7.4.4 Results of CI-SSNN-Left Travel Time Prediction Model for Left-Turn Movement on Testing Set

The CI-SSNN-Left model is now applied on a testing set which has different travel time patterns than those presented in the training set. The R^2 value for a plot between the predicted versus actual average travel times as shown in Figure 7.26, is 0.73. This value of R^2 for the testing set is lower than that of the training set. One reason of the lower R^2 value is attributed to the inability of artificial neural networks to extrapolate. This factor was acknowledged in Section 7.2.2 also where the CI-SSNN-Left model results for travel time estimation were discussed. It is expected that if training and testing sets have similar travel time ranges, this model's accuracy can be significantly improved.

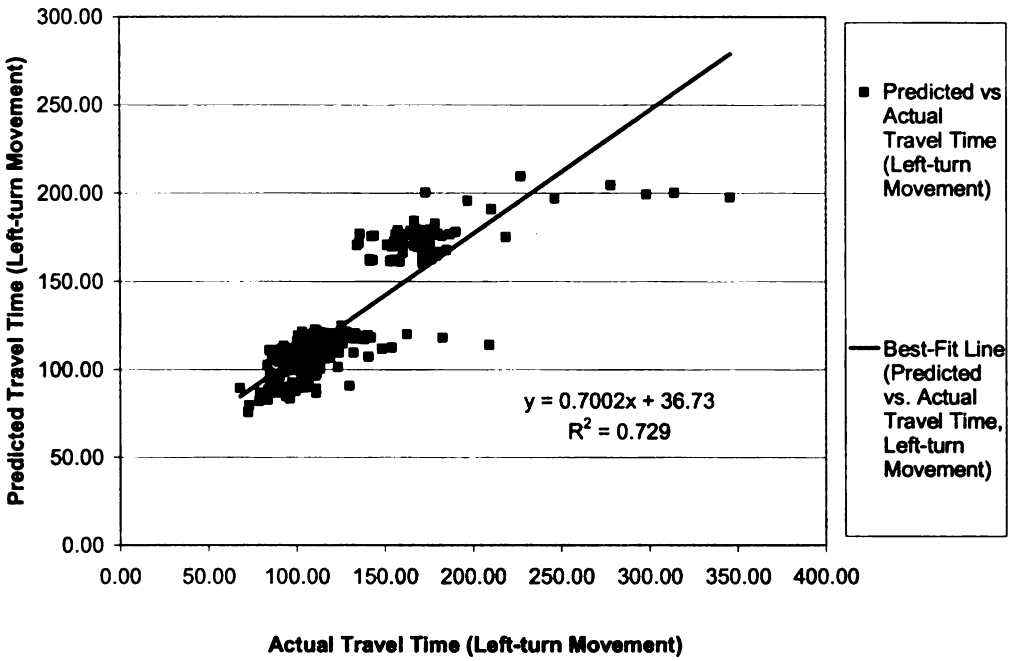

Figure 7.26 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Left-Turn Movement.

The APE plots for testing sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23, respectively, are presented in Figures 7.27, 7.28, 7.29, and 7.30. These plots show that APE distribution is mostly less than 15% error, except for links 7-8, 9-10, 16-17, 19-20, 24-25, 13-18, and 18-23. The plot of predicted and actual travel time patterns with departure time-periods for different arterials is presented in Figures 7.31, 7.32, 7.33, and 7.34. It is seen that the predicted travel time follows the actual travel time fluctuations in most of the arterials in the testing set. Also, predicted travel time lies within the same range as that of the actual travel time. It is also observed that there is a large error for some future departure time-periods as shown in Figure 7.31 and 7.34. The reason is the same as mentioned earlier that the artificial neural networks are incapable of performing well when the extrapolation is required.

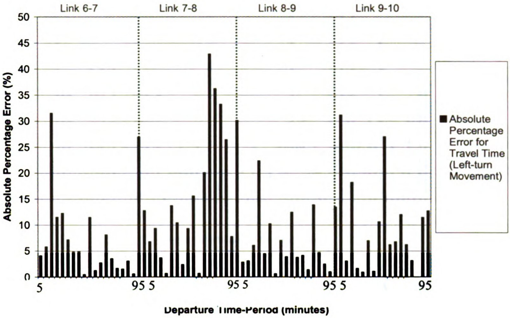

Figure 7.27 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Testing Set 1.

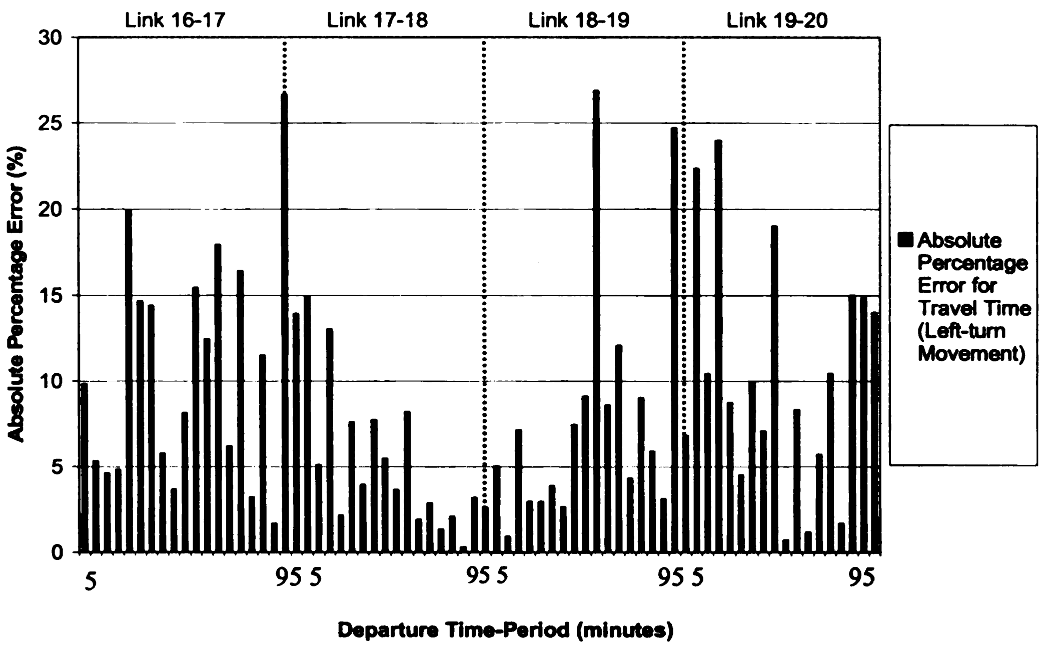

Figure 7.28 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Testing Set 2.

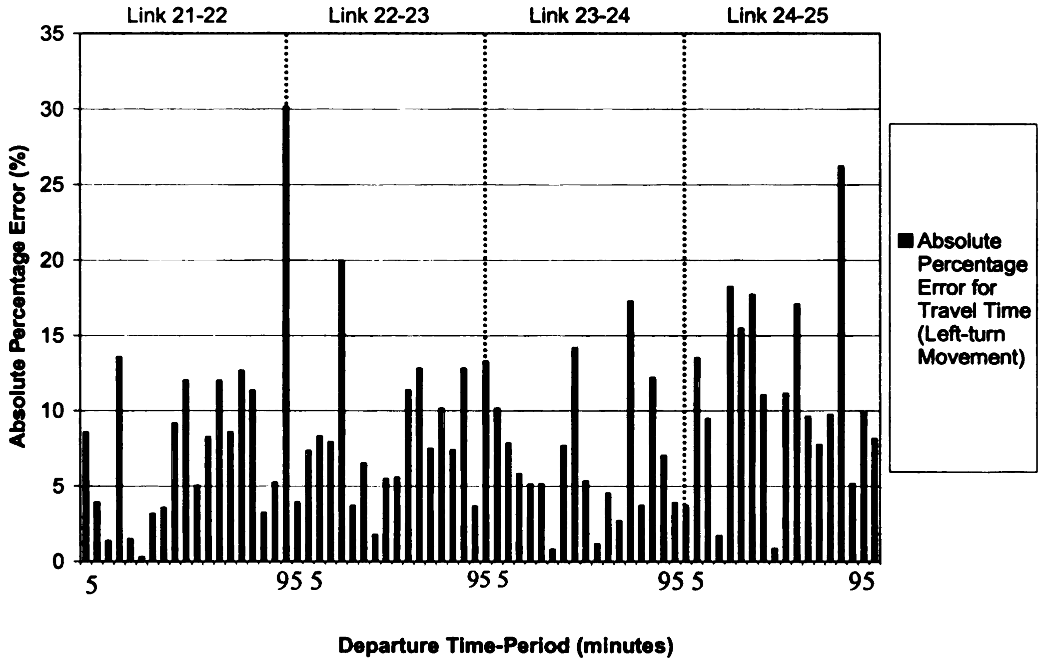

Figure 7.29 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Testing Set 3.

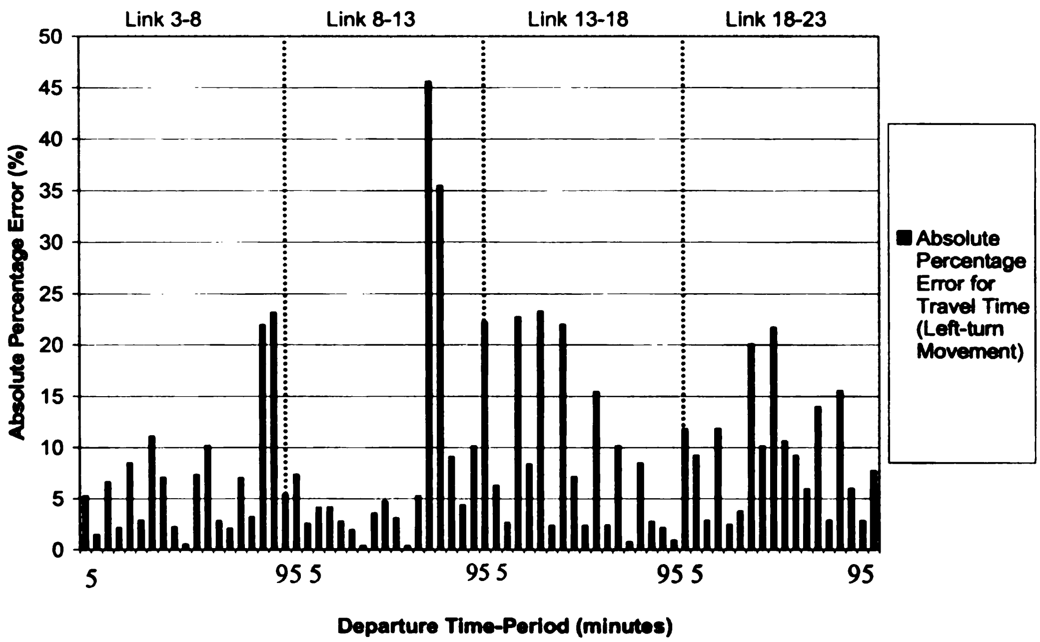

Figure 7.30 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Left-Turn Movement on Testing Set 4.

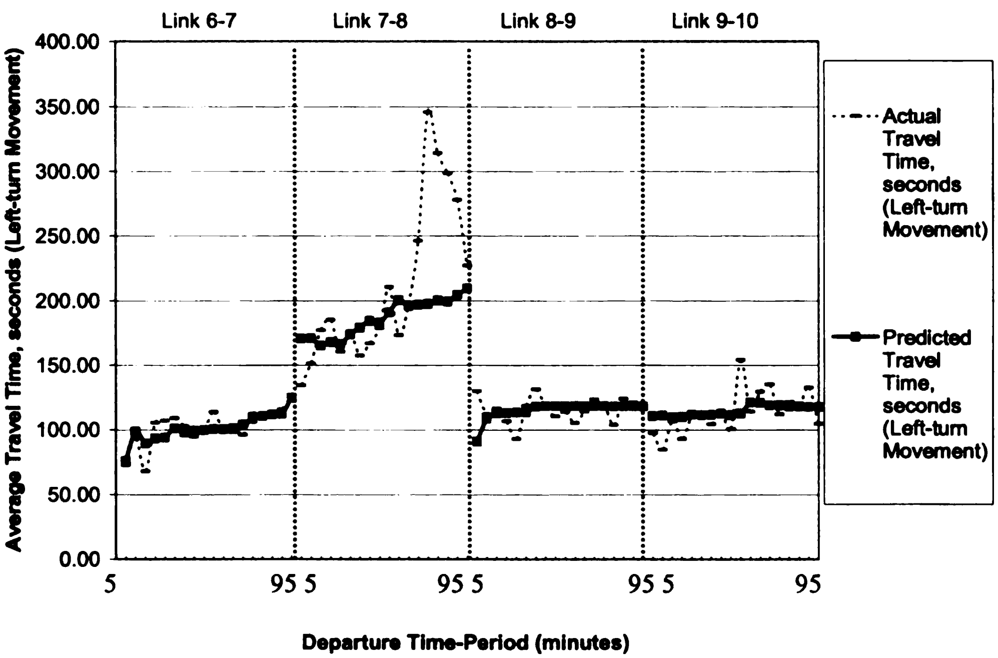

Figure 7.31 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Left-Turn Movement on Testing Set 1.

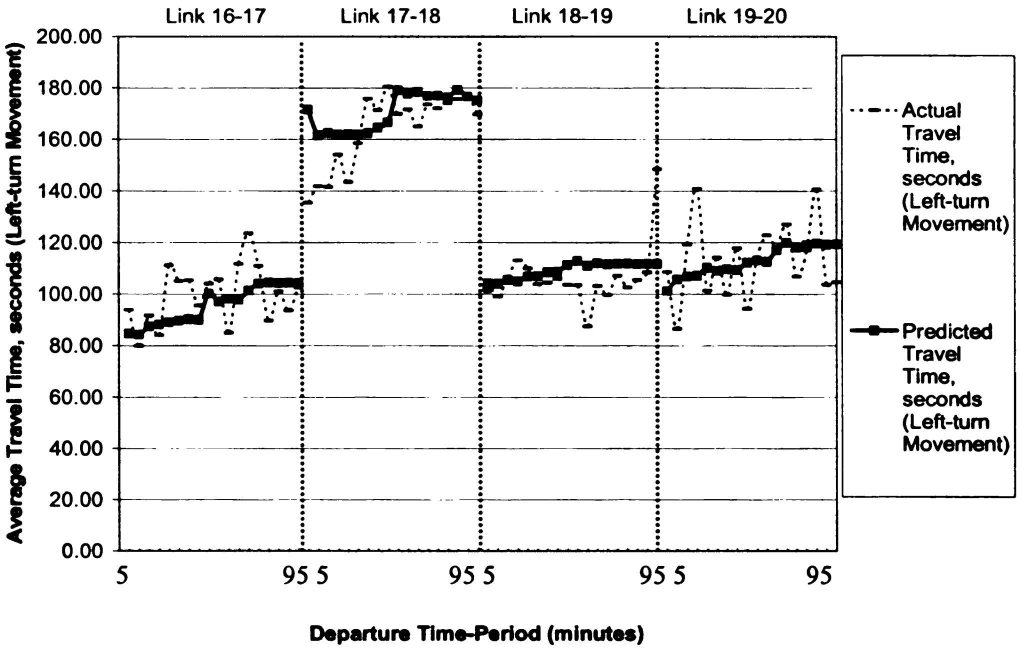

Figure 7.32 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Left-Turn Movement on Testing Set 2.

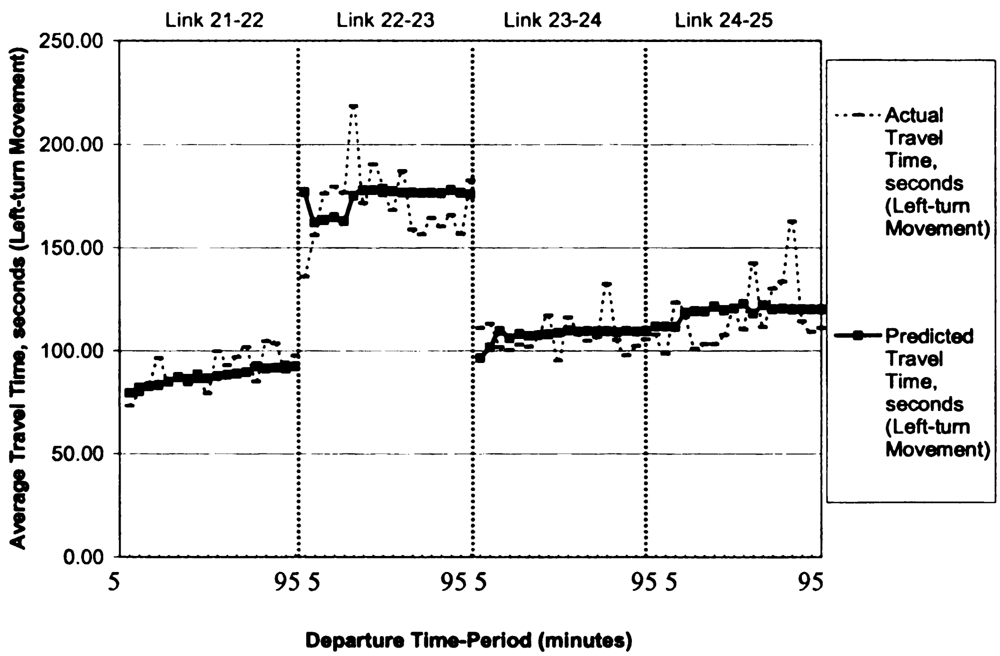

Figure 7.33 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Left-Turn Movement on Testing Set 3.

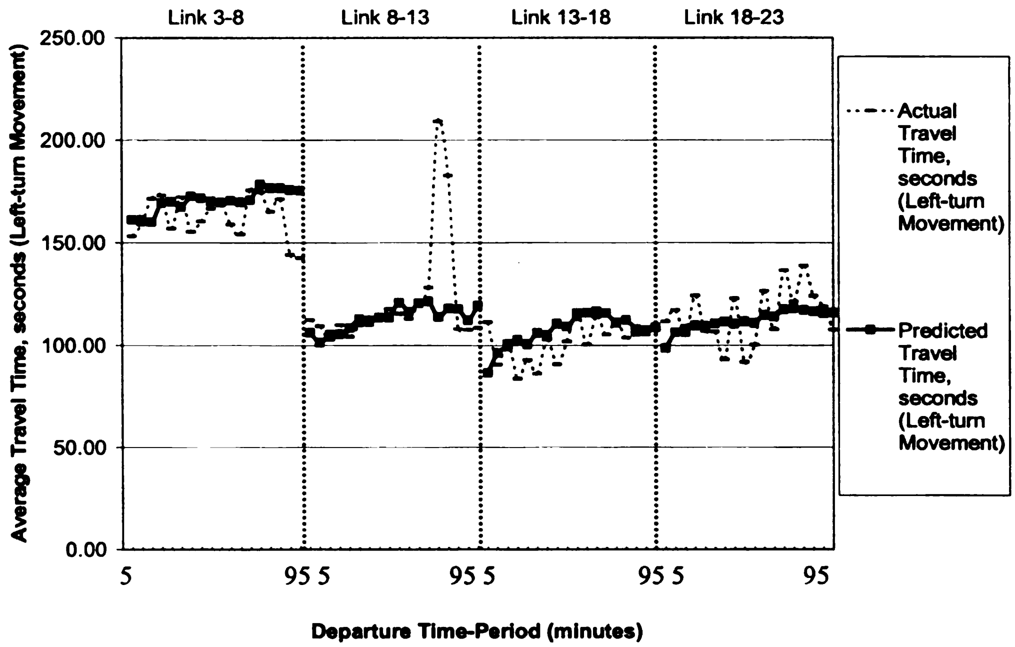

Figure 7.34 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Left-Turn Movement on Testing Set 4.

The detailed analysis of the CI-SSNN-Left model on the testing set is done by noting suitable performance measures. The performance measures values presented in Table 7.12 show that a MAPE value of 8.8% is obtained for entire testing set. Also, MAE, RMSE, and MAPE value of arterial 6-10 is higher as compared to other arterials. It is also because of the same reason that average travel time value was greater than 250 seconds for arterial 6-10 in testing set. Since, the maximum value of travel time encountered in training set was 250 seconds, so, CI-SSNN-Left failed to predict extrapolated values.

Table 7.12 – Performance Measure of CI-SSNN-Left Model for Travel Time Prediction for Left-Turn Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	15.8	30.0	9.8
16-20	10.1	12.8	8.6
21-25	10.7	13.7	8.5
3-23	10.8	17.6	8.4
Total Testing Set	11.8	19.8	8.8

7.4.5 Results of CI-SSNN-Right Model for Travel Time Prediction for Right-Turn Movement on Training Set

This section presents the efficiency of the CI-SSNN-Right model for training and its predictive performance on the testing set for the average travel time prediction of right-turning vehicles on an arterial link. The plot between the predicted travel time against actual travel time for right-turn movement is shown in Figure 7.35. The R^2 value is 0.88 which reflects an efficient learning of the model. The APE distribution for training sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23, respectively, is shown in Figures 7.36, 7.37, 7.38 and 7.39. The APE values for the arterials during most of the future departure time-periods are less than 15%. The arterials that have an APE value more than 15% for some future departure time-periods are 6-7, 16-17, 18-19, 23-24, 24-25, 8-13, and 13-18. The probable reason for the poor performance of model on these arterials is the stochastic nature of the CORSIM simulation.

The performance measures of the CI-SSNN-Right model on the training set is shown in Table 7.13. The MAPE value for the entire training set is 9.6%. This shows an efficient learning ability of the CI-SSNN-Right model. It is expected that this model should perform with comparable accuracy when applied on a testing set.

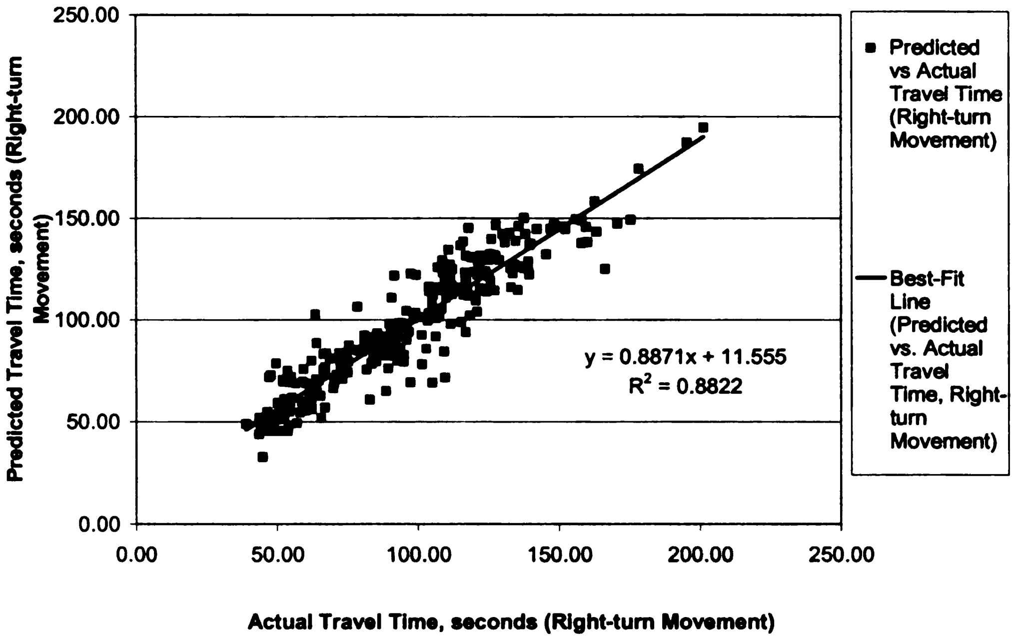

Figure 7.35 – Training Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Right-Turn Movement.

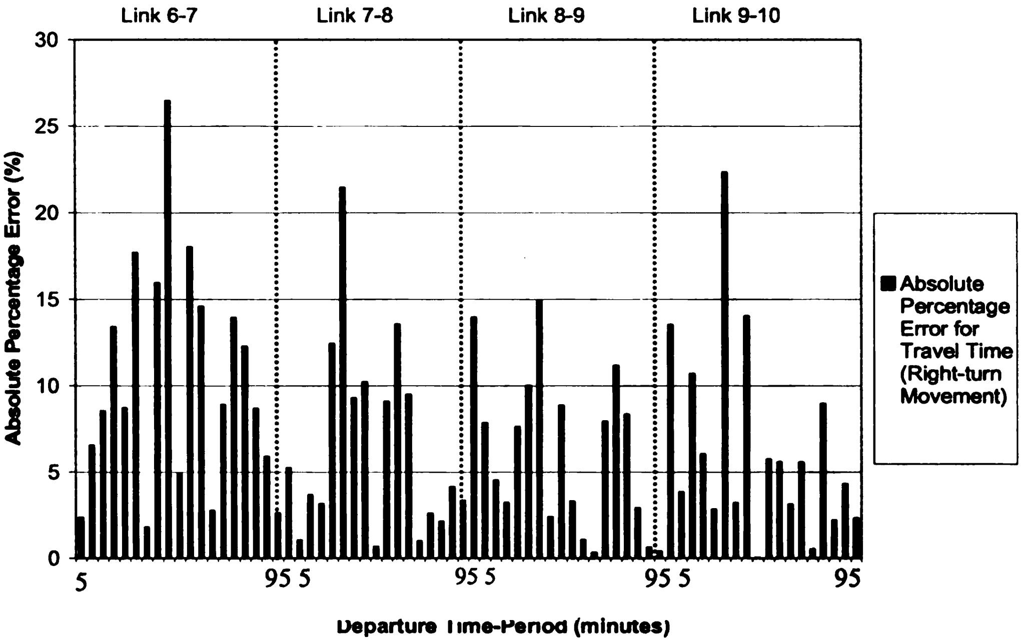

Figure 7.36 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Training Set 1.

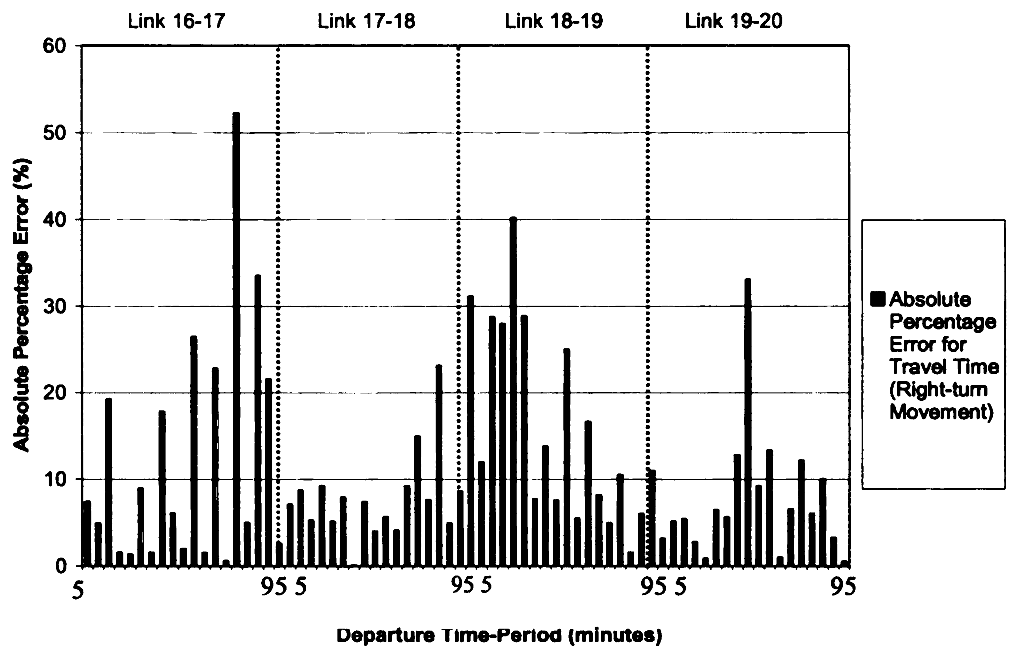

Figure 7.37 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Training Set 2.

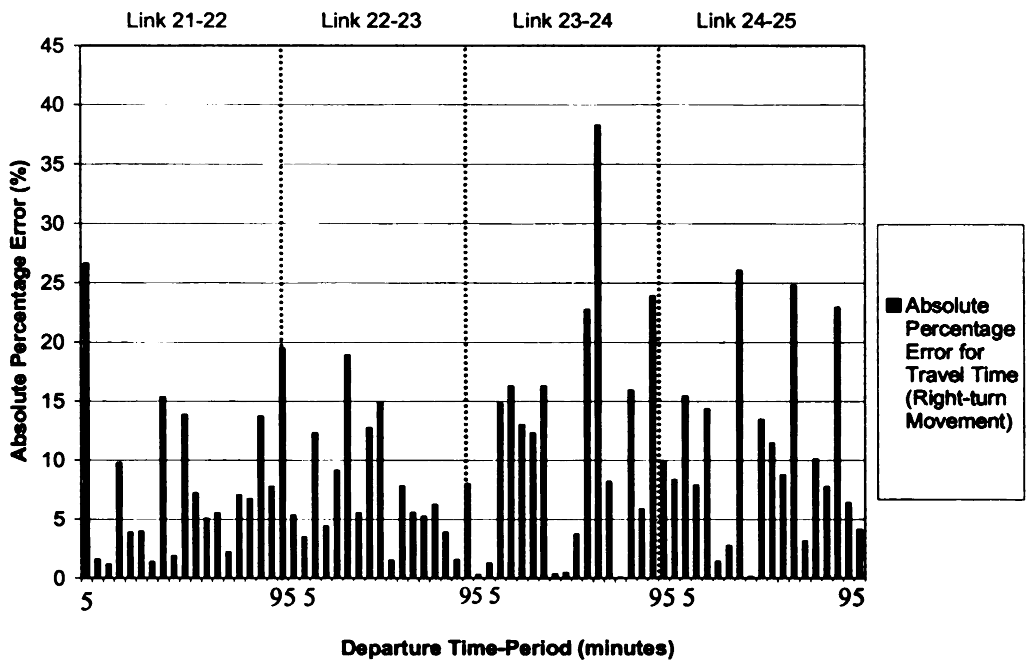

Figure 7.38 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Training Set 3.

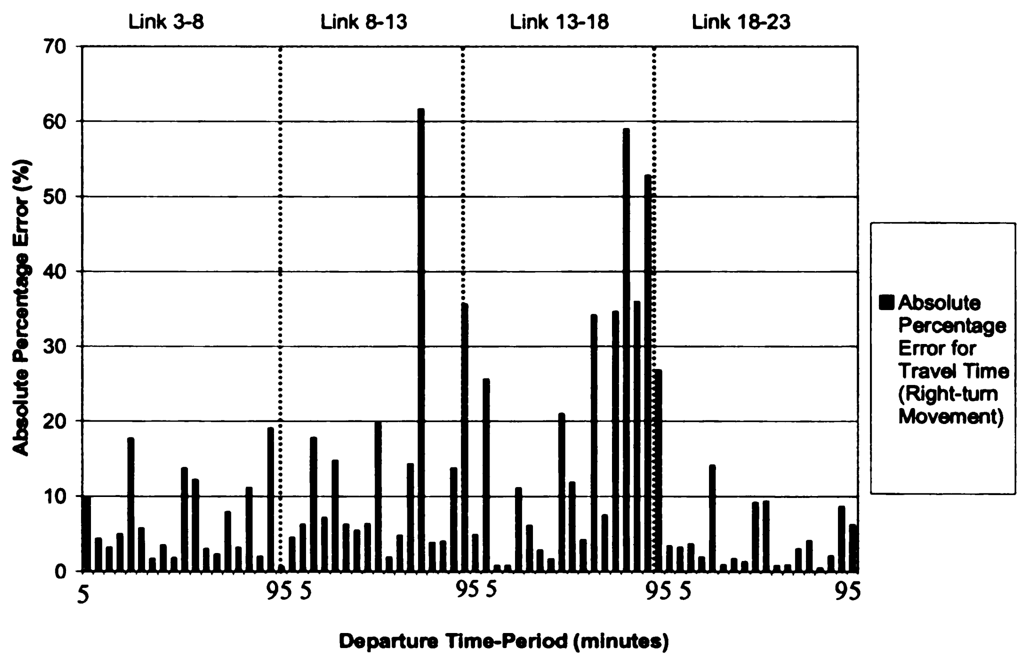

Figure 7.39 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Training Set 4.

Table 7.13 – Performance Measure of CI-SSNN-Right Model for Travel Time Prediction for Right-Turn Movement on Training Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	6.5	8.5	7.3
16-20	9.3	11.8	11.0
21-25	9.0	12.1	9.3
3-23	8.8	12.8	10.7
Total Training Set	8.4	11.4	9.6

7.4.6 Results of CI-SSNN-Right Travel Time Prediction Model for Travel Time Prediction for Right-Turn Movement on Testing Set

After training of the CI-SSNN-Right model, it is applied on a testing set to evaluate the generalization ability of the model. The plot of the predicted versus actual travel time is shown in Figure 7.40. This plot shows a R^2 value of 0.81, which is comparable to that obtained in the training set.

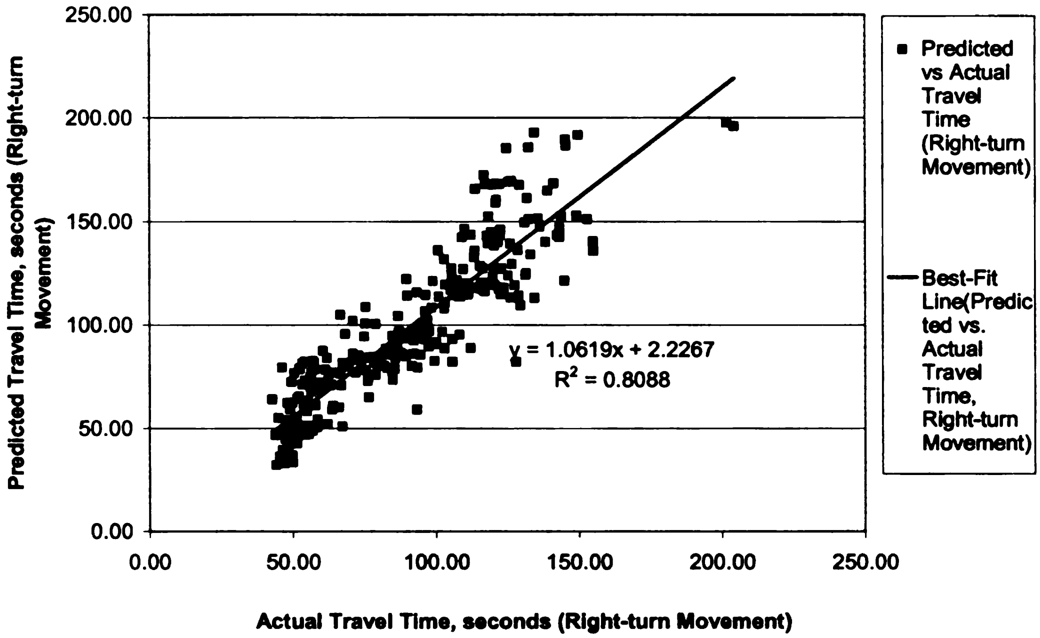

Figure 7.40 – Testing Set Scatter Plot of Predicted versus Actual Travel Time for Travel Time Prediction for Right-Turn Movement.

The APE distribution for testing sets 1, 2, 3, and 4, which represent arterials 6-10, 16-20, 21-25, and 3-23, respectively, is shown in Figures 7.41, 7.42, 7.43, and 7.44. The APE distribution for the predicted travel time shows that most of the arterials in the testing set have less than 20% percentage error value for predicted travel time. The arterial links which have higher percentage error for some future departure time-periods are identified as 6-7, 7-8, 8-9, 16-17, 18-19, 19-20, 21-22, 22-23, 24-25, 8-13, and 3-18. These arterials show on average higher percentage error than 20% value. It indicates that the travel time prediction models developed here perform relatively poor for right-turning movements as compared to through and left-turning movements. The reason for the relatively poor performance of prediction models for right-turning movements is because the right-turning traffic shares the same lanes and signal control state as that of through movements. The interference caused by the through movement traffic to the right-turn movements makes their travel time prediction difficult.

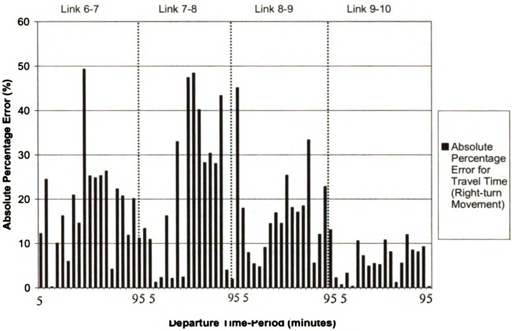

Figure 7.41 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Testing Set 1.

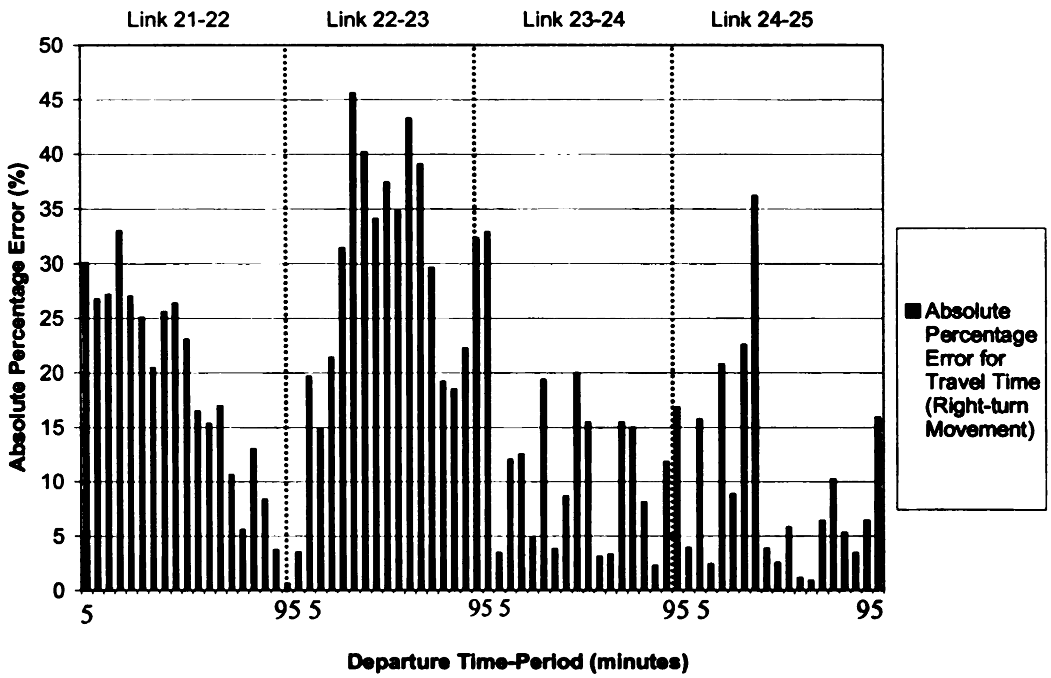

Figure 7.42 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Testing Set 2.

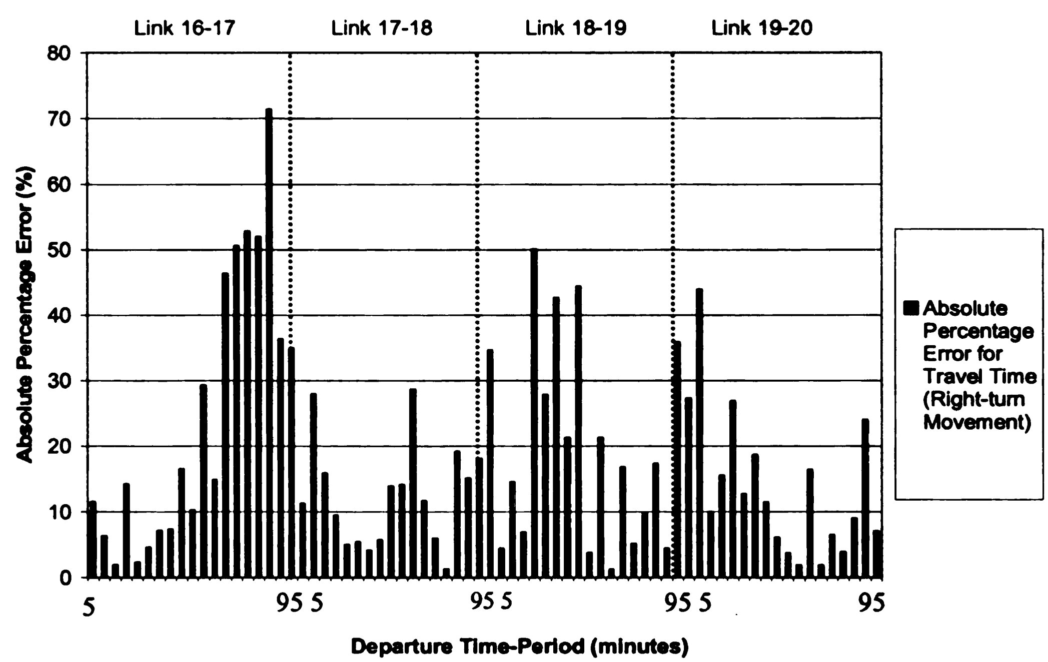

Figure 7.43 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Testing Set 3.

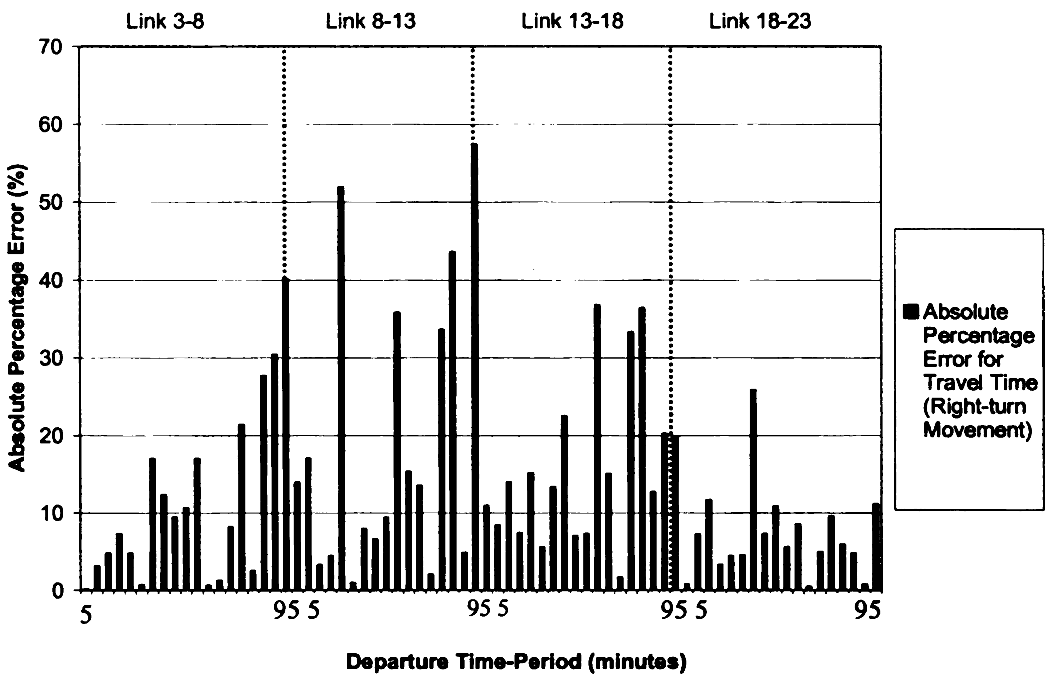

Figure 7.44 – Absolute Percentage Error of Predicted Travel Time versus Departure Time for Travel Time Prediction for Right-Turn Movement on Testing Set 4.

The plots showing pattern of the predicted and actual travel time are presented in Figures 7.45, 7.46, 7.47, and 7.48. It is shown that the predicted travel time is following the actual travel time pattern except for the arterials which have higher APE distribution as shown earlier.

The performance measure values of the CI-SSNN-Right model on the testing set are presented in Table 7.14. The MAPE value for the testing set is 15.9%. It was also observed in the APE distribution plots presented earlier that the percentage error values are above 20% for a significant number of arterials. This indicates the relatively poor predictive performance of travel time modeling when applied to the right-turn movements.

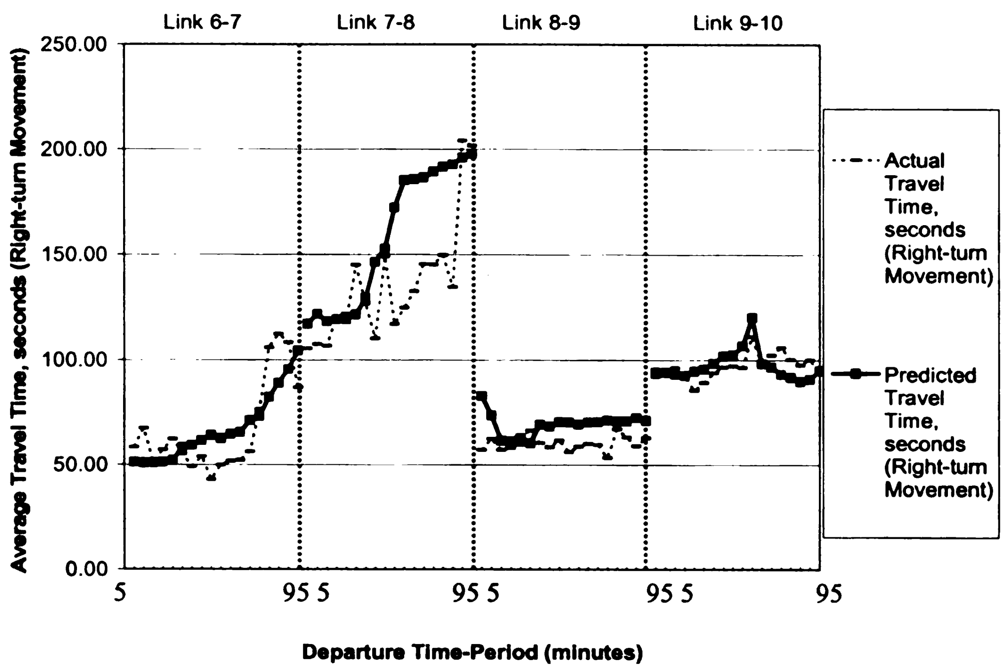

Figure 7.45 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Right-Turn Movement on Testing Set 1.

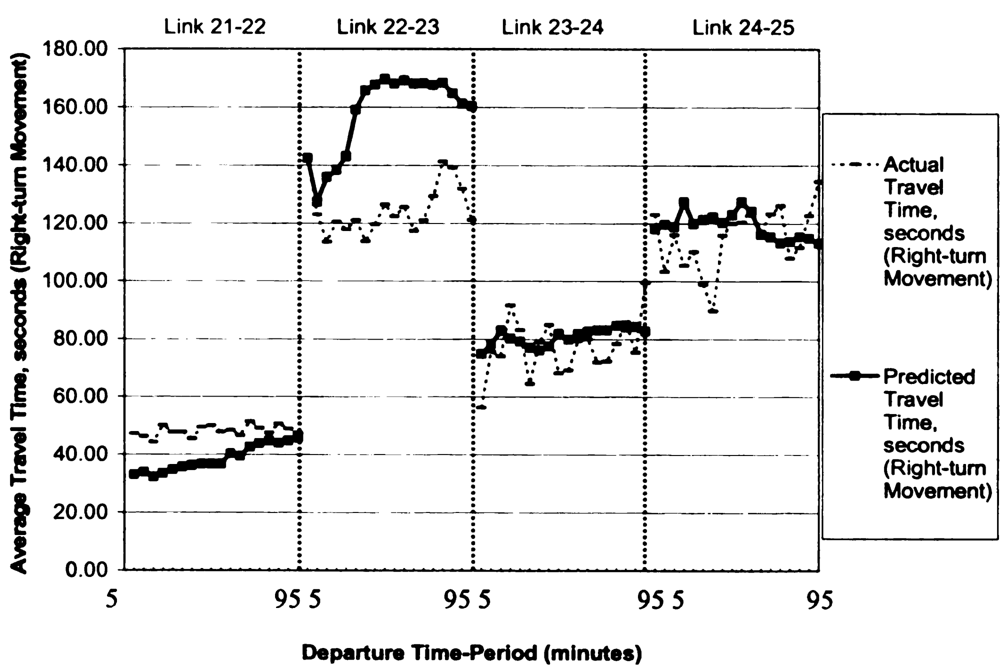

Figure 7.46 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Right-Turn Movement on Testing Set 2.

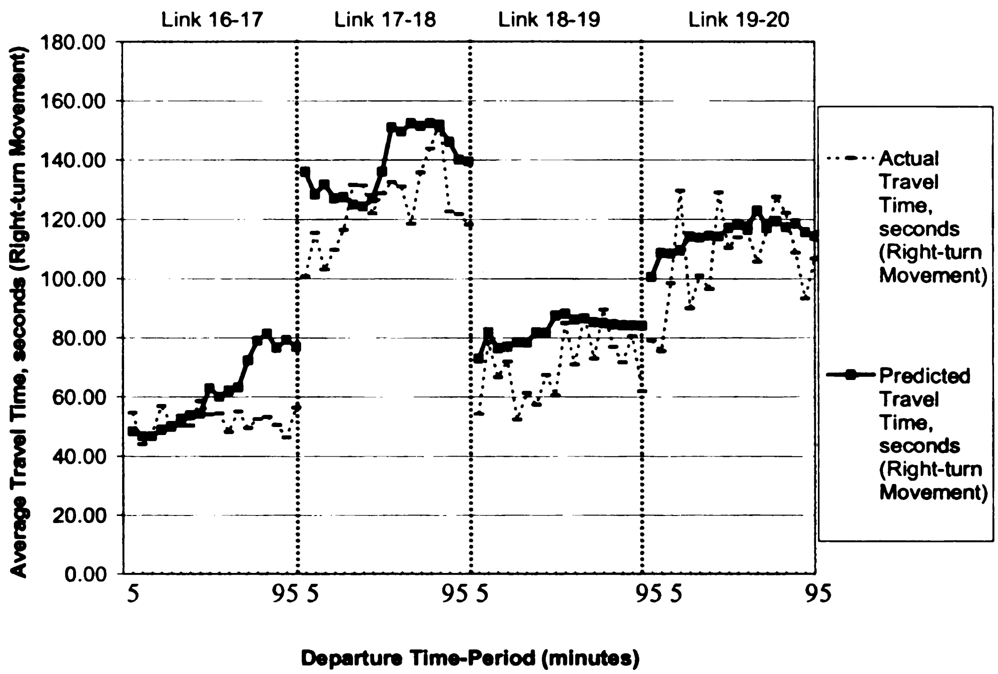

Figure 7.47 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Right-Turn Movement on Testing Set 3.

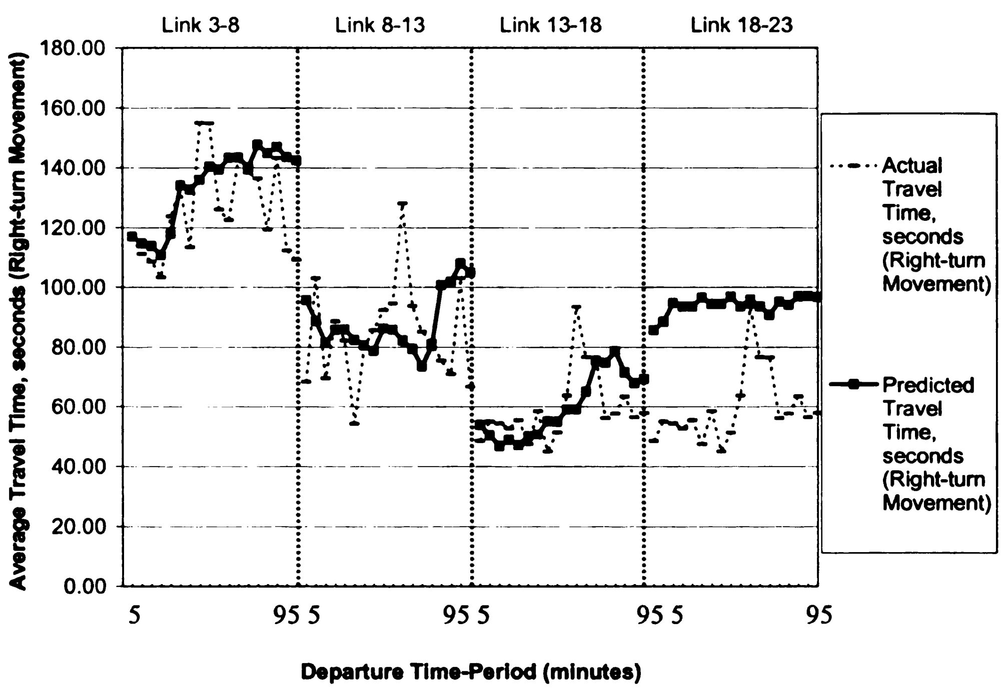

Figure 7.48 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Right-Turn Movement on Testing Set 4.

Table 7.14 – Performance Measure of CI-SSNN-Right Model for Travel Time Prediction for Right-Turn Movement on Testing Set

Arterial	MAE (seconds)	RMSE (seconds)	MAPE (%)
6-10	13.5	19.5	15.3
16-20	13.6	16.5	17.9
21-25	15.4	20.7	16.9
3-23	10.9	14.9	13.3
Total Testing Set	13.4	18.0	15.9

7.5 Results of Travel Time Prediction on Arterial Route

The development of the travel time prediction models for specific traffic movement on arterial links leads to an easy extension to any selected route in the network. The example route which is selected in this study has intersection 16 as its origin and intersection 25 as its destination. This example route was shown earlier in Figure 6.6. The path adopted for this route covers links 16-17, 17-18, 18-23, 23-24, and 24-25. The way of extending the CI-SSNN models developed earlier for this route is discussed in Chapter 6. Recalling that the travel time prediction models developed earlier for specific traffic movement is applied on each link comprising the route. In this way, the travel time prediction can be done for arterial route once the models were developed for links.

The results obtained on the testing set are enough to show the predictive performance of CI-SSNN models applied to this route. Figure 7.49 shows the pattern of the actual and predicted average travel time on the route with departure time-periods in future. The first future departure time-period is starting from 5 minutes and ending at 10 minutes. Similarly, the last future departure time-period starts at 90 minutes and ends at 95 minutes. It is seen that the predicted average travel time follows the same pattern of the actual travel time.

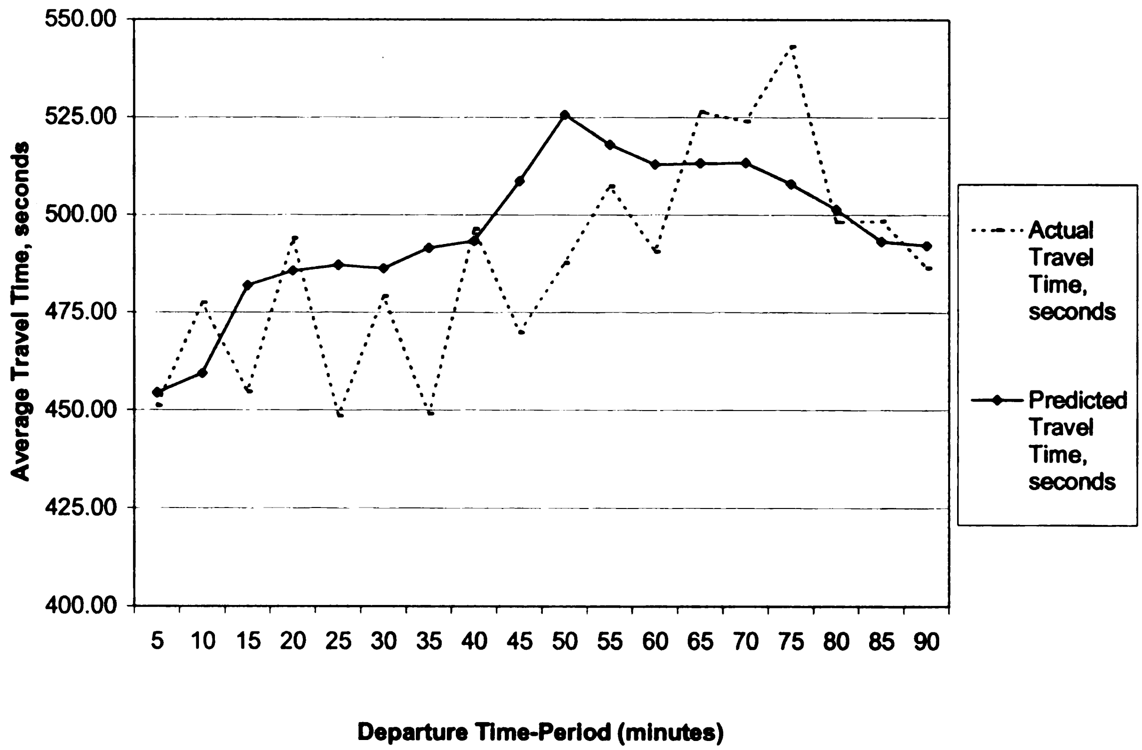

Figure 7.49 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction for Route 16-25 on Testing Set.

A detailed pattern of the actual and predicted travel time is presented in Figure 7.50 analyzing arterial links that comprise the example route 16-25. This figure shows that the predicted travel time closely follows the actual travel time values for arterial links considering appropriate traffic movement.

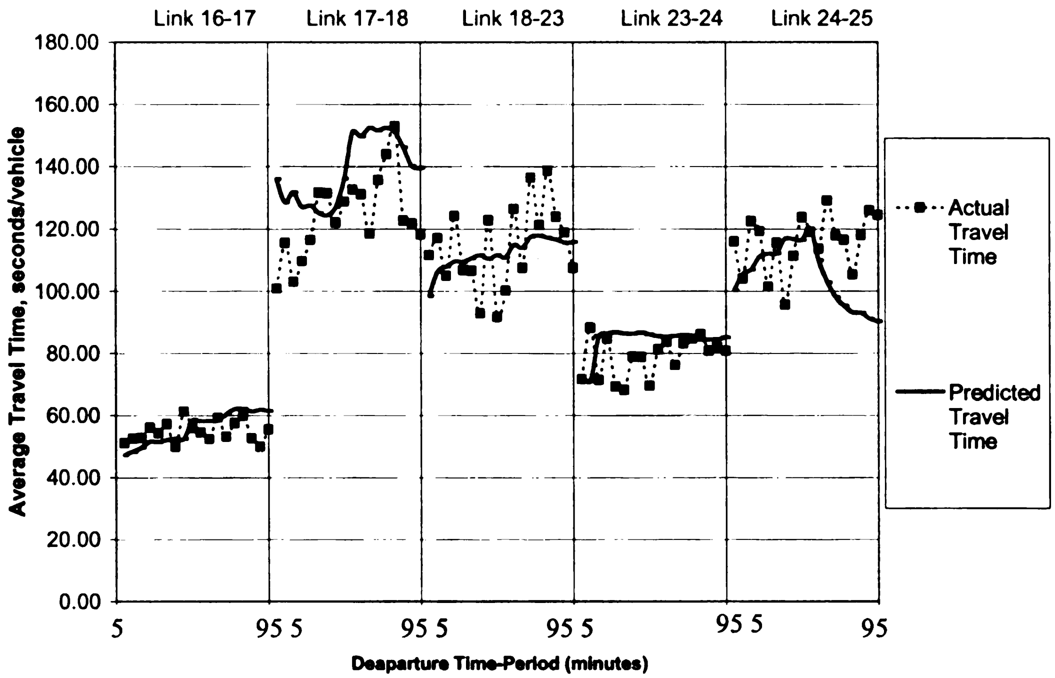

Figure 7.50 – Pattern of Actual and Predicted Travel Time for Travel Time Prediction on arterial links comprising the Route 16-25 on Testing Set.

A further investigation is required to see the Absolute Percentage Error (APE) distribution for the route. Figure 7.51 shows that APE values for all the departure time-periods are within 10% value. Table 7.15 lists values of the measures of performance for this route which shows that the MAPE value is 3.8% for this route.

It is noted that the MAPE value for the travel time prediction on example route is much lower than those for arterial links presented earlier. The reason is that the average travel time for the route is obtained in this study by adding the travel time for each link considering the appropriate traffic movement. At some of the links comprising the route, the travel time may have been predicted as higher than actual travel time, and for the other links, the predicted travel time may have been lower than actual travel time. The addition of the travel time on the links comprising the route balances out the negative and positive errors and hence provides more accurate value of travel time for the route. It is worth mentioning that the way of finding out the travel time on the route as adopted in this study is valid as the appropriate traffic movement required to traverse the route on each link is considered.

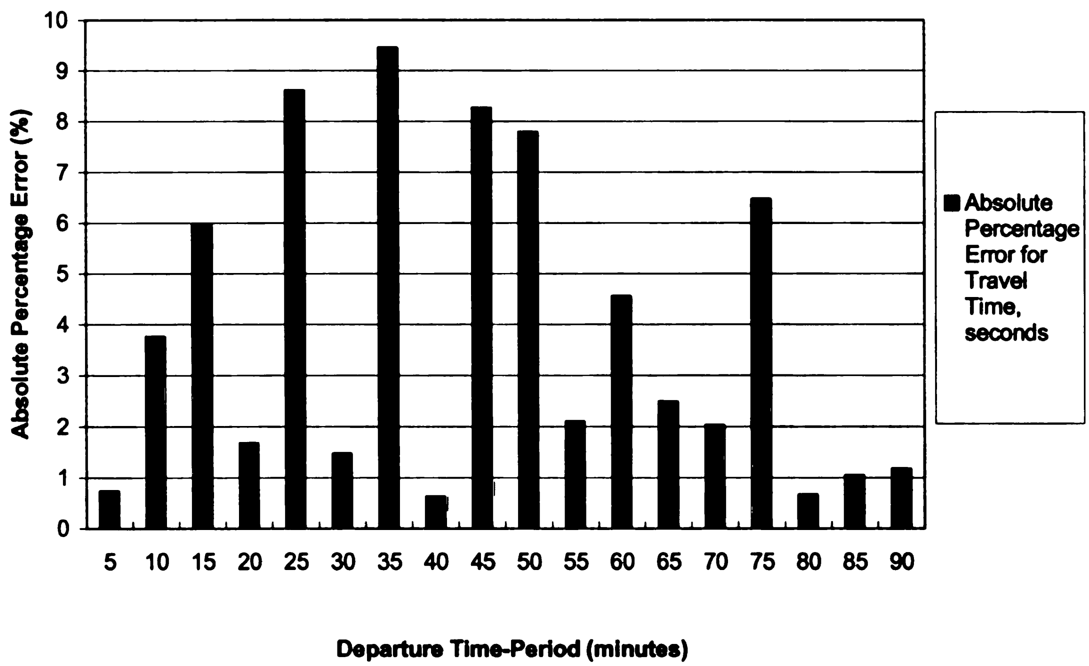

Figure 7.51 – Absolute Percentage Error Distribution for Travel Time Prediction for Route 16-25 on Testing Set.

Table 7.15 – Performance Measure of CI-SSNN Model for Travel Time Prediction for Route 16-25 on Testing Set

Route	MAE (seconds)	RMSE (seconds)	MAPE (%)
16-25 (testing set)	18.4	23.2	3.8

Chapter 8

Conclusions and Future Research

The conclusions and findings in this study are summarized. The limitations of the study and scope of future research for field implementation of the proposed methodology is discussed.

8.1 Conclusions and Findings

The conclusions and significant findings from this research on travel time estimation and prediction are as follows:

1. Travel time estimation and prediction on urban arterial networks is a complex dynamic process with non-linear relationship between traffic parameters like average flow, average speed, queue length, geometrics, etc. and the average travel time of current and future departure time-periods.
2. Travel time prediction should be combined with the estimation of travel time as the estimated travel time for current departure time-period acts as an important input variable along with known traffic parameters for travel time prediction modeling.
3. The proposed models of travel time estimation and short-term prediction are using traffic parameters that can be easily obtained or collected in the field by traffic

detection technologies like loop detectors and image processing based techniques. The developed models estimate and predict the travel time using inputs as average flow rate, queue length, geometrics, and signal control parameters only. The reliance of these models on easily available data makes the models economical and feasible for field deployment for real-time applications.

4. The developed models explicitly take into account turning movements, geometrics as well as signal control scheme on arterial networks. This is a significant contribution of this study because of the complexity involved in inclusion of all these parameters.

5. The state-space dynamics of traffic behavior in an urban arterial is captured successfully by using a generic class of Recurrent Neural Networks (RNN) termed the State-Space Neural Networks (SSNN). The use of SSNN provides flexibility and efficiency while at the same time modeling the non-linear dynamical traffic system.

6. A pre-processing of observable traffic parameters is required to minimize the input parameters for SSNN models. A statistical technique called as the Conditional Independence (CI) graph is introduced for the first time with artificial neural networks in this research. The analysis of conditional independence and interaction among observable traffic parameters and the average travel time increases the understanding of the process as well as improves the predictive performance of SSNN.

7. The SSNN model based on the CI graphs is compared with the typical SSNN model used by researchers in this study. It shows that the accuracy of the CI

graphs based SSNN model is comparable to that of typical SSNN models which do not use the inferences made by the CI graphs. The CI graphs based SSNN model is dependent upon lesser input traffic variables as compared to typical SSNN model. The constraint of limited amount of data obtained in any experiment or in the field makes the CI graphs based SSNN models a better choice than the SSNN models. This application of the CI graphs is recommended not only for the travel time estimation and prediction modeling as done in this research but also to other traffic modeling problems.

8. Three SSNN models based on the CI graphs are developed for travel time estimation for the specific traffic movement of through, left, and right-turn movement on arterial links. Similarly, three SSNN models based on the CI graphs are developed for short-term travel time prediction for arterial links. Each single model is able to capture travel time fluctuations for uncongested, congested, and the transition stage between these two traffic conditions on arterial links. Thus, each single SSNN model can perform successfully on arterial networks throughout the day.
9. A single SSNN model based on the CI graph is proposed for travel time estimation and prediction on urban arterials for specific movement which is efficient for both congested and non-congested traffic conditions occurring throughout a day.
10. The CI graph based SSNN models developed for arterial links are extended to obtain results on any route in arterial network. Thus, the developed travel time

models can be applied on a single arterial link, an entire arterial or a route in a network.

11. The Mean Absolute Percentage Error (MAPE) for proposed average travel time estimation models for arterial links ranges from 6.5% to 15.0%. The MAPE value for average travel time prediction models ranges from 6.9% to 17.9%. The MAPE value for the route considered in this study is 3.8%.
12. The Mean Absolute Percentage Error for average travel time estimation and prediction for right-turn movement is found more challenging that that of through and left-turn movements. The reasons for complexity in right-turn movement travel time modeling is the high interference among through and right-turn movement vehicles on an arterial.
13. The travel time estimation and prediction modeling approach in this study is more generalized as compared to the studies done to date. The proposed models account for traffic variables like geometrics, turning movements, signal control scheme, which were either assumed constant or not explicitly taken into account in earlier studies. The proposed models provide prediction for through, left, and right-turning movements on arterial links and routes, which is a significant contribution of this study.

8.2 Limitations

The limitations of proposed modeling approach are listed as follows –

1. The proposed methodology does not include some factors like weather effects, incidents etc. These factors are required for a real-time implementation of the model in field. The flexibility of the modeling approach allows the inclusion of these factors provided that the data is available.
2. The modeling of travel time estimation and short-term prediction is based on a hypothetical arterial network created in a microscopic traffic simulation. The implementation of the models in a real world setting requires calibration and validation with site data. However, enough variability has been included in experimental set-up that accounts for a wide range of geometric conditions and traffic demands.
3. The accuracy of average travel time estimation and prediction models for right-turning movements is not as good as through and left-turning movements. A high interference between right-turning traffic and through traffic on arterial links is observed at signalized intersections in this study. This makes the travel time modeling for right-turning movements more challenging than other traffic movements. Thus, the travel time modeling for right-turning movement requires consideration for improvement of the accuracy.

8.3 Scope of Future Research

There are some recommendations and scope of future research which are listed below:

1. The coupling of Conditional Independence graph with Artificial Neural Networks (ANN) is recommended not only for travel time problem but also to other engineering problems where ANN is used as a potential modeling or prediction technique.
2. The short-term travel time prediction in this study focuses on 5 minutes time-period ahead in future. The travel time prediction for further time-periods in the future can be done using the same methodology.
3. The real-time implementation of travel time estimation and short-term prediction model should be done using the same methodology. The author realizes that issues related to data-collection in the field, data-fusion, and dissemination of travel time information will pose another challenge to handle.

APPENDIX

MATLAB Code for State-Space Neural Network

The Appendix presents the MATLAB code for State-Space Neural Network Model. This code is written for travel time prediction of thru movement on arterials. Similar code can be used for other SSNN models changing the parameters.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

display('TRAVEL TIME PREDICTION FOR THRU MOVEMENT');

load traininout.m ;                    %loading input data file
P = traininout(:,:);
P = rot90(P,1);
ptr1 = P( 2:9,:);ptr1 = con2seq(ptr1);           %link 1 input

ttr1 = P(1,:);ttr1= con2seq(ttr1);       %link 1 output-travel time


for i = 1:288
ptr{1,i} = ptr1{1,i};
end

for i = 1:288
ttr{1,i} = ttr1{1,i};
end

        %Loading up State Space Neural Network called here as net1
disp('Loading up network net1 ...');

load net1;
net1 = init(net1);                  %Initializing the network

                               %Training Parameters
net1.adaptParam.epochs = 10000;
```

```
net1.adaptParam.goal = 0.001;
net1.adaptParam.max_fail = 10000
                         %Five Validations to check when error rises
net1.adaptParam.mem_reduc = 1;
                    %Full Jacobian Calculated, no memory restrictions
net1.adaptParam.min_grad = 1e-10;
net1.adaptParam.mu = 0.9;                   %momentum constant
net1.adaptParam.mu_dec = 0.001;
net1.adaptParam.mu_inc = 10;
net1.adaptParam.mu_max = 1e10;
net1.adaptParam.show = 100;
net1.adaptParam.time = inf;
net1.adaptparam.lr = 0.001;                 % Learning Rate
net1.performFcn = 'mse';          %Performance = 'Mean Square Error'


                                  %Training
disp('Training ...');

[net1,tr]= adapt(net1,ptr,ttr);

save net1.MAT net1 ;

                         %Conversion of ttr to concurrent matrix
ttr = seq2con(ttr);

                              %Results - Network Training
a2 = sim(net1,ptr);
a2 = seq2con(a2);

figure(1)
title('Regression Analysis of Training Results: Travel Time Thru'); hold on
[m,b,r] = postreg(a2{1,1},ttr{1,1}); hold off
perftrg1 = mse(a2{1,1} - ttr{1,1})
perftrg2 = mae(a2{1,1} - ttr{1,1})


load testinout.m;                           %loading testing file
Q = testinout(:,:);
Q = rot90(Q,1);
test1 = Q(2:9,:); test1 = con2seq(test1);

for i = 1:288
test.P{1,i} = test3{1,i};
end
```

```
testT1 = Q(1,:);testT1 = con2seq(testT1);

for i = 1:288
test.T{1,i} = testT1{1,i};
end

test.T = seq2con(test.T);

                              %Results - Network Testing

a3 = sim(net1,test.P);
a3 = seq2con(a3);

figure(5)
title('Regression Analysis of Testing Results: Travel Time Thru'); hold on
[m,b,r] = postreg(a3{1,1},test.T{1,1}); hold off
perftst1 = mse(a3{1,1} - test.T{1,1})
perftst2 = mae(a3{1,1} - test.T{1,1})

format short;
disp('Network Architecture');
disp('Number of Layers');
disp(net1.numlayers+1);
nneurons = 5;
slayer = size(net1.layers);
for i=1:slayer(1,1)
  nneurons =[nneurons net1.layers{i}.size];
end


                              %saving input weights files
iw1 = net1.IW{1}';
iw2 = net1.IW{2}';
iw3 = net1.IW{3}';
iw4 = net1.IW{4}';


save iw1.dat iw1 -ascii -tabs;
save iw2.dat iw2 -ascii -tabs;
save iw3.dat iw3 -ascii -tabs;
save iw4.dat iw4 -ascii -tabs;

                              %saving layerweights files
lw1 = net1.LW{1,1}';
lw2 = net1.LW{2,2}';
lw3 = net1.LW{3,3}';
```

```
lw4 = net1.LW{4,4}';
lw41 = net1.LW{4,1}';
lw42 = net1.LW{4,2}';
lw43 = net1.LW{4,3}';


save lw1.dat lw1 -ascii -tabs;
save lw2.dat lw2 -ascii -tabs;
save lw3.dat lw3 -ascii -tabs;
save lw4.dat lw4 -ascii -tabs;
save lw41.dat lw41 -ascii -tabs;
save lw42.dat lw42 -ascii -tabs;
save lw43.dat lw43 -ascii -tabs;


                %forming training and testing output files to save
trgouttthru = a2{1}';
testouttthru = a3{1}';

                                %saving output in .dat files
save trgouttthru.dat trgouttthru -ASCII -DOUBLE -TABS;
save testouttthru.dat testouttthru -ASCII -DOUBLE -TABS;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

REFERENCES

Abu-Lebdeh, G., and Benekohal, R.F. 2000. "Genetic Algorithms for Traffic Signal Control and Queue Management of Oversaturated Tow-Way Arterials. In *Journal of the Transportation Research Board*, TRR # 1727, TRB, National Research Council, Washington, DC, pp. 61-67.

Abu-Lebdeh, G., and Benekohal, R.F. 2003. "Design and Evaluation of Dynamic Traffic Management Strategies for Congested Conditions". In *Transportation Research-A*, Volume 37, Issue 2, Feb. 2003, pp. 109-127.

Ben-Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H.N., and Mishalani, R. 2001. Network State Estimation and Prediction for Real-Time Traffic Management. *Networks and Spatial Economics*. 1, pp. 293-318.

Brockwell, P.J., and Davis, R.A. *Introduction to Time Series and Forecasting*. 2nd Edition. 2002. Springer-Verlag. New York.

Chen, H., and Abu-Lebdeh, G. 2006. Development of a Framework for an Integrated Dynamic Signal-Dynamic Speed Traffic Management Algorithm for Signalized Networks. In *TRB 2006 Annual Meeting CD-ROM*. Washington DC.

Chien, S. I., and Kuchipudi, C. M. 2003. Dynamic Travel Time Prediction with Real-Time and Historic Data. *Journal of Transportation Engineering*, Vol. 129, No. 6, pp. 608-616.

Elman, J. 1990. Finding Structure in Time. *Cognitive Science*, Vol. 14, pp. 179-211.

Gang-Len, C. and Su ,C. 1995. Predicting Intersection Queue with Neural Network Models. *Transportation Research*, 3(C), pp. 175-191.

Haykin, S. *Neural Networks: A Comprehensive Foundation*. 1999. Prentice-Hall, Upper Saddle River, N.J.

Honkela, A. 2001. Nonlinear Switching State-Space Models. MS Thesis. The Helsinki University of Technology. Finland.

Ishak, S., and Al-Deek, H. 2002. Performance Evaluation of Short-Term Time-Series traffic Prediction Model. *Journal of Transportation Engineering*, Vol. 128, No. 6, pp. 490-498.

Jordan MI. 1986. Serial Order: A Parallel Distributed Processing Approach. Technical Report TR-8604. Institute for Cognitive Science. The University of California, San Diego. USA.

Kutner, M.H., Nachtsheim, C.J., and Neter, J. *Applied Linear Regression Models*. 2004. McGraw-Hill/Irwin. New York. USA.

Ledoux, C. 1997. An Urban Traffic Flow Model integrating Neural Networks. *Transportation Research, Part C*. Vol. 5. No. 5. pp. 287-300.

Lin, W.H., Kulkarni, A., and Mirchandani, P. 2004. Short-Term Arterial Travel Time Prediction for Advanced Traveler Information Systems. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, Vol. 8, No. 3, pp. 143 - 154.

Liu, H., van Zuylen, H., van Lint, H., Salomons, M. 2006. Urban Arterial Travel Time Prediction with State-Space Neural Networks and Kalman Filters. In *TRB 2006 Annual Meeting CD-ROM*. Washington DC.

Longfoot, J. 1991. An Automatic Network Travel Time System-ANNTS. In *Proc. of Society of Automotive Engineers*. Vol. 2. pp. 1053-1061.

Mandic, D.P., and Chambers, J.A. *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. 2001. John Wiley & Sons. New York.

Mark, C. D., and Sadek, A.W. 2004. Learning Systems for Predicting Experiential Travel Times in the Presence of Incidents: Insights and Lessons Learned. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1879*, TRB, National Research Council, Washington, D.C., pp. 51-58.

May, A.D. *Traffic Flow Fundamentals*. 1990. Prentice Hall. NJ.

Medsker, L.R., and Jain, L.C. *Recurrent Neural Networks: Design and Applications.* 1999. The CRC Press International Series on Computational Intelligence. CRC Press.

Owen, L.E., Zhang, Y., Rao, L., and McHale, G. 2000. Traffic Flow Simulation using CORSIM. In *Proceedings of the 2000 Winter Simulation Conference.*

Rajaraman, G. 1998. Evaluation and Refinement of Dynamic Travel Time Model For Signalized Streets. MSCE Thesis. The Purdue University. USA.

Rivals, I., and Personnaz, L. 1996. Black-Box Modeling with State-Space Neural Networks. In *Neural Adaptive Control Technology*, R.Zbikowski and K.J. Hunt eds., World Scientific, pp. 237-264.

Schrank, D., and Lomax, T. 2005. The 2005 Urban Mobility Report. Texas Transportation Institute, The Texas A&M University System. http://mobility.tamu.edu

Sisiopiku, V.P., and Rouphail, N.M. 1994. Toward the Use of Detector Output for Arterial Link Travel Time Estimation: A Literature Review. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1457*, TRB, National Research Council, Washington, D.C., pp. 158-165.

Sisiopiku, V.P., Rouphail, N.M, and Santiago, A. 1994. Analysis of Correlation between Arterial Travel Time and Detector Data from Simulation and Field Studies. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1457*, TRB, National Research Council, Washington, D.C., pp. 166-173.

Stathopoulos, A., and Karlaftis, M.G. 2003. A Multivariate State Space Approach for Urban Traffic Flow Modeling and Prediction. *Transportation Research Part C* 11, pp.121-135.

Traffic Detector Handbook. 1990. Technical Report FHWA-IP-90-002. 2nd Edition. Federal Highway Administration.

van Lint, J.W.C., Hoogendoorn, S.P., and Van Zuylen, H.J. 2002. Freeway Travel Time Prediction with State-Space Neural Networks: Modeling State-Space Dynamics with Recurrent Neural Networks. In *Transportation Research Record: Journal of the*

Transportation Research Board, No. 1811, TRB, National Research Council, Washington, D.C., pp. 30-39.

van Lint, J.W.C. 2004. Reliable Travel Time Prediction for Freeways. PhD Thesis. Delft University Press, Delft, The Netherlands.

Vlahogianni, E.I., Golias, J.C., and Karlaftis, M.G. 2004. Short-Term Traffic Forecasting: Overview of Objectives and Methods. *Transport Reviews*, Vol. 24, No. 5, pp. 533-557.

Wey, W. and Jayakrishnan, R. 2004. Applying a Network Programming Algorithm to a Network wide Traffic Control Scheme: Simulation-Based Performance Studies. *Transportation Research Record 1882*. pp. 105-112.

Whittaker, J., Garside, S., and Lindveld, K. 1997. Tracking and Predicting a Network Traffic Process. *International Journal of Forecasting*. Vol. 13, pp. 51-61.

Whittaker, J. *Graphical Models in Applied Multivariate Statistics*. 1990. Wiley, Chichester.

Wu, C. H., Ho, J.M., and Lee, D.T. 2004. Travel-Time Prediction with Support Vector Regression. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, Issue 4, pp. 276-281.

Zhang, X., and Rice J. 2003. Short-Term Travel Time Prediction. *Transportation Research, Part C*. Vol. 11, pp. 187-210.

Zwet, E., and Rice, J. 2004. A Simple and Effective Method for Predicting Travel Times on Freeways. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No. 3, pp. 200-207.