USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No 018PY01

Network Origin-Destination Demand Estimation using Limited Link Traffic Counts: Strategic Deployment of Vehicle Detectors through an Integrated Corridor Management Framework

By

Srinivas Peeta, Principal Investigator
Professor of Civil Engineering
Purdue University
peeta@purdue.edu

and

Shou-Ren Hu, Co-Principal Investigator
Assistant Professor of Transportation and Communication Management Science
National Cheng Kung University, Taiwan
shouren@mail.ncku.edu.tw

and

Chun-Hsiao Chu
Assistant Professor of Tourism
Aletheia University, Taiwan
au5907@mail.au.edu.tw

and

Han-Tsung Liou
Graduate Research Assistant of Transportation and Communication Management Science
National Cheng Kung University, Taiwan
yonex17@hotmail.com

Report Submission Date: October 15, 2009

DISCLAIMER

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No 018PY01 Final Report, October 2009

Network Origin-Destination Demand Estimation using Limited Link Traffic Counts: Strategic Deployment of Vehicle Detectors through an Integrated Corridor Management Framework

Introduction

In typical road traffic corridors, freeway systems are generally well-equipped with traffic surveillance systems such as vehicle detector (VD) and/or closed circuit television (CCTV) systems in order to gather timely traffic information for traffic control and/or management purposes. However, other highway facilities in the corridor, especially arterials and surface streets in the vicinity of the freeway, mostly lack detector/sensor systems. Yet, most traffic management and control methods/frameworks in the literature assume the availability of time-dependent traffic measures (such as counts, flows, speeds, etc.) on all links of the corridor. Hence, there is a critical disconnect between the practical reality and methodological expectations in terms of detection capabilities. This research seeks to develop a mechanism to strategically deploy vehicle detectors to infer network origin-destination (O-D) demands using limited link traffic count data. It leads to the problem of the identification of "optimal" locations for installing detectors so that maximum system observability is achieved with a limited monetary budget. From an integration standpoint, it addresses the question of where to locate detectors on the non-freeway facilities so that, in conjunction with the installed detectors on freeways, the entire corridor can be managed effectively by obtaining the maximum possible accurate information on traffic conditions.

The primary goal of the first stage of this project is to address the network sensor location problem (NSLP) directly so as to obtain the unobserved link flows given the minimum subset of observed link flows provided by passive counting sensors. It circumvents the data needs (in terms of turning movement proportions or prior O-D structure) or assumptions (on traffic assignment rules) associated with the O-D demand estimation problem where the NSLP is a sub-problem. A simple and efficient linear algebra based method is proposed to solve the NSLP. Given the link-path incidence matrix to represent the network topology, the concept of linear independence in relation to a set of links is used to identify the minimum subset of network links to equip with vehicle sensors so as to estimate the flows on all links. This subset of links constitutes the "basis" of the vector space of the link-path incidence matrix, and the proposed approach is labeled the basis link method. The approach does not require any assumption on road users' route choice behavioral rules and/or traffic assignment principles. Also, by solving the NSLP independently rather than as a sub-problem of a specific application, it allows applicability to many link-based applications in transportation planning and traffic management, such as pavement management systems, congestion pricing, and link strengthening for disaster response. It can also serve as a platform to address broader problems such as network O-D estimation or link travel time estimation.

Findings

This study proposes a basis link method to address the network sensor location problem under steady-state traffic conditions. A fundamental contribution of this research is the illustration of a direct mapping between the basis link flows and the non-basis link flows, which can be obtained from the network structure represented by the link-path incidence matrix. Based on the theoretical investigation and numerical analysis, several findings are listed below.

1. Given the network structure represented by its link-path incidence matrix, a theoretical minimum subset of network links provided by the reduced row echelon form (RREF) algorithm does exist, and a direct mapping between the basis link flows and the non-basis link flows is also theoretically proved.

2. The study illustrates the possibility of multiple solutions in terms of the set of basis links. However, as shown by one of the Lemmas, that does not affect the uniqueness in terms of the inferred link flows.

3. The empirical analysis highlights the primacy of the network topology in determining the set of basis links. It also indicates the possibility of an upper bound on the number of basis links based on the topology, suggesting that it may not be necessary to equip every link with sensors from a planning perspective.

4. While the number of basis links is related to the network scale in terms of the number of links/nodes or used paths under different network topologies, there is no direct correlation between the number of links/nodes or used paths in the network and the percentage of links to install sensors on.

5. The sensitivity analysis suggests no direct relationship between the percentage of links to be equipped with sensors and the number of links/nodes in a network, though a positive correlation between the number of basis links and the number of links/nodes or paths is generally observed.

Recommendations

This research has proposed a linear algebraic approach for the determination of the minimum subset of equipped links to infer the flows on the unobserved links. The proposed basis link method provides network full observability without requiring any assumptions in terms of the knowledge of O-D flows, path flows, user route choice behavior, or traffic assignment rules. This property has broader implications in terms of potentially aiding in solving a range of problems (such as O-D demand estimation, travel time estimation, traffic assignment) in both the static and dynamic contexts. Therefore, a straightforward research direction is to estimate O-D demands for a general network, in an integrated manner, based on the (full) link flow information provided by the proposed basis link approach. Further, the existence of multiple solutions provides some flexibility for traffic agencies in choosing the links to install sensors on. That is, a traffic agency may prefer to install sensors on some links because of their importance based on one or more criteria related to objectives such as minimizing deployment costs, reducing traffic impacts, or the relative importance of a link (based on the facility type or its criticality for disaster response, etc.). In such instances, priority rankings provided by the agency can be seamlessly adapted with the proposed basis link method. This is another immediate research issue that is worthy of further investigation.

Contacts

For more information:

Srinivas Peeta
Principal Investigator
Civil Engineering
Purdue University
peeta@purdue.edu

Shou-Ren Hu
Co-Principal Investigator
National Cheng Kung University, Taiwan
shouren@mail.ncku.edu.tw

NEXTRANS Center
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No 018PY01

Network Origin-Destination Demand Estimation using Limited Link Traffic Counts: Strategic Deployment of Vehicle Detectors through an Integrated Corridor Management Framework

By

Srinivas Peeta, Principal Investigator
Professor of Civil Engineering
Purdue University
peeta@purdue.edu

and

Shou-Ren Hu, Co-Principal Investigator
Assistant Professor of Transportation and Communication Management Science
National Cheng Kung University, Taiwan
shouren@mail.ncku.edu.tw

and

Chun-Hsiao Chu
Assistant Professor of Tourism
Aletheia University, Taiwan
au5907@mail.au.edu.tw

and

Han-Tsung Liou
Graduate Research Assistant of Transportation and Communication Management Science
National Cheng Kung University, Taiwan
yonex17@hotmail.com

Report Submission Date: October 15, 2009

PURDUE UNIVERSITY Discovery Park

ACKNOWLEDGMENTS

The authors would like to thank the NEXTRANS Center, the USDOT Region V Regional University Transportation Center at Purdue University, and the National Cheng Kung University, Taiwan, for supporting this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Background and motivation

Information on link flows in a vehicular traffic network is critical for developing long-term planning and/or short-term operational management strategies. In the literature, most studies to develop such strategies typically assume the availability of measured link traffic information on all network links, either through manual survey or advanced traffic sensor technologies. In practical applications, the assumption of installed sensors on all links is generally unrealistic due to budgetary constraints. It motivates the need to estimate flows on all links of a traffic network based on the measurement of link flows on a subset of links with suitably equipped sensors. This study, addressed from a budgetary planning perspective, seeks to identify the smallest subset of links in a network on which to locate sensors that enables the accurate estimation of traffic flows on all links of the network under steady-state conditions. Here, steady-state implies that the path flows are static.

1.2 Study objectives

The primary goal of this paper is to address the network sensor location problem (NSLP) directly so as to obtain the unobserved link flows given the minimum subset of observed link flows provided by passive counting sensors. It circumvents the data needs (in terms of turning movement proportions or prior O-D structure) or assumptions (on traffic assignment rules) associated with the O-D demand estimation problem where the NSLP is a sub-problem.

Also, by solving the NSLP independently rather than as a sub-problem of a specific application, it allows applicability to many link-based applications in transportation planning and traffic management, such as pavement management systems, congestion pricing, and link strengthening for disaster response. It can also serve as a platform to address broader problems such as network O-D estimation or link travel time estimation.

1.3 Organization of the research

The remainder of the paper is organized as follows. Chapter 2 introduces the basis link method in the NSLP context. Chapter 3 discusses the associated solution algorithm and various properties of the method. Chapter 4 illustrates the characteristics of the method and related insights by testing different network topologies/configurations. Finally, some concluding comments are provided in Chapter 5.

CHAPTER 2. BACKGROUND AND PROBLEM STATEMENT

This chapter provides some background information about the NSNP. Section 2.1 states the NSLP in the context of network link traffic flow and/or path flow estimation. Section 2.2 reviews the related work in the literature concerning the network sensor location problems. Section 2.3 describes the NSLP specifically investigated in this research.

2.1 Background

Link flow data in a vehicular traffic network represents valuable information to address long-term planning and/or short-term operational needs. For instance, link traffic flow measurements can be used to infer a trip origin-destination (O-D) demand table for the network (for example, Maher, 1983; Cascetta and Nguyen, 1988; Cascetta *et al*., 1993; Ashok and Ben-Akiva, 2002). While the various methods have different degrees of capabilities to predict the O-D demand estimates, most of the early literature assumes that traffic flows are available for each network link or can be collected at specific locations, such as traffic counting stations on a screen line (Wu and Chang, 1996) or cordon line (Chang and Tao, 1999). Similarly, deployable dynamic traffic assignment (Peeta and Ziliaskopoulos, 2001) frameworks proposed to manage the dynamics of traffic congestion typically assume the availability of link traffic flows on all network links in the determination of the assignment strategies or related consistency-checking procedures (Peeta and Bulusu, 1999; Ben-Akiva *et al*., 2001; Zhou and Mahmassani, 2005). In summary, an implicit assumption for most problems requiring such link flow data is that the network is equipped with an extensive advanced traffic management system that enables the collection of link flow data on all links.

The assumption of a network-wide traffic sensor system may not be realistic for practical applications due to the budgetary constraints of traffic management agencies. Nevertheless, solution methods for many planning and operational problems associated with traffic networks implicitly assume that sensors are installed on all network links. An urban network of moderate size can entail substantial costs to deploy a large number of sensors. Since the quantity and quality of the collected traffic flow information significantly affects the estimation accuracy and reliability of network traffic flow estimates, there is a trade-off between the prediction accuracy of network traffic flow estimates and the cost associated with the extent of deployment of a sensor system. It motivates the need to address the problem of optimal sensor locations under a limited budget: Can we identify a minimum subset of sensor installed links and their locations for accurate vehicular flow estimates throughout the network ? This problem or variants thereof can be broadly labelled as the network sensor location problem (NSLP). Over the past decade, the NSLP has been typically addressed as a sub-problem of broader problems related to O-D demand estimation, time-dependent link travel time estimation, and operational consistency-seeking procedures.

2.2 Network sensor location problems

In the literature, the NSLP has been mostly addressed as a sub-problem of the broader O-D demand estimation problem (Yang and Zhou, 1998; Bianco *et al*., 2001; Gan *et al*., 2005; Yang *et al*., 2006; Ehlert *et al*., 2006), rather than as an independent problem in the context of link-based applications. Thereby, it has been used to determine the minimum number of sensor locations (for example, counting stations), or the optimal locations for a given number of sensors, to estimate the O-D demand. Such problems typically assume the availability of the turning proportions at a node (e.g. Bianco *et al*., 2001), or a link usage proportion matrix (obtained using an appropriate traffic assignment procedure) which captures the proportion of O-D trips of a given path that traverse a specific link (e.g. Gan *et al*., 2005). This data is not readily available in many field applications. However, such assumptions enable the formulation of the broader O-D estimation problem as an integer program (Yang and Zhou, 1998; Gan *et al*., 2005; Yang *et al*., 2006), a mixed integer program (Ehlert *et al*., 2006), or a mathematical program which minimizes the cost associated with sensor installation (Bianco *et al*., 2001). The associated NSLP sub-problem has then been typically solved using a combination of column generation procedures and branch-and-bound heuristics, or genetic algorithms.

As part of the O-D estimation problem, Yang *et al*. (2006) develop models and algorithms to address two screen-line based sensor location problems: how to select the optimal locations of a given number of counting stations to separate as many O-D pairs as possible, or how to determine the minimum number of counting stations and their locations to separate all O-D pairs? They require the satisfaction of pre-defined O-D covering and link independence rules (Yang and Zhou, 1998) as constraints in the formulation of the problems. The O-D covering rule states that the traffic sensors on a road network should be located such that a certain portion of trips between any O-D pair are observable. The link independence rule states that the sensor locations should ensure the linear independence of the traffic counts on the chosen links. They require the availability of a link-path incidence matrix and a historical O-D structure to solve these problems.

The NSLP is an analog of the "observability" problem in linear system of equations. Castillo *et al*. (2007, 2008a, 2008b) address the observability problem using the algebraic techniques of the null-space method. Castillo *et al*., 2008c proposed a Bayesian updating approach to solve the observability problem in a traffic network, where the minimum subset of links is determined to equip with vehicle sensors so as to infer the O-D flows and/or unequipped link flows. The Bayesian network model is also used to determine the optimal number and locations of the link counters based on a maximum correlation criterion. However, these approaches require either a known matrix relating link and O-D flows (e.g., the ***F*** matrix in Castillo *et al*., 2008a, 2008b) or route choice probabilities given by some traffic assignment rules (e.g., the stochastic user equilibrium principle adopted in Castillo *et al*., 2008c) to formulate the flow conservation equation. Further, prior knowledge on O-D flows and model parameters is assumed in the initiation of the solution procedure. The unknown O-D and unobserved link flows are obtained using O-D and link flows collected at some strategic links equipped with vehicle sensors or advanced data collection systems. The network flow estimation results are generally better than those obtained through only link flow observations, since O-D flows collected, for instance, via the license plate recognition technique provide more information for path flow reconstruction and O-D flow estimation (Castillo *et al*., 2008d).

In summary, the algebraic null-space method or Bayesian network model solve the network observability problem as a sub-problem of the broader O-D demand estimation problem, assuming some known user route choice decisions (e.g. Castillo *et al*., 2008c) and/or prior O-D demand structures (e.g., Castillo *et al*., 2008d). However, the assumptions of prior knowledge on model parameters and/or O-D demands can constrain the applicability of these approaches in practice.

Some studies (e.g., Castillo *et al*., 2008d; Gentili and Mirchandani, 2005) have suggested that the use of active sensors along with techniques such as license plate recognition provides more information to determine network path and/or O-D flows. However, it is very expensive to deploy a comprehensive infrastructure to actively collect traffic flow and path information of the equipped vehicles. Also, the main purpose of locating active sensors in a traffic network is to obtain sufficient information on flows on specified paths and/or O-D pairs (e.g., Gentili and Mirchandani, 2005), with or without link flow information provided by passive counting sensors.

2.3 NSLP problem statement

Given the network structure (link-path incidence matrix), we seek to identify the minimum subset of links on which to locate sensors so as to infer the flows on all links under steady-state conditions. Here, steady-state implies that the path flows are static. Hence, the objective is to address the NSLP in a planning context by focusing on long-term steady-state conditions, motivated by the limited budget available to purchase and install sensors. Thereby, the problem is static in nature, and the estimated link flows (for example, the AADT or ADT) are based on the average flows on the sensor-equipped links over a period of time.

The primary goal of this paper is to address the NSLP directly so as to obtain the unobserved link flows given the minimum subset of observed link flows provided by passive counting sensors. It circumvents the data needs (in terms of turning movement proportions or prior O-D structure) or assumptions (on traffic assignment rules) associated with the O-D demand estimation problem where the NSLP is a sub-problem. A simple and efficient linear algebra based method is proposed to solve the NSLP. Given the link-path incidence matrix to represent the network topology, the concept of linear independence in relation to a set of links is used to identify the minimum subset of network links to equip with vehicle sensors so as to estimate the flows on all links. This subset of links constitutes the “basis” of the vector space of the link-path incidence matrix, and the proposed approach is labeled the basis link method. The approach does not require any assumption on road users’ route choice behavioral rules and/or traffic assignment principles. Also, by solving the NSLP independently rather than as a sub-problem of a specific application, it allows applicability to many link-based applications in transportation planning and traffic management, such as pavement management systems, congestion pricing, and link strengthening for disaster response. It can also serve as a platform to address broader problems such as network O-D estimation or link travel time estimation.

CHAPTER 3. BASIS LINK METHOD

This chapter illustrates an algebraic-based approach to deal with the NSNP. Section 3.1 introduces the concepts of "basis" in a vector space and the "rank" of a matrix, and illustrate their linkage in the context of the problem. Section 3.2 introduces the notion of using a link-path incidence matrix to represent a network structure in the context of the proposed basis link method to solve the NSLP. This method is used to identify the basis links, and then derive the non-basis link flows through information contained in the basis link flows measured using the sensors.

3.1 The basis

"Basis" is a key concept associated with a vector space (Friedberg *et al*., 2003).

Definition 1. *A basis* $\boldsymbol{\beta}$ *for a vector space* $\boldsymbol{V}$ *is a linearly independent subset of* $\boldsymbol{V}$ *that generates* $\boldsymbol{V}$.

By definition, the dimension l of $\boldsymbol{V}$ is the cardinality of a basis $\boldsymbol{\beta}$ of $\boldsymbol{V}$. Then, any linearly independent subset of $\boldsymbol{V}$ that contains exactly l vectors is a basis for $\boldsymbol{V}$.

A matrix space is also a vector space. The rank of a matrix is defined as below:

Definition 2. *If* $\boldsymbol{A} \in \boldsymbol{M}_{m\times n}(\boldsymbol{H})$, *the rank of* $\boldsymbol{A}$, *denoted rank*($\boldsymbol{A}$), *is the rank of the linear transformation* $\lambda_{\boldsymbol{A}} : \boldsymbol{H}^n \to \boldsymbol{H}^m$, where $\boldsymbol{H}$ is some field.

The *rank* of a matrix has the following properties:

(i) The rank of any matrix equals the maximum number of its linearly independent columns; that is, the rank of a matrix is the dimension of the subspace generated by its columns which is called column space.

(ii) The rank of any matrix equals the maximum number of its linearly independent rows; that is, the rank of a matrix is the dimension of the subspace generated by its rows which is called row space.

(iii)The rows and columns of any matrix generate subspaces of the same dimension numerically equal to the rank of the matrix; that is, in a specific matrix the rank of the column space equals the rank of the row space.

(iv)Let $\boldsymbol{A}$ be an (m×n) matrix of rank r. Then $r \leq m$, and $r \leq n$.

(v) rank($\boldsymbol{A}^T$) = rank($\boldsymbol{A}$)

If $\boldsymbol{\beta} = \{\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_l\}$ is a basis for $\boldsymbol{V}$ and the matrix $\boldsymbol{B} = [\boldsymbol{B}_1\ \boldsymbol{B}_2 \cdots \boldsymbol{B}_l]$, then any member $\boldsymbol{v} \in \boldsymbol{V}$ can be written uniquely in the form $\boldsymbol{v} = \boldsymbol{Bw}$, where $\boldsymbol{w}^T = \{w_1, w_2, \ldots, w_l\}$ is a vector of scalar coefficients. Thereby, it is useful to relax the nomenclature and call the matrix $\boldsymbol{B}$ along with the set $\boldsymbol{\beta}$ a basis for $\boldsymbol{V}$ (Stewart, 1998). Then, any matrix $\boldsymbol{A}_{m\times n}$ in the vector space $\boldsymbol{V}$ can be represented by linear combinations of the elements $\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_l$ in $\boldsymbol{B}$. The maximum number of linearly independent column (or row) vectors in $\boldsymbol{A}_{m\times n}$, which represents the rank of $\boldsymbol{A}_{m\times n}$, is equal to the number of linearly independent vectors of $\boldsymbol{\beta}$, which is its cardinality l. This relates the rank of a matrix $\boldsymbol{A}_{m\times n}$ to the cardinality of a basis $\boldsymbol{\beta}$.

3.2 Link-path incidence matrix and basis links

A link-path incidence matrix is a 0-1 matrix that describes the network structure through the spatial relationships between the paths and links of that network. This matrix can be represented through a set of column or row vectors. If $\boldsymbol{L}_{m\times n}$ denotes the link-path incidence matrix with m paths and n links, it can be expressed as:

$$\boldsymbol{L}_{m\times n} = [\boldsymbol{L}_1\ \boldsymbol{L}_2 \cdots \boldsymbol{L}_j \cdots \boldsymbol{L}_n]$$

where $\boldsymbol{L}_j$ is the j^{th} column vector of dimension (m×1). The basis of the vector space associated with $\boldsymbol{L}_{m\times n}$ consists of l linearly independent column vectors, and the links corresponding to these columns are called the *basis links*. The remaining links in the network are called the *non-basis links*. If the flows on the basis links are observed using sensors, then by definition of basis, the flows on all links can be inferred through linear combinations of the basis link flows. This conceptual platform is used here to address the NSLP.

CHAPTER 4. SOLUTION ALGORITHM

Chapter 4 introduces the solution algorithm for the NSLP. To solve the NSLP, Section 4.1 illustrates the concept of "reduced row echelon form" (RREF) (Friedberg *et al*., 2003) to identify the basis links of the link-path incidence matrix. Section 4.2 discusses the properties of the RREF. An example is provided in to illustrate the determination of the basis links of $\boldsymbol{L}_{m\times n}$ using its RREF. In Section 4.3, the proof for inferring the non-basis link flows from the basis link flows is provided. Section 4.4 explores implications of multiple solutions in the context of the set of basis links.

4.1 Reduced row echelon form

A matrix is said to be in its "reduced row echelon form" if it satisfies the following three conditions (Friedberg *et al*., 2003):

I. Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).

II. The first nonzero entry in each row is the only nonzero entry in this column.

III. The first nonzero entry in each row is 1 and it appears in a column to the right of the leading 1 in any preceding row. By definition, if the first non-zero number in a row is 1, it is called the leading 1.

The RREF for the link-path incidence matrix can be obtained using the Gaussian elimination algorithm. The associated steps are as follows (Anton, 1984):

Step 1: Locate the leftmost column of $\boldsymbol{L}$ that does not consist entirely of zeros.

Step 2: Swap the top row with another row, if needed, to bring a nonzero entry to the top of the column found in *Step 1*.

Step 3: If the entry that is now at the top of the column found in *Step 1* is γ, multiply the first row by $1/\gamma$ in order to introduce a leading 1.

Step 4: Add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zeros in that column.

Step 5: Cover the top row in the matrix and begin again with *Step 1* applied to the sub-matrix that remains. Continue in this way until the entire matrix is in row-echelon form.

Step 6: Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's. The resulting matrix represents the RREF of $\boldsymbol{L}$.

Let $\boldsymbol{L}$ be an $(m\times n)$ matrix of rank r ($r > 0$) with column vectors $\boldsymbol{L}_1, \boldsymbol{L}_2,\ldots, \boldsymbol{L}_n$, and let $\boldsymbol{T}$ be the RREF of $\boldsymbol{L}$. Denote the column vectors of $\boldsymbol{T}$ by $\boldsymbol{t}_1, \boldsymbol{t}_2,\ldots, \boldsymbol{t}_n$. The RREF $\boldsymbol{T}$ has the following properties (Friedberg *et al*., 2003):

(a) The number of nonzero rows in $\boldsymbol{T}$ is r.

(b) For each $k = 1, 2, \ldots, r$, there is a column vector $\boldsymbol{t}_{j_k}$ of $\boldsymbol{T}$ such that $\boldsymbol{t}_{j_k} = \boldsymbol{e}_k$, where $\boldsymbol{e}_k$ is an $(m\times 1)$ unit column vector whose k^{th} row element is 1. $\boldsymbol{t}_j$ is the j^{th} column vector in $\boldsymbol{T}$.

(c) The column vectors of $\boldsymbol{L}$, numbered $\boldsymbol{L}_{j_1}, \boldsymbol{L}_{j_2}, \ldots, \boldsymbol{L}_{j_r}$, are linearly independent and denote the basis of the vector space associated with $\boldsymbol{L}$.

(d) The reduced row echelon form of a matrix is *unique*.

Next, re-arrange $\boldsymbol{T}$ so that its first r columns are the linearly independent unit column vectors; $\boldsymbol{t}_{j_k} = \boldsymbol{e}_k$ where $k = 1, 2, \ldots, r$. Then, for consistency, we also re-arrange the column vectors $\boldsymbol{L}_{j_1}, \boldsymbol{L}_{j_2}, \ldots, \boldsymbol{L}_{j_r}$ to be the first r columns in $\boldsymbol{L}$. Based on the re-arranged $\boldsymbol{T}$ and $\boldsymbol{L}$ matrices, the following property holds (Friedberg *et al*., 2003):

(e) For each $j = 1, 2,\ldots, n$, if the j^{th} column vector of $\boldsymbol{T}$ is $\alpha_1\boldsymbol{e}_1 + \alpha_2\boldsymbol{e}_2 + \ldots + \alpha_r\boldsymbol{e}_r$, then the j^{th} column vector of $\boldsymbol{L}$ is $\alpha_1\boldsymbol{L}_{j_1} + \alpha_2\boldsymbol{L}_{j_2} + \cdots + \alpha_r\boldsymbol{L}_{j_r}$, where $\alpha_1, \alpha_2, \ldots, \alpha_r$ are the linear combination coefficients.

We now state Lemma 1 assuming the re-arranged ***T*** and ***L*** matrices.

Lemma 1. *Any column vector* $\boldsymbol{t}_j$ $(j = 1,2,\ldots,n)$ *in* ***T*** *can be represented by a linear combination of a set of r unit column vectors whose linear combination coefficients* α *are the column elements in* $\boldsymbol{t}_j$ *corresponding to the r non-zero rows of* ***T****. That is, an* (m×1) *column vector* $\boldsymbol{t}_j = [\alpha_1\, \alpha_2 \cdots \alpha_r\; 0 \cdots 0]^T$ *in* ***T*** *can be represented as* $\boldsymbol{t}_j = \alpha_1 \boldsymbol{e}_1 + \alpha_1 \boldsymbol{e}_2 + \cdots + \alpha_r \boldsymbol{e}_r$.

Proof.

By property (a) of the RREF, the number of nonzero rows in ***T*** is *r*, the rank of ***L***. By property (iv) of the rank of ***L***, $r \le m$, and $r \le n$; and by property (iii) of the rank of ***L***, the rank of the column space is equal to that of the row space. Let us assume that the row rank is *r* and $m \le n$. Then, by definition $r \le m$. There are two possibilities: $r = m$ or $r < m$.

(1) The rank of ***L*** is equal to the number of rows in ***L*** ($r = m$):

When the rank of ***L*** is equal to the number of rows in ***L***, it implies the first *r* column unit vectors in ***T*** constitute an (r×r) square unit matrix. Then, by the definition of unit column vector, any column vector $\boldsymbol{t}_j = [\alpha_1\, \alpha_2 \cdots \alpha_r]^T$ in ***T*** can be represented as $\boldsymbol{t}_j = \alpha_1 \boldsymbol{e}_1 + \alpha_1 \boldsymbol{e}_2 + \cdots + \alpha_r \boldsymbol{e}_r$.

(2) The rank of ***L*** is less than the number of rows in ***L*** ($r < m$):

When the rank of ***L*** is less than the number of rows in ***L***, by property (a) of the RREF, the number of nonzero rows in ***T*** is *r*. It implies that the bottom (m-r) rows of ***T*** are zero rows. Then, any column vector $\boldsymbol{t}_j$ in ***T*** has the form $\boldsymbol{t}_j = [\alpha_1\, \alpha_2 \cdots \alpha_r\; 0 \cdots 0]^T$, with the last ($m$-$r$) elements being zeros. Also, based on the re-arrangement of ***T*** and property (b), there is a sub-matrix in ***T*** which is an (m×r) unit matrix whose column vectors are *r* unit column vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, …. , $\boldsymbol{e}_r$ of dimension (m×1). Therefore, any column vector $\boldsymbol{t}_j = [\alpha_1\ \alpha_2 \cdots \alpha_r\ 0 \cdots 0]^T$ in $\boldsymbol{T}$ can be represented as $\boldsymbol{t}_j = \alpha_1 \boldsymbol{e}_1 + \alpha_1 \boldsymbol{e}_2 + \cdots + \alpha_r \boldsymbol{e}_r$.

This logic can be repeated if $n \le m$, that is, $r \le n$.

This completes the proof. ■

4.2 RREF and the basis links

Property (c) in Section 4.1 indicates how the basis for ***L*** can be identified after determining its RREF, ***T***. That is, the RREF can be used to identify the set of basis links. Now, we demonstrate the procedure for identifying the basis links for a given network using the RREF.

Table 4.1 illustrates the link-path incidence matrix ***L*** for the network shown in Figure 4.1. The network has 10 nodes and 10 links. Node 1 is the origin node, and nodes 9 and 10 are the destination nodes. By using the Gaussian elimination algorithm described in the previous section, the RREF ***T*** of ***L*** is obtained, and is shown in Table 4.2.

According to property (a) of the RREF, the rank of the link-path incidence matrix is 3. Hence, for this network, we need only 3 vehicle sensors for inferring the link flow information on all links. Hence, only 30% of the links need to be installed vehicle sensors. Further, by property (c), the first, second, and ninth columns (denoted by $link_{1-2}$, $link_{2-3}$, and $link_{8-9}$) are linearly independent, and the corresponding links (shown using dashed arrows in Figure 4.1) are the basis links. Therefore, we can use the information of these three links to completely describe the network structure. By property (e) of the RREF, the column vector corresponding to a specific non-basis link in ***L*** can be represented by a linear combination of the column vectors associated with the basis links whose coefficients are identical to the linear combination coefficients used to obtain the non-basis column vectors in the RREF of the link-path incidence matrix[1].

[1] Gentili and Mirchandani (2005) use the rank of a link-path incidence matrix ***B*** to verify if a system of linear equations is determinable, and to determine the number and locations of active sensors on specific links in order to provide *additional path flow* information to infer the complete flow estimates on all paths. By contrast, the rank of a given link-path incidence matrix ***L*** in the proposed basis link method is used to determine the number of (independent) basis link vectors in order to provide link flow estimates on the

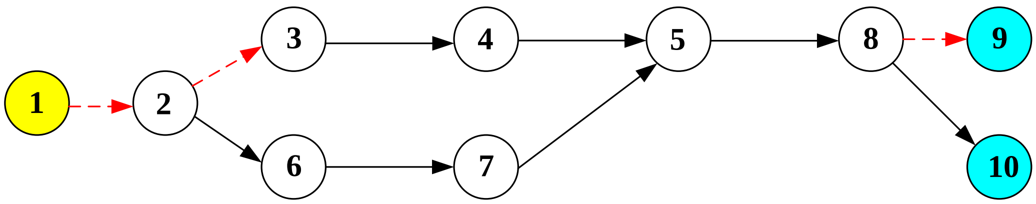


Figure 4.1. Example Network.

Table 4.1. Link-path incidence matrix of the network in Figure 4.1.

Link number	1	2	3	4	5	6	7	8	9	10
link / path	1-2	2-3	2-6	3-4	4-5	5-8	6-7	7-5	8-9	8-10
1-2-6-7-5-8-9	1	0	1	0	0	1	1	1	1	0
1-2-3-4-5-8-9	1	1	0	1	1	1	0	0	1	0
1-2-6-7-5-8-10	1	0	1	0	0	1	1	1	0	1
1-2-3-4-5-8-10	1	1	0	1	1	1	0	0	0	1

Table 4.2. RREF of the link-path incidence matrix

Link number	1	2	3	4	5	6	7	8	9	10
link / path	1-2	2-3	2-6	3-4	4-5	5-8	6-7	7-5	8-9	8-10
1-2-6-7-5-8-9	**1**	0	1	0	0	1	1	1	0	1
1-2-3-4-5-8-9	0	**1**	-1	1	1	0	-1	-1	0	0
1-2-6-7-5-8-10	0	0	0	0	0	0	0	0	**1**	-1
1-2-3-4-5-8-10	0	0	0	0	0	0	0	0	0	0

Table 4.3. Link-path incidence matrix when $link_{5-8}$ is denoted as link 1

Link number	1	2	3	4	5	6	7	8	9	10
link / path	5-8	2-3	2-6	3-4	4-5	1-2	6-7	7-5	8-9	8-10
1-2-6-7-5-8-9	1	0	1	0	0	1	1	1	1	0
1-2-3-4-5-8-9	1	1	0	1	1	1	0	0	1	0
1-2-6-7-5-8-10	1	0	1	0	0	1	1	1	0	1
1-2-3-4-5-8-10	1	1	0	1	1	1	0	0	0	1

unequipped links via linear combination of the *observed link flows*.

Table 4.4. RREF of the link-path incidence matrix where $link_{5-8}$ is denoted as link 1

Link number	1	2	3	4	5	6	7	8	9	10
link / path	5-8	2-3	2-6	3-4	4-5	1-2	6-7	7-5	8-9	8-10
1-2-6-7-5-8-9	**1**	0	1	0	0	1	1	1	0	1
1-2-3-4-5-8-9	0	**1**	-1	1	1	0	-1	-1	0	0
1-2-6-7-5-8-10	0	0	0	0	0	0	0	0	**1**	-1
1-2-3-4-5-8-10	0	0	0	0	0	0	0	0	0	0

Property (d) indicates that the RREF of the link-path incidence matrix is unique. However, it does not imply that the basis for the vector space represented by a link-path incidence matrix is unique because the order of the basis links in the link-path incidence matrix is arbitrary. For example, if we swap the positions of $link_{1-2}$ and $link_{5-8}$ to obtain a different link-path incidence matrix (shown in Table 4.3) for the same network, its corresponding RREF is shown in Table 4.4. Here, the basis links are $link_{5-8}$, $link_{2-3}$, and $link_{8-9}$. This result is consistent with the network structure because the flows on $link_{1-2}$ and $link_{5-8}$ are identical. Note that the RREF in Table 4.4 is identical to that in Table 4.2. However, it represents a special case because some columns in $\boldsymbol{L}$ have identical elements; it also illustrates the notion of multiple solutions for the set of basis links. In general, the RREF for different link-path incidence matrices will be different.

4.3 Inferring non-basis link flows from basis link flows

For a traffic network under steady-state conditions, by the definition of link-path incidence matrix, the (1×n) link flow matrix $\boldsymbol{F}$ can be obtained as the product $\boldsymbol{P}^T\boldsymbol{L}$[2], where $\boldsymbol{P}$ is the static path flow matrix of the network of dimension (m×1).

[2] This is similar to the flow conservation equation described in Castillo *et al*. (2008a; 2008b) where a known matrix (**F**) relating link and O-D flows is obtained from the network topology.

Let $\boldsymbol{P}_{m\times 1} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix}$, $\boldsymbol{L}_{m\times n} = \begin{bmatrix} L_{11} & L_{12} & L_{13} & \cdots & L_{1n} \\ L_{21} & L_{22} & L_{23} & \cdots & L_{2n} \\ L_{31} & L_{32} & L_{33} & \cdots & L_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & L_{m3} & \cdots & L_{mn} \end{bmatrix}$.

Then,

$$\boldsymbol{F}_{1\times n} = \boldsymbol{P}^T \boldsymbol{L} = \left[\sum_{i=1}^{m} p_i L_{i1} \quad \sum_{i=1}^{m} p_i L_{i2} \quad \cdots \quad \sum_{i=1}^{m} p_i L_{in} \right] \qquad (1)$$

where L_{ij} is 0 or 1; and p_i is the flow on the i^{th} path.

Theorem 1. *The non-basis link flows in the link flow matrix **F** can be obtained through a linear combination of the basis link flows whose coefficients are identical to the elements of the corresponding non-basis column vectors in the RREF of the link-path incidence matrix **L**.*

Proof.

Suppose r is the rank of $\boldsymbol{L}_{m\times n}$; then, based on the discussion in Section 2, there are r basis links. Let $\boldsymbol{L}^{\boldsymbol{B}}{}_{m\times r}$ be the set of column vectors corresponding to the basis links in the network; $\boldsymbol{L}^B = [\boldsymbol{L}_1^B\ \boldsymbol{L}_2^B \cdots \boldsymbol{L}_r^B]$, where $\boldsymbol{L}_j^B$ is the j^{th} (m×1) basis column vector in $\boldsymbol{L}$. Let $\boldsymbol{L}^{\boldsymbol{NB}}{}_{m\times(n\text{-}r)}$ be the set of column vectors corresponding to the non-basis links in the network; $\boldsymbol{L}^{NB} = [\boldsymbol{L}_1^{NB}\ \boldsymbol{L}_2^{NB} \cdots \boldsymbol{L}_{n-r}^{NB}]$, where $\boldsymbol{L}_j^{NB}$ is the j^{th} (m×1) non-basis column vector in $\boldsymbol{L}$. Then, $\boldsymbol{L}$ can be rewritten as:

$$\boldsymbol{L}_{m\times n} = [\boldsymbol{L}^B\ \boldsymbol{L}^{NB}] \qquad (2)$$

Also, let $\boldsymbol{F}_{1\times r}^B$ be the matrix of basis link flows; $\boldsymbol{F}^B = [F_1^B\ F_2^B \cdots F_r^B]$, where F_j^B is the j^{th} basis link flow. $F_j^B = \sum_{i=1}^{m} p_i L_{ij}^B = \boldsymbol{P}^T \boldsymbol{L}_j^B$. Let $\boldsymbol{F}_{1\times(n-r)}^{NB}$ be the matrix of non-basis link flows;

$\boldsymbol{F}^{NB} = [F_1^{NB} F_2^{NB} \cdots F_{n-r}^{NB}]$, where F_j^{NB} is the j^{th} non-basis link flow. $F_j^{NB} = \sum_{i=1}^{m} p_i L_{ij}^{NB} = \boldsymbol{P}^T \boldsymbol{L}_j^{NB}$. Then, $\boldsymbol{F}$ can be rewritten as: $\boldsymbol{F}_{1\times n} = [\boldsymbol{F}^B \; \boldsymbol{F}^{NB}]$, and by extension:

$$\boldsymbol{F}_{1\times n} = [\boldsymbol{F}^B \; \boldsymbol{F}^{NB}] = \boldsymbol{P}^T [\boldsymbol{L}^B \; \boldsymbol{L}^{NB}] \tag{3}$$

We will now relate the non-basis column vectors in $\boldsymbol{L}$ to its basis column vectors. By the definition of basis, the j^{th} non-basis column vector of $\boldsymbol{L}$ (that is, $\boldsymbol{L}_j^{NB}$) can be expressed as the linear combination of the r linearly independent basis column vectors. From properties (c) and (e) of the RREF, the associated linear combination coefficients are identical to the linear combination coefficients used to obtain the j^{th} column vector of $\boldsymbol{T}$. Lemma 1 illustrates that these coefficients ($\alpha_{j_k j}$) are the column elements in the j^{th} non-basis column vector of $\boldsymbol{T}$. Thereby:

$$\boldsymbol{L}_j^{NB} = \sum_{j_k=1}^{r} \alpha_{j_k j} \boldsymbol{L}_{j_k}^B, \; j = 1, 2, \ldots, n-r \tag{4}$$

where the scalar $\alpha_{j_k j}$ is the linear combination coefficient corresponding to the j_k^{th} row in the j^{th} non-basis column vector in $\boldsymbol{T}$. We will now express F_j^{NB} in terms of $F_{j_k}^B$:

$$\begin{aligned}
F_j^{NB} &= \sum_{i=1}^{m} p_i L_{ij}^{NB}, \; j = 1, 2, \ldots, n-r \\
&= \mathbf{P}^T \mathbf{L}_j^{NB} \\
&= \mathbf{P}^T \sum_{j_k=1}^{r} \alpha_{j_k j} \mathbf{L}_{j_k}^B \\
&= \sum_{j_k=1}^{r} \alpha_{j_k j} \mathbf{P}^T \mathbf{L}_{j_k}^B \\
&= \sum_{j_k=1}^{r} \alpha_{j_k j} F_{j_k}^B, \quad j = 1, 2, \ldots, n-r.
\end{aligned} \tag{5}$$

This completes the proof.

Theorem 1 illustrates a key characteristic of the proposed basis link method in solving the NSLP. It indicates that a direct mapping exists between the basis link flows and the non-basis link flows which can be obtained from the network structure (represented by the link-path incidence matrix).

For the example network in Section 3.2, Theorem 1 indicates that sensors are required on only 3 of the 10 links so as to establish flows on all links. As can be seen in Figure 4.1, if flows are observed on $link_{2\text{-}3}$, there is no need to observe them on $link_{3\text{-}4}$ and $link_{4\text{-}5}$.

From Theorem 1, a direct observation is that links with the same column vector elements in the RREF of the link-path incidence matrix have the same flows, as discussed in Lemma 2.

Lemma 2. *Links which have identical column vector elements in the RREF* $\boldsymbol{T}$ *of* $\boldsymbol{L}$ *have identical flows.*

Proof.

Based on the re-arranged $\boldsymbol{T}$, we have $\boldsymbol{T} = [\boldsymbol{T}^B\ \boldsymbol{T}^{NB}]$. From Theorem 1:

$$F_j^{NB} = \sum_{b=1}^{r} \alpha_{bj} F_b^B,\ j = 1, 2, ..., n-r. \tag{6}$$

Suppose the column elements in link l are identical to those of link j. By property (d) of the RREF, since RREF $\boldsymbol{T}$ of $\boldsymbol{L}$ is unique, if we swap links l and j the RREF is identical to the previous one. There are two possibilities:

(1) When link l is a basis link and link j is a non-basis link:

Since link l is a basis link, its RREF column vector is a unit column vector with the l^{th} row element being 1. As the column vector elements in both links l and j are identical, it means that $\alpha_{bj} = 0, \forall b$ except $b = l,\ l \le r,$ and $\alpha_{lj} = 1$. From (6):

$\Rightarrow F_j^{NB} = F_l^{B}$.

(2) When both links *l* and *j* are non-basis links:

Since the column vector elements in both the non-basis links are identical, it implies that $\alpha_{bj} = \alpha_{bl}, \forall b$. From (6):

$\Rightarrow F_j^{NB} = F_l^{NB}$.

This completes the proof.

The inference in Lemma 2 can be obtained by simply observing $\boldsymbol{L}$ directly. This is because column vectors with identical elements in the RREF also imply that the corresponding column vectors in $\boldsymbol{L}$ are identical. Since $\boldsymbol{F} = \boldsymbol{P}^T \boldsymbol{L}$, and $\boldsymbol{P}$ is the static path flow vector, column vectors in $\boldsymbol{L}$ which are identical will imply identical flows for the corresponding links.

It should be noted here that the network topology is a key determinant of the number of sensors to be installed. As a starting point, since the rank of a link-path incidence matrix is not greater than the number of rows or columns, the number of the basis links is not greater than the number of paths or links of the network. However, in a general traffic network where the number of paths is typically greater than the number of links, if the *rank* of the link-path incidence matrix is equal to the number of links in the network, the minimum subset of basis links will imply all links of the network. This is not a limitation of the proposed approach, but a reflection of the primacy of the network topology in the NSLP context. Further, as will be illustrated in Section 4.5, there may be an upper bound on the number of basis links that is governed by the network topology irrespective of the total number of links in the network. That is, there will be practical instances where the number of basis links will always be less than the total number of network links. This suggests that it is beneficial to use the basis link method independent of the scale of the network.

Another aspect to be noted is the possibility of the existence of multiple solutions in terms of the set of basis links. From a practical standpoint, it would imply a different subset of links on which sensors should be installed. However, this will still imply a unique set of link flows for the network links as shown in Lemma 3.

Lemma 3. *The network link flows inferred by different sets of basis links are unique.*

Proof.

For a given network, assume that $\boldsymbol{B}$ represents the set of basis links. It allows the partitioning of $\boldsymbol{L}$ into $[\boldsymbol{L}^B\ \boldsymbol{L}^{NB}]$.

From (3),

$$\boldsymbol{F}_1 = \boldsymbol{P}^T\boldsymbol{L} = \boldsymbol{P}^T[\boldsymbol{L}^B\ \boldsymbol{L}^{NB}].$$

If multiple solutions exist for this network in terms of the set of basis links, let ***B'*** represents another set of basis links. Then:

$$\boldsymbol{F}_2 = \boldsymbol{P}^T\boldsymbol{L}' = \boldsymbol{P}^T[\boldsymbol{L}^{B}{'}\ \boldsymbol{L}^{NB}{'}]$$

This implies that at least one basis link *j* in $\boldsymbol{B}$ has shifted to the set of non-basis links ***NB'***. However, the elements in $\boldsymbol{L}$ of the column vector associated with the shifted link remain unchanged. That is:

$$\boldsymbol{P}^T\boldsymbol{L}_j^B = \boldsymbol{P}^T\boldsymbol{L}_j^{NB}{'} = F_j.$$

This is true for all links that shift between the basis link set and the non-basis link set.

Hence, flows are unique irrespective of whether the associated link is in the set of basis links or non-basis links.

This completes the proof.

4.4 Discussion on multiple solutions

As illustrated through the example in Section 4.2, multiple solutions in terms of the set of basis links can exist for a given network structure represented by the link-path incidence matrix $\boldsymbol{L}$. Since the ordering of the column vectors in $\boldsymbol{L}$ is arbitrary, the column positions of the links in $\boldsymbol{L}$ can decide the set of basis links due to the steps of the Gaussian elimination algorithm for the RREF. Hence, if multiple solutions exist, they can be determined by simply varying the specific locations of link column vectors in $\boldsymbol{L}$. In the algorithm, the leftmost columns are processed first to identify basis links. Hence, if traffic agencies prioritize links in some order of importance, they can be considered seamlessly in the proposed basis link method by assigning the higher priority links to the leftmost columns in ***L***. Next, we identify some characteristics of ***L*** or ***T*** that imply multiple solutions in terms of the basis link set.

*4.4.1. Identical column elements in **L*** or ***T***

By Lemma 2, any pair of columns with identical coefficients in ***L*** or ***T*** will have the same link flows. Hence, a non-basis link with a column vector in ***L*** or ***T*** that is identical to that of a basis link can be swapped with the basis link to enter the basis. This implies multiple solutions in terms of the set of basis links.

*4.4.2. Swappability of the column pairs in **T***

If a pair of columns in ***T*** does not have identical column elements, multiple solutions exist if they are swappable. Here *swappability* means that after swapping any column pair, the non-basis link becomes a basis link, and the basis link becomes a non-basis link. To illustrate swappability, we will derive one condition under which it exists as an example.

Let the column elements for a pair of swappable links be different. Let us assume that after the swap, the column elements of the j^{th} non-basis link (which was previously the l^{th} basis link) are identical to the column elements in the original j^{th} non-basis link (which is now the l^{th} basis link). If the column elements are $[\alpha_{1j}\ \alpha_{2j} \cdots \alpha_{lj} \cdots \alpha_{rj}]^T$, then from (5):

$$F_j^{NB} = \sum_{b=1}^{l-1} \alpha_{bj} F_b^B + \alpha_{lj} F_l^B + \sum_{b=l+1}^{r} \alpha_{bj} F_b^B \qquad (7)$$

$$F_j^B = \sum_{b=1}^{l-1} \alpha_{bj} F_b^B + \alpha_{lj} F_l^{NB} + \sum_{b=l+1}^{r} \alpha_{bj} F_b^B \qquad (8)$$

Subtracting Eq. (8) from Eq. (7):

$$F_j^{NB} - F_l^{B} = \alpha_{lj}(F_l^{B} - F_j^{NB})$$
$$\Rightarrow \alpha_{lj} = -1$$

Moreover, since link *l* is the basis link whose column elements are a unit column vector with the l^{th} row element being 1, $\alpha_{ll} = 1$. Hence, if $\alpha_{ll} = 1$ and $\alpha_{lj} = -1$, multiple solutions exist if the remaining column elements for the non-basis link are identical for the pair of swappable links.

CHAPTER 5. NUMERICAL ANALYSIS AND INSIGHTS

Chapter 5 studies the capability of the proposed basis link method for the NLSP and reveals the implications for field applications. Section 5.1 analyzes five test networks to demonstrate the applicability of the proposed basis link method to solve the NSLP and to derive related insights. The networks considered consist of the following topologies: parallel highway network, fishbone network, radial network, complete network, and a network proposed by Yang and Zhou (1998) labeled Yang's network here. Section 5.2 performs sensitivity analysis on the effects of network topology and number of O-D pairs/paths on the minimum subset of links to be installed with vehicle sensors.

5.1 Effect of network topology

5.1.1. Parallel highway network

The parallel highway network shown in Figure 5.1 is analyzed using the basis link method. It consists of 4 O-D pairs, 14 links, and 9 nodes. Nodes 1 and 2 are trip origin nodes, and nodes 8 and 9 are the destination nodes. Table 5.1 illustrates the link-path incidence matrix for the network. The Gaussian elimination algorithm is used to obtain its RREF (Table 5.2). The basis links correspond to the 9 shaded columns in the table. They include links 1, 2, 3, 4, 5, 7, 9, 11, and 13. They represent the links on which to install vehicle sensors. Hence, about 64% of the links need to be equipped with sensors to estimate the flows on all links in the parallel highway network under steady-state conditions.

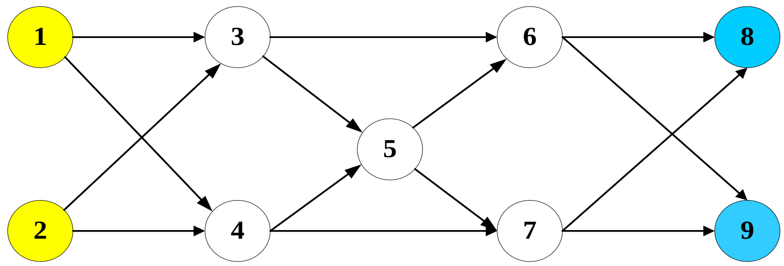


Figure 5.1. The Parallel Highway Network.

Table 5.1. Link-path incidence matrix of the parallel highway network

Link number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
link / path	1-3	1-4	2-4	2-3	3-5	3-6	4-5	4-7	5-6	5-7	6-8	6-9	7-8	7-9
1-3-6-8	1	0	0	0	0	1	0	0	0	0	1	0	0	0
1-3-5-7-8	1	0	0	0	1	0	0	0	0	1	0	0	1	0
1-4-7-8	0	1	0	0	0	0	0	1	0	0	0	0	1	0
1-3-6-9	1	0	0	0	0	1	0	0	0	0	0	1	0	0
1-3-5-6-9	1	0	0	0	1	0	0	0	1	0	0	1	0	0
1-4-7-9	0	1	0	0	0	0	0	1	0	0	0	0	0	1
2-4-5-6-8	0	0	1	0	0	0	1	0	1	0	1	0	0	0
2-4-7-8	0	0	1	0	0	0	0	1	0	0	0	0	1	0
2-3-6-8	0	0	0	1	0	1	0	0	0	0	1	0	0	0
2-4-5-6-9	0	0	1	0	0	0	1	0	1	0	0	1	0	0
2-4-7-9	0	0	1	0	0	0	0	1	0	0	0	0	0	1
2-3-6-9	0	0	0	1	0	1	0	0	0	0	0	1	0	0

Table 5.2. RREF of the parallel highway network

Link number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
link / path	1-3	1-4	2-4	2-3	3-5	3-6	4-5	4-7	5-6	5-7	6-8	6-9	7-8	7-9
1-3-6-8	**1**	0	0	0	0	1	0	0	0	0	0	1	0	0
1-3-5-7-8	0	**1**	0	0	0	0	0	1	0	0	0	0	0	1
1-4-7-8	0	0	**1**	0	0	0	0	1	0	0	0	0	0	1
1-3-6-9	0	0	0	**1**	0	1	0	0	0	0	0	1	0	0
1-3-5-6-9	0	0	0	0	**1**	-1	0	0	0	1	0	-1	0	1
1-4-7-9	0	0	0	0	0	0	**1**	-1	0	1	0	0	0	0
2-4-5-6-8	0	0	0	0	0	0	0	0	**1**	-1	0	1	0	-1
2-4-7-8	0	0	0	0	0	0	0	0	0	0	**1**	-1	0	0
2-3-6-8	0	0	0	0	0	0	0	0	0	0	0	0	**1**	-1
2-4-5-6-9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2-4-7-9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2-3-6-9	0	0	0	0	0	0	0	0	0	0	0	0	0	0

5.1.2. Fishbone network

The second network analyzed is the fishbone-shape network shown in Figure 5.2. It contains 4 O-D pairs, 18 links, and 10 nodes. Nodes 1 and 2 are the origin nodes, and nodes 9 and 10 are the destination nodes. The RREF of the link-path incidence matrix of the fishbone network is obtained by the Gaussian elimination algorithm and identifies 12 basis links: 1, 2, 3, 4, 5, 6, 8, 9, 10, 13, 14, and 17. Hence, in this 18-link network, only 12 sensors are needed to obtain the flow information on all links. Thereby, only 67% of the network links need to be equipped with sensors.

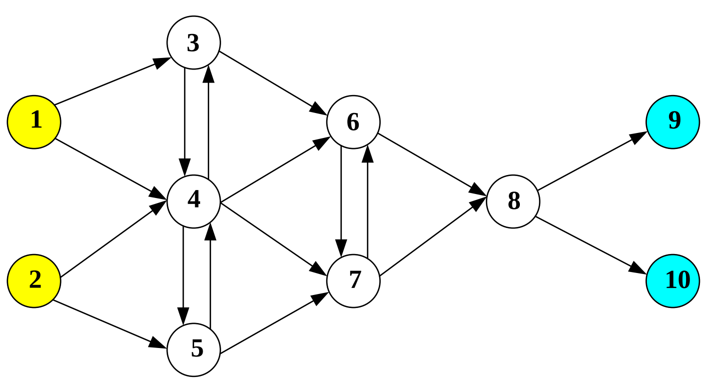


Figure 5.2. The Fishbone Network.

5.1.3. Radial network

Figure 5.3 shows a radial network with 7 O-D pairs, 20 links, and 7 nodes. Nodes 1, 2, and 7 are origin nodes, and nodes 1, 4, and 7 are destination nodes. The RREF associated with this network is also obtained by the Gaussian elimination algorithm. It indicates that sensors ought to be installed on 14 of the 20 network links. Therefore, 70% of the network links need to be equipped with sensors.

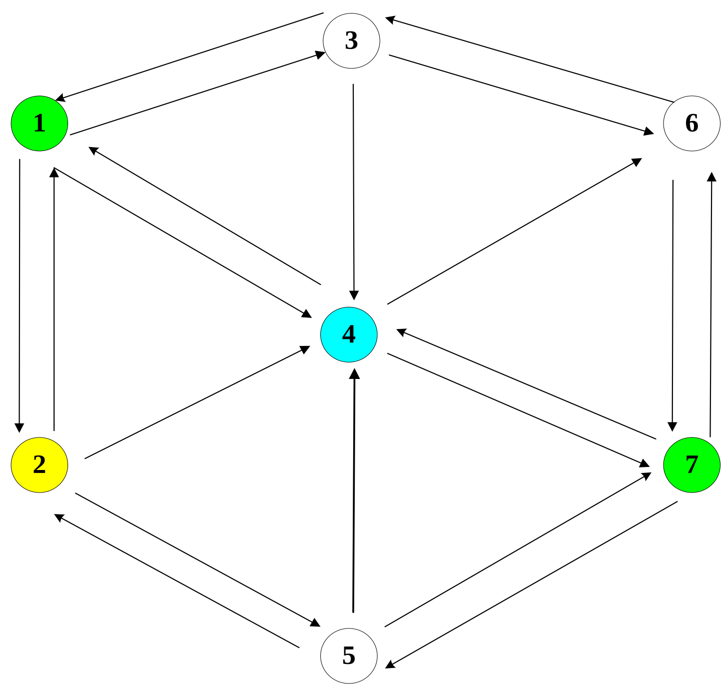


Figure 5.3. The Radial Network.

5.1.4 Complete network

A bidirectional complete network has $n \times (n\text{-}1)$ links, where n is the number of nodes. Figure 5.4 shows a complete network consisting of 6 O-D pairs, 30 links, and 6 nodes. Nodes 1 and 2 are the origin nodes, and nodes 3, 4, and 5 are the destination nodes. The RREF of the link-path incidence matrix of the complete network identifies 18 basis links. Thereby, 60% of the links need to be equipped with vehicle sensors in order to infer the flows on the unequipped links.

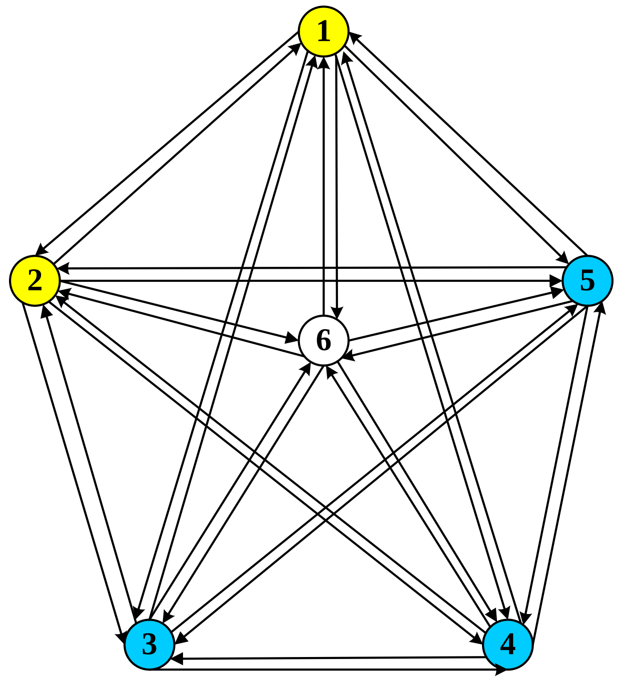


Figure 5.4. The Complete Network.

5.1.5. Yang's network

To test the applicability of the proposed basis link method to more general network cases, a larger network proposed by Yang and Zhou (1998) is adopted (see Figure 5.5). The network consists of 182 O-D pairs, 76 links, and 24 nodes. The shaded nodes in Fig. 6 represent both trip origins and destinations. For each O-D pair, one to five paths exist depending on their spatial locations, resulting in 368 paths. The RREF of the link-path incidence matrix of Yang's network identifies 62 basis links. Hence, 81% of the links need to be equipped with sensors to estimate the flows on all links; they are shown in Figure 5.6.

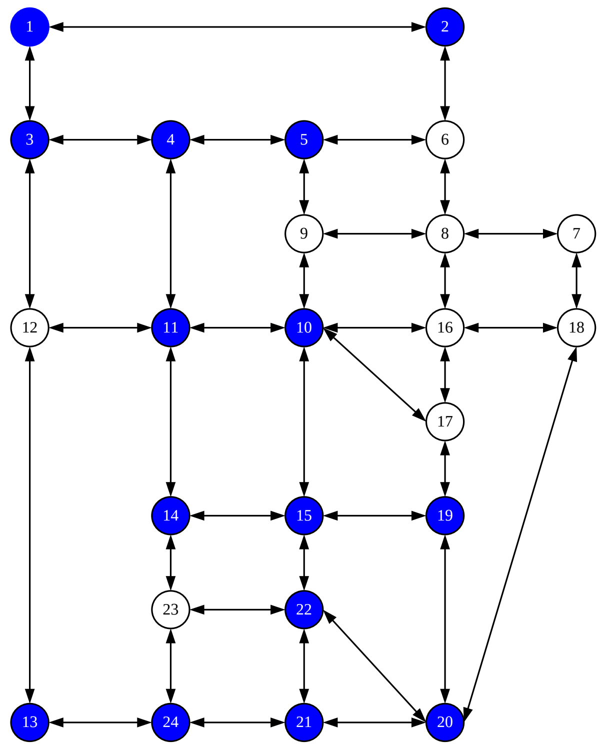


Figure 5.5. Yang's Network.

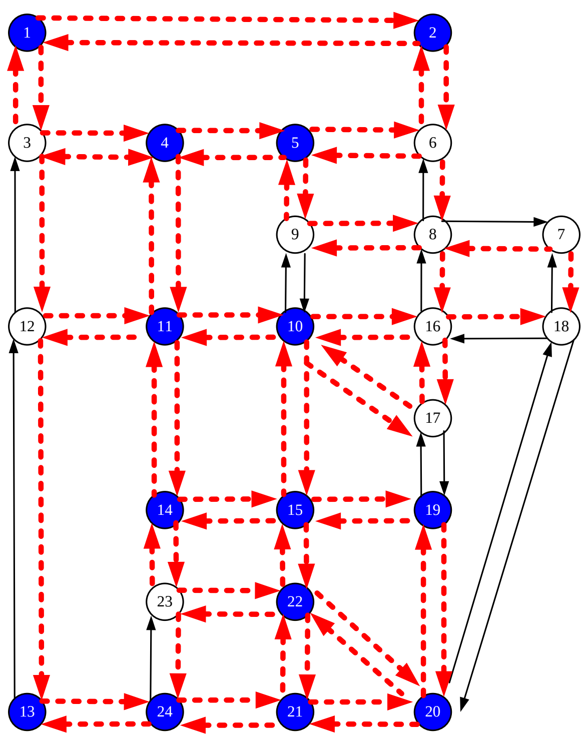


Figure 5.6. Spatial Locations of the Basis Links on Yang's Network.

5.1.6. Insights

Table 5.3 summarizes the results of the basis link method to identify the minimum subset of links to equip with sensors for the five test networks. It indicates that while the number of basis links is related to the network scale in terms of the number of links/nodes or used paths under different network topologies, there is no direct correlation between the number of links/nodes or used paths in the network and the percentage of links to install sensors on. That is, the network topology, in terms of how the links and nodes are connected in the physical structure of the network, is a key determinant of the set of basis links. This is logically consistent because the link-path incidence matrix, which implies the network structure, is the foundation for the proposed basis link method. Table 9 also indicates the computational CPU time for each test network. The computational time is primarily to obtain the RREFs of the link-path incidence matrices associated with the test networks. The RREFs were obtained using MATLAB V6.0 on an Intel Centrino Duo T5600 1.83 GHz Windows XP SP2 OS platform with 2GB memory.

Table 5.3. Comparison of the five test networks

Item / Network	Number of O-D pairs	Number of nodes	Number of paths (m)	Number of links (n)	Number of basis links (r)	% of links to be equipped with sensors	CPU time (sec)
Parallel highway network	4	9	12	14	*9*	64%	0.01
Fishbone network	4	10	22	18	*12*	67%	0.02
Radial network	7	7	22	20	*14*	70%	0.03
Complete network	6	6	30	30	*18*	60%	0.06
Yang's network	182	24	368	76	*62*	81%	1.60

5.2 Sensitivity analysis

This section explores the effect of network topology and number of O-D pairs/paths on the number of basis links. The minimum subset of links and the percentage of links to be installed with sensors are determined in the following contexts: (1) effect of network scale in terms of the number of links/nodes, (2) effect of network topology in terms of network connectivity, (3) effect of number of O-D pairs, (4) effect of number of used paths, and (5) effect of the network degree. The sensitivity analysis is based on the fishbone and radial network structures, and modifications thereof.

5.2.1. Effect of network scale

We investigate the effects of network scale by adding more links and nodes to the original fishbone network. Two modified fishbone-shape network structures are analyzed, in addition to the original fishbone network.

(1) Modified fishbone network I: Figure 5.7 illustrates the modified fishbone network I. It is obtained by adding two more nodes: $node_{11}$ and $node_{12}$, and four links: $link_{6-11}$, $link_{11-9}$, $link_{7\text{-}12}$, and $link_{12-10}$. Therefore, the new network contains 22 links and 12 nodes. The origin nodes (1, 2) and destination nodes (9, 10) remain unchanged.

(2) Modified fishbone network II: The modified fishbone network II, shown in Figure 5.8, is obtained by extending the modified fishbone network I by adding four more links: $link_{3\text{-}11}$, $link_{11-8}$, $link_{5-12}$, and $link_{12-8}$. Therefore, the new network contains 26 links and 12 nodes. The origin and destination nodes are unchanged.

Table 5.4 compares the three fishbone networks in terms of the number of basis links and the percentage of links to be equipped with vehicle sensors. It indicates that the number of basis links generally increases as the number of links/nodes is increased. However, the percentage of links to be equipped with sensors for a larger network (the modified fishbone network II) is not necessarily larger than that for a smaller network (the modified fishbone network I). It reiterates the conclusions in Section 5.1.6 on the role of network topology.

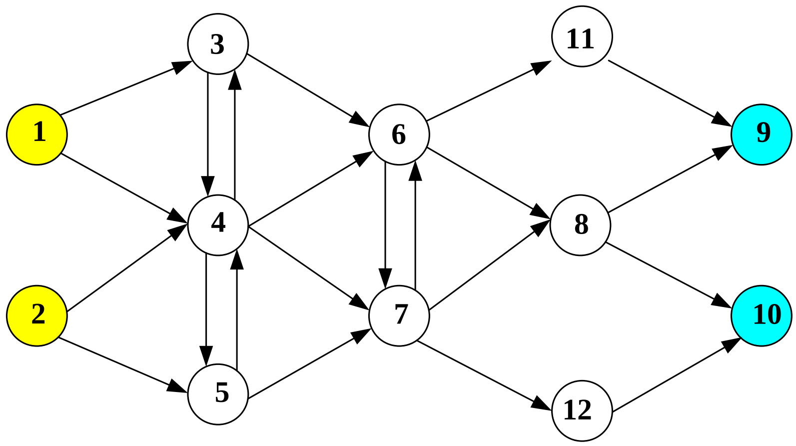


Figure 5.7. Modified Fishbone Network I.

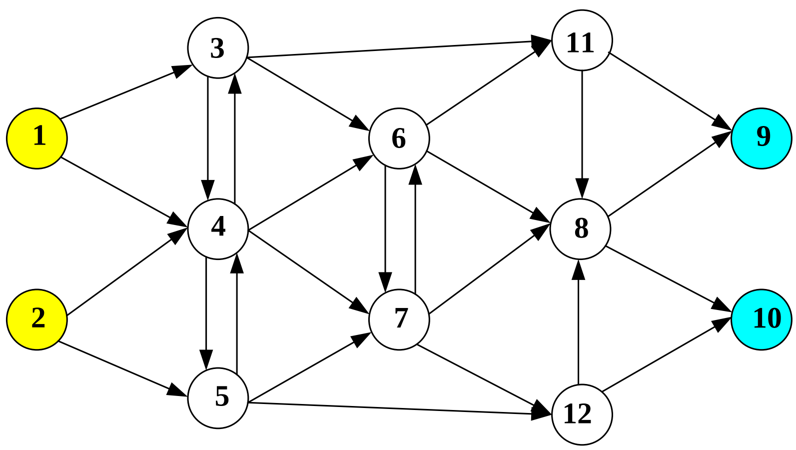
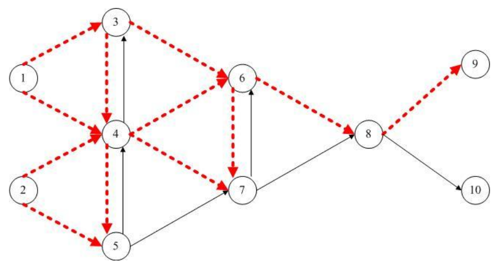
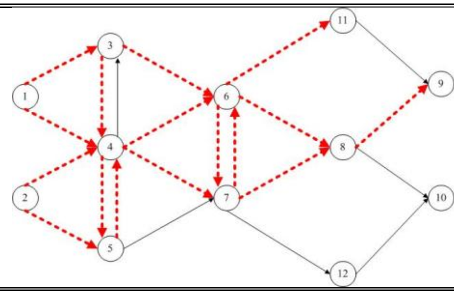
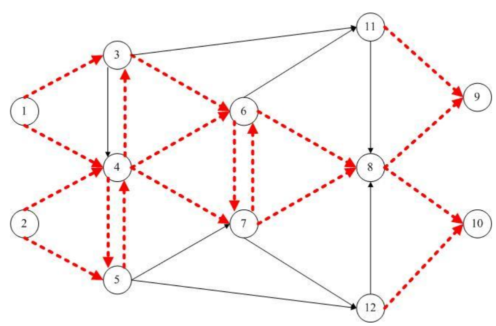


Figure 5.8. Modified Fishbone Network II.

Table 5.4. Effect of network scale

Network	Network configuration and the basis links	Number of nodes	Number of paths (m)	Number of links (n)	Number of basis links (r)	% of links to be equipped with sensors
Fishbone Network		10	22	18	12	67%
Modified Fishbone Network I		12	45	22	17	77%
Modified Fishbone Network II		12	41	26	18	69%

5.2.2. Effect of network connectivity

We explore the effect of network connectivity by changing the in- and out-degrees of the original fishbone network, while retaining the numbers of links and nodes of the original fishbone network. The new network structure, shown in Figure 5.9, is labeled the modified fishbone network III. Table 5.5 compares the original fishbone network and the modified fishbone network III in the context of the basis link method. It indicates that there is no clear correlation between the number of links/nodes and the number of basis links. Instead, it illustrates that the connectivity implied by the network topology affects the number of basis links, confirming the insights of Section 5.1.6. It also hints at a relationship between the number of paths being used and the number of basis links; this will be addressed in Section 5.2.4.

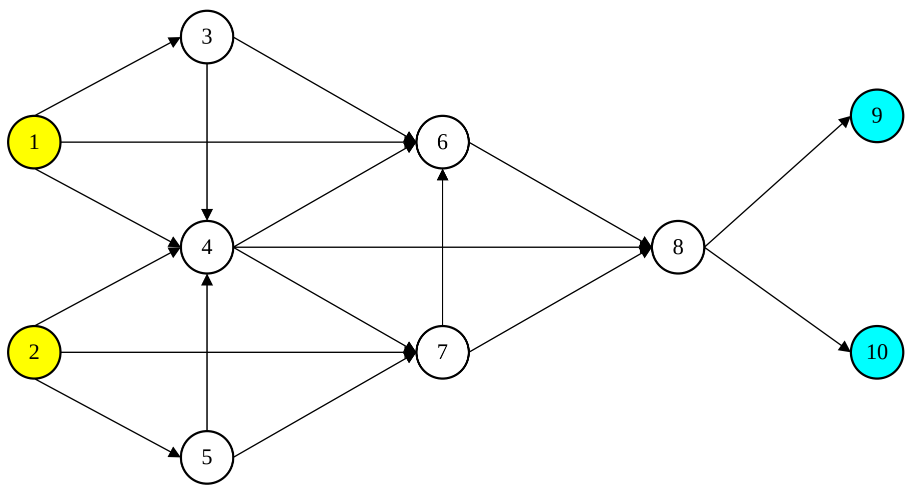
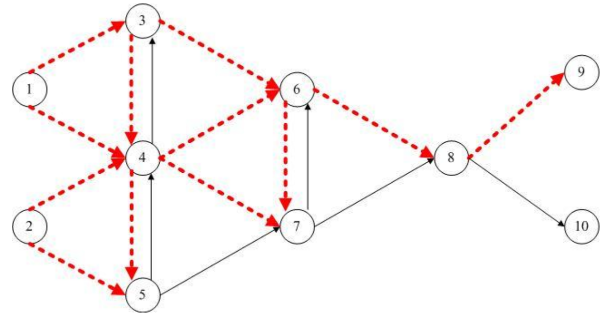
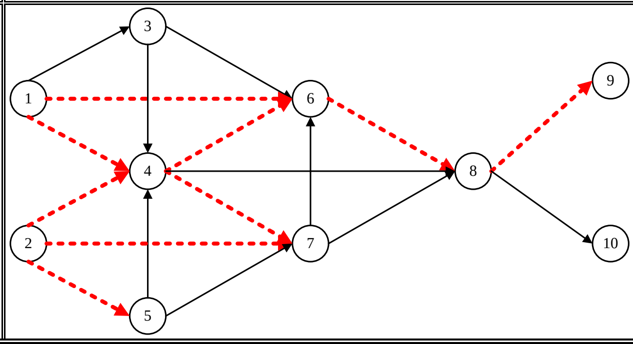


Figure 5.9. Modified Fishbone Network III.

Table 5.5. Effect of network connectivity

Network	Network configuration and the basis links	Number of nodes	Number of paths (m)	Number of links (n)	Number of basis links (r)	% of links to be equipped with sensors
Fishbone Network		10	22	18	*12*	67%
Modified Fishbone Network III		10	16	18	*9*	50%

5.2.3. Effect of the number of O-D pairs

The effect of the number of O-D pairs on the percentage of basis links is investigated by incrementing the number of O-D pairs by 1 in the range of 1 – 8 for the original fishbone network, and assuming 3 paths per O-D pair. Table 5.6. indicates that the percentage of links to be equipped with sensors increases with the number of O-D pairs up to a point, beyond which there is no effect. That is, this percentage has an upper bound of 78%, implying that the network topology may make it unnecessary to observe additional link flows to infer them for the entire network. The results also indicate a correlation between the number of paths and the number of basis links.

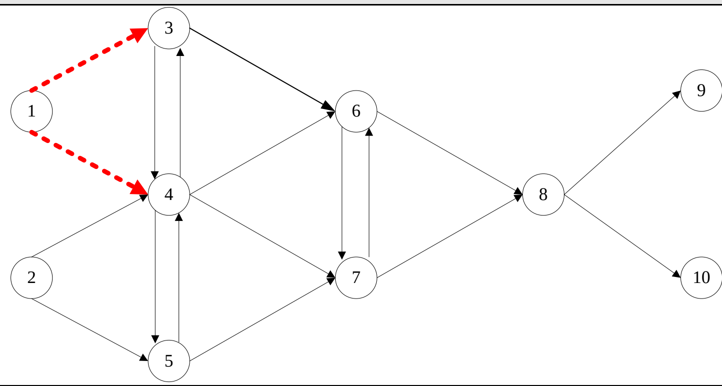
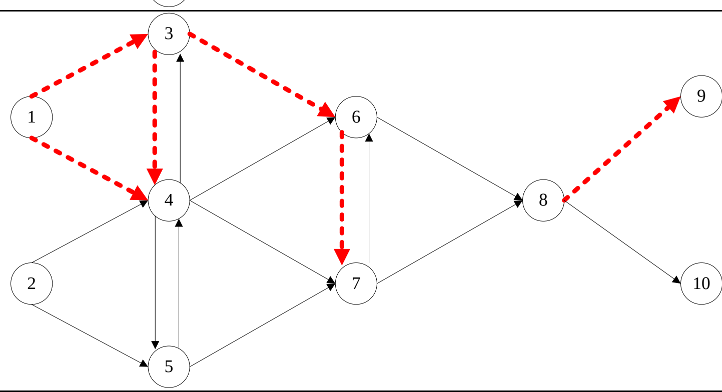
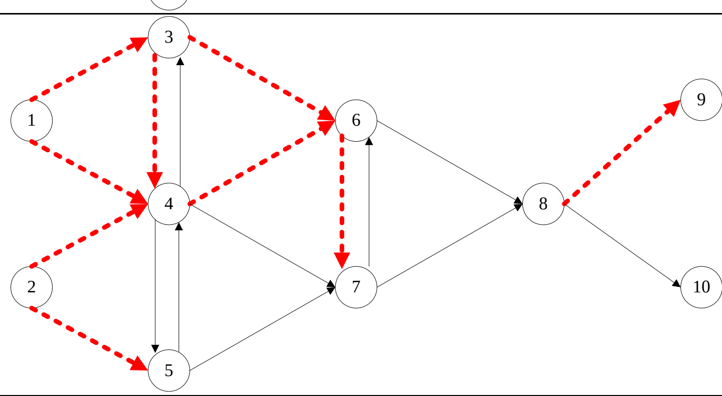
5.2.4. Effect of the number of paths

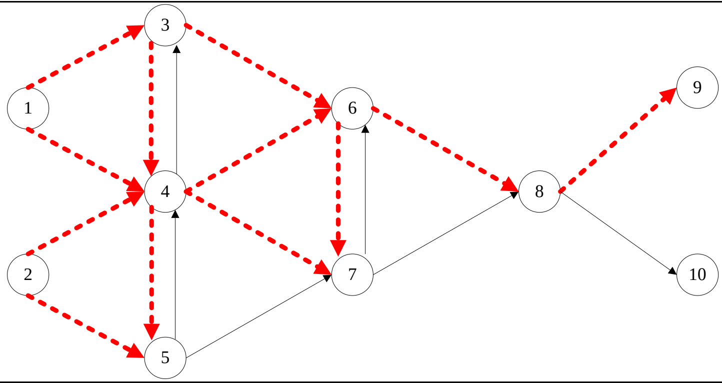
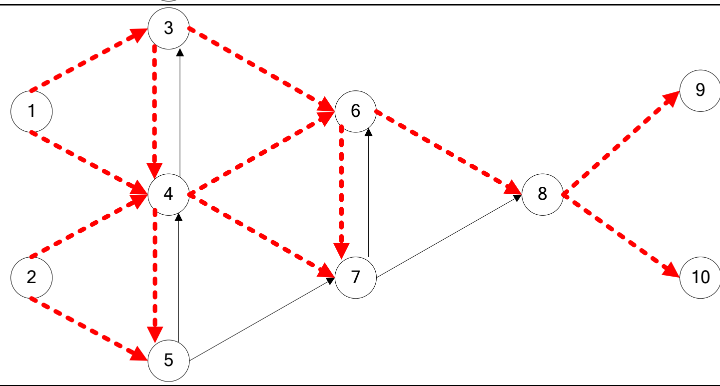
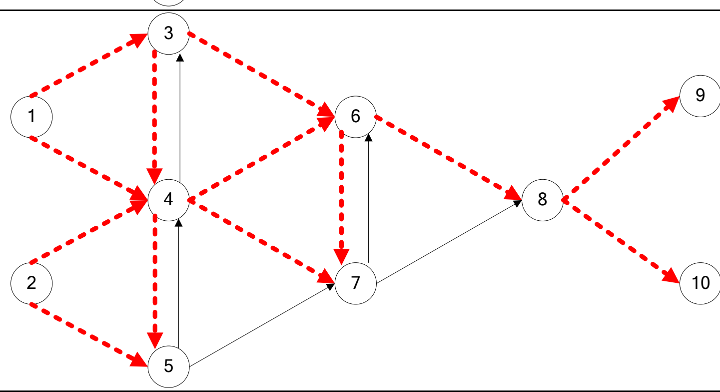
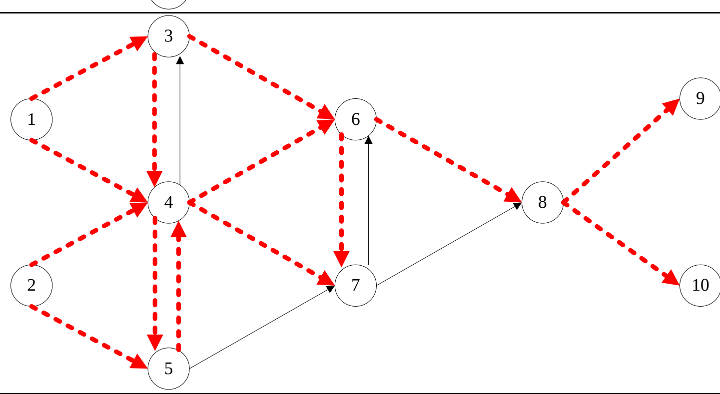
Here, the number of paths per O-D pair for each of the 4 O-D pairs of the original fishbone network is assumed to take values of 4, 8, 12, 14, and 16, resulting in a total of 16, 32, 48, 56, and 64 paths, respectively, for the entire network. The numerical results, shown in Table 5.7, indicate that the percentage of links to be equipped with sensors increases with the number of paths up to an upper bound of 78%. It reinforces the insights of Section 5.2.3 related to the role of network topology on the upper bound.

5.2.5. Effect of the network degree

The effect of the network degree is examined using complete and incomplete networks. The complete network shown in Figure 5.4 is adopted as the baseline case (scenario 1); the number of basis links is 16 and the percentage of links to be equipped with sensors is 60%. By retaining this network shape and the number of nodes, two incomplete networks consisting of 20 and 16 links are evaluated (scenarios 2 and 3, respectively). The results, shown in Table 5.8, indicate that the number of basis links decreases with a decrease in the network degree represented by the number of connected arcs. However, the percentage of links to be equipped with sensors for the complete network is 60%, which is less than those for the two incomplete networks (70% and 81%, respectively). This suggests that complete connectivity offers more opportunities for relationships involving link flows, thereby reducing the percentage of links that need to be equipped.

Table 5.6. Effect of the number of O-D pairs

Scenario	Network configuration and basis links	Number of O-D pairs	Number of paths	Number of basis links	% of links to be equipped with sensors
1		*1*	3	2	11%
2		*2*	6	6	33%
3		*3*	9	9	50%

4		*4*	12	12	67%
5		*5*	15	13	72%
6		*6*	18	13	72%
7		*7*	21	14	78%

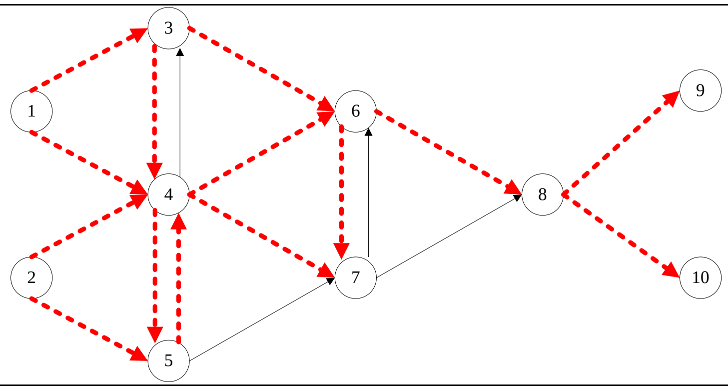
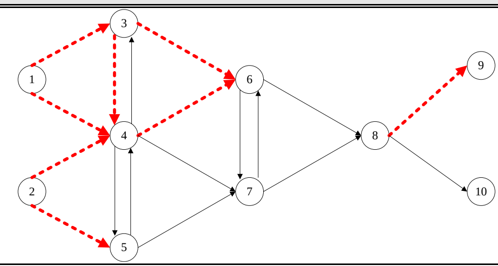
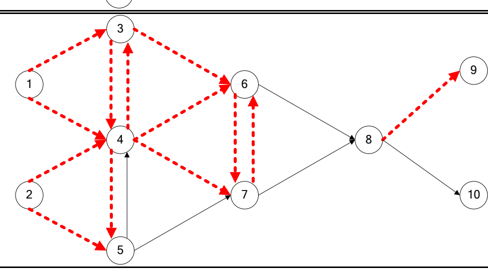
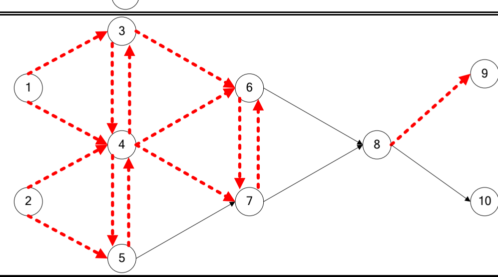
8		*8*	24	14	78%

Table 5.7. Effect of the number of paths

Scenario	Network configuration and basis links	Number of paths	Number of basis links	% of links to be equipped with sensors
1		*16*	8	44%
2		*32*	13	72%
3		*48*	14	78%

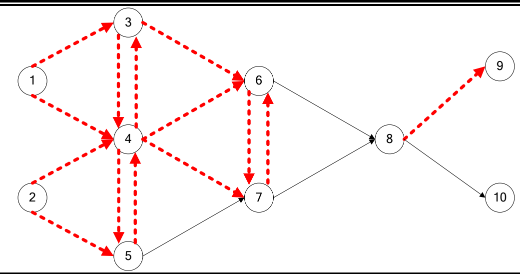
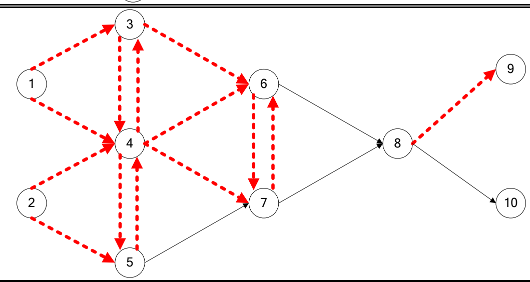
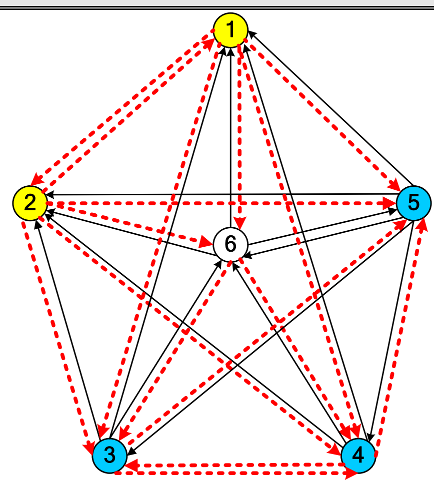
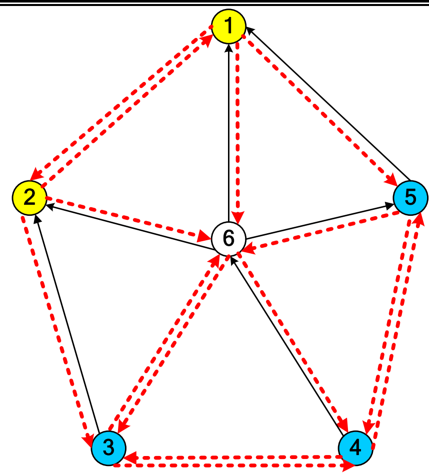
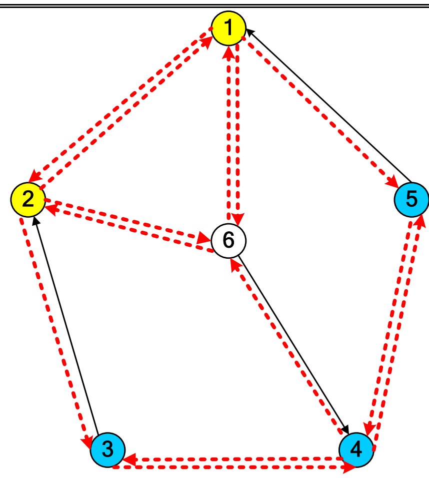
4		*56*	14	78%
5		*64*	14	78%

Table 5.8. Effect of the network degree

Scenario	Network configuration and basis links	Number of paths	Number of links	Number of basis links	% of links to be equipped with sensors
1		30	30	18	60%

2		30	20	14	70%
3		27	16	13	81%

5.2.6. Insights

The sensitivity analysis suggests no direct relationship between the percentage of links to be equipped with sensors and the number of links/nodes in a network, though a positive correlation between the number of basis links and the number of links/nodes or paths is generally observed. The key determinant for the minimum subset of links to be equipped with sensors so as to infer the flows on all links is the network topology represented by the link-path incidence matrix. The results also indicate that there may be an upper bound on the number of basis links that is governed by the network topology independent of the scale of the network. A practical benefit of such upper bounds is that it may be unnecessary to install sensors on all links, leading to cost savings for the traffic agency.

CHAPTER 6. CONCULSIONS

This chapter summarizes the research, highlights its contributions, and proposes directions for future research.

6.1 Summary

This study addresses the two primary objectives:

1. Addresses the network sensor location problem (NSLP) directly so as to obtain the unobserved link flows given the minimum subset of observed link flows provided by passive counting sensors.

2. Analyzes the corresponding theoretical aspects related to the proposed basis link method in the determination of the minimum subset of network links to infer unobserved link flows.

This study proposes a basis link method to address the network sensor location problem under steady-state traffic conditions. A fundamental contribution of this research is the illustration of a direct mapping between the basis link flows and the non-basis link flows, which can be obtained from the network structure represented by the link-path incidence matrix. Based on the theoretical investigation and numerical analysis, several findings are listed below.

1. Given the network structure represented by its link-path incidence matrix, a theoretical minimum subset of network links provided by the reduced row echelon form (RREF) algorithm does exist, and a direct mapping between the basis link flows and the non-basis link flows is also theoretically proved.

2. The study illustrates the possibility of multiple solutions in terms of the set of basis links. However, as shown by one of the Lemmas, that does not affect the uniqueness in terms of the inferred link flows.

3. The empirical analysis highlights the primacy of the network topology in determining the set of basis links. It also indicates the possibility of an upper bound on the number of basis links based on the topology, suggesting that it may not be necessary to equip every link with sensors from a planning perspective.

4. While the number of basis links is related to the network scale in terms of the number of links/nodes or used paths under different network topologies, there is no direct correlation between the number of links/nodes or used paths in the network and the percentage of links to install sensors on.

5. The sensitivity analysis suggests no direct relationship between the percentage of links to be equipped with sensors and the number of links/nodes in a network, though a positive correlation between the number of basis links and the number of links/nodes or paths is generally observed.

6.2 Future research directions

This research has proposed a linear algebraic approach for the determination of the minimum subset of equipped links to infer the flows on the unobserved links. The proposed basis link method provides network full observability without requiring any assumptions in terms of the knowledge of O-D flows, path flows, user route choice behavior, or traffic assignment rules. This property has broader implications in terms of potentially aiding in solving a range of problems (such as O-D demand estimation, travel time estimation, traffic assignment) in both the static and dynamic contexts. Therefore, a straightforward research direction is to estimate O-D demands for a general network, in an integrated manner, based on the (full) link flow information provided by the proposed basis link approach. Further, the existence of multiple solutions provides some flexibility for traffic agencies in choosing the links to install sensors on. That is, a traffic agency may prefer to install sensors on some links because of their importance based on one or more criteria related to objectives such as minimizing deployment costs, reducing traffic impacts, or the relative importance of a link (based on the facility type or its criticality for disaster response, etc.). In such instances, priority rankings provided by the agency can be seamlessly adapted with the proposed basis link method. This is another immediate research issue that is worthy of further investigation.

REFERENCES

Anton, H. (1984). Elementary Linear Algebra, John Wiley & Sons, Inc., N.Y.

Ashok, K., Ben-Akiva, M.E. (2002). Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows. Transportation Science, 36 (2), p. 184-198.

Bekhor, S., Ben-Akiva, M.E., Ramming, M.S. (2006). Evaluation of choice set generation algorithm for route choice models. Annals of Operations Research. 144 (1), p. 235-247.

Ben-Akiva, M.E., Bottom, J.A., Ramming, M.S. (2001). Route guidance and information systems. Journal of Systems and Control Engineering, 215 (4), p. 317-324.

Bianco, L., Confessore, G., Reverberi, P. (2001). A network based model for traffic sensor location with implications on O/D matrix estimates. Transportation Science, 35 (1), p. 50-60.

Cascetta, E., Inaudi, D., Marquis, G. (1993). Dynamic estimators of origin-destination matrices using traffic counts. Transportation Science, 27 (4), p. 363-373.

Cascetta, E. and Nguyen, S. (1988). A unified framework for estimating or updating origin-destination matrices from traffic counts. Transportation Research Part B, 22 (6), p. 437-455.

Castillo, E., Conejo, A.J., Pruneda, R.E., Solares, C. (2007). Observability in linear systems of equations and inequalities: Applications. Computers and Operations Research, 34(6), p. 1708-1720.

Castillo, E., Conejo, A.J., Menéndez, J.M., Jiménez, P. (2008a). The observability problem in traffic network models. Computer-Aided Civil and Infrastructure Engineering, 23 (3), p. 208-222.

Castillo, E., Jiménez, P., Menéndez, J.M., Conejo, A.J. (2008b). The observability problem in traffic models: algebraic and topological methods. IEEE Transactions on Intelligent Transportation Systems, 9 (2), p. 275-287.

Castillo, E., Menéndez, J.M., Sanchez-Cambronero, S. (2008c). Traffic estimation and optimal counting location with path enumeration using Bayesian networks. Computer-Aided Civil and Infrastructure Engineering, 23 (3), p. 189-207.

Castillo, E., Menéndez, J.M., Jiménez, P. (2008d). Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations. Transportation Research Part B, 42 (5), p. 455-481.

Chang, G.L. and Tao, X. (1999). An integrated model for estimating time-varying network origin-destination distributions. Transportation Research Part A, 33 (5), p. 381-399.

Ehlert, A., Bell, M.G.H., Grosso, S. (2006). The optimisation of traffic count locations in road networks. Transportation Research Part B, 40 (6), p. 460-479.

Friedberg, S.H., Insel, A.J., Spence, L.E. (2003). Linear Algebra, Pearson Education, Upper Saddle River, N.J.

Gan, L., Yang, H., Wong, S.C. (2005). Traffic counting location and error bound in origin-destination matrix estimation problems. Journal of Transportation Engineering, 131 (7), p. 524-534.

Gentili, M. and Mirchandani, P.B. (2005). Locating active sensors on traffic networks. Annals of Operations Research, 136 (1), p. 229-257.

Maher, M.J. (1983). Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. Transportation Research Part B, 17 (6), p. 435-447.

Peeta, S. and Bulusu, S. (1999). A generalized singular value decomposition approach for consistent online dynamic traffic assignment. Transportation Research Record, 1667, p. 77-87.

Peeta, S. and Ziliaskopoulos, A. (2001). Foundations of dynamic traffic assignment: The past, the present and the future. Networks and Spatial Economics, 1 (3/4), p. 233-266.

Stewart, G.W. (1998). Matrix Algorithms Volume 1: Basic Decompositions, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Wu, J. and Chang, G.L. (1996). Estimation of time-varying origin-destination distributions with dynamic screenline flows. Transportation Research, Part B 30 (4), p. 277-290.

Yang, H. and Zhou, J. (1998). Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research Part B, 32 (2), p. 109-126.

Yang, H., Yang, C., Gan, L. (2006). Models and algorithms for the screen line-based traffic-counting location problems. Computers and Operations Research, 33 (3), p. 836-858.

Zhou, X. and Mahmassani, H.S. (2005). Recursive approaches for online consistency checking and OD demand updating for real-time dynamic traffic assignment operation. Transportation Research Record, 1923, p. 218-226.