1 **An Introduction to the Network Weight Matrix**

2 **Alireza Ermagun** (*Corresponding Author*)

3 Ph.D. Candidate

4 University of Minnesota

5 Department of Civil, Environmental, and Geo- Engineering

6 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA

7 ermag001@umn.edu

8 **David Levinson**

9 Professor

10 University of Minnesota and University of Sydney

11 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA

12 dlevinson@umn.edu

13 Paper submitted for:

14 Presentation at 96th Annual Transportation Research Board Meeting, January 2017

15 Standing Committee on Statistical Methods (ABJ80)

16 5311 (errors:1) words + 8 figures + 4 tables

17 July 31, 2016

1 ABSTRACT

2 This study introduces the network weight matrix as a replacement for the spatial weight matrix
3 to measure the spatial dependence between links of a network. This matrix stems from the con-
4 cept of betweenness centrality and vulnerability in network science. The elements of the matrix
5 are a function not simply of proximity, but of network topology, network structure, and demand
6 configuration. The network weight matrix has distinctive characteristic, which are capable of re-
7 flecting spatial dependence between traffic links: (1) The elements are allowed to have negative and
8 positive values, which capture competitive and complementary nature of links, (2) The diagonal
9 elements are not fixed to zero, which takes the self-dependence of a link upon itself into consider-
10 ation, and (3) The elements not only reflect the spatial dependence based on the network structure,
11 but they acknowledge the demand configuration as well. We verified the network weight matrix by
12 modeling traffic flows in a 3×3 grid test network with 9 nodes and 24 directed links connecting
13 72 origin-destination (OD) pairs. The results disclose models encompassing the network weight
14 matrix outperform both models without spatial components and models with the spatial weight
15 matrix. This leads inexorably to the conclusion that the network weight matrix represents a more
16 accurate and defensible spatial dependency between traffic links, and thereby augments traffic flow
17 prediction.

18 **Keywords:** Spatial Weight Matrix; Spatial Econometrics; Traffic Flow; Traffic Network

1 INTRODUCTION

2 “Everything is related to everything else, but near things are more related than distant things.” –
3 Waldo Tobler, 1970 (1).

4 In 1970, Tobler was the introduced what he called “the first law of geography” (1). Ever
5 since, this law has become the mainspring of spatial dependence at the core of regional science and
6 geography (2). This dependence is ascertained by a notion of relative space, which accentuates the
7 role of distance in spatial dependence. The strength and structure of potential dependence between
8 spatial units are typically determined by a spatial weight matrix, which is an integral part of spatial
9 econometrics (3). A whole host of complex and nuanced approaches have gradually emerged for
10 constructing spatial weight matrices. In spite of attempted remedies, a satisfactory solution has not
11 yet been obtained for the proper creation of spatial weight matrices (4).

12 Following the recent prominence of spatial dependence, a growing interest has examined
13 spatial autocorrelation phenomena in natural, formal, social, and applied science. Spatial depen-
14 dence permeates geology, epidemiology, criminology, sociology, geography, and economics, to
15 name a few. Transportation science is no exception, and is affected by this wave. Obviously and
16 consistent with physical models, there is spatial dependence between traffic links in a network
17 (5, 6). Methods to develop this with more statistical rigor have progressed over time. For instance,
18 Kamarianakis and Prastacos (7) use a first- and second-order neighbors weight matrix in space-
19 time autoregressive integrated moving average (STARIMA) model to forecast short-term relative
20 velocity on major arterials of Athens, Greece. Henceforth, only a few studies have attempted to
21 measure the extent of spatial autocorrelation between neighboring traffic links. Min et al. (8) use
22 a first-order weight matrix in generalized STARIMA model to forecast short-term traffic flow on
23 Beijing, China. Cheng et al. (5) scrutinize the spatio-temporal autocorrelation structure on 22
24 traffic links of central London, England. Proposed models assume that the neighboring links are
25 positively correlated, and thereby the positive off-diagonal elements are implemented for the spa-
26 tial weight matrices. In this paper we show this is a restrictive assumption that misestimates the
27 flow of traffic links, given the competitive and complementary nature of links.

28 The current study introduces a network econometrics approach using a network weight
29 matrices to bridge the gap between two strands of literature: network science and geospatial anal-
30 ysis. Having this introduction, the remainder of this paper is structured as follows. The following
31 section reviews the notion of the spatial weight matrix and its common structures in multifaceted
32 disciplines. We then introduce network econometrics. Following the definition of network econo-
33 metrics, we discuss betweenness centrality property of the network. We subsequently propose the
34 network weight matrix. We finally verify the network weight matrix and conclude the paper by
35 broaching a number of recommendations, which suggests future research.

36 SPATIAL WEIGHT MATRIX: A REVIEW OF EXISTING APPROACHES

37 In geography and spatial science, spatial weight matrices are foundational building blocks of spatial
38 analysis. A spatial weight matrix for a set of m units is an $m \times m$ matrix, which its components,
39 W_{ij} , regularly satisfy two major rules: (1) $W_{ij} \geq 0$ and (2) $\sum_{j=1}^m W_{ij} = 1$, for all $i = 1, 2, \dots, m$ (4).
40 Theoretically, W_{ij} expresses the strength of potential spatial dependence between unit i and unit
41 j , and the greater the W_{ij} the more dependency. Although the spatial weight matrix has capacity
42 to capture the self-dependence of unit i upon itself, the matrix is typically employed with a zero
43 diagonal matrix (4). Non-diagonal elements, however, are measured by a number of methods.

44 Pertaining to theory, Aldstadt and Getis (9) argue that three viewpoints are classically in-

spired spatial analysts to structure a spatial weight matrix: (1) Benefiting from theoretical notion of spatial dependence, (2) Benefiting from geometric indicators, and (3) Benefiting from descriptive expression of the data. Pertaining to measurement methods, the first two viewpoints embrace two main categories: a distance and a boundary approach, respectively. Each category encompasses a wide spectrum of methods, including use of a contiguity matrix, fixed distance, or inverse distance. The third viewpoint derives its theoretical support from a familiar statistics and geostatistics rule: “Let the data speak for themselves” (10). This viewpoint is superior to alternatives that are incapable of detecting the complexities of spatial dependence within the data (9). Use of the empirical variogram function is a well-known instance of this viewpoint. There is prolific literature discussing assorted structural forms of spatial weight matrices. Getis and Aldstadt (3), for instance, summarize the previous attempts to create a spatial weight matrix. We hence eschew digging into proposed forms of spatial weight matrix in detail, and confine it to list the widely-used spatial weight matrices in Table 1.

Rules are made to be broken. Notwithstanding the positive elements and row standardization rules of spatial weight matrix are predominant, the literature has witnessed a few exceptions. In 2013, Bhattacharjee and Jensen-Butler (11) estimated a spatial weight matrix under structural constraint. They assume the matrix is unknown and potentially asymmetric with both positive and negative off-diagonal elements. Likewise, in 2015, Chung and Hewings (12) introduced a spatial weight matrix encompassing both negative and positive elements to examine the competitive and complementary relationship between regional economies. The matrix also has nonzero diagonal elements to capture the self-dependent nature of regions.

Knowledge of the spatial weight matrix is still rudimentary. The literature on creation of spatial weight matrix leaves room to grow in extracting spatial dependence from data. While spatial analysts acknowledge the value of theoretical conceptualization toward examining the spatial dependence of phenomena that manifests itself in a spatial weight matrix, empirically convenient approaches are succeeding theory (3). Consequently, creating a spatial weight matrix that is theoretically defensible has become troublesome in geography and spatial science. This research aims to fill the lacuna by introducing a novel network weight matrix. Constructing the network weight matrix requires understanding network topology and infrastructure. Accordingly, we represent the notion of link centrality and link vulnerability in the following sections as fundamental properties of a network.

TABLE 1 : Spatial weight matrix configurations

Weight Method	Formulation	Definition
Boundary Approach		
Contiguity	$w_{ij} = \begin{cases} 1 : l_{ij} > 0 \\ 0 : l_{ij} = 0 \end{cases}$	l_{ij} : length of shared boundary
Shared boundary	$w_{ij} = \frac{l_{ij}}{\sum_{i \neq j} l_{ij}}$	l_{ij} : length of shared boundary
Distance Approach		
Radial Distance	$w_{ij} = \begin{cases} 1 : 0 \leq d_{ij} \leq d \\ 0 : d_{ij} > d \end{cases}$	d_{ij} : distances between spatial unit d : distance threshold
Power Distance	$w_{ij} = d_{ij}^{-\alpha}$	α : any positive exponent
Exponential Distance	$w_{ij} = \exp(-\alpha d_{ij})$	α : any positive value
Double-Power Distance	$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{d}\right)^k\right] : 0 \leq d_{ij} \leq d \\ 0 : d_{ij} > d \end{cases}$	d_{ij} : distances between spatial unit d : distance threshold

1 NETWORK TOPOLOGY: CENTRALITY

2 In this section, we provide a brief historical overview of the centrality concept and measurements.
 3 Let us begin with a simple definition of network (also called a graph in the mathematics literature).
 4 A network encompasses a set of edges and vertices (also known as links and nodes, respectively).
 5 Mathematically, a network is represented by an adjacency matrix, which in the simplest case is a
 6 square, binary, and symmetric matrix with the order of number of vertices. The element of this
 7 matrix is one, if there is an edge between corresponding vertices; otherwise, it is zero (13). One
 8 variant is the directed network, in which links have a particular path direction between two nodes.
 9 Large networks coalesce with numerous routes that use a single edge between two vertices.
 10 Flows on edges are not arbitrarily distributed over a network, but rather exhibit dependency. This
 11 leads to questions, including““

- 12 • What are the most central links of a network?
- 13 • What links have the most influence over others?
- 14 • Which link connections are most crucial to the functioning of a network?
- 15 • Will removing a link break down the network into smaller networks? If so, which links
- 16 are those?

17 These questions are of primary interest since they concern crucial subjects, including net-
 18 work resilience and reliability (14). To answer these questions, the concept of centrality was
 19 introduced in the 19th century (15). In the discipline of network science, the definition of cen-
 20 trality originates from two distinct structural perspectives. The first perspective assigns the role
 21 of centrality based on closeness to every other link in a network. This perception stems from the
 22 idea that the closest link has more access to other links (16). The second perspective is moti-
 23 vated by the idea that links are central to the degree they stand between other links on the path of
 24 flow (17). Accordingly, the medial character of such links facilitates the flow between origins and
 25 destinations.

26 Over the past decade, a number of centrality measurements have emerged to capture the
 27 importance of the links (18). The two most widely used measurements are degree centrality and
 28 betweenness centrality. Concretely, the following terms define and formulate these below.

- 29 1. Degree centrality, which is also known as connectivity (a term which is also defined with
 30 other meanings in the literature, so we will not use it here), measures the number of links
 31 that connect to a specific link (19). In a network with l links, this is formally defined
 32 by Equation 1. In this equation, $r(i, j) = 1$ if link i and j are connected; otherwise,
 33 $r(i, j) = 0$.

$$C_D(i) = \sum_{j=1}^l r(i, j) \quad (1)$$

- 34 2. Betweenness centrality is formulated in light of the hypothesis that flow is passed from
 35 an origin to a destination only along a network path linking them. This concept is the
 36 first proposed by Freeman (17) in social network science. Pursuant to this hypothesis,

the betweenness of link i in a network with n nodes is measured as Equation 2. In this equation, P_{od} stands for the number of shortest paths between o and d , and P_{oid} represents the number of shortest paths between o and d that pass through link i .

$$C_B(i) = \sum_{o=1}^{n-1} \sum_{d=1}^{n-1} \frac{P_{oid}}{P_{od}} \quad (2)$$

The degree centrality is incapable of quantifying the network weight of a link for two main reasons. First, it is probable a link with high degree centrality is poorly connected to all other links in a network. Second, a link with few direct connections is not necessarily less important, since it may play a “cut-link” role in a network, and elimination of this breaks down the network into two disconnected pieces. This role is captured by betweenness centrality. As a result, we use the measure of betweenness centrality to quantify the network weight of links.

NETWORK INFRASTRUCTURE: VULNERABILITY

The vulnerability concept examines whether and to what extent degradation of a link affects the network topology and performance, irrespective of the probability of degradation. The network performance is typically assessed by network characteristics, including accessibility, mobility, connectivity, circuitry, and centrality. Taylor and D’Este(20) broadly discussed the concept of vulnerability, and defined vulnerability by using the notion of accessibility. We define the vulnerability, however, by using the notion of betweenness centrality as follows:

“The criticality of the link is determined by the change in the betweenness of other links upon the elimination of the link.”

Consider a network $G(N, L)$, where N and L stand for a set of n nodes and a set of l links. Let us define $C_B(l, G(N, L))$ as betweenness centrality of link l in network $G(N, L)$, and $C_B(l, G(N, L - \{j\}))$ as betweenness centrality of link l following the elimination of link j from network $G(N, L)$. At the link level, the change in betweenness centrality of link l , stemming from the elimination of link j , indicates the spatial influence of link j on link l . This shift reveals the degree of spatial dependence between link j and link l . Accordingly, we introduce $\Delta C_B(l, j)$ term denoting the change in betweenness centrality of link l , when network link j fails. In more formal terms, the $\Delta C_B(l, j)$ term is formulated in Equation 3.

$$\Delta C_B(l, j) = C_B(l, G(N, L)) - C_B(l, G(N, L - \{j\})) \quad (3)$$

At the network level, we formulate the criticality index of the link j in a network per Equation 4. The criticality index of link j indicates the change in betweenness centrality of network following the removal of link j from the network.

$$I_c(j) = \sum_{l=1}^L C_B(l, G(N, L)) - \sum_{l=1}^{L-1} C_B(l, G(N, L - \{j\})) \quad (4)$$

Having the concepts of link centrality and link vulnerability, we discuss the creation process of network weight matrix in the coming section. Specifically, we benefit from Equation 3 to construct the network weight matrix.

1 AN INTRODUCTION TO NETWORK WEIGHT MATRIX

2 The competitive and complementary principle of economic says: (1) two goods are competitive if
 3 an increase in the price of the first goods increases the demand of the second goods, and (2) two
 4 goods are complementary if an increase in the price of the first goods decreases the demand of the
 5 second goods (21). This is true in traffic network analysis as well. Consider two links i and j in
 6 the network $G(N, L)$. We define the competitive and complementary links in a traffic network with
 7 the following terms:

8
 9 **Definition I:** Link i is complementary to link j , when an increase in the cost of link i de-
 10 creases the flow of link j .

11
 12 **Definition II:** Link i competitive with link j , when an increase in the cost of link i increases
 13 the flow of link j .

14
 15 The traditional spatial weight matrices used in spatial analysis are unable to capture the
 16 competitive nature of links, as all the components are positive. They are also simply structured in
 17 light of either geometric or proximity indicators. These matrices suffer from a lack of robust traffic
 18 theory, and thereby are incapable to reflect actual spatial dependence between traffic links. We
 19 remedy this deficiency by introducing the network weight matrix, which allows the data to speak
 20 for themselves. The network weight matrix is an algorithm, which is built from the concept of
 21 vulnerability in network science. The $\Delta C_B(ij)$ term stems from Equation 3 as the corresponding
 22 component of network weight matrix. Accordingly, the components of the network weight matrix
 23 are allowed to take both negative and positive values. In addition, the weights in the matrix are a
 24 function not simply of proximity, but of network topology and demand configuration. To structure
 25 the network weight matrix, we postulate four major hypotheses:

26
 27 **Hypothesis I:** The spatial dependence of competitive links is negative, and more critical
 28 competitive links have a greater absolute spatial dependence weight in a network.

29
 30 **Hypothesis II:** The spatial dependence of complementary links is positive, and more criti-
 31 cal complementary links have a greater spatial dependence weight in a network.

32
 33 **Hypothesis III:** Near links are not necessarily more related than distant links.

34
 35 Let us consider two distinct toy networks $G(N, L)$, as classroom instances for understand-
 36 ing the creation process of the network weight matrix. For simplicity, we test different demand
 37 configurations to show the impact of demand configuration on network weight of links and test our
 38 hypotheses.

39
 40 **Example I:** The first instance encompasses five nodes and six links with equal travel costs.
 41 We consider two distinct demand configurations in this example. The first includes four routes,
 42 which pass the flow by directional links from node 1 to node 5. The second comprises two routes,
 43 which pass the flow by directional links from node 1 to node 3. Figure 1 depicts the toy network
 44 and its distinct paths for these configurations.

45 The network weight matrix for each demand configurations is structured in two steps:

Step I: The betweenness centrality of all links is calculated when all links are in the network,

Step II: Each link is removed from the network (only one at a time) and betweenness centrality is recomputed for the network with the missing link. The results are shown in Table 2.

Step III: Equation 3 is used to compute each element of the network weight matrix. The network weight matrices for two distinct demand configurations are depicted in Figure 2.

FIGURE 1 : A toy network and its link path graph for Example I

For illustration, looking at the very first row of the network weight matrix for 1-5 OD pair in Figure 2, it indicates there is no spatial dependence between link (1,2) and other links, when flow passes from origin 1 to destination 5. It is empirically true as removing link (1,2) does not change the path of flow on the network. Likewise, looking at the second row of this matrix, it shows the negative spatial dependence between link (1,3) and both link (1,2) and link (2,3). This reveals the

TABLE 2 : Betweenness centrality calculation for Example I

Network Scenario	1-5 OD Pair						1-2 OD Pair					
	(1,2)	(1,3)	(2,3)	(3,4)	(3,5)	(4,5)	(1,2)	(1,3)	(2,3)	(3,4)	(3,5)	(4,5)
All links	0	1	0	0	1	0	0	1	0	0	0	0
Without (1,2)	0	1	0	0	1	0	0	1	0	0	0	0
Without (1,3)	1	0	1	0	1	0	1	0	1	0	0	0
Without (2,3)	0	1	0	0	1	0	0	1	0	0	0	0
Without (3,4)	0	1	0	0	1	0	0	1	0	0	0	0
Without (3,5)	0	1	0	1	0	1	0	1	0	0	0	0
Without (4,5)	0	1	0	0	1	0	0	1	0	0	0	0

1-5 OD Pair							1-2 OD Pair						
Links	(1,2)	(1,3)	(2,3)	(3,4)	(3,5)	(4,5)	Links	(1,2)	(1,3)	(2,3)	(3,4)	(3,5)	(4,5)
(1,2)	0	0	0	0	0	0	(1,2)	0	0	0	0	0	0
(1,3)	-1	1	-1	0	0	0	(1,3)	-1	1	-1	0	0	0
(2,3)	0	0	0	0	0	0	(2,3)	0	0	0	0	0	0
(3,4)	0	0	0	0	0	0	(3,4)	0	0	0	0	0	0
(3,5)	0	0	0	-1	1	-1	(3,5)	0	0	0	0	0	0
(4,5)	0	0	0	0	0	0	(4,5)	0	0	0	0	0	0

FIGURE 2 : Network weight matrices for Example I

competitive nature of links in line with our first hypothesis. It is empirically true as removing link (1,3) shifts the path of flow from link (1,3) to link (1,2), and thereby link (2,3) in the network. As expected, the network weight matrix is an asymmetric matrix since the toy network is directional. It is then not surprising that link (1,3) affects link (1,2), but not vice versa. The nature of this is illustrated by the centrality role of link (1,3) in passing flow between the given OD pair. As shown in Figure 2, the network weight matrices are dissimilar for two demand configurations. It emphasizes the essential role of demand configuration in spatial dependence between traffic links.

Example II: The second instance encompasses four nodes and five links. The travel costs of links (1,2), (2,3), (1,3), and (3,4) are equal, while the travel costs of link (2,4) is $\sqrt{3}$ times of other links. Similar to Example I, we test two distinct demand configurations in this example. The first includes three routes, which pass the flow by directional links from node 1 to node 4. The second comprises two routes, which pass the flow by directional links from node 2 to node 4. Figure 3 represents the toy network and its distinct paths for these configurations.

We construct the network weight matrix for each demand configurations as two steps illustrated in Example I. We show the results of step I and step II in Table 3 and Figure 4, respectively. The network weight matrix for 1-4 OD pair in Figure 4 reveals the competitive and complementary nature of traffic links. Looking at the second row of this matrix, we find there is a negative spatial dependence between link (1,3) and both link (1,2) and link (2,4). It means removing link (1,3) shifts the original flow from path 2 to path 3, which is shorter than path 1. According to Hypothesis II, the spatial dependence between link (1,3) and link (3,4) is positive, which demonstrates the complimentary nature of these links. It is illustrated by the fact that no flow passes through link (3,4) when link (1,3) is removed from the network.

FIGURE 3 : A toy network and its link path graph for Example II

TABLE 3 : Betweenness centrality calculation for Example II

Network Scenario	1-4 OD Pair					2-4 OD Pair				
	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)
All links	0	1	0	0	1	0	0	0	1	0
Without (1,2)	0	1	0	0	1	0	0	0	1	0
Without (1,3)	1	0	0	1	0	0	0	1	1	0
Without (2,3)	0	1	0	0	1	0	0	0	1	0
Without (2,4)	0	1	0	0	1	0	0	1	0	1
Without (3,4)	1	0	0	1	0	0	0	0	1	0

1-4 OD Pair						2-4 OD Pair					
Links	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)	Links	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)
(1,2)	0	0	0	0	0	(1,2)	0	0	0	0	0
(1,3)	-1	1	0	-1	1	(1,3)	0	0	0	0	0
(2,3)	0	0	0	0	0	(2,3)	0	0	0	0	0
(2,4)	0	0	0	0	0	(2,4)	0	0	-1	1	-1
(3,4)	-1	1	0	-1	1	(3,4)	0	0	0	0	0

FIGURE 4 : Network weight matrices for Example II

As assumed, the network weight matrix, W_{ij} , comprises both negative and positive elements, which supports the competitive and complementary nature of traffic links. In addition, the more centrality of a link, the greater the network weight. These findings are in light of our first and second hypotheses. We also infer near links are not necessarily more related than distant links. For instance, link (1,3) in Figure 1 is the same distance from links (1,2), (2,3), (3,4), and (4,5). In the case of 1-5 OD pair, while there is a spatial dependency between link (1,3) and both links (1,2) and (2,3), the dependency with links (3,4) and (4,5) is zero. Likewise, link (3,4) and link (3,5) are connected. The dependency between these links is -1 when demand flows from node 1 to node 5, and it is zero when demand flows from node 1 to node 2. These examples not only show how the topology of a network affects the spatial dependence between links, but how the configuration of demand, which shapes the flow on the network, manipulates this dependency.

The traffic network is a complex system, and this complexity forms a convoluted spatial dependence structure between traffic links. The complementary and competitive nature of links is not simply captured by looking only at the structure of the network. For instance, one may infer link (1,3) and link (3,5) in Example I are complementary since they are a series in the network. Likewise, the same is inferred for link (1,3) and link (3,4) in Example II. While the former is not the case, the latter is correct. The demand configuration causes this complexity and deference. Unquestionably, the traditional spatial weight matrix is incapable of reflecting such a convoluted spatial dependence structure, as it simply supposes the spatial dependence is a function of structural proximity. The following section verifies and validate the network weight matrix in a more complex simulation-based problem.

NETWORK WEIGHT MATRIX VERIFICATION: THE SIMULATION-BASED PROBLEM

To assess the capability of network weight matrix over spatial weight matrix, we consider a 3×3 grid test network with 9 nodes and 24 directed links connecting 72 OD pairs. We represent the schematic of the test network in Figure 5.

Pertaining to the traffic network theory, travel demand in a network is either fixed or allowed to vary as a function of travel costs. Likewise, the travel costs for each traffic link is either fixed or an increasing function of link flows (22). To solve the test network problem, we postulate a fundamental assumption: The traffic flow on each link is significantly lower than link capacity, and thereby no congestion exists in the network. Consequently, we relax both variable travel

FIGURE 5 : Simulation-based test problem

demand and variable link travel costs. The traffic assignment problem is then solved by the all-or-nothing assignment condition, since both the user equilibrium and the system-optimal assignment conditions collapse to the all-or-nothing assignment in the case of fixed link travel costs.

For simulation, we assume 100 units of demand for each OD pairs, and randomly assign the cost of travel to each link. For the travel cost assignment, we draw a rational number between 1 and 1.5 from a normal distribution. This threshold is large enough to differ between traffic links, and small enough to prevent the removal of a link and its corresponding routes because of large cost differentials. The trips between origins and destinations are implicitly assigned to the minimum travel cost path. To conduct the network and flow analysis, we developed a script in the AIMSUN traffic simulation software package. To assure the stability of the analysis, we repeat the link cost assignment process and the flow analysis 100 times with different random seeds for the link costs.

We construct the network weight matrix and spatial weight matrix of the test network. We use Equation 3 to create the network weight matrix. To calculate the elements of the network weight matrix, we developed a code in Matlab to use Dijkstra's algorithm (23) to find the shortest travel cost between each OD pair. To configure the spatial weight matrix, we use the first-order adjacency form that has been used in the literature of traffic flow analysis (5, 7). In an undirected network, this is a binary matrix, in which the non-zero elements signify two links directly connected by a node. In a directed network, the traffic flows in one direction, and a link influences its downstream links. Hence, the spatial weight matrix is asymmetric as it allows the impact in one direction (5). We show both matrices in Figure 6 for comparison. In all 100 example repetitions, the spatial weight matrix is a fixed matrix as it is formed based on the network structure. The network weight matrix, however, changes as it is a function of not only the network structure, but the link cost and thus the flow pattern. In the following subsection, we compare the network weight matrix and spatial weight matrix in spatial analysis. We also validate the capability of models comprising the network weight matrix to predict the traffic flow.

		Network Weights Matrix																							
Links		(1,2)	(2,1)	(3,2)	(2,3)	(1,4)	(4,1)	(2,5)	(5,2)	(6,3)	(3,6)	(4,5)	(5,4)	(5,6)	(6,5)	(7,4)	(4,7)	(8,5)	(5,8)	(6,9)	(9,6)	(7,8)	(8,7)	(8,9)	(9,8)
(1,2)		6	0	0	1	-6	0	4	-1	-1	0	-5	0	-1	0	0	-1	0	1	0	0	-1	0	0	0
(2,1)		0	2	0	0	0	-2	-2	0	0	0	0	-2	0	0	0	0	0	0	0	0	0	0	0	0
(3,2)		0	1	8	0	0	-1	4	-3	2	-6	0	-1	0	-7	0	0	-1	0	0	1	0	0	0	-1
(2,3)		0	0	0	2	0	0	-2	0	-2	0	0	0	-2	0	0	0	0	0	0	0	0	0	0	0
(1,4)		-2	0	0	0	2	0	-2	0	0	0	0	-2	0	0	0	0	0	0	0	0	0	0	0	0
(4,1)		0	-6	-2	0	0	6	0	-4	-2	0	-1	4	0	1	1	0	0	0	0	-1	-1	0	0	1
(2,5)		4	-2	4	-4	-6	0	14	0	0	-8	-3	2	2	-6	0	-1	0	1	0	0	-1	0	0	0
(5,2)		-2	0	-2	0	0	-2	0	4	-2	0	1	0	-2	0	-1	0	1	0	0	0	1	0	0	0
(6,3)		0	0	2	-6	0	0	0	-8	8	0	0	0	4	-2	0	0	-2	0	0	2	0	0	0	-2
(3,6)		0	0	-2	0	0	0	-2	0	0	2	0	0	-2	0	0	0	0	0	0	0	0	0	0	0
(4,5)		-4	0	0	-1	0	-4	-2	1	1	0	5	0	2	0	0	-1	0	0	1	0	-1	0	-1	0
(5,4)		0	-7	-3	0	-3	4	2	-2	-3	0	0	14	0	2	-3	4	3	-5	0	-1	0	-7	0	1
(5,6)		0	0	0	-8	0	0	2	-6	4	-4	1	0	14	0	0	-1	2	-3	4	-2	-1	0	-6	0
(6,5)		0	-1	-3	0	0	1	-2	0	-3	0	0	2	0	5	0	1	0	1	-2	0	0	-1	0	-2
(7,4)		0	0	0	0	0	0	0	0	0	0	0	-2	0	0	2	0	-2	0	0	0	-2	0	0	0
(4,7)		-1	0	0	0	1	0	-1	0	0	0	-2	4	0	0	0	7	0	-7	0	0	1	-6	0	0
(8,5)		0	0	-1	0	0	0	0	1	-1	0	-5	2	2	-3	-7	0	13	0	0	-6	4	-3	-3	3
(5,8)		1	0	1	0	-1	0	2	0	0	-1	0	-2	0	1	0	-3	0	5	-2	0	-3	0	0	-2
(6,9)		0	0	-1	0	0	0	-1	0	0	1	1	0	4	-1	0	-1	0	-5	6	0	-1	0	-6	0
(9,6)		0	0	1	0	0	0	0	-1	1	0	0	0	-2	0	0	0	-3	0	0	3	0	0	0	-3
(7,8)		0	0	0	0	0	0	0	0	0	0	-7	0	-1	0	-6	1	4	-2	-1	0	7	0	1	0
(8,7)		0	0	0	0	0	0	0	0	0	0	0	-2	0	0	0	-2	-2	0	0	0	0	2	0	0
(8,9)		0	0	0	0	0	0	0	0	0	0	0	0	-2	0	0	0	-2	0	-2	0	0	0	2	0
(9,8)		0	0	0	0	0	0	0	0	0	0	0	-1	0	-5	0	-1	3	-1	0	-5	0	1	0	5

		Spatial Weights Matrix																							
Links		(1,2)	(2,1)	(3,2)	(2,3)	(1,4)	(4,1)	(2,5)	(5,2)	(6,3)	(3,6)	(4,5)	(5,4)	(5,6)	(6,5)	(7,4)	(4,7)	(8,5)	(5,8)	(6,9)	(9,6)	(7,8)	(8,7)	(8,9)	(9,8)
(1,2)		0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(2,1)		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(3,2)		0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(2,3)		0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(1,4)		0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
(4,1)		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(2,5)		0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0
(5,2)		0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(6,3)		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(3,6)		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
(4,5)		0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
(5,4)		0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
(5,6)		0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
(6,5)		0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
(7,4)		0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
(4,7)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
(8,5)		0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
(5,8)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
(6,9)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
(9,6)		0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
(7,8)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
(8,7)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
(8,9)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
(9,8)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0

FIGURE 6 : Network and spatial weight matrices of the test network

1 Calibration and Model Specification

2 The traffic flow on a link is not only related to the cost of the link, but the cost of other links in
 3 the network. This relation is captured by the network weight matrix. We assume reflecting the
 4 cost of other traffic links has the potential to improve the traffic flow prediction. In this section, we
 5 develop three distinct models to test this assumption.

6
 7 **Model I:** This model simply considers a linear relationship between traffic flow on each
 8 link and its corresponding travel cost.

9
 10 **Model II:** This model includes not only a linear relationship between traffic flow on each
 11 link and its corresponding travel cost, but the cost of other traffic links. To capture this spatial
 12 relationship, we test the first order adjacency spatial weight matrix as shown in Figure 6.

13
 14 **Model III:** This model includes not only a linear relationship between traffic flow on each
 15 link and its corresponding travel cost, but the cost of other traffic links. To capture this spatial
 16 relationship, we test the network weight matrix as shown in Figure 6.

17
 18 We develop a bivariate linear regression (BLR) for Model I, and spatial cross-regressive
 19 for Model II and Model III. The spatial cross-regressive model, also called spatial lag of X (SLX),
 20 takes the spatial lags of explanatory variables into consideration. This model allows us to examine
 21 the effects of travel cost spillovers. We use ordinary least squares (OLS) estimation, as it is the
 22 best linear unbiased estimator (BLUE) of the coefficients for both types of models (24). Comparing
 23 these models enable us to assess and verify the power of the proposed network weight matrix in
 24 traffic analysis. A point worthy of attention is that we do not aim to delve into an analysis of traffic
 25 flow and travel cost, but we intend to test whether and to what extent the network weight matrix
 26 augments spatial analysis of a traffic network. Hereafter, we assume a simple linear relationship
 27 between traffic flow and travel cost.

28 We represent the results of the models over 100 iterations in Table 4. Pertaining to Model
 29 1, the travel cost of a link is negatively correlated with the flow on that link, which is in line with
 30 our hypothesis. This variable is significant at the 90 percent confidence interval in all iterations.
 31 The goodness-of-fit of the model fluctuates between 0.089 and 0.583 with the average value of
 32 0.290. Pertaining to Model II, the travel cost of a link has a negative correlation with flow on that
 33 link. The coefficient of the spatial term varies from negative to positive values, and it is statisti-
 34 cally insignificant in 83 percent of the cases. This indicates there is no steady spatial dependency
 35 between flow of a link and the cost of other links in our test network. In addition, comparing
 36 the goodness-of-fit of Model I and Model II does not show any model improvement. On average,
 37 adding the spatial term in the model augments the adjusted R^2 of the model by only 1 percent. We
 38 conclude that the spatial dependency that is captured by spatial weight matrix is unsuccessful in
 39 explaining the spatial relationship between traffic links.

40 Pertaining to Model III, the coefficient of the travel cost has a negative sign and is almost
 41 always statistically significant at the 99 percent confidence interval. The goodness-of-fit of the
 42 model ranges from 0.435 to 0.950 with the average value of 0.713. The results disclose that using
 43 the weight matrix augments the adjusted R^2 of the model, on average, by about 40 percent. To give
 44 the reader a sense of comparing the general goodness-of-fit of the models in all iterations, we graph
 45 the value of the adjusted R^2 of the models in Figure 7. As shown, there is not a significant difference

TABLE 4 : Linear regression and spatial cross-regressive models for 100 iterations

2*Models	Model I			Model II			Model III		
	Min	Max	Avg.	Min	Max	Avg.	Min	Max	Avg.
Intercept	1495.74	3222.17	2177.22	-1177.67	-101.88	-553.36	1462.79	3351.64	2073.791
t-test	2.82	7.04	4.50	-0.77	-7.90	-3.66	1.99	7.25	4.09
Coefficient	-2126.79	-717.94	-1272.20	-321.46	-119.27	-206.05	-2194.71	-703.99	-1278.24
t-test	-1.80	-5.76	-3.26	-4.33	-21.03	-8.16	-1.71	-6.48	-3.24
Coefficient	-	-	-	-	-	-	-94.07	282.95	49.56
t-test	-	-	-	-	-	-	0.09	3.79	0.64
Adj. R^2	0.089	0.583	0.290	0.435	0.950	0.713	0.062	0.669	0.304

1 between Model I and Model II, since in different iterations one is better than the other. Model III,
2 however, is significantly performs better than the other models in all cases. Drilling down further,
3 we calculate the ratio of the adjusted R^2 of the models to observe the magnitude of difference
4 between the goodness-of-fit of the models. We depict the cumulative density function of these
5 ratios in Figure 8. According to this figure, Model II outperforms Model I in 70 percent of cases.
6 This superiority, however, is less than 1.5 times in 65 percent of situations. In all cases, Model 3
7 outperforms both Model I and Model II. As shown, Model IIIperforms 2 times and 4 times better
8 than the other models in 70 percent and 20 percent of the cases, respectively. This leads inexorably
9 to the conclusion that the network weight matrix represents a more accurate spatial dependency
10 between traffic links, and thereby improves traffic flow prediction.

FIGURE 7 : Adjusted R^2 of models in 100 iterations

11 CONCLUDING REMARKS

12 The spatial dependence between traffic links is potentially important but largely unexplored in
13 traffic network analysis. The acceptance of the first law of geography governs this belief that
14 adjacent traffic links exhibit a positive spatial dependence, and it decays with increase of distance.
15 A set of structures for the spatial weight matrix have emerged to take the spatial dependence of
16 adjacent links into account. As of today, traffic scientists have forced spatial weight matrices to

FIGURE 8 : Cumulative density function for the ratio of Adjusted R^2 of models in 100 iterations

1 have positive elements, as it is presumed adjacent traffic links have a positive dependency. This
 2 has led to overlooking the competitive nature of traffic links in a network, and thereby has misled
 3 traffic flow analysis.

4 To examine the existent spatial dependence in a traffic network, we propose a novel network
 5 weight matrix. This matrix stems from the concept of betweenness centrality and vulnerability
 6 in network science. We structure the network weight matrix to take into consideration not only
 7 network topology and infrastructure, but demand configuration. This study introduces a new way
 8 of viewing spatial traffic flow dependence in the context of network econometrics. The network
 9 weight matrix introduced in the current research has distinctive characteristic, which are capable
 10 of reflecting actual spatial dependence between traffic links:

- 11 • The elements are allowed to have negative and positive values, which capture competitive
 12 and complementary nature of links.
- 13 • The diagonal elements are not fixed to zero, which takes the self-dependence of link i
 14 upon itself into consideration.
- 15 • The elements not only reflect the spatial dependence based on the network topology and
 16 the costs of links, but they acknowledge the demand configuration as well.

17 The network weight matrices are theoretically defensible as they are based on the traffic
 18 theory. As the elements of the network weight matrix more closely reflect the dependence struc-
 19 ture of the traffic links, the weight matrix becomes more accurate and defensible. Not only do they
 20 capture the competitive and complementary nature of links, they embed additional network dy-
 21 namics such as cost of links and demand configuration. This study begins a new line of research in
 22 examining actual spatial dependence between traffic links by introducing a network weight matrix.

23 The following topics are recommended for further exploration:

- How does the network weight matrix illustrate the spatial correlation of traffic components in a network?
 - How does the network weight matrix augment the prediction power of the traditional traffic forecasting models, including STARIMA and generalized STARIMA models?
 - How does the performance of the network weight matrix change in a congested traffic network?
- Although we tested the network weight matrix in a traffic network problem, it has potential for implementation in other disciplines of spatial and network science. We believe the network weight matrix has further applications in models of physical flow, but also social networks for which links may be either competitive or complementary with each other.

REFERENCES

- [1] Waldo R Tobler. A computer movie simulating urban growth in the detroit region. *Economic geography*, 46:234–240, 1970.
- [2] Luc Anselin. *Spatial econometrics: methods and models*, volume 4. Springer Science & Business Media, 2013.
- [3] Arthur Getis and Jared Aldstadt. Constructing the spatial weights matrix using a local statistic. In *Perspectives on spatial data analysis*, pages 147–163. Springer, 2010.
- [4] Francois Bavaud. Models for spatial weights: a systematic look. *Geographical analysis*, 30(2):153–171, 1998.
- [5] Tao Cheng, James Haworth, and Jiaqiu Wang. Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems*, 14(4):389–413, 2012.
- [6] Wanli Min and Laura Wynter. Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, 19(4):606–616, 2011.
- [7] Yiannis Kamarianakis and Poulicos Prastacos. Forecasting traffic flow conditions in an urban network: comparison of multivariate and univariate approaches. *Transportation Research Record: Journal of the Transportation Research Board*, (1857):74–84, 2003.
- [8] Xinyu Min, Jianming Hu, and Zuo Zhang. Urban traffic network modeling and short-term traffic flow forecasting based on gstarima model. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 1535–1540. IEEE, 2010.
- [9] Jared Aldstadt and Arthur Getis. Using amoeba to create a spatial weights matrix and identify spatial clusters. *Geographical Analysis*, 38(4):327–343, 2006.
- [10] Peter Gould. Letting the data speak for themselves. *Annals of the Association of American Geographers*, 71(2):166–176, 1981.
- [11] Arnab Bhattacharjee and Chris Jensen-Butler. Estimation of the spatial weights matrix under structural constraints. *Regional Science and Urban Economics*, 43(4):617–634, 2013.

- 1 [12] Sungyup Chung and Geoffrey JD Hewings. Competitive and complementary relationship
2 between regional economies: a study of the great lake states. *Spatial Economic Analysis*,
3 10(2):205–229, 2015.
- 4 [13] Frank Harary. The determinant of the adjacency matrix of a graph. *Siam Review*, 4(3):202–
5 210, 1962.
- 6 [14] Benny Sudakov and Van H Vu. Local resilience of graphs. *Random Struct. Algorithms*,
7 33(4):409–433, 2008.
- 8 [15] Alex Bavelas. A mathematical model for group structures. *Human organization*, 7(3):16–30,
9 1948.
- 10 [16] Alex Bavelas. Communication patterns in task-oriented groups. *Journal of the acoustical*
11 *society of America*, 1950.
- 12 [17] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages
13 35–41, 1977.
- 14 [18] Stephen P Borgatti. Centrality and network flow. *Social networks*, 27(1):55–71, 2005.
- 15 [19] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*,
16 1(3):215–239, 1978.
- 17 [20] Michael AP Taylor and Glen M D’Este. *Transport network vulnerability: a method for*
18 *diagnosis of critical locations in transport infrastructure systems*. Springer, 2007.
- 19 [21] Shinichi Ichimura. A critical note on the definition of related goods. *The Review of Economic*
20 *Studies*, 18(3):179–183, 1950.
- 21 [22] DE Boyce and BN Janson. A discrete transportation network design problem with com-
22 bined trip distribution and assignment. *Transportation Research Part B: Methodological*,
23 14(1):147–154, 1980.
- 24 [23] Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathe-*
25 *matik*, 1(1):269–271, 1959.
- 26 [24] James P LeSage. What regional scientists need to know about spatial econometrics. *Available*
27 *at SSRN 2420725*, 2014.