A Discrete Optimization Approach for Locating Automatic Vehicle Identification Readers for the Provision of Roadway Travel Times

Hanif D. Sherali[1], Jitamitra Desai[2], Hesham Rakha[3], Ihab El-Shawarby[4]

November, 2002

Acknowledgement: This work was supported by a grant from the *Virginia Department of Transportation*, the *National Science Foundation* under Grant No. DMI – 0094462, and the *ITS Implementation Center*.

[1] Professor. Grado Department of Industrial and Systems Engineering., Virginia Tech, Blacksburg, VA.
[2] Graduate Research Assistant. Grado Department of Industrial and Systems Engineering., Virginia Tech, Blacksburg, VA.
[3] Assistant Professor. Charles Via Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA.
[4] Visiting Scholar. Virginia Tech Transportation Institute, Virginia Tech, Blacksburg, VA.

ABSTRACT

This paper develops an algorithm for optimally locating surveillance technologies with an emphasis on Automatic Vehicle Identification tag readers by maximizing the benefit that would accrue from measuring travel times on a transportation network. The problem is formulated as a quadratic 0-1 optimization problem where the objective function parameters represent benefit factors that capture travel time variability along specified trips. The INTEGRATION software is utilized to derive these benefit factors for four freeway section types that include merge, diverge, weaving, and bottleneck sections. The approach also develops two composite functions that estimate travel time variability along a trip that may constitute any of the four identified segments. The simulation results are recorded as generic look-up tables that can be used for any freeway section for the purpose of computing the associated benefit factor coefficients. An optimization approach based on the Reformulation-Linearization Technique coupled with semidefinite programming concepts is designed to solve the formulated reader location problem. Computational results are presented using data pertaining to a freeway section in San Antonio, Texas, as well as synthetic test cases, to demonstrate the efficacy of the proposed approach.

Keywords. Travel time, surveillance technologies, Automatic Vehicle Identification (AVI), reader location, simulation, Reformulation-Linearization Technique (RLT), semidefinite programming.

1. Introduction

The principal focus of traffic management and control systems is to ensure the safe and efficient movement of traffic on roadways. This requires surveillance, detection, and monitoring of freeways, arterial roads, and intersections to glean specific information regarding the flow of traffic. An effective surveillance/detection system provides a foundation of information on which Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS) depend. For example, such surveillance systems can be utilized to estimate roadway travel times that can be displayed to the public via Variable Message Signs (VMS), Highway Advisory Radio (HAR), and/or on the World Wide Web (WWW).

This paper stems from the interest of the Virginia Department of Transportation (VDOT) to explore Automatic Vehicle Identification (AVI) systems for estimating roadway travel times that can be displayed via VMS, HAR, and WWW technologies. Such a system relies on a combination of passive tags attached to vehicles and electronic interrogators (or readers) mounted on overhead structures. The system detects the passage of a vehicle at each reading location within the network by monitoring the signal sent by the interrogator's antenna. A vehicle is detected each time an altered signal is reflected back to the antenna. Since each vehicle tag returns a unique signal, specific vehicles can then be identified. Upon detection of a vehicle, the system transmits the detection information, via a modem, to the AVI data processing system center, where an algorithm matches tag reads to estimate travel times between readers. One such algorithm developed by Dion and Rakha [*1*] is unique in three aspects. First, it is designed to handle both stable (mean constant) and unstable (varying mean) traffic conditions. Second, the algorithm can be successfully applied for low levels of market penetration (less than 1%). Third, the algorithm works for both freeway and signalized arterial roadways. This algorithm utilizes a robust data-filtering procedure by identifying valid data within a dynamically varying search window. The size of the search window varies as a function of the number of observations within the current interval, the number of observations in the previous sampling intervals, and the number of consecutive observations outside the search window.

It should be noted that the location procedures that are developed are not restricted to use with AVI tag readers; instead the paper develops procedures for locating surveillance technology for the estimation of roadway travel times regardless of the type of technology. While the use of spot speed measurements (e.g. loop detectors) could be utilized within the proposed framework, the main thrust of the paper is on spatial measurements of travel times.

The use of such electronic technologies to study traffic characteristics began in the mid-nineteen eighties. However, the use of simulation as a tool in Advanced Transportation Management Systems (ATMS) has come to the fore in recent times [*2-4*]. Simulation experiments have been utilized to study the correlation between link travel times and detector outputs [*5*], and an algorithm for re-identifying vehicles between two consecutive detector stations has also been proposed in the literature [*6*]. The relationship between detector location and the ability of a system to monitor traffic behavior and to estimate link travel characteristics using CORSIM (a *micro-simulation* program) for *arterials* was investigated in detail by Thomas [*7-8*], and is deemed to be a pioneering work in the area of using simulation tools for detector location. In this paper, we study the location of detectors for *freeways* using an optimization approach, with the objective of capturing the maximum variability in travel times, where certain related benefit parameters in the model formulation are estimated using a simulation study.

A central problem that arises in the deployment of this emerging AVI surveillance technology is to determine the number and location of detection stations or readers that would best cover the network under consideration by providing a maximum degree of information about traffic variability in the network, subject to certain resource limitation constraints. Towards this end, one needs to first of all assess the variability in traffic conditions over the transportation network. We perform this task via a simulation analysis using the INTEGRATION package [*9-11*]. Generic look-up tables for mean and variance of travel times along freeway sections are developed depending on the traffic demand and roadway characteristics. Two composite trip-based functions are developed to assess the benefit that might accrue from measuring travel times along a trip that is comprised of several such basic freeway sections, where the benefit reflects the ability to capture information regarding the variability in travel times. Based on this construct, a mathematical model is formulated for determining an optimal number of readers, and their prescribed locations (vehicle detection stations), so as to maximize the total benefit accruing from their usage. These two tasks constitute the principal thrust of the present paper.

The remainder of this paper is organized as follows. In Section 2, we describe the construction of the graph on which the optimization problem is to be solved, and subsequently present a formulation for the reader location problem. Section 3 deals with the determination of the benefit factors b_{ij} that constitute the objective function parameters. The solution of the Reader Location problem using the Reformulation-Linearization Technique is advocated in Section 4, and results for several simulated test cases are presented in Section 5. Finally, Section 6 and Section 7 conclude the paper and provide some recommendations for further research.

2. Formulation of the Reader Location Problem

Consider the following conceptualization of the problem. Suppose that we are given a certain transportation road network that identifies the various freeways. Furthermore, on this network, suppose that we identify various *Origin-Destination* (O-D) pairs (i, j) belonging to some set A for which we might be interested in measuring travel times, if economically feasible to do so. We assume that for each O-D pair, there exists some unique route so that if a reader is installed at each of the end nodes i and j, then the corresponding travel time data obtained would pertain to this route. Furthermore, the routes connecting these O-D pairs may or may not share common sections of roadways (links). Naturally, we would be able to measure travel times for any $(i, j) \in A$ if and only if we locate a reader at each of the upstream and downstream end-points i and j.

Accordingly, given such an urban road network for a particular region, we construct a suitable graph in order to formulate the underlying reader location problem. This graph, denoted $G(N, A)$, has a node set N given by the union of the end-points of the various O-D pairs, and has a set of directed arcs $(i, j) \in A$, each of which represents a particular O-D pair and its corresponding route as identified above. Note that if we are interested in measuring travel times on O-D pairs (i, j) as well as (j, i) for any pair of nodes i and j, we would include both (i, j) and (j, i) in A.

As far as the desirability of covering any link $(i, j) \in A$ is concerned, it is relatively more beneficial to cover a link that might represent an O-D pair that enjoys significant usage, as well as exhibits a greater variability in travel times. The testing of other objective functions is possible but beyond the scope of this paper. Accordingly, suppose that we develop some suitable mechanism for prescribing a benefit factor b_{ij} for covering link $(i, j) \in A$. (A method for deriving such factors is described in Section 3 of this paper.) Furthermore, let C_j denote the site-specific cost of installing a reader at location $j \in N$, and let R denote the maximum number of available readers. In addition, suppose that we also have a maximum budgetary limitation B. Then, defining a binary variable

$$y_j = \begin{cases} 1, \text{ if a reader is located at node } j, \\ 0, \text{ otherwise}, \qquad \forall\, j \in N, \end{cases} \tag{1}$$

we can formulate the *Reader Location* problem (**RL**) as follows.

$$\begin{aligned} \textbf{RL}: \quad & \text{Maximize} \quad \sum\sum_{(i,j)\in A} b_{ij} y_i y_j & \text{(1a)} \\ & \text{subject to :} \quad \sum_{j\in N} y_j \le R & \text{(1b)} \\ & \sum_{j\in N} C_j y_j \le B & \text{(1c)} \\ & y \text{ binary.} & \text{(1d)} \end{aligned}$$

The objective function (1a) seeks to maximize the total coverage benefit. Constraint (1b) asserts that the total number of readers should not exceed the available maximum number R, constraint (1c) imposes a budgetary restriction on the total acquisition plus installation costs, and constraint (1d) represents the binary restrictions on the decision variables y. Observe that if the maximum possible number of readers that could be installed subject to the budget restrictions (1c) is itself less than or equal to R, (i.e., the sum of the R + 1 cheapest readers exceeds B), then constraint (1b) is redundant and can be omitted. On the other hand, if the maximum total installation cost for R readers is no more than B (i.e., the sum of the R most expensive readers is less than or equal to B), then (1c) is redundant and can be deleted. In this latter case, because of the nature of the objective function, we can impose (1b) as an equality constraint.

3. Determination of the Benefit Factors (b_{ij})

In the reader location problem **RL**, the objective function parameters are the benefit factors for the corresponding arcs in the constructed graph G. In general, benefit accrues from garnering as much information about the variability in travel times as possible. Consequently, a first attempt at formulating the problem is to identify an objective function that maximizes the travel time variability coverage of the surveillance system. Specifically, if the travel time on a link is nearly the same throughout the day (with possibly only seasonal variations), then it is of little value to install special devices to measure real-time travel times given that drivers will be familiar with the traffic conditions. Instead, off-line sampling techniques could be used to estimate typical travel times in this case. On the other hand, it is far more important to capture real-time information about the movement of traffic on a link having a considerable variation in travel time during different periods of the day.

For example, Figure 1 demonstrates the variability of travel times along a freeway segment along I-35 in San Antonio, TX. This figure clearly demonstrates both a high level of variability in travel times from one day to another, in addition to a high level of variability in travel times by time-of-day. Consequently, the need to predict travel times becomes more desirable in such cases. It should be noted that the travel times drop below the free-speed travel time because the confidence limits were computed considering two standard deviations of a log-normal distribution.

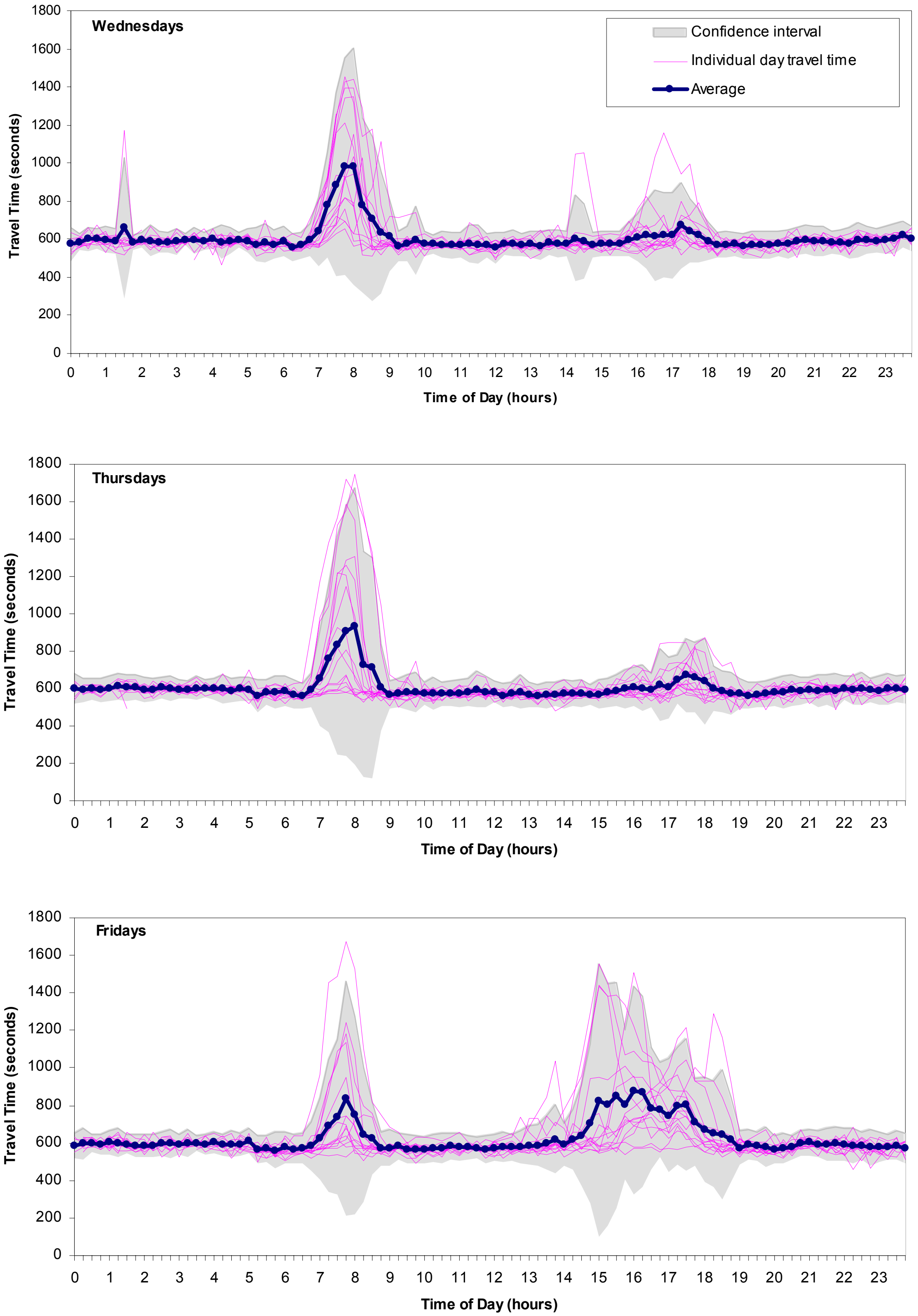

Figure 1: Observed travel times on I-35 South between the AVI stations at O'Connor Rd. and New Braunfels Rd. between June 11th and September 30th, 1998

To illustrate our proposed methodology for assessing such benefit factors, we focus on a transportation network that is comprised of freeway segments in the form of a chain graph, or more generally, having a tree structure. A node is identified in this network wherever the roadway geometry changes (e.g. an entering ramp, an exit ramp, or a change in the number of lanes), and the roadway between any two adjacent nodes is modeled as a link (arc). Links can be easily divided into smaller links with no impact on the formulation if a tighter spacing of readers is being contemplated. The disaggregation of a link into a series of shorter links that the location of the AVI tag readers can be investigated in a denser fashion, although this would increase the computational load burden on the solution algorithm. Each segment of this transportation network between some origin-destination pair for which we are interested in measuring travel times might be composed of several such links. These segments effectively translate to the arcs in the model graph G, with the union of the end-points of these segments translating to the nodes of G, as described above.

Furthermore, we assume that any particular link is composed of four basic sections, namely an On-Ramp or Merge section, an Off-Ramp or Diverge section, a Weaving section, and/or a Bottleneck section. Figure 2 illustrates these four sections, where the specifications given are in accordance with the Highway Capacity Manual (Transportation Research Board, Washington D. C, 2000). If the benefit factors can be estimated for these individual sections, then the benefit factor for any link can be determined by using a suitable combination of these basic factors as delineated in the sequel. In addition, the following principal attributes pertaining to a section, or a link, affect travel times.

1. Expected demand on the link (veh/h or V/C ratio);
2. Demand distribution (ratio of vehicles on the freeway to vehicles on the corresponding on-ramp, or exit, or weaving, or bottleneck sections);
3. Number of lanes on the freeway;
4. Number of lanes on the section (on-ramp, exit, weaving, bottleneck);
5. Auxiliary (Acceleration/Deceleration/Weaving) lane length (m).

Travel time characteristics were determined using simulation for various combinations of each of the above attributes. While the authors acknowledge that field data would have been more desirable to develop these travel time variability factors, however these data were not available at the time of this study. Furthermore, it would be very difficult to find roadway and traffic conditions that would cover the wide range of scenarios that were simulated. Consequently, the approach used a state-of-the-art microscopic stochastic traffic simulation and assignment model, namely the INTEGRATION software. Specifically, for each resulting combination, the simulation process produces a series of travel time values for which the mean and standard deviation can be computed, yielding the coefficient of variation defined as COV = standard deviation / mean. This statistic provides a normalized measure of the variability of travel time for a particular section of roadway under specified traffic demand characteristics. The fact that the COV is a normalized measure of variability allows this factor to be aggregated with the COVs for other sections that might experience much longer average travel times.

Below, we discuss the analysis conducted for an on-ramp section in concert with the above five traffic flow and roadway attributes. A similar analysis was conducted for each of the other three sections including diverge, weaving, and bottleneck sections [*13*].

3.1 Overview of the INTEGRATION Model

Prior to discussing the specifics of the simulation analyses, the INTEGRATION software used for this purpose [*9 – 12*] is described briefly. This software, which was conceived as an integrated simulation and traffic assignment model, performs simulations by explicitly tracking the movement of individual vehicles at a deci-second level of resolution. This microscopic approach allows the software to model car-following, lane-changing, and gap acceptance behavior at a high level of fidelity. The model's microscopic simulation also permits considerable flexibility in representing spatial variations in traffic conditions, in addition to considering time variations in traffic demands. The INTEGRATION software was selected because its validity has been tested significantly and demonstrated in the context of several real-life applications, such as for example, the modeling of the entire Salt Lake City region, Hwy 401 in Toronto, and the Columbia Pike in Arlington, VA.

The INTEGRATION 2.30 software allows the user to alter the random number seed within the simulation model to capture the stochastic nature of traffic. This stochastic nature in captured in two ways. First, the variability in vehicle departures from origin-destination zones can be varied using a fully random (negative exponential vehicle headways) to partially random headways (shifted negative exponential distribution). Second, the software allows for the modeling of stochastic differences in driver car-following and lane changing behavior using a speed variability factor that mimics the macroscopic platoon dispersion behavior. A subsequent publication will compare the platoon dispersion behavior to macroscopic behavior and field data.

3.2 Simulation Analysis for On-Ramp Sections

As shown in Figure 2, sections of a freeway that lie within 450m, both upstream and downstream, of an on-ramp characterize an on-ramp or merge section. While performing the simulation runs, the on-ramp section was modeled using the INTEGRATION software with nodes 1 and 2 as origin nodes and node 5 as the destination node. The freeway was modeled as consisting of two lanes, whereas the on-ramp consists of just one lane, which is typical of freeways in urban and rural areas. That is to say, links (1, 3) and (4, 5) were designated to have two lanes each, but link (2, 3) has just one lane. Note that link (3, 4) was modeled as having three lanes because an acceleration lane was provided for traffic coming in from the on-ramp, where the length of this link depends on the acceleration lane length, which was varied as part of the analysis. The freeway capacity or on-ramp capacity was assumed to be 2,000 veh/h/lane based on field loop detector data gathered from different locations in North America, as illustrated in Figures 3 and 4. While the field data do demonstrate isolated flow observations in the range of 2300 veh/h, however, these observations represent random representations that are not sustained for periods of 15 minutes as is required for estimating roadway capacity. The speed-flow relationship of Figure 4 was input to the INTEGRATION software.

The study considered different lengths of the acceleration lane ranging from 0m (no acceleration lane) to 200m in increments of 50m, resulting in five distinct on-ramp sections. The analysis was performed for each of these five independent on-ramp sections.

To characterize the traffic demand, the *V/C* ratio for the freeway was varied from 0.25 to 1.00 in increments of 0.125, thereby yielding seven different demand patterns for the freeway. Note that this demand originates at node 1, and has its destination node as node 5. Similarly, the on-ramp *V/C* ratio was varied from 0.1 to 0.9 in increments of 0.2, yielding five different demand patterns for the on-ramp, with the origin node being node 2 and the destination node being node 5. This produced 35 different (freeway *V/C* ratio, on-ramp *V/C* ratio) pairs. The simulation logic generated the corresponding traffic demand using a negative exponential distribution for the inter-arrival times of vehicles, and hence, the arrivals themselves follow a Poisson distribution. The random number seed within INTEGRATION, used to generate realizations from these distributions, was varied from 1 to 5 in steps of 1. This yielded five random simulation runs for each of the 35 pairs of freeway and on-ramp demands, and the resulting mean travel times for each of the five runs were used to compute the required average and standard deviation of travel time for each pair of demand patterns. The COV values thus obtained have been recorded in look-up tables (see [*13*] for details of these results), so that given a traffic demand, and the type of basic freeway section under consideration, the corresponding COV value can be obtained from these look-up tables.

In summary, a total of 875 runs were executed in computing the on-ramp benefit functions. These 875 runs covered 5 acceleration lane scenarios, 7 freeway demand scenarios, 5 on-ramp demand scenarios, and 5 random number seed scenarios. Similar procedures were utilized for the off-ramp, weaving, and bottleneck sections.

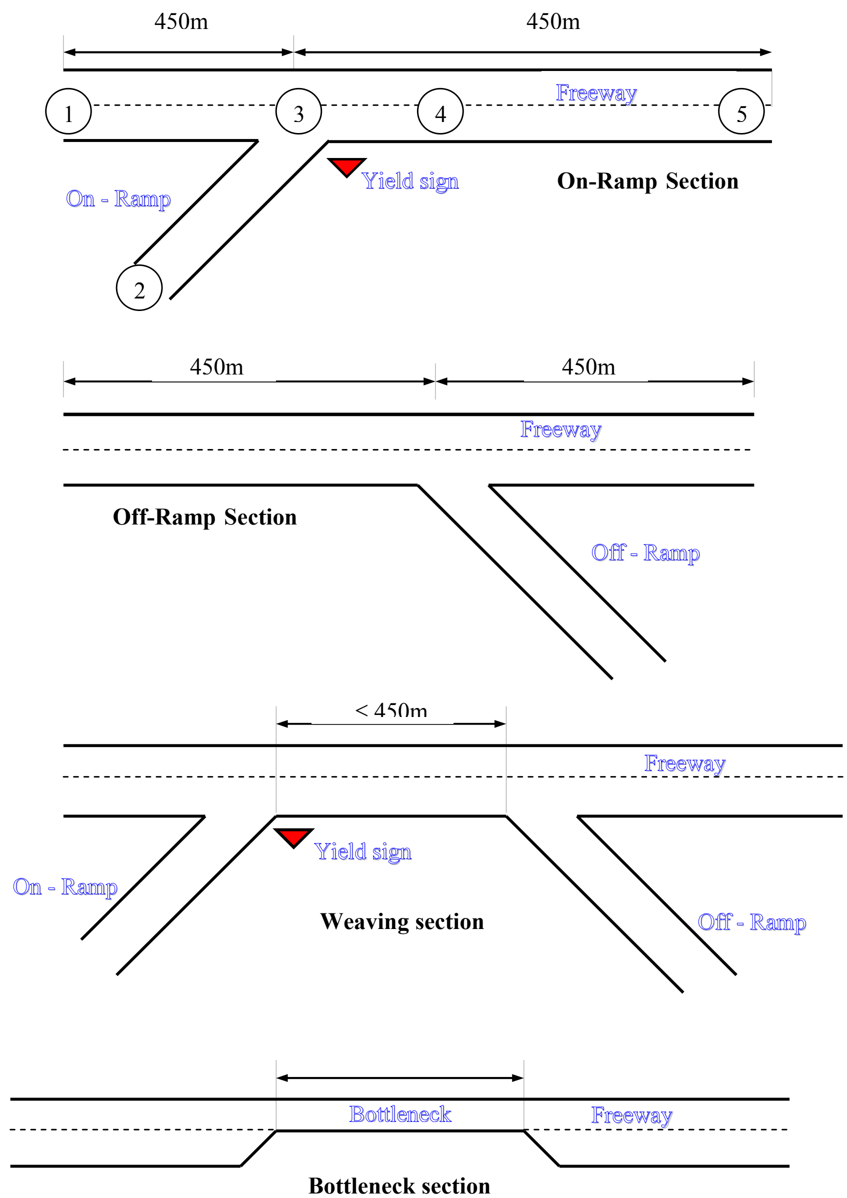

Figure 1: Basic freeway sections considered in the analysis

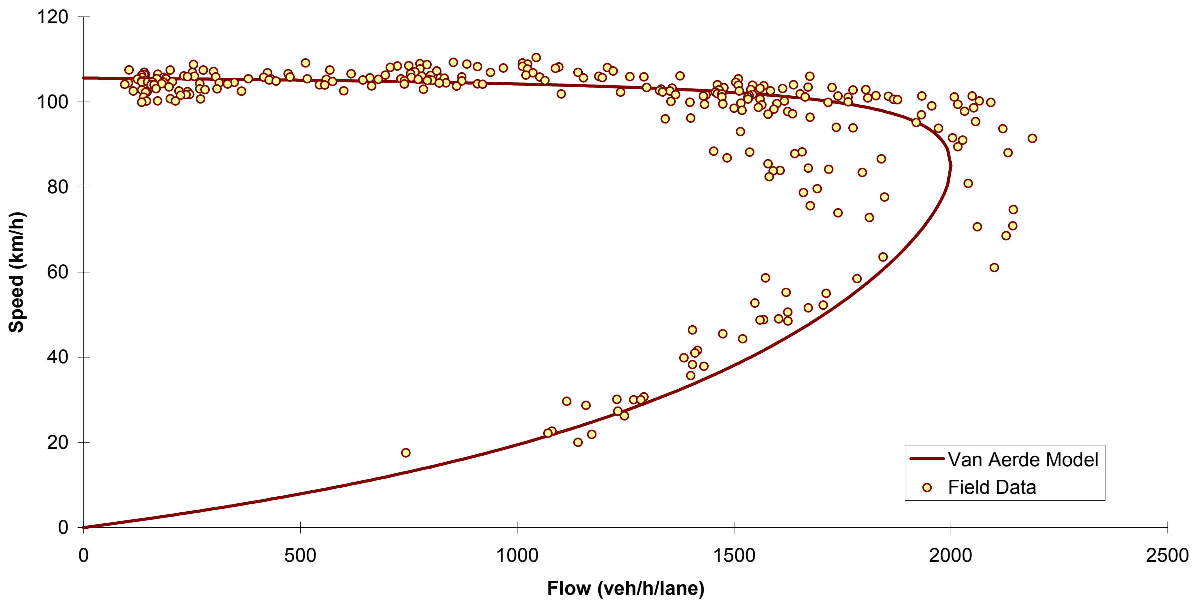

Figure 2: Sample speed-flow relationship (Hwy401, Toronto)

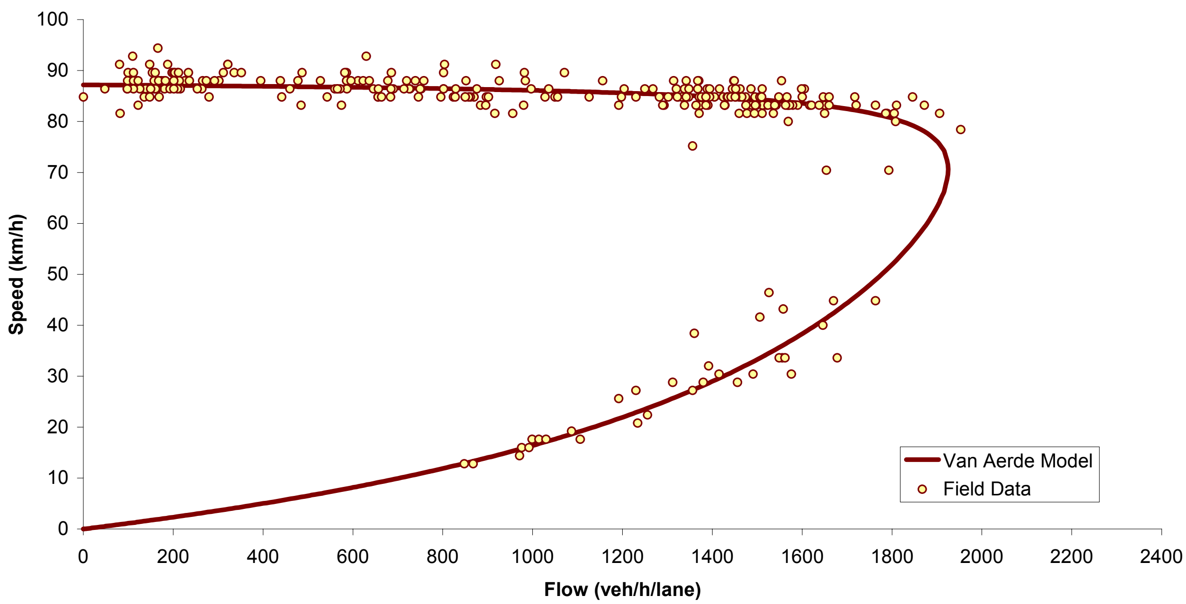

Figure 3: Sample speed-flow relationship (I-4 Orlando, Florida)

For the purpose of illustration, the surface plots of the Freeway Average Travel Time and the Freeway COV for an on-ramp section having an acceleration lane length of 50m are shown in Figures 5 and 6, respectively. These figures clearly demonstrate that the average travel time increases as the on-ramp and freeway demand increases, as would be expected. The results also indicate that the variability in traffic conditions is higher when the *V/C* ratios on the on-ramp and freeway section are in the range of 70 percent when conditions are typically relatively highly congested.

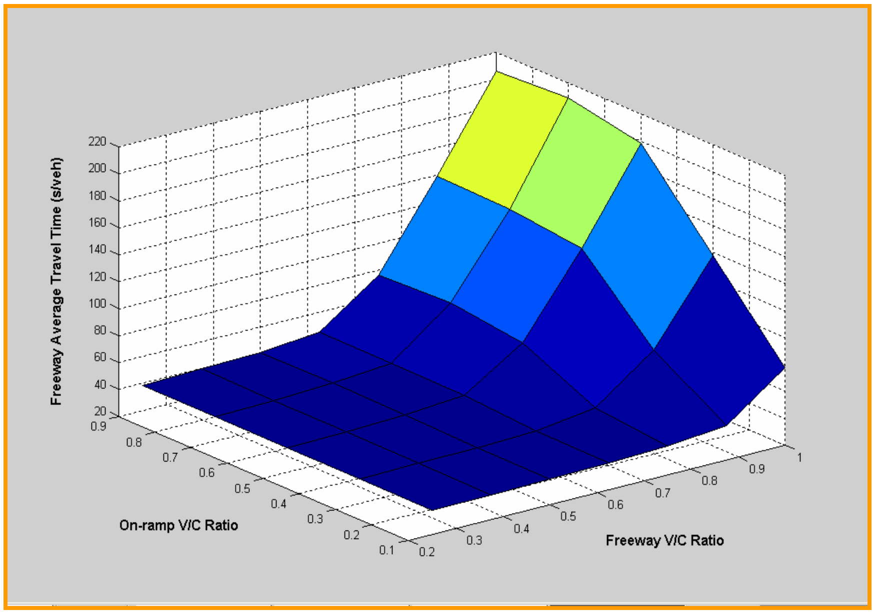

Figure 4: Freeway average travel time for an on-ramp section having an acceleration lane length of 50m

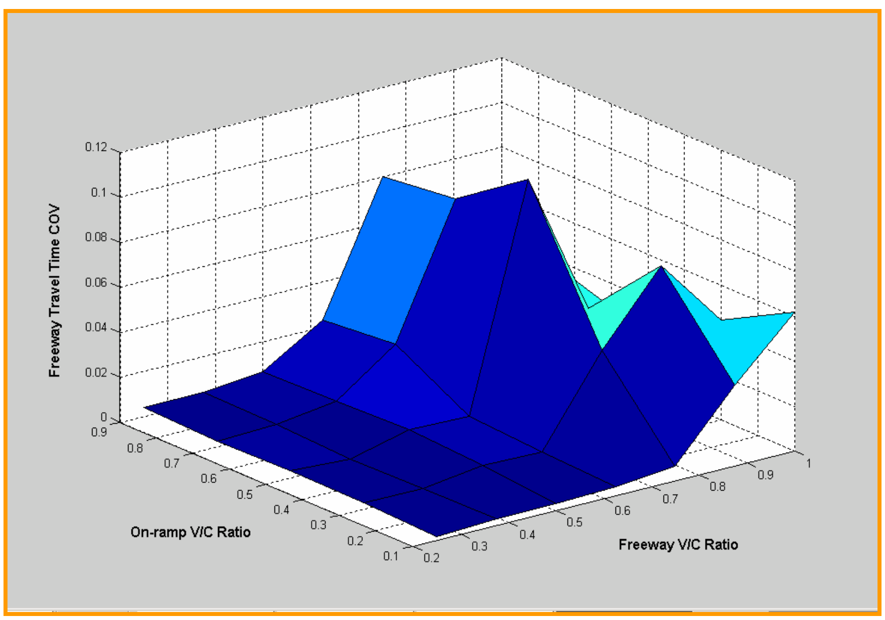

Figure 5: Freeway travel time COV for an on-ramp section having an acceleration lane length of 50m

Note that we have focused above on determining the COV values for a given demand pattern that might exist within some particular time interval of the day. This demand characteristic, however, would typically vary over different intervals of the day, as illustrated in Figure 1. Consequently, it is proposed that the COV be computed for the distribution of travel times between such intervals of the day. The net COV value used for the particular link in our analysis could then be taken as the maximum of these within-intervals and between-intervals COV values. Figure 1 illustrates that typically travel time variability within a time interval is higher than the travel time variability between intervals.

3.3 Derivation of Trip Benefit Functions

As described above, the roadway between any O-D pair, say (p, q), can be composed of several links $l = 1,\ldots, m$. We now derive an expression for the composite COV for the entire route between p and q, given the coefficient of variation $COV_l = \sigma_l/\mu_l$, for each individual link $l = 1, \ldots, m$, where, μ_l is the expected value of the travel time for link l and σ_l is the standard deviation of the travel time for link l. Let us define

$$COV_{\max} \equiv \max_l \{COV_l\} \text{ and } COV_{\min} \equiv \min_l \{COV_l\};$$

Note that since, $\sum_{l=1}^{m} \mu_l^2 = \sum_{l=1}^{m} \left(\frac{\sigma_l^2}{COV_l^2} \right)$, and that

$$\frac{1}{COV_{\min}^2} \sum_{l=1}^{m} \sigma_l^2 \geq \sum_{l=1}^{m} \mu_l^2 \geq \frac{1}{COV_{\max}^2} \sum_{l=1}^{m} \sigma_l^2 , \quad \text{we have}$$

$$COV_{\max} \geq \left\{ \sum_{l=1}^{m} \sigma_l^2 \middle/ \sum_{l=1}^{m} \mu_l^2 \right\}^{1/2} \geq COV_{\min} .$$

Hence, motivated by these bounding inequalities, one composite COV function that we propose is as follows:

$$b_{p-q}^{(1)} \quad = \quad \left\{ \sum_{l=1}^{m} \sigma_l^2 \Big/ \sum_{l=1}^{m} \mu_l^2 \right\}^{1/2} . \tag{2a}$$

As another alternative, suppose that we assume the independence of travel time behavior on the individual links. This assumption might not be completely valid since travel times on adjacent links might typically bear some correlation in practice especially during the formation of queues. Nevertheless, for the sake of deriving benefit factor coefficients that tend to reflect variability in travel times, we might adopt this simplifying assumption. This leads to an alternative composite COV function as:

$$b_{p-q}^{(2)} = \left\{ \sum_{l=1}^{m} \sigma_l{}^2 \right\}^{1/2} \Big/ \sum_{l=1}^{m} \mu_l \, . \tag{2b}$$

Observe from (2a) and (2b) that, in general,

$$b_{p-q}^{(1)} \quad \geq \quad b_{p-q}^{(2)} \; ,$$

i.e., alternative (2a) uniformly yields higher benefit estimates than (2b). In our computations, we will examine both these composite functions and study the sensitivity of the prescribed reader locations to the choice of using b_{p-q} as given by either of these two functions.

4. Solution of the Reader Location Problem

The Reader Location Problem (**RL**) is a linearly constrained mixed-integer zero-one quadratic programming problem [*14 –15*]. Adams and Sherali [*15*] review alternative approaches for this class of problems, and propose a Reformulation-Linearization Technique (RLT) that is further refined in Sherali and Adams [*16-18*]. We employ this RLT approach in concert with semidefinite cuts as developed by Sherali and Fraticelli [*19*] in order to solve Problem **RL**.

The RLT method can be used to transform the mixed-integer quadratic program **RL** into an equivalent zero-one mixed-integer linear programming problem. The emphasis on deriving this linearization is to achieve as tight a linear programming (LP) relaxation as possible. This approach enhances problem solvability by way of providing an equivalent linear representation whose continuous LP relaxation yields a tight upper bound on the problem. The RLT methodology has been shown to be a competitive procedure, in the aforementioned references, to solve problems of the type **RL**.

Accordingly, let us apply the RLT methodology as prescribed below to convert the problem **RL** into an equivalent linearized representation **RL'**. (We assume here that neither (1b) nor (1c) is implied by the other. Whenever this assumption does not hold, the corresponding redundant constraint would then be deleted from the analysis below.)

Step 1: Multiply the constraint (1b) by y_i, and (1- y_i), $\forall\, i \in N$ to yield (3a) and (3b), respectively.

$$\sum_{j \in N} y_j\, y_i \ \le\ R\, y_i \qquad \forall\, i \in N, \tag{3a}$$

$$\sum_{j \in N} y_j\,(1-y_i) \ \le\ R(1-y_i) \quad \forall\, i \in N. \tag{3b}$$

Remark: Note that when (1b) is imposed as an equality constraint (as for example in the absence of (1c)), then we simply need to multiply this equality with each variable y_i, $\forall\, i \in N$, at this step, and also retain this original constraint in the reformulated problem (refer [*16*]).

Step 2: Multiply the constraint (2c) by y_i and (1- y_i), $\forall\, i \in N$, respectively.

$$\sum_{j \in N} C_j\, y_j\, y_i \ \le\ B\, y_i \qquad \forall\, i \in N, \tag{3c}$$

$$\sum_{j \in N} C_j\, y_j (1-y_i) \ \le\ B\,(1-y_i) \qquad \forall\, i \in N. \tag{3d}$$

Step 3: Multiply the constraints $0 \le y_j \le 1$ by y_i and $(1-y_i)$, $\forall\, i \in N$, where $i < j$ to avoid duplicate combinations, for each $j \in N$:

$$0 \ \le\ y_i\, y_j \ \le\ y_i \text{, and} \tag{3e}$$

$$0 \ \le y_j - y_i\, y_j \ \le 1 - y_i, \ \ \forall\, i < j\ ,\ \{\, i\ ,\ j\} \in N. \tag{3f}$$

Step 4: In the system (3), set $y_j^{\,2} = y_j$, $\forall\ \ j \in N$ (since y is supposed to be binary valued), and substitute $w_{ij} = y_i\, y_j$, $\forall\, \{i,\ j\} \in\ N,\ i\ <\ j.$ This yields the following reformulated linear mixed-integer program, which we call **RL'**.

$$\textbf{RL'}: \quad \text{Maximize} \quad \sum_{\substack{(i,j)\in A \\ i<j}} \sum b_{ij} w_{ij} + \sum_{\substack{(i,j)\in A \\ i>j}} \sum b_{ij} w_{ji} \tag{4a}$$

subject to:

$$\sum_{j\in N} y_j - R \le \sum_{\substack{j\in N \\ j>i}} w_{ij} + \sum_{\substack{j\in N \\ j<i}} w_{ji} + y_i(1 - R) \le 0 \quad \forall\, i \in N \tag{4b}$$

$$\sum_{j\in N} C_j y_j - B \le \sum_{\substack{j\in N \\ j>i}} C_j w_{ij} + \sum_{\substack{j\in N \\ j<i}} C_j w_{ji} + y_i(C_i - B) \le 0 \quad \forall\, i \in N \tag{4c}$$

$$y_j - w_{ij} \ge 0 \qquad \forall\, \{i, j\} \in N,\ i < j \tag{4d}$$

$$y_i - w_{ij} \ge 0 \qquad \forall\, \{i, j\} \in N,\ i < j \tag{4e}$$

$$- y_i - y_j + w_{ij} \ge -1 \qquad \forall\, \{i, j\} \in N,\ i < j \tag{4f}$$

$$w_{ij} \ge 0,\ \forall\, \{i, j\} \in N,\ i < j \tag{4g}$$

$$y_j \ \text{binary},\ \forall\, j \in N. \tag{4h}$$

Observe that if we denote $|N| = n$, then Problem **RL'**, has $n(n-1)/2$ continuous variables, n binary variables, and $n(3n+5)/2$ structural constraints in addition to the logical restrictions (4g) and (4h). In **RL'**, we have replaced the products of the original variables, $y_i y_j$, by a new variable, w_{ij}, and this operation has derived a higher dimensional equivalent linear mixed-integer programming representation of **RL**. Note that for any binary y, constraints (4d) - (4g), enforce that $w_{ij} = y_i y_j \ \forall\, \{i, j\} \in N,\ i < j$, and hence the equivalence. Moreover, following the reduced RLT representation guidelines recommended by Sherali *et al*. [*18*], we will also explore the use of an alternative representation **RL"**, in which the left-hand inequalities in (4b) and (4c), as well as the inequalities in (4f), are deleted. In this case, the original constraints (1b) and

(1c) are no longer implied, and are therefore added back into the model. The motivation for this is that the objective function is trying to drive the w_{ij} variables to as high values as possible, and hence, the constraints that restrict this tendency are most relevant. This would reduce the number of structural constraints in **RL"** to $n(n+1)$. Likewise, we will also experiment with using the most basic linearized form of the proposed model, referred to as **RL_{base}**, comprised of (4a), (1b), (1c) and (4d), (4e), (4g) and (4h). (Note that the constraints (4f) can be deleted in this case because of the nature of the objective function, while ensuring that $w_{ij} = y_i y_j$ holds true, $\forall$ $i < j$, for any binary solution y.)

Furthermore, from a continuous relaxation point of view, incorporating certain additional valid inequalities that further tighten this relaxation and help derive stronger (smaller) upper bounds is an avenue worth exploring. With this motivation, we propose the generation of a class of valid inequalities or cutting planes derived using semidefinite programming concepts. These cuts are referred to as *semidefinite cuts* [*19*]. To present the concept underlying these semidefinite cuts, consider the $n \times n$ symmetric matrix $[y y^T]_L$, where $[\cdot]_L$ denotes linearization of $[\cdot]$ under the substitution defined in *Step 4*. Note that $[y y^T]$ is positive semidefinite (PSD) since $\alpha^T [y y^T] \alpha = [(\alpha^T y)^2] \geq 0, \quad \forall$ $\alpha \in R^n$ [*20*]. However, if $(\bar{y}, \bar{w})$ is an optimal solution to the continuous relaxation of **RL'**, (denoted $\overline{\mathbf{RL}}$**'**, say), and if $\bar{y}$ has fractional components so that $\bar{w}_{ij}$ does not necessarily equal $\bar{y}_i \bar{y}_j, \forall\, i < j$, the matrix $[y y^T]_L$ need not be PSD for $(y, w) = (\bar{y}, \bar{w})$. But since feasibility to **RL'** requires $[y y^T]_L$ to equal $[y y^T]$, and therefore be PSD, we can validly impose the constraints

$$\alpha^T [y y^T]_L \alpha \quad \geq \quad 0 \qquad \forall\ \alpha \in R^n,\ \|\alpha\| = 1. \tag{5}$$

Given a fractional solution $(\bar{y}, \bar{w})$ to $\overline{\textbf{RL}}$**'**, we adopt the polynomial-time scheme prescribed by Sherali and Fraticelli [*19*] to generate violated members of (5), in order to further tighten the continuous relaxation of **RL'**. In our implementation, we let this process generate such cuts as long as the difference in the objective function values between two consecutive continuous relaxations is greater than 0.01.

5. Computational Results

For the purpose of illustrating the methodology, we consider the *Interstate*-35 *North* Freeway, also known as the North Panam Freeway, in San Antonio, Texas. This freeway serves the Northeast Corridor and provides access to several industrial outlets in San Antonio. The total length of this freeway in the San Antonio jurisdiction is 20 miles. The route is entirely urban and suburban, and the majority of the adjacent land-use consists of warehouse, light industry, and heavy commercial development. Several new major projects are proposed or are under construction in this corridor. This route is also the southern continuation of the San Antonio-Austin Corridor, and is part of the "NAFTA Superhighway". An 8-mile segment of this freeway, lying between Shin Oak Drive and Pine Street, is chosen as shown in Figure 7. Different O-D paths are identified for possible travel time measurement and the corresponding reader location optimization problem is then solved on the associated graph *G* using the proposed alternative model representations. This segment of the freeway contains 7 on-ramps, 9 off-ramps, and 7 weaving sections and presently has 7 readers located at the different numbered station points shown in Figure 7. In particular, station points # 42, and # 45 are located at overpasses.

Table 1 displays the site-specific costs C_j for the nodes 1,…, 8, and the computed COV values for each of the links along this 8-mile freeway segment, defined for consecutive node pairs, as designated by the rows in the table. The COV values for these basic sections have been derived using the analysis presented in Section 3. Given these COV values, the composite benefit factors, $b_{p-q}^{(1)}$ or $b_{p-q}^{(2)}$, for all the possible $\binom{8}{2} = 28$ O-D pairs (p, q), $p, q \in \{1,\ldots,8\}$, $p < q$, can be computed using the formulae (2a) or (2b), respectively.

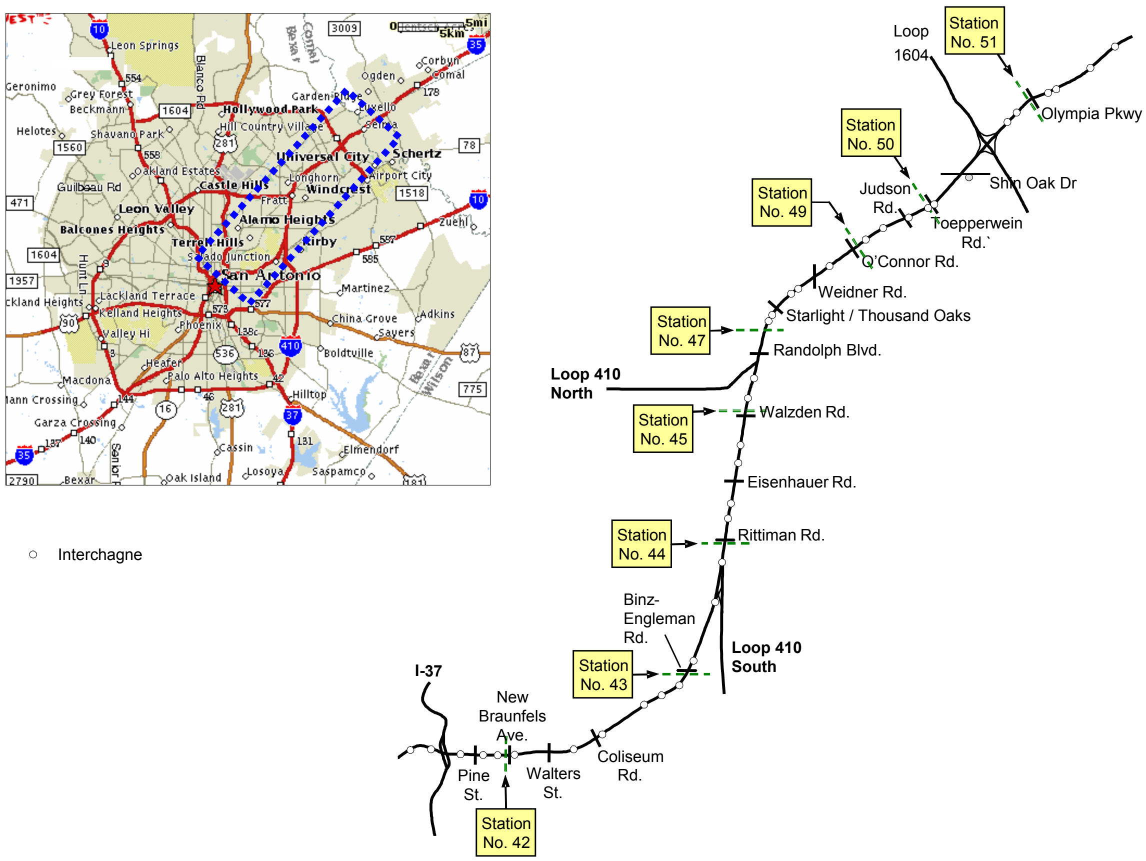

Figure 6: Study corridor

<table>
<tr><th>Node j</th><th>Cost C_j</th><th>Overpass</th><th>AADT[5] (veh/day)</th><th>COV</th></tr>
<tr><td rowspan="2">1</td><td rowspan="2">6.32</td><td rowspan="2">Shin Oak Drive</td><td rowspan="3">107,000</td><td rowspan="3">0.3341</td></tr>
<tr></tr>
<tr><td rowspan="2">2</td><td rowspan="2">9.16</td><td rowspan="2">Judson</td></tr>
<tr><td rowspan="2">107,000</td><td rowspan="2">0.2400</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">7</td><td rowspan="2">O’ Connor</td></tr>
<tr><td rowspan="2">162,000</td><td rowspan="2">1.5500</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">3.63</td><td rowspan="2">George Beach</td></tr>
<tr><td rowspan="2">108,000</td><td rowspan="2">0.1187</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2">9.11</td><td rowspan="2">Loop 410</td></tr>
<tr><td rowspan="2">120,000</td><td rowspan="2">0.5354</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2">1.24</td><td rowspan="2">Walters</td></tr>
<tr><td rowspan="2">131,000</td><td rowspan="2">0.2230</td></tr>
<tr><td rowspan="2">7</td><td rowspan="2">3.68</td><td rowspan="2">New Braunfels</td></tr>
<tr><td rowspan="3">190,000</td><td rowspan="3">0.2768</td></tr>
<tr><td rowspan="2">8</td><td rowspan="2">5.15</td><td rowspan="2">Pine</td></tr>
<tr></tr>
</table>

Table 1: Average traffic data for I-35

We also assume that the total available budget is 30, and that the maximum number of readers available for installation is 5. Note that these parameters ensure that neither constraint (1b) nor (1c) is implied by the other, because the sum of the six cheapest readers is 27.01 < 30, and also the sum of the five most expensive readers exceeds 30. For this problem instance, the different proposed variants of the reader location problem were solved using the composite benefit factor function $b^{(1)}_{p-q}$, yielding the results shown in Table 2. All computations have been performed on a Pentium III, 833 MHz, 128 MB RAM computer, using the mathematical programming software, AMPL – CPLEX 7.0.

[5] Annual Average Daily Traffic

Problem	Constraints included	Number of nodes enumerated	% gap between LP and IP
RL_{base}	(4a), (1b), (1c), (4d), (4e), (4g) and (4h)	13	35.17
RL' (without cuts)	(4a) – (4h)	9	9.22
RL' (with cuts)	(4a)-(4h) and semidefinite cuts	9	9.22
RL"	(4a)-(4h) w/o the LHS inequalities in (4b) and (4c), and w/o (4f), but with (1b) and (1c)	4	19.83

Table 2: Comparison of results for variations of the reader location problem

Table 2 reveals that in terms of the tightness of the relaxations, the enhanced models **RL'** and **RL"** (with or without the cuts) yield significantly reduced gaps between the LP relaxation value at the root node and the final IP value, with **RL'** yielding relatively stronger relaxations, as shown in the last column of the table. This leads to the enumeration of fewer nodes for these problems as compared with the base formulation to solve the problem. In this case, there was no significant impact of the semidefinite cuts since the first order RLT relaxation itself produced an extremely tight linear underlying representation of the problem. The optimal reader locations turned out to be at the nodes {1, 3, 4, 5, 7}. For this sample test case, all problems were solved within 0.02 CPU seconds. The optimal reader locations when $b_{p-q}^{(2)}$ was used were identical to the above case.

5.1 Additional Simulated Test Cases

To further test the proposed approach, we constructed some additional test cases by incorporating the existing station reader locations as potential nodes in the problem.

Hence, the corresponding graph G has a total of 13 nodes. The number of available readers was taken to be 7 to mimic the existing situation. This scenario was examined using three synthetic traffic demand patterns. Also, the number of nodes as reflected by the density (proportion of non-zero coefficients) of the b_{ij} matrix was varied to determine the sensitivity of the reader locations to this feature. In this analysis, we used 1 and 0.75 as the two different densities of the b_{ij} matrix, respectively corresponding to 91 and 68 O-D pairs. The sensitivity of the reader locations was also tested for the two composite benefit factor functions proposed in Section 3, leading to a total of $3\times2\times2 = 12$ problem instances. For all the test cases, we ensured that the parameters R and B did not render either of the constraints (1b) or (1c) redundant. Table 3 displays the results obtained for these test cases.

For most of the test cases, the optimal reader locations turned out to be at the nodes {3, 4, 6, 7, 8, 10, 11}, utilizing all the seven readers. The only exception was for Problem 7 where the readers were optimally located at the nodes {3, 4, 5, 6, 7, 8}, using only 6 readers. (The budget constraint happened to be restrictive in this case.)

Relative to the base formulation, the enhanced model **RL'** (without cuts) significantly reduced the LP-IP gap at the root node of the enumeration tree by 81.44% at an average, leading to a substantial reduction in the number of branch-and-bound nodes enumerated (82.32% at an average). The incorporation of cuts further tightened the model representation, reducing the LP-IP gap and the number of nodes enumerated in relation to the base case by 88.93%, and 88.05%, respectively. On the other hand, although the

Problem (Demand Loading Type, b_{ij} Density, Composite Function)	RL_{base}		RL' (no cuts)		RL' (with cuts)		RL"	
	LP - IP % gap	Nodes Enumerated	LP - IP % gap	Nodes Enumerated	LP - IP % gap	Nodes Enumerated	LP - IP % gap	Nodes Enumerated
1 (1, 1, 1)	39.88	300	7.86	12	3.26	9	19.78	11
2 (1, 0.75, 1)	31.94	76	7.85	28	3.68	18	21.48	12
3 (1, 1, 2)	23.61	48	2.04	14	1.34	10	9.70	2
4 (1, 0.75, 2)	27.86	24	5.36	13	3.33	6	15.97	12
5 (2, 1, 1)	10.62	202	3.13	16	2.67	9	16.87	5
6 (2, 0.75, 1)	29.55	86	8.13	29	3.76	24	20.92	15
7 (2, 1, 2)	22.26	34	2.42	15	1.76	14	9.31	2
8 (2, 0.75, 2)	22.65	26	3.10	11	2.01	6	14.08	6
9 (3, 1, 1)	8.68	229	0	11	0	6	0	5
10 (3, 0.75, 1)	25.98	101	5.62	22	3.90	15	18.61	15
11 (3, 1, 2)	20.10	42	3.37	20	2.76	15	8.01	2
12 (3, 0.75, 2)	22.09	20	4.06	19	3.16	10	15.10	9
Averages	**23.76**	**99**	**4.41**	**17.5**	**2.63**	**11.83**	**14.15**	**8**

Table 3: Results for the additional simulated test cases

simpler enhancement **RL"** reduced the LP-IP gap in relation to the base case to a lower extent (by 40.44%), it resulted in a somewhat greater (91.91%) reduction in the number of nodes enumerated. (Evidently, the CPLEX-MIP software was able to manipulate the relatively simpler structure of **RL"** more efficiently.) Furthermore, in general, the relative gaps between the LP and the IP solution values were somewhat greater for the lower density graphs, usually lower for the type 2 benefit factor, and also varied with the demand pattern. All the test problems were solved in a fraction of a second (within 0.02 seconds of CPU time).

5.2 Larger Random Test Cases

To assess the performance of the proposed model and algorithms on relatively larger problem instances, we generated four additional test cases. For these cases, the COV values for each of the links were randomly generated, and then the corresponding benefit factor coefficients were computed as discussed in Section 3. From the results shown in Table 4, it can be observed that the tight linear programming representation provided by model **RL'** reduces the computational effort considerably. Moreover, the predominant effect of the semidefinite cuts can be easily verified in the case of the larger sized problems where the formulation **RL'** using these cuts uniformly resulted in the smallest LP-IP gap (the percentage difference between the optimal value for the linear programming relaxation at node zero, and the actual integer programming optimal value), the fewest number of nodes enumerated, and the least CPU time over all the test cases. Observe that for this set of larger problems, the simpler enhancement **RL"** was not as effective as the formulation **RL'**, although yet better than the base model. Hence, overall, we recommend the enhanced formulation **RL'** using semidefinite cuts for implementation.

# of Nodes		100	200	300	400	Averages
RL_{base}	**LP-IP % gap**	45.06	34.22	42.74	55.88	**44.475**
	Nodes enumerated	618	988	1277	2327	**1302.5**
	CPU time (s)	1.60	12.8	102.3	492.02	**152.18**
RL' (without cuts)	**LP-IP % gap**	15.87	7.78	12.66	19.89	**14.05**
	Nodes Enumerated	176	312	744	1230	**615.5**
	CPU time (s)	0.14	6.34	76.40	386.40	**117.32**
RL' (with cuts)	**LP-IP % gap**	2.81	4.64	12.86	14.71	**8.755**
	Nodes Enumerated	46	183	577	866	**418**
	CPU time (s)	0.04	4.21	70.00	355.23	**107.37**
RL''	**LP-IP % gap**	35.97	37.27	45.32	48.88	**41.86**
	Nodes Enumerated	277	533	1037	1944	**947.75**
	CPU time (s)	0.20	10.02	97.78	421.0	**132.25**

Table 4: Comparisons of results for larger random test cases

6. Summary and Conclusions

In this paper, we have addressed the issue of optimally locating AVI tag readers to capture as much travel time variability information as possible. An optimization model was developed for this purpose, using an objective function that suitably measures the benefits accruing from the usage of the established readers. To derive the corresponding sections, namely, for on-ramps, off-ramps, weaving sections, and bottleneck sections, under a wide range of traffic conditions. The expected travel time, and the corresponding coefficients of variation were computed for each of the scenarios examined by the simulation model. These simulation results (recorded in [*13*]) are generic in nature upon normalizing the flow by dividing by the roadway capacity, and thus can potentially be used for any given freeway.

Having performed this analysis, we designed a mechanism for deriving the benefit factors for the different O-D links that might be comprised of several such basic sections, in order to compose the objective function of the reader location model. The resulting quadratic 0-1 integer program was then enhanced using the Reformulation-Linearization Technique to derive a tight equivalent linear mixed-integer programming representation. This was further augmented with semidefinite cuts, and subsequently solved using a branch-and-bound methodology implemented within CPLEX. Computational results using simulated test cases based on a section of I-35 North in San Antonio, Texas, as well as larger randomly generated synthetic test problems, revealed that the first order RLT representation **RL'** produced bounds that were relatively close to optimality, and that the generation of the semidefinite cuts improved this representation further, enabling the efficient solution of realistically sized problems.

In conclusion, it is worth noting that physically establishing AVI tag readers is a costly venture, since each reader costs $15,000/lane to install when an overhead pass already exists, and $70,000/lane if an overhead structure needs to be erected along with the reader itself. With such large monetary investments being involved, it is imperative to identify appropriate, beneficial locations for these readers. We have shown that optimization techniques can be gainfully employed to determine effective allocations of such types of scarce physical and monetary resources. The methodology prescribed in this paper to solve such a reader location problem provides an impetus for the effective deployment of this, and other similar, surveillance technologies in Advanced Traffic Information Systems.

7. Recommendations for Further Research

The proposed formulation and solution approach serve as a first step to solving the problem of locating surveillance equipment. As is the case with any research effort further research is required in a number of areas, as follows:

a. Collect field data and validate the coefficient of variation factors that were developed using the INTEGRATION simulation model.
b. Conduct further research to investigate including other considerations in the formulation of the objective function. Potential factors to consider may include the number of O-D pairs that pass a specific roadway section.
c. Apply the model to some realistic networks in order to investigate the applicability of the proposed approach, and study the sensitivity of results to different input parameters, including cost factors and budget constraints.
d. Investigate the potential use of alternative solution methods, including the use of a Genetic Algorithm solution procedure.
e. Develop an automated decision support system that would provide reader locations for generic transportation network specifications.

Bibliography

1. Dion F., and H. Rakha, "Estimating Spatial Travel Times using Automatic Vehicle Identification Data," Paper submitted to *Transportation Research*: *Part B*.

2. Kuehne, R. D., J. Palen, C. Gardner, and S. Ritchie, "Loop-based Travel Time Measurement," *Proceedings of the International Conference on Applications of Advanced Technologies in Transportation Engineering*, ASCE, 167-174, Reston, VA, 1998.
3. Yang, Q., H. N. Koutsopoulos, and M. E. Ben-Akiva, "Simulation Laboratory for Evaluating Dynamic Traffic Management Systems," *Transportation Research Record*, 1710, 122-130, 2000.
4. Bergan, A. T., L. Henion, M. Krukar, and B. Taylor, "Electronic License Plate Technology: Automatic Vehicle Location and Identification," *Canadian Journal of Civil Engineering*, 15(6), 1035-1042, December, 1988.
5. Sisiopiku, V. P., N. M. Rouphail, and A. Santiago, "Analysis of Correlation between Arterial Travel Time and Detector Data from Simulation and Field Studies," *Transportation Research Record*, 1457, 166-173, December, 1994.
6. Coifman, B., "Vehicle Re-identification and Travel Time Measurement in Real-Time on Freeways using Existing Loop Detector Infrastructure," *Transportation Research Record*, 1643, 181-191, November, 1998.
7. Thomas, G. B., "Optimal Detector Location on Arterial Streets for Advanced Traveler Information Systems," Doctoral Dissertation, Arizona State University, Tempe, AZ, December, 1998.
8. Thomas, G. B., "The Relationship Between Detector Location and Travel Characteristics on Arterial Streets," *ITE Journal - Institute of Transportation Engineers*, 69(10), 36-42, October, 1999.

9. Van Aerde M., B. Hellinga, M. Baker, and H. Rakha, "INTEGRATION: An Overview of Current Simulation Features," *Transportation Research Board 75th Annual Meeting*, Washington DC, January, 1996.

10. Van Aerde, M., and H. Rakha, "Comparison of Simulation Modules of TRANSYT and INTEGRATION Models", *Transportation Research Record*, 1566, 1-7, 1996.

11. Rakha H., and K. Ahn, "The INTEGRATION Framework for Estimating Mobile Source Emissions," Submitted for publication in the *Journal of Advanced Transportation*, 2002.

12. M. Van Aerde and Associates. *INTEGRATION Release 2.30 for Windows, User's Guides*, M. Van Aerde and Associates, Kingston, Ontario, Canada, 2002.

13. Desai, J., "A Discrete Optimization Approach to Solve the Reader Location Problem for Estimating Travel Times", M. S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2002.

14. Rader, D., "Quadratic 0-1 Knapsack Problem," *INFORMS Philadelphia Conference*, November, 1999.

15. Adams, W. P., and H. D. Sherali, "A Tight Linearization and an Algorithm for Zero-One Quadratic Programming Problems", *Management Science*, 32(10), 1274-1290, 1986.

16. Sherali, H. D., and W. P. Adams, "A Hierarchy of Relaxations Between the Continuous and Convex Hull Representations for Zero-One Programming Problems", *SIAM Journal on Discrete Mathematics*, 3(3), 411- 430, 1990.

17. Sherali, H. D., and W. P. Adams, "A Hierarchy of Relaxations and Convex Hull Characterizations for Mixed-Integer Zero-One Programming Problems", *Discrete Applied Mathematics*, 52, 83-106, 1994.

18. Sherali, H. D., J. C. Smith, and W. P. Adams, "Reduced First-level Representations Via the Reformulation-Linearization Technique: Results, Counter-Examples, and Computations", *Discrete Applied Mathematics*, 101, 247-267, 2000.

19. Sherali, H. D., and B. M. P. Fraticelli, "Enhancing RLT Relaxations Via a New Class of Semidefinite Cuts", *Journal of Global Optimization*, 22, 233-261, 2002.

20. Bazaraa, M. S., H. D. Sherali, and C. M. Shetty, "*Nonlinear Programming: Theory and Algorithms*", John Wiley and Sons, Inc., 2nd edition, New York, N.Y., 1993.