Optimal Sensor Placement on Highway Networks: A Traffic Dynamics Based Approach

Sebastian A. Nugroho
Department of Electrical Engineering and Computer Science
University of Michigan
1301 Beal Ave., Ann Arbor, MI 48109
Email: snugroho@umich.edu

Suyash C. Vishnoi, Corresponding Author
Department of Civil, Architectural, and Environmental Engineering
The University of Texas at Austin, Texas
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Email: scvishnoi@utexas.edu

Ahmad F. Taha
Department of Civil and Environmental Engineering
Vanderbilt University
2301 Vanderbilt Pl, Nashville TN 37235
Email: ahmad.taha@vanderbilt.edu

Christian G. Claudel
Department of Civil, Architectural, and Environmental Engineering
The University of Texas at Austin, Texas
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Email: christian.claudel@utexas.edu

Word Count: 5513 words + 1 table(s) (250 words per table) = 5763 words

Submission Date: August 1, 2021

ABSTRACT

This paper investigates the practical engineering problem of traffic sensors placement on stretched highways with ramps. Since it is virtually impossible to install bulky traffic sensors on each highway segment, it is crucial to find placements that result in optimized network-wide, traffic observability. Consequently, this results in accurate traffic density estimates on segments where sensors are *not* installed. The substantial contribution of this paper is the utilization of control-theoretic observability analysis—jointly with integer programming—to determine traffic sensor locations based on the nonlinear dynamics and parameters of traffic networks. In particular, the celebrated asymmetric cell transmission model is used to guide the placement strategy jointly with observability analysis of nonlinear dynamic systems through Gramians. Thorough numerical case studies are presented to corroborate the proposed theoretical methods and various computational research questions are posed and addressed. The presented approach can also be extended to other models of traffic dynamics.

Keywords: Traffic sensor placement, highway traffic networks, asymmetric cell transmission model, observability Gramian, convex integer programming.

1 INTRODUCTION

With the development of intelligent transportation systems technologies, numerous sensing methods for traffic data collection have become popular (*1, 2*). Fixed sensors such as induction loops and magnetometers allow network operators to obtain high-quality measurements of vehicle density besides other information such as vehicle speeds and flow (*1*). Sensors, however, are generally expensive to install and maintain which makes them infeasible for installation throughout the network and covering all segments. This poses a problem for traffic management operations and controls that require knowledge of the traffic state on all segments. Such operations include traffic control tasks such as ramp metering, and variable speed limit control. While it may only be feasible to collect data from a fixed number of segments in the network, it is still possible to obtain accurate estimates of the state of traffic on all the segments if the sensors are strategically placed.

The problem of placing traffic sensors on highway networks has been divided into two main categories in the literature, one involving estimation and the other involving observability. Under the former category, the objective is to determine the optimal placement of sensors that minimizes the estimation error for unmeasured quantities such as travel time, OD matrix, link flows. Under observability, the literature is further divided based on the type of observability that is considered, full or partial observability. A fully observable system is one in which all the states are observable given the available measurements from the sensors. Different observability problems considered in this category are link flow observability, route flow observability and OD flow observailibty. Partial observability implies that not all states are observable given the set of available sensor measurements. This concept is attractive in cases where the number of sensors required to achieve full observability is too large. Interested readers can refer to (*2–7*) and references therein for a literature review of the aforementioned categories of sensor placement problem. In the current work, we focus on determining the optimal sensor placement to achieve full observability of the system while maintaining a balance between the number of sensors and the degree of observability of the system. Note that the degree of observability is a quantitative measure of the quality of state estimates that can be obtained by using a certain sensor placement configuration and is not related to the idea of partial observability. Unlike the aforementioned studies that utilize the relationship between the various flows in the network such as link and path flows, this study uses a traffic dynamics model to determine the relationship between various state variables. Also, the states considered here are traffic densities instead of flows.

Some studies that are closer to the work presented in this paper are (*8, 9*) which also consider a traffic dynamics model and applies a control theoretic approach to determine optimal sensor locations on highway segments. These studies linearize the nonlinear traffic dynamics around a steady-state traffic flow. The major drawback of such approaches is that the linearized dynamics are valid only around the specific states. Furthermore, these studies consider the Greenshield's fundamental diagram (*10*) which is inferior in modeling traffic flows when compared with the triangular and trapezoidal fundamental diagrams which are often considered in the implementation of the *cell transmission model (CTM) (11, 12).* (*1*) does deal with optimal sensor placement and density reconstruction while considering the CTM model but to simplify things it still linearizes the model and thus has the same drawback as above.

This paper investigates the sensor placement problem for highway networks with ramps having nonlinear traffic dynamics. The traffic model is built using the *asymmetric CTM* (ACTM) (*13*) (a close variant of the CTM) and the triangular fundamental diagram, and thus is also an improvement over (*8, 9, 14*) in terms of modeling traffic dynamics, to remove the dependence of optimal

sensor placements on assumptions about traffic states. The traffic sensor placement is addressed via observability analysis based on the nonlinear traffic model, which is different from the ones used in the aforementioned papers.

In the context of sensor placement for nonlinear systems, four different approaches have been proposed recently. The first approach leverages empirical observability Gramian to find the best set of sensors that maximizes some metrics on the Gramian matrix—see (*15–17*). Another approach is proposed in (*18*) where the observability Gramian around a certain initial state is constructed through a moving horizon estimation (MHE) framework. The third approach, developed in (*19*), introduces a randomized algorithm for dealing with the sensor placement problem and accordingly, theoretical bounds for eigenvalue and condition number of observability Gramian are proposed. The last and most recent approach is established in (*20*) where the authors make use of the numerous observer designs for some classes of nonlinear systems posed as semidefinite programs (SDP).

In this paper, based on the nonlinear traffic model developed using ACTM, we formulate the sensor placement problem using traffic observability analysis.The resulting problem—categorized as convex IP—is solved via an integer branch-and-bound (BnB) algorithm and therefore, optimal traffic sensor locations can be obtained. Notice that the use of nonlinear traffic model and observability analysis result in traffic sensor locations that are valid for various traffic conditions.

The paper is organized as follows. In Section 2, we discuss the modeling of stretched highway with ramps using ACTM, which result in a nonlinear state-space form. Next, Section 3 discusses our strategy for addressing traffic sensor placement based on system's observability. The proposed approach is extensively tested in Section 4 for solving the traffic sensor placement problem via numerous case studies. Finally, Section 5 concludes the paper.

Paper's Notation: Let R, R^n, and $R^{p\times q}$ denote the set of real numbers, real-valued row vectors with size of n, and p-by-q real matrices respectively. For any vector $z \in R^n$, $\|z\|_2$ denotes its Euclidean norm, i.e. $\|z\|_2 = \sqrt{z^\top z}$, where $z^\top$ is the transpose of z. The symbol $\otimes$ denotes the Kronecker product where $\det(A)$ and $\text{trace}(A)$ return the determinant and trace of matrix A. For any set M, $|M|$ denotes its cardinality.

2 NONLINEAR DISCRETE-TIME MODELING OF TRAFFIC NETWORKS WITH RAMPS

This section presents the discrete-time modeling of traffic dynamics on a stretched highway with arbitrary number and location of ramps. To that end, here we utilize the *Lighthill-Whitman-Richards* (LWR) Model (*21, 22*) for traffic flow. In this paper, the relationship between traffic density and flux is given by the triangular-shaped fundamental diagram which has been extensively used in the literature (*11*) and is given as

$$q\left(\rho(t,d)\right) = \begin{cases} v_f \rho(t,d), & \text{if } 0 \le \rho(t,d) \le \rho_c \\ w_c\left(\rho_m - \rho(t,d)\right), & \text{if } \rho_c \le \rho(t,d) \le \rho_m. \end{cases} \tag{1}$$

where t and d denote the time and distance; $\rho(t,d)$ denotes the traffic density, $q(t,d)$ denotes the traffic flux, v_f denotes the free-flow speed, w_c denotes the congestion wave speed and ρ_m denotes the maximum density. To represent the traffic dynamics as a series of difference, state-space equations we discretize the LWR Model with respect to both space and time (this is also referred to as the Godunov discretization). Thus, the highway of length L is divided into segments (cells) of equal length l and the time horizon is divided into steps of duration T. The segments

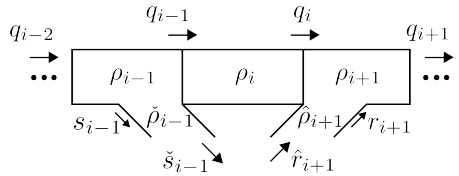


FIGURE 1: **Consecutive segments of the highway with on-ramp and off-ramp.**

form both the highway and the attached ramps. Throughout the paper, the segments forming the highway are referred to as mainline segments. We assume that the highway is split into N mainline segments.

To ensure computational stability, the Courant-Friedrichs-Lewy condition (CFL) given as $v_f T l^{-1} \leq 1$ has to be satisfied. Since each segment is of the same length l, then we have $\rho(t,d) = \rho(kT,l)$, where $k \in N$ represents the discrete-time index. For simplicity of notation, from here on $\rho(kT,l)$ is simply written as $\rho[k]$.

As mentioned earlier, traffic is modeled using the ACTM which is originally given in (*11, 13, 23*). The ACTM is a variant of the CTM that departs from the CTM in its treatment of asymmetric merge junctions such as the on-ramp-highway junctions. Unlike the CTM, it assumes separate allocations of the available space on the highway for traffic from each merging branch, which allows for comparatively simple flow conservation equations at those merges than the original CTM. A variation to the original ACTM is introduced in the modeling of the ramps which here are treated as normal segments rather than point queues as in the original approach. The traffic model used in this work is close to the one used in (*24*). Interested readers can refer to the same for mathematical derivations of the traffic model.

The discrete-time flow conservation equation for a mainline segment with both on- and off-ramp can then be written as

$$\rho_i[k+1] = \rho_i[k] + \frac{T}{l}\big(q_{i-1}[k] + r_i[k] - q_i[k] - s_i[k]\big) \tag{2}$$

where ρ_i is the density of Segment i, q_i is the traffic flow from Segment i to Segment $i+1$, r_i is the flow from the on-ramp segment attached to Segment i into Segment i, and s_i is the flow from Segment i into the off-ramp attached to it. Equations for segments with only one or no ramp can be written by removing the respective flow terms. The flows and conservation equations for the ramps can also be written similar to the mainline segments. An illustration for the highway model is given in Fig. 1. In the figure $\hat{\rho}$ and $\check{\rho}$ denote the density of on-ramps and off-ramps respectively, and $\check{s}$ and $\hat{r}$ denote the flow leaving the off-ramp and the flow entering the on-ramp respectively.

The flows are calculated using the demand and supply formulae which can be found in (*24*). Demand of a segment refers to the maximum traffic flux that can leave the segment if there is infinite capacity ahead and supply refers to the maximum traffic flux that can enter the segment if there is infinite traffic wanting to enter. In this paper we assume that the upstream demand at Segment 1 and downstream supply at Segment N are known, denoted by f_{in} and f_{out} respectively. Similarly, we assume that all the on-ramp demands $\hat{f}$ and off-ramp supplies $\check{f}$ are also known. Let the set of mainline segments, on-ramps and off-ramps be denoted as Ω, Ω_I and Ω_O respectively with $N = |\Omega|$, $N_I = |\Omega_I|$ and $N_O = |\Omega_O|$. Then the input vector can be defined as $u[k] = [f_{\text{in}}[k]\ f_{\text{out}}[k]\ \ldots\ \hat{f}_j[k]\ \ldots\ \check{f}_l[k]\ \ldots]^\top \in R^{2+N_I+N_O}$, $j \in \Omega_I$ and $l \in \Omega_O$ and the state vector

can be defined as $x[k] := [\rho_i[k] \ \dots \ \hat{\rho}_j[k] \ \dots \ \check{\rho}_l[k] \ \dots]^\top \in R^{N+N_I+N_O}$, $i \in \Omega$, $j \in \Omega_I$ and $l \in \Omega_O$.

The evolution of traffic density can be written in a compact state-space form as follows

$$x[k+1] = Ax[k] + Gf(x,u,k) + B_{\mathrm{u}}u[k], \tag{3}$$

where $A \in R^{n\times n}$ for $n := N + N_I + N_O$ represents the linear dynamics of the system, $B_{\mathrm{u}} \in R^{n\times m}$ for $m := 2 + N_I + N_O$ represents the way external inputs affecting the system, $f : R^n \times R^m \to R^g$ is a vector valued function representing nonlinearities in (2), and $G \in R^{n\times g}$ is a matrix representing the distribution of nonlinearities.

The above result is important as it allows us to perform optimal sensor placement for the traffic dynamical system defined by (1)–(2), as well as to develop a robust state observer. This also allows for other control-theoretic studies to use this state-space model for other traffic engineering applications. The next section presents our strategy to address the optimal sensor placement problem for the nonlinear traffic dynamical system derived above.

3 OBSERVABILITY-BASED SENSOR PLACEMENT

In this section we discuss our approach for addressing the traffic sensor placement problem. The traffic dynamics with measurements can be expressed as

$$x[k+1] = Ax[k] + Gf(x,u,k) + B_{\mathrm{u}}u[k] \tag{4a}$$

$$y[k] = \Gamma Cx[k]. \tag{4b}$$

In the above model, we introduce $y \in R^p$ to represent the vector of measurements, which corresponds to all of the highway segments equipped with sensors measuring the density. The matrix $C \in R^{p\times n}$ in (4) is useful to determine the placement of traffic sensors, whereas $\Gamma := \mathrm{Diag}(\gamma)$ with $\gamma \in \{0,1\}^p$ represents the selection of sensors—that is, $\gamma_i = 1$ if the i-th highway segment is measured (this consequently leads to nonzero row i in C) and $\gamma_i = 0$ otherwise.

Having described system (4) with sensor placement, we now state the paper's major computational objective: from a given possible set of traffic sensors $\mathscr{G}_\gamma$ such that $\gamma \in \mathscr{G}_\gamma$, find the best (or *optimal*) sensor configuration γ^* such that system (4) is observable, i.e., the system's initial state $x_0 := x[0]$ can be uniquely determined from a finite set of measurements.

In this paper, we opt to formulate the traffic sensor placement using the concept of observability through MHE framework developed in (*18*). The framework utilizes a series of past measurement data to estimate the initial state at each time window. The reasons of pursuing this approach are two-fold. First, as argued in (*18*), this approach is considerably more scalable than using empirical observability Gramian and second, we experience numerical issues in applying the method from (*20*).

With that in mind, we consider the first N observation window (or N discrete measurements) of system (4) expressed in the equation below

$$h(\gamma, x_0) := \tilde{y} - g(\gamma, x_0). \tag{5}$$

In (5), the mapping $g : N^p \times R^n \to R^{Np}$ is given as

$$g^{\top}(\gamma,x_0) := \Big[(\Gamma C x[0])^{\top} \; (\Gamma C x[1])^{\top} \ldots \; (\Gamma C x[N-1])^{\top} \Big]. \tag{6}$$

Note that $g(\cdot)$ defined above is indeed a function of x_0 since for any k such that $0 < k \leq N-1$ then we have

$$x[k] = A^k x[0] + \sum_{j=0}^{k-1} A^{k-1-j} \left(G f(x,u,j) + B_{\mathrm{u}} u[j] \right). \tag{7}$$

The term $\tilde{y} \in R^{Np}$ in (5) denotes the stacked N measurements data constructed as

$$\tilde{y}^{\top} := \left[\tilde{y}[0]^{\top} \; \tilde{y}[1]^{\top} \; \ldots \; \tilde{y}[N-1]^{\top} \right]. \tag{8}$$

Obviously, for every initial condition of the system x_0, it holds that $h(\gamma,x_0) = 0$ such that $\tilde{y} = g(\gamma,x_0)$. Hence, for a fixed γ, the mapping $g(\cdot)$ maps the initial state into the N measurements output. The observability of system (4) with respect to $g(\cdot)$ is formally defined as follows (*25*).

Definition 1: The system (4) with a prescribed γ is uniformly observable if, for all admissible inputs, there exists a finite $N > 0$ such that the mapping $g(\gamma,x_0)$ defined in (6) is injective (one-to-one) with respect to x_0. Note that, from this Definition, if $g(\gamma,x_0)$ is injective with respect to x_0, then x_0 can be uniquely determined from the set of measurements $\tilde{y}$. A sufficient condition for $g(\cdot)$ to be injective is that the Jacobian of $g(\cdot)$ around x_0, denoted by $J_w(\cdot)$, is full rank (*25*). It can also be shown that if system (4) has no nonlinear counterparts, then the Jacobian matrix $J_w(\cdot)$ reduces to the N-step observability matrix for the linear dynamics.

In this work, we use the concept of observability Gramian to quantify the system's osbervability for a given set of sensors. The observability Gramian for the nonlinear system (4) with respect to γ around x_0 can be constructed as

$$W_o(\gamma,x_0) := J_w^{\top}(\gamma,x_0) J_w(\gamma,x_0), \tag{9}$$

where $W_o(\cdot) \in R^{n \times n}$ and $J_w(\cdot) \in R^{Np \times n}$ is given as (*18*)

$$J_w(\gamma,x_0) := \left[I \otimes \Gamma C \right] \times \begin{bmatrix} I \\ \dfrac{\partial x[1]}{\partial x[0]} \\ \vdots \\ \dfrac{\partial x[N-1]}{\partial x[0]} \end{bmatrix}, \tag{10}$$

where for system (4), by applying chain rule, for $0 < k \leq N-1$ the k-th partial derivative can be obtained from

$$\frac{\partial x[k]}{\partial x[0]} = \prod_{j=0}^{k-1} A + G \frac{\partial f}{\partial x}(x,u,j), \tag{11}$$

where the term $\frac{\partial f}{\partial x}(\cdot) \in R^{g \times n}$ denotes the Jacobian of $f(\cdot)$ with respect to x evaluated at discrete time j. This requires $x[j]$ and $u[j]$ to be known, which can be obtained from simulating (4) for up

to j-th discrete time. In the context of sensor placement problem, the objective is to find the best γ satisfying $\gamma \in \mathscr{G}_\gamma$ which maximizes the observability of system (4) based on the observability Gramian (9).

Some measures that quantify observability based on Gramian matrix include the rank, smallest eigenvalue, condition number, trace, and determinant—see reference (*16*). In this work, we maximize the determinant and trace of the observability Gramian. The trace of observability Gramian measures the average observability of the system in all directions in state space. Therefore, a larger value of the trace indicates an increase in overall observability (*17*). Similar to the trace, determinant of observability Gramian matrix also measures the system's observability in every directions in state space. Nonetheless, as pointed out in (*15*), determinant has the ability to capture information from all elements of the Gramian matrix as well as taking into account information redundancy for the case when multiple sensors are considered. It is also suggested in (*16*) that the determinant is a better measure for quantifying observability than the trace, since trace tends to overlook (near) zero eigenvalues. Despite their differences, both metrics are popular in sensor placement literature through empirical observability Gramian; see (*15–17*). The resulting sensor placement problem is given as follows

$$(\mathbf{P1}) \quad \kappa = \underset{\gamma}{\text{minimize}} \quad \begin{cases} -\det\left(W_o(\gamma, \hat{x}_0)\right), \\ -\text{trace}\left(W_o(\gamma, \hat{x}_0)\right), \end{cases} \tag{12a}$$

$$\text{subject to} \quad \gamma \in \mathscr{G}_\gamma,\ \gamma \in \{0,1\}^p. \tag{12b}$$

Notice that **P1** only takes the integer variables γ as the optimization variable while $\hat{x}_0$ is fixed. Since the constraint and objective function are convex, then **P1** is categorized as a convex IP, which can be solved optimally via a BnB algorithm. After the observability of the system has been determined from the solutions of **P1**, then the measurement equation (4) can be reformulate into $y[k] = \tilde{C}x[k]$ where $\tilde{C} \in R^{n \times \tilde{p}}$ is the reduced state-to-output matrix that corresponds to the nonzero rows of $\Gamma^* C$, where γ^* is the optimal solution of **P1**. The traffic density estimation is then performed based on $y[k] \in R^{\tilde{p}}$.

Another metric of interest to us is the relative error ζ which represents the difference between actual initial state x_0 and the *computed* initial state $\tilde{x}_0$, which is obtained from solving the following nonlinear least-square problem (*18*)

$$(\mathbf{P3}) \quad \underset{\tilde{x}_0}{\text{minimize}} \ \|\tilde{y}_\gamma - g_\gamma(\tilde{x}_0)\|_2^2 \tag{13a}$$

$$\text{subject to} \quad 0 \leq \tilde{x}_0 \leq 1 \times \rho_m. \tag{13b}$$

In (13a), $\tilde{y}_\gamma$ is obtained from simulating the undisturbed traffic dynamics (4) with a prescribed set of sensors γ from $k = 0$ to $k = N$ and initial state x_0. Likewise, the nonlinear mapping $g_\gamma(\cdot)$ is constructed as in (6) with the same sensor combination γ.

After $\tilde{x}_0$ is obtained as the solution of **P3**, ζ is computed as

$$\zeta := \frac{\|\tilde{x}_0 - x_0\|_2}{\|x_0\|_2}. \tag{14}$$

Note that ζ appraises the quality of sensor placement, as smaller ζ suggests that the given sensor

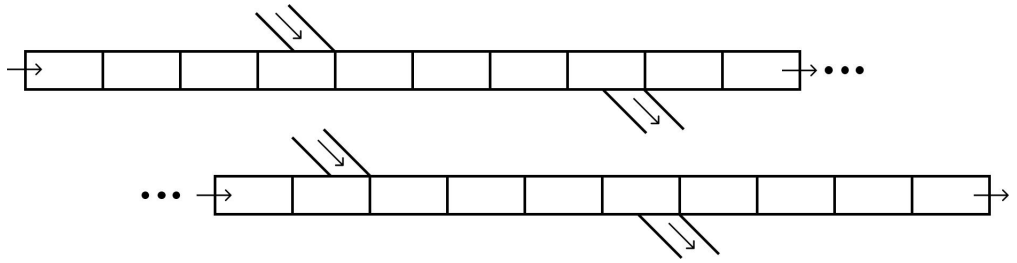


FIGURE 2: **Structure of the highway used for case-study; consists of 20 mainline segments of length 250 m each with 2 on-ramps and 2 off-ramps alternatively placed on every fourth segment. Arrows indicate the direction of traffic.**

configuration is capable to provide a more accurate estimate of the initial state. While this is reasonable if **P3** returns an *optimal* solution, as **P3** is a nonconvex optimization problem, it is highly unlikely that optimal solutions can be obtained.

4 CASE STUDY: RESULTS AND ANALYSIS

This section demonstrates the proposed approach for determining traffic sensors location. Specifically, in this numerical study we attempt to answer the following questions:

- *Q1*: How does the observation window length relate to traffic observability and initial state estimation?
- *Q2*: How computationally efficient are the determinant and trace observability metrics in terms of solving the sensor placement problem?
- *Q3*: How do actual and presumed initial states affect the resulting sensors' location? Are sensor placements robust to changes in initial states?
- *Q4*: How reliable is the state estimation when using sensor locations obtained from utilizing determinant and trace observability metrics?
- *Q5*: How does the theory-driven placement of sensors compare to randomized and uniform sensor placement strategies?

All simulations are performed using MATLAB R2020a running on 64-bit Windows 10 with 2.2GHz Intel$^{\text{R}}$ Core$^{\text{TM}}$ i7-8750H CPU and 16 GB of RAM. Throughout the section, all highways are configured with $v_f = 28.8889$ m/s (65 mph), $w_c = 6.6667$ m/s (15 mph), $\rho_c = 0.0249$ vehicles/m (40 vehicles/mile), $\rho_m = 0.1333$ vehicles/m, and $l = 250$ m. The discrete-time step is chosen to be $T = 1$ sec. The highway, referred to the rest of the section as *Highway A*, stretches for approximately 3.1 miles with 20 segments on the mainline, 2 on-ramps, and 2 off-ramps which are placed alternatively on every fourth segment making a total of 24 segments. The structure of the highway is depicted in Fig. 2.

4.1 Observability Analysis for Traffic Sensor Placement

Herein, we perform a numerical analysis on sensor placement approach through traffic network's observability discussed in Section 3. The objective of traffic sensor placement problem translates to finding the set of r highway segments that must be equipped with traffic sensors such that the entire highway traffic is observable, which is carried out by solving **P1**. Realize that **P1** is classified as a convex *integer programming* (IP) since $\hat{x}_0$, the presumed initial state, is fixed. Here, **P1** is solved using a greedy approach presented in (*26*) due to numerical issues presented by available exact

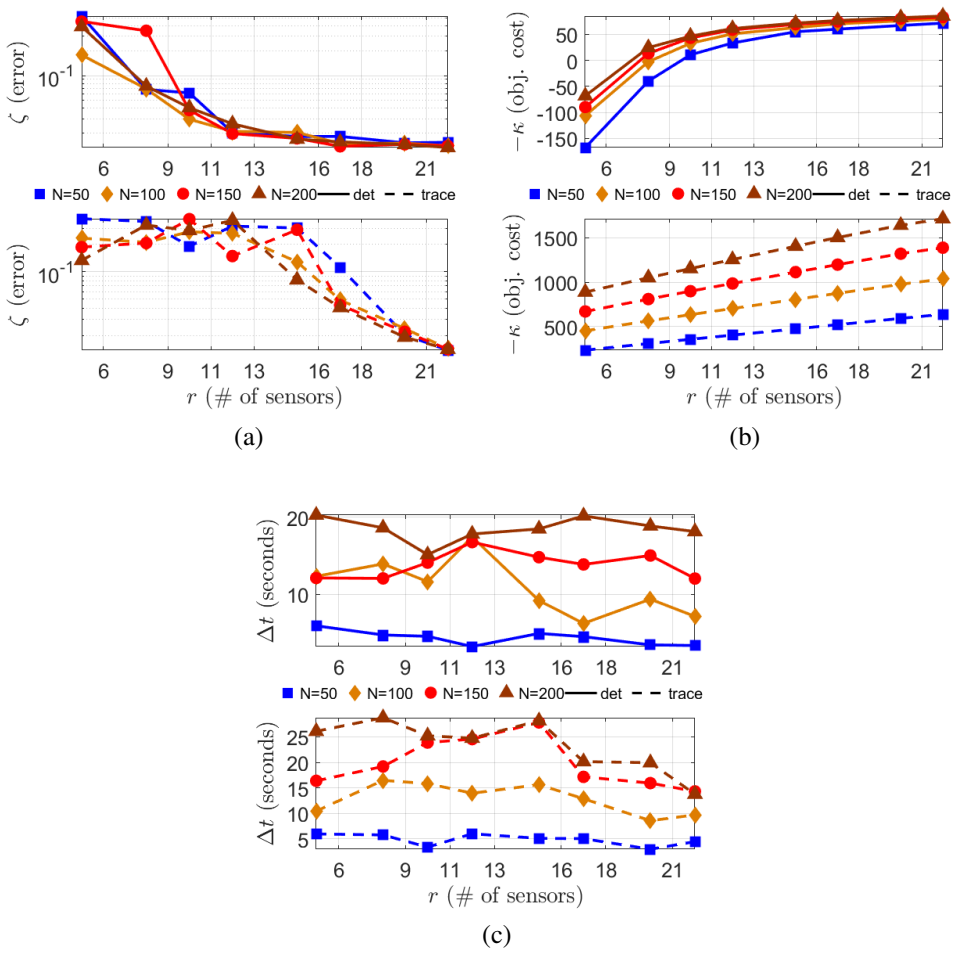


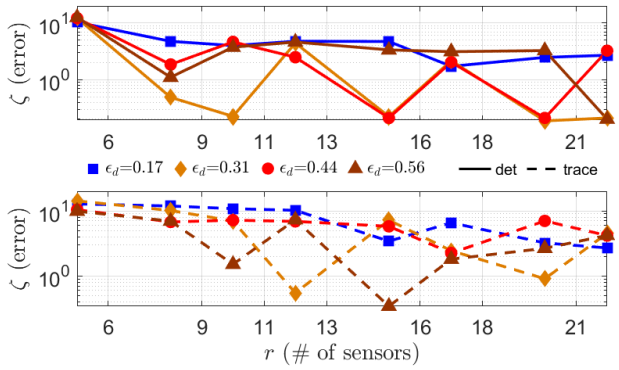
FIGURE 3: **Results on observability analysis for different number of sensor allocations: (a) the relative error ζ, (b) inverse optimal value of P1, denoted by $-\kappa$, and (c) total computational time, which includes the overall time spent for solving P1 and P3.**

solvers.

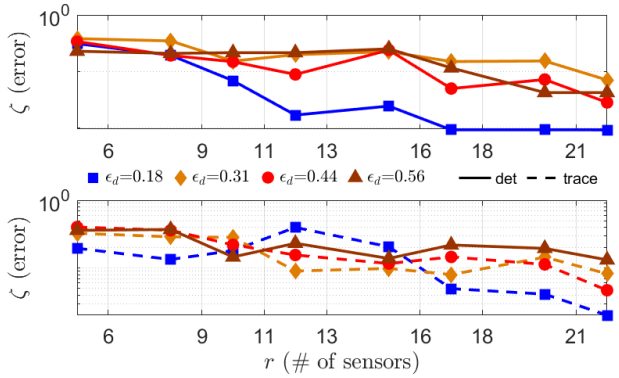
In the first instance of observability analysis, problem **P1** is solved with fixed actual initial state x_0 and presumed initial state $\hat{x}_0$, which are generated randomly in $(0, \rho_m]^n$. **P1** is then solved with different observation windows N for varying number of allocated sensors r. In this case, $r = \lceil p \times r_p\% \rceil$ where $\lceil \cdot \rceil$ is the ceiling function, $p = n = 24$, and $r_p\%$ represents the percentage of sensor's allocation (from 20% to 90%). The number of allocated sensors is a convex constraint such that, with respect to (12b), $\gamma \in \mathscr{G}_\gamma$ is equivalent to $\sum_i \gamma_i = r$.

We solve **P1** for two observability metrics—the trace and determinant metrics discussed in Section 3. Here, the determinant is used in the form of logarithmic determinant based on the suggestion in (*16*). Logarithmic determinant has a concave shape while determinant does not.

In this numerical study, we put our interest on comparing the relative error ζ, the inverse optimal value of **P1**, denoted by $-\kappa$, and total computational time Δt. Problem **P3** is solved by using the MATLAB data-fitting function `lsqnonlin`, which implements a *trust region reflective* algorithm (*27*). The optimality tolerance is set to be 10^{-6} and an initial guess equals to $\hat{x}_0$. To



(a)



(b)

FIGURE 4: **Relative error ζ for (a) various presumed initial state $\hat{x}_0$ and (b) actual initial state x_0.**

reduce the chance of variability, the actual x_0 and presumed $\hat{x}_0$ initial states are fixed for this particular test. The results of this numerical study are depicted in Fig. 3. In particular, it can be seen from Fig. 3a that larger observation window yields smaller relative error. Notice that this behavior is reflected better from the trace but less evident from the determinant—the variations are presumed to be caused by suboptimal solutions obtained from solving **P3**. Fig. 3b also indicates this behavior: larger observation window yields better optimal value **P1**, which in turn implies that the observability measure is directly proportional with observation window. This corroborates first principles in control theory; more available data (i.e., a larger observation window) results in a better system observability. The corresponding total computational time for both **P1** and **P3** are reported in Fig. 3c. The resulting sensor's location for observation window $N = 100$ and $N = 200$ are shown in Tab. 1. For the same number of sensors and observation windows, the optimal sensor locations for determinant and trace objective functions are different. This observation is expected as both metrics measure different aspect on observability Gramian.

Next, we asses the variability of sensor placement due to the differences in actual x_0 and presumed $\hat{x}_0$ initial states. The numerical experiment is carried out as follows. First, we set x_0 to be fixed and use different values for $\hat{x}_0$ which are randomly generated inside $(0, \rho_m]^n$ in such a way that their Euclidean distances, computed as $\varepsilon_d := \|\hat{x}_0 - x_0\|_2$, are unique to each other. For this scenario, the results are given in Fig. 4a, which shows the resulting relative error. In another scenario, $\hat{x}_0$ is fixed while x_0 varies—see Fig. 4b for the results. It can be seen from both figures that, in

TABLE 1: **Sensor locations for two different observation lengths and observability metrics. The notation $(x)+\{y\}$ means that the set of sensors $\{y\}$ is appended to that of row (x) from the same column.**

r	det		trace	
	$N=100$	$N=200$	$N=100$	$N=200$
5	$\{1,7,12,21,24\}$	$\{1,12,21,23,24\}$	$\{12,21,22,23,24\}$	$\{12,21,22,23,24\}$
8	$(5)+\{18,22,23\}$	$(5)+\{6,16,22\}$	$(5)+\{7,8,16\}$	$(5)+\{6,7,8\}$
10	$(8)+\{4,15\}$	$(8)+\{9,18\}$	$(8)+\{6,15\}$	$(8)+\{14,15\}$
12	$(10)+\{3,10\}$	$(10)+\{3,14\}$	$(10)+\{5,14\}$	$(10)+\{5,16\}$
15	$(12)+\{6,16,20\}$	$(12)+\{4,8,19\}$	$(12)+\{3,4,13\}$	$(12)+\{3,4,13\}$
17	$(15)+\{8,14\}$	$(15)+\{2,10\}$	$(15)+\{1,2\}$	$(15)+\{1,2\}$
20	$(17)+\{2,11,19\}$	$(17)+\{5,7,15\}$	$(17)+\{9,10,17\}$	$(17)+\{10,12,17\}$
22	$(20)+\{5,13\}$	$(20)+\{13,17\}$	$(20)+\{11,12\}$	$(20)+\{9,11\}$

general, varying $\hat{x}_0$ while fixing x_0 yields higher relative error magnitudes than fixing $\hat{x}_0$ and varying x_0. Also when $\hat{x}_0$ is fixed, ζ experiences less variations than in the other scenario. However, these variations are most likely attributed to the suboptimal solutions obtained from solving **P3** via `lsqnonlin`. It is also observed that larger distance ε_d does not necessarily result in larger relative error ζ, suggesting that the proposed method is rather resilient towards the values of actual and presumed initial states.

4.2 Traffic Density Estimation with Various Sensor Allocations

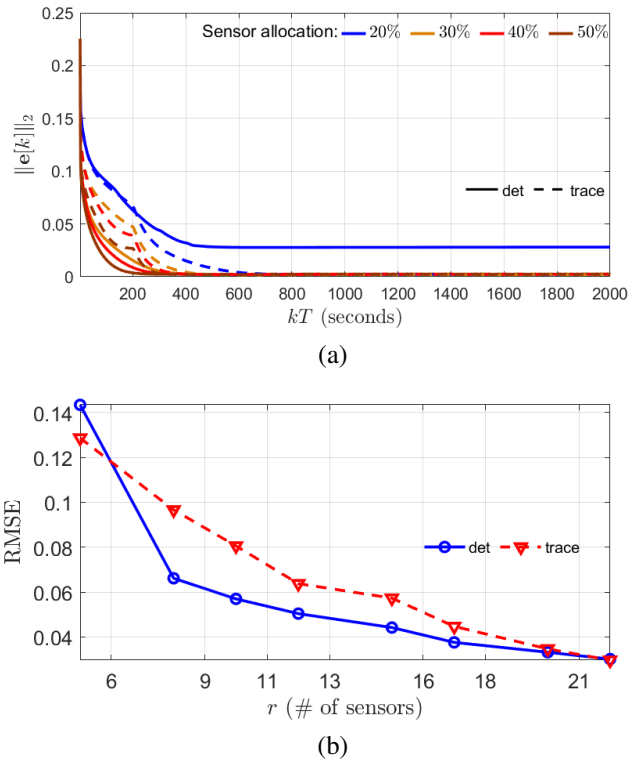
This section investigates different traffic sensor placements—obtained from the previous section—for traffic density estimation purpose. To that end, we implement the Extended Kalman Filter (EKF) such as in (*28*). Herein we simulate Highway A with final time $k_f = 2000$ and invoke Gaussian noise with covariance matrices $Q = \nu I$ and $R = \nu I$ where $\nu = 10^{-3}$ with corresponding unknown input matrices $B_{\mathrm{w}} = \begin{bmatrix} B_{\mathrm{u}} & O \end{bmatrix}$ and $D_{\mathrm{w}} = \begin{bmatrix} O & I \end{bmatrix}$, which simulate process and measurement noise.

In this part of numerical study, we are focused on finding out whether different observability metrics used to solve **P1** have explicit impact on the quality of traffic density estimation, since according to Tab. 1, different observability metrics return distinctive sensor locations. With that in mind, we compare the resulting estimation errors $e[k]$ and the root-mean-square error (RMSE)—which is computed from

$$\mathrm{RMSE} = \sum_{i=1}^{n} \sqrt{\frac{1}{k_f} \sum_{k=0}^{k_f} (e_i[k])^2}. \tag{15}$$

The sensor locations used in this test, for both observability metrics, are collected from Tab. 1 that correspond to observation window $N = 200$.

The results of this numerical test are illustrated in Fig. 5. It is indicated from Fig. 5b that, as more sensors are utilized, the estimation error is decreasing. However, it is of importance to note that using determinant as the observability metric yields better state estimation results than doing so with trace, since the RMSE for determinant is slightly smaller. This showcases the prominence of determinant as opposed to trace to quantify observability. The trajectories of estimation error for different number of sensors are given in Fig. 5a, from which it can be observed that utilizing more



(a)

(b)

FIGURE 5: **State estimation results with EKF for different allocations of sensor: (a) the estimation error and (b) the RMSE.**

sensors yields smaller and faster convergence on estimation error—notice that this observation is in accordance with RMSE shown in Fig. 5b. It is also observed from Fig. 5a that, in general, the determinant returns smaller estimation error compared to the trace. A slight exception occurs when the sensor allocation is at 20% (also see Fig. 5b). This is suspected to be caused by suboptimal sensor placement from the greedy algorithm—see Tab. 1.

4.3 Optimal, Randomized, and Uniform Sensor Placement

In the last part of numerical test, we analyze the results of traffic density estimation with optimal, randomized, and uniform sensor placement. The result of this numerical experiment are provided in Fig. 6. The figure showcases the RMSE of traffic density estimation with optimal, randomized, and uniform sensor placement. For the uniform case, the sensors are firstly placed at odd locations and will begin to be placed at even locations only when all odd locations are used up. Also note that, to compensate randomization on sensor placement, the simulations for this case (randomized sensor placement) are performed 10 times for every r and the results are averaged. It is observed that the optimal sensor placements return significantly smaller RMSE than the randomized and uniform ones, regardless of observability metric being used. These results suggest that using optimal sensor placement, regardless of observability metric, gives better traffic density estimation than doing so with randomized sensor placement. The resulting traffic density estimations are depicted in Fig. 7.

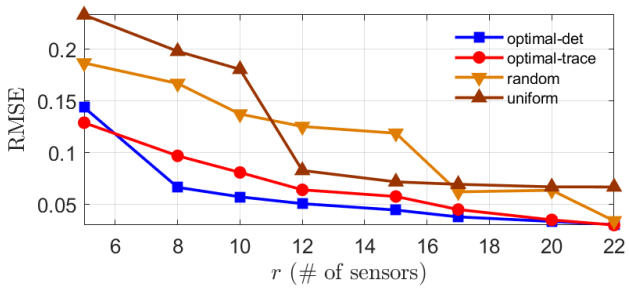
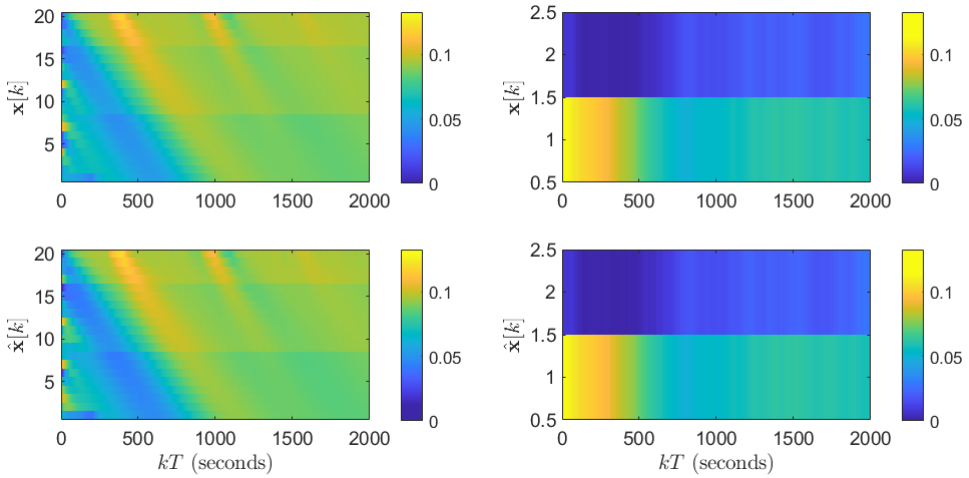


FIGURE 6: **RMSE for optimal, randomized and uniform sensor placement for Highway A. The simulations with randomized sensor placement are performed** 10 **times and the presented results correspond to the average values.**



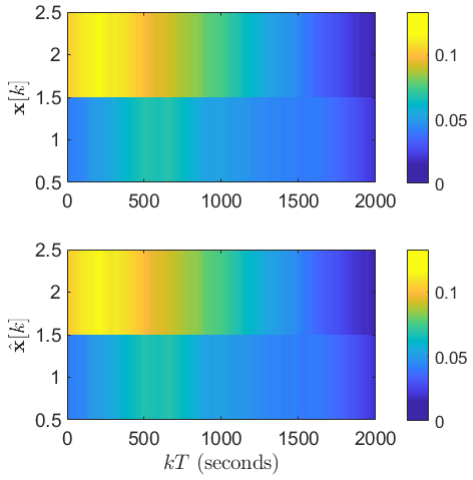


FIGURE 7: **Comparison between real** $x[k]$ **and estimated** $\hat{x}[k]$ **traffic densities on (a) mainline segments, (b) on-ramps, and (c) off-ramps for highway A with** 40% **measurements—sensor locations are obtained form solving P1 with trace observability metric.**

5 CONCLUSIONS, PAPER LIMITATIONS, AND FUTURE WORK

Given the thorough computational analysis in the previous section, the following observations are made, thereby answering the posed research questions in Section 4:

- *A1:* Increasing observation window yields smaller initial state estimation errors while traffic network's quantifiable observability is directly proportional with observation window. This corroborates control-theoretic first principles.
- *A2:* Using both trace and determinant observability metrics to solve the sensor placement problem takes an equivalent amount of time using the greedy algorithm approach.
- *A3:* The resulting sensor placement obtained from solving **P1** can be used over a wide range of initial operating conditions.
- *A4:* When used to determine optimal sensor locations, the determinant is better than the trace in term of state estimation error and hence offers better network-wide observability.
- *A5:* The optimal sensor placement outperforms randomized and uniform sensor placement in estimating traffic density.

The approach in this paper has its own set of limitations. First, we consider a time-invariant traffic model where in reality, some parameters are in fact time-varying, which include critical density, split ratio, free-flow speed, and congestion wave speed. We also do not consider the capacity drop phenomenon or the stochasticity of traffic in the model presented in this paper. Second, the sensor placement strategy does not consider the effect of measurement noise. To that end, future work will include solving the traffic sensor placement problem while considering a time-varying nonlinear traffic model incorporating capacity drop and stochasticity and taking measurement noise into account, which is resulting in a robust sensor placement. Finally, we point out that the presented placement approach in the paper can be extended or utilized to other models in transportation systems beyond stretched highways, assuming that a nonlinear state-space representation of the dynamics is possible.

ACKNOWLEDGMENTS

We acknowledge the financial support from the National Science Foundation through Grants 1636154, 1728629, 1917164, and 1917056.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: S.A. Nugroho, S.C. Vishnoi, A.F. Taha and C.G. Claudel; analysis and interpretation of results: S.A. Nugroho, S.C. Vishnoi, A.F. Taha and C.G. Claudel; draft manuscript preparation: S.A. Nugroho, S.C. Vishnoi, A.F. Taha and C.G. Claudel. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Lovisari, E., C. Canudas de Wit, and A. Kibangou. Density/Flow reconstruction via heterogeneous sources and Optimal Sensor Placement in road networks. *Transportation Research Part C*, Vol. 69, No. C, 2016, pp. 451–476.
2. Gentili, M. and P. Mirchandani. Locating sensors on traffic networks: Models, challenges and research opportunities. *Transportation Research Part C*, Vol. 24, 2012.
3. Viti, F., M. Rinaldi, F. Corman, and C. M. Tampère. Assessing partial observability in network sensor location problems. *Transportation research. Part B: methodological*, Vol. 70, 2014, pp. 65–89.
4. Castillo, E., Z. Grande, A. Calviño, W. Y. Szeto, and H. K. Lo. A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks. *Journal of Sensors*, Vol. 2015, 2015, pp. 1–26.
5. Salari, M., L. Kattan, W. H. Lam, H. Lo, and M. A. Esfeh. Optimization of traffic sensor location for complete link flow observability in traffic network considering sensor failure. *Transportation Research Part B*, Vol. 121, 2019, pp. 216–251.
6. Eisenman, S. M., X. Fei, X. Zhou, and H. S. Mahmassani. Number and location of sensors for real-time network traffic estimation and prediction: Sensitivity analysis. *Transportation Research Record*, Vol. 1964, No. 1, 2006, pp. 253–259.
7. Fei, X., H. S. Mahmassani, and S. M. Eisenman. Sensor coverage and location for real-time traffic prediction in large-scale networks. *Transportation Research Record*, Vol. 2039, No. 1, 2007, pp. 1–15.
8. Agarwal, S., P. Kachroo, and S. Contreras. A dynamic network modeling-based approach for traffic observability problem. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 4, 2016, pp. 1168–1178.
9. Contreras, S., P. Kachroo, and S. Agarwal. Observability and sensor placement problem on highway segments: A traffic dynamics-based approach. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 3, 2015, pp. 848–858.
10. Greenshields, B., J. Bibbins, W. Channing, and H. Miller., A study of traffic capacity. In *Highway research board proceedings*, National Research Council (USA), Highway Research Board, 1935, Vol. 1935.
11. Gomes, G., R. Horowitz, A. A. Kurzhanskiy, P. Varaiya, and J. Kwon. Behavior of the Cell Transmission Model and effectiveness of ramp metering. *Transportation Research Part C: Emerging Technologies*, Vol. 16, No. 4, 2008, pp. 485 – 513.
12. Daganzo, C. F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, Vol. 28, No. 4, 1994, pp. 269 – 287.
13. Gomes, G. and R. Horowitz. Optimal freeway ramp metering using the Asymmetric Cell Transmission Model. *Transportation Research Part C: Emerging Technologies*, Vol. 14, No. 4, 2006, pp. 244–262.
14. Contreras, S., S. Agarwal, and P. Kachroo. Quality of traffic observability on highways With Lagrangian sensors. *IEEE Transactions on Automation Science and Engineering*, Vol. 15, No. 2, 2018, pp. 761–771.
15. Serpas, M., G. Hackebeil, C. Laird, and J. Hahn. Sensor location for nonlinear dynamic systems via observability analysis and MAX-DET optimization. *Computers & Chemical Engineering*, Vol. 48, 2013, pp. 105 – 112.

16. Qi, J., K. Sun, and W. Kang. Optimal PMU placement for power system dynamic state estimation by using empirical observability gramian. *IEEE Transactions on power Systems*, Vol. 30, No. 4, 2015, pp. 2041–2054.
17. Singh, A. K. and J. Hahn. Sensor location for stable nonlinear dynamic systems: Multiple sensor case. *Industrial & engineering chemistry research*, Vol. 45, No. 10, 2006, pp. 3615–3623.
18. Haber, A., F. Molnar, and A. E. Motter. State observation and sensor selection for nonlinear networks. *IEEE Transactions on Control of Network Systems*, Vol. 5, No. 2, 2018, pp. 694–708.
19. Bopardikar, S. D., O. Ennasr, and X. Tan. Randomized sensor selection for nonlinear systems With application to target localization. *IEEE Robotics and Automation Letters*, Vol. 4, No. 4, 2019, pp. 3553–3560.
20. Nugroho, S. A. and A. F. Taha., Sensor placement strategies for some classes of nonlinear dynamic systems via Lyapunov Theory. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 4551–4556.
21. Lighthill, M. J. and G. B. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A*, Vol. 229, No. 1178, 1955, pp. 317–345.
22. Richards, P. I. Shock waves on the highway. *Operations research*, Vol. 4, No. 1, 1956, pp. 42–51.
23. Gomes, G. C. *Optimization and microsimulation of on -ramp metering for congested freeways*. Ph.D. thesis, 2004.
24. Vishnoi, S. C., S. A. Nugroho, A. F. Taha, C. G. Claudel, and T. Banerjee., Asymmetric Cell Transmission Model-Based, Ramp-Connected Robust Traffic Density Estimation under Bounded Disturbances. In *2020 American Control Conference, ACC 2020, Denver, CO, USA, July 1-3, 2020*, IEEE, 2020, pp. 1197–1202.
25. Hanba, S. On the "uniform" observability of discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, Vol. 54, No. 8, 2009, pp. 1925–1928.
26. Nugroho, S. A. *Control node and sensor selection in dynamical systems*. Ph.D. thesis, 2021.
27. Coleman, T. F. and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, Vol. 6, No. 2, 1996, pp. 418–445.
28. Wang, Y. and M. Papageorgiou. Real-time freeway traffic state estimation based on Extended Kalman filter: a general approach. *Transportation Research Part B: Methodological*, Vol. 39, No. 2, 2005, pp. 141–167.