On Cybersecurity of Freeway Control Systems: Analysis of Coordinated Ramp Metering Attacks

Jack Reilly*
Graduate Student
Department of Civil and Environmental Engineering
University of California Berkeley
652 Sutardja Dai Hall
Berkeley CA 947201710
Phone: (916) 768-1755
Email: jackdreilly@berkeley.edu

Sebastien Martin
Graduate Student
Massachusetts Institute of Technology

Mathias Payer
Assistant Professor
Department of Computer Science
Purdue University

Alexandre M. Bayen
Chancellor Associate Professor
Director, Institute of Transportation Studies
Department of Electrical Engineering and Computer Sciences
Department of Civil and Environmental Engineering
University of California, Berkeley

* - Corresponding Author

words + 7 figures + 1 tables

ABSTRACT

This article focuses on cybersecurity of transportation systems and investigates their vulnerability to attacks on the sensing and control infrastructure. An array of different attack points, classified into *physical*, *close-proximity*, and *virtual* layers, are reviewed and investigated. We construct two benchmark *scenarios* which exploit these vulnerabilities to identify the potential harm of a traffic control system compromise. A more in-depth analysis is then presented on the takeover of a series of networked onramp metering traffic lights. The analysis is conducted using a methodology for precise and intelligent onramp metering attacks based on finite-horizon optimal control techniques and multi-objective optimization. The methodology is demonstrated in simulation for two examples of high-level attack objectives: *congestion-on-demand*, which aims to create precise pockets of congestion, and *catch-me-if-you-can*, which attempts to aid a fleeing vehicle from chasing pursuants.

INTRODUCTION

Public traffic infrastructure is arriving in the cyber age with increasing connectivity between the different segments of roadways. For example, freeways are commonly instrumented with loop detectors that allow for real-time monitoring of roadway speeds (1). Estimates of road traffic conditions are then fed directly into onramp traffic light metering algorithms which regulate traffic flow to improve congestion (2). Finally, these metering algorithms can be coordinated and controlled by a remote command and monitoring center, leading to a regional network of interconnected sensors and controllers (3).

Increased efforts to build systems which understand and utilize the interconnectivity are evidenced by *integrated-corridor-managament* (ICM) projects such as *Connected Corridors* (4) and mobile applications which use GPS probe data to improve navigation (5).

This connectivity offers great potential to better analyze, control and manage traffic but also poses a significant security risk. A compromise at any level of the traffic control infrastructure can lead to both direct access of an attacker to alter traffic lights and changeable message signs, and indirect access via spoofing of sensor readings, which may *trick* the control algorithms to respond to false conditions.

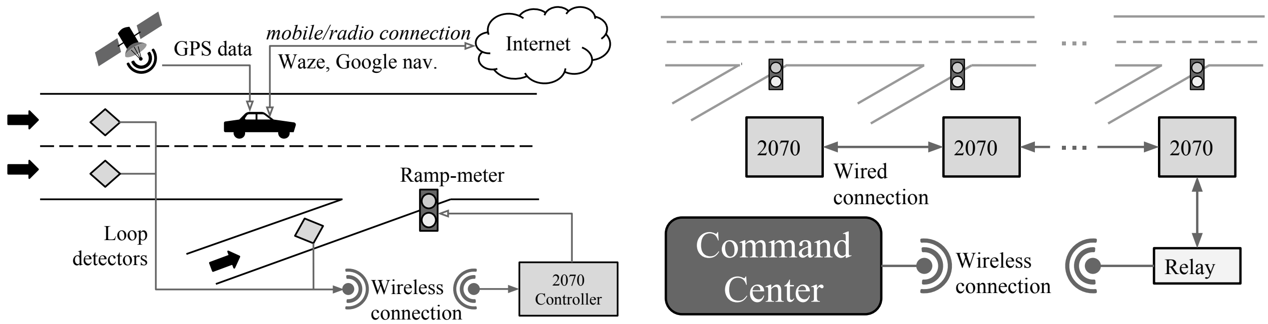
A number of traffic-related atacks of infrastructure systems have already been demonstrated in the past few years. A man-in-the-middle attack on GPS coordinate transmissions from mobile navigation applications showed it is possible to trick navigation services into inferring non-existent jams (6), while a similar attack used a fleet of mobile phone emulators to mimic the presence of many virtual vehicles on a roadway (7). A popular type of vehicle-detection sensor was revealed to use a type of wireless protocol vulnerable to data injection attacks, and a demonstration showed that the access point could be tricked into receiving arbitrary readings (8). Cyber attacks on a centralized command center remain a serious threat given the frequent discovery of networking vulnerabilities, such as the Heartbleed bug (9). Even insider attacks on command centers have precedent as two Los Angeles traffic engineers in 2009 were found guilty of intentionally creating massive delays by adjusting signal times at key intersections (10).

Given the existence of such vulnerabilities and the scale at which they can be exploited, understanding the nature and costs of such attacks becomes paramount to public safety. In this article, we present a systematic approach to analyzing the topic of traffic control system vulnerabities and their potential impact.

To do so, we begin by constructing a taxonomy of different vulnerabily locations in traffic control systems, defining three distinct layers: physical, close-proximity, and virtual. Difficulty, impact, and cost values are also associated with each potential attack. We motivate our classifications by presenting two scenarios that combine a number of attacks to accomplish a high-level goal.

We then focus our analysis on an in-depth exploration of freeway attacks using coordinated, ramp metering.. We show using the developed method that ramp metering control permits an attacker to achieve very precise congestion patterns. An attacker can then consider high-level objectives, such as permitting a fleeing vehicle to escape pursuants on a particular freeway stretch. To achieve this, we develop a methodology based on adjoint computations and finite-horizon optimal control for finding optimal metering rates to create a desired disruption on the freeway. We additionally give an overview of multi-objective optimization and discuss how such an approach is useful for solving high-level attack objectives which contain many conflicting sub-goals.

Two detailed applications of the multi-objective optimal control approach to ramp metering attacks are then given. The first application shows how ramp metering can allow an attacker to cause congestion in precise locations and at precise moments in time along a freeway. The second application finds a strategy to solve the aforementioned problem of allowing a fleeing vehicles to escape pursuants. Numerical results are presented, as well as a discussion of the benefits of the



(a) Local freeway control system.

(b) Global freeway control system.

FIGURE 1 The physical roadway, sensors, connected vehicles and controllers near a freeway/onramp junction in Figure 1(a) form a cyber-physical network we refer to as a local freeway control system. In Figure 1(b), the local controllers are wired together, then connected to a command center via a relay box to form the global control system. This article analyzes vulnerability locations associated with each component.

multi-objective optimization method. We conclude with some future areas of study for traffic system security as well as extensions of the multi-objective optimization approach to other transporation applications.

TRAFFIC SYSTEM VULNERABILITIES

The Freeway Control System

Modern freeways encompass control and monitoring mechanisms which permit traffic management to mitigate congestion and improve traffic flow in real-time. While the exact combination of sensors, controllers and transmitters differ from location to location, this article chooses one particular instatiation of a freeway control system, which we find to be representative. Figure 1(a) shows a control system installed near a junction of a freeway and an onramp. We consider three elements of the control system:

- Sensors, used to gather information about the freeway state. For example, loop detectors are used to acquire the flow of vehicles along the freeway and onramps/offramps, while the trajectory of vehicles equipped with GPS (or containing GPS-powered smartphone applications) can be used for estimating real-time traffic conditions (5).
- Actuators, used to influence the evolution and efficiency of the freeway. The most common actuation strategy is *ramp metering*, where traffic lights installed on freeway onramps control the influx of vehicles to the mainline. Other actuators include variable speed limit control (11) and variable message signs. For the purposes of this article, the ramp meters are the only actuators we will consider.
- Local controllers, such as *2070* boxes (12) and the older *170* boxes (13), which allows interaction between the sensors and ramp meters.

We assume control boxes are wired to the nearby metering light and have a wireless connection to nearby sensors. Vehicles with navigation devices such as TomTom (14) automatically analyze radio-broadcasted traffic reports from traffic control centers to improve their navigating functionality.

In order to allow coordinated control and sensing across a freeway stretch with many on-ramps, the local control systems are connected to allow for a more global configuration. Figure 1(b) depicts our representative global communication architecture. The local control boxes are wired together along the freeway to form the actuation network, with intermediary *relay boxes* allowing for an uplink and downlink to a remote *command center*. The command center contains instrumentation and personnel for monitoring traffic conditions and setting the metering lights accordingly.

Infrastructure Weaknesses

Infrastructure is built up of several layers and each layer poses individual security risks, starting from tampering with the actual devices, cables or wireless signals, to attacking the software of deployed devices or attacking the command center. Attackers can leverage vulnerabilities in the infrastructure to control or disrupt these connected systems. Individual attacks can thereby target the physical layer, the communication layer, at the layer of the control center, or any combination thereof.

Direct physical access: The physical layer is the lowest attackable layer and involves direct access to individual wires, opening and accessing the control box, or tampering with individual sensors. Physical attacks involve clipping, tampering, removing, or replacing of wires or hardware. For instance, copper wire theft near freeways is a common occurence (15, 16). Such attacks need low sophistication, are easy to carry out, and are hard to protect against as each device must be physically protected given that software-based protection is not effective against physical attacks. On the other hand, the attack is costly as (i) direct physical access is needed, (ii) the attacker is exposed, and (iii) the attack does not scale (i.e., each piece of equipment is attacked individually). Examples of such an attack in Figure 1(a) include clipping or removing wires between sensors and the *2070* controller, tampering with individual sensors, the ramp meter, or the *2070* controller.

Proximity access (locality): Figure 1(b) depicts multiple control boxes chained together to form a corridor where actuators have a coordinated plan between the different control boxes. An attack on the communication layer forges, removes, replaces, or inserts attacker-controlled measurements into the control system, which may then make further decisions based on forged data. An attacker can either replace or add sensors to the current sensor network to inject new measurements or attack the software running on sensors and/or actuators to take over control. Both aspects of the attack are feasible; the first aspect needs additional hardware and an attacker that delivers the hardware, the second aspect needs to find a software vulnerability with a security analysis of the existing devices. These attacks need higher sophistication and knowledge but no longer need direct hardware access to the existing sensors and scales to some extent.

Networked/virtual access: Remote connections from the physical freeway infrastructure to the command center defines another layer with potential vulnerabilities. An attack on this layer can be done by forging or controlling messages from/to the command center and possibly even compromises the command center itself. For this scenario an attacker needs to find software vulnerabilities in the software running in the command center. Direct access to these centers is usually not given and this attack therefore is highly sophisticated (or needs insider access). This attack is the hardest possible attack as command centers and back links are usually guarded but allows a great scaling effect as many control boxes can be controlled directly.

Table 1 gives a (partial) list of vulnerabilities in our freeway control system along with classifications for each attack.

Attack Scenarios

We will consider two fictional but realizable attack scenarios and study their consequences on the compromised network. The first scenario involves indirect control of the freeway, through spoofing

Attack Description	Access	Control	Complexity	Cost
copper theft/clipping wires	physical	low	low	low
replacing a single sensor/actuator	physical	low	low	low
attacking a single sensor/actuator	locality	low	medium	low
replacing a single control box	physical	medium	medium	medium
replacing a set of sensors/actuator	physical	medium	medium	medium
attacking a set of sensors/actuator	locality	low	medium	low
replacing a corridor of control boxes	physical	high	medium	medium
attacking a corridor of control boxes	network	high	high	medium
attacking the control center	network	high	high	high
spoofing GPS data	network	medium	high	medium
attacking navigation software	network	medium	medium	medium

TABLE 1 List of possible infrastructure attacks with access to different layers that is needed, level of control that the attacker gains, sophistication of the attack, and cost.

the sensors, to achieve a local objective. The second scenario involves complete control of the ramp meters to achieve a global objective along a larger stretch of freeway.

Indirect Attack: VIP-lane
The objective of the attacker is to clear a predetermined section of a regularly congested freeway. The attacker decides to drop low-cost wireless transmitters near the *2070* controllers of the freeway section[1]. As the actual loop-detector sensors communicate with the control box wirelessly, the attacker will be able to override the loop-detector signals and send false data that indicates a fully congested freeway. This will indirectly affect the ramp meters, which will respond by limiting onramp flow and thus clearing the mainline of the freeway. The attacker will then transmit false GPS location data via a set of hacked cellphones to trick navigation software into believing the freeway is congested. Approaching vehicles using navigation software will then be rerouted around the fake congestion which leads to a further reduction in incoming flow. The net effect of the attack is a congestion-free commute for the attacker: a private VIP lane created purely by indirect, sensor-based attacks.

Direct Attack: catch-me-if-you-can
The objective of the attacker is to escape from pursuants along a large section of freeway. In order to achieve this objective, a full control of the ramp meters is used. One approach is to hack the command center itself, with the downside being the expensiveness and complexity of such an attack (see Table 1). Another solution is to begin by hacking of the *2070* boxes, and since all the *2070* boxes are networked along the freeway (see Figure 1(b)), a single hacked box can serve as a means of compromising the other nearby boxes, leading to a cascading attack. The attacker can then acquire full control of all the *2070* boxes, and in turn, the ramp metering lights.

Once full control is obtained, precise control must be applied to achieve the desired objective. The remainder of this article describes how a freeway can be controlled just by varying the metering lights in a coordinated fashion, and how the *catch-me-if-you-can* scenario can be achieved.

[1]see our link (17) for a Youtube video depiction

THEORY FOR COORDINATED FREEWAY ATTACKS

An attacker can negatively influence the performance of the freeway network or achieve some criminal goal by setting the metering lights to a particular configuration. Such an attack can be carried out by leveraging a discrete dynamical freeway model to compute metering rates using finite-horizon optimal control and multi-objective optimization techniques.

Freeway Model

We model the freeway as a sequence of n mainline links (labeled $1, \ldots, n$), where both an onramp and offramp are present between consecutive links[2]. Flow dynamics along a link i is modeled using a discretized version of the *Lighthill-Whitham-Richards* (18, 19) (LWR) partial differential equation. The continuous LWR equation takes the following form:

$$\frac{\partial \rho_i(t,x)}{\partial t} + \frac{\partial f(\rho_i(t,x))}{\partial x} = 0, \tag{1}$$

with $\rho_i(t,x)$ representing the *density* of vehicles at a particular point in space and time, and f capturing the relationship between the density and *flow* of vehicles, a relationship referred to as a *fundamental diagram* of traffic. We assume f has the following triangular form (20):

$$f(\rho) = \min\left(v\rho, w(\rho^{\max} - \rho), f^{\max}\right),$$

where $v, w, \rho^{\max}$ and $f^{\max}$ are characteristics of the particular freeway section.

Our discrete model is adapted from (3, 21) and was chosen for its suitability to ramp metering applications. Following (3), we discretize Equation (1) into cells of spatial size $\triangle x$ and temporal size $\triangle t$ using a *Godunov-based* or *cell-transmission-model* (CTM) scheme (20, 22, 23). The resulting discrete model has T time-steps, N spatial cells, and N onramps and offramps. The state of cell $i \in [1, N]$ at time $k \in [1, T]$ is given by $\rho[i,k]$, while the number of vehicles on the adjacent onramp is given by $l[i,k]$. The states of cell and onramp i are advanced from time k to $k+1$ according to the following equations:

$$\delta[i,k] = \min\left(v\rho[i,k], f^{\max}\right) \tag{2}$$

$$\sigma[i,k] = \min\left(w\left(\rho^{\max} - \rho[i,k]\right), f^{\max}\right) \tag{3}$$

$$d[i,k] = \min\left(l[i,k]/\triangle t, r^{\max}\right) \tag{4}$$

$$f^{\text{in}}[i,k] = \min\left(\sigma[i,k], d[i-1,k] + \beta[i,k]\,\delta[i,k]\right) \tag{5}$$

$$f^{\text{out}}[i,k] = \begin{cases} \delta[i,k] & \text{if } \frac{p f^{\text{in}}[i+1,k]}{\beta[i,k](1+p)} \geq \delta[i,k] \\ \frac{f^{\text{in}}[i+1,k] - d[i+1,k]}{\beta[i,k]} & \text{if } \frac{f^{\text{in}}[i+1,k]}{1+p} \geq d[i+1,k] \\ \frac{p f^{\text{in}}[i+1,k]}{(1+p)\beta[i,k]} & \text{otherwise} \end{cases} \tag{6}$$

$$r[i,k] = f^{\text{in}}[i,k] - \beta[i,k]\, f^{\text{out}}[i,k] \tag{7}$$

$$\rho[i,k+1] = \rho[i,k] + \frac{\triangle t}{\triangle x}\left(f^{\text{in}}[i,k] - f^{\text{out}}[i,k]\right) \tag{8}$$

$$l[i,k+1] = l[i,k] + \triangle t\left(D[i,k] - r[i,k]\right) \tag{9}$$

Equations (2)-(9) model the merging of onramp and mainline flows, as well as the propogation of congestion waves across the freeway network.

[2]Spatial cells which do not have an adjacent onramp (or offramp), one can set the vehicle demand to zero (set the offramp turning ratio to zero).

Onramp Metering Model We introduce a control parameter $u_i[k] \in [0,1]$, a scaling factor on the demand of onramp i at time-step k and represents the influence of onramp traffic lights on the discrete model. We augment Equation (4) to include the introduced control:

$$d[i,k] = u[i,k] \min\left(l[i,k]/\triangle t, r^{\max}\right) \tag{10}$$

Finite-Horizon Optimal Control and the Adjoint Method.

Using the model in Section 4.1, we seek a method to compute a coordinated ramp metering policy $u[i,k]$ over all space $i \in [1,N]$ and time $k \in [1,T]$, which minimizes (or reduces) some specified objective. We cast the problem as a finite-horizon optimal control problem, and present a methodology, referred to as the *adjoint method*, for solving such constrained optimization problems.

Generally speaking, we consider the minimization of some objective that is a function of both the control variables and the *state* variables. The state variables are assumed a deterministic function of the control variables. Let $\mathbf{u}$ be the concatenation of all metering control parameters $u[i,k]$ and let ρ be the concatenation of all state variables (variables not controlled directly, e.g. density and queue length variables). After concatenating all the discrete Equations (2)-(9) and moving all terms to the left-hand side, one can succinctly express the discrete, controllable dynamical system by:

$$H(\mathbf{u}, \rho) = 0. \tag{11}$$

Given some objective function $J(\mathbf{u}, \rho)$, our goal is now to find the optimal $\mathbf{u}^*$ which solves the following constrained *finite-horizon optimal control* problem:

$$\min_{u} J(\mathbf{u}, \rho) \tag{12}$$

$$\text{subject to: } \textit{Equation } (11). \tag{13}$$

Gradient Methods via the Adjoint Method As J and H may be non-convex functions of the control and state, it is not always possible to efficiently find the global optimum of J in Problem (12)-(13). Thus, we use a first-order gradient descent approach as a means of reducing the objective value.

We now need to compute the gradient of J with respect to the control variables $\mathbf{u}$ subject to the H constraints. With the partial derivative[3] expressions of H and J, we can compute the gradient of J with respect to $\mathbf{u}$:

$$\nabla_{\mathbf{u}} J\left(\mathbf{u}', \rho'\right) = \frac{\partial J(\mathbf{u}', \rho')}{\partial \rho} \frac{d\rho}{d\mathbf{u}} + \frac{\partial J(\mathbf{u}', \rho')}{\partial \mathbf{u}} \tag{14}$$

or in abbreviated notation:

$$\nabla_{\mathbf{u}} J = J_\rho d_{\mathbf{u}} \rho + J_{\mathbf{u}} \tag{15}$$

It is often prohibitively expensive to compute $d_{\mathbf{u}}\rho$ explicitly. Therefore, as the gradient of H with respect to $\mathbf{u}$ is always zero (since the right hand size is constant for feasible $\mathbf{u}, \rho$):

$$\nabla_{\mathbf{u}} H = H_\rho d_{\mathbf{u}} \rho + H_{\mathbf{u}} = 0, \tag{16}$$

we can add it to Equation (15) with a Lagrange-like multiplier λ:

[3] The partial derivative terms are not always defined in terms of classical derivatives. We omit this technical detail to simplify the presentation and instead refer the reader to (24, 25, 26).

$$\nabla_{\mathbf{u}} J = J_\rho d_{\mathbf{u}}\rho + J_{\mathbf{u}} + \lambda^T \left(H_\rho d_{\mathbf{u}}\rho + H_{\mathbf{u}}\right) \tag{17}$$
$$= \left(J_\rho + \lambda^T H_\rho\right) d_{\mathbf{u}}\rho + \left(J_{\mathbf{u}} + \lambda^T H_{\mathbf{u}}\right) \tag{18}$$

The adjoint method chooses the λ value to set the first term to zero (and eliminate $d_{\mathbf{u}}\rho$), and arrive at the following expressing for $\nabla_{\mathbf{u}} J$:

$$\nabla_{\mathbf{u}} J = \left(J_{\mathbf{u}} + \lambda^T H_{\mathbf{u}}\right) \tag{19}$$
$$\text{such that: } H_\rho^T \lambda = -J_\rho \tag{20}$$

The λ variable is commonly referred as the *discrete adjoint variable* (24, 27), while the system of equations in (20) is referred as the *discrete adjoint system*. It is shown in (3) that for freeway traffic network applications, the adjoint method leads to gradient computations which scale linearly with the size of the network and time-horizon, making it especially suitable for real-time applications.

Several Objectives: Interactive Multi-objective Optimization

Some objective are hard to state: as a consequence, traducing a goal into an objective function to minimize is not always an easy thing to do. A solution is to divide the objective into multiple and smaller sub-objectives that are easier to state.

For example, in the *catch-me-if-you-can* scenario the attacker wants to escape from his chasers. Hence the attacker wants to cross the freeway as quickly as possible, but also wants to slow down the chasers. As a consequence, we have two simpler but competing objectives.

Such a situation with multiple, competing objectives can be described as a *multi-objective optimization problem*.

Multi-objective Optimization and Pareto Front

Definition 4.1 (Multi-objective optimization problem)**.** Given $N \in \mathbb{N}$, let $(f_i(\mathbf{u}, \rho))$ be a set of objective functions describing the goal of a freeway attack. The *multi-objective optimization problem* we consider is the following simultaneous minimization problem:

$$\min_{x \in X} \; (f_1(x), f_2(x), \ldots, f_N(x)) \tag{21}$$

As we want to minimize a vector and not a scalar, we need to define how a solution of equation (21) can be "better" than another.

Definition 4.2 (Pareto front)**.** An solution $x \in X$ is said to *Pareto dominate* another solution x' if:

- $\forall i \leq N \quad f_i(x) \leq f_i(x')$
- $\exists j \leq N \quad f_j(x) < f_j(x')$

A solution $x \in X$ is called *Pareto optimal* if there is no other solution x' that dominates it. The set of all Pareto-optimal solutions is called the *Pareto front*, $P \subseteq X$

Hence, we consider Pareto-optimal solutions to be the solutions of Equation (21).

Decision Maker
There are many ways to find a Pareto-optimal solution. For example if we have three objective functions, we can minimize f_1 first, minimize f_2 on the subset $\arg\min_{x \in X} f_1(x)$ and finally minimize f_3 on the remaining subset to obtain a Pareto-optimal solution. But we could also do the same in any order, with potentially very different results. Thus, the Pareto front can sometimes be very large and hard to explore.

As a consequence, we need to be able to identify the most desirable solutions within the potentially large Pareto front. As all Pareto-optimal solutions are equally considered solutions of Equation (21), human expertise is needed to select the preferred solutions.

The *Decision Maker* (DM) represents the human whose expertise will help solve the multi-objective optimization problem. We assume that the DM is able to discriminate any solution on the Pareto front. As a consequence, the DM has a hidden objective function: $u(\mathbf{u}, \rho)$, the *utility function*, which can only be indirectly observed through probing the DM. With u, we can reformulate the multi-objective optimization problem as:

$$\min_{x \in P} \; u(x) \tag{22}$$

The DM is essential to most multi-objective optimization techniques, and there are several ways to interact with him:

- He can evaluate his utility function u on any given Pareto-optimal solution.
- He can give more general preferences on the Pareto front, for example a preference for one of the objective functions, or for a given subset of the Pareto front.

Finite-horizon oOptimal Control and Multi-objective Optimization
Scalarization In order to find Pareto-optimal solutions, we will reduce the problem to the common scalar minimization problem, which can be solved with the optimal control tools of Section 4.2. This process is called *scalarization*. As our particular scalarization, we use a linear combination of the individual objective functions:

$$f(x) = \sum_{i \leq N} a_i f_i(x). \tag{23}$$

The DM can favor a specific objective f_i over other objectives by increasing the a_i coefficient.

As a consequence, we can explore at least a subset (with the hope that this subset is representative of the entire Pareto front) of the Pareto front by minimizing a linear combination of the objective functions.

A Posteriori Method Equation (23) allows one to sample the Pareto front by exploring the space of the coefficients which can provide to the DM a representative subset of Pareto-optimal solutions. The DM can then chose *a posteriori* his preferred solutions. And as such this method is called an *a posteriori method*.

This method can be computationally costly, but provides a good overview of the Pareto front. In particular, it gives an estimation of the lower and upper bounds of each objective function. Thus one can scale each objective function to take values only between 0 and 1, allowing the different objectives to be easily compared.

Interactive Method Unlike with the a posteriori method, *Interactive methods* are based upon a repeated interaction with the Decision Maker.

1. The DM gives an indication of how to compute the next Pareto-optimal solution — for example an idea for the next set of coefficients (a_i) to use and his evaluation of the previous simulation.
2. The interactive scalarization process uses that indication to create a scalar objective — for example using Equation (23), we obtain a scalar objective with the set of coefficients given by the DM.
3. The finite-horizon optimal control method is used to solve the corresponding optimization problem, and gives the result to the DM.

This process is repeated until the DM is satisfied with the results.

The important part of the interactive method is the kind of indications that can be given by the DM, and how the indications and the simulation history will be used in the scalarization process. Section 5.2 gives an example of an interactive method.

ATTACKS

We will now apply the tools of *adjoint-based finite-horizon optimal control* and *multi-objective optimization* from Section 4 to two examples of attacks. The first attack highlights the precision of coordinated ramp metering attacks, while the second showcases the benefits of multi-objective optimization.

First Attack: *congestion-on-demand*

Congestion-on-demand describes a class of objectives where an attacker wishes to create congestion patterns of a precise nature. This can be done by constructing objectives which maximize total-travel-time over the desired region in space and time, and minimize total-travel-time everywhere else.

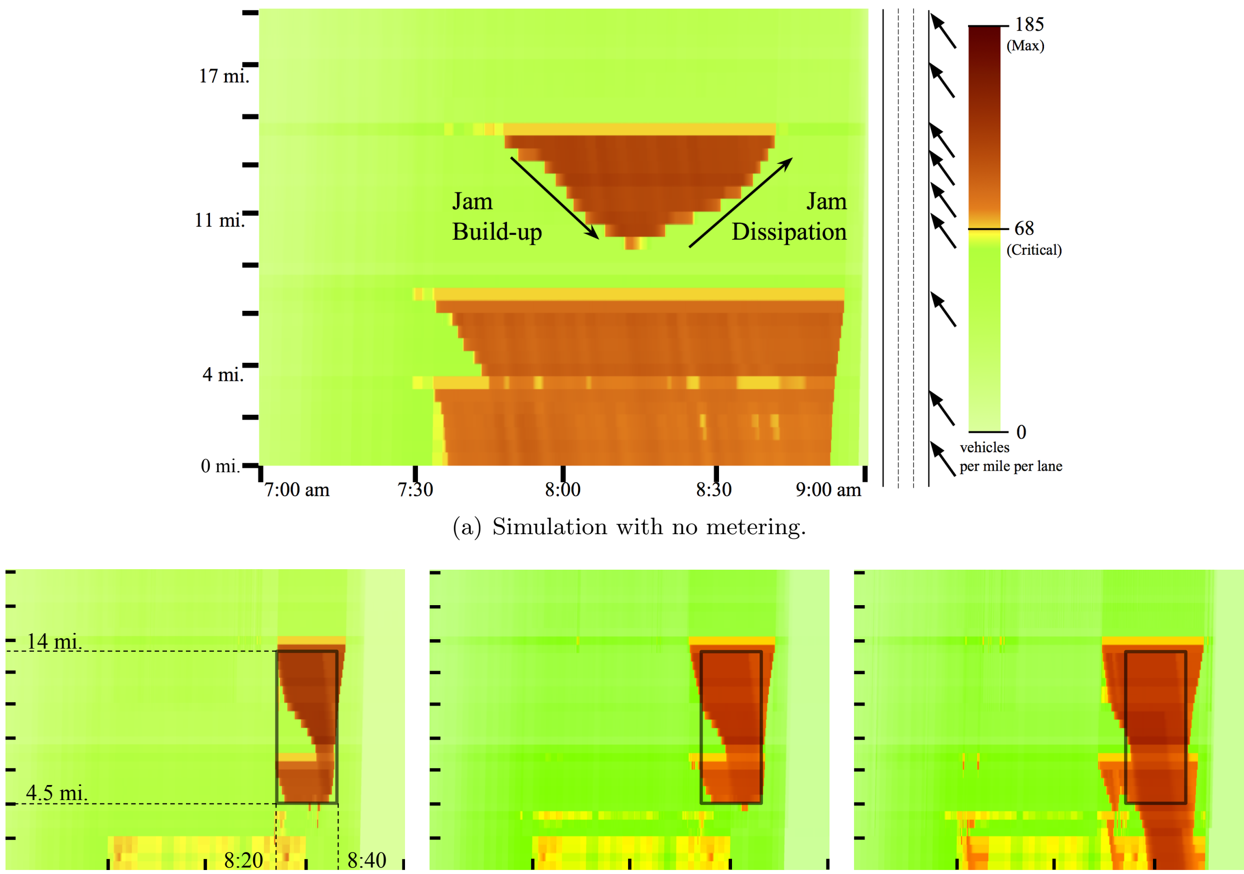
The attacks for the first example, *box objective* (to be described), use a macroscopic freeway model of a 19.4 mile stretch of the I15 South Freeway in San Diego California. The model was split into 125 links with 9 onramps and was calibrated (28, 29) using loop-detector measurements available through the PeMS loop-detector system (1).

Figure 2(a) is a *Space-time diagram* of the I15 freeway. It plots a color representation of traffic density ρ for every time and location. Given the relationship between ρ, the velocity v and the flow f (see Section 4.1), the space-time diagram gives a good indication of the entire freeway state. There is no ramp metering control applied to the simulation in Figure 2(a), i.e. the ramp meters are always set to green.

Examples

Box Objective The *box objective* creates a box of congestion in the space-time diagram, i.e. congestion will be created on a precise segment of the freeway during a precise time interval.

As we have two competing goals (maximize congestion in the box, minimize congestion elsewhere), we apply the multi-objective optimization procedure in Section 4.3. Indeed, we have



(a) Simulation with no metering.

(b) Trade-off: $\alpha = 0.3$

(c) Trade-off: $\alpha = 0.5$

(d) Trade-off: $\alpha = 0.9$

FIGURE 2 Figure 2(a) depicts a space-time diagram of vehicle densities on 19.4 mile stretch of I15 Freeway with no ramp metering. The box objective, and example of *congestion-on-demand*, is applied in Figures 2(b)-2(d). The user specifies a "desired" traffic jam between postmile 4.5 and 14, for a duration of 20 minutes between 8:20 and 8:40. For this, the α parameter enables the proper design of tradeoffs in the objective.

the following two objective functions:

$$f_1(\mathbf{u}, \rho) = -\sum_{(i,k)\in \text{Box}} \rho[i,k] \tag{24}$$

$$\text{and } f_2(\mathbf{u}, \rho) = \sum_{(i,k)\notin \text{Box}} \rho[i,k] \tag{25}$$

To solve this multi-objective problem, we will follow the method described in Section 4.3 and balance our two objectives using a linear combination. As we limit ourselves to one degree of freedom, we introduce a single parameter $\alpha \in [0, 1]$ and minimize the following objective function:

$$J_\alpha(\mathbf{u}, \rho) = \alpha\, f_1(\mathbf{u}, \rho) \,+\, (1-\alpha)\, f_2(\mathbf{u}, \rho)\,, \tag{26}$$

where α is a trade-off parameter: $\alpha = 1$ is complete priority on the congestion inside the box, while $\alpha = 0$ is complete priority on limiting density outside the box.

The results of the box objective are presented in Figures 2(b)-2(d). We give space-time diagrams for three different values of the parameter α. The box of the objective is shown as a black frame with an actual size of 10 miles and 20 minutes. As the trade-off moves from $\alpha = 0.3$ to 0.9, there is a clear increase in the congestion within the box, at the expense of allowing the congestion to spill outside the desired bounds. In fact, Figure 2(d) ($\alpha = 0.9$) activates the bottleneck near the top-left of the box earlier than Figure 2(b) ($\alpha = 0.3$) to congest the middle portion of the box, which leads to a propagation of a congestion wave outside the bottom-right of the box.

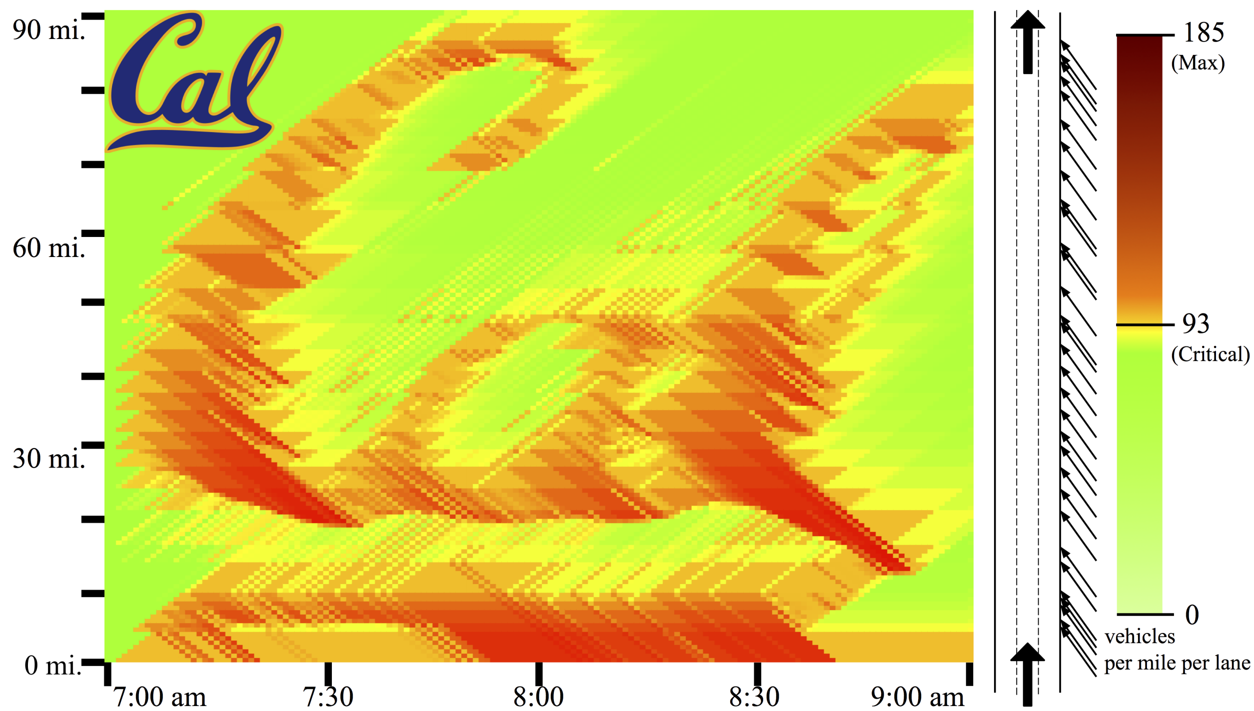


FIGURE 3 Space-time diagram obtained following a *congestion-on-demand* attack with a Cal logo as the objective. The attack was simulated on a 90 miles and 33-onramp freeway, for a 2 hours simulation time and using coordinated ramp metering.

<u>Arbitrary Patterns</u> Any congestion pattern may be created if the network has enough control ramps. Indeed, we can choose the negative and positive coefficients of the congestion-on-demand method carefully to match a desired pattern. We give an example in Figure 3: with the proper ramp metering strategy, we are able to create a space-time diagram resembling the Cal logo. See (30) for a video simulation of the Cal attack.

Attack 2: *catch-me-if-you-can*

We will now show that the use of the multi-objective optimization methods introduced in Section 4.3 can allow the design of even more realistic and hard to define attacks. We will consider the example of a vehicle chase, presented in Section 3.3.2. Some vehicles are pursuing the driver along the freeway, while the driver wishes to escape. This objective is distinct from the *congestion-on-demand* attack, as our desired congestion pattern cannot immediately be imagined beforehand and is highly dependent upon the eventual path of the driver.

This attack cannot easily be translated into a scalar objective function. Therefore, we translate it into a multi-objective problem (see Section 4.3). We can split this attack into four simpler and sometimes conflicting goals, each goal associated with an objective function to minimize:

1. The followers (everyone behind the driver) should cross the freeway section as slowly as possible — Minimizing f_1 will maximize the traffic density of all freeway sections behind the driver's trajectory.

2. In particular, those vehicles directly behind the driver should be impeded with increased priority — Minimizing f_2 will maximize the traffic density difference between the cells of the driver's trajectory and the cells immediately behind.

3. As to not arrouse suspicious from monitoring traffic managers, most other travel times should be reduced — Minimizing f_3 will reduce the total travel time of all the vehicles on the freeway to avoid unneccessary congestion.

4. The driver should quickly exit the freeway — Minimizing f_4 will reduce the driver's travel time, to allow him to cross the freeway as quickly as possible and escape his followers.

We have four objective functions. In practice, presenting the results is clearer with only three functions, and we have chosen to keep only f_1, f_2 and f_3 in this article, as f_4 was not essential for producing interesting results. We will use the linear scalarization technique presented in Section 4.3, and chose three coefficients $a_1, a_2, a_3 \in \mathbb{R}_+$, so that $\sum_{i=1}^{3} a_i = 1$. The objective function we want to optimize is then the following:

$$J(\mathbf{u}, \rho) = \sum_{i=1}^{3} a_i \, f_i(\mathbf{u}, \rho) \tag{27}$$

Implementation

Graphical Representation The space-time diagram in Figure 4 for a 21 miles freeway with 6 adjacent onramps and a 20 minutes simulation time, is an example output of the optimal control scalarization method. Such plots are useful for the DM to discern between "good" and "bad" metering rates. The driver's trajectory is represented in blue, while the trajectory of three pursuants (a, b, c) are also depicted losing ground on the driver.

Ternary Graph The triangle in Figure 4 depicts the chosen set of coefficients a_i. The red dots represents the weighted average of the three corners of an equilateral triangle: the closer the red circle is to the a_i corner, the closer a_i is to 1. This is called a *ternary graph.* The top edge will always be a_1, and the right and left a_2 and a_3 respectively. In this example, we can see that the dominant coefficients are a_1 and a_2. As a consequence, we have an significant congestion behind the driver, forming immediately behind him.

A posteriori Method - Grid Exploration Our approach for the a posteriori method is to automatically "explore the triangle of coefficients" to help the *Decision Maker* find a preferred solution. Figure 5 presents the result of the a posteriori method. We plot the values of each objective function for the optimal solution associated with all sets of a_i coefficients. The lowest values of each f_i are always reached with the highest values of a_i (where f_i has been normalized to take values between 0 and 1; see Section 4.3). Any non-monotonicity in the graphs are attributed to early terminations of the optimizer's gradient descent or convergence to sub-optimal local minima. The conflicting nature of the objectives is apparent. Figure 5(b) shows that f_1 is penalized more by high a_3 values

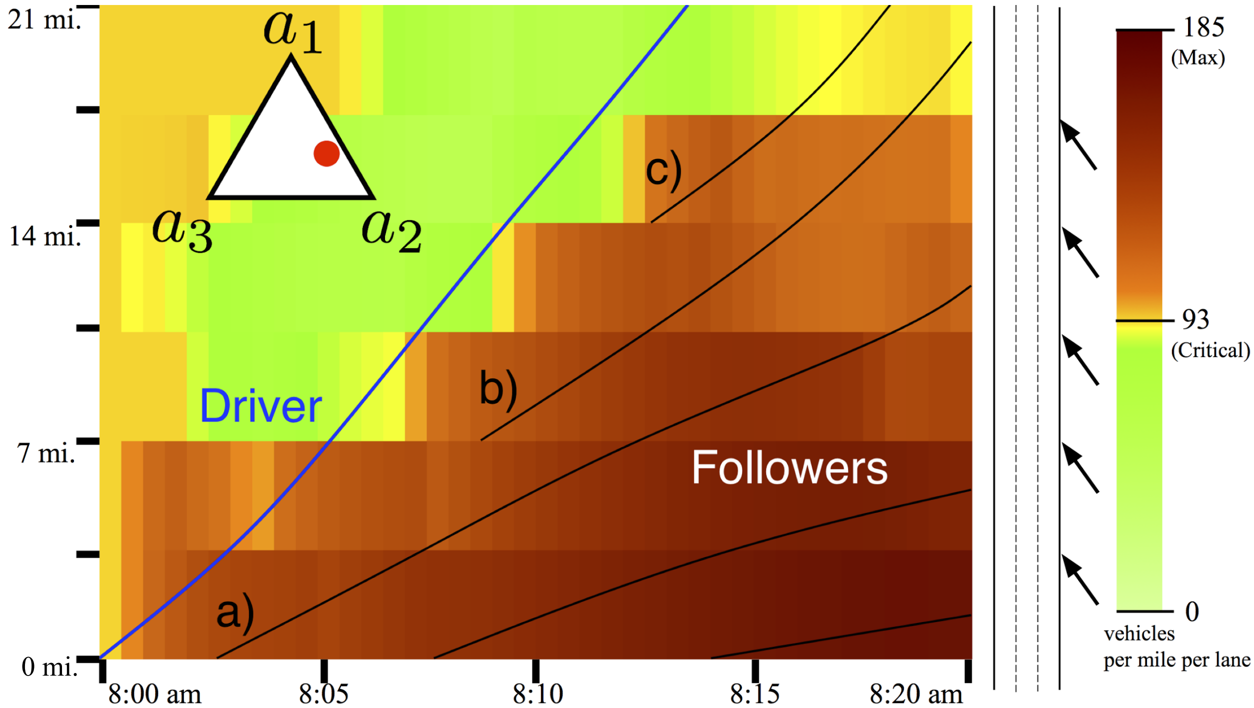


FIGURE 4 Space-time diagram with a ternary graph representing the a_1, a_2, a_3 coefficients (here 30%, 55% and 15% respectively) used for the scalarization process in the catch-me-if-you-can example. The trajectory of the driver (blue line) appears to always gain distance in relation to pursuants further upstream (black lines). Best viewed in color.

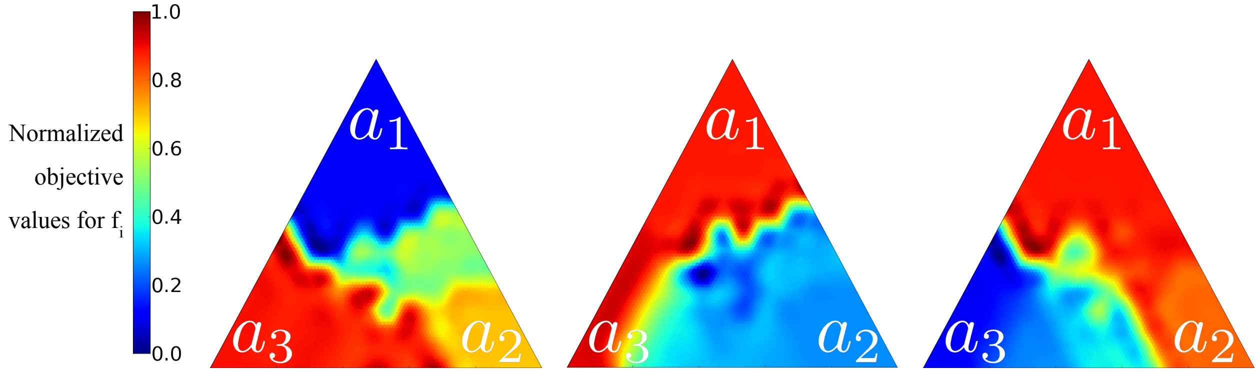
than by high a_2 values, i.e. lowering the total travel time at the expense of congesting the region behind the driver.

The a posteriori method provides the DM with a global overview of the Pareto front, enabling him to immediately locate a desired solution, or at least identify interesting starting points in the Pareto front. For example, Figure 5 gives an indication that the center regions of the triangles have large variations and should be explored further.

Interactive Method A web application (diagram in Figure 6[4] was developed to allow a full exploration of the interactive method. The DM first selects his desired coefficients (a_i) by clicking on the appropriate spot within triangle b). Then, after a scalarization using the particular coefficients and an optimization of the resultant objective, the interface plots the space-time diagram of the resulting simulation in window a), along with the driver's trajectory. Any other vehicle's trajectory can be visualized by clicking at the starting point of the desired trajectory. To enhance the exploration process, the interactive program also chooses two random (but nearby) sets of coefficients and plots their simulation in c1) and c2).

Figure 7 shows an overview of the results obtained while using the interactive interface. The first column shows simulations for the corners of the ternary graph, i.e. only one objective is active at a time. The results are intuitive in that optimizing f_1 (Figure 7.1) produces congestion everywhere behind the driver, optimizing f_2 (Figure 7.2) creates a distinct increase in congestion

[4]Web application demo available at (30)



(b) Scaled values of f_1 (c) Scaled values of f_2 (d) Scaled values of f_3

FIGURE 5 A grid exploration over the ternary graph. An optimization was conducted for a grid of coefficients regularly spaced on the ternary graph. The resulting scalarized objective is decomposed into the constituent objectives (normalized between 0 an 1) and plotted on separate summary ternary graphs.

behind the driver, and optimizing f_3 (Figure 7.3) maintains critical density everywhere, equivalent to maximizing throughput at maximum freeway speeds.

The second column (Figures 7.A-C) shows an interactive shift from favoring f_3 (minimize travel times) to favoring f_2 (trajectory boundary congestion). The shift progressively limits congestion formation, and intelligently removes more congestion *ahead* of the driver, as to not impact the delay of pursuant vehicles.

The last column of Figure 7 demonstrates how the interactive process allows for fine-tuning of the balance of the objectives. Figure 7.a appears to be overly congested in the driver's trajectory. An interactive progression towards lower total travel times concludes with a desirable congestion boundary in Figure 7.c.

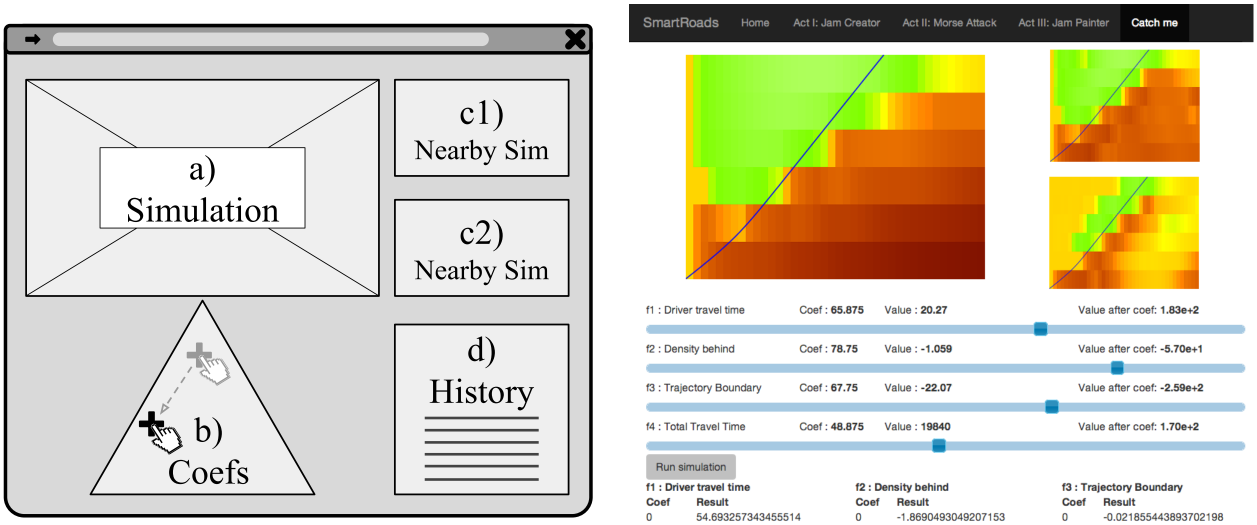
CONCLUSION

This article presents an overview of freeway traffic control systems and their vulnerability to physical and cyber-attacks. The impact of an attack is understood via the response of the control system, with direct attacks on the metering lights being potentially more effective than indirect attacks on the sensing infrastructure. Coordinated ramp metering attacks, being the highest level compromise, are extensively analyzed using methods from the fields of optimal control and multi-objective optimization. Detailed numerical simulations of coordinated ramp metering attacks were conducted to demonstrate the hazards of such compromises and the utility of optimal control tools in not only the hands of traffic managers, but also of adversaries.

As future work, we will develop methods that leverage knowledge of freeway dynamics to detect when a compromise of the traffic control system has occured and how to mitigate the potential harm. For instance, as already demonstrated on water SCADA systems (31), one can detect when sensor readings lie outside those expected given the dynamical assumptions and classify such a sensor as faulty or compromised.

REFERENCES

[1] Jia, Z., C. Chen, B. Coifman, and P. Varaiya, The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, IEEE, 2001, pp. 536–541.



(a) Diagram of web application functionality.

(b) Actual web application (30) created for the purpose of this article, the triangle is replaced by 4 sliders, to match the 4 objective functions

FIGURE 6 Interface of the interactive optimization system used to solve the multi-objective optimization problem to produce the attacks presented in the article.

[2] Papageorgiou, M., H. Hadj-Salem, and J. Blosseville, ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record*, Vol. 1320, 1991, pp. 58–64.

[3] Reilly, J., W. Krichene, M. L. Delle Monache, S. Samaranayake, P. Goatin, and A. Bayen, Adjoint-based optimization on a network of discretized scalar conservation law PDEs with applications to coordinated ramp metering. *Journal of Optimization Theory and Applications (under review)*, 2014.

[4] Miller, M. A. and A. Skabardonis, *San Diego I-15 Integrated Corridor Management (ICM) System: Stage II (Analysis, Modeling, and Simulation*. California PATH Program, Institute of Transportation Studies, University of California at Berkeley, 2010.

[5] Work, D. B., S. Blandin, O. P. Tossavainen, B. Piccoli, and A. M. Bayen, A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress*, Vol. 2010, No. 1, 2010, p. 1.

[6] Jeske, T., Floating car data from smartphones: What google and waze know about you and how hackers can control traffic. *Proc. of the BlackHat Europe*, 2013.

[7] Tufnell, N., Students hack Waze, send in army of traffic bots. *wired.co.uk*, 2014. URL: `http://www.wired.co.uk/news/archive/2014-03/25/waze-hacked-fake-traffic-jam`.

[8] Zetter, K., Hackers Can Mess With Traffic Lights to Jam Roads and Reroute Cars. *wired.com*, 2014.

[9] Codenomicon, *The Heartbleed Bug*, 2014. URL: `www.heartbleed.com`.

[10] Grad, S., Engineers who hacked into L.A. traffic signal computer, jamming streets, sentenced. *Los Angeles Times*, 2009.

[11] Muralidharan, A. and R. Horowitz, Optimal control of freeway networks based on the link node cell transmission model. In *American Control Conference (ACC)*, IEEE, 2012, pp. 5769–5774.

[12] AASHTO, ITE, and NEMA, *Model 2070 Controller Standard Version 03*, 2012.

[13] FHWA, *Type 170 Traffic Signal Controller System - Microcomputer Based Intersection Controller*. Federal Highway Administration, 1978.

[14] *TomTom*, 2014. URL: `http://www.tomtom.com/`.

[15] Sutton, J., Copper wire stolen from traffic signal, street lights in Oklahoma City. *The Oklahoman*, 2014.

[16] Rosenberg, M., Underground copper wire heist causes San Jose freeway flood. *San Jose Mercury News*, 2014.

[17] *SmartRoads Video Trailer*, 2014. URL: `https://www.youtube.com/watch?v=aWI3WK2STJc`.

[18] Lighthill, M. J. and G. B. Whitham, On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, Vol. 229, No. 1178, 1955, p. 317.

[19] Richards, P., Shock waves on the highway. *Operations research*, Vol. 4, No. 1, 1956, pp. 42–51.

[20] 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.

[21] Delle Monache, M. L., J. Reilly, S. Samaranayake, W. Krichene, P. Goatin, and A. M. Bayen, A PDE-ODE model for a junction with ramp buffer. *SIAM Journal on Applied Mathematics*, Vol. 74, No. 1, 2014, pp. 22–39.

[22] Godunov, S. K., A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, Vol. 89, No. 3, 1959, pp. 271–306.

[23] Lebacque, J. P., The Godunov scheme and what it means for first order traffic flow models. In *Internaional symposium on transportation and traffic theory*, 1996, pp. 647–677.

[24] Gugat, M., M. Herty, A. Klar, and G. Leugering, Optimal Control for Traffic Flow Networks. *Journal of Optimization Theory and Applications*, Vol. 126, No. 3, 2005, pp. 589–616. doi:10.1007/s10957-005-5499-z.

[25] Giles, M. B. and S. Ulbrich, Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions. *SIAM Journal on Numerical Analysis*, Vol. 48, No. 3, 2010, pp. 905–921.

[26] Ulbrich, S., A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM journal on control and optimization*, Vol. 41, No. 3, 2002, pp. 740–797.

[27] Giles, M. B. and N. A. Pierce, An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion*, Vol. 65, No. 3-4, 2000, pp. 393–415. doi:10.1023/A:1011430410075.

[28] Dervisoglu, G., A. Kurzhanskiy, G. Gomes, and R. Horowitz, Macroscopic freeway model calibration with partially observed data, a case study. In *American Control Conference (ACC), 2014*, IEEE, 2014, pp. 3096–3103.

[29] Muralidharan, A. and R. Horowitz, Imputation of Ramp Flow Data for Freeway Traffic Simulation. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2099, No. -1, 2009, pp. 58–64.

[30] Reilly, J. and S. Martin, *SmartRoads Website*, 2014. URL: `http://traffic.berkeley.edu/smartroads`.

[31] Amin, S., X. Litrico, S. Sastry, and A. M. Bayen, Cyber security of water scada systemspart I: analysis and experimentation of stealthy deception attacks. *Control Systems Technology, IEEE Transactions on*, Vol. 21, No. 5, 2013, pp. 1963–1970.

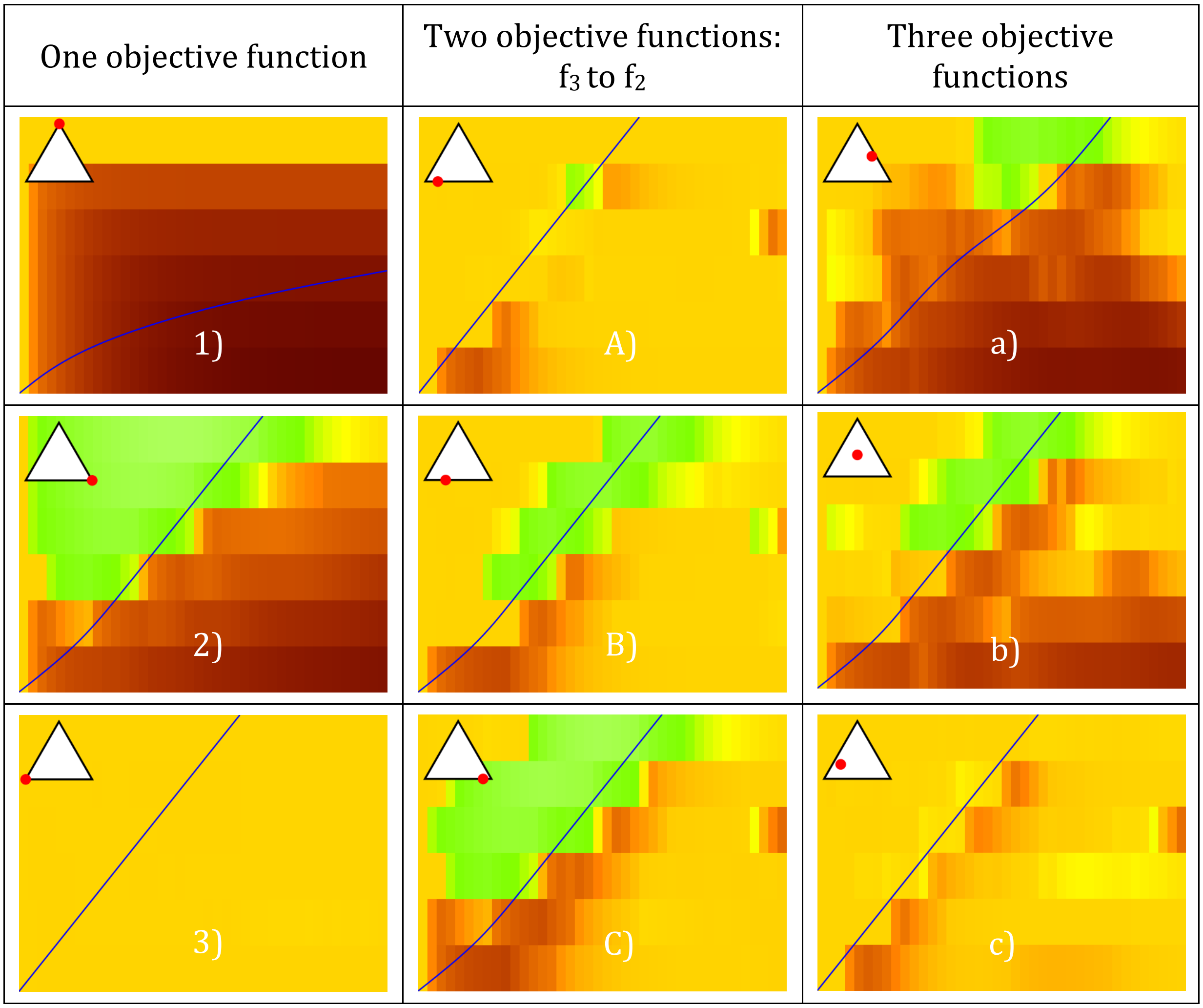


FIGURE 7 **Summary of *catch-me-if-you-can* simulations generated via the interactive method. Column 1 shows optimizations over individual objectives. Column 2 shows a transititon from favoring f_3 to favoring f_2. Column 3 shows a progression across all three objectives.**