Stealthy Attack Detection of controlled ramp meters in freeway networks

Charalambos Menelaou, Kangkang Zhang, Stelios Timotheou, Christos G.Panayiotou, and Thomas Parisini

***Abstract*— The increasing reliance on intelligent transportation systems (ITS) for traffic management has simultaneously heightened the potential for cybersecurity threats. Malicious cyber attacks on such systems can lead to operational inefficiencies, heightened congestion, and compromised safety. This work introduces a stealthy integrity attack tailored for ramp metered freeway systems. The attack model is conceptualized as a closed-loop dynamical system, formulated as an optimization problem based on the Cell Transmission Model. This work further proposes a distributed detection mechanism designed to identify attack residuals, which, by incorporating an adaptive threshold scheme, can detect the presence of an attack. To demonstrate the efficacy of our detection scheme, we present a scenario illustrating its application in freeway road networks.**

I. Introduction

The integration of Information and Communication Technology advancement into intelligent transportation system enables the development of traffic management strategies. The majority of them offers a promising avenue for addressing congestion issues, as evidenced by the plethora of solutions found in the literature [1]. These solutions include system-level controls, such as road-based traffic management schemes, and vehicle-level controls, like vehicle-based traffic management systems, highlighting the multifaceted approach needed to manage large-scale road transportation systems [2].

The advent of intelligent transportation systems (ITS) has transformed traffic management strategies, offering unprecedented control over traffic flow and efficiency. Despite the plethora of solution approaches and the advancements in traffic management strategies and their demonstrated successes in curbing congestion, the literature reveals a significant gap in addressing the vulnerability of these systems to malicious attacks. However, the vast growth of traffic management systems has introduced new vulnerabilities, particularly to cyber threats that aim to disrupt system operations. On that account the study by [3] uniquely identifies the potential for high-level attacks within freeway networks but falls short of offering concrete solutions for such threats. Moreover, the focus within the literature of cybersecurity has predominantly been on connected vehicles, with limited exploration into the broader implications of attacks on traffic management infrastructures [4], [5]. This oversight in current research highlights a pressing need for a comprehensive investigation into the security of traffic management systems. This work aims to bridge this gap by focusing on the identification and mitigation of potential cyber threats to ramp metering operations within freeway networks. Specifically, it examines scenarios where attackers target ramp meters to create congestion, thereby affecting the overall network performance. The goal is to develop a resilient framework capable of detecting and countering such stealthy attacks, ensuring the continued efficacy, safety, and reliability of transportation networks.

Ramp metering emerges as a key strategy within this context, aimed at enhancing freeway throughput by managing the entry rate of vehicles from on-ramps to the freeway mainline. This method operates under the premise of admission control and utilizes traffic-responsive strategies, which rely on real-time data to orchestrate metering actions. The effectiveness of ramp metering, whether through uncoordinated or coordinated approaches, is well-documented, with studies demonstrating its capability to manage the flow from single or multiple consecutive on-ramps effectively [6], [7], [8]. A notable implementation of coordinated ramp metering on the Monash Freeway in Melbourne, Australia, has showcased the potential of such strategies to significantly reduce travel times and congestion [9], [10].

In addressing traffic congestion within freeway networks, Model Predictive Control (MPC) strategies, underpinned by the Cell Transmission Model (CTM), have gained prominence. MPC's forward-looking nature allows for the optimization of current traffic states by accounting for future traffic flow projections, making it particularly suited for tasks such as the coordination of ramp metering actions [11], [12], [13]. The work presented in [13] further elucidates that under specific conditions, the challenges associated with non-convex nonlinear MPC problems can be circumvented through approximations, thus broadening the applicability and effectiveness of MPC in traffic management within a real-time framework.

However a transportation system is a typical cyber-physical system which integrates computation and communication cyber technologies and control techniques in the physical system. Such a highly complex integration not

This work is supported by the European Union (i. ERC, xURANUS, No. 101088124, and ii. Horizon 2020 Teaming, KIOS CoE, No. 739551), and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Strategy. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

C. Menelaou, S. Timotheou and C.G. Panayiotou, are with the KIOS CoE, and the Dept. of Electrical and Computer Engineering, University of Cyprus. (cmenel02@ucy.ac.cy, timotheou.stelios@ucy.ac.cy, and christosp@ucy.ac.cy)

K. Zhang and T. Parisini are with the Dept. of Electrical and Electronic Engineering, Imperial College London. T. Parisini is also with the Dept. of Engineering and Architecture, University of Trieste. (kzhang5@ic.ac.uk and t.parisini@gmail.com)

only vulnerable to traditional safety issues such as physical component faults, but also to cyber security challenges such as the cyber attacks studied in the more recent literature (see, e.g., [14], [15], [16], [17]). From a control engineering perspective, cyber attacks are divided into two main classes: denial of service (DoS) attacks and integrity (or deception) attacks. DoS attacks compromise the *availability* of data communication networks but are inevitably exposed to anomaly detectors [15]. *Integrity attacks* compromise the *integrity* of communication networks by injecting false data, but at the same time, keep deception to anomaly detectors. As the most common type of attacks, integrity attacks are highly destructive and challenging to detect, and are the target attack type in this paper.

To the best of our knowledge, in the first work that develops a comprehensive approach to address stealthy cyber attacks within ramp metering systems, a crucial component of freeway network traffic management. Unlike previous studies, which predominantly focus on the optimization of traffic flow and congestion reduction through intelligent control strategies, this work aims in identifying and mitigating stealthy attacks targeting these ramp-metering controlled systems. Specifically, the introduction of a stealthy attack scenario, designed to exploit the operational dynamics of ramp metering, underscores a potential critical risk in intelligent transportation systems. Furthermore, the development of a distributed detection scheme opposes a significant advancement towards enhancing the resilience of traffic management infrastructures against stealthy attacks. This dual focus not only bridges a vital gap in the existing literature but also sets a new precedent for the integration of attack detection mechanisms into the design and operation of intelligent traffic management solutions. In this regard this work contributions are:

1) The proposing of a novel stealthy attack scenario tailored for ramp metering systems within freeway networks. This attack scenario is designed to be covert, allowing it to bypass traditional security measures and disrupt the traffic flow without immediate detection.
2) In response to the stealthy attack, this work develop a distributed detection mechanism specifically designed to identify and counteract the *attack*. By employing a distributed approach, our system ensures robustness and resilience, significantly reducing the likelihood of a successful attack by quickly isolating and mitigating its effects.

Section II presents the modeling framework used in this work, while Section III provides the mathematical formulation and an efficient solution approach for the ramp metering problem. Next, Section IV describes the scenario where an attacker aims to perform an "intelligent" attack scheme by inducing congestion within a specific cell. Sections V and VI propose the attack detection methodology, based on a bank of distributed observers and rigorously characterizing the class of attacks that can be detected by the detection methodology, respectively. Subsequently, Section VII evaluates the proposed detection methodology, and finally, Section VIII concludes this work and explores potential avenues for future research.

II. Asymmetric Cell Transmission model

The Cell Transmission Model (CTM) [18], is the discrete analog of the well-known Lighthill-Whitham-Richards (LWR) continuum flow model, which established for its simplicity, widespread use, and flexibility in modeling traffic dynamics on freeway networks. As a first-order macroscopic model, the CTM utilizes a fundamental diagram (FD) to represent traffic dynamics along a freeway network by dividing the road into homogeneous cells $i \in \mathcal{I} = \{1, \dots, I\}$, each of equal length L. The time dimension is discretized into time-steps of duration T_s, and the traffic dynamics within each cell $i \in \mathcal{I}$ are mathematically described by the FD using the macroscopic parameters of the free-flow speed v_i^f (km/h), the backward wave propagation speed w_i (km/h), the capacity Q_i (veh/h), and the jam density ρ_i^j. The free-flow speed represents the average speed at which vehicles travel through congestion-free cells, whereas the backward wave propagation speed denotes the speed at which congestion queues propagate upstream in congested cells. Moreover, the capacity of cell $i \in \mathcal{I}$ is defined as the maximum flow that can be accommodated, while the jam density is the maximum number of vehicles that can be queued within a cell.

In this work a modification of CTM is considered [13], the asymmetric CTM (ACTM) which accounts for nonhomogeneous cells. Contrary to the homogeneous cells in the traditional CTM, each cell $i \in \mathcal{I}$ has length L_i, such that, $L_i \geq v_f T_s$.This model permits each cell to incorporate at most one onramp and one offramp, with all onramps being metered and located upstream of any offramp, as illustrated in Figure 1. The cells are numbered from 1 to I, beginning with the upstream-most section, and the time is discretized into K time-steps ($k \in \mathcal{K} = \{1, \dots, K\}$). Traffic dynamics in ACTM follow a trapezoidal-shaped Fundamental Diagram (FD), outlining three regimes: a) free-flow, b) maximum capacity, and c) congestion. The transitions between these regimes are distinct by critical densities $\rho_i^{C_1}$ and $\rho_i^{C_2}$, where $\rho_i^{C_1} = Q_i/v_i^f$ and $\rho_i^{C_2} = \rho_i^J - Q_i/w_i$, as shown in Figure 2. In ACTM, the flow variables $q_i(k)$, $r_i(k)$, and $s_i(k)$ (veh/h) are utilized to represent the mainstream outflow, the metered onramp flow, and the offramp flow of cell i during the closed interval $[kT_s, (k+1)T_s)$, respectively. Specifically, $q_i(k)$ denotes the number of vehicles leaving cell i, $r_i(k)$ denotes the metered flow of vehicles entering cell i from an onramp, and $s_i(k)$ refers to the flow directed towards an offramp within the specified time interval. Furthermore, the parameters $q_0(k)$ and $r_i^s(k)$, for all cells $i \in \mathcal{I}$, signify the supply demands for the mainstream and the onramps, respectively. The offramp flow is related to the mainstream outflow by a known split ratio $\beta_i(k) \in [0, 1]$, such that,

$$s_i(k) = \beta_i(k)\big(s_i(k) + q_i(k)\big) \equiv \frac{\beta_i(k)}{1 - \beta_i(k)} q_i(k). \qquad (1)$$

Additionally, the variables $\rho_i(k)$ (veh/km) and $l_i(k)$ (veh)

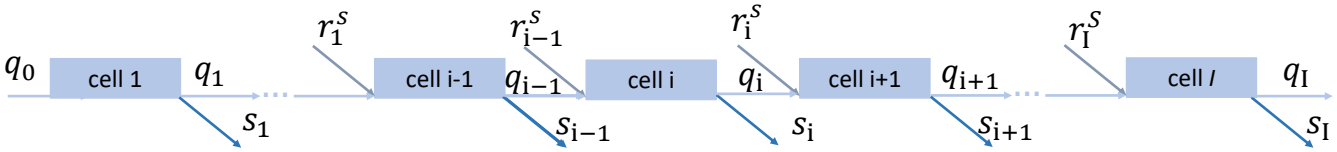

Fig. 1. A typical network devided into I cells.

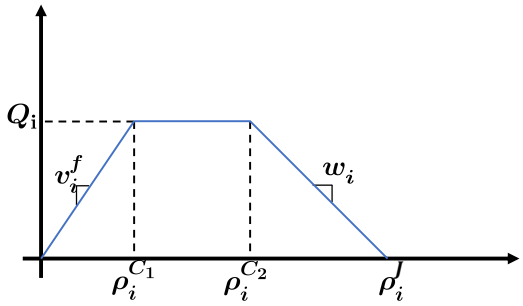

Fig. 2. A typical flow-density fundamental diagram.

represent the density of vehicles within cell i and the number of vehicles queued at the onramp of cell i, respectively. The mainstream flow is calculated as the minimum of what can be sent by the upstream section assuming maximum speed and what can be received by the downstream section, such that

$$\begin{aligned} &q_i(\rho_i(k), \rho_{i+1}(k), k) \\ &= \min\Big(\big(1-\beta_i(k)\big)v_i^f \rho_i(k), w_{i+1}\big(\rho_{i+1}^J - \rho_{i+1}(k)\big) \\ &\quad - r_{i+1}(k), Q_i, Q_{i+1} - r_{i+1}(k), \big((1-\beta_i(k))/\beta_i(k)\big)Q_i\Big). \end{aligned} \tag{2}$$

Notably, for the terminal cell, identified as $i = I$, upstream traffic constraints are not applicable due to its position at the end of the freeway network (see Figure 1). Furthermore, the flow of vehicles from metered onramps, denoted as $r_i(k)$, is subject to several constraints. These include the physical capacity of the mainstream, the instantaneous supply and demand at the onramp, and the optimal control decisions encapsulated by the metered demand $r_i^*(k)$. The onramp flow $r_i(k)$ is governed by

$$\begin{aligned} &r_i(\rho_i(k), k) \\ &= \min\left(r_i^s(k) + \frac{1}{T_s} l_i(k), Q_i, w_i\big(\rho_i^J - \rho_i(k)\big), r_i^*(k)\right). \end{aligned} \tag{3}$$

The conservation of vehicle densities within the freeway, for all cells $i \in \mathcal{I}$ at discrete time steps $k \in \mathcal{K}$, is crucial for modeling the evolution of traffic patterns. This is mathematically defined as

$$\rho_i(k+1) = \rho_i(k) + \frac{T_s}{L_i}\big(q_{i-1}(k) - q_i(k)\big) + \eta_i(k), \tag{4}$$

where $q_{i-1}(k)$ represents the flow into cell i from its upstream neighbor; for the first cell in the system, $i = 1$, $q_0(k)$ represents the entry flow into the freeway. The term $\eta_i(k)$ is introduced to model external disturbances impacting the traffic flow, such as sensor inaccuracies, weather conditions, or other unforeseen events that could influence the density and flow within each cell. To facilitate the modeling of these disturbances within the traffic flow dynamics, we have the assumption below.

Assumption 1: There exists a known positive scalar function $\bar{\eta}_i(k)$ such that

$$|\eta_i(k)| \le \bar{\eta}_i(k), \ \forall\ k \in \mathcal{K}. \tag{5}$$

Additionally, since the onramps, rates can be controlled to capture the interplay between supply, demand, and actual ramp flow, the length of onramp queues can be defined as

$$l_i(k+1) = l_i(k) + T_s\big(r_i^s(k) - r_i(k)\big). \tag{6}$$

III. Ramp Metering Problem

Ramp metering is a critical control strategy aimed at optimizing freeway traffic flow by regulating the rate at which vehicles are allowed to enter the mainstream. This is achieved with the goal of minimizing the Total Travel Time (TTT) for all vehicles across the network. Practically, ramp metering is realized through traffic lights installed at the beginning of each onramp.

A. Objective Function

The objective of minimizing TTT aggregates the total number of vehicles across all cells and onramp queues for every time step. Mathematically, the TTT objective function, J_{TTT}, is defined as:

$$J_{TTT} = \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} \Big(L_i \rho_i(k) + l_i(k)\Big). \tag{7}$$

B. Ramp Metering Control Problem Formulation

Despite the diversity of approaches available for formulating and solving the ramp metering problem, the Model Predictive Control (MPC) approach is distinguished for its efficiency and practicality. The MPC problem is addressed at intervals of $M = N^C/T_s$ time-steps, where N^C represents the control horizon. Specifically, at the time-step $k = M(p-1)$, the current measured traffic states $\bar{\rho}_i(k)$ and the supply demands $q_0(k)$ and $r_i^s(k)$ are utilized to solve the p-th MPC optimization problem over the time horizon $\mathcal{K}_p = \{M(p-1)+1, \ldots, M(p-1)+N^P\}$, with N^P denoting the prediction horizon and satisfying $N^P \ge N^C$. The optimization aims to determine the optimal metered

flows for each onramp, $r_i^*(k)$ for $i \in \mathcal{I}$, to minimize the Total Travel Time (TTT) metric, by solving the problem:

$$(\mathrm{P}_1) \quad \min \; J_{TTT}^{MPC}(p) = \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_p} \Big(L_i \rho_i(k) + l_i(k) \Big) \tag{8a}$$

s.t. Freeway dynamics: (1) − (3) and (6),

$$\rho_i(k+1) = \rho_i(k) + \frac{T_s}{L_i}\big(q_{i-1}(k) - q_i(k)\big), \forall k \in \mathcal{K}_p \tag{8b}$$

$$\text{Initialization: } \rho_i(k) = \bar{\rho}_i(k), k = M(p-1). \tag{8c}$$

Note that constraint (8a) adapts the cell densities dynamics to current traffic conditions, integrating predictive control to anticipate mainstream density disturbances. The decision variable of the Problem (P_1) is the optimal ramp meters $r_i^*(k)$ for each $i \in \mathcal{I}$. Finally, Problem (P_1) is a Nonlinear Program (NLP) due to the discontinuous functions within constraints (2) and (3).

C. Ramp Metering Problem Solution Approach

Despite the original ramp metering formulation of Problem P_1 being a nonlinear program, work in [13] has demonstrated that the nonlinear constraints of (2) and (3) can be relaxed into a set of linear inequalities, thus simplifying the problem without compromising its integrity. Accordingly, (2) is relaxed into the following five linear inequalities:

$$q_i(k) \le \big(1 - \beta_i(k)\big) v_i^f \rho_i(k), \tag{9a}$$

$$q_i(k) \le w_{i+1}\big(\rho_{i+1}^J - \rho_{i+1}(k)\big) - r_{i+1}(k), \tag{9b}$$

$$q_i(k) \le Q_i, \tag{9c}$$

$$q_i(k) \le Q_{i+1} - r_{i+1}(k), \tag{9d}$$

$$q_i(k) \le \big((1-\beta_i)/\beta_i\big) Q_i. \tag{9e}$$

To relax the discontinuous function of (3), we set the ramp metering rate equal with the obtained control decisions, i.e., $r_i^*(k) \equiv r_i(k)$, and replace (3) with the following three linear inequalities:

$$r_i(k) \le r_i^s(k) + \frac{1}{T_s} l_i(k), \tag{10a}$$

$$r_i(k) \le Q_i, \tag{10b}$$

$$r_i(k) \le w_i\big(\rho_i^J - \rho_i(k)\big). \tag{10c}$$

Hence, a feasible solution to Problem P_1 to minimize the Total Travel Time (TTT) metric, by solving the problem:

$$(\mathrm{P}_2) \quad \min \; J_{TTT}^{MPC}(p) = \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_p} \Big(L_i \rho_i(k) + l_i(k) \Big) \tag{11a}$$

s.t. Freeway dynamics: (1), (6), (9) − (10) and (8b)

$$\text{Initialization: } \rho_i(k) = \bar{\rho}_i(k), k = M(p-1). \tag{11b}$$

Similarly as above the decision variable of the Problem (P_1) is the optimal ramp meters $r_i^*(k)$ for each $i \in \mathcal{I}$.

IV. Stealthy Attack For Ramp Metering Problem

In this section, we explore a type of covert attack targeting specific cells of the underlying freeway system. Such an "intelligent" attack can cause congestion within the targeted cell (this cell is referred to as the attacked cell hereafter), but at the same time, is stealthy with respect to the sensor measurements of this cell. To this end, the background is set below. In this work, we consider the case that only one cell is under attacked with the attack event starting at the time-step k_0 where $k_0 \in \mathcal{K}$. Then, the *disclosure, disruption resources and system knowledge* defined in [15] available to the attacker is given below.

- *Disclosure Resources.* For the attacked cell $i^a \in \mathcal{I}^a$, the attacker has estimates of the demands of the mainstream and onramps $q_0(k)$ and $r_i^s(k)$ for any $k \ge k_0$. Also, the attacker can read the sensor measurement of the attacked cell $\rho_{i^a}(k)$ and the control signal $r_{i^a}^*(k)$ of the upstream onramp traffic lights for any $k \ge k_0$.
- *Disruption Resources.* The attacker is able to disrupt the demand for affecting the upstream onramp traffic lights (denoted by $\bar{r}_{i^a}^*$) and the transmitted sensor measurement $\bar{\rho}_{i^a}$ of the attacked cell. In particular, the attacker can inject the false data $a_{i^a}^r(k)$ into $r_{i^a}^*(k)$ and the false data $a_{i^a}^\rho(k)$ into the transmitted sensor measurement $\bar{\rho}_{i^a}(k)$, i.e.,

$$\bar{r}_{i^a}^*(k) = r_{i^a}^*(k) + a_{i^a}^r(k), \tag{12a}$$

$$\bar{\rho}_{i^a-1}(k) = \rho_{i^a-1}(k) + a_{i^a-1}^\rho(k), \quad \forall\, k \ge k_0. \tag{12b}$$

- *System Knowledge.* The attacker has perfect knowledge of the freeway dynamical system (4)-(6), and expressions in Eqs. (1)-(3).

Note that, in the absence of attacks the nominal freeway system and queue length dynamics are given as follows

$$\rho_i^n(k+1) = \rho_i^n(k) + \eta_i(k) + \frac{T_s}{L_i}\big(q_{i-1}(\rho_{i-1}^n, \rho_i^n, k) - q_i(\rho_i^n, \rho_{i+1}^n, k)\big), \tag{13}$$

$$l_i^n(k+1) = l_i^n(k) + T_s\big(r_i^s(k) - r_i^n(k)\big), \tag{14}$$

where ρ_i^n and l_i^n are the vehicle density and the length of the onramp in the nominal case, corresponding to ρ_i and l_i, respectively.

We proceed with presenting the attack model. An "intelligent" stealthy attack can be implemented through an MPC scheme where the attacker has access to the true state of the network. To achieve such an attack, the attacker maintains its own traffic model used within an MPC scheme to derive the attack signals $a_i^\rho(k)$ and $a_i^r(k)$, $\forall k \in \mathcal{K}$. In this case, variables $\breve{\rho}_i(k)$ and $\breve{l}_i(k)$ denote the state of freeway density and onramp queue length that the attacker maintains. Hence, the attacker models the dynamics of the considered freeway system as

$$\breve{\rho}_i(k+1) = \breve{\rho}_i(k) + \frac{T_s}{L_i}\left(q_{i-1}(k) - q_i(k)\right), \tag{15}$$

$$\breve{l}_i(k+1) = \breve{l}_i(k) + T_s\left(r_i^s(k) - r_i(k)\right), \tag{16}$$

where the variables $q_i(k)$ and $r_i(k)$, $\forall k \in \mathcal{K}$, are defined according to Eqs. (1) - (3). By modelling the true dynamics of the freeway system, the goal of the "intelligent" attack is to maximize the density at the attacked cell while minimizing it elsewhere. This involves the introduction of the coefficient $\zeta_i(k)$ within the attacker's objective function such that:

$$J_{\text{ATTACK}} = \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} \Big(\zeta_i(k) L_i \breve{\rho}_i(k) + \breve{l}_i(k)\Big). \tag{17}$$

where a positive coefficient $\zeta_i(k)$ aims to reduce the density at a cell, thereby mitigating congestion, while conversely, a negative coefficient $\zeta_i(k)$ encourages congestion within the attacked cell by increasing the experienced TTT. Hence, the attack signals $a_i^{\rho}(k)$ and $a_i^{r}(k)$, $\forall k \in \mathcal{K}$, can be derived by solving the following problem:

$$(\mathrm{P}_3) \quad \min \; J^{MPC}_{ATTACK}(p) = \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}_p} \Big(\zeta_i(k) L_i \breve{\rho}_i(k) + \breve{l}_i(k)\Big) \tag{18a}$$

s.t. Freeway dynamics: $(1), (9)-(10), (15)$ and (16),

$$a^{\rho}_{i^a}(k) = \breve{\rho}_{i^a}(k) - \rho_{i^a}(k), k \in \mathcal{K}_p \tag{18b}$$

$$a^{r}_{i^a}(k) = r_{i^a}(k) - r^*_{i^a}(k), k \in \mathcal{K}_p \tag{18c}$$

$$\text{Initialization: } \breve{\rho}_i(k) = \bar{\rho}_i(k) - a_i^{\rho}(k),\ k = M(p-1),$$
$$\zeta_i(k), \rho_i(k), r_i^*(k)\, \forall k \in \mathcal{K}_p. \tag{18d}$$

V. Distributed Attack Detection Scheme

In this section, we present an attack detection methodology, based on a bank of distributed observers as shown in Fig. 3, to detect the underlying covert attacks. The i-th local observer produces an estimate $\hat{\rho}_i(k)$ of the density $\rho_i(k)$ for the i-th cell ($i \in \mathcal{I}$), using the information $\hat{\rho}_{i-1}$ and $\hat{\rho}_{i+1}$ provided by the observers from its neighbour.

The idea of using the distributed methodology, inspired by [17], is that the attack occurring on the onramp of the h-th cell can backwards propagate congestion to the i-th cell. The deviation between the true density ρ_i and the density ρ_i^n in the nominal case does not equal zero or close to zero. Hence, the i-th local anomaly detector for the i-th cell can detect such a deviation, thereby revealing the cover attack occurring at the h-th cell. The details are given in the sequel. To this end, we suppose that the communication between neighbouring cells is not disrupted by the attacker and external disturbances.

Considering the dynamical system (4), the distributed state observer and the corresponding detection residual for the i-th cell ($i \in \mathcal{I}$) is designed as

$$\hat{\rho}_i(k+1) = \hat{\rho}_i(k) + \gamma_i\big(\bar{\rho}_i(k) - \hat{\rho}_i(k)\big) + \frac{T_s}{L_i}\big(q_{i-1}(\hat{\rho}_{i-1}, \hat{\rho}_i, k) - q_i(\hat{\rho}_i, \hat{\rho}_{i+1}, k)\big), \tag{19a}$$

$$\epsilon_i(k) = \bar{\rho}_i(k) - \hat{\rho}_i(k), \tag{19b}$$

where γ_i is the observer gain, such that $\tilde{\gamma}_i = 1 - \gamma_i$ satisfies $|\tilde{\gamma}_i| < 1$. The $q_i(\hat{\rho}_i, \hat{\rho}_{i+1}, k)$ are constructed based on the estimation $\hat{\rho}_i(k)$ and $\hat{\rho}_{i+1}(k)$ where $\hat{\rho}_{i+1}(k)$ is the estimate of ρ_{i+1} generated by the observer of the $i+1$-th cell. Specifically, $q_i(\hat{\rho}_i, \hat{\rho}_{i+1}, k)$ are obtained by replacing $\rho_i(k)$ and $\rho_{i+1}(k)$ in Eq. (2) with $\hat{\rho}_i(k)$ and $\hat{\rho}_{i+1}(k)$, respectively. Note that from (2), we have

$$\begin{aligned} &|q_i(\rho_{i-1}, \rho_i) - q_i(\hat{\rho}_{i-1}, \hat{\rho}_i)| \\ &\le \alpha_{i,1}|\rho_{i-1} - \hat{\rho}_{i-1}| + \alpha_{i,2}|\rho_i - \hat{\rho}_i|, \end{aligned} \tag{20}$$

where $\alpha_{i,1} = \alpha_{i,2} = v_f$. The initial condition is chosen as $\rho_i(0) = 0$. By this selection, there exists a positive scalar $\sigma_{0,i}$ such that $|\rho_i(0) - \hat{\rho}_i(0)| \le \sigma_{0,i}$ for any $i \in \mathcal{I}$. Then, the boundedness property of the observer (19) is presented in the following theorem.

Theorem 1: Consider the system (4) satisfying Assumption 1, and the observer (19). In the nominal scenario, the state estimation error $\rho_i(k) - \hat{\rho}_i(k)$ is bounded for all $i \in \mathcal{I}$ and $k \in \mathcal{K}$. Moreover, there exists a nonnegative scalar function $\bar{\epsilon}_i(k)$ such that the residual $\epsilon_i(k)$ satisfies

$$|\epsilon_i(k)| \le \bar{\epsilon}_i(k),\ \forall\ i \in \mathcal{I},\ \forall\ k \in \mathcal{K}, \tag{21}$$

where $\bar{\epsilon}_i(k)$ is given by

$$\begin{aligned} \bar{\epsilon}_i(k) \triangleq\ & (\tilde{\gamma}_i)^k \sigma_{i,0} + \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} \bar{\eta}_i(j) \\ &+ \frac{T_s}{L_i} \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} \big(\alpha_{i-1,1}\bar{\epsilon}_{i-1}(j) \\ &+ (\alpha_{i-1,2} + \alpha_{i,1})\bar{\epsilon}_i(j) + \alpha_{i,2}\bar{\epsilon}_{i+1}(j)\big). \end{aligned} \tag{22}$$

Proof: Note that in the nominal case, $\bar{\rho}_i(k) = \rho_i(k)$. Let $\tilde{\rho}_i(k)$ denote the deviation between $\rho_i(k)$ and $\hat{\rho}_i(k)$, i.e., $\tilde{\rho}_i(k) \triangleq \rho_i(k) - \hat{\rho}_i(k)$. Then, based on (4) and (19), the error system is obtained as follows:

$$\begin{aligned} \tilde{\rho}_i(k+1) =\ & \tilde{\gamma}_i \tilde{\rho}_i(k) + \eta_i(k) \\ &+ \frac{T_s}{L_i}(q_{i-1}(\rho_{i-1}, \rho_i, k) - q_{i-1}(\hat{\rho}_{i-1}, \hat{\rho}_i, k)) \\ &- \frac{T_s}{L_i}(q_i(\rho_i, \rho_{i+1}, k) - q_i(\hat{\rho}_i, \hat{\rho}_{i+1}, k)), \end{aligned} \tag{23a}$$

$$\epsilon_i(k) = \tilde{\rho}_i(k). \tag{23b}$$

Since $|\tilde{\gamma}_i| < 1$, the system (23) has the bounded-input-bounded-output property. From Assumption 1, $\eta_i(k)$ is bounded, and thus, $\tilde{\rho}_i(k)$ is bounded. Moreover, ϵ_i can be explicitly written as

$$\begin{aligned} \epsilon_i(k) = \tilde{\rho}_i(k) =\ & (\tilde{\gamma}_i)^k \tilde{\rho}_i(0) + \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} \eta_i(j) \\ &+ \frac{T_s}{L_i} \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} (q_{i-1}(\rho_{i-1}, \rho_i, k) - q_{i-1}(\hat{\rho}_{i-1}, \hat{\rho}_i, k)) \\ &- \frac{T_s}{L_i} \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} (q_i(\rho_i, \rho_{i+1}, k) - q_i(\hat{\rho}_i, \hat{\rho}_{i+1}, k)) \end{aligned}$$

Based on the triangle inequality, from Assumption 1 and (26), we have

$$|\epsilon_i(k)| \le (\tilde{\gamma}_i)^k \sigma_{i,0} + \sum_{j=0}^{k-1} (\tilde{\gamma}_i)^{k-1-j} \bar{\eta}_i(j)$$

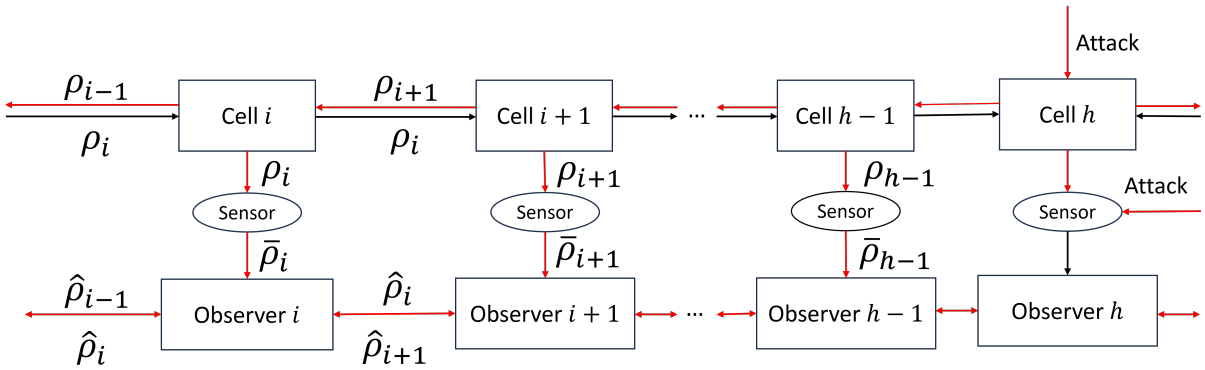

Fig. 3. A typical cyber-physical system with attacks and faults.

$$
\begin{aligned}
&+\frac{T_s}{L_i}\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}(\alpha_{i-1,1}|\tilde{\rho}_{i-1}(j)|+\alpha_{i-1,2}|\tilde{\rho}_i(j)|)\\
&+\frac{T_s}{L_i}\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}(\alpha_{i,1}|\tilde{\rho}_i(j)|+\alpha_{i,2}|\tilde{\rho}_{i+1}(j)|)\\
&\le(\tilde{\gamma}_i)^k\sigma_{i,0}+\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}\bar{\eta}_i(j)\\
&+\frac{T_s}{L_i}\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}(\alpha_{i-1,1}\bar{\epsilon}_{i-1}(j)+\alpha_{i-1,2}\bar{\epsilon}_i(j))\\
&+\frac{T_s}{L_i}\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}(\alpha_{i,1}\bar{\epsilon}_i(j)+\alpha_{i,2}\bar{\epsilon}_{i+1}(j)).
\end{aligned}
$$

Hence, $|\epsilon_i(k)|\le\bar{\epsilon}_i(k)$ where $\bar{\epsilon}_i(k)$ is given in (22). ■

Based on the threshold ϵ_i, the attack detection decision is made based on the following principle: *An attack occurs if for some $i\in\mathcal{I}$, there exists some time k_d such that $|\epsilon_i(k_d)|>\bar{\epsilon}_i(k_d)$.*

VI. Detectability Analysis

Intuitively, an attack occurring at the $i+1$-th cell can be detected by the i-th local observer if the interconnection used by the observer $q_i(\hat{\rho}_i,\hat{\rho}_{i+1},k)$ is very different from the real interconnection $q_i(\rho_i,\rho_{i+1},k)$. In this respect, we define the following mismatching function:

$$\Delta q_i(k)=q_i(\hat{\rho}_i,\hat{\rho}_{i+1},k)-q_i(\rho_i,\rho_{i+1},k),\ \forall\ i\in\mathcal{I}. \quad (24)$$

Note that such a mismatching function is zero in the nominal case, and is nonzero for some $i\in\mathcal{I}$ in the presence of the attack. Then, the following theorem characterizes the class of detectable attacks.

Theorem 2: Consider the system (4) satisfying Assumption 1, and the observer (19). A covert attack generated by (8) is detectable if there exists some time k_d and some $i\in\mathcal{I}$ such that the mismatching function in (26) satisfies

$$
\begin{aligned}
&\frac{T_s}{L_i}\sum_{j=0}^{k_d-1}(\tilde{\gamma}_i)^{k_d-1-j}\left(|\Delta q_{i-1}(j)+\Delta q_i(j)|\right)\\
&>\bar{\epsilon}_i(k_d)+(\tilde{\gamma}_i)^{k_d}\tilde{\rho}_i(0)+\sum_{j=0}^{k_d-1}(\tilde{\gamma}_i)^{k_d-1-j}\left(\bar{\eta}_i(j)\right). \quad (25)
\end{aligned}
$$

Proof: In the presence of the attack, based on the defined mismatching function (26), the error system in (23) is written by

$$
\begin{aligned}
\tilde{\rho}_i(k+1)&=\tilde{\gamma}_i\tilde{\rho}_i(k)+\eta_i(k)\\
&\quad+\frac{T_s}{L_i}\left(\Delta q_{i-1}(k)+\Delta q_i(k)\right), \quad (26a)\\
\epsilon_i(k)&=\tilde{\rho}_i(k). \quad (26b)
\end{aligned}
$$

The solution of ϵ_i can be obtained by

$$
\begin{aligned}
&\epsilon_i(k)=\tilde{\rho}_i(k)=(\tilde{\gamma}_i)^k\tilde{\rho}_i(0)\\
&+\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}\left(\frac{T_s}{L_i}(\Delta q_{i-1}(k)+\Delta q_i(k))+\eta_i(j)\right).
\end{aligned}
$$

Based on the triangle inequality and Assumption 1, we have

$$
\begin{aligned}
&|\epsilon_i(k)|\ge(\tilde{\gamma}_i)^k\tilde{\rho}_i(0)\\
&+\sum_{j=0}^{k-1}(\tilde{\gamma}_i)^{k-1-j}\left(\frac{T_s}{L_i}(-|\Delta q_{i-1}(j)+\Delta q_i(j))|+\bar{\eta}_i(j)\right).
\end{aligned}
$$

In order to detect the attack, $|\epsilon_i(k_d)|>\bar{\epsilon}_i(k_d)$ must hold, which requires condition (25). Hence, the theorem is proved. ■

VII. Simulation Results

To evaluate the performance of the proposed distributed attack detection scheme, we consider a simulation scenario involving a homogeneous 3-lane freeway stretch divided into 10 cells, as depicted in Figure 1. Traffic flow is simulated entering the freeway through the mainline input and from on-ramps located at cells 2 and 8, while exiting occurs at the mainline output and off-ramps of all cells. This setup aims to mimic a typical 3-hour morning peak traffic pattern. The simulations are conducted based on the ACTM dynamics considering the following parameters: $v_i^f=100$ km/h, $T_s=10$ s, $w_i=30$ km/h, $Q_i=6000$ veh/h, and $\rho_i^J=300$ veh/km.

Figure 4 illustrates the density space-time diagrams resulting from the considered demand scenarios (a) without and (b) with ramp metering capabilities. These diagrams provide a visual representation of the instantaneous density across each region throughout the simulation duration as

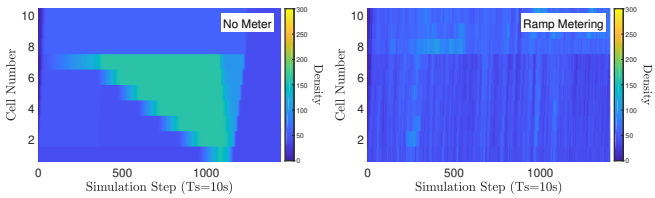

Fig. 4. The instantaneous density of each cell measured at each simulation time-step for the considered demand scenario: (a) without and (b) with metered ramps.

measured on each cell. The scenario depicted in Figure 4 is the one without metered ramps that represents the baseline condition in which traffic flow from on-ramps is unregulated. As expected, particularly during the morning commute, congestion occurs, underscoring the challenges of managing traffic influx without control measures. Conversely, in the scenario with metered ramps Figure 4 (b), ramp metering controls are strategically employed to regulate the vehicle influx from on-ramps, as per the solution approach outlined in P_2. This proactive management aims to optimize freeway throughput and minimize congestion. Notably, this strategy successfully maintains all cells within a congestion-free regime, clearly demonstrating the effectiveness of ramp metering in preventing traffic congestion under peak demand conditions.

A. Attack Scenario

An attack is assumed to occur on the ramp meter of cell 8, with the attacker having the resources to inject attack signals to the sensor readings of cells 7 and 8 (stealthy attack). This scenario is implemented based on the stealthy attack methodology presented in Problem P_3. The impact of this stealthy attack on the freeway traffic management system is analyzed by observing the instantaneous density as measured and perceived by three distinct entities within the system: (a) the freeway controller, which is under the influence of a stealthy attack, (b) the proposed observer, and (c) the attacker, who possesses accurate knowledge of the system's real state.

Figure 5 provides a representation of how a stealthy attack can alter the perceived traffic density within a freeway system, illustrating the discrepancies from three critical viewpoints: the freeway controller, an independent observer, and the attacker. Figure 5 (a) shows how the stealthy attack biases the freeway controller's measurements, potentially leading to decisions that exacerbate rather than mitigate congestion. This misrepresentation emphasizes the importance of having a reliable detection mechanism to prevent misguided responses to manipulated data. Conversely, the density estimated by the observer, as shown in Figure 5 (b), is clear that potentially can offer a more accurate estimate of the actual traffic conditions, but is evident that still the observer is unable to capture the true state of the system, especially at the attack. This perspective is crucial for maintaining a clear

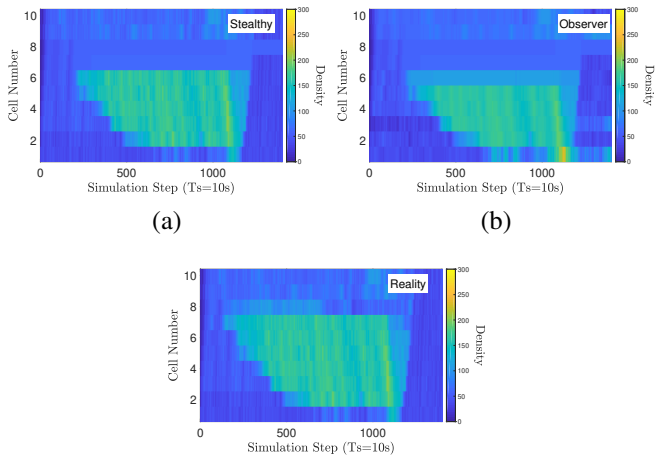

Fig. 5. The instantaneous density as measured by (a) the freeway controller under a stealthy attack, (b) the observer, and (c) the attacker, representing the real state of the system.

understanding of the freeway's operational state, unaffected by the stealthy signals. Finally, Figure 5 (c), unveils the genuine state of the system, thus highlighting the discrepancy between manipulated and actual traffic conditions.

To further validate the efficiency of our proposed detection mechanism, we delve into its ability to distinguish an attack through the comparison between the adaptive threshold and the observed residuals. This analysis is particularly focused on a scenario where the attacker targets cell 8. Figure 6 showcases the relationship between the value of the residual (represented by the blue line) and the adaptive threshold (depicted by the orange line) over time for each cell. Looking at the figure we can observe that there is no significant discrepancies in the cells under stealthy attack (cells 7 and 8), where the observed data are not diverging from what the system anticipates. On the other hand, in cell 6, where the attacker cannot inject an attack signal the value of the residual surpasses the adaptive threshold, triggering an alarm that indicates the presence of an attack. Hence, the ability of the system to accurately identify these variances not only attests to the robustness of the proposed framework but also reinforces its capability to safeguard the integrity and smooth operation of the traffic management system under various scenarios.

VIII. Conclusions

This work proposes a novel approach to identify stealthy cyber attacks targeting ramp metering systems in freeway networks. By presenting a novel stealthy attack scenario and proposing a robust, distributed detection mechanism, this work takes significant steps towards enhancing the resilience of traffic management infrastructures against such attacks. The effectiveness of our proposed detection mechanism, validated through rigorous simulations, demonstrates its potential in quickly identifying the impacts of stealthy attacks, thereby ensuring the continuous flow and safety of traffic.

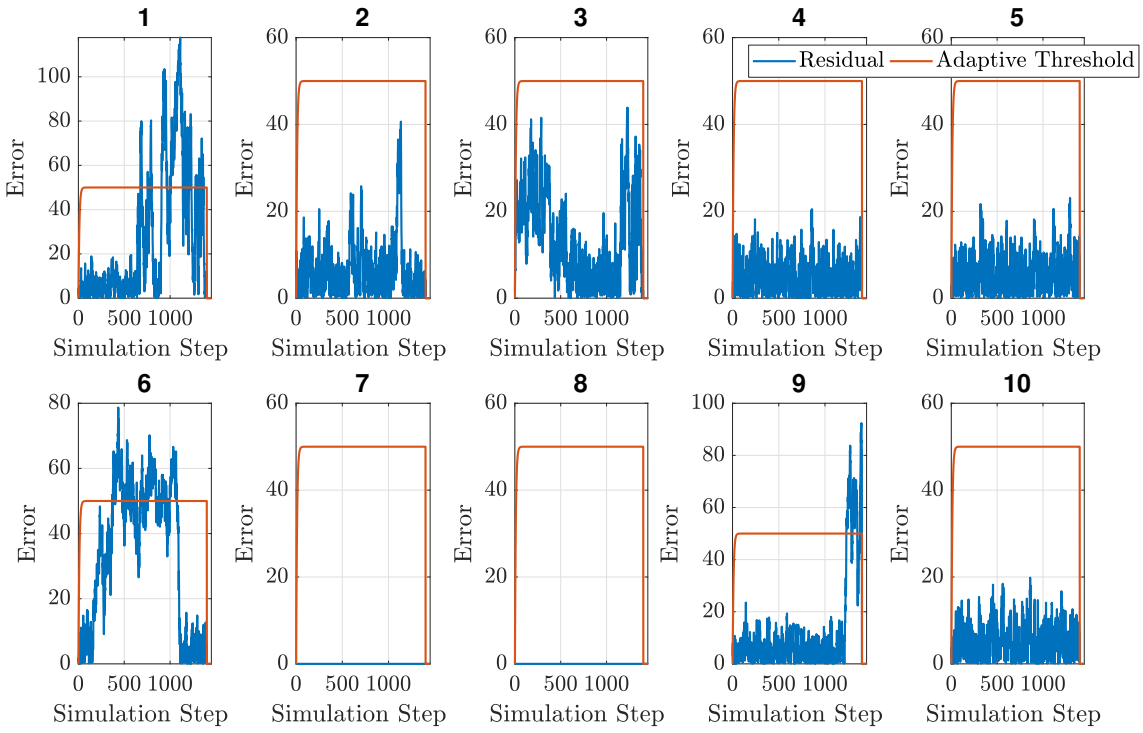

Fig. 6. The discrepancy between the adaptive threshold and the residual during a stealthy attack.

Looking forward, this research opens up new avenues for exploring the integration of advanced cybersecurity measures within the broader framework of intelligent transportation system operation.

REFERENCES

[1] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.

[2] S. Siri, C. Pasquale, S. Sacone, and A. Ferrara, "Freeway traffic control: A survey," *Automatica*, vol. 130, p. 109655, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0005109821001758

[3] J. Reilly, S. Martin, M. Payer, and A. M. Bayen, "Creating complex congestion patterns via multi-objective optimal freeway traffic control with application to cyber-security," *Transportation Research Part B: Methodological*, vol. 91, pp. 366–382, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261516303307

[4] I. Pekaric, C. Sauerwein, S. Haselwanter, and M. Felderer, "A taxonomy of attack mechanisms in the automotive domain," *Computer Standards & Interfaces*, vol. 78, p. 103539, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920548921000349

[5] C. Dong, H. Wang, D. Ni, Y. Liu, and Q. Chen, "Impact evaluation of cyber-attacks on traffic flow of connected and automated vehicles," *IEEE Access*, vol. 8, pp. 86 824–86 835, 2020.

[6] R. Horowitz and P. Varaiya, "Control design of an automated highway system," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 913–925, 2000.

[7] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal motorway traffic flow control involving variable speed limits and ramp metering," *Transportation Science*, vol. 44, no. 2, 2010.

[8] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE transactions on intelligent transportation systems*, vol. 3, no. 4, pp. 271–281, 2002.

[9] I. Papamichail, M. Papageorgiou, V. Vong, and J. Gaffney, "Heuristic ramp-metering coordination strategy implemented at monash freeway, australia," *Transportation Research Record*, vol. 2178, no. 1, pp. 10–20, 2010.

[10] I. Papamichail and M. Papageorgiou, "Traffic-responsive linked ramp-metering control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 111–121, 2008.

[11] E. Camacho and C. B. Alba, *Model predictive control*. Springer Science & Business Media, 2013.

[12] I. Papamichail, A. Kotsialos, I. Margonis, and M. Papageorgiou, "Coordinated ramp metering for freeway networks–a model-predictive hierarchical control approach," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 3, pp. 311–331, 2010.

[13] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 4, pp. 244–262, 2006.

[14] A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *28th Int. Conf. Distrib. Comput. Syst. Workshops*. IEEE, 2008, pp. 495–500.

[15] A. Teixeira, I. Shames, H. Sandberg, and K. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[16] S. Dibaji, M. Pirani, D. Flamholz, A. Annaswamy, K. Johansson, and A. Chakrabortty, "A systems and control perspective of CPS security," *Annu. Rev. Control*, vol. 47, pp. 394–411, 2019.

[17] A. Barboni, H. Rezaee, F. Boem, and T. Parisini, "Detection of covert cyber-attacks in interconnected systems: A distributed model-based approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3728–3741, 2020.

[18] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0191261594900027