Detecting Data Spoofing in Connected Vehicle based Intelligent Traffic Signal Control using Infrastructure-Side Sensors and Traffic Invariants

Junjie Shen, Ziwen Wan, Yunpeng Luo, Yiheng Feng*, Z. Morley Mao† and Qi Alfred Chen
University of California, Irvine, *Purdue University, †University of Michigan

***Abstract*—Connected Vehicle (CV) technologies are under rapid deployment across the globe and will soon reshape our transportation systems, bringing benefits to mobility, safety, environment, etc. Meanwhile, such technologies also attract attention from cyberattacks. Recent work shows that CV-based Intelligent Traffic Signal Control Systems are vulnerable to data spoofing attacks, which can cause severe congestion effects in intersections. In this work, we explore a general detection strategy for infrastructure-side CV applications by estimating the trustworthiness of CVs based on readily-available infrastructure-side sensors. We implement our detector for the CV-based traffic signal control and evaluate it against two representative congestion attacks. Our evaluation in the industrial-grade traffic simulator shows that the detector can detect attacks with at least 95% true positive rates while keeping false positive rate below 7% and is robust to sensor noises.**

I. Introduction

The Connected Vehicle (CV) technologies enable Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communications and can help the vehicles and infrastructure make more informed driving/control decisions. They can benefit our transportation system by improving mobility, reducing safety risks and greenhouse gas emissions, etc. As a result, government agencies across the globe are competing to push for CV deployments [1]–[3]. Particularly, the US is one of a few early adopters that has been testing the CV applications in US cities since 2016 [1]. In general, CV applications can be categorized into *vehicle-* and *infrastructure-side*. The vehicle-side applications aim to maximize fuel efficiency, enable better perception, etc. The infrastructure-side applications focus on improving the control decisions that will be executed in the infrastructure. The most representative infrastructure-side application is CV-based traffic signal control, which designs to improve traffic mobility by assigning signal timing plans that prioritize lanes with longer queues to reduce the total vehicle delays. Due to the enormous efforts for CV deployment, CV applications need to consider an inevitable transition period where CVs and Regular Vehicles (RVs) coexist on the road [4]. It is projected to take over 20 years to reach 95% market penetration rate (i.e., 95% vehicles are CVs) [4].

The widespread deployment of CV applications has a significant impact on the traffic safety and operations, thus making them to be valuable targets of cyberattacks. Among them, the most representative attack targets *real-world* infrastructure-side CV application [5], where they discover that CV-based traffic signal control is vulnerable to data spoofing attacks. The authors assume that the attacker can compromise the On-Board Unit on a CV to send malicious vehicle states to the signal controller to cause traffic congestions in the intersection. To exploit the vulnerabilities, they design two congestion attacks that target the full deployment and transition periods respectively. The two attacks are demonstrated to be very effective on the USDOT Intelligent Traffic Signal System (I-SIG) [6], a system already under testing in real-world intersections in US cities.

To defeat such data spoofing attacks, we explore a general spoofing detection strategy that cross validates the cyber-layer vehicle states using the physical-layer ones to identify the spoofers. In our design, we use the readily available *infrastructure-side sensors* [7]–[9] to obtain the physical states of CVs. However, the infrastructure-side sensors suffer from a fundamental limitation in the detection range compared to the CV communication range. This leads to challenges when dealing with vehicles out of the sensor range. To address it, we leverage well-established *traffic models* in transportation systems, which empirically describe the normal vehicle driving behaviors. We take the traffic models as the *traffic invariants* to estimate the physical states of the vehicles out of sensor range and assign trust scores based on the difference between reported and estimated states. Finally, a threshold-based anomaly detector is applied to identify suspicious CVs that have large negative impacts on the signal plan.

We evaluate the detector in an industrial-grade traffic simulator, PTV Vissim [10]. In the offline detection setting, our detector can strike a good balance between True Positive Rate (TPR) and False Positive Rate (FPR)–it can achieve at least 95% TPRs under all penetration rates while maintaining a low FPR of 7%. Specifically, when CVs are fully deployed, our detector shows a perfect detection with 100% TPR and 0% FPR. We also evaluate the robustness of our detector to the infrastructure-side sensor detection noises. The results indicate that our detector can tolerate even $3\times$ normal sensor detection noises. We then systematically explore the online detection capability of our detector. Results show that our detector remains effective even in online setting, where in the worst case, the FPR is only increased by 5% when the TPR maintains at 98.1% compared to offline detection. In summary, this work makes the following contributions:

- We explore a general spoofing defense for infrastructure-side CV applications based on the discrepancy between the cyber- and physical-layer vehicle states leveraging infrastructure-side sensors and traffic models.
- We implement the detector for CV-based intelligent traffic signal systems and evaluate the detection effectiveness against two congestion attacks. Results show that the de-

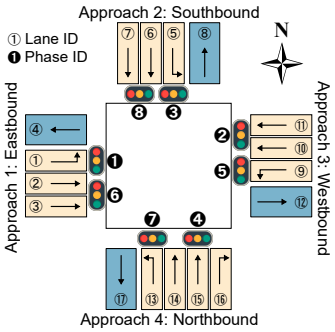

Fig. 1: An 8-phase intersection.

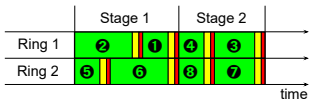

Fig. 2: Signal timing plan for 8-phase intersection. Circled numbers with solid background denote phase IDs.

tector can achieve >95% true positive rate when the false positive rate is <=7% and is robust to sensor noises.

- We explore detection timeliness and effectiveness when it is deployed as an online detector. Results show that online detection increases the false positive rate by only 5% when maintaining a high true positive rate in the worst case.

II. Background and Threat Model

A. CV-based Traffic Signal Control

Traffic signal control is one of the important applications of CV technologies designed to improve transportation mobility by reducing vehicle delays in the intersection. It determines the signal timing plan based on the real-time vehicle states (including vehicle ID, location, speed, heading, etc.) periodically broadcast from the CVs in the intersection. Specifically, CVs send their vehicle states in a standard Basic Safety Message format, which is transmitted over the CV communication network, i.e., Dedicated Short-Range Communication (DSRC) [11] or Cellular Vehicle-to-Everything (C-V2X) [12]. Among the open efforts in CV deployment in the US [1], the Intelligent Traffic Signal System (I-SIG) [6] is the only traffic signal system designed for the mobility of general urban intersections [13]. It has been tested in US cities and shown high effectiveness at reducing vehicle delays [14].

Fig. 1 shows the configuration of a common major arterial intersection. As shown in the figure, the traffic lanes are grouped by different *phases*, each with a dedicated traffic signal. As a major arterial intersection, it has two concurrent phase sequences (or *rings*) that do not interfere with each other and can be planned simultaneously. However, the phases in the same ring need to be planned sequentially due to the confliction. The 8 phases are separated into two *stages*, which are also conflicting with each other. The CV-based traffic signal controller is invoked at the end of each stage to calculate an *allocation of green lights* to the phases, which is called a *signal timing plan*. Fig. 2 shows a typical timing plan for 8-phase intersections. The green blocks indicate the allocated green light durations. The yellow and red blocks are two predefined durations for the yellow light and red clearance light (to accommodate for potential red light runners). The inputs to the signal controller are the latest CV states received at the end of each stage, which we refer them as a CV or traffic *snapshot* in this work.

B. Congestion Attacks on I-SIG

Recent work [5] discovered two vulnerabilities in the I-SIG system that can be exploited using data spoofing attacks. The first one is in queue length estimation, where the I-SIG system estimates the number of vehicles stopping in a queue based on the farthest stopped CV in the transition periods, i.e., when the Penetration Rate (PR) is smaller than 95%. They find that such a design can be exploited by the attacker to inject a fake long queue in the estimation, which we term ***queue length attack***. The second vulnerability targets the arrival time estimation, where I-SIG estimates when will a vehicle arrive at the stop line or stop behind a queue. An attacker can exploit this by setting the CV with a slow speed to cause a late arrival time estimation, which we term ***arrival time attack***. In both attacks, the I-SIG system allocates an *unnecessary* long green time for traffic lanes that are not busy and thus starve the other lanes that need prioritization. The attacks are evaluated in an industrial-grade traffic simulator PTV Vissim [10]. Results show that the two attacks can effectively cause severe congestion effects in full deployment and transition periods of CVs, respectively.

C. Traffic Modeling

Traffic models as the traffic invariants. Traffic modeling is a well-established topic in transportation engineering, which aims to precisely describe the relationships between vehicles and infrastructure with mathematical equations [15]. Due to the attractive property of describing normal traffic behaviors, we repurpose traffic models as the traffic invariants for security enforcement, e.g., benign vehicles will generally behave according to the traffic models. Specifically, we leverage *microscopic* traffic models as they describe individual vehicle behaviors when reacting to the actions of other vehicles.

Newell's car-following model. Car-following models [16]–[18] are the most typical microscopic model type that describes how should a vehicle follow its leading vehicle. The Newell's car-following model [17] is a simple but effective model, which is built upon the assumption that the follower vehicle's trajectory looks similar to its leading vehicle's trajectory, except with a time delay and a space translation. More formally, the model describes the following relationships:

$$\begin{aligned} s &= v_f \cdot \tau + d \\ x_{\text{follow}}(t+\tau) &= x_{\text{lead}}(t) - d, \end{aligned} \tag{1}$$

where s is the spacing the following vehicle keeps to the leading vehicle; x_{follow} and x_{lead} are the longitudinal position offsets of the following and leading vehicles; τ and d are two empirically determined parameters describing the reaction time and stopping distance of a normal driver, which are typically between 1.0–1.7 and 6.0–9.6, respectively [19], [20]; v_f is the free-flow speed, which is the speed limit of the road.

D. Threat Model

We assume same threat model as the congestion attacks proposed by Chen et al. [5], where the attacker is able to compromise the in-vehicle CV communication device, i.e., the On-Board Unit, in their own vehicle to send malicious BSM messages. We do not assume the attacker can spoof the vehicle

identifier in the BSM messages, which are protected by the Security Credential Management System. Thus, the attacker needs to use the original certificate associated with the physical vehicle in order to get the messages correctly authenticated.

III. Related Work

Security analyses of CV applications. Previous works [21]–[23] have studied the security threats on the CV communication networks, such as DoS, spamming, masquerading, replay attacks, which greatly damage the availability, authenticity, and confidentiality of the network. On the other side, domain-specific attacks have been demonstrated in different CV application scenarios. Chen et al. [5] show that CV-based Intelligent Traffic Signal Systems are highly vulnerable to congestion attacks if the attacker can compromise the OBU device. Amoozadeh et al. [24] demonstrate that message falsification attack cause significant instability in the Cooperative Adaptive Cruise Control (CACC) vehicle stream. Abdo et al. [25] perform a detailed analysis on CACC and present 4 different attack scenarios. Furthermore, Huang et al. [26] analyze the impact of falsified CV data and propose a black-box attack on the CV-based traffic signal control system.

Defenses against data spoofing. Sun et al. [27] propose a verification scheme approach that utilizes the angle-of-arrival and frequency-of-arrival to detect spoofing attacks. Such a defense requires extra hardware and the presence of enough number of reflectors in the driving environment. Guo et al. [28] propose a collaborative intrusion detection system, which leverages the sensor data from onboard sensors of neighboring CVs. Liu et al. [29] propose a blockchain-based framework to build trust and defeat spoofing attacks in CV applications. Both works assume specific hardware or software updates in the CVs, which may require enormous efforts due to a large number of CVs. In comparison, our detector does not impose such requirements on the CVs. Instead, we reuse the readily-available infrastructure-side sensors to establish the physical root-of-trust for the detection (§V-A). In this work, we approach from a novel angle by constructing and propagating trust based on infrastructure-side sensors and traffic invariants, which are complementary to existing defenses.

IV. Challenges of Applying Infrastructure-side Sensors for Data Spoofing Detection

Data spoofing attacks are cyber-layer attacks, where attackers send dishonest information of their physical states over the communication channels. Thus, a natural strategy to detect such attacks is to cross validate with physical-layer information. One readily available physical-layer information source is the infrastructure-side sensors, e.g., cameras [7], [30] and LiDARs [9], which perceive vehicle positions and speeds in the physical world and can serve as the ***physical root-of-trust***. Currently, the infrastructure-side sensors are already performing vehicle detection and tracking in red-light enforcement [31] and traffic monitoring [32]. Thus, we reuse them as a cost-effective solution for data spoofing detection.

A. Defense Challenges

Fundamental limitation of infrastructure-side sensors: detection range. Although infrastructure-side sensors can provide accurate detection of the CVs to validate their reported states, their detection ranges are often much more constrained than the CV communication ranges. As illustrated in Fig. 3, the effective detection ranges of traffic cameras are usually at ∼100 meters [8], [33], [34], while CV communication channels (e.g., DSRC and C-V2X) can cover much larger ranges (typically >300 meters [35]). This thus leaves opportunities for the attackers since they can simply spoof CV locations beyond the sensor detection range to evade direct detection. For example, the spoofed CVs are usually located at the end of each intersection approaches with a distance of ∼300 meters to the center of the intersection in the congestion attacks [5]. In fact, this is a fundamental limitation of sensors compared to cyber-layer communication–extending the sensing range (e.g., installing and synchronizing with additional sensors) is often much more costly and difficult than extending cyber-layer communication ranges (e.g., using signal relay devices or opting to longer range communication protocols such as C-V2X). Because of this fundamental limitation, two defense challenges need to be addressed in order to leverage the infrastructure-side sensors for effective data spoofing detection in the CV context.

Challenge 1: How to systematically propagate the trust from the sensor range to the CV communication range? With the help of infrastructure-side sensors, it is straightforward to verify the reported states of the CVs within the sensor range and establish trust for the ones that match the detection results. But for the CVs outside the sensor range, there is no direct way to measure their states. Despite that, the trusted CVs within the sensor range can provide useful information to verify the positions of the farther CVs, and hence a systematic solution is required to facilitate such trust propagation.

Challenge 2: How to infer the RV states outside the sensor range? Since it is estimated that the deployment of CV technology needs more than 20 years to reach at least 95% market penetration rate [4], there will be an inevitable transition period where CVs and RVs co-exist on the road. During such period, the RVs that are outside of sensor range may disrupt the trust propagation and cause mis- or false detections since the detector is not aware of these RVs. Thus, gaining the knowledge of these RV states is important for accurate spoofing detection.

V. Defense Design

A. Design Overview

In our design, we define the *trust* of a CV based on its integrity, i.e., CVs that report a state far away from its ground truth state will be assigned with lower trust (or higher suspicion). Our detector measures the trust of each CV in a traffic snapshot (i.e., the received CV states that a CV application is used for decision-making) and pinpoint the ones that have the lowest trust and the largest impact on the CV application performance. As shown in Fig. 4, the detector takes

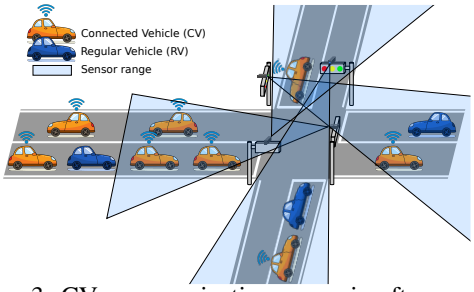

Fig. 3: CV communication range is often much larger than infrastructure-side sensor range.

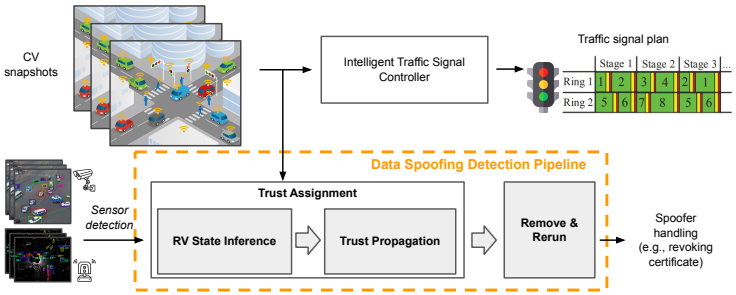

Fig. 4: Defense design overview.

the CV snapshot and the corresponding sensor detection results as input and outputs the suspicious CVs that are likely to be spoofers for further handling. The detection process involves two major steps: *Trust Assignment* and *Remove-and-Rerun*.

Trust Assignment (TA). In this step, we start from our physical root-of-trust, i.e., sensor detection results, to assign *suspicious scores* to the CVs in the sensor range by comparing their reported states with the detection results. Next, we propagate the trust out to the CV range in the order of CVs' reported distances to the sensor range. Since there is no direct way to measure the physical states of CVs out of the sensor range (Challenge 1), we estimate the CV states based on our ***traffic invariants***, i.e., the traffic models, which are empirically derived mathematical equations describing the vehicle driving behaviors under various traffic conditions (§II-C). For example, the car-following models can be used to estimate a vehicle's spacing and velocity based on its leading vehicle. We then use the estimated state as a proxy to the CV's ground truth state to calculate the suspicious score.

Since traffic model's accuracy depends on the availability of surrounding vehicle information, it is thus imperative to address Challenge 2 to infer their states in the current CV snapshot. To achieve that, we look into the "future" sensor frames when an RV first enters the sensor range to learn which vehicles are its neighbors. Based on the neighboring vehicles in a future time window, we can thus apply traffic models to infer an RV's state in the current CV snapshot. When the detection process is deployed as an offline analysis, such "future" sensor frames are always available. However, when deployed as an online detection, we have to delay the detection for a certain duration to wait for more sensor frames to come. In this case, since the future time window is limited, the number of recoverable RVs will also be reduced. However, in practice, even a relatively short future time window (e.g., 6 seconds as shown in §VIII-A) is sufficient to cover the majority of RVs. Although the detection is delayed in the online analysis, this is not much a problem for traffic signal control since the attack effect often takes time to build up, e.g., the I-SIG congestion attacks take ∼18 minutes to build up the congestion effects [5].

After the trust assignment, we rank the CVs based on their suspicious scores. As will be shown in §VI-B, the spoofing CVs are always ranked with top suspicious scores. In practice, we can aggregate the suspicious score rankings from multiple CV snapshots to more accurately pinpoint the spoofing CVs. Nevertheless, in our design, we apply a Remove-and-Rerun step to further improve the detection accuracy.

Remove-and-Rerun (RnR). In RnR, we re-execute the CV application *with* and *without* a suspicious CV to confirm its impact on the attack objective. The intuition behind this is that the attacker's goal is to disrupt the CV application to cause adverse effects on some CV application metrics, which can often be quantified in the application itself. For example, the congestion attacks on I-SIG are designed to increase the total delay of the vehicles in the intersection, which is exactly what I-SIG is optimized for. Such an attack objective driven approach can effectively distinguish attack CVs from benign ones among the most suspicious CVs.

B. Trust Assignment

The TA step essentially checks the cyber-layer CV states against their physical-layer states to locate dishonest CVs. It calculates a suspicious score s_i for each CV in the current snapshot as follows:

$$\begin{aligned} s_i =& \|x_i^c - x_i^p\| \\ x_i^c \in& X^c, x_i^p \in X^p, i \in \{1, ..., n\}, \end{aligned} \tag{2}$$

where x_i^c and x_i^p represent CV's reported and physical states (detailed later); n is the total number of CVs in the current snapshot; $\|\cdot\|$ denotes the absolute difference between two states, e.g., Euclidean distance. Based on the suspicious scores, we then rank the CVs to obtain the top-K suspicious ones X':

$$X' = \underset{X' \subset X^c, |X'|=K}{\arg\max} \{s_i \mid i = 1, ..., n\}. \tag{3}$$

We define the vehicle state as a vector $x = [t, l, r, v, h]$, where t is the timestamp; l is the traffic lane ID (e.g., as in Fig. 1); r is the distance to the intersection center; v is the vehicle speed; h is the vehicle heading. Since traffic signal controllers need to know the intersection geometry in order to plan for dynamic traffic situations, they often have a pre-built map that supports querying these state elements from vehicle BSMs. For example, I-SIG by default will convert the BSMs into such information before optimizing for the signal plan.

Among cyber- and physical-layer states, x_i^c is readily available from CVs' continuous broadcasting. For x_i^p, we rely on the physical root-of-trust (i.e., sensor detection) and traffic invariants (i.e., traffic models) to obtain and infer the physical states of CVs in and out of the sensor range. The former is out of scope of this work since vehicle detection and tracking is a well-studied topic in computer vision [33], [34] and already has many commercial products on the market that

can provide real-time detection [7], [30], [32], [36]. For the latter, it involves two sub-steps: *RV state inference* and *trust propagation*, and we will detail them in §V-B1 and §V-B2.

Car-following model as the traffic invariant. Since car-following models describe the inter-vehicle spacing that a vehicle will maintain given the leading vehicle's speed, we use them as the traffic invariant to infer the *ideal* position that a vehicle will be located in the current lane based on its leading vehicle. Specifically, we apply the widely-used Newell's car-following model (Eq. 1 in §II-C) in our design for its simplicity. Given a leading vehicle state $[t, l, r, v, h]$, we can estimate the following vehicle state at time t as follows:

$$\mathcal{M} : [t, l, r, v, h] \rightarrow [t, l, r + (v \cdot \tau + d), v, h]. \tag{4}$$

This conforms to the Newell's model that (1) the follower drives at the same speed as the leader, and (2) the follower's spacing is adjusted based on the speed.

1) RV State Inference: For RVs out of sensor range, we infer their states based on their *future* leading vehicles in the sensor range. More concretely, if an RV appears in the sensor detection in any of the future frames between t_0 and $t_0 + T$, we can thus infer the RV's physical state at time t_0 as follows:

$$x_j^{t_0} = \begin{cases} \mathcal{M}(x_{\text{lead}}^{t_0}), & \text{if } \exists\ t \in \{t_0, ..., t_0 + T\}, \|x_j^t - C\| < R \\ \varnothing, & \text{otherwise} \end{cases} \tag{5}$$
$$j \in \{1, ..., m\},$$

where x_j is the RV state at time; x_{lead} is the leading vehicle in the sensor frame; m is the total number of RVs; t_0 is the time of the CV snapshot to check for spoofing activity; $\mathcal{M}(\cdot)$ denotes the state estimation function based on the car-following model; C and R are the geographic center and radius of the sensor range, respectively. Depending on the time window (or the delay) allowed in the detector, it is possible that an RV will not appear in the sensor range. In such case, our detector will simply not be aware of this RV.

Identifying leading vehicle. To find the leading vehicle for a target vehicle, we iterate over all available vehicle states at t and look for the one that 1) belongs to the same lane and 2) is in front of and closest to the target vehicle. In cases when the distance to the closest leading vehicle is greater than $v_f \cdot \tau + d$, we regard the target vehicle as in free-flow traffic and exclude the leading vehicle since it should have negligible impact on target vehicle's driving behavior.

Handling RVs without leading vehicles. When there is no leading vehicle or the leading vehicle is too far away, we then estimate the RV state at t_0 based on its kinematics, assuming that the RV maintains the same speed between t_0 and t.

2) Trust Propagation: With more RV states made available out of sensor range, we can now more accurately estimate the physical state of the CVs and propagate the trust from our physical root-of-trust. Similar to the RV state inference, we apply the traffic invariant, i.e., the Newell's car-following model, to estimate the physical state of the CV i based on its leading vehicle as follows:

$$x_i^p = \begin{cases} \mathcal{M}(x_{\text{lead}}), & \text{if } \exists\ x_{\text{lead}} \\ [t_i^c, l_i^c, d_i^c, v_f, h_i^c], & \text{otherwise.} \end{cases} \tag{6}$$

Different from RV state inference, even when there is no leading vehicle, we are still aware of the existence of CV i. Thus, instead of ignoring it, we set its state the same as its reported cyber-layer state *except its speed as the free-flow speed of the lane* since we know for sure there is no leading vehicle and hence the CV should be driving at the free-flow speed in the normal case.

Suspicious score calculation. After obtaining the physical states from the sensor detection or inference, we then calculate a suspicious score for each CV, which is defined as the *distance* between the cyber- and physical-layer states (Eq. 7). More concretely, the suspicious score calculation depends on the availability of physical state elements as below:

$$s_i = \begin{cases} |d_i^c - d_i^p|, & \text{if } \exists\ d_i^p \\ |(v_i^c - v_i^p) \cdot \tau + d|, & \text{otherwise.} \end{cases} \tag{7}$$

When the CV's physical distance to the intersection is available, we calculate the suspicious score directly based on the difference to the one in the cyber-layer state. When only the speed element is available in the inferred physical speed, we plug the speed difference into Newell's model to obtain a spacing penalty such that the suspicious score is a distance measurement and hence comparable to the above case.

C. Remove and Rerun

We perform RnR on all top-K suspicious CVs (Eq. 3). For each CV k in X', we exclude it from the current CV snapshot and re-execute the I-SIG application to obtain the new total delay. Since the attacker aims to increase the total delay and ultimately cause congestion in the intersection, removing a spoofing CV from the snapshot would likely to reduce the total delay. Formally, we perform the RnR as follows:

$$x_k^c = \begin{cases} \text{spoofer}, & \text{if } \frac{\mathcal{F}(X^c) - \mathcal{F}(X^c \setminus \{x_k^c\})}{\mathcal{F}(X^c)} \geq \epsilon \\ \text{benign}, & \text{otherwise,} \end{cases} \tag{8}$$

where $\mathcal{F}(\cdot)$ denotes the I-SIG application that calculates the total delay; x_k^c is one of the top-K CVs; ϵ is an empirically determined anomaly threshold for spoofing detection. Since total delay varies under different traffic demands, we calculate a total delay reduction percentage based on the one with all CVs rather than an absolute value.

VI. Defense Effectiveness Evaluation

In this section, we evaluate our detector against the two congestion attacks [5] in an offline setting, where it is performed as a post-processing step on the saved traffic snapshots.

A. Evaluation Methodology

Real-world intersection configuration. In our evaluation, we use PTV Vissim to generate the traffic snapshots for the I-SIG system and execute the produced signal plans. To faithfully reproduce the simulation setup, we obtained the detailed Vissim configurations they used when evaluating the attack. Specifically, we simulate a real-world intersection as shown in Fig. 1. In addition, we also use the same traffic demand (i.e., vehicle arrival rate) and turning ratio (i.e., vehicle lane changing probability) in each lane that the authors collected from real-world intersection [5].

TABLE I: Suspicious score rankings of the attack CVs in Trust Assignment. The numbers in the parentheses are the CV snapshots that we rank the attack CV in Top-K and the total number of CV snapshots in the simulation, respectively.

PR	Attack	Top-1 Rate	Top 3 Rate	Top-5 Rate
100%	Arrival time attack	91.2% (93/102)	99.0% (101/102)	100.0% (102/102)
75%	Queue length attack	89.3% (200/224)	99.1% (222/224)	100.0% (224/224)
50%		90.7% (194/214)	99.5% (213/214)	100.0% (214/214)
25%		92.8% (180/194)	99.5% (193/194)	100.0% (194/194)

Infrastructure-side sensor detection assumption. In this section, we assume perfect camera detections, i.e., the vehicle positions and speeds can be accurately detected in the sensor range. We will relax this assumption later in §VII to evaluate the robustness against sensor noises. We set the sensor range to 91 meters (300 ft) in the evaluation since this is a common specification for traffic cameras [8], [36] and prior works also report similar capabilities [33], [34].

Evaluation metrics. To quantify the detection performance, we show the *True Positive Rates (TPRs)* and *False Positive Rates (FPRs)* under the attacked and benign CV snapshots, respectively. To generate such CV snapshots, we simulate each Vissim seed twice–with and without attack. Each simulation lasts for 4000 seconds, which consists of about 70–140 snapshots depending on the congestion level. In the attacked simulations, the attacker may choose not to spoof a snapshot if she cannot find any spoofing location to increase the total delay. We also plot the *Receiver Operating Characteristic (ROC) curves* to systematically show the TPR and FPR under different anomaly thresholds ϵ (§V-C). To validate the effectiveness of trust assignment, we report the *Top-K rate* to show the percentage of CV snapshots that rank the spoofing CV among the top-K most suspicious CVs.

B. Results

Accuracy of trust assignment. Table I shows the Top-K rates in the attacked snapshots. As shown, the TA step is quite effective at finding the attack CVs, where it *always ranks the attack CV among the top-5 most suspicious CVs across all PR settings*. Therefore, we set $K = 5$ in the following RnR step (thus denoted as RnR-5) as this empirically ensures that the attack CV, if any, is among the ones that will be validated. Moreover, over 89% of the total CV snapshots rank the attack CVs at Top-1. The reason that some attacked CVs are not ranked at the top is mainly that in these snapshots some benign CVs exhibit driving behaviors that are not considered in the car-following model (e.g., lane changing) such that the TA mistakenly assigns high suspicious scores to these benign CVs. This indicates that a simpler detector that purely relies on TA is unlikely to be effective. Yet, the TA is a crucial part of the detection process as it helps narrow down the detection scope for RnR to accurately pinpoint the attack CVs.

Effectiveness of the complete detector pipeline. We now evaluate the performance of our complete detection pipeline (TA and RnR). Since the benign CV snapshots are also involved in the evaluation, we need to select the anomaly threshold in RnR (§V-C) such that it would not incur many false positives while still maintaining a good detection accuracy. Table II shows the best TPRs that can be achieved under different FPR levels by varying the anomaly threshold. Our detector can achieve a *perfect detection with 100% TPR and 0% FPR* when the CVs are fully deployed (i.e., PR = 100%). Also, when the FPR is 7%, the detector can achieve at least 95% TPRs in all PR settings. Benefitting from the RnR-5, the complete detector pipeline further improves the detection accuracy on top of TA.

TABLE II: Attack detection effectiveness (TA + RnR-5).

PR	Attack	Detection effectiveness				
		TPR (FPR=7%)	TPR (FPR=5%)	TPR (FPR=3%)	TPR (FPR=1%)	TPR (FPR=0%)
100%	Arrival time attack	100%	100%	100%	100%	100%
75%	Queue length attack	100%	99.1%	96.0%	70.1%	61.6%
50%		99.5%	99.1%	98.1%	78.5%	0.5%
25%		95.4%	85.6%	82.0%	69.6%	0%

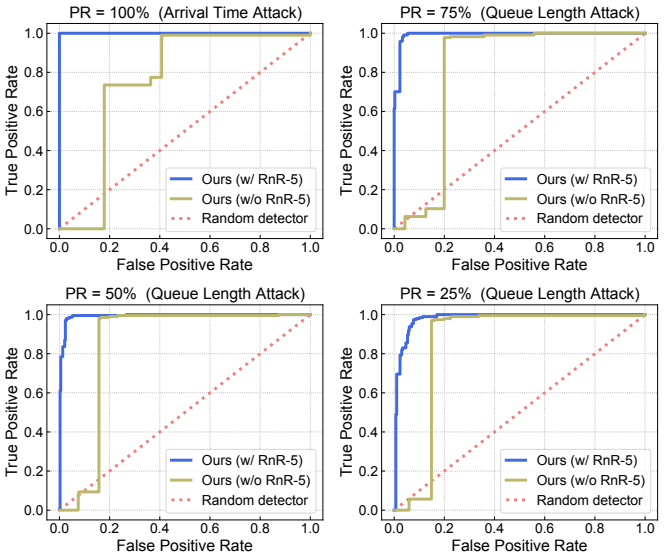

Fig. 5: Attack detection ROC curves of our detector with and without the RnR-5.

Nevertheless, the detection performance in the lower PR settings is generally worse, where the TPR drops to 85.6% when FPR is 5%. This is mostly caused by the congestion effects in the attacked snapshots, where a benign CV stops in the middle of an empty lane waiting to cut into a queue in the adjacent lane. This makes the I-SIG queue estimation mistakenly considers many RVs in the empty lane and thus allocates an unnecessary long green time for it. In such cases, removing this benign CV from the snapshot will actually reduce the total delay. In practice, such cases might not be much of a concern since one can use attack detection from multiple snapshots to reduce the false positives.

Detection effectiveness under different anomaly thresholds. To systematically evaluate the detection performance, we plot the detection ROC curves by varying the anomaly thresholds. To highlight the importance of RnR-5, we also

plot the ROC curves of a simpler detector design *without* RnR-5. We use a threshold-based anomaly detection design to classify CVs as attackers if their suspicious scores are above the threshold. Fig. 5 shows the ROC curves of the detector with and without the RnR-5. As shown, incorporating the RnR-5 greatly improves the detection performance in all PR settings. Moreover, the ROC curves indicate that our detector can generally well-distinguish the attack and benign CVs.

VII. Robustness to Sensor Noises

Experimental setup. In our TA design (§V-B), sensor detection noises in the position and velocity will directly affect the suspicious score calculation. Lu et al. [33] quantifies that the average position errors are only 0.8m and 1.7m within 50m and 120m distances from the camera, respectively. And the position errors are mostly longitudinal, i.e., in the direction of the lane. For the velocity error, they find that the estimated speed has an average error of 1.47 m/s to the actual speed. We model the camera detection noises based on their findings. More concretely, we inject random errors sampled from Gaussian distributions $e_{\text{pos}} \sim N(0, \sigma^2_{\text{pos}})$ and $e_{\text{vel}} \sim N(0, \sigma^2_{\text{vel}})$ to the detected position and velocity for all vehicles (CVs and RVs) in the sensor range. We select $\sigma_{\text{pos}} = 1.7$ m and $\sigma_{\text{vel}} = 1.47$ m/s in the evaluation. In addition, we also evaluate larger noise levels by scaling the error amounts to $2 \times \{\sigma_{\text{pos}}, \sigma_{\text{vel}}\}$ and $3 \times \{\sigma_{\text{pos}}, \sigma_{\text{vel}}\}$, respectively.

Results. Fig. 6 shows the ROC curves with/without camera detection noises. As shown, the detection performance is barely affected when PR is 100% or 75%, and is only slightly worse in the lower PR settings. This is mainly because the camera detection errors are relatively much smaller than the distance between the spoofed CV location and the location estimated from the traffic invariant. For example, even with 3× position errors, the error standard deviation merely equals to a common vehicle length (4–5m). In comparison, to induce a large total delay increment, the spoofed CV location is usually >18m away from the estimated location.

VIII. Online Detection Exploration

Timing overhead. Our detector is assumed to run on the Road-Side Unit that hosts the CV applications. To ensure that we do not assume any unreasonable computation capability, we measure the timing overhead on a Raspberry Pi. The TA and RnR-5 steps take at maximum 9.05 sec and 6.13 sec, respectively. Therefore, we estimate the maximum required duration of the detection process as 15.58 sec.

Detection timeliness requirements. The I-SIG system runs at the end of each signal stage (§II-A) to plan for the next signal timing plan. Thus, the online attack detection needs to be finished within one stage in order to keep up with the pace of I-SIG. Specifically, we assume a minimum green light as 7 sec [37] for each phase. Each phase will also go through a yellow light (typically 3 sec) and red clearance period (to account for red light runners, typically 1 sec). Hence, one signal stage (i.e., two sequential phases) will take at least 22 sec. Excluding the computation delay, it leaves 6 sec for the future time window. Therefore, we set the future time window as 6 sec in the online detection evaluation.

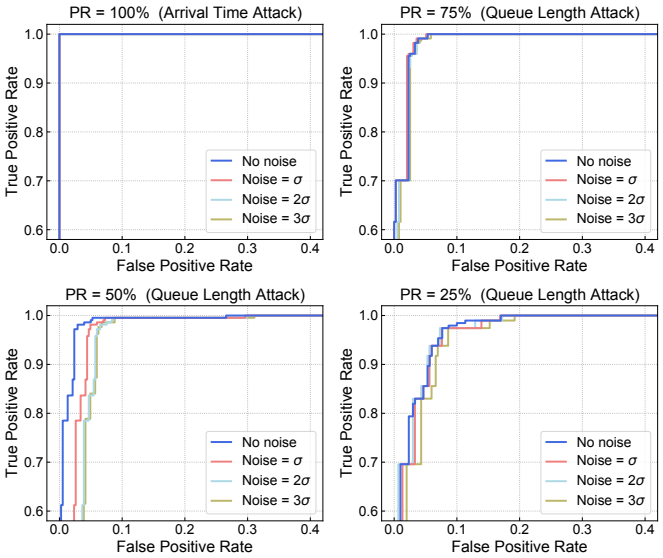

Fig. 6: Attack detection ROC curves (zoom-in view) under different levels of camera detection noises ($\sigma = \{\sigma_{\text{pos}}, \sigma_{\text{vel}}\}$).

A. Online Detection Effectiveness

Fig. 7 shows the comparison of the offline and online detection ROC curves. Since the future time window only affects the RV state inference (§V-B1), online detection has no impact in the full deployment period (PR = 100%) and has little impact when PR is 75% where the number of RVs are limited. Interestingly, limiting the future time window has a larger impact on PR = 50% compared to PR = 25%. This is because when PR = 25%, although the number of RVs increases, the number of CVs decreases correspondingly. Therefore, the detector is less affected due to the smaller number of CVs that need to be validated. Nevertheless, even when PR = 50%, the FPR only increases by 5% if we aim to keep the same TPR at 98.1%. Such a good online detection performance is likely because a future time window of 6 sec can already cover many RVs to enter the sensor range.

IX. Conclusion

In this work, we explore a general defense solution against data spoofing attacks on infrastructure-side CV applications. Building upon the general principle that using physical-layer information to cross validate cyber-layer information, we leverage the readily-available infrastructure-side sensors, such as traffic cameras, to estimate the physical-layer CV states. However, such infrastructure-side sensors suffer from a fundamental limitation that the sensor range is generally much smaller than the CV communication range. To address this, we take traffic models as traffic invariants to infer the vehicle states that are out of sensor range. We implement and evaluate our detector against two representative data spoofing attacks that aim to cause congestion in the intersections by exploiting the CV-based traffic signal control. Our results show that the detector can effectively identify the spoofer with high detection accuracy and is robust to sensor noises. We also demonstrate that an online detection setting only slightly degrades the detection performance.

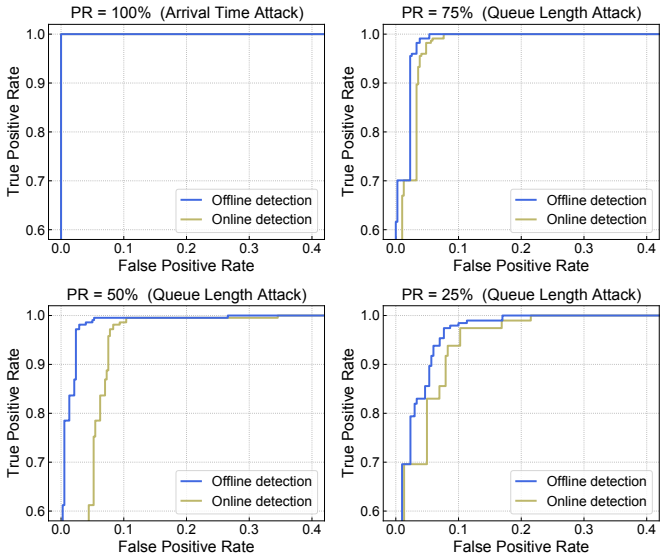

Fig. 7: Attack detection ROC curves (zoom-in view) of the offline and online detection setups.

X. Acknowledgement

This work was supported in part by the National Science Foundation (NSF) under grants CNS-1850533, CNS-1929771, CNS-2145493, and USDOT under grant 69A3552047138.

References

[1] U.S. Department of Transportation (USDOT), "USDOT Connected Vehicle Pilot Deployment Program." https://www.its.dot.gov/pilots/.

[2] Ryan Wu, "C-V2X automotive tech brings enhanced safety and efficiency to China's roads." https://www.qualcomm.com/news/onq/2021/03/02/c-v2x-brings-enhanced-safety-and-efficiency-chinas-roads, 2021.

[3] ETAuto, "Europe to add 69 million connected cars during 2020-25: Report." https://auto.economictimes.indiatimes.com/news/auto-technology/europe-to-add-69-million-connected-cars-during-2020-25-report/79350675, 2020.

[4] J. Volpe, "Vehicle-infrastructure integration (VII) initiative benefit-cost analysis Version 2.3," *USDoT Tech. Rep*, 2008.

[5] Q. A. Chen, Y. Yin, Y. Feng, Z. M. Mao, and H. X. Liu, "Exposing Congestion Attack on Emerging Connected Vehicle based Traffic Signal Control," in *NDSS*, 2018.

[6] L. Head, "The Multi Modal Intelligent Traffic Signal System (MMITSS): A Connected Vehicle Dynamic Mobility Application," in *Mid Year Meeting, Traffic Signal Systems Committee, Transportation Research Board, MMITSS. I-95.06 (20)*, vol. 20, 2016.

[7] Cubic, "GRIDSMART — Single Camera Solution for Traffic Management." https://gridsmart.com/products/gridsmart-solution/.

[8] Honda Technology, "Honda Demonstrates New "Smart Intersection" Technology." https://csr.honda.com/2018/10/04/honda-demonstrates-new-smart-intersection-technology/, 2018.

[9] H. Xu, Z. Tian, J. Wu, H. Liu, J. Zhao, *et al.*, "High-Resolution Micro Traffic Data from Roadside LiDAR Sensors for Connected-Vehicles and New Traffic Applications," tech. rep., University of Nevada, Reno. Solaris University Transportation Center, 2018.

[10] PTV Group, "PTV Vissim - Traffic Simulation Software." https://www.ptvgroup.com/en/solutions/products/ptv-vissim/.

[11] J. B. Kenney, "Dedicated Short-Range Communications (DSRC) Standards in the United States," *Proceedings of the IEEE*, 2011.

[12] A. Papathanassiou and A. Khoryaev, "Cellular V2X as the Essential Enabler of Superior Global Connected Transportation Services," *IEEE 5G Tech Focus*, vol. 1, no. 2, pp. 1–2, 2017.

[13] "Connected Vehicle Applications for Mobility." https://www.its.dot.gov/pilots/pilots_mobility.htm.

[14] K. Ahn, H. Rakha, and D. K. Hale, "Multi-Modal Intelligent Traffic Signal Systems (MMITSS) impacts assessment," tech. rep., United States. Department of Transportation, 2015.

[15] A. D. May, *Traffic Flow Fundamentals*. 1990.

[16] R. W. Rothery, "Car Following Models," *Trac Flow Theory*, 1992.

[17] G. F. Newell, "A simplified car-following theory: a lower order model," *Transportation Research Part B: Methodological*, 2002.

[18] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transportation Research Part B: Methodological*, 1981.

[19] S. Ahn, M. J. Cassidy, and J. Laval, "Verification of a Simplified Car-Following Theory," *Transportation Research Part B: Methodological*, vol. 38, no. 5, pp. 431–440, 2004.

[20] X. Wu, H. X. Liu, and N. Geroliminis, "An Empirical Analysis on the Arterial Fundamental Diagram," *Transportation Research Part B: Methodological*, vol. 45, no. 1, pp. 255–266, 2011.

[21] S. Hu, Q. A. Chen, J. Sun, Y. Feng, Z. M. Mao, and H. X. Liu, "Automated Discovery of Denial-of-Service Vulnerabilities in Connected Vehicle Protocols," in *USENIX Security*, 2021.

[22] W. Whyte, J. Petit, V. Kumar, J. Moring, and R. Roy, "Threat and Countermeasures Analysis for WAVE Service Advertisement," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 1061–1068, IEEE, 2015.

[23] A. Laszka, B. Potteiger, Y. Vorobeychik, S. Amin, and X. Koutsoukos, "Vulnerability of Transportation Networks to Traffic-Signal Tampering," in *ICCPS*, IEEE, 2016.

[24] M. Amoozadeh, A. Raghuramu, C.-N. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, "Security Vulnerabilities of Connected Vehicle Streams and Their Impact on Cooperative Driving," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 126–132, 2015.

[25] A. Abdo, S. M. B. Malek, Z. Qian, Q. Zhu, M. Barth, and N. Abu-Ghazaleh, "Application Level Attacks on Connected Vehicle Protocols," in *RAID*, 2019.

[26] S. E. Huang, Q. A. Chen, Y. Feng, Z. M. Mao, W. Wong, and H. X. Liu, "Impact Evaluation of Falsified Data Attacks on Connected Vehicle Based Traffic Signal Control Systems," in *AutoSec*, 2021.

[27] M. Sun, Y. Man, M. Li, and R. Gerdes, "SVM: Secure Vehicle Motion Verification with a Single Wireless Receiver," in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 65–76, 2020.

[28] P. Guo, H. Kim, L. Guan, M. Zhu, and P. Liu, "VCIDS: Collaborative intrusion detection of sensor and actuator attacks on connected vehicles," in *International Conference on Security and Privacy in Communication Systems*, pp. 377–396, Springer, 2017.

[29] X. Liu, B. Luo, A. Abdo, N. Abu-Ghazaleh, and Q. Zhu, "Securing Connected Vehicle Applications with an Efficient Dual Cyber-Physical Blockchain Framework," *arXiv preprint arXiv:2102.07690*, 2021.

[30] Transoft, "TrafxSAFE Connect - Real-time Road Safety Monitoring Platform." https://safety.transoftsolutions.com/trafxsafe-connect/.

[31] W. Hu, A. T. McCartt, and E. R. Teoh, "Effects of red light camera enforcement on fatal crashes in large US cities," *Journal of safety research*, vol. 42, no. 4, pp. 277–282, 2011.

[32] TrafficVision, "Traffic Intelligence from Video." http://www.trafficvision.com/.

[33] D. Lu, V. C. Jammula, S. Como, J. Wishart, Y. Chen, and Y. Yang, "CAROM–Vehicle Localization and Traffic Scene Reconstruction from Monocular Cameras on Road Infrastructures," in *ICRA*, IEEE, 2021.

[34] K. Cordes, N. Nolte, N. Meine, and H. Broszio, "Accuracy evaluation of camera-based vehicle localization," in *ICCVE*, IEEE, 2019.

[35] Douglas Gettman, "DSRC and C-V2X: Similarities, Differences, and the Future of Connected Vehicles." https://www.kimley-horn.com/dsrc-cv2x-comparison-future-connected-vehicles/, 2020.

[36] Teledyne FLIR, "Teledyne FLIR TrafiSense AI: AI-Powered Thermal Traffic Sensor." https://www.flir.com/products/trafisense-ai/.

[37] T. Urbanik, A. Tanaka, B. Lozner, E. Lindstrom, K. Lee, S. Quayle, S. Beaird, S. Tsoi, P. Ryus, D. Gettman, *et al.*, *Signal Timing Manual*, vol. 1. Transportation Research Board Washington, DC, 2015.