On the Cybersecurity of Traffic Signal Control System With Connected Vehicles

Yiheng Feng, Shihong Ed Huang, Wai Wong, Qi Alfred Chen, Z. Morley Mao, and Henry X. Liu, *Member, IEEE*

***Abstract*— Connected vehicle (CV) technology brings both opportunities and challenges to the traffic signal control (TSC) system. While safety and mobility performance could be greatly improved by adopting CV technologies, the connectivity between vehicles and transportation infrastructure may increase the risks of cyber threats. In the past few years, studies related to cybersecurity on the TSC systems were conducted. However, there still lacks a systematic investigation that provides a comprehensive analysis framework. In this study, our aim is to fill the research gap by proposing a comprehensive analysis framework for the cybersecurity problem of the TSC in the CV environment. With potential threats towards the major components of the system and their corresponding impacts on safety and efficiency analyzed, data spoofing attack is considered the most plausible and realistic attack approach. Based on this finding, different attack strategies and defense solutions are discussed. A case study is presented to show the impact of the data spoofing attacks towards a selected CV based TSC system and corresponding mitigation countermeasures. This case study is conducted on a hybrid security testing platform, with virtual traffic and a real V2X communication network. To the best of our knowledge, this is the first study to present a comprehensive analysis framework to the cybersecurity problem of the CV-based TSC systems.**

***Index Terms*— Traffic signal control system, cybersecurity, connected vehicles, security testing platform.**

I. Introduction

TRAFFIC signal control (TSC) system plays a critical role in urban transportation operations by regulating conflicting traffic flows to ensure safety and efficiency. With the development of connected vehicle (CV) technologies, vehicles, and transportation infrastructure are able to transmit information through wireless communications. In the CV environment, the TSC system receives CV trajectory data and optimizes signal parameters while vehicles receive traffic signal status to assist trajectory planning. Although many existing studies showed that safety, mobility, and fuel efficiency could be greatly improved by adopting CV technologies, the connectivity may increase the risks of cyber threats.

The study of the cybersecurity problem of the TSC system falls into a broader field of industrial control system cybersecurity [1]. The system can be roughly divided into three levels: network communication level, operating system level, and application level, which are similar to a computer system [2]. The network communication level security mainly considers CV communication technologies such as Dedicated Short Range Communication (DSRC) or cellular Vehicle-to-Everything (C-V2X) and their associated credential encryption mechanisms. The Security Credential Management System (SCMS) applied in the current CV system [3] is one of the examples at this security level. The operating system level security refers to the traffic controller's hardware, firmware, and basic operational functionalities. For instance, the Malfunction Monitoring Unit (MMU) prevents signals of conflicting movements from turning to green at the same time. The MMU can be implemented in either hardware or firmware, which provides basic safety and security assurance. The application level refers to different types of signal control algorithms, varying from fixed-time, actuated to adaptive. Protections from all three levels are necessary because redundant security mechanisms raise difficulties for attackers.

This paper focuses on the security analysis of the application level of the TSC system. While this generally understudied area was explored by a few existing studies, there lacks systematic analysis of the cybersecurity problem of the TSC system in the CV environment. This paper aims to propose an analysis framework consisting of three major components: attack strategies, defense solutions, and a security testing platform that integrates both attack and defense services with a CV-based TSC system. The threat model of the TSC system is first discussed, with a brief introduction of the existing literature. Based on the threat model, data spoofing attacks, which refer to an injection of intentionally modified CV data into a TSC system, are identified as the most realistic attack strategy. Details of the analysis framework are then introduced. To assess the risk of the attacks, major factors including the attacker's objective, budget, and knowledge about the

Manuscript received 28 October 2020; revised 18 December 2021; accepted 2 February 2022. Date of publication 16 February 2022; date of current version 12 September 2022. This work was supported in part by the U.S. National Science Foundation through Secure and Trustworthy Cyberspace (SaTC) under Grant 1930041 and in part by the Mcity—University of Michigan. The Associate Editor for this article was S. Olariu. *(Corresponding author: Yiheng Feng.)*

Yiheng Feng is with the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: feng333@purdue.edu).

Shihong Ed Huang is with the Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: edhuang@umich.edu).

Wai Wong is with the Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch 8140, New Zealand (e-mail: wai.wong@canterbury.ac.nz).

Qi Alfred Chen is with the Department of Computer Science, University of California at Irvine, Irvine, CA 92617 USA (e-mail: alfchen@uci.edu).

Z. Morley Mao is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: zmao@umich.edu).

Henry X. Liu is with the Department of Civil and Environmental Engineering and the University of Michigan Transportation Research Institute, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: henryliu@umich.edu).

Digital Object Identifier 10.1109/TITS.2022.3149449

system are considered and two defense solution directions are discussed. To test the effectiveness of both attack and defense models, a cybersecurity testing platform is designed and built at Mcity, a closed testing facility for connected and automated vehicles (CAVs). Finally, a comprehensive case study is conducted to demonstrate the cybersecurity analysis framework with a selected CV-based TSC system.

The main contributions of this paper are: 1) develop the first comprehensive analysis framework for the cybersecurity problem of the TSC system in the CV environment; 2) investigate different threat models, attack strategies, and defense solutions with different CV-based TSC systems; 3) integrate transportation engineering domain knowledge into the cybersecurity analysis of the TSC system; and 4) demonstrate the proposed analysis framework using a prototype evaluation system with a mix of simulation and real-world facilities.

Note that on Nov. 18th, 2021, the Federal Communications Commission (FCC) adopted new rules for the 5.9GHz band spectrum. The upper 30 MHz bandwidth is allocated for automobile safety using cellular vehicle-to-everything (C-V2X) technology instead of DSRC [4]. The switch of the radio technology from DSRC to C-V2X does not impact the threat model, defense strategies, and analysis results presented in this paper, since the both technologies have the identical application, network and security layers [5]. This paper focuses on the security analysis at the application level of the TSC system (i.e., how the traffic control system utilizes data received to make control decisions), which can be applied to different radio technologies at the radio access layer.

The rest of the paper is organized as follows. Section 2 analyzes the threat models of a TSC system and reviews the related studies. Section 3 introduces the cybersecurity analysis framework. Section 4 presents the case study that implements both attack strategies and defense solutions in a CV-based TSC system. Section 5 discusses how the proposed study can be improved from different perspectives and Section 6 concludes the paper.

II. Threat Model

In this section, potential threats towards the major components of a TSC system and their corresponding impacts on the safety and efficiency of the traffic signal operations are examined. The most relevant threat model for CV-based TSC systems is subsequently proposed.

Threat models are highly dependent on the architecture of the system. Typically, a TSC system consists of three major components: signal controllers,[1] traffic sensors, and a central management system. The signal controllers are located within signal cabinets and execute signal timing plans according to different control logic. The fixed time signal plans are programmed within the controller and no sensor input is needed. For actuated signal control, the controllers receive data from traffic sensors and adopt simple logic such as gap out and max out to adjust signal timing. For adaptive signal control systems (e.g., SCOOT [6], and SCATS [7]), the sensor data are first transmitted to the central management system, which generates all or partial signal timing parameters for all intersections in the network. The signal timing parameters are then distributed to the local signal controllers for execution. Existing studies showed that some components may be more vulnerable to cyber-attacks (e.g., spoofing of user identity, data spoofing, denial of service (D.o.S), etc.) than others.

Ghena *et al.* [8] considered the threat model as "infiltrating the traffic network through its wireless infrastructure". In this setting, local controllers were connected to the central management system through a wireless network (e.g., 5.8G Hz radios or 900 MHz radios). Their paper demonstrated that the wireless network could be penetrated using a radio produced by the same vendor. Once on the network, the attacker could take partial control of the signal controller through either debug port of the operating system or through the National Transportation Communications for Intelligent Transportation Systems Protocol, or NTCIP commands, to modify signal timing parameters directly. Assuming that the signal controllers could be directly accessed, Laszka *et al.* [9] and Ghafouri *et al.* [10] formulated mathematical programming problems to investigate the impact of such attacks on fixed-timing signals. Heuristic and decomposition algorithms were developed to quantify the impacts at the network level. Under the same assumption, Reilly *et al.* [11] presented a study on attacking freeway ramp meters (considered as the signal control on the freeway) to generate arbitrarily complex congestion patterns. The attacker might have different objectives varying from causing network-wide congestion to escaping from police pursuit. Recently, Perrine *et al.* [1] studied the district-wide impact of cyber-attacks towards traffic signals. In this study, it was assumed that the attackers could disable part of the traffic signals in the network by replacing normal signal timing plans with flashing red (operated as a four-way stop sign intersection). A dynamic traffic assignment (DTA) model was then incorporated to model the changes in route choice with disrupted traffic signals. The downtown Austin network was used for demonstration and simulation results showed that the total delay increased by more than 400% by only disabling a few intersections in the network. All the above threat models assumed that the signal controllers or the central management system could be accessed directly and manipulated freely. In reality, most of the signal controllers and central management systems are connected through wired connections (e.g., fiber network) within agencies' private networks. As a result, directly hacking into the private network may require physical access to the devices or penetrating transportation agencies' security firewalls.

Comparing with signal controllers and the central management system, traffic sensors have to be installed outdoors for data collection purposes. Traditional loop-detectors have lower risks towards cyberattacks because they don't transmit data through wireless communications, which is the most common media to penetrate the system. In contrast, new types of vehicle detectors, which utilize wireless technology to detect travelers (e.g., Bluetooth, and Wi-Fi) are proved to be vulnerable [12]. If the detectors are compromised, the attacker could send

[1]The signal controller here not only includes the controller hardware, but also the control algorithms (e.g., fixed-time, actuated, and adaptive).

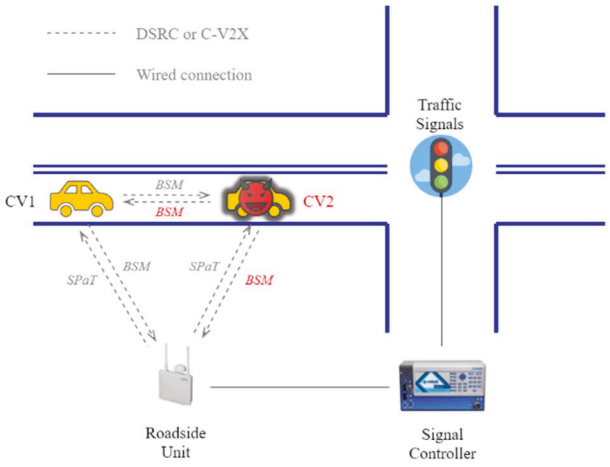


Fig. 1. Threat model - CV data spoofing attack.

any arbitrary traffic data to the TSC system and influence the control decisions. A study conducted by Feng *et al.* [13] demonstrated that the entire corridor could be impacted even if only one vehicle detector is under attack sending falsified detection data (i.e., adding fake vehicle calls or canceling real vehicle calls) under actuated control. Vehicle queues would propagate backward from the intersection under attack and eventually create network gridlock.

Most studies discussed previously assumed that devices from the transportation infrastructure (i.e., controllers, or detectors) could be compromised. One main reason is that in the traditional signal control framework, vehicles are not involved in the decision-making process. They are only passively detected by traffic sensors. Nevertheless, when vehicles are equipped with CV technologies (i.e., onboard unit (OBU)), they could serve as remote sensors that proactively provide information to the TSC system. In such cases, CVs broadcast Basic Safety Messages (BSMs), which comprise information such as the location, speed, acceleration, and heading of their host vehicles. Since these BSMs data are the reflections of the traffic states, they contain fruitful information for traffic signal optimization. Besides all the benefits of utilizing CV data in the TSC system, this feature brings a new threat model: a CV is compromised and sends spoofed data to influence the signal timing plans, as shown in Fig. 1. The roadside unit (RSU) receives BSMs from all the nearby CVs (e.g., CV1) and broadcasts the Signal Phasing and Timing (SPaT) messages. The signal controller generates and executes signal timing plans based on the received CV data. Assume CV2 is compromised and broadcasting falsified BSMs. The TSC system would utilize information from both CV1 and CV2 to optimize the traffic signal parameters. Apparently, the resultant timing plan would deviate from the optimal solution due to the falsified information, which reduces the efficiency of the intersection operation. Because the spoofing data attack is launched from the vehicle side and could only influence the signal control decision, rather than direct manipulation, this is considered as an indirect attack.

We believe this CV-based data spoofing attack is a realistic and plausible attack mechanism for the following reasons:

- It is expected that in the near future, there would be a significant number of CVs on the road serving as remote sensors, and an increasing number of CV-based TSC systems will be deployed since CV technology has shown its great potential in improving mobility. Both the government and private sectors consider such technology as an essential component for the future transportation system. For example, the U.S. Department of Transportation (USDOT) launched several projects targeting testing and evaluation of CV technologies including Safety Pilot Model Deployment (SPMD) [14], CV Pilot Deployment Projects [15], and Smart City Challenge [16].
- The threat model leverages vulnerabilities from the vehicle side, instead of tampering with any infrastructure devices. Transportation agencies can take any precautions to protect the integrity of the TSC system from the infrastructure side, but vehicle owners could still be malicious. The attackers have arbitrary access to their vehicles and do not need to expose themselves in a public environment. Studies showed that it is possible to hack into an OBU by exploiting software vulnerabilities. For instance, Veeraraghave [17] found an admin user vulnerability which gives root user privileges, with which critical security information such as the SCMS certificates can be accessed. Then the attacker could modify the content of the V2X messages of a compromised vehicle with a legitimate identity, and thus becomes an inside attack. A more comprehensive survey on security issues and state-of-the-art defense works in V2X can be found in [18].
- All messages in the communication network are in broadcast mode. That means every message (i.e., BSM and SPaT) that are received by other CVs and the infrastructure could also be received by the attacker. The attackers could acquire complete input and output information of the control system. The data spoofing attacks could be launched anywhere remotely as long as the compromised vehicles are within the communication range, which reduces the chance of being identified.

III. Cybersecurity Analysis Framework

In this section, the cybersecurity framework of the CV-based TSC system, as shown in Fig. 2, is introduced. It consists of three main components: risk assessment, defense solutions, and a security testing platform that can be used for evaluating the impact of different attack strategies and the benefits of various defense solutions. Note that both risk assessment and defense solutions will be analyzed based on the threat model (i.e., data spoofing) identified in the previous section.

A. Risk Assessment

According to the Transportation System Section Cybersecurity Framework Implementation Guidance [19], risk assessment is the first step in implementing the National Institute of Standards and Technology (NIST) cybersecurity framework. Risks could be estimated as impact multiplies likelihood [20]. The likelihood of cyber-attacks is very difficult to estimate because no known existing attacks are targeting CV-based TSC systems. This is largely because such systems are still under the research and development stage. As pointed out in

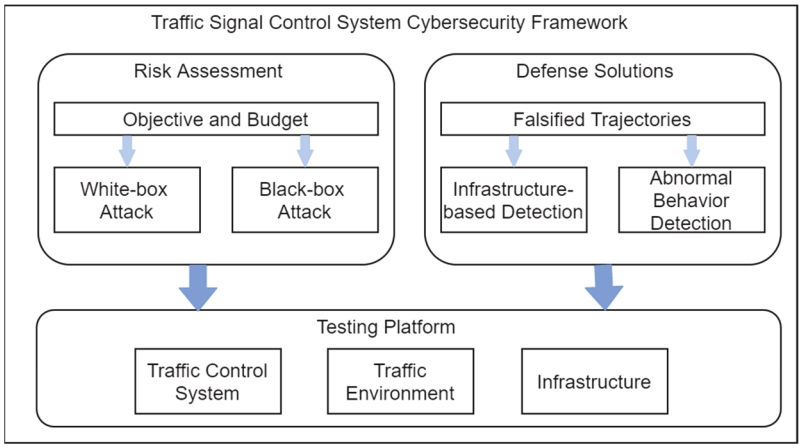


Fig. 2. Signal control system cybersecurity framework.

Section 2, it is expected that as more CV-based TSC systems would be deployed in the near future, it is necessary to evaluate the impact based on hypothetical cyberattacks before the real-world implementations.

The impact of cyberattacks is typically determined by the attacker's objective, budget, and knowledge about the system.

1) Objective: Feng *et al.* [13] and Reilly *et al.* [11] listed a few attacker's objectives toward the TSC system from different perspectives:

- To cause safety issues. Since data spoofing attacks could not directly manipulate the signal timing parameters in the controller, causing safety issues (e.g., enabling conflicting phases to be green at the same time) would be difficult. Most of the safety-critical functionalities are hardcoded at the hardware or firmware level. Even if the input data is compromised, the controller is not able to generate unsafe timing plans.
- To obtain personal benefits. Personal benefit is a common goal for the attackers such as reducing travel time or escaping from a police pursuit. By manipulating the signal timing, the attackers may cause congestion in some approaches and free-flow conditions in the other approaches. This type of attack may cause disturbance to the traffic operations on a small scale near the attacker's vehicle for a short time period. Both spatial and temporal impacts are limited.
- To compromise network mobility. With this objective, attackers try to compromise the mobility performance of the system with falsified data, which makes the timing plans deviate from their optimality and thus generate excessive vehicle delay. Data spoofing attacks do not require the attack vehicles to be on the roadway. As a result, the attacks could last for a long time period and thus cause significant consequences to the traffic operation at the intersections under attack.

With different objectives, the impact could be represented by certain performance measures such as the number of potential crashes/conflicts (safety goal), total additional vehicle delays in the network (mobility goal), personal travel time/delay (personal goal), or a combination of those objectives.

2) Budget: The attacker has a limited budget, which means the number of CVs that could be compromised at the same time is limited. This requires the attackers to optimize their attacking strategies and target the most vulnerable locations of the system. The vulnerabilities are represented in different ways based on different scales, such as critical intersections in the traffic network [1] and critical traffic features (e.g., queue length, travel time) at one intersection [21]. In general, the severity of consequences would increase with an attacker's budget. However, the "benefit" is marginalized after some threshold. For example, [1] reveals that the total system travel time almost stopped increasing when the number of compromised traffic signals is beyond 20. This is because when the network already becomes congested due to attacks, new disruptions may not lead to worse system performance.

3) Knowledge About the System: The attacker's knowledge about the system determines the attack strategy. Existing attack studies mostly adopt the "white-box" scenario, which assumes the attackers acquire full knowledge about the TSC system in terms of control algorithms, model parameters, and implementation details. For example, a previous study by Chen *et al.* [22] assumes that the attackers know the source code of the signal control model and a comprehensive security analysis can be performed. The attackers can exhaustively try all the data spoofing options by analyzing the system design and implementation to understand the upper bound of the attack's effectiveness. Other attack studies with the assumption of direct manipulation of signal timing plans [1], [9], [11], [23]–[25] also belong to the "white-box" scenario. White-box attacks can be utilized by transportation agencies to identify the most critical vulnerabilities and prioritize the countermeasures. The second category is the "black-box" scenario, which assumes that the attackers only have limited knowledge of the TSC system. A recent study proposed a black-box attack scenario towards CV-based TSC systems with the first step to learning the control logic and then launching attacks based on the learned logic [21]. In the field of information engineering, such black-box analysis approaches have been widely used for identifying web application vulnerabilities [26], [27].

In summary, the impact of data spoofing attacks is determined by the attacker's objective, budget, and knowledge about the TSC system. If a trajectory spoofing attack event is donated as $\boldsymbol{B}$, then the risk of such attacks for a given TSC

system can be formulated as

$$R(\Omega) = \int_{\Omega} F(\boldsymbol{B}) p(\boldsymbol{B}) d\boldsymbol{B} \tag{1}$$

where $F(\boldsymbol{B})$ is the impact of the attack; $p(\boldsymbol{B})$ is the likelihood of the attack, and $\boldsymbol{\Omega}$ is a high-dimensional feasible attack space. The dimension is determined by the number of spoofed data elements considered in a trajectory (e.g., position, speed, acceleration, and heading) and their feasible ranges. The impact can be associated with attacker's budget and the likelihood is related to the attacker's knowledge about the system. Finally, the objective of the attacker is to maximize the risk.

B. Defense Solutions

The defense solutions should be generic to safeguard the system from different attack strategies and minimize risks. Since data spoofing attacks try to generate falsified vehicle trajectories (i.e., series of BSMs with spoofed GPS locations and vehicle speeds), the defense solutions should focus on detecting and filtering out falsified vehicle trajectories. Two defense solution directions are discussed in this section. One straightforward approach is to use additional data sources such as infrastructure-based sensors to cross-validate unknown trajectories. For example, Canepa and Claudel [28] formulated a mixed-integer linear feasibility problem to detect falsified trajectories. Detector data provide initial conditions and boundary conditions for the model. The information (e.g., average speed) brought by falsified trajectories may influence the estimation of traffic state, making the original mixed-integer linear problem infeasible. Shoukry *et al.* [29] also used legacy loop detectors to estimate macroscopic traffic states. A set of honest vehicles are then identified, whose velocity values are consistent with the macroscopic traffic states. These approaches, however, have limited applicability due to the requirement of additional data sources. A recent work by Suo and Sarma [30] proposed a security protocol called proof-of-travel, which determines the trust level of a vehicle based on its "reputation". The "reputation" is calculated from the infrastructure component by verifying its digital signatures along with its spatial movement. However, this approach cannot detect insider attackers who already have legitimate certificates.

The second approach is to model the defense problem as a misbehavior detection problem. As defined in Equation 1, the risk of attack B depends on both impact and likelihood. From the defender's perspective, the objective and budget of the attacker, the strategies of how the falsified data are constructed, and the likelihood of the attacks, are usually unknown. Without further information, it is generic and reasonable to assume that the system impact of any data spoofing attack and their occurrence probabilities are the same (i.e., the likelihood follows a high-dimensional uniform distribution) and non-negative. Based on this assumption, minimizing the risks is equivalent to minimizing the size of the high-dimensional space $\boldsymbol{\Omega}$. To this end, a generic and upgradable trajectory-based hierarchical defense (TBHD) solution was proposed to confine the feasible trajectory space (i.e., normal behaviors) to a smaller space at each level of defense. The TBHD consists of three hierarchies. Level 1 is a pointwise checking that examines whether data elements in the received BSMs fall within their feasible ranges. Level 2 is a multiple-point checking that examines whether the consecutive BSMs of one CV obey the laws of physics (e.g., relations between position, speed, and acceleration). Level 3 is a complete trajectory checking that examines whether the entire trajectory follows traffic flow models and/or is consistent with the estimated traffic state. Fig. 3 shows the feasible trajectory space with different defense levels. Ω_1, Ω_2, and Ω_3 represent the feasible trajectory spaces under defense levels 1, 2, and 3 respectively. Since the TBHD framework can significantly reduce the size of the high dimensional trajectory space level by level, it can greatly reduce the system impact from data spoofing attacks. More details of TBHD can be found in [31]. An example of the TBHD model will be introduced in the comprehensive study.

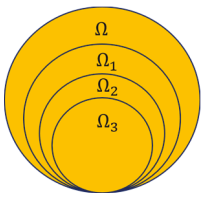


Fig. 3. Feasible trajectory space with different defense levels.

Note that the design of TBHD is flexible, different checking rules and models can be applied to each level. The higher levels of defense models require higher computation resources and longer time to execute but are also capable of detecting more sophisticated attacks. The benefit of the hierarchical structure design is that the defender could choose to enable defense levels based on available resources and estimated capability of the attackers.

C. Security Testing Platform

Due to the complex nature of the real TSC system and the sensitivity of the cybersecurity research, it is very difficult and unrealistic to directly implement the attack and defense methods at real world intersections. Therefore, developing a testing platform that supports evaluating different attack and defense models with desired TSC systems is highly beneficial. As illustrated in Fig. 2, the security testing platform contains three main components: a TSC system, a traffic environment, and the communication infrastructure that supports V2X communications. The proposed testing platform is implemented at Mcity, which is a closed CAV testing facility at the University of Michigan.[2]

Fig. 4 shows an overview of the testing platform. Note that this is a hybrid testing platform where the traffic is generated in microscopic simulation but the communication network utilizes real hardware devices (e.g., RSUs). For a detailed description of the testing platform, readers can refer to [32]. The traffic network of Mcity is built and calibrated in VISSIM, which is used to generate virtual traffic and proxy

[2]https://mcity.umich.edu/

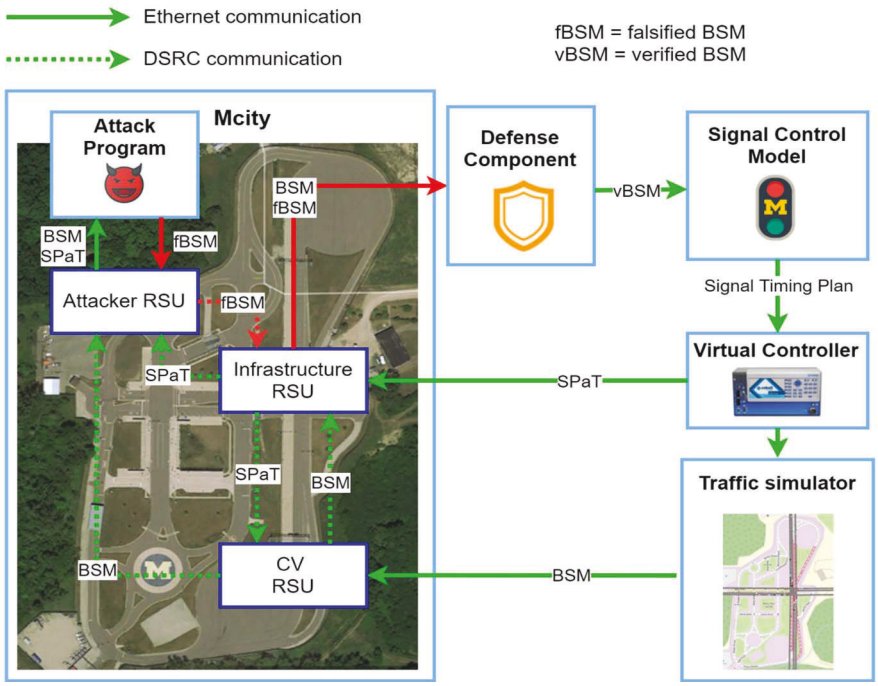


Fig. 4. Testing platform overview.

BSMs. Three RSUs in Mcity are employed for the testing. The CV RSU is used to broadcast BSMs from simulated vehicles, which mimics the V2V communication network. The infrastructure RSU connects to the TSC system. It broadcasts SPaT messages obtained from virtual signal controllers and receives BSMs from virtual vehicles. In a normal operation condition, the received BSMs will be utilized by the Signal Control Model for signal optimization. The Attacker RSU serves as the communication device for the attacker. It receives BSMs and SPaT and forwards them to the Attack Program, which generates falsified BSMs. The falsified BSMs are then broadcast through the Attack RSU. The falsified BSMs are mixed with the normal BSMs when the infrastructure RSU forwards both types of messages to the Defense Component. The Defense Component executes the detection algorithms to verify the normal BSMs and filter out the falsified BSMs. Finally, only verified BSMs are forwarded to the TSC system for generating optimal signal timing plans, which are then executed in the Virtual Controller. This design aims to mimic the real-world operation environment to a great extent in order to accurately evaluate the cybersecurity of different CV-based TSC systems.

In the next section, we will elaborate on how the proposed framework is applied to evaluate the cybersecurity of a selected TSC system with specific attack and defense models.

IV. A Comprehensive Case Study

In this section, a comprehensive case study on a CV based adaptive signal control model, Intelligent Signal (i.e., I-SIG), from the Multi-Modal Intelligent Traffic Signal System (MMITSS) Project [33], [34] is presented. The signal control model I-SIG will first be introduced. Specific attack and defense models are subsequently presented. Finally, the impact of data spoofing attacks on the operation of I-SIG and the effectiveness of the defense model is presented.

A. Signal Control System

Intelligent Signal Control (I-SIG) is a CV-based adaptive signal control model. Assuming the signal phasing is based on the dual-ring barrier structure, I-SIG optimizes the sequence and green duration of each phase using CV trajectory data. At the beginning of each barrier, I-SIG takes a snapshot of the trajectories received from all the CVs within the communication range. The trajectory data is converted to an arrival table, which contains the estimated time of arrival (ETA) information in each signal phase. Based on the arrival table, I-SIG solves a dynamic programming (DP) based optimization problem to find the optimal signal timing plan with the objective of either minimizing total delay or queue length. When the penetration rate of CV is low, a traffic state estimation algorithm named Estimation of Vehicle Location and Speed (EVLS) is executed to estimate the location and speed of regular vehicles based on CV trajectories, so that the complete ETA information can be obtained. I-SIG plans as many stages as needed so that all the vehicles can be properly served. After obtaining the optimal signal plan, I-SIG only executes the first stage (i.e., the four phases in the current barrier) of the plan and arranges the phase sequence of the second stage. When a new barrier starts, I-SIG repeats this optimization process. For more details about I-SIG, please refer to [35], [36].

B. Attack Model and Falsified Trajectory Generation

We follow the structure of the risk assessment (i.e., the attacker's objective, budget, and knowledge about the system)

to illustrate the attack model. It is assumed that the objective of the attacker is to increase the total vehicle. It is also assumed that there is only one attacker and she can only compromise one vehicle. This implies that one falsified trajectory can be added to the communication network at the same time. The reason for this setting is to increase the difficulty of launching attacks with prominent consequences. Regarding the knowledge of the system, a two-step "black-box" [21] attack scenario towards the I-SIG system is constructed. In the first step, the attacker tries to learn the control logic through observations using a surrogate model. Based on the surrogate model, in the second step, the attacker launches falsified data attacks to influence the control systems to make sub-optimal control decisions. It is found that the I-SIG model is vulnerable to two types of attacks: ETA attack and phantom queue attack. The ETA attack leverages the fact that I-SIG uses ETAs in the arrival table to evaluate delay. As a result, if the falsified trajectory could generate an abnormal ETA when the I-SIG takes a snapshot of the trajectories, the abnormal ETA could extend the green phases to an unnecessarily long time period. This will greatly increase the delay of red phases. The phantom queue attack leverages the traffic state estimation algorithm (i.e., EVLS) under lower CV penetration rates. The EVLS algorithm applies a shockwave model to estimate the queue length based on the stopped CVs. If the falsified trajectory could pretend to be a stopped CV, then a phantom queue is created. Subsequently, the signal optimization model allocates a longer green time to discharge the queue that does not exist. This will also increase the delay in other phases. For more details about the vulnerability analysis on I-SIG, please refer to [22].

In this case study, it is assumed that the attacker is aware of the vulnerabilities and she aims to generate a falsified trajectory $\boldsymbol{B} \in \Omega$ that fulfills the attack goals. The falsified trajectory generation is formulated as an optimization problem **P1**.

$$\max_{a(t)} \sum_{t=t_s}^{t_e} \left(d_l\left(t-\tau^w\right)-d^w-d\left(t\right)\right)^2 \tag{2}$$

$$\text{s.t. } g\left(d\left(t\right), v\left(t\right), a\left(t\right)\right)=0 \tag{3}$$

$$v\left(t\right)=d\left(t\right)-d(t-\Delta t); \quad a\left(t\right)=v\left(t\right)-v(t-\Delta t) \tag{4}$$

$$0 \le v\left(t\right) \le v_f; \quad a_{min} \le a\left(t\right) \le a_{max} \tag{5}$$

$$d_f\left(t\right)+d_0 \le d\left(t\right) \le d_l\left(t\right)-d_0 \tag{6}$$

$$v\left(t_s\right)=v_s; \quad d\left(t_s\right)=d_s \tag{7}$$

The objective of the optimization problem is to minimize the cumulative square error in car-following behavior. In this case, Newell's first-order car following model [37] is selected where τ^w and d^w are time and distance displacement in the model. d_l and d are the locations of the leading vehicle and falsified vehicle. t_s and t_e are the start and end time of the falsified trajectory. The objective function tries to mimic a normal car-following behavior to reduce the probability of being detected. Newell's model is chosen as an example. In reality, the attacker could use the received BSM data to calibrate a human-like car-following model as the objective function. Eq. (3) represents a general attack goal. The function $g\left(\cdot\right)$ could take different forms with different attack goals (e.g., ETA attack and phantom queue attack), which would be illustrated in the experiment design. Eqs. (4-5) represents the vehicle dynamics and boundaries of speed v and acceleration a. Eq. (6) guarantees that the falsified trajectory keeps a safe distance d_0 from the leading vehicle $d_l\left(t\right)$ and following vehicle $d_f\left(t\right)$ at any time t. Finally, Eq. (7) is the initial condition, so that the falsified trajectory enters the intersection area from a certain distance (e.g., communication range) with a certain speed.

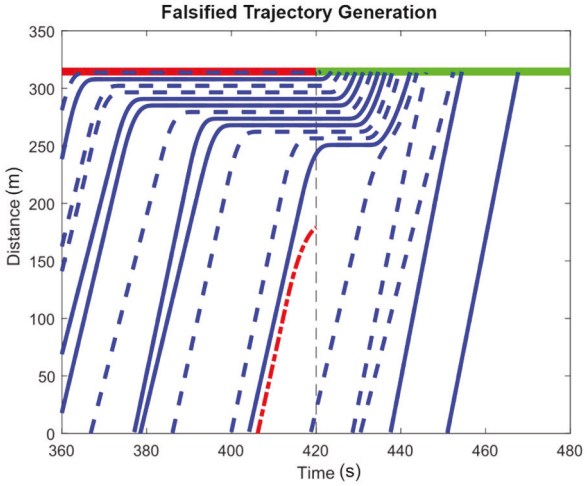


Fig. 5. Falsified trajectory generation with EAT attack goal (blue solid lines: observed cv trajectories; blue dashed lines: unobserved vehicle trajectories; red dashed line: falsified trajectory).

Fig. 5 reveals an example of the falsified trajectory with ETA attack as the goal under 50% CV penetration rate. From the signal control system's perspective, the solid blue curves represent normal trajectories that are observed (i.e., CVs) and the dashed blue curves represent normal trajectories that are not observed (i.e., non-CVs). The red dashed line represents the observed falsified trajectory generated by the attacker. It is generated at the boundary of the communication range (assuming 300m) and follows its leading vehicle at first. It then slows down to achieve the attack goal around time 420s (i.e., reach a predefined ETA).

C. Defense Model

As shown in Fig. 5, the objective of the defense model is to identify the red dashed line as a falsified trajectory. To make the defense problem more challenging, it is assumed that the attacker generates a falsified trajectory using the optimization problem illustrated in the previous section. The objective and constraints guarantee the generated falsified trajectory to be as "real" as possible. As a result, simple misbehavior detection methods, such as L1 and L2 of the TBHD are not able to identify the anomaly. More sophisticated models are necessary. To address this challenge, a data-driven approach to identify falsified trajectories is proposed, as the L3 model in the TBHD framework. The key idea is to calculate the similarities between the trajectories using a distance metric. Inspired by the word embedding model from the Natural Language

Processing (NLP) community [38], a trajectory embedding model is developed. First, each trajectory data point is encoded to a word based on different traffic features such as speed, range, and rate range. A neural network is used to train the context word pairs (i.e., trajectory points contains similar traffic state) and find a vector representation of each trajectory data point. The trajectory data that carry similar traffic state information would have similar vector representations. The distance metric then could be represented by the distance between vectors (e.g., L2 norm). A distance matrix can be calculated for each pair of trajectories. It is assumed that the behaviors of the falsified trajectories are different from the normal trajectories as shown in Fig. 5. As a result, the distance between a falsified trajectory and a normal trajectory should be greater than the distance between two normal trajectories. A hierarchical clustering algorithm is then adopted to group the trajectories and identify the falsified ones. More details can be found in [39].

D. Experiment Setup

A VISSIM simulation model is built based on the highway intersection at Mcity. The traffic demand for each movement is 350 vehicles per hour. The free flow speed is 60 km/h.

1) Signal Control Model: The traffic signals at this intersection include eight phases, following the dual-ring barrier structure. The phase sequence is fixed with a leading left-turn. I-SIG is adopted as the signal control model. The minimum green time is set to be 5 seconds and the maximum green time is set to be 30 seconds for each phase. The yellow interval is 2 seconds and the all-red clearance time is 2 seconds.

2) Attack Model: For each signal cycle, one falsified trajectory is generated at westbound through movement. It is assumed that, by analyzing historical BSM and SPaT data, the attacker knows when the signal optimization is executed [22], which is defined as the time of interest (ToI). The attacker starts generating the falsified trajectory T seconds before the ToI (i.e., t_s) and stops generating the falsified trajectory after ToI (i.e., t_e). The falsified trajectory is generated 300 meters from the stop bar to match the communication range with a resolution of 10 Hz. Each trajectory data point is encoded to the BSM and broadcast via the Attack RSU, as shown in Fig. 4. Two types of attacks are considered: ETA attack and phantom queue attack.

In the ETA attack, the attack goal in Eq. (3) is presented as follows:

$$v\left(t_e\right) > v_{stop}; \quad \frac{\left(d_s - d\left(t_e\right)\right)}{v\left(t_e\right)} = ETA \tag{8}$$

where v_{stop} is a speed threshold (2m/s), below which the vehicle is considered as a stopped vehicle. d_s is the location of the stop bar. In this case, ETA is set to be 64 seconds, the maximum time during which a vehicle could be served within the current barrier. In other words, to serve the vehicle with an ETA of 64s, both phases in the barrier need to be extended to the maximum green time.

In the phantom queue attack, the attack goal in Eq. (3) is shown as follows:

$$v\left(t_e\right) = 0; \quad \left(d_s - d\left(t_e\right)\right) \times k_j = Queue \tag{9}$$

where k_j is the jam density. The value of the queue is set to be 15 vehicles. Assuming that the saturation flow rate is 2 seconds per vehicle, 15 vehicles take 30 seconds to discharge, which equals to the maximum green time. Thus, a further increase in the queue value not only cannot maximize the impact but instead increases the probability of being detected.

3) Defense Model: The defense model, as shown in Fig. 4, is executed before the received trajectories are fed into the signal control model. All trajectories go through the defense model and are labeled either normal or fake using the trajectory embedding and hierarchical clustering methods. Only trajectories labeled as 'normal' would be sent to the signal control model for optimization.

E. Experiment Results

The simulation resolution is set to be 10 Hz, which is consistent with the transmission rate of BSM and SPaT. First, 10 hours of simulation under normal signal operations are executed from which all vehicle trajectories are collected. These vehicle trajectories are used to train the trajectory embedding model. Since the detection experiments are performed with simulated traffic, simulated vehicle trajectories are used to train the model. If the detection model is implemented at real world intersections, then local traffic pattern and driving behavior data from real drivers should be collected to train the model. Then three experimental scenarios are considered. In scenario 1, the CV market penetration rate is set to be 100%, and ETA attacks are launched. In scenario 2, the CV market penetration rate is set to be 50%, and ETA attacks are launched. In scenario 3, the CV market penetration rate is set to be 50%, and the phantom queue attacks are launched. For each scenario, three experiments are conducted: (i) normal signal operation, (ii) operation with attack, and (iii) operation with both attack and defense. Each experiment lasts for 1 hour, with the first 5 minutes taken as the warm-up time.

An example of the attack and defense process is shown in Fig. 6 under 100% CV penetration rate. Fig. 6(a) shows the falsified trajectory generated under the ETA attack goal and Fig. 6(b) shows the clustering result. For easy identification, the falsified vehicle is always labeled as 1 in all cases. Other vehicles are labeled (starting from 2) based on the time they enter the communication range. The red dashed line (set as 0.3 in this case) represents the threshold to cut the dendrogram in the hierarchical clustering algorithm, which is a calibrated value from data. Because of the abnormal behavior (slowing down when the front vehicle is still far away), the distance metric between the falsified trajectory and other trajectories is greater. Therefore, the clustering method is able to identify the falsified vehicle. Similarly, Fig. 6(c) shows the falsified trajectory generated under the phantom queue attack goal, which is also identified by the clustering algorithm, as shown in Fig. 6(d). When there is no attack goal (Fig. 6(e)), the falsified trajectory simply travels at free-flow speed. The clustering algorithm fails to identify in this case (Fig. 6(f)).

All experiment results are summarized in Table II. Average vehicle delay is used as the performance index because the attacker's target is to compromise the efficiency of the signal

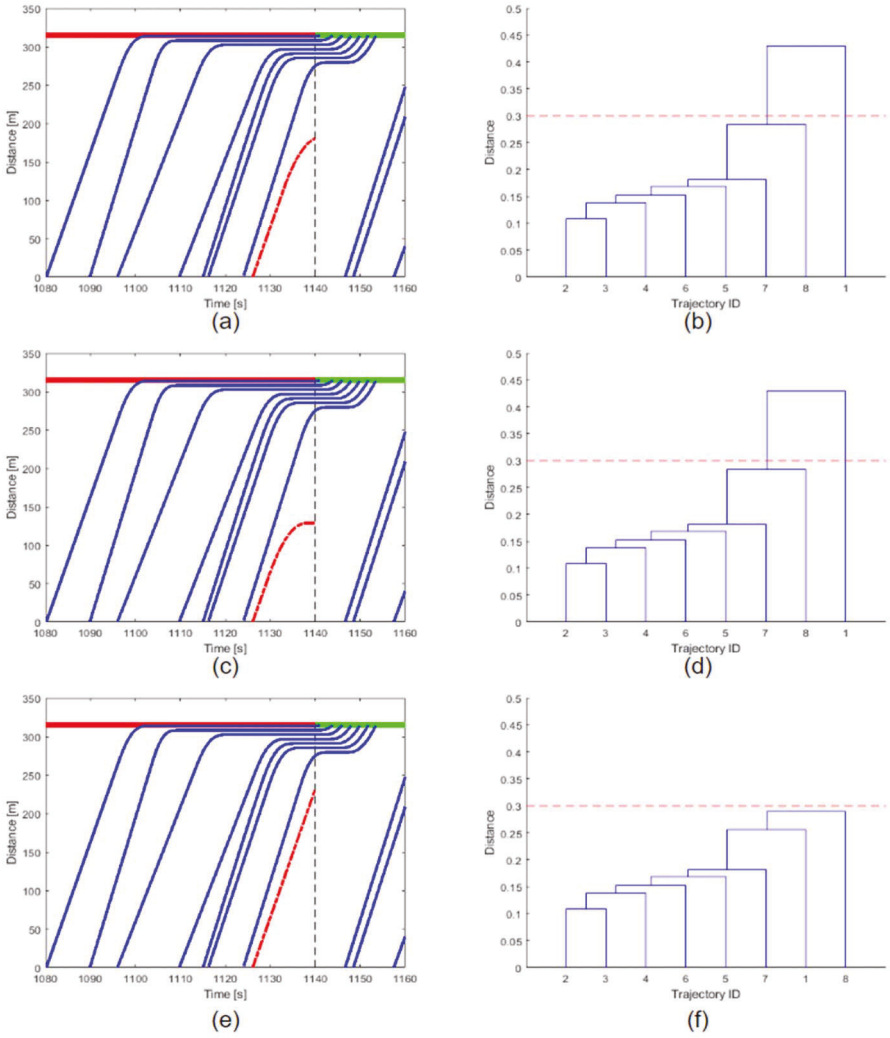


Fig. 6. An illustrative example of the attack and defense process. (Subfigures (a), (c), (e) show the falsified trajectories under ETA attack goal, phantom queue attack goal, and no attack goal; Subfigures (b), (d), (f) show the classification results).

TABLE I

AVERAGE VEHICLE DELAY COMPARISON FOR EACH EXPERIMENT

Scenario	PR	Attack goal	Experiment	Description	Average delay [s/veh]	Percentage Increase in Delay
1	100%	ETA = 64 s	1	Normal operation	39.50	-
			2	Attack w/o defense	48.58	23.0%
			3	Attack w/ defense	41.01	3.8%
2	50%	ETA = 64 s	4	Normal operation	43.38	-
			5	Attack w/o defense	51.03	17.6%
			6	Attack w/ defense	44.80	3.3%
3	50%	QUEUE = 15 veh	7	Normal operation	43.38	-
			8	Attack w/o defense	47.16	8.7%
			9	Attack w/ defense	45.26	4.3%

operations. It can be seen that in all scenarios, attacks with different goals all increase the system delay. The ETA attacks are more effective, which causes an additional 23.0% delay in scenario 1 and 17.6% in scenario 2. The phantom queue attack, although less severe, also increases an additional 8.7% vehicle delay. The main reason for the difference is that the ETA attack is able to extend both phases in a barrier to their maximum values while the phantom queue attack is only able to extend the leading phase to its maximum value [22]. Note that in all scenarios, only one falsified trajectory is allowed to be added per signal cycle, so that the attack is quite significant. When the defense model is implemented, the increase in delay is significantly reduced, as shown in experiments 3, 6, and 9. In other words, the proposed defense

TABLE II
AVERAGE VEHICLE DELAY COMPARISON FOR EACH EXPERIMENT

Experiment	Cycles	Attacks launched	Attack success rate	Falsified Trajectories Tested by defense	Detection rate	Regular Trajectories Tested by Defense	False alarm rate
3	42	22	52.4%	21	100.0%	2486	1.1%
6	42	32	76.2%	30	100.0%	1086	4.9%
9	41	30	73.2%	21	95.2%	1071	3.7%

model can successfully safeguard the system and detect most of the falsified trajectories.

To better understand the performance of the attack and defense models, the attack success rate, detection rate, and false alarm rate of experiments 3, 6, and 9 are presented in Table. In experiment 3, a total of 42 cycles are recorded and 22 attacks are launched. The attack success rate is 52.4%. The low success rate indicates that almost half of the time, the falsified trajectory generation model could not always find feasible solutions during the generation period. A total number of 21 falsified trajectories go through the defense component and all of them are successfully identified (1 falsified trajectory is not identified by the defense model because there are not enough 'normal' trajectories being observed in the cycle). The false alarm rate is as low as 1.1%. In experiments 6 and 9, 32, and 30 attacks are launched, respectively. The attack success rates increase to 76.2% and 73.2%, respectively. This matches our expectations because there is more room for generating falsified trajectories in time and space when the CV penetration rate is lower. Note that the attacker (as well as the TSC system) is only able to observe CVs. In experiments 6 and 9, 30 and 21 falsified trajectories are identified by the defense model, respectively. The detection rate is 100% for experiment 6 and 95.2% for experiment 9. The false alarm rates are 4.9% for experiment 6 and 3.7% for experiment 9. The increased false alarm rate is due to the low CV penetration rate, in which the relations (e.g., car-following) between CV trajectories contain more uncertainties. Results from Table also explain the slight increase of delay in experiments 3, 6, and 9 when the defense model is applied. First, some real trajectories are incorrectly labeled as falsified, therefore they are not properly served by I-SIG. Second, a few falsified trajectories are not identified when there are not enough observations.

V. DISCUSSIONS

In this study, we consider data spoofing attacks using compromised CVs as the threat model and introduce both attack strategies and defense solutions towards a CV based TSC system. In this section, we further discuss attack types, attack goals, and defense strategies under different operational environments that contribute to multiple aspects of the cyber-security of TSC systems.

A. Attack Types

The data spoofing attack is not the only type of attack that could be launched towards the TSC system. Wang *et al.* [40] summarized a list of attack methods, including impersonation/masquerading, Sybil, spoofing, cheating with sensor messages, denial of service (DoS), black hole, reply, delay and suppression, and collusion. The spoofing attack considered in this paper requires a legal identity to transmit falsified messages. As a result, the attacker's capability may be limited by the high cost (e.g., the cost of purchasing vehicles/devices that possess legal identities). Other types of attacks may cause a more severe impact at a lower cost. For instance, the DoS attack could cause congestion in the communication channel. Consequently, the TSC system may receive fewer BSMs and allocate shorter green time than needed. A comprehensive study of different types of attacks targeting the TSC system is indispensable.

B. Attack Goals

There are different components in the TSC systems including data collection, traffic state estimation, optimization model, and implementation details. Therefore, attack surfaces and goals may be TSC system-specific. One attack model that is effective to a certain TSC system may be less effective to another. Taking I-SIG as an example, it uses ETA and number of approaching vehicles (including the queuing vehicles) as the critical traffic features in determining the signal timing plan, other TSC systems may use other traffic features such as travel time [41], queue length [42], vehicle delay [43], and flow rate [44] for signal optimization. Therefore, attack goals have to be changed based on the vulnerability analysis of each TSC system.

C. Defender Strategies

The proposed defense model performs satisfactorily on detecting falsified trajectories if an attacker has a goal. However, if an attacker does not have any goal, as shown in Fig. 6(e) and Fig. 6(f), the proposed method may not work. Fortunately, when an attacker has no specific goal or less aggressive goals (e.g., set ETA to a more conservative value), the impact on the TSC system typically is also less severe. In other words, an attacker has to make a balance between the attack aggressiveness and the probability been detected. The same principle applies to the defender. In the hierarchical clustering algorithm, if the threshold is set to a high value, then both the detection rate and false alarm rate decrease. Because the dissimilarity between the normal trajectories and falsified trajectory has to be greater than the threshold to be identified. If the threshold is set to a low value, the detection rate increases, but so does the false alarm rate. A high false alarm

rate would mistakenly filter too many normal trajectories, which in return reduces the data quality.

D. Operational Environment

The operational environment (e.g., CV penetration rate and traffic volume) also plays an important role in determining attack and defense strategies. Generally, CV-based TSC systems work better under a higher CV penetration rate. This is evidenced by numerous previous studies such as [36], [41], [45], and our experiments as well. The average vehicle delay in experiment 1 (PR = 100%) is about 10% lower than the average delay in experiments 4 and 7 (PR = 50%). Moreover, the attack success rate is relatively low when the CV penetration rate is high. This is because the feasible space for falsified trajectories generation shrinks with more observed CVs. The false alarm rate of the defense model under a higher CV penetration rate is also lower because of fewer patterns in the observed trajectories. Overall, higher CV penetration rates indicate a safer and more efficient operating environment. Traffic volume also influences the performance in a similar way. Under the same CV penetration rate, higher traffic volumes mean more observed CVs, which increases the attack difficulty and improves the performance of the defense model. Other operational environment parameters such as GPS error and communication delay would also influence the model performance, which requires further investigation.

E. Other TSC Applications

This work only studies the cybersecurity problem of adaptive traffic signal control systems considering passenger vehicles. The TSC system contains other components such as signal priority and pedestrian signal. Lately, connected vehicle technology has been also applied to these areas, which may bring security issues. For example, in the impersonation attack, the attacker's vehicle could pretend to be an emergency vehicle and request signal priority to achieve personal benefits, while other vehicles suffer from long delays. Other recent studies [45]–[49] focus on cooperative driving, in which traffic signal parameters and CAV trajectories are optimized simultaneously. Under such circumstances, the data spoofing attack may not only affect traffic signal operations but also CAV trajectory planning.

VI. Conclusion

In this paper, the cybersecurity problem of the TSC system in a CV environment is systematically investigated. Potential threats of the major components of a TSC system are analyzed and the data spoofing attack is considered the most realistic and plausible threat model. A cybersecurity analysis framework is then proposed including risk assessment, defense solutions, and a testing platform implemented in real-world transportation infrastructure. A comprehensive case study is presented to show how the framework is applied to a selected TSC system. Experiment results show the impact of the data spoofing attack on increasing system delay and the effectiveness of the defense model in detecting and filtering falsified trajectories. Lastly, the discussion section summarizes the current study and layouts several directions for future research.

Acknowledgment

The views presented in this paper are those of the authors alone.

References

[1] K. A. Perrine, M. W. Levin, C. N. Yahia, M. Duell, and S. D. Boyles, "Implications of traffic signal cybersecurity on potential deliberate traffic disruptions," *Transp. Res. A, Policy Pract.*, vol. 120, pp. 58–70, Feb. 2019, doi: 10.1016/j.tra.2018.12.009.

[2] B. W. Lampson, "Computer security in the real world," *Computer*, vol. 37, no. 6, pp. 37–46, Jun. 2004, doi: 10.1109/MC.2004.17.

[3] W. Whyte, A. Weimerskirch, V. Kumar, and T. Hehn, "A security credential management system for V2 V communications," in *Proc. IEEE Veh. Netw. Conf.*, Dec. 2013, pp. 1–8, doi: 10.1109/VNC.2013.6737583.

[4] (Nov. 18, 2020). *FCC Modernizes 5.9 GHz Band to Improve Wi-Fi and Automotive Safety*. Federal Communications Commission. Accessed: Nov. 23, 2021. [Online]. Available: https://www.fcc.gov/document/fcc-modernizes-59-ghz-band-improve-wi-fi-and-automotive-safety

[5] K. Ansari, "Joint use of DSRC and C-V2X for V2X communications in the 5.9 GHz ITS band," *IET Intell. Transp. Syst.*, vol. 15, no. 2, pp. 213–224, 2021.

[6] P. B. Hunt, "SCOOT–A traffic responsive method of coordinating signals," *Transp. Road Res. Lab. TRRL*, vol. 1, no. 1, pp. 1–41, Jan. 1981.

[7] A. G. Sims and K. W. Dobinson, "The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits," *IEEE Trans. Veh. Technol.*, vol. 29, no. 2, pp. 130–137, May 1980.

[8] B. Ghena, W. Beyer, A. Hillaker, J. Pevarnek, and J. A. Halderman, "Green lights forever: Analyzing the security of traffic infrastructure," in *Proc. 8th USENIX Conf. Offensive Technol.*, Berkeley, CA, USA, 2014, p. 7. [Online]. Available: http://dl.acm.org/citation.cfm?id=2671293.2671300

[9] A. Laszka, B. Potteiger, Y. Vorobeychik, S. Amin, and X. Koutsoukos, "Vulnerability of transportation networks to traffic-signal tampering," in *Proc. ACM/IEEE 7th Int. Conf. Cyber-Phys. Syst. (ICCPS)*, Apr. 2016, pp. 1–10, doi: 10.1109/ICCPS.2016.7479122.

[10] A. Ghafouri, W. Abbas, Y. Vorobeychik, and X. Koutsoukos, "Vulnerability of fixed-time control of signalized intersections to cyber-tampering," in *Proc. Resilience Week (RWS)*, Aug. 2016, pp. 130–135, doi: 10.1109/RWEEK.2016.7573320.

[11] 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," *Transp. Res. B, Methodol.*, vol. 91, pp. 366–382, Sep. 2016, doi: 10.1016/j.trb.2016.05.017.

[12] M. Prigg. (May 2014). *Has New York's Traffic Light System Been HACKED*. Accessed: Jun. 29, 2020. [Online]. Available: http://www.dailymail.co.U.K./sciencetech/article-2617228/New-Yorks-traffic-lights-HACKED-technique-work-world.html

[13] Y. Feng, S. Huang, Q. A. Chen, H. X. Liu, and Z. M. Mao, "Vulnerability of traffic control system under cyberattacks with falsified data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 1, pp. 1–11, Dec. 2018, doi: 10.1177/0361198118756885.

[14] D. Bezzina and J. R. Sayer, "Safety pilot model deployment: Test conductor team report," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 812 171, 2015.

[15] USDOT. *Intelligent Transportation Systems–Connected Vehicle Pilot Deployment Program*. Accessed: Jun. 29, 2020. [Online]. Available: https://www.its.dot.gov/pilots/

[16] (Sep. 14, 2015). *FACT SHEET: Administration Announces New 'Smart Cities' Initiative to Help Communities Tackle Local Challenges and Improve City Services*. Accessed: Jun. 29, 2020. [Online]. Available: https://obamawhitehouse.archives.gov/the-press-office/2015/09/14/fact-sheet-administration-announces-new-smart-cities-initiative-help (accessed Jun. 29, 2020).

[17] S. Janaarthanan, "Security analysis of vehicle to vehicle arada locomate on board unit," M.S. thesis, Iowa State Univ., Ames, IA, USA, 2019.

[18] A. Ghosal and M. Conti, "Security issues and challenges in V2X: A survey," *Comput. Netw.*, vol. 169, Mar. 2020, Art. no. 107093, doi: 10.1016/j.comnet.2019.107093.

[19] U.S. Department of Homeland Security. (2015). *Transportation Systems Sector Cybersecurity Framework Implementation Guide*. Accessed: Jun. 29, 2020. [Online]. Available: https://www.cisa.gov/publication/tss-cybersecurity-framework-implementation-guide

[20] National Cooperative Highway Research Program, *Transportation Research Board, and National Academies of Sciences, Engineering, and Medicine, Security 101: A Physical Security Primer for Transportation Agencies*, Transportation Research Board, Washington, DC, USA, 2009, doi: 10.17226/22998.

[21] S. Huang, Y. Feng, W. Wong, Q. A. Chen, Z. M. Mao, and H. X. Liu, "Impact evaluation of falsified data attacks on connected vehicle based traffic signal control systems," in *Proc. 3rd Int. Workshop Automot. Auto. Vehicle Secur.*, 2021, p. 25.

[22] Q. A. Chen, Y. Yin, Y. Feng, Z. M. Mao, and H. X. Liu, "Exposing data spoofing attacks on CV-based traffic signal control," presented at the 25th Netw. Distrib. Syst. Secur. Symp. (NDSS), 2018.

[23] J. M. Ernst and A. J. Michaels, "Framework for evaluating the severity of cybervulnerability of a traffic cabinet," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2619, no. 1, pp. 55–63, Jan. 2017, doi: 10.3141/2619-06.

[24] A. A. Ganin, A. C. Mersky, and A. S. Jin, "Resilience in intelligent transportation systems (ITS)," *Transp. Res. C, Emerg. Technol.*, vol. 100, pp. 318–329, Oct. 2019, doi: 10.1016/j.trc.2019.01.014.

[25] C.-C. Yen, D. Ghosal, M. Zhang, C.-N. Chuah, and H. Chen, "Falsified data attack on backpressure-based traffic signal control algorithms," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Dec. 2018, pp. 1–8, doi: 10.1109/VNC.2018.8628334.

[26] J. Bau, E. Bursztein, D. Gupta, and J. Mitchell, "State of the art: Automated black-box web application vulnerability testing," in *Proc. IEEE Symp. Secur. Privacy*, May 2010, pp. 332–345, doi: 10.1109/SP.2010.27.

[27] A. Doupé, M. Cova, and G. Vigna, "Why Johnny can't pentest: An analysis of black-box web vulnerability scanners," in *Detection of Intrusions and Malware, and Vulnerability Assessment*, Berlin, Germany: Springer, 2010, pp. 111–131, doi: 10.1007/978-3-642-14215-4_7.

[28] E. S. Canepa and C. G. Claudel, "Spoofing cyber attack detection in probe-based traffic monitoring systems using mixed integer linear programming," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2013, pp. 327–333, doi: 10.1109/ICCNC.2013.6504104.

[29] Y. Shoukry, S. Mishra, Z. Luo, and S. Diggavi, "Sybil attack resilient traffic networks: A physics-based trust propagation approach," in *Proc. ACM/IEEE 9th Int. Conf. Cyber-Phys. Syst. (ICCPS)*, Apr. 2018, pp. 43–54, doi: 10.1109/ICCPS.2018.00013.

[30] D. Suo and S. E. Sarma, "Proof-of-travel: A protocol for trustworthy V2I communication and incentive designs," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Dec. 2020, pp. 1–4, doi: 10.1109/VNC51378.2020.9318330.

[31] W. Wong, S. Huang, Y. Feng, Q. A. Chen, Z. M. Mao, and H. X. Liu, "Trajectory-based hierarchical defense model to detect cyber-attacks on transportation infrastructure," presented at the Transp. Res. Board 98th Annu. Meeting Transp. Res. Board, 2019. Accessed: Mar. 23, 2021. [Online]. Available: https://trid.trb.org/view/1572995

[32] Y. Feng, C. Yu, S. Xu, H. X. Liu, and H. Peng, "An augmented reality environment for connected and automated vehicle testing and evaluation," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1549–1554, doi: 10.1109/IVS.2018.8500545.

[33] USDOT. *Intelligent Transportation Systems–Dynamic Mobility Applications (DMA)*. Accessed: Jun. 29, 2020. [Online]. Available: https://www.its.dot.gov/research_archives/dma/bundle/mmitss_plan.htm

[34] K. Ahn, H. A. Rakha, K. Kang, and G. Vadakpat, "MMITSS simulation model development and assessment," *Transp. Res. Rec., J. Transp. Res. Board*, to be published.

[35] S. Sen and K. L. Head, "Controlled optimization of phases at an intersection," *Transp. Sci.*, vol. 31, no. 1, pp. 5–17, 1997.

[36] Y. Feng, K. L. Head, S. Khoshmagham, and M. Zamanipour, "A real-time adaptive signal control in a connected vehicle environment," *Transp. Res. C, Emerg. Technol.*, vol. 55, pp. 460–473, Jun. 2015, doi: 10.1016/j.trc.2015.01.007.

[37] G. F. Newell, "A simplified car-following theory: A lower order model," *Transp. Res. B, Methodol.*, vol. 36, no. 3, pp. 195–205, 2002, doi: 10.1016/S0191-2615(00)00044-8.

[38] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*. Accessed: Jul. 25, 2020. [Online]. Available:

[39] S. E. Huang, Y. Feng, and H. X. Liu, "A data-driven method for falsified vehicle trajectory identification by anomaly detection," *Transp. Res. C, Emerg. Technol.*, vol. 128, Jul. 2021, Art. no. 103196, doi: 10.1016/j.trc.2021.103196.

[40] P. Wang, X. Wu, and X. He, "Modeling and analyzing cyberattack effects on connected automated vehicular platoons," *Transp. Res. C, Emerg. Technol.*, vol. 115, Jun. 2020, Art. no. 102625, doi: 10.1016/j.trc.2020.102625.

[41] J. Lee, B. Park, and I. Yun, "Cumulative travel-time responsive real-time intersection control algorithm in the connected vehicle environment," *J. Transp. Eng.*, vol. 139, no. 10, pp. 1020–1029, Oct. 2013, doi: 10.1061/(ASCE)TE.1943-5436.0000587.

[42] C. Priemer and B. Friedrich, "A decentralized adaptive traffic signal control using V2I communication data," in *Proc. 12th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2009, pp. 1–6, doi: 10.1109/ITSC.2009.5309870.

[43] J. Wu, D. Ghosal, M. Zhang, and C.-N. Chuah, "Delay-based traffic signal control for throughput optimality and fairness at an isolated intersection," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 896–909, Feb. 2018, doi: 10.1109/TVT.2017.2760820.

[44] J. Zheng *et al.*, "Traffic signal optimization using crowdsourced vehicle trajectory data," presented at Transp. Res. Board 97th Annu. Meeting Transp. Res. Board, 2018. Accessed: Jun. 29, 2020. [Online]. Available: https://trid.trb.org/view/1496953

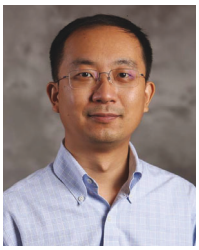
[45] H. Liu, X.-Y. Lu, and S. E. Shladover, "Traffic signal control by leveraging cooperative adaptive cruise control (CACC) vehicle platooning capabilities," *Transp. Res. C, Emerg. Technol.*, vol. 104, pp. 390–407, Jul. 2019, doi: 10.1016/j.trc.2019.05.027.

[46] Z. F. Li, L. Elefteriadou, and S. Ranka, "Signal control optimization for automated vehicles at isolated signalized intersections," *Transp. Res. C, Emerg. Technol.*, vol. 49, pp. 1–18, Dec. 2014, doi: 10.1016/j.trc.2014.10.001.

[47] C. Yu, Y. Feng, H. X. Liu, W. Ma, and X. Yang, "Integrated optimization of traffic signals and vehicle trajectories at isolated urban intersections," *Transp. Res. B, Methodol.*, vol. 112, pp. 89–112, Jun. 2018, doi: 10.1016/j.trb.2018.04.007.

[48] Y. Feng, C. Yu, and H. X. Liu, "Spatiotemporal intersection control in a connected and automated vehicle environment," *Transp. Res. C, Emerg. Technol.*, vol. 89, pp. 364–383, Apr. 2018, doi: 10.1016/j.trc.2018.02.001.

[49] Y. Guo, C. Xiong, J. Ma, and X. Li, "Joint optimization of vehicle trajectories and intersection controllers with connected automated vehicles: Combined dynamic programming and shooting heuristic approach," *Transp. Res. C, Emerg. Technol.*, vol. 98, pp. 54–72, Jan. 2019, doi: 10.1016/j.trc.2018.11.010.



Yiheng Feng received the B.S. and M.E. degrees from the Department of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2005 and 2007, respectively, and the Ph.D. degree in systems and industrial engineering from the University of Arizona in 2015. He is currently an Assistant Professor with the Lyles School of Civil Engineering, Purdue University. His research interests include traffic signal systems control and security and CAV testing and evaluation.

Shihong Ed Huang received the bachelor's degree in traffic engineering from Beijing Jiaotong University, China, in 2013, the master's degree in civil engineering from the University of Michigan, Ann Arbor, USA, in 2016, and the Ph.D. degree from the Department of Civil and Environmental Engineering, in August 2020. His current research interests include traffic control and optimization, traffic flow theory, and cybersecurity of transportation systems in a connected vehicle environment.

Wai Wong received the bachelor's degree (Hons.) in civil engineering and the Ph.D. degree in transportation and traffic engineering from the Department of Civil Engineering, The University of Hong Kong, in 2013 and 2017, respectively. He is currently a Lecturer with the Department of Civil and Natural Resources Engineering, University of Canterbury, New Zealand. With his strong interests in the future transportation system and smart city development, his research work mainly focuses on CV systems, transportation big data analytics, intelligent transportation system (ITS), cybersecurity in CV environment, traffic flow theory, and statistical modeling.

Qi Alfred Chen received the Ph.D. degree from the University of Michigan in 2018. He is currently an Assistant Professor with the Department of Computer Science, University of California at Irvine, Irvine. His current research interests span software and AI security, systems security, network security, and security problems at the AI and software stacks in autonomous CPS and IoT systems (e.g., autonomous driving and intelligent transportation). He was a recipient of NSF Career Award and ProQuest Distinguished Dissertation Award with the University of Michigan.

Z. Morley Mao received the B.S., M.S., and Ph.D. degrees from the University of California at Berkeley, Berkeley, CA, USA. She is currently a Professor with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. She was a recipient of the NSF Career Award, the Sloan Fellowship, and the IBM Faculty Partnership Award. She has been named the Morris Wellman Faculty Development Professor.

Henry X. Liu (Member, IEEE) received the bachelor's degree in automotive engineering from Tsinghua University, China, in 1993, and the PhD. degree in civil and environment engineering from the University of Wisconsin-Madison in 2000. He is currently a Professor in the Department of Civil and Environmental Engineering and the Director of Mcity at the University of Michigan, Ann Arbor. He is also a Research Professor at the University of Michigan Transportation Research Institute and the Director for the Center for Connected and Automated Transportation (USDOT Region 5 University Transportation Center). From August 2017 to August 2019, Prof. Liu served as DiDi Fellow and Chief Scientist on Smart Transportation for DiDi Global, Inc., one of the leading mobility service providers in the world. Prof. Liu conducts interdisciplinary research at the interface of transportation engineering, automotive engineering, and artificial intelligence. Specifically, his scholarly interests concern traffic flow monitoring, modeling, and control, as well as testing and evaluation of connected and automated vehicles. Prof. Liu is the managing editor of Journal of Intelligent Transportation Systems.