Detection of Cyberattacks at Connected Signalized Intersections: from the Physical Layer Perspective of Cyber-Physical Systems

Journal:	*IEEE Transactions on Intelligent Vehicles*
Manuscript ID	T-IV-24-01-0083
Manuscript Type:	Full Length Article
Date Submitted by the Author:	05-Jan-2024
Complete List of Authors:	Gu, Yingfan; University of Cincinnati, Department of Civil & Architectural Engineering & Construction Li, Zhixia; University of Cincinnati, Department of Civil and Architectural Engineering and Construction Management Zhang, Yunpeng; University of Houston, Department of Information Science Technology Tiwari, Shivam; University of Houston, Department of Information Science Technology Wei, Heng; University of Cincinnati, Department of Civil and Architectural Engineering and Construction Management Zhang, Guohui; University of Hawai'i at Manoa, Department of Civil and Architectural Engineering and Construction Management Ma, Muting; The University of Alabama, Culverhouse College of Business Baidya, Sabur; University of Louisville, Department of Computer Science and Engineering
Keywords:	System integration, safety and security;System design, modeling and deployment, vehicle system, human factors, Intelligent Transport Systems
Abstract:	Detecting cyberattacks is crucial for securing the connected transportation system. Previous research has predominantly focused on the detection from the perspective of cyber layer. This type of detection carries the risk that the cyber layer detection itself may become a target of the attacks. In contrast, standalone sensor-data-based algorithms that avoid communication to detect cyberattacks based on patterns of vehicle trajectories from the physical layer perspective of a cyber-physical system is promising to become a high-efficiency alternative to avoid the algorithm being attacked. This study develops a vehicle trajectory-oriented machine-learning algorithm to detect cyberattacks. Vehicle trajectories and driving behaviors are captured using a driving simulator under normal and cyberattack scenarios. Cyberattack detection models were constructed using Hidden Markov Models (HMMs), namely HMM-4-C. Whether a vehicle is being attacked or not is considered as a hidden state, and the trajectory data serve as the observation sequence. Expectation Maximization (EM) algorithm is applied to estimate the likelihood parameters of HMM-4-C. The results indicate that 98% of cyberattack scenarios in our research were successfully detected using HMM-4-C. Furthermore, we first propose the Effective Detection Period (EDP) and Effective Detection Distance (EDD) from the perspective of a cyberattack protector based on the potential real-world setting. The proposed method effectively detects cyberattacks at connected signalized intersections, which can be used to develop proactive cyberattack detection strategies in conjunction with cyber- perspective

	detection algorithms.

Response to Reviewers' Comments

T-IV-23-09-2962

We would like to take this opportunity to thank the editor and all the reviewers for their insightful and valuable comments and suggestions that help us further improve the manuscript. We really appreciate all the comments. Based on the review comments, we have thoroughly revised the manuscript. **In this document, we are providing our detailed response to each of the comments in this document.** Thank you very much again for your time and expertise in helping improve the manuscript. After the approval from the editor and the editorial office, we are invited to submit this revised manuscript as a new paper. In this submission, *we used red font in the manuscript to reflect the additions/changes/revisions we made to the manuscript in this version.*

Reviewer #1:

Comment 1: In Abstract, it is stated that "the proposed method effectively detects future cyberattacks". How is this conclusion reached since it is not discussed in detail in the paper? What is a "future cyberattack"?

Author Response: Thank you for bringing up this comment. We agree that there is an unclear statement regarding that "the proposed method effectively detects future cyberattacks". The "future cyberattack" means the cyberattack on the following vehicle in the cyberattack scenarios. For greater clarity and precision, we have fully updated the abstract.

Author Action: Really appreciate the reviewer's comment about the abstract. The revised abstract has been updated in the manuscript. We have revised the abstract to ensure greater clarity and precision in our manuscript. The abstract in the manuscript is: "The proposed method effectively detects cyberattacks at connected signalized intersections, which can be used to develop proactive cyberattack detection strategies in conjunction with cyber- perspective detection algorithms."

Comment 2: Section II "Literature Review" requires improvement. There are many typos. When introducing authors of a reference, the authors should use "et al." if there are more than two authors, instead of only listing the first author. Please examine such related studies as Detecting Data Spoofing in Connected Vehicle based Intelligent Traffic Signal Control using Infrastructure-Side Sensors and Traffic Invariants; Resilient event-triggered control of connected automated vehicles under cyber-attacks; A Survey of Cyber Attacks on Cyber Physical Systems: Recent Advances and Challenges

Author Response: Thank you for bringing this up. We agreed and updated the literature section to compare the methodology in this research with other existing studies and correct the typos. We also have referenced more researchers' work in the literature, further enriching our review and providing a comprehensive context for our study. These additions aim to highlight the unique contributions of our work and situate it more clearly within the existing body of research.

Author Action: Really appreciate the reviewer's comment about the literature review. The revised literature has been updated in the manuscript[1]–[3].

[1] N. Zhao, X. Zhao, N. Xu, and L. Zhang, "Resilient Event-Triggered Control of Connected Automated Vehicles Under Cyber Attacks," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 12, pp. 2300–2302, Dec. 2023, doi: 10.1109/JAS.2023.123483.

[2] J. Shen, Z. Wan, Y. Luo, Y. Feng, Z. M. Mao, and Q. A. Chen, "Detecting Data Spoofing in Connected Vehicle based Intelligent Traffic Signal Control using Infrastructure-Side Sensors and Traffic Invariants," in *2023 IEEE Intelligent Vehicles Symposium (IV)*, Anchorage, AK, USA: IEEE, Jun. 2023, pp. 1–8. doi: 10.1109/IV55152.2023.10186689.

[3] W. Duo, M. Zhou, and A. Abusorrah, "A Survey of Cyber Attacks on Cyber Physical Systems: Recent Advances and Challenges," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 5, pp. 784–800, May 2022, doi: 10.1109/JAS.2022.105548.

Comment 3: After reviewing all related studies, please summarize the differences between them and this work to highlight the contributions of this paper.

Author Response: We appreciate the reviewer's comment about highlight the contributions of this manuscript. After reviewing all related studies, we have identified several research gaps pertinent to our work and are striving to make meaningful contributions to current research in the field.

Author Action: Really appreciate the reviewer's comment about the literature. The revised literature section has been updated in the manuscript. The research gaps have been presented in the manuscript after reviewing all related studies as follows:

- "Recent research has enhanced our understanding of the security risks and communication detection methods in connected vehicle streams. However, there has been less emphasis on cyberattacks in vehicle-to-infrastructure (V2I) scenarios, such as those at connected signalized intersections. Cyberattacks can cause greater damage in scenarios involving connected signalized intersections compared to platooning because the resulting uncontrolled vehicle collisions can vary in type (Head-on, T-bone, etc.), rather than being predominantly rear-end collisions as in platooning situations. Therefore, cyberattacks at connected signalized intersections should be more emphasized and systematically studied."

- "Current researchers focused on congestion as the primary objective of attackers, such as queue length attack and arrival time attack. In this scenario, attackers need to be informed about specific details such as phase sequence, minimum, and maximum green time. However, this may not be the most efficient method for causing damage from the attacker's perspective."

- "Research focusing on spoofing attacks from the driver's side is scarce. These attacks can lead to more serious consequences than those targeting delays, because the driver may engage in behaviors that violate traffic rules."

- “The research utilized microscopic traffic simulation software such as VISSIM instead of real driving data. Although VISSIM is designed to simulate real-world driving conditions as closely as possible, it cannot capture all the nuances of human driving behavior. Real driving data includes unpredictable elements such as sudden braking, acceleration, and diverse driver reactions, which are difficult to model precisely.”

- “Researchers have primarily focused on detection methods at the cyber layer, such as communication detection algorithms, rather than on the physical layer. Understanding the behavior of human drivers during cyberattack events is crucial in comprehensively evaluating the impact of such attacks and developing effective countermeasures.”

Comment 4: Fig. 9 and Algorithm 2 require detailed introductions.

Author Response: Thank you for bringing this up. We have added more detailed introductions for Fig. 9 and Algorithm 2 by including the step-by-step introductions of modeling and detection. Moreover, we have improved Fig. 9 by changing the line color and adding the corresponding equations.

Author Action: Really appreciate the reviewer’s comment regarding this equation. We have updated the manuscript by adding the detailed introductions as follows:

“As shown in Fig. 9, we first collect all the trajectory data during the experiment. We performed k -medoids clustering to generate different patterns of trajectory groups. Then we split the data into two parts: training datasets and testing datasets. Training datasets were used to build the HMM-4-C model using the EM algorithm (**Algorithm 1**). We built three detection models by using corresponding training datasets. For the testing datasets, we combined them together and anonymized the data. Then, **Algorithm 2** was applied using the built models and anonymized data. Specifically, for each testing dataset, the forward and backward algorithm was used to calculate its probability under the current model. Therefore, we obtained three different probabilities for a testing dataset, each corresponding to one of the three models. The detection result of the testing dataset is the maximum probability calculated by the forward and backward algorithm under the corresponding model.”

Comment 5: Some figures are not clear, such as Figs. 8 and 9.

Author Response: Thank you for bringing this up. We have revised Figs. 8 and 9 including both the color and equations. Moreover, we have added the detailed introduction about the Figs. 8 and 9.

Author Action: Really appreciate the reviewer’s comment regarding this equation. We have updated Figs. 8 and 9. of the manuscript as follows:

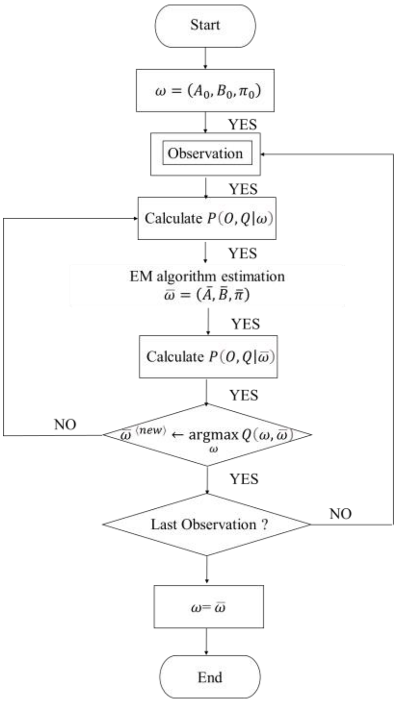

Start
$\omega = (A_0, B_0, \pi_0)$
YES
Observation
YES
Calculate $P(O, Q|\omega)$
YES
EM algorithm estimation
$\bar{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$
YES
Calculate $P(O, Q|\bar{\omega})$
YES
NO
$\bar{\omega}^{(new)} \leftarrow \underset{\omega}{\operatorname{argmax}} Q(\omega, \bar{\omega})$
YES
Last Observation ?
NO
$\omega = \bar{\omega}$
End

Fig. 8. The EM algorithm.

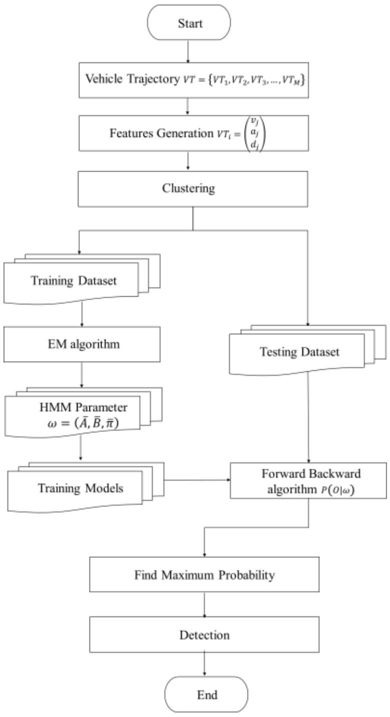

Start
Vehicle Trajectory $VT = \{VT_1, VT_2, VT_3, \dots, VT_M\}$
Features Generation $VT_i = \begin{pmatrix} v_j \\ a_j \\ d_j \end{pmatrix}$
Clustering
Training Dataset
EM algorithm
Testing Dataset
HMM Parameter
$\omega = (\bar{A}, \bar{B}, \bar{\pi})$
Training Models
Forward Backward
algorithm $P(O|\omega)$
Find Maximum Probability
Detection
End

Fig. 9. HMM-4-C model and detection.

Comment 6: What is the scalability of the proposed method? As the discussed attack scenario is relatively simple, can the detection method be applied to more complex and practical cases? In addition, the method detects a cybercriminal after an attack is successfully launched. It can lead to a situation where the attack is exposed after vehicles have mistakenly entered the intersection, i.e., after a disastrous event has caused. How to avoid such a situation? Can the method detect attacks before they cause damage? Furthermore, experiments show that the method has a high false detection rate in Stop scenarios. The authors need to elaborate on the significance of the proposed method in practical scenarios.

a. **What is the scalability of the proposed method? As the discussed attack scenario is relatively simple, can the detection method be applied to more complex and practical cases?**

Author Response: Thank you for bringing this up. The reason we select this relatively "simple" scenario is that red-light countdown (RLCD) is a typical and widely used ITS application [4], [5], as well as emerging connected vehicle application [6]–[8]. In addition, from the attacker's perspectives, RLCD application is relative an efficient and simple way to attack the transportation system and cause the relative serious consequences (i.e., red light running and T-born collision). Therefore, cyberattack on RLCD is relative practical and easily happened on the connected vehicle environment.

[4] Y.-C. Chiou and C.-H. Chang, "Driver responses to green and red vehicular signal countdown displays: Safety and efficiency aspects," *Accident Analysis & Prevention*, vol. 42, no. 4, pp. 1057–1065, Jul. 2010, doi: 10.1016/j.aap.2009.12.013.

[5] E. Uhlemann, "Connected-Vehicles Applications Are Emerging [Connected Vehicles]," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 25–96, Mar. 2016, doi: 10.1109/MVT.2015.2508322.

[6] "Taking Audi's Red Light Countdown System for a Spin," PCMAG. Accessed: Nov. 13, 2023. [Online]. Available: https://www.pcmag.com/opinions/taking-audis-red-light-countdown-system-for-a-spin

[7] R. Stumpf, "Audi's new tech can help you beat red lights," Popular Science. Accessed: Nov. 13, 2023. [Online]. Available: https://www.popsci.com/technology/new-audi-tech-provides-traffic-light-updates/

[8] J. C. Wolf, J. Ma, B. Cisco, J. Neill, B. Moen, and C. Jarecki, "Deriving Signal Performance Metrics from Large-Scale Connected Vehicle System Deployment," *Transportation Research Record*, vol. 2673, no. 4, pp. 36–46, Apr. 2019, doi: 10.1177/0361198119838520.

b. In addition, the method detects a cybercriminal after an attack is successfully launched. It can lead to a situation where the attack is exposed after vehicles have mistakenly entered the intersection, i.e., after a disastrous event has caused. How to avoid such a situation? Can the method detect attacks before they cause damage?

Author Response: Thank you for bringing this up. The method is built to detect the trajectory exist already. we focus on the detection on the first step, which is the first vehicle detection. The potential case for this method is that when the cyberattack happens, attacked vehicle mistakenly enter into the intersection or using the detectable driving behaviors. The method detects the cyberattack scenarios and make a warning to the following vehicles, as well as target the intersection as a cyberattacked intersection.

Author Action: Really appreciate the reviewer's comment regarding this question. We have updated the manuscript by adding the explanation as follows:

"In the physical layer detection setting, successful detection might often occur only after a vehicle physically and mistakenly enters the intersection, that is, after a disastrous event has already occurred. Therefore, infrastructure-side sensor detection faces a challenge in this regard. We will implement the advanced hybrid detection model, incorporating data fusion to focus on enhancing accuracy and speeding the detection to address the issue.**"**

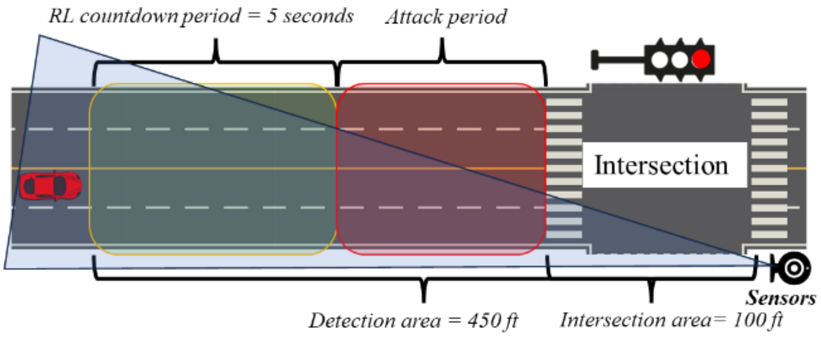
c. Furthermore, experiments show that the method has a high false detection rate in Stop scenarios. The authors need to elaborate on the significance of the proposed method in practical scenarios.

Author Response: Thank you for bringing this up. The high false detection rate in Stop scenarios came from the relative similarity of Stop with some trajectories in Cyberattack scenario. Therefore, stop scenario exists difficult on the discriminate with Cyberattack scenario. The limited detection results of HMM-4-C, as shown in Fig. 22, primarily arise from the relative similarities in trajectory features between the cyberattack and stop scenarios. These similarities are more pronounced than those with the pass situation, as shown in Figs. 10 and 11. This can lead to confusion in the classification process. According to the aforementioned trajectory diagrams, people tend to stop when facing a cyberattack. The 6.75% of the Stop testing datasets misdetection arises because the features of cyberattack trajectories are more similar to the stop scenario than to the pass scenario.

Author Action: Really appreciate the reviewer's comment regarding this point. We have updated the manuscript by adding the describe about misdetection.

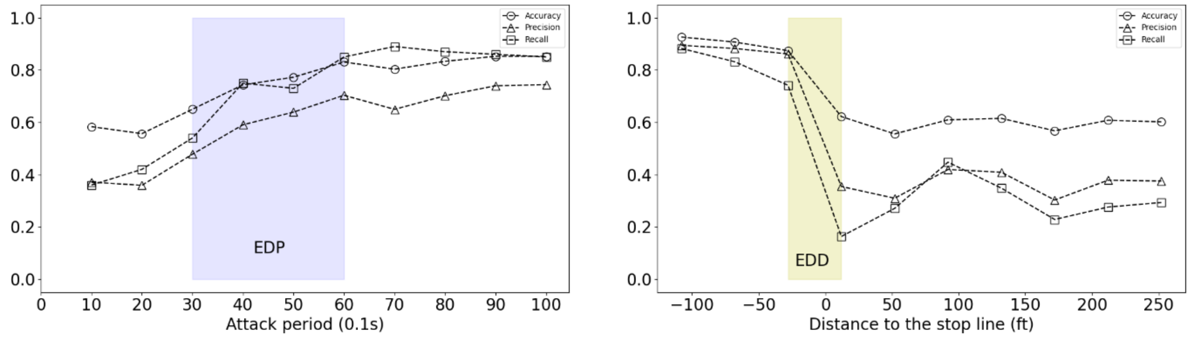
"HMM-4-C showed higher detection accuracy for all three scenarios, with the Pass scenario achieving 99.83% true positive rate and the Stop and Cyberattack scenarios achieving 90.42% and 98.00% true positive rates, respectively. However, there were some misdetections, the 6.75% of the Stop testing datasets misdetection arises because the features of cyberattack trajectories are more similar to the stop scenario than to the pass scenario."

In addition, we propose the real-world application of evaluating the effectiveness of this detection method. Specificly, we propose a potential real-world setting for infrastructure-side detection. The sensors are installed at the intersection and oriented towards the direction of approaching vehicles. In the case, vehicle trajectory can be captured in real-time. The trajectory data is capture as soon as the vehicle enters the detection area.



Real-world physical layer cyberattacks detection

We further conduct sensitivity analyses of time-period-based detection and distance-based detection to explore their potential for future real-world applications.



Comment 7: There are many presentation errors and typos. For example,

a. Abbreviations used only once or twice should be removed, such as OBU and TTC. Some abbreviations are not defined, such as CAV.

b. "For example, Two traffic engineers …", "Two" => "two". Similar typos can be found in

the paper, such as "To detect falsified trajectory data, The author …";

c. "as there isn't an ample amount of …", "isn't" => "is not";

d. "HMMs are based on a probabilistic framework which provides naturally handles uncertainties in cyberattack trajectories";

e. "DeBruhl have" => "DeBruhl et al. have";

f. "Figure 4 presents the hidden state and observation processes", "Figure 4" => "Figure 5";

g. "αt(i) called the forward parameter" => "αt(i) is called the forward parameter";

h. "Fig. 11 depict the speed …", "depict" => "depicts";

i. "The "V" type speed curves in the Pass scenario in Fig. 10a", "Fig. 10a" => "Fig. 11a";

j. "cyberattacked trajectory change in frequently and abruptly".

Author Response: Thank you for bringing this up. We have corrected all the errors and typos including both the presentation format and theoretical equation.

Author Action: Really appreciate the reviewer's comment for bringing theses up. The typos have been corrected by following in the manuscript:

a. CAV stands for "Connected & Automated Vehicles". We have added the full name for CAV in the manuscript. For the abbreviations of "OBU" and "TTC", we have revised them to "on-board unit" and "time to collision" respectively. Moreover, we have also revised 'RSU' and 'BSMs' for greater conciseness." The revision ensures that each abbreviation is clearly matched with its full form and maintains grammatical consistency.

b. The sentence has been revised to be "two traffic engineers hacked"

c. The sentence has been fully revised.

d. "HMMs are based on a probabilistic framework which provides naturally handles uncertainties in cyberattack trajectories"; It has been revised based on the content to be more clearly for the readers.

e. "DeBruhl have" => "DeBruhl et al. have"; "et al." has been added in all the literature for a more formal and accurate format.

f. "Figure 4 presents the hidden state and observation processes". "Figure 4" has been corrected to "Figure 5";

g. "αt(i) called the forward parameter". We have revised as: "αt(i) is called the forward parameter";

h. "Fig. 11 depict the speed …", "depict" has been corrected to "depicts";

i. "The "V" type speed curves in the Pass scenario in Fig. 10a", "Fig. 10a" has been corrected to "Fig. 11a";

j. "cyberattacked trajectory change in frequently and abruptly". The sentence has been fully revised in the manuscript.

Detection of Cyberattacks at Connected Signalized Intersections: from the Physical Layer Perspective of Cyber-Physical Systems

Yingfan Gu, Zhixia Li, Yunpeng Zhang, Shivam Tiwari, Heng Wei, Guohui Zhang, Muting Ma and Sabur Baidya

[1]***Abstract*—Detecting cyberattacks is crucial for securing the connected transportation system. Previous research has predominantly focused on the detection from the perspective of cyber layer. This type of detection carries the risk that the cyber layer detection itself may become a target of the attacks. In contrast, standalone sensor-data-based algorithms that avoid communication to detect cyberattacks based on patterns of vehicle trajectories from the physical layer perspective of a cyber-physical system is promising to become a high-efficiency alternative to avoid the algorithm being attacked. This study develops a vehicle trajectory-oriented machine-learning algorithm to detect cyberattacks. Vehicle trajectories and driving behaviors are captured using a driving simulator under normal and cyberattack scenarios. Cyberattack detection models were constructed using Hidden Markov Models (HMMs), namely HMM-4-C. Whether a vehicle is being attacked or not is considered as a hidden state, and the trajectory data serve as the observation sequence. Expectation Maximization (EM) algorithm is applied to estimate the likelihood parameters of HMM-4-C. The results indicate that 98% of cyberattack scenarios in our research were successfully detected using HMM-4-C. Furthermore, we first propose the Effective Detection Period (EDP) and Effective Detection Distance (EDD) from the perspective of a cyberattack protector based on the potential real-world setting. The proposed method effectively detects cyberattacks at connected signalized intersections, which can be used to develop proactive cyberattack detection strategies in conjunction with cyber- perspective detection algorithms.**

***Index Terms*—Connected vehicle, cyberattack, machine learning, HMMs, HMM-4-C, vehicle trajectories, driving behaviors.**

A. INTRODUCTION

The emergence of connected vehicles and connected and autonomous vehicle technologies is expected to pose cybersecurity issues to the urban transportation system. [1]–[4]. Connected vehicles are equipped with an on-board unit that broadcasts basic safety messages. The roadside unit receives the messages and relays the information to other vehicles and infrastructure [5]–[8]. As vehicles become more and more connected to the wireless networks, this technological advancement expands the attack surface for cybercriminals to execute sophisticated cyberattacks [9]–[11].

The frequency of cyberattack events is on the rise [12]–[14]. For example, two traffic engineers hacked into the signal system, and reduced traffic flows by programming the signals because of the labor protest [15]. New York City's wireless vehicles were attacked by a cybersecurity expert using a very cheap wireless device [16]. In Texas, hackers broke into a traffic sign and changed the messages on the digital signs of a road closure message with "Zombies Ahead" [17]. It is catastrophic if cyberattacks occur within the transportation system, with hackers potentially causing disastrous events [18], [19]. Furthermore, the Advanced Traffic Management System (ATMS) will be a critical component of the USA transportation system shortly. With the increasing use of advanced Traffic Management Systems, the United States is becoming an increasingly popular target for cyberattacks [20], [21].

Existing methodologies has predominantly focused on the perspective of cyber layer detection, like sensor intrusions and communication anomalies [22]–[24]. This type of detection, such as using algorithmic recognition of network attacks, carries the risk of also becoming a target when attacks occur [25]. Compared to the cyber layer perspectives, using a standalone algorithm to analyze vehicle trajectories and driving behaviors patterns based on sensor data from the physical layer of a cyber-physical system is promising to be useful for detecting irregularities that could indicate a cyberattack [26]. Previous research has suggested that vehicle trajectories could be useful for studying cybersecurity [27]. If we could establish a standalone detection algorithm that is based on vehicle trajectories under cyberattacks from sensor data, early-stage warning for the cyberattack scenarios at the very beginning when the cyberattack happens would become feasible. This method that avoids communication is promising to become a high-efficiency alternative cyberattack detection

This work is partly supported by the USDOT Tier 1 UTC Transportation Cybersecurity Center for Advanced Research (CYBER-CARE). *(Corresponding author: Zhixia Li).*

Yingfan Gu is with the Civil & Architectural Engineering & Construction Department, University of Cincinnati, Cincinnati, Ohio 45221, USA (guyf@mail.uc.edu).

Zhixia Li is with the Civil & Architectural Engineering & Construction Department, University of Cincinnati, Cincinnati, Ohio 45221, USA (lizx@ucmail.uc.edu).

Yunpeng Zhang is with the College of Technology, Information and Logistics Technology Department, University of Houston-Main Campus, Houston, Texas 77204, USA (yzhan226@central.uh.edu).

Shivam Tiwari is with the College of Technology, Information and Logistics Technology Department, University of Houston-Main Campus, Houston, Texas 77204, USA (stiwari4@CougarNet.uh.edu).

Heng Wei is with the Civil & Architectural Engineering & Construction Department, University of Cincinnati, Cincinnati, Ohio 45221, USA (heng.wei@uc.edu).

Guohui Zhang is with Department of Environmental and Construction Engineering, University of Hawaii at Manoa, Honolulu, HI 96822 USA (email:guohui@hawaii.edu)

Muting Ma s with the Culverhouse College of Business, The University of Alabama, Tuscaloosa, Alabama, 35487, USA (mma10@ua.edu).

Sabur Baidya is with the Department of Computer Science and Engineering, The University of Louisville, Louisville, Kentucky, 40292, USA (shbaid01@louisville.edu).

method to avoid the detection algorithm being attacked. However, there has been a lack of such trajectory-data-based detection algorithms, as there lacks such ground truth trajectory data collected under real cyberattack at transportation facilities to train and validate the detection models.

In addition, to accurately detect cyberattack, we need to model vehicle trajectories in an efficient manner. Existing method to detect trajectory include: Convolutional Neural Network (CNN), Gaussian Mixture Model (GMM) or Recurrent Neural Network (RNN) [28]–[30]. However, neural network-based methods may not always be optima, as they require a large amount of historical data for training. This might not be suitable in the context of cyberattacks, as there isn't an ample amount of existing data available. Research has proved the effectiveness of Long Short-Term Memory (LSTM) in the area of pedestrian trajectory prediction [31] and uncrewed aerial vehicles detection [32], [33]. It can potentially remember patterns over long sequences. During cyberattacks, the upcoming trajectory and behavior might be stochastic. Therefore, there is need of a suitable method to overcome the challenges of inadequate data and abrupt trajectory turning to accurately detect cyberattack events. Hidden Markov models (HMMs) can therefore be a potential tool because the upcoming trajectory and behavior are stochastic and depend on the last state, which can be naturally described by a hidden Markov process [34]. This might make them appropriate for estimating sudden changes in situations when an attack occurs [35], [36] and beneficial in cyberattack situations where data is limited.

In summary, HMMs may offer the following advantages in cyberattack trajectory pattern recognition: (1) Compared to neural networks, HMMs can be computationally more efficient. (2) HMMs might require less data to estimate their parameters, making them preferable in scenarios with limited data. Additionally, the literature review did not find HHMs being used to construct a cyberattack model. Therefore, the objectives of this research are to address two key issues. The first issue is the scarcity of available cyberattack scenarios, vehicle trajectories, and driver behavior for current researchers. To address this, we create a comprehensive cyberattack scenario within a connected vehicle environment. The second issue pertains to the absence of a method for detecting cyberattack scenarios. In response to this, we propose a state-of-the-art, physical layer-based method named HMM-4-C to detect cyberattacks.

II. LITERATURE REVIEW

Cyberattacks increasingly threaten the stability of the transportation system and cause serious problems [37]–[39]. Previous research primarily centered on vehicle platoon dynamics under cyberattacks. Wang et al. [40] presented an extended car platoon model by introducing increased weight parameters to characterize different security attacks. Cui et al. [41] evaluated the effects of cyberattacks under cooperative adaptive cruise control (CACC) platoon scenarios and demonstrated that cyberattacks may influence the stability of CACC platoon. Amoozadeh et al [42]. found that cyberattacks can cause instability of the CACC vehicle stream. Li et al. [43] tested attack factors and communicated positions and speeds in a longitudinal connected and automated vehicles (CAV) using a collision index. Wang et al. [44] developed the CAV dynamics platoon under a cyberattack scenario using different acceleration, speed, and position. The results indicate that a cyberattack could potentially lead to serious disruptions in vehicle platoons. Khattak et al. [45] conducted three types of cyberattacks, message falsification, dedicated denial of service, and spoofing attacks, and used the time to collision to reveal that traffic stream and CAV string are unstable under cyberattacks. In summary, due to the severe problems that cyberattacks can inflict on current transportation systems, researching effective detection methods has become critically important.

Detecting cyberattacks under connected vehicle environment has become an important research area [46]. In vehicle platoons scenarios, detection methods primarily rely on vehicle-to-vehicle communication detection [47], such as denial-of-service (DoS) attacks [48]. Biron et al. [49] presented a detection scheme for DoS attack in connected vehicles. It simulated the attack by assuming that the attacker keeps the communication network busy with fake requests, rendering the network unable to respond to legitimate user requests. Mousavinejad et al. [50] presented a distributed attack detection algorithm for vehicle platooning by using predesign criteria of each vehicle state. After that, Zhang et al. [51] established several relationships between attack parameters and platoon system performance under DoS attacks. Xiao et al. [52] introduced a detection method based on neural network that verifying the time-stamps of the data packets under the presence of intermittent communication attacks.

Recent research has enhanced our understanding of the security risks and communication detection methods in connected vehicle streams. However, there has been less emphasis on cyberattacks in vehicle-to-infrastructure scenarios, such as those at connected signalized intersections [53]. Cyberattacks can cause greater damage in scenarios involving connected signalized intersections compared to those in vehicle platooning. This is because the resulting uncontrolled vehicle collisions can vary in type, such as head-on or T-bone collisions, rather than predominantly rear-end collisions, as is common in platooning situations. Therefore, cyberattacks at connected signalized intersections should be more emphasized and systematically studied. In addition, researchers have primarily focused on detection methods at the cyber layer, such as communication detection algorithms, rather than on the physical layer. Understanding the behavior of human drivers during cyberattack events is crucial in comprehensively evaluating the impact of such attacks and developing effective countermeasures.

Spoofing attacks on traffic control systems at connected signalized intersections are increasingly attracting scholarly attention. For example, Feng et al [54]. manipulated the detector data in actuated and adaptive signal control to simulate the spoofed cyberattacks to maximize system delay. Chen et al [55]. conducted the congestion attack on the Intelligent Traffic Signal System (I-SIG) by spoofing of data from a single attack vehicle generates cyberattacks on connected signalized intersections. However, current researchers have focused on congestion as the primary objective of attackers, considering attacks such as those on queue length and arrival time [53]. In this scenario, attackers need to be informed about specific details such as phase sequence, minimum, and maximum green time. However, this may not be the most efficient method for causing damage from the attacker's perspective. Additionally, research on spoofing attacks targeting the driver's side is scarce. These attacks can lead to more serious consequences than those targeting delays, because the driver may easily engage in behaviors that violate traffic rules.

Recent attention has focused on the development of detection methods for cyberattacks at connected signalized intersections. Huang et al. [6] simulated the attack by using falsified traffic trajectories and introduced a fake vehicle that arrived and stopped, presenting false data to the infrastructure. The proposed method involves training a neural network and conducting a hierarchical clustering algorithm. DeBruhl et al. [56] have designed attacks and abnormal behaviors to simulate cyberattack situations, and the successful detection of such attacks depends on the vehicle model switching when the misbehaviors occur. Shen et al [53]. implemented a congestion attack on I-SIG and corresponding defense action by estimating the trustworthiness of CVs based on readily-available infrastructure side sensors. However, the research utilized microscopic traffic simulation software such as VISSIM instead of real driving data. Although VISSIM is designed to simulate real-world driving conditions as closely as possible, it might not capture all the nuances of human driving behavior. Real driving data includes unpredictable elements such as sudden braking, acceleration, and diverse driver reactions, which are difficult to model precisely. Moreover, obtaining actual human driving data during cyberattacks is challenging. Most researchers used the hypothetical vehicle trajectory and assumed the vehicle movement under the event [52], [57]. When the attack occurred, the trajectories were strictly based on a pattern without humans getting involved. Previous research used assumed vehicle trajectory data and benchmark datasets for studying attack detection [58]–[60]. Therefore, this research aims to contribute to the literature by conducting a human-involved driving experiment under cyberattack conditions.

III. METHODOLOGY

A. The architecture of experimental scenarios

Red-light countdown (RLCD) is an emerging connected vehicle application that displays the remaining time of the red light on the driver’s side [61]–[64]. When the vehicle approaches a connected signalized intersection, the driver receives the RLCD message, which allows the driver to adjust their speed to cross the intersection. The application can substantially reduce the loss of speed by enabling the driver to adjust the speed of vehicles to go through the intersection, which can improve energy efficiency by maintaining a steady speed with fewer stops. In this scenario, countdown messages will be sent from the traffic signal controller to the roadside unit. The roadside unit will then broadcast the message to the on-board unit in the vehicle’s cabinet.

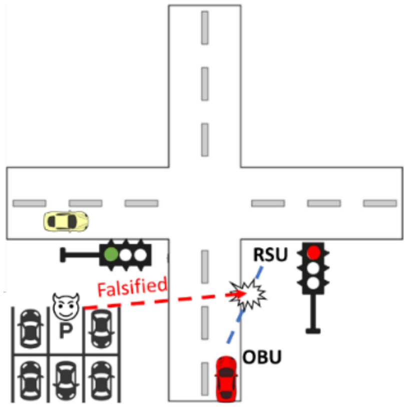


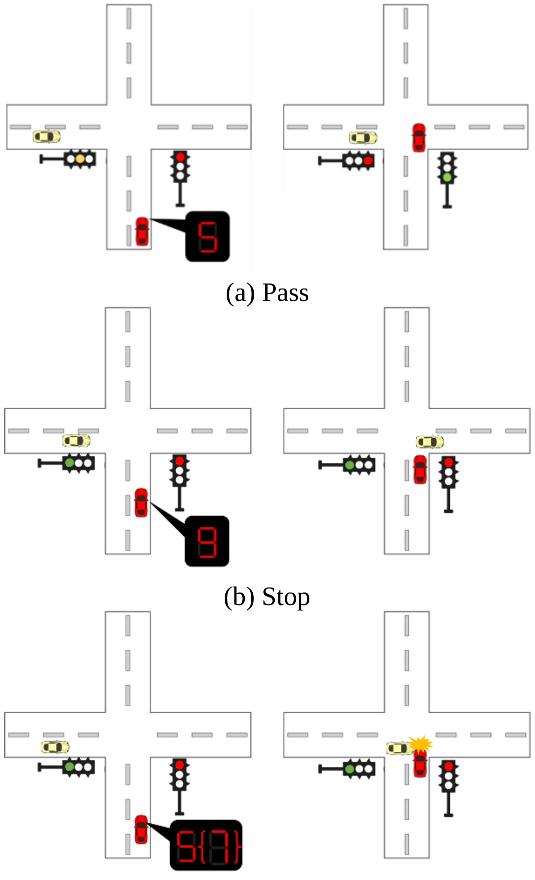
Fig. 1. Signal countdown timer attack model.

From the perspective of an attacker, it is relatively difficult to manipulate the traffic signal controller to alter signal phasing and timing to attack the transportation system, given that the controller is physically located in a roadside box. However, launching a spoofing attack via the communication between the roadside unit and the on-board unit is much easier. Attacking the RLCD application is a relatively efficient and simple way because it can potentially lead to serious consequences, such as the red-light running [65]. The red-light running is the most dangerous and could lead to more serious consequences. Therefore, a spoofing attack on the RLCD application is the most likely scenario that the attacker would consider.

Spoofing attack on the RLCD application is relatively practical and can easily occur in the connected vehicle environment. In this case, attacks may compromise the system by altering or blocking messages between vehicles and infrastructure [66]. If a cyberattack occurs, the driver will receive a falsified RLCD message. This may lead the driver to maintain their speed when the vehicle enters the intersection, potentially resulting in a severe accident. Fig. 1 illustrates the mechanism of cyberattacks on the RLCD application.

Generally, when a vehicle approaches an intersection with a red light, the driver will encounter one of two scenarios: a short-duration red light and a long-duration red light based on the vehicle’s speed. A short-duration red light means that the driver may be able to drive through the intersection without needing to come to a full stop, as the light may turn green before the vehicle enters the intersection. We define this situation as the ‘Pass scenario’ as shown in Fig. 2a. A long-duration red light means the driver must stop entirely before the stop line. This situation is defined as the ‘Stop scenario,’ as illustrated in Fig. 2b.

Besides the above scenarios, attackers may compromise the system by altering communication [67], [68]. If a cyberattack occurs, the driver will receive a falsified message indicating that there is a short-duration red light. The driver may continue to maintain the speed as the vehicle enters the intersection. However, the actual duration is 12 seconds, with 7 seconds remaining, which differs from the falsified message of 5 seconds displayed on the dashboard. We define it as a ‘Cyberattack scenario,’ as illustrated by Fig. 2c.



(a) Pass

(b) Stop

(c) Cyberattack

Fig. 2. RLCD scenarios at connected signalized intersection.

B. *Experiment procedure*

To collect the human-involved driving data, we simulated this specific scenario in a laboratory environment. An experiment was designed to implement cyberattack scenarios under the connected vehicle environment. The miniSim driving simulator was applied in the research, presented by the National Advanced Driving Simulator. The driving simulator creates an ideal environment for collecting vehicle trajectories. Thirty-two participants with valid driving licenses were publicly recruited for this study. The demographic information is summarized in Table 1.

TABLE I

SUMMARY OF DEMOGRAPHICS FOR PARTICIPANTS (N=32)

Variable	Category	N	Percent (%)
Age	21~25 years old	22	68.8%
	26~30 years old	6	18.8%
	31~35 years old	3	9.4%
	36~40 years old	1	3.1%
Gender	Female	10	31.3%
	Male	22	68.7%

Fig. 3 provides an overview of the route designed for the experiment, with the route indicated in red and the scenario zone in yellow. Each participant was required to drive through seven intersections three times. The connected vehicle application was implemented in all intersections. Fig. 4 shows that the driver received the message indicating the remaining RLCD (“Red light remains in 5 seconds, 4 seconds… until 1 second”) as they approached each intersection.

To minimize the impact of psychological expectations for the cyberattack scenario, the RLCD application in other intersections were kept functioning normally. The cyberattack scenario was only introduced at one intersection marked by the yellow area in Fig. 3. During the first time, this intersection was set to a short-duration RLCD (5 seconds), indicating the ‘Pass scenario’. During the second time, it was set to the ‘Stop scenario’, with a long-duration RLCD (12 seconds). For the third time, a cyberattack scenario we set up where cybercriminals falsified the RLCD message. The long-duration RLCD message (12 seconds) was changed to a short-duration message (5 seconds). The red light was still on when the drivers reached the stop line at this intersection. Since the frequency of transmitting under the connected vehicle environment is 10Hz [69], we configured the data acquisition system to capture the vehicle's trajectory parameters at a sampling frequency of 10Hz. In the experiment, a speed limit of 35 mph was adopted.

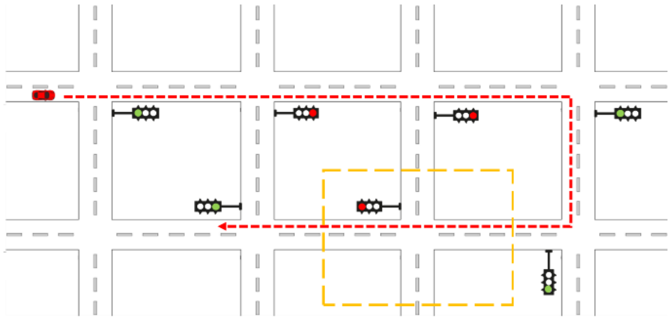
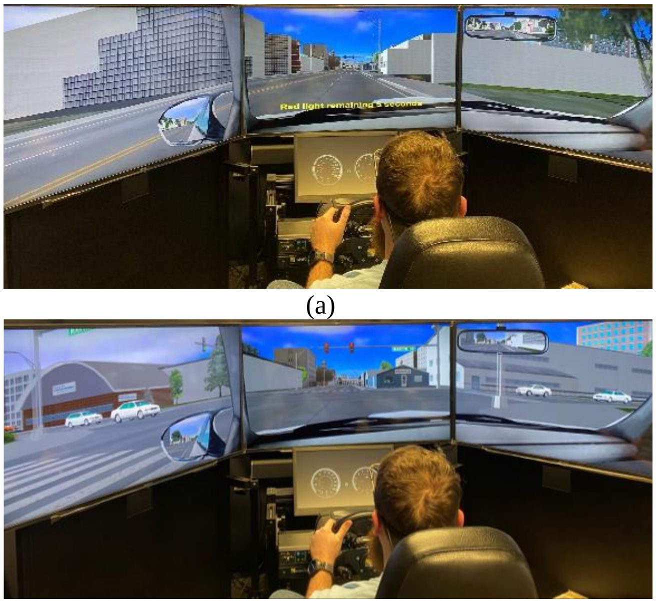


Fig. 3. Designed route of the experiment.



(a)

(b)

Fig. 4. Display RLCD message at cyberattack scenairo.

C. *Hidden Markov models for vehicle trajectory (HMM-4-C) under the cyberattack*

This study treats vehicle movement as a time-series Markov process [70], [71], with different hidden states corresponding to the different scenarios in the cyberattack context. The time series of vehicle trajectories under each state is considered an observation [36], [72]. During the cyberattacks, the hidden state space is:

$$S = \{S_1, S_2, S_3\} \tag{1}$$

Where S_1 , S_2 and S_3 represents the Pass, Stop and Cyberattack scenario, respectively. The observations are the sequence of vehicle trajectory VT, recorded at a 10Hz time interval. M is the number of the observation.

$$VT = \{VT_1, VT_2, VT_3, \dots, VT_M\} \tag{2}$$

For a given observation sequence VT_i at time i, it can be represented by a matrix of size $N \times M$, N is number of recorded features. For vehicle trajectories, the characteristic features of the motion state are its kinetic state and direction of motion. In a cyberattack situation, changes in the kinetic state and motion direction indicate the impact of the cyberattack on vehicle trajectories. It will lead the attacked vehicle drive into the intersection or drive out of the road to causes catastrophic accident. In this research, we focus on changes in the axis direction connected to the intersection. The axis direction of velocity, acceleration and the distance to the stop line are regarded as the cyberattack features. M is the number of time intervals for which the observations are recorded. Therefore, VT_i can be represented by

$$VT_i = \begin{pmatrix} v_j \\ a_j \\ d_j \end{pmatrix}, \forall i = 1,2,3, \dots, M, \forall j = 1,2,3, \dots, N \tag{3}$$

Where v_j represents the velocity; a_j represents the acceleration; d_j represents the distance to the stop line.

Then, we denoted $\{Q\}$ as the hidden state segment during the time the vehicle is approaching the intersection. q_t means the one of hidden state in $\{S\}$ at time t when the vehicle is approaching. $\{O\}$ is denoted as the observations segment. The observation o_t is the observed vehicle trajectory in $\{VT\}$ for a simplified annotation. Figure 5 presents the hidden state and observation processes of HMM-4-C in this study.

$$Q = \{q_1, q_2, q_3, \dots, q_T\} \tag{4}$$

$$O = \{o_1, o_2, o_3, \dots, o_T\} \tag{5}$$

Equation (6) represents the probability of each state given the state attained in the previous time step [70], [71], which is

$$P(q_t | q_{t-1}, o_{t-1}, \cdots, q_1, o_1) = P(q_t | q_{t-1}) \tag{6}$$

In addition, the probability of each observation is independent of other observations and states and only depends on the current state.

$$P(o_t | q_t, o_{t-1}, \cdots, q_1, o_1) = P(o_t | q_t) \tag{7}$$

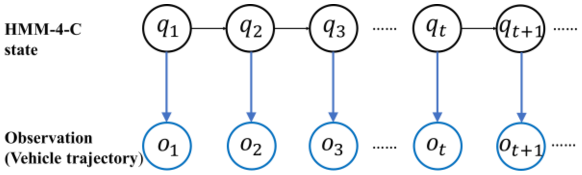


Fig. 5. The hidden state process and observation process of HMM-4-C.

The parameters of HMM-4-C can be represented by $\omega = (A, B, \pi)$ [6], [71], where A is the transition probability, B is the emission probability and π is the Initial probability of the state.

The transition probability A indicates hidden state transition likelihood matrix that can be represented by

$$A = [a_{ij}] \tag{8}$$

Each a_{ij} represent the probability of moving from hidden state i to state j.

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i),$$
$$\forall i, j = 1,2,3; \forall t = 1,2,3, \dots, T \tag{9}$$

The emission probability B represents the probability of each observation given a hidden state, and it can be represented by

$$B = [b_j(k)] \tag{10}$$

Each $b_j(k)$ means the probability of an observation k being generated from a given the current state j.

$$b_j(k) = P(o_t = VT_k | q_t = S_j),$$
$$\forall j = 1,2,3; \forall k = 1,2,3, \dots, M; \forall t = 1,2,3, \dots, T \tag{11}$$

The Initial probability of state π means the probability that the model will start in a state that can be represented by

$$\pi = [\pi_i] \tag{12}$$

$$\pi_i = P(q_1 = S_i), \forall i = 1,2,3 \tag{13}$$

Therefore, the probability of the state series $Q = \{q_1, q_2, q_3, \dots, q_T\}$ can be represented by

$$P(Q, \omega) = \pi_{q_1} \prod_{t=1}^{T-1} a_{q_t q_{t+1}} \tag{14}$$

The probability of the observation from the state series $Q = \{q_1, q_2, q_3, \dots, q_T\}$ is

$$P(O | Q, \omega) = \prod_{t=1}^{T} b_{q_t}(o_t) \tag{15}$$

Then, the joint probability distribution of O and Q is,

$$P(O, Q | \omega) = \pi_{q_1} b_{q_1}(o_1) \prod_{t=1}^{T-1} a_{q_t q_{t+1}} b_{q_{t+1}}(o_{t+1}) \tag{16}$$

Finally, sum all the Q , we can get the probability of specifical vehicle trajectory observation under the model $\omega = (A, B, \pi)$.

$$P(O | \omega) = \sum_{q_1, q_2, q_3, \dots, q_T} \pi_{q_1} b_{q_1}(o_1) \prod_{t=1}^{T-1} a_{q_t q_{t+1}} b_{q_{t+1}}(o_{t+1}) \tag{17}$$

Nevertheless, the calculation of the possibility in Equation (17) is relatively computationally expensive due to its high time complexity ($O(TN^T)$). Therefore, we used the forward and backward algorithm to solve this problem [73]–[75]. The algorithm can calculate the probability under the model by a given sequence of observations at a low complexity ($O(N^2T)$).

We define two variables $\alpha_t(i)$ and $\beta_t(i)$. $\alpha_t(i)$ is called the forward parameter that means the probability of past observations in a given state q_t at time t. $\beta_t(i)$ represents back parameter that means the probability of the future observations in a given state q_t at time t. A schematic diagram of the forward and back parameters is presented in Fig. 6, while Fig. 7 shows the computation process of the joint event using the forward and back algorithm.

This algorithm can calculate the probability that a model generated a sequence of observations. For a known $\omega =$

(A, B, π), at time t, if the observation series is $o_1, o_2, o_3, \dots, o_t$, the forward possibility is

$$\alpha_t(i) = P(o_1, o_2, o_3, \dots, o_t, q_t = S_i|\omega) \qquad (18)$$

Then, the forward formula can be represented by the base case $\alpha_1(i)$ and inductive step. Following the definition of the parameters of the HMMs, debase case $\alpha_1(i)$ is

$$\alpha_1(i) = \pi_i b_i(o_1), \forall i = 1,2, \dots, N \qquad (19)$$

Sum of all the different probabilities of getting to state j times the emission. The inductive step is

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(o_{t+1}),$$
$$\forall j = 1,2, \dots, N, \forall t = 1,2, \dots, T-1 \qquad (20)$$

Then, the final step is that sum all the probabilities which will end up the state given the observation sequence. The probability of the observation O by given the model ω is

$$P(O|\omega) = \sum_{i=1}^{N} \alpha_T(i) \qquad (21)$$

For the backward possibility, at time t, the observation is $O_w = (o_{t+1}, o_{t+2}, \dots, o_T)$. The backward algorithm is

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \dots, o_T|q_t = S_i, \omega) \qquad (22)$$

The base case and inductive step are

$$\beta_T(i) = 1, \forall i = 1,2, \dots, N \qquad (23)$$

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}\, b_j(o_{t+1}) \beta_{t+1}(j),$$
$$\forall i = 1,2, \dots, N, \forall t = T-1, T-2, \dots, 1 \qquad (24)$$

The probability of the observation O by given the model ω is

$$P(O|\omega) = \sum_{i=1}^{N} \pi_j b_j(o_1) \beta_1(j) \qquad (25)$$

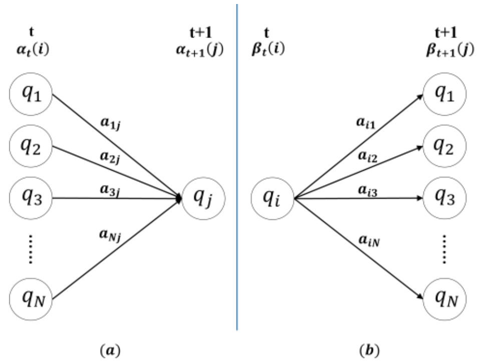


Fig. 6. Forward and backward algorithm in HMM-4-C. (a) forward parameter. (b) backward parameter.

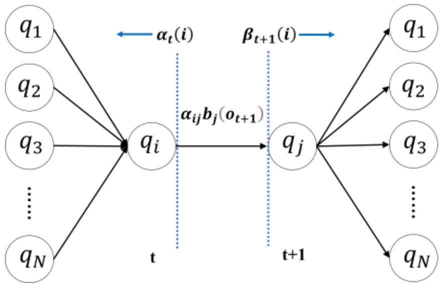


Fig. 7. The computation process of the joint event.

D. EM algorithm (Baum-Welch algorithm)

In this research, the basic parameters $\omega = (A, B, \pi)$ are not given. We need to use the observed trajectory datasets to estimate the parameters. Here, we used Expectation-Maximization (EM) algorithm [73], [76], [77] to estimate $\omega = (A, B, \pi)$. In the other words, for a given set of trajectory observation, we need to find parameters $\omega = (A, B, \pi)$ to maximize the likelihood of generating that the sequence of observation. Therefore, in this section, we estimate three different parameters ω_1, ω_2 and ω_3 for the three different scenarios using the EM algorithm.

In each model, we first define initial set of parameters $\omega_0 = (A_0, B_0, \pi_0)$. The objective is to find the optimal set of parameters $\bar{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$ that maximize our observations from training datasets.

$$\omega_0 = (A_0, B_0, \pi_0) \qquad (26)$$
$$\bar{\omega} = (\bar{A}, \bar{B}, \bar{\pi}) \qquad (27)$$

Therefore, we resort to looping the expectation step (E-step) and the maximization step (M-step) in EM algorithm to update the parameters $\omega = (A, B, \pi)$ [77]–[80]. Following the EM algorithm, we evaluate the expectation of the complete data log-likelihood function, under the posterior distribution of hidden states based on the current estimate of the parameters ω, the likelihood function of E step is

$$Q(\omega, \bar{\omega}) = \sum_{I} P(O, Q|\bar{\omega}) \log P(O, Q|\omega) \qquad (28)$$

Depend on Equation (14)-(17), the likelihood function be expressed as

$$Q(\omega, \bar{\omega}) = \begin{matrix} \sum_{I} \left(\log \pi_{i_1}\right) P(O, Q|\bar{\omega}) + \\ \sum_{I} \left(\sum_{t=1}^{T-1} \log a_{i_t i_{t+1}}\right) P(O, Q|\bar{\omega}) + \\ \sum_{I} \left(\sum_{t=1}^{T} \log b_{i_t}(o^t)\right) P(O, Q|\bar{\omega}) \end{matrix} \qquad (24)$$

Where T is the length of O and o^t is the t th observation.

The analytical solution for maximize the likelihood function given by Equation (24) involves setting the derivatives to zero subject to the following constraints.

$$\sum_{j=1}^{N} \alpha_{ij} = 1 \qquad (25)$$

$$\sum_{k=1}^{M} b_j(k) = 1 \qquad (26)$$

$$\sum_{i=1}^{N} \pi_i = 1 \qquad (27)$$

Therefore, the final formula $\omega^{\langle new \rangle} = \left(A^{\langle new \rangle}, B^{\langle new \rangle}, \pi^{\langle new \rangle}\right)$ is

$$a_{ij}{}^{\langle new \rangle} = \frac{\sum_{t=1}^{T-1} P\left(O, q_t = S_i, q_{t+1} = S_j \middle| \omega^{\langle old \rangle}\right)}{\sum_{t=1}^{T-1} P\left(O, q_t = S_i \middle| \omega^{\langle old \rangle}\right)} \qquad (28)$$

$$b_j(k)^{\langle new\rangle} = \frac{\sum_{t=1}^{T} P\left(O, q_t = S_i \middle| \omega^{\langle old\rangle}\right) Q(o_t = VT_k)}{\sum_{t=1}^{T} P\left(O, q_t = S_i \middle| \omega^{\langle old\rangle}\right)} \quad (29)$$

$$\pi_i{}^{\langle new\rangle} = \frac{P\left(O, q_1 = S_i \middle| \omega^{\langle old\rangle}\right)}{P\left(O \middle| \omega^{\langle old\rangle}\right)} \quad (30)$$

In the M step, we obtain a new estimate of the parameters $\overline{\omega}^{\langle new\rangle}$ by maximizing $Q(\omega, \overline{\omega})$ subject to the constraints, which is,

$$\overline{\omega}^{\langle new\rangle} \leftarrow \underset{\omega}{\operatorname{argmax}}\, Q(\omega, \overline{\omega}) \quad (31)$$

We then loop the algorithm until it finds the $\overline{\omega}$ that maximizes the log-likelihood function, and generate the new $\overline{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$, which indicates a new model. Algorithm 1 provides the process of the EM algorithm. Figs. 8 provide the steps of the EM algorithm.

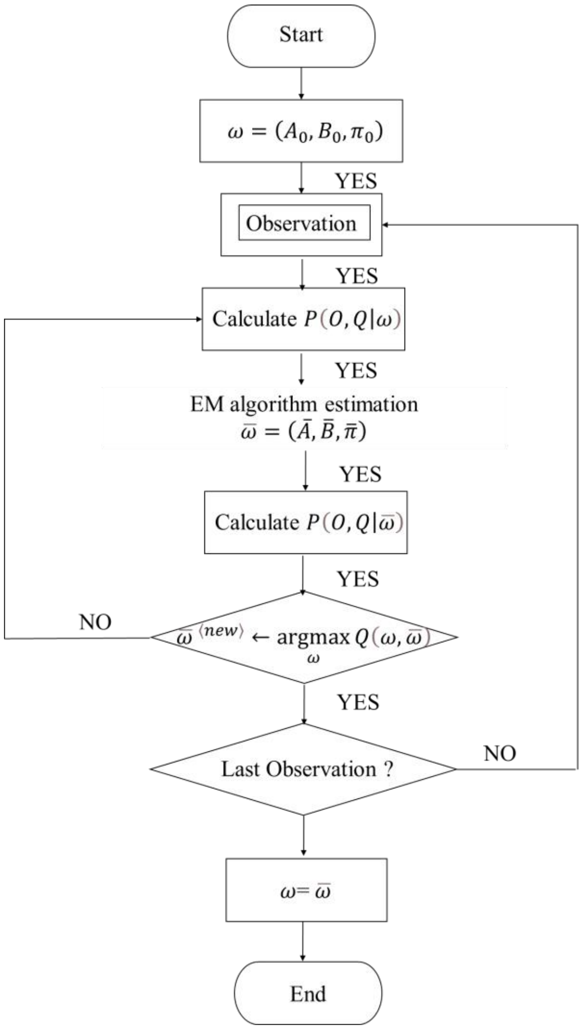

Start
$\omega = (A_0, B_0, \pi_0)$
YES
Observation
YES
Calculate $P(O, Q|\omega)$
YES
EM algorithm estimation
$\overline{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$
YES
Calculate $P(O, Q|\overline{\omega})$
YES
NO
$\overline{\omega}^{\langle new\rangle} \leftarrow \underset{\omega}{\operatorname{argmax}}\, Q(\omega, \overline{\omega})$
YES
Last Observation ?
NO
$\omega = \overline{\omega}$
End

Fig. 8. The EM algorithm

```
Algorithm 1: EM algorithm
Input: Observation sequence $O$, Initial $\omega_0 = (A_0, B_0, \pi_0)$
Output: Updated parameters of HMM-4-C, $\overline{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$
While until convergence do
    $a_{ij}{}^{\langle new\rangle} = \frac{\sum_{t=1}^{T-1} P(O, q_t = S_i, q_{t+1} = S_j | \omega^{\langle old\rangle})}{\sum_{t=1}^{T-1} P(O, q_t = S_i | \omega^{\langle old\rangle})}$
    $b_j(k)^{\langle new\rangle} = \frac{\sum_{t=1}^{T} P(O, q_t = S_i | \omega^{\langle old\rangle}) Q(o_t = VT_k)}{\sum_{t=1}^{T} P(O, q_t = S_i | \omega^{\langle old\rangle})}$
    $\pi_i{}^{\langle new\rangle} = \frac{P(O, q_1 = S_i | \omega^{\langle old\rangle})}{P(O | \omega^{\langle old\rangle})}$
return $\overline{\omega} = (\bar{A}, \bar{B}, \bar{\pi})$
```

E. Clustering

Trajectories naturally exhibit different patterns even under the same scenarios. The diversity comes from different human drivers. If we only use a single pattern to train the model, it may lead to overfitting to that specific pattern and fail to generalize to others. To avoid this, the training datasets should to be picked from a variety of patterns even within the same scenarios.

Typically, trajectory-based clustering research employs k-means method. In this research, we used k -medoids (**Algorithm 2**) clustering method because it more robust as compared to k-means. The difference from k -means is that it uses actual data points as the centers of the clusters rather than the mean of the points in the cluster [81]. This makes k-medoids more robust to noise and outliers compared to k-means. In the subsequent section, we undertake a comparative analysis of detection models, examining their performance both with and without the integration of clustering, to elucidate its impact in the context of cyberattack trajectories research.

```
Algorithm 2: Iterative $k$ -Medoids Clustering
Input: Trajectory datasets;
Output: Clustering results;
Set $\widehat{M} = \widetilde{VT}$
while $New\ Clusters \neq Old\ Clusters$ do
    $[C, M] = kmedoids(\widetilde{VT}, \widehat{M})$
    for $j$ = 1 to length ($M$) do
        $\widetilde{M}_j = \left\{ M_j - \max(M_j) < \frac{\max(M_j)(100 - \tau)}{100} \right.$
        $\widehat{M}_j = \underset{\forall M_j \in \widetilde{M}_j}{\operatorname{argmax}} \sum_{\forall i} M_j(i)$
    end for
    $\widehat{M} = \{ \widehat{M}_1, \dots, \widehat{M}_n \}$
end while
$M^* = \{ \widehat{M}_1, \dots, \widehat{M}_k \}$, where $|C_k| \geq p$
```

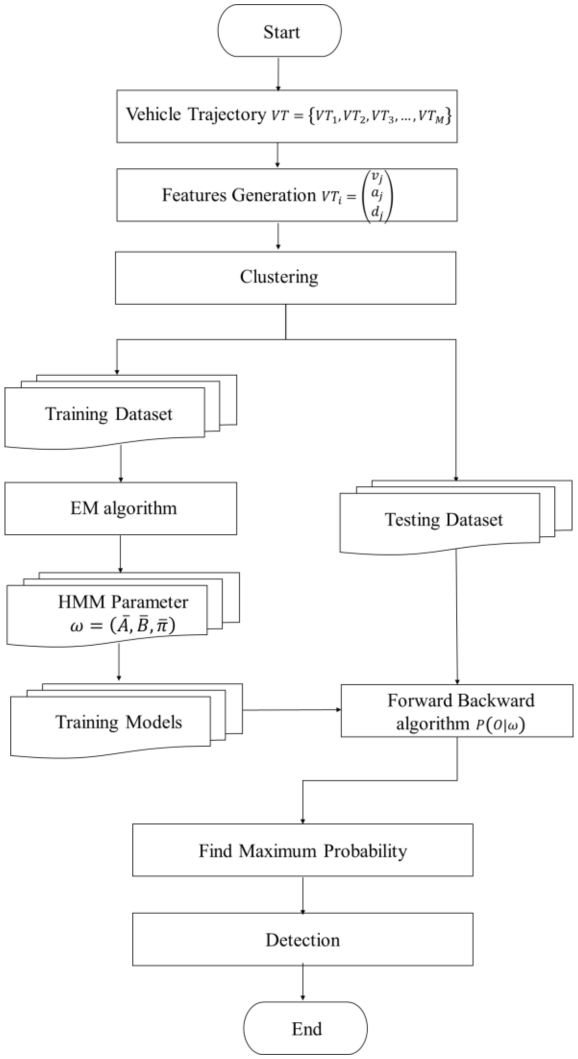

Start
Vehicle Trajectory $VT = \{VT_1, VT_2, VT_3, \ldots, VT_M\}$
Features Generation $VT_i = \begin{pmatrix} v_j \\ a_j \\ d_j \end{pmatrix}$
Clustering
Training Dataset
EM algorithm
HMM Parameter
$\omega = (\bar{A}, \bar{B}, \bar{\pi})$
Training Models
Testing Dataset
Forward Backward
algorithm $P(O|\omega)$
Find Maximum Probability
Detection
End

Fig. 9. HMM-4-C model and detection.

F. *Training and Testing Datasets*

The vehicle trajectory datasets were divided into two parts: the training and testing datasets. As shown in Fig. 9, we first collected all the trajectory data during the experiment. Then we split the data into two parts: training datasets and testing datasets. Training datasets were used to build the HMM-4-C model using the EM algorithm (**Algorithm 1**). We built three detection models by using corresponding training datasets. For the testing datasets, we combined them together and anonymized the data. Then, **Algorithm 3** was applied using the built models and anonymized data. Specifically, for each testing dataset, the forward and backward algorithm was used to calculate its probability under the current model. Therefore, we obtained three different probabilities for a testing dataset, each corresponding to one of the three models. The detection result of is the maximum probability calculated by the forward and backward algorithm under the corresponding model. All the detection results, including true positive (TP), false negative (FN), false positive (FP), and true negative (TN) are recorded for creating the confusion matrix.

```
Algorithm 3: HMM-4-C Detection
  Input: Testing datasets;
  Output: detection results;
  for i = each testing dataset VT do
      for j = each Model do
          //calculate the probability of specific test sequences under different models
          Prob (j) = Forward and backward Algorithm (Model(j), VT(i) )
          Prob of VT(i) = Append (Prob (j))
      end for
      //choose the maximum probability
      Most likely = maximum (Prob of VT (i))
      detection/recognition = index (maximum (Prob of VT (i)))
  end for
  return detection results;
```

G. *Bidirectional Long Short-Term Memory (Bi-LSTM)*

To evaluate HMM-4-C proposed in the research, a neural network known as Bidirectional Long Short-Term Memory (Bi-LSTM) was used to comparation. Bi-LSTM considers both past and future data in its learnings, enhances its learning ability than LTSM. Detailed model of Bi-LSTM can be found in the following literature [82]–[84]. To evaluate the performance of HMM-4-C, same trajectory datasets were also modeled using the Bi-LSTM. To prepare the vehicle trajectory datasets for analysis, the vehicle trajectory datasets were also divided into the same training and testing datasets. The trajectory datasets were modeled using the Bi-LSTM and compared with HMM-4-C in the results section.

H. *Evaluation of the model's performance*

In the research, we conduct accuracy, precision, recall as well as F1 score to describe the detection performance. All these methods can be calculated using the values in the confusion matrix. Accuracy, precision and recall are commonly used in the fields of machine learning, particularly in classification and detection tasks [85]–[87]. Accuracy measures the proportion of true results among the total number of cases examined. Precision assesses how many of the items identified as positive by the model are actually positive. Recall measures how many of the actual positives the model captures by labeling them as positive. Often, there's a trade-off between recall and precision. Improving recall may reduce precision and vice versa. To consider both precision and recall, one can use the F-measure score, also named F1 score is used to both consider both precision and recall. Specifically, F1 score is defined by the results of precision and recall. Performance measurement methods are defined in Equation 32-35.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{32}$$

$$Precision = \frac{TP}{TP + FP} \tag{33}$$

$$Recall = \frac{TP}{TP + FN} \tag{34}$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (35)$$

Where TP is True Positive numbers, TN is True Negative numbers, FP is False Positive numbers FP and FN is False Negative numbers.

IV. RESULTS AND DISCUSSIONS

A. Simulator Experiment Result: Trajectory Analysis

The driving simulator has collected the trajectory data when they drive through the three different scenarios, namely, “Pass”, “Stop”, and “Cyberattack”. The time distance diagrams are presented in Fig. 10.

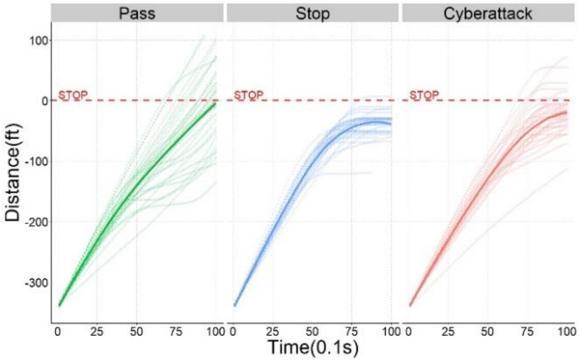
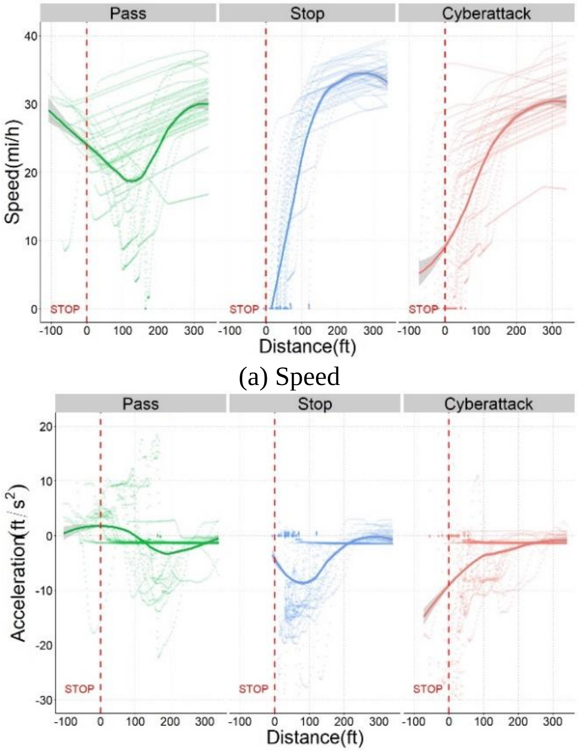


Fig. 10. Time distance diagram in three different scenarios.

Thirty-two vehicle trajectories are combined to generate loess regression curves in each scenario. The time 0 denotes the moment at which the RLCD message was displayed to the drivers. The distance 0 indicates the location of the stop line.



(a) Speed

(b) Acceleration

Fig. 11. The feature of speed and acceleration with distance under three scenarios.

In the Pass scenario, participants faced a short-duration RLCD. Most of the participants drove through the intersection at a consistent speed. A few participants slowed down with little uncertainty when facing the red light before approaching the intersection. In the Stop scenario, participants faced a long-duration red light and had to come to a complete stop before getting into the intersection. Most participants stopped a distance away from the stop line. In the Cyberattack scenario, seven out of thirty-two participants drove into the intersection and caused severe car accidents. The participants had not realized the red light was still on after the RLCD ended, and the vehicle kept approaching the intersection. As a result, facing a cyberattack changed driving behaviors and vehicle trajectories. These unusual driving behaviors provided unique features for a feasible detection.

Fig. 11 depicts the speed and acceleration diagram for each scenario. In the Pass scenario, fifteen drivers maintained a stable deceleration rate to pass the intersection, showing the straight speed curves in Fig. 11b. It indicates that these drivers did not use the brake or gas pedal during that time but kept a uniform deceleration rate. The rest of the participants slowed their speed when they approached the intersection, even though they had a short-duration RLCD message. When the countdown ended, the light turned green. They were required to increase their speed to cross the intersection when approximately 100 feet away from the stop line. Eventually, speed curves in Fig. 11a show the "V" type. The "V" type speed curves correspond to acceleration curves in Fig. 11b. The acceleration curves show a negative value of 10 ft/s^2 followed by an increase to a positive value of 10 ft/s^2 to accelerate and pass through the intersection. In the Stop scenario, participants continuously used brakes and reduced the speed when they approached the. Most participants began using the brakes at around 200ft away from the stop line. The deceleration rate generally came to zero when the vehicle came to a stop.

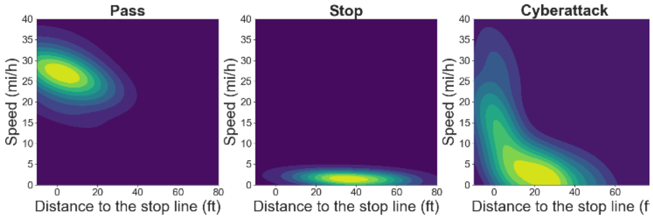


Fig. 12. Velocity measurements prior to the stop line across three different scenarios

Fig. 12 shows the distribution of velocities prior to the stop line under three scenarios. Diverse driving behaviors are observed under the cyberattack. This indicates that drivers exhibit the diversity of driving behaviors under cyberattack scenarios compared to the other two scenarios The time space diagram under cyberattack is shown in Fig. 13. Attack period shown in the figure means the period following the initial five seconds of falsification. The deceleration curves show that participants applied a high value close to negative 30ft/s^2. The hard braking action suggests the unexpected events that occurred at that time. Fig. 14 detailly shows that seven participants were attacked successfully and continued to enter the intersection when the red light was on, resulting in serious accidents. The rest of participants attempted to stop by applying the brakes abruptly, using a significant and characteristic deceleration rate.

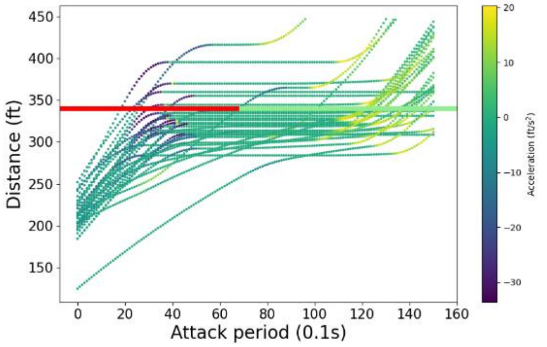


Fig. 13. The time space diagram under cyberattack scenario.

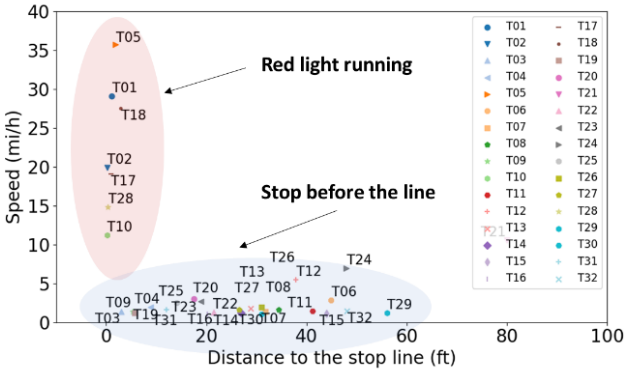


Fig. 14. Diversity in participants' behaviors under cyberattack scenario.

B. Modeling and clustering research

In total, we collected 96 trajectory datasets from the 32 participants. Driving behaviors are diverse under cyberattack situation, as shown in Figs. 13 and 14. Several participants applied hard breaking while the rest of them did not. Additionally, some drivers entered into the intersection while others stopped before the stop line.

The diversity comes from the different driving behaviors under cyberattack scenario. In this case, we perform k-medoids clustering based on the braking behaviors. The result of k-medoids clustering is presented in Fig. 15.

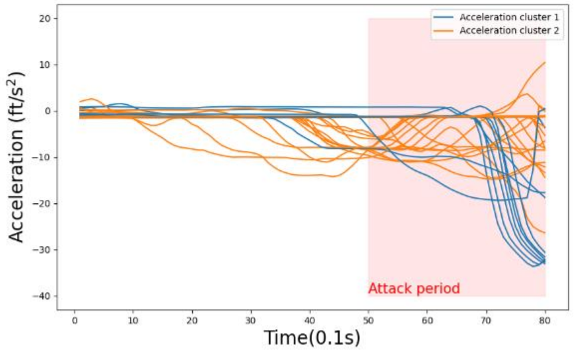


Fig. 15. Acceleration based k-medoids clustering.

People who used hard breaking were selected through k-medoids clustering. The trajectories of these drivers exhibit relatively similar patterns. In this research, we selected two groups which based on the diversity of behaviors: the red light running and unexpected stopping.

After clustering the data, we try to analyze the models' performance before and after clustering. We performed k-medoids clustering to generate different patterns of trajectory groups. In the clustering model, we acquired 20 training datasets, each randomly selected from the clustered datasets. Three different models were then trained using corresponding selected training datasets. The remaining 12 trajectories in each scenario were used as testing datasets to evaluate the performance of the model. In the no-clustering model, the training datasets were selected randomly.

In the testing step, we tested 36 anonymous testing datasets, including 12 datasets from each scenario. The corresponding scenario with the maximum probability was considered as the detection result. Hundreds of training and testing steps were performed using different random seeds. Ultimately, we collected all the detection results to compose the confusion matrix. The matrix includes all the detection results, including true positive (TP), false negative (FN), false positive (FP), and true negative (TN), as shown in Fig. 16.

The comparative analysis results clearly indicate that the HMM-4-C model exhibits improved performance following the application of clustering. Clustering serves as a pivotal method for enhancing the detection of cyberattacks at the physical layer. The result highlights the efficacy of clustering as a robust tool in behavioral analysis within cyber-physical systems.

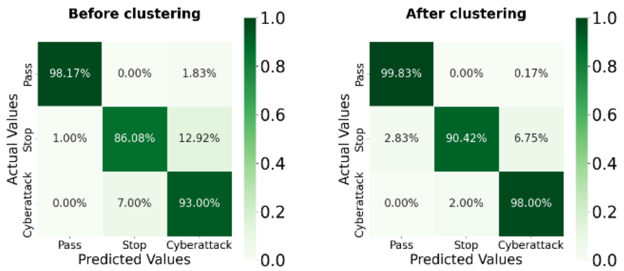


Fig. 16. Detection results of before and after clustering.

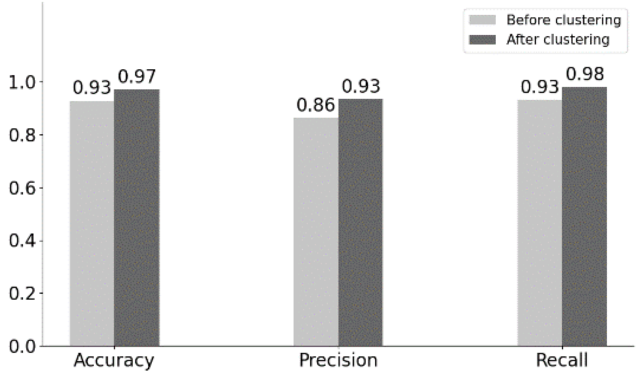


Fig. 17. Performance comparison before and after clustering.

C. *Model scalability and real-world application*

The real-world application is a crucial aspect of evaluating the effectiveness of a detection method. We further conduct sensitivity analyses of time-period-based detection and distance-based detection to explore their potential for future real-world applications. Fig. 18 illustrates the relationship between the model's performance and the length of the attack period. Attack period at 10 means that we use only one second of data length for modeling, specifically, the seconds following the five RLCD message.

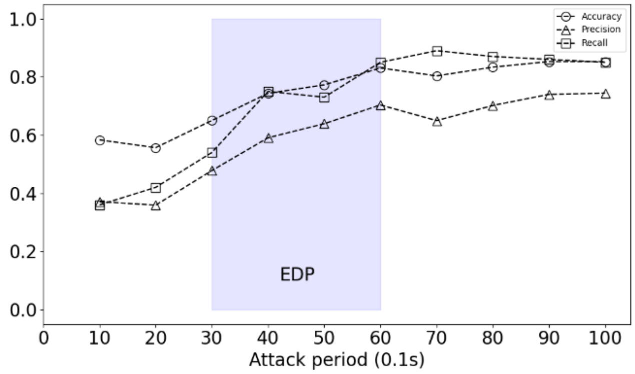


Fig. 18. Model performance with time-period-based dataset inputs.

The fiugre shows that the period from three to six seconds following the RLCD message significantly improve the the performance of HMM-4-C. During the attack period, the model's performance does not show significant improvement in the first three seconds. The reason may be that the participants naturally exhibit the perception-reaction time (PRT), which normally takes two or three seconds to respond the unexcepted situation. Therefore, the initial time period may not provide sufficient information for the physical layer-based cyberattack detection method in real world. From the perspective of protectors, to effectively detect data spoofing attacks on intelligent traffic signal in connected vehicle environment, we should focus more on the period following the normal signal timing plan plus the PRT to capture the physical layer features for detection. We define this period as the Effective Detection Period (EDP).

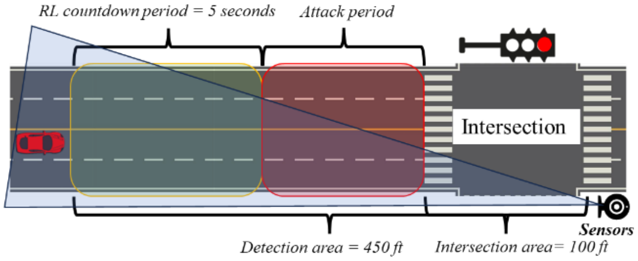


Fig. 19. Real-world physical layer cyberattacks detection.

Moreover, infrastructure-side sensors technologies are more practical for current cyberattack detection based on the physical layer [53]. Therefore, we conduct a sensitivity analysis to explore the relationship between detection distance and the accuracy of detection.

Fig. 19 shows the potential real-world setting for infrastructure-side detection. The sensors are installed at the intersection and oriented towards the direction of approaching vehicles. In the case, vehicle trajectory can be captured in real-time. The trajectory data is capture as soon as the vehicle enters the detection area. However, the sensors can only capture trajectory data from the very beginning up to the current time. Certainly, the more trajectory data that is captured, the higher the quality of the model that can be built, but this requires even more time. From the perspective of cyberattack protectors, the faster a cyberattack is detected, the safer the intersection will be. Fig. 20 illustrates the relationship between detection distance and model performance in this setting. As observed in figure, detection distance plays a significant role within approximately 50 feet near the stop line. We can define this range as the Effective Detection Distance (EDD).

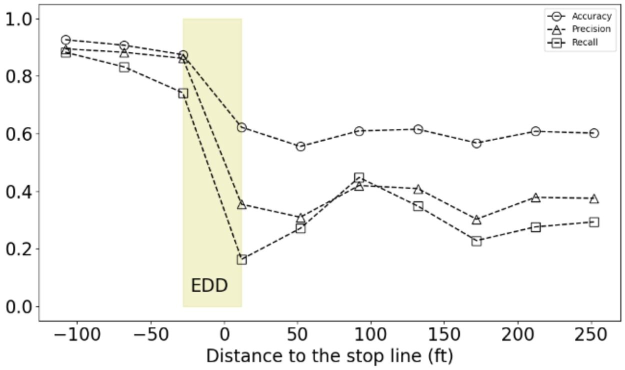


Fig. 20. Model performance with distance-based detection.

D. *Performance of HMM-4-C*

In order to compare two models, we performed a hundred epochs for training Bi-LSTM model until achieving stability. The training accuracy and loss of training process are shown in Fig. 21.

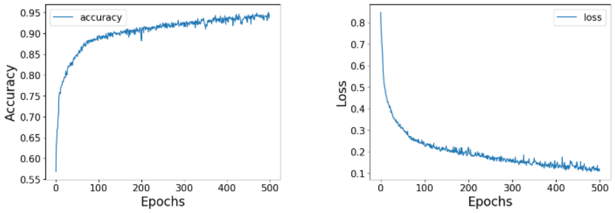


Fig. 21. The training accuracy and loss of the Bi-LSTM.

Two models used the same datasets for both training and testing. The final confusion matrixes of two models are shown in Fig. 22.

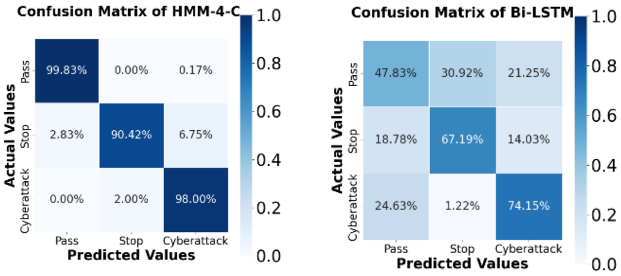


Fig. 22. Final Detection results.

HMM-4-C showed higher detection accuracy for all three scenarios, with the Pass scenario achieving 99.83% true positive rate and the Stop and Cyberattack scenarios achieving 90.42% and 98.00% true positive rates, respectively. However, there were some misdetections, the 6.75% of the Stop testing datasets misdetection arises because the features of cyberattack trajectories are more similar to the stop scenario than to the pass scenario. The detection results from the confusion matrix of Bi-LSTM show a relative lower accuracy than HMM-4-C. The Pass scenario achieving 47.83% true positive rate and the Stop and Cyberattack scenarios achieving 67.19% and 74.15% true positive rates, respectively. In the end, the detection results show that 98% of cyberattack trajectories were successfully detected using HMM-4-C. HMM-4-C performed better than the Bi-LSTM in detecting cyberattacked trajectories.

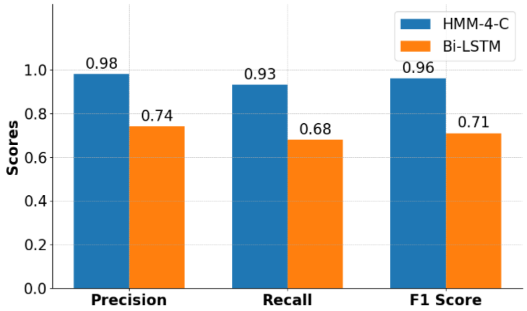


Fig. 23. The comparison of cyberattack detection results in the two models.

The results of precision, recall and F1 score are presented in Fig. 23. In the term of cyberattack detection results, HMM-4-C correctly predicted positive instances 98% of the time when it claimed something was positive, while the Bi-LSTM model was correct 74% of the time. The application of detection cyberattack events always require high precision, HMM-4-C would be preferable based on the provided results in the context. Moreover, HMM-4-C correctly identified 93% of the actual positive events, while the Bi-LSTM model identified 68% of them.

Given the F1 scores of the two models, it's evident that HMM-4-C outperformed the Bi-LSTM on the cyberattack datasets, with an F1 score of 0.96 versus 0.71, respectively. These indicates the high detection results achieved by the model demonstrates the potential of using HMM-4-C and trajectory data for detecting anomalous driving behaviors, which can contribute to the development of more robust and secure transportation systems.

From the perspective of detection results, HMM-4-C achieves better outcomes than Bi-LSTM. In many areas, Bi-LSTM and other deep learning methods have outperformed HMMs [88]–[90]. However, in this research, the higher detection accuracy is attributed to the characteristics features under the cyberattack scenario. One reason is that HMMs are inherently designed for sequences and can effectively capture short-term temporal dependencies between states [91], [92]. HMMs based methods are based on a probabilistic framework which provides a natural way to handle uncertainties in vehicle trajectories. Since cyberattacks occur abruptly and within seconds, the participants react by applying the brakes suddenly, as shown in Figs. 10 and 11. This leads to significant and characteristic state changes over a short period. The trajectory patterns for detecting cyberattacks are mostly local (short sequences) [93]. Therefore, HMM-4-C is better suited than the deep learning methods, which often captures longer dependencies.

In addition, HMMs can work effectively with smaller datasets than other deep learning methods [94]. In our research, training HMM-4-C requires iterative procedures i.e., the EM algorithm, which only converge with 20 examples. As mentioned by previous literature, HMMs outperform Bi-LSTMs when the training dataset is insufficient [95]. Most deep learning models, like Bi-LSTMs, typically require larger datasets to train effectively and generalize well [96]. The intricate architecture allows them to model long-term dependencies and complex relationships in the data. This capability is a double-edged sword: while it lets the model capture sophisticated patterns, it also demands more data to prevent overfitting and to train effectively. For cyberattacks with smaller dataset sizes of vehicle trajectories, HMMs are more suitable in current situation.

V. CONCLUSIONS

This study introduces a novel HMMs based method called HMM-4-C for detecting cyberattacks at connected signalized intersections. The findings validate that Markov Chain-based detection methods have an edge in detecting cyberattacks. This method provides a proof of the concept from a newer angle to achieve cyberattack detection from the physical layer in the cyber-physical system instead of addressing the issue from the traditional perspective of cyber layer. The research has the potential to predict and proactively detect cyberattacks at connected signalized intersections. It can be used in coordination with cyber-based methods to further confirm the occurrences of cyberattacks in the early stages.

This current study has following limitations. The training and testing datasets were based on driving simulator. Future research will take a further step by involving field road driving to validate the findings. In the physical layer detection setting, successful detection might often occur only after a vehicle physically and mistakenly enters the intersection, that is, after a disastrous event has already occurred. Therefore, infrastructure-side sensor detection faces a challenge in this regard. We will implement the advanced hybrid detection model, incorporating data fusion to focus on enhancing accuracy and speeding the detection to address the issue.

VI. REFERENCES

[1] J. F. Powers, *Cyber terrorism and extremism as threat to critical infrastructure protection*. Ljubljana; Tampa: Ministry of Defense, Republic of Slovenia : Institute for Corporative Security Studies : Joint Special Operations University, 2020. Accessed: Dec. 21, 2021. [Online]. Available: https://dk.mors.si/IzpisGradiva.php?id=1121

[2] E. S. Canepa and C. G. Claudel, "Spoofing cyber attack detection in probe-based traffic monitoring systems using mixed integer linear programming," in *2013 International Conference on Computing, Networking and Communications*

(ICNC), San Diego, CA: IEEE, Jan. 2013, pp. 327–333. doi: 10.1109/ICCNC.2013.6504104.

[3] S. K. Khan, N. Shiwakoti, P. Stasinopoulos, and Y. Chen, "Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions," *Accident Analysis & Prevention*, vol. 148, p. 105837, Dec. 2020, doi: 10.1016/j.aap.2020.105837.

[4] B. Sheehan, F. Murphy, M. Mullins, and C. Ryan, "Connected and autonomous vehicles: A cyber-risk classification framework," *Transportation Research Part A: Policy and Practice*, vol. 124, pp. 523–536, Jun. 2019, doi: 10.1016/j.tra.2018.06.033.

[5] M. Gupta, J. Benson, F. Patwa, and R. Sandhu, "Secure V2V and V2I Communication in Intelligent Transportation Using Cloudlets," *IEEE Transactions on Services Computing*, vol. 15, no. 4, pp. 1912–1925, Jul. 2022, doi: 10.1109/TSC.2020.3025993.

[6] S. E. Huang, Y. Feng, and H. X. Liu, "A data-driven method for falsified vehicle trajectory identification by anomaly detection," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103196, Jul. 2021, doi: 10.1016/j.trc.2021.103196.

[7] S. Jung, J. Kim, M. Levorato, C. Cordeiro, and J.-H. Kim, "Infrastructure-Assisted On-Driving Experience Sharing for Millimeter-Wave Connected Vehicles," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7307–7321, Aug. 2021, doi: 10.1109/TVT.2021.3094806.

[8] A. P. Mdee, M. T. R. Khan, J. Seo, and D. Kim, "Security Compliant and Cooperative Pseudonyms Swapping for Location Privacy Preservation in VANETs," *IEEE Transactions on Vehicular Technology*, pp. 1–15, 2023, doi: 10.1109/TVT.2023.3254660.

[9] I. Butun, P. Österberg, and H. Song, "Security of the Internet of Things: Vulnerabilities, Attacks and Countermeasures," *IEEE Commun. Surv. Tutorials*, vol. 22, no. 1, pp. 616–644, 2020, doi: 10.1109/COMST.2019.2953364.

[10] H. Zhang *et al.*, "Emerging Trends in Intelligent Vehicles: The IEEE TIV Perspective," *IEEE Trans. Intell. Veh.*, vol. 8, no. 8, pp. 3983–3995, Aug. 2023, doi: 10.1109/TIV.2023.3291457.

[11] G. Bendiab, A. Hameurlaine, G. Germanos, N. Kolokotronis, and S. Shiaeles, "Autonomous Vehicles Security: Challenges and Solutions Using Blockchain and Artificial Intelligence," *IEEE Trans. Intell. Transport. Syst.*, vol. 24, no. 4, pp. 3614–3637, Apr. 2023, doi: 10.1109/TITS.2023.3236274.

[12] M. Chowdhury, M. Islam, and Z. Khan, "Security of Connected and Automated Vehicles," *arXiv preprint arXiv:2012.13464*, 2020.

[13] J. Haddad and B. Mirkin, "Resilient perimeter control of macroscopic fundamental diagram networks under cyberattacks," *Transportation Research Part B: Methodological*, vol. 132, pp. 44–59, Feb. 2020, doi: 10.1016/j.trb.2019.01.020.

[14] J. Petit and S. E. Shladover, "Potential Cyberattacks on Automated Vehicles," *IEEE Trans. Intell. Transport. Syst.*, pp. 1–11, 2014, doi: 10.1109/TITS.2014.2342271.

[15] S. Grad, "Engineers who hacked into L.A. traffic signal computer, jamming streets, sentenced," LA Times Blogs - L.A. NOW. Accessed: Apr. 17, 2022. [Online]. Available: https://latimesblogs.latimes.com/lanow/2009/12/engineers-who-hacked-in-la-traffic-signal-computers-jamming-traffic-sentenced.html

[16] M. Prigg, "New York's traffic lights HACKED," Mail Online. Accessed: Apr. 17, 2022. [Online]. Available: https://www.dailymail.co.uk/sciencetech/article-2617228/New-Yorks-traffic-lights-HACKED-technique-work-world.html

[17] J. R. Miller, "Hackers Crack Into Texas Road Sign, Warn of Zombies Ahead," Fox News. Accessed: Apr. 17, 2022. [Online]. Available: https://www.foxnews.com/story/hackers-crack-into-texas-road-sign-warn-of-zombies-ahead

[18] D. Morris, G. Madzudzo, and A. G. Perez, "Cybersecurity and the auto industry: the growing challenges presented by connected cars," *IJATM*, vol. 18, no. 2, p. 105, 2018, doi: 10.1504/IJATM.2018.092187.

[19] M. T. Whitty, "Is There a Scam for Everyone? Psychologically Profiling Cyberscam Victims," *Eur J Crim Policy Res*, vol. 26, no. 3, pp. 399–409, Sep. 2020, doi: 10.1007/s10610-020-09458-z.

[20] Y. Benmessaoud, L. Cherrat, and M. Ezziyyani, "Real-Time Self-Adaptive Traffic Management System for Optimal Vehicular Navigation in Modern Cities," *Computers*, vol. 12, no. 4, Art. no. 4, Apr. 2023, doi: 10.3390/computers12040080.

[21] A. M. de Souza, N. L. S. da Fonseca, and L. Villas, "A fully-distributed advanced traffic management system based on opportunistic content sharing," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6. doi: 10.1109/ICC.2017.7997071.

[22] F. van Wyk, Y. Wang, A. Khojandi, and N. Masoud, "Real-Time Sensor Anomaly Detection and Identification in Automated Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 1264–1276, Mar. 2020, doi: 10.1109/TITS.2019.2906038.

[23] M. N. Mejri, N. Achir, and M. Hamdi, "A new security games based reaction algorithm against DOS attacks in VANETs," in *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Jan. 2016, pp. 837–840. doi: 10.1109/CCNC.2016.7444896.

[24] M. S. Faughnan *et al.*, "Risk analysis of Unmanned Aerial Vehicle hijacking and methods of its detection," in *2013 IEEE Systems and Information Engineering Design Symposium*, Apr. 2013, pp. 145–150. doi: 10.1109/SIEDS.2013.6549509.

[25] J. Cui, L. S. Liew, G. Sabaliauskaite, and F. Zhou, "A review on safety failures, security attacks, and available countermeasures for autonomous vehicles," *Ad Hoc Networks*, vol. 90, p. 101823, Jul. 2019, doi: 10.1016/j.adhoc.2018.12.006.

[26] W. Duo, M. Zhou, and A. Abusorrah, "A Survey of Cyber Attacks on Cyber Physical Systems: Recent Advances and Challenges," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 5, pp. 784–800, May 2022, doi: 10.1109/JAS.2022.105548.

[27] L. Karim and A. Boulmakoul, "Trajectory-based Modeling for Fraud Detection and Analytics: Foundation and Design," in *2021 IEEE/ACS 18th International Conference on Computer Systems and Applications (AICCSA)*, Tangier, Morocco: IEEE, Nov. 2021, pp. 1–7. doi: 10.1109/AICCSA53542.2021.9686920.

[28] K. Kumaran Santhosh, D. P. Dogra, P. P. Roy, and A. Mitra, "Vehicular Trajectory Classification and Traffic Anomaly Detection in Videos Using a Hybrid CNN-VAE Architecture," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 11891–11902, Aug. 2022, doi: 10.1109/TITS.2021.3108504.

[29] I. Kalinov *et al.*, "WareVision: CNN Barcode Detection-Based UAV Trajectory Optimization for Autonomous Warehouse Stocktaking," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6647–6653, Oct. 2020, doi: 10.1109/LRA.2020.3010733.

[30] Y. Liu, K. Zhao, G. Cong, and Z. Bao, "Online Anomalous Trajectory Detection with Deep Generative Sequence Modeling," in *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, Apr. 2020, pp. 949–960. doi: 10.1109/ICDE48307.2020.00087.

[31] C. Zhang, Z. Ni, and C. Berger, “Spatial-Temporal-Spectral LSTM: A Transferable Model for Pedestrian Trajectory Prediction,” *IEEE Transactions on Intelligent Vehicles*, pp. 1–14, 2023, doi: 10.1109/TIV.2023.3285804.

[32] T. Hickling, N. Aouf, and P. Spencer, “Robust Adversarial Attacks Detection based on Explainable Deep Reinforcement Learning for UAV Guidance and Planning,” *IEEE Transactions on Intelligent Vehicles*, pp. 1–14, 2023, doi: 10.1109/TIV.2023.3296227.

[33] Y. Xue and W. Chen, “Multi-Agent Deep Reinforcement Learning for UAVs Navigation in Unknown Complex Environment,” *IEEE Transactions on Intelligent Vehicles*, pp. 1–14, 2023, doi: 10.1109/TIV.2023.3298292.

[34] N. Lin, C. Zong, M. Tomizuka, P. Song, Z. Zhang, and G. Li, “An Overview on Study of Identification of Driver Behavior Characteristics for Automotive Control,” *Mathematical Problems in Engineering*, vol. 2014, pp. 1–15, 2014, doi: 10.1155/2014/569109.

[35] C.-E. Wu, W.-Y. Yang, H.-C. Ting, and J.-S. Wang, “Traffic pattern modeling, trajectory classification and vehicle tracking within urban intersections,” in *2017 International Smart Cities Conference (ISC2)*, Sep. 2017, pp. 1–6. doi: 10.1109/ISC2.2017.8090791.

[36] Y. Zhao, S. Shen, and H. X. Liu, “A hidden Markov model for the estimation of correlated queues in probe vehicle environments,” *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103128, Jul. 2021, doi: 10.1016/j.trc.2021.103128.

[37] Q. He, X. Meng, and R. Qu, “Towards a Severity Assessment Method for Potential Cyber Attacks to Connected and Autonomous Vehicles,” *Journal of Advanced Transportation*, vol. 2020, pp. 1–15, Sep. 2020, doi: 10.1155/2020/6873273.

[38] S. F. Meyer, R. Elvik, and E. Johnsson, “Risk analysis for forecasting cyberattacks against connected and autonomous vehicles,” *J Transp Secur*, vol. 14, no. 3–4, pp. 227–247, Dec. 2021, doi: 10.1007/s12198-021-00236-4.

[39] Á. Török, Z. Szalay, G. Uti, and B. Verebélyi, “Modelling the effects of certain cyber-attack methods on urban autonomous transport systems, case study of Budapest,” *J Ambient Intell Human Comput*, vol. 11, no. 4, pp. 1629–1643, Apr. 2020, doi: 10.1007/s12652-019-01264-8.

[40] P. Wang, G. Yu, X. Wu, H. Qin, and Y. Wang, “An extended car-following model to describe connected traffic dynamics under cyberattacks,” *Physica A: Statistical Mechanics and its Applications*, vol. 496, pp. 351–370, Apr. 2018, doi: 10.1016/j.physa.2017.12.013.

[41] L. Cui, J. Hu, B. B. Park, and P. Bujanovic, “Development of a simulation platform for safety impact analysis considering vehicle dynamics, sensor errors, and communication latencies: Assessing cooperative adaptive cruise control under cyber attack,” *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 1–22, Dec. 2018, doi: 10.1016/j.trc.2018.10.005.

[42] M. Amoozadeh *et al.*, “Security vulnerabilities of connected vehicle streams and their impact on cooperative driving,” *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 126–132, Jun. 2015, doi: 10.1109/MCOM.2015.7120028.

[43] Y. Li, Y. Tu, Q. Fan, C. Dong, and W. Wang, “Influence of cyber-attacks on longitudinal safety of connected and automated vehicles,” *Accident Analysis & Prevention*, vol. 121, pp. 148–156, Dec. 2018, doi: 10.1016/j.aap.2018.09.016.

[44] P. Wang, X. Wu, and X. He, “Modeling and analyzing cyberattack effects on connected automated vehicular platoons,” *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102625, Jun. 2020, doi: 10.1016/j.trc.2020.102625.

[45] Z. H. Khattak, B. L. Smith, and M. D. Fontaine, “Impact of cyberattacks on safety and stability of connected and automated vehicle platoons under lane changes,” *Accident Analysis & Prevention*, vol. 150, p. 105861, Feb. 2021, doi: 10.1016/j.aap.2020.105861.

[46] 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*, Linz Austria: ACM, Jul. 2020, pp. 65–76. doi: 10.1145/3395351.3399348.

[47] X. Liu, B. Luo, A. Abdo, N. Abu-Ghazaleh, and Q. Zhu, “Securing Connected Vehicle Applications with an Efficient Dual Cyber- Physical Blockchain Framework,” in *2021 IEEE Intelligent Vehicles Symposium (IV)*, Nagoya, Japan: IEEE, Jul. 2021, pp. 393–400. doi: 10.1109/IV48863.2021.9575869.

[48] N. Zhao, X. Zhao, N. Xu, and L. Zhang, “Resilient Event-Triggered Control of Connected Automated Vehicles Under Cyber Attacks,” *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 12, pp. 2300–2302, Dec. 2023, doi: 10.1109/JAS.2023.123483.

[49] Z. Abdollahi Biron, S. Dey, and P. Pisu, “Real-Time Detection and Estimation of Denial of Service Attack in Connected Vehicle Systems,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3893–3902, Dec. 2018, doi: 10.1109/TITS.2018.2791484.

[50] E. Mousavinejad, F. Yang, Q.-L. Han, X. Ge, and L. Vlacic, “Distributed Cyber Attacks Detection and Recovery Mechanism for Vehicle Platooning,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3821–3834, Sep. 2020, doi: 10.1109/TITS.2019.2934481.

[51] D. Zhang, Y.-P. Shen, S.-Q. Zhou, X.-W. Dong, and L. Yu, “Distributed Secure Platoon Control of Connected Vehicles Subject to DoS Attack: Theory and Application,” *IEEE Trans. Syst. Man Cybern, Syst.*, vol. 51, no. 11, pp. 7269–7278, Nov. 2021, doi: 10.1109/TSMC.2020.2968606.

[52] S. Xiao, X. Ge, Q.-L. Han, and Y. Zhang, “Secure Distributed Adaptive Platooning Control of Automated Vehicles Over Vehicular Ad-Hoc Networks Under Denial-of-Service Attacks,” *IEEE Transactions on Cybernetics*, pp. 1–13, 2021, doi: 10.1109/TCYB.2021.3074318.

[53] J. Shen, Z. Wan, Y. Luo, Y. Feng, Z. M. Mao, and Q. A. Chen, “Detecting Data Spoofing in Connected Vehicle based Intelligent Traffic Signal Control using Infrastructure-Side Sensors and Traffic Invariants,” in *2023 IEEE Intelligent Vehicles Symposium (IV)*, Anchorage, AK, USA: IEEE, Jun. 2023, pp. 1–8. doi: 10.1109/IV55152.2023.10186689.

[54] Y. Feng, S. Huang, Q. A. Chen, H. X. Liu, and Z. M. Mao, “Vulnerability of Traffic Control System Under Cyberattacks with Falsified Data,” *Transportation Research Record*, vol. 2672, no. 1, pp. 1–11, Dec. 2018, doi: 10.1177/0361198118756885.

[55] 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 *Proceedings 2018 Network and Distributed System Security Symposium*, San Diego, CA: Internet Society, 2018. doi: 10.14722/ndss.2018.23222.

[56] B. DeBruhl, S. Weerakkody, B. Sinopoli, and P. Tague, “Is your commute driving you crazy?: a study of misbehavior in vehicular platoons,” in *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, New York New York: ACM, Jun. 2015, pp. 1–11. doi: 10.1145/2766498.2766505.

[57] W. Wei, H. Song, H. Wang, and X. Fan, “Research and Simulation of Queue Management Algorithms in Ad Hoc Networks Under DDoS Attack,” *IEEE Access*, vol. 5, pp. 27810–27817, 2017, doi: 10.1109/ACCESS.2017.2681684.

[58] A. A. Alsulami, Q. Abu Al-Haija, A. Alqahtani, and R. Alsini, “Symmetrical Simulation Scheme for Anomaly Detection in Autonomous Vehicles Based on LSTM Model,” *Symmetry*, vol. 14, no. 7, p. 1450, Jul. 2022, doi: 10.3390/sym14071450.

[59] M. Basnet and M. H. Ali, “A Deep Learning Perspective on Connected Automated Vehicle (CAV) Cybersecurity and Threat Intelligence,” p. 21, 2021.

[60] S. Iqbal, P. Ball, M. H. Kamarudin, and A. Bradley, “Simulating Malicious Attacks on VANETs for Connected and Autonomous Vehicle Cybersecurity: A Machine Learning Dataset,” p. 12, 2022.

[61] S. Biswas, I. Ghosh, and S. Chandra, “Influence of signal countdown timer on efficiency and safety at signalized intersections,” *Can. J. Civ. Eng.*, vol. 44, no. 4, pp. 308–318, Apr. 2017, doi: 10.1139/cjce-2016-0267.

[62] J. Henry, “Honda, Ohio aim to make smart-mobility corridor even smarter,” Automotive News. Accessed: Apr. 18, 2022. [Online]. Available: https://www.autonews.com/mobility-report/honda-ohio-aim-make-smart-mobility-corridor-even-smarter-more-connected-cars

[63] K.-F. Wu, M. N. Ardiansyah, and W.-J. Ye, “An evaluation scheme for assessing the effectiveness of intersection movement assist (IMA) on improving traffic safety,” *Traffic Injury Prevention*, vol. 19, no. 2, pp. 179–183, Feb. 2018, doi: 10.1080/15389588.2017.1363891.

[64] R. Stumpf, “Audi’s new tech can help you beat red lights,” Popular Science. Accessed: Nov. 13, 2023. [Online]. Available: https://www.popsci.com/technology/new-audi-tech-provides-traffic-light-updates/

[65] Y.-C. Chiou and C.-H. Chang, “Driver responses to green and red vehicular signal countdown displays: Safety and efficiency aspects,” *Accident Analysis & Prevention*, vol. 42, no. 4, pp. 1057–1065, Jul. 2010, doi: 10.1016/j.aap.2009.12.013.

[66] S. Parkinson, P. Ward, K. Wilson, and J. Miller, “Cyber Threats Facing Autonomous and Connected Vehicles: Future Challenges,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 2898–2915, Nov. 2017, doi: 10.1109/TITS.2017.2665968.

[67] M. Islam, M. Chowdhury, H. Li, and H. Hu, “Cybersecurity Attacks in Vehicle-to-Infrastructure Applications and Their Prevention,” *Transportation Research Record*, vol. 2672, no. 19, pp. 66–78, Dec. 2018, doi: 10.1177/0361198118799012.

[68] T. Mecheva and N. Kakanakov, “Cybersecurity in Intelligent Transportation Systems,” *Computers*, vol. 9, no. 4, p. 83, Oct. 2020, doi: 10.3390/computers9040083.

[69] E. Adams *et al.*, “Development of DSRC device and communication system performance measures recommendations for DSRC OBE performance and security requirements.,” no. FHWA-JPO-17-483, May 2016, [Online]. Available: https://rosap.ntl.bts.gov/view/dot/31627

[70] L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” *PROCEEDINGS OF THE IEEE*, vol. 77, no. 2, p. 30, 1989.

[71] L. R. Rabiner and B. H. Juang, “An Introduction to Hidden Markov Models,” p. 12, 1986.

[72] Y. Li, F. Wang, H. Ke, L. Wang, and C. Xu, “A Driver’s Physiology Sensor-Based Driving Risk Prediction Method for Lane-Changing Process Using Hidden Markov Model,” *Sensors*, vol. 19, no. 12, p. 2670, Jun. 2019, doi: 10.3390/s19122670.

[73] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, “A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains,” *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164–171, 1970.

[74] P. A. Devijver, “Baum’s forward-backward algorithm revisited,” *Pattern Recognition Letters*, vol. 3, no. 6, pp. 369–373, Dec. 1985, doi: 10.1016/0167-8655(85)90023-6.

[75] K. Xie, K. Ozbay, H. Yang, and C. Li, “Mining automatically extracted vehicle trajectory data for proactive safety analytics,” *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 61–72, Sep. 2019, doi: 10.1016/j.trc.2019.07.004.

[76] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum Likelihood from Incomplete Data Via the EM Algorithm,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977, doi: 10.1111/j.2517-6161.1977.tb01600.x.

[77] T. K. Moon, “The expectation-maximization algorithm,” *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, Nov. 1996, doi: 10.1109/79.543975.

[78] A. Churbanov and S. Winters-Hilt, “Implementing EM and Viterbi algorithms for Hidden Markov Model in linear memory,” *BMC Bioinformatics*, vol. 9, no. 1, p. 224, Dec. 2008, doi: 10.1186/1471-2105-9-224.

[79] S. Jeong, Y. Kang, J. Lee, and K. Sohn, “Variational embedding of a hidden Markov model to generate human activity sequences,” *Transportation Research Part C: Emerging Technologies*, vol. 131, p. 103347, Oct. 2021, doi: 10.1016/j.trc.2021.103347.

[80] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, “The variational approximation for Bayesian inference,” *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, Nov. 2008, doi: 10.1109/MSP.2008.929620.

[81] A. Mcfadyen, M. O’Flynn, T. Martin, and D. Campbell, “Aircraft trajectory clustering techniques using circular statistics,” in *2016 IEEE Aerospace Conference*, Big Sky, MT, USA: IEEE, Mar. 2016, pp. 1–10. doi: 10.1109/AERO.2016.7500601.

[82] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, Jul. 2005, doi: 10.1016/j.neunet.2005.06.042.

[83] H. Sun, M. Chen, J. Weng, Z. Liu, and G. Geng, “Anomaly Detection for In-Vehicle Network Using CNN-LSTM With Attention Mechanism,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10880–10893, Oct. 2021, doi: 10.1109/TVT.2021.3106940.

[84] H. Zhang, Z. Nan, T. Yang, Y. Liu, and N. Zheng, “A Driving Behavior Recognition Model with Bi-LSTM and Multi-Scale CNN,” in *2020 IEEE Intelligent Vehicles Symposium (IV)*, Oct. 2020, pp. 284–289. doi: 10.1109/IV47402.2020.9304772.

[85] S. Sivaraman and M. M. Trivedi, “A General Active-Learning Framework for On-Road Vehicle Recognition and Tracking,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 267–276, Jun. 2010, doi: 10.1109/TITS.2010.2040177.

[86] S. Kumar, K. Singh, S. Kumar, O. Kaiwartya, Y. Cao, and H. Zhou, “Delimitated Anti Jammer Scheme for Internet of Vehicle: Machine Learning Based Security Approach,” *IEEE Access*, vol. 7, pp. 113311–113323, 2019, doi: 10.1109/ACCESS.2019.2934632.

[87] B. He, R. Ai, Y. Yan, and X. Lang, “Accurate and robust lane detection based on Dual-View Convolutional Neutral Network,” in *2016 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2016, pp. 1041–1046. doi: 10.1109/IVS.2016.7535517.

[88] T. Zia and U. Zahid, “Long short-term memory recurrent neural network architectures for Urdu acoustic modeling,” *Int J Speech Technol*, vol. 22, no. 1, pp. 21–30, Mar. 2019, doi: 10.1007/s10772-018-09573-7.

[89] Y. Lee, H. Jeon, and K. Sohn, “Predicting Short-Term Traffic Speed Using a Deep Neural Network to Accommodate

Citywide Spatio-Temporal Correlations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1435–1448, Mar. 2021, doi: 10.1109/TITS.2020.2970754.

[90] K. Lee, M. Eo, E. Jung, Y. Yoon, and W. Rhee, "Short-Term Traffic Prediction With Deep Neural Networks: A Survey," *IEEE Access*, vol. 9, pp. 54739–54756, 2021, doi: 10.1109/ACCESS.2021.3071174.

[91] T. Fernando, S. Denman, A. McFadyen, S. Sridharan, and C. Fookes, "Tree Memory Networks for modelling long-term temporal dependencies," *Neurocomputing*, vol. 304, pp. 64–81, Aug. 2018, doi: 10.1016/j.neucom.2018.03.040.

[92] Y. Qi and S. Ishak, "A Hidden Markov Model for short term prediction of traffic conditions on freeways," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 95–111, Jun. 2014, doi: 10.1016/j.trc.2014.02.007.

[93] Q. He, X. Meng, R. Qu, and R. Xi, "Machine Learning-Based Detection for Cyber Security Attacks on Connected and Autonomous Vehicles," *Mathematics*, vol. 8, no. 8, p. 1311, Aug. 2020, doi: 10.3390/math8081311.

[94] M. Levi, Y. Allouche, and A. Kontorovich, "Advanced Analytics for Connected Car Cybersecurity," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Jun. 2018, pp. 1–7. doi: 10.1109/VTCSpring.2018.8417690.

[95] S. Lefèvre, A. Carvalho, and F. Borrelli, "A Learning-Based Framework for Velocity Control in Autonomous Driving," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 32–42, Jan. 2016, doi: 10.1109/TASE.2015.2498192.

[96] Y. Zhang and Z. Lu, "Exploring semi-supervised variational autoencoders for biomedical relation extraction," *Methods*, vol. 166, pp. 112–119, Aug. 2019, doi: 10.1016/j.ymeth.2019.02.021.

Yingfan Gu is currently pursuing the Ph.D. degree with the Civil & Architectural Engineering & Construction Department at the University of Cincinnati, Ohio. His research interests include intelligent transportation systems, trajectory modeling, cybersecurity, and machine learning.

Zhixia Li, Ph.D. received his BE in Electrical and Communication Engineering in 2003, and Ph.D. in Civil Engineering in 2011. He is Associate Professor with University of Cincinnati, Department of Civil and Architectural Engineering and Construction Management. His research areas span in connected and autonomous vehicle, intelligent transportation systems, traffic safety, human factors, GIS-Transportation, sustainable transportation, and traffic simulation.

Yunpeng Zhang received his Ph.D. degree in Computer Science from Northwestern Polytechnical University, Xi'an, China. He is currently an Assistant Professor at the University of Houston, Houston. His research focuses on developing novel security and intelligence techniques to ensure cyber/physical system reliability, security, and performance in multiple industries, including energy, healthcare, smart cities, commerce, transportation, finance, government, defense, the internet of things, etc. Dr. Zhang is familiar with the state-of-the-art research and technologies related to cyber and physical security, artificial intelligence, cryptography, access control, intrusion detection, blockchain, trust management, intelligent monitoring, deep learning, and data analysis.

Shivam Tiwari Joined University of Houston in August 2021. He received his Master in Computer Science from University of Houston in TX, USA in May 2023 and Bachelors in Technology from SRM University, Chennai, TN, India.

Heng Wei is a Professor of Transportation Systems and Engineering in Department of Civil & Architectural Engineering & Construction Management at The University of Cincinnati (UC). He has a wide spectrum of research interests and expertise in advanced transportation systems, including intelligent transportation systems (ITS), travel demand in the context of Smart City, traffic data monitoring and management, traffic operation and safety, transportation conformity, computing and information technologies in transportation infrastructure systems, microscopic traffic simulation modeling, Connected Automated Vehicle (CAV) affected traffic flow theory and behavior modeling, artificial intelligent techniques in transportation, and geographic information system (GIS) application in transportation. Dr. Wei received his M.S. and Ph.D. degrees from The University of Kansas, B.S. and M.S. degrees from Beijing University of Technology, China, all in Civil Engineering.

Guohui Zhang received a Ph.D. degree from the University of Washington (UW), Seattle, in 2000. He is currently a Professor with the Department of Civil and Environmental Engineering, the University of Hawaii at Manoa, Honolulu. His primary research areas include transportation system resilience analysis, large-scale transportation systems modeling, sustainable traffic network infrastructure design, planning, and operation, traffic sensing and senor data analytics, artificial intelligence in transportation, connected and autonomous vehicle systems, traffic-impacted public health, cyber-transportation systems, and transportation safety and security.

Muting Ma received the Ph.D. degree from University of Louisville. He is currently a postdoctoral researcher at the University of Alabama. His main research interests include algorithm design, operation research, and optimization, as well as applications on Connected and Autonomous Vehicles (CAVs).

Sabur Baidya (M’ 20) received the Ph.D. degree in Computer Science from the University of California, Irvine, CA, USA, in 2019, and the M.S. degree in computer science from the University of Texas at Dallas, in 2013. He is currently an Assistant Professor of Computer Science and Engineering in the J.B. Speed School of Engineering at the University of Louisville (UofL), USA. Prior to that he was a postdoctoral scholar in the Electrical and Computer Engineering department at the University of California San Diego (UCSD). He was also a visiting summer researcher in the WINLAB at Rutgers University and has had prior working experience with IBM, Cisco Systems, Huawei Research Lab, and Nokia Bell Labs. His research interests include the areas of the Internet of Things, wireless networks, intelligent and autonomous systems, and edge-cloud computing. He is a member of the ACM, and IEEE Communication Society and IEEE Robotics and Automation Society.