Enhanced Cyber-physical Security of Steering Stability Control System for Four-Wheel Independent Drive Electric Vehicles

Lulu Guo
School of Electrical and Computer Engineering
University of Georgia
Athens, Georgia 30602
Email: lulu.guo@uga.edu

Bowen Yang
School of Electrical and Computer Engineering
University of Georgia
Athens, Georgia 30602
Email: bowen.yang@uga.edu

Jin Ye
School of Electrical and Computer Engineering
University of Georgia
Athens, Georgia 30602
Email: jin.ye@uga.edu

***Abstract*—In this paper, we present a residual-based anomaly detection method to enhance the cyber-physical security of the steering stability control system (SSCS) in a four-wheel independent drive electric vehicle. With the approach of the linear quadratic regulator, the SSCS is developed through the yaw moment generated by the torque deviation between the motors, the goal of which is to improve the lateral stability of the vehicle body. To prevent the vehicle against cyber-physical attacks, e.g., integrity attacks, we propose a residual-based anomaly detection method. Compared to traditional residual-based anomaly detection, the presented method can deal with threats on both control inputs and sensor measurements by combining physics-based and learning-based approaches. Simulation results have shown the effectiveness of the proposed detection method.**

I. Introduction

In recent years, four-wheel independent drive electric vehicles (4WDEVs) have shown significant advantages in the short drive chain, compact structure, and fast generated torque. Due to the quick response to the traction of drive motors, this distributed-driven powertrain platform provides great potential to improve both longitudinal and lateral performances of the vehicle through various approaches, for instance, fuzzy logic control [1], robust control [2], and those optimization-based methods [3], [4]. In most of these steering stability control systems (SSCSs), vehicle stability is pursued by adding a yaw moment through torque split-based technologies. Although the active steering stability control can significantly enhance the yaw stability of the vehicle [5]–[10], once malicious cyber-physical attacks infect the vehicle, the stability control system may instead lead to severe consequences, for instance, disabling brakes, turning off headlights, taking over steering [10]–[12], and real incidents in Cherokee Jeep [13] and Tesla [14].

This concern requires more attention in modern cars because the number and complexity of embedded electronic control units (ECUs) are increasing rapidly. Therefore, the cyber-physical security of the SSCS for 4WDEVs should be addressed. Up to date, there have been many publications making efforts to prevent the network in the vehicle against cyber-physical attacks, such as secure controller area network (CAN) [15], but they alone cannot ensure the safety of the car, especially in the control level. Once the system is affected by cyber-physical threats, the core problem is how to protect the control system, and the first step is to identify the potential attacks and alert the driver for driving safety.

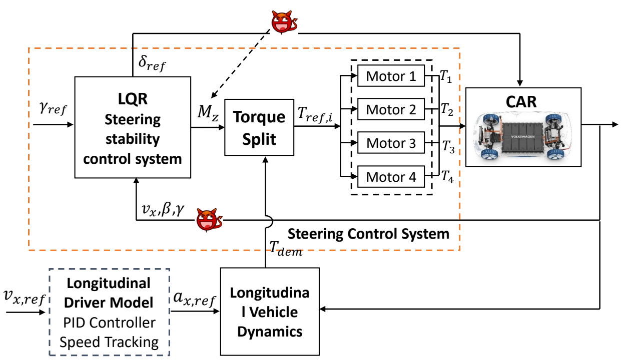

Fig. 1. Diagram of the steering stability control system of the 4WDEV.

In general, the detection methods of a cyber-physical control system can be categorized into two schemes: physics-based methods and data-based methods. In both schemes, the main idea of threat identification is to calculate the residual $r_k = |y(k) - \hat{y}(k)|$, and then $r_k \geq \tau$ (k represents k^{th} time instant, and τ is a threshold) is considered as a proxy for the presence of attacks, such as the works in [16], [17]. In most of the physics-based methods, the residual is obtained by estimating the system outputs through state observers, while in a data-based detector, it is predicted by using regression techniques, e.g., machine learning, deep learning, etc [16]–[20].

In this paper, for a 4WDEV, we design an SSCS and analyze cyber-physical security, based on which we develop a residual-based anomaly detection method to improve the cyber-physical security of 4WDEVs. When calculating the residual, besides the control-theoretical observer, we also use a deep-learning network to predict the system outputs, in which the controller behavior is considered, and thus cyber-attacks on control inputs can be identified, while in most of the physics-based methodologies, only threats on sensor measurements can be identified. The paper is organized as follows. In section II, the longitudinal and lateral control systems are described, based on which the residual-based detector is designed in Section III. Then, comparison results are shown in section IV to validate the effectiveness of the proposed threat detection, and conclusions are given in section V.

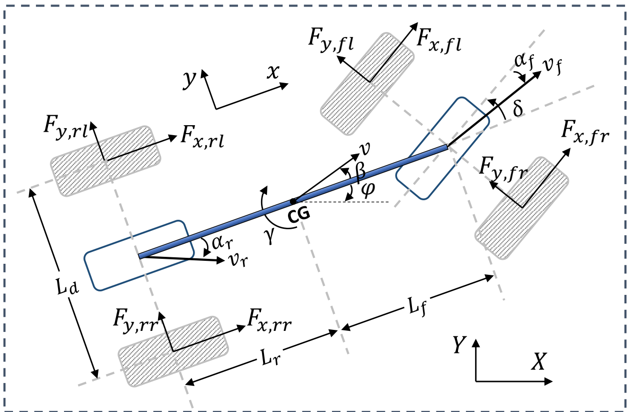

Fig. 2. Diagram of the steering stability control system of the 4WDEV.

II. Vehicle modeling and system description

Fig. 1 presents the lateral and longitudinal control systems in a 4WDEV, in which the longitudinal driver model is developed with a proportional-integral-derivative (PID) controller to track the given speed profile. The total torque demand T_{dem} can be derived from the longitudinal vehicle dynamics in [21]. In the paper, we consider a typical two-degree-of-freedom yaw plane vehicle model for simplification (see Fig. 2), a detailed description of which can be found in [22]. In the SSCS, the controller is designed to track the required yaw rate γ_{ref}, which can be derived by the driver's action. The vehicle dynamics can be described as follows:

$$\dot{\beta}(t) = \frac{F_{yf}(t) + F_{yr}(t)}{v_x M} - \gamma(t), \tag{1a}$$

$$\dot{\gamma}(t) = \frac{L_f F_{yf}(t) - L_r F_{yr}(t) + M_z(t)}{I_z}, \tag{1b}$$

where β represents the lateral slip angle; γ is the yaw rate of the vehicle body; v_x is the instantaneous vehicle speed; F_{yf} and F_{yr} represent the resultant lateral forces of the front and rear tires, respectively; M is the vehicle mass; L_f and L_r are the distances from the center of mass to the front and rear axles, respectively; M_z is the additional yaw moment from the difference between longitudinal tire forces. In the above equations, F_{yf} and F_{yr} are determined by the tire characteristics, road conditions, tire sideslip angle, front-wheel steering angle, etc. Due to high complexity of the tires, the tire model to establish the lateral forces is typically simplified to an empirical formula through experimental data. For the steering stability control, we use a linear tire model: $F_y(t) \approx -2C_y\alpha(t)$, where α is front-wheel steering angle; C_y is the cornering stiffness of the tires. Then, the state-space equation is formulated as follows:

$$\dot{X} = AX + BU \tag{2a}$$

$$A = \begin{bmatrix} -\frac{2C_{y,f}+2C_{y,r}}{v_x M} & \frac{2C_{y,r}L_r - 2C_{y,f}L_f}{v_x^2 M} - 1 \\ \frac{2C_{y,r}L_r - 2C_{y,f}L_f}{I_z} & \frac{-2C_{y,r}L_r^2 - 2C_{y,f}L_f^2}{I_z v_x} \end{bmatrix} \tag{2b}$$

$$B = \begin{bmatrix} \frac{2C_{y,f}}{v_x M} & 0 \\ \frac{2C_{y,f}L_f}{I_z} & \frac{1}{I_z} \end{bmatrix} \tag{2c}$$

where the state of the system is defined as $X = [\beta, \gamma]^T$; the control input is $U = [\delta, M_z]^T$.

Then, a linear quadratic regulator (LQR) controller is designed with $U^{opt} = -KX + U_r$, where U_r brings the outputs to the desired point; K is the feedback gain matrix to minimize the cost function

$$J = \int_0^\infty [(X - X_{ref})^T Q (X - X_{ref}) + U^T R U] dt \tag{3}$$

and is derived by the solution of the Riccati equations:

$$PA + A^T P - PBR^{-1}B^T P^T + Q = 0, \tag{4a}$$

$$A^T - PBR^{-1}B^T\xi + QX_{ref} = 0. \tag{4b}$$

Then $K = R^{-1}B^T P$ and $U_r = R^{-1}\xi$. Here $X_{ref} = [\beta ref, \gamma_{ref}]^T$ is the reference of the system state; Q and R are the positive weighting matrices. Because during steering, low lateral slip angle indicates better stability and comfortability, we set the desired lateral slip angle as $\beta_{ref} = 0$.

III. Attack taxonomy and Residual-based detection methodology

A. Modeling of the cyber-physical attacks

To model the attack taxonomy, we assume that the attacker can illegally get access to the CAN bus, modify the sensor measurements, and hijack the stability control system. For convenient expression, as introduced in [23], we consider a general control architecture, which has three components: the plant (physical phenomena of interest including the actuators), sensors (or observers) to obtain the system outputs, denoted as y, and control commands u. Let $\widetilde{u}$ and $\widetilde{y}$ represents the signal under attacks. In the case of integrity attacks, a malicious attacker can either physically or remotely gain access to the signals and generate false data. Then, as shown in Fig. 3, three attack scenarios are considered in the work: (i) integrity attacks on y, expressed as $\widetilde{y} \neq y$ and $u_a = u_{nom}$, where u_{nom} represents the results of the normal controller; (ii) integrity attacks on u, as $\widetilde{y} = y$, $u = u_{nom}$, and $u_a \neq u$; besides the above two cyber-attacks on signals in CAN bus, we consider (iii) integrity attacks on the controller, as $\widetilde{y} = y$, $u = \widetilde{u}_c \neq u_{nom}$, and $u_a = u$. According to the SSCS described above, the most dominating feedback signals that might be attacked include the system states β and γ. Specific definitions of these attacks are summarized in Table. I.

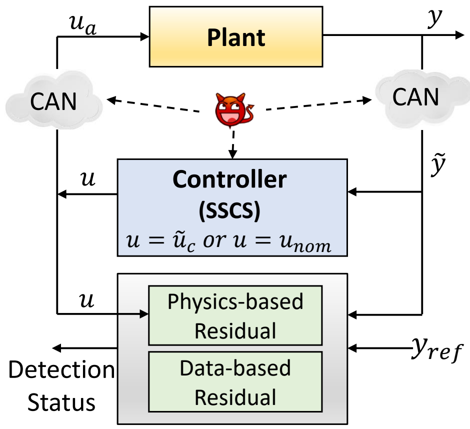

Fig. 3. Diagram of the residual-based detection method.

TABLE I
DEFINITIONS OF CYBER-PHYSICAL ATTACKS

Attack	Case	Definition
(i)	1-4	$\widetilde{\gamma} = \kappa_s \gamma,\ \kappa_s \in \{0, 0.5, 1.5, 2\}$
	5-8	$\widetilde{\gamma} = \gamma + p,\ p \in \{-0.05, -0.02, 0.02, 0.05\}$
	9-12	$\widetilde{\beta} = \kappa'_s \beta,\ \kappa'_s \in \{0, 0.5, 1.5, 2\}$
	13-16	$\widetilde{\beta} = \beta + p',\ p' \in \{-0.005, -0.002, 0.002, 0.005\}$
(ii)	17-21	$\widetilde{u}_a = \kappa_u u,\ \kappa_u \in \{0.4, 0.6, 0.8, 1.2, 1.4\}$
(iii)	22-26	$\widetilde{u}_c = \kappa_c u_{nom},\ \kappa_c \in \{0.4, 0.6, 0.8, 1.2, 1.4\}$

B. Residual-based detection methodology

Based on the system dynamics and operating data, we propose a residual-based cyber-attack detection methodology for the SSCS. Firstly, by using the system model, a state observer is developed to estimate the states of the system, as follows:

$$\dot{\hat{X}} = A\hat{X} + L(X - \hat{X}) - BU \tag{5}$$

where L is the observer gain that ensures observer stability. Consider that the output estimation error at time instance k is $r_k^{phy} = g(X_i, \hat{X}_i)$, $i = k - l + 1, \ldots, k$ (l represents the window size of detection, and the superscript phy denotes physics-based), which is the function of the estimation and measurement values. Then, r_k^{phy} can be one of the residuals to identify the cyber-physical threats. Based on the sequences $\hat{X}_i$ and X_i over the window size of detection, the residual signal at time instance k can be defined as

$$r_k^{phy} = \sum_{i=k-l+1}^{k} \|X_i - \hat{X}_i\|/l. \tag{6}$$

Typically, for physics-based detection, the residual $r_k > \tau^{phy}$ (here τ^{phy} is a threshold) is considered as a proxy for the presence of attacks, and the detector will trigger an alarm.

Notice that the control input U is directly used in the state estimation, the detection criteria r_k^{phy} cannot identify cyber-attacks on the controller. Hence, besides the physics-based t_k^{phy}, we develop a data-based residual by using a deep network - long short-term memory (LSTM), which has been used widely in many domain becasue of its superior ability in capturing the dynamics of a system [24]–[27]. In the deep LSTM, the sequential observations for states prediction are chosen as $x = [\gamma, \beta, v_x, \gamma_{ref}]^T$ in normal driving conditions. Then, in off-line training process, the predicted system states are calculated based on their past N_d values, vehicle speed v_x,

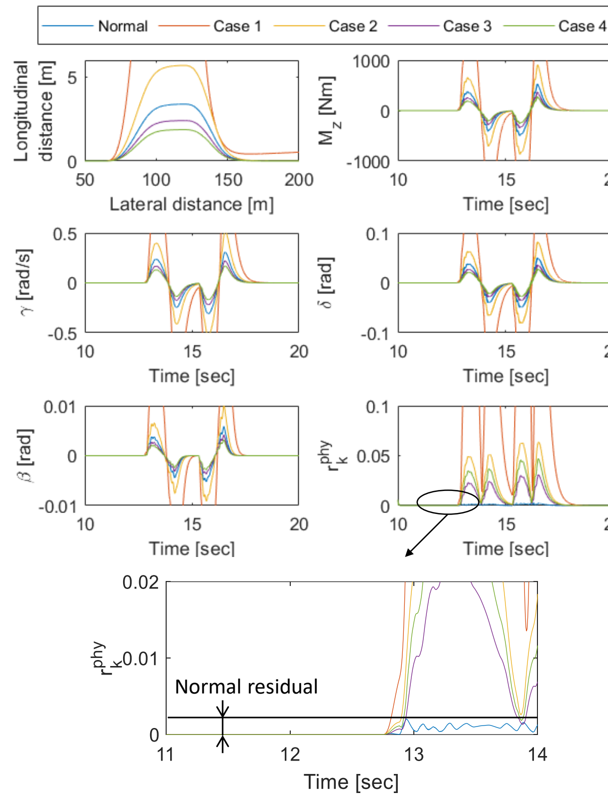

Fig. 4. Simulation results of Cases 1-4 (Physics-based residual r_k^{phy}).

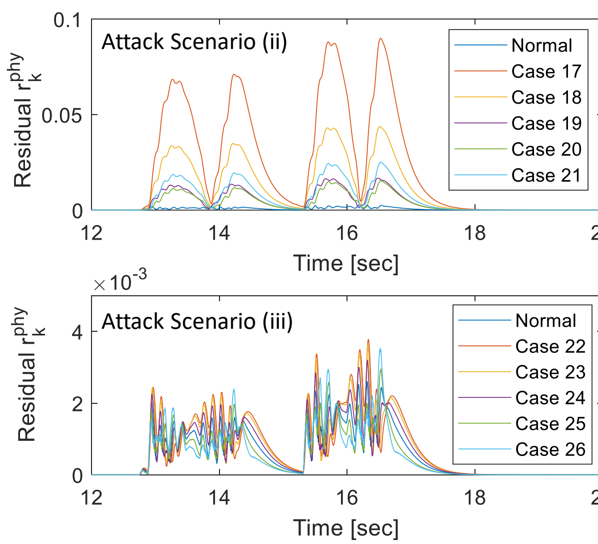

Fig. 5. Results of attack scenarios (ii) and (iii) (Physics-based residual r_k^{phy}).

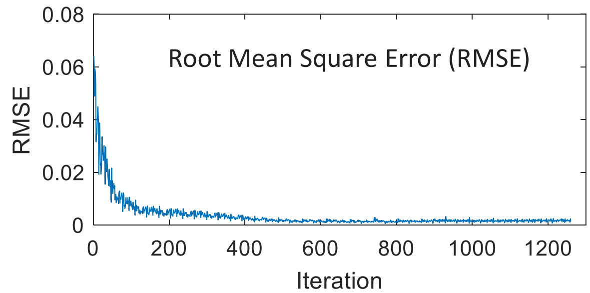

Fig. 6. Training accuracy of the LSTM, where $RMSE = \|Y - \hat{Y}\|/2$.

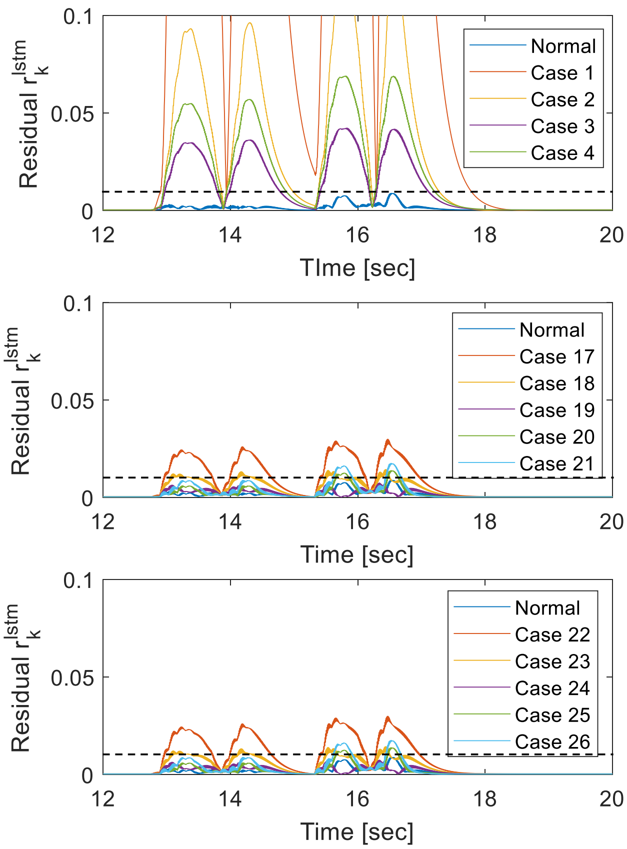

Fig. 7. Results of attack scenarios (i), (ii), and (iii) (Data-based residual r_k^{lstm}), where the black dot line represents the threshold τ^{lstm}.

and the reference γ_{ref}. The moving horizon split-window N_d is used to generate the raw data matrix, which is formulated as $X_{input} = [x(k - N_d + 1), x(k - N_d + 2), ..., X(k)]$. Here k represents the current time instance. The output of the network corresponding to the input X_{input} is defined as $Y_{output} = [\gamma(k+1), \beta(k+1)]^T$. A standard cost function for training the network is to minimize the empirical loss of its model predictions, as follows:

$$\arg \min_{W} Loss(\hat{Y}_{output}, Y_{output}) + \lambda R(W), \tag{7}$$

where R measures the complexity of a model, W is the weight to be optimized, and λ is a trade-off hyper-parameter.

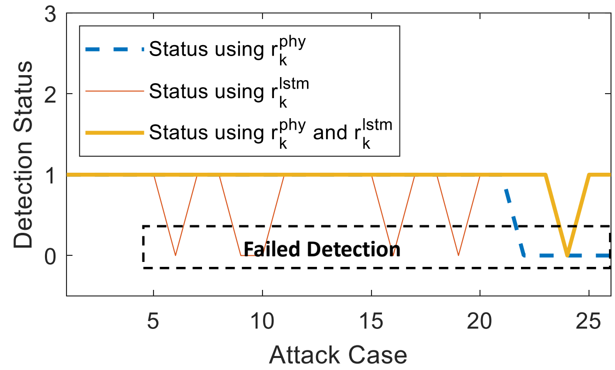

Fig. 8. Comprehensive detection results of the defined attacks.

After obtaining the well-trained network, based on the given inputs X_{input} in real time, the predicted value is defined as $\hat{X}_i^{lstm} = \hat{Y}_{output}$, based on which the data-based residual is calculated by

$$r_k^{lstm} = \sum_{i=k-l+1}^{k} \|X_i - \hat{X}_i^{lstm}\|/l. \tag{8}$$

Then, $r_k^{lstm} > \tau^{lstm}$ (here τ^{lstm} is a threshold) is also used to detect the cyber-physical attacks.

Finally, both the two residuals are adopted to identify the cyber-physical threats, as follows:

$$status = 1 \text{ if } r_k^{phy} > \tau^{phy} \text{ or } r_k^{lstm} > \tau^{lstm}, \tag{9}$$

and if not, $status = 0$, where $status$ represents the status of system security; 0 and 1 represent the normal and abnormal condition, respectively. In addition to detection status, by using the two physics-based and data-based residuals, one can also make a preliminary threat localization: attacks on the controller or others, which can benefit to threat diagnosis and mitigation. For example, if $status = 1$ and $r_k^{phy} < \tau^{phy}$, then the possibility of the presence of controller attacks should be considered, which can be further identified with more criteria, e.g., performance indexes like tracking accuracy, torque ripple, and frequency of steering to evaluate the performance degradation caused by control failure.

IV. Performance Evaluation

In this section, we present the results of the specified attack cases in Table I under a double-lance change driving condition. The entering scenario is on straight running with an initial speed of 80 km/h. The road surface is assumed to be flat and smooth, with a friction coefficient of 0.7.

For the sake of precise observation of the effectiveness of the obtained residuals, we present specific results of integrity attacks on the system state, control inputs, and controller, respectively. Figs. 4-5 show the results of residual r_k^{phy}, from which we can observe that sensor data integrity attacks may have a significant influence on the system performance, even cause instability. The results indicate that r_k^{phy} can adequately reflect the impact of cyber-attacks; thus, in the case of attacks on system states and control inputs, the threat can be identified timely by physics-based residual. However, for the cyberattacks on the controller, such as attack scenario (iii), this kind of observer-based detection would not be sufficient for threat identification. This is because, in the design process of the observer, the control inputs are directly used in the system model (see Equation (5)), which is also the actual control instruction to the actuator. Therefore, the defined physics-based residual cannot reflect the performance degradation of the controller, as shown in Fig. 5.

To illustrate the effectiveness of the data-based residual, with a double-lance change driving condition, we obtain 3996 normal observations. Among the total training data, we randomly choose 80% of the data to train the network and the rest 20% data to validate the model accuracy. The training process is given in Fig. 6, which illustrates the accuracy of the trained network. Then, we present the results of r_k^{lstm} in Cases 1-4, Cases 17-21, and Cases 22-26, as shown in Fig. 7. From these results, we can see that the residual r_k^{lstm} is useful in identifying those cyber-physical attacks causing visible effects, especially for controller attacks.

It worth noting that, compared to the developed model-based residual, the obtained r_k^{lstm} through LSTM shows less ability to distinguish normal conditions and control inputs attacks. Even though the data-based residual can be used to identify the potential threats, it highly depends on the number of training data. Thus, a certain driving condition that differs greatly from the training set may lead to misjudgment when only the data-based residual is used. Therefore, in real-world applications, it is necessary to combine both the physics-based and data-based residuals. Finally, the comprehensive detection results, $status$, are presented in Fig. 8, wherein, thresholds τ^{phy} and τ^{lstm} are set to 0.0025 and 0.01, respectively. From the results, we can see that the combined residuals can improve the accuracy of threat detection. Besides, due to the fixed detection thresholds, τ^{phy} and τ^{lstm}, the detection accuracy can not be up to 100%. This is because, in real-world applications, the driving conditions are changing, and then the prediction error (in normal situations) of the designed estimator, either observer-based or data-based, is time-varying. Moreover, considering the uncertainty of driving road and system modeling, an adaptive threshold should be introduced, which will be one of the future works of cyber-physical attack detection and diagnosis.

V. Conclusions and Future works

In this paper, a steering stability control system is designed, based on which a residual-based anomaly detector is developed to identify cyber-physical attacks. Besides the control-theoretical observer, we also use a deep-learning network to predict the system output, with which a data-based residual is calculated to be an auxiliary identification index. Simulation results have shown the effectiveness of the proposed residual-based detection methodology and validated the necessity to combine the physics-based and data-based results.

Acknowledgment

This work was supported in part by the National Science Foundation Program (Award number ECCS-1946057).

References

[1] F. Li, J. Wang, and Z. Liu, "Motor torque based vehicle stability control for four-wheel-drive electric vehicle," in *2009 IEEE Vehicle Power and Propulsion Conference*. IEEE, 2009, pp. 1596–1601.

[2] F. Jia, Z. Liu, H. Zhou, and T. Teng, "A robust control invariant set approach to yaw stability of four-wheel drive electric vehicle," *IFAC-PapersOnLine*, vol. 51, no. 31, pp. 325–330, 2018.

[3] B. Ren, H. Chen, H. Zhao, and L. Yuan, "Mpc-based yaw stability control in in-wheel-motored ev via active front steering and motor torque distribution," *Mechatronics*, vol. 38, pp. 103–114, 2016.

[4] B. Huang, S. Wu, S. Huang, and X. Fu, "Lateral stability control of four-wheel independent drive electric vehicles based on model predictive control," *Mathematical Problems in Engineering*, vol. 2018, 2018.

[5] C. Hodge, K. Hauck, S. Gupta, and J. C. Bennett, "Vehicle cybersecurity threats and mitigation approaches," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2019.

[6] G. K. Rajbahadur, A. J. Malton, A. Walenstein, and A. E. Hassan, "A survey of anomaly detection for connected vehicle cybersecurity and safety," in *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018, pp. 421–426.

[7] M. Levi, Y. Allouche, and A. Kontorovich, "Advanced analytics for connected car cybersecurity," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE, 2018, pp. 1–7.

[8] X. Shao, C. Dong, and L. Dong, "Research on detection and evaluation technology of cybersecurity in intelligent and connected vehicle," in *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*. IEEE, 2019, pp. 413–416.

[9] C. Watney and C. Draffin, "Addressing new challenges in automotive cybersecurity," 2017.

[10] Y. Xun, J. Liu, N. Kato, Y. Fang, and Y. Zhang, "Automobile driver fingerprinting: A new machine learning based authentication scheme," *IEEE Transactions on Industrial Informatics*, 2019.

[11] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham *et al.*, "Experimental security analysis of a modern automobile," in *2010 IEEE Symposium on Security and Privacy*. IEEE, 2010, pp. 447–462.

[12] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, T. Kohno *et al.*, "Comprehensive experimental analyses of automotive attack surfaces." in *USENIX Security Symposium*, vol. 4. San Francisco, 2011.

[13] A. Greenberg. (2015, July 21), "Hackers remotely kill a jeep on the highway—with me in it. wired. [online]." *Available: https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/.*

[14] "Cyber cttacks in connected cars: what tesla did differently to win. https://blog.appknox.com/cyber-attacks-in-connected-cars/."

[15] A. Weimerskirch and R. Gaynier, "An overview of automotive cybersecurity: Challenges and solution approaches." in *TrustED@ CCS*, 2015, p. 53.

[16] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on scada systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1396–1407, 2014.

[17] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding sensor outputs for injection attacks detection," in *53rd IEEE Conference on Decision and Control, IEEE*, December 2014, pp. 5776–5781.

[18] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 580–591, 2014.

[19] M. J. Desforges, P. J. Jacob, and J. E. Cooper, "Applications of probability density estimation to the detection of abnormal conditions in engineering," in *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 212, no. 8, 1998, pp. 687–703.

[20] F. Pasqualetti, F. Dörfler, and F. Bullo, "Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 2195–2201.

[21] L. Guzzella and A. Sciarretta, "Vehicle propulsion systems, introduction to modeling and optimization. 2005."

[22] R. Rajamani, *Vehicle dynamics and control*. Springer Science & Business Media, 2011.
[23] J. Giraldo, D. Urbina, A. Cardenas, J. Valente, M. Faisal, N. O. T. J. Ruths, H. S. Sandberg, and R. Candell, "A survey of physics-based attack detection in cyber-physical systems," *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, p. 76, 2018.
[24] S. Merity, N. S. Keskar, and R. Socher, "Regularizing and optimizing lstm language models," *arXiv preprint arXiv:1708.02182*, 2017.
[25] Z. Zhao, W. Chen, X. Wu, P. C. Chen, and J. Liu, "Lstm network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.
[26] F. J. Ordóñez and D. Roggen, "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, 2016.
[27] L. Gao, Z. Guo, H. Zhang, X. Xu, and H. T. Shen, "Video captioning with attention-based lstm and semantic consistency," *IEEE Transactions on Multimedia*, vol. 19, no. 9, pp. 2045–2055, 2017.