A Two-Stage Extended Kalman Filter-Based Approach Against FDI Cyber-Attack in Intelligent and Connected Vehicles

Bin Sun[1][a], Shichun Yang[1][b], Yu Wang[1], Jiayi Lu[1] and Yaoguang Cao[1,2][c]
[1]*School of Transportation Science and Engineering, Beihang University, Beijing, China*
[2]*State Key Lab of Intelligent Transportation System, Beihang University, Beijing, China*
{*sunbin2022, yangshichun, lujiayi, caoyaoguang*}*@buaa.edu.cn, wangyu_200105@163.com*

Keywords: Cybersecurity, False Data Injection, Two-Stage Extended Kalman Filter, Intelligent and Connected Vehicles.

Abstract: With the widespread integration of artificial intelligence and telecommunication technologies in vehicles, the challenge of cybersecurity in Intelligent and Connected Vehicles (ICVs) has gained significant attention. A typical and high-risk cyber-attack technique involves False Data Injection (FDI) into sensors through the network, resulting in deviations in subsequent planning and control algorithm outcomes. Existing approaches suffer from limited robustness, being suitable only for simple models or requiring extensive data for the training model, which limits their practicality. Therefore, this paper proposes a method based on a Two-stage Extended Kalman Filter (TSEKF), which not only detects cyber-attacks but also restores the vehicle's true motion state, thereby enhancing the robustness of vehicle ego state perception. The experimental results demonstrate that the proposed method exhibits strong performance across various motion scenarios, offering an effective solution for the safe operation of ICVs.

1 INTRODUCTION

Intelligent and Connected Vehicles (ICVs) are a future trend, integrating advanced technologies and improving travel efficiency, but also introducing security risks due to external information exchange. Protecting vehicle sensors from cyber-attacks is essential for safe operation(Mwanje et al., 2024).Common cyber attacks include FDI, DRA, and DoS. FDI is the most typical, injecting false data into vehicle sensors, leading to inaccurate algorithms and potential safety risks (Ju et al., 2022). This paper focuses on detecting FDI attacks.

Cyber-attack detection for ICVs can be classified into three main categories. The first category is model-based attack detection methods, which design an observer based on the vehicle's dynamic model, assuming that modeling and measurement uncertainties have upper bounds. An attack is detected when the measurement innovation exceeds a threshold. The work in (He et al., 2021) focuses on sensor attack detection using a saturation-like observer. The methods proposed in (Dutta et al., 2018), and (Abdollahi Biron et al., 2016) utilize a sliding mode observer for attack detection, characterized by a simple design process and some robustness to modeling uncertainties. However, observer-based methods are primarily suited for deterministic system models, offering simplicity but limited performance under imperfect communication.The second category involves attack detection based on machine learning techniques. These methods utilize machine learning to achieve attack detection. In (Ju et al., 2020), Support Vector Machines (SVM) is employed to detect speed and position sensor attacks during vehicle following. Hsiao-Chung Lin et al. developed an Atta detection model using a pre-trained VGG16 deep learning classifier to learn attack behavior features and classify threats (Lin et al., 2022). Several scholars (Hossain et al., 2020), (Lokman et al., 2019), (Han et al., 2018), (Javed et al., 2021) have used foundational classifiers to detect anomalies in CAN messages, such as Decision Trees, Logistic Regression and Support Vector Classifiers. Wei Lo et al. implemented a hybrid network combining Convolutional Neural Networks (CNN) and Long Short-Term Memory networks to automatically extract spatial and temporal features from vehicular network traffic for attack detection (Lo et al., 2022). However, machine learning-based detection faces two major challenges: the inability to

[a] https://orcid.org/0009-0008-4998-0974
[b] https://orcid.org/0000-0003-3426-7988
[c] https://orcid.org/0000-0002-6107-2425

Sun, B., Yang, S., Wang, Y., Lu, J. and Cao, Y.
A Two-Stage Extended Kalman Filter-Based Approach Against FDI Cyber-Attack in Intelligent and Connected Vehicles.
DOI: 10.5220/0013137900003941
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 11th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2025)*, pages 301-308
ISBN: 978-989-758-745-0; ISSN: 2184-495X
Proceedings Copyright © 2025 by SCITEPRESS – Science and Technology Publications, Lda.

quantify cyber-attacks and potential significant performance degradation in the presence of measurement noise. Additionally, effective detection and mitigation of cyber-attacks require substantial amounts of real data, which involves considerable effort in data collection and labeling. The last category is hybrid methods that integrate model-based approaches with artificial intelligence. The literature (Wang et al., 2020) presents a method that cascades CNN with a χ^2 detector, where CNN first detects and removes anomalous sensor data, followed by the χ^2 detector to identify undetected anomalies. Although this approach shows performance improvements, the detection results rely heavily on CNN training. Guo et al. designed a machine learning method that combines battery dynamics and vehicle kinematic models to detect cyber-attacks on electric vehicles in various driving scenarios (Guo et al., 2021). Linxi Zhang et al. merge traditional rule-based intrusion detection techniques with emerging machine learning methods, striking a balance between detection accuracy and efficiency (Zhang and Ma, 2022). Although each of the three aforementioned methods has its strengths and weaknesses in attack detection, they all primarily emphasize anomaly detection in the data, neglecting the estimation of the vehicle's true state.

The Kalman filter helps estimate vehicle states in noisy settings. However, it struggles with inaccurate models or biased sensor data. The Two-stage Kalman Filter (TSKF) was initially proposed to addresses random biases (Keller and Darouach, 1997). This paper introduces a Two-stage Extended Kalman Filter (TSEKF) for nonlinear systems to detect and mitigate cyber-attacks in ICVs. It identifies cyber-attacks and estimates vehicle states via a dual-stage process. The performance of the method is validated for different vehicle motions.

The remainder of the paper is organized as follows. Section II introduces the vehicle dynamics modeling. Section III discusses the modeling of cyber-attacks and the detection methods using the TSEKF algorithm. Section IV describes the experiment results of the TSEKF in various motion states. Finally, Section VI summarizes the entire paper.

2 VEHICLE DYNAMICS

In the extened Kalman Filter prediction step, a twin-track model is utilized to model the lateral and longitudinal vehicle dynamics, thereby enhancing the precision of vehicle state estimation (Henning and Sawodny, 2016). The model consists of a nonlinear dynamic state equation combined with a linear output

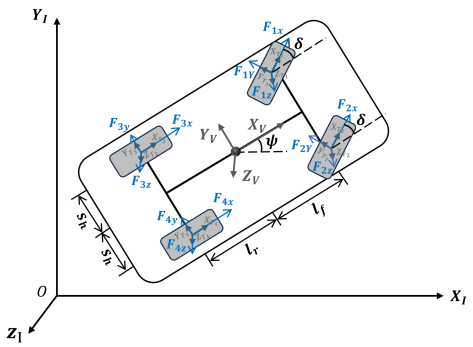


Figure 1: Double-track model of vehicle lateral dynamic modal.

equation:

$$\begin{aligned} \dot{x} &= r(x,u), \quad x(0) = x_0 \\ y &= Cx + F_y(x)f \end{aligned} \quad . \tag{1}$$

The state vector of the dynamic equations is:

$$x = [\ v_{x,V} \quad v_{y,V} \quad \dot{\psi}_V \quad a_{x,I} \quad a_{y,I}\ ]^T, \tag{2}$$

with longitudinal velocity $v_{y,V}$, lateral velocity $v_{x,V}$, yaw rate $\dot{\psi}_V$, longitudinal acceleration $a_{x,I}$ and lateral acceleration $a_{y,I}$.The subscript for the state variables indicates the reference coordinate system for measurements: I for the inertial coordinate system, V for the vehicle coordinate system, and T_i for the tire coordinate system. The tire numbering is as follows: 1 for the left front, 2 for the right front, 3 for the left rear, and 4 for the right rear. The coordinate system is shown in 1.

The control input vector of the model of the dynamic equations is:

$$\begin{aligned} u = \big[\delta \ \ T_{b1,T_1} \ \ T_{b2,T_2} \ \ T_{b3,T_3} \ \ T_{b4,T_4} \\ T_{d1,T_1} \ \ T_{d2,T_2} \ \ T_{d3,T_3} \ \ T_{d4,T_4}\big] \end{aligned}, \tag{3}$$

with the front steering angle δ, the barking moment T_{bi,T_i} and driving moment T_{di,T_i} of four wheels.

2.1 Non-Linear Tier Model

The torque balance equations for the four wheels are as follows:

$$\theta_{tw} \cdot \dot{\omega}_{i,T_i} = -r_{(tire)} \cdot F_{ix,V} - T_{bi,T_i} + T_{di,T_i}, \tag{4}$$

where θ_{tw} is the moment of inertia of an individual tire around its rotational axis, $r_{(tire)}$ is the tire radius, and $F_{xi,V}$ is the longitudinal ground force acting on the tire in the vehicle driving direction.

The tire model used in this study is a simplified version of the Magic Formula Tire (MFT) model (Dieter et al., 2018), based on the MFT 5.2 model proposed by Pacejka (Pacejka, 2012), and includes a

complete set of parameters. The fundamental equations are as follows:

$$\begin{aligned} F_{i,V} &= F_{i(\max)} \sin(C \arctan(B \tfrac{\|s_i\|}{\mu})) \begin{bmatrix} k_x & 0 \\ 0 & 1 \end{bmatrix} \tfrac{s_i}{\|s_i\|} \\ F_{i(\max)} &= D\mu F_{iz}(1 + k_{F_z} \tfrac{F_{Z0} - F_{iz}}{F_{Z0}}). \end{aligned} \tag{5}$$

The parameters B, C, D, k_x and k_{F_z} represent the initial slope, saturation shape, peak tire force, longitudinal scaling factor, and tire load coefficient, respectively. These parameters describe the characteristic relationship curve between slip and tire force, but do not have physical units. The maximum transferable force $F_{i(\max)}$ increases with tire load F_{iz}, which can be calculated as in reference (Dieter et al., 2018). However, as the load continues to increase, the rate of increase of the maximum force gradually decreases, a relationship determined by the coefficient k_{F_z} (Best, 2014).

The inputs for the tire model (5) are the longitudinal slip k_i and the lateral slip angle α_i. For each tire, these are represented as:

$$s_i = \begin{bmatrix} k_i \\ \alpha_i \end{bmatrix}. \tag{6}$$

The longitudinal slip k_i is calculated as:

$$k_i = \frac{r_{\text{tire}}\omega_{i,T_i} - v_{xi,T_i}}{v_{xi,T_i}}. \tag{7}$$

The side slip angle α_i is defined by

$$\alpha_i = \begin{cases} \delta - \arctan\left(\frac{v_{y,V} + l_f\dot{\psi}_V}{v_{x,V} - s_h\dot{\psi}_V}\right), i = 1 \\ \delta - \arctan\left(\frac{v_{y,V} + l_f\dot{\psi}_V}{v_{x,V} + s_h\dot{\psi}_V}\right), i = 2 \\ -\arctan\left(\frac{v_{y,V} - l_r\dot{\psi}_V}{v_{x,V} - s_h\dot{\psi}_V}\right), i = 3 \\ -\arctan\left(\frac{v_{y,V} - l_r\dot{\psi}_V}{v_{x,V} + s_h\dot{\psi}_V}\right), i = 4 \end{cases}. \tag{8}$$

As shown in Figure 1, l_f and l_r represent the distances from the vehicle's center of gravity to the front and rear axles, respectively, s_h is half of the vehicle's width.

2.2 Vehicle Dynamic Model

The general equation for the horizontal motion of a vehicle can be expressed as follows:

$$\begin{aligned} \dot{v}_{x,V} &= \tfrac{1}{m}\sum_{i=1}^{4} F_{ix,V} + \dot{\psi}_V v_{y,V} \\ \dot{v}_{y,V} &= \tfrac{1}{m}\sum_{i=1}^{4} F_{iy,V} - \dot{\psi}_V v_{x,V} \\ \ddot{\psi}_V &= \tfrac{1}{\theta_{zz}}\sum_{i=1}^{4} M_{iz,V} \ . \end{aligned} \tag{9}$$

All kinetic quantities, tire forces, and moments are illustrated in Figure 1. In equation (9), m represents the vehicle's mass, θ_{zz} denotes the moment of inertia of the entire vehicle about the z axis, while $F_{ix,V}$ and $F_{iy,V}$ are the horizontal components of the respective tire forces in the x or y direction. $M_{iz,V}$ represents the moments generated by these forces around the vehicle's center of gravity (COG).

It is important to note that the additional translational acceleration caused by the rotating reference frame has been accounted for in equation (10), but the accelerations in inertial coordinates are calculated as:

$$\begin{aligned} a_{x,I} &= \tfrac{1}{m}\sum_{i=1}^{4} F_{ix,V} \\ a_{y,I} &= \tfrac{1}{m}\sum_{i=1}^{4} F_{iy,V} \end{aligned}, \tag{10}$$

The complete model is derived by combining equations (9) with a nonlinear tire model (5),

$$f(x,u) = \begin{bmatrix} \frac{1}{m}\sum_{i=1}^{4} F_{ix,V} + \dot{\psi}_V v_{y,V} \\ \frac{1}{m}\sum_{i=1}^{4} F_{iy,V} - \dot{\psi}_V v_{x,V} \\ \frac{1}{\theta_{zz}}\sum_{i=1}^{4} M_{iz,V} \\ \frac{1}{\Delta t}\left(\frac{1}{m}\sum_{i=1}^{4} F_{ix,V} - a_{x(\Delta t),I}\right) \\ \frac{1}{\Delta t}\left(\frac{1}{m}\sum_{i=1}^{4} F_{iy,V} - a_{y(\Delta t),I}\right) \end{bmatrix}. \tag{11}$$

The parameters used in the vehicle dynamic model and MFT tire model proposed in this study are listed in Table 1. These parameters were obtained from the open database of the simulation software.

3 CYBER-ATTACK MODELING AND TSEKF-BASED APPROACH

The nonlinear vehicle dynamics model established in Chapter 2 can be discretized using the 4th-order Runge-Kutta method:

$$\begin{aligned} \mathbf{x}_k &= f_d(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{y}_k &= C\mathbf{x}_k + \mathbf{v}_k \end{aligned}, \tag{12}$$

with the state vector $\mathbf{x}_k \in \mathbb{R}^n$, control input $\mathbf{u}_k \in \mathbb{R}^m$, and observation vector $\mathbf{y}_k \in \mathbb{R}^p$. Here, $\mathbf{w}_k \in \mathbb{R}^p$, $\mathbf{w}_k \sim \mathcal{N}(0, Q)$ and $\mathbf{v}_k \in \mathbb{R}^p$, $\mathbf{v}_k \sim \mathcal{N}(0, R)$ denote process and measurement noise, respectively, both of which are assumed to follow a normal distribution with a mean of zero. The covariance matrices Q and R represent the process noise and measurement noise covariance, respectively. Sensor noise captures the random uncertainty present during the measurement process. Process noise reflects external influences, such

Table 1: Modelling Parameters.

Symbol	Description	Value	Unit
μ	Friction coefficient of road surface	0.8	-
F_{z0}	Nominal tire load	5150	N
k_x	Longitudinal scaling factor	1.1743	-
k_{Fz}	Degenerative tire load factor	0.1342	-
B	Initial slope parameter	10.4962	-
C	Shape factor for saturation region	1.5402	-
D	Influences maximum tire force peak	1.1006	-
m	Vehicle mass	2100	kg
θ_{zz}	Moment of inertia for Vehicle	2549	$kg \cdot m^2$
θ_{tw}	Moment of inertia for tires	2.1	$kg \cdot m^2$
l_f	Distance cog to front	1.27	m
l_r	Distance cog to rear	1.37	m
r_{tire}	tire radius	0.34	m
s_h	Half track-width	0.81	m
g	gravitational acceleration constant	9.81	m/s^2

as wind force and road surface irregularities. Additionally, internal non-parametric input uncertainties, such as actuator delays, are neglected; we assume that the response of actuators, like steering, is rapid. Similar uncertainties for other parameters are also disregarded.

In equation (13), the state observation matrix C is chosen as follows:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (13)$$

3.1 Cyberthreat Modelling

FDI is one of the most prevalent forms of cyber-attacks, where attackers, unaware of system parameters or previous event data, directly introduce erroneous information into the original data. This manipulation can result in significant discrepancies in speed, acceleration, and other critical parameters, leading downstream decision-making processes to rely on incorrect motion states for path planning. Consequently, the safety constraints derived from these decisions are based on flawed states, rendering them ineffective and posing substantial safety risks of ICVs. FDI attack can be modeled as follows (Tan et al., 2017):

$$y_k = \begin{cases} Cx_k + v_k & k < \tau \\ Cx_k + F_y b_k + v_k & k \geq \tau \end{cases}. \quad (14)$$

Here, the subscript k indicates the value at a given discrete time k; b represents the bias/attack vector added to the measurements; F denotes the observation matrix of the attack vector, illustrating how b influences the system's observations; and τ is the moment when the attack is activated.

Thus, considering the cyber-attack, the dynamics described by equation (13) can be updated to

$$\begin{aligned} \mathbf{x}_k &= f_d(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{y}_k &= C\mathbf{x}_k + F \cdot \mathbf{b}_k + \mathbf{v}_k \end{aligned}, \quad (15)$$

with cyber threat $\mathbf{b}_k \in \mathbb{R}^{nb}$.

We assume that the cyber-attack is stationary during the attack duration and can be modeled as a Gaussian random process

$$b_{k+1} = b_k + w_{f,k}, \quad (16)$$

where $\mathbf{w}_{b,k} \in \mathbb{R}^{n_b}$, $\mathbf{w}_{b,k} \sim \mathcal{N}(0, Q_b)$ represents the uncertainty of the attack magnitude.

For cyber-attacks, this paper considers FDI in the form of erroneous data in longitudinal velocity, lateral velocity, yaw rate, longitudinal acceleration, and lateral acceleration, specifically:

$$b = \begin{bmatrix} b_{v_x} & b_{v_y} & b_{\psi} & b_{a_x} & b_{a_y} \end{bmatrix}^T. \quad (17)$$

The output matrix of the attack is related to the state variables, specifically as follows:

$$F_y(x) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (18)$$

The measurement output when $b = 0$ is:

$$y = \begin{bmatrix} v_{x,V} & v_{y,V} & \dot{\psi}_V & a_{x(\Delta t),I} & a_{y(\Delta t),I} \end{bmatrix}^T. \quad (19)$$

3.2 Two-Stage Extended Kalman Filter

By dividing state estimation and attack estimation into two stages, the TSEKF effectively reduces the computational complexity associated with high-dimensional parameters. The first stage focuses on estimating the true state of the vehicle's motion, while the second stage concentrates on estimating the magnitude of the cyber-attack. Consequently, the first stage can be redesigned based on different vehicle model parameters, and the second stage can be optimized to reduce false positive rates and shorten attack detection times. These two stages operate relatively independently, with no mutual interference, making

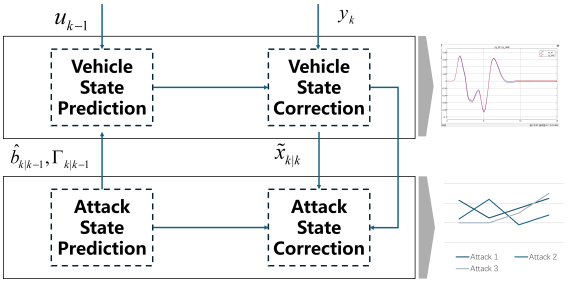


Figure 2: The proposed TSEKF algorithm.

the algorithm particularly well-suited for application in the field of intelligent and connected vehicles.

The overall framework of the algorithm is illustrated in Figure 2. In the first stage of vehicle motion state prediction, the inputs $\mathbf{u}_k$, the attack prior estimate $\hat{b}_{k|k-1}$, and state transformation $\Gamma_{k|k-1}$ are used to predict the state. Then, the measurement values $\mathbf{y}_k$ are utilized to compute the corrected motion state estimate. This corrected state estimate $\hat{\mathbf{x}}_{k+1|k+1}$ is subsequently used in the second stage for detecting cyber-attacks, yielding a corrected attack estimate $\hat{b}_{k|k}$, that serves as the prior estimate for the next iteration. The prediction of cyber-attack detection $\hat{b}_{k|k-1}$ leverages the prior estimate of the attack and determines the next state transition $\Gamma_{k|k-1}$.

The subscript $k|k-1$ indicates that at time step k, measurements up to time step $k-1$ are used. Utilizing the nonlinear motion equations (13) and combining them with the current cyber-attack estimate, the prior state is predicted for the vehicle's motion state:

$$\hat{x}_{k|k-1} = f_d(\hat{x}_{k-1|k-1}, u_{k-1}, \widehat{b}_{k-1|k-1}). \tag{20}$$

The prior estimate of the cyber-attack is computed based on the assumed stationary characteristics of the cyber-attack:

$$\hat{b}_{k|k-1} = \hat{b}_{k-1|k-1}. \tag{21}$$

According to the principles of the Extended Kalman Filter (EKF), the approximate linearized state transition Jacobian matrix for the nonlinear system is obtained through numerical computation using finite difference methods:

$$A_k = \frac{\partial f_d}{\partial x}. \tag{22}$$

According to equation (23), the self-covariance of the vehicle motion state estimation error is calculated as follows:

$$P_{xx,k+1|k} = A_k P_{xx,k|k} A_k^T + Q. \tag{23}$$

The self-covariance of the estimation error for the cyber-attack, based on equation (22) and the uncertainty of the cyber-attack process, is given by the following result:

$$P_{bb,k+1|k} = P_{ff,k|k} + Q_f. \tag{24}$$

The state and attack cross-covariance matrix can be computed using the state transition matrix and the static assumption of the cyber-attack, resulting in:

$$P_{xb,k+1|k} = A_k P_{xb,k|k}. \tag{25}$$

To eliminate the cross-covariance from equation (26), the state transformation should be processed as follows:

$$\begin{aligned} \begin{bmatrix} \widetilde{x}_{k|k-1} \\ \tilde{b}_{k|k-1} \end{bmatrix} &= \begin{bmatrix} I_n & -\Gamma_{k|k-1} \\ 0 & I_b \end{bmatrix} \begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{b}_{k|k-1} \end{bmatrix} \\ &= \begin{bmatrix} \hat{x}_{k|k-1} - \Gamma_{k|k-1} \hat{f}_{k|k-1} \\ \tilde{b}_{k|k-1} \end{bmatrix} \end{aligned}. \tag{26}$$

By applying the state transformation, the two phases of the EKF can be effectively separated. The new system state $\tilde{\mathbf{x}}$ is a linear combination of the system state and the attack state, while the attack estimation state $\tilde{b} = \hat{b}$ remains unchanged. Here, I_n and I_b are appropriately dimensioned identity matrices. According to reference (May et al., 2023), it can be expressed in an iterative calculation as a form independent of the cross-covariance:

$$\Gamma_{k|k-1} = A_{k-1} \Gamma_{k-1|k-1} P_{bb,k-1|k-1} P_{bb,k|k-1}^{-1}. \tag{27}$$

When the measurement value is y_k obtained, the corrected kinematic state estimate is given by:

$$\tilde{x}_{k|k} = \tilde{x}_{k|k-1} + K_{\tilde{x},k}(y_k - F_{y,k}\hat{b}_{k|k-1} - C_{k+1}\hat{x}_{k+1|k}), \tag{28}$$

with Kalman gain for state estimate $K_{\tilde{x},k}$, and calculated by

$$K_{\tilde{x},k} = P_{\tilde{x},k|k-1} C_k^T S_{\tilde{x},k}^{-1}. \tag{29}$$

Similar to the classic EKF algorithm, in equation (31), $S_{\tilde{x},k}$ represents the covariance of the observation error

$$S_{\tilde{x},k} = C_k P_{\tilde{x},k|k-1} C_k^T + R. \tag{30}$$

The posterior estimate of the cyber-attack is given by:

$$\hat{b}_{k|k-1} = \hat{b}_{k|k-1} + K_{b,k}(v_k - F_{y,k}\hat{b}_{k|k-1}), \tag{31}$$

with

$$v_k = y_k - C_k(\tilde{x}_{k|k-1} + \Gamma_{k|k-1}\hat{b}_{k|k-1}). \tag{32}$$

Similarly, the Kalman gain for attack estimate $K_{b,k}$ is

$$K_{b,k} = P_{bb,k|k-1} S_{f,k+1}^{-1}, \tag{33}$$

where β_k and $S_{f,k+1}$ are calculated as follows:

$$\beta_k = F_{y,k} + C_k \gamma_{k|k-1} \tag{34}$$

$$S_{f,k+1} = \beta_{k+1} P_{ff,k+1|k} \beta_{k+1}^T + S_{x,y,k+1}. \tag{35}$$

Finally, the error covariance for the next time step is updated using the following calculation:

$$P_{\tilde{x}\tilde{x},k|k} = (I - K_{\tilde{x}\tilde{x},y,k} C_k) P_{\tilde{x}\tilde{x},k|k-1} \tag{36}$$

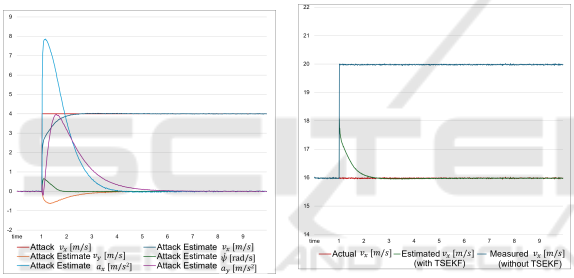
$$P_{bb,k|k} = (I - K_{f,k}\beta_k) P_{bb,k|k-1}. \tag{37}$$

4 EXPERIMENT RESULTS AND DISCUSSION

Various driving maneuvers were constructed in the simulation environment, including linear drive and circular drive, to simulate vehicle dynamics in single-directional movement, as well as coupled lateral and longitudinal movements. Different types of data injection attacks were designed to evaluate the performance of the proposed TSEFK algorithm. The covariance of process noise Q, measurement noise R, and attack noise Q_b was configured as diagonal matrices, with the value of diagonal elements set as 3×10^{-4}, 1×10^{-5} and 2×10^{-3}, respectively.

4.1 Linear Drive

Two driving maneuvers, with constant speed and acceleration, were designed to test the algorithm performance under cyber-attacks targeting longitudinal speed and longitudinal acceleration, respectively.

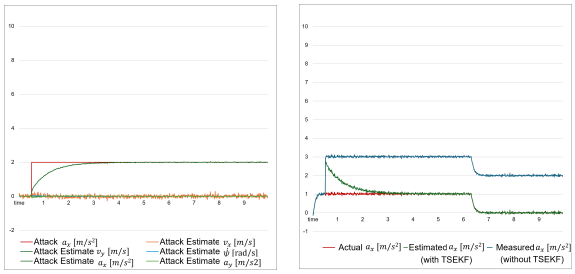


(a) Cyber-attack Estimate (b) Vehicle State

Figure 3: Linear Drive.

For the constant-speed scenario, the vehicle speed was set at 50 km/h ($\approx 16m/s$), , reflecting typical urban traffic. At $t = 1s$, an FDI attack was initiated on the longitudinal speed of the vehicle, with the attack magnitude $b_{v_x} = 4(m/s)$. Figure (3a) illustrates the algorithm estimation of cyber-attack magnitude. When the cyber-attack (solid red line) occurs, the proposed algorithm immediately detects it, with the estimated attack magnitude on longitudinal speed (solid green line) converging to the actual attack value within approximately 1 second. However, it should be noted that when the cyber-attack is initiated, there is significant fluctuation in the estimated attack magnitudes for both longitudinal and lateral acceleration. This is because acceleration is the derivative of speed; thus, any substantial deviation in speed measurements may not only reflect direct attacks on speed, but may also indicate an attack in acceleration. Over time, the estimated values for acceleration converge toward zero.

Figure (3b) presents the actual state, the measured state (without TSEFK), and the estimated state generated by the proposed algorithm (with TSEFK). Following a cyber-attack, the measured value displays an immediate deviation. Although the estimated value increases slightly, they rapidly reduce and converge to the actual state within approximately 1 second.



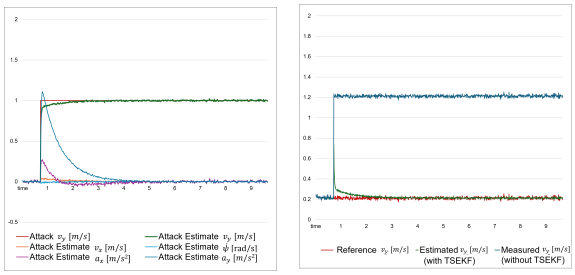
(a) Cyber-attack Estimate (b) Vehicle State

Figure 4: Linear Accelerated Drive.

For the acceleration scenario, the vehicle speed was also set at 50 km/h ($\approx 16m/s$), while an FDI attack was initiated in the longitudinal acceleration, with the attack magnitude $b_{a_x} = 2(m/s^2)$. The attack can be detected immediately and its magnitude stabilizes within 1 second. Unlike in Figure (3a), the estimated attack magnitudes for other variables remain close to zero. As explained previously, acceleration is the derivative of other variables and determines their rate of change. Therefore, when acceleration is attacked, it exhibits a deviation independently, unaffected by deviations in other variables. In Figure (4b), the vehicle estimated state return to the actual state after a slight increase.

4.2 Circular Drive

We used the circular motion to test the algorithm's evaluation of the vehicle's lateral motion state indicators, with the steering angle fixed at 0.1 rad and the vehicle speed set to the previously mentioned $50km/h$. At $t = 0.7s$, a FDI attack was initiated at the

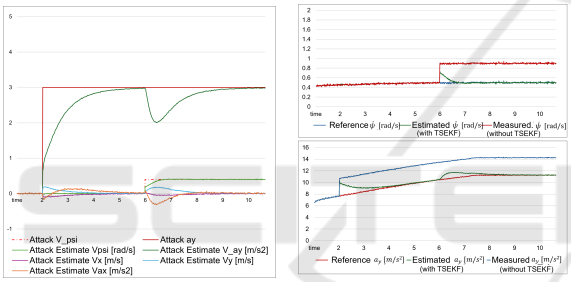


(a) Cyber-attack Estimate (b) Vehicle State

Figure 5: Circular Drive.

lateral speed. Similarly to the linear constant speed scenario, Figure (5a) shows that following the attack (solid red line) at the lateral speed, the corresponding estimated attack value (solid green line) rises swiftly and converges to the actual attack magnitude within approximately 1 second. Similarly, an attack on lateral speed causes fluctuations in the estimated attack value for acceleration, but these ultimately converge to zero around 2 seconds. In Figure (5b), the algorithm's estimated values experience a slight increase before converging to the actual lateral speed.

To further validate the algorithm's performance under combined motion conditions, we introduced an acceleration scenario to the constant-speed circular motion. The steering angle was kept constant at 0.1 rad, while the vehicle speed increased uniformly from $50km/h$ to $70km/h$. At $t = 2s$, an FDI attack with $b_{a_y} = 3(m/s)$ was initiated in the lateral acceleration, followed by another FDI attack with $b_{\psi} = 0.4(rad/s)$ on the yaw rate at $t = 6s$.



(a) Cyber-attack Estimate (b) Vehicle State

Figure 6: Circular Accelerated Drive.

In Figure(6a), when only a lateral acceleration attack is applied (solid red line), similar to the longitudinal acceleration attack, the attack magnitude is effectively tracked. This causes only minor fluctuations in the estimated attack values for other variables. The actual lateral acceleration of the vehicle, in the absence of attack, is also accurately estimated. However, when a cyber-attack is applied to the yaw rate (dashed red line), it induces a change in the estimated attack values not only for yaw rate but also for acceleration. This occurs because lateral acceleration is one of the factors causing yaw rate change; thus, when the yaw rate deviates due to a cyber-attack, the algorithm may also infer that lateral acceleration has been attacked, resulting in a perceived shift. However, over time, the estimated value for the lateral acceleration attack converges to its actual value. This fluctuation also affects the vehicle state estimation, as shown in Figure (6b): for the acceleration state estimation, there is a noticeable deviation from the vehicle's true state between 6 and 7 seconds, though this deviation remains much smaller than that of the direct measurement. For the yaw rate state estimation, after a slight increase, it swiftly converges to the vehicle's actual state.

In summary, the TSEKF algorithm effectively detects cyber-attack magnitudes and estimates vehicle state within 1–2 seconds under various driving conditions. It performs better in estimating attacks and states when targeting acceleration, though attacks on speed-related variables cause some fluctuations in the acceleration estimates. Despite this, the estimated values remain significantly closer to the true state compared to direct measurements without TSEKF.

5 CONCLUSION

In this paper, we propose a TSEKF-based cybersecurity approach to address FDI attacks in ICVs. By separating state and attack estimation into two stages, the method reduces computational complexity while achieving accurate state estimation and attack detection across various vehicle states. The experimental results demonstrate strong stability and robustness against FDI attacks under different driving conditions. In the first stage, the algorithm accurately estimates vehicle motion states despite attacks, while in the second stage it effectively detects and estimates sensor data deviations, providing robust protection. This research expands the application of Kalman filters in ICV security and lays a foundation for advanced cyber-attack protection mechanisms.

In summary, the proposed method offers a promising solution to ensure the safety of ICV and counter cyber threats. Future work could explore the integration of additional attack patterns and multisensor fusion techniques to enhance protection in complex scenarios.

ACKNOWLEDGEMENTS

This work is supported by National Key R&D Program of China (Grant No.2023YFB3107400), National Key R&D Program of China (Grant No.2022YFB2503300) and the National Natural Science Foundation of China (No. U22A202101).

REFERENCES

Abdollahi Biron, Z., Dey, S., and Pisu, P. (2016). Sensor fault diagnosis of connected vehicles under imperfect communication network. In *Dynamic Sys-*

tems and Control Conference, volume 50695, page V001T16A003. American Society of Mechanical Engineers.

Best, M. C. (2014). A new empirical 'exponential' tyre model. *International Journal of Vehicle Design*. Publisher: Inderscience Publishers Ltd.

Dieter, S., Manfred, H., and Roberto, B. (2018). Vehicle dynamics: modeling and simulation.

Dutta, R. G., Yu, F., Zhang, T., Hu, Y., and Jin, Y. (2018). Security for Safety: A Path Toward Building Trusted Autonomous Vehicles. In *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–6. ISSN: 1558-2434.

Guo, L., Ye, J., and Yang, B. (2021). Cyberattack Detection for Electric Vehicles Using Physics-Guided Machine Learning. *IEEE Transactions on Transportation Electrification*, 7(3):2010–2022.

Han, M. L., Kwak, B. I., and Kim, H. K. (2018). Anomaly intrusion detection method for vehicular networks based on survival analysis. *Vehicular Communications*, 14:52–63.

He, X., Hashemi, E., and Johansson, K. H. (2021). Distributed control under compromised measurements: Resilient estimation, attack detection, and vehicle platooning. *Automatica*, 134:109953.

Henning, K.-U. and Sawodny, O. (2016). Vehicle dynamics modelling and validation for online applications and controller synthesis. *Mechatronics*, 39:113–126.

Hossain, M. D., Inoue, H., Ochiai, H., Fall, D., and Kadobayashi, Y. (2020). LSTM-Based Intrusion Detection System for In-Vehicle Can Bus Communications. *IEEE access : practical innovations, open solutions*, 8:185489–185502.

Javed, A. R., ur Rehman, S., Khan, M. U., Alazab, M., and G, T. R. (2021). CANintelliIDS: Detecting In-Vehicle Intrusion Attacks on a Controller Area Network Using CNN and Attention-Based GRU. *IEEE Transactions on Network Science and Engineering*, 8(2):1456–1466.

Ju, Z., Zhang, H., Li, X., Chen, X., Han, J., and Yang, M. (2022). A Survey on Attack Detection and Resilience for Connected and Automated Vehicles: From Vehicle Dynamics and Control Perspective. *IEEE Transactions on Intelligent Vehicles*, 7(4):815–837.

Ju, Z., Zhang, H., and Tan, Y. (2020). Distributed Deception Attack Detection in Platoon-Based Connected Vehicle Systems. *IEEE Transactions on Vehicular Technology*, 69(5):4609–4620.

Keller, J.-Y. and Darouach, M. (1997). Optimal two-stage Kalman filter in the presence of random bias. *Automatica*, 33(9):1745–1748. Publisher: Elsevier.

Lin, H.-C., Wang, P., Chao, K.-M., Lin, W.-H., and Chen, J.-H. (2022). Using Deep Learning Networks to Identify Cyber Attacks on Intrusion Detection for In-Vehicle Networks. *Electronics*, 11(14):2180. Publisher: Multidisciplinary Digital Publishing Institute.

Lo, W., Alqahtani, H., Thakur, K., Almadhor, A., Chander, S., and Kumar, G. (2022). A hybrid deep learning based intrusion detection system using spatial-temporal representation of in-Vehicle network traffic. *Vehicular Communications*, 35:100471. Publisher: Elsevier.

Lokman, S.-F., Othman, A. T., and Abu-Bakar, M.-H. (2019). Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review. *EURASIP Journal on Wireless Communications and Networking*, 2019(1):184.

May, M. P., Henning, K.-U., and Sawodny, O. (2023). Experimental validation of sensor fault estimation for vehicle dynamics with a nonlinear tire model. *Control Engineering Practice*, 141:105725.

Mwanje, M. D., Kaiwartya, O., Aljaidi, M., Cao, Y., Kumar, S., Jha, D. N., Naser, A., and Lloret, J. (2024). Cyber security analysis of connected vehicles. *IET Intelligent Transport Systems*.

Pacejka, H. B. (2012). Semi-Empirical Tire Models. *Tire and Vehicle Dynamics*, pages 149–209. Publisher: Elsevier.

Tan, R., Nguyen, H. H., Foo, E. Y. S., Yau, D. K. Y., Kalbarczyk, Z., Iyer, R. K., and Gooi, H. B. (2017). Modeling and Mitigating Impact of False Data Injection Attacks on Automatic Generation Control. *IEEE Transactions on Information Forensics and Security*, 12(7):1609–1624.

Wang, Y., Masoud, N., and Khojandi, A. (2020). Real-time sensor anomaly detection and recovery in connected automated vehicle sensors. *IEEE transactions on intelligent transportation systems*, 22(3):1411–1421. Publisher: IEEE.

Zhang, L. and Ma, D. (2022). A Hybrid Approach Toward Efficient and Accurate Intrusion Detection for In-Vehicle Networks. *IEEE access : practical innovations, open solutions*, 10:10852–10866.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS