Inter-Message Correlation for Intrusion Detection in Controller Area Networks

Sang Uk Sagong[0000−0003−2463−094X], Radha Poovendran[0000−0003−0269−8097], and Linda Bushnell[0000−0002−8751−2409]

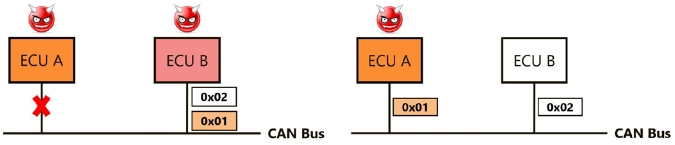
University of Washington, Seattle, USA
{sagong,rp3,lb2}@uw.edu

Abstract. Electronic Control Units (ECUs) exchange data via in-vehicle network protocols such as the Controller Area Network (CAN) protocol. These protocols do not encrypt data or authenticate messages since they were designed for an isolated network. Many studies have developed Intrusion Detection Systems (IDSs) that fingerprint each ECU to secure the CAN protocol. These IDSs, however, cannot detect an attack in which an adversary spoofs sensor measurements or control signals in a message without changing the transmitter of that message. In order to detect such attacks, we develop a motion-based IDS (MIDS) that exploits the correlation between messages that convey the same information of a vehicle's movement, such as vehicle speed. We also introduce a new metric to quantify the effectiveness of MIDS. We evaluate MIDS using CAN data from two real vehicles by demonstrating that MIDS can detect the attacks on the CAN bus or ECUs.

Keywords: Intrusion Detection System · Message Data · Correlation · Controller Area Network.

1 Introduction

Data for controlling a vehicle is exchanged among Electronic Control Units (ECUs) via in-vehicle network protocols such as the Controller Area Network (CAN) [4], FlexRay [3], and Local Interconnect Network [5]. In these protocols, a message is not encrypted or authenticated because the in-vehicle networks were designed to be isolated from external networks. Modern vehicles, however, are equipped with ECUs that have outward-facing interfaces (e.g., Bluetooth, Wi-Fi, and cellular network), which may introduce attack surfaces [8, 15]. It is difficult to incorporate cryptographic primitives such as data encryption or message authentication in the existing in-vehicle protocols due to the lack of backward compatibility with legacy systems [16]. As a consequence, anomaly-based Intrusion Detection Systems (IDSs) have been developed to detect attacks on the CAN bus by tracking abnormal deviations in physical properties of the bus or ECUs on the bus [9, 11, 13, 17, 19]. Frequently used physical properties are message frequency [13], clock skew of an ECU [9, 19], entropy of the CAN bus [17], and voltage levels of the CAN bus [11].



(a) Masquerade attack **(b)** Data falsification attack

Fig. 1: Attack models. In a masquerade attack, the transmitter of a spoofed message 0x01 is changed from ECU A to ECU B. In a data falsification attack, however, the spoofed message 0x01 is transmitted from the compromised ECU A that is the original transmitter of message 0x01 after the attack occurs.

Consider two ECUs A and B that transmit messages with IDs 0x01 and 0x02, respectively. As illustrated in Fig 1a, the compromised ECU B injects a spoofed message 0x01 instead of ECU A in a masquerade attack [9, 19]. The anomaly-based IDSs that fingerprint each ECU can detect the masquerade attack because the transmitter of message 0x01 is changed, which causes deviations in physical properties. An adversary, however, may spoof data of message 0x01 using the compromised ECU A as shown in Fig. 1b. A data falsification attack can bypass these anomaly-based IDSs since the transmitter of message 0x01 is not changed after the attack. In order to detect the data falsification attack, we propose a motion-based IDS (MIDS). We make the following contributions in this paper:

- We propose MIDS that exploits the correlation between messages that contain speed-related data, i.e., wheel speed, vehicle speed, and odometer data.
- We analyze the detection probability of MIDS by deriving its bounds under the data falsification attack.
- We introduce a new metric called ϵ-Deviation Index to quantify the effectiveness of MIDS.
- We demonstrate MIDS using data from two real vehicles. Our hardware evaluation shows that MIDS can detect the data falsification attack.

The rest of the paper is organized as follows. Section 2 reviews the related work, and Section 3 summarizes a background on the CAN protocol. Section 4 presents the adversary model. MIDS is proposed in Section 5, and Section 6 derives bounds on the detection probability of MIDS. Section 7 presents the experimental evaluation. Section 8 concludes the paper.

2 Related Work

Multiple IDSs that detect cyber attacks using abnormal deviations in traffic through the CAN bus have been proposed [13, 17]. Based on the fact that most of the messages in the CAN protocol are transmitted with a fixed length and period, an IDS that detects the existence of spoofed messages using a frequency of message occurrence is proposed [13]. The authors of [17] proposed the entropy-based IDS that exploits coincidence among a set of messages. The entropy-based IDS, however, can be bypassed if an adversary replicates the structure and pattern of legitimate traffic [9].

In order to detect attacks that replicate the legitimate traffic, IDSs that fingerprint ECUs by exploiting physical properties of ECUs have been developed [9–11]. The authors of [9] proposed the clock-based IDS (CIDS) that detects an attack by tracking a sudden change in the clock skew of ECUs. It has been demonstrated that the cloaking attack bypasses the CIDS by mimicking the clock skew [19]. Since the voltage characteristics of ECUs are determined by a CAN transceiver's hardware, they are unique to each ECU and difficult to replicate. The voltage-based IDSs are proposed, which detect an attack by exploiting the voltage characteristics [10, 11]. An adversary may compromise a legitimate ECU and inject spoofed messages that convey false sensor measurements or control signals using that compromised ECU. These IDSs cannot detect this data falsification attack because the transmitter of the spoofed messages remains the same after the attack, which does not induce any deviations in physical properties.

The authors of [12] demonstrated that there is the correlation between measurements from different types of sensors under the normal operation of a vehicle because physical movement affects multiple sensors simultaneously. They, however, did not propose a concrete structure of an IDS or analyze the detection probability of the IDS under the data falsification attack.

3 Preliminaries

In this section, we review the CAN protocol and explain how a sensor measurement or control signal can be extracted from a message.

3.1 CAN Protocol Background

The CAN protocol is a multi-master broadcast bus network in which any ECU can transmit messages and receive all ongoing messages through the CAN bus. An ECU that accesses the CAN bus first transmits a message. If two or more ECUs attempt to send messages simultaneously, the message with the smallest ID is transmitted first through a content-based collision detection process called *arbitration*. For example, consider two ECUs A and B that try to send their messages with IDs 0x001 and 0x010, respectively. ECU B recognizes that a higher priority message is being transmitted and stops transmitting its message through an arbitration.

A data frame in the CAN protocol is composed of seven fields as illustrated in Fig. 2, and the length of the data field can be varied from 1 to 8 bytes. The data field is not encrypted, and there does not exist a field for message authentication. If a message is not transmitted and received correctly due to external electromagnetic interference or malfunction of CAN transceivers, the ECU retransmits that message after an error frame is transmitted.

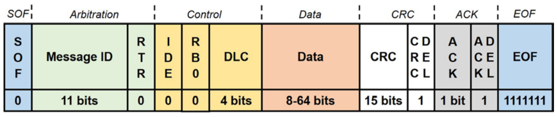


Fig. 2: Structure of a data frame in the CAN protocol.

3.2 Extracting Sensor Measurements and Control Signals

An ECU encapsulates multiple sensor measurements or control signals in a message after converting them to a bit sequence between 1 and 8 bytes. In order to compute correlation coefficients, these sensor measurements and control signals have to be extracted from the messages. We find actual values of the sensor measurements and control signals as follows. First, we check a unit of the targeted data that we want to extract. Second, we check the message ID and bit location in the data field that contain the targeted data. This information can be found either by reverse engineering or in an interface control document of a vehicle. Third, we convert a bit sequence to a decimal number. If the Most Significant Bit (MSB) is a sign bit, the decimal number is negative when the MSB is 1. Fourth, we multiply a conversion ratio to that decimal number and then add an offset to get the actual value. For instance, Message 0x0B0 of a 2010 Toyota Camry contains measurements of a wheel speed sensor, which is located from the 1^{st} bit to 16^{th} bit [18]. Also, the wheel speed is measured in km/h, and the conversion ratio and offset for the wheel speed are 0.1 and 0, respectively.

4 Adversary Model

In this section, we describe an adversary model in which an adversary compromises a legitimate ECU and launches a data falsification attack. The adversary may have physical access to the CAN bus through the on-board diagnostics (OBD-II) port that is mandated for all automobiles in EU [21] and US [7]. Then, the adversary uploads its malicious code to a targeted ECU using a pass-thru device such as Hyundai Global Diagnostic System [1] and Volkswagen VAG-COM Diagnostic System [20] to compromise the ECU. The adversary can also remotely compromise an ECU without physical access to the CAN bus [8, 15]. We consider two cases: (1) the adversary compromises an ECU having a telematics unit, and (2) the adversary compromises an ECU without a telematics unit. In the first case, the adversary remotely accesses the operating system of an ECU having a telematics unit to figure out a particular code that handles wireless connectivity by reverse engineering. By exploiting that code, the adversary executes its malicious code on that ECU [8]. In the second case, the adversary lets a remotely compromised ECU, which is compromised in the first case, upload the malicious code to a targeted ECU without a telematics unit as a pass-thru device does. As a result, the adversary can compromise any ECUs in the CAN bus [8].

Once an ECU is compromised, the adversary may manipulate periods and data fields of messages. Since an attack that changes periods of messages can

Table 1: Frequently used symbols

Notation	Variable
n_{normal}	number of samples in normal data
n_{attack}	number of samples per attack experiment
w	window size
$\bar{a}$	vector containing measurements from sensor a
$\bar{b}$	vector containing measurements from sensor b
$\bar{d}$	disturbance generated according to a Gaussian distribution $N(0, \sigma^2)$
$\bar{b'}$	$\bar{b} + \bar{d}$
σ_a	standard deviation of $\bar{a}$
σ	standard deviation of $\bar{d}$
P_d	detection probability of MIDS
σ_{lim}	minimum standard deviation of $\bar{d}$, at which $P_d > 1 - \epsilon$
σ_{bypass}	maximum standard deviation of $\bar{d}$, below which P_d=0
σ_{detect}	minimum standard deviation of $\bar{d}$, above which P_d=1
$\varDelta_{bypass}$	decrement of correlation coefficient per sample for σ_{bypass}
$\varDelta_{detect}$	decrement of correlation coefficient per sample for σ_{detect}
ρ	vector of correlation coefficient between $\bar{a}$ and $\bar{b}$ (or $\bar{b'}$)
ρ'	normalized ρ using μ_ρ and σ_ρ
ρ_0	last correlation coefficient before an attack
ρ_{attack}	first correlation coefficient after an attack
ρ_{ref}	vector containing all previous correlation coefficients such that $\rho' < \gamma$
μ_ρ/σ_ρ	mean/standard deviation of elements of ρ_{ref}
L^+/L^-	upper/lower control limits of CUSUM
κ	sensitivity threshold of CUSUM
$\varGamma$	detection threshold of CUSUM
γ	update threshold of CUSUM

be easily detected using the existing IDSs [9–11, 13, 17], we consider two types of the data falsification attacks. In the first type of attack, the adversary adds a disturbance to the legitimate data, where the disturbance is generated according to a zero-mean Gaussian distribution in order to analyze effects of deviation of the disturbance. In the second type of attack, the adversary manipulates one of the wheel speed sensor values to be increased, in which a non-zero-mean disturbance is added, while other sensor values remain the same. The physical properties of ECUs do not deviate from the normal behavior after the data falsification attack, and the data falsification attack bypasses the existing IDSs.

5 Motion-based Intrusion Detection System

In this section, we propose MIDS, an IDS that exploits the correlation between messages. MIDS may exploit any pair of messages that contain data of the same physical movement of a vehicle. Without loss of generality, we use speed-related data for MIDS in this paper. Frequently used symbols are summarized in Table 1.

Table 2: Correlation coefficients of four wheel speeds, vehicle speed, and steering wheel angle.

Data 1	Data 2	Correlation Coefficient
Wheel Speed 1	Wheel Speed 2	0.9999
Wheel Speed 1	Wheel Speed 3	0.9999
Wheel Speed 1	Wheel Speed 4	0.9999
Wheel Speed 2	Wheel Speed 3	0.9998
Wheel Speed 2	Wheel Speed 4	0.9999
Wheel Speed 3	Wheel Speed 4	1.0000
Wheel Speed 1	Vehicle Speed	0.9999
Wheel Speed 1	Steering Wheel Angle	-0.5324

5.1 Correlation Between Messages

Consider $\bar{a}$ and $\bar{b}$ with the same length of n, where a_i and b_i denote the i^{th} samples of $\bar{a}$ and $\bar{b}$, respectively. The correlation coefficient ρ between $\bar{a}$ and $\bar{b}$ can be computed as follows:[1]

$$\rho = \frac{\sum_{i=1}^{n}(a_i - \mu_a)(b_i - \mu_b)}{\sqrt{\sum_{i=1}^{n}(a_i - \mu_a)^2}\sqrt{\sum_{i=1}^{n}(b_i - \mu_b)^2}}, \tag{1}$$

where μ_a and μ_b denote means of $\bar{a}$ and $\bar{b}$, respectively. Multiple sensors are implemented in a vehicle to measure the same movement such as vehicle speed, and measurements from these sensors are highly correlated with each other during normal operation [12]. We use data collected from a 2010 Toyota Camry, which contains four wheel speeds, vehicle speed, and steering wheel angle [18]. Table 2 shows correlation coefficients for possible pairs of four wheel speeds, vehicle speed, and steering wheel angle, using Eq. (1). All four wheel speeds and vehicle speed are highly correlated with each other (i.e., $\rho > 0.99$), while the steering wheel angle is not correlated with wheel speed 1 because the vehicle speed and steering wheel angle are independently controlled.

5.2 Proposed MIDS

MIDS consists of a correlation computer and cumulative sum (CUSUM) detector as illustrated in Fig. 3. In this section, we explain how MIDS computes correlation coefficients and tracks deviations in them to detect an attack.

Correlation Computer The correlation computer calculates correlation coefficients between sensor measurements and control signals in messages. An ECU may not save all ongoing messages through the CAN bus due to its limited memory size and computation capability. Computing correlation coefficients using all past data, as defined in Eq. (1), is not a viable option in practice. MIDS

[1] We use Pearson correlation coefficient because sensor measurements of the same vehicle movement are in the linear relationship (Chapter 5.6 in [14]).

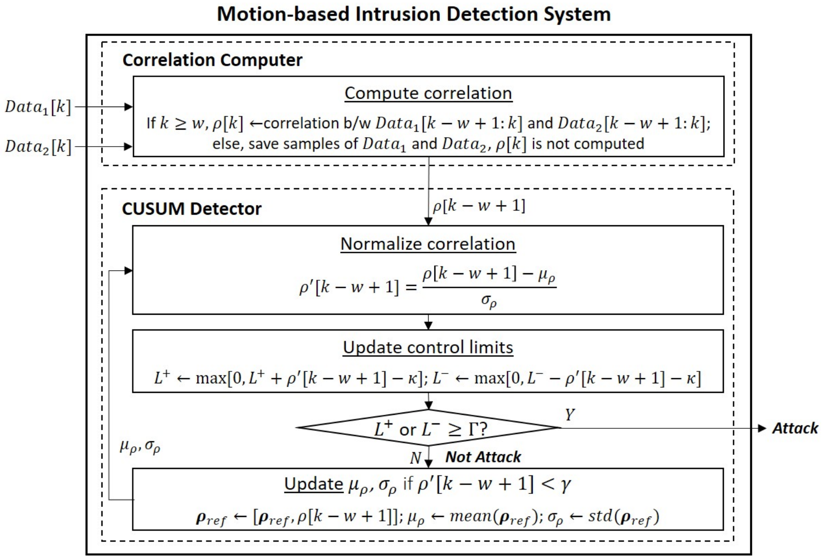

Motion-based Intrusion Detection System
Correlation Computer
$Data_1[k]$
$Data_2[k]$
Compute correlation
If $k \geq w$, $\rho[k]$ ←correlation b/w $Data_1[k-w+1:k]$ and $Data_2[k-w+1:k]$;
else, save samples of $Data_1$ and $Data_2$, $\rho[k]$ is not computed
CUSUM Detector
$\rho[k-w+1]$
Normalize correlation
$\rho'[k-w+1] = \frac{\rho[k-w+1]-\mu_\rho}{\sigma_\rho}$
Update control limits
$L^+ \leftarrow \max[0, L^+ + \rho'[k-w+1]-\kappa]$; $L^- \leftarrow \max[0, L^- - \rho'[k-w+1]-\kappa]$
L^+ or $L^- \geq \Gamma$?
Y
Attack
μ_ρ, σ_ρ
N
Not Attack
Update μ_ρ, σ_ρ if $\rho'[k-w+1] < \gamma$
$\boldsymbol{\rho}_{ref} \leftarrow [\boldsymbol{\rho}_{ref}, \rho[k-w+1]]$; $\mu_\rho \leftarrow mean(\boldsymbol{\rho}_{ref})$; $\sigma_\rho \leftarrow std(\boldsymbol{\rho}_{ref})$

Fig. 3: Structure of MIDS.

computes correlation coefficients using the w most recent data samples in real time, which is updated by a sliding window. The window size is determined by ECU's memory size and computation capability. If the number of data samples is less than w (i.e., $k < w$), MIDS saves the data samples in its memory but does not compute a correlation coefficient. When $k \geq w$, correlation coefficients are computed and fed to the CUSUM detector as described in Fig. 3.

CUSUM Detector MIDS uses the CUSUM method that computes cumulative sums of deviations from the normal value in order to detect a sudden change. The CUSUM method is widely used to track the drift of values that steadily increase or decrease (Chapters 2.1 and 2.2 in [6]). MIDS keeps the normal behavior of correlation coefficients by tracking the mean μ_ρ and standard deviation σ_ρ of ρ_{ref} that is a vector containing all previous ρ's from the normal data. For every incoming ρ, MIDS computes the normalized correlation coefficient $\rho' = \frac{\rho - \mu_\rho}{\sigma_\rho}$. Using ρ', the upper and lower control limits, L^+ and L^-, are updated as follows:

$$L^+ \leftarrow \max\left[0, L^+ + \rho' - \kappa\right],\ L^- \leftarrow \max\left[0, L^- - \rho' - \kappa\right],$$

where both L^+ and L^- are initially zero and κ is the sensitivity threshold. MIDS declares an attack when either control limit exceeds the detection threshold $\varGamma$. MIDS appends the current ρ to ρ_{ref} if $\rho' < \gamma$, where γ is the update threshold. Then, μ_ρ and σ_ρ are updated and used to normalize the next incoming ρ. We set γ, κ, and $\varGamma$ to such values that make zero false-positive probability in the normal data [9, 19].

6 Theoretical Analytics of MIDS

In this section, we derive standard deviations of a disturbance at which the attack is not detected by MIDS or always detected by MIDS. Then, we define a performance metric. Symbols are referred in Table 1.

Bounds on Detection Probability of MIDS Consider $\bar{a}$ and $\bar{b}$ that are measurements from two different sensors. An adversary spoofs $\bar{b}$ by adding a disturbance $\bar{d}$ to yield $\bar{b}' := \bar{b} + \bar{d}$, where $\bar{d}$ is generated according to a Gaussian distribution $N(0, \sigma^2)$. $\bar{b}'$ can be approximated as $\bar{a} + \bar{d}$ because $\bar{b} \simeq \bar{a}$. If $\bar{b}$ is disturbed from the l^{th} element in the window with size w, using Eq. (1), the correlation coefficient ρ between $\bar{a}$ and $\bar{b}'$ can be approximated to be

$$\rho \simeq \frac{\sigma_a^2 + \sum_{i=l}^{w} a_i d_i}{\sigma_a \sqrt{\sigma_a^2 + 2\sum_{i=l}^{w} a_i d_i + \sum_{i=l}^{w} d_i^2}}. \tag{2}$$

For the analysis, we assume that ρ linearly decreases as l decreases, which is experimentally observed in Section 7. Also, μ_ρ and σ_ρ are not updated because $\rho' \geq \gamma$ during the attack as explained in Section 5.

Theorem 1. *A data falsification attack is not detected by MIDS if* $\sigma < \sigma_{bypass}$*, where* $\sigma_{bypass} = \sigma_a \sqrt{\frac{1}{(\rho_0 - \kappa\sigma_\rho)^2} - 1}$.

Proof. When all elements of $\bar{b}'$ are disturbed (i.e., l=1), $\sum_{i=l}^{w} a_i d_i$ is zero because the mean of $\bar{d}$ is zero and $\bar{a}$ and $\bar{d}$ are independent. As a result, $\rho \simeq \frac{\sigma_a^2}{\sigma_a\sqrt{\sigma_a^2+\sigma^2}} = \frac{\sigma_a}{\sqrt{\sigma_a^2+\sigma^2}}$, which is the minimum value of ρ after the attack. Since ρ decreases linearly, a decrement of ρ per sample for σ_{bypass}, Δ_{bypass}, can be derived as

$$\Delta_{bypass} = \frac{\rho_0 - \frac{\sigma_a}{\sqrt{\sigma_a^2+\sigma^2}}}{w}. \tag{3}$$

We find the condition under which both control limits, L^+ and L^-, stay zero. The normalized correlation coefficient at the k^{th} data sample, $\rho'(k-w+1)$, can be derived in terms of Δ_{bypass} as

$$\rho'(k-w+1) = \frac{\rho_0 - (k - n_{normal})\Delta_{bypass} - \mu_\rho}{\sigma_\rho} = -\frac{(k - n_{normal})\Delta_{bypass}}{\sigma_\rho},$$

because $\rho_0 \simeq \mu_\rho$ when there is no attack. L^+ and L^- stay zero if $\frac{(k-n_{normal})\Delta_{bypass}}{\sigma_\rho}$ is less than κ for all $k \leq (n_{normal} + w)$, which indicates that the data falsification attack is not detected. At the last attack sample that corresponds to the minimum value of ρ (i.e., $k = n_{normal} + w$), Δ_{bypass} becomes

$$\Delta_{bypass} = \frac{\kappa\sigma_\rho}{w}. \tag{4}$$

Substituting Eq. (3) into Eq. (4) gives $\sigma_{bypass} = \sigma_a \sqrt{\frac{1}{(\rho_0 - \kappa\sigma_\rho)^2} - 1}$. □

Theorem 2. *MIDS can always detect a data falsification attack if* $\sigma > \sigma_{detect}$*, where* $\sigma_{detect} = \sigma_a \sqrt{\frac{1}{(\rho_0 - j\sigma_\rho(\mathit{\Gamma}+\kappa))^2} - 1}$.

Proof. Using Eq. (2), ρ decreases from ρ_0 to $\frac{\sigma_a}{\sqrt{\sigma_a^2 + {\sigma_{detect}}^2}}$ at the j^{th} attack sample in the window if σ is set to σ_{detect}. $\frac{\sigma_a}{\sqrt{\sigma_a^2 + {\sigma_{detect}}^2}} \simeq 0$ because $\sigma_{detect} >> \sigma_a$. As a consequence, j can be approximated as $\left\lceil \frac{\rho_0}{\rho_0 - \rho_{attack}} \right\rceil$, where $\lceil\,\rceil$ denotes a ceiling operator. A decrement of ρ per sample for σ_{detect}, $\mathit{\Delta}_{detect}$, can be derived as

$$\mathit{\Delta}_{detect} = \frac{\rho_0 - \frac{\sigma_a}{\sqrt{\sigma_a^2 + {\sigma_{detect}}^2}}}{j}. \tag{5}$$

An attack is always detected if either L^+ and L^- becomes larger than $\mathit{\Gamma}$ at the first attack sample. Notice that $L^- > \mathit{\Gamma}$ when $\rho_{attack} \leq \mu_\rho - \sigma_\rho\left(\mathit{\Gamma} + \kappa\right)$. As a result, $\mathit{\Delta}_{detect}$ can be also represented as follows:

$$\mathit{\Delta}_{detect} = \sigma_\rho\left(\mathit{\Gamma} + \kappa\right). \tag{6}$$

Substituting Eq. (5) into Eq. (6) gives $\sigma_{detect} = \sigma_a \sqrt{\frac{1}{(\rho_0 - j\sigma_\rho(\mathit{\Gamma}+\kappa))^2} - 1}$. □

Performance Metric We introduce a metric that formalizes how the detection probability is close to σ_{bypass}. Let $P_d(\sigma)$ denote the detection probability of MIDS when the standard deviation of the disturbance is σ. We define σ_{lim} as:

$$\sigma_{lim}(\epsilon) = \min\left[\sigma : P_d(\sigma) > 1 - \epsilon\right].$$

We define the ϵ-Deviation Index (ϵ-DI) as ϵ-DI$= \sigma_{lim}(\epsilon) - \sigma_{bypass}$. A smaller value of ϵ-DI signifies a more effective detector and less freedom for the attacker at choosing a standard deviation of the disturbance.

7 Evaluation

In this section, we demonstrate that MIDS is effective in detecting a data falsification attack by using data from two real vehicles.

7.1 Testbeds

Chevrolet Camaro A 2016 Chevrolet Camaro is used in our experiment in a controlled environment as shown in Fig 4a. A Vector GL1000 logger is connected to the CAN bus of the vehicle via the OBD-II port to collect all ongoing messages through the CAN bus as shown in Fig 4b. We collect data while the vehicle is being driven on the road for 68 minutes in order to validate that MIDS detects a data falsification attack in a practical environment. We also collect data for 24 minutes while the Camaro is on a chassis dynamometer. This experiment

(a) Chevrolet Camaro **(b)** Vector GL1000 logger

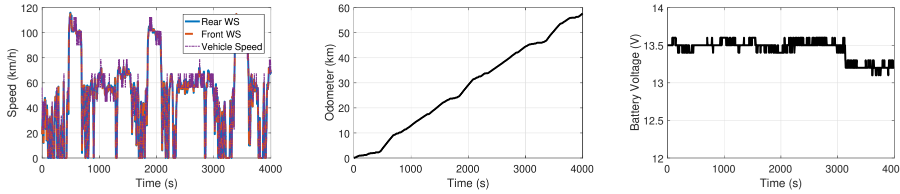
Fig. 4: A 2016 Chevrolet Camaro is used to validate MIDS. All messages transmitted through the CAN bus are collected using a Vector GL1000 logger.

emulates an attack in which an adversary spoofs the rear wheel speed to large values because only the rear wheels rotate on the chassis dynamometer (i.e., adding non-zero-mean disturbance). We find messages that contain rear wheel speed, front wheel speed, odometer data, and battery voltage data using the interface control document.[2]

Toyota Camry In order to demonstrate that MIDS is applicable to other vehicles, we use the CAN data logged from a 2010 Toyota Camry [18]. The data is collected using a Dearborn Group Gryphon S3 and the Hercules software through a wireless link. Messages 0x0B0 and 0x0B2 contain measurements from all four wheel speed sensors (i.e., two wheel speeds in each message). Because each wheel that corresponds to respective wheel speed sensor is not reverse engineered in [18], we indicate two wheel speeds in Message 0x0B0 as wheel speeds 1 (from 1st bit to 16th bit) and 2 (from 17th bit to 32nd bit) and the other two wheel speeds in Message 0x0B2 as wheel speeds 3 (from 1st bit to 16th bit) and 4 (from 17th bit to 32nd bit). Message 0x610 (from 17th bit to 24th bit) contains the vehicle speed [2].

CAN Data Preprocessing In order to compute correlation coefficients using Eq. (2), the number of the message that contains sensor measurements or control signals has to be the same as that of the other message. The number of two different messages, however, can be different if their periods are different. For instance, periods of Messages 0x0B0 (wheel speed) and 0x610 (vehicle speed) in the Camry are 10ms and 500ms, respectively. As a result, the length of the wheel speed is 50 times longer than that of the vehicle speed. We generate more samples for the vehicle speed by interpolation to make its length be the same as that of the wheel speed when computing a correlation coefficient between the wheel speed and vehicle speed in Table 2.

[2] We assume that MIDS uses a pair of data that are transmitted from two different ECUs where only one ECU is compromised, thus data from only one ECU is spoofed.



(a) Front wheel, rear wheel, and vehicle speeds **(b)** Odometer **(c)** Battery voltage

Fig. 5: Front wheel, rear wheel, and vehicle speeds, odometer, and battery voltage of the Camaro when it is driven on the road.

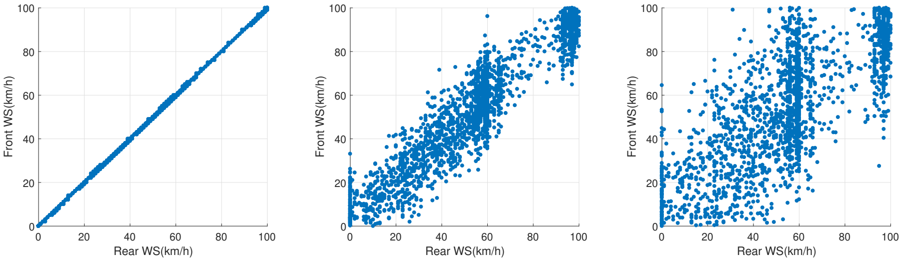
The rear wheel speed, front wheel speed, odometer, and battery voltage of the Camaro are presented in Fig. 5 when it is driven on the road. For the odometer, we set the initial value to 0km to make Fig. 5b show the distance that the Camaro is driven in that data collection. We compute the vehicle speed by differentiating odometer values and verify that wheel speeds closely match with the vehicle speed. The battery voltage is within a normal range (12-14V). We also extract four wheel speeds and vehicle speed of the Camry as well.

7.2 Example of MIDS

For an illustration of MIDS, we observe the correlation coefficients and control limits of the CUSUM detector under a single execution of a data falsification attack on the Camaro. We generate attack data by adding a disturbance to the front wheel speed, where the disturbance is generated according to a Gaussian distribution with a zero mean and standard deviation σ. In this experiment, we set window size to 1500, update threshold γ to 4, sensitivity threshold κ to 7, and detection threshold Γ to 5 in order to avoid false alarm.

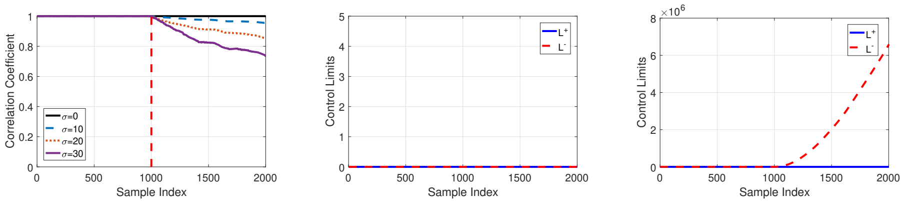
Fig. 6 shows scatter plots between the front and rear wheel speeds using 2500 samples. As demonstrated in Fig. 6a, the front wheel speed is almost the same as the rear wheel speed without an attack. Figs. 6b and 6c present scatter plots under the attack when σ is 10 and 20, respectively. As increasing σ, points are scattered in a wider area, which indicates that the front wheel speed becomes less correlated with the rear wheel speed.

MIDS computes the correlation coefficients between the front and rear wheel speeds for 250 seconds to track the normal behavior of the correlation coefficients before a data falsification attack occurs. Then, the attack data is fed to MIDS. Fig. 7a demonstrates that the correlation coefficient ρ is greater than 0.99 before the attack, which indicates that the front wheel speed is highly correlated with the rear wheel speed. After the attack occurs, ρ decreases to 0.9566, 0.8514, and 0.7246 at the sample index of 2000 when σ is set to 10, 20, and 30, respectively. This result shows that ρ decreases more if a larger disturbance is added to the legitimate data. When a small disturbance is added (σ=1), the deviation in ρ is



(a) Scatter plot without attack **(b)** Scatter plot for σ=10 **(c)** Scatter plot for σ=20

Fig. 6: Scatter plots between the front and rear wheel speeds of the Camaro for various values of σ.

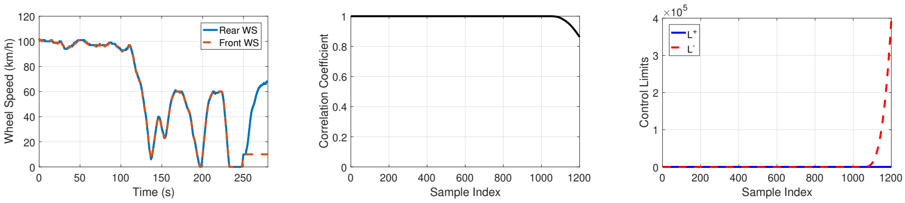


(a) Correlation coefficient for various values of σ **(b)** Control limits when σ=1 **(c)** Control limits when σ=20

Fig. 7: Example of MIDS. Using the Camaro data, the correlation coefficient and control limits are demonstrated. The attack starts from the 1001st sample that is indicated by a red dashed line in (a).

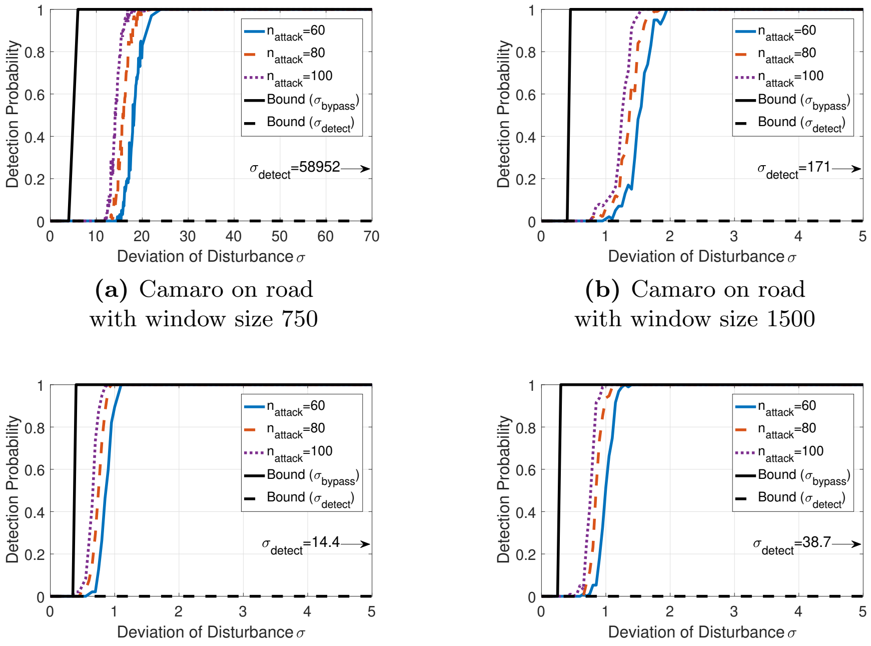
too small to increase either control limits. Consequently, L^+ and L^- stay zero as shown in Fig. 7b. When σ is set to 20, ρ suddenly drops, which makes the lower control limit increase right after the attack as demonstrated in Fig. 7c. Since the lower control limit becomes larger than $\mathit{\Gamma}$=5, MIDS detects the attack.

Using the data collected from the Camaro on the chassis dynamometer, Fig. 8a shows the front and rear wheel speeds of the Camaro when an adversary spoofs the rear wheel speed between 250s and 270s (between the sample indexes of 1000 and 1200) in order to cause a wheelspin. During the attack, the front wheel speed is set to 10km/h while the rear wheel speed data is manipulated to be increased from 10km/h to 70km/h. In this experiment, we set window size to 1500, γ to 4, κ to 7, and $\mathit{\Gamma}$ to 5. As demonstrated in Fig. 8b, the correlation coefficient decreases to 0.83 at the sample index of 1200 after the attack. Fig. 8c demonstrates that the lower control limit increases above $\mathit{\Gamma}$=5, which demonstrates that MIDS successfully detects the data falsification attack.



(a) Front and rear wheel speeds **(b)** Correlation coefficient **(c)** Control limits

Fig. 8: Example of MIDS under a data falsification attack. The adversary spoofs the rear wheel speed after 250s (after the sample index of 1000).



(a) Camaro on road with window size 750 **(b)** Camaro on road with window size 1500

(c) Camry with window size 750 **(d)** Camry with window size 1500

Fig. 9: Detection probability of MIDS. The black solid line and black dashed line indicate σ_{bypass} and σ_{detect}, respectively.

7.3 Example of Detection Probability of MIDS

We select pairs of speed-related values for MIDS in each vehicle. The front and rear wheel speeds are used for the Camaro, and a disturbance is added to the front wheel speed. For the Camry, wheel speeds 1 and 3 are used where a disturbance is added to the wheel speed 3. MIDS is fed with 2500 samples of the normal data, followed by n_{attack} samples of the attack data in each experiment. MIDS is successful if it detects an attack and is failed otherwise. 100 and 95 non-overlapping segments of size n_{attack} are prepared from the attack data to emulate 100 and 95 independent attack experiments for the Camaro and Camry, respectively. We compute the detection probability of MIDS, which is the percentage of experiments where MIDS is successful. For this evaluation, we set γ to 4, κ to 7, and $\mathit{\Gamma}$ to 5.

Fig. 9 demonstrates the detection probability of MIDS for the Camaro and Camry. We set n_{attack} to 60, 80, and 100 and window size to 750 and 1500. The black solid line and black dashed line indicate σ_{bypass} and σ_{detect}, respectively. Due to the limitation of the space, we present values of σ_{detect} with a black arrow in Fig. 9. For a window size of 1500 and ϵ=0.05, ϵ-DI decreases from 1.23 to 1.05 as increasing n_{attack} from 80 to 100 in the Camaro. The same trend is observed from the Camry that ϵ-DI decreases from 0.80 to 0.68 when n_{attack} is increased from 80 to 100. ρ decreases to a smaller value when MIDS exploits more spoofed messages per attack experiment, which makes MIDS be more effective in detecting the attack.

The correlation coefficient fluctuates less if more samples are used for computing ρ because the impact of an outlier is reduced. As a consequence, MIDS has more strict criteria (smaller σ_ρ) in detecting an anomaly in ρ if a larger window size is used. For ϵ=0.05 and n_{attack}=60, ϵ-DI reduces from 18 to 1.5 when the window size is increased from 750 to 1500 as shown in Figs. 9a and 9b. For the Camry, Figs. 9c and 9d, however, demonstrate that ϵ-DI increases from 1.05 to 1.10 as the window size increases from 750 to 1500. Even though MIDS has smaller σ_ρ when the window size is 1500, ρ drops more when the window size is 750, which dominates the impact of smaller σ_ρ. As a result, MIDS becomes more effective when a smaller window size is used in the Camry.

8 Conclusion

In this paper, we investigated a limitation of the existing anomaly-based IDSs under a data falsification attack. We proposed a motion-based IDS (MIDS) that exploits the correlation between messages. MIDS computes the correlation coefficients between two sensor measurements or control signals, and it tracks sudden deviations in the correlation coefficients to detect the data falsification attack. We derived standard deviations of the disturbance, below which the attack is not detected by MIDS and above which the attack is always detected by MIDS. In order to quantify the effectiveness of MIDS, we presented the ϵ-DI that is a range of deviation of the disturbance that an adversary may introduce without being detected. We demonstrated that MIDS can detect the data falsification attack on the Camaro and Camry by using the wheel speeds. Our work suggests that a defending mechanism exploiting the correlation between messages will increase security assurance to automobiles.

Acknowledgment

This work was supported by ONR grant N00014-16-1-2710 P00002.

References

1. Hyundai Global Diagnostic System. https://hyundai.service-solutions.com/en-US/Pages/Home.aspx, accessed: 2018-06-25
2. University of Tulsa Crash Reconstruction Research Consortium. http://tucrrc.utulsa.edu, accessed: 2019-05-02
3. International Standard ISO 17458 Road Vehicles-FlexRay Communication System, Part 1 General Information and Use Case Definition (2013)
4. International Standard ISO 11898-1 Road Vehicles-Controller Area Network (CAN), Part 1 Data Link Layer and Physical Signaling (2015)
5. International Standard ISO 17987 Road Vehicles-Local Interconnect Network (LIN), Part 1 General Information and Use Case Definition (2016)
6. Basseville, M., Nikiforov, I.: Detection of Abrupt Changes: Theory and Application. Prentice Hall Englewood Cliffs (1993)
7. California Air Resource Board: HD OBD Regulatory Documents (2012)
8. Checkoway, S., et al.: Comprehensive Experimental Analyses of Automotive Attack Surfaces. In: USENIX Conference on Security. pp. 77–92 (2011)
9. Cho, K., Shin, K.: Fingerprinting Electronic Control Units for Vehicle Intrusion Detection. In: USENIX Conference on Security Symposium. pp. 911–927 (2016)
10. Cho, K., Shin, K.: Viden: Attacker Identification on In-Vehicle Networks. In: ACM Conference on Computer and Communications Security. pp. 1109–1123 (2017)
11. Choi, W., et al.: VoltageIDS: Low-Level Communication Characteristics for Automotive Intrusion Detection System. IEEE Transactions on Information Forensics and Security **13**(8), 2114–2129 (2018)
12. Ganesan, A., et al.: Exploiting consistency among heterogeneous sensors for vehicle anomaly detection. In: SAE Technical Paper (2017)
13. Hoppe, T., et al.: Security Threats to Automotive CAN Networks - Practical Examples and Selected Short-Term Countermeasures. In: International Conference on Computer Safety, Reliability, and Security. pp. 235–248 (2008)
14. Leon-Garcia, A.: Probability, Statistics, and Random Processes For Electrical Engineering. Pearson (2011)
15. Miller, C., Valasek, C.: Remote Exploitation of An Unaltered Passenger Vehicle. In: Black Hat USA (2015)
16. Murvay, P., Groza, B.: Source Identification Using Signal Characteristics in Controller Area Networks. IEEE Signal Processing Letters **21**(4), 395–399 (2014)
17. Müter, M., Asaj, N.: Entropy-based Anomaly Detection for In-vehicle Networks. In: IEEE Intelligent Vehicles Symposium. pp. 1110–1115 (2011)
18. Ruth, R., et al.: Accuracy of Event Data in the 2010 and 2011 Toyota Camry during Steady State and Braking Conditions. SAE International Journal of Passenger Cars-Electronic and Electrical Systems **5**, 358–372 (2012)
19. Sagong, S., et al.: Cloaking the Clock: Emulating Clock Skew in Controller Area Networks. In: ACM/IEEE International Conference on Cyber-Physical Systems. pp. 32–42 (2018)
20. Ross Tech: Volkswagen VAG-COM Diagnostic System. http://www.ross-tech.com/vag-com/index.html, accessed: 2018-06-25
21. The European Union: Directive 98/69/EC of the European Parliament and of the Council (1998)