AVTPnet: Convolutional Autoencoder for AVTP anomaly detection in Automotive Ethernet Networks

Natasha Alkhatib, Maria Mushtaq, Hadi Ghauch, Jean-Luc Danger

Telécom Paris, IP Paris, Palaiseau, France
{natasha.alkhatib, maria.mushtaq, hadi.ghauch, jean-luc.danger}@telecom-paris.fr

Abstract—Network Intrusion Detection Systems are well considered as efficient tools for securing in-vehicle networks against diverse cyberattacks. However, since cyberattacks are always evolving, signature-based intrusion detection systems are no longer adopted. An alternative solution can be the deployment of deep learning based intrusion detection system (IDS) which play an important role in detecting unknown attack patterns in network traffic. To our knowledge, no previous research work has been done to detect anomalies on automotive ethernet based in-vehicle networks using anomaly based approaches. Hence, in this paper, we propose a convolutional autoencoder (CAE) for offline detection of anomalies on the Audio Video Transport Protocol (AVTP), an application layer protocol implemented in the recent in-vehicle network Automotive Ethernet. The CAE consists of an encoder and a decoder with CNN structures that are asymmetrical. Anomalies in AVTP packet stream, which may lead to critical interruption of media streams, are therefore detected by measuring the reconstruction error of each sliding window of AVTP packets. Our proposed approach is evaluated on the recently published "Automotive Ethernet Intrusion Dataset", and is also compared with other state-of-the art traditional anomaly detection and signature based models in machine learning. The numerical results show that our proposed model outperforms the other methods and excel at predicting unknown in-vehicle intrusions, with 0.94 accuracy. Moreover, our model has a low level of false alarm and miss detection rates for different AVTP attack types.

Index Terms—AVTP, Anomaly Detection, Automotive Ethernet, Neural Network, In-Vehicle Network

I. INTRODUCTION

Since the advent of powerful electronic components such as sensors and actuators as well as a robust in-vehicle infrastructure for efficient data exchange between them, driving has become safer (i.e. 360-degree surround view parking assistance, and collision avoidance systems) [1] and more pleasant (i.e. infotainment features) [2] [3] during the last several decades. Ethernet, a flexible and scalable networking technology in communication systems, is recently standardized and adopted for in-vehicle communication [4] [5] between different Electronic Component Units (ECU). In fact, it fulfills basic automotive requirements which existing in-vehicle protocols LIN, CAN, and FlexRay are not designed to cover, including reduced connectivity costs, cabling weight and support for high data bandwidth.

To ensure low-latency and high-quality transmission of time-critical and prioritized streaming data for high-end infotainment and ADAS systems, the IEEE 1722 audio-video transport protocol (AVTP) [6] is adopted. In fact, AVTP specifies a protocol for audio, video, and control data transportation on a Time-Sensitive Networking (TSN) capable network [7]. As a result, we believe that AVTP protocol will be a critical protocol for automotive ethernet based in-vehicle network in motor vehicles.

Despite the advantages of automotive ethernet, the drive toward connectivity has significantly expanded the attack surfaces of automobiles, making automotive ethernet-based in-vehicle networks increasingly susceptible to cyberattacks, posing significant security and safety issues [8]. In fact, automotive ethernet can be attacked by exploiting its vulnerabilities [9] [10]. These security breaches can affect protocols working on top of automotive ethernet including AVTP protocol and might therefore lead to the interruption of a media stream.

To address this, intrusion detection systems (IDS) should be used in addition to specific security measures as an extra layer of protection. These systems can be classified based on their analyzed activity (i.e., monitoring a network or a host activity logs) and their detection approach (i.e., signature-based or anomaly-based detection). Deep learning models, usually referred to as anomaly-based intrusion detection techniques, are in general neural network models with a large number of hidden layers. These models can learn extremely complicated non-linear functions, and their hierarchical layer structure allows them to acquire meaningful feature representations from incoming data. Researchers have explored deep learning techniques for in-vehicle intrusion detection on Controller Area Network (CAN) bus protocol since 2015 [12] [13]. However, only few work has been addressed to study the efficiency of developing deep learning based IDS for systems employing automotive ethernet due to the lack of relevant and public datasets. Among them, Alkhatib et al. [11] proposed a deep learning-based sequential model for offline intrusion detection on Scalable Service-Oriented Middleware over IP (SOME/IP) application layer protocol on top of automotive ethernet. Moreover, Jeong et al. [25] presented an intrusion detection method for detecting audio-video transport protocol (AVTP) stream injection attacks in automotive ethernet-based networks.

In this paper, we propose an anomaly detection based IDS Convolutional Autoencoders (CAE) for detecting anomalies on the AVTP protocol. The CAE consists of an encoder and a decoder with convolutional neural network (CNN) structures that are asymmetrical, i.e. the encoder architecture isn’t similar to the decoder architecture. Anomalies in AVTP packet stream...
are therefore detected by applying the reconstruction error of each sliding window of AVTP packets transformed to images as an anomaly score.

The main contributions of this paper are as follows:

- We propose a novel anomaly based intrusion detection method to detect unknown cyberattacks on AVTP protocol used in ethernet based in-vehicle network for media streaming. To the best of our knowledge, we are the first to leverage anomaly based approaches for intrusion detection on the recently adopted in-vehicle network protocol.
- We develop an IDS to evaluate the performance of our intrusion detection method by using a recently published real dataset called "Automotive Ethernet Intrusion Detection" [24]. Moreover, we have also created a new attack type, called "replay attack" from the normal packet traces of the published dataset.
- We have compared the performance of our model with other state-of-the art traditional anomaly detection and signature based models in machine learning for different attack types. The numerical results show that our proposed model outperform their performance and excel at predicting unknown in-vehicle intrusions, with 0.94 accuracy. Moreover, our model has a very low level of false alarm and miss detection rates for different AVTP attack types.

Towards this end, our paper is organized into six Sections. In Section 2, we present an overview of media stream transportation using AVTP network protocol. Section 3 discusses the detection of network anomalies using convolutional autoencoder and relevant research work. In Section 4, we present the considered AVTP dataset and the different considered attacks. The process for building our suggested anomaly based architecture, CAE, its training and its evaluation are presented in Section 5. We discuss our experimental results in Section 6. Finally, we conclude our paper with future work direction.

II. TRANSMISSION OF MEDIA STREAMS USING AVTP

Traditional in-vehicle networks are mostly based on bus technology that can not keep up with the growing communication demands of the self-driving car. In fact, they cannot meet the in-vehicle network requirements for high bandwidth, reliability and real-time communication expectations. Automotive Ethernet, a novel in-vehicle network communication technology, is implemented to ensure an appropriate level of quality of service (QoS) which is essential for time-critical automotive applications.

Audio Video Bridging (AVB) over Ethernet, a set of technical standards, provides improved synchronization, low-latency, and reliability for switched Ethernet networks between multimedia devices. Recently, a lot of automotive products such end-nodes device (i.e., speakers, cameras, digital signal processors) and network hub (i.e., AV Bridges) support it. In fact, end-nodes can be a talker, listener or both. A talker is the transmitter of a data stream or the source of the AVB system and a listener is the receiver of the destination of the AVB system. These end-nodes are connected by an AVB Bridge which acts as a switch that receives time-critical data from the AVB talker and forwards it to the AVB listener. This interconnection between these three components, as presented in Fig. 1, is called AVB Ethernet Local Area Network (LAN).

As previously mentioned, AVB has diverse sub-standards to support time-critical in-vehicle applications such as IEEE 802.1 Qav, IEEE 802.1 Qat, IEEE 802.1 AS and IEEE 1722. In our work, we are only considering automotive cyberattacks on IEEE 1722 which is a stream transmission protocol in charge of transporting control data and audio and video streams. As depicted in Fig. 2, the IEEE 1722 packet and its content are sent through an Ethernet frame. The IEEE 802.1Q header is also included in the Ethernet packet. Furthermore, the priority information encapsulated within is critical for the functioning of AVB QoS concept. Moreover, only AVB listener members that share the same AVB talker’s VLAN tag can receive the audio/video stream. In the case of The IEEE 1722, the ethertype field’s hexadecimal value is 0X22F0.

In terms of IEEE 1722 streaming packets, the header, the stream ID, the "Presentation time," payload information, and the payload itself are all included therein. The data type of the A/V stream is specified in the header which also includes its sequence number needed by AVB listeners to detect missing packets. The MAC address of the talker is used to produce the stream ID, which identifies a single data stream. The format of the data within the payload is directly related to the field of payload information. The AVBTP is a time presentation which
III. STATE-OF-THE-ART ANOMALY-BASED IDS USING CONVOLUTIONAL AUTO-ENCODER

Intrusion Detection Systems (IDSs) are considered as an efficient tool to guarantee the confidentiality, integrity and availability of network data. In fact, network intrusions can be detected and identified by comparing their attack signature to a dataset which contains a pre-defined list of cyberattack patterns. This approach is called signature-based intrusion detection. However, a regular updating of signature databases is not practicable because of the constant evolution of innovative attack tactics. An alternative solution could be the adoption of anomaly-based IDSs which find pattern in the data that deviates from other observation and indicates the presence of malicious activities in the network traffic.

Deep learning techniques are increasingly used to address the development of complex anomaly detection based IDSs [15]. One of the most commonly studied feature learning techniques is the use of autoencoder (AE) neural networks which can be adapted to detect anomalies in high-dimensional data and for different data types including images/videos, sequence data and graph data. In this work, we will leverage convolutional auto-encoders (CAE) to detect anomalies in an AVTP stream transformed into images. We describe the adopted method in Section IV.

The autoencoder AE, introduced by Rumelhart et al. [16], seeks to learn a low-dimensional feature representation space suitable for reconstructing the provided data instances. During the encoding process, the encoder maps the original data onto low-dimensional feature space, while the decoder tries to retrieve the original data from the projected low-dimensional space. Reconstruction loss functions are used to learn the parameters of the encoder and decoder networks. Its reconstruction error value must be minimized during the training of normal instances and therefore used during testing as an anomaly score. In other words, compared to the typical data reconstruction error, anomalies that differ from the majority of the data have a large data reconstruction error. The following equations govern the behavior of an AE:

$$z_k = \phi_c(x_R; \Theta_c), \forall R \in \{1, ..., p\}. \quad (1)$$

$$\hat{x}_R = \phi_d(z_k; \Theta_d), \forall K \in \{1, ..., m\}, \forall m \leq p. \quad (2)$$

$$\{\Theta_c^*, \Theta_d^*\} = \arg\min_{\Theta_c, \Theta_d} \frac{1}{n} \sum_{k=1}^{n} \|x_R - \phi_d(\phi_c(x_R; \Theta_c); \Theta_d)\|^2 \quad (3)$$

$$s_x = \|x_R - \phi_d(\phi_c(x_R; \Theta_c^*); \Theta_d^*)\|^2 \quad (4)$$

where n is the number of samples, $\phi_c$ is the encoding network with the parameters $\Theta_c$, $\phi_d$ is the decoding network with the parameters $\Theta_d$, and $s_x$ is a reconstruction error-based anomaly score to $x$, known as the mean squared error [15].

In addition, the convolutional auto-encoder CAE design employs convolutional and pooling layers for encoding and deconvolutional and unpooling layers for decoding. Due to the weight-sharing, CAE is capable of absorbing localized frequent patterns in a 2D structure [21]. Therefore, researchers have been widely using them for the anomaly detection of concrete defects [21], in automated video surveillance [18], on system logs [19], and on network application protocols such as HTTP [20].

IV. DATASET DESCRIPTION

In order to detect intrusions on Audio Video Transport Protocol (AVTP), we have used the "Automotive Ethernet Intrusion Dataset" dataset [24] created by Jeong et al. [25] and which contains benign and malicious AVTP packet captures from their physical automotive Ethernet testbed.

| Dataset                      | # Normal packets | # Injected packets | Size   |
|------------------------------|------------------|--------------------|--------|
| Indoor Original 1           | 139,440          | N/A                | 63.3 MB|
| Indoor Injected 1           | 139,440          | 65,988             | 93.3 MB|
| Indoor Injected 2           | 307,020          | 130,906            | 198.8 MB|
| Replay attack (new)         | 139,440          | 78,880             | 99.1 MB|

The datasets are recorded in the PCAP file format and, therefore, are viewed using prevalent programming libraries and packet analyzers (such as Wireshark). In fact, the dataset contains four benign (attack-free) packet captures and four malicious ones collected in different environments. The malicious packet captures represent the message injection of arbitrary stream AVTP data units (AVTPDUs) into the IVN since the attacker’s goal is to output a single video frame, at a terminal application connected to the AVB listener, by injecting previously generated AVTPDUs during a certain period. For our experiment, we have only considered the AVTP packets collected indoor, presented in Table I. Furthermore, we have created a new type of attack using Python [26], called replay attack, which represents an attacker injecting previously sent legitimate AVTP packets, as seen in Table I.

A. Dataset Preprocessing

We have pre-processed the AVTP dataset, provided in pcap format, using Python programming language [26] and its efficient libraries.

- **Step 1 (AVTP packet labeling):** In order to evaluate the performance of our proposed IDS, the AVTP dataset must be labeled. In order to differentiate if a certain AVTP packet is injected or normal. Hence, we have used the packet crafting library Scapy [27] for decoding and labeling. A packet is labeled as injected if it belongs to the "single-MPEG-frame.pcap" pcap file, otherwise it is considered normal.
• Step 2 (Sliding window process): After labeling the AVTP dataset, it needed to be properly reshaped into images for use in our convolutional autoencoder-based IDS. Each packet in our dataset has 438 bytes/features, each of which has a decimal value between 0 and 255. After tuning, the most suitable number of bytes used to detect anomalies is the first 48 bytes of each AVTP packet. We then used Feature-based Sliding Window (FSW) \([28]\) to partition all data records into square windows, where each window is a 48x48x1 three-dimensional matrix and the slide size is 1, as presented in formula \((5)\).

\[
SW_k = \begin{pmatrix}
  a_{k,1} & a_{k,2} & \cdots & a_{k,48} \\
  a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{k+48,1} & a_{k+48,2} & \cdots & a_{k+48,48}
\end{pmatrix}
\]

\((5)\)

We use \(SW_k\) to denote \(k\)th sliding window, \(a_{m,n}\) is an AVTP packet feature such that \(0 \leq a_{m,n} \leq 255\) and \(k \leq m \leq k + 48, 1 \leq n \leq 48\).

• Step 3 (Sliding window labeling): After grouping our data records into sliding windows, we had to label each AVTP window by “1” if it contains any injected or replayed packets, otherwise it is labeled as “0”.

After pre-processing, the AVTP dataset, presented in Table II is ready to be leveraged for anomaly detection. Since we are training our anomaly detector through semi-supervised learning approach, the dataset “Indoor Original 1” with normal data records was used for training while datasets “Indoor Injected 1” and “Indoor Injected 2” and “Replay attack” were used for testing. For reproducibility, readers can access the pre-processed datasets, provided in pickle format on our github repository \([29]\).

| Dataset                | # Normal SW | # Injected SW | Size     |
|------------------------|-------------|---------------|----------|
| Indoor Original 1      | 139,393     | N/A           | 2.6 GB   |
| Indoor Injected 1      | 49,377      | 156,004       | 3.8 GB   |
| Indoor Injected 2      | 106,822     | 331,057       | 8.1 GB   |
| Replay attack (new)    | 256         | 268,017       | 5.1 GB   |

### V. PROPOSED ANOMALY-BASED CAE

The purpose of this work is to investigate the usefulness of a CAE’s reconstruction error as an anomaly detector for AVTP packet captures. Firstly, AVTP packet captures and features are extracted from the AVTP dataset, as described in section IV which are then fed as input data to the CAE. A CAE is trained on normal events, and the AVTP packets’ frame reconstruction error (FRE) is used to classify both normal and abnormal events. Following that, the Balanced Accuracy, False Positive Rate (FPR) and False Negative Rate (FNR) are used to quantify classification performance.

\[MFRE_{normal} = \frac{1}{n \times 48 \times 48} \sum_{i=1}^{n} \sum_{j=1}^{48} \sum_{k=1}^{48} (x_{ijk} - y_{ijk})^2\]  

\((6)\)

where \(n\) represents the total number of normal instances used for training, \(x_{ijk}\) and \(y_{ijk}\) represent the value of the feature.
\( k \in \{1, \ldots, 48\} \) for the packet \( j \in \{1, \ldots, 48\} \) in the original frame \( i \in \{1, \ldots, n\} \) and the reconstructed frame, respectively. To optimize the loss function, we use Adam optimizer [30].

### B. Evaluation metrics

For measuring the performance of our anomaly based IDS, we use the balanced accuracy (Balanced Acc), the false positive rate (FPR), and the false negative rate (FNR).

The Balanced Accuracy (Balanced Acc) is usually used in binary classification, especially when dealing with imbalanced data. It’s the arithmetic mean of true positive rate (TPR) and true negative rate (TNR).

\[
\text{Balanced Acc} = \frac{TPR + TNR}{2} \tag{7}
\]

The True Positive Rate (TPR) measures the proportion of correctly predicted anomalies out of total AVTP anomaly prediction made by the model.

\[
TPR = \frac{TP}{TP + FN} \tag{8}
\]

The True Negative Rate (TNR) measures the proportion of correctly predicted normal AVTP window frames out of total normal AVTP window frame prediction made by the model.

\[
TNR = \frac{TN}{TN + FP} \tag{9}
\]

The false positive rate (or “false alarm rate”) usually refers to the expectancy of the false positive ratio. In our case, it represents the ratio of normal AVTP window frames predicted as abnormal.

\[
FPR = \frac{FP}{TN + FP} \tag{10}
\]

The false negative rate (or “miss rate”) represents the ratio of abnormal AVTP window frames predicted as normal frames.

\[
FNR = \frac{FN}{TP + FN} \tag{11}
\]

Where: TP = True Positive; FP = False Positive; TN = True Negative; FN = False Negative.

Hence, the model is assumed to have a large predictive power in detecting cyberattacks on AVTP protocol if the balanced accuracy is near to 1.0 and its false positive and negative rates approach zero.

Additionally, in order to determine the most suitable threshold for discriminating between normal and abnormal AVTP frames, multiple thresholds were tested, using the MFRE distribution from normal instances as a reference to identify the threshold that achieves the best performance. The experiments include thresholds based on the mean frame reconstruction error (MFRE) \((\mu_{FRE})\) and \(k\) standard deviations \((\sigma_{FRE})\) from normal examples. In our case, \(k = \{1, 2\}\), therefore, the following thresholds were considered: \(\theta_1 = \mu_{FRE}\), \(\theta_2 = \mu_{FRE} + \sigma_{FRE}\), \(\theta_3 = \mu_{FRE} + 2\sigma_{FRE}\).

### VI. Results

The CAE model proposed in this work is trained using Keras [31], a deep learning framework built on top of TensorFlow 2 [32]. All experiments were carried out on a dedicated GPU server Tesla V100S PCIe 32GB, powered by NVIDIA Volta™, with 32GB of RAM.

For anomaly detection, the threshold used to identify anomalies has a significant impact on model performance, including balanced accuracy, FPR, and FNR. In this paper, we optimize the model in terms of thresholds in order to obtain the model’s optimal performance. As depicted in Table IV, the model’s balanced accuracy, FPR and FNR varies when modifying the threshold \(\theta\). It has the highest balanced accuracy (0.94 in average) and false positive and negative rates approaching zero value, for different attack types when the threshold value is \(\theta = \mu + \sigma\). In other words, our proposed model classifies well the normal and abnormal AVTP packet frames.

In addition, we compare our proposed CAE to other machine-learning based anomaly detection algorithms such as the Isolation Forest [33] and the Local Outlier Factor [34]. Moreover, we also compare it to a reproduced version of the CNN model proposed by Jeong et al. [25] and which is trained using supervised learning approach with a normal and abnormal "Indoor Injected 1” packet trace, through 3-fold cross-validation. It is observed, in Table IV, that the Isolation Forest and Local outlier Factor models, usually used for anomaly detection, had a poor performance in capturing the spatio-temporal information. As seen, they were unable to classify well the different classes and had in average high false negative and positive rates. However, the CNN model

| Model                       | Indoor Injected 1 | Indoor Injected 2 | Replay Attack |
|-----------------------------|-------------------|-------------------|---------------|
|                             | Balanced Acc  | FPR    | FNR    | Balanced Acc | FPR    | FNR    | Balanced Acc | FPR    | FNR    |
| CAE, \( \theta = \mu \) (proposed) | 0.87  | 0.25 | 0.008 | 0.87  | 0.25 | 0.007 | 0.85  | 0.01 | 0.28 |
| CAE, \( \theta = \mu + \sigma \) (proposed) | 0.86  | 0.25 | 0.02  | 0.87  | 0.25 | 0.02  | 0.77  | 0.25 | 0.21 |
| CAE, \( \theta = \mu + 2 \sigma \) (proposed) | 0.99  | 0  | 0.04  | 0.98  | 0  | 0.04  | 0.85  | 0.01 | 0.27 |
| Isolation Forest [33]       | 0.57  | 0.34 | 0.51  | 0.58  | 0.6 | 0.45  | 0.52  | 0.85 | 0.5  |
| Local Outlier Factor [34]   | 0.28  | 0.64 | 0.8   | 0.57  | 0 | 0.85  | 0.5   | 0  | 0.99 |
| CNN for AVTP [25]           | 0.82  | 0.32 | 0.004 | 0.82  | 0.32 | 0.005 | 0.73  | 0.34 | 0.19 |

**TABLE IV**: COMPARISON OF CAE WITH OTHER MODELS
significantly achieves better performance and reasonable balanced accuracy than the traditional algorithms. Moreover, our model had a 12% gain in balanced accuracy over the CNN model in detecting the replay attack, which wasn’t used for its training. Hence, our reconstruction-based model, trained using a semi-supervised approach, has proved to be efficient in detecting different and unknown cyberattacks in an AVTP communication.

VII. CONCLUSIONS AND FUTURE WORK

Anomaly detection in in-vehicle network protocols, especially in automotive ethernet, is a burgeoning study area. With the development of realistic datasets which represents automotive cyberattacks on this protocol, we are able to develop anomaly-based detection models using deep learning and to evaluate them. In this paper, we suggested a CAE architecture for learning normal AVTP communication behavior to identify cyberattacks on this protocol. The numerical results show that our proposed model outperform the state-of-the art machine learning based model performance and excel at predicting unknown in-vehicle intrusions, with balanced accuracy value approaching 0.94 and low false positive and negative rates for different attacks and packet traces. For future work, we aim to test the performance of diverse sequential models such LSTM and Transformer on different cyberattacks types applied on AVTP protocol and in real-time scenarios.

REFERENCES

[1] Lu, Meng, et al. "ADAS applications for improving traffic safety." 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583). Vol. 4. IEEE, 2004.

[2] Parada-Loira, Francisco, Elisardo González-Aguila, and José L. Alba-Castro. "Hand gestures to control infotainment equipment in cars." 2014 IEEE Intelligent Vehicles Symposium Proceedings. IEEE, 2014.

[3] Tashev, Ivan, et al. "Commute UX: Voice enabled in-car infotainment system." (2009).

[4] P. Hank, S. Müller, O. Vermees and J. Van Den Keybus, "Automotive Ethernet: Invehicle networking and smart mobility," 2013 Design, Automation and Test in Europe Conference & Exhibition (DATE), 2013, pp. 1735-1739, doi: 10.7873/DATEx.2013.349.

[5] Bello, Lucia Lo. "The case for ethernet in automotive communications." ACM SIGBED Review 8.4 (2011): 7-15.

[6] "IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks," in IEEE Std 1722-2016 (Revision of IEEE Std 1722-2011) , vol., no., pp.1-233, 16 Dec. 2016, doi: 10.1109/IEEESTD.2016.7782716.

[7] Migge, Jörn, et al. "Insights on the Performance and Configuration of AVB and TSN in Automotive Ethernet Networks." Proc. Embedded Real-Time Software and Systems (ERTS 2018) (2018).

[8] Khan, Shah Khalid, et al. "Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions." Accident Analysis & Prevention 148 (2020): 105837.

[9] Ethernet Security. https://www.etas.com

[10] D. Wang and S. Ganesan, "Automotive Network Security," 2021 IEEE International Conference on Electro Information Technology (EIT), 2021, pp. 193-196, doi: 10.1109/EIT51626.2021.9491889.

[11] Alkhatib, Natasha, Hadi Ghauch, and Jean-Luc Danger. "SOME/IP Intrusion Detection using Deep Learning-based Sequential Models in Automotive Ethernet Networks." arXiv preprint arXiv:2108.08262 (2021).

[12] M. Kang and J. Kang, "A Novel Intrusion Detection Method Using Deep Neural Network for In-Vehicle Network Security," 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), 2016, pp. 1-5, doi: 10.1109/VTCSpring.2016.7504089.

[13] A. Taylor, S. Leblanc and N. Japkowicz, "Anomaly Detection in Automobile Control Network Data with Long Short-Term Memory Networks," 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2016, pp. 130-139, doi: 10.1109/DSAA.2016.20.

[14] Matheus, Kirsten, and Thomas Königseder. Automotive ethernet. Cambridge University Press, 2021.

[15] Chalapathy, Raghavendra, and Sanjay Chawla. "Deep learning for anomaly detection: A survey." arXiv preprint arXiv:1901.03407 (2019).

[16] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." nature 323.6086 (1986): 533-536.

[17] Chow, Jun Kang, et al. "Anomaly detection of defects on concrete structures with the convolutional autoencoder." Advanced Engineering Informatics 45 (2020): 101105.

[18] Ribeiro, Manassés, André Eugênio Lazzaretti, and Heitor Silvério Lopes. "A study of deep convolutional auto-encoders for anomaly detection in videos." Pattern Recognition Letters 105 (2018): 13-22.

[19] Y. Cui, Y. Sun, J. Hu and G. Sheng, "A Convolutional Auto-Encoder Method for Anomaly Detection on System Logs," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2018, pp. 3057-3062, doi: 10.1109/SAMC.2018.00519.

[20] S. Park, M. Kim and S. Lee, "Anomaly Detection for HTTP Using Convolutional Autoencoders," in IEEE Access, vol. 6, pp. 70884-70901, 2018, doi: 10.1109/ACCESS.2018.2881003.

[21] Masci, Jonathan, et al. "Stacked convolutional auto-encoders for hierarchical feature extraction." International conference on artificial neural networks. Springer, Berlin, Heidelberg, 2011.

[22] Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." Proceedings of the 25th international conference on Machine learning. 2008.

[23] Krizhevsky, Alex, and Geoffrey E. Hinton. "Using very deep autoencoders for content-based image retrieval." ESANN. Vol. 1. 2011.

[24] Seonghoon Jeong, Boosun Jeon, Boheung Chung, Huy Kang Kim, February 4, 2021. "Automotive Ethernet Intrusion Dataset", IEEE DataPort, doi: https://dx.doi.org/10.21227/1y3r-q009.

[25] Jeong, Seonghoon, et al. "Convolutional neural network-based intrusion detection system for AVTP streams in automotive Ethernet-based networks." Vehicular Communications 29 (2021): 100338.

[26] Python. https://www.python.org.

[27] Scapy. https://scapy.net.

[28] X. Zhang, X. Cui, X. Cheng and L. Zhang, "A Convolutional Encoder Network for Intrusion Detection in Controller Area Networks," 2020 16th International Conference on Computational Intelligence and Security (CIS), 2020, pp. 366-369, doi: 10.1109/CISS2066.2020.00084.

[29] AVTP IDS. https://github.com/Alkhatibnatasha/AVTP-IDS

[30] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

[31] Keras https://keras.io.

[32] Tensorflow https://www.tensorflow.org.

[33] Liu, Fei Tony, Ting, Kai Ming and Zhou, Zhi-Hua. “Isolation-based anomaly detection.” ACM Transactions on Knowledge Discovery from Data (TKDD) 6.1 (2012): 3.

[34] Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000, May). "LOF: identifying density-based local outliers. In ACM sigmod record. ACM SIGMOD Review 8.4 (2011): 7-15.

[35] Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." Proceedings of the 25th international conference on Machine learning. 2008.