IDS in the CAN bus using Statistical Analysis and Neural Networks

Araya Kibrom Desta, Ismail Arai, Kazutoshi Fujikawa
Nara Institute of Science and Technology, Inet-Lab

1. Introduction

Present-day vehicles are equipped with multiple Electronic Control Units(ECUs), each of which communicate with one another using a protocol called Controller Area Network(CAN).
Even though, CAN provides its own share of benefits in modernizing automobiles, it also opens a new security hole in the automotive industry. CAN bus doesn't use any mechanism for encrypting or authenticating CAN packets. Security researchers have been able to exploit this security hole to remotely control some critical car components. As a counter measure against this drawback two methodologies of defense, Prevention and Detection, have been proposed. But due to the low processing power of ECUs and a desire of unaltering the CAN de facto standard, we are mainly focusing on a mechanism to detect intrusions inside the CAN bus.

2. Purpose of the Research

The main purpose of this research is to detect cyber attacks inside CAN bus by using different statistical anomaly detection methods and Long Short Term Memory(LSTM) Recurrent Neural Networks(RNN). We will train a neural network to predict subsequent packets, using data from sequences of previously seen messages on the CAN bus. The error difference found by evaluating the actual value and the predicted value will be used for detecting anomalies in a sequence of CAN Packets.

Each CAN packet, along side with other information, has an arbitration ID and timing information. Using this two information and the fact that CAN packets appear in the CAN bus at a fixed frequency[1], we aim to detect malicious message sequences in a fixed time window. Given the available information from the CAN bus and knowledge of attack signatures, we have evaluated some statistical methods that can effectively identify CAN attacks.

3. Replication of existing Researches

To evaluate the methods, we have collected 10 minutes of CAN data from a real car.

1. <u>One Sample/Universal t-test method[2]</u>

This anomaly Detection approach works by calculating statistical data about on going network traffic and comparing them with historical values (μ_{Ht}) measured during training. In every 1 second, we collected the following information for each arbitration ID.

- ✓ ID: arbitration ID
- ✓ N_p : the number of packets in the flow
- ✓ μ_t: the average time difference between successive packets
- ✓ σ^2_t: the variance of the time difference between successive packets
- ✓ T_t: the time difference test value

$$T_t = \frac{\mu_t - \mu_{Ht}}{\sqrt{\frac{\sigma^2}{N_p}}}$$

After we calculated *Tt*, the corresponding p value is solved , and if the p-value is less than a predefined threshold(0.26), the detector alerts the driver to take appropriate measures.

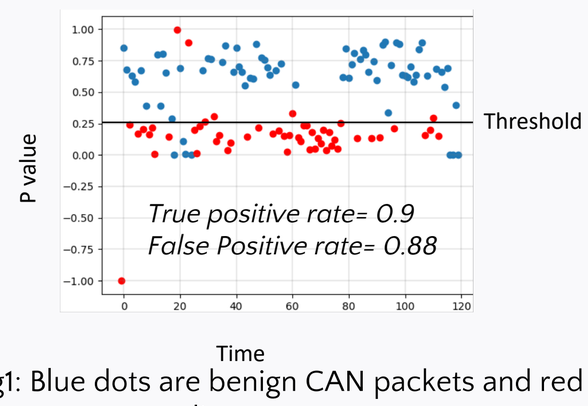

Fig1: Blue dots are benign CAN packets and red dots represent anomalies.

<u>2. IDS using One Class Support Vector Machine[2]</u>,
we trained OCSVM against N_p, μ_t and σ^2_t

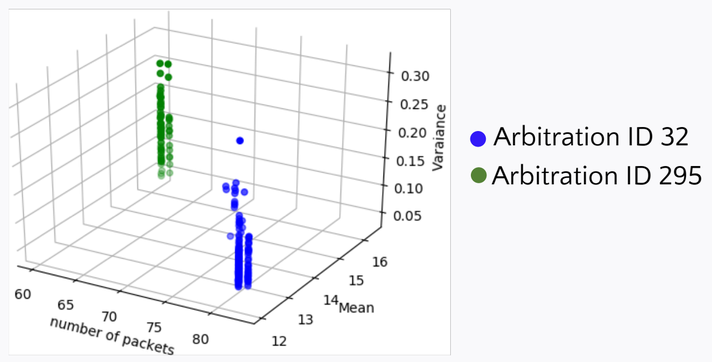

Fig2: data view for IDs 32 and 295

The training gave us good results for both insertion and drop attacks.

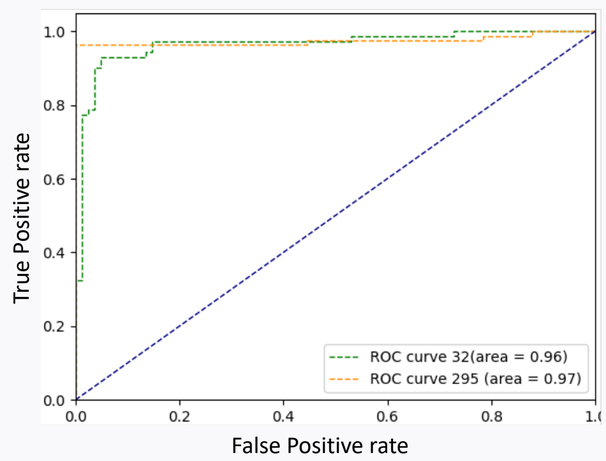

Fig3: ROC curve for arbitration IDs 32 and 295

<u>References</u>

[1] Song, **Intrusion detection system based on the analysis of time intervals of CAN messages for in-vehicle network**. International Conference on Information Networking. Vol. 2016-March IEEE Computer Society, 2016. pp. 63-68
[2] A. Taylor, N. Japkowicz, and S. Leblanc, "Frequency-based anomaly detection for the automotive CAN bus," in Proc. 2015 World Congress on Industrial Control Systems Security (WCICSS), Dec. 2015, pp. 45–49.

4. Discussion

OCSVM method is more efficient in detecting very short insertion and drop attacks at an acceptable rate, it is even possible to get more practical positive false alarm rates with a higher training data compared to t-test method. In case of the t-test method, its performance can be improved by selecting an optimal threshold value. The training data we used to experiment all the methods are only periodic CAN messages, but all the aforementioned methods fail to detect any anomaly sent with a non-periodic arbitration ID.

5. Future Work

Most of the methods described here mainly focus on periodicity and timing information of CAN packets. None of the methods used any information from the data portion of the packets. And we believe the optimal time that a user should be notified about an intrusion should be in about 1 second. But, for periodic messages which appear in the CAN bus in average of later than 1 second it is impossible to notify the driver before the intrusions cause much more damage.

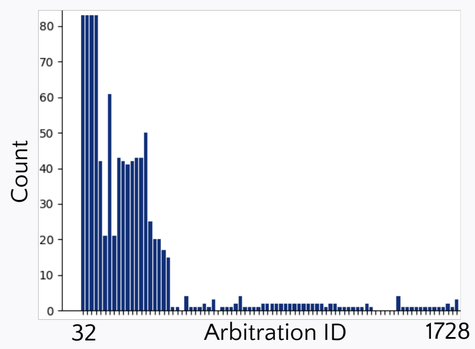

Fig3: Packet count of all arbitration IDs in one second.

In our research we are trying to improve some of the methodologies described here and we will also use sequences of CAN packet data portion to identify anomalies.

A simple strategy for general sequence learning is to use RNN with LSTM[3].

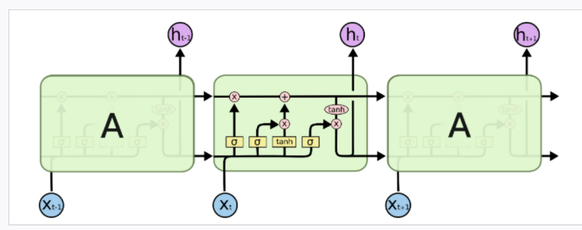

Fig5: Module in an LSTM[4]

With this technology, we will predict data frames using the trained network and depending on how close our prediction was with the actual received data frame, we will determine whether a data sequence has an anomaly or not.

The advantage of using RNN over the statistical methods is, unlike statistical methods RNN have a better anomaly detection capability for a short term anomalies. RNN will also be able to detect anomalies arriving in the CAN bus with non-periodic arbitration IDs.
Furthermore, We will continue to tweak LSTM for better accuracy results and try to continue digging on how to use RNN for identifying anomalies that appear during abnormal car states, like intrusion detection during car crashes. Intrusion detection during this state can be more difficult due to abrupt data packet information changes.

<u>References</u>

[3] F.A. Gers, J. Schmidhuber, and F. Cummins. Learning to forget: continual prediction with LSTM. Neural Computation, 12(10):2451–2471, 2000.
[4] http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Contact Information

Araya.kibrom_Desta.js3@is.naist.jp
{Ismail,fujikawa}@itc.naist.jp