SIL765: Networks and System Security

Semester II, 2021-2022

Assignment-4

March 25, 2022

Problem-1: Anomaly Detection (100 Marks)

An anomaly-based intrusion detection system (IDS) can detect attacks by detecting variations in the benign behavior. In this assignment, we will look into the Controller Area Network (CAN) bus communication which is utilized by different electronic components of a vehicle to coordinate their actions. For instance, for enabling the cruise (speed) control mode in a vehicle, the engine, the accelerator and the brake need to coordinate. The CAN bus provides a communication medium on which each electronic component broadcasts a message which is received and processed by all other electronic components.

However, an attacker can take control of an electronic component and inject malicious packets on the CAN bus. These malicious packets could be crafted to launch a Denial of Service (DoS) attack which prevents the regular communication on the CAN bus. The attacker could also try to impersonate another electronic unit and send the impersonated packets, e.g., the attacker could impersonate a brake unit and send a message which implies engaging the brake. In this problem, you will design two IDSs: one to detect a DoS attack and another to detect an impersonation attack on the CAN bus. For more details about how to detect such attacks, you can refer to the following paper. Also, please feel free to refer to any other reading material to understand the basics of CAN bus, e.g., the structure of a CAN packet and how packets are communicated on the CAN bus.

Reference

Michael R. Moore, et al., "Modeling inter-signal arrival times for accurate detection of CAN bus signal injection attacks: A data-driven approach to in-vehicle intrusion detection," Proceedings of the 12th Annual Conference on Cyber and Information Security Research, 2017.

Dataset

Please access the following Google drive link and download the two datasets (in the CSV format): one for a DoS attack (`DoS_dataset.csv`) and another for an impersonation attack (`gear_dataset.csv`).
`https://drive.google.com/drive/folders/1EBR8M_OXh7TBympzpaB0c0COHgq8vaoT?usp=sharing`.
These two datasets have been obtained from the following link.
`https://ocslab.hksecurity.net/Dataset/CAN-intrusion-dataset`.

Each dataset was obtained by logging the CAN traffic from a real vehicle while the corresponding attack was being performed. The data is organized in 12 columns as follows.

1. Timestamp (in seconds): This refers to the time at which the packet is recorded from the bus.

2. Message Identifier (in HEX): This indicates what the payload is about. In a benign environment, this identifier can be utilized to recognize the sender as well.

3. Data Length (Integer): This refers to the length of the payload in bytes. It can be any number between 0 to 8. Note that parsing the information in the following columns depend on this number.
4. Data[0] (in HEX): It refer to the first byte of the payload.
5. Data[1] (in HEX): It refer to the second byte of the payload.
6. Data[2] (in HEX): It refer to the third byte of the payload.
7. Data[3] (in HEX): It refer to the fourth byte of the payload.
8. Data[4] (in HEX): It refer to the fifth byte of the payload.
9. Data[5] (in HEX): It refer to the sixth byte of the payload.
10. Data[6] (in HEX): It refer to the seventh byte of the payload.
11. Data[7] (in HEX): It refer to the eighth byte of the payload.
12. Manually-Added Correct Label (T or R): T represents to an injected message while R represents a normal message.

Tasks

1. Find out how many different message identifiers have been recorded on this CAN bus.
2. Find out if the benign messages (with the label R) are periodic?
3. Following the above reference reading material, design the two IDSs which use the inter-message arrival time to detect the DoS attack and the impersonation attack. For each IDS, use any machine learning model which can perform the two-class (T and R) classification after training. Note that the actual payload data will not be utilized to detect any anomaly.
4. Divide the dataset into an appropriate training set and a validation set, and report the detection rate and the false alarm rate of the designed IDSs.

Submission

- `ids_dos`: This should contain the code corresponding to the IDS which detects the DoS attack.
- `ids_impersonation`: This should contain the code corresponding to the IDS which detects the impersonation attack.
- `readme`: This should be the pdf file containing all the necessary details about your solution. For instance, it should explain the steps to build and execute your code. It should have the screenshots of terminals to demonstrate that your code works as desired. It should contain discussions about the detection rate and the false alarm rate.