**A Novel Approach for Accurate State Prediction for Connected Automated Vehicles (CAVs) under Attack**

1. **Introduction**

**To be added…**

1. **Data**

For this exercise, we generated multiple vehicle paths with systems of 40 sensors which may be attacked or un-attacked and 1 protected sensor measuring x-position of the vehicle at each timestep along the path. These paths are simulated using SUMO. The x-position generated from SUMO for multiple paths are treated as protected sensors and the rest of the 40 sensors are generated by adding noise of a constant standard deviation to the protected sensors. Same random noise is also added to protected sensor. Multiple attack paths are generated with 25%, 50%, 75% and 90% attacked sensors to test performance of the proposed approach under different edge conditions. The way you generate the attacks is very complicated and not described. We may need to do this for any reports/papers.

1. **Proposed Approach**

The preliminary exercise showed that reconstruction-based anomaly detection methods like – LSTM Auto Encoders and TadGANs (Time series anomaly detection using Generative Adversarial Networks) are not alone sufficient to detect paths different from what as anomalous. Based on current understanding, such anomaly detection methods are unable to detect anomalies that have possible paths that fit the model the anomaly detection method has learned but which do not follow the actual path the tracked object follows. Thus we proposed an approach that can overcome the shortcomings of anomaly detection methods. Please find below a visual representation of the approach. This approach has four main components, the first being the sensor wise anomaly detection using methods like TadGAN and the three other are mentioned below–

* 1. **Sensor Attack Identification Model/ Agreement check (SAIT/Agreement Check)**

Using this we intend to find attacked sensors in a system of multiple sensors measuring the state of the vehicle. It is a two level of agreement check.

* + 1. Level 1: Difference from protected sensor check
    2. Level 2: Histogram comparison between un-attacked and attacked system
  1. **Attack imputation model (AIM)**

Using this, we intend to impute the attached sensor values with a more reasonable value which are aligned with an un-attacked system. Currently, we use SAITs model to impute the attacked sensor values.

* 1. **Final State Prediction Model (FSPM/fusion)**

Once we have identified the attacked sensor and impute its values with a more reasonable value, we then use it to predict the final state of the vehicle using all sensors (features) available.

A diagram of a system

Description automatically generated

1. **Study Objectives**

This study is primarily intended to test and evaluate the scenarios in which the proposed approach performs well and under what extreme conditions the proposed approach fails to accurately identify the next state of a given vehicle in a CAVs ecosystem. The primary, secondary and exploratory objectives of the analysis are listed below.

Are some of the tasks below repeated?

* 1. **Primary Objectives**

In alignment with what has been mentioned above, following are the primary objectives of the study.

* + 1. Evaluate the performance of proposed approach on multiple paths when vehicles are under attack (<=25% of sensors are attacked) by calculating RMSE of final state prediction with ground truth. The RMSE should be the average over many runs (10000) with different noise realizations.
    2. Compare the performance of proposed approach and anomaly detection models (ADM) on multiple paths when vehicles are under attack (<=25% of sensors are attacked) by comparing average RMSE of final state prediction from both approaches.
  1. **Secondary Objectives**
     1. Evaluate performance to find attacked sensors (Prob(decide attacked|unattacked) and Prob(decide unattacked|attacked)) using SAIM (level 1 + level 2) with ADM to find attacked sensors when 25%, 50%, 75% and 90% of sensors are attacked. Compare with ADM.
     2. Compare the performance of proposed approach and ADM on multiple paths when vehicles are under different levels of attacks (50%, 75% or 90% of sensors attacked) by comparing average RMSE of final state prediction from both approaches.
     3. Evaluate the performance of SAIM level 1 + Attack imputation model + Final State Prediction Model (FSPM) under different levels of attacks (25%, 50%, 75% or 90% of sensors attacked) with the ground truth. IS SAIM level 1 + Attack imputation model + Final State Prediction Model (FSPM) DIFFERENT FROM our approach (see 4.2.2)?
     4. Evaluate the performance of SAIM + Attack imputation model + Final State Prediction Model (FSPM) under different levels of attacks (25%, 50%, 75% or 90% of sensors attacked) with the ground truth.
     5. Evaluate performance of proposed approach (end-to-end) with a scenario where only un-attacked sensors are used to with Final State Prediction Model (FSPM). Also add comparison to Kalman filter and CRB.

The detailed endpoints for each primary objective and secondary objective can be found below.

* 1. **Primary Endpoint**
     1. Given paths with less than 25% of attacked sensors, predict the final state using the proposed approach and calculate difference in prediction over each timestep and each sensor w.r.t the true path. Do the same for all paths when un-attacked.
        1. Plot a histogram of differences for all timesteps from proposed approach predictions and un-attacked system. The comparisons must be for the same noise sequences.
        2. Perform a z-test to identify the statistical similarity between mean (what does mean here mean?) of differences obtained from proposed approach predictions and un-attacked system for multiple paths for all timesteps and sensors.
     2. Plot the averaged RMSE of predicted fused state over time and averaged RMSE over time of predicted fused state when only TadGAN (ADM) is used for anomaly detection, for 25% of sensors as attacked.
  2. **Secondary Endpoint**
     1. For different levels of attacks (25%, 50%, 75%, 90%) compare the accuracy, precision and recall of proposed approach (SAIM) and ADM to identify attacked sensor for multiple paths. Please find the table below.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Methods | Proposed Approach | | | | ADM (TadGAN) | | | |
| Percentage of attacked sensors | 25% | 50% | 75% | 90% | 25% | 50% | 75% | 90% |
| Accuracy |  |  |  |  |  |  |  |  |
| Precision |  |  |  |  |  |  |  |  |
| Recall |  |  |  |  |  |  |  |  |
| F1-score |  |  |  |  |  |  |  |  |

* + 1. Perform both 4.3.2. for higher level of attacks 50%, 75% and 90%.
    2. Plot the final fused path over time while using only SAIM level 1 of agreement check for different number of attacked sensors and compare it with actual path.
    3. Plot the final fused path over time while using only SAIM level 1 + level 2 of agreement check for different number of attacked sensors and compare it with actual path.
    4. Plot the final fused path over time while using end-to-end approach for different number of attacked sensors and compare it with the path obtained from the fusion of just un-attacked sensors.