Unsupervised Learning Algorithms for Denial of Service Detection in Connected Vehicles

Connected Cars Context
► 237 millions connected cars in the world in 2021 and over 400 millions by 2025.
► Global ITS (Intelligent Transportation System) market from $22.91 billion in 2021 to $42.80 billion in 2028
► Cybersecurity of the connected cars is highly critical.
► Denial of Service (DoS) attacks jam and flood the network by mass of malicious messages.

► Proposed DoS detection system architecture:
- CAM DoS attack generated by one or more malicious vehicles.
- CAM messages contain information (position, velocity, etc.).
- CAM messages are sent to a central or distributed servers for analysis.
- Clustering-based detection algorithms are applied to separate benign and malicious data.

DoS Clustering-based Model Detection
Several Clustering algorithms are applied:
► Centroid-based clustering (e.g. K-Means, CLARA)
► Density-Based clustering (e.g. DBSCAN)
► Probabilistic Model-based Clustering (e.g. GMM). EM (Expectation-Maximization) algorithm to estimate the parameters of GMM, it alternates between two steps:
- E step: compute the log likelihood expectation using the current parameter estimation.
- M-step : estimate of the parameters by maximizing the expected log likelihood found in the (E) step.

Results and Analysis
► Evaluation based on Vehicular Reference Misbehavior (VeReMi) dataset, containing the message logs for each vehicle in the simulation:
- Position, velocity, acceleration and heading vectors.
- Position noise, velocity noise, acceleration noise.
- Reception timestamp, transmission time, sender, ID message, etc.
► GMM produces the best results in terms of F-measure. It is approximately 95% for the three subsets (DoS (D1), random DoS (D2) and disruptive DoS (D3)), with high precision and recall values.