Identifying Novel CAN Bus Attack Types at Runtime

NC STATE UNIVERSITY Computer Science

Sumaita Sadia Rahman
ssrahman@ncsu.edu

Caio Batista de Melo
cbatist4@ncsu.edu

ICDM 2025
November 12–15, 2025
Washington DC, USA
IEEE International Conference on Data Mining

Problem & Motivation

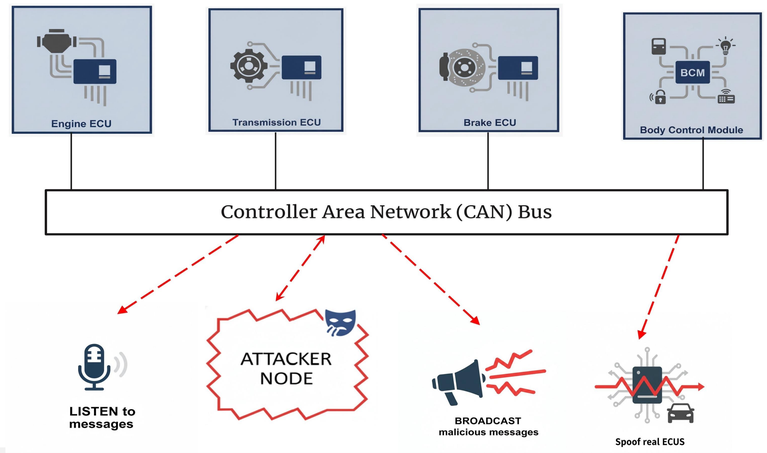


Fig 1: A simplified CAN bus schematic showing an attacker node's ability to listen to, broadcast, and spoof messages between ECUs.

- Messages on CAN bus do not have security features, and the network is vulnerable to attacks
- Significant work has focused on detecting intrusions on the CAN bus; however, few frameworks have considered post-detection classification
- Existing classification assume that system engineers know all possible attack types at design time. But **unfortunately, that is not always true**

Goal

Allow vehicles to identify novel CAN bus attack types

Our Approach

- Clustering-based metrics identify the number of different attack types in the dataset
- Our framework can detect a different number of types detected as more data is added
- Novelty signal: rerun clustering when a new attack block is added; if optimal clusters increase or silhouette score increases by a relevant threshold, we may have encountered a novel attack type

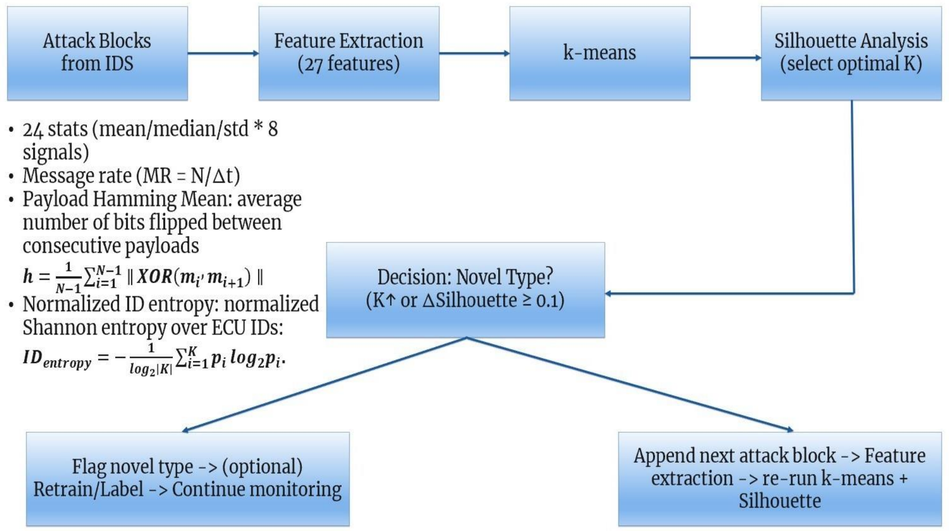


Fig 2: Runtime pipeline: attack blocks are summarized into 27 features, clustered, and re-clustered as new blocks are added; a novel attack type is indicated when the optimal number of clusters increases, or the Silhouette score increases by ≥ 0.1.

Clustering & Metrics

- **K-means** to group attack block
- **Silhouette**: values in [−1, +1] range; **0.7+ indicates strong clusters**
- **Adjusted Rand Index (ARI)**: values in [−1, +1] range; **+1 indicates identical** groupings, **0** indicates close to random groupings.
- **Data:** two literature datasets, Car-Hacking and Survival-IDS

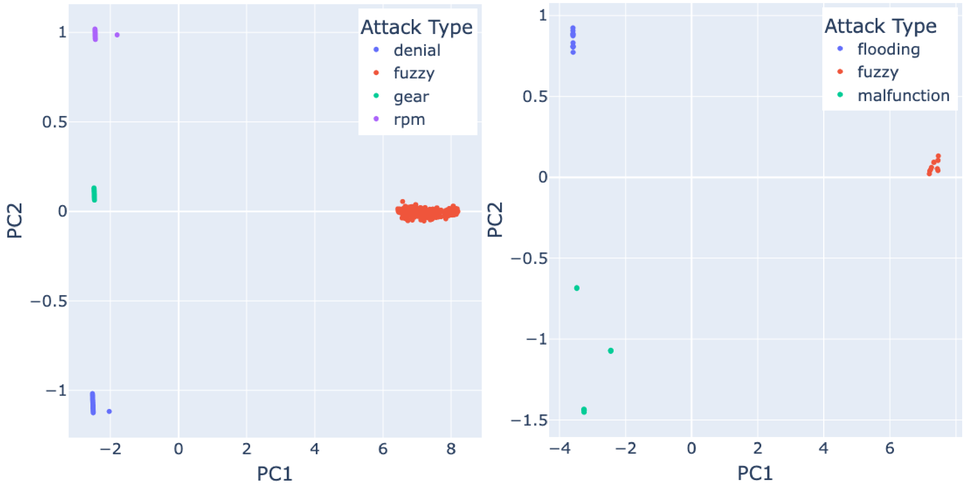
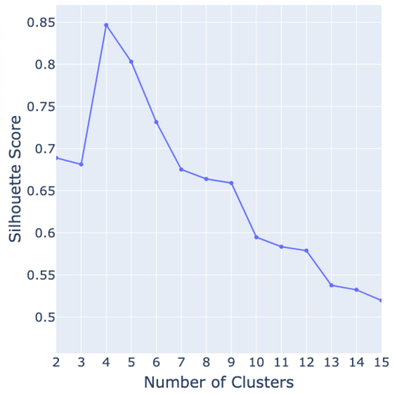


Fig 3: PCA visualization for Car-Hacking (left) and Survival-IDS (right)



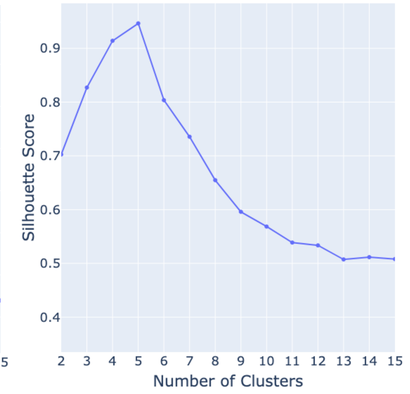


Fig 4: Silhouette Analysis for Car-Hacking (left) and Survival-IDS (right)

Results

Dataset/Attack	Clusters[1]	ARI[1]	Blocks Added	ARI[2]
Car-Hacking Denial of service	3	1.00	5	1.00
Fuzzy	4	0.99	1	1.00
Drive gear spoofing	3	1.00	19	1.00
RPM gauge spoofing	3	1.00	4	1.00
Survival-IDS Flooding	4	0.65	1	0.67
Fuzzy	4	0.72	2	0.75
Malfunction	2	1.00	1	1.00

Key Insights & Next Steps

- **Car-Hacking**: Perfect Grouping → ARI = +1
- **Survival-IDS:** Optimal K = 5 (vs. 3 types); *malfunction blocks were separated* ; with more data, malfunction points could be more concentrated
- **Novelty detection:** Method quickly identifies a novel attack type not seen previously, after at most two data points in most cases
- **Takeaway:** A clustering-based methodology that effectively (i) identifies when a novel attack type is encountered and (ii) groups same-type attacks.
- **Next steps:** Expand analysis to more datasets, connect with state-of-the-art detection system to evaluate latency and footprint at runtime