

Poster: Mind the Structural Gap - Evaluating Automotive CAN Datasets as Interchangeable Benchmarks for Intrusion Detection

Isha Pali
UQ-IITD Research Academy
i.pali@uq.edu.au

Naipeng Dong
University of Queensland
n.dong@uq.edu.au

Shubham Thakur
IIT Delhi
jtm242078@iitd.ac.in

Vireshwar Kumar
IIT Delhi
viresh@cse.iitd.ac.in

Vikas Maurya
IIT Delhi
vikasmauryaphd@gmail.com

Dan Dongseong Kim
University of Queensland
dan.kim@uq.edu.au

Abstract

Network intrusion detection systems (NIDS) proposed for the automotive CAN are routinely evaluated on public datasets under the implicit assumption that these datasets represent comparable notions of benign traffic. We challenge this assumption through a measurement study of six public CAN datasets: CH, OTIDS, CAN-MIRGU, ROAD, SynCAN, and X-CANIDS. Characterizing each along four structural dimensions ID space size, transition density, payload entropy, and inter-ID correlation, we find variation of up to $10\times$ in ID space and $2\times$ in transition density, and different structurally isolated clusters in terms of payload entropy and transition density. These findings show that public CAN datasets are *not* interchangeable benchmarks, and that evaluation results are strongly conditioned on dataset structure.

1 Introduction

Automotive Controller Area Network (CAN) lacks encryption and authentication, making in-vehicle networks vulnerable to spoofing, replay, fuzzing, and denial-of-service attacks. Consequently, network intrusion detection systems (NIDS) have become the primary defense mechanism in modern vehicles. CAN security research routinely evaluates NIDS using public datasets such as CH [5], OTIDS [3], CAN-MIRGU [4], ROAD [6], SynCAN [1], and X-CANIDS [2], often treating them as interchangeable benchmarks despite major differences in capture conditions and traffic characteristics.

In this work, we perform a structural measurement study of six public CAN datasets using four metrics: ID space size, transition density, payload entropy, and inter-ID correlation. Our analysis reveals substantial structural disagreement across datasets: ID space varies by up to $10\times$, transition density around $2\times$, and inter-ID correlation ranges from near-independent to moderately coupled traffic regimes. These findings demonstrate that public CAN datasets are *not interchangeable benchmarks*, and that NIDS evaluation results are strongly conditioned on dataset structure.

2 Proposed Metrics

We compute four properties from the *benign-only* traffic.

ID Space Size: This indicates unique CAN arbitration IDs present in each dataset. This may help determine whether there are issues with the message diversity/dimensionality within the dataset.

Transition Density (TD): Transition density measures the proportion of possible ID-to-ID transitions that are observed in the data, defined as $TD = |\mathcal{T}_{\text{obs}}|/N^2$, where $|\mathcal{T}_{\text{obs}}|$ is the number of observed transitions and N is the number of unique IDs. Higher TD values indicate a more densely connected transition network.

Payload Entropy (H): Shannon entropy of raw payload bytes within a time window, computed as $H = -\sum_i p_i \log_2 p_i$, where p_i is the probability of byte value i .

Inter-ID Correlation (ρ_{ID}): Mean Pearson correlation between message sequences of different IDs, computed as $\rho_{\text{ID}} = \frac{1}{|\mathcal{P}|} \sum_{(i,j) \in \mathcal{P}} \rho_{ij}$, where $\mathcal{P}$ is the set of ID pairs and ρ_{ij} is their Pearson correlation.

3 Results and Discussion

Table 1 summarises structural properties, data availability, and attack coverage across six widely used public CAN-Bus datasets.

ID Space Size: The number of unique CAN IDs varies widely across datasets, from just 10 IDs in SynCAN to 106 in ROAD. Larger ID spaces introduce richer communication patterns and more complex transition structures.

Transition Density and Inter-ID Correlation: Transition density (TD) ranges from 0.475 in CH to 0.850 in SynCAN. However, Figure 1 shows that TD alone does not fully describe CAN traffic structure. Datasets with similar TD can still exhibit very different inter-ID correlation (ρ_{ID}) patterns and communication behaviour. For example, OTIDS and ROAD both exhibit relatively high TD, yet differ substantially in terms of inter-ID correlation. OTIDS combines high TD with weak correlation, indicating permissive and weakly

Table 1: Comparison of public CAN Bus datasets across structural properties, data availability, and attack coverage. (Notations: † = float signals rather than raw payload bytes. ✓ = available/present, × = unavailable/absent, △ = partial availability, ✓* = questionable labelling, (s) = simulated injection.)

Dataset	Year	Type	Structural Properties				Data Availability				Attack Coverage					
			IDs	TD	H	ρ_{ID}	Sig.	Raw	DBC	Labels	DoS	Fuzz	Spoof	Replay	Masq.	Susp.
CH	2018	Real	27	0.475	5.08	0.357	×	✓	×	✓	✓	✓	✓	×	×	×
OTIDS	2017	Real	45	0.677	4.87	0.058	×	✓	×	△	✓	✓	×	×	✓*	×
MIRGU	2024	Real	56	0.503	5.29	0.539	×	✓	×	✓	✓	✓	✓	✓	✓(s)	✓(s)
ROAD	2024	Real	106	0.617	5.54	0.526	✓	✓	✓	✓	×	✓	✓	×	✓	×
SynCAN	2020	Synth.	10	0.850	6.71†	0.003	✓	×	×	✓	✓	×	✓	✓	×	✓
X-CANIDS	2023	Real	64	0.569	5.58	0.080	✓	✓	✓	✓	×	✓	×	✓	✓	✓

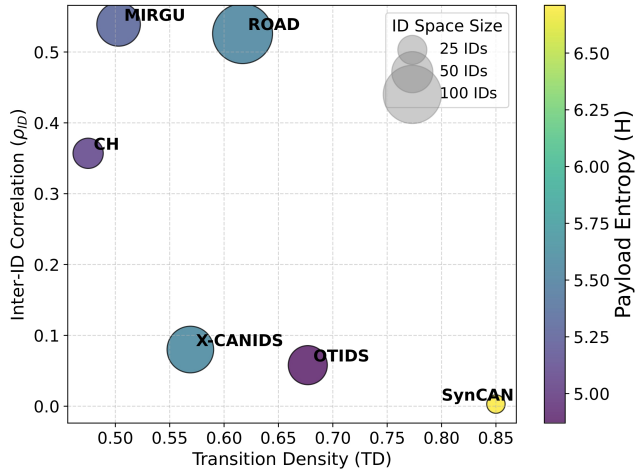


Figure 1: Structural diversity of public CAN Bus datasets.

constrained transitions, whereas ROAD and MIRGU exhibit much stronger inter-ID correlation, reflecting coordinated activity and tighter sequential structure. In contrast, SynCAN forms a clear outlier with extremely high TD but near-zero correlation, consistent with its synthetic and weakly coupled signal generation.

Payload Entropy (H): Payload entropy ranges from 4.87 bits in OTIDS to 6.71 bits in SynCAN. Higher entropy weakens statistical separation for entropy-based detectors. SynCAN reports unusually high entropy because it models floating-point signals rather than raw CAN payload bytes, placing it in a different data regime from real CAN traces. Table 1 further summarises metadata availability and attack coverage across datasets, revealing additional axes of incomparability

4 Conclusion and Future Work

We presented a structural analysis of six public CAN Bus NIDS datasets. The datasets exhibit substantial disagreement, including 10× variation in ID space, 2× variation in transition density, and large differences in message sequence behaviour and payload entropy. Our findings show that public CAN

datasets are not interchangeable benchmarks. Differences in traffic structure, metadata availability, attack coverage, and captured realism suggest that NIDS evaluated on different datasets may effectively be solving different detection problems. Going forward, we plan to carefully analyze how these structural differences translate to NIDS performance variance across datasets.

Use of AI Policy

AI tools were used for text refinement and code assistance. All ideas, analysis, and conclusions are original to the authors.

References

- [1] M. Hanselmann, T. Strauss, K. Dormann, and H. Ulmer. SynCAN dataset, 2020. Retrieved May 2026 from <https://github.com/etas/SynCAN/blob/master/README.md>.
- [2] S. Jeong, S. Lee, et al. X-canids: Signal-aware explainable intrusion detection system for controller area network-based in-vehicle network. *IEEE Transactions on Vehicular Technology*, 73:3230–3246, 2023.
- [3] H. Lee, S. H. Jeong, and H. K. Kim. OTIDS: A novel intrusion detection system for in-vehicle network by using remote frame. In *15th Annual Conference on Privacy, Security and Trust (PST)*, pages 57–5709. IEEE, 2017.
- [4] S. Rajapaksha, G. Madzudzo, et al. CAN-MIRGU: A comprehensive can bus attack dataset from moving vehicles for intrusion detection system evaluation. In *Symposium on Vehicle Security and Privacy (VehicleSec)*, San Diego, CA, USA, 2024. NDSS Symposium.
- [5] H. M. Song, J. Woo, and H. K. Kim. In-vehicle network intrusion detection using deep convolutional neural network. *Vehicular Communications*, 21:100198, 2020.
- [6] M. E. Verma, M. D. Iannacone, et al. ROAD: The real ornl automotive dynamometer controller area network intrusion detection dataset. *arXiv:2012.14600*, 2020.