Systematic Evaluation of Automotive Intrusion Detection Datasets

Arash Vahidi, Thomas Rosenstatter, Nishat I Mowla

thomas.rosenstatter@ri.se

RISE Research Institutes of Sweden

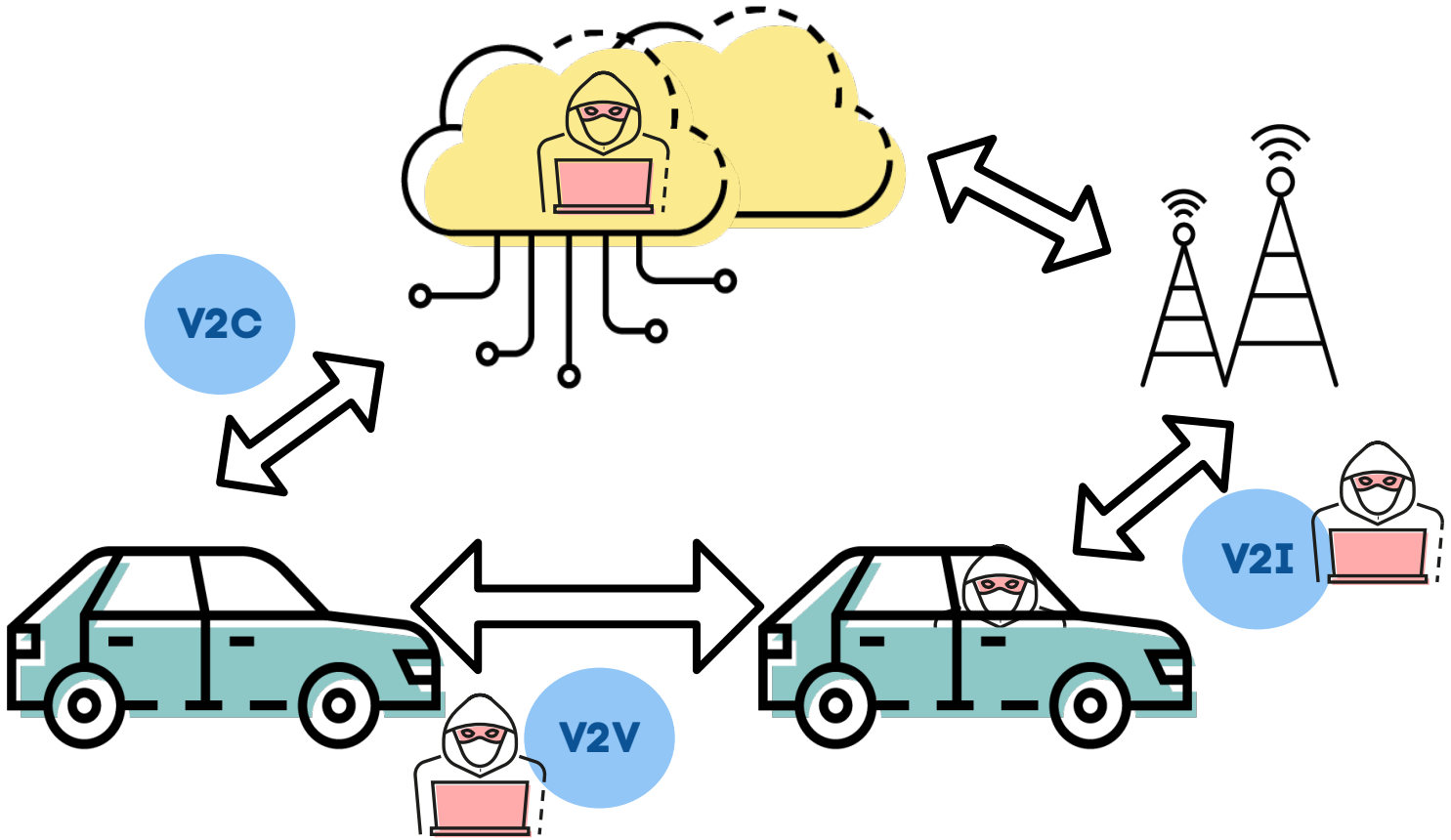
ACM Computer Science in Cars Symposium | CSCS 2022

RI. SE

CYRE V

Cyber Resilience for Vehicles

Vehicle connectivity



Motivation

5.1.1. The Approval Authority or the Technical Service shall verify by means of document checks that the vehicle manufacturer has taken the necessary measures relevant for the vehicle type to:

(a) Collect and verify the information required under this Regulation through the supply chain so as to demonstrate that supplier-related risks are identified and are managed;

(b) Document risks assessment (conducted during development phase or retrospectively), test results and mitigations applied to the vehicle type, including design information supporting the risk assessment;

(c) Implement appropriate cyber security measures in the design of the vehicle type;

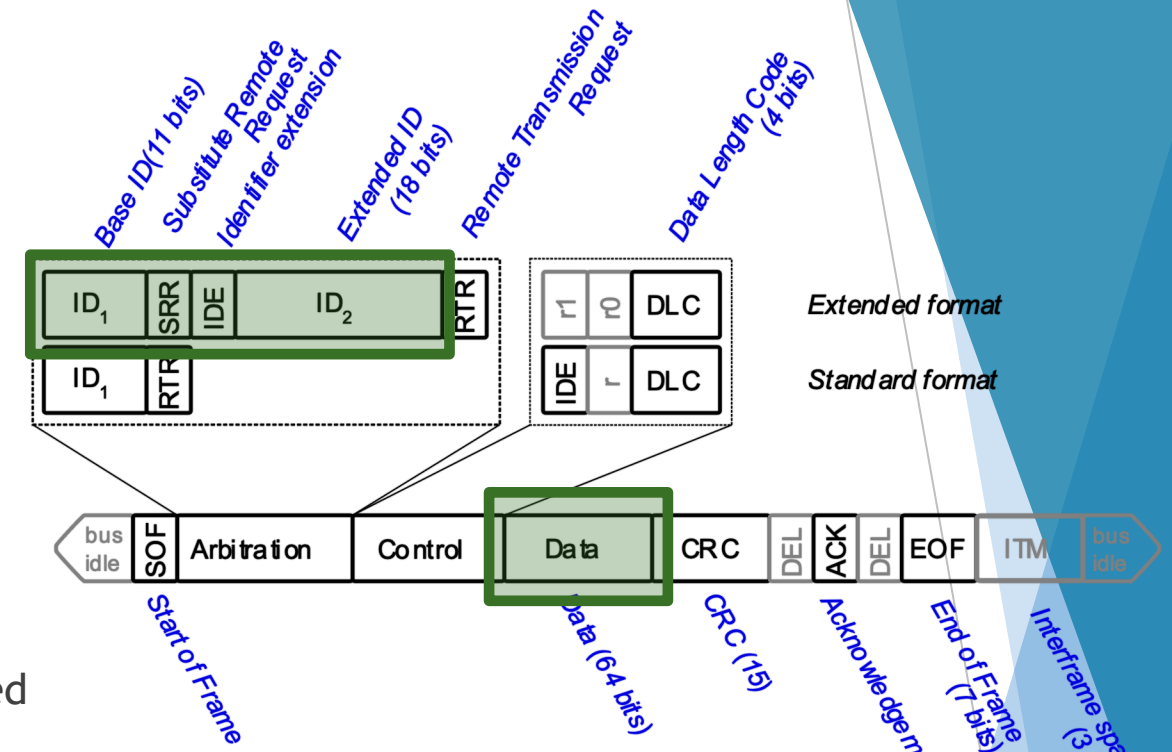
(d) Detect and respond to possible cyber security attacks;

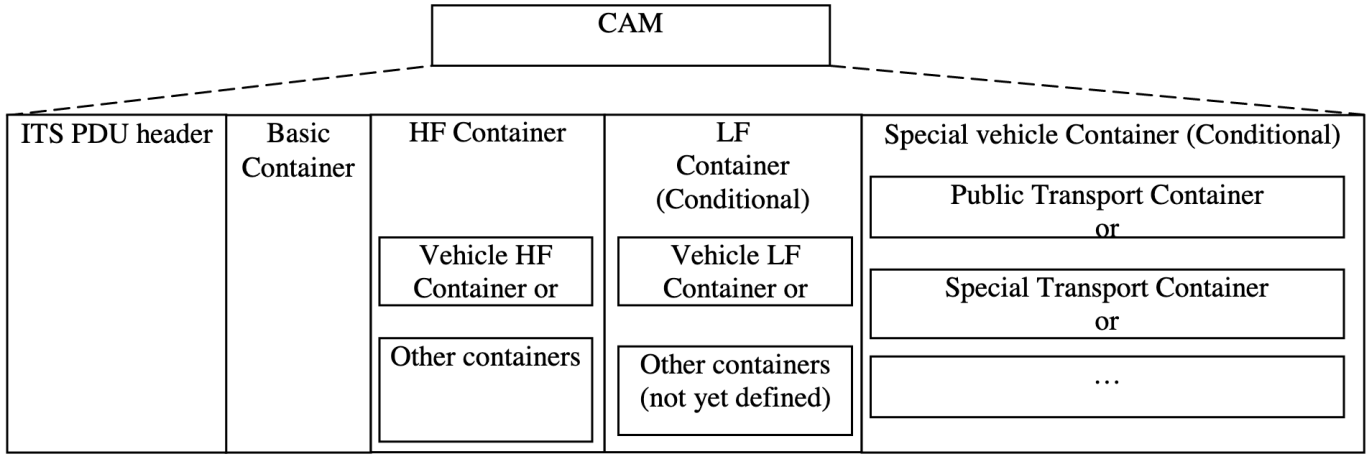
(e) Log data to support the detection of cyber-attacks and provide data forensic capability to enable analysis of attempted or successful cyber-attacks.

UN Regulation No. 155, page 6

Motivation

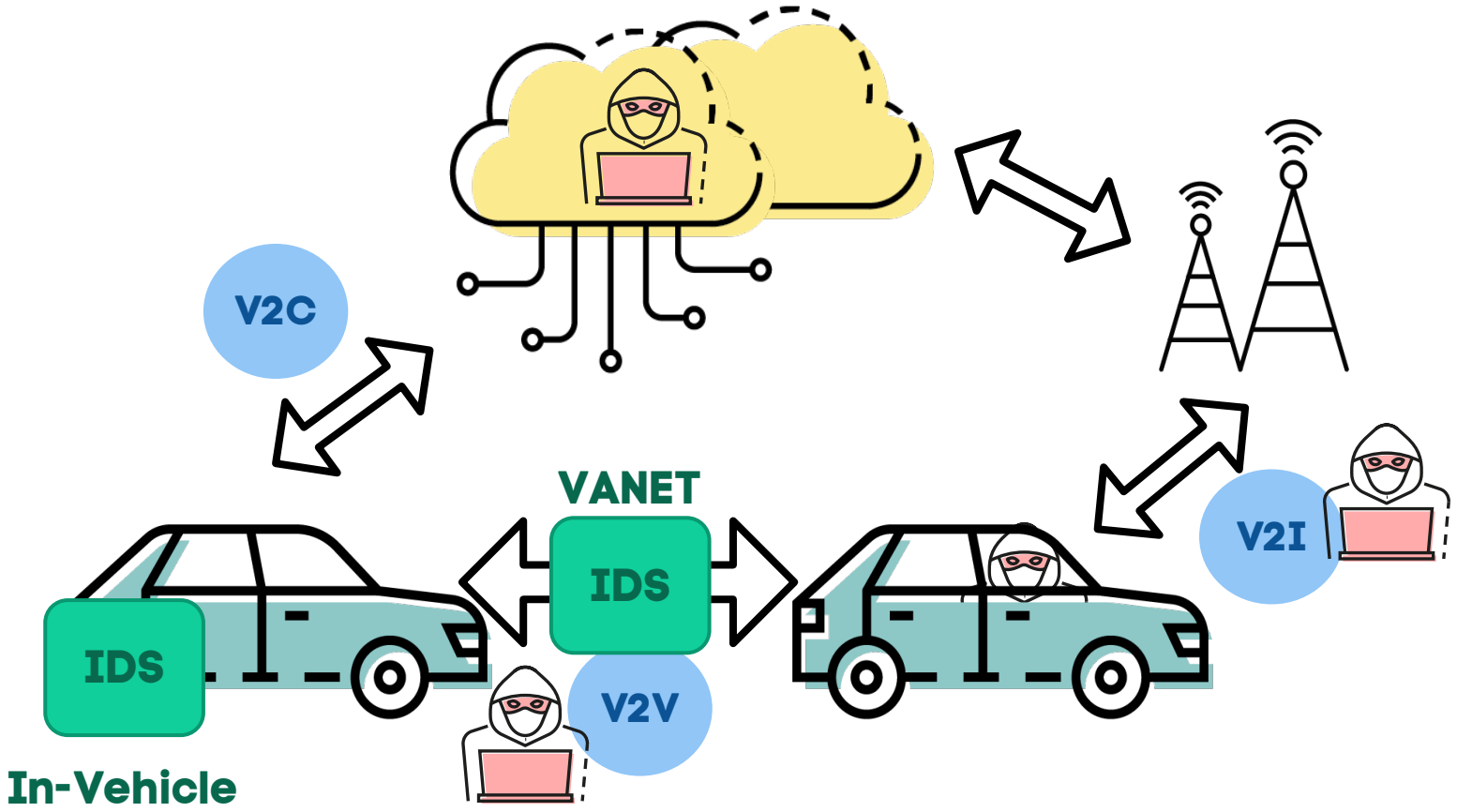
- IDSs are required (UN Reg. No. 155)
- Available IDS datasets are often
  - Not realistic
  - Parts of the frame are removed or modified





ETSI EN 302 637-2 V1.3.1, General structure of a CAM

Intrusion Detection in Vehicles



How to evaluate datasets?

- **Data readiness levels** by Lawrence in 2017 and later adapted by Castelijns et al. in 2020

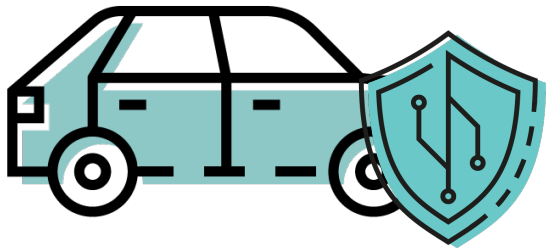
Band	Weight	Deficiency
C	40	Parseability
	25	Data storage
	15	Decoding
	10	Data formats
	10	Disjoint datasets
B	20	Column types
	30	Missing values
	20	Inconsistent data entries
	10	Duplicated records
	20	Meaningful values
A	20	Interpretable values
	20	Feature scaling
	20	Outlier detection
	30	Feature selection
	10	Coverage gap
AA	40	Legal violations
	40	Security risks
	20	Bias detection
AAA	-	None

Analysed Datasets

- In-vehicle datasets (10)
  - 10x intrusion detection purpose
- VANET datasets (7 incl. simulations)
  - 3x intrusion detection purpose
  - 4x misbehaviour detection
- Other possibly relevant datasets (3)
  - 1x driving behaviour
  - 1x vehicle trajectories
  - 1x V2V communication

Proposed Deficiencies for IDS Datasets

Band	Weight	Deficiency
C	30	Dataset documentation
	30	Objective
	20	Parseability
	20	Dataset age
B	40	Format correctness and consistency
	20	Dataset size
	20	Completeness
	20	Label inclusion and correctness
A	20	Class balance
	30	Attack documentation
	20	Security coverage
	30	Attack realism
AA	40	Dataset realism and diversity
	20	Feature context and documentation
	20	Difficulty
	20	Transformation and anonymisation
AAA	-	None



Band C

Band	Weight	Deficiency
C	30	Dataset documentation
	30	Objective
	20	Parsability
	20	Dataset age

Band B

Band	Weight	Deficiency
B	40	Format correctness and consistency
	20	Dataset size
	20	Completeness
	20	Label inclusion and correctness

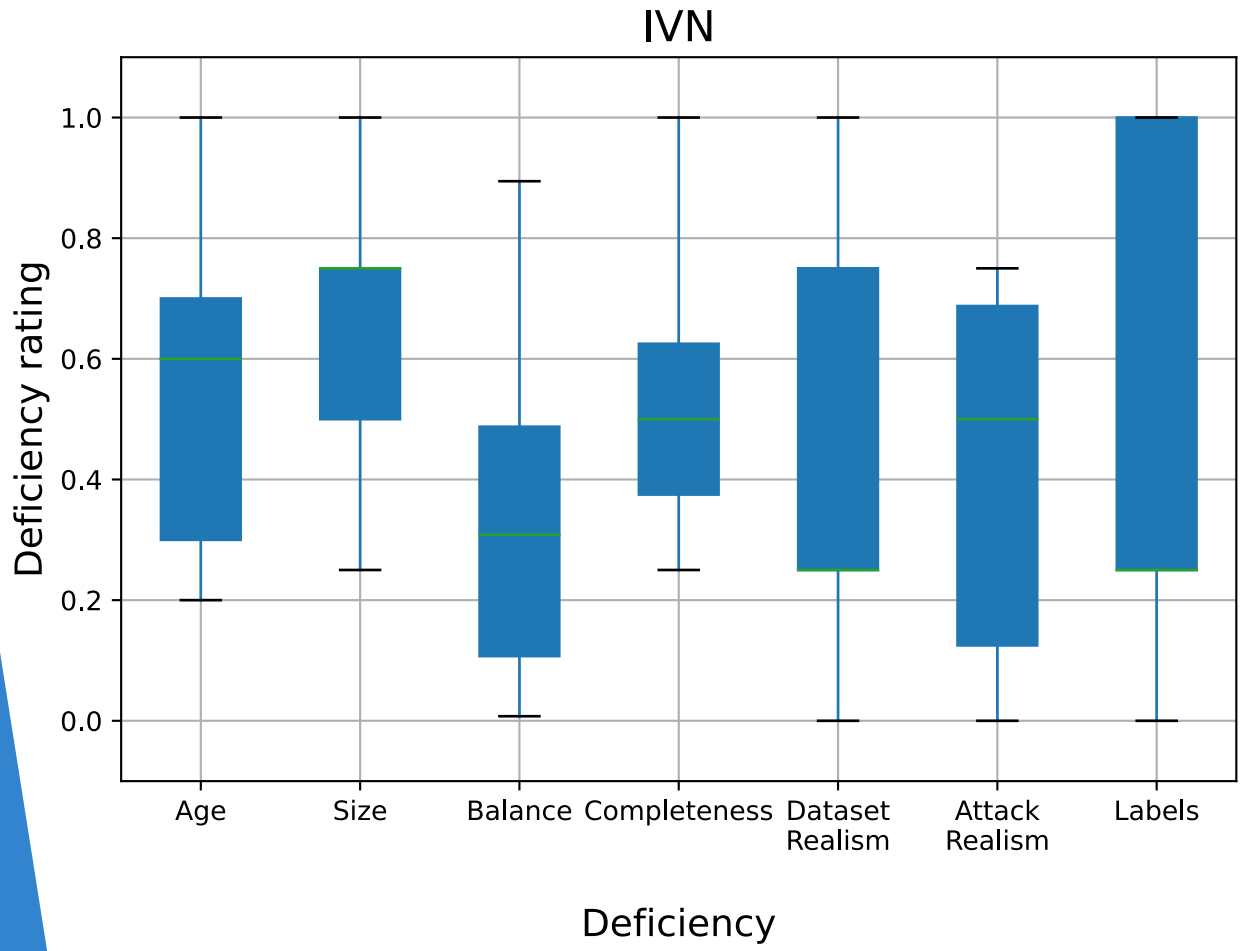
Band A

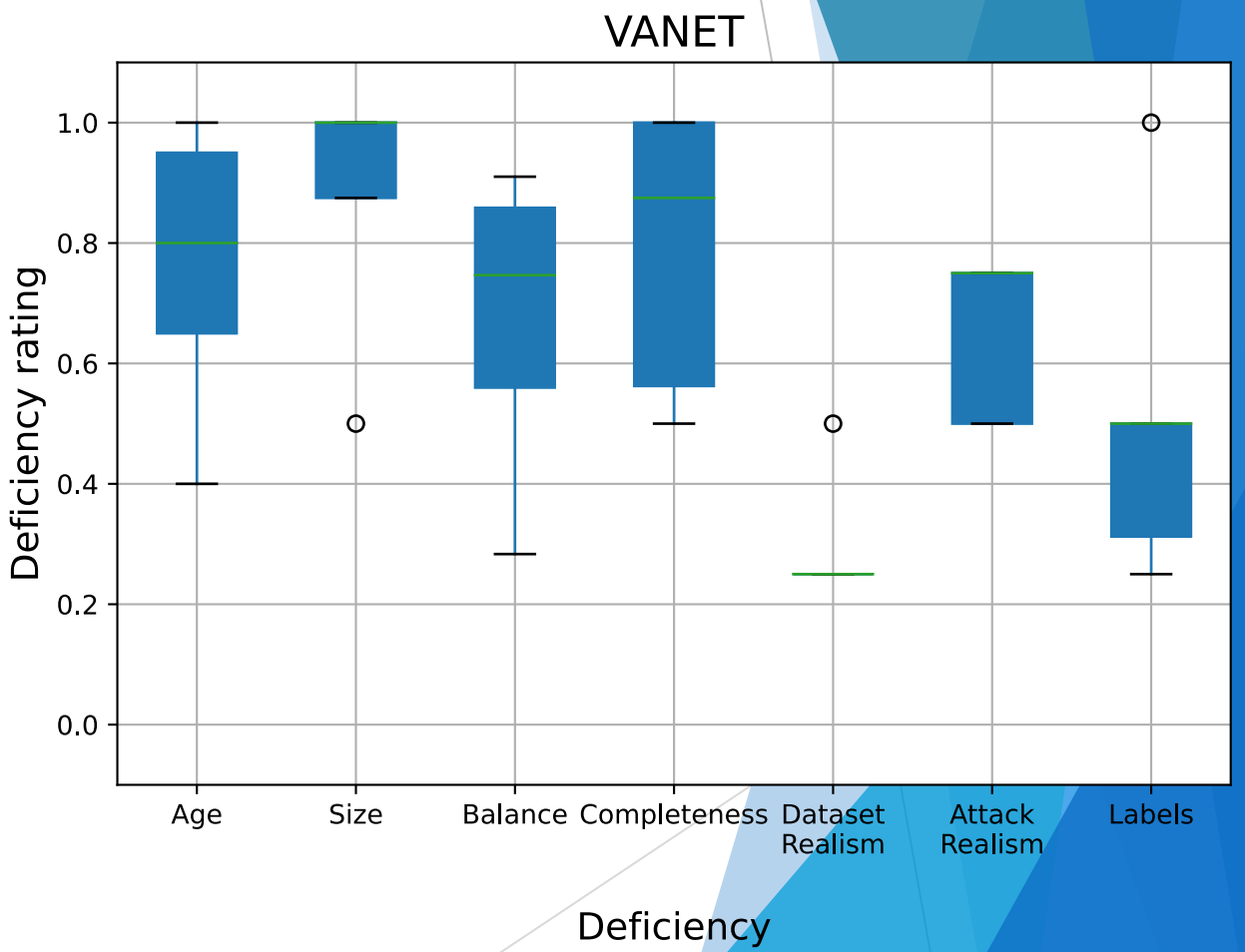
Band	Weight	Deficiency
A	20	Class balance
	30	Attack documentation
	20	Security coverage
	30	Attack realism

Band AA

Band	Weight	Deficiency
AA	40	Dataset realism and diversity
	20	Feature context and documentation
	20	Difficulty
	20	Transformation and anonymisation

Results


IVN
Deficiency rating
1.0
0.8
0.6
0.4
0.2
0.0
Age
Size
Balance
Completeness
Dataset Realism
Attack Realism
Labels
Deficiency


VANET
Deficiency rating
1.0
0.8
0.6
0.4
0.2
0.0
Age
Size
Balance
Completeness
Dataset Realism
Attack Realism
Labels
Deficiency

Results: Security Coverage

Dataset	Threat						Capability				
	S	T	R	I	D	E	Re	Wr	Sup	Rep	Dir
In-Vehicle datasets											
HCRL OTIDS	S	-	-	-	D	E	Re	Wr	Sup	Rep	-
HCRL Car-Hacking Dataset	S	T	-	-	D	E	Re	Wr	Sup	-	-
HCRL Survival	S	-	-	-	D	E	Re	Wr	-	-	-
TU Eindhoven v2	S	-	-	I	D	E	Re	Wr	Sup	-	-
SIMPLE	S	T	-	-	D	E	Re	Wr	Sup	-	-
SynCAN	S	T	-	-	D	E	Re	Wr	Sup	-	-
ORNL	S	T	-	-	D	E	Re	Wr	Sup	Rep	-
CrySyS	S	T	-	-	D	-	Re	Wr	-	-	-
Hisingen	S	T	-	-	D	E	Re	Wr	Sup	Rep	-
Bi2022	S	-	-	-	D	E	Re	Wr	-	-	-
VANET datasets											
Belenko2018	S	-	-	-	D	-	Re	Wr	Sup	-	-
VeReMi	S	T	-	-	-	-	Re	Wr	-	-	-
VeReMi Extension	S	T	-	-	D	-	Re	Wr	-	-	-
Lastinec2019	S	T	-	-	-	-	Re	Wr	-	-	-
VDoS-LRS	-	-	-	-	D	-	Re	Wr	-	-	-
VDDD	-	-	-	-	D	-	Re	Wr	-	-	-
Iqbal2021	S	T	-	-	-	-	Re	Wr	-	-	-

Recommendations

- Only one IVN dataset and two VANET datasets reached band A (score 85 as threshold)
- Information about context, attack type and attacker capabilities is often missing or ambiguous
  - **Improvement of documentation**
- Labels are sometimes missing, incomplete or incorrect
  - **Improve labelling**
- Attacks were often not very realistic
  - **Include more varied realistic attacks**

- **Assessment of the dataset before publication**

Systematic Evaluation of Automotive Intrusion Detection Datasets

- Propose a method to evaluate security datasets
- Adapt the **data readiness levels** to security datasets
- Evaluate available automotive datasets for in-vehicle and VANET
- Only one IVN dataset and two VANET datasets reached band A (score 85 as threshold)

References

- N. D. Lawrence, ‘Data Readiness Levels’, *arXiv:1705.02245 [cs]*, May 2017, Accessed: Mar. 10, 2022. [Online]. Available: http://arxiv.org/abs/1705.02245

- L. A. Castelijns, Y. Maas, and J. Vanschoren, ‘The ABC of Data: A Classifying Framework for Data Readiness’, in *Machine Learning and Knowledge Discovery in Databases*, vol. 1167, P. Cellier and K. Driessens, Eds. Cham: Springer International Publishing, 2020, pp. 3–16. doi: 10.1007/978-3-030-43823-4_1.