Francesco Cerasuolo

Continuous and Adaptive Learning for Traffic Analysis in the New Internet Era

Tutor: Prof. Antonio Pescapè

Cycle: XXXVIII

Year: Second

Candidate's information

- **MSc degree:** MSc degree in Computer Engineering from University of Naples Federico II
- **DIETI Research group/laboratory:** Traffic Group/ARCLab
- **PhD start date – end date:** 01/11/2022-31/10/2025
- **Scholarship type:** Unina
- **Periods abroad:** 3/10/2024-3/03/2025, University of Edinburgh, Prof. Paul Patras

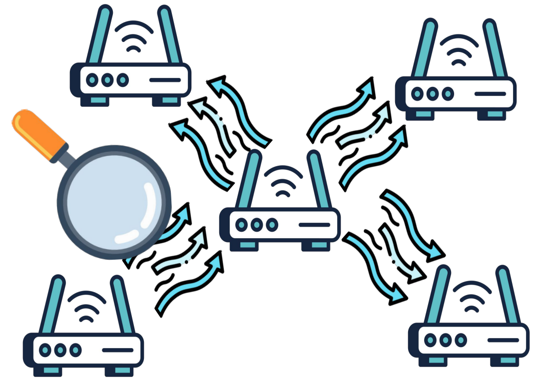
Summary of study activities

- **Ad hoc PhD courses / schools**
  - Ethics&AI
  - Hands-on Network Intrusion Detection via Machine and Deep Learning
  - Strategic Orientation for STEM Research & Writing
- **PhD School**
  - TMA PhD School, Dresden University of Technology
- **Seminars**
  - Robotics Meet AI & 5G - The future is now
  - Economic Fitness: Concepts, Methods and Applications
  - Open Science and Open Access
  - Sustainable IT: Strategies and best practices for a green engineering future
  - Media Forensics in the era of Generative AI
  - Introduction to large language models: evolution and the current state
  - Social Network Analysis Methods and Applications

Summary of study activities

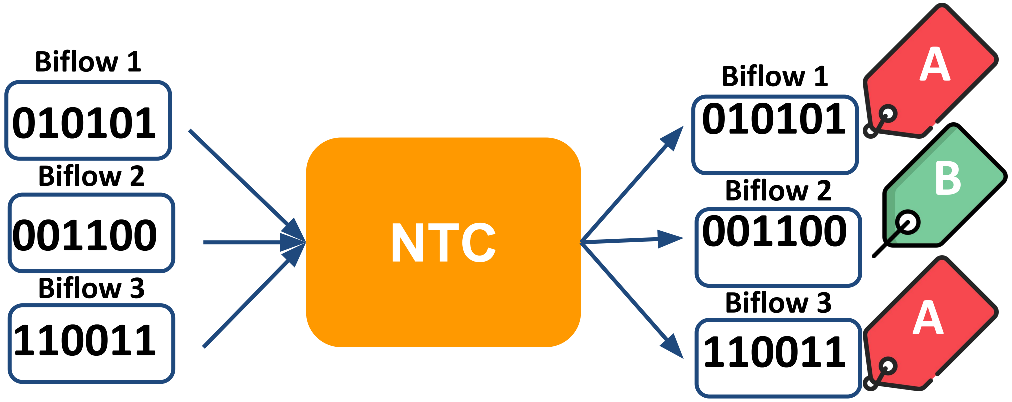
- **Conferences**
  - **IEEE International Conference on Big Data** Workshop “Machine Learning for Securing IoT Systems Using BigData”, 15-18 December 2023, Sorrento, Italy
  - **20th Italian Networking Workshop**, 22-24 January 2024, Madonna di Campiglio, Italy
  - **Italian Conference on CyberSecurity (ITASEC) Conference**, 8-12 April 2024, Salerno, Italy
  - **Network Traffic Measurement and Analysis (TMA) Conference**, 21-24 May 2024, Dresden, Germany
  - 1st International Workshop on Trustworthy and eXplainable Artificial Intelligence for Networks (TX4Nets), **IFIP/IEEE Networking**, 3-6 June 2024, Thessaloniki, Greece
  - **IEEE Symposium on Computers and Communications (ISCC)**, 26-29 June 2024, Paris, France

Research area



Network Traffic Analysis (NTA)

- Collecting and inspecting network data
- Understand and enhance performance



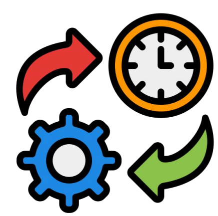
Network Traffic Classification (NTC)

- Associate a label to each traffic object (e.g., (bidirectional) flow, session, burst, etc.)

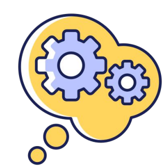
Main **challenges** of this domain:

Encrypted Traffic

Increasing traffic



Ever-changing nature of traffic



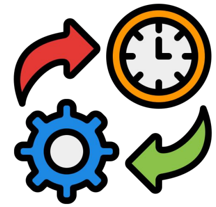
Trustworthiness of decisions

Challenges

Main **challenges** of this domain:

Encrypted Traffic

Increasing traffic



Ever-changing nature of traffic

Trustworthiness of decisions

ML- and DL-based solutions increasingly explored to effectively manage traffic classification tasks

Lifelong Learning allows models to continually learn, adapt, retain past knowledge, and save resources

XAI clarifies the opaque nature of DL models and decision process

Different **area of use:**

Mobile Apps

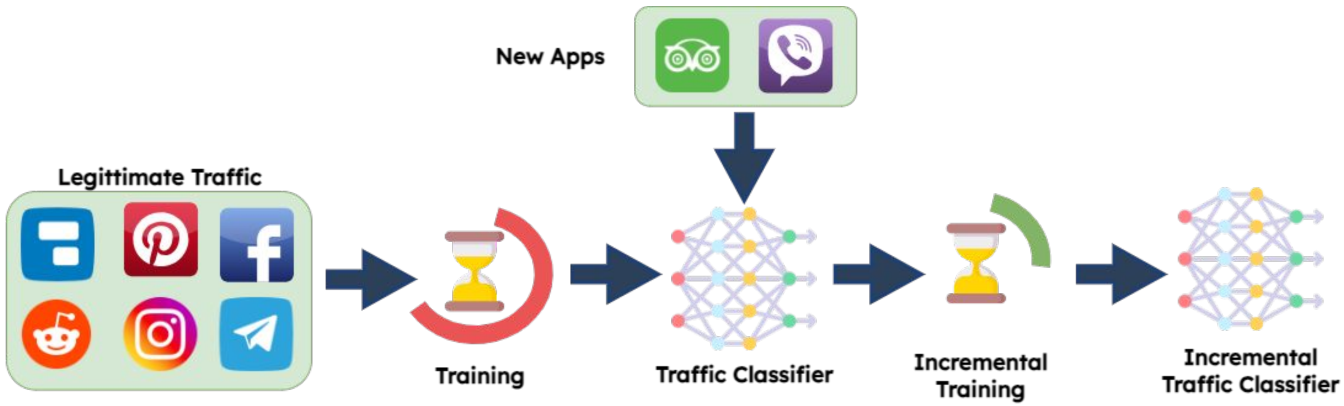
(ML/DL-based) **Traffic Classifiers** to identify app/services that generate traffic

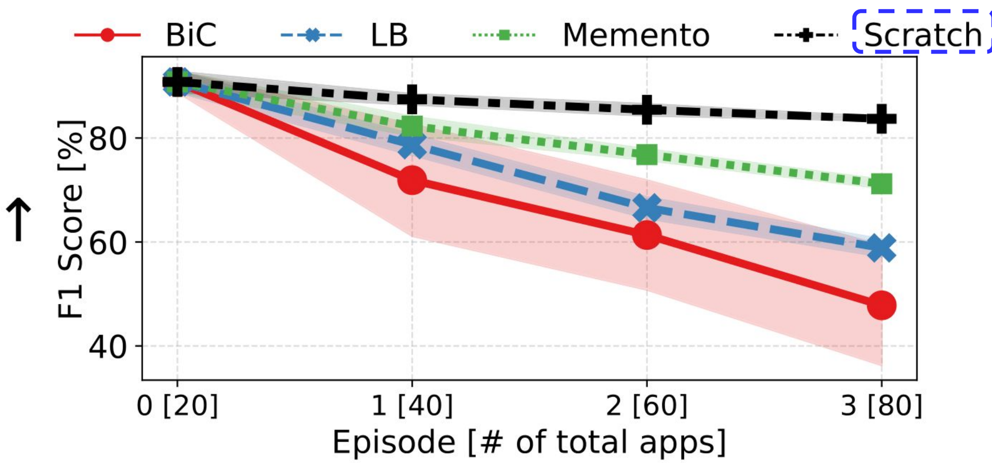
Security

(ML/DL-based) **Network Intrusion Detection Systems (NIDS)** distinguish legitimate traffic and malicious one and *(optionally)* attack classes

Incremental Mobile Apps Classification

New mobile apps are released each day. Hence, I define a **new incremental approach (`Memento`)** that **outperform state-of-the-art** for mobile app traffic classification

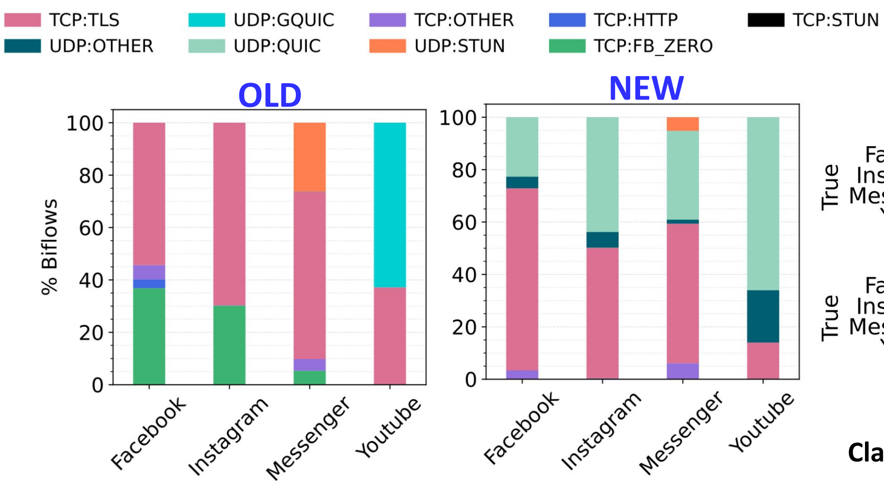


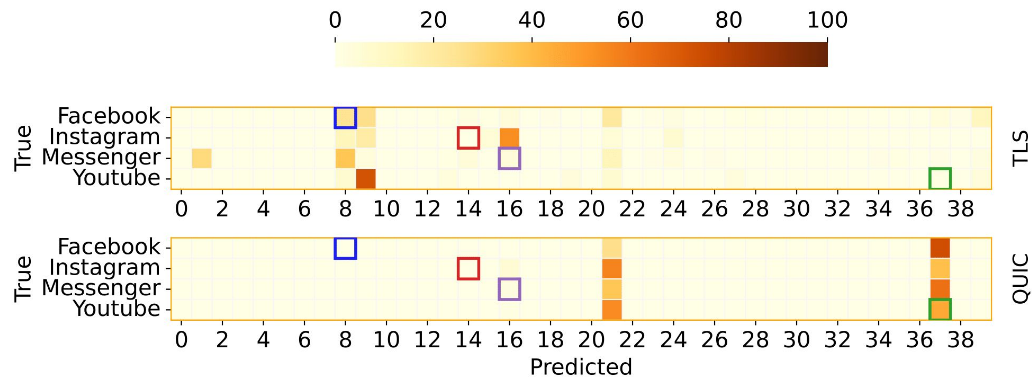


It involves no incremental process—viz. training with all data at once

Incremental Mobile Apps Classification

Traffic patterns of know apps also change over time. To classify this app, there is the need for **updating existing traffic classifiers effectively and efficiently**





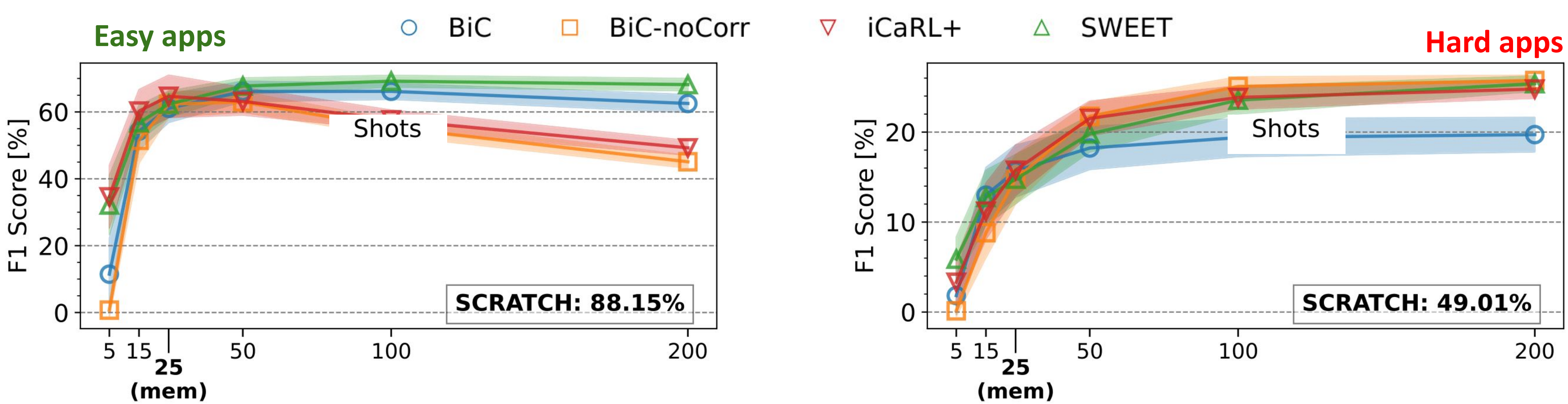
Classification results for a classifier trained on old traffic and tested on new one

This phenomenon is known as ***concept drift***. As found, new protocols adoption significantly changes traffic shapes of already known apps.

Incremental Mobile Apps Classification

Collecting sample from new apps is a difficult operation and specifically the labeling operation require **domain expertise** and **human effort**

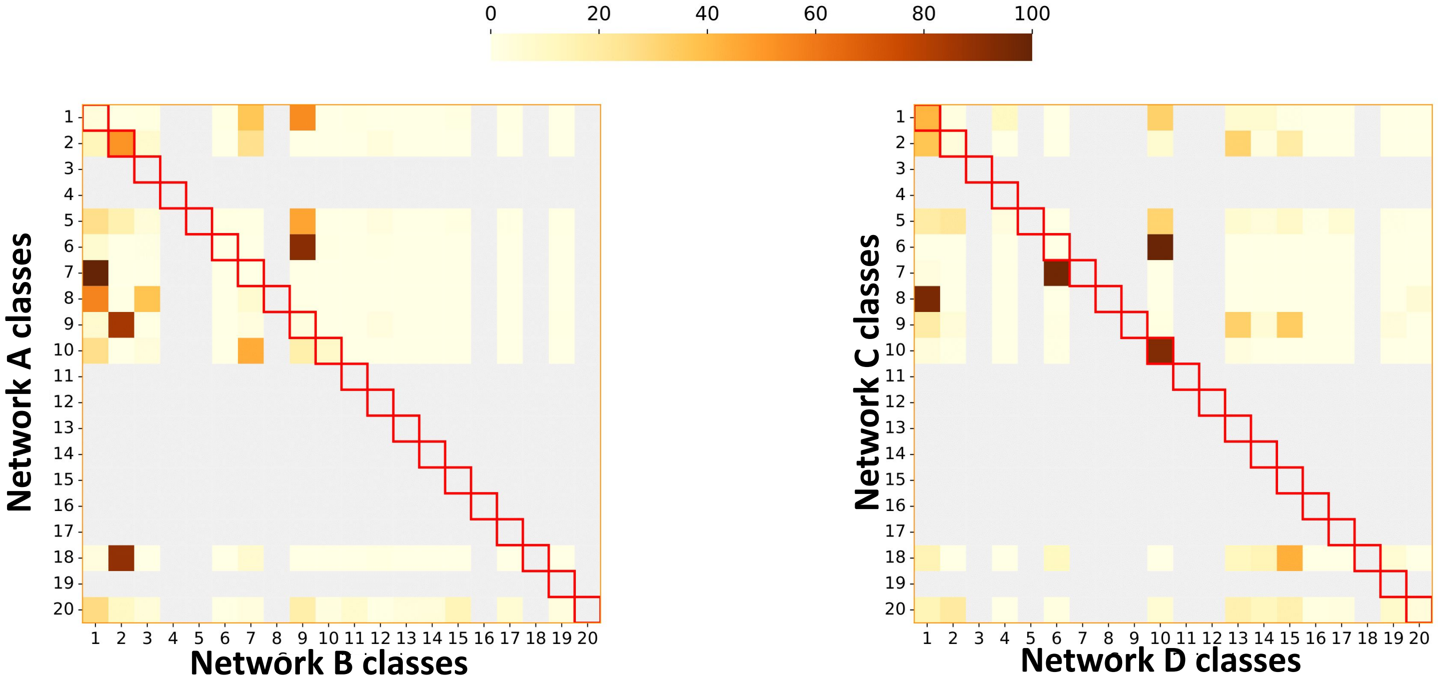
Hence, I devise a new few-shot incremental approach **SWEET**, including *adaptive traffic augmentation strategies*



information technology
electrical engineering

Incremental NIDS for Cybersecurity

In cybersecurity domain, different networks include **different types of attacks** and **different devices** generating benign traffic

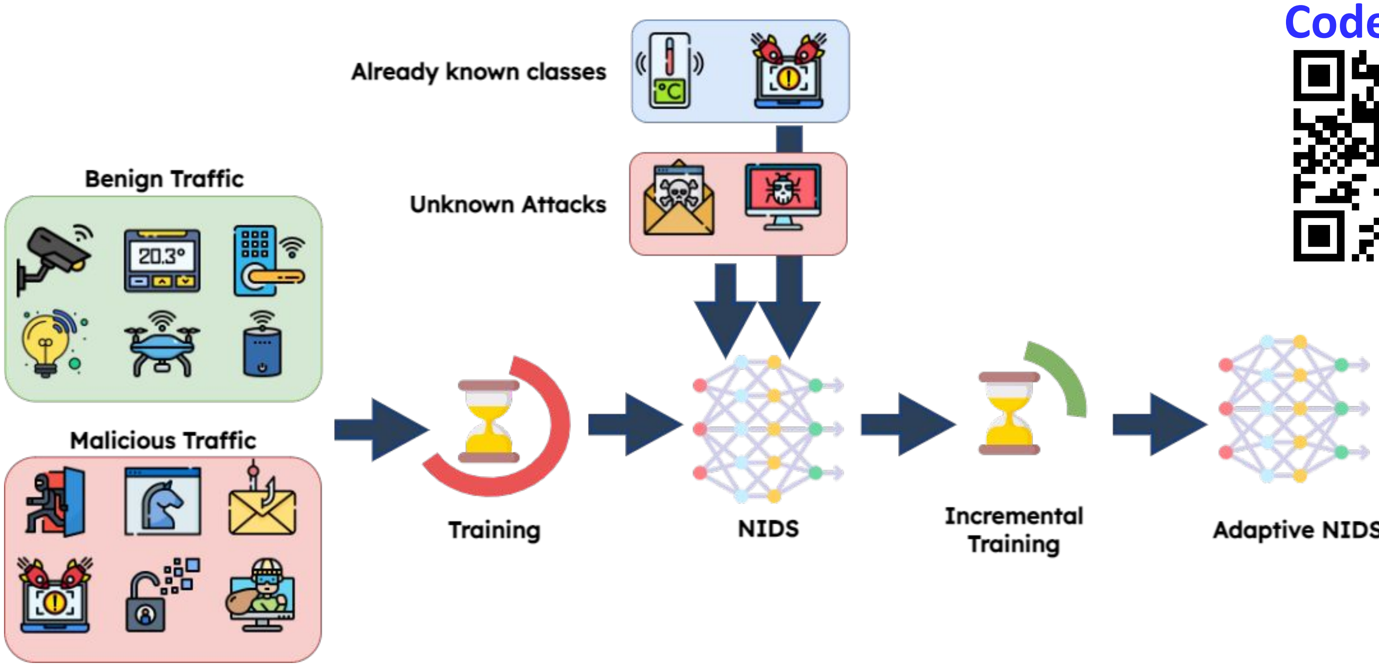


** confusion matrices from training NIDS on network traffic from a network and testing on traffic from a different network*

Incremental NIDS for Cybersecurity

Furthermore, we experience **0-day attacks**. Hence, I identify two subset of new classes to be added: *(i) already known classes* and *(ii) unknown attacks*. I design a new pipeline including both a **Domain Adaptation** and a **Class Incremental** procedure.

Code here!

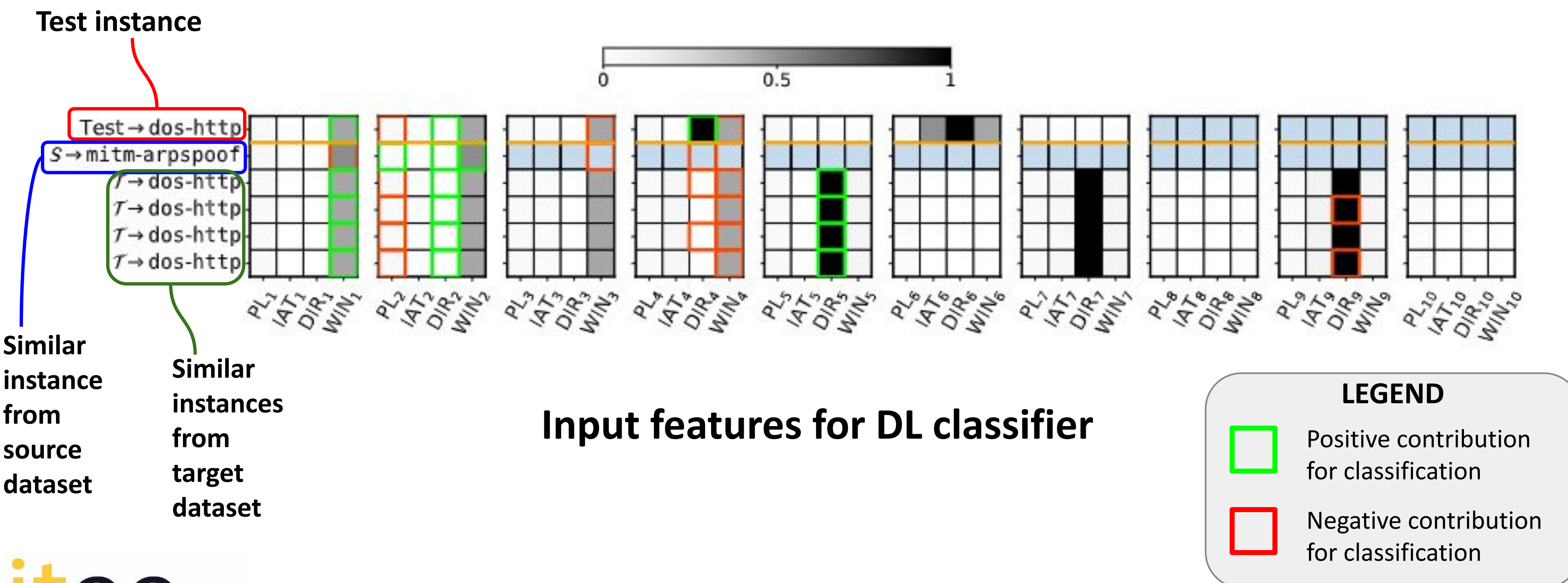


As a result, **incremental approaches** are able to reach classification performance **>70% F1 Score on both datasets**

Incremental NIDS for Cybersecurity

Mixing XAI techniques (viz. **similarity-based explanation** and **SHAP**), we discover that:

- the **target dataset has strong impact**
- Incremental NIDS **does not rely on minimal-distance** or **nearest-neighbor classification**



Input features for DL classifier

information technology electrical engineering PhD

Research products

[J1]	F. Cerasuolo, A. Nascita, G. Bovenzi, G. Aceto, D. Ciuonzo, A. Pescapè, D. Rossi, *MEMENTO: A Novel Approach for Class Incremental Learning of Encrypted Traffic,* **Computer Networks**, 245, p.110374
[J2]	F. Cerasuolo, G. Bovenzi, D. Ciuonzo, A. Pescapè, *Adaptable, Incremental, and Explainable Network Intrusion Detection Systems for Internet of Things*, **Engineering Applications of Artificial Intelligence** - *under 2nd round of review*
[J3]	F. Cerasuolo, G. Bovenzi, D. Ciuonzo, A. Pescapè, *Attack-Adaptive Network Intrusion Detection Systems for IoT Networks through Class Incremental Learning*, **Computer Networks** - *under review*
[C1]	A. Nascita, F. Cerasuolo, G. Aceto, D. Ciuonzo, V. Persico, A. Pescapè. *Explainable Mobile Traffic Classification: the case of Incremental Learning,* **Proceedings of the 2023 on Explainable and Safety Bounded, Fidelitous, Machine Learning for Networking**, pp. 25-31. 2023.
[C2]	F. Cerasuolo, G. Bovenzi, C. Marescalco, F. Cirillo, D. Ciuonzo, A. Pescapè, *Adaptive Intrusion Detection Systems: Class Incremental Learning for IoT Emerging Threats,* **2023 IEEE International Conference on Big Data (BigData)** (pp. 3547-3555)

Research products

[C3]	F. Cerasuolo, G. Bovenzi, V. Spadari, D. Ciuonzo, A. Pescapè, *Explainable Few-Shot Class Incremental Learning for Mobile Network Traffic Classification,* **IEEE Global Communications Conference (GLOBECOM), 2024**
[C4]	V. Spadari, F. Cerasuolo, G. Bovenzi, A. Pescapè, *An MLOps Framework for Explainable Network Intrusion Detection with MLflow,* **29th IEEE Symposium on Computers and Communications (ISCC) 2024**
[C5]	F. Cerasuolo, I. Guarino, V. Spadari, G. Aceto, A. Pescapè, *XAI for Interpretable Multimodal Architectures with Contextual Input in Mobile Network Traffic Classification,* **2024 IFIP Networking Conference (IFIP Networking) (pp. 757-762)**
[C6]	F. Cerasuolo, I. Guarino, G. Bovenzi, G. Antichi, A. Pescapè, *When Online Social Network Mobile Apps Meet QUIC: Characterization and Classification,* **International Conference on Passive and Active Network Measurement, 2025** - *under review*
[D1]	Mobile App Traffic Dataset, **MIRAGE-QUIC**

MIRAGE-QUIC here!

Next Year

- **Multimodal architectures** to more effectively harness the diversity present in network traffic

- **Memory free approaches** for more scalable and efficient incremental learning

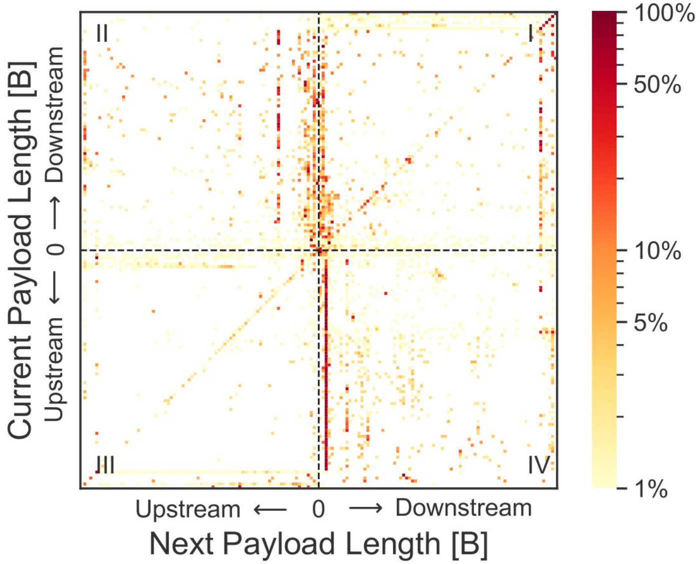
- Class Incremental Learning in **federated network**

- **Class removal** from a classifier

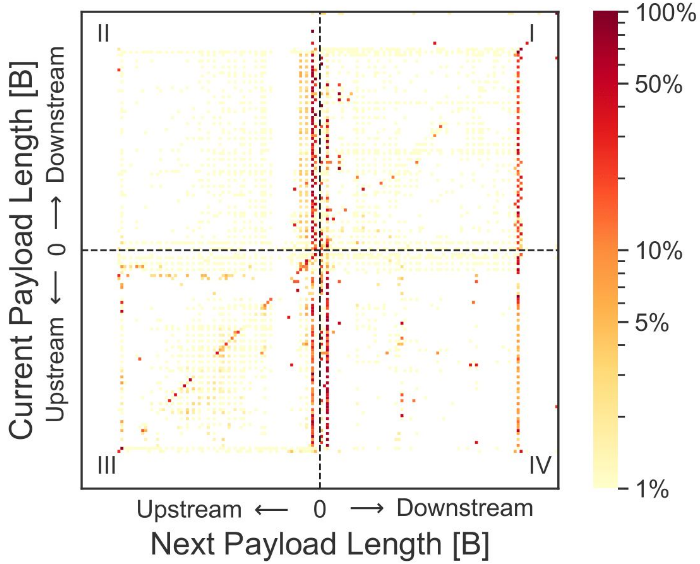
Thank you for the attention!

Backup Slides

Concept Drift 1/2



TLS

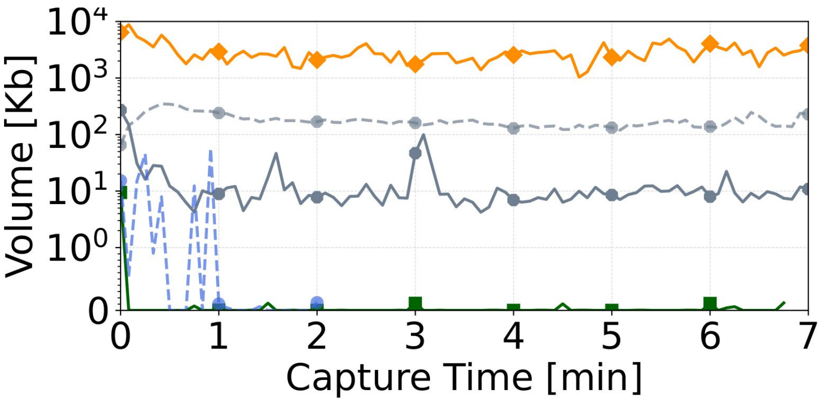


QUIC

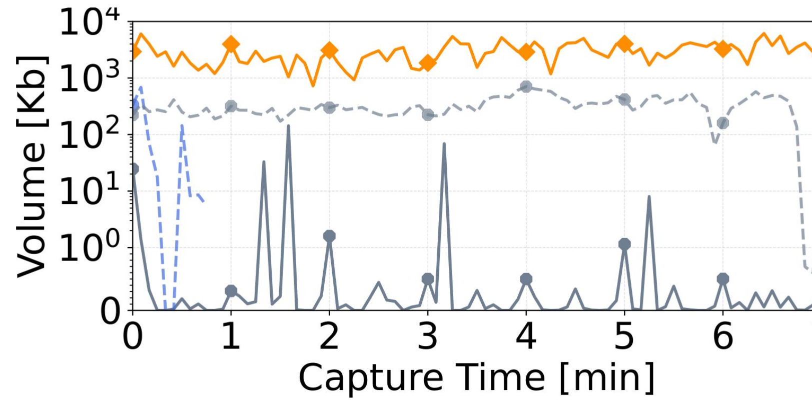
itee PhD
information technology
electrical engineering

Concept Drift 2/2

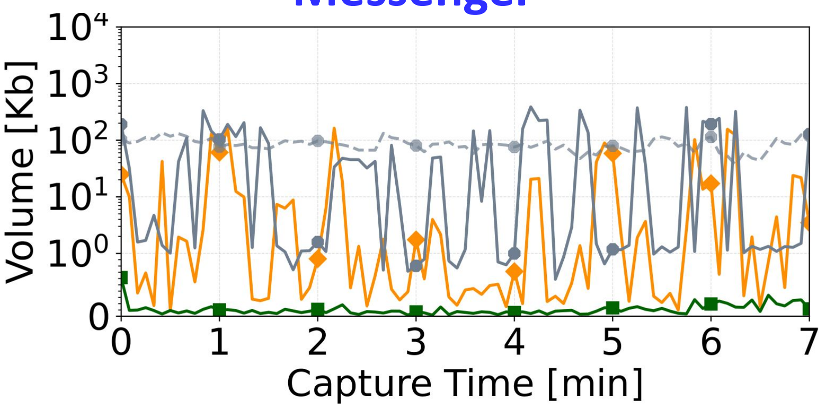
Facebook



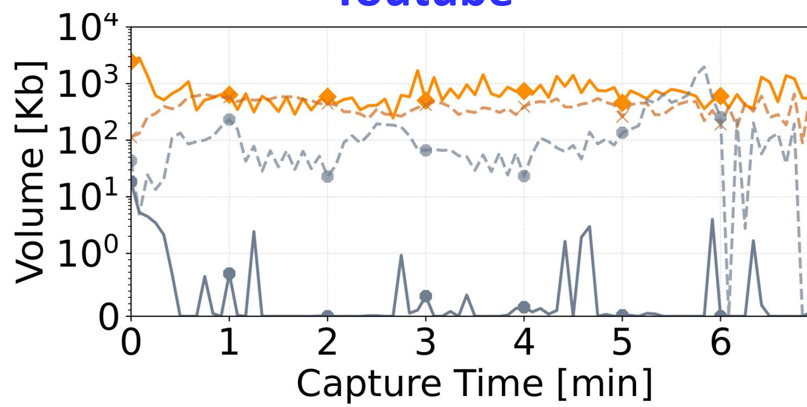
Instagram



Messenger

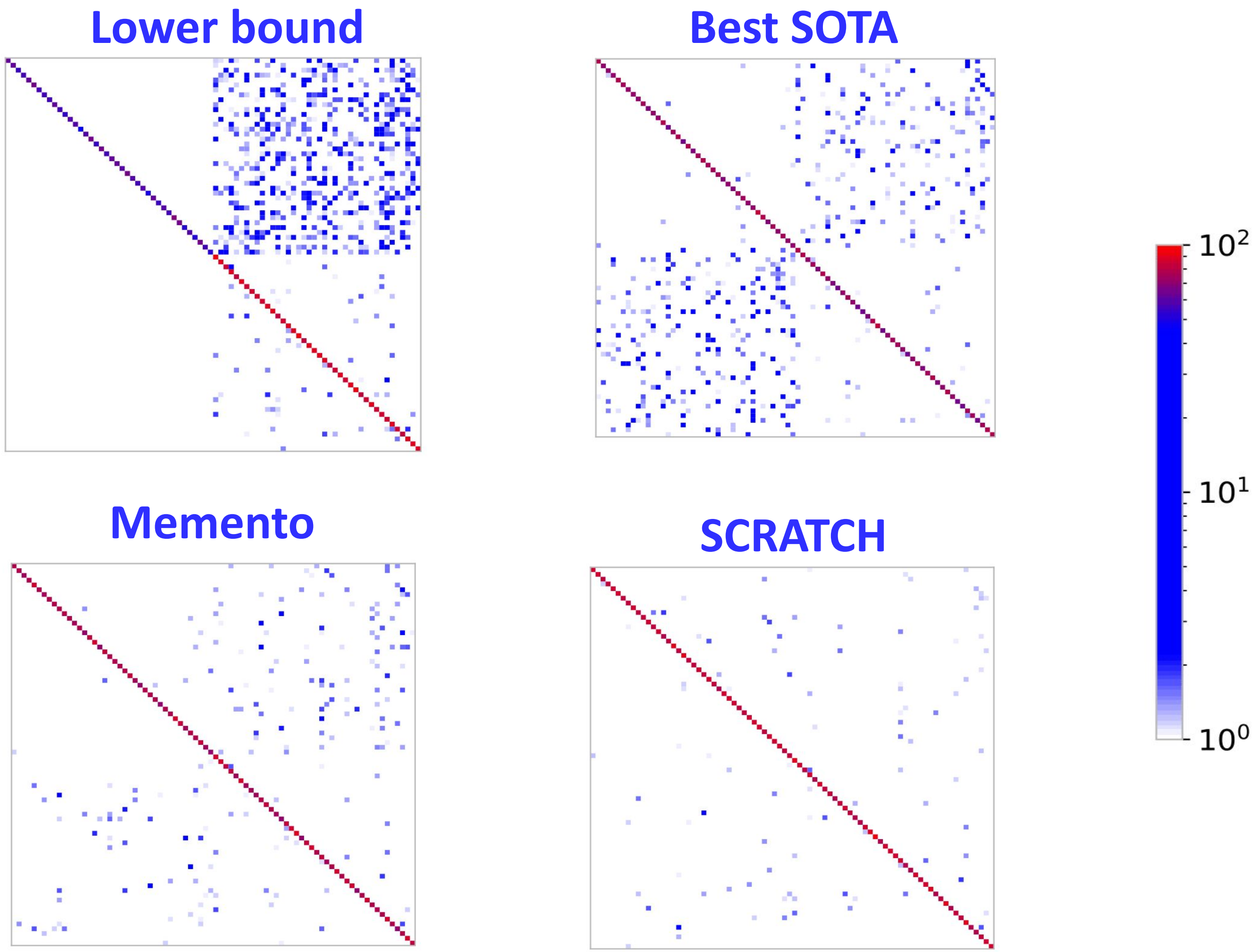


Youtube



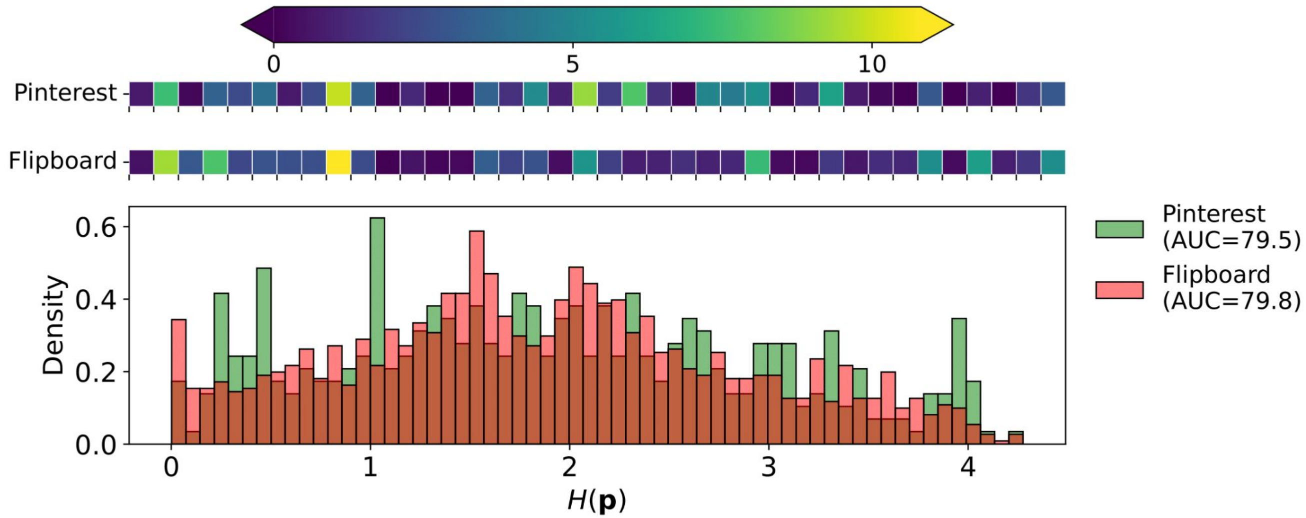
itee PhD
information technology
electrical engineering

Memento

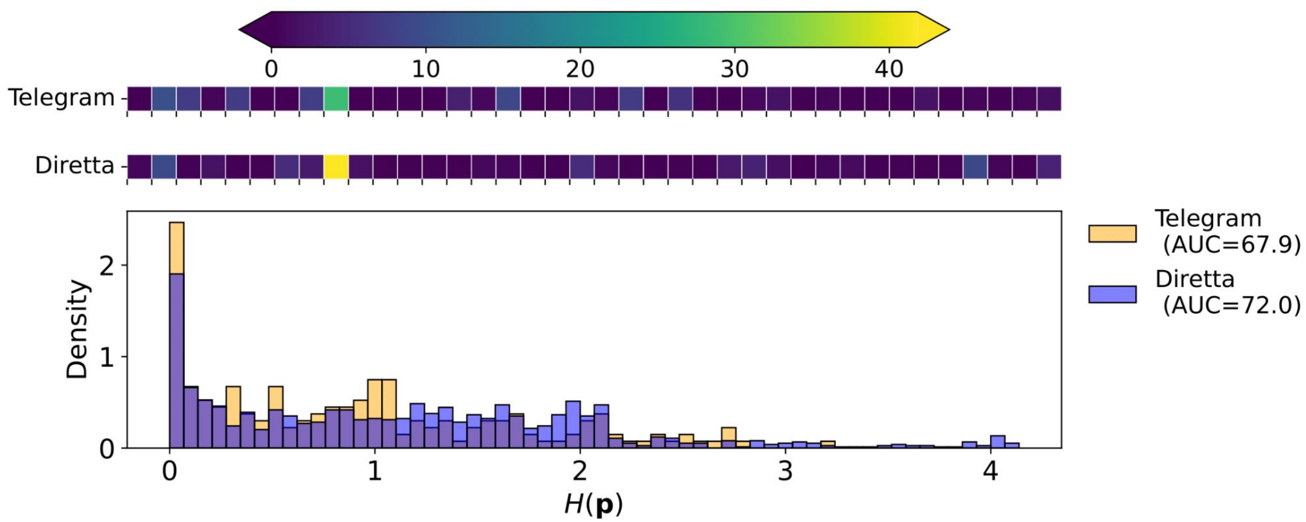


Sweet

HARD APPS

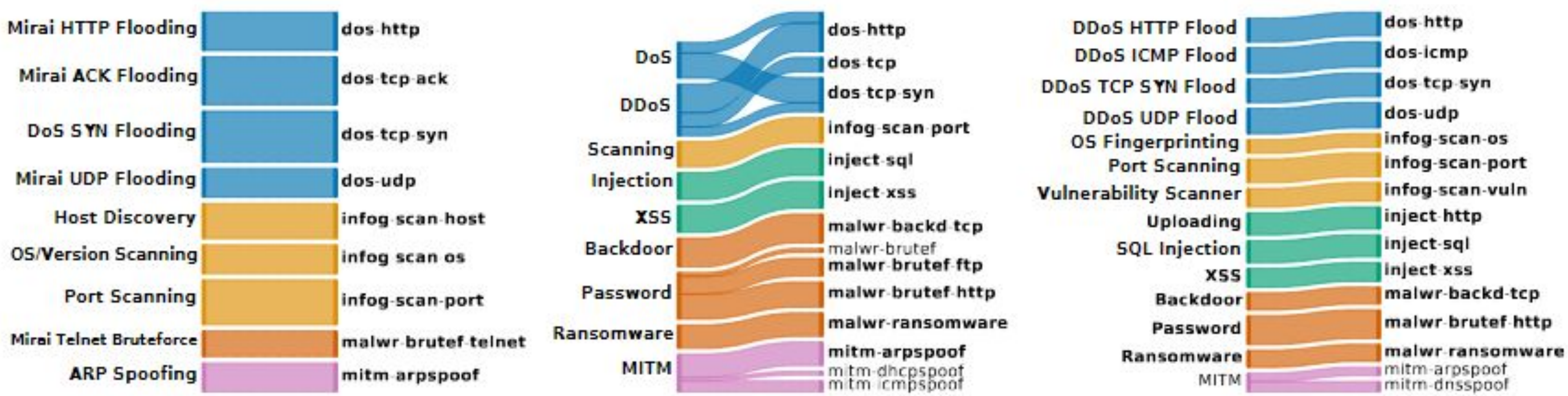


EASY APPS



Adaptive NIDS

Relabeling procedure to uniform the three considered datasets



(a) IoT-NID (b) TON_IoT (c) Edge-IIoT

Adaptive NIDS

From bigger dataset: BiC is the best

From smaller dataset: FT-Mem is the best

S	T	Appr.	F1 [%]							SCRATCH NIDS
			C^{src}_{unc}	C^{src}_{com}	C^{src}	C^{tgt}_{unc}	C^{tgt}_{com}	C^{tgt}	C^{all}	
Edge-IIoT	IoT-NID	FT	0.00	6.90	2.96	71.28	94.19	85.03	37.15	77.71
		FT-Mem	50.37	74.38	60.66	74.40	92.38	85.18	**70.88**	
		BiC	59.66	94.77	74.71	69.21	39.70	51.51	65.04	
	TON_IoT	FT	0.00	21.78	14.00	30.64	97.81	81.02	44.93	80.83
		FT-Mem	44.72	70.52	61.31	33.61	97.81	81.76	**70.75**	
		BiC	31.92	76.90	60.84	31.75	24.68	26.45	44.97	
IoT-NID	Edge-IIoT	FT	3.53	13.67	9.61	61.70	95.81	76.32	48.53	77.71
		FT-Mem	57.07	61.43	59.69	65.37	95.79	78.41	70.61	
		BiC	60.91	80.50	72.67	67.03	86.72	75.47	**74.30**	
	TON_IoT	FT	2.25	18.74	10.49	73.89	83.27	77.80	47.21	85.69
		FT-Mem	76.82	70.85	73.84	79.44	83.29	81.05	77.77	
		BiC	79.74	71.94	75.84	77.94	81.64	79.48	**77.83**	
TON_IoT	Edge-IIoT	FT	0.13	34.11	25.62	38.30	79.05	64.50	46.55	80.83
		FT-Mem	17.98	83.03	66.77	68.59	79.00	75.28	**71.35**	
		BiC	35.62	94.57	79.83	63.38	50.89	55.35	66.65	
	IoT-NID	FT	0.00	35.32	14.72	77.96	79.01	78.48	43.70	85.69
		FT-Mem	65.30	68.93	66.81	82.61	79.12	80.87	**73.20**	
		BiC	76.15	75.86	76.03	73.82	46.04	59.93	68.71	

N.B.: $C^{src} = C^{src}_{unc} \cup C^{src}_{com}$, $C^{tgt} = C^{tgt}_{unc} \cup C^{tgt}_{com}$, and $C^{all} = C^{src} \cup C^{tgt}$.

Adaptive NIDS

Target dataset strong impact for correctly classified samples

Table 7
[%] of correctly classified biflows having the top-1, top-3, and top-5 neighbor from the incremental training set ($D = D^{mem} \cup D^{tgt}$) which belongs to the same dataset for a NIDS trained on IoT-NID and adapted to TON_IoT with BiC.

Test Dataset	Classes	Same Dataset Neighbors [%]		
		Top-1	Top-3	Top-5
IoT-NID	C^{src}_{unc}	80.03	82.88	80.02
	C^{src}_{com}	96.13	95.95	95.63
TON_IoT	C^{tgt}_{unc}	70.92	74.16	76.01
	C^{tgt}_{com}	83.90	85.63	86.10

itee PhD
information technology
electrical engineering