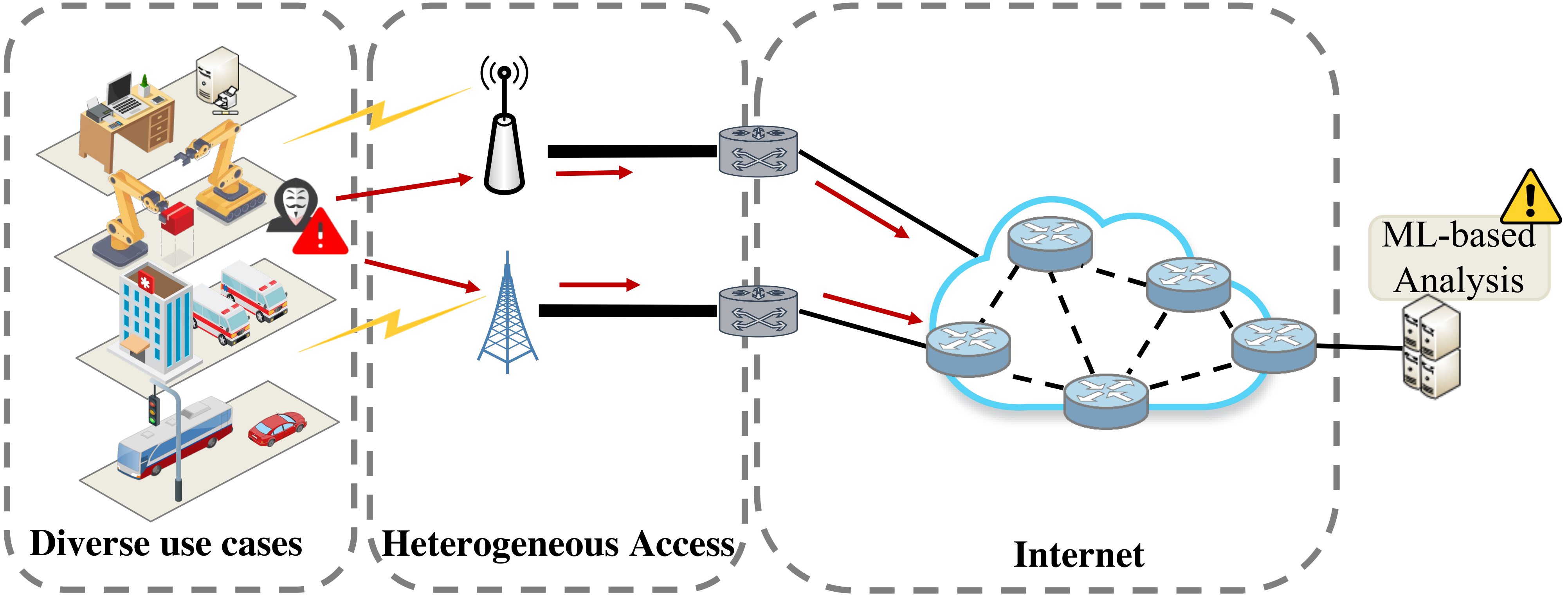
Advanced Threat Defense with In-Network Traffic Analysis for IoT Gateways

Mingyuan Zang*, Changgang Zheng†, Noa Zilberman†, and Lars Dittmann*

*Technical University of Denmark, †University of Oxford

Internet of Things (IoT) Network

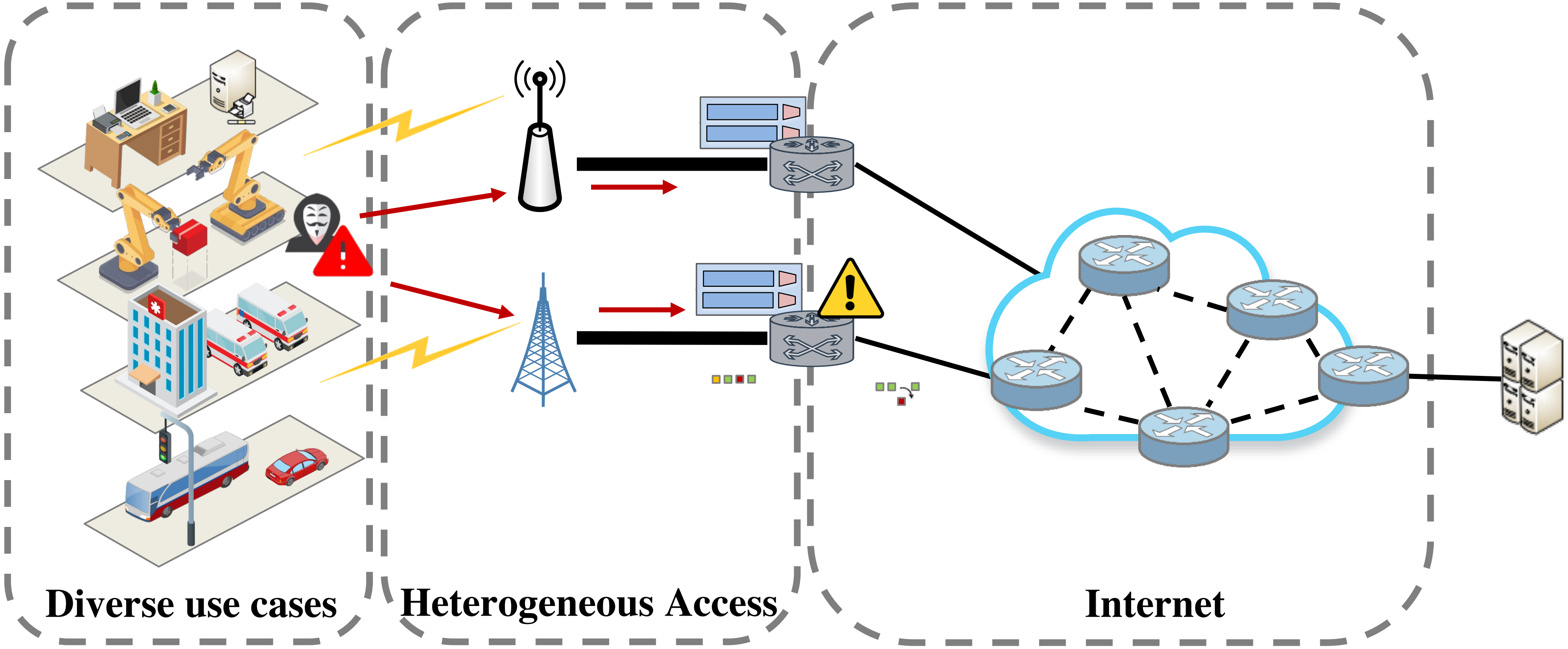


5G/6G's extremely low latency requirements + emerging attack variants in IoT

→ Fast spreading threats with changing patterns

How to continuously learn and swiftly mitigate emerging threat patterns in IoT network?

Internet of Things (IoT) Network

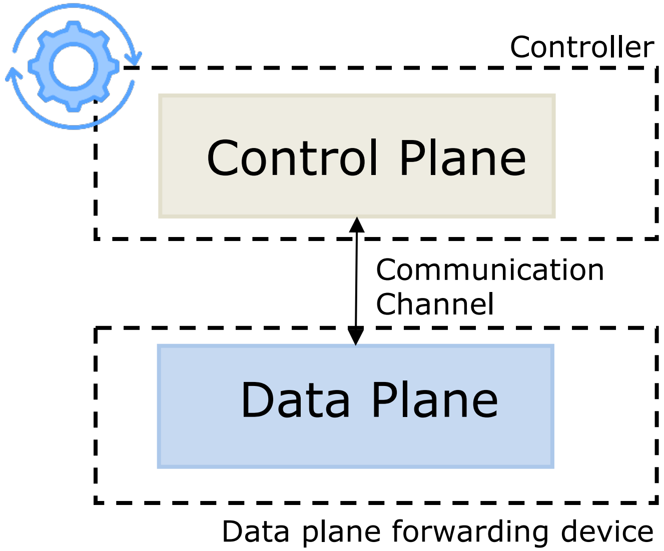


Programmable data planes enable in-network ML-based mitigation

What is in-network ML (inference)?

From Software-Defined Networking (SDN) To Programmable Data Plane

SDN OpenFlow switch



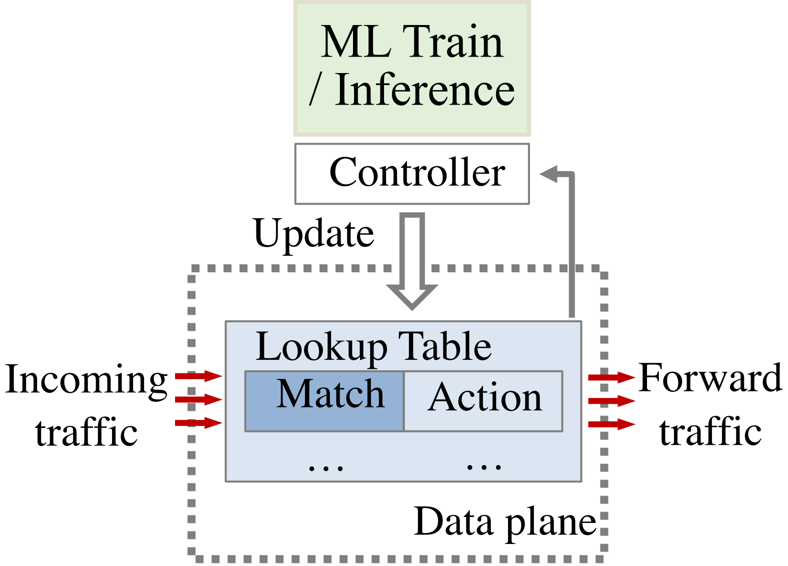
Controller

Control Plane

Communication Channel

Data Plane

Data plane forwarding device



ML Train / Inference

Controller

Update

Lookup Table

Match | Action

… …

Incoming traffic

Forward traffic

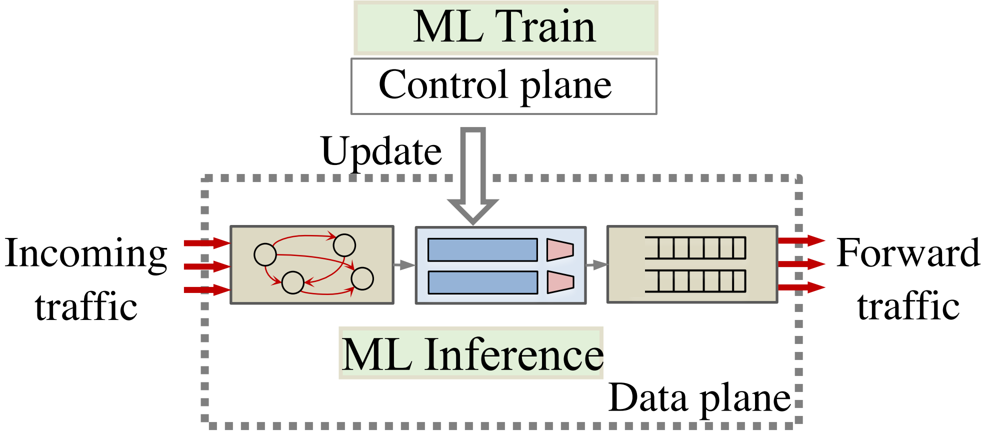
Data plane

ML deployed at controller/cloud

From Software-Defined Networking (SDN) To Programmable Data Plane

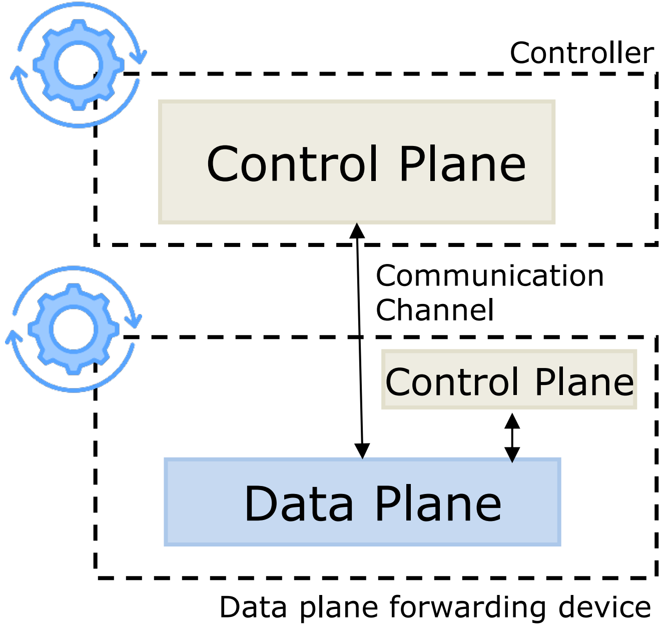
*"This is how I **know** to process packets …"*

*"This is how I **want** to process packets"*



- ✓ Flexible packet parsing
- ✓ Immediate action to anomaly
- ✓ Runtime reconfigurable

Programmable switch

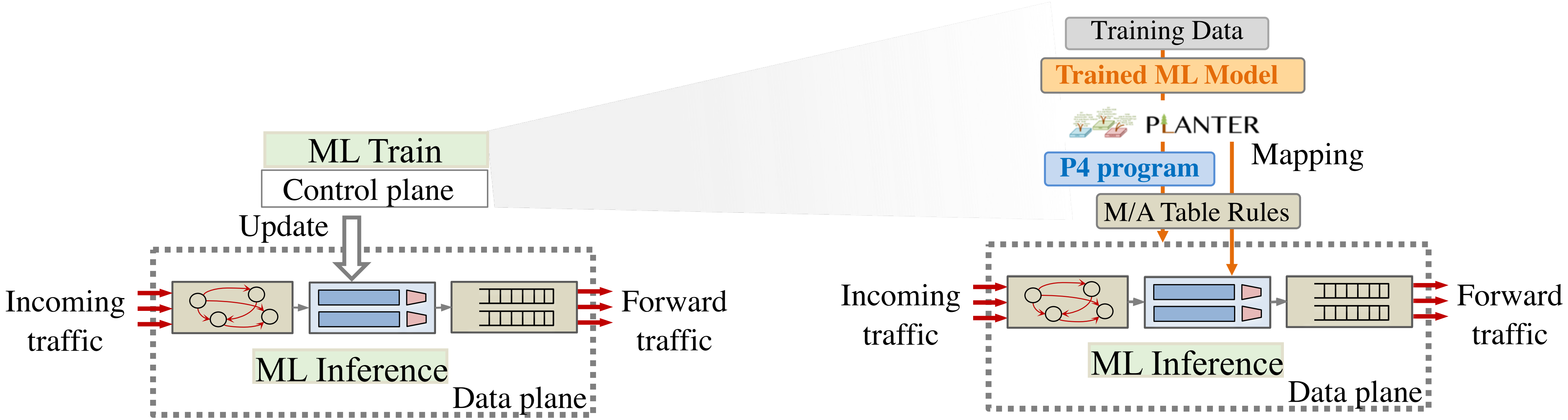


Offload ML inference to the data plane

Efficient In-Network ML Inference

In-network ML inference in Planter [1]

- A trained model → a series of inference operations on programmable pipeline (Match-Action table rules)
- Support common-used model: Bayes, SVM, DT, NN, …



[1] C. Zheng et al., "Automating In-Network Machine Learning," arXiv preprint arXiv:2205.08824, 2022

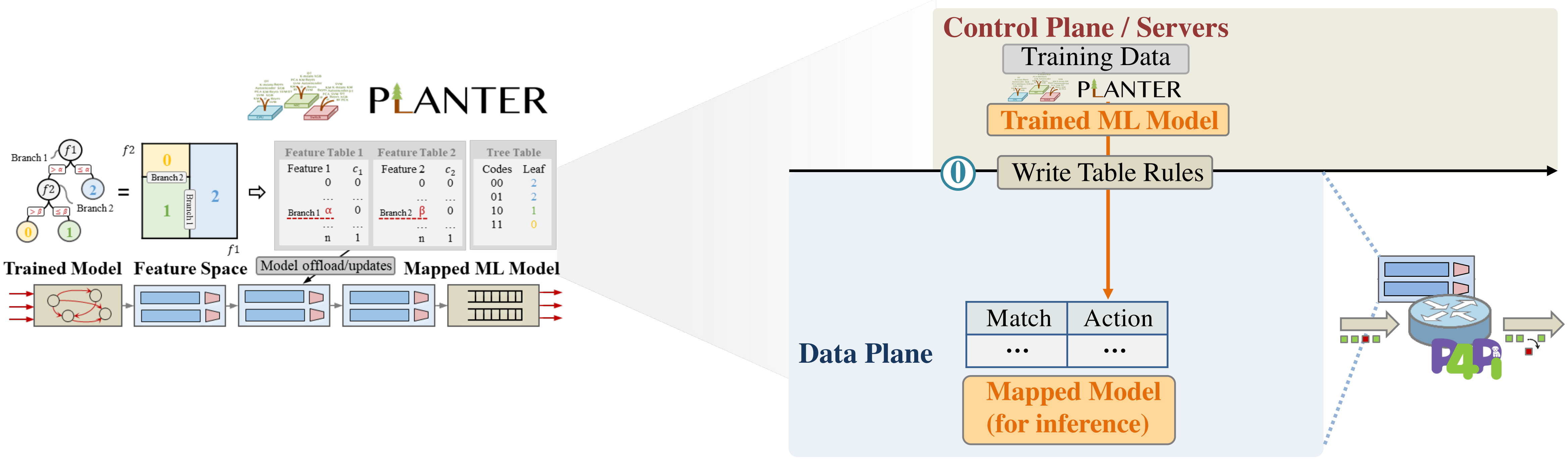
How to apply in-network ML inference to IoT gateway without affecting data plane service?

Proposed Design – P4Pir

Step 0: In-network ML inference in IoT gateway

- Tree model (Decision Tree/Random Forest) inferred in gateway data plane

  Initialize the mapped model within the processing pipeline (Match-Action table rules)



Feature Table 1	
Feature 1	c_1
0	0
...	...
Branch 1 α	0
...	...
n	1

Feature Table 2	
Feature 2	c_2
0	0
...	...
Branch 2 β	0
...	...
n	1

Tree Table	
Codes	Leaf
00	2
01	2
10	1
11	0

Match	Action
···	···

Proposed Design – P4Pir

Step 1: In-band feature extraction and fast mitigation

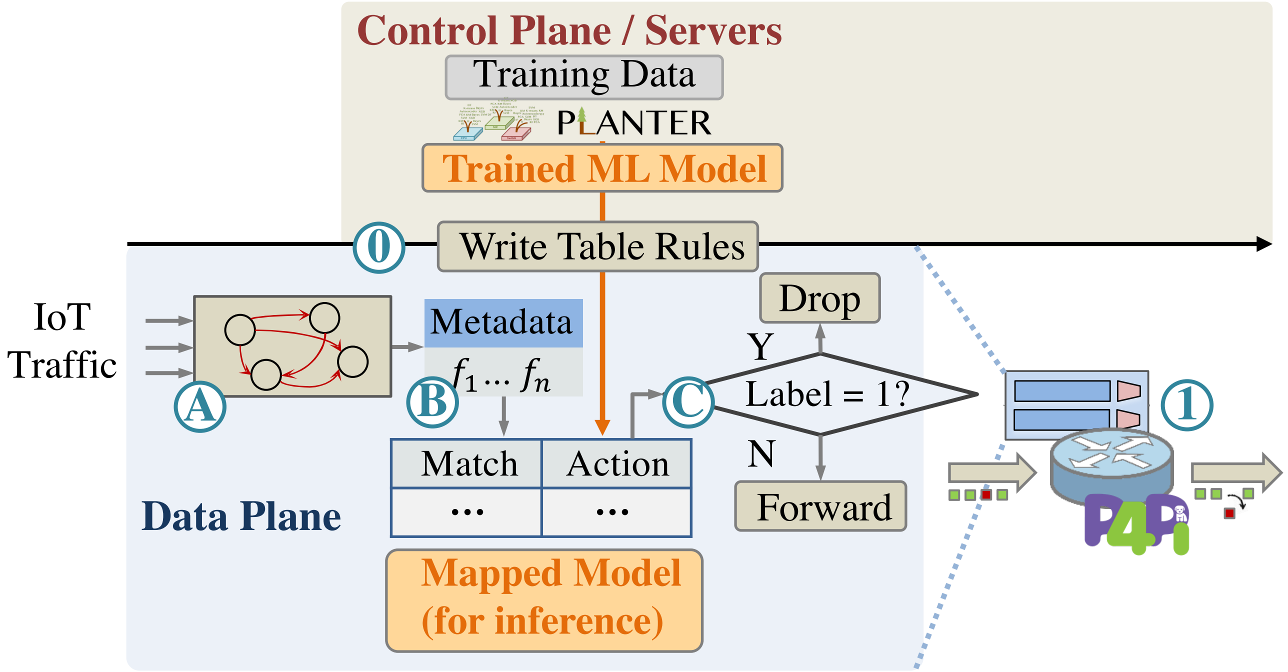
- Customized packet parsing and feature extraction

  Extracted features » in-network ML inference

- Threat mitigation based on inference results in data plane

  Benign (label = 0) → forward

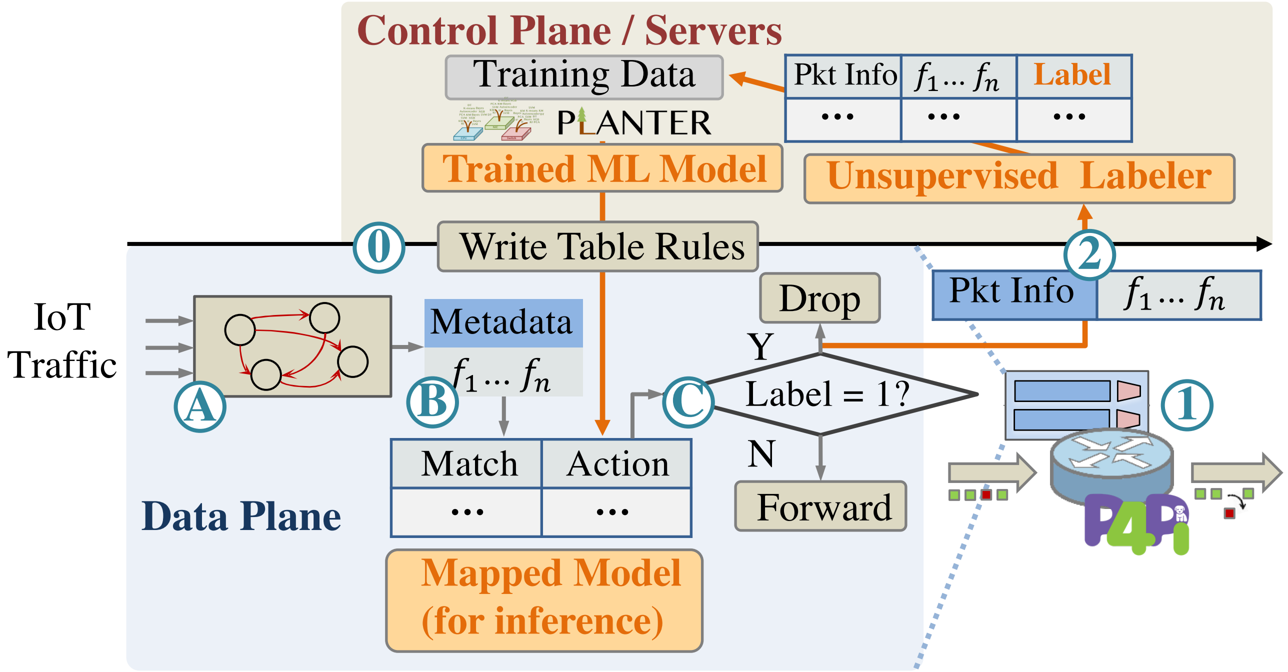
  Malicious (label = 1) → drop



Proposed Design – P4Pir

Step 2: Proactive logging and unsupervised labeling for IoT traffic

- Proactive logging of extracted features in digests
- Unsupervised-based *iForest* algorithm to automate log labeling

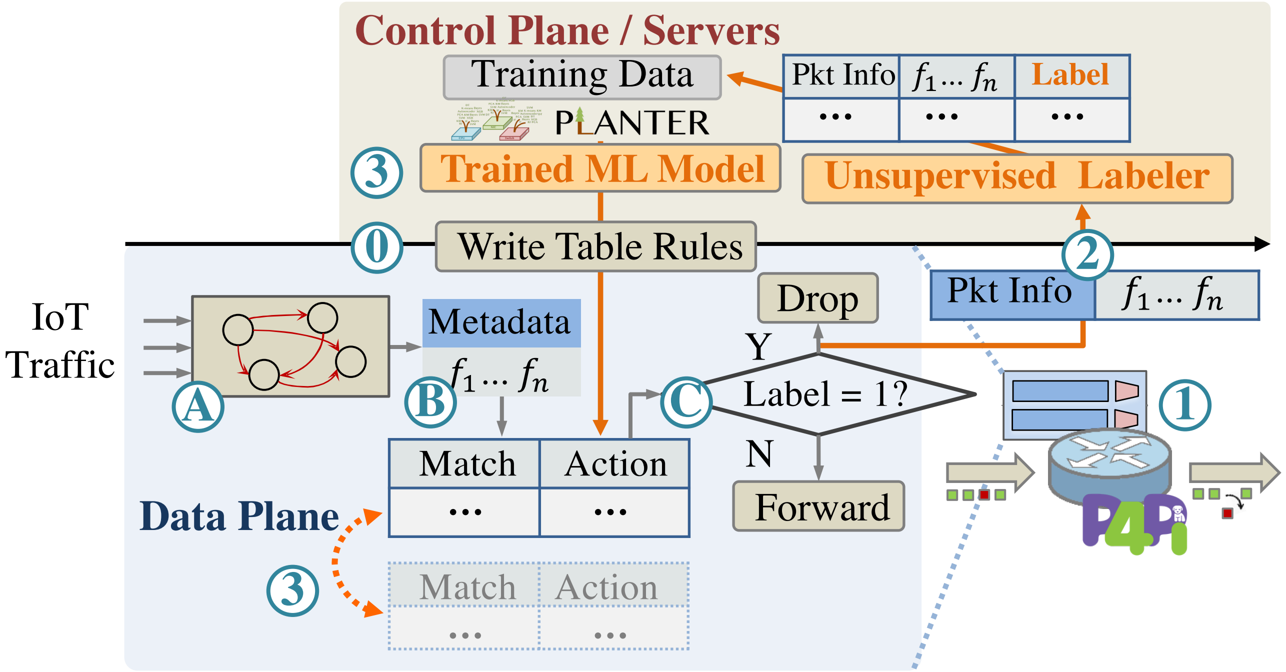


Proposed Design – P4Pir

Step 3: Continuous update for in-network model

- Shadow table modifications for hitless updates of in-network model

  Runtime update the retrained model without disrupting data plane functions



Evaluation Results

- **Prototype**

  P4Pi: Raspberry Pi 4 Model B + BMv2 programmable switch

- **Performance**

  >30% accuracy ↑, real-time mitigation, negligible jitter, 8% ↑ on CPU utilization

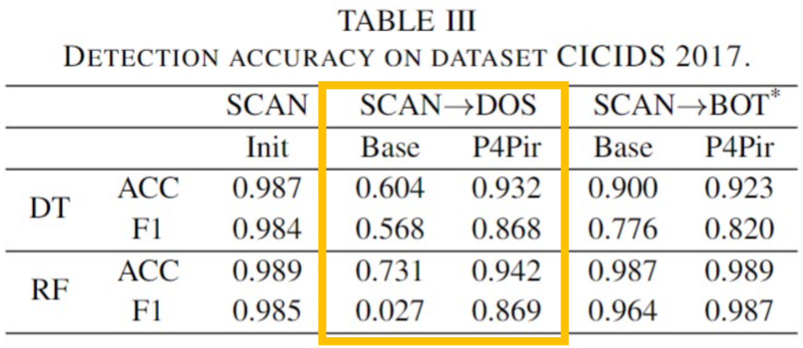


TABLE III
DETECTION ACCURACY ON DATASET CICIDS 2017.

		SCAN	SCAN→DOS		SCAN→BOT*	
		Init	Base	P4Pir	Base	P4Pir
DT	ACC	0.987	0.604	0.932	0.900	0.923
	F1	0.984	0.568	0.868	0.776	0.820
RF	ACC	0.989	0.731	0.942	0.987	0.989
	F1	0.985	0.027	0.869	0.964	0.987

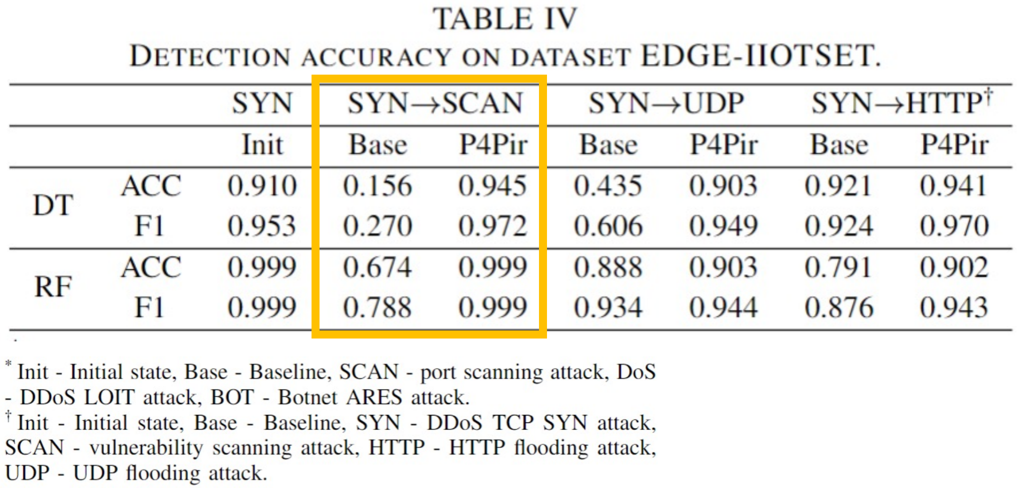
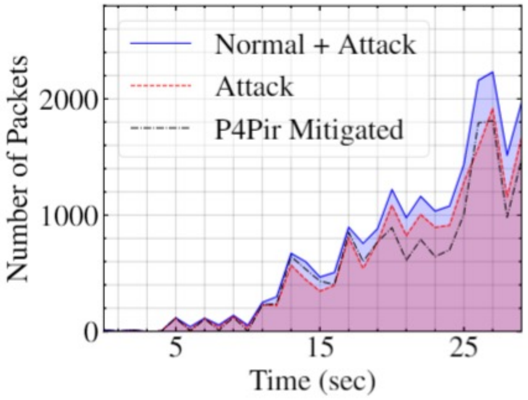


TABLE IV
DETECTION ACCURACY ON DATASET EDGE-IIOTSET.

		SYN	SYN→SCAN		SYN→UDP		SYN→HTTP†	
		Init	Base	P4Pir	Base	P4Pir	Base	P4Pir
DT	ACC	0.910	0.156	0.945	0.435	0.903	0.921	0.941
	F1	0.953	0.270	0.972	0.606	0.949	0.924	0.970
RF	ACC	0.999	0.674	0.999	0.888	0.903	0.791	0.902
	F1	0.999	0.788	0.999	0.934	0.944	0.876	0.943

* Init - Initial state, Base - Baseline, SCAN - port scanning attack, DoS - DDoS LOIT attack, BOT - Botnet ARES attack.
† Init - Initial state, Base - Baseline, SYN - DDoS TCP SYN attack, SCAN - vulnerability scanning attack, HTTP - HTTP flooding attack, UDP - UDP flooding attack.



(a) Mitigation performance.

Conclusion

We present P4Pir, an in-network ML-based analysis solution to defend against emerging threats on IoT gateway:

- Accurate ML-based traffic analysis inferred within the IoT gateway
- Swift mitigation of malicious traffic within forwarding data plane
- Continuous learning of emerging traffic patterns with runtime model updates

Further work:

- Distributed deployment of P4Pir

  e.g. Federated learning… FLIP4 [1]

Mingyuan

Questions?

Changgang

Questions?

[1] M. Zang et al., “Federated Learning-Based In-Network Traffic Analysis on IoT Edge,” IFIP Networking 2023 - Sec4IoT, 2023