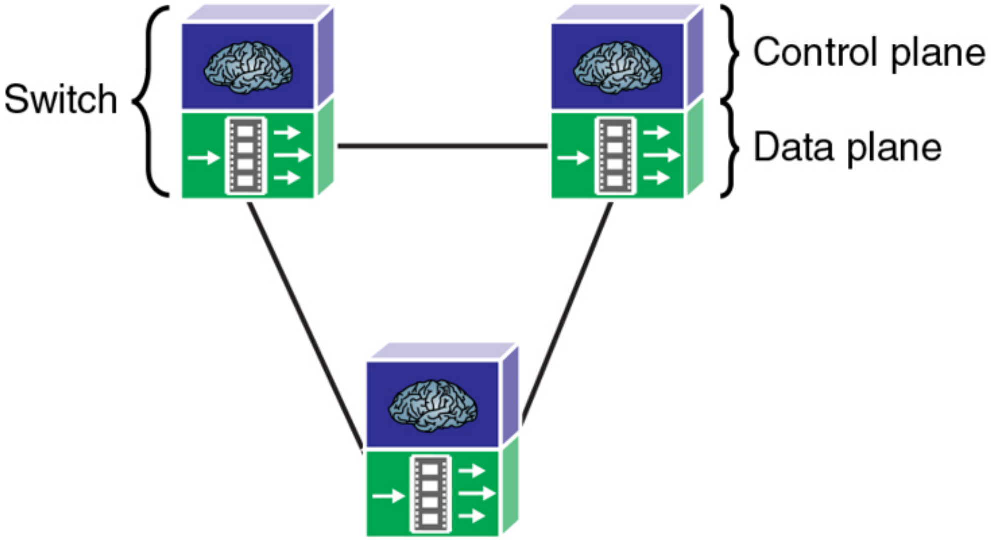
Role of Programmable Data Plane in Attack Detection

Dr. Mah-Rukh Fida
Senior Lecturer,
School of Business, Computing and Social Science
University of Gloucestershire, UK
mrukh@glos.ac.uk

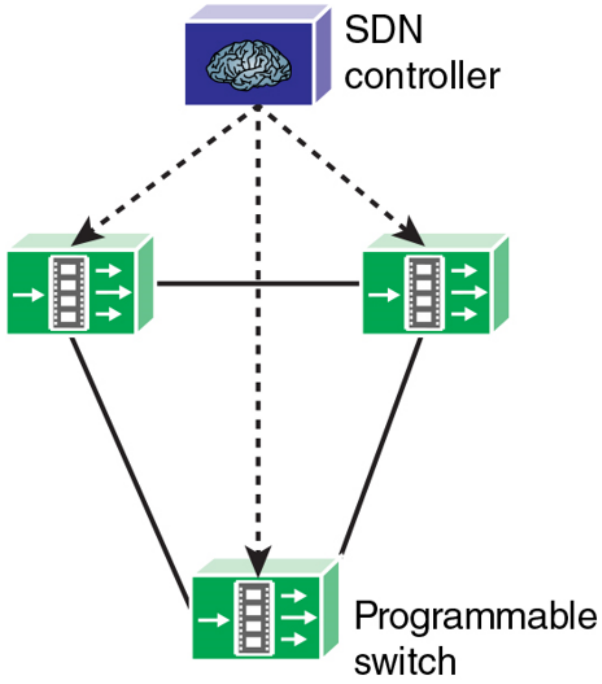
Agenda

1. Progress from Closed Traditional Networking to Programmable Networking
2. PDP: Challenges and Way-outs
3. Attack Detection
    - Statistical Measures
        - Entropy
        - Heavy Hitters
    - ML Methods
        - Packet Mirroring
        - In-network Monitoring
        - Hybrid Approach
4. Targeted research questions

1. Traditional to Software Defined Networking (SDN)



Traditional networks



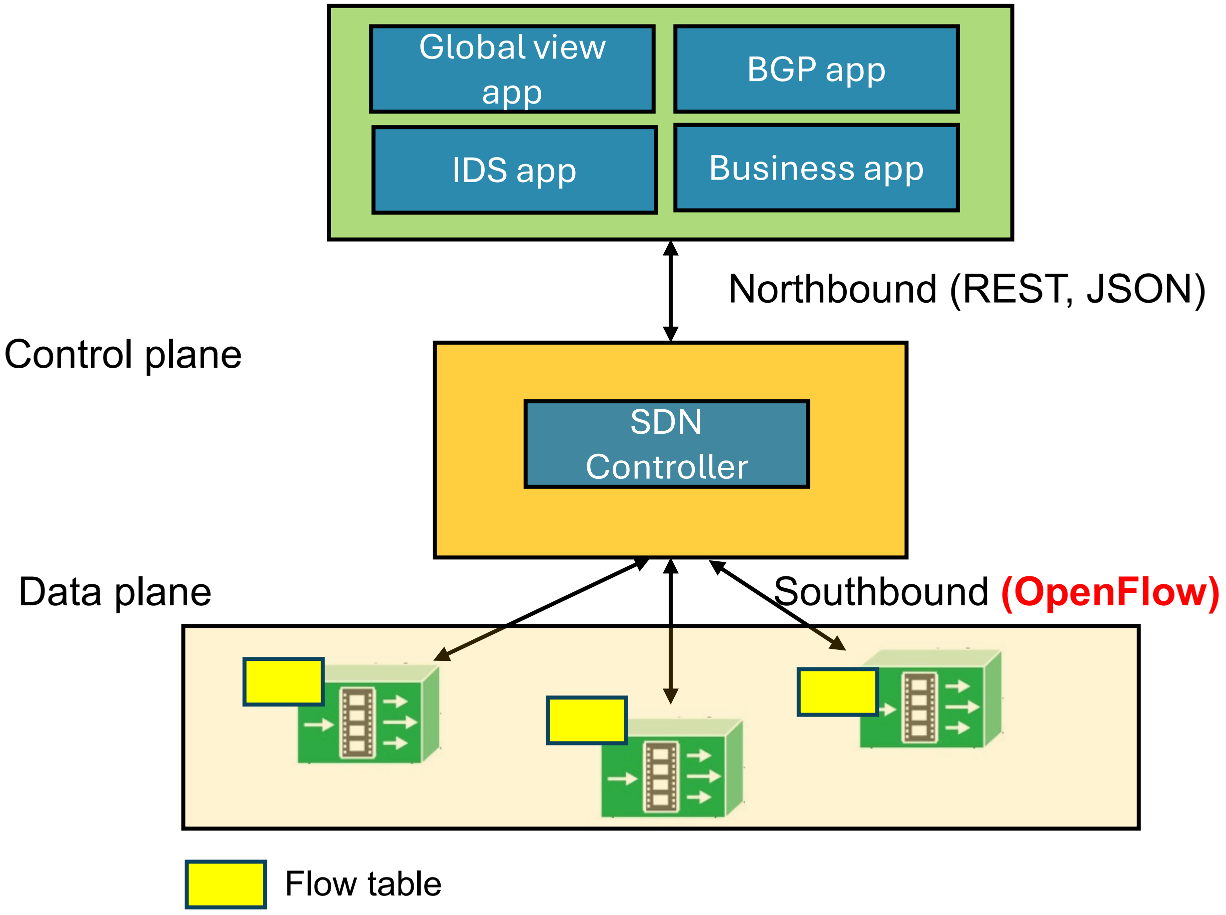
Software defined networks

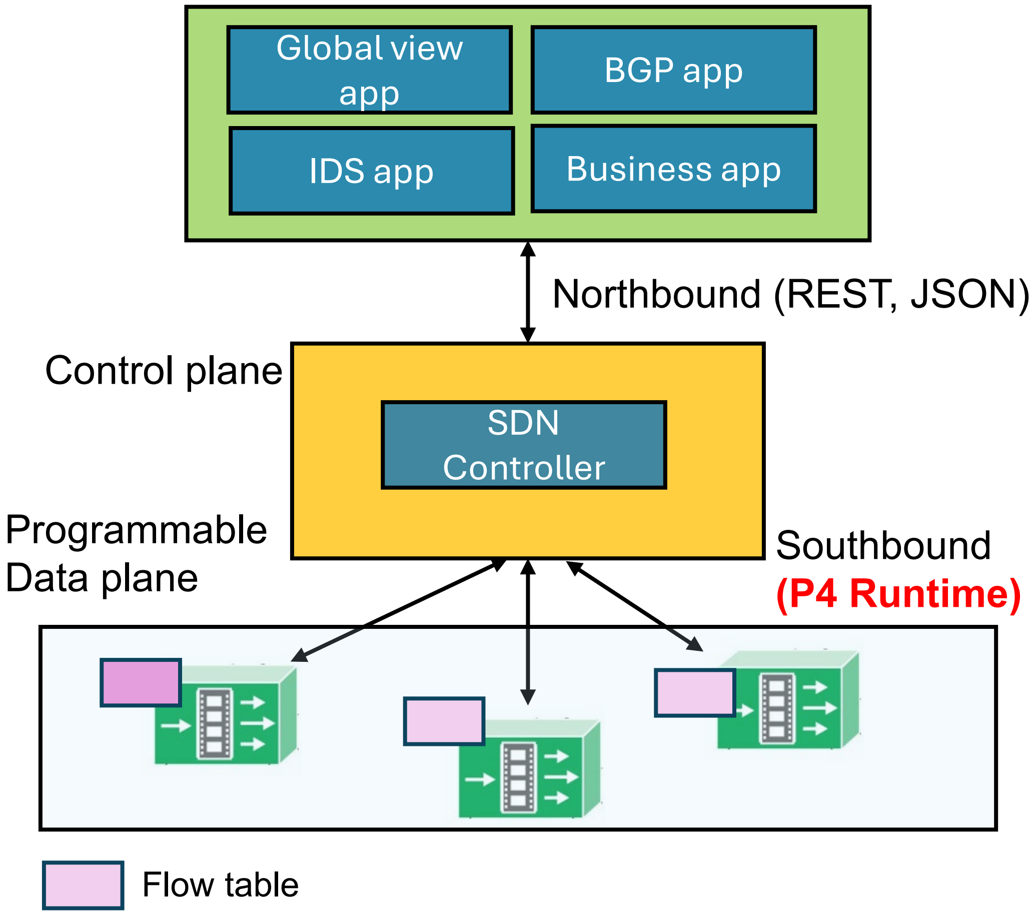
—— Packet flow

- - - - Packet-forwarding rules

W. Stallings, “Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud” Addison Wesley, 2017.

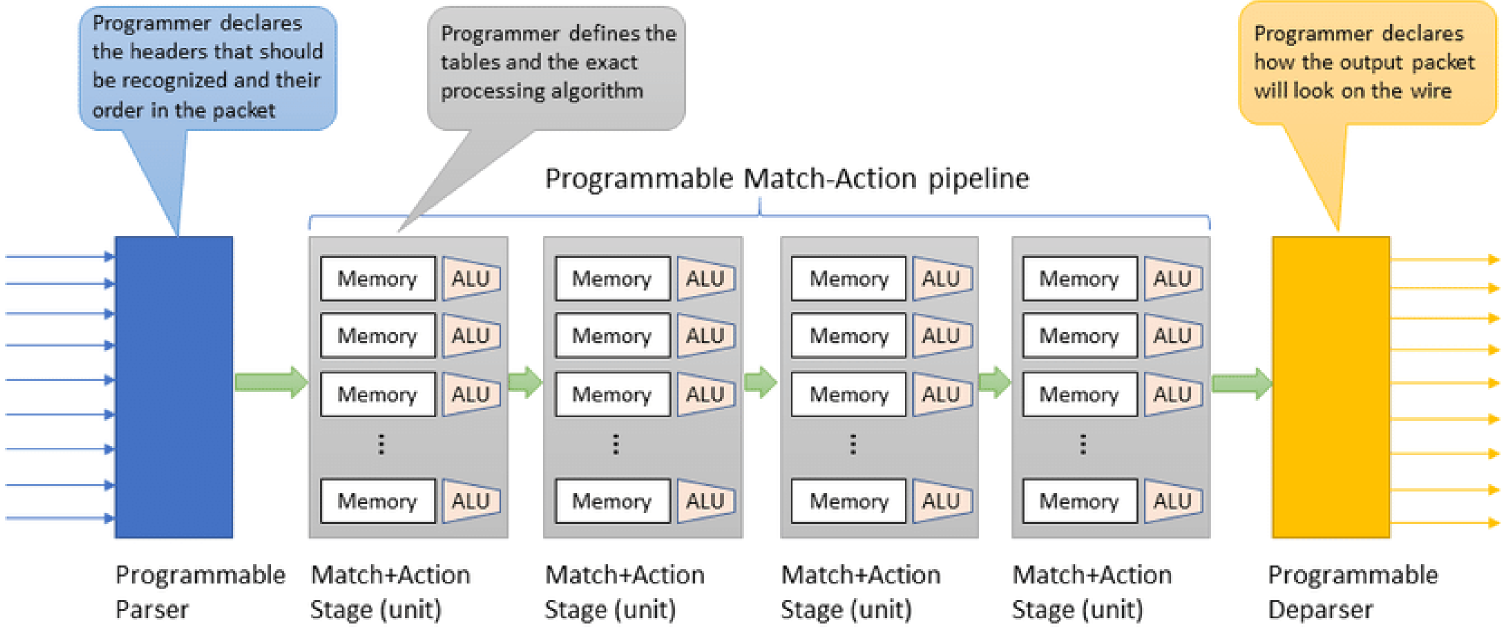
1. SDN using OpenFlow





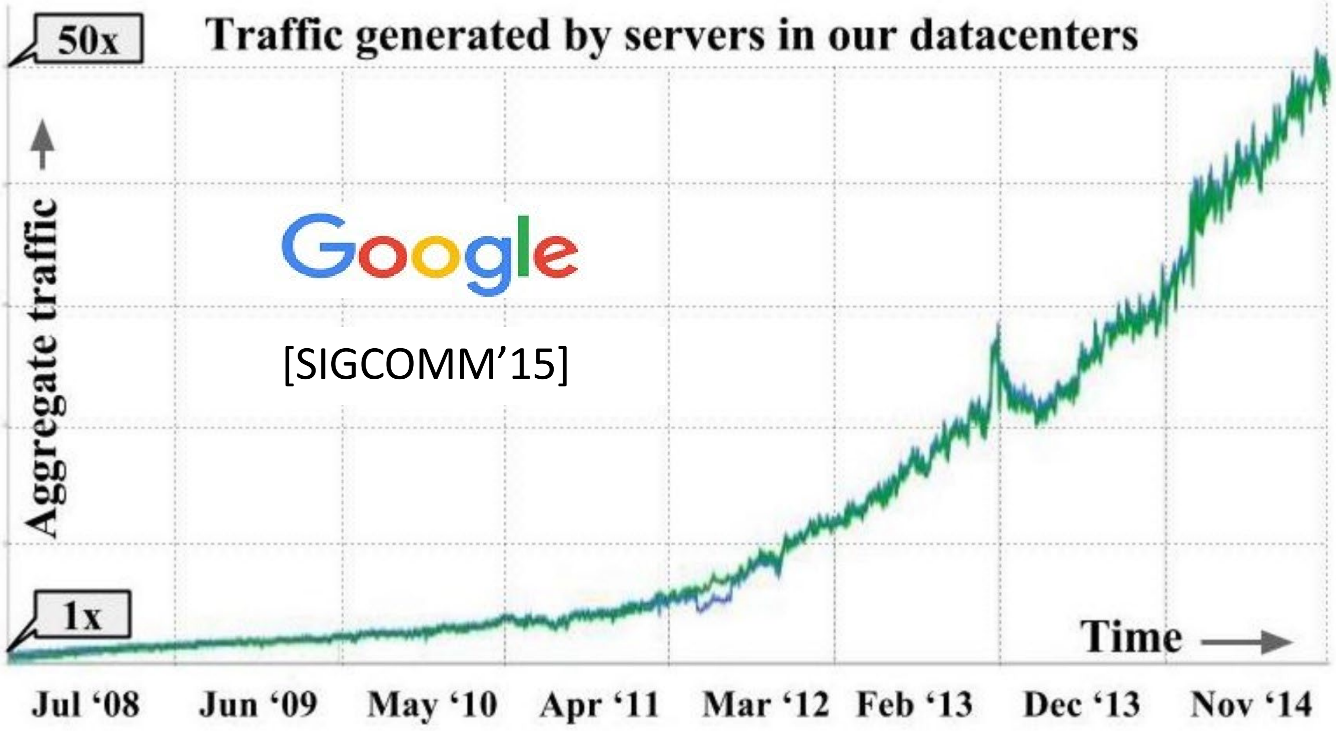
1. P4 programmable data plane (PDP)

- P4[1] programmable switches permit a programmer to program the data plane
  - Define and parse new protocols
  - Customize packet processing functions
  - Measure events occurring in the data plane with high precision
  - Offload applications to the data plane


Programmer declares the headers that should be recognized and their order in the packet
Programmer defines the tables and the exact processing algorithm
Programmer declares how the output packet will look on the wire
Programmable Match-Action pipeline
Memory
ALU
Programmable Parser
Match+Action Stage (unit)
Match+Action Stage (unit)
Match+Action Stage (unit)
Match+Action Stage (unit)
Programmable Deparser

[1] P4 stands for stands for Programming Protocol-independent Packet Processors

2. PDP Challenge I:
Growing Data vs Limited Memory



[SIGCOMM'15]

Year	Mem (MB)
2012	10-20
2014	30-60
2016	50-100

SilkRoad [SIGCOMM'17]

Significant data growth

Slow memory growth

2. PDP Challenge II: Program Complexity vs Limited Programmability

Load balancing

Security

Routing

Network Telemetry

Congestion Control

– No floating-point operations

– Statistical operations such as division, logs etc not allowed. No loops

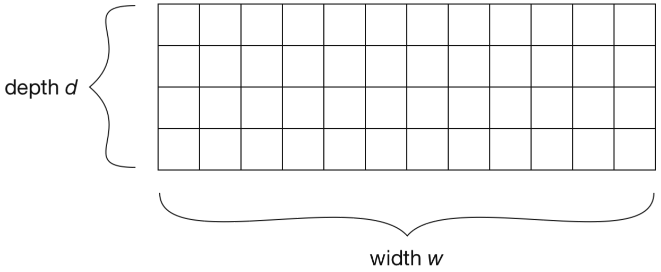
– Limited state sharing across stages

Diverse Use cases

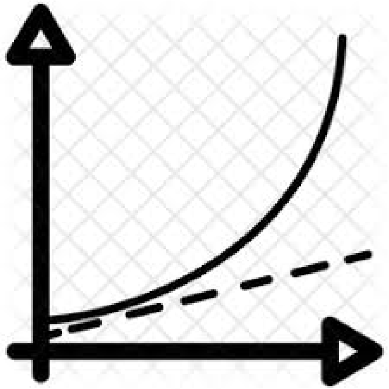
Limited Programmability

2. PDP: Dealing with Challenges

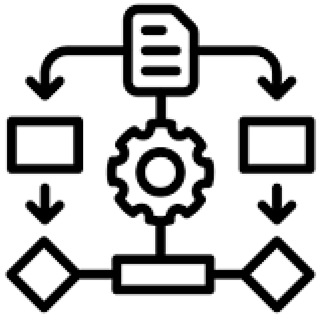
Sampling



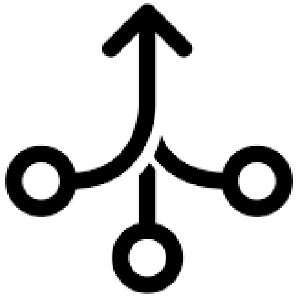
Sketches



approximation

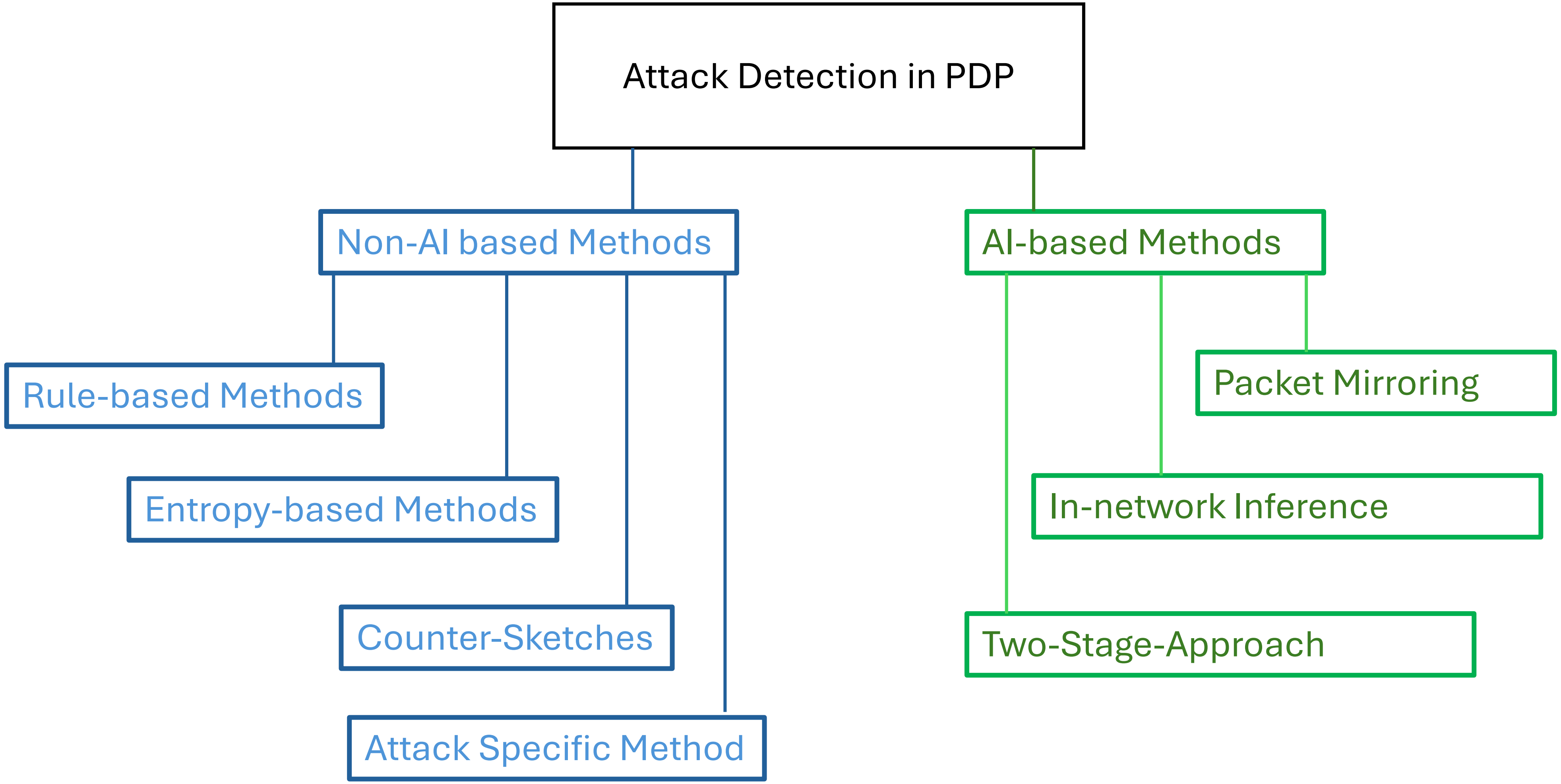


Distributed algorithms



Aggregation

3. Attack Detection in PDP



Attack Detection in PDP

Non-AI based Methods

Rule-based Methods

Entropy-based Methods

Counter-Sketches

Attack Specific Method

AI-based Methods

Packet Mirroring

In-network Inference

Two-Stage-Approach

3. Attack Detection in PDP

Non AI-based Methods

- These include methods that rely on known information about malicious services/nodes and the pattern of malicious attack
- These can target only a single pattern of an attack

AI-based Methods

- These methods target diverse patterns followed by either
- a single attack,
- a category of attack type or
- multiple attack categories
- It can be signature-based or an anomaly-based method

3.1 Non-AI based AD: Rule-based Methods

G. K. Ndonda et al. (2018) integrates P4 programmable switches within SDN framework to alleviate the burden on dedicated IDS in Industrial Control Systems (ICS).

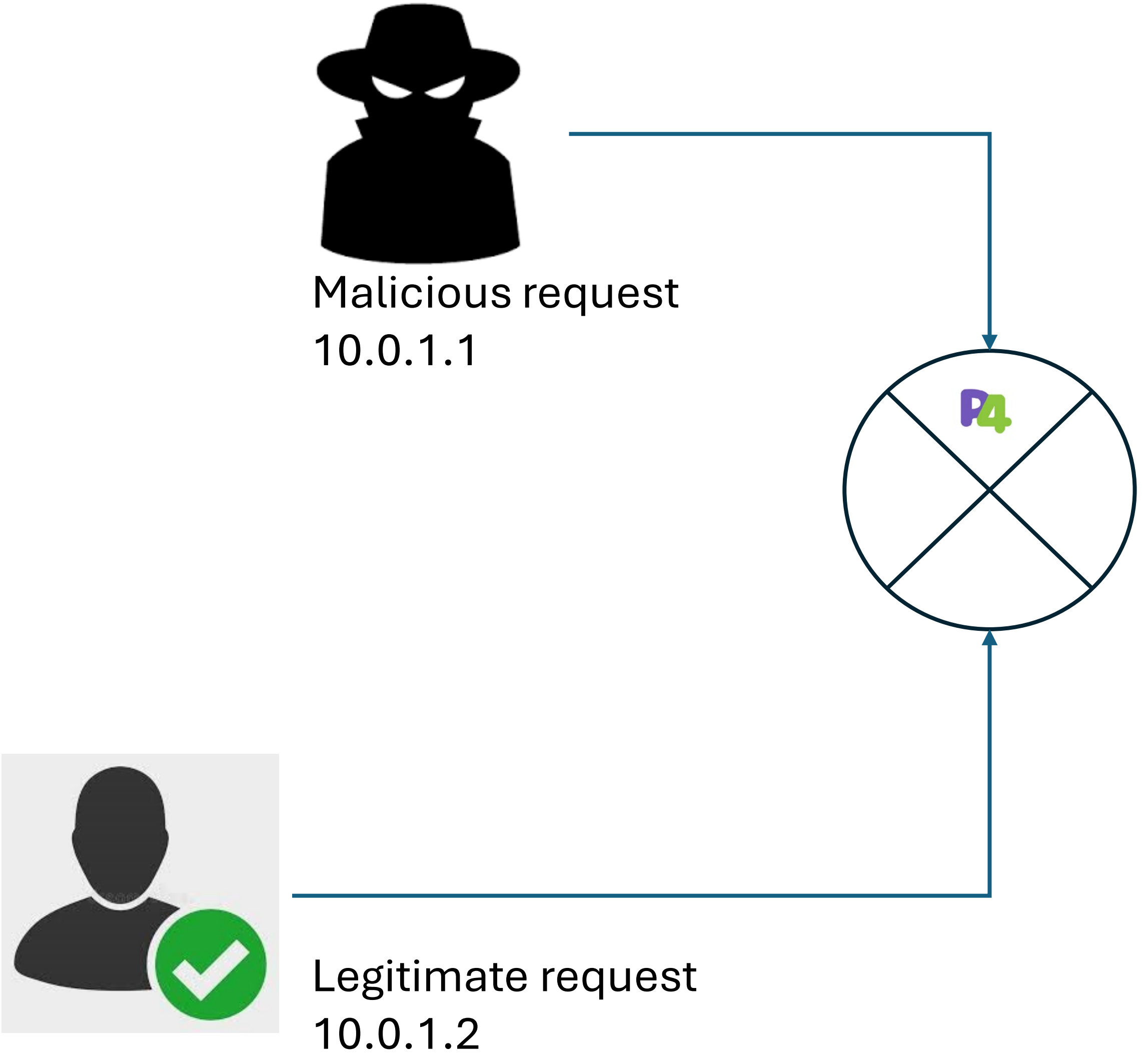
In this setup, match-action table rules contain **whitelisted Modbus** flow identifiers. If a traffic flow does not match the whitelist or exceeds an allowed packet length (a safeguard against buffer overflow attacks), it is forwarded to the IDS for deep packet inspection.

If the traffic is deemed legitimate, the SDN controller is notified and updates the whitelist in the switches accordingly.

Similarly, a blocking table on the P4 switch drops malicious flows.

Gorby Kabasele Ndonda and Ramin Sadre. *A two-level intrusion detectionsystem for industrial control system networks using P4*. In 5th InternationalSymposium for ICS & SCADA Cyber Security Research 2018 5, pages 31–40,2018

3.1 Non-AI based AD: Rule-based Methods

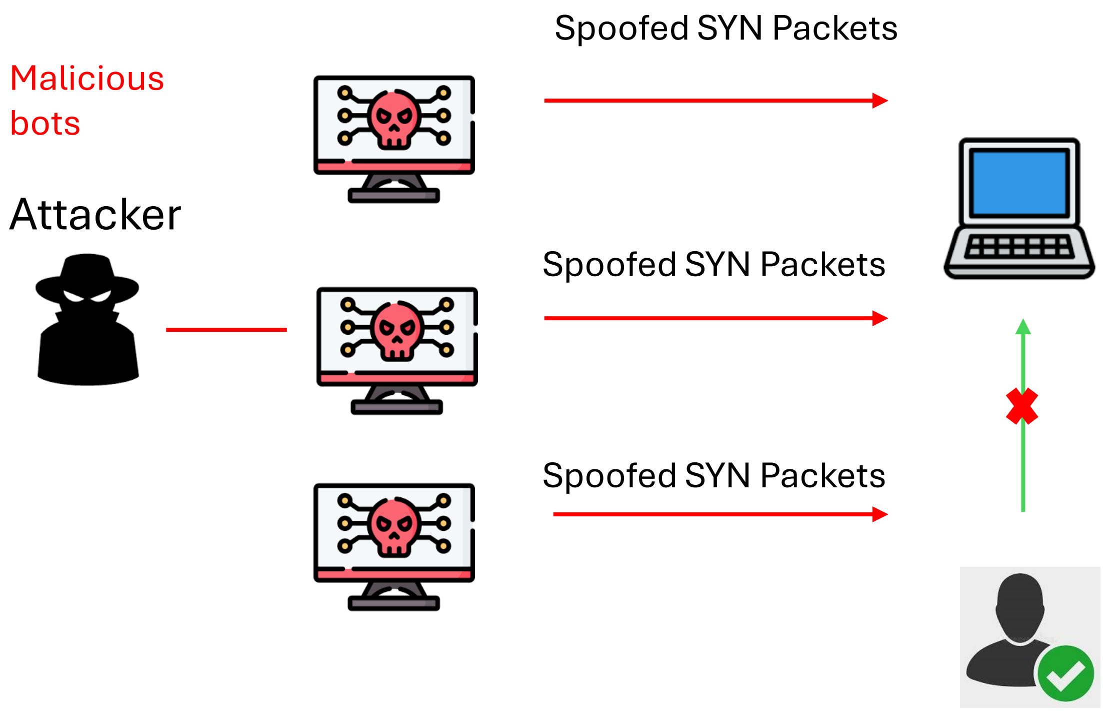


Malicious request
10.0.1.1

Legitimate request
10.0.1.2

Key	Action	Action Data
10.0.1.1/32	Forward	DesMAC:00:00:00:00:01:01 Address, Port = 1
10.0.1.2/32	Drop	
*	No Action/Mirror	

3.1 Non-AI based AD: Attack Specific Method



Spoofed SYN Packets

Malicious bots

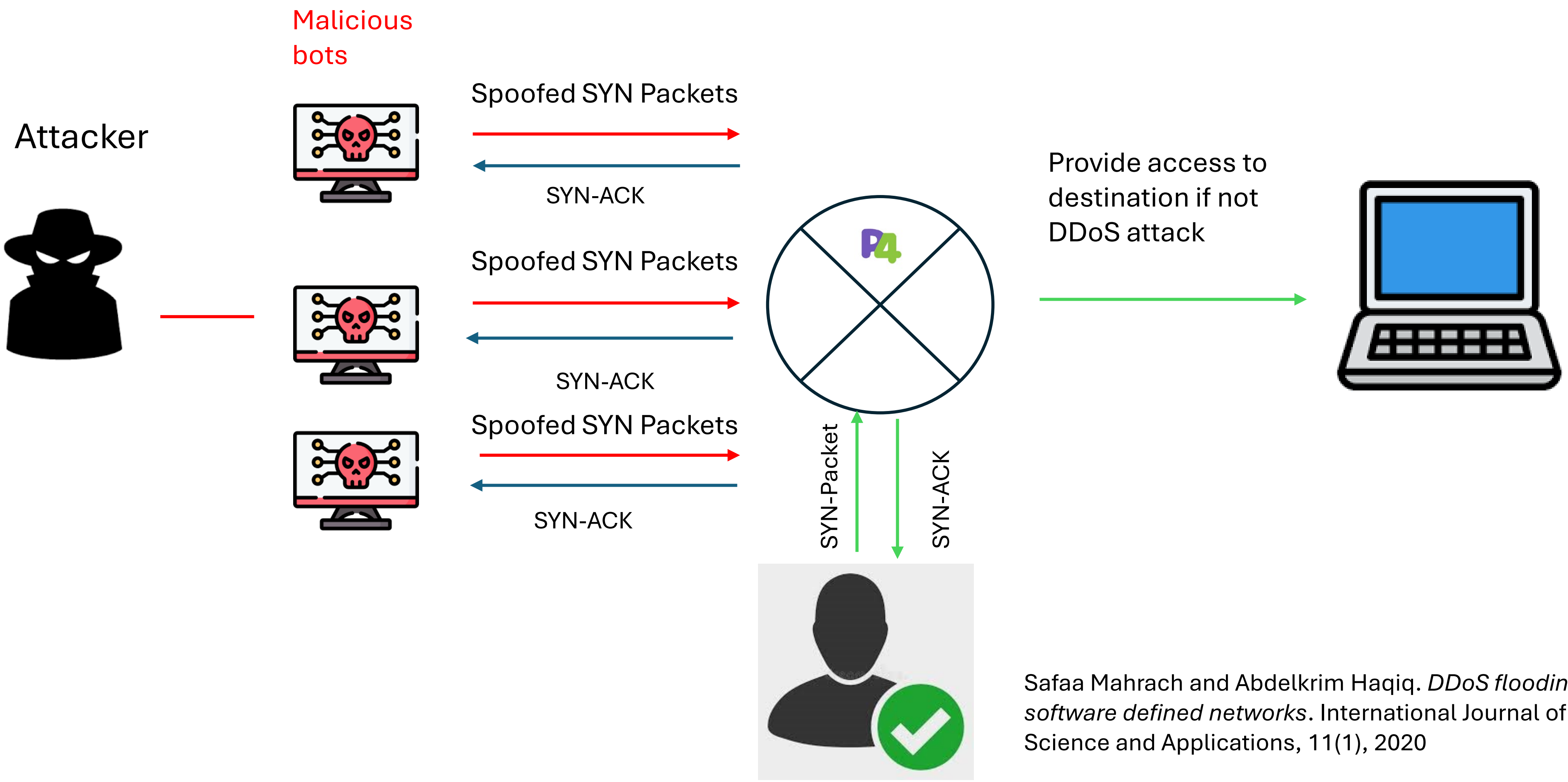
Attacker

Spoofed SYN Packets

Spoofed SYN Packets

TCP SYN Flood Attack Targeting a Destination

3.1 Non-AI based AD: Attack Specific Method



Safaa Mahrach and Abdelkrim Haqiq. *DDoS flooding attack mitigation in software defined networks*. International Journal of Advanced Computer Science and Applications, 11(1), 2020

3.1 Non-AI based AD: Entropy-based Methods

Entropy is a concept from information theory.

- It measures the uncertainty associated with a variable.
- It describes the degree of dispersion or concentration of a distribution.

$$H(X) = log_{2(m)} - \frac{1}{m} S(X)$$

$$S(X) = \sum_{x=1}^{N} f_x log_2(fx)$$

m: number of packets in a window
N=|X|, distinct set of IP addresses
f_x: frequency of an ip address x.
H(X) = 0, occurs when all the addresses are the same.
Dispersed distribution result in higher entropy values reaching the maximum H(X), when all addresses are distinct.

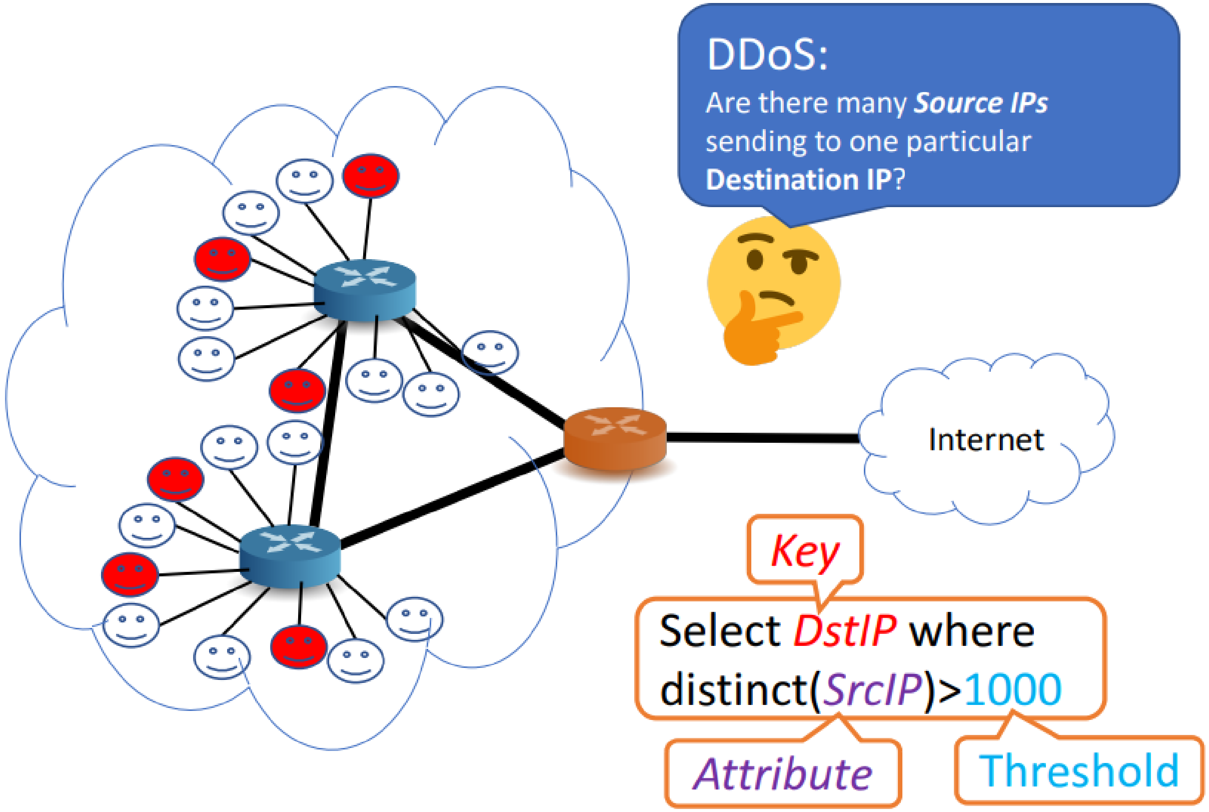
Shannon Entropy

Alexandre da Silveira Ilha, ˆAngelo Cardoso Lapolli, Jonatas Adilson Marques, and Luciano Paschoal Gaspary. ***Euclid:*** *A fully in-network, P4-based approach for real-time DDoS attack detection and mitigation*. IEEE Transactions on Network and Service Management, 18(3):3121–3139, 2020

3.1 Non-AI based AD: Entropy-based Methods

- EUCLID calculates entropies separately for the sets of source and destination IP addresses.

- In a given window, if the entropy of the source IP address exceeds a predefined threshold(derived from the EWMA and EWMMD of the source IP) while the entropy of the destination IP falls below a threshold, the packet stream is flagged as anomalous.



DDoS:
Are there many ***Source IPs*** sending to one particular **Destination IP**?

Select *DstIP* where distinct(*SrcIP*)>1000

Alexandre da Silveira Ilha, ˆAngelo Cardoso Lapolli, Jonatas Adilson Marques, and Luciano Paschoal Gaspary. ***Euclid:*** *A fully in-network, P4-based approach for real-time DDoS attack detection and mitigation*. IEEE Transactions on Network and Service Management, 18(3):3121–3139, 2020

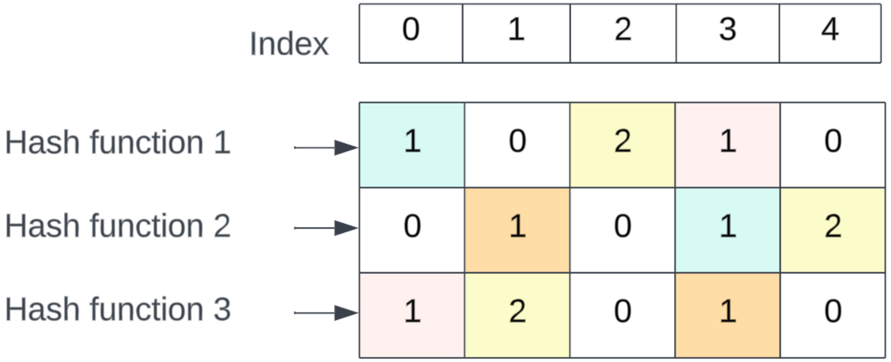
3.1 Non-AI based AD: Counter-Sketches Methods

Counter sketches such as Count-Min Sketch are data streaming algorithms, to track count-related information of flows, using stateful registers of P4.

In the data plane, sketches serve as counters based on universal streaming techniques, while the control plane provides APIs and libraries for applications to run estimation queries on the collected data.

UnivMon (2016) uses counter data structure to capture various traffic dimensions, such as source IPs, destination ports, destination IP appearing in different flows or number of packets per flow.

Applications include detecting heavy hitters, identifying DDoS attacks targeting specific destination IPs, and change detection—determining flows that contribute most to traffic variations over consecutive time intervals



Index	0	1	2	3	4
Hash function 1	1	0	2	1	0
Hash function 2	0	1	0	1	2
Hash function 3	1	2	0	1	0

A count-min sketch

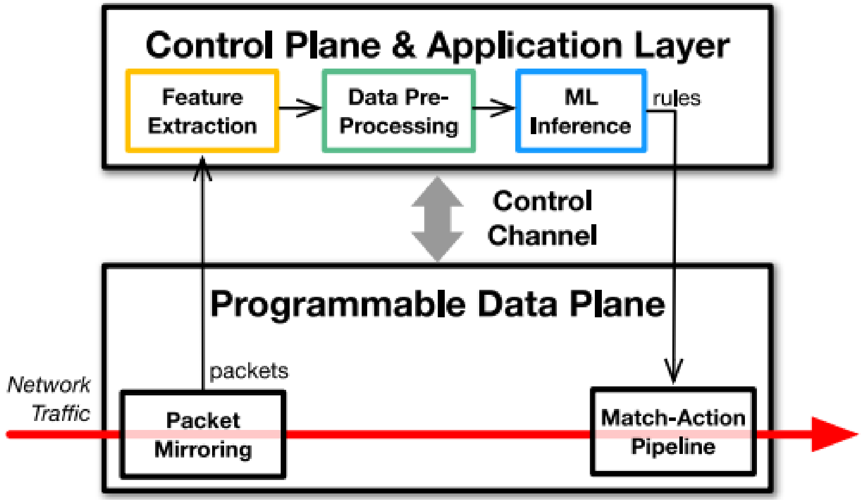
Zaoxing Liu, Antonis Manousis, Gregory Vorsanger, Vyas Sekar, andVladimir Braverman. One sketch to rule them all: Rethinking network flow monitoring with univmon. In Proceedings of the 2016 ACM SIGCOMM Conference, pages 101–114, 2016

3.2 AI-Based Methods: Packet Mirroring

- By employing packet mirroring (D. C Roberto et al 2024), a duplicate copy of the network traffic is transmitted from the data plane to the control plane, where it undergoes traffic feature extraction and pre-processing before ML algorithms can be executed upon it.

- Unfortunately, the channel between data and control planes often comes with severe bandwidth and latency bottlenecks.

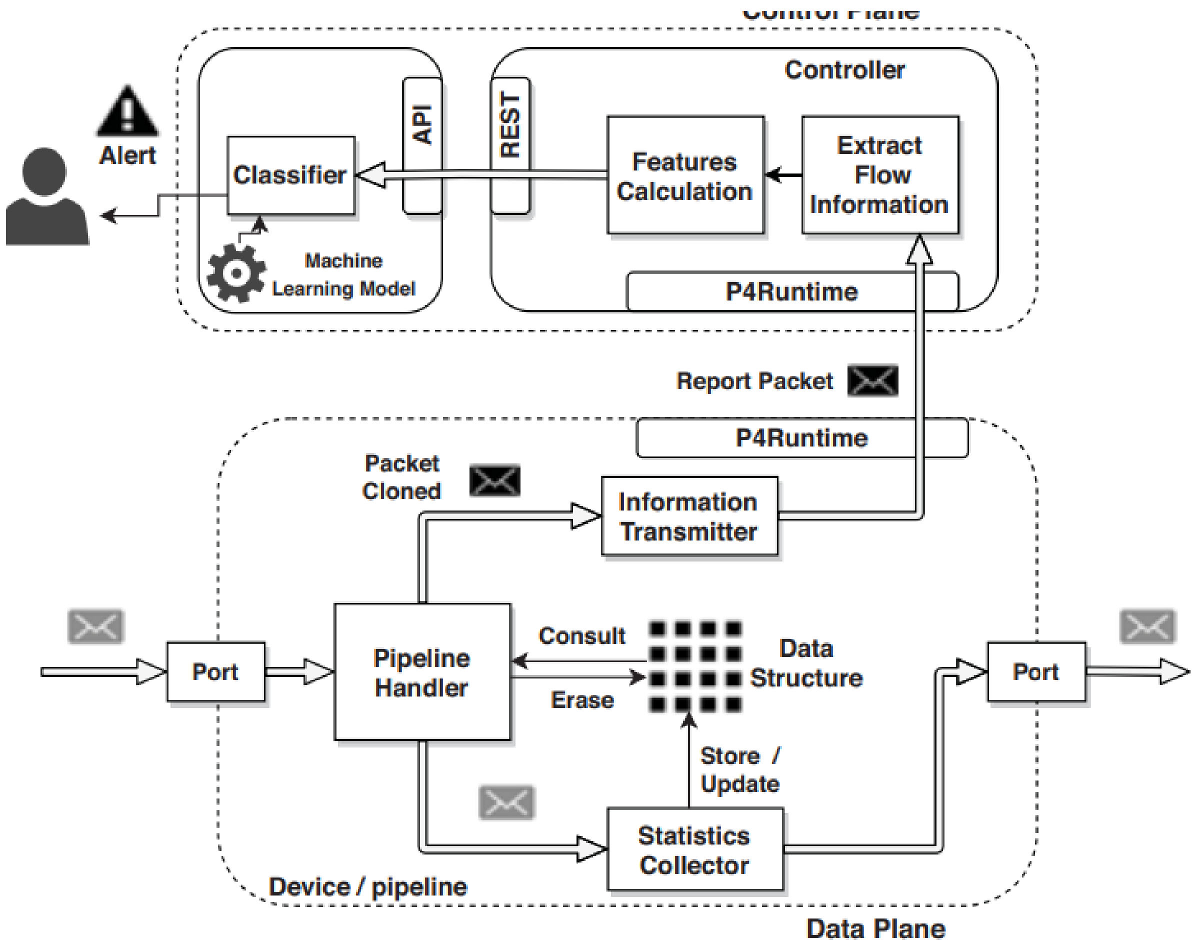
- These inherent limitations imply that it may be overwhelmed by the sheer volume of mirrored data, thereby jeopardising ordinary SDN network operations and compromising prompt response to ongoing network attacks.



Doriguzzi-Corin, Roberto, Luis Augusto Dias Knob, Luca Mendozzi, Domenico Siracusa, and Marco Savi. *Introducing packet-level analysis in programmable data planes to advance network intrusion detection*. Computer Networks 239 (2024): 110162.

3.2 AI-Based Methods: Packet Mirroring

- ORACLE (2021) introduces a DDoS detection mechanism with coordination between the data and control planes.
- The data plane in ORACLE is solely responsible for storing traffic information at a per-flow granularity using hash-based data structures. Periodically, digested traffic data is sent to the control plane.
- The control plane computes feature from the sent information about the flows. The classification of traffic based on these computed features is then carried out at the control plane.
- In ORACLE the data-plane and the control-plane operations are based on time window. The control plane then applies a trained binary classifier on the computed features



Sebasti ́an G ́omez Mac ́ıas, Luciano Paschoal Gaspary, and Juan FelipeBotero. *Oracle: An architecture for collaboration of data and control planes to detect DDoS attacks*. In 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM), pages 962–967,2021

3.2 AI-Based Methods: In-network Inference

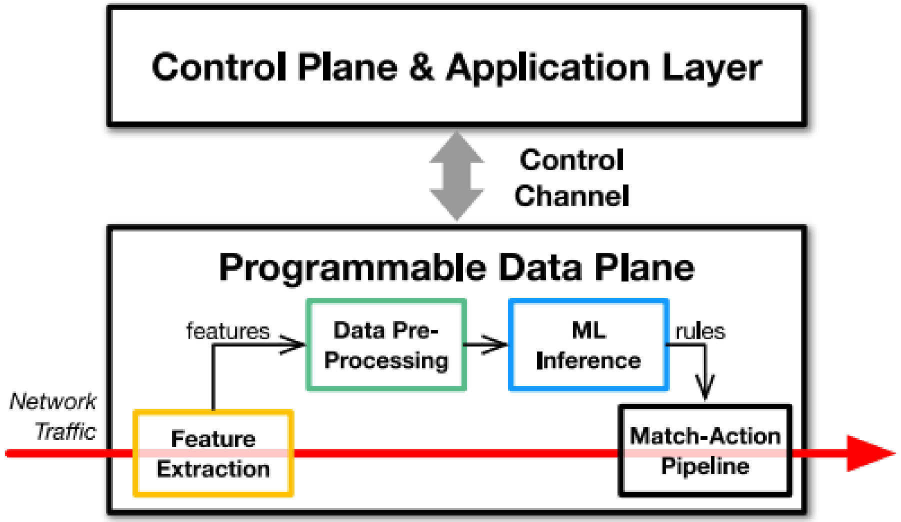
Here feature extraction as well as ML inference is done at the data plane.

Pros:

- Inference at line rate
- Early detection and mitigation

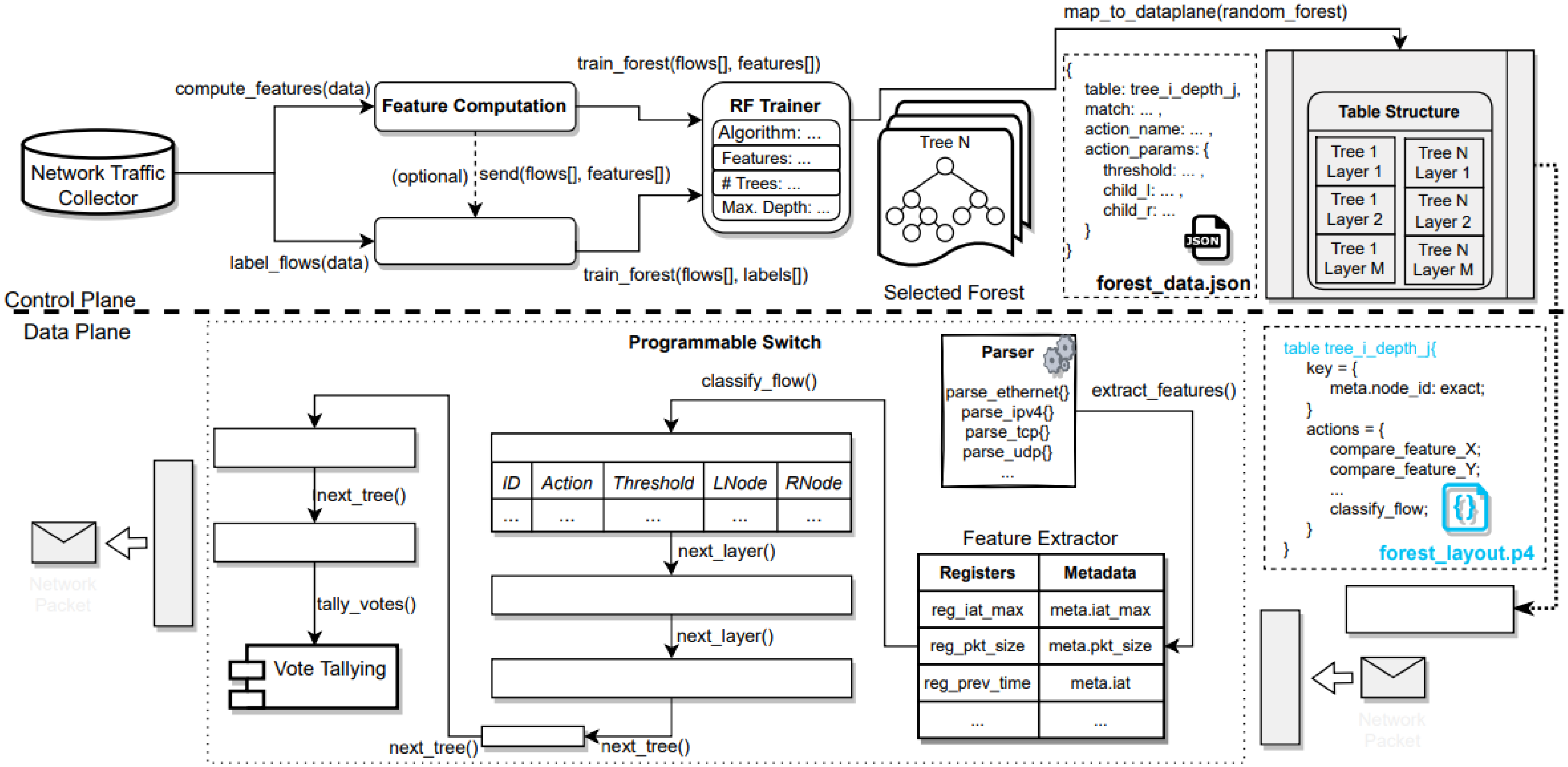
Cons

- Limited memory capacity (for match-action tables)
- Limited arithmetic operations
- Simple ML models which may not result in accurate inference



3.2 AI-Based Methods: In-network Inference

The Architecture of BACKORDERS: the main modules run in the control plane (RF Mapper) and in the data plane/programmable switch (feature extractor and online classifier)



RF Trainer
Algorithm: ...
Features: ...
Trees: ...
Max. Depth: ...

```
{
  table: tree_i_depth_j,
  match: ... ,
  action_name: ... ,
  action_params: {
    threshold: ... ,
    child_l: ... ,
    child_r: ...
  }
}
```

forest_data.json

Table Structure	
Tree 1 Layer 1	Tree N Layer 1
Tree 1 Layer 2	Tree N Layer 2
Tree 1 Layer M	Tree N Layer M

ID	Action	Threshold	LNode	RNode
...	...	...	...	...

Parser

parse_ethernet{}
parse_ipv4{}
parse_tcp{}
parse_udp{}
...

Registers	Metadata
reg_iat_max	meta.iat_max
reg_pkt_size	meta.pkt_size
reg_prev_time	meta.iat
...	...

```
table tree_i_depth_j{
  key = {
    meta.node_id: exact;
  }
  actions = {
    compare_feature_X;
    compare_feature_Y;
    ...
    classify_flow;
  }
}
```

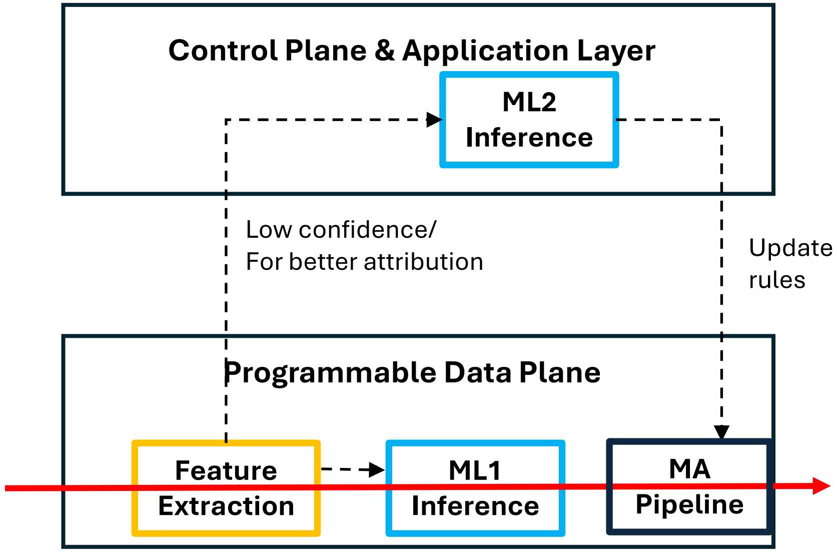
forest_layout.p4

Bruno Coelho and Alberto Schaeffer-Filho. *Backorders: using Random forests to detect DDoS attacks in programmable data planes*. In Proceedings of the 5th International Workshop on P4 in Europe, pages1–7, 2022

3.2 AI-Based Methods: Two-Stage Approach

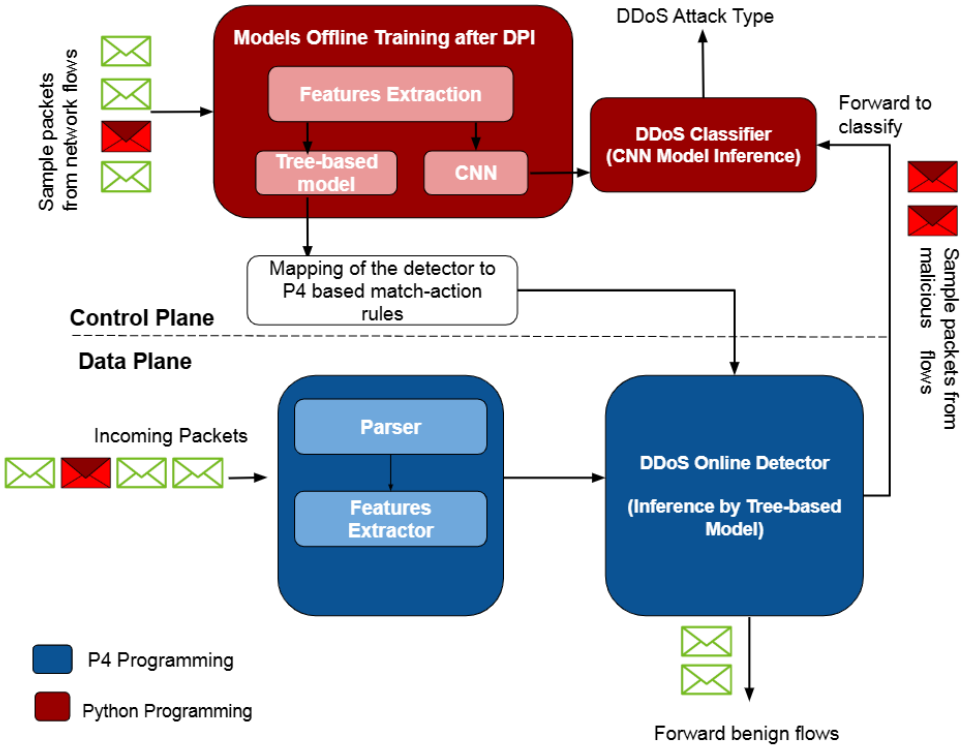
- Control plane is Intelligent but Slow
- Data Plane is quick but limited in resources

- Why not use both in conjunction:
  - Distributed inference?
  - Statistical feature computation at control plane



3.2 AI-Based Methods: Two-Stage Approach

- DDoShield (2024), aimed to not only detect but attribute different types of DDoS attacks, including
  - Volumetric (UDP Flood, ICMP Flood, and TCPFlood)
  - Protocol based (ACK Flood, SYN Flood, PUSH Flood, RST/FIN Flood, ACK Fragmentation, UDP Fragmentation etc)
  - Application based (HTTP Flood, Slowloris)
- Detector of Attack: At data plane with a lightweight ML (Decision Tree)
- Classification of Attack Flows: ResNet
- Rule Update for Blocking the attack Type

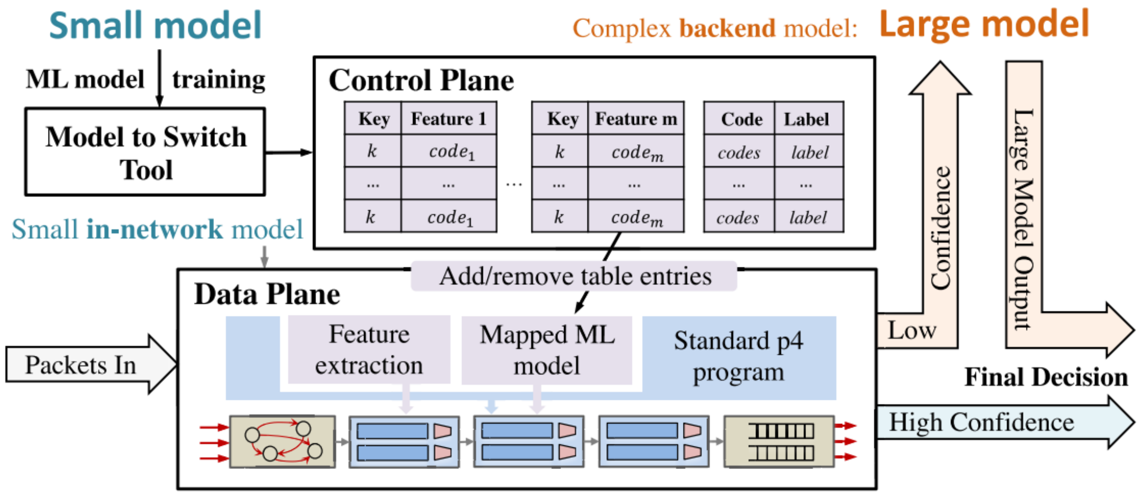

Sample packets from network flows
Models Offline Training after DPI
Features Extraction
Tree-based model
CNN
DDoS Attack Type
DDoS Classifier (CNN Model Inference)
Forward to classify
Sample packets from malicious flows
Mapping of the detector to P4 based match-action rules
Control Plane
Data Plane
Incoming Packets
Parser
Features Extractor
DDoS Online Detector
(Inference by Tree-based Model)
P4 Programming
Python Programming
Forward benign flows

Architecture of DDoS Shield

Azza H Ahmed, Mah-Rukh Fida, and Ameer Shakayb Arsalaan. *Ddoshield: In-network defensive architecture against volumetric and non-volumetric DDoS attacks*. Authorea Preprints, 2024

3.2 AI-Based Methods: Two-Stage Approach

- C. Zheng et al. (2024) proposed, Ilsy, a two-tier machine learning (ML) framework combining a lightweight Random Forest (RF) at network switches for classification tasks such as anomaly detection.
- Additionally, a more comprehensive back-end model, such as an RF with 200 trees and an extensive feature set, is deployed on a server. The primary objective of Ilsy is to minimize the load and latency on the back-end server.
- It achieves this by directly forwarding benign packets with high detection confidence to their destination ports, while redirecting anomalous packets or those with low confidence to the server for detailed packet inspection.
- Using a 2-stage deployment, Ilsy reduces the load on the backend, achieves high throughput, low latency, and high ML performance



Key	Feature 1
k	*code*$_1$
...	...
k	*code*$_1$

Key	Feature m
k	*code*$_m$
...	...
k	*code*$_m$

Code	Label
codes	*label*
...	...
codes	*label*

The high-level architecture of Ilsy.

C. Zheng *et al.*, *Ilsy: Hybrid In-Network Classification Using Programmable Switches*, in *IEEE/ACM Transactions on Networking*, vol. 32, no. 3, pp. 2555-2570, June 2024, doi: 10.1109/TNET.2024.3364757

4. Targeted research questions:

In-Network Attack Detection Models

- Most Used
  - Decision Tree
  - Random Forest
  - XGBoot
- Less frequently used
  - Binarized Neural Network

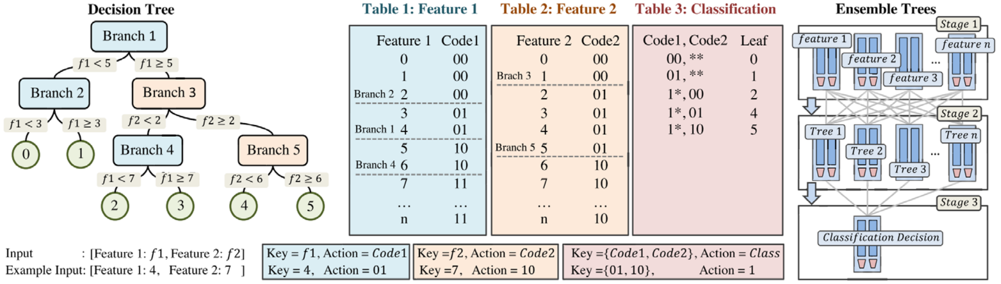


Table 1: Feature 1

	Feature 1	Code1
	0	00
	1	00
Branch 2	2	00
	3	01
Branch 1	4	01
	5	10
Branch 4	6	10
	7	11
	…	…
	n	11

Table 2: Feature 2

	Feature 2	Code2
	0	00
Brach 3	1	00
	2	01
	3	01
	4	01
Branch 5	5	01
	6	10
	7	10
	…	…
	n	10

Table 3: Classification

Code1, Code2	Leaf
00, **	0
01, **	1
1*, 00	2
1*, 01	4
1*, 10	5

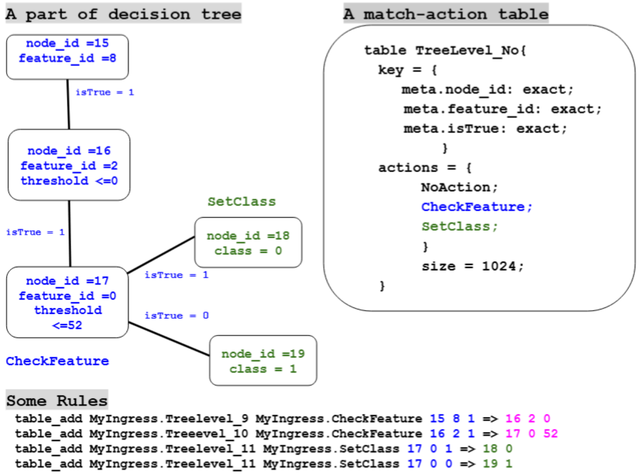
Input : [Feature 1: $f1$, Feature 2: $f2$]
Example Input: [Feature 1: 4, Feature 2: 7]

Key = $f1$, Action = $Code1$
Key = 4, Action = 01

Key =$f2$, Action = $Code2$
Key =7, Action = 10

Key ={$Code1$, $Code2$}, Action = $Class$
Key ={01, 10}, Action = 1

Mapping of a Decision Tree and Ensemble Model to Stages (source *)



```
table TreeLevel_No{
  key = {
     meta.node_id: exact;
     meta.feature_id: exact;
     meta.isTrue: exact;
          }
  actions = {
        NoAction;
        CheckFeature;
        SetClass;
        }
        size = 1024;
  }
```

Some Rules

```
table_add MyIngress.Treelevel_9 MyIngress.CheckFeature 15 8 1 => 16 2 0
table_add MyIngress.Treeevel_10 MyIngress.CheckFeature 16 2 1 => 17 0 52
table_add MyIngress.Treelevel_11 MyIngress.SetClass 17 0 1 => 18 0
table_add MyIngress.Treelevel_11 MyIngress.SetClass 17 0 0 => 19 1
```

Part of a Decision Tree, with equivalent rules to run by a match-action table (source &)

* C. Zheng *et al*., *IIsy: Hybrid In-Network Classification Using Programmable Switches*, in *IEEE/ACM Transactions on Networking*, vol. 32, no. 3, pp. 2555-2570, June 2024, doi: 10.1109/TNET.2024.3364757

& Azza H Ahmed, Mah-Rukh Fida, and Ameer Shakayb Arsalaan. *Ddoshield: In-network defensive architecture against volumetric and non-volumetric DDoS attacks*. Authorea Preprints, 2024

4. Targeted research questions:

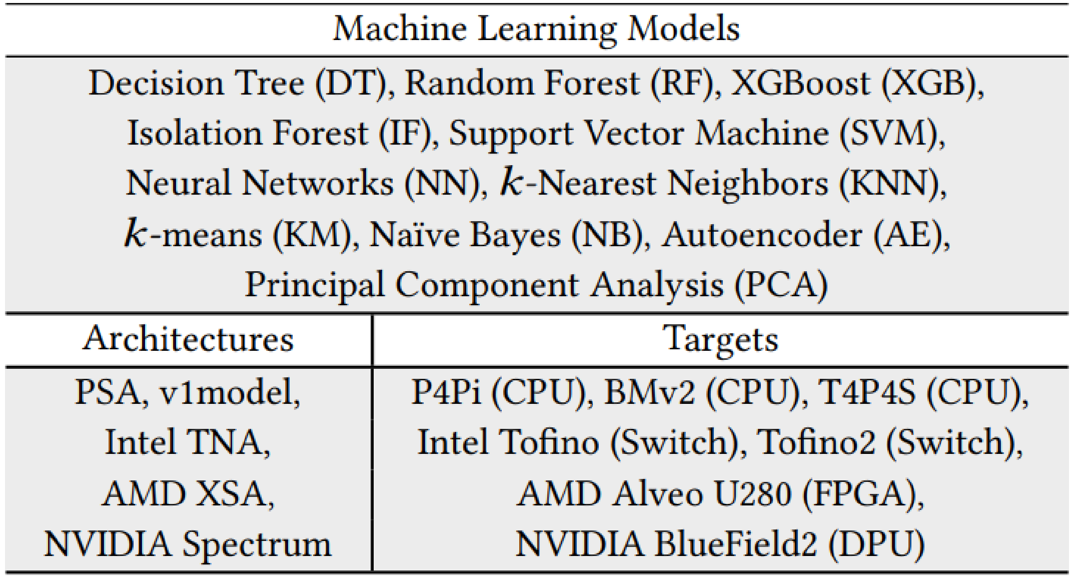
- Similar some other areas of research are:
  - Whether to use packet or flow-based features?
  - At what phase of the flow, to initiate detection?
  - How to ensure high throughput, high accuracy with low burden on memory?
  - How to approximate statistical computation on derived features?
  - How to detect emerging attacks as well as known attacks?
  - Applicability of PDP in attack detection at IoT, Cellular, Cloud and Data Centres along with SDN.

What is happening?

In-Network Attack Detection Models: Planter

- Growing use of Planter (see *)

- Planter (C. Zheng et al., 2024) is a modular framework for one-click implementation of in-network ML algorithms



<table>
<tr><th colspan="2">Machine Learning Models</th></tr>
<tr><td colspan="2">Decision Tree (DT), Random Forest (RF), XGBoost (XGB), Isolation Forest (IF), Support Vector Machine (SVM), Neural Networks (NN), k-Nearest Neighbors (KNN), k-means (KM), Naïve Bayes (NB), Autoencoder (AE), Principal Component Analysis (PCA)</td></tr>
<tr><th>Architectures</th><th>Targets</th></tr>
<tr><td>PSA, v1model, Intel TNA, AMD XSA, NVIDIA Spectrum</td><td>P4Pi (CPU), BMv2 (CPU), T4P4S (CPU), Intel Tofino (Switch), Tofino2 (Switch), AMD Alveo U280 (FPGA), NVIDIA BlueField2 (DPU)</td></tr>
</table>

Models, architectures and targets supported by Planter

* Zang, Mingyuan, et al. "Federated In-Network Machine Learning for Privacy-Preserving IoT Traffic Analysis." *ACM Transactions on Internet Technology* 24.4 (2024): 1-24.
* Hong, Xinpeng, Changgang Zheng, and Noa Zilberman. "In-network machine learning for real-time transaction fraud detection." *ECAI 2024*. IOS Press, 2024. 2902-2909.

Zheng, Changgang, et al. "Planter: Rapid prototyping of in-network machine learning inference." *ACM SIGCOMM Computer Communication Review* 54.1 (2024): 2-21.

Summary

- Networking moved from Traditional to SDN (OpenFlow) to SDN(P4)
- Attack Detection in PDP:
  - Non-AI based: Targeting a single characteristics
  - AI-based: Using multiple characteristics
- Tree-based: Mostly
- Areas of research:
  - Use sophisticated ML/DL -> Accuracy
  - Reduce latency, improve throughput
  - Line rate processing, decrease memory load
  - Cater diverse attacks
  - Application in emerging technologies