Machine Learning for SDN security: improving intrusion and vulnerability detection

AI for Good

Evren Tuna
6G R&D Engineer, ULAK Communications Inc.

29.05.2023

HISTORY

ULAK COMMUNICATIONS INC.

The Beginning of the ULAK Project
February 2013

The Beginning of the SDN Project
November 2015

Foundation of ULAK Communication by SSTEK
April 2017

The First Live Site of ULAK
May 2018

The Beginning of ULAK 5G Project
May 2018

MAYA is in Commercial Use
May 2019

The ULAK's 1000. Site
January 2020

The First Live Site Abroad
October 2021

The Beginning of ULAK 6G Project
January 2022

The Beginning of ULAK 6G R&D
March 2022

2000+ Installed+ Contracted ULAK MAYA SD-WAN Sites
July 2022

3000+ On Air Contracted ULAK LTE Base Stations
July 2022

Young but Experienced

Massive but Agile

Advanced but Excited as day-one

PRODUCTS SOLUTIONS

Radio Access Network

- LTE-A Base Station
- 5G Base Station
- OAM

Core Network

- ÇINAR 5G Core
- EPC for Private LTE

SDN

- SD-WAN (MAYA)
- SD-DC

Vertical Solutions

- V2X – Vehicle to Everything
- Mission Critical Communications
- Public Safety
- Private Networks

What is SDN?

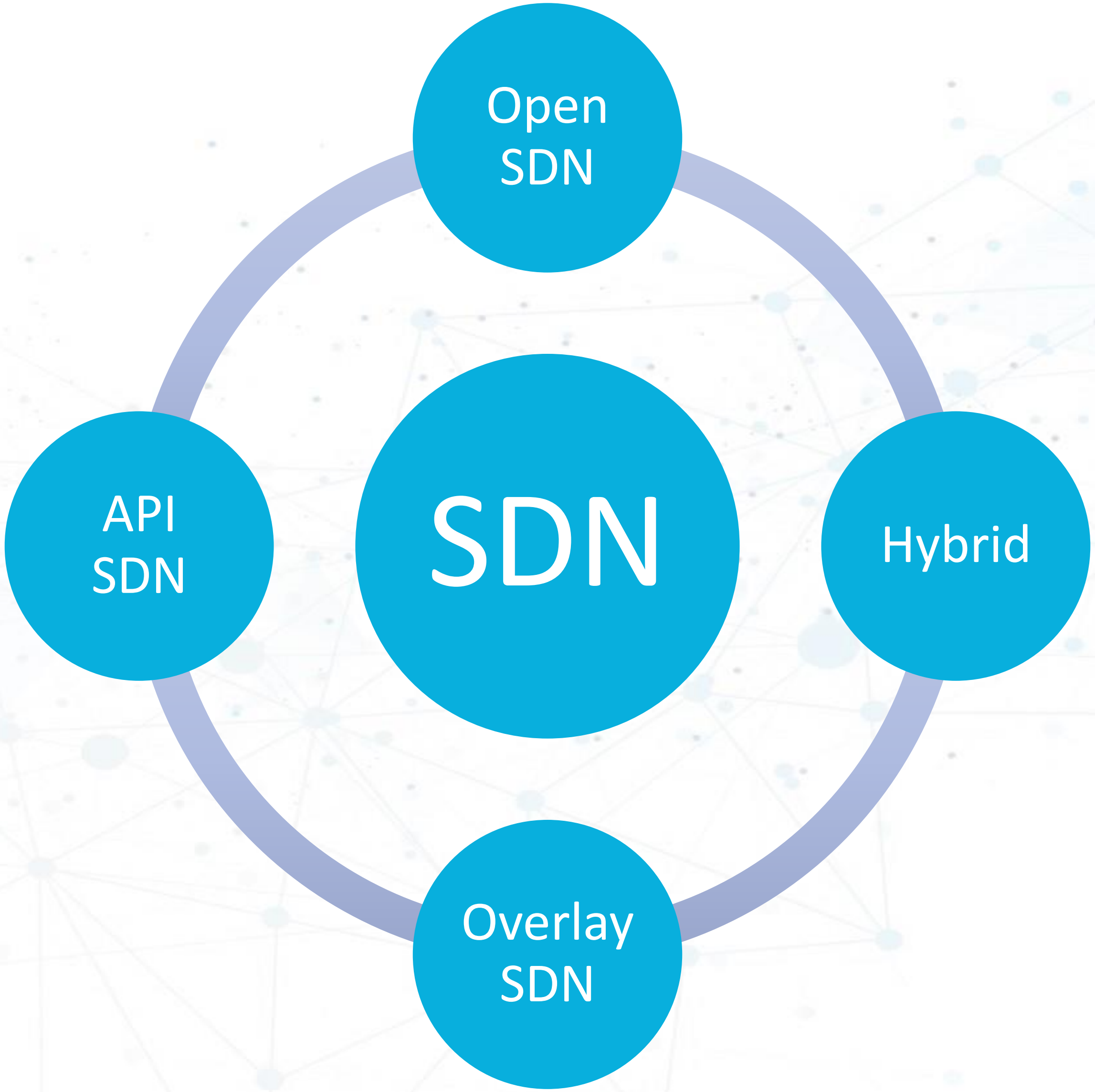
- To make a network more flexible and easier to manage
- Centralizes management by abstracting the control plane from the data forwarding function
- Delivers a centralized, programmable network

Controller	Southbound APIs	Northbound APIs

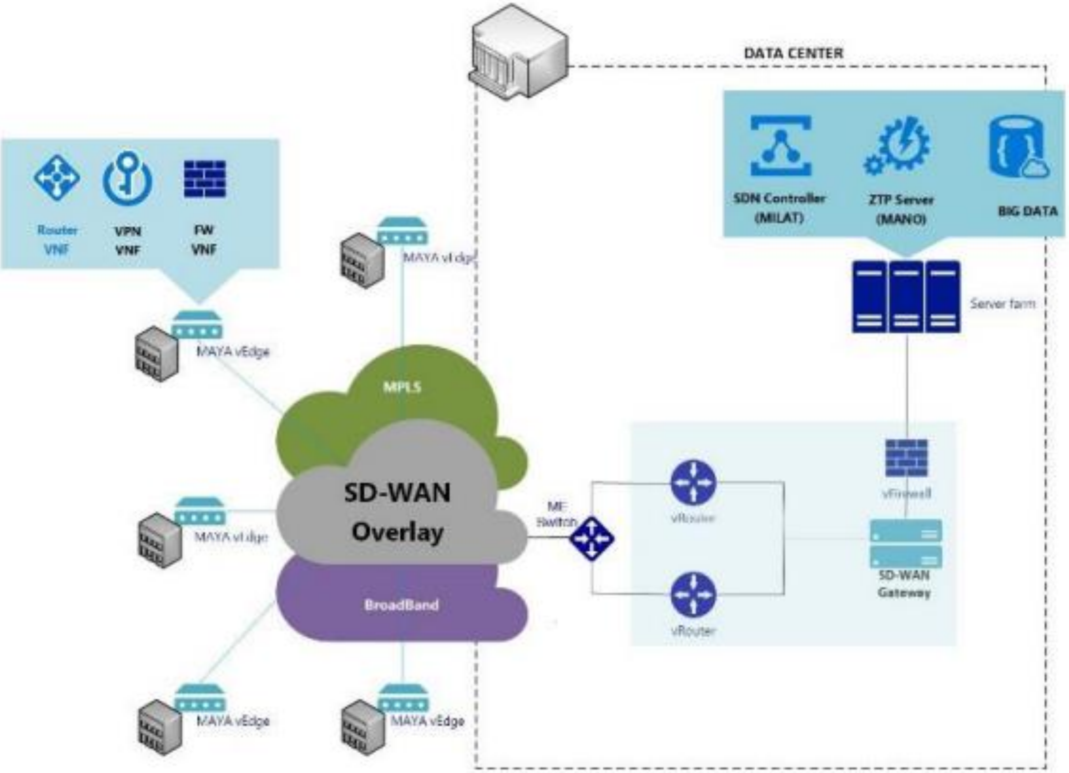
Benefits of SDN

- Ease of network control
- Agility
- Flexibility
- Greater control over network security
- Simplified network design and operation

Types of SDN



MAYA SD-WAN



- MAYA SDN Controller
- MAYA ZTP Server
- MAYA BIG-DATA Platform
- MAYA vEDGE
- MAYA SD-WAN Gateway

MAYA SD-WAN

ULAK COMMUNICATIONS INC.

Routing

- BGPv4
- BGP Multipath
- OSPFv2
- RIPv1 and RIPv2
- SLA Aware Routing
- Application Aware Routing
- Static Routing
- Multicast*
- Segment Routing**
- Hybrid WAN

Tunnelling

- VxLAN
- IPSec VPN
- SSL VPN
- L2 over GRE
- L2TPv3
- DMVPN
- Multiprotocol Label Switching (MPLS)**
- Full Mesh / Partial Mesh Hub and Spoke VPN deployments

QoS

- Application Aware QoS
- Bandwidth Limitation
- Bandwidth Dedication
- Random Early Detection (RED)
- Weighted Random Early Detection (WRED)
- DSCP Classification
- Two-Way Active Measurement Protocol (TWAMP)**

Network Services

- DHCP Server/Client/Relay
- NAT mapping
- DNS Forwarding
- Dynamic DNS

Security

- Stateful Inspection Firewall
- Granular Access Control
- ICMP Type Filtering
- 802.1x
- RADIUS / LDAP/TACACS+
- Web Proxy
- URL/Content Filtering

Zero Touch Provisioning

- Device Authentication
- Remote Configuration
- Remote VNF Provisioning
- Zero Touch Replacement
- Configuration Backup/Restore
- Group Based Deployment

Monitoring & Management

- Topology Monitoring
- Auto Discovery
- Topology Health status
- Network Inventory
- Netflow,Sflow,Flow Statistics
- SNMPv2,v3,SysLog
- NBI(Rest API,Web Socket)
- Multi User /Role Support WEB Based GUI
- Multi Tenancy
- SPAN/RSPAN

** Development in progerss*

*** SDN Controller feature*

Attacks & Vulnerabilities in SDN

ULAK COMMUNICATIONS INC.

Unauthorized Access

Attacks & Vulnerabilities in SDN

- Unauthorized Access
- Data Leakage

Attacks & Vulnerabilities in SDN

- Unauthorized Access
- Data Leakage
- Data Modification

Attacks & Vulnerabilities in SDN

ULAK COMMUNICATIONS INC.

- Unauthorized Access
- Data Leakage
- Data Modification
- Malicious/Compromised Applications

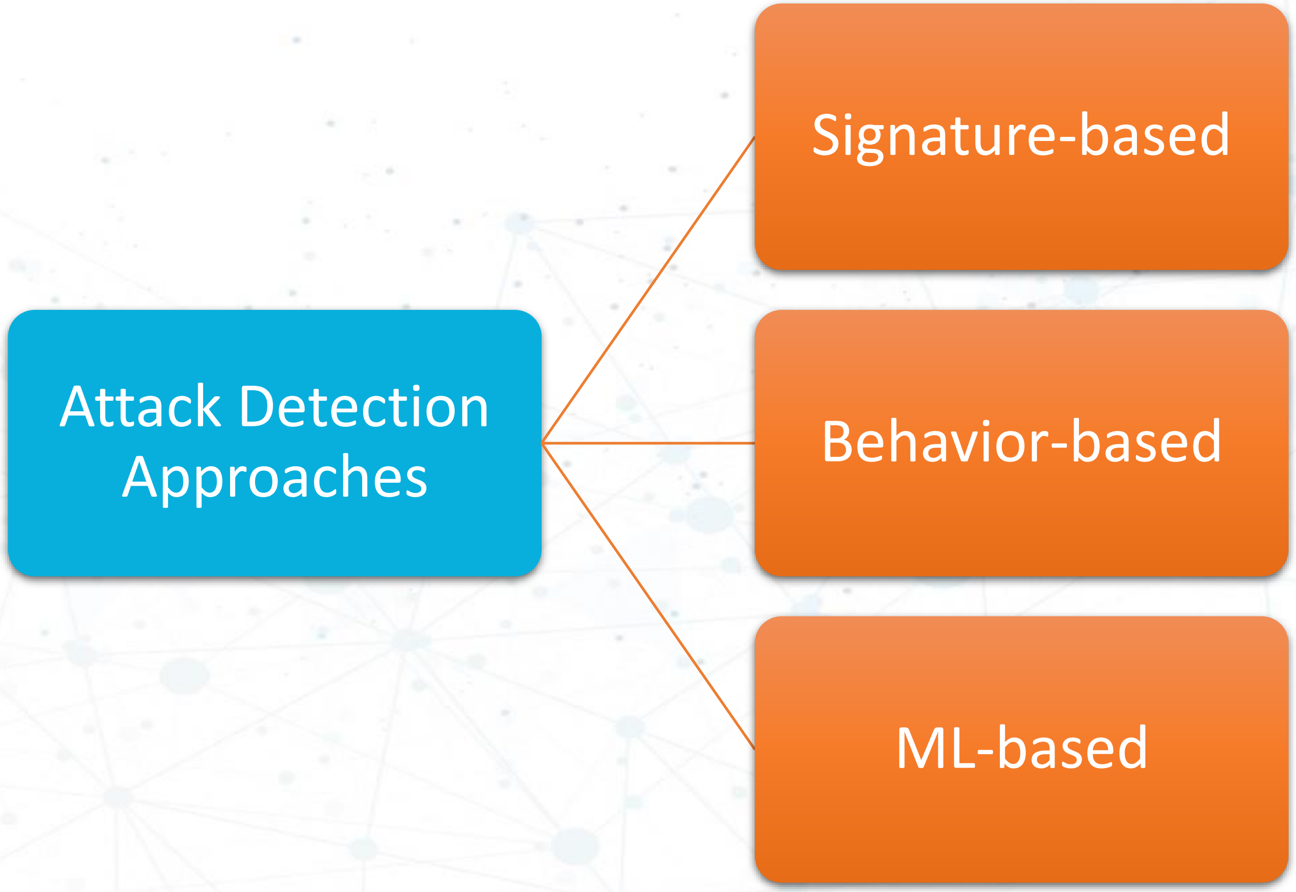
Attacks & Vulnerabilities in SDN

- Unauthorized Access
- Data Leakage
- Data Modification
- Malicious/Compromised Applications
- Denial of Service

Attacks & Vulnerabilities in SDN

- Unauthorized Access
- Data Leakage
- Data Modification
- Malicious/Compromised Applications
- Denial of Service
- Configuration Issues

ML-based SDN


Attack Detection Approaches
Signature-based
Behavior-based
ML-based

Signature-based

- It monitors the incoming network traffic to identify sequences and patterns that match a specific attack signature.

Pros:

- Easy-to-implement method on the network
- Proactively detects specific attacks in advance
- Operates quickly
- Significantly low false positive rate

Cons:

- A slight change in the signature can lead to the failure of anomaly detection
- Cannot detect previously unknown attacks
- Constant updating of signatures to combat new types of attacks

Behavior-based

- Rather than searching for specific patterns associated with certain attack types, it monitors behaviors that may be correlated with attacks.

Pros:

- High sensitivity in anomaly detection
- Does not require signature updates for new attack types
- Does not need to expose itself externally for improvement and development

Cons:

- It may consider ongoing attacks as normal traffic
- Due to the analysis required for implementation on the network, it demands high effort for deployment

ML-based

- Using ML, it analyzes data and monitors network traffic for network breaches.

Pros:

- Detect new types of attacks
- Independent on external sources for improvement and development
- It is open to performance enhancements
- It does not require pre-analysis of the network
- It is easy to implement with SDN

Cons:

- The false positive rate may be high
- The processing requirements are relatively higher compared to signature-based detection

ML-based SDN

Supervised Learning	Unsupervised Learning	Reinforcement Learning	Semi-supervised Learning
• K-Nearest Neighbour • Decision Tree • Random Forest • Neural Network • Support Vector Machine • Bayes' Theory • Hidden Markov Model	• K-Means • Self-Organizing Map	• Reinforcement Learning • Deep Reinforcement Learning • RL-based Game Theory	

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: Coarse-grained intrusion detection
Method: DT, RF

Input: 10 features
Output: 2 classes: normal and anomaly

- C. Song et al., “Machine-learning based threat-aware system in software defined networks,” in Proc. IEEE ICCCN, Vancouver, BC, Canada, Jul./Aug. 2017, pp. 1–9.

Objective: Coarse-grained intrusion detection
Method: HMM

Input: 5 features
Output: 2 classes: normal and anomaly

- T. Hurley, J. E. Perdomo, and A. Perez-Pons, “HMM-based intrusion detection system for software defined networking,” in Proc. IEEE ICMLA, Anaheim, CA, USA, Dec. 2016, pp. 617–621.

Objective: Coarse-grained intrusion detection
Method: SVM

Input: IP address, Transport port
Output: 2 classes: normal and anomaly

- A. S. da Silva, J. A. Wickboldt, L. Z. Granville, and A. Schaeffer-Filho, “ATLANTIC: A framework for anomaly traffic detection, classification, and mitigation in SDN,” in Proc. IEEE NOMS, Istanbul, Turkey, Apr. 2016, pp. 27–35.

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: Coarse-grained intrusion detection
Method: SVM

Input: 3 features
Output: 2 classes: normal and anomaly

- M. Nobakht, V. Sivaraman, and R. Boreli, "A host-based intrusion detection and mitigation framework for smart home IoT using OpenFlow," in Proc. IEEE ARES, Salzburg, Austria, Aug./Sep. 2016, pp. 147–156.

Objective: Coarse-grained intrusion detection
Method: DT, BayesNet, decision table, Naïve Bayes

Input: 4 features
Output: 2 classes: normal and anomaly

- S. Nanda, F. Zafari, C. DeCusatis, E. Wedaa, and B. Yang, "Predicting network attack patterns in SDN using machine learning approach," in Proc. IEEE NFV-SDN, Palo Alto, CA, USA, Nov. 2016, pp. 167–172.

Objective: Coarse-grained intrusion detection
Method: Deep NN

Input: 6 features
Output: 2 classes: normal and anomaly

- T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in Proc. IEEE WINCOM, Fes, Morocco, Oct. 2016, pp. 258–263.

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: Coarse-grained intrusion detection
Method: Recurrent NN

Input: 6 features
Output: 2 classes: normal and anomaly

- T. Tang, S. A. R. Zaidi, D. McLernon, L. Mhamdi, and M. Ghogho, “Deep recurrent neural network for intrusion detection in SDN-based networks,” in Proc. IEEE NetSoft, Montreal, QC, Canada, 2018, pp. 1–5.

Objective: Fine-grained intrusion detection
Method: SVM

Input: 23 features
Output: 5 classes: normal, DoS, U2R, R2L, Probe

- P. Wang, K.-M. Chao, H.-C. Lin, W.-H. Lin, and C.-C. Lo, “An efficient flow control approach for SDN-based network threat detection and migration using support vector machine,” in Proc. IEEE ICEBE, Macau, China, Nov. 2016, pp. 56–63.

Objective: Fine-grained intrusion detection
Method: RF

Input: 41 features
Output: 5 classes: normal, DoS, U2R, R2L, Probe

- N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, “A deep learning approach to network intrusion detection,” IEEE Trans. Emerg. Topics Comput. Intell., vol. 2, no. 1, pp. 41–50, Feb. 2018.

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: DDoS attack detection
Method: SOM

Input: 6 features
Output: 2 classes: normal and DDoS

- R. Braga, E. Mota, and A. Passito, “Lightweight DDoS flooding attack detection using NOX/OpenFlow,” in Proc. IEEE LCN, Denver, CO, USA, Oct. 2010, pp. 408–415.

Objective: DDoS attack detection
Method: Naïve Bayes, k-NN, k-means, k-medoids

Input: -
Output: 5 classes: normal and DDoS

- L. Barki, A. Shidling, N. Meti, D. G. Narayan, and M. M. Mulla, “Detection of distributed denial of service attacks in software defined networks,” in Proc. IEEE ICACCI, Jaipur, India, Sep. 2016, pp. 2576–2581.

Objective: DDoS attack detection
Method: Deep NN

Input: 20 features
Output: 2 classes: normal and DDoS

- C. Li et al., “Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN,” Int. J. Commun. Syst., vol. 31, no. 5, 2018.

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: DDoS attack detection
Method: Deep NN

Input: 68 features
Output: 8 classes: normal and 7 types of DDoS

- Q. Niyaz, W. Sun, and A. Y. Javaid, “A deep learning based DDoS detection system in software-defined networking (SDN),” arXiv preprint arXiv:1611.07400, 2016.

Objective: Application fault detection
Method: ML approaches

- L. J. Jagadeesan and V. Mendiratta, “Programming the network: Application software faults in software-defined networks,” in Proc. IEEE ISSREW, Ottawa, ON, Canada, Oct. 2016, pp. 125–131.

Objective: Firewall performance optimization
Method: Neural network, HMM

- Z. Din and J. de Oliveira, “Anomaly free on demand stateful software defined firewalling,” in Proc. IEEE ICCCN, Vancouver, BC, Canada, Jul. 2017, pp. 1–9.

* J. Xie *et al*., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

ML-based SDN

Objective: DDoS attack detection
Method: XGBoost

Input: 41 features
Output: 2 classes: normal and DDoS

- Z. Chen, F. Jiang, Y. Cheng, X. Gu, W. Liu and J. Peng, "XGBoost Classifier for DDoS Attack Detection and Analysis in SDN-Based Cloud," *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Shanghai, China, 2018, pp. 251-256, doi: 10.1109/BigComp.2018.00044.

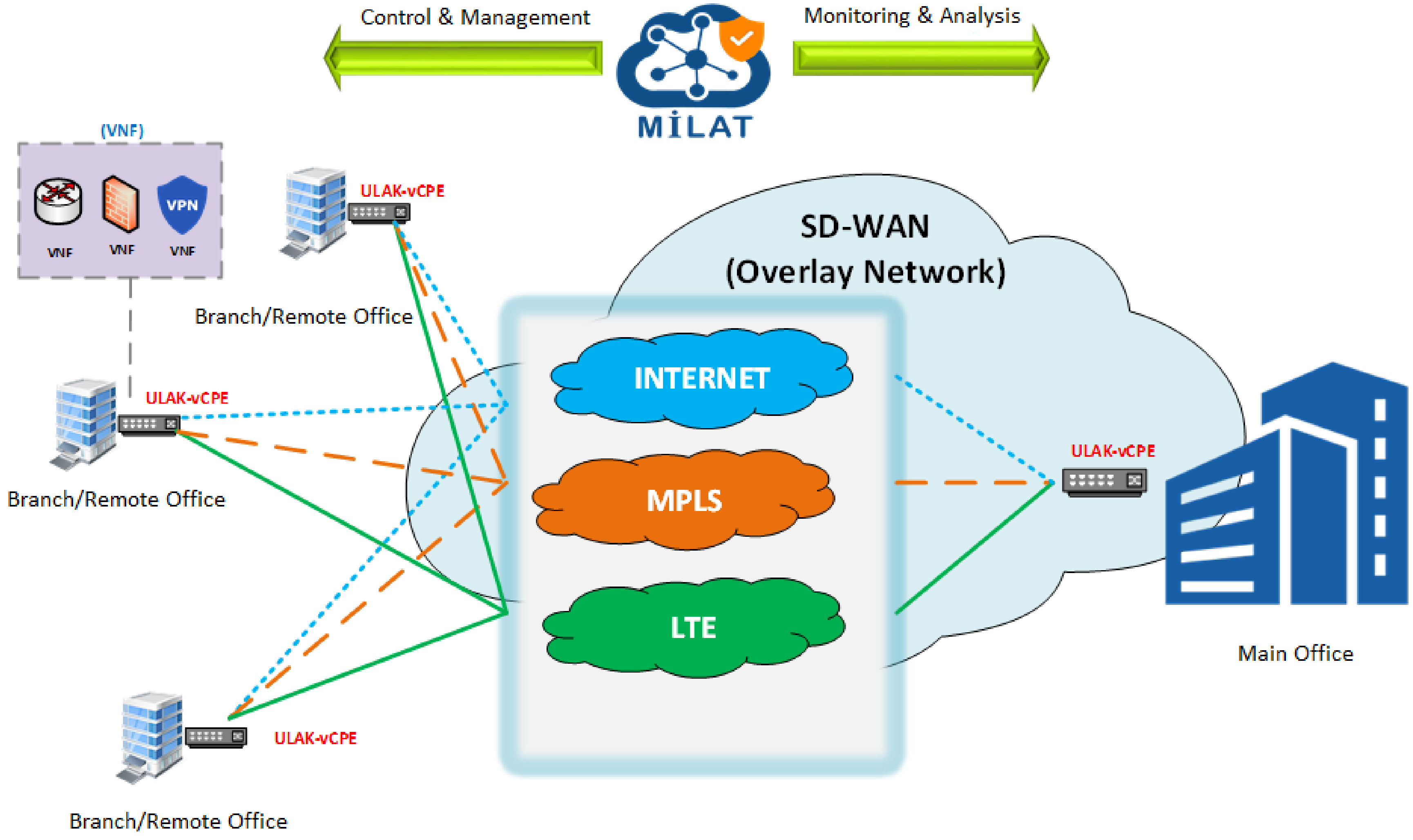
Objective: DDoS attack detection
Method: CNN, RNN

Input: 80 features
Output: 2 classes: normal and DDoS

- S. Haider *et al*., "A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks," in *IEEE Access*, vol. 8, pp. 53972-53983, 2020, doi: 10.1109/ACCESS.2020.2976908.

- Canadian Institute for Cybersecurity → https://www.unb.ca/cic/
- Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization”, 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018.

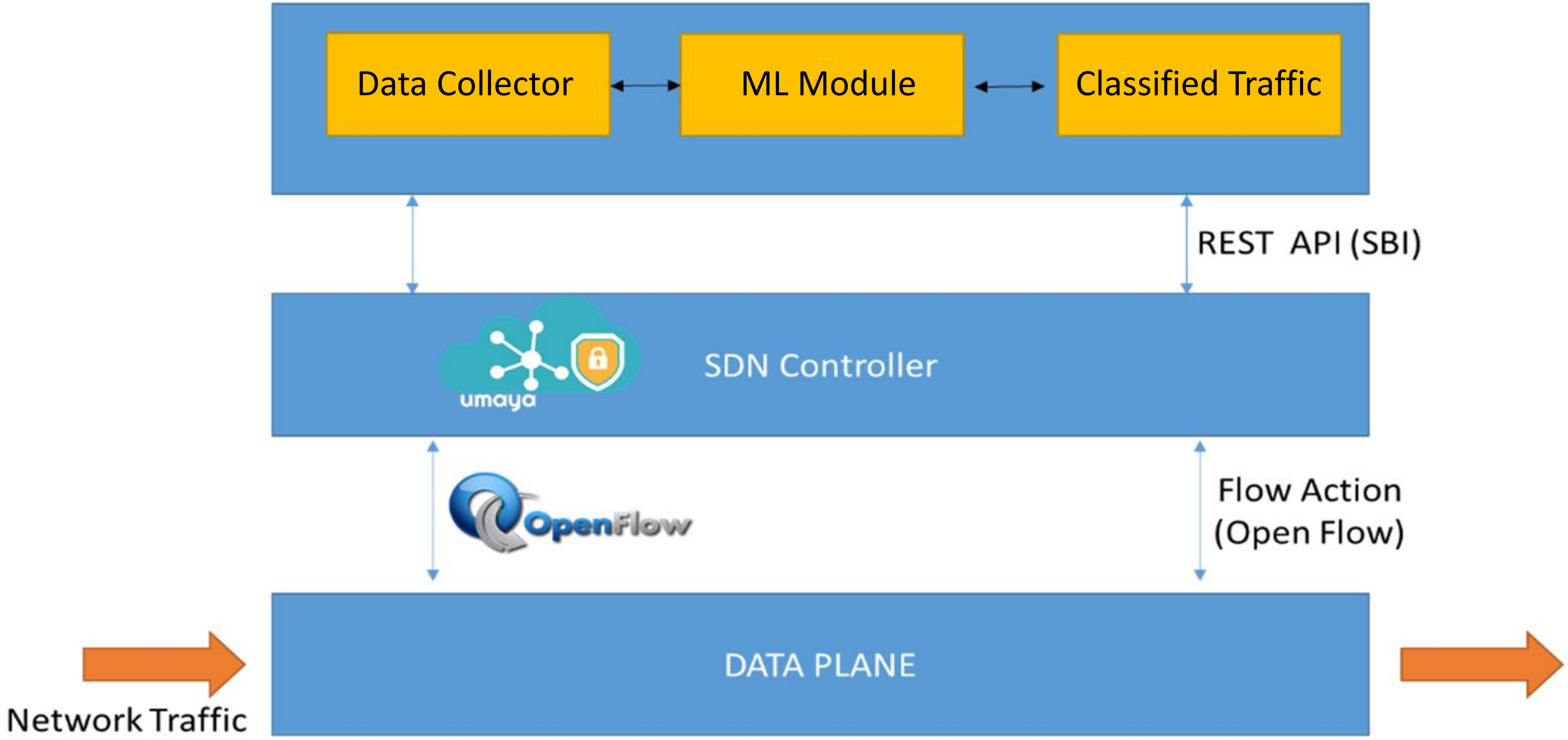
Problem Definition



Problem Definition

- We expect to get a security solution that operates on an SDN infrastructure and is capable of detecting attacks based on metrics obtained from traffic flows.
- We provide necessary measures to enhance the SDN solution.
- This solution is targeted towards addressing a real-world problem that we are focusing on.
- You can envision it as being integrated with the ULAK Maya SDN platform. Our aim is to draw attention to this research area and highlight its importance.

Problem Definition



Dataset

- **Classification problem**
  - Normal, DDoS, Malware, Web-based
- **Dataset**
  - A time-labelled dataset
  - Training and validation sets
  - Test set will not be shared
- **Inputs**
  - 79 features
- **Outputs**
  - Normal, DDoS, Malware, Web-based

Evaluation Criteria

- Classification problem
- The challenge teams will share their technical reports and codes with ULAK challenge team.
- The technical report is expected to be in pdf format. Participants are expected to explain their solution, including the outcomes of their models.

Evaluation Criteria

- There will be no restriction for the ML models. Participants are required to provide an explanation for their choice of ML model.
- Additionally, the performance of the selected ML model should be compared with at least three baseline ML models.
- The selection of baseline models should be well-known and aligned with prior art, and the specific choice of baseline models can be determined by the participants.
- Labels:
    - DoS Hulk, BENIGN, DDoS, PortScan, DoS GoldenEye, FTP-Patator, DoS slowloris, DoS Slowhttptest, SSH-Patator, Web Attack – XSS, Web Attack - Brute Force, Web Attack – Sql, Injection, Bot, Infiltration, Heartbleed

Evaluation Criteria

- For each label, the criteria must be as follow:
    - The criteria will cover a maximum False Positive value, i.e., 10%.
    - Accuracy must be at least 90%.
    - Recall and precision be at least 90%.
    - The performance evaluation considers K-fold cross validation.

Evaluation Criteria

- The ML model with less complexity is preferred when two models achieve similar performance.
- ULAK Comm. will use another test dataset to evaluate the model performances.
- Challenge teams are free to use any tools or APIs.

Timeline

- Registration: 29 May - 25 August 2023
- Submission deadline: 31 August 2023
- Evaluation: 31 October
- Grand Challenge finale (awards): 13 Dec.

- Link: https://challenge.aiforgood.itu.int/match/matchitem/81

Challenge Team

Evren Tuna
6G R&D Engineer

evren.tuna@ulakhaberlesme.com.tr

Abdullah Mekki
Network & Systems Engineer

abdullah.mekki@ulakhaberlesme.com.tr

Thank You!

Evren Tuna

evren.tuna@ulakhaberlesme.com.tr

/evrentuna