

SmartX Intelligent Sec: A Security Framework Based on Machine Learning and eBPF/XDP

Talaya Farasat
University of Passau
Passau, Germany

JongWon Kim
GIST
Gwangju, South Korea

Joachim Posegga
University of Passau
Passau, Germany

ABSTRACT

Information and Communication Technologies (ICT) infrastructures are becoming increasingly complex day by day, facing numerous challenges to support the latest networking paradigms. Security is undeniably a critical component for the effective functioning of these advanced ICT infrastructures. By considering the current network security challenges, we propose SmartX Intelligent Sec, an innovative intelligent security framework. SmartX Intelligent Sec leverages a combination of the lightweight extended Berkeley Packet Filter/eXpress Data Path (eBPF/XDP) for efficient network packet capturing and filtering malicious network traffic, and a Bidirectional Long Short-Term Memory (BiLSTM) classifier for network threat detection. Our real-time prototype demonstrates that SmartX Intelligent Sec offers comprehensive automation features, enabling continuous network packet capturing, effective network threat detection, and efficient filtering of malicious network traffic. This framework ensures enhanced security and operational efficiency for modern ICT infrastructures.

KEYWORDS

SmartX Intelligent Sec, eBPF/XDP, BiLSTM

1 INTRODUCTION

Network security remains a significant challenge, especially with the increasing sophistication of cyber threats like Distributed Denial of Service (DDoS) attacks. These attacks can severely disrupt critical infrastructure, leading to extensive financial losses and operational downtime. Traditional security measures often fall short in countering the volume and complexity of modern threats, necessitating the adoption of more advanced solutions.

Currently, eBPF/XDP (a light-weight, powerful virtual machine inside linux kernel) has attracted significant interest in the research community and has been used for a variety of use cases, e.g., network virtualization, load balancing, and network packet filtering/network threat mitigation.

Meanwhile, machine learning also offers powerful tools for enhancing network security that far exceed traditional methods in terms of speed, accuracy, and adaptability. As cyber threats become more sophisticated, the usage of machine learning into network security will continue to play a crucial role in detecting network threats and identifying various network anomalies.

In this paper, we propose SmartX Intelligent Sec, an innovative intelligent security framework leveraging powerful, lightweight eBPF/XDP tool in conjunction with latest machine learning/deep learning techniques. SmartX Intelligent Sec employs eBPF/XDP for continuous network packet capturing and filtering of malicious network traffic and integrates with BiLSTM for detecting network threats. Overall, this combination (eBPF/XDP with BiLSTM) allows

SmartX Intelligent Sec to provide comprehensive automation features for continuous network packet capturing, network threat detection, and filtering malicious network traffic.

SmartX Intelligent Sec has been evaluated both in terms of network threat detection accuracy and the efficacy of filtering malicious network traffic. SmartX Intelligent Sec, BiLSTM-based network threat detection module demonstrates outstanding performance, with an accuracy of 99.3%, F-score of 99.3%, and ROC-AUC of 99.9%. Furthermore, our real-time SmartX Intelligent Sec prototype demonstrates that eBPF/XDP efficiently filters/drops 2,295,337 malicious network packets (in 15 sec. of DDoS attack), affirming its capability in actively mitigating networks threats.

2 RELATED WORK

Previous studies explore eBPF-based network monitoring systems and ideas about integration of machine learning techniques with eBPF/XDP. The significant performance improvement with eBPF/XDP for fast network packet processing and saving resources as compared to traditional methods has already been highlighted [14, 26, 27, 34]. Ben-Yair et al. [29] propose to use machine learning with eBPF for detecting anomalies. Similarly in [8, 12, 16, 17, 30, 38], simple machine learning models are used in conjunction with eBPF based technology for the purpose of detecting various network intrusions. In [13, 15, 37], eBPF is used to develop solutions against Denial-of-Service attacks, but they do not use machine learning.

After integrating simple machine learning models with eBPF/XDP, researchers also focus on utilizing complex and more effective deep learning models with eBPF and XDP. Zhang et al. [35] deploys eBPF in conjunction with LSTM for forecasting network situations. Sakuraba et al. [33] use eBPF with LSTM specifically for monitoring routers. Kostopoulos [31] employs eBPF/XDP with simple neural networks for detection and prevention from network attacks. Recently, Karma IDS [6] has been introduced, leveraging eBPF for packet capturing and LSTM for network threat detection. Our work is closely related to Karma IDS, which employs eBPF solely for packet capturing. We use eBPF/XDP, allowing our system not only to capture packets but also to filter out malicious packets. By combining eBPF and XDP, SmartX Intelligent Sec can not only capture network packets but also filter and block malicious network traffic directly. This dual approach provides security by mitigating threats in real-time and improves performance by reducing the overhead associated with traditional packet processing methods. Moreover, instead of using LSTM, we evaluate its advanced version, BiLSTM for network threat detection. It is noteworthy our BiLSTM-based network threat detection module achieves an accuracy of 99.3% (better than Karma IDS). As BiLSTM trains in both forward and backward directions, indeed it can be a better option than a simple LSTM or with simple neural networks.

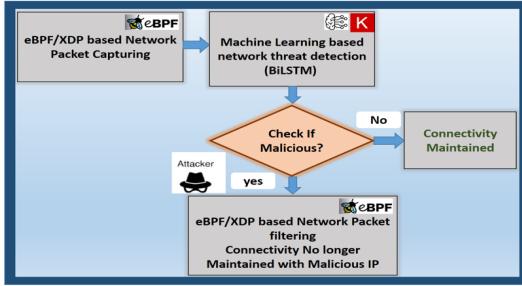


Figure 1: SmartX Intelligent Sec Design

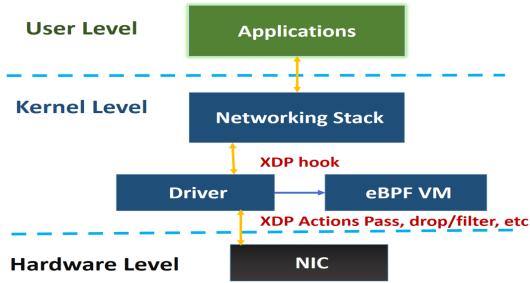


Figure 2: Linux network stack with eBPF/XDP

3 SMARTX INTELLIGENT SEC DESIGN

The design of SmartX Intelligent Sec is illustrated in Figure 1. Initially, the eBPF/XDP tool captures the network packets. These network packets are then analyzed by the BiLSTM-based machine learning tool to detect network threats. If the network traffic is identified as malicious, the eBPF/XDP tool filters/drops the malicious network packets, and the connection with that malicious IP is terminated. Conversely, if the network traffic is deemed normal, the connectivity is maintained.

3.1 Network Packet Capturing

Linux eBPF is a light-weight, powerful virtual machine that provides a set of libraries which allows dynamic injection of codes from user space into various kernel events, see Figure 2. Meanwhile, XDP provides special hooks for eBPF kernel programs to efficiently pass, drop/filter, and redirect network packets received at networking ports [25]. eBPF/XDP programs can be attached to the network driver with three different points, i.e., skb (generic), native, and hw (program offloaded to the hardware). The fastest one, i.e., hw or offload mode allows the XDP program runs on the NIC itself. However, for this mode, SmartNIC is required. In native mode, the XDP hook is called in the driver (driver support needed) before the kernel allocates the socket buffer. In skb or generic mode, XDP hook is called after the packet DMA and socket buffer allocation. So, in this case, the processing performance is significantly lower than in other modes. We configure the XDP tools [5] and use specifically xdp-dump (use native mode) for network packet capturing.

3.2 Network Threat Detection

For network threat detection, while we primarily focus on advanced deep learning techniques such as LSTM and BiLSTM, we first validate various simple machine learning algorithms to verify their performance accuracies.. So, we train various machine learning algorithms and evaluate their performance. The process begins with the selection of a suitable dataset, which is crucial for training machine learning models. In this section, first, we highlight our dataset and its key features. Next, we provide an overview of the different machine learning algorithms used in our study, explaining our selected corresponding hyper-parameters. Finally, we evaluate the performance of these algorithms, comparing metrics such as Accuracy, F-Score, and ROC to determine the best-performing model for network threat detection.

3.2.1 Dataset. We utilize our dataset available here [7, 20]. We configure CICFlowMeter-V4.0 [2] to convert all .pcap files, encompassing both normal and malicious network traffic, into network flows. CICFlowMeter-V4.0 generates over 80 network traffic features. We use SelectKBest algorithm from scikit-learn [4], selecting the top 20 features, see Table 1. Our dataset comprises a total of 2.5 million network flows, balanced with 125,000 normal flows and 125,000 malicious flows. We frame this problem as a binary classification problem and use two labels, i.e., 0 for normal and 1 for malicious. We use Python scikit-learn train_test_split to create train and test datasets and use 70% data for training and 30% data for testing.

3.2.2 Machine learning Algorithms. We develop various machine learning models using the Python scikit-learn library, including Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Logistic Regression (LR), Decision Tree, Extra-Trees, and Multilayer Perceptron (MLP). We implement Long Short-Term Memory (LSTM) and Bidirectional LSTM (biLSTM) by using the Python Keras framework. These models are chosen due to their proven effectiveness in addressing network threat detection challenges, as evidenced by prior research [9, 11, 19, 21–23, 28, 32, 39–42].

SVM works by finding the optimal hyperplane that maximizes the margin between two classes of data. This ensures the greatest separation between the classes. SVM can be utilized with various kernel functions, but we choose a linear kernel for its simplicity and speed in binary classification tasks. All other parameters are kept at their default values.

GNB is a probabilistic algorithm based on Bayes' Theorem, which updates prior probabilities into posterior probabilities by incorporating the information, called likelihoods provided by the observed data. In this paper, we consider one of the extensions of naive bayes which is called Gaussian Naive Bayes which follows the Gaussian distribution. We implement a binary GNB classifier with all default parameters.

LR assigns data points to discrete classes by transforming its output with the logistic sigmoid function returning a probability value that can be mapped to discrete classes. We implement our binomial logistic regression, setting the solver to 'liblinear' while keeping all other parameters at their default values.

Decision Tree is a flowchart-like structure with a root node, internal nodes, branches, and leaf nodes. The root node represents

Features	Description
Flow Duration	Length of connection in seconds
Fwd Pkt Len Max	Maximum size of packets in the forward direction
Fwd Pkt Len Mean	Average size of packets in the forward direction
Fwd Pkt Len Std	Standard Deviation of the size of the packets in the forward direction
Flow IAT Mean	Mean inter-arrival time between packet flows
Flow IAT Max	Maximum inter-arrival time between packet flows
Fwd IAT Mean	Forward inter-arrival time, the time between two packets sent forward direction Mean
Fwd IAT Tot	Forward inter-arrival time, the time between two packets sent forward direction Total
Fwd IAT Std	Forward inter-arrival time, the time between two packets sent forward direction Standard deviation
Fwd Header Len	Total bytes used for a header in the forward direction
Pkt Len Min	Minimum length of the packets flow
Pkt Len Var	Minimum inter-arrival time of packet
SYN Flag Cnt	Number of packets with SYN Flag
Fwd Seg Size Avg	Average segment size in forward direction
Subflow Fwd Pkts	Average number of packets in sub-flow in a forward direction
Active Max	The amount of time a flow was active before becoming idle Max
Idle Std	The amount of time a flow was idle before becoming active Standard deviation
Idle Mean	The amount of time a flow was idle before becoming active Mean
Idle Min	The amount of time a flow was idle before becoming active Min
Idle Max	The amount of time a flow was idle before becoming active Max

Table 1: Features used in training of Machine Learning Algorithms

the entire dataset, while each internal node represents an input feature to perform a test. Branches emanating from a node represent the outcome of the test, that is, all possible feature values. Each leaf node represents a class label or a class probability distribution. A decision tree is learned by splitting the dataset into smaller subsets and tree is built by recursively splitting each derived subset until further splitting does not improve predictions. We create our Decision Tree model as a binary classifier, setting the `max_depth` parameter to 3 while keeping all other parameters at their default values.

Extra Trees is an ensemble machine learning algorithm that combines the results of multiple de-correlated decision trees collected in a “forest” to output its classification results. We build our Extra Trees model as a binary classifier, setting `max_depth` to 3 while keeping all other parameters at their default values.

MLP is a class of feedforward artificial neural networks that maps input data sets to a set of appropriate outputs. It is the most basic type of neural network. As scikit-learn is also capable of basic deep learning modeling, we use this library to generate our MLP model. We use all default parameters except `max_iter` which we set as 300.

LSTM employs a special form of RNNs. Traditional RNNs are designed to store inputs in order to predict the outputs. But they are not performed well when several discrete time lags occur between the previous inputs and the present targets. To overcome this limitation, LSTM is capable of connecting time intervals to

form a continuous memory. We use the Python Keras framework (backend Tensorflow) [3] to generate the LSTM model as a binary classifier. Our LSTM model contains an input layer, three hidden layers with 50 LSTM cells, a dropout layer with 20% dropout rate, and an output layer with a single node. We use Adam optimizer and he model is trained for 50 epochs with the default learning rate = 0.001.

BiLSTM in contrast with LSTM, processes data in both forward and backward directions using two separate hidden layers. These hidden layers are connected to the same output layer, enabling the model to traverse the input data twice. Bidirectional networks have been shown to outperform unidirectional LSTMs in network threat detection and various other fields [42]. We implement our BiLSTM model using the Python Keras framework with TensorFlow as the backend. The model features an input layer, a hidden layer with 64 LSTM units (BiLSTM creates from these LSTMs will be concatenated), and an output layer for binary classification. All other parameters are consistent with those used in the LSTM model.

Performance Evaluation: We evaluate the performance of our machine learning models based on Accuracy, F-score, and Receiver Operating Characteristic (ROC) metrics, see Table 2. True positive (TP) and true negative (TN) are the samples that are correctly classified. False negative (FN) and false positive (FP) are misclassified samples. Accuracy is the ratio of the number of correct predictions over the entire dataset. The F-score takes both false positives and false negatives into account. It is the weighted average harmonic

Models	Accuracy	F-Score	ROC
SVM	83.06	85.3	94.4
GNB	81.9	84.3	83.0
LR	75.6	72.7	88.3
Decision Tree	82.3	84.8	92.8
Extra Trees	86.4	87.2	95.5
MLP	93.3	93.6	96.6
LSTM (with GPU)	95.8	95.9	98.9
BiLSTM (with GPU)	99.3	99.3	99.9

Table 2: Performance Evaluation of Machine Learning Algorithms

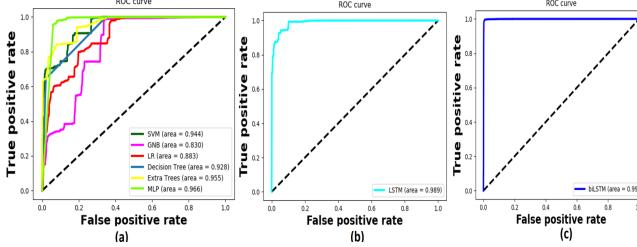


Figure 3: ROC result of different classifiers

value of precision and recall. Precision is the rate of true positives within all positives. Recall, also called sensitivity, is a measurement for the rate of positives that were correctly identified in comparison to the total number of actual positives. The ROC curve is the graphical representation of the true positive rate (TPR) against the false positive rate (FPR) rate at different classification thresholds.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$F - Score = \frac{Precision * Recall}{Precision + Recall} \quad (2)$$

where,

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

Our results indicate that deep learning techniques, such as LSTM and BiLSTM, outperform traditional machine learning algorithms. BiLSTM achieves the highest performance among all other machine learning algorithms, with an accuracy of 99.3%, an F-score of 99.3%, and an ROC of 99.9%. Consequently, we choose BiLSTM for our real-time SmartX Intelligent Sec prototype.

Additionally, we measure the training times of two deep learning models (LSTM and BiLSTM) on both CPU and GPU, highlighted in Table 3. Our results indicate that training on an NVIDIA Tesla T4 GPU is significantly more efficient compared to an Intel® Xeon® D-2183IT CPU. We plan to implement real-time or incremental training in the future. If real-time training will be considered in future, we note that BiLSTM will require more time than LSTM but offers superior performance. This makes it crucial to decide whether to prioritize time efficiency or performance.

	LSTM	BiLSTM
CPU (Intel® Xeon® D-2183IT)	5 min. 45 sec.	13 min. 33 sec.
GPU (Tesla T4)	1 min. 9 sec.	2 min. 43 sec.

Table 3: Time comparison between CPU and GPU

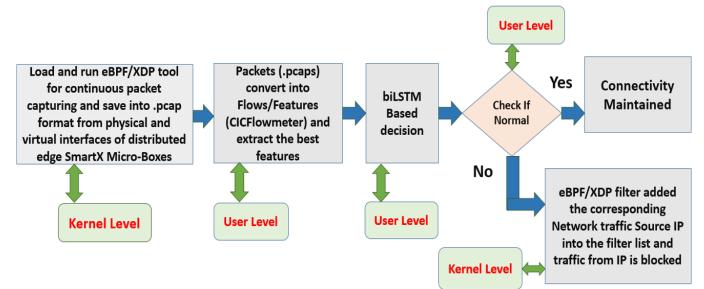


Figure 4: SmartX Intelligent Sec Real Time Implementation Working

3.3 Network Packet Filtering

We utilize the configured XDP tool [5], specifically xdp-filter, for network packet filtering. The eBPF/XDP tool, known for its efficiency, achieves high filter/drop rates, handling tens of millions of packets per second on a single CPU core [5]. When the BiLSTM-based network threat detection module identifies incoming traffic as malicious, xdp-filter immediately captures the source IP of the malicious traffic. Consequently, all traffic from that malicious source IP is dropped or filtered, and connectivity with that source IP is terminated.

3.4 Benefits of using eBPF/XDP

eBPF/XDP filters out malicious traffic, retrieving only normal packets and thereby significantly mitigating network threats. Modern infrastructures often include numerous virtualized and containerized cloud nodes. Under heavy networking traffic, capturing and filtering packets from these networking ports can consume substantial compute and storage resources, potentially disrupting legitimate services. Therefore, lightweight software functions are essential for managing various networking ports flexibly while minimizing resource consumption. In such scenarios, eBPF/XDP-based packet capturing and filtering of malicious traffic is an optimal solution due to its efficiency and low resource usage.

4 SMARTX INTELLIGENT SEC PROTOTYPE IMPLEMENTATION

Aligned with future Internet testbeds like the Global Environment for Networking Innovation (GENI), the Gwangju Institute of Science and Technology (GIST) has launched the OF@TEIN testbed/playground [1, 10, 18, 24, 25, 36]. The OF@TEIN playground provides a dynamic array of distributed physical and virtual resources for users and developers to learn operational and development issues and conduct various experiments. It is an overlay, multi-site playground spanning heterogeneous underlay networks, connecting around 14 sites across 10 countries. The playground

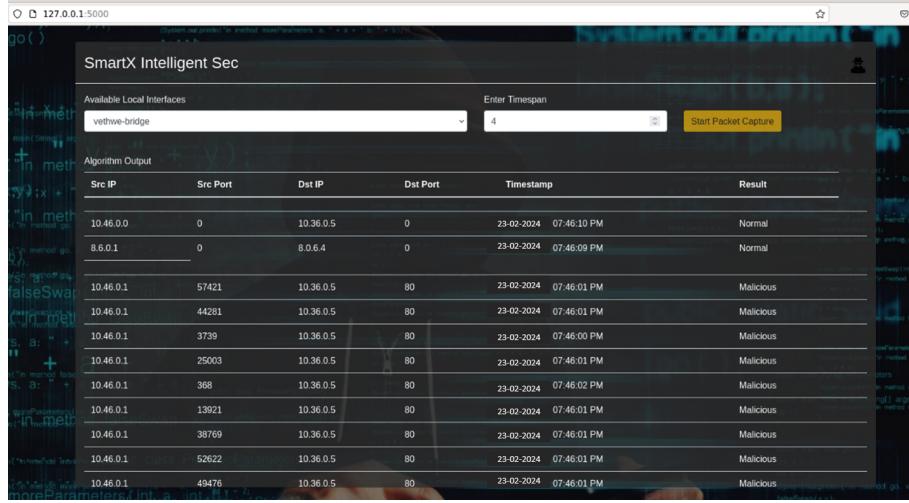


Figure 5: SmartX Intelligent Sec Real Time Implementation

```

Loaded on interfaces:
xdpfilt_alw_all
    vethwe-d datapath (skb mode)
xdpfilt_alw_all
    vethwe-bridge (skb mode)
xdpfilt_alw_all
    vxlan-6784 (skb mode)

Enabled features
tcp, udp, ipv6, ipv4, ethernet, allow
tcp, udp, ipv6, ipv4, ethernet, allow
tcp, udp, ipv6, ipv4, ethernet, allow

Filtered ports:
Mode          Hit counter

Filtered IP addresses:
Mode          Hit counter
10.46.0.1      2295337

Filtered MAC addresses:
Mode          Hit counter

```

Figure 6: Status of XDP-filter (Attacking Container Pod-IP (10.46.0.1) has been added to filtered IP addresses list)

features "SmartX Micro-Boxes," commodity server-based hyper-converged resources (compute/storage/networking) distributed at multi-site edge locations. These Micro-Boxes support experiments around Cloud computing, Software-defined Networking (SDN), and Network Function Virtualization (NFV). Equipped with cloud-native (containerized) functionalities, the Micro-Boxes function as Kubernetes-orchestrated workers with SDN-coordinated connectivity to other Micro-Boxes. The infrastructure is managed by the "SmartX Playground Tower," which automates the building, operation, and utilization of the playground. The Playground Tower oversees the multi-site playground operation, leveraging several entities, including a Visibility Center, to ensure efficient management.

We implement a real-time prototype of SmartX Intelligent Sec over the OF@TEIN playground, as illustrated in Figure 4 and 5. First, eBPF/XDP captures real-time network traffic from the selected interface (physical or virtual) at the kernel level. Users can specify the packet capturing duration (up to 10 seconds), making the time

complexity flexible. For instance, if the user selects 4 seconds, packets are captured for that period, then processed by CICFlowmeter to extract flow features and select the top 20 features. These features are then analyzed by BiLSTM for network threat detection. If the network traffic is normal, packets continue to move from kernel space to user space. Conversely, if the traffic is identified as malicious, XDP-filter immediately adds the source IP to the filter list, dropping packets from that source at the kernel level with an extremely high drop rate, capable of handling tens of millions of packets per second on a single CPU core.

Real Time Testing We also simulate a DDoS attack, as illustrated in Figure 5. SmartX Intelligent Sec is deployed on the Kubernetes veth interface of the GIST-1 SmartX Edge Micro-Box, part of the OF@TEIN playground. We set the packet capture duration to 4 seconds. During this time, eBPF/XDP captures packets and forwards them to the subsequent modules for further processing, as depicted in Figure 4.

To simulate the DDoS attack, we use a testing machine to launch an attack for 15 seconds on the GIST-1 SmartX Edge Micro-Box

using a Docker image of hping3 (TCP SYN flood), with the attacker Pod having the IP 10.46.0.1. When the attacker Pod is activated, eBPF/XDP begins capturing packets for the first 4 seconds. These packets are then displayed as malicious by SmartX Intelligent Sec (detected by BiLSTM).

Subsequently, the eBPF/XDP filter immediately adds the attacker Pod IP 10.46.0.1 to its filter list. As a result, all malicious DDoS traffic from the attacker Pod IP 10.46.0.1 is completely filtered/dropped. The eBPF/XDP filter status in Figure 12 shows that 2,295,337 packets have been filtered/blocked from the attacker Pod IP 10.46.0.1.

5 CONCLUSION, LIMITATIONS, AND FUTURE WORK

We utilize the highly effective BiLSTM deep learning algorithm for network threat detection in conjunction with the fast, efficient, and lightweight eBPF/XDP for network packet capturing and filtering of malicious traffic. This powerful combination addresses contemporary network security challenges robustly. Our real-time prototype demonstrates that SmartX Intelligent Sec provides comprehensive automation features for continuous network packet capturing, network threat detection, and filtering malicious network traffic. However, a limitation of our current work is that the BiLSTM-based network threat detection module operates at the user level. As future work, we plan to shift this module to the kernel space to enhance efficiency further. Additionally, we aim to explore real-time and incremental training methods inside the Linux kernel space.

6 ACKNOWLEDGEMENT

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2019-0-01842, Artificial Intelligence Graduate School Program (GIST)).

REFERENCES

- [1] 2015. OF@ TEIN: A community effort towards open/shared SDN-Cloud virtual playground. *Proceedings of the Asia-Pacific Advanced Network (APAN)* 40 (2015), 22–28. <https://doi.org/10.7125/40.4>
- [2] 2022. CICflowMeter: <https://github.com/ahashkari/CICflowMeter>.
- [3] 2022. <https://keras.io/>.
- [4] 2022. scikit-learn: <https://scikit-learn.org/stable/>.
- [5] 2024. eBPF/XDP: <https://github.com/xdp-project/xdp-tools>.
- [6] 2024. An Intrusion Detection System using eBPF and LSTM: <https://dev.to/pree2111/karma-ids-497p>.
- [7] 2024. SmartX Dataset For Network Threat Detection: https://drive.google.com/drive/folders/11DDvarZg9YOvzPfK2e78CLP7S_a9bCXm.
- [8] Nemalikanti Anand, Saifulla M A, and Pavan Kumar Aakula. 2023. High-performance Intrusion Detection System using eBPF with Machine Learning algorithms. Researcg Gate.
- [9] Emin Anarim, Ramin F Fouladi, and Cemil Eren Kayatas. 2016. Frequency based DDoS attack detection approach using naive Bayes classification. IEEE 39th International Conference on Telecommunications and Signal Processing, 104–107. <https://doi.org/10.1109/TSP.2016.7760838>
- [10] Muhammad Usman Aris Cahyadi Risdianto and JongWon Kim. 2019. SmartX Box: Virtualized Hyper-Converged Resources for Building an Affordable Playground. *Electronics* 8(11) (2019), 1242. <https://doi.org/10.3390/electronics811242>
- [11] Hakan Aydin, Zeynep Orman, and Muhammed Ali Aydin. 2022. A long short-term memory (LSTM)-based distributed denial of service (DDoS) detection and defense system design in public cloud network environment. *Elsevier Computers & Security* 118 (2022), 102725. <https://doi.org/10.1016/j.cose.2022.102725>
- [12] Maximilian Bachl, Joachim Fabini, and Tanja Zseby. 2021. A flow-based IDS using Machine Learning in eBPF. arXiv preprint arXiv:2102.09980 (2021).
- [13] Young Eun Choe, Jun-Sik Shin, Seunghyung Lee, and JongWon Kim. 2020. eBPF/XDP based network traffic visualization and dos mitigation for intelligent service protection. Springer International Conference on Emerging Internetworking, Data and Web Technologies, 458–468. https://doi.org/10.1007/978-3-030-39746-3_47
- [14] Gustavo de Carvalho Bertoli, Lourenço A Pereira, Cesar Marcondes, and Osamu Saotome. 2020. Evaluation of netfilter and eBPF/XDP to filter TCP flag-based probing attacks. Proceedings of XXII symposium on operational applications in defense area (SIGE), 35–39.
- [15] Henri Maxime Demoulin, Isaac Pedisich, Nikos Vasilakis, Vincent Liu, Boon Thau Loo, and Linh Thi Xuan Phan. 2019. Detecting Asymmetric Application-layer Denial-of-Service Attacks In-Flight with FineLane. Proceedings of the 2019 USENIX Conference on USENIX Annual Technical Conference, 693–708.
- [16] Marinos Dimolianis, Adam Pavlidis, and Vasilis Maglaris. 2021. Signature-Based Traffic Classification and Mitigation for DDoS Attacks Using Programmable Network Data Planes. *IEEE Access* 9 (2021), 113061–113076. <https://doi.org/10.1109/ACCESS.2021.3104115>
- [17] Marinos Dimolianis, Adam Pavlidis, and Vasilis Maglaris. 2021. SYN Flood Attack Detection and Mitigation using Machine Learning Traffic Classification and Programmable Data Plane Filtering. 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN), 126–133. <https://doi.org/10.1109/ICIN51074.2021.9385540>
- [18] JongWon Kim et al. 2013. OF@TEIN: An OpenFlow-enabled SDN Testbed over International SmartX Rack Sites, Vol. 36. Proceedings of the Asia-Pacific Advanced Network (APAN), 17–22. <https://doi.org/10.7125/APAN.36.3>
- [19] Talaya Farasat and Akmal Khan. 2021. Detecting and analyzing border gateway protocol blackholing activity. *Wiley International Journal of Network Management* 31 (2021), e2143. <https://doi.org/10.1002/nem.2143>
- [20] Talaya Farasat, JongWon Kim, and Joachim Posegga. 2024. Advancing Network Security: A Comprehensive Testbed and Dataset for Machine Learning-Based Intrusion Detection. <http://arxiv.org/abs/2410.18332>. <https://doi.org/10.48550/arXiv.2410.18332>
- [21] Talaya Farasat and Joachim Posegga. 2024. Machine Learning Techniques for Python Source Code Vulnerability Detection. <https://arxiv.org/abs/2404.09537>. <https://doi.org/10.1145/3626232.3658637>
- [22] Talaya Farasat, Muhammad Ahmad Rathore, Akmal Khan, JongWon Kim, and Joachim Posegga. 2023. Machine Learning-based BGP Traffic Prediction. 2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Exeter, UK, 1925–1934. <https://doi.org/10.1109/TrustCom60117.2023.00262>
- [23] Talaya Farasat, Muhammad Ahmad Rathore, Akmal Khan, Sun Park, and Jong Won Kim. 2021. BGP traffic volume forecasting using LSTM framework. Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies.
- [24] Talaya Farasat, Muhammad Ahmad Rathore, and JongWon Kim. 2023. Integrating Machine Learning for Network Threat Detection with SmartX Multi-Sec Framework. IEEE International Conference on Information Networking (ICOIN), Bangkok, Thailand. <https://doi.org/10.1109/ICOIN56518.2023.10049012>
- [25] Talaya Farasat, Muhammad Ahmad Rathore, and JongWon Kim. 2023. Securing Kubernetes Pods communicating over Weave Net through eBPF/XDP from DDoS attacks. ACM Proceedings of the Thirteenth Conference on Data and Application Security and Privacy, USA.
- [26] Jorge Gallego-Madrid, Irene Bru-Santa, Alvaro Ruiz-Rodenas, Ramon Sanchez-Iborra, and Antonio Skarmeta. 2024. Machine learning-powered traffic processing in commodity hardware with eBPF. *Computer Networks* 243 (2024).
- [27] Takanori Hara and Masahiro Sasabe. 2023. On Practicality of Kernel Packet Processing Empowered by Lightweight Neural Network and Decision Tree. In IEEE International Conference on the Network of the Future (NOF).
- [28] Mohamed Idhammad, Karim Afdel, and Mustapha Belouch. 2018. Semi-supervised machine learning approach for DDoS detection. *Springer Applied Intelligence* 48 (2018), 3193–3208.
- [29] Pavel Rogovoy Ido Ben-Yair and Nezer Zaidenberg. 2019. AI and eBPF based performance anomaly detection system. ACM Proceedings of the International Conference on Systems and Storage. <https://doi.org/10.1145/3319647.3325842>
- [30] Nikos Kostopoulos, Stavros Korintis, Dimitris Kalogerias, and Vasilis Maglaris. 2021. Mitigation of DNS Water Torture Attacks within the Data Plane via XDP-Based Naive Bayes Classifiers. IEEE 10th International Conference on Cloud Networking (CloudNet), 133–139. <https://doi.org/10.1109/ICIN51074.2021.9385540>
- [31] Stamatis Kostopoulos. 2024. Machine learning-based near real time intrusion detection and prevention system using eBPF.
- [32] Kshira Sagar Sahoo, Bata Krishna Tripathy, Kshirasagar Naik, Somula Ramasubbaredy, Balamurugan Balusamy, Manju Khari, and Daniel Burgos. 2020. An Evolutionary SVM Model for DDOS Attack Detection in Software Defined Networks. *IEEE Access* 8 (2020), 132502 – 132513. <https://doi.org/10.1109/ACCESS.2020.3009733>
- [33] Minato Sakuraba, Junichi Kawasaki, Takuya Miyasaka, and Atsushi Tagami. 2023. An Anomaly Detection Approach by AIML in IP Networks with eBPF-Based Observability. IEEE 24st Asia-Pacific Network Operations and Management

- Symposium (APNOMS).
- [34] Dominik Scholz, Daniel Raumer, Paul Emmerich, Alexander Kurtz, Krzysztof Lesiak, and Georg Carle. 2018. Performance Implications of Packet Filtering with Linux eBPF. In IEEE 30th International Teletraffic Congress (ITC 30), 209–217.
 - [35] Wenwen Sun, Yi Li, and Shaopeng Guan. 2022. Linux Network Situation Prediction Model Based on eBPF and LSTM. IEEE 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 551–556. <https://doi.org/10.1109/ISKE54062.2021.9755426>
 - [36] Muhammad Usman and JongWon Kim. 2019. SmartX Multi-View Visibility Framework for unified monitoring of SDN-enabled multisite clouds. *Wiley Transactions on Emerging Telecommunications Technologies* (2019), e3819. <https://doi.org/10.1002/ett.3819>
 - [37] Huub van Wieren. 2019. Signature-Based DDoS Attack Mitigation: Automated Generating Rules for Extended Berkeley Packet Filter and Express Data Path. Master's thesis, University of Twente.
 - [38] Zicheng Wang, Tiejin Chen, Qinrun Dai, Yueqi Chen, Hua Wei, and Qingkai Zeng. 2024. When eBPF Meets Machine Learning: On-the-fly OS Kernel Compartmentalization. <https://arxiv.org/abs/2401.05641>.
 - [39] Yuanyuan Wei, Julian Jang-Jaccard, Fariza Sabrina, Amardeep Singh, Wen Xu, and Seyit Camtepe. 2021. AE-MLP: A Hybrid Deep Learning Approach for DDoS Detection and Classification. *IEEE Access* 9 (2021), 146810 – 146821. <https://doi.org/10.1109/ACCESS.2021.3123791>
 - [40] Yi-Chi Wu, Huei-Ru Tseng, Wuu Yang, and Rong-Hong Jan. 2011. DDoS detection and traceback with decision tree and grey relational analysis. *International Journal of Ad Hoc and Ubiquitous Computing* 7 (2011), 121–136. <https://doi.org/10.1504/IJAHUC.2011.038998>
 - [41] Satyajit Yadav and S. Selvakumar. 2015. Detection of application layer DDoS attack by modeling user behavior using logistic regression. IEEE International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 1–6. <https://doi.org/10.1109/ICRITO.2015.7359289>
 - [42] Xiangzhe Zhang, Zhaoyuan Liu, and Jiaqing Bai. 2019. An improved method of DDoS attack detection for controller of SDN. IEEE International Conference on Computer and Communication Engineering Technology (CCET), 249–253. <https://doi.org/10.1109/CCET48361.2019.8989356>