DDoS Attack Detection via Privacy-aware Federated Learning and Collaborative Mitigation in Multi-domain Cyber Infrastructures

Marinos Dimolianis
National Technical University of Athens (NTUA)
Zografou, Greece
mdimolianis@netmode.ntua.gr

Dimitrios K. Kalogeras
Institute of Computer & Communication Systems (ICCS)
Zografou, Greece
dkalo@noc.ntua.gr

Nikos Kostopoulos
National Technical University of Athens (NTUA)
Zografou, Greece
nkostopoulos@netmode.ntua.gr

Vasilis Maglaris
National Technical University of Athens (NTUA)
Zografou, Greece
maglaris@netmode.ntua.gr

***Abstract*—Interconnected cyber infrastructures, accessible via the Internet, are a common target of DDoS attacks intending to downgrade their operations and services. Collaborative protection mechanisms are prime candidates to defend against massive attacks but, although collaborations were instrumental in the Internet success story, this is largely not extended to multi-domain cyber security. Notably, collaborative DDoS detection is hindered by data privacy legislations, while mitigation is limited to operations of stand-alone rigid firewalls. Motivated by these shortcomings, we propose a Federated Learning schema for collaborative privacy-aware DDoS detection. Coordination is orchestrated by a third trusted party that aggregates machine learning models proposed by collaborators based on their private attack and benign traces, without exchanging sensitive data. Attacks detected via the privacy-aware federated model are subsequently mitigated by efficient and scalable firewalls, implemented within the eXpress Data Path (XDP) data plane programmability framework. Our approach was evaluated using production traffic traces in terms of packet classification accuracy and packet processing performance. We conclude that our proposed Federated Learning framework enabled collaborators to accurately classify benign and attack packets, thereby improving individual domain accuracy. Furthermore, our data plane programmable firewalls promptly mitigated large-scale attacks in emulated federated cyber infrastructures.**

***Index Terms*—Federated Learning, Federated Clouds, Multi-domain Networks, Multi-domain DDoS Protection, Programmable Data Planes, eXpress Data Path (XDP)**

I. Introduction

Distributed Denial of Service (DDoS) attacks are a major threat targeting critical resources of cyber infrastructures to degrade the level of their offered services. These attacks are delivered to victim domains via interconnected networks, e.g. Autonomous Systems (AS's) of the global Internet and overwhelm the link capacity and processing resources of targeted networks.

Although collaborative protection mechanisms among AS's were instrumental in the Internet success story to defend against large-scale attacks, their extension to multi-domain cyber environments for coordinated DDoS detection is not straightforward. This is mainly hindered by strict data privacy legislations, i.e. GDPR [1]. Federated Learning (FL) [2] is a promising approach to address such privacy regulations by allowing collaborating parties to cooperatively train Machine Learning (ML) models without exposing private data. FL has been proposed for various use cases like word prediction [2], healthcare applications [3] and image recognition [4]. Few efforts [5], [6] consider collaborative DDoS detection without addressing the needs of multi-domain cyber infrastructures. These environments require flexible and efficient methods to identify large-scale attacks, while exchanging attack information should rely on tools and data packet exchange protocols, e.g. BGP, that are widely adopted among collaborating domains.

In contrast to collaborative DDoS detection, collaborative mitigation is widely employed in production environments. DDoS attacks are mitigated by filters enforced by collaborating domains. These filters are typically implemented in routing devices and discard either all traffic (BGP blackholing [7]) or the malicious portion via a limited number of flow-based rules (e.g. BGP Flowspec [8]). Our early work [9], [10] illustrated that DDoS mitigation relying on flow-based filtering schemes is less effective than signature-based ones, where filtering rules are defined by combining data packet fields of multiple protocol layers. Thus, we leverage programmable data planes (eXpress Data Path – XDP [11]) to design a signature-based filtering mechanism tailored to evolving federated cyber infrastructures.

In this paper, we extend our early work on signature-

This work was partially supported by the European Commission H2020 Framework Programme (Grant Agreement No. 856726, GN4-3), the Special Account for Research Funding of NTUA and the FELICE H2020 Project (Grant ID 101017151)

based DDoS protection [9], [10] to collaborative multi-domain cyber infrastructures. Our schema detects malicious packet signatures using Multi-layer Perceptrons (MLPs); these are cooperatively trained by a centralized coordinating entity via FL techniques that do not expose private benign and attack data of collaborators. Subsequently, malicious packets are filtered in XDP-enabled [11] programmable firewalls deployed within the victim network infrastructure. For large-scale attacks, mitigation can be activated on-demand in collaborating transit domains (e.g. transit AS's).

The remainder of this paper is structured as follows: In Section II we discuss related efforts on collaborative DDoS protection and outline our key contributions; Section III presents a high-level overview of our mechanism and its core design principles; Section IV provides implementation details for the proposed DDoS protection framework; Section V presents experimental evaluations for DDoS detection accuracy and mitigation performance on DNS Amplification attacks. Section VI summarizes our work and suggests future directions.

II. Related Work & Contributions

DDoS detection and mitigation for collaborative network domains, e.g. AS's, have been widely investigated. The former refers to mechanisms that allow network domains to share data for enhancing their attack detection capabilities. The latter refers to filters raised on-demand by collaborators to drop the attack traffic. Related efforts are analyzed in subsections II-A and II-B accordingly; in subsection II-C Federated Learning schemes for DDoS protection are presented. Finally, in subsection II-D we summarize our key contributions.

A. Collaborative DDoS Detection

In [12], network traffic is monitored in disperse points of multiple network domains in an attempt to concurrently detect attacks targeting subnetworks. Attacks are identified by concurrent alerts generated by collaborating network domains. In [13], Internet Service Providers (ISPs) collaborate to detect ongoing DDoS attacks; based on predefined static rules, they exchange belief scores for suspected DDoS attacks. In [14], security events are exchanged between collaborating ISPs to validate ongoing attacks and provide appropriate countermeasures. The main focus of this work is on the communication process between collaborators. In [15], an effort for creating a European Federation of ISPs, Internet Exchanges (IX) and Academic Networks is made; the members are exchanging attack traffic characteristics via a centralized platform without exposing victim IP addresses for privacy concerns.

B. Collaborative DDoS Mitigation

BGP blackholing [7] is the most common way for collaborative DDoS filtering. Victim domains request from upstream/peer networks to drop all traffic destined to them to protect their internal infrastructures. Although this safeguards transit links, benign traffic is also dropped.

A collaborative approach for filtering reflection-based DDoS attacks is proposed in [16]. The mitigation process is coordinated via a third trusted party that collects sensitive data from collaborating entities and applies threshold-based filtering rules to multi-domain network environments. Contrary to our proposed framework, [16] cannot be extended to non-reflection based attacks without exposure of sensitive data, while instead of using fixed thresholds, our approach relies on machine learning algorithms to generalize from training data.

In [17], a collaborative schema for DDoS mitigation in SDN-domains is proposed. Upon the detection of the attack, specialized reports including the identified malicious sources and the victim IPs are generated; these are transferred to network domains located across the attack path, that enforce filtering rules based on the reputation of the victim domain. We extended [17] in [18], in which signaling, coordination, and orchestration of the collaborative mitigation is based on Blockchain technology; the proposed framework was tailored to federated trusted environments of Tier 1 ISPs [19].

C. Federated Learning for DDoS attacks

In [5], a DDoS detection and mitigation framework for Internet of things (IoT) environments is proposed. IoT nodes collaborate to train a common Machine Learning (ML) model via the Federated Averaging technique to accurately detect malicious traffic. This is subsequently filtered in a distributed fashion at multiple IoT nodes. In [6], a DDoS detection schema based on Federated Averaging is presented. It uses flow-based features to identify various DDoS attack types; DDoS mitigation was considered out of their scope. Similarly, solutions for collaborative DDoS detection based on FL are proposed in [20], [21], [22]. Moreover, in [23], a multi-task FL model is proposed. It concurrently performs DDoS detection, VPN/Tor traffic recognition and network application identification. This reduces the management overhead of individual ML models while respecting network data privacy. However, these works do not consider DDoS attack mitigation and are not directly applicable to interconnected cyber infrastructures.

D. Key Contributions

We present below the key contributions of this work:

- In related efforts, collaborators exchange either coarse-grained data for DDoS detection [12], [14], or predefined static rules [13], [15]; they also focus only on attack data [14], [15], [17], [18]. In contrast, our FL scheme (i) enables for DDoS detection using both benign and attack data without exposing private information and (ii) creates ML models with generalization capabilities able to identify "unseen" (not trained with) benign and attack packets.
- Most FL schemes [5], [6], [23] simulate multi-domain data by splitting single datasets into multiple parts. Instead, we employ production network traffic aggregated by collaborating domains, e.g. AS's, to perform fully realistic experiments.
- Typical filtering mechanisms employed in collaborative DDoS mitigation [17], [18] have the following drawbacks: they (i) support packet filtering based on limited

packet field combinations and (ii) pose limitations on the supported number of rules. In contrast, we consider an XDP-based data plane programmable firewall that enables packet filtering based on arbitrary packet field combinations and scales its performance with the number of cores.

III. Design Principles & High Level Overview

A. Design Principles

We present below the core design principles of our schema:

- *Collaborative DDoS Detection via Federated Learning*: Network traffic is classified as malicious or benign based on pre-agreed supervised learning models (e.g. common MLP architecture and feature definitions) trained via the Federated Averaging technique [2]. Thus, collaborating domains converge to federated ML models without sharing private data. This enables them to learn from *foreign* benign and attack packets without sharing their content.
- *DDoS Mitigation via cloud-native and scalable programmable firewalls based on the XDP framework*: Similarly to [24], we employ data plane programmability frameworks (XDP) to design high-performance Commercial off-the-shelf (COTS) firewalls for cyber infrastructures. In contrast to legacy router based filters, these can be programmed to match and promptly block arbitrary packet field combinations while scaling their resources on-demand in an NFV-compliant fashion.
- *Upstream propagation of DDoS filtering requests*: Our scheme enables the dissemination of signature-based filtering rules among participating domains, e.g. collaborating Autonomous Systems (cAS's). These rules are distributed to collaborators based on methods that are native to multi-domain environments, i.e. using BGP sessions. These can be used to effectively block attacks before reaching the victim domain, extending the limited filtering capabilities of blackholing or flow-based protection mechanisms.

B. High-level Overview

A high-level design of the proposed architecture for collaborative DDoS protection is depicted in Fig. 1. Malicious actors launch DDoS attacks attempting to overwhelm the network bandwidth and/or processing resources of a host IP/subnet located in the victim domain, e.g. victim AS (vAS). Both malicious and benign traffic reach vAS via interconnected domains e.g. cAS(1), cAS(2). Monitoring (packet-based) data are exported by network devices e.g. edge routers and organized in packet signatures; these are in turn used as input to the DDoS Detection app. There, pre-trained Multilayer Perceptrons (MLPs) classify packet signatures as malicious or benign (step i). The MLP training process has been conducted via FL techniques that enable distributed and privacy-preserving learning amongst collaborating AS's (cAS's). The training process is orchestrated by the Collaboration Manager (step a) in pre-agreed time-periods. We assume that the FL training process will be completed before any of the collaborating parties is attacked. Furthermore, our proposed schema adopts a cross-silo federated learning architecture among a moderate number of highly available and fault tolerant cyber infrastructures, e.g. georedundant data centers.

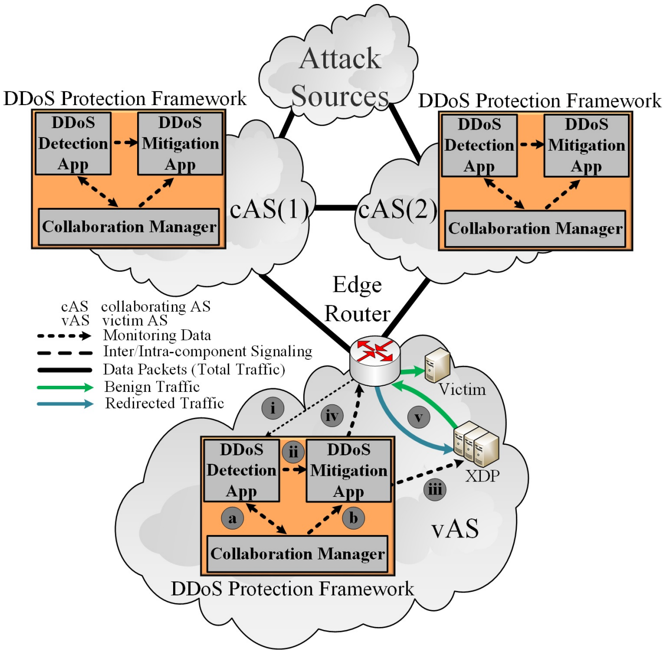


Fig. 1. Collaborative DDoS Protection Architecture

The DDoS Detection app conveys to the DDoS Mitigation app the identified malicious signatures and the corresponding victim IP/subnet (step ii). In turn, a Firewall Instance (FI) is created (step iii) that uses the identified malicious signatures as filtering rules. After FI instantiation, the DDoS mitigation app notifies the edge router to redirect traffic destined to the victim to the corresponding FI (step iv). Malicious traffic is dropped while benign traffic is bounced back and forwarded to its original destination (step v).

The DDoS Detection app based on traffic/system metrics, e.g. increased link utilization, can request help from upstream/peer domains to protect its network/compute resources. The Collaboration Manager identifies adjacent cAS's that forward attack traffic [17] and populates the identified malicious signatures coupled with the victim IP address. cAS's, willing to filter malicious traffic, receive the requested signatures and signal their own DDoS Mitigation app (step b) to on-demand mitigate the offending traffic.

IV. Collaborative DDoS Detection & Mitigation Architecture

Our framework consists of three distinct applications (apps): (i) DDoS Detection, (ii) DDoS Mitigation, and (iii) Collaboration Manager. These are detailed in subsections below.

A. DDoS Detection Application

The DDoS Detection app retrieves packet-based data from external monitoring mechanisms and identifies malicious

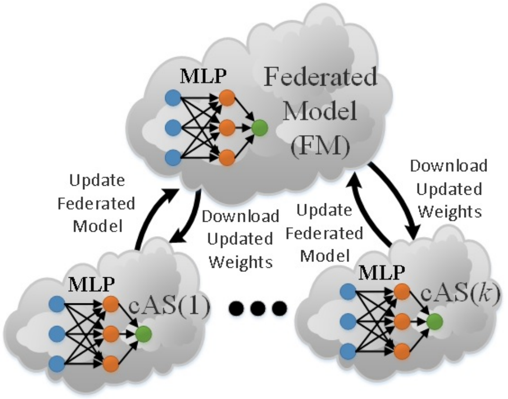


Fig. 2. Federated Learning for collaborating AS's

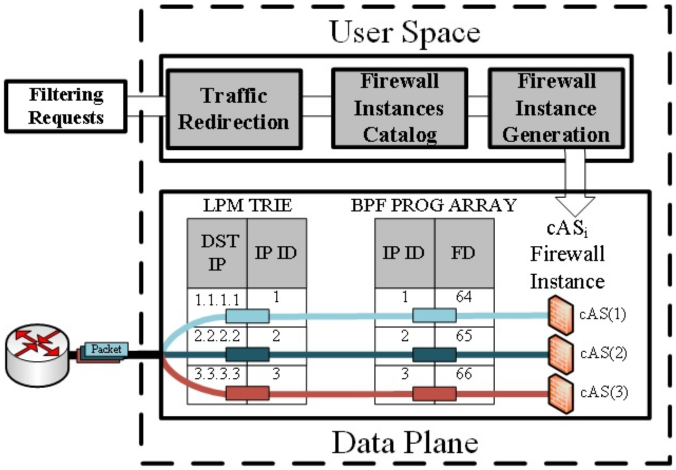
packet signatures. Signature classification is conducted by MLPs trained via FL techniques.

Monitoring data are collected within time-windows and aggregated based on preselected packet fields, forming packet signatures. Packet signatures may be represented by a vector X = [x_1 x_2 … x_i], where x_i corresponds to packet field value i. Vectors X are used as input to MLPs, that classify them as malicious or benign. Malicious signatures are organized per destination IP address to generate filtering rules at the DDoS Mitigation app of the vAS.

The accuracy of the MLP model affects significantly the identification of malicious packets and the subsequent mitigation, since filtering rules are based on the identified malicious signatures. To improve the accuracy of the MLP model without compromising privacy, we considered a collaborative learning approach based on Federated Averaging [2]. Prerequisite for training a Federated Model (FM) is the use of a common MLP model coordinated by a third trusted party (Fig. 2). We consider that FM may reside in a neutral independent coordinator. Such understanding is common in Internet architectures e.g. major Internet eXchanges (IXes) [25]. Indicatively, an IX on top of the offered interconnection services may offer an FL coordination service to its customers. Therefore, as an independent entity, it would ensure high availability and guarantee customer privacy against inference attacks [26].

Initially, packet fields (features) relevant to an attack vector must be selected [10]. To reduce training times and the FM complexity, inconsequential features may be eliminated. This could be achieved by leveraging on feature importance methods or interpretable machine learning techniques [27].

Participating domains agree on common MLP hyperparameters (e.g. FM architecture, learning rates). The training process starts with an initial FM weight vector. In each training iteration (round) a new set of weights w_{FM} is evaluated and distributed amongst the k collaborating AS's. Each collaborator i = 1…k uses w_{FM} as initial weights and subsequently updates its local weights w_i based on its private training data N_i. These are conveyed to the FM third party coordinator that calculates



DST IP	IP ID
1.1.1.1	1
2.2.2.2	2
3.3.3.3	3

IP ID	FD
1	64
2	65
3	66

Fig. 3. DDoS Mitigation Application Architecture

the new weights w_{FM} based on the following equation:

$$w_{FM} = \sum_{i=1}^{k} \frac{N_i}{N} w_i,\ where\ N = \sum_{i=1}^{k} N_i$$

A round is completed after w_{FM} calculation with new weights distributed to the cAS's. Finally, each cAS adopts the FM update that achieves the highest accuracy on its local validation dataset. In case collaborators share their local accuracies per round, a common FM may be universally adopted once the (weighted) average accuracy for all participants reaches an acceptable level.

B. DDoS Mitigation Application

The DDoS Mitigation app receives requests for ongoing attacks either from the DDoS Detection app of the victim or the Collaboration Manager of cAS's. Subsequently, this app may raise appropriate mitigation countermeasures.

Typical filtering mechanisms e.g. Access Control Lists (ACLs), can match packets based on combinations of multiple but predefined packet fields. These rules are stored in network devices with stringent memory limitations [28]. Thus, offloading DDoS filtering to an external firewall should (i) support any packet field combination (signature) that can block malicious packets, (ii) have no limit on the number of filtering rules, and (iii) allow dynamic filtering rules creation, read, update, and deletion (CRUD).

We implemented the mitigation app based on the XDP data plane programmability framework. Note that our XDP-based firewall conforms to the NFV paradigm and could be offered as a VNF service in cloud infrastructures, similarly to [24]. As the XDP framework performance depends on the number of processors, our firewall is able to adjust to high-speed DDoS attacks by dynamically assigning additional processor instances (vCPU's).

XDP memory structures for storing packet signatures are Berkeley Packet Filter (BPF) Maps; these do not allow ternary packet field matching, i.e. using wildcards on packet fields. Therefore, for developing an XDP firewall program that supports various types of signatures, a BPF MAP per signature type would be required. This would (i) degrade the total performance due to multiple memory lookups [10], [29] (proportional to the signature types) and (ii) introduce downtime since for each BPF Map addition, XDP programs require reloading.

The DDoS Mitigation app was designed to conform with the aforementioned XDP limitations. As depicted in Fig. 3, it is based on a user space and a data plane program. The former manages signatures installation while the latter performs packet filtering. The user space program receives filtering requests from vAS and/or cAS's e.g. victim IP/network, signatures. If there are no signatures, a unique identifier IP ID is created (Firewall Instances Catalog). Packet signatures are transformed into XDP programs, i.e. Firewall Instances (FIs), via appropriate Jinja templates [30] (Firewall Instance Generation). Each FI parses packet fields and their values that form the requested signatures. Subsequently, it contains if-then-else conditions to match and drop malicious packets. Each generated FI is indexed by a unique File Descriptor (FD) and can be accessed, updated or deleted dynamically, without affecting already deployed FIs. After FI instantiation, the user space program signals the edge router to redirect the network traffic destined to the victim IP/subnet.

The data plane program receives the redirected packets, parses their destination IP, and performs a lookup on an LPM (Longest Prefix Match) TRIE BPF Map; this matches IP addresses/subnets to their corresponding IP ID. Subsequently, the IP ID is used as input to a special memory structure BPF PROG ARRAY, that passes the packet to its corresponding FI. According to the FIs signatures, malicious packets are blocked while benign packets are appropriately forwarded. Note that if the total number of signatures increases significantly, a packet signature reduction process may be enforced as proposed in our previous work [10].

C. Collaboration Manager

The Collaboration Manager (CM) is an application that (i) handles filtering requests for/from collaborators and (ii) orchestrates the FL training process based on the private benign and attack traces of individual collaborators, without exchanges of any sensitive data.

CM employs the BGP protocol to serialize and convey filtering requests. We needed to overcome the limitation of the predefined packet fields imposed by BGP Flowspec. To that end, victim's CM BGP Speaker initializes a BGP session with collaborators CM advertising the support of the Content-URI address family [31], similar to [17]. This allows the advertisement of specialized BGP Update messages that include URIs pointing to the requested filtering rules (signatures).

Note that, the use of BGP enables our scheme to leverage on well-established tools, e.g. Resource Public Key Infrastructure (RPKI), could be extended to check collaborators eligibility on announcing IP prefixes/addresses of the aforementioned requested rules.

CM coordinates also the Federated Averaging training process. This is an offline procedure between the collaborators and a neutral third party hosting the Federated Model. CM retrieves the weights from each training round and publishes them to the FM via a message broker, i.e. RabbitMQ [32]. Subsequently, it receives the generated weights calculated as the average of collaborators weights. The proposed message broker enables for collaborators authentication, inter-collaborators private agreements (e.g. sharing accuracy results on their local datasets) and reliable weights delivery.

Note that typical FL use cases [2], [5] consider as collaborators low throughput devices. In our case the size of MLPs weights that are exchanged between cAS's have negligible impact on the high-throughput links that interconnect them.

V. Experimental Evaluation

We implemented all software applications of the proposed architecture and deployed them in our laboratory testbed. The DDoS Detection app was based on PyTorch and PySyft libraries. The Collaboration Manager was based on Ryu's SDN Controller BGP Speaker [33] and RabbitMQ message broker [32]. The DDoS Mitigation app was deployed as a VNF within a Virtual Machine (VM); the hypervisor physical machine was equipped with an Intel i7-2600 CPU and a 10G Netronome SmartNIC [34] (XDP-enabled). This was directly connected to another VM with high-speed packet generation capabilities [35].

To assess the detection accuracy and mitigation performance of our mechanism, we considered DNS Amplification attacks. In subsection V-A we analyze the employed DNS datasets. In subsection V-B we compare the classification accuracy of the proposed Federated Model to non-collaborative approaches. In subsection V-C we showcase the packet processing performance of our mitigation mechanism.

A. Datasets

Amplification attacks are among the DDoS attacks with the highest impact on cyber infrastructures. Therefore, we focused our experiments on a commonly encountered attack vector, the DNS Amplification attack. Our scheme could be extended to other DDoS attack vectors, e.g. NTP Amplification attacks and TCP SYN Floods, by appropriately selecting the feature set of the machine learning models. However, our evaluation focuses on assessing the benefits of FL for collaborative DDoS protection. Therefore, we performed our experiments based on a single DDoS vector, i.e. DNS Amplification attacks.

As benign traffic, we used DNS traffic traces from a 10G link between the WIDE Japanese backbone and DIX-IE Internet Exchange [36]. Benign DNS traffic was aggregated per destination AS based on BGP data [37]. In turn, AS's were sorted in descending order based on the total received packets; dataset B(i) contains benign traffic destined to AS's ranked by incoming traffic.

As malicious traffic, we used seven DNS Amplification attacks contained in the Booters dataset [38], henceforth referred to as A(i). Attacks in A(1), A(2), A(3), A(6) and

TABLE I
PACKET FEATURES FOR DNS PACKET CLASSIFICATION

Packet Fields (Features)		
ip.length	*dns.flags.checkdisable*	*dns.count.answers*
udp.length	*dns.flags.authoritative*	*dns.count.auth_rr*
dns.qry.name	*dns.flags.truncated*	*dns.count.add_rr*
dns.qry.type	*dns.flags.recdesired*	*dns.flags.recavail*
-	*dns.flags.authenticated*	-

A(7) generated type ANY DNS responses. By contrast, in A(4) and A(5), attackers generated type A DNS responses. Specifically, A(4) contains responses for a single domain name that resolved into a very large number of IP addresses. A(5) corresponds also to a type A attack, in which attackers could not generate responses with heavy payload.

B. DDoS Detection Accuracy

In this subsection, we evaluate the classification accuracy of our FL approach and compare it to individual approaches. Specifically, we considered 7 collaborating AS's, henceforth referred to as cAS(i), where i=1...7. Each cAS(i) has access to its own private traffic mix M(i) that contains attack dataset A(i) and benign dataset B(i).

We trained each cAS(i) model individually based on dataset M(i) using an MLP of 13 input neurons, 27 (13x2+1) hidden and a single output node, as suggested in [39]. MLP weights were updated based on the Adam [40] algorithm. The features employed for the MLP (see Table I) are based on the packet fields presented in our previous work [10]. Note that we ignored packet fields whose values are (i) identical in attack and benign packets or (ii) arbitrarily generated in the utilized datasets (e.g. DNS ID, TCP sequence number). These types of features (i), (ii) are not able to enhance the classification accuracy of ML models and can be safely ignored upon collaborators agreements.

The Federated Model (FM) was trained using the same MLP architecture with weights conveyed from all collaborators. The hyperparameters for cAS(i) models and FM were tuned based on grid search [41], using validation datasets comprising of 30% of datasets M(i).

We evaluated the trained models using as test datasets A(i) and B(i). The metrics considered in our evaluation were the True Positive Rate (TPR) and the True Negative Rate (TNR) achieved on the testing datasets. TPR is defined as the percentage of attack packets that are classified as malicious, while TNR is the percentage of benign packets that are correctly classified as benign.

In Fig. 4, we depict the average TPR and TNR achieved by each AS(i) and the FM for all datasets, A(1)-A(7) and B(1)-B(7) accordingly. Note that we excluded A(5) from the average TPR calculation, since it introduced insignificant malicious traffic (˜6 Mbps). The FM achieves on average the highest combination of TPR and TNR amongst individual cAS's models, as shown in Fig. 4.

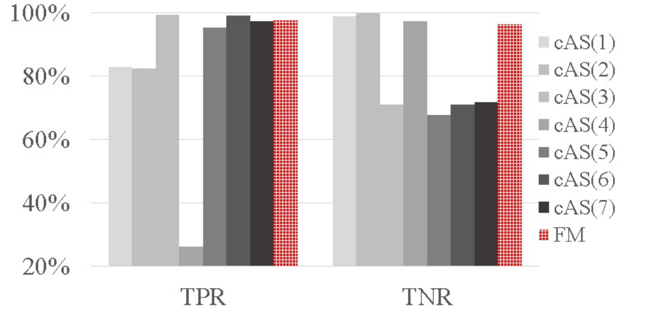


Fig. 4. Average TPR and TNR of Individual Collaborator Models and Federated Model on all datasets

Notably, the individual model of cAS(4) was unable to achieve a high TPR on the testing dataset; this model was trained on data that significantly diverged from the testing dataset that assembled the individual attack datasets of collaborators. However, despite contributing poorly to the collaboration scheme, cAS(4) is able to generalize to the testing dataset using the FL model. Therefore, provided that most collaborators contribute significantly to the training process, we observe that our FL scheme enables all collaborating parties to identify benign and attack packets that as individuals might misclassify them.

C. Packet Filtering Performance

In this subsection, we assess the packet filtering performance of the DDoS Mitigation app. We evaluate its packet processing performance considering CPU scalability capabilities and the number of supported Firewall Instances (FIs).

We replayed DNS traffic consisting of packets that can be dropped by a single signature, i.e. *dns.qry.type=255 and dns.qry.name=<Root>*, per FI. This signature can block all the attack traffic contained in datasets A(1), A(2) and A(3); note that this type of signature can be generated based on the signature reduction technique that was presented in [10].

We launched concurrent attacks ranging from 10 to 1000 that target different collaborators with accumulated throughput of 10 Million packets per second (Mpps). To evaluate the packet processing performance, we counted the number of packets that were processed and dropped. This enables us to assess our firewall, as a service offered to collaborating AS's. In Fig. 5, we evaluate the scalability of our firewall in terms of the deployed FIs implemented as a VNF within a VM and evaluated under a vertical scaling scenario with 1, 2 and 3 CPU cores.

The packet processing performance of our mechanism scales almost linearly with the number of cores. Such behavior is also validated in [11], [42]. As expected, increasing the number of collaborators decreases the overall packet processing rate of our firewall. Specifically, this is reduced linearly between 10 and 200 FIs and from that point it remains the same. The enhanced performance for the small number of FIs is attributed to level one (L1) instruction cache hits [43] while after a specific number of FIs the L1 instruction cache misses do not affect the overall performance. These conclusions

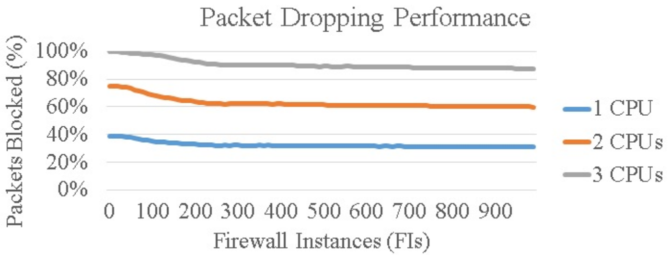


Fig. 5. DDoS Mitigation app Scalability

were validated using the perf tool [44] that provides CPU performance statistics for user-defined time intervals.

In total, our approach can handle successfully up to 1000 concurrent attacks targeting an equal number of collaborators. These correspond to the number of concurrent blackholed IP prefixes announced in a large European IX [45]. Thus, the proposed firewall can serve as a scalable and efficient mechanism for large-scale attack mitigation in federated cyber infrastructures.

VI. Conclusions

In this paper we proposed a collaborative DDoS protection framework for interconnected cyber infrastructures. Our approach leveraged on the Federated Learning (FL) paradigm for collaborative and privacy-aware DDoS detection without requiring exchanges of sensitive data among collaborating entities. Attack mitigation was based on VNF-compliant scalable and programmable firewalls that were instantiated on-demand by victims. Specifically, our schema analyzed, within time windows, packet-based data forming signatures. These were used as input to supervised Machine Learning models, trained cooperatively via the Federated Averaging technique. Suspicious traffic was redirected to programmable (XDP-based) firewalls to be filtered out. During massive attacks, our schema enabled victims to raise filtering requests on collaborating domains to block them, presumably early in attack paths.

Our framework was evaluated both in terms of detection accuracy and mitigation performance for typical DNS Amplification DDoS attacks, which are considered among the most devastating DDoS attack vectors. The conducted experiments considered benign and malicious production traffic in cyber infrastructure environments. The FL approach enabled collaborators to accurately classify benign and attack packets improving their individual accuracy. Based on the achieved packet processing performance, the proposed programmable firewall provides a scalable filtering mechanism for evolving federated cyber infrastructures.

As future work, our signature-based approach will be extended via multi-task learning techniques [23] to concurrently recognize multiple attack vectors. To reduce training times and the complexity of the generated Federated Model, we will explore federated feature selection mechanisms [46]. Moreover, we plan to incorporate trust-based schemes [47] to improve performance, robustness and security of FL. Furthermore, we plan to compare our XDP-based mitigation mechanism with other data plane programmability techniques, such as P4 [48] and DPDK [49]. We will extend our framework to the collaborative and privacy-aware detection and mitigation of other cyber threats, e.g. for traffic generated by Domain Generation Algorithms (DGA's) [50]. Finally, some components of our proposed schema may be subject to single point of failure. These concerns may be mitigated by distributing the operation of these components, e.g. via leveraging on Blockchain technology [18].

References

[1] "General Data Protection Regulation - GDPR", Accessed September 30, 2022. [Online]. Available: https://gdpr-info.eu/.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson and B. Aguera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", in Proceedings of the International Conference on Artificial Intelligence and Statistics, April 2017, pp. 1273–1282.

[3] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis and W. Shi, "Federated Learning of Predictive Models from Federated Electronic Health Records", International Journal of Medical Informatics, Volume 112, pp. 59–67, April 2018.

[4] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson and M. Jirstrand, "A Performance Evaluation of Federated Learning Algorithms", in Proceedings of the Workshop on Distributed Infrastructures for Deep Learning, December 2018, pp. 1–8.

[5] J. Li, L. Lyu, X. Liu, X. Zhang and X. Lyu, "FLEAM: A Federated Learning Empowered Architecture to Mitigate DDoS in Industrial IoT", IEEE Transactions on Industrial Informatics, Volume 18, no. 6, pp. 4059–4068, June 2021.

[6] Q. Tian, C. Guang, C. Wenchao and W. Si, "A Lightweight Residual Networks Framework for DDoS Attack Classification Based on Federated Learning", in Proceedings of the Conference on Computer Communications Workshops, May 2021, pp. 1-6.

[7] D. Turk, "RFC 3882 - Configuring BGP to Block Denial-of-Service Attacks", Accessed September 30, 2022. [Online]. Available: https://datatracker.ietf.org/doc/rfc3882/.

[8] P. Marques, N. Sheth, R. Raszuk, B. Greene, J. Mauch and P. McPherson, "RFC 5575 - Dissemination of Flow Specification Rules", Accessed September 30, 2022. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc5575.

[9] M. Dimolianis, A. Pavlidis and V. Maglaris, "SYN Flood Attack Detection and Mitigation using Machine Learning Traffic Classification and Programmable Data Plane Filtering", in 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN), March 2021, pp. 126–133.

[10] M. Dimolianis, A. Pavlidis and V. Maglaris, "Signature-Based Traffic Classification and Mitigation for DDoS Attacks Using Programmable Network Data Planes", IEEE Access, Volume 9, pp. 113061–113076, August 2021.

[11] T. Høiland-Jørgensen, J. D. Brouer, D. Borkmann, J. Fastabend, T. Herbert, D. Ahern and D. Miller, "The eXpress Data Path: Fast Programmable Packet Processing in the Operating System Kernel", in Proceedings of the International Conference on Emerging Networking EXperiments and Technologies, December 2018, pp. 54–66.

[12] Y. Chen, K. Hwang and W. S. Ku, "Collaborative Detection of DDoS Attacks over Multiple Network Domains", IEEE Transactions on Parallel and Distributed Systems, Volume 18, no. 12, pp. 1649–1662, December 2007.

[13] J. François, I. Aib and R. Boutaba, "FireCol: A Collaborative Protection Network for the Detection of Flooding DDoS Attacks", IEEE/ACM Transactions on Networking, Volume 20, no. 6, pp. 1828–1841, April 2012.

[14] J. Steinberger, B. Kuhnert, A. Sperotto, H. Baier and A. Pras, "Collaborative DDoS Defense using Flow-based Security Event Information", in Proceedings of the Network Operations and Management Symposium, April 2016, pp. 516–522.

[15] C. Hesselman and R. Yazdani, "CONCORDIA Cyber security cOmpeteNCe fOr Research anD InnovAtion DDoS Clearing House for Europe Cross-sector Pilot Demo", Accessed September 30, 2022. [Online]. Available: https://www.sidnlabs.nl/downloads/2deJudioEsd0oFWufTXdV9/099fa8c92f7d601e0669bec73b2fa272/NEW-20200123-CONCORDIA-T3.2-demo-review-final.pdf.

[16] D. Wagner, D. Kopp, M. Wichtlhuber, C. Dietzel, O. Hohlfeld, G. Smaragdakis and A. Feldmann, "United We Stand: Collaborative Detection and Mitigation of Amplification DDoS Attacks at Scale", in Proceedings of the SIGSAC Conference on Computer and Communications Security, November 2021, pp. 970-987.

[17] K. Giotis, M. Apostolaki and V. Maglaris, "A Reputation-based Collaborative schema for the Mitigation of Distributed Attacks in SDN Domains", in Proceedings of the Network Operations and Management Symposium, April 2016, pp. 495–501.

[18] A. Pavlidis, M. Dimolianis, K. Giotis, L. Anagnostou, N. Kostopoulos, T. Tsigkritis, I. Kotinas, D. Kalogeras, and V. Maglaris, "Orchestrating DDoS Mitigation via Blockchain-based Network Provider Collaborations", Knowledge Engineering Review, Volume 35, pp. 1–17, April 2020.

[19] "The Global Leaders' Forum launches Communications Blockchain Network (CBN) - Deutsche Telekom Global Carrier", Accessed September 30, 2022. [Online]. Available: https://globalcarrier.telekom.com/newsroom/news/news-pages/global-leaders-forum-launches-communications-blockchain-network.

[20] J. Li, Z. Zhang, Y. Li, X. Guo and H. Li, "FIDS: Detecting DDoS Through Federated Learning Based Method", in Proceedings of the International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), October 2021, pp. 856-862.

[21] R. Doriguzzi-Corin and D. Siracusa, "FLAD: Adaptive Federated Learning for DDoS Attack Detection", arXiv, May 2022.

[22] J. Zhang, P. Yu, L. Qi, S. Liu, H. Zhang and J. Zhang, "FLDDoS: DDoS Attack Detection Model based on Federated Learning", in Proceedings of the International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), October 2021, pp. 635-642.

[23] Y. Zhao, J. Chen, D. Wu, J. Teng and S. Yu, "Multi-Task Network Anomaly Detection using Federated Learning", in Proceedings of the International Symposium on Information and Communication Technology, December 2019, pp. 273–279.

[24] N. Van Tu, J. H. Yoo and J. W. K. Hong, "Building Hybrid Virtual Network Functions with eXpress Data Path", in Proceedings of the International Conference on Network and Service Management, October 2019, pp. 1-9.

[25] "DE-CIX – Deutscher Commercial Internet Exchange", Accessed September 30, 2022. [Online]. Available: https://www.de-cix.net/.

[26] L. Lyu, H. Yu and Q. Yang, "Threats to Federated Learning: A Survey", arXiv, March 2020.

[27] C. Molnar, "Interpretable Machine Learning: A Guide for Making Black Box Models Explainable", Accessed September 30, 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book/.

[28] "Cisco ASR 9000 Series Aggregation Services Router Routing Configuration Guide - Implementing BGP Flowspec", Accessed September 30, 2022. [Online]. Available: https://www.cisco.com/c/en/us/td/docs/routers/asr9000/software/asr9k_r5-2/routing/configuration/guide/b_routing_cg52xasr9k/b_routing_cg52xasr9k_chapter_011.html.

[29] S. Miano, R. Doriguzzi-Corin, F. Risso, D. Siracusa and R. Sommese, "Introducing SmartNICs in Server-Based Data Plane Processing: The DDoS Mitigation Use Case", IEEE Access, Volume 7, pp. 107161-107170, August 2019.

[30] "Jinja Documentation", Accessed September 30, 2022. [Online]. Available: https://jinja.palletsprojects.com/en/3.0.x/.

[31] A. Narayanan, S. Previdi and B. Field, "BGP Advertisements for Content URIs", Accessed September 30, 2022. [Online]. Available: https://slideplayer.com/slide/8149985/.

[32] "RabbitMQ Message Broker", Accessed September 30, 2022. [Online]. Available: https://www.rabbitmq.com/.

[33] "Ryu Component-based Software Defined Networking Framework", Accessed September 30, 2022. [Online]. Available: https://github.com/faucetsdn/ryu.

[34] "Netronome Agilio SmartNICs", Accessed September 30, 2022. [Online]. Available: https://www.netronome.com/products/agilio-cx/.

[35] "PF RING – ntop", Accessed September 30, 2022. [Online]. Available: https://www.ntop.org/products/packet-capture/pf_ring/.

[36] K. Cho, K. Mitsuya and A. Kato, "Traffic Data Repository at the WIDE Project", in Proceedings of the USENIX Annual Technical Conference, June 2000.

[37] "BGP Routing Table Analysis - IPv4 Prefixes and their Origin ASNs", Accessed September 30, 2022. [Online]. Available: https://thyme.apnic.net/current/.

[38] J. J. Santanna, R. Van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Z. Granville and A. Pras, "Booters - An Analysis of DDoS-as-a-Service Attacks", in Proceedings of the International Symposium on Integrated Network Management, May 2015, pp. 243–251.

[39] C. Siaterlis and V. Maglaris, "Detecting Incoming and Outgoing DDoS Attacks at the Edge using a Single Set of Network Characteristics", in Proceedings of the Symposium on Computers and Communications, June 2005, pp. 469–475.

[40] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization", in Proceedings of the 3rd International Conference on Learning Representations, May 2015, pp. 1–15.

[41] "Hyperparameter Optimization - Grid Search", Accessed September 30, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Hyperparameter_optimization#Grid_search.

[42] O. Hohlfeld, J. Krude, J. H. Reelfs, J. Rüth and K. Wehrle, "Demystifying the Performance of XDP BPF", in Proceedings of the Conference on Network Softwarization, June 2019, pp. 208–212.

[43] J. L. Hennessy and D. A. Patterson, "Computer Architecture - A Quantitative Approach 5th edition". Morgan Kaufmann/Elsevier, 2012.

[44] "perf (Linux) - Wikipedia", Accessed September 30, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Perf_(Linux).

[45] M. Nawrocki, J. Blendin, C. Dietzel, T. C. Schmidt and M. Wählisch, "Down the Black Hole: Dismantling Operational Practices of BGP Blackholing at IXPs", in Proceedings of the Internet Measurement Conference, October 2019, pp. 435–448.

[46] P. Cassarà, A. Gotta and L. Valerio, "Federated Feature Selection for Cyber-Physical Systems of Systems", arXiv, September 2021.

[47] A. Gholami, N. Torkzaban and J. S. Baras, "On the Importance of Trust in Next-Generation Networked CPS Systems: An AI Perspective", arXiv, April 2021.

[48] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese and D. Walker, "P4: Programming Protocol-Independent Packet Processors", ACM SIGCOMM Computer Communication Review, Volume 44, no. 3, pp. 87-95, July 2014.

[49] "DPDK: Data Plane Development Kit", Accessed September 30, 2022. [Online]. Available: https://www.dpdk.org/.

[50] D. Plohmann, K. Yakdan, M. Klatt, J. Bader and E. Gerhards-Padilla, "A Comprehensive Measurement Study of Domain Generating Malware", in Proceedings of the USENIX Security Symposium, August 2016, pp. 263-278.