FedITN: Toward a Novel Dataset for Collaborative Intrusion Detection in Federated Heterogeneous IT Networks

Léo Lavaur
IMT Atlantique, IRISA, Cyber CNI
leo.lavaur@imt-atlantique.fr

Fabien Autrel
IMT Atlantique, IRISA
fabien.autrel@imt-atlantique.fr

***Abstract*—**

***Index Terms*—dataset, federated machine learning, intrusion detection systems, collaboration, heterogeneity**

I. Introduction

A. Use case

Information Technology (IT) networks, the Industrial Internet of Things (IIoT), and smart healthcare are examples of application domain where intrusion detection is relevant. Each comes with its set of constraints, and Federated Intrusion Detection Systems (FIDSs) currently struggle to handle heterogeneity [1]. As most use case involves an IT component, and we focus on detecting intrusion in IT infrastructures. Because of virtualization technologies, it is also simple to instantiate, and can produce enough heterogeneity **[?]**to impair FIDSs detection efficiency.

We first consider a setup of four organizations involved in a collaboration system based on Federated Learning (FL). In FL-enabled intrusion detection, each participant trains a Machine Learning (ML) model on data that has been collected locally. This model is then used to detect intrusion in the network. In parallel, the model is trained online on new data. Each client owns an enterprise-level infrastructure of varying size, from 4 to 20 nodes. In the context of the generation of this dataset, clients share both, feature selection and models (or at least part of it, *e.g.* the last layers of a Deep Neural Network (DNN)). However, there are differences amongst clients in terms of data distribution, network segmentation, services, and volume.

In this use case, participants are companies or organizations that have a digital presence and host web services, *e.g.* public web server. We do not consider means of production, while it may be done in future works.

B. Motivation

In previous work [1], we reviewed the state of the art of FL. We highlighted five current issues, and the associated research directions, namely: (1) performance, (2) adaptability and scalability, (3) security and privacy, (4) resiliency, (5) and transferability.

Transferability is a key direction for FIDS [1], particularly in Cross-silo Federated Learning (CS-FL) settings [2]. In their work, Kairouz *et al.* [2] show that CS-FL can lead to high heterogeneity in both, data and model, which degrades performance. Adaptability and scalability are especially relevant in Cross-device Federated Learning (CD-FL) [2], as a high number of clients can also negatively impact the model's performance [3], *e.g.* with client dropouts, or asynchronous client responses [4]. Alongside performance [5], [6], security, privacy, and resiliency (*i.e.* self-defense and self-healing) are all highly discussed topics in the literature of FL [7], [8], and thus apply to FIDS [1].

A majority of the related works that focus on IT networks rely on known datasets for intrusion detection, such as NSL-KDD [9], UNSW-NB15 [10], and CIC-IDS2017 [11]. To fit federated settings, datasets are artificially partitioned and distributed amongst participants. While this approach can provide an approximation to the heterogeneity in real-world settings, it is limited to sample distribution **[?]**. However, heterogeneity in IT networks can be found in terms of network size **[?]**, network segmentation **[?]**, services—e.g. NFS or Samba for file-sharing **[?]**, service version **[?]**, and data-collection location **[?]**. Sun *et al.* [6], [12] provide a dataset for distributed data collection, but focus on a specific use case of IT networks, with only two types of anomalies ("SMB attacks" and "TCP SYN flood"). Furthermore, data is already preprocessed and only provided as images, which restricts the range of algorithms that can be used afterward.

We propose different approach to represent heterogeneity in IT networks, by providing a novel dataset with multiple topologies generating benign data, and a labelled set of attacks that overlap between the different organizations. This new dataset allows evaluating multiple key research directions for FIDS [1]:

(a) *performance against heterogeneity*: the ability for the federation to maintain high performance with heterogeneous participants;
(b) *knowledge transfer between clients*: the ability for one client to recognize patterns unknown in his local data;
(c) *model adaptability*: the ability of a local model to evolve in time, and to adapt to new devices in the local network;
(d) *generation capability*: the ability for a local model to correctly characterize behavior for similar but different services;
(e) *vulnerability to adversarial samples*: the ability of a local

model to correctly prevent evasion or poisoning attacks based on adversarial learning.

comment:

Note: the last item is not directly related to FIDS [1], but to the works of Hassan and Thomas (TSP) on adversarial attacks. FL is particularly vulnerable to those, and a such dataset can be used for that IF we have their input on the topic.

C. Highlights of the paper

Our contributions are threefold:

I. We introduce a new dataset named FedITN, with covers all classes of network-related attacks, based on recent datasets and two MITRE matrices: ATT&CK® [13] and D3FEND® [14]. The dataset contains normal and attack traffic, and can be used for a variety of research directions.

II. We evaluate our dataset and compare it with known datasets from the literature, namely NSL-KDD [9], UNSW-NB15 [10], and CIC-IDS2017 [11].

III. We reproduce state-of-the-art approaches and show the impact of real-world heterogeneity on federation performance.

This paper is structured as follows. Section II reviews related works. Section III displays the methodology used to generate the dataset, defines the terminology, and details the specification process. Section IV describes the dataset and its content: considered attacks, collected and selected features, architectures, and played attack scenarios. Finally, Section V proposes a comparison of the dataset with and the ones in literature, by implementing state-of-the-art FIDS approaches.

II. Building reproducible experiments for FIDSs

A. Existing datasets

The Kharon project produced a dataset for malware analysis and classification [15]. The dataset contains seven malware and provides flow grahs and detailed analysis for each one.

Tarasov *et al.* [16] propose a dataset for deduplication analysis, showing that while the topic already existed in the literature, datasets in these papers were either (a) private and thus not available to the public, (a) hard to reproduce, or (a) synthetic dataset lacking details. In total, three quarters of the reviewed datasets were not usable for cross-evaluation. This situation is comparable to the situation of FIDS where most datasets are either not applicable, synthetic, or lack generalization capability.

B. Reproducible experiments

The authors of [17] propose a reproducible experiment testbed for log-based intrusion detection. Their work defines the

III. Methodology

This section covers the methodology developed to build this dataset. The dataset has to check three objectives: (i) be representative of the real-world heterogeneity of IT networks in federated settings; (ii) be comparable to existing datasets of the literature; (iii) implement realistic attacks and attack scenarios.

A. Definitions

This section goes through the terminology used in the dataset.

1) Attack:

2) Scenario:

3) Topology:

B. Considered attacks and classification

As pointed out in introduction of this section III, the dataset has to conceal two objectives: be comparable to the literature, and implement realistic attacks. Therefore, we start by comparing the most used datasets in the literature—namely NSL-KDD [9], UNSW-NB15 [10], and CIC-IDS2017 [11]—to extract relevant attack classifications. These classifications are put in perspective of the MITRE ATT&CK® [13] and D3FEND [14] matrices to provide insight on attacks, and methods for their detection. This approach ensures that the dataset is comparable with the literature, and therefore that experiments performed on it are relevant. More details on the attack classifications can be found in Section IV-A.

To provide realistic data in terms of attacks, we implement the entire attack stack in the topologies, from the attacker machine to the target, and including any necessary services, *e.g.* a DNS server for a DNS amplification attack. To that end, we rely on a cyber-range infrastructure that implement services on dedicated Virtual Machines (VMs), and complete attack scenarios through a set of executable actions. Listing 1 gives an example of action available in the topologies.

```
arpspoof -i {interface} -t {target} -r {gateway} 2>
↪  /dev/null & trap 'kill $!' INT;
↪  /tmp/download_swap/run.sh '{interface}' '{file}'
↪  '{regex}'; kill $a
```

Listing 1. Example action – ARP spoofing and file swapping

The fact that attacks are executed in a running, realistic environment ensures that the collected data is representative of real-world deployments. A list of the considered attacks and their classification can be found in Section IV-A.

C. Topologies

To illustrate the heterogeneity of IT networks, we design three topologies of companies or organizations having digital activities. All three topologies are based on the same underlying network segmentation scheme:

- a demilitarized zone (DMZ) for hosting public-facing services;

- multiple local area networks (LANs) for different activities;
- a wide area network (WAN) that represent the Internet.

However, while they cover similar use cases and are comparable, the topologies are made different on a set of defined parameters. We consider the following: architecture (network size, network segmentation, and addressing scheme), services (for a usage), service version, and data-collection location.

(i) **architecture** – network architecture relates to the number of hosts, the number of local networks, or the addressing scheme. While some feature selection strategies discard identification means, like MAC and IP addresses [18], [19], others rely on the network topology to extract information [20].

(ii) **services** – different services may rely on various protocols, producing data of various types and behaving in various ways as a result. For instance, a service based on TCP will result in connection setup and a lot of back-and-forth traffic, whereas a service based on UDP will result in a more constant stream of data. Furthermore, an update of a service may result in modification of its behavior, which means that the collected data will depend on the service's version.

(iii) **probe location** – even in a specific architecture, the location where data is collected can have an impact training data. For instance, as compared to a star-shaped topology with numerous subnets, a topology with a single primary gateway that captures traffic and several services on the same network will result in a different dataset.

A fourth topology is provided to act as a novice participant, who has weak security policies and only a few devices. This topology allows testing extreme scenarios:

- *knowledge transfer* – his local traffic is not enough to build an exhaustive dataset;
- *adaptability* – while expanding his topology, new services should not raise alerts;
- *participant weighting* – as a novice, his contribution should not be given the same confidence as experts'.

The list of services comes from the considered attack list, and represent the union of all services that where either (a) targeted by an attack in the list; (b) required for an attack to succeed (*e.g.* DNS for DNS amplification); (c) required for the topology to live (*e.g.* NTP for time synchronization); or (d) deemed relevant as per the application domain (IT use case). However, following the objective of making each topology different, some services have been replaced in some topologies by alternatives providing the same feature. For instance, a Surricata Intrusion Detection System (IDS) in one topology can be replaced by Zeek in another.

Life in the topologies is provided by a set of traffic generators covering a wide range of protocols: DNS, HTTP, FTP, LDAP, Ping, and others. The traffic generators are also implemented as actions that we can deploy on any machine. Furthermore, as our topologies are *virtualized* (as opposed to *emulated*), machines run full Operating Systems (OSs) that also generate traffic, *e.g.* to fetch updates, synchronize time, perform backups. Details on the four topologies are available in Section IV-C.

IV. Dataset

This section defines the dataset and its content. As introduced in Section III, the dataset is divided in four parts, one four each participant to the federation. Each participant possess his own topology, different from the others. The following sections describe the different aspects of the datasets, namely attacks (Section IV-A), features (Section IV-B), topology architectures (Section IV-C), and scenarios (Section IV-D).

A. Attacks

The dataset currently contains 55 attacks, implemented as 55 corresponding actions in the test bed. These attacks span over height classes: bruteforce, Denial of Service (DoS), Adversarial-in-the-Middle (AitM), port scan, vulnerability scan, web, remote exploits, and complex.

comment:

This shall be updated after the mapping with ATT&CK and D3FEND is done, as the feature recommendation for each attack will influence the way we group them afterward.

1) Bruteforce:

2) Denial-of-Service:

3) Adversarial-in-the-Middle:

4) Port scan:

5) Vulnerability scan:

6) Web:

7) Remote exploits:

8) Complex:

B. Features

Features are generated with CICFlowMeter [19]. This allows the dataset to be evaluated on the same features used in the literature, which make it particularly comparable. In fact, as algorithms do not require modifications, we reduce the risk of making changes that impact performance.

CICFlowMeter produces network flows from PCAP files or live network captures. It characterizes flows with six parameters: flow ID, source IP, destination IP, source port, destination port, and protocol. In addition to each entry, CICFlowMeter also computes more than 80 analysis features that can be used afterward to detect anomalies. It covers statistical per-flow network traffic features, such as number of packets, session duration, or packet length.

C. Architecture

1) Topology-01: The first topology represents an advanced organization, mature in its security, and able to monitor and protect its own network. The topology is composed of five subnets (DMZ, users, administration, local servers, and monitoring) plus a WAN representing the Internet.

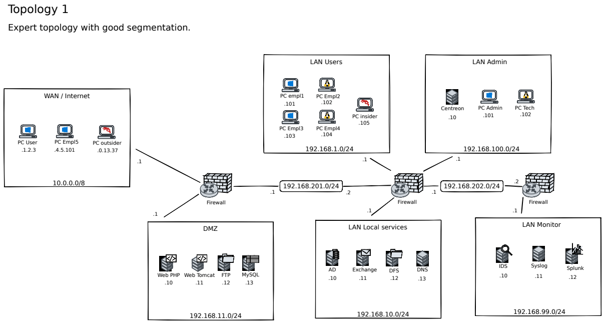

Fig. 1. Representation of Topology-01

TABLE II
Statistics about Topology-01

Property	Value
# LANs	5
# machines (hosts/server/net)	10 / 12 / 3
Monitoring	Internal (SIEM)
Attackers	2

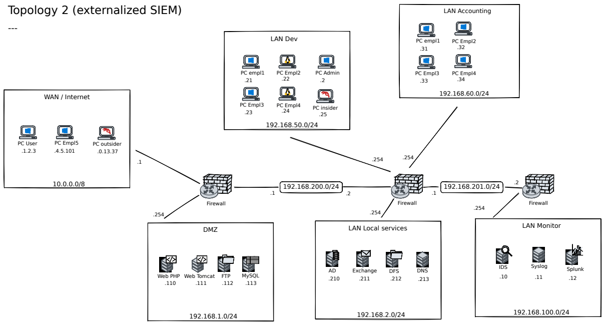

Fig. 2. Representation of Topology-02

2) Topology-02: Topologies 2 and 3 are similar to the first one, but differ by externalizing the monitoring. In addition to the IDS outgoing traffic that is generated, the second topology also differs by having two user LANs: one for devlopers and one for an accounting department.

TABLE III
Statistics about Topology-02

Property	Value
# LANs	5
# machines (hosts/server/net)	13 / 11 / 3
Monitoring	Externalized
Attackers	2

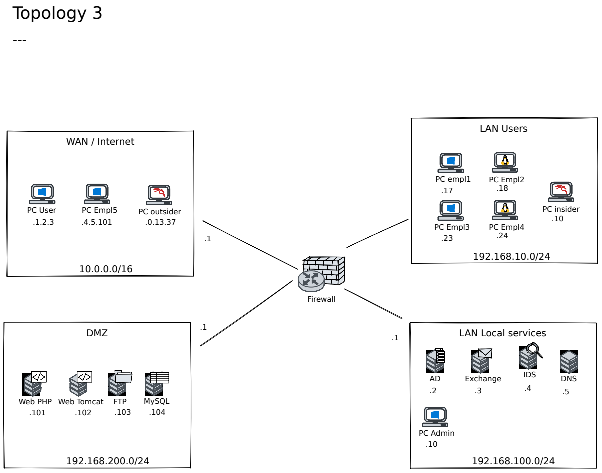

Fig. 3. Representation of Topology-03

3) Topology-03: The third topology is similar to the second one, but smaller and simpler. The only LANs are DMZ, users, and local servers. Monitoring is also externalized.

TABLE IV
Statistics about Topology-03

Property	Value
# LANs	3
# machines (hosts/server/net)	9 / 8 / 1
Monitoring	Externalized
Attackers	1

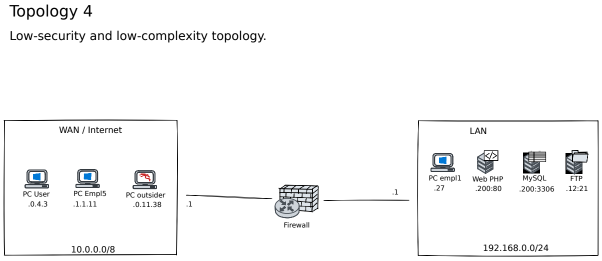

Fig. 4. Representation of Topology-04

4) Topology-04: The last topology represents a novice organization. Compared to the others, only few security features are in place. Only one LANs is present, for both users and exposed services.

TABLE V
Statistics about Topology-04

Property	Value
# LANs	1
# machines (hosts/server/net)	4 / 3 / 1
Monitoring	None
Attackers	1

D. Scenarios

V. Evaluation

Coming soon.

A. Data validity and comparison with other datasets

B. State-of-the-Art approaches on heterogeneous data

comment:

Compare different approaches on this data when compared to CICIDS2017, and possibly NSL-KDD and UNSW-NB15. Expected results would show how sota approaches falter on really heterogeneous data.

C. Show-case experiment

comment:

Make a small experiment showing the interest of real different clients:

- show the addition of a device based on pretrained collaborative model

VI. Conclusion

The quest for representative training and evaluation data is a constant challenge in the Machine Learning (ML) community. This is also relevant in the field of intrusion detection, as research tends to focus on ML for these tasks. However, known of the existing datasets have been generated for federated, heterogeneous context.

Therefore, in this paper, we proposed a novel dataset for intrusion detection in federated context. FedITN contains data from four different topologies representing four clients with various levels of cyber-maturity. The dataset contains normal and attack traffic, and can be used for a variety of research, for local and federated algorithms.

With our evaluation, we showed that **[...]**.

Future works will focus on refining the statistical representation of the dataset to better capture the heterogeneity of the real world. New attacks will be added locally, as well as attacks that can be correlated amongst clients, such as botnet campaigns or Advanced Persistent Threats (APTs). Finally, variations of the feature processing will be proposed for the same data alongside network flows, such as graph representation [20].

References

[1] L. Lavaur, M.-O. Pahl, Y. Busnel, and F. Autrel, "The evolution of federated learning-based intrusion detection and mitigation: A survey," *IEEE Transactions on Network and Service Management*, Special Issue on Network Security Management, 2022.

[2] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, H. Eichner, S. E. Rouayheb, D. Evans, J. Gardner, Z. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konečný, A. Korolova, F. Koushanfar, S. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, M. Raykova, H. Qi, D. Ramage, R. Raskar, D. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu, and S. Zhao, "Advances and open problems in federated learning," *arXiv:1912.04977 [cs, stat]*, 8, 2021. arXiv: 1912.04977. [Online]. Available: http://arxiv.org/abs/1912.04977 (visited on 04/01/2022).

[3] S. Rathore, B. Wook Kwon, and J. H. Park, "BlockSecIoTNet: Blockchain-based decentralized security architecture for IoT network," *Journal of Network and Computer Applications*, December 2018 2019, Publisher: Elsevier Ltd. [Online]. Available: https://doi.org/10.1016/j.jnca.2019.06.019.

[4] S. A. Rahman, H. Tout, C. Talhi, and A. Mourad, "Internet of things intrusion detection: Centralized, on-device, or federated learning?" *IEEE Network*, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9183799/ (visited on 06/01/2021).

[5] W. Schneble and G. Thamilarasu, "Attack detection using federated learning in medical cyber-physical systems," presented at the International Conference on Computer Communications and Networks, 2019.

[6] Y. Sun, H. Ochiai, and H. Esaki, "Intrusion detection with segmented federated learning for large-scale multiple LANs," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Glasgow, United Kingdom: IEEE, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9207094/ (visited on 10/01/2021).

[7] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, 2021, Publisher: Elsevier B.V. [Online]. Available: https://doi.org/10.1016/j.future.2020.10.007.

[8] T. D. Nguyen, P. Rieger, H. Yalame, H. Möllering, H. Fereidooni, S. Marchal, M. Miettinen, A. Mirhoseini, A.-R. Sadeghi, T. Schneider, and S. Zeitouni, "FLGUARD: Secure and private federated learning," *arXiv:2101.02281 [cs]*, 21, 2021. arXiv: 2101.02281. [Online]. Available: http://arxiv.org/abs/2101.02281 (visited on 05/18/2021).

[9] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Issue: Cisda, IEEE, 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5356528/.

[10] S. A. V. Jatti and V. J. Kishor Sontif, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," *International Journal of Recent Technology and Engineering*, 2, 2019, Publisher: IEEE.

[11] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications, 2018. [Online]. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006639801080116 (visited on 10/14/2021).

[12] Y. Sun, H. Esaki, and H. Ochiai, "Adaptive intrusion detection in the networking of large-scale LANs with segmented federated learning," *IEEE Open Journal of the Communications*

Society, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9296578/ (visited on 10/04/2021).

[13] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "MITRE ATT\&CK®: Design and philosophy," 2020. [Online]. Available: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf (visited on 07/07/2022).

[14] P. E. Kaloroumakis and M. J. Smith, "Toward a knowledge graph of cybersecurity countermeasures," 2021.

[15] N. Kiss, J.-F. Lalande, M. Leslous, and V. Viet Triem Tong, "Kharon dataset: Android malware under a microscope," 2016.

[16] V. Tarasov and A. Mudrankit, "Generating realistic datasets for deduplication analysis," 2021.

[17] R. Uetz, C. Hemminghaus, L. Hackländer, P. Schlipper, and M. Henze, "Reproducible and adaptable log data generation for sound cybersecurity experiments," in *Annual Computer Security Applications Conference*, Virtual Event USA: ACM, 6, 2021. [Online]. Available: https://dl.acm.org/doi/10.1145/3485832.3488020 (visited on 08/08/2022).

[18] G. Draper-Gil, A. H. Lashkari, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of encrypted and VPN traffic using time-related features:" in *Proceedings of the 2nd International Conference on Information Systems Security and Privacy*, Rome, Italy: SCITEPRESS - Science, 2016. [Online]. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0005740704070414 (visited on 10/15/2021).

[19] A. Habibi Lashkari, G. Draper Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features:" in *Proceedings of the 3rd International Conference on Information Systems Security and Privacy*, Porto, Portugal: SCITEPRESS - Science and Technology Publications, 2017. [Online]. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006105602530262 (visited on 10/15/2021).

[20] L. Leichtnam, E. Totel, N. Prigent, and L. Mé, "Sec2graph: Network attack detection based on novelty detection on graph structured data," in *Detection of Intrusions and Malware, and Vulnerability Assessment*, C. Maurice, L. Bilge, G. Stringhini, and N. Neves, Eds., Series Title: Lecture Notes in Computer Science, Cham: Springer International Publishing, 2020. [Online]. Available: http://link.springer.com/10.1007/978-3-030-52683-2_12 (visited on 06/10/2022).

TABLE I

Available attacks in the dataset

Attack	**Category**	**Target**	**ATT&CK Technique**	**ATT&CK Tactic**
Bruteforce FTP	Bruteforce	FTP Server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce lopin form	Bruteforce	Web Server w/ login form	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce MySQL	Bruteforce	MySQL server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce RDP	Bruteforce	Windows Host w/ RDP server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce SMB	Bruteforce	Windows Host w/ SMB server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce SSH	Bruteforce	SSH server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce Telnet	Bruteforce	Telnet server	Password Guessing (T1110.001)	Credential Access (TA0006)
Bruteforce VNC	Bruteforce	VNC server	Password Guessing (T1110.001)	Credential Access (TA0006)
DNS amplification	DoS	Any host	Reflection Amplification (T1498.002)	Impact (TA0040)
ICMP IGMP flood	DoS	Any host	Direct Network Flood (T1498.001)	Impact (TA0040)
PUSH ACK flood	Dos	Any host	Direct Network Flood (T1498.001)	Impact (TA0040)
R.U.D.Y.	DoS	Web Server w/ form	Service Exhaustion Flood (T1499.002)	Impact (TA0040)
slowloris	DoS	Web Server	Service Exhaustion Flood (T1499.002)	Impact (TA0040)
SYN flood	DoS	Any host	OS Exhaustion Flood (T1499.001)	Impact (TA0040)
TCP killer	DoS	Any host	Application or System Exploitation (T1499.004)	Impact (TA0040)
TCP RST flood	DoS	Any host	Direct Network Flood (T1498.001)	Impact (TA0040)
UDP flood	DoS	Any host	Direct Network Flood (T1498.001)	Impact (TA0040)
ZIP bomb	DoS	Any host	ARP Cache Poisoning (T1557.002) Transmitted Data Manipulation (T1565.002) OS Exhaustion Flood (T1499.001)	Credential Access (TA0006) Collection (TA0009) Impact (TA0040)
ARP Poisoning	MitM	Any host	ARP Cache Poisoning (T1557.002)	Credential Access (TA0006) Collection (TA0009)
Binary file swapping	MitM	Any host	ARP Cache Poisoning (T1557.002) Transmitted Data Manipulation (T1565.002)	Credential Access (TA0006) Collection (TA0009) Impact (TA0040)
DNS Spoofing	MitM	Any host		
JS injection	MitM	Any host	ARP Cache Poisoning (T1557.002) Transmitted Data Manipulation (T1565.002)	
SSH transparent proxy	MitM	Any host		
Firewalk	Port scan	Firewall		
Intense Scan	Port scan	Any host		
Intense Scan plus UDP	Port scan	Any host		
Intense Scan, all TCP ports	Port scan	Any host		
Ping Scan	Port scan	Any host		
Quick Scan	Port scan	Any host		
Quick Scan plus	Port scan	Any host		
DNS Zone transfer	Network attacks	DNS server		
LLMNR/NBT-NS/mDNS Poisoning	Network attacks	Any host		
RDP arbitrary code execution	Network attacks	Windows Host w/ RDP server		
SSL Vulnerability analyze	Network attacks	Any host w/ SSL		
Phishing Email	Social Engineering	Mail server		
Phishing Email w/ attach.	Social Engineering	Mail server		
SMB scan	Vulnerability scan	Windows Host w/ SMB server		
RawCopy (hive/files dump)	Microsoft Windows	Windows host w/ SMB		
EternalBlue	Microsoft Windows	Windows host w/ SMB		
Windows disable keyboard	Microsoft Windows	Windows host		
Ransomware	System attacks	Windows host (maybe others)		
SSH user enumeration	System attacks	nnix host w/ OpenSSH7.7		
Directory BF	Web vuln scan	web server		
Hearthbleed vuln scan	Web vuln scan	web server		
Shellshock vuln scan	Web vuln scan	web server		
Web vuln scan	Web vuln scan	web server		
XSS vuln scan	Web vuln scan	web server		
Dump database	SQL injection	vulnerable web server w/ DB		
Enumerate database	SQL injection	vulnerable web server w/ DB		
RFI exploitation	Web attacks	web server w/ PHP		
Shellshock	Web attacks	web server		
Tomcat backdoor uplaod	Web attacks	tomcat server		
Tomcat manager login bruteforce	Web attacks	tomcat server		
Wordpress vuln scan	Web attacks	wordpress server		