CYBER SECURITY
COOPERATIVE
RESEARCH
CENTRE

SCReeD Dataset Declaration Form

Dataset name:* DoS/DDoS-MQTT-IoT

Dataset version:* 1

Dataset URL:* https://github.com/CSCRC-SCREED/DoS_DDoS_MQTT_IoT

Creation date:* 18 May 2023

Last update:* 18 May 2023

Author(s):* Alaa Alatram, Leslie F. Sikos, Mike Johnstone, Patryk Szewczyk, James Jin Kang

Author(s) affiliation(s): Edith Cowan University

Author contact(s):* l.sikos@ecu.edu.au

Keywords:* IoT, MQTT, DoS, DDoS, Cybersecurity Dataset, Machine learning

Description/background:* Adversaries may exploit a range of vulnerabilities in Internet of Things (IoT) environments. These vulnerabilities are typically exploited to carry out attacks, such as denial-of-service (DoS) attacks, either against the IoT devices themselves, or using the devices to perform the attacks. These attacks are often successful due to the nature of the protocols used in the IoT. One popular protocol used for machine-to-machine IoT communications is the Message Queueing Telemetry Protocol (MQTT). Countermeasures for attacks against MQTT include testing defenses with existing datasets. However, there is a lack of real-world test datasets in this area. For this reason, this paper introduces a DoS/DDoS-MQTT-IoT dataset—that contains various DoS/DDoS attack scenarios using MQTT traffic—to help develop and test countermeasures against such attacks. To this end, a physical IoT testbed was constructed and a large volume of IoT data was generated that included standard MQTT traffic as well as 10 DoS scenarios. The usability of the dataset has been evaluated via machine learning.

Dataset funding: N/A

Attribute details:* No. | Message Type | QoS Level | QoS Level | Requested QoS | Epoch Time | Protocol | Source | Frame length on the wire | Time delta from previous displayed frame | Time since reference or first frame | Frame length on the wire | Stream index | iRTT | Time since first frame in this TCP stream | TCP Segment Len | Calculated window size | Syn | Reset | Acknowledgment | Clean Session Flag | Keep Alive | User Name

Length | Password Length | Retain | Clean Session Flag | Will Retain | Will Flag | Will Message Length | Will Topic Length | Topic Length | Msg Len | Info

Intended target (if specified): N/A

Format:* Microsoft Excel Spreadsheet

License:* Creative Commons Attribution 4.0 License/Open Access

Standard compliance: N/A

Type:* CSV

Size:* 41.46 GB

Availability: Public

Data status: Available

Data provenance:* SOURCES: A realistic IoT physical testbed was developed to generate a novel MQTT dataset based on actual data transmissions between publishers and subscribers through an MQTT broker. Additionally, these datasets had to contain both normal and abnormal data.
COLLECTION METHODOLOGY: The physical testbed was designed into several networks, so IP addresses need to be allocated to each network. Since the physical testbed is to be implemented. Different algorithms were written for each group of sensors. Python language was used for the algorithms to design realistic scenarios for gathering data from surrounding environments. The following presents the algorithms that were implemented for the sensors: • Algorithm 1.0 Water Level Detection for Tank; • Algorithm 2.0 Reading Voltage for Solar Power System; • Algorithm 3.0 Controlling Air Con Based on Reading the Temperature; • Algorithm 4.0 Carbon Monoxide (CO) Gas Detector; • Algorithm 5.0 Flame Detection System; • Algorithm 6.0 Smoke Gas Detection; • Algorithm 7.0 Vibration Detection System; • Algorithm 8.0 Motion Detection System; • Algorithm 9.0 Touch Detection Sensor; • Algorithm 10.0 Sound Detection System; • Algorithm 11.0 Barometric Pressure Measurement System. In addition to designing algorithms for sensors to generate normal data, Python language also will be used to implement scenarios for DoS and DDoS attacks. The abnormal data comprising five scenarios for each DoS and DDoS attack are: • CONNECT Flooding Attack (BF_DoS); • CONNECT Flooding Attack (BF_DDoS); • Delayed CONNECT Flooding Attack (Delay_DoS); • Delayed CONNECT Flooding Attack (Delay_DDoS); • Invalid Subscription Flooding Attack (Sub_DoS); • Invalid Subscription Flooding Attack (Sub_DDoS); • CONNECT Flooding with WILL Payload Attack (WILL_DoS); • CONNECT Flooding with WILL Payload Attack (WILL_DDoS); • TCP SYN Flooding Attack (SYN_DoS); • TCP SYN Flooding Attack (SYN_DDoS). A dataset was generated consisting of different DoS and DDoS attacks. Based on the techniques described above, realistic testbeds and network traffic techniques have been used to obtain the DoS/DDoS-MQTT-IoT dataset.

Source computing infrastructure:* Testbed

Accompanying program(s)/script(s): N/A

Software installer or VM for replication: N/A

Generated or captured via:* Hardware

Category/categories:* Packet capture

Published in: Computer Networks (Elsevier)

Open research question(s) (if any): N/A

Potential use case(s) or application area(s): Cyber Security

Data access control:* Global access

Data retention period:* N/A

Data validation/checksum:* 7C480670939258C96CAAA92C6A03A2B3 (MD5)

GDPR compliance:* Yes

Consent: Yes

Ethics approval: N/A

Ethics considerations: N/A

** denotes mandatory field*

☒ I confirm that I have read, understood, and agreed to the submission guidelines, policies, and submission declaration.

☒ I confirm that the contributors of the dataset have no conflict of interest to declare.

☒ I agree to take public responsibility for my dataset's contents.

Signature of Corresponding Author Date: 09 May 2024

(signed on behalf of all contributors)