The BoT-IoT dataset Source Files

The raw network pacekts (Pcap files) of the BoT-IoT dataset were created by application of the tshark tool, in the Cyber Range Lab of the Australina Center for Cyber Security (ACCS), and incorporates a combination of normal and abnormal traffic. Simulated network traffic was generated through Ostinato tool and Node-red (for non-IoT and IoT respectively).The dataset's source files are provided in different formats, such as the original pcap files, the generated argus files and finally in csv format. The files were separated, based on attack category and subcategory, to better assist in labeling process.

Table 1 Directory structure of Bot-IoT dataset

ARGUS: ○ DDoS ○ DDoS_HTTP ○ DDoS_TCP ○ DDoS_UDP ○ DoS ○ DoS_HTTP ○ DoS_TCP ○ DoS_UDP ○ Scan ○ OS ✦ 1 ✦ 2 ✦ 3 ✦ 4 ○ Service ○ Theft ○ Data_Exfiltration ○ Keylogging
Dataset: ○ 5% ○ 10-best features ✦ 10-best Training-Testing split ○ All features ○ Entire Dataset ○ Features Explanation
Ground Truth

PCAPs:

- DDoS ○ DDoS_HTTP ○ DDoS_TCP ○ DDoS_UDP
- DoS ○ DoS_HTTP ○ DoS_TCP ○ DoS_UDP
- Scan

- OS
  - 1
  - 2
  - 3
  - 4
- Service

- Theft
  - Data_Exfiltration
  - Keylogging

Those who wish to make use of Bot-IoT dataset have to cite the following paper that elaborates on its creation.

1. *Paper to be added, at a later date.*

For more information about the dataset, please contact the authors:

1. Nickolaos Koroniotis: e-mail (n.koroniotis@student.adfa.edu.au)
2. Nour Moustafa: e-mail (nour.moustafa@unsw.edu.au)