Report

First, install: sklearn, matplotlib, and scapy.

Now download the dataset from:

https://mettl-miscellaneous-public.s3.ap-south-1.amazonaws.com/client_public_data/369004/botnet/Botnet_Detection_Dataset.7z

Extract the dataset in the directory that contains main.py

Now you see Botnet_Detection_Dataset folder in the directory.

Botnet_Detection_Dataset/Benign contains pcap files of normal flows.
Botnet_Detection_Dataset/Botnet contains pcap files of botnets.

To extract all flows, simply run:

```
python3 main.py generate_flows Botnet_Detection_Dataset
```

or replace Botnet_Detection_Dataset with the desired dataset path.
The flows will be stored inside all_flows.pkl

To filter flows, run:

```
python3 main.py filter_flows
```

This way, the flows related to (local IP, remote IP) pairs with less than 4 flows are dropped. Also TCP flows with incompleted handshakes are dropped. Flows with more that 100 packets are dropped too.

Command output:

```
number TCP flows:  142869
number of incompleted TCP handshaked flows (removed):  74984
result of filtering the traffic of (local IP, remote IP) pairs which have less than four flows:
removed 334503 flows
filtered 1799 flows which had more than 100 packets
saved flows into filtered_flows.pkl
```

To extract two level flow dependencies:

```
python3 main.py 2_lvl_dep
```

The result will be stored in:
T_dep=1-N_dep=5-S_dep_th=0.5-two_level_flow_dependencies.pkl

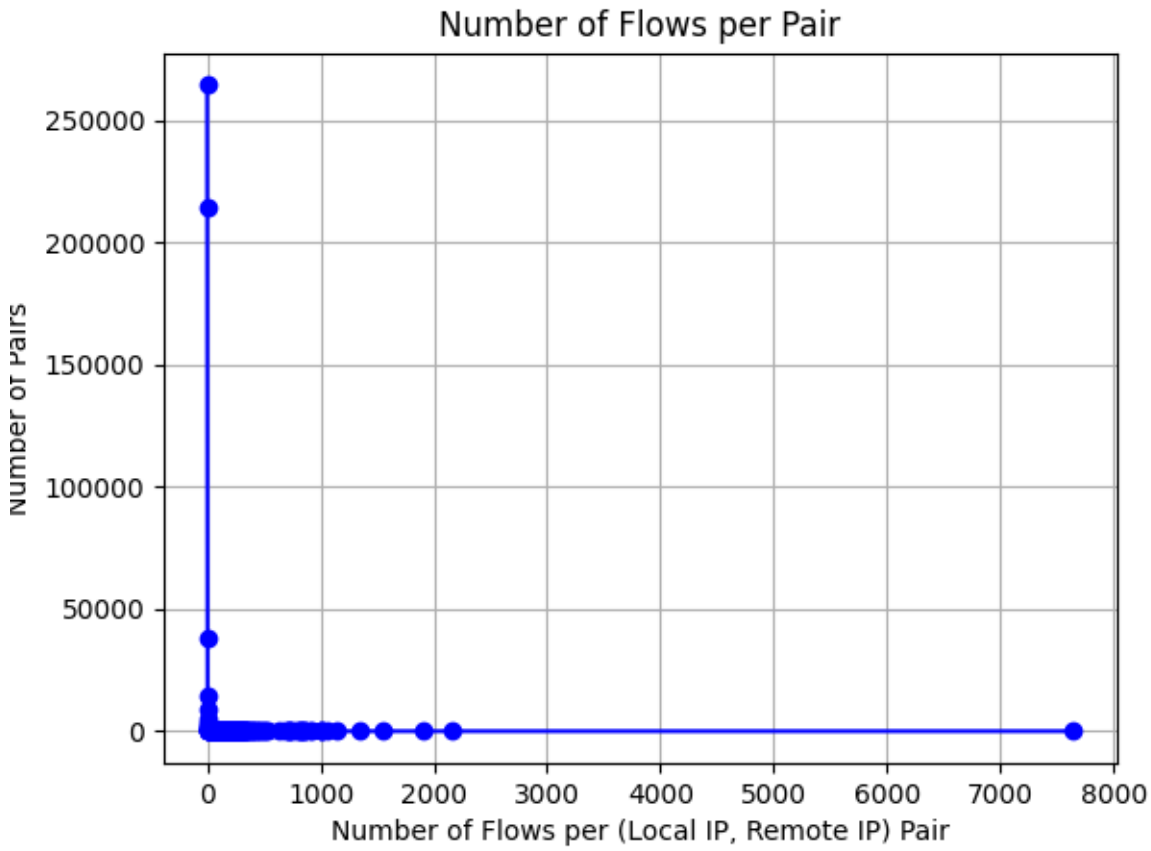
To extract three level flow dependencies:
python3 main.py k_lvl_dep 3 T_dep=1-N_dep=5-S_dep_th=0.5-two_level_flow_dependencies.pkl

The result will be stored inside:

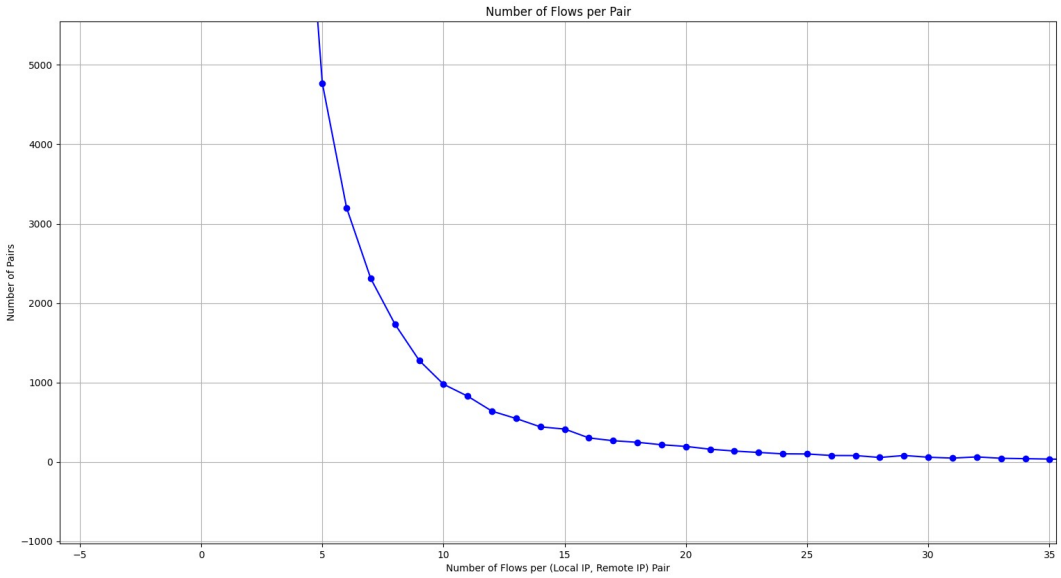
3-level-flow-dependencies.pkl

To perform clustering:
python3 main.py compute_dist 3 3-level-flow-dependencies.pkl

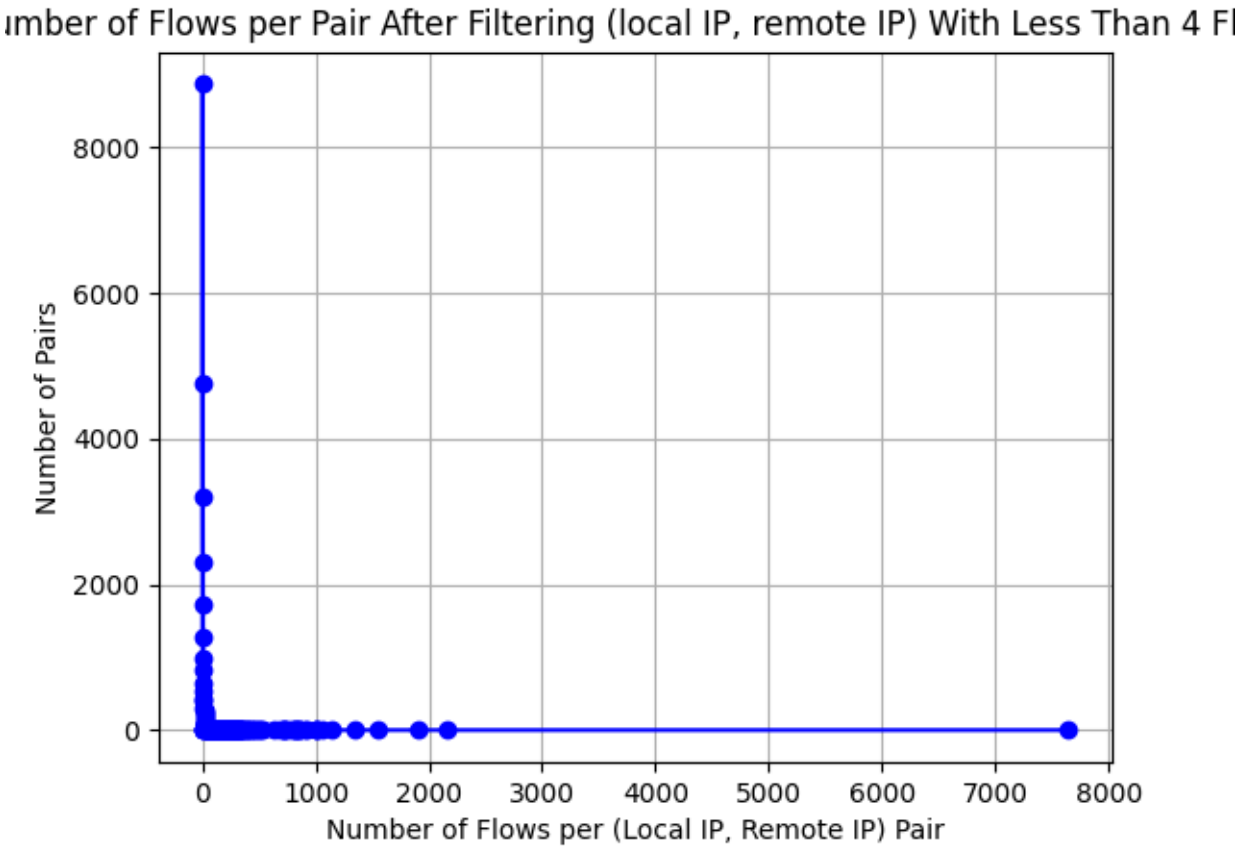
Results:



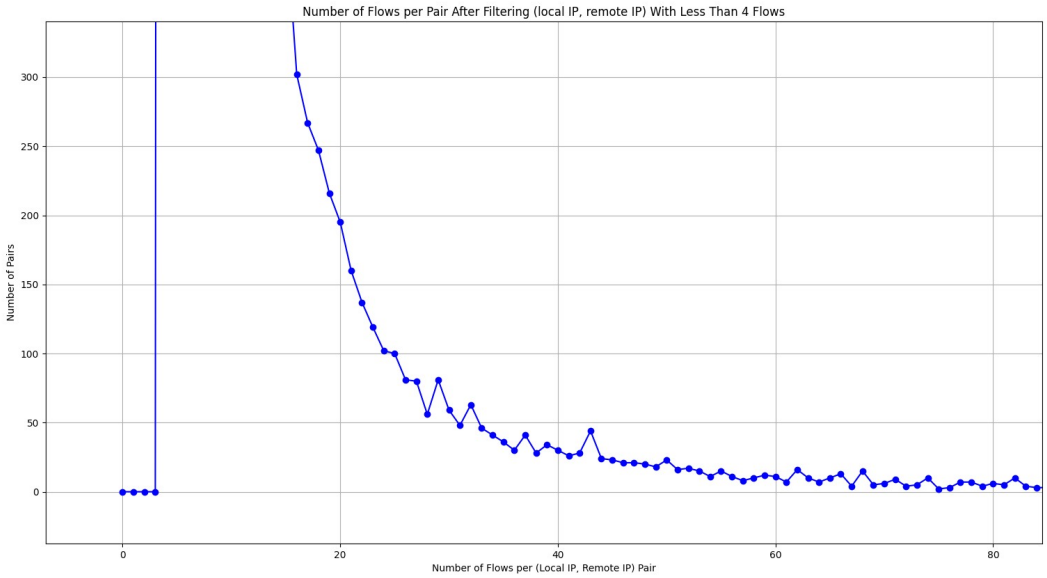
zoomed:



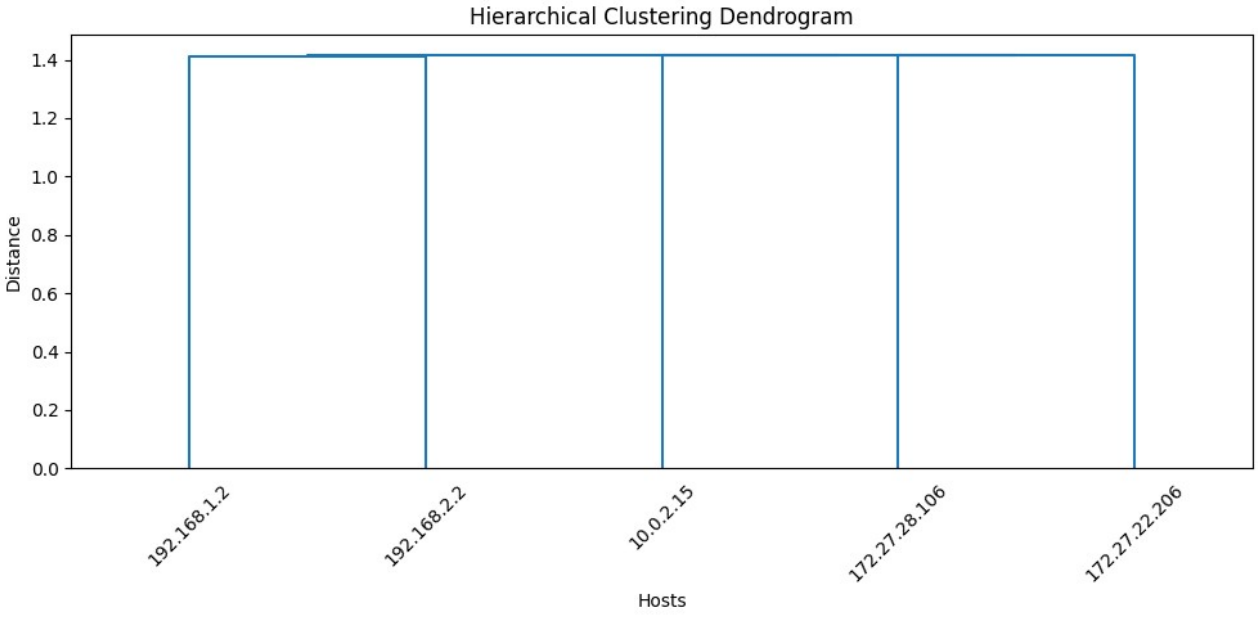
after filtering the flows:



zoomed:



clustering:



p2pbox_1 192.168.1.2
p2pbox_2 192.168.2.2
torrent 172.27.28.106
storm:
66.154.80.101
66.154.80.105
66.154.80.111
66.154.80.125
66.154.83.107
66.154.83.113
66.154.83.138
66.154.83.80
66.154.87.39

66.154.87.41
66.154.87.57
66.154.87.58
66.154.87.61
zeus:
10.0.2.15
vinchuka:
172.27.22.206