This material is based upon work supported by the Defense
Advanced Research Projects Agency (DARPA) under Contract No.
HR001119C0102. The opinions, findings, and conclusions stated
herein are those of the authors and do not necessarily reflect
those of DARPA.

Comparing telescope subnets

Introduction

One of the prevailing assumptions about IBR is that it is approximately the same everywhere, at least for dark IPv4 space. Scanning through the entire IPv4 address space is a fast and easy-to-automate operation, so there is relatively little benefit to scanners limiting their scans to subnets they expect to respond to their probes. Therefore the characteristics of subnet scans should be similar, at least in basic terms across different subnets. There may be some differences because some scanners divide up the work of scanning the Internet among several scanning hosts, and those hosts are not always running exactly the same software – but these differences, although interesting, should not change the basic statistics.

(Of course, any scanners who *do* prune dark space out of their scans, will not,

by definition, appear in the telescope data. . .)

This directory contains two scripts that show how to use the pre-processed data from the pcap files (created by the tools in `../pcap-ingestion` and `../meanies`) to look for gross differences in scanner behavior for different subnets. These are simple scripts, and their primary utility is to show to use the data and the other tools, such as `firecracker`, to do this kind of analysis.

Running the scripts

`compare-subnets.sh`

The `compare-subnets.sh` creates plots showing the differences between a given list of /24 subnets in the telescope for a given span of time.

Like most of the other scripts in this package, `compare-subnets.sh` takes its parameters from a parameter file, environment variables, or values on its commandline. See `../../params/example.params` for a description of the parameters. The required parameters are `DATANAME`, `FCSVDIR`, `SUBNETS`, and `DATAEXPR`. The meaning of `DATANAME` and `FCSVDIR` are the same as for other scripts, but `SUBNETS` and `DATEEXPR` are unique to the scripts in this directory.

`SUBNETS` is a list of the prefixes of the /24 subnets to compare. For example, `SUBNETS` could be `10.2.3.0 10.2.4.0` to compare subnets 10.2.3.0/24 and 10.2.4.0/24.

`DATEEXPR` is a shell-style regular expression that matches the names of the hours that you want to compare. For example, if `DATEEXPR` is "2023-12-10" then all of the hours from December 10, 2023 are used.

The script creates three plots, named `plot0.pdf`, `plot1.pdf`, and `plot2.pdf`, in the current directory. These plots show the eight most common, second eight most common, and third eight most common protocol/destination port combinations, by subnet, for the time period specified. If the plots look the same for each subnet, then the subnets are similar, but otherwise there may be interesting differences.

`compare-meanies.sh`

The `compare-meanies.sh` script is similar to the `compare-subnets.sh` script, but does not create plots (because the Meanie behavior makes comparing different protocol/port combinations almost meaningless).

Like `compare-subnets.sh`, `compare-meanies.sh` requires the `DATANAME`, `SUBNETS` and `DATEEXPR` parameters. It also requires the `MEANIEDIR` parameter, which should be the same as the parameter used by `../meanies/find-meanies.sh` to pre-process the Meanie data.

The output of `compare-meanies.sh` is the total count of all the Meanie packets from all of the given subnets, and the total for each individual subnet.