Lab Instructions

Install Python and the dependent libraries

This lab is based on Python, which you will have to download and install on your lab environment – a laptop is sufficient for the exercises.

Python can be downloaded from here:
https://www.python.org/downloads/

We will also need to install a number of Python libraries to perform some of the lab exercises, so the following install commands will need to be run from the command line:

```
pip install scikit-learn
pip install splunk-sdk
pip install pandas
pip install numpy
pip install networkx
pip install matplotlib
```

Install Jupyter Notebooks

The labs themselves are all Python notebooks that should be run in Jupyter Notebooks. You will need to install Jupyter Notebooks from here:
https://jupyter.org/install

Once installed you can run the notebook environment running the following from the command line:

```
jupyter notebook
```

This will open Jupyter Notebooks in your web browser, and you can browse to the lab notebooks within the notebook environment in the browser.

Install Splunk

Next up we are going to install and run Splunk in the environment. Splunk can be downloaded from here:
https://www.splunk.com/en_us/download.html
From the download page there will be links to instructions for installing and running Splunk. Essentially the downloaded tar.gz file will need to be unzipped into your applications or opt directory and once installed you will need to run the following form the command line to start Splunk:

```
/splunk/bin/splunk start
```

You will be prompted to set a username and password the first time you start Splunk as well.

Install Splunk apps

Once Splunk is installed you will need to install the following apps and their dependencies into your Splunk instance:
https://splunkbase.splunk.com/app/3559/
https://github.com/splunk/botsv2
Each of the apps contain instructions on their dependencies and how to install them.

Download CIDDS data

The CIDDS data can be found here:
https://www.hs-coburg.de/forschung/forschungsprojekte-oeffentlich/informationstechnologie/cidds-coburg-intrusion-detection-data-sets.html
We are going to use the netflow logs from the CIDDS-001 dataset, which you will need to download and unzip.

Import CIDDS data into Splunk

Once unzipped there are four source data files that we are going to ingest from the CIDDS dataset:

CIDDS-001/traffic/OpenStack/CIDDS-001-internal-week1.csv
CIDDS-001/traffic/OpenStack/CIDDS-001-internal-week2.csv
CIDDS-001/traffic/OpenStack/CIDDS-001-internal-week3.csv
CIDDS-001/traffic/OpenStack/CIDDS-001-internal-week4.csv

To import this data into Splunk you. Need to set up a file monitor to read data from the CIDDS-001/traffic/OpenStack directory and set the sourcetype to be csv and ingest the data into an index called cidds.

More instruction on how to do this can be found here:
https://docs.splunk.com/Documentation/Splunk/8.1.3/Data/MonitorfilesanddirectorieswithSplunkWeb