**Introduction:**

Our proposed semester project is, **“To develop a prolog program that can utilize Netflow output to calculate the risk in IT Infrastructure”**. Our prolog program will be able to import data from Netflow output, and together other network facts and policy rules provided in our prolog program, calculate the risk to an IT Infrastructure’s network.

**What is a Netflow?**

Netflow was originally developed by Cisco, and provides a means to observe traffic on a network in detail. What makes Netflow useful is that it groups network traffic into flows. Each flow is determined by default based on the following seven values:

1. Ingress interface
2. Source IP address
3. Destination IP address
4. IP Protocol,
5. Source port for UDP or TCP
6. Destination port for UDP or TCP
7. IP type of service

**How Netflow helps in traffic monitoring?**

* This means that communications between hosts on the network can be observed, and Netflow is intelligent enough to stop a flow that is using TCP when connection is closed by one of the hosts.
* Netflow provides network administrators and cyber security personnel the ability to see where traffic is going on a computer network.
* This allows for the identification of hosts that are utilizing more network resources than is permitted by policy, allows for detection of network misconfigurations (i.e. hosts that should not be able to communicate are communicating over the network), allows detection of malicious external hosts that are communicating with hosts within the organization’s network and the detection of DDoS attacks on the organization’s network.

**Motivation:**

For our project, we plan, “**To leverage this data about traffic flow on the network that Netflow provides to produce a prolog security tool to help with identifying some of the cyber risks of a given network”**.

As Netflow only provides data on traffic movement on a network, our security tool will only be able calculate risk that can be determined from measures related to the information obtained from Netflow’s output. Since Netflow is developed and supported by Cisco, and many organizations rely on Cisco’s routers and switches to route and regulate network traffic on the organization’s network, our security tool would help said organizations to better understand the risk to their network and provide clearer information to the organization as the organization determines where to invest to reduce risk to the organization’s network.

**Objective: (What we want to achieve)**

1. Read in the netflow output from file to Prolog as facts
2. Write prolog program to analyze data from netflow
3. Develop measures to determine the risk in the network utilizing Netflow data
4. Program measures into Prolog
5. Prolog program will output network risk based on the measures we specify in the program using input from Netflow

**Deliverables: (We want to submit)**

* Project Proposal Deadline April 6
  + A presentation about the overview of our project
* Project Progress Report: April 20
  + Completing objectives 1 and 2
  + Initiating objective 3
* Project Pre-Demo Report: May 1
  + Completing all the objectives for a small data set
* Project Final Demos and Documentation: May 11-12 (6-9pm).
  + Complete all the objectives for a medium-large scale data set
  + Documentation about running the project
  + Be ready with the code for demonstrating in the class

**List of tools:**

1. Strawberry prolog/SWI prolog
2. Microsoft Office for generating documentation
3. Netflow traffic generator if we are unable to obtain sample Netflow output from other means and we have the following links to initiate our data gathering process once the professor gives us a feedback
   1. [**http://stackoverflow.com/questions/32511699/netflow-sample-data-sets**](http://stackoverflow.com/questions/32511699/netflow-sample-data-sets)
   2. [**https://github.com/ntop/ntopng**](https://github.com/ntop/ntopng)
   3. [**https://www.flowtraq.com/product/flow-exporter/**](https://www.flowtraq.com/product/flow-exporter/)
   4. [**https://www.plixer.com/products/flowalyzer/**](https://www.plixer.com/products/flowalyzer/)

**Methodologies: (To complete the project)**

1. Investigate data available in output from Netflow to determinewhat KRI, KPX, and KPI’s that can be determined from the Netflow data.
2. Analyze and categorize the Netflow data into prolog facts and rules.
3. Determine how to code Prolog to read the data in the Netflow output into the Prolog program as the facts and rules we produced in #2.
4. Determine the functional logic that our Prolog program will use to solve for the risk in the network systems we will be analyzing based on the KRI, KPX, and KPI’s determined in #1 and the facts and rules from the Netflow data determined in #2.
5. Code the logic for our program into Prolog, and test our program with a small data set to correct any bugs or errors we observe.
6. Scale up our test dataset to for our final project submission, resolving in bugs or errors in our program that we observe.
7. Produce documentation explaining how to execute our Prolog program, the format that the Netflow data needs to be in and any additional information or instructions useful for running or interpreting our program.

**References:**

1. <http://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_white_paper0900aecd80406232.html>
2. <http://www.cisco.com/c/dam/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_case_study0900aecd80311fc2.pdf>
3. <https://en.wikipedia.org/wiki/NetFlow>