Activity Plot

Diana Paterson
Presentation Summary:

- This presentation will introduce a new visualization for Netflow data called Activity Plot.
- Activity Plot displays host activity as a function of time using a small selection of colors to capture a limited number of activity categories.
- This presentation will:
  - Provide background for the implementation, and review related work
  - Detail the raw data
  - Review the structure of the activity plot
  - Describe a prototype implementation
  - Present two case studies
  - Offer future work and conclusions
Background: Existence Plots

- Introduced by Phil Groce and Jeff Janies at FloCon 2008
Activity Plot – Activity Viewer

- A visualization for displaying host activity as a function of time.
- Activity can be host related, time related, simple, complex etc.
- Individual hosts are plotted against time in a simple two dimensional grid.
Raw Data

- NetFlow data collected using the SiLK tool’s flow collector.
- Payload data was not collected or examined.
- Two UNIX scripts used to process the data and generate the input information for the visualization.
  - These scripts use set and bag tools to generate the data required to determine the activity classifications.
Sample Network Description

- Four /24s address spaces.
- Hosts consist of a combination of standard traditional users and temporary development/experimental machines.
- Addresses have been anonymized to protect the identity of both the network and the external hosts.
- The visualization presented contains 1 month (the first month captured) of network activity.
Visualization: Activity Plot
Visualization structure:

- Grid format used to display activity and non-activity.
- 14 days of data visible at one time at an hourly resolution.
- Days of the week used in the time labels.
- Small number of activities visualized.
- Background highlighting used to group hours in a day.
Implementation: Activity Viewer
Case Study 1: Lease Switching?

- Host 192.168.22.68 and 192.168.22.66
- Significant changes in the time based activity patterns of both hosts.
- Further investigation with the SiLK tools supports the conjecture.
Case Study 2: Malicious Activity

- February 28, 2007 at 17:00 host 192.168.22.82 started using the same port as a client and a server.
Case Study 2: Malicious Activity

- Significant changes in the time based activity patterns with respect to its prior activity.
- Further examination with the tool revealed that the host had activity on 9000 ports during that first hour.
- Most of the activity was client based directed at another internal server.
- Further analysis with the SiLK tools showed that ~9000 ports were involved in a typical scanning activity of another internal host.
Conclusions

- The prototype was able to visualize interesting networking features.
- It has been decided to include the visualization in the FloVis tool set.
- Visualization is generic and can represent other types of categorical information (e.g., not just servers/clients)
Future work 1

- Investigate scalability options (e.g., aggregation)
- Addition of basic filtering on time, IP, port and/or activity.
- Negative and positive filtering.
- Implementation of different activities to be plotted e.g. volume, scan response, etc.
- Further encoding of other features within categories.

🌟
Future Work 2

- More annotation.
- Bird’s eye view.
- User study to:
  - obtain feedback,
  - gain insight into the things the user likes, dislikes, would like to see added, feels that is missing.