NVisionIP: An Animated State Analysis Tool for Visualizing NetFlows

Ratna Bearavolu Kiran Lakkaraju William Yurcik

National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign
{ratna,kiran,byurcik}@ncsa.uiuc.edu

Abstract—**In this paper, we describe a NetFlow visualization tool, NVisionIP, which provides network administrators increased situational awareness of the state of networked devices within an IP address space. It does this by providing three increasingly detailed views of the state of devices in an entire IP address space to subnets to individual machines. Operators may use NVisionIP to transparently view NetFlow traffic without filtering or may selectively filter and interactively query NVisionIP for unique views given experience or relevant clues.**

Index Terms— **NetFlows, Visualization, Network Security**

I. Introduction

What is the state of devices on your large and complex network? This is a question management commonly poses to network administrators and up to now the answer has been problematic. IDS sensors give binary alarms for signature-matches or anomalous traffic, if no alarms then there is no state information about the devices on the network. Scans test for software vulnerabilities but this is more about predicting posture to future attacks than knowledge of current state. Network device monitoring devices like MRTG[1] and the Flowscan[2] may display traffic levels by service as well as aggregate traffic load levels – while this is certainly useful for managing traffic congestion and detecting high volume events, there are no details about device state and small events are obscured.

While NetFlows provide an excellent source of information concerning the behavior of the network, the sheer magnitude of NetFlow logs often makes it difficult to gain an understanding of that behavior. In this paper we present a tool, NVisionIP [1,3-5,9-11], that uses NetFlows to visually represent activity on an entire IP address space. NVisionIP presents information at three different levels allowing operators to select which level to use.

II. System Architecture

The NVisionIP system architecture is comprised of three modules: Data Retrieval Module, Computation Module and Visualization Module. As shown in Figure 1, the three modules interact using a Mediator object [2]. By using a Mediator object, we avoid direct referencing of a module by other modules, thus providing the flexibility of modifying the modules independently. The Data Retrieval Module reads in the NetFlow files, preprocesses them, and places them in a table structure for the Computation module to use. For every IP address in the input table, the Computation Module calculates various statistics as shown in Figure 2. These statistics are then passed to the Visualization Module that presents information to a user.

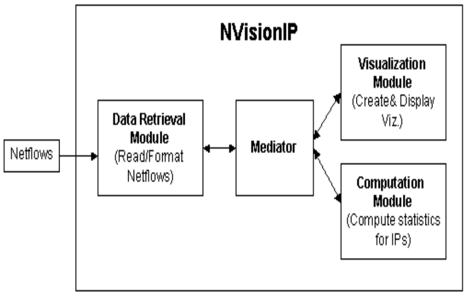

NVisionIP
Netflows
Data Retrieval Module (Read/Format Netflows)
Mediator
Visualization Module (Create& Display Viz.)
Computation Module (Compute statistics for IPs)

Figure 1. NVisionIP System Architecture

- *Number of times IP address part of a flow*
- *Number of times IP address destination of a flow*
- *Number of times IP address source of a flow*
- *Number of ports used by IP address*
- *Number of destination ports used by IP address*
- *Number of source ports used by IP address*
- *Number of protocols used by IP address*
- *Number of bytes transmitted to/from IP address*
- *Number of bytes transmitted to/from IP address*

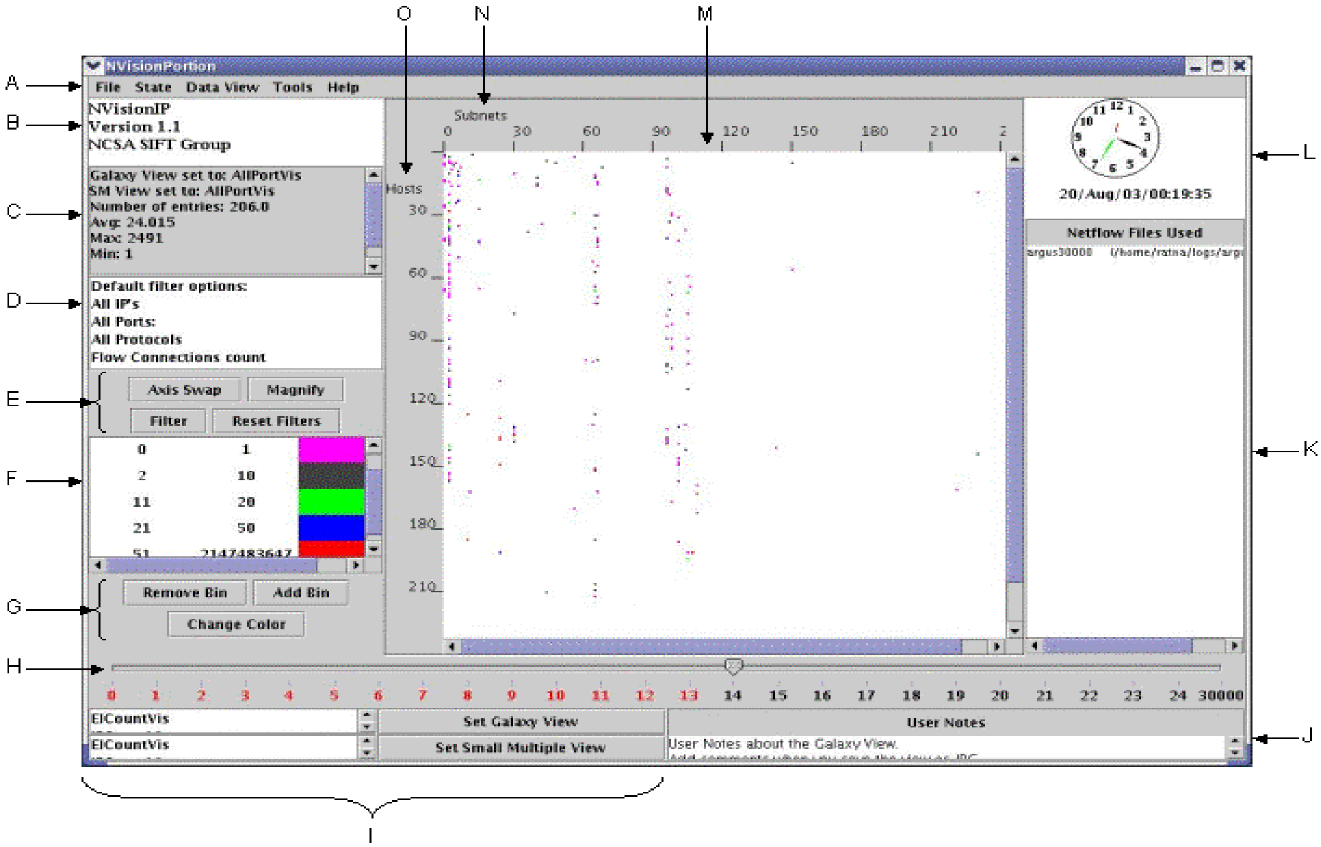
Figure 2. Statistics Derived from NetFlows

III. How to use NVisionIP

NVisionIP can be downloaded here:
<http://security.ncsa.uiuc.edu/distribution/NVisionIPDownLoad.html>

NVisionIP builds on the concept of “overview, browse, drill-down to details-on-demand” championed by Shneiderman [6] and Tufte [8] to support three different views: (1) a Galaxy View (GV) - a high-level view of an entire network, (2) a Small Multiple View (SMV) - a subnet view of traffic from multiple machines within the network, and (3) a Machine View (MV) information about flows into/out of a single machine. Overview plus detail breaks up content into comprehensible pieces while also allowing for simultaneous comparisons of different views which may reveal interrelationships [7].

[1] Multi Router Traffic Grapher <http://mrtg.hdl.com/mrtg.html>
[2] <http://net.doit.wisc.edu/~plonka/FlowScan/>



NVisionIP
Version 1.1
NCSA SIFT Group

Galaxy View set to: AllPortVis
SM View set to: AllPortVis
Number of entries: 206.0
Avg: 24.015
Max: 2491
Min: 1

Default filter options:
All IP's
All Ports:
All Protocols
Flow Connections count

0	1	🟪
2	10	⬛
11	20	🟩
21	50	🟦
51	2147483647	🟥

20/Aug/03/00:19:35

Netflow Files Used
argus30000 /home/ratna/logs/arg

ElCountVis
ElCountVis

User Notes about the Galaxy View.

A. Menu Bar
B. NVisionIP version number and other information
C. Summary of statistics being displayed on the grid
D. Current filter options being applied to the grid – default are displayed in the picture
E. Axis Swap – To swap the subnet and hosts axis; Magnify – Apply a magnifier around the cursor area to enhance the grid; Filter – Apply user specified filter to the data; Reset Filter – Remove any filters applied to the data.
F. Legend mapping a range of numbers to a color
G. Functionalties to customize the legend (F)
H. Slider showing different intervals of data, 30000 indicates the total number of netflows the user selected to process. In the image, the user moved the slider bar to interval 14. This implies that (1200*14) netflows have been processed and displayed

I. Options for the user to select which attribute of the machines the Galaxy view and Small Multiple view should display. In the image, it is set to show number of unique ports used by the machines (See Panel C)
J. Notes area for the user to make comments about the GV. The notes get saved, when the user saves the GV as an image.
K. List of netflow files that are loaded into the system
L. Shows the time stamp of the last netflow processed among the selected intervals.
M. XY grid that shows a class B network
N. Subnets axis
O. Hosts axis

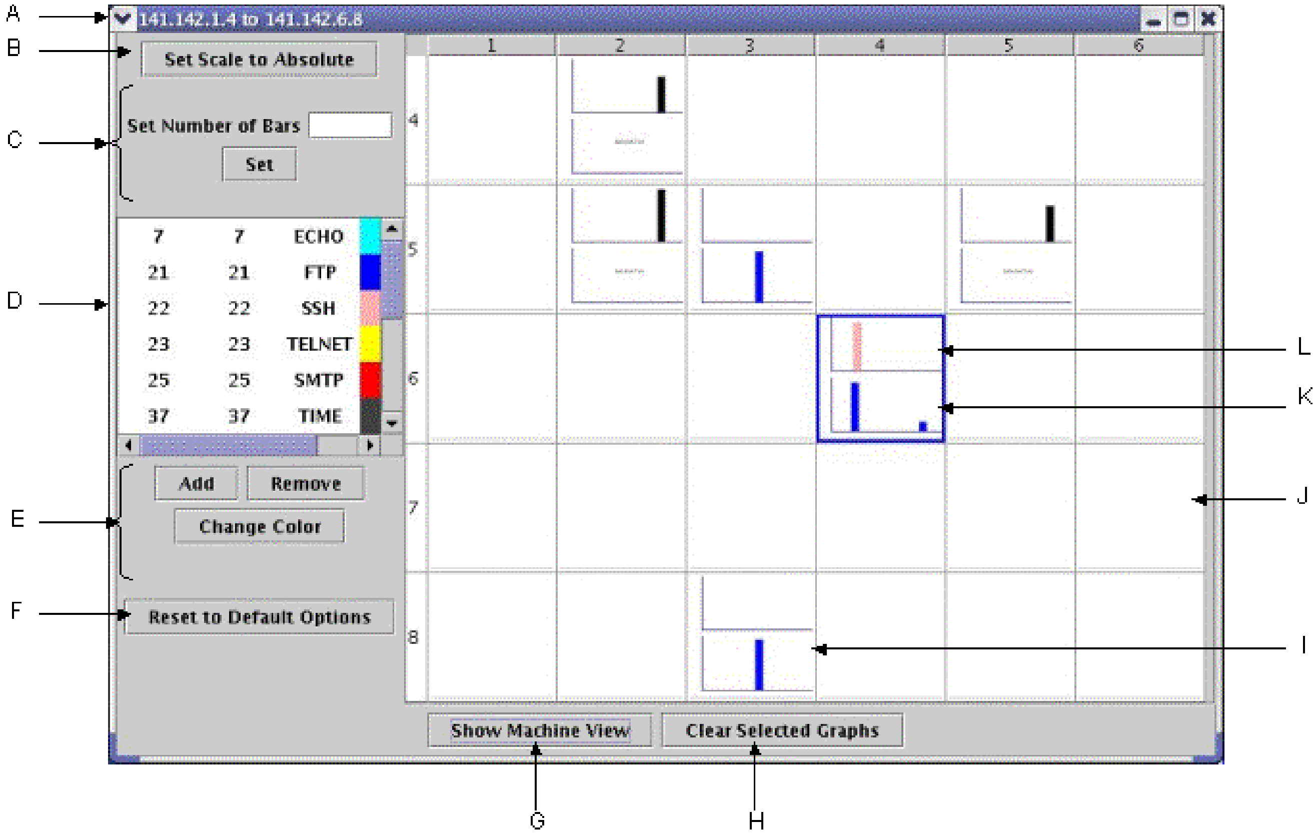
Figure 3. NVisionIP Galaxy View (GV)

Figure 3 shows the GV, a 2D graph of a Class B network where the hosts are on the X-axis and subnets are on the Y-axis—this orientation can be changed by the *Swap Axis Button* (Figure 3:Label E). A single point on the graph represents an IP address on the network being monitored. For example, point (50, 70) on the graph represents the IP address 141.142.50.70. An IP address displays information about any of the statistics from Figure 2 using a color code. The mapping from number to color is provided by the customizable binning legend (Figure 3:Label F). The default statistic configured in the Galaxy view is the number of unique ports used by the host. The *Set Galaxy View Button* (Figure 3: Label I) can change this view.

To process input data, the user selects NetFlow files to be visualized using the menu file pull down. NVisionIP divides input files into intervals of equal numbers of flows (user selects number of intervals). The last number on the slider bar reflects the total number of NetFlows loaded.

When the user moves the bar across the data intervals, NVisionIP provides the option of either viewing the results as a summation (cumulative view) or piece-wise (animated view). For example, if the user chooses a cumulative view and moves the slider (Figure 3:Label H) from interval 0 to 4, then the results that are displayed for all the IPs are the summation of the values (Ports, Protocols, Count, etc) of NetFlows from interval 0 to 4. If the user had selected the animation view and moved the slider from interval 0 to 4, the GV would show a 4 frame animation—each frame representing activity during one time interval. The animation shows how IP device state changes over time, providing a temporal feel for device state on the network.

To learn more about the GV filtering option see [3,5]. GV magnification and storage options are described within the application itself.



7	7	ECHO	🟦
21	21	FTP	🟦
22	22	SSH	🟥
23	23	TELNET	🟨
25	25	SMTP	🟥
37	37	TIME	⬛

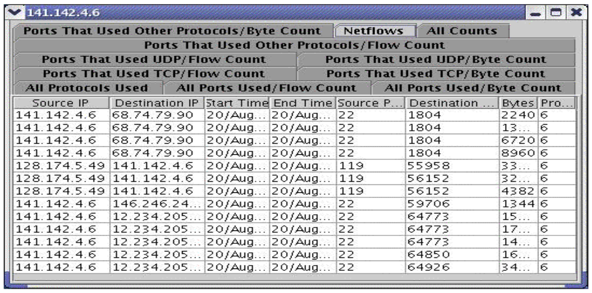
A. Range of IPs being displayed
B. Sets the scale of the bar charts axis – to absolute or relative (default)
C. Sets the number of bars to displayed for every machine
D. Legend mapping port/Protocol number to a color
E. Functionalities to customize the legend
F. Resets the options to default (Scale, Number of bars, Port/Protocol mapping to default color)

G. Shows the machine view of a selected machine
H. Clears any selected machine (graph)
I. Machine 141.142.3.8 in SMV
J. Machine 141.142.6.7 – An empty panel, as no information is contained by NVisionIP
K. Lower graph – showing other ports (all ports – special ports (from the legend)) used by 141.142.4.6
L. Upper graph – showing the special ports used by 141.142.4.6

Figure 4. NVisionIP Small Multiple View (SMV)

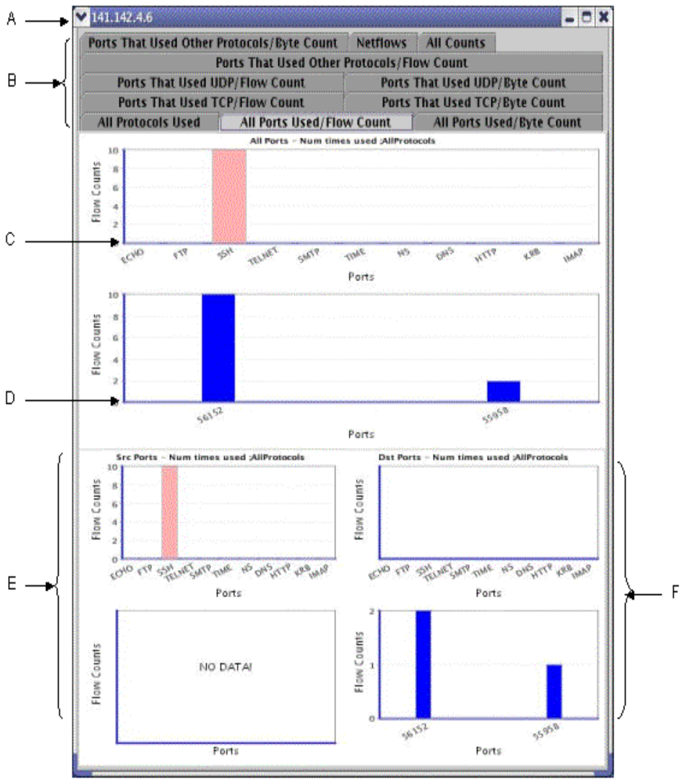
The SMV provides information about adjacent devices in an IP address space. The primary purpose of the SMV is facilitating quick browsing of subnets within an address space for information about the ports and protocols used by each IP device. The user can scan and compare activity across the subset of machines selected using a mouse to highlight the region of interest. Each square in the SMV grid (Figure 4:Label I) represents a device with an IP address. Each square is divided into two histograms: (1) the top histogram represents traffic from well-known ports, and (2) the bottom histogram represents traffic on active ports above port 1024 ordered from most to least active. At a glance, a user sees and compares port activity of different devices. If a machine uses an unusual port, this will be immediately visible. Similar to GV, the user can define the colors associated with the particular ports/protocols. Also, the user can define what ports/protocols are considered "of special interest" using the interface in (Figure 4:Label E).

The MV is the most detailed view simultaneously displaying all statistics from a single machine. Figure 6 shows the eleven tabs that a user may select to view different information from a particular machine. The eleven tabs seen in the MV hold information on the statistics from Figure 2 plus the raw NetFlows source data used to generate the information for the machine being examined (Figure 5). Each tab consists of 6 sets of histograms as shown in Figure 6: lower left is source activity leaving the IP device, lower right is destination activity entering the IP device, and the upper half is an aggregate of traffic activity both entering and leaving each IP device.


141.142.4.6
Ports That Used Other Protocols/Byte Count
Netflows
All Counts
Ports That Used Other Protocols/Flow Count
Ports That Used UDP/Flow Count
Ports That Used UDP/Byte Count
Ports That Used TCP/Flow Count
Ports That Used TCP/Byte Count
All Protocols Used
All Ports Used/Flow Count
All Ports Used/Byte Count

Source IP	Destination IP	Start Time	End Time	Source P...	Destination ...	Bytes	Pro...
141.142.4.6	68.74.79.90	20/Aug...	20/Aug...	22	1804	2240	6
141.142.4.6	68.74.79.90	20/Aug...	20/Aug...	22	1804	13...	6
141.142.4.6	68.74.79.90	20/Aug...	20/Aug...	22	1804	6720	6
141.142.4.6	68.74.79.90	20/Aug...	20/Aug...	22	1804	8960	6
128.174.5.49	141.142.4.6	20/Aug...	20/Aug...	119	55958	33...	6
128.174.5.49	141.142.4.6	20/Aug...	20/Aug...	119	56152	32...	6
128.174.5.49	141.142.4.6	20/Aug...	20/Aug...	119	56152	4382	6
141.142.4.6	146.246.24...	20/Aug...	20/Aug...	22	59706	1344	6
141.142.4.6	12.234.205...	20/Aug...	20/Aug...	22	64773	15...	6
141.142.4.6	12.234.205...	20/Aug...	20/Aug...	22	64773	17...	6
141.142.4.6	12.234.205...	20/Aug...	20/Aug...	22	64773	14...	6
141.142.4.6	12.234.205...	20/Aug...	20/Aug...	22	64850	16...	6
141.142.4.6	12.234.205...	20/Aug...	20/Aug...	22	64926	34...	6

Figure 5. NetFlows Raw Data Tab Within MV


A
B
C
D
E
F
141.142.4.6
Ports That Used Other Protocols/Byte Count
Netflows
All Counts
Ports That Used Other Protocols/Flow Count
Ports That Used UDP/Flow Count
Ports That Used UDP/Byte Count
Ports That Used TCP/Flow Count
Ports That Used TCP/Byte Count
All Protocols Used
All Ports Used/Flow Count
All Ports Used/Byte Count
All Ports - Num times used :AllProtocols
Flow Counts
10
8
6
4
2
ECHO
FTP
SSH
TELNET
SMTP
TIME
NS
DNS
HTTP
KRB
IMAP
Ports
56152
55958
Src Ports - Num times used :AllProtocols
Dst Ports - Num times used :AllProtocols
NO DATA!
0
1
2

A. IP address of the machine
B. Different tabs that display information of this machine.
C. All the special ports vs., number of flows it was used in (flow count) this machine
D. All the other ports (All ports *minus* special ports) used by this machine
E. Set of bar charts showing ports that were used as source ports
F. Set of bar charts showing ports that were used as destination ports

Figure 6: NVisionIP Machine View (MV)

IV. Network Management

NVisionIP allows a network administrator to transparently monitor flows to/from each device on a network in order to learn the behavior of the network being managed. This is a different approach than alarming a network with sensors searching for signatures or thresholds for specified events. However, given clues an operator can also interactively configure GV filters to target suspicious activity for further inspection. Examples of activity NVisionIP has been used to detect includes:

- activity on unallocated parts of an address space indicating malicious scans or backscatter from attacks elsewhere
- DoS attacks into/out of a network
- devices infected with worms scanning to propagate showing a large number of connections attempts
- services conforming to official organizational policies
- unusual activity on ports not seen before
- large byte transfers to/from unexpected devices (malware)

The reader is referred to [1,3,5,10] to gain deeper insight into how NVisionIP has been found to help security engineers discover network security attacks.

V. Summary

NVisionIP is designed to help network administrators visually monitor the status of networked devices on IP address spaces. By presenting information visually on one screen with drill-down levels of detail - Galaxy View, Small Multiple View, and Machine View – a user may determine relationships between events at different levels transparently or with the help of filtering. The NVisionIP animation feature within the GV helps users understand how network devices change state over time. The end result is a situational awareness of the current state of networked devices on large and complex IP address spaces as well as a history of how devices came to their current state. The ability to view and interact with device state information on an entire logical IP address spaces is a new capability - to the knowledge of the authors NVisionIP is the only tool that currently provides this capability.

References

1. R. Bearavolu, K. Lakkaraju, W. Yurcik, and H. Raje, "A Visualization Tool for Situational Awareness of Tactical and Strategic Security Events on Large and Complex Computer Networks" *IEEE Military Communications Conference (Milcom)*, 2003.
2. E. Gamma, R. Helm, R. Johnson and J. Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley, 1994.
3. K. Lakkaraju, W. Yurcik, Adam J. Lee, R. Bearavolu, Y. Li and X. Yin, " NVisionIP: NetFlow Visualizations of System State for Security Situational Awareness", *Workshop on Visualization and Data Mining for Computer Security (VizSEC/DMSEC*) in conjunction with *11th ACM Conf. on Computer and Communications Security (CCS)*, 2004.
4. K. Lakkaraju, W. Yurcik, R. Bearavolu, and A.J. Lee, "NVisionIP: An Interactive Network Flow Visualization Tool for Security", *IEEE International Conference on Systems, Man and Cybernetics (SMC*), 2004.
5. K. Lakkaraju, R. Bearavolu and W. Yurcik, "NVisionIP—A Traffic Visualization Tool for Security Analysis of Large and Complex Networks", *Intl. Multiconference on Measurement, Modeling, and Evaluation of Computer-Communication Systems (Performance TOOLS),* 2003.
6. B. Shneiderman and C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edition*, Addison-Wesley, 2005.
7. J. Tidwell, *UI Patterns and Techniques. Retrieved.* <http://time-tripper.com/uipatterns/Overview_Plus_Detail>
8. E. R. Tufte, *The Visual Display of Quantitative Information 2nd edition*, CT:Graphics Press, 2001.
9. W. Yurcik, "The Design of an Imaging Application for Computer Network Security Based on Visual Information Processing," *SPIE Defense and Security Symposium/Visual Information Processing XIII*, 2004.
10. W. Yurcik, K. Lakkaraju, J. Barlow, and J. Rosendale, "A Prototype Tool for Visual Data Mining of Network Traffic for Intrusion Detection", *3rd IEEE International Conference on Data Mining (ICDM) Workshop on Data Mining for Computer Security (DMSEC),* 2003.
11. W. Yurcik, J. Barlow, K. Lakkaraju, and M. Haberman, "Two Visual Computer Network Security Monitoring Tools Incorporating Operator Interface Requirements," *ACM CHI Workshop on Human-Computer Interaction and Security Systems (HCISEC)*, 2003.