Guiding Security Analysis through Visualization

VAST 2011 Mini Challenge #2 Award: "High Potential for Scalability"

Lane Harrison * Wenwen Dou † Aidong Lu ‡ William Ribarsky § Xiaoyu Wang ¶

Computer Science
UNC-Charlotte

ABSTRACT

We present a multiple views visualization for the security data in the VAST 2010 Mini Challenge 2. The visualization is used to monitor log event activity on the network log data included in the challenge. Interactions are provided that allow analysts to investigate suspicious activity and escalate events as needed. Additionally, a database application is used to allow SQL queries for more detailed investigation.

Index Terms: H.5.2 [Information Interfaces & Presentations]: User Interfaces - Graphical User Interfaces (GUI); I.3.6 [Methodology and Techniques]: Interaction Techniques

1 INTRODUCTION

The scenario of the 2011 VAST Mini Challenge 2 was to develop a network monitoring visualization to "give insight as quickly and clearly as possible in order to minimize the burden on the CNO [computer network operations] team". This application was to be developed for the fictitious All Freight Corporation, a shipping company which operates soley in the United States.

Several data sources were provided representing network operations on the All Freight network for several days. These included network infrastructure and business priority information, firewall log data, intrusion detection system (IDS) logs, Windows system event logs, and more. See the challenge website for a complete list.

Given this data, participants were to design a visualization system to increase situational awareness. Furthermore, participants were to use their system to answer a series of questions regarding suspicious events, how early the events could be detected with their system, and the resulting recommendations of changes to make to the network.

The aim of the tool our team developed for this challenge was to provide an automatically-updating overview of network event activity that analysts could monitor for anomalous activity. However, we found it necessary to provide interactions to help analysts verify suspicious activity before moving to other security tools.

2 DESIGN PROCESS AND ANALYST INPUT

Part of the design process included a meeting with a security analyst who worked in a national bank's headquarters. The meeting lasted one hour and was held at the bank in their security operations center. The meeting was informative, as the analyst explained how each type of data in the challenge is used and analyzed in a large corporate network.

*e-mail: ltharri1@uncc.edu
†e-mail: wdou1@uncc.edu
‡e-mail: aidong.lu@uncc.edu
§e-mail: ribarsky@uncc.edu
¶e-mail: xwang25@uncc.edu

We presented our initial design, which was based on VisAlert [2]. VisAlert uses a radial display to show network events similar to the data in the challenge. The analyst expressed preference for a non-radial display, which was noted in final design Additionally, the analyst pointed out that the network infrastructure is valuable and should be emphasized, as many corporate networks do not have this information.

The analyst also demonstrated ArcSight [1], which they use for their log management and day-to-day analysis. ArcSight maintains a database of all event activity, which is useful in correlating events. Using ArcSight, the analyst was able to create time-based histograms of network event activity as well as basic link graphs that can show activity patterns between machines. However, the main display of arcsight consists of a table view showing events as they appear and a histogram of activity over time. Our final design expands on this by providing multiple charts and an overview of the network infrastructure.

Based on our meeting, we developed a simple multiple views visualization. To simplify the data management, the firewall, system event, and intrusion detection logs were imported into a MySQL database. Because the system event and intrusion detection logs were not in a database readable format, Perl scripts were implemented to convert them into comma separated value (CSV) format before importing. Navicat lite, a database management application, was used to provide access to the database. During the analysis, Navicat was used to execute queries as followups to suspicious activity.

3 VISUAL REPRESENTATION

The resulting visualization made use of the IDS, system, and firewall log data as well as the network infrastructure information. While the system can update automatically as new events occur, interactions are provided to assist analysts in classifying activity as malicious or benign.

First, the network infrastructure is preserved as much as possible in a node-link diagram. This will be referred to as the network view. Nodes in the network change color based on which machines are involved in a user-selected portion of the IDS and system log charts. This allows the analyst to discover which machines were involved in particular events.

Below the network view are two stacked bar charts. The top chart shows intrusion detection log events over time, which each stack division representing a different type of event. On the bottom is the system event stacked bar chart. Like the intrusion detection chart, each stack division in the system event chart represents a different type of event. The charts show all the events from the previous three hours.

Finally, on the right side is a simple text view that shows the types of intrusion detection events, system events, and firewall services used for a user-selected node in the current displayed timerange. The services query usually took several seconds to run, so it was only executed when a user clicked on a node.

The focus of this visualization was to increase situational aware-

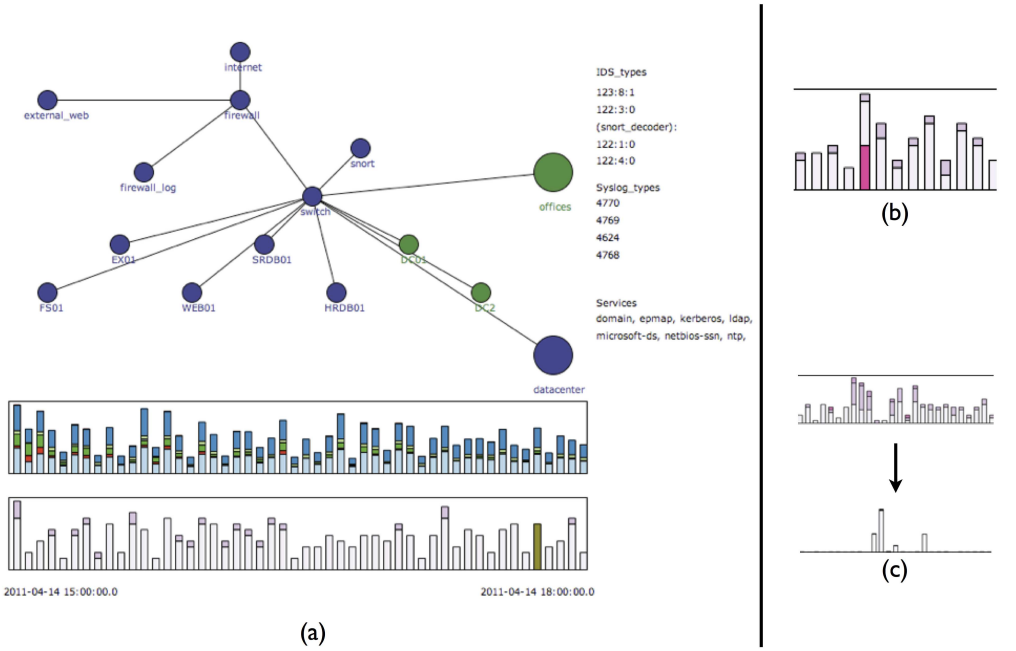

Figure 1: (a) An overview of visualization used. (b) The dark red-pink color is one rarely seen and warranted investigation. (c) Shows a transition in IDS event frequency. The bottom timestep has a few periods during which bursts of activity occur. This led to the discovery of the office machines exhibiting malicious behavior.

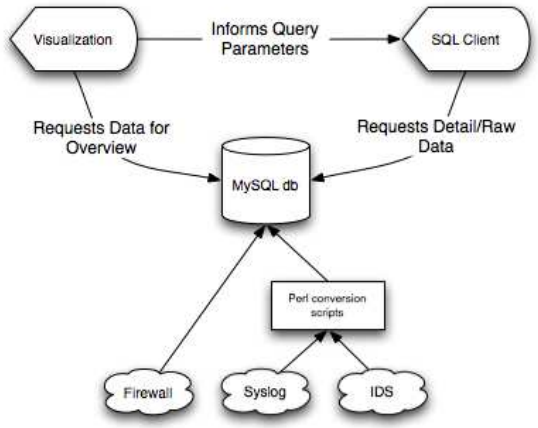

Figure 2: An overview of the analytical process and system.

ness and provide interactions that can help guide analysts in classifying and correlating suspicious activity. In the next section we describe the analysis strategy we used to find malicious activity in the All Freight network data (also see figure 2).

4 Analytical Process

Initially, we focused on monitoring the visualization for abnormal or suspicious looking activity. When suspicious activity was encountered, the visualization interactions were used to determine which machines were involved, what log events were triggered, and what specific services were used (found in the firewall data).

Since we did not have access to a security professional for the analysis portion of the challenge, we made use of the links provided in the challenge description to determine what certain log events meant. Furthermore, when we encountered an unfamiliar firewall service, we would search the web to find if it was associated with malicious activity.

Examples of visual events which triggered further investigation include significant changes in event frequency and unexpected events. A change in frequency is reflected visually by a significant change in network activity in a given time range when compared to previous activity. An unexpected event typically was shown as a color previously not seen in the barchart. Both of these are described in figure 1.

Since previous research has indicated that security analysts are comfortable with SQL and command-line tools [3], we made use of SQL queries to view the data in detail. The visualization proved useful in guiding the query by narrowing the time-range, IP addresses/range, and ports or services to include in the search.

5 Conclusion

A security log visualization approach is described that focuses on preserving network infrastructure information and providing interactions that help network defenders verify whether events are benign or malicious. Further investigation is supported by providing access to the raw data via SQL queries which can be guided by the visualization. Future iterations should incorporate other data provided in the challenge, such as the Nessus report and business priority information.

References

[1] ArcSight ESM - Enterprise Security Manager. http://www.arcsight.com/products/products-esm/.

[2] S. Foresti, J. Agutter, Y. Livnat, S. Moon, and R. Erbacher. Visual correlation of network alerts. *Computer Graphcs and Applications (CG&A)*, pages 1–12, Feb 2006.

[3] R. Thompson, E. Rantanen, W. Yurcik, and B. Bailey. Command line or pretty lines?: comparing textual and visual interfaces for intrusion detection. *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, Apr 2007.