Exploring the Design Space for Cyber Alerts in Context

Michelle Dowling* Virginia Tech
Lyndsey Franklin† Pacific Northwest National Lab
Mi Feng‡ Worcester Polytechnic Institute
Meg Pirrung§ Pacific Northwest National Lab
Robert Jasper¶ Pacific Northwest National Lab
Joseph Cottam‖ Pacific Northwest National Lab
Leslie Blaha** Pacific Northwest National Lab

Abstract

During knowledge elicitations with cyber analysts, we uncovered a need for tools that help analysts understand threat alerts in a context of baseline “normal” behaviors. We used an iterative design process to create a prototype alert management system with which we can explore the critical design space for effective baseline visualizations. We report herein on the design elements of this user interface, and present associated expert feedback from a design evaluation in the associated poster. We summarize our initial insights into the design of effective baseline visualizations that could be integrated into a larger visualization-support cyber alert management workflow desired by cyber analysts.

Index Terms: H.5.2 [Information Interfaces and Presentation]: user Interfaces—User-Centered Design H.5.2 [Information Interfaces and Presentation]: user Interfaces—Evaluation/methodology

1 Introduction

Supporting cyber analyst workflow requires addressing issues with mixing raw and aggregate data, interpreting the results of automation and intelligently handling alerts of differing severity. This abstract describes work in progress towards specifying the design space for a visualization-supported workflow for cyber alert management. A need for such a workflow was derived from expert knowledge elicitations and task analyses with four cyber analysts at the Pacific Northwest National Laboratory [3]. This elicitation identified in four key tasks in threat investigation workflows: alerting, thresholding, threat hunting, and reporting. All four tasks could benefit from visualizations capturing network activity patterns, highlighting known and novel threats, and separating threats from normal network behaviors.

The elicitation revealed that many existing visualization tools ignore the importance of context. Quickly capturing and understanding baseline network activity was the “most lamented” shortcoming of existing tools according to our pool of cyber analysts. Baselines provide critical contexts for understanding the behaviors that trigger an alert investigation. Current practices rely on experts either bringing their own mental models of baseline behavior derived through experience or developing contexts on the fly through time-intensive studying of log files or other relevant data sources. Importantly, in cyber analytics that leverage machine learning, the raw log data provide only part of the relevant baseline information. The data representation within the machine learning and the resulting outputs should also be integrated into the interactive visualization system.

To better study how we can support cyber analytic workflow with relevant baseline visualizations, we have designed a prototype alert management system to address three critical needs that were brought up during the interviews: 1) providing a method for understanding the semantics of the rules that generated an alert, 2) providing a baseline for what activities are considered “normal,” and 3) providing an interactive visualization of the alert data. The analysts desire visualizations similar to those provided by Splunk [4] but with interactions to enable further exploration of raw network data.

This abstract describes key design considerations for visualizing network alerts in the context of baseline data. The corresponding poster illustrates the full preliminary design space and provides both the critical feedback and initial evaluation results from testing with expert cyber analysts. Future work includes completing the relevant design space for a full system supporting effective integration of human and machine intelligence for a visualization-supported threat management workflow.

2 Overall Interface Design

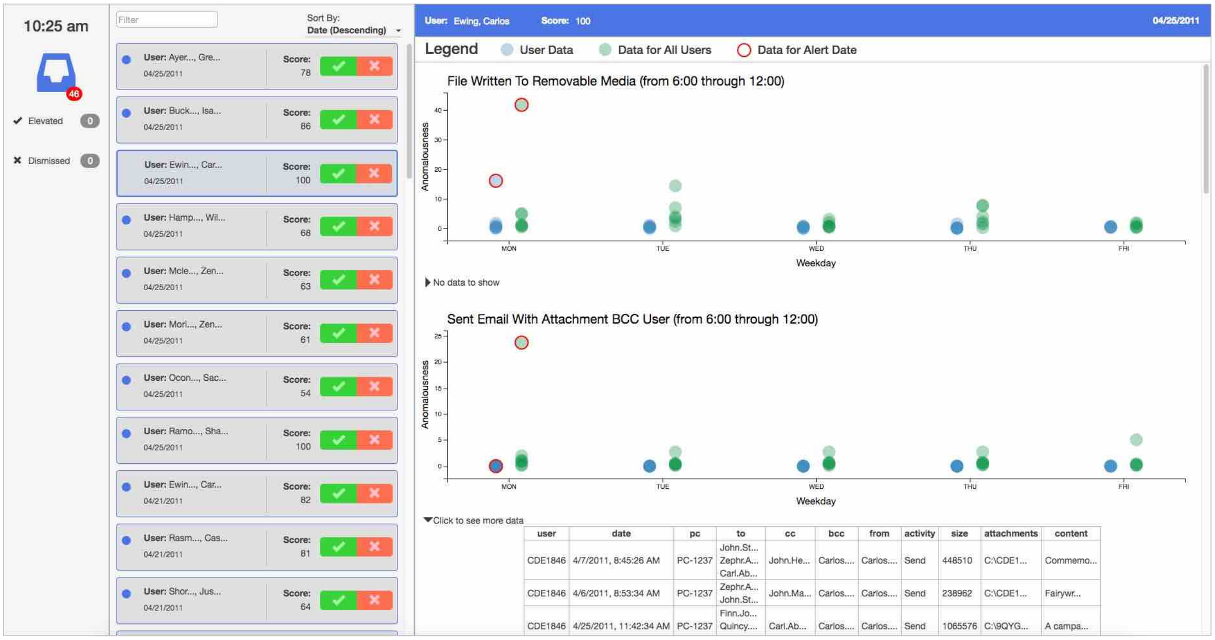
Our prototype interface emphasizes the triage of incoming alerts; one example is shown in Figure 1. To minimize the need to learn a completely new tool, we chose to use the familiar inbox metaphor to represent the incoming alerts. This metaphor also supports new alerts arriving without losing context of the existing alerts in the system. Alerts are listed in the second column in Figure 1, displayed as an overview of the user and date associated with each alert. Alert disposition in this system is accomplished through a simple sorting of the alerts into relevant inbox folders. This might indicate an analyst’s determination that an alert is a false positive, is in need of further observation, or requires an immediate elevation to full investigation. We note that in the interface in Figure 1, we are using the CMU CERT simulated data set of insider threat activities [2], and we have depicted a binary Elevate/Dismiss decision. Multi-class decisions are left to future implementations, particularly one’s leveraging interactive machine learning to aid in tracking and recommending class assignments.

3 Contextualizing Alerts with Interactive Baseline Visualizations

When an alert is selected, the right-hand read window is updated to display details associated with the given alert—much like clicking on an email to view its contents. These alert details are a representation of the alert event data and relevant contextual data corresponding to the given alert. This is the portion of the interface dedicated to capturing and contextualizing alert data. The key challenge to designing this visualization is that multiple types and resolutions of data must be made available through the interactive visualization. These include the following:

- Scores or alert values, indicating degree of rule violation or severity of alert
- Aggregate or embedding space features leveraged by machine learning or statistical analyses
- Raw data logs
- Relationships of machine learning features and raw data to alert triggers or scores

*e-mail: dowlingm@vt.edu
†e-mail: lyndsey.franklin@pnnl.gov
‡e-mail: mfeng2@wpi.edu
§e-mail: meg.pirrung@pnnl.gov
¶e-mail: robert.jasper@pnnl.gov
‖e-mail: joseph.cottam@pnnl.gov
**e-mail: leslie.blaha@pnnl.gov



User: Ewing, Carlos **Score:** 100 04/25/2011

User	Date	Score
Ayer..., Gre...	04/25/2011	78
Buck..., Isa...	04/25/2011	86
Ewin..., Car...	04/25/2011	100
Hamp..., Wil...	04/25/2011	68
Mcle..., Zen...	04/25/2011	63
Mori..., Zen...	04/25/2011	61
Ocon..., Sac...	04/25/2011	54
Ramo..., Sha...	04/25/2011	100
Ewin..., Car...	04/21/2011	82
Rasm..., Cas...	04/21/2011	81
Shor..., Jus...	04/21/2011	64

user	date	pc	to	cc	bcc	from	activity	size	attachments	content
CDE1846	4/7/2011, 8:45:26 AM	PC-1237	John.St... Zephr.A... Carl.Ab...	John.He...	Carlos....	Carlos....	Send	448510	C:\CDE1...	Commemo...
CDE1846	4/6/2011, 8:53:34 AM	PC-1237	Zephr.A... John.St...	John.Ma...	Carlos....	Carlos....	Send	238962	C:\CDE1...	Fairywr...
CDE1846	4/25/2011, 11:42:34 AM	PC-1237	Finn.Jo... Quincy....	Carl.Ab...	Carlos....	Carlos....	Send	1065576	C:\9QYG...	A campa...

Figure 1: Screenshot of the prototype inbox interface for a visualization-supported cyber threat management workflow.

One of the tricky things in representing machine learning features and output is that machine learning often involves transformations and embeddings of the data, particularly in cases of non-numeric data like cyber sources (e.g., [1]). Thus, capturing the mapping between the features and raw data requires particular design consideration and testing. For example, for the CERT data demonstrated in Figure 1, we leveraged the deep autoencoder from [5] to find anomalous patterns. This algorithm relies on features defined as counts of events in the data logs, with each feature aggregated over 6-hour windows. The 6-hour aggregates are further summarized to daily averages, which are shown in the scatterplot. Finding the right level of aggregation requires context awareness and testing with experts.

The activity of a given network user on any day can be flagged as anomalous, as represented in the output score. The anomalousness scores, transformed into percentiles, are included in the list of alerts in the inbox, providing the analyst the machine's perceived importance of the alert-triggering activity. Given the large number of available features in the machine learning representation, we provided by default the three most highly weighted activities contributing to the anomalousness score. Figure 1 illustrates two features, each with a dedicated scatterplot (the third being lower on the screen). The scatterplots are where we begin to provide more context for the analyst, taking advantage of the fact that comparable activity representations can be computed for a specific user/asset associated with an anomaly and for all other all assets in the network.

Using our prototype in Figure 1 to illustrate each point, we propose that an effective baseline visualization should:

- **Highlight alert-triggering activity** to direct the analyst to the unusual patterns. Anomalous activity values corresponding to the day of the alert are circled in red.

- **Provide activity history/context for the asset or user triggering the alert.** User activity is illustrated in blue dots in the scatterplots for a retrospective period of time. We here show the user's activity over the past 6 weeks. The vertical position of the dots is determined by the anomalousness of that activity on that day. The opacity of the dots is fixed to allow points to overlap while still communicating this information.

- **Provide activity history/context for other assets/users who should exhibit similar profiles.** We show a second set of green dots in each scatterplot for all users. The data plotted for all the network users shows their associated activities and activity weights over the same time frame, using same mapping of height and opacity as the alert-triggering user's dots.

- **Enable trace-back to the relevant raw data.** To display raw data associated with these scatterplots, we provided a table beneath each scatterplot. This table is available on demand, through a "see more data" option listed below each plot. Linking and brushing were enabled between the two to allow analysts to thoroughly explore the data.

Acknowledgments

The authors wish to thank the cyber analysts who participated in our interviews. This research was sponsored the U.S. Department of Energy through the Analysis in Motion Initiative at Pacific Northwest National Laboratory. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies, expressed or implied, of the U.S. Government.

References

[1] D. Arendt. Hot topics: Information retrieval for network security. In *Proceedings of* 10th *International Symposium on Visualization for Cyber Security*. Atlanta, GA, 2013.

[2] D. M. Cappelli, A. P. Moore, and R. F. Trzeciak. *The CERT guide to insider threats: how to prevent, detect, and respond to information technology crimes (Theft, Sabotage, Fraud)*. Addison-Wesley, 2012.

[3] L. Franklin, M. Pirrung, M. Dowling, M. Feng, and L. Blaha. Toward a visualization-supported workflow for cyber alert management using threat models and human-centered design. In *Proceedings of the* 14th *International Symposium on Visualization for Cyber Security*, 2017.

[4] Splunk, Inc. Splunk, 2017.

[5] A. Tuor, S. Kaplan, B. Hutchinson, N. Nichols, and S. Robinson. Deep learning for unsupervised insider threat detection in structured cybersecurity data streams. In *Artificial Intelligence for Cybersecurity Workshop at AAAI*, 2017.