VAST Challenge 2013: Situation Awareness and Prospective Analysis

Mark Whiting[1], Kristin Cook[1], Celeste Lyn Paul[2], Kirsten Whitley[2], Georges Grinstein[3], Bohdan Nebesh[2], Kristen Liggett[4], Michael Cooper[1], John Fallon[3]

[1] Pacific Northwest National Laboratory
[2] Department of Defense
[3] University of Massachusetts Lowell
[4] Air Force Research Laboratory

ABSTRACT

The 2013 Visual Analytics Science and Technology (VAST) Challenge presented three distinct challenge problems. Mini-Challenge 1 (MC1) asked participants to use visual analytics to predict the success of new movies. Mini-Challenge 2 (MC2) focused on the design of a situation awareness display for monitoring the health, performance, and security of a large computer network. Mini-Challenge 3 (MC3) requested participants to identify the timeline of important network events in two weeks of network data for a fictitious marketing company. The VAST Challenge received 31 final submissions, with MC1 receiving 106 interim submissions over the challenge period. Participants came from 11 different countries, and 14 awards were given.

Keywords: Visual analytics, human information interaction, sense making, evaluation, metrics, contest.

Index Terms: H.5.2 [Information Interfaces & Presentations]: User Interfaces – Evaluation/methodology.

1 INTRODUCTION

Now in its eighth year, the Visual Analytics Science and Technology (VAST) Challenge [1] aims to advance visual analytics through a series of competitions. In the VAST Challenge, researchers and software developers put themselves in the role of analysts to determine if their tools and techniques can address the specified problems effectively. VAST Challenge problems provide realistic tasks and data sets which live on after the completion of each year's challenge and are used for education, software evaluation, and demonstration of new techniques.

The VAST Challenge 2013 offered participants a variety of mini-challenges against which to exercise their visual analytics prowess. Mini-Challenge 1 (MC1), also known as "Box Office VAST," invited participants to use diverse data and create visualizations to support predictions of opening weekend box office gross and viewer ratings for movies. MC1 was the first VAST Challenge focusing on prediction, the first to include continuous participation and feedback, and one of the first to provide an opportunity for iterative improvements throughout its duration. Teams submitted entries each week prior to a movie's opening weekend, and they were provided with feedback on their visualizations, analyses, and accuracy from analysts and visual analytics researchers.

Mini-Challenge 2 (MC2) was a design-focused challenge, in which contestants were asked to create an innovative large display to support situation awareness in a large computer network operations center. This task used a unique format among VAST Challenge tasks, as no data was provided. The contestants received only a description of the network operations center and an outline of the goals for the improved displays. Participants were asked to take risks and envision creative solutions.

Mini-Challenge 3 (MC3) completed a three-year evolution of cyber network visual analytics tasks. Participants were provided with multiple types of network metadata representing two weeks of operation of the fictitious Big Marketing Company. Participants were challenged to identify a timeline of network issues and develop hypotheses about the issues they observed. MC3 provided participants a new opportunity to ask questions of the fictitious corporation's system administrators to gain information they needed to put the observed events in context.

The VAST Challenge committee assessed this year's challenges as difficult offerings for the visual analytics community. Two of the mini-challenges (MC1 and MC2) were very different than any previous VAST challenge offerings. MC3 was deliberately designed to be significantly more complex than the 2012 offerings, as requested by participants in the VAST Challenge 2012 workshop. The mini-challenges required participants to demonstrate their skills in data acquisition (notably in MC1) and ingest (MC1 and MC3), to create innovative visual designs (particularly MC2), to answer challenging questions (particularly MC3), and to provide clear explanations in both written and video form.

2 SCOPE OF VAST CHALLENGE 2013

VAST Challenge 2013 consisted of three independent mini-challenges. Teams were invited to submit to one, two, or all three mini-challenges. Each final submission required both a written entry with ample illustrations, and an explanatory video, which was useful for illustrating interactions important to the solution.

2.1 Challenge Tasks

The three individual mini-challenge tasks are summarized below. Descriptions of the tasks are posted at http://www.vacommunity.org/vastchallenge2013/. All mini-challenge materials are archived in the Visual Analytics Benchmark Repository [2].

2.1.1 Mini-Challenge 1: Box Office VAST

The theme of Box Office VAST was prospective visual analytics. Participants were asked to predict how well a set of movies would perform in the U.S. market in terms of box office "take" (ticket sales) and in movie viewer ratings for their opening weekends. A key requirement of the challenge was that contestants used visual

email: {mark.whiting, kris.cook, michael.cooper}@pnnl.gov, {clpaul, banebes}@tycho.ncsc.mil, grinstein@cs.uml.edu, kristen.liggett@wpafb.af.mil, john_fallon@student.uml.edu

analytics to support their movie analysis. The challenge developers created an entire website (http://www.boxofficevast.org) dedicated to supporting this mini-challenge.

Data for MC1 was not provided all at once as in previous challenges. Data came from three sources: IMDb, the Internet Movie Database; Twitter, an online microblogging platform; and Bitly, a link-sharing service. Each of these data sets was periodically updated. The IMDb data was updated each Friday. Twitter and Bitly data concerning a movie was updated each day from the day the movie was announced as being part of the mini-challenge, through the Thursday before the opening weekend.

Movies in the U.S. typically open on a specific weekend. The challenge developers selected movies for certain opening weekends about two weeks in advance. Data was made available for those movies to be used in making predictions. Predictions and visualizations were due the Thursday before the opening weekend. On the Monday after the opening weekend, preliminary box office take and viewer ratings were posted. On the Friday following the opening, official IMDb numbers were posted. Evaluations included how close predictions came to actuals and reviews of the visualizations submitted. Recognitions were posted on a leaderboard, and submitters received comments from visualization and information analysts to help improve their efforts. On average, there were 3 movies to predict per week, and the mini-challenge ran for 25 weeks..

Movie prediction challenges have previously been offered in the data mining and machine learning communities. The specific goals of this mini-challenge included featuring visualization as a prominent part of the analysis, integrating and using structured and unstructured data, and innovating ways a user could interact with the software environment to enhance predictions.

2.1.2 Mini-Challenge 2: Innovative Design

MC2 tested participants' skills in visual design, with a strong emphasis on creativity and innovation. In the MC2 scenario, the fictitious Big Enterprise is searching for a design for their future cyber situation awareness display. The company's network operations team will use this display to understand the health, security, and performance of their very large computer network. Participants were provided a description of the mini-challenge goals, background on the general concept of situation awareness, and an overview of the current operation of the Big Enterprise network operations center. The description of the current operations was kept intentionally brief and at a high-level, with the intent of discouraging participants from focusing on adaptation of current techniques to a specific use case. Instead, it was hoped that the relatively limited information would free participants to think creatively and establish their own boundaries and assumptions.

Participants were given the following guidance regarding priorities for creating their solutions:

- The sky is the limit: Big Enterprise is looking for bigger and bolder ideas than the status quo.
- The bigger the better: Represent the biggest network you can.
- The bigger the network, the more stuff happens: Represent the greatest amount and diversity of activity you can in the design.
- Managing complexity is essential: Balance network scale and network activity with effectiveness of the visual design.

Participants were asked to provide a high-resolution image of their full situation awareness display along with storyboards illustrating expected interactions.

2.1.3 Mini-Challenge 3: Complex Computer Network Analysis

MC3 was the culmination of a three-year series of VAST Challenge tasks focusing on situation awareness in the cyber security domain. In 2011, the challenge focused on network security for the fictitious All Freight Corporation, a network consisting of approximately 150 computers [3]. This dataset consisted of three days of data from multiple data sources, including intrusion detection system, firewall and syslog log files. The 2012 challenge tasks focused on the element of scale [4], particularly in Mini-Challenge 1, which required contestants to provide situation awareness of network health data for a million node network. The 2012 Mini-Challenge 2 contained log data for a 5000 computer network, providing an order of magnitude increase in network size over 2011.

This year, MC3 focused once again on providing situation awareness of a large computer network, but this task represented an increase in complexity from previous years in two important ways. First, it provided data from two weeks of network operation, which is a much longer time span than the previous challenges in the series. Second, the embedded scenario included over 30 ground truth events, compared to four or five events in previous years. The data sources were NetFlow [5], Big Brother [6] (health and status metrics) and intrusion prevention system logs. Participants were asked to infer issues on the network primarily from patterns and anomalies they observed in the network metadata and health data.

A true reflection of cyber data, MC3 presented participants with a low-level view of activity in which human intent is obscured. To help gain greater insight into the clues participants saw in the fictitious network's data, they were encouraged to ask up to five questions of the fictitious company's system administrators. In response, the VAST Challenge committee provided information that helped them verify hypotheses and reveal an underlying story.

2.2 Review Process

The VAST Challenge committee recruited reviewers with expertise either in visual analytics, information analysis, or application domains. Experts in design-related disciplines were also recruited to review MC2. Subject matter experts were recruited from the pool of previous VAST Challenge reviewers and their social networks. A total of 66 reviewers participated, each providing from 1 to 10 reviews. Each submission received 4 to 7 anonymous peer reviews. All reviewers were given the opportunity to recommend entries for award consideration.

Peer review questions varied across the individual challenges. However, in all cases, reviewers provided both ratings and explanatory comments. MC1 review areas of interest included creativity, visualization and interaction effectiveness, use of structured and unstructured data, identification of key trends, and improvements over time. MC2 review questions focused on creativity, effectiveness, the degree to which the design would scale to large networks, and overall quality.

Although MC3 was in many ways a continuation of the 2012 MC2, the review criteria were somewhat different. In 2012, reviewers were asked to evaluate the plausibility of the answers provided, rather than the accuracy of the solutions. This proved to be very challenging for the peer reviewers. In 2013, accuracy reviews were reinstated, but all accuracy reviews were performed

by a small subcommittee of people very familiar with the data. These accuracy reviews identified the degree to which the submissions identified the events embedded in the data, but the accuracy reviewers also gave credit to submissions that identified other valid events in the data that were not intentionally embedded as part of the scenario. The accuracy review scores were not provided to the peer reviewers, but they were provided to the VAST Challenge Committee for use in the award determination process.

MC3 review criteria also considered the degree to which the solution supported situation awareness for "loud" (or more obvious) and subtle events, the quality of the analysis, the clarity of the submission, the quality of the visualizations and interactions, the potential of the solution to scale and to support real-time situation awareness, novelty, and overall submission quality.

The VAST Challenge Review Committee held separate one-day meetings to determine awards for each of the three challenges. During the meeting, the committee considered the reviewer award recommendations and finalized the list of awards and honorable mentions based on all reviewer scores and comments.

3 VAST Challenge 2013 Results

The submissions recognized for an award in 2013 are listed in Table 1. Additional information about the Challenge entries can be found in the Challenge papers included in the VAST electronic proceedings.

3.1 Mini-Challenge 1 Awards

Box Office VAST included interim recognitions for the participating teams and final entry awards. During the challenge, an Analyst's Recognition went to the Inria's CinemaViz team, led by Charles Perin, and to the University of Konstanz's Team Kurboknopf, led by Daniel Seebacher. A UX/HCI Recognition went to Team VADER from Arizona State University and to Team Turboknopf. These recognitions were noted on the leaderboard on the boxofficevast.org site, and their visualizations can be seen on boxofficevast.org site under the Featured Visualizations tab.

The three award winners for MC1 were comprehensively successful, as well as featuring individual outstanding qualities. The awards focused on specific notable qualities that made these submissions stand out. Arizona State University was recognized for their visual approaches to blending structured and unstructured data in their tools and analyses. The University of Konstanz Moovis team, led by Michael Hund, demonstrated very effective human analyst-prediction tool interactions in their submission. The other University of Konstanz team (Seebacher) submission showed an outstanding visual design across their toolset. The University of Stuttgart's Team Prolix, led by Steffen Koch, was also recognized with an honorable mention for their tool's ability to facilitate in-depth visual exploration across a broad range of data.

3.2 Mini-Challenge 2 Awards

Since Mini-Challenge 2 focused on out of the box thinking for design, creativity awards were given to two submissions from two different Purdue teams. *Solar Wheels*, led by Jack Shen-Kuen Chang, was recognized for the approach the team took to both responding to the large scale needs presented by the MC and for both the interaction style that would allow users to observe the entire situation in a large display and also walk-up to a more detailed presentation of information. *SpringRain*, led by Marlen Promann, was recognized for its different take on a situation awareness interface—one that provided a "data within a waterfall" metaphor, while also preserving the ability to quickly and simply alert system administrators to problem areas. Three Honorable Mentions were also awarded. SAS Institute's *NOCturne*, led by Jordan Benson, was recognized for specific visualizations used for the MC2 problem, including a linked connectivity matrix and topology graph showing the situation at major network locations. The University of Konstanz team, led by Fabian Fischer, designed a dashboard featuring detection and adaptation of the display when an analyst turned their attention toward it. MIT Lincoln Laboratory's display design, led by Diane Staheli, featured various approaches to depicting event relationships within the Big Enterprise network.

Table 1: VAST Challenge 2013 Awards

Mini-Challenge 1
Arizona State University, *Box Office VAST – Team VADER*: Excellent Visual Analysis of Structured and Unstructured Data
University of Konstanz (Hund Team), *MooVis: A Visual Analytics Tool for the Prediction of Movie Viewer Ratings and Boxoffice*, Excellent Interactive Analysis
University of Konstanz (Seebacher Team), *Similarity-Driven Visual-Interactive Prediction of Movie Ratings and Box Office Results*, Effective Visual Design
University of Stuttgart, *Prolix - Visual Prediction Analysis for Box Office Success*, Honorable Mention for In-Depth Visual Exploration of Features
Mini-Challenge 2
Purdue University, *Solar Wheels*: Outstanding Creative Design Award
Purdue University, *SpringRain*: Outstanding Creative Design Award
SAS Institute, *NOCturne*: Honorable Mention - Interesting Visualization Technique
University of Konstanz, *Adaptive User-Aware Dashboard Design*: Honorable Mention - Interesting Visualization Technique
MIT Lincoln Lab, *Situational Awareness Display Design*: Honorable Mention for Visualization of Event Relationships
Mini-Challenge 3
Central South University, *Big Marketing - Visual Analytics for Network Security Situation Awareness*: Outstanding Comprehensive Solution
Peking University, *AnNetTe - Visual Analytics for Network Security Data*: Outstanding Situation Awareness Award
University of Konstanz, *Using VACS for Visual Exploration of VAST Challenge 2013*: Honorable Mention for Intriguing Visualization
Johns Hopkins University Applied Physics Laboratory, *Visual Analytics for Network Situation Awareness*: Honorable Mention for Intriguing Visualization
Penn State University, *Leveraging 'Visualization Functions' in Hypothesis-based Collaboration on Cyber Analysis*: Honorable Mention for Noteworthy Collaborative Analysis Strategy

3.3 Mini-Challenge 3 Awards

Awards for MC 3 were given to teams that invested considerable effort in tacking the sophisticated problem presented this year. The comprehensive solution award given to Central South University, led by Ying Zhao, reflected that team's handling of multiple views, multi-source data, with the ability to handle the hierarchies of data involved in the challenge. Peking University and the University of Stuttgart's collaborative team, led by Siming Chen, received a situation awareness award for their entry *AnNetTe*, with which they were able to identify twelve significant events within the MC3 data and characterize the activities over the two weeks of data particularly well. Two teams provided submissions that, while not sufficiently developed to win

an award outright, were creative and intriguing to garner Honorable Mentions. The University of Konstanz team, led by Fabian Fischer, described a system of interactive visualizations called Visual Analytic Suite for Cyber Security (VACS), usable with a powerwall display. The Johns Hopkins University, Applied Physics Laboratory offered Galaxy, an event visualization tool that they adapted for this cyber challenge. Their node space and link space tools were atypical visualizations of network activity. Penn State University's team, led by Chen Zhong, also received an Honorable Mention for their application of "Visualization Functions" to enhance collaborations and make analytics more agile.

4 Discussion

Each of the 2013 VAST Mini-Challenge tasks represented a departure from previous years in some ways. The participants demonstrated creativity and ambition in their proposed solutions to these difficult challenges. This section summarizes the VAST Challenge Review Committee's observations about the submissions for this year's challenges.

4.1 General Observations

VAST Challenge 2013 received 31 submissions across the three challenges. Table 2 compares the number of submissions over the life of the VAST Challenge. MC1 had 106 interim submissions over the course of the challenge, far beyond any previous participation count. This represents significant effort by its participants, and it was encouraging to see that teams were willing to actively work on the mini-challenge over several months. When considering only final submissions, the number of entries in each of the 2013 mini-challenge tasks was lower than expected. There are many possible reasons for this decrease, however since each mini-challenge was so different from the others there may be no explanation applicable across all MCs. The new participation format for MC1 may have deterred some participants. The design-only format of MC2 was quite different from previous mini-challenge tasks, and new challenges have resulted in lower than average participation in the past. MC3 required significant data processing and analysis which may have limited participation.

Student teams performed exceedingly well in 2013. Student groups captured three of the four awards on MC2 and MC3, and were also well-represented on MC1. It was noteworthy that the student teams exercised creativity in their submissions.

We encourage risk-taking on these mini-challenges. The VAST Challenge provides an opportunity for participants to think broadly and creatively, even if the solution may be impractical today. The committee strongly believes that creativity and innovation are essential to producing the breakthroughs that will lead to high-impact future solutions.

4.2 MC1 Observations

Several of the goals for Mini-Challenge 1 were met by the participants. Submissions demonstrated that interactive visualization substantially benefitted the prediction process. Visual interaction was used in all stages: identifying features, exploring & modifying model performance, exploring sensitivity, and explaining results. Entries also showed how structured data can be combined with unstructured data in analysis environments supporting a predictive task. Data trends were identified and used in the prediction process. Several teams were able to explore subtle feature relationships, such as variation by genre or by release date.

Additional successes by the contestants included their ability to adapt to various changing conditions over the weeks and incorporating flexibility into their tools. During one week, the contestants were provided no tweet data, due to a problem with the data feed capture software. The number of interim submissions for that week did not decrease, and the participants used this as an opportunity to adapt their tool to a problem akin to what might be faced in the real world.

The mini-challenge creators did not focus on accuracy as a measure of success for the submissions, as we severely restricted the data to be used in the challenge and real world prediction tools would be using a much broader and larger set of data from which to predict. However, the submissions did provide reasonably accurate predictions in general, often comparing favorably with both professional movie site predictions and the actual opening weekend figures.

A successful mini-challenge also generates many follow-on questions, as was the case with Box-Office VAST. For example, the challenge designers had several questions concerning how we might weight early predictions more favorably than later ones and how we should consider accuracy against submission time. Another set of questions concerned data drivers for the predictions. Featured actors often drive movie successes, but, for example, Johnny Depp was not enough to drive the *Lone Ranger's* opening weekend over a mere $29 million.

Table 2: VAST Challenge Submission Counts

Submissions	2006	2007	2008	2009	2010	2011	2012	2013
Mini-Challenge 1	-	-	22	22	14	30	27	10*
Mini-Challenge 2	-	-	13	17	22	8	13	10
Mini-Challenge 3	-	-	12	5	17	13	-	11
Mini-Challenge 4	-	-	20	-	-	-	-	-
Grand Challenge	6	7	6	5	5	5	-	-
Total	***6***	***7***	***73***	***49***	***58***	***56***	***40***	***31***

**Mini-Challenge 1 for 2013 received 106 interim submissions.*

4.3 MC2 Observations

This was the first year that the VAST Challenge included a purely design-oriented challenge. As such, it was a learning experience for the designers to determine how best to present the challenge task and for the committee to evaluate the submissions received. Ultimately, it was decided to evaluate the submissions using a very simple set of four criteria as described in section 2.2. We are grateful to the many reviewers who provided us with detailed suggestions regarding the review process.

The submitting teams clearly took diverse approaches. The most successful teams were the ones who integrated designers into their team. As with the other mini-challenges, there were many "safe" designs that iterated on the status quo. While progress can and has been made using this approach, the design challenge was specifically aimed at rewarding those who chose to take risks and leap ahead.

Many team solutions struggled with designing for a large display. For example, the visualizations were appropriate, but the

use and size of text elements were often too small to be useful. Situation awareness is still a difficult challenge for the visualization community. There has been considerable improvement over the course of three years involving cyber situation awareness challenges, but there is still a struggle between "awareness" and "analysis".

4.4 MC3 Observations

This mini-challenge was clearly very complex for participants to address, just as it was complex to produce. However, the challenge materials were downloaded by 273 unique email addresses by the submission closing date.

The goal for the mini-challenge was to provide situation awareness, yet we found that most participants preferred to focus on providing in-depth analysis solutions rather than a situation awareness view. This may be a partial consequence of the fact that data is provided to the participants as a static set rather than in a pseudo-real time stream. Regardless, situation awareness for network security remains an opportunity area for research.

There were some participants who did a particularly noteworthy job of reporting their results by pulling together an analytic story. These submissions were much simpler for reviewers to understand and appreciate. This analytic story is also an important element of effective situation awareness. We encourage future participants to concentrate on techniques for providing their submissions in a way that expresses the narrative of what is happening, rather than listing facts without context. While there were several outstanding and highly creative submissions received, the committee also encourages more traditional solutions that demonstrate strong design principles. Unfortunately, some of the submissions that presented more traditional approaches did not clearly express the rationale for their design choices. A thorough description of design choices would have provided the reviewers with a much better appreciation for the quality of the solution.

This challenge represents the first time that the participants have been given the opportunity to ask questions of the committee in order to gain additional insight into the story behind the data. The committee was surprised at the relatively few questions received, but even more surprised by the disparity between the types of questions anticipated and the questions received. The committee had expected participants to identify the pattern of attacks in the data and ask questions that might help explain who was interested in attacking the network. We also expected participants to ask questions about the data exfiltrations evidenced in the data. However, participant questions typically focused on data characterization (the identity of a specific computer, for example). There are several potential reasons for this, including the complexity of the challenge and the lack of deep familiarity with the cyber security domain. We also observed that this year's participants looked for the types of issues embedded in the 2011 and 2012 challenges, such as the presence of unidentified computers on the network. However, in this year's challenge, such issues were only distractions from the serious problems affecting the network.

VAST Challenge datasets are generally constructed to contain a wide range of clues ranging from the extremely obvious to the very subtle. The most successful submissions are generally those that could go beyond the obvious to identify the relevant subtle signals. This year, many participants pointed out more obvious signals such as denial of service attacks and port scans. However, these obvious issues are also very common and fairly innocuous. Very few participants actually identified the subtle but much more significant issues such as the exfiltration of a couple of very large data files via FTP. These subtle signals were much more important in this dataset, but participants were often distracted by the more obvious signals.

Given the complexity of the network security field, it is clear that engagement with domain experts can make a significant impact on participant success when tackling a specialized field. Having a domain expert as a member of the team or as a consultant can help participants prioritize tasks and data.

4.5 VAST Challenge as a Research Community Resource

The VAST Challenge Participant Workshop has been held during IEEE VIS conferences for the past several years. These workshops highlight outstanding submissions to the challenge and provide a forum for community discussion of needs and opportunities related to the challenge problems. These workshops also have a direct bearing on the scope and scale of future VAST Challenges. For example, this year's MC3 is a direct outcome of the 2012 workshop, during which participants asked to be challenged with greater complexity in the scenarios.

In addition, the VAST Challenge has become a source of data and problems used in curricula in multiple countries. These datasets are also used to evaluate visual analytics tools and provide a resource for researchers and tool providers.

5 Toward VAST Challenge 2014

Plans are already getting underway for VAST Challenge 2014, which is expected to include a diverse array of challenges. Although the scope of the 2014 is not finalized, the committee hopes to present the community with an interrelated set of new challenge tasks that deal with textual, network, and geospatial data.

Mini-challenges 2 and 3 mark the completion of a three-year cycle focused specifically on cyber security situation awareness. We anticipate a continued focus in the cyber security domain, but with a shift toward "moving targets" for computer network defense.

Acknowledgements

The committee acknowledges the U.S. Department of Defense, Pacific Northwest National Laboratory, Central Intelligence Agency, United States Air Force, and the University of Massachusetts Lowell for supporting the VAST Challenge.

The committee also wishes to thank Wendy Cowley, Adam Roberts, Aaron Pickett, Jim Talbott, John Burnette, Russ Burtner, Ian Roberts, and Kelly Sandretto of Pacific Northwest National Laboratory; Catherine Plaisant at the University of Maryland, and Patrick Stickney of the University of Massachusetts Lowell.

References

[1] K. Cook, G. Grinstein, and M. Whiting. The VAST Challenge: history, scope, and outcomes: An introduction to the Special Issue. *Information Visualization*. Online publication July 12, 2013, doi:10.1177/1473871613490678. http://ivi.sagepub.com/content/early/2013/07/02/1473871613490678. Accessed 2013-09-01.

[2] Visual Analytics Benchmark Repository, http://hcil.cs.umd.edu/localphp/hcil/vast/archive/index.php. Accessed 2013-09-01.

[3] G. Grinstein, K. Cook, P. Havig, et al. VAST 2011 Challenge: cyber security and epidemic. *Proceedings of VAST 2011* (Providence, RI, 23–28 October 2011) Piscataway, NJ: IEEE Computer Society Press.

[4] K. Cook, G. Grinstein, M. Whiting, et al. VAST Challenge 2012: visual analytics for big data. *Proceedings of VAST 2012* (Seattle, WA, 14–

19 October 2012), pp. 251–255. Piscataway, NJ: IEEE Computer Society Press.

[5] M. Lucas. *Network flow analysis*. No Starch Press, 2010.

[6] S. MacGuire. Big Brother: A Web-based Unix Network Monitoring and Notification System, Sys Admin: 1–6, October 1996.