VAST 2011 Challenge: Cyber Security and Epidemic

Georges Grinstein
University of Massachusetts Lowell

Kristin Cook
Pacific Northwest National Laboratory

Paul Havig
Air Force Research Laboratory

Kristen Liggett
Air Force Research Laboratory

Bohdan Nebesh
Department of Defense

Mark Whiting
Pacific Northwest National Laboratory

Kirsten Whitley
Department of Defense

Shawn Konecni
University of Massachusetts Lowell

ABSTRACT

The 6th Visual Analytics Science and Technology (VAST) Challenge posed three related mini-challenges for participants to solve using a combination of visual analytics software and their own analytic reasoning abilities. Teams could solve one, two or all three mini-challenges and assess the overall situation to enter the Grand Challenge. Mini-challenge One (MC1) involved the characterization of the spread of an epidemic using given maps, geospatial and text data gathered from microblog tweets. Mini-challenge Two (MC2) involved the development and use of situation awareness data to identify issues of concern in the computer networking operations at a major freight shipping company. Mini-challenge Three (MC3) involved the exploration of a corpus of news articles to examine terrorist threats to a metropolitan area. The Grand Challenge was to determine whether the epidemic spread, the network events, and the potential terrorist groups identified in the mini-challenges were related. Participants were asked to analyze the data and provide solutions and explanations for the various challenges. The Challenge data sets were downloaded by nearly 600 people by the time submissions closed. The Challenge received 56 submissions, drew participants from 11 different countries, and gave 12 varied awards.

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

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

1 BACKGROUND

Now in its sixth year, the objective of the VAST Challenge [1] is to provide researchers with realistic tasks and data sets for evaluating their software, as well as to advance the field in solving more complex problems. The VAST Challenge is designed to help researchers understand how their software would be used in a novel analytic task and determine if their data transformations, visualizations, and interactions would be beneficial for particular analytic tasks. Researchers and software providers have repeatedly used the data sets from throughout the life of the VAST Challenge as benchmarks to demonstrate and test the capabilities of their systems. The ground truth that is embedded in the data sets has helped researchers evaluate and strengthen the utility of their visualizations.

2 VAST 2011 CHALLENGE SCOPE

The VAST 2011 Challenge consisted of three related mini-challenges (MC1, MC2, and MC3) and one Grand Challenge (GC). Each mini-challenge consisted of a data set, instructions, and questions to be answered. The GC required participants to integrate the information from all three data sets and write a brief summary and explanation of the overall situation.

The VAST 2011 scenario featured various identifiable terrorist activities, an epidemic, and a freight company's network security logs. All of the events in the scenario occurred in the fictional city of Vastopolis during the first half of 2011. MC1 consisted of text (tweets) which participants needed to process to identify the symptoms and details of an epidemic. There were two different sets of illnesses, a waterborne illness and an airborne illness. The participants were asked to locate and pinpoint the source of the epidemic, to describe the method of transmission of the epidemic, and determine if deployment of treatment resources outside of the affected area was necessary. MC2 provided over 8GB of network logs, including vulnerability scans, firewall logs, operating system security logs, intrusion detection system logs, and optional packet capture data. Participants were asked to develop a situation awareness visualization encompassing this data and to identify major network events transpiring over a three-day window. MC3 required participants to analyze a corpus of over 4,000 news articles to determine if there were any imminent terrorist threats.

The data for MC1 and MC3 were developed by the IVPR[2] at the University of Massachusetts Lowell. The first task for the MC1 scenario was the identification of a city with a river running through its center. Next, the tweet data set was created using a mixture of real tweets collected from Twitter along with a set of synthetic tweets generated with controlled tokens and synonyms using simple dictionaries. These were processed to remove foul language and embed map (latitude-longitude) and time information matching associated weather data. MC3's data set was created from a corpus of old news articles filtered to remove proper nouns and other text (dates and unique headers) that would give away the data set's true origin. There were about fifty additional articles injected into the data set that contained both ground truth and secondary misleading scenarios.

The MC2 data set was developed by Pacific Northwest National Laboratory. This data was created by developing a synthetic network which simulated the architecture of the fictitious freight company. The data were produced by simulating activity, including attacks, on this network over the course of the three day operating period.

grinstein@cs.uml.edu; kris.cook@pnnl.gov; Paul.havig@wpafb.af.mil; Kristen.liggett@wpafb.af.mil; danko1@gmail.com; mark.a.whiting@pnnl.gov; skoneci@yahoo.com

IEEE Symposium on Visual Analytics Science and Technology
October 23 - 28, Providence, Rhode Island, USA
978-1-4673-0013-1/11/$26.00 ©2011 IEEE

[2] Institute for Visualization and Perception Research

3 VAST 2011 Challenge Submissions

Teams were asked to provide a video and a concise process description as to how they arrived at their conclusions and how the various visualizations and tools helped in the analysis.

Participants submitted 5 GC entries and 51 MC entries. Table 1 shows a comparison of the number of submissions over the life of the VAST Challenge.

	2006	2007	2008	2009	2010	2011
MC1 Submissions	-	-	22	22	14	30
MC2 Submissions	-	-	13	17	22	8
MC3 Submissions	-	-	12	5	17	13
MC4 Submissions	-	-	20	-	-	-
GC Submissions	6	7	6	5	5	5
Total Submissions	6	7	73	49	58	56

Table 1. Summary of Number of Submissions by Year

The number of entries this year was on par with those of the past years. This was especially rewarding given the diversity of data included in the challenge. Most interesting is that, as of publication time, the dataset has been downloaded 671 times, as compared to 537 in 2010. In addition, this year's challenge had more than twice the number of student entries as non-student entries (38 vs. 18). To successfully compete in the GC, participants were required to transform, visualize, and analyze data from all three mini-challenges. The analytic tasks were diverse, ranging from situation awareness to identification of geospatial and temporal trends to criminal investigation.

4 Review Process

The VAST Challenge Review Committee recruited reviewers from throughout the visualization and analysis communities. Several subject matter experts learned about the Challenge through a blog entry published by an analyst educator. In all, 56 reviewers participated, with reviewers providing between one and six reviews each. Three to six external peer reviewers, including at least one subject matter expert, reviewed each entry. The reviewers were given an opportunity to recommend submissions for specific awards.

Each reviewer was given electronic access to the solutions for their assigned submissions. Reviewers were asked to rate the analytic process, the visualizations, the interactions, and the novelty of the submission. Reviewers were also asked to evaluate the accuracy of each team's solution. However, as the tasks and data sets for this challenge were more realistically complex, accuracy was not the only measure of interest. For example, the groups and events associated with the terrorist threat of MC3 relied on finding thirteen critical articles in a corpus of over 4,000 articles. To appropriately identify all of the network events embedded in MC2, participants needed to jointly analyze all sources provided and discriminate between innocuous anomalies and important network attacks. In both cases, all teams were able to discover non-trivial information in the data sets and several teams achieved close to accurate solutions. Interestingly, as in the past, some teams found other patterns not anticipated by the developers of the data set.

The VAST Challenge Review Committee held a two-day meeting to determine awards. The Challenge committee members each took responsibility for reading and summarizing the submitted reviews for one or more of the mini-challenges. The committee reviewed and evaluated the award recommendations from the reviewers and identified additional appropriate awards.

As in previous years, the awards were not pre-established. Instead, the committee identified awards recognizing the best qualities in the submissions. Awards were given for overall quality, analytic processes, innovative approaches, clarity of explanation, and potential for scalability. All teams receiving an award were given the opportunity to contribute two-page summaries for the proceedings. As in the past, all submissions and publications will be available at the Visual Analytics Benchmark Repository [2]. All teams will receive certificates of participation and are invited to the VAST Challenge Participants' Workshop at the 2011 IEEE VisWeek Conference to demonstrate their software and approach.

5 Summary of VAST Challenge 2011 Awards

Several trends were noted in this year's submissions. While a number of teams wrote custom software to address the challenges, as had occurred in the past challenges, several teams developed software using visualization toolkits. It was common for teams to use existing software, including commercial software, to address all or parts of the Challenge. Of particular interest was the fact that a tool that was used for a previous year's Challenge as a research prototype was used in this year's challenge by several teams as an established "off-the-shelf" tool.

Also notable in this year's Challenge were a few entries that took advantage of new form factors for innovative analysis. One entry made use of a tablet device, while at least two others made use of very large display environments.

The level of data preprocessing performed by the teams was notable on MC2. Although the data provided for this mini-challenge was relatively small compared to real-world environments, the size and diversity of the data sources necessitated that teams develop strategies for data management and multi-type data analysis. In the previous Challenges potential submitters asked the VAST Challenge committee to provide preprocessing in the form of extracted entities, text processing similar to that required for MC1 and MC3. This year no such requests were received.

MC2 asked submitters to provide situation awareness visualizations, which would permit users to see at a glance the health of their network and the presence of emerging issues. However, most of the submissions provided visualizations and interactions more oriented toward forensic analysis than situation awareness. This represents an opportunity for further development and future Challenge tasks.

Awards were given for the novel use of specific tools (for example, the use of word clouds for filtering other visualization), for outstanding analysis, for novel extensions to mobile devices and to large screen workspaces to support collaboration, for informative use of statistics and evidence in a report, for innovative tool adaptation, and for scalability.

Challenge	Student Team	Non-Student Team	Total Awards
MC1	3	2	5
MC2	3	0	3
MC3	2	1	3
GC	1	0	1
Total Awards	9	3	12

6 Participant Discussion Workshop Session

Participant workshops have been held during VisWeek every year since 2008. This workshop combines invited speakers with group discussions and an opportunity for teams to demonstrate their solutions. A participant workshop is being planned for VisWeek 2011 to continue this tradition and provide an opportunity for the teams to interact with one another.

7 The Path Forward

This VAST Challenge marks the sixth year of the event. This event has consistently attracted significant participation. As stated above, there were more than 671 downloads this year. We have learned a great deal about scenario and data set generation [3, 4]. We know the data sets are being used from email requests, downloads, and citations. Classes and software companies continue to use them and thus they represent a valuable asset as benchmark data sets for the visual analytics community.

Preparing the data sets and running the Challenge are labor-intensive activities. The committee along, with numerous students and staff members, worked not just on the synthesis but also on organizing the reviews, judging, identifying the awards, setting up and running the workshop. The value is clear but the future of the challenges relies upon community support in order to continue.

8 Acknowledgments

The Committee acknowledges the US Department of Defense for helping support the VAST Challenge. Members of the committee were also supported in part by the National Science Foundation (0947343 and 0947358).

The Committee also wishes to thank Catherine Plaisant and Manas Desai at the University of Maryland; Adam Roberts, Bill Pike, Jean Scholtz, and John Burnette of Pacific Northwest National Laboratory; and all the students of the University of Massachusetts Lowell, especially John Fallon, Patrick Stickney, Heather Byrne, Baochen Sun, Yen-Fu Luo and Loura Costello.

9 References

[1] VAST Challenge: http://www.cs.umd.edu/hcil/vastchallenge/

[2] Visual Analytics Benchmark Repository: http://hcil.cs.umd.edu/localphp/hcil/vast/archive/

[3] Costello, L., Grinstein, G., Plaisant, C. and Scholtz, J., Advancing User-Centered Evaluation of Visual Analytic Environments through Contests, *Information Visualization* 8 (2009) 230–238

[4] Whiting, M., Haack, J., and Varley, C. 2008. Creating realistic, scenario-based synthetic data for test and evaluation of information analytics software. In *Proc. of BELIV'08*, ACM, New York, NY