VAST 2008 Challenge: Introducing Mini-Challenges

Georges Grinstein[1]
University of Massachusetts Lowell
Catherine Plaisant
University of Maryland

Sharon Laskowski and Theresa O'Connell
National Institute of Standards and Technology

Jean Scholtz and Mark Whiting
Pacific Northwest National Laboratory

ABSTRACT
Visual analytics experts realize that one effective way to push the field forward and to develop metrics for measuring the performance of various visual analytics components is to hold an annual competition. The VAST 2008 Challenge is the third year that such a competition was held in conjunction with the IEEE Visual Analytics Science and Technology (VAST) symposium. The authors restructured the contest format used in 2006 and 2007 to reduce the barriers to participation and offered four mini-challenges and a Grand Challenge. Mini Challenge participants were to use visual analytic tools to explore one of four heterogeneous data collections to analyze specific activities of a fictitious, controversial movement. Questions asked in the Grand Challenge required the participants to synthesize data from all four data sets. In this paper we give a brief overview of the data sets, the tasks, the participation, the judging, and the results.

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

The objectives of the VAST 2008 Challenge [1] remain similar to the objectives of the VAST 2006 and 2007 contests [2,3,4]: to support researchers in their efforts to move visual analytics discoveries and applications into practice through an innovative evaluation forum. These contests and challenges, organized by the authors, also help in developing and testing metrics and evaluation methods for visual analysis environments.

2 VAST 2008 CHALLENGE

The VAST 2008 Challenge was restructured to encourage participation, by providing four mini-challenges in addition to one Grand Challenge. Participation in 2008 overwhelmingly exceeded previous years' numbers (six in 2006; seven in 2007). In 2008 teams from 28 organizations submitted 73 entries to the mini challenges and Grand Challenge. Furthermore, we decided that we would present several awards identifying excellent work, rather than simply determining overall winners.

The VAST 2008 Challenge scenario concerned a fictitious, controversial socio-political movement. Participants were provided with an excerpt from the movement's manifesto and the following four data sets, one for each mini-challenge:

- cell phone records over a 10 day period
- a chronicle of migrant boat journeys with passenger lists, launch and landing sites and landing/interdiction status
- a catalog of wiki edits to a page discussing the movement
- geospatial data of an evacuation from a building in which a bomb exploded.

The National Visualization and Analytics Center (NVAC) Threat Stream Generator project team at Pacific Northwest National Laboratory developed the data sets. Each set was embedded with non-trivially discoverable ground truth [5].

Each mini-challenge consisted of a data set, instructions, and a number of questions to be answered. Participants could enter one or more of the mini-challenges. The Grand Challenge task required participants to pull together information from all four data sets and produce an analysis of the movement's activities based on its beliefs.

2.1 VAST 2008 Challenge Entries

We had six Grand Challenge (GC) entries and 67 mini-challenge entries. The breakdown of entries into the mini-challenges was:

- 22 cell phone entries (CP Mini)
- 13 migrant boat entries (MB Mini)
- 12 wiki edit entries (WE Mini)
- 20 evacuation trace entries (ET Mini)

Twenty eight different organizations from 13 countries submitted entries. Thirteen were student teams.

2.1 Judging

The judges for the challenge consisted of the VAST 2008 Challenge committee members and several professional analysts. All participants received feedback on their entries which will aid these participants in producing better quality entries in 2009. Feedback was divided into four categories:

- accuracy (based on ground truth)
- process descriptions
- analysis
- visualizations that were used to perform the analysis

The ground truth in the datasets enabled us to provide a number of measures of accuracy. We also provided feedback on each team's analyses and visualizations that were used in their process.

No specific awards for the VAST 2008 Challenges were predefined. We decided to provide awards based on aspects of the submissions that the judges felt were noteworthy. Award winners were given the opportunity to submit two-page papers for

[1]grinstein@cs.uml.edu; plaisant@cs.umd.edu; sharon.laskowski@nist.gov; theresa.oconnell@nist.gov; jean.scholtz@pnl.gov; mark.a.whiting@pnl.gov

inclusion in the VAST proceedings. We made only one award per team, thus a team receiving a Grand Challenge award did not get an award for outstanding work in a mini challenge. Hence not all mini challenges have an equal number of awards. Other exemplary work was noted in the feedback provided to the teams.

Three grand challenge entry teams were selected to participate in an interactive session during VisWeek 2008:

- Oculus Info. Inc.
- Palantir Technologies
- Pennsylvania State University – Northeastern Visualization and Analytics Center (NEVAC) Team.

These teams were required to have fully functioning and robust software capable of ingesting a new data set (similar to the Grand Challenge 2008 data set) within 30 minutes. They had the opportunity to work with an analyst for two hours with the goal of solving this new challenge.

The other challenge awards were:

TEAM	AWARD
Beijing University of Posts and Telecommunications	Social Network Accuracy (CP Mini)
Fraunhofer Institute	Tool Integration (ET Mini)
Oculus Info Inc.	Support for Diverse Analytic Techniques (GC)
Palantir Technologies	Interactive Visual Analytic Environment (GC)
Pennsylvania State University	Data Integration (GC)
Southern Illinois University Edwardsville	Innovative Trace Visualization (ET Mini)
SPADAC Inc.	Analysis Summary (MB Mini)
University of Bari	User Testing to Obtain Consensus (ET Mini)
University of California, Davis	Intuitive Social Network Graphs (CP Mini)
University College Dublin	Node-Link Animation (CP Mini)
University of Maryland - SocialAction	Time Visualizations of Cell Phone Activity (CP Mini)
Vision Systems & Technology, Inc.	Effective Toolkit Integration (CP Mini)
VRVis Research Center	Simple and Effective Integrated Display (MB Mini)

3 PARTICIPANT DISCUSSION SESSION

The VAST 2008 Challenge committee organized a discussion session at VisWeek 2008 for the VAST 2008 Challenge participants to discuss the results of the mini-challenges and the Grand Challenge. We anticipate that sharing information about what worked and what did not will help the visual analytic research community to make even more progress in the coming years. We also used this opportunity to involve the participants in planning for next year's Challenge.

4 Lessons Learned

Accuracy becomes more difficult to judge as our tasks become more realistic and the data sets become more complex. For some of the mini-challenges, accuracy was less important than the supporting evidence provided within analyses.

Teams were asked to provide analysis both in the mini-challenges and in the Grand Challenge. An analysis of the situation differs from just reporting the facts. While we acknowledge that many teams did not include or have access to analysts, it is necessary for researchers designing analysts' tools to understand the requirements for producing an analytic report. We intend to provide more references and examples for teams on analysis and what makes a good analytic product. Some of this year's submissions will also serve as examples.

Many submissions described tools developed specifically for the challenge, some possibly developed after extensive study of the data. This is an opportunity for newcomers to become familiar with analytic problems and tasks, and some of these tools were quite innovative.

Reviewing the submissions manually was a monumental task. We were able to automatically score the social network results but clearly more automated evaluation is needed.

5 THE PATH FORWARD

We plan to continue with the challenge format for 2009. The data sets, tasks, and ground truth from the past contests and the 2008 challenge are available to the public [6]. The interactive session data sets are available to educators for class use (please contact us via our website [6]). We will be refining our evaluation criteria during the year with the goal of simplifying the evaluation process.

ACKNOWLEDGMENTS

This work was supported in part by the National Visualization and Analytics Center™ (NVAC™) located at the Pacific Northwest National Laboratory in Richland, WA. The Pacific Northwest National Laboratory is managed for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC05-76RL01830. It was also supported in part by the Intelligence Advanced Research Projects Activity (IARPA). Members of the committee are also supported in part by the National Science Foundation and the National Institute of Standards and Technology.

We also wish to thank Bill Pike, Jereme Haack, Carrie Varley, John Hayes, and Cindy Henderson of Pacific Northwest National Laboratory; the University of Massachusetts Lowell students who worked to process the data and various evaluation computations, especially Adem Albayrak, Loura Costello, and Heather Byrne; and John Grantham at the National Institute of Standards and Technology.

Certain commercial equipment, instruments, materials, services or companies are identified in this paper in order to specify adequately the experimental procedure. This in no way implies endorsement or recommendation by the National Institute of Standards and Technology.

REFERENCES

[1] VAST 2008 Challenge: www.cs.umd.edu/hcil/VASTchallenge08
[2] VAST 2006 Contest: www.cs.umd.edu/hcil/VASTcontest06
[3] VAST 2007 Contest: www.cs.umd.edu/hcil/VASTcontest07
[4] Grinstein, G., Plaisant, C.; Laskowski, S; O'Connell, T; Scholtz, J., and Whiting, M. *VAST 2007 Contest – Blue Iguandon*. The 2007 Visual Analytics Science and Technology Symposium Proceedings, pp. 231-232.
[5] Whiting, M., Haack, J., and Varley, C. 2008. Creating realistic, scenario-based synthetic data for test and evaluation of information analytics software. In Proceedings of BELIV'08 (Florence). ACM, New York, NY
[6] VAST Challenge Portal: http://vac.nist.gov