

VAST 2006 Contest – A Tale of Alderwood

Georges Grinstein¹
University of
Massachusetts Lowell

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

Catherine Plaisant
University of Maryland

Jean Scholtz
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 first Visual Analytics Science and Technology (VAST) contest was held in conjunction with the 2006 IEEE VAST Symposium. The competition entailed the identification of possible political shenanigans in the fictitious town of Alderwood. A synthetic data set was made available as well as tasks. We summarize how we prepared and advertised the contest, developed some initial metrics for evaluation, and selected the winners. The winners were invited to participate at an additional live competition at the symposium to provide them with feedback from senior analysts.

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

Index Terms: H.5.2 [Information Interfaces & Presentations]: User Interfaces – Graphical User Interfaces (GUI)

1. BACKGROUND

There are currently no standard metrics for measuring the performance of visualizations and other visual analytic components. Our community is interested in the development of high level metrics to measure the impact of such tools (Fig. 1).

In order to attribute improvements to visualizations, lower level metrics need to be developed to allow comparison across a number of different types of visualizations. An effective way to investigate many different types of visualization is by running a contest in conjunction with a conference [1]. Asking contest participants to use visualizations to explore a specified data set and perform specified tasks provides samples of visualizations and gives researchers a chance to show their work and get feedback.

As both the visualizations and the analytic product are evaluated for the contest, this also provides feedback for developing and refining the metrics. Some other conference-related competitions or evaluations have been very successful (e.g. the Text REtrieval Conference [2], Knowledge Discovery

and Data Mining Cup [3], the Critical Assessment of Microarray Data Analysis [4], or the IEEE InfoVis contest [5]).

The VAST 2006 contest [6] was held in conjunction with the Visual Analytics Science and Technology (VAST) 2006 International Symposium, October 31-November 2. One of the objectives of the contest was to make the research community aware of realistic tasks and data used in analytic work. Applying research tools to tasks and data could enhance the research team's understanding of the impact their tool could have on the analysis process. The contest was open to all, except for the contest organizers and judges. Both academic and commercial competitors could participate. Contestants could use any existing commercial product or research prototype, and of course could combine tools. Forming teams was strongly encouraged.

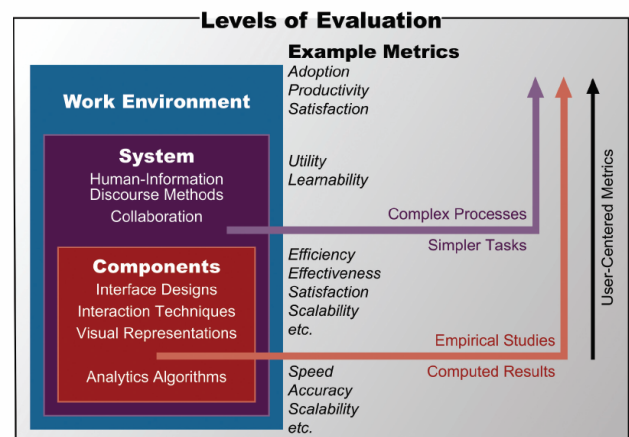


Figure 1: The three levels of evaluation and example metrics (Figure 6.1 in [7])

2. CONTEST MATERIALS

The data set was developed by staff at the Pacific Northwest National Laboratory. It consisted of about 1200 news stories from the Alderwood Daily News (a fictitious town and paper) plus other items collected by the previous fictitious investigators including some photos, maps of Alderwood and vicinity (in bitmap form), some files with other mixed materials (spreadsheet with voter registry information, phone call log) and a few pages of background information. For more information see: <http://www.cs.umd.edu/hcil/VASTcontest06/>. The goal for the contestants was to determine if inappropriate activities were taking place and report on hypotheses and conclusions, including identification of people, places, and events. For each question answered, the relevant documents and other materials from the dataset used to obtain the answers had to be identified and the process used to answer these questions described. An answer template was provided. The data set and tasks were released in February 2006 with a submission deadline of July

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

15th, so participants had several months to prepare their submissions.

3. ADVERTISING AND DISSEMINATION

The announcement was posted on numerous IEEE and ACM sites as well as KDD and European visualization or graphics groups. Personal emails were sent out as well to individuals that managed distribution lists or had an interest in visualization. Last year's IEEE Visualization participant list also received the announcement. Web search engines were provided with key words for quick search returns.

4. SUBMISSIONS AND JUDGING

We received six submissions. All contestants worked hard preparing their entries and all teams will be given certificates of participation. Submissions which included a video describing the tools used and the process of discovery were reviewed by external judges with content and visual analytics expertise. Since we had a synthetic dataset "ground truth" was known which allowed us to compare it with the teams' answers, and to experiment with some quantitative metrics. Scoring was based on the correctness of answers to the questions and the evidence provided, the subjective assessment of the quality of the displays, interactions and support for the analytical process, the clarity of the explanations provided, and the quality of the submitted video.

Two full day judging sessions took place in separate locations. A total of 10 judges reviewed the materials submitted by six teams. Experts in visual analytics, human-computer interaction and visualization participated, as well as professional analysts. The judges reviewed the correctness of the answers, the evidence provided, the quality of the explanations of the process used, and the description of how the tool facilitated the analysis. Videos proved very useful in clarifying the verbal descriptions of the processes used. Each team of judges wrote a summary of their conclusions and suggested awards. Finally, the contest chairs discussed the results and made the final award decisions.

5. WINNERS AND THE LIVE CONTEST

There were 3 winners.

- First place, Corporate Category:
Proulx, P., et al., Oculus Info Inc.
- First Place, Student Category:
Adams, S. and Singhal, K., Georgia Inst. of Tech.
- Second Place, Corporate Category:
Lankenau, R. et al., SSS Research, Inc.

The three teams have been invited to participate in a special interactive contest before the symposium during which professional analysts will interact with the systems and provide feedback. For this live contest (3 hours), a similar but smaller set of materials will be used. Each team will consist of two participants, one of which will be an assigned senior analyst, who will be trained on the team's tool usage. The team will then analyze the new contest data. The correctness of their answers and the number of subplots they are able to locate will be a factor in the overall scoring. The participating analyst and contest judges will assess the process of arriving at the answers during the live contest and use a rating scale for the subjective assessment. The research teams will be allowed to drive the

application to eliminate the analysts' having to learn all the details of how to operate the software.

These three winners deserve special recognition for the effort they undertook to participate in this contest and will receive a awards. They were also invited to present their work during the poster session. Their materials will be posted on the contest website and linked to the Information Visualization Benchmark Repository [8].

6. CONCLUSIONS

We need to develop an approach to facilitate partnering. We thought that this would happen with groups having complementary expertise (e.g. groups with text analysis or reasoning tools might want to seek partners with user interface expertise) but it did not. We had few entries but enough to make the evaluation exercise a valuable experience.

The data sets and results are archived. Our hope is that in the future the visual analytics community can organize a TREC-like series of yearly evaluations or competitions. We believe it is valuable to have both classic contest submissions as well as a live one. The data set complexity could be increased over the years by a combination of adding more data types, providing larger volumes of data, adding uncertainty, increasing deception, and increasing the complexity of scenarios.

We use Thomas and Cook's levels as a guideline for evaluation [6]. The metrics need to be refined and validated. A meta analysis could then be performed to determine any correlations between qualitative and quantitative measures. We hope that future contests will appeal to larger and more diverse audiences than just the visual analytics community and that the competition will encourage partnerships and collaborations.

ACKNOWLEDGMENTS

This work was supported in part by the Disruptive Technology Office and by the National Visualization and Analytics CenterTM (NVACTM) 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.

REFERENCES

- [1] Plaisant, C., Fekete, J., Grinstein, G., Promoting Insight Based Evaluation of Visualizations: From Contest to Benchmark Repository. *HCIL Technical Report # 2004-30*. www.cs.umd.edu/hcil/pubs/tech-reports.shtml
- [2] TREC: <http://trec.nist.gov/>
- [3] KDD Cup: www.acm.org/sigs/sigkdd/kddcup/index.php
- [4] CAMDA: www.camda.duke.edu/camda06
- [5] Infovis 2004 contest : www.cs.umd.edu/hcil/iv04contest
- [6] VAST 2006 Contest: www.cs.umd.edu/hcil/VASTcontest06
- [7] Thomas, J. and Cook, K. (Eds.) *Illuminating the Path: The Research and Development Agenda for Visual Analytics* IEEE CS Press (2005), <http://nvac.pnl.gov/agenda.stm>
- [8] Information Visualization Benchmark Repository: www.cs.umd.edu/hcil/InfovisRepository/