

VAST 2007 Contest – Blue Iguanodon

Georges Grinstein¹
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 second Visual Analytics Science and Technology (VAST) contest was held in conjunction with the 2007 IEEE VAST Symposium. In this contest participants were to use visual analytic tools to explore a large heterogeneous data collection to construct a scenario and find evidence buried in the data of illegal and terrorist activities that were occurring. A synthetic data set was made available as well as tasks. In this paper we describe some of the advances we have made from the first competition held in 2006.

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

We are using the VAST contest to help in developing metrics and evaluation methodologies for visual analysis environments. Competitions of this sort have been useful in other domains such as The Text REtrieval Conference (TREC) [1], Knowledge Discovery and Data Mining Cup [2], the Critical Assessment of Microarray Data Analysis [3], and the IEEE InfoVis contest [4].

The VAST 2006 and 2007 contests [5, 6] were each held in conjunction with the Visual Analytics Science and Technology (VAST) 2006 and 2007 International Symposia.

One of the objectives of the contests was to make the research community aware of realistic tasks and data used in analytic work. The contest requires participants to not only develop or select tools and visualizations to use but to apply these tools to solve an analytic problem, namely finding some illegal or terrorist activity within a collection of multimedia data which for now focused mostly on text.

2. Contest Methodology

The participants were given a data set developed by the NVAC Threat Steam Generator project team at Pacific Northwest

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

National Laboratory. This data set contained:

- about 1500 news stories simulating an online news archive
- two blog excerpts with entries extracted over certain time segments
- a few pictures (in jpg format)
- a few small databases (in XLS and CVS format)
- a few pages of background information (in .DOC or PDF format).

Information concerning the plots was embedded in data of different formats, so contestants had to deal with the heterogeneity to best assemble their stories. The 2007 dataset differed from last year's by including new data types (e.g. blogs, cartoons), several major subplots instead of one, and information gaps, requiring teams to identify places in their analysis where their knowledge was incomplete.

Participants were asked to find the major plots embedded in this data set along with identifying people involved in illegal and/or terrorist activities. Participants were also asked to find the time frame for the activities and to list the various important events. They then were asked to write a debrief that described the situation and make some recommendations for further investigations based on theories developed in the analysis.

We requested a process description, as well as a video, describing how the tools were used in the analysis. Screen shots of different visualizations were to highlight insights provided by the tools.

There were seven entries distributed to the judges and two parallel meetings held over two days in separate locations to evaluate the entries. The judges consisted of experts in visual analytics, human-computer interaction and visualization and professional analysts. The university entries were judged separately from the commercial entries and the evaluations merged through conference calls.

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 various tools facilitated the analysis. Videos proved very useful in clarifying and enhancing the verbal descriptions of the processes. Each team of judges wrote a summary of their conclusions and suggested awards. Lastly, the contest chairs discussed the results and made the final award decisions. Accuracy for the *who*, *what*, *where* and *when* questions was scored based on ground truth in the synthetic data set.

3. DIFFERENCES BETWEEN THE 2006 AND 2007 CONTESTS

First of all, the participants this year had the advantage of viewing the submissions from last year. They also had access to last year's data to practice on until the 2007 contest data was released early in March. We are convinced that this helped as the analysis done by all the teams was much improved over the entries received in 2006. Secondly, we developed more systematic scoring guidelines for this year. We drafted rating criteria for the utility of the system and criteria for the quality of the visualizations. We also provided two versions of the data set.

Participants could choose whether to use raw data or to use preprocessed data (whose entities had already been extracted).

In 2006 we announced the winners and provided participants with the ground truth. For this year's contest, in addition to providing ground truth, each participant received detailed comments on their entries. These comments were based on the accuracy of the answers, the debrief, utility of the system, the quality of the visualizations, and the description of the process. We think that participants will find this feedback useful in refining their systems and in preparing future entries.

4. Lessons Learned

Many of the entries used the process description to tell us about their tool, focusing on describing its functionality. While we did want to understand the tool, we were more interested in the description of the usage of the tool in the context of solving the problem, not just a description of its features. We intend to clarify this and will also encourage participants next year to review this year's winning entries.

We provided two forms of data this year and had intended to use this to provide two categories (if not more) for entries. We had seven submissions this year (up from six in 2006) but we felt that this was not enough to further subdivide the university entries and commercial entries.

The utility and visualization quality ratings were difficult for the analysts to use. On the other hand their comments about the debriefings and the utility of the system were extremely valuable. The committee members focused as well on providing ratings on the quality of the visualizations. These rating criteria were developed during the year and were tested on previous submissions to determine their applicability. This year the participating teams received only comments, but next year we intend to refine our rating criteria and provide quantitative scores as well as qualitative data for the entries.

5. Contest Winners

The University winner for 2007 was Georgia Tech with student members Carsten Görg, Zhicheng Liu, Neel Parekh, Kanupriyah Singhal and faculty advisor John Stasko. The Corporate Division winner for 2007 was Oculus Info whose members included Lynn Chien, Annie Tat, Patricia Enns, Winnifred Kuang, Tom Kapler and Bill Wright.

A great deal of work went into submitting an entry and we feel that it deserves recognition. Whether a winner or not, all participating teams this year have been invited to submit posters to VAST and will have 2 page papers in the proceedings.

However for the winners we provide more extensive visibility. The winners were invited to participate in a closed interactive session at the Symposium. At this time they were given a new, smaller synthetic data set to ingest into their system. They were given two hours with an analyst who worked with them using the system to solve another analytic problem similar to that used in the contest. The winners were invited to contribute to a short journal paper for Computer Graphics and Applications. The winners participated in a panel at the VAST symposium where they discussed their experiences in the contest and in the interactive session.

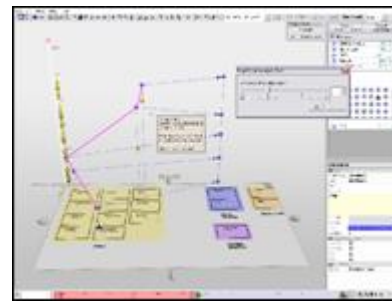


Figure 1. The GeoTime link analysis tool from Oculus

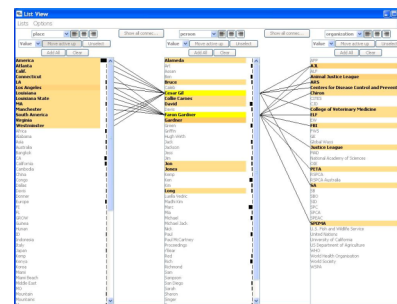


Figure 2. The List view showing connections from Georgia Tech

6. The Path Forward

We have been encouraged to see that many people have downloaded the 2006 and 2007 data sets. The Threat Stream Generator project team is considering various directions to go in developing the 2008 data set. This may include adding more data types, providing larger volumes of data, adding uncertainty, increasing deception, and increasing the complexity of the scenarios. The goal is to appeal to more groups and attract more entries.

We will be refining our evaluation criteria during the year. We will analyze our results to determine correlations between the qualitative and quantitative measures. Our goal is to use the quantitative scores along with the qualitative comments to arrive more rigorously at overall scores for the VAST 2008 contest.

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 Disruptive Technology Office.

We also wish to thank Jereme Haack, Carrie Varley, and Cindy Henderson of the Pacific Northwest National Laboratory Threat Steam Generator project team and the University of Massachusetts Lowell students who worked to process the data and various evaluation computations: Ryan Beaven, Chris Deveau, Curran Kelleher and Christine Mathai.

REFERENCES

- [1] TREC: <http://trec.nist.gov/>
- [2] KDD Cup: www.acm.org/sigs/sigkdd/kddcup/index.php
- [3] CAMDA: www.camda.duke.edu/camda06
- [4] Infovis 2004 contest: www.cs.umd.edu/hcil/iv04contest
- [5] VAST 2006 Contest: www.cs.umd.edu/hcil/VASTcontest06
- [6] VAST 2007 Contest: <http://www.cs.umd.edu/hcil/VASTcontest07>