Visual Analytics at Boeing

Dave Kasik
Senior Technical Fellow
Visualization and Interactive Techniques
Outline

- Overview of analytics
- Introduction to visual analytics
- Boeing’s approach
- Specific examples
- Challenges
Analytics Overview

- Analysis vs. Analytics
- When dealing with digital data, analysis is the detailed examination of any size and number of data collections
- Analytics is the science behind analysis
- In practice, analytics requires understanding the
  - Cognitive processes
  - Data acquisition
  - Tools
  - Techniques & methods
  - Results communication patterns

that let people obtain an optimal or reasonable decision based on existing data.
The Data Age

• Moore’s Law: Transistor capacity doubles every 24 months

• The human brain’s volume doubles only every $3 \times 10^7$ months

• The amount of data stored digitally was estimated at 180 exabytes in 2006.

• By 2011 there will be at least 1800 exabytes (1.8 zetabytes)
Types of Analysis

- Situation awareness
  - Command and control
- Tracking and visibility
  - Determine status
- Causal chain analysis
  - Determine why something happened
- Hypothesis testing
  - Explore possible explanations
- Detecting anomalies and correlations
  - Prevent event occurrences
- Prediction/Forecasting
  - Improve quantities ordered
- What-if studies
  - Explore alternatives
- Summarizing
  - Communicate results
Tools and Techniques

- Numerous tools and techniques
  - Statistical analysis
  - Text and data mining
  - Control rooms
  - Operations centers
  - Dashboards
  - Alerting
  - Fish-bone charts
  - Matrix analysis
  - Neural nets
  - Competing hypotheses
  - Modeling & simulation
  - Pareto analysis
  - …

- Visual analytics is a new way of analysis
Variations of Non-Geometric Data Visualization

- Automated Reports
  - [Cognos, Crystal Reports, Business Objects]
- Information Visualization
  - [Tufte, ManyEyes]
- Excel, PowerPivot
  - [Pivot tables, Histograms]
- Visual Analytics
  - [Active data exploration, Highly interactive]

  Best on numeric data
  - [Statistics, Clustering]

  Numeric
  - Text
  - Streaming Video
Automated Reports

Images produced with Cognos software
Information Visualization

Napoleon’s March to Moscow

John Snow, Cholera, 1854
Excel

- In a class by itself as the most widely available producer of visual reports

Score Summary

Categories

Count of Communication

Communication

Program

Utilization Level

1 - 5
1 - 4
1 - 3
1 - 2
1 - 1
Visual Analytics
Focus on Visual Analytics

The science of analytical reasoning facilitated by interactive visual interfaces

- **Goals**
  - Design interactive visual interfaces that allow innate “visual intelligence” to find meaningful patterns in datasets
  - Synthesize information and derive insight from massive, dynamic and conflicting data

- **Analytic foundation**
  - *Illuminating the Path*, nvac.pnl.gov (free download), 2005

Detect the expected and discover the unexpected
Corporate Players

- **Startups trying to make $$**
  - Tableau Software, Spotfire, QlikView, Centrifuge
    - Best for numeric data
  - Starlight (Future Point Systems)
    - OK for both numeric and text
  - Geotime (Occulus)
    - Time dependent geo-located events
  - Analysts’ Notebook
    - Tracks analysis process
  - IN-SPIRE (Pacific Northwest National Labs)
    - Excellent for text and hypothesis testing
- **Thomson Data Analyzer**
  - Focuses on text in intellectual property
- **Other startups and research systems**

- **Larger companies starting to notice**
  - IBM (Cognos, ManyEyes, SPSS)
  - SAS Institute
  - Oracle (BusinessObjects)
  - Microsoft (Sharepoint 2010 includes PowerPivot)
  - Others
What Is Boeing Doing?

- **Fund directed university research**
  - Offset-funded projects at
    - Simon Fraser University (Vancouver BC)
    - University of British Columbia (Vancouver BC)
    - Dalhousie University (Halifax NS)
  - Joint Brazil – Canada – Boeing project
    - Mobility and natural user interfaces applied to visual analytics
  - Strategic universities
    - Stanford (Jeff Heer)
    - University of Illinois, Urbana-Champaign (Jiawei Han)

- **Push the state-of-the-art internally**
  - Text mining (Anne Kao, BR&T)
  - Advanced analytics study group

- **Transition the technology**
  - Connect with IT implementers
  - Have trained analysts work directly with people who have problems
    - Internal resources for sensitive data
    - Pair interns with Boeing SME’s
Bird Strike Project

- The threat by the numbers:
  - Approximately 1 bird strike per 2,000 flights
  - Around 20 strikes per day on jet transport category aircraft alone
  - About 1 in 10 strikes are damaging
  - The reported costs average $123 million per year
  - However, 80% of strikes go unreported, and the true cost could be as high as $615 million per year

- Other factors:
  - Increasing traffic
  - Bigger, quieter engines
  - Twin-engine configurations
  - Increasing bird populations
Time and Location Analysis: Tableau
Computing System Performance

- Recommend for Everett and Charleston choices between:
  - Panasonic C1 vs. TechPC Laptop
    - Task Performance
    - 3D Interaction Performance
  - Wired vs. Wireless
  - Thick vs. Thin
    - Login to Velocity and Thin Client Performance
    - Task Performance
    - 3D Interaction Performance
### Overall Everett vs. Charleston, Wired - Wireless

**Observations:**
- Server variation causes Labor-on spike (last test, orange)
- NOTE: BP18 gets rid of this problem.

<table>
<thead>
<tr>
<th>Test Case</th>
<th>SOI</th>
<th>Location</th>
<th>Connection</th>
<th>Avg Time-secs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>40-26</td>
<td>88-30</td>
<td>Wired</td>
<td>200</td>
</tr>
<tr>
<td>7.6</td>
<td>40-26</td>
<td>88-30</td>
<td>Wirelesss</td>
<td>180</td>
</tr>
<tr>
<td>21.0</td>
<td>40-26</td>
<td>88-30</td>
<td>Wired</td>
<td>160</td>
</tr>
<tr>
<td>31.7</td>
<td>40-26</td>
<td>88-30</td>
<td>Wireless</td>
<td>140</td>
</tr>
<tr>
<td>66.6</td>
<td>40-26</td>
<td>88-30</td>
<td>Wired</td>
<td>120</td>
</tr>
</tbody>
</table>

- Charleston wired and wireless essentially equivalent
- Everett wireless consistently 25% faster than Charleston wireless overall, BUT
Thick vs. Thin – Task Average for all SOIs

- XenApp poor 3D rotation performance for two largest SOIs.
- HPRGS and XenDesktop OK rotation performance for all and preferred.
- Dedicated blades ~30% quicker than XenApp, tablet/TechPC for 4 key tests

Observations:
- Server load (Labor-on) caused red variation, esp. for 66.6 test
  - NOTE: BP18 gets rid of this problem.
- XenDesktop makes relative performance between Charleston and Everett closer because of smaller network footprint
  - HPRGS ~15-20% faster in Everett
  - XenDesktop ~10% faster in Everett
Visual Analysis Lessons Learned

- Summary: analysis is not for everyone

1. External Data Sources
   - Getting access
   - Determining efficient access methods

2. Search for Information

3. Structure Loop

4. Shoebox
   - Keeping the shoebox current

5. Read & Extract

6. Search for Relations

7. Evidence File

8. Schematize

9. Search for Evidence

10. Schema

11. Build Case

12. Search for Support

13. Hypothesis
   - Identifying appropriate tools
   - Understanding data values
   - Posing questions & problems

14. Tell Story

15. Reevaluate

16. Presentation
   - Communicating results

- Summary: analysis is not for everyone
Collaboration Lessons Learned

- University relationships are mutually beneficial
  - Challenging real world problems for research
  - Fast path to technology transition as well as research papers
  - Job possibilities for students
Summary

- Develop a better understanding of analytic tasks
- Leads to better selection of proper tools and techniques
- Moving advanced technology into practice is a contact sport