SparkClouds: Visualizing Trends in Tag Clouds

Bongshin Lee, Nathalie Henry Riche, Amy K. Karison and Sheelagh Carpendale

Presented by

Weiya Chen

02 February 2012
Table

- Introduction
- Designing sparkClouds
- Controlled experiment
- Results
- Conclusion
Introduction

Tag cloud

→ Tag frequency by font size
→ Data source evolve

State of the Union Address, 2002 vs. 2011

President Bush, January 29, 2002

President Obama, January 25, 2011
Show trends?

sparkline

sparkClouds = tag clouds + sparkline
Designing sparkClouds

• Tag clouds as used in sparkClouds
  → Font size Encoding
  → Layout

• Sparklines
  → Data Encoding
  → Representing Zero
• Unifying Tag and Sparkline

→ Alignment
→ Overlays
→ Mirror
→ Foreground and background colors
Controlled experiment

• Test 4 visualizations across 6 different tasks
  SparkClouds
  Multi Line Graph
  Stacked Bar Chart
  Parallel Tag Cloud

• Datasets
  75 most frequent words from 5 books
  12 time points
  four options chosen from top 25 words

• Measurement
  Task completion time
  Accuracy
  participants’ subjective preferences
• Multi Line Graph
• Stacked Bar Chart
• Parallel Tag Cloud (PTC)
<table>
<thead>
<tr>
<th>#</th>
<th>Type</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Specific Data</td>
<td>What topic was ranked as the most frequent topic (top-ranked) during month X?</td>
</tr>
<tr>
<td>2</td>
<td>Specific Data</td>
<td>During which month did topic &quot;A&quot; occur most frequently?</td>
</tr>
<tr>
<td>3</td>
<td>Topic Trends</td>
<td>Which of the following topics experienced the longest continuous increase/decrease in rank during the year (e.g., for the most months in a row)?</td>
</tr>
<tr>
<td>4</td>
<td>Topic Trends</td>
<td>Which of the following topics experienced the largest increase/decrease in rank from month Y to month Z?</td>
</tr>
<tr>
<td>5</td>
<td>Overview</td>
<td>Which of the following topics was the most frequent topic (top-ranked) for the greatest number of months during the year?</td>
</tr>
<tr>
<td>6</td>
<td>Overview</td>
<td>Which of the following topics was ranked among the top 25 topics most consistently over the year (e.g., was ranked at all during the most months)?</td>
</tr>
</tbody>
</table>
Results

• Mean overall task time for each visualization

![Bar chart showing task time in seconds for different visualizations](chart.png)
• Mean task completion time by visualization for each task
• **Errors**

![Bar chart showing the total number of errors for different tasks and categories. The chart includes categories such as SparkCloud, MultiLine, ParallelCloud, and StackedBar. Each category is divided into subcategories represented by different colors: T1, T2, T3, T4, T5, and T6. The y-axis represents the total number of errors ranging from 0 to 14.](image-url)
• Subjective preferences
Conclusion

• SparkClouds inherit advantages from both sparklines and tag clouds
  → provide people with good overall trends
  → using little additional space
  → compact and aesthetically pleasing layout

• Future work
  → importance of stability
  → compromise between stability and visual clutter
Questions