MatrixWave: Visual Comparison of Event Sequence Data

Jian Zhao
Leo Zhicheng Liu
Mira Dontcheva
Aaron Hertzmann
Alan Wilson
Events

registration

ticket

flight
Event Sequence Data

- Personal life activities
- Webserver logs
- Usability study logs

...
Webserver Logs

**homepage**

**program**

**workshop**

**homepage**

**housing**

**google search**

**registration**
Questions

• How many users visit a particular page?
• How soon do users leave the website?
• What paths do users mainly follow?
• ...
Sankey Diagram

Step 1

Step 2

Step 3

Step 4

Step 5

Webpage A

Webpage B

Webpage A to Webpage B

[Source: Google Analytics]
[Source: Google Analytics]
Problems

Visually encode large and dense event sequence data

Visually compare two related event sequence datasets
A → D → B
A → B → A → D
A → B → B → B
B → B
C → A → B → C

step 1  step 2

step 3  step 4
A → D → B
A → B → A → D
A → B → B → B
B → B
C → A → B → C

step 1

step 2

A
B
C
D
A → D → B
A → B → A → D
A → B → B → B
B → B
C → A → B → C
A → D → B
A → B → A → D
A → B → B → B
B → B
C → A → B → C
Few visits to this page  Many visits to this page
A → D → B
A → B → A → D
A → B → B → B
B → B → X
C → A → B → C

step 1

step 2

step 3

step 4

Drop-off link
5 steps, ~20 nodes, ~40 links
5 steps, ~100 nodes, ~250 links
Problems

Visually encode large and dense event sequence data

Visually compare two related event sequence datasets
Visual Comparison

Sensor X

Sensor Y

juxtaposition (side by side)

[Glazer et al. 2011]
Visual Comparison

juxtaposition (side by side)  superposition (overlay)

[Glazer et al. 2011]
Visual Comparison

- Juxtaposition (side by side)
- Superposition (overlay)
- Explicit encoding (difference)

[Glazer et al. 2011]
Juxtaposition in MatrixWave

Each node/link
- Value in data 1
- Value in data 2
- Change between them
Superposition and Explicit Encoding

Few visits in the second dataset -100% 100%  More visits in the second dataset

Few visits to this page

Few visitors traversed this link

Many visits to this page

Many visitors traversed this link
Superposition and Explicit Encoding

Few visits in the second dataset  -100% 100%  More visits in the second dataset

Few visits to this page  Many visits to this page

Few visitors traversed this link  Many visitors traversed this link
Demo
Design Alternatives - Nodes
User Study

Adapted Sankey

MatrixWave
User Study

• 12 participants
  – 8 males and 4 females

• 8 comparison tasks
  – 4 types: node, link, step, and path

• 2 techniques
  – Sankey and MatrixWave

• Dataset
  – first 6 steps of the 1000 most common event sequences on two different days on Adobe.com
Results (Node)

Node comparison tasks

- Task 1: *which page has larger volume?*
- Task 2: *which page has greater change?*
Results (Link)

Link comparison tasks

- Task 3: *which page has more increasing links?*
- Task 4: *which link has the largest volume?*

Accuray

- MatrixWave
- Sankey

Completion time (seconds)
Results (Step)

Step comparison tasks

- Task 5: *is the overall traffic increasing?*
- Task 6: *which step has a larger overall volume?*
Results (Path)

Path comparison tasks
- Task 7: which path has the largest volume?
- Task 8: is the overall path traffic increasing?

Accuracy
- MatrixWave
- Sankey

Completion time (seconds)
Results

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>MatrixWave</th>
<th>Sankey</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Completion time (seconds)

<table>
<thead>
<tr>
<th></th>
<th>MatrixWave</th>
<th>Sankey</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
User Preference

- **strongly prefer MatrixWave**
- **prefer MatrixWave**
- **neutral**
- **prefer Sankey**
- **strongly prefer Sankey**

(number of participants)
Future Work

• Further evaluation with domain experts in realistic work settings
• Comparison of multiple (>2) datasets
• Visualization of temporal information