Overview

- What is a trail?
- How do we get trail data?
  - Characterize trail as network data
- Trails and Loom
  - Visualization
  - Networks from trails
  - Finding similar trails
What is a Trail?

• A **trail** is a **trace** of the **movement** of something over time

• For example, the movement of an attachment through a series of email communications creates a **trail**

• What are some other examples of trails?
  – People moving from place to place – geospatial trails
  – Twitter hashtags
  – ...

Geospatial Trails

• Usually geospatial trails represent agents travelling in continuous space and time.

• Network data: discrete node and discrete time.

[Diagram showing continuous space vs discrete location node, and continuous time vs discrete time with labels 'Aggregate', 'Slice', and 'vs']
Geospatial Trails

<table>
<thead>
<tr>
<th>Time</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017, June 7, 9 am</td>
<td>Green St.</td>
</tr>
<tr>
<td>2017, June 7, 10 am</td>
<td>Design District</td>
</tr>
<tr>
<td>2017, June 7, 11 am</td>
<td>Chinatown Gate</td>
</tr>
<tr>
<td>2017, June 7, 12 am</td>
<td>16 th st.</td>
</tr>
</tbody>
</table>

Trails visualization

- **ORA Over-time visualizer**
  - Benefit: Can see changes in network structure over time
  - Drawback: For sparse trail data, not very effective

- **ORA GIS Visualizer**
  - Benefit: Can see the spatial distribution of trails
  - Drawback: Lose the temporal information

- **Loom**
  - Benefit: Can see the temporal distribution and the places travelled to
  - Drawback: Spatial distances, where they exist, are not preserved
What we’ll do

- Import a “DynamicMetaNetwork” with spatial information
- Visualization
  - Understand the benefits and drawbacks of different visualizations of trail data
    - ORA Over-time visualizer
    - ORA GIS visualizer
    - Loom
- Finding Similar trails
  - Use Loom to cluster trails
- Obtain networks from trails

Import a dynamic meta-network

- Same as importing a regular meta-network
  - Drag-and-drop
  - File->Open Meta Network
- Import TrailsDataset.xml
Importing

The Data

- Our trail:
  - Locations are our nodes
  - Agents are what is moving between them

- Lets explore the data
  - In ORA
  - Networks over time visualizer
  - Geospatial visualizer
ORA Main Window

Networks Over Time Visualizer
Networks Over Time Visualizer

Geospatial Visualizer
Geospatial Visualizer

[Diagram of Geospatial Visualizer interface]

Geospatial Visualizer

[Diagram of Geospatial Visualizer interface]
Geospatial Visualizer

[Image of Geospatial Visualizer interface]

Geospatial Visualizer

[Image of Geospatial Visualizer interface]
• **Visualization** over time is hard
  - State of the art revolves around animation
  - Loom allows us to visualize trails over time in a static, understandable environment

• Trails may have similar **patterns**, but these are difficult to observe
  - Loom allows us to cluster similar trails together

• We can get **networks from trails**, for example, who is connected by the given attachment?
  - Loom allows us to easily export such networks to ORA
What we’ll do

• Import a “DynamicMetaNetwork” with spatial information

• Visualization
  • Understand the benefits and drawbacks of different visualizations of trail data
    • ORA Over-time visualizer
    • ORA GIS visualizer
    • Loom

• Finding Similar trails
  • Use Loom to cluster trails

• Obtain networks from trails

Why cluster?

• Why are we interested in trails and trail clustering?
  • Gain information by analyzing agents across space and time together.
  • Interested in grouping agents that display same behavior across time. E.g. visit the same locations across time.
Feature vector representation using PFSA

\[ \beta \alpha \alpha \beta \alpha \beta \alpha \beta \alpha .... \]

\[ \begin{array}{c}
\alpha \\
\beta \\
\alpha \beta \\
\beta \alpha \\
\beta \beta \\
\end{array} \]

Depth = 1

\[ \begin{array}{c}
\alpha \alpha \\
\alpha \beta \\
\beta \alpha \\
\beta \beta \\
\beta \alpha \alpha \\
\beta \alpha \beta \\
\beta \beta \alpha \\
\beta \beta \beta \\
\end{array} \]

Depth = 2

\[ \begin{pmatrix}
\pi_{11} & \pi_{12} & 0 & 0 \\
0 & 0 & \pi_{23} & \pi_{24} \\
\pi_{31} & \pi_{32} & 0 & 0 \\
0 & 0 & \pi_{43} & \pi_{44} \\
\end{pmatrix} \]

State Transition matrix

State Probability Vector

Clustering of Trails using PFSA

- Each trail is now represented by a numerical feature vector, the state probability vector of the derived PFSA (the model of the generative process).
- To look at joint spatiotemporal behavior we now cluster the agent trails based on their feature vectors.
- This is done using a two step process.
  - A coarse clustering step: Trails are initially grouped coarsely according to the locations visited, irrespective of the frequency of the visits.
  - A cluster refining step: The coarse clusters are each then clustered using agglomerative clustering to derive groups of trails which visit “similar” locations with “similar” frequencies.
Refining the Coarse Clustering

Depth = 1

Depth = 2

Depth = 3

Viewing time sequences

- Each cluster contains trails with similar patterns in the sequences of locations visited
- Thus extract the longest common subsequence amongst all the trails belonging to a cluster.
What we’ll do

• Import a “DynamicMetaNetwork” with spatial information
• Understand the benefits and drawbacks of different visualizations of trail data
  – ORA Over-time visualizer
  – ORA GIS-visualizer
  – Loom
• Use Loom to cluster similar trails
  – The high level concept
  – The details
• Obtain networks from trails

Generating Networks from Trails

• We can better understand how different cities relate via championships by getting networks out of them

What we’ll do
• Generate the networks
• View them in ORA
• Use ORA Network Visualizer
Exporting the Matrices

What we now have

- ORA uses all of the trails and outputs a single meta-network
  - **Colocation** – An edge is created between the trophies if they ever existed at the same place at the same time
  - **Visit Matrix** – An edge is created between city and trophy if the city ever won that trophy
  - **Transition** – An edge is created between cities if a trophy ever traveled from one to the other in consecutive years
Summary

- We discussed what a trail was – a trace of the movement of something through a network over time
- We used an example dataset and looked at trail data three different ways – in the Networks Over Time visualizer, the GIS visualizer and Loom
- We talked about how to find similar trails in Loom
- We looked at how we can get new, interested networks out of our trail data