Sequential pattern mining for analyzing visitor trajectories
Nyoman Juniarta, Miguel Couceiro, Amedeo Napoli, Chedy Raïssi

To cite this version:
Nyoman Juniarta, Miguel Couceiro, Amedeo Napoli, Chedy Raïssi. Sequential pattern mining for analyzing visitor trajectories. ISWS 2018 - International Semantic Web Research Summer School 2018, Jul 2018, Bertinoro, Italy. hal-01890429

HAL Id: hal-01890429
https://inria.hal.science/hal-01890429
Submitted on 8 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
1. Introduction

Hecht Museum is an archaeological museum in Haifa, Israel.

In this work, we present our work on the grouping and mining of trajectory patterns of visitors in this museum.

2. Dataset and visitor behaviors

Within the framework of CrossCult project, we are working on a dataset of the trajectories of 254 visitors in Hecht Museum. Each trajectory contains a list of visited items. An example is illustrated in the table below.

Table I. An example of one visitor trajectory

<table>
<thead>
<tr>
<th>Start</th>
<th>Finish</th>
<th>Item</th>
<th>Room</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:10:09</td>
<td>14:11:14</td>
<td>102</td>
<td>1</td>
</tr>
<tr>
<td>14:12:45</td>
<td>14:20:19</td>
<td>402</td>
<td>4</td>
</tr>
<tr>
<td>14:22:10</td>
<td>14:25:42</td>
<td>407</td>
<td>4</td>
</tr>
</tbody>
</table>

Based on his/her movement, a visitor can be grouped as one of four defined behaviors: ant, grasshopper, butterfly, and fish, as illustrated in the figures below.

3. Workflow

Dataset of visitor paths → Grouping/clustering → Sequential pattern minings → Labeled dataset → Finding JEPs for each label → Mapping labels into behaviors

4. Clustering

Each trajectory is modeled as a sequence of items.

The trajectory in Table I becomes:

(102, 202, 301, 402) → (102, 202, 301)

Distance between any two sequences is measured by simACS [1].

5. Sequential pattern mining

We apply two algorithms to obtain two lattices:

- **MFCS** (mining frequent contiguous sequences) [2]
  
  \[(102, 202, 301, 402) \rightarrow (102, 202, 301)\]

- **MRGS** (mining rare general sequences) [3]
  
  \[(102, 202, 301, 402) \rightarrow (102, 202, 402)\]

6. Jumping emerging pattern (JEP)

From the two lattices, there are some concepts whose extent contains visitors from the same label. The intent of such concepts is a JEP.

7. Conclusion

Our results highlight some interesting patterns that may define visitor behaviors. This can help museum researchers to analyze and evaluate the placement of items and the visiting styles. Furthermore, these patterns can be analyzed to build a recommendation system for future visitors.

8. References

