Sequential pattern mining for analyzing visitor trajectories

Nyoman Juniarta, Miguel Couceiro, Amedeo Napoli, Chedy Raïssi

Sequential pattern mining for analyzing visitor trajectories

Nyoman Juniarta, Miguel Couceiro, Amedeo Napoli, and Chedy Raïssi
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
nyoman.juniarta@loria.fr

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.

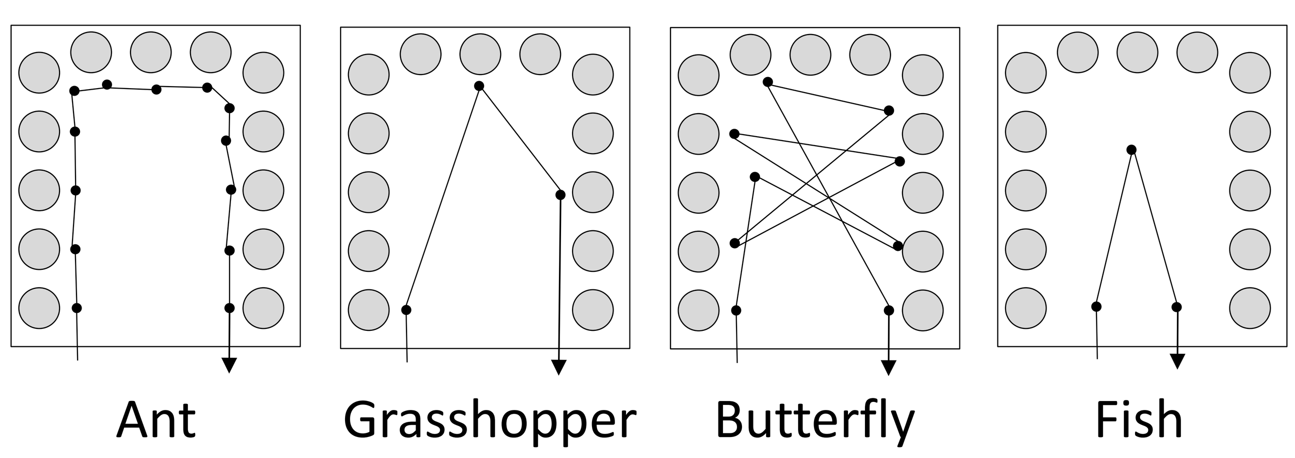
2. Dataset and visitor behaviors

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.

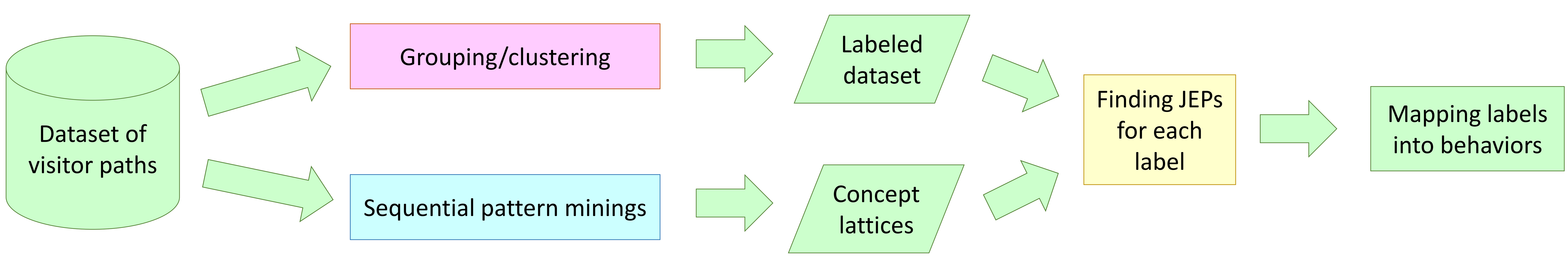
Table I. An example of one visitor trajectory

Start	Finish	Item	Room
14:10:09	14:11:14	102	1
14:12:45	14:20:19	402	4
14:22:10	14:25:42	407	4

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



4. Clustering

Each trajectory is modeled as a **sequence of items**.

The trajectory in Table I becomes:
(102, 402, 407)

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

5. Sequential pattern mining

We apply two algorithms to obtain two lattices:

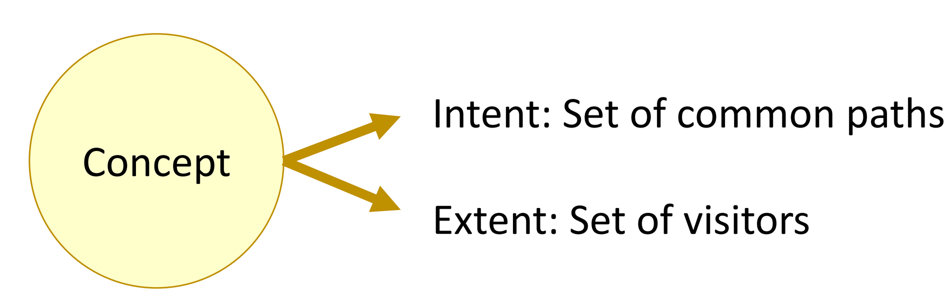
- MFCS (mining frequent **contiguous** sequences) [2]

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

- MRGS (mining rare **general** sequences) [3]

  (**102, 202**, 301, **402**) → (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.

6. Results

Patterns	Count
(1, 1, 1)	33
(1, 7)	13
(1, 1)	66

Patterns	Count
(1, 3)	38
(3, 1)	9
(4, 7)	31
(7, 4)	11

	Visitors	Patterns	Cluster	Behavior
MFCS	70, 107, 121, 133, 201, 202	(1, 1, 402)	A	ant
MFCS	103, 165, 188	(701, 707)	B	ant
MFCS	4, 8, 32	(101, 102, 101)	C	ant/butterfly
MFCS	46, 47	(101, 602)	D	grasshopper
MFCS	89, 163	(602, 203)	D	grasshopper
MRGS	71, 79	(701, 504)	D	grasshopper
MRGS	97, 98	(701, 406)	D	grasshopper
			E	fish

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

[1] E. Egho et al., 2015, On measuring similarity for sequences of itemsets. Data Mining and Knowledge Discovery 29(3), 732–764.
[2] A. Buzmakov et al., 2016, On mining complex sequential data by means of FCA and pattern structures. International Journal of General Systems 45(2), 135–159.
[3] V. Codocedo et al., 2017, A proposition for sequence mining using pattern structures. Proceedings of ICFCA, 106–121.