JAMES COOK UNIVERSITY
AUSTRALIA

Spatio-Temporal Data Mining
&
Trajectory Data Mining

Reading

JAMES COOK UNIVERSITY AUSTRALIA

- Di Wang, Tomio Miwa and Takayuki Morikawa (2020). Big Trajectory Data Mining: A Survey of Methods, Applications, and Services https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7472055/

- Yu Zheng (2015) Trajectory Data Mining: An Overview, ACM Transactions on Intelligent Systems and Technology, 6(3): 29

- Microsoft 2007. https://www.microsoft.com/en-us/research/project/trajectory-data-mining/

Outline

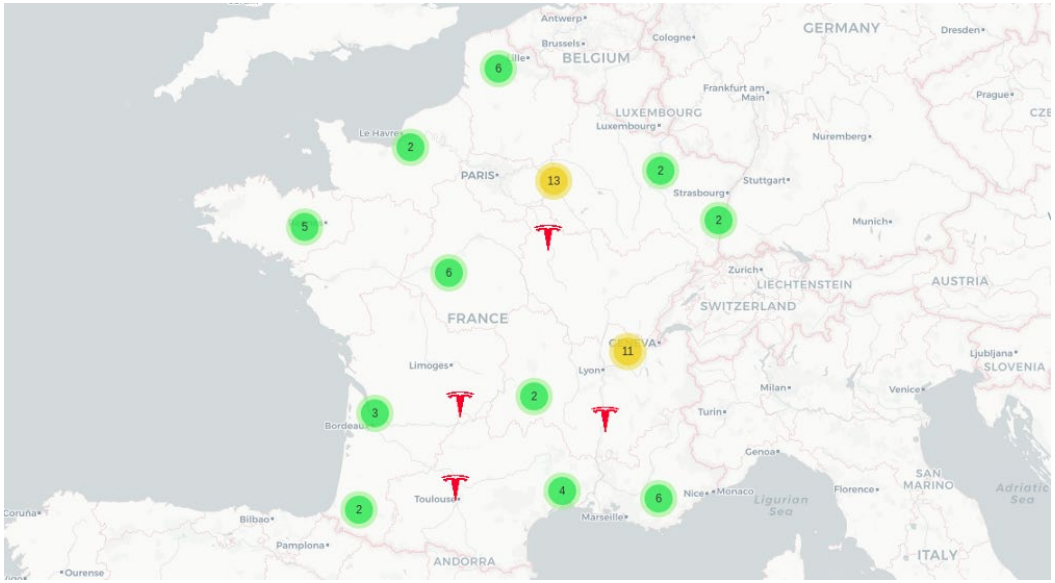
JAMES COOK UNIVERSITY AUSTRALIA

- Geospatial/ Spatial-temporal data mining
  - Definitions, techniques & use cases
- Trajectory data mining

What is Spatial?

JAMES COOK UNIVERSITY AUSTRALIA

- Relating to or existing in space only
- Take a look at https://blog.locale.ai/

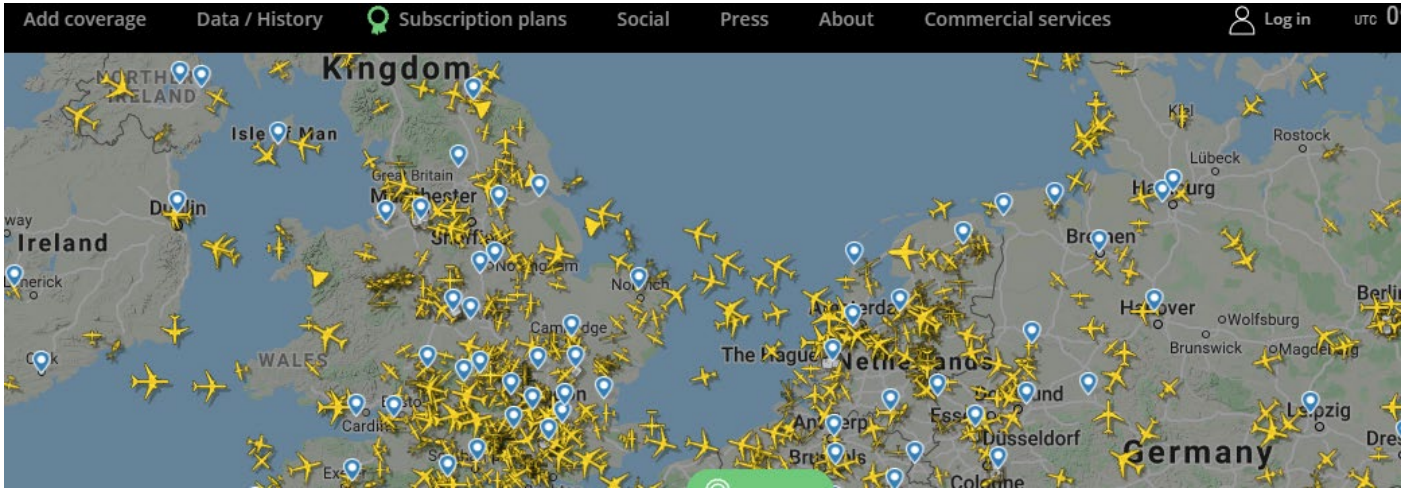


What is Spatio-Temporal?

JAMES COOK UNIVERSITY AUSTRALIA

- Relating to *both space and time*
- Spatial (location) and temporal (time) attached
- Changes and movements over time

https://www.flightradar24.com/51.5,-0.12/6



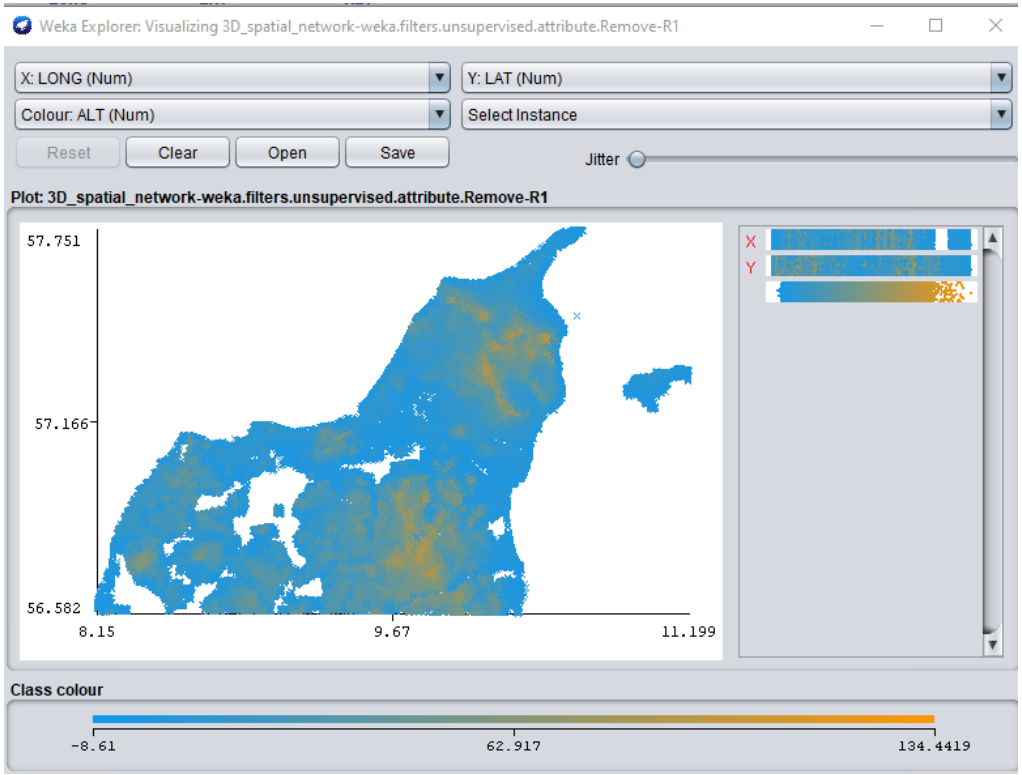
Geo-Spatial data

JAMES COOK UNIVERSITY AUSTRALIA

- An eg. of the data

VOLCANX020	NAME	LOCATION	LATITUDE	LONGITUDE
509	Baker	US-Washington	48.7767982	-121.8109970
511	GlacierPeak	US-Washington	48.1118011	-121.1110001
513	Rainier	US-Washington	46.8698006	-121.7509995
515	St.Helens	US-Washington	46.1997986	-122.1809998

Geospatial data

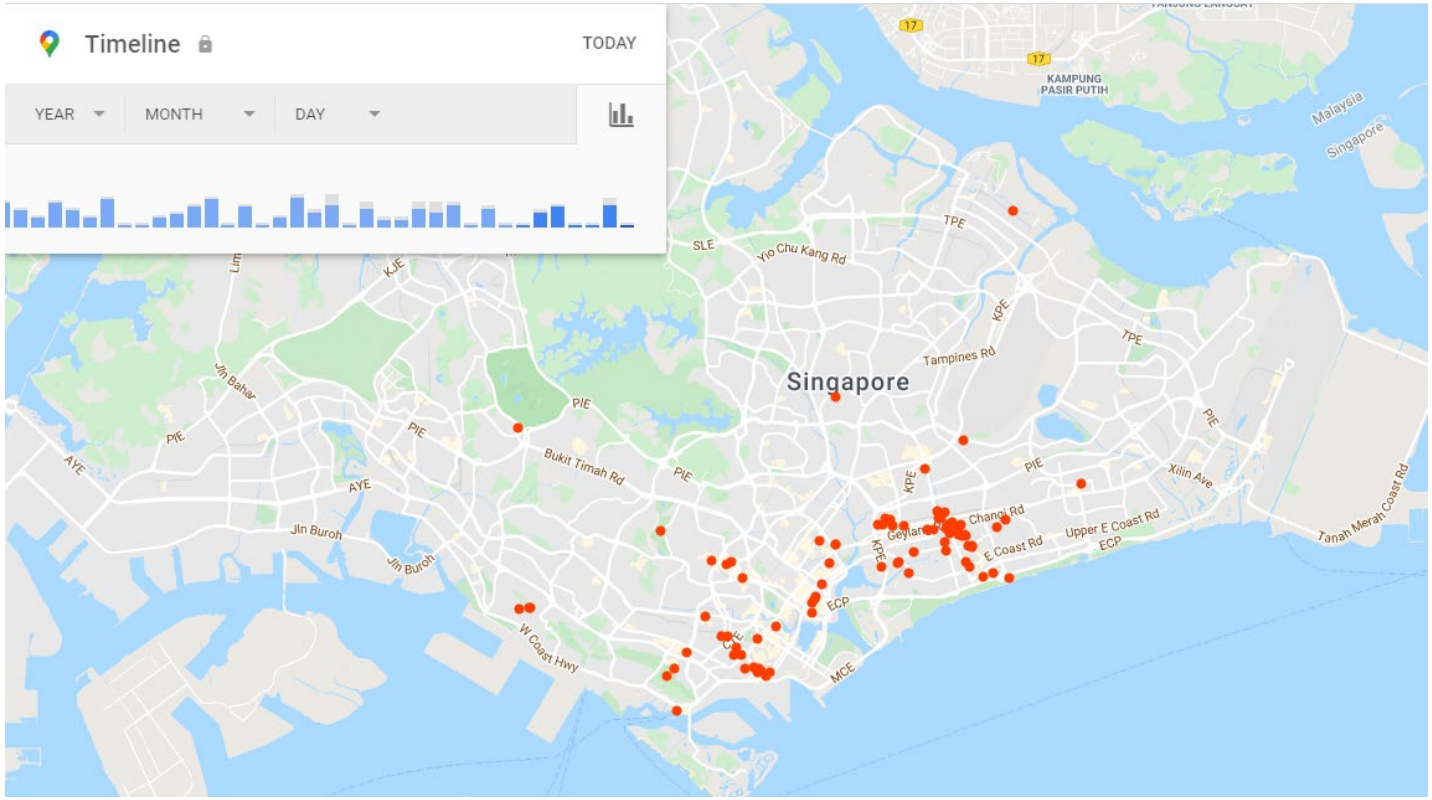


We have looked at it before!
3d_spatial.arff in Lab 4.1

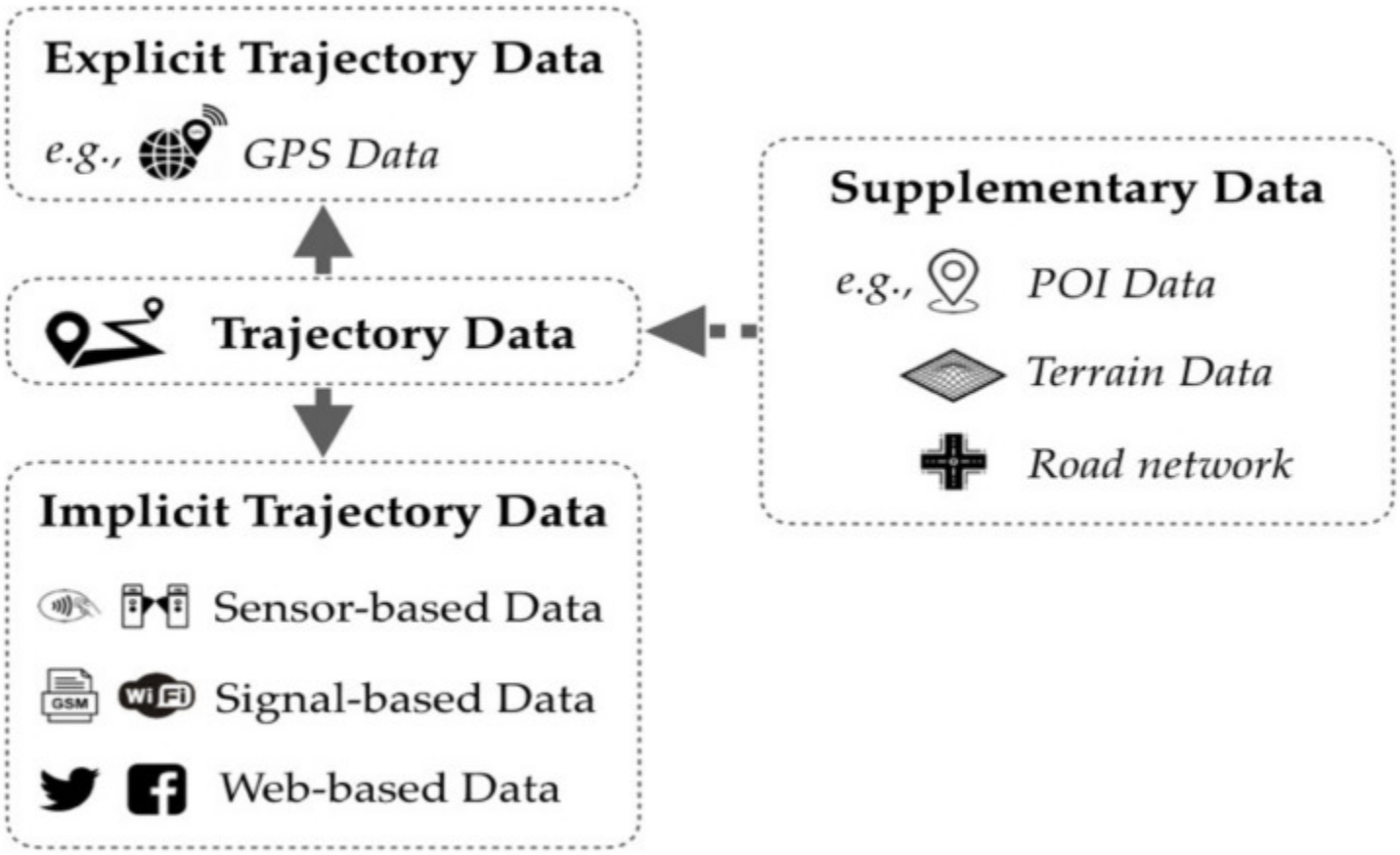
Spatial-temporal Data

JAMES COOK UNIVERSITY AUSTRALIA

- Best example: google maps timeline



Data sources



Use of Spatio-Temporal Mining

JAMES COOK UNIVERSITY AUSTRALIA

- Space and time are ubiquitous aspects of reality
- We are living in a space with time dimension…
- Thus basically all human (things) related data are spatio-temporal in nature
- Advances in automatic (semi-automatic) data generators (sensors, RFID tags, GPS receivers, mobiles etc) result in MASSIVE spatio-temporal data
- It is believed that more than 95% of business data are spatial or spatio-temporal

Trajectory Data Mining

JAMES COOK UNIVERSITY AUSTRALIA

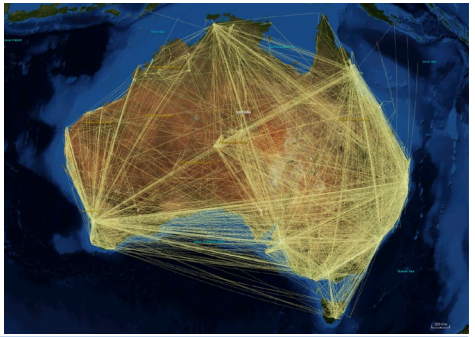
- Geospatial -> space only
- Geospatial temporal -> space and time
  - *Trajectory data mining* => + ***Movement***
  - "a trace generated by a moving object within a certain spatiotemporal context and is generally represented by a series of chronologically ordered points." (Zhang 2014)

Trajectory Examples

JAMES COOK UNIVERSITY AUSTRALIA

- Vehicle trajectories (cars, buses, trucks etc)
- Animal movements (birds, sharks etc)
- People movements (tourists, photo-takers, students etc)
- Mouse click movements (HCI, software design etc)

Can you name other egs?



Understanding Movements

JAMES COOK UNIVERSITY AUSTRALIA

- Animal movements
  - Cows frequent visits to shades, but rare visits to grazing areas => indication of sickness?
  - Bees periodic visits to hive from flowers => useful for beekeeping
- Human movements
  - Frequent visits to fast food restaurants but rare visits to gyms/parks/beaches => indication of health risk
  - Frequent visits to Indian/Korean/Japanese restaurants => Asian?

Moving Objects

JAMES COOK UNIVERSITY
AUSTRALIA

Questions to ask: Where, when, why, what?

Raw GPS Trajectory

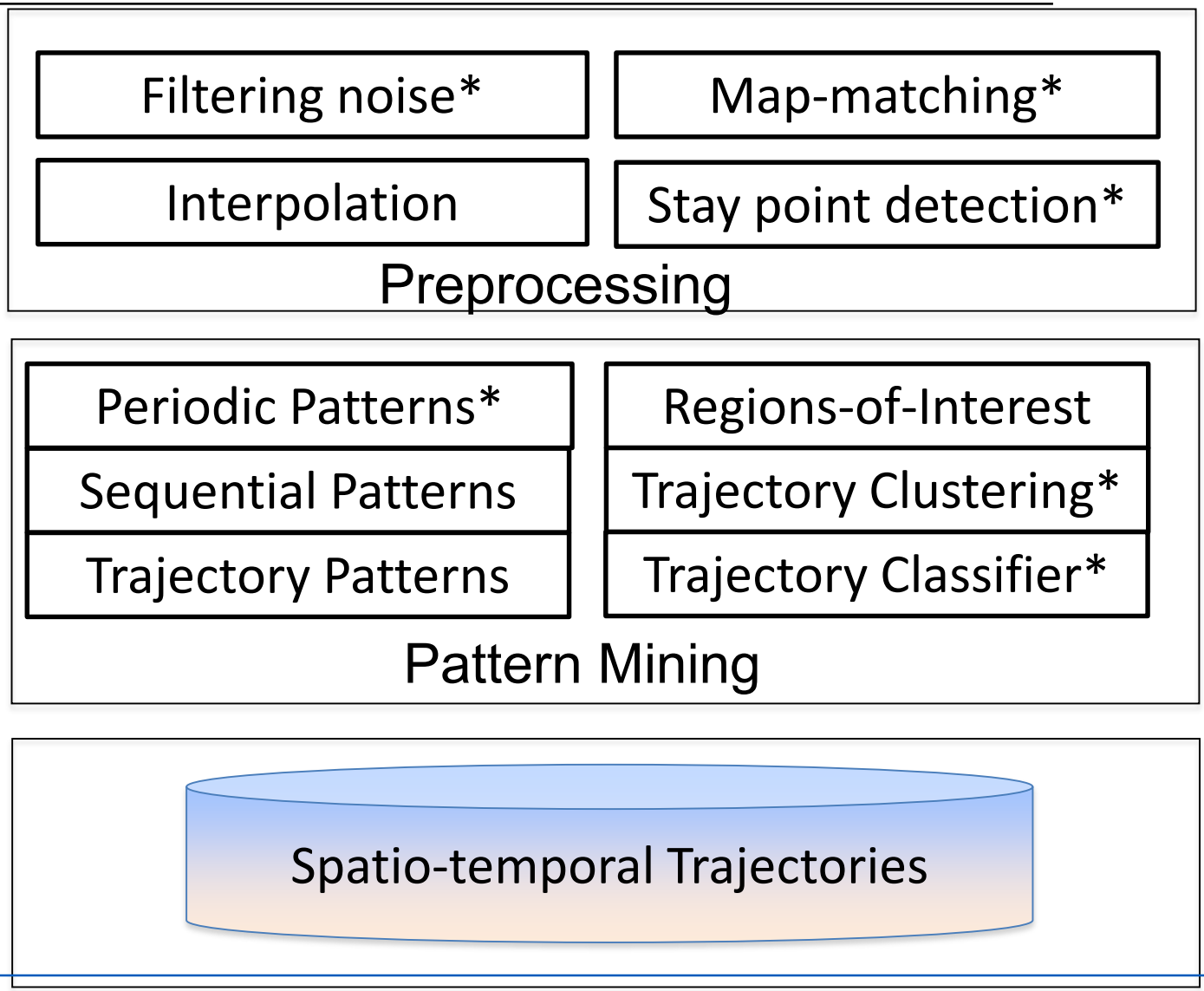
JAMES COOK UNIVERSITY AUSTRALIA

dog_walk.txt

```
1, -16.898353, 145.694882, 2016-11-06T18:13:07, MOVING
1, -16.898339, 145.694865, 2016-11-06T18:13:08, MOVING
1, -16.898307, 145.694836, 2016-11-06T18:13:09, MOVING
1, -16.898271, 145.694818, 2016-11-06T18:13:10, MOVING
1, -16.898265, 145.694808, 2016-11-06T18:13:11, MOVING
1, -16.89825, 145.694795, 2016-11-06T18:13:12, MOVING
1, -16.898232, 145.69477999999998, 2016-11-06T18:13:13, MOVING
1, -16.898222, 145.694772, 2016-11-06T18:13:14, MOVING
1, -16.898207, 145.694757, 2016-11-06T18:13:15, MOVING
1, -16.898203, 145.694745, 2016-11-06T18:13:16, MOVING
1, -16.898203, 145.69472999999996, 2016-11-06T18:13:17, MOVING
1, -16.898204000000003, 145.694716, 2016-11-06T18:13:18, MOVING
1, -16.898209, 145.69470400000003, 2016-11-06T18:13:19, MOVING
1, -16.898217, 145.69469, 2016-11-06T18:13:20, MOVING
1, -16.898229, 145.694677, 2016-11-06T18:13:21, MOVING
1, -16.898242, 145.694668, 2016-11-06T18:13:22, MOVING
1, -16.898253, 145.694661, 2016-11-06T18:13:23, MOVING
1, -16.898261, 145.694652, 2016-11-06T18:13:24, MOVING
1, -16.898271, 145.694641, 2016-11-06T18:13:25, MOVING
1, -16.898281, 145.69463, 2016-11-06T18:13:26, MOVING
1, -16.89829, 145.694623, 2016-11-06T18:13:27, MOVING
1, -16.898296999999996, 145.694616, 2016-11-06T18:13:28, MOVING
1, -16.898301, 145.69460499999997, 2016-11-06T18:13:29, MOVING
1, -16.898305, 145.69459499999996, 2016-11-06T18:13:30, MOVING
1, -16.898311, 145.694584, 2016-11-06T18:13:31, MOVING
1, -16.898320000000002, 145.694573, 2016-11-06T18:13:32, MOVING
1, -16.898327, 145.69456300000002, 2016-11-06T18:13:33, MOVING
1, -16.898335, 145.69455099999996, 2016-11-06T18:13:34, MOVING
1, -16.898344, 145.694542, 2016-11-06T18:13:35, MOVING
1, -16.89835, 145.694535, 2016-11-06T18:13:36, MOVING
1, -16.898355, 145.694528, 2016-11-06T18:13:37, MOVING
1, -16.898359, 145.694518, 2016-11-06T18:13:38, MOVING
1, -16.898362, 145.694509, 2016-11-06T18:13:39, MOVING
1, -16.898368, 145.69450200000003, 2016-11-06T18:13:40, MOVING
1, -16.898381, 145.694499, 2016-11-06T18:13:41, MOVING
1, -16.898396, 145.69449400000002, 2016-11-06T18:13:42, MOVING
1, -16.898412, 145.694487, 2016-11-06T18:13:43, MOVING
1, -16.898426, 145.69448, 2016-11-06T18:13:44, MOVING
1, -16.898434, 145.694472, 2016-11-06T18:13:45, MOVING
1, -16.898445000000002, 145.694463, 2016-11-06T18:13:46, MOVING
1, -16.898452, 145.69445299999998, 2016-11-06T18:13:47, MOVING
1, -16.898452, 145.694439, 2016-11-06T18:13:48, MOVING
1, -16.898452, 145.69442800000002, 2016-11-06T18:13:49, MOVING
1, -16.898455, 145.694417, 2016-11-06T18:13:50, MOVING
1, -16.898466, 145.694408, 2016-11-06T18:13:51, MOVING
1, -16.898479, 145.694403, 2016-11-06T18:13:52, MOVING
1, -16.898491, 145.694398, 2016-11-06T18:13:53, MOVING
1, -16.898503, 145.69439, 2016-11-06T18:13:54, MOVING
1, -16.898512999999998, 145.694382, 2016-11-06T18:13:55, MOVING
1, -16.89852, 145.694372, 2016-11-06T18:13:56, MOVING
1, -16.898524, 145.69436400000004, 2016-11-06T18:13:57, MOVING
1, -16.898529, 145.69435500000003, 2016-11-06T18:13:58, MOVING
1, -16.898535000000003, 145.69434699999996, 2016-11-06T18:13:59, MOVING
1, -16.898542, 145.694338, 2016-11-06T18:14:00, MOVING
1, -16.898548, 145.69432999999998, 2016-11-06T18:14:01, MOVING
1, -16.898555999999996, 145.69432099999997, 2016-11-06T18:14:02, MOVING
1, -16.898565, 145.694314, 2016-11-06T18:14:03, MOVING
1, -16.898575000000005, 145.69430600000004, 2016-11-06T18:14:04, MOVING
1, -16.898584, 145.694296, 2016-11-06T18:14:05, MOVING
1, -16.898592, 145.694288, 2016-11-06T18:14:06, MOVING
1, -16.898606, 145.694274, 2016-11-06T18:14:07, MOVING
```

How the data looks like?

Overview of Trajectory Data Mining



TDM Noise Filtering

JAMES COOK UNIVERSITY AUSTRALIA

- What?
  - The process of fitting raw trajectory recordings onto an underlying map structure before data mining
- How?
  - Very different from ‘structured data’
  - The idea is how do you combine location (map) with time data?
  - Noisy with GPS etc.

Issues with GPS Trajectories

JAMES COOK UNIVERSITY AUSTRALIA

- Spatial uncertainties
- Errors and noisy
- Irregular
- Could be too densely recorded or too coarsely recorded

→ Preprocessing

Trajectory Simplification

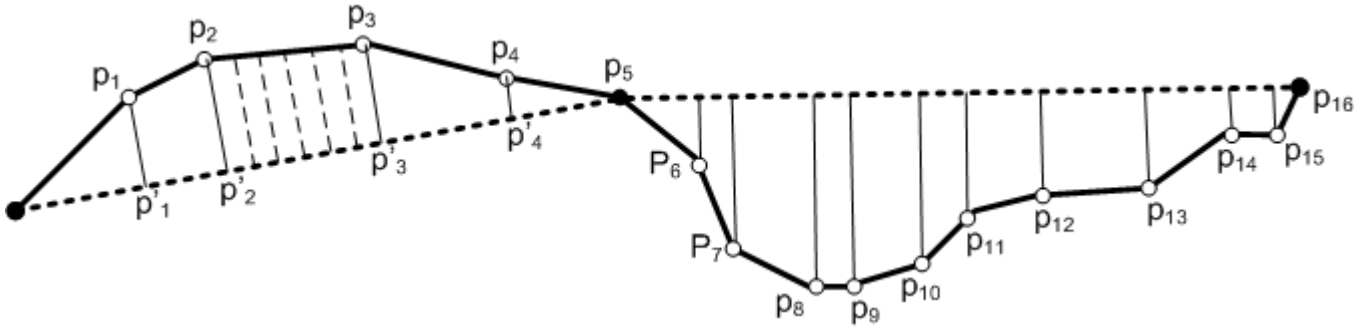
JAMES COOK UNIVERSITY AUSTRALIA

- Aim
  - Reduce the complexity of an input trajectory
  - Sensors capture as much movement details as possible by oversampling but still want to preserving the motion of the tracked entity
- Performance metrics
  - Reduce processing time
  - Reduce Error measure
- What error measure?
  - Criteria include perpendicular Euclidean distance and time synchronized Euclidean distance

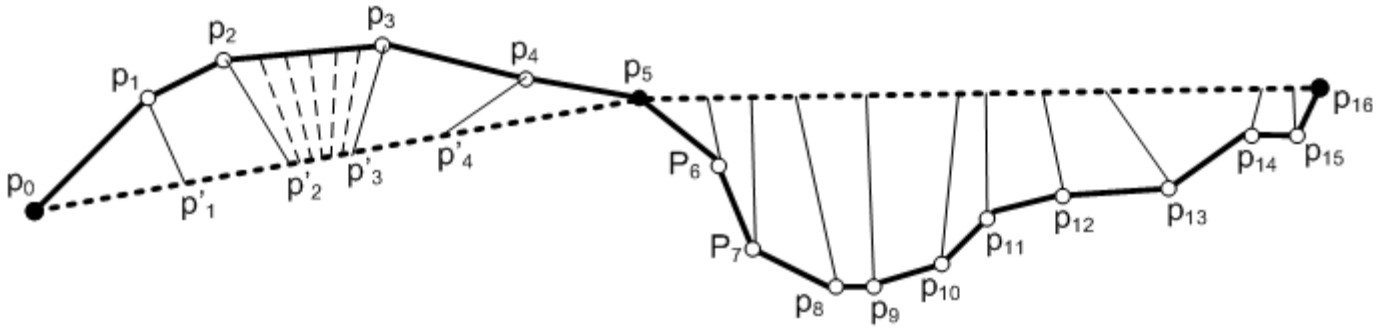
Illustration of Error Measures

JAMES COOK UNIVERSITY AUSTRALIA

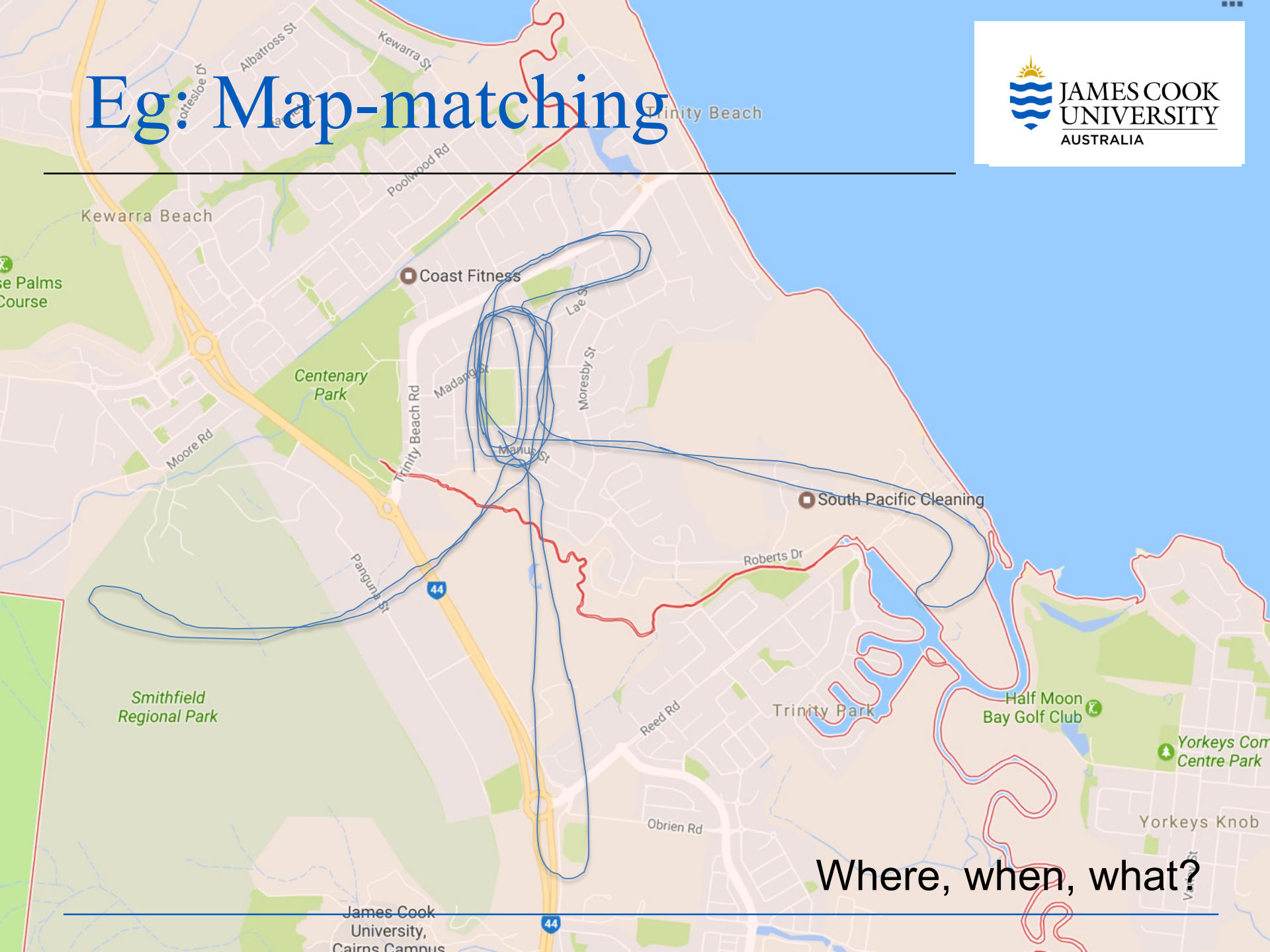
- Perpendicular Euclidean Distance



- Time Synchronized Euclidean Distance

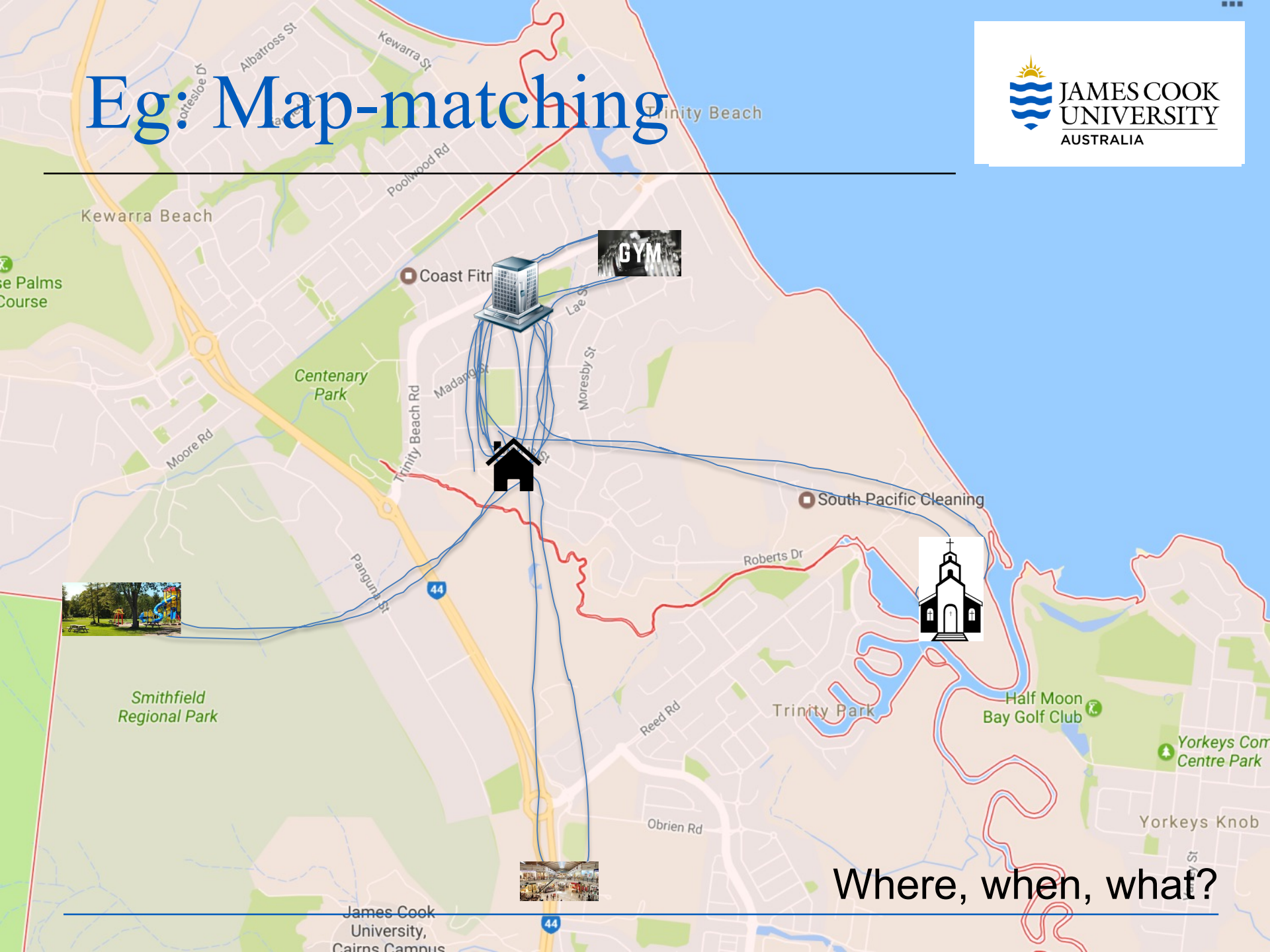


Eg: Map-matching



Where, when, what?

Eg: Map-matching



Where, when, what?

Eg: Map-matching

JAMES COOK UNIVERSITY AUSTRALIA

Monday, Wednesday 12pm

Monday, Wednesday 1pm

9am 5pm

Sunday 11am

Sunday 12pm

Saturday 4pm

Saturday 5pm

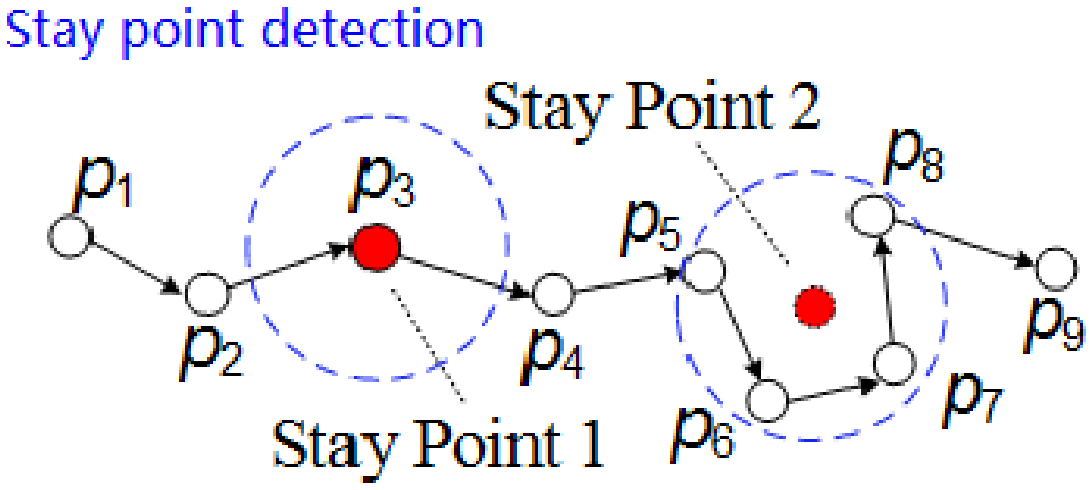
Thursday 8pm

Thursday 6pm

Where, when, what?

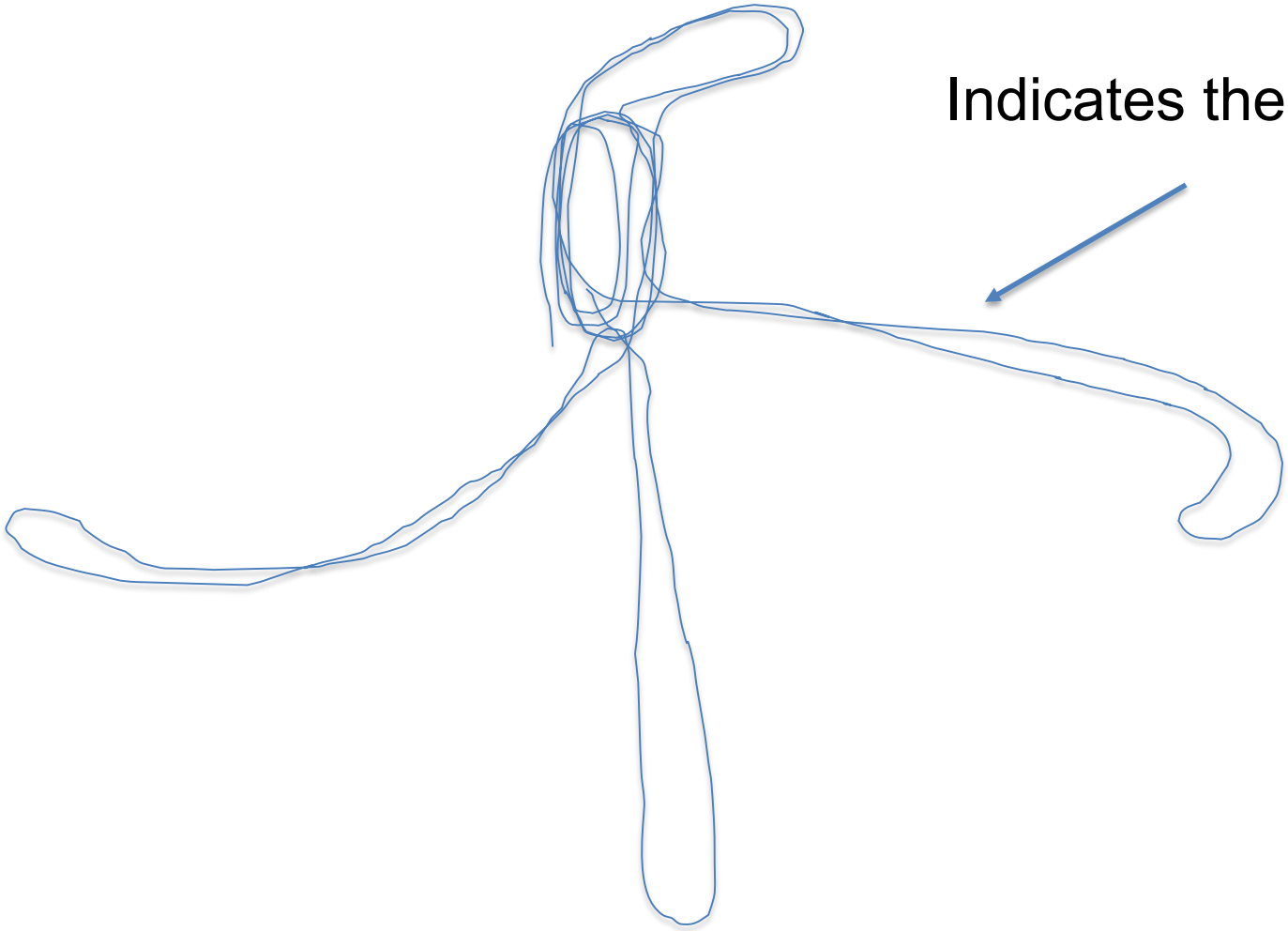
Stay Point Detection

- The identification of a location a moving object has stayed for a while within a certain distance threshold
- These stay points can indicate interesting insights for eg. at a restaurant/ shopping mall.
- Uses clustering technique studied earlier eg DBSCAN



Stop/Move Representation

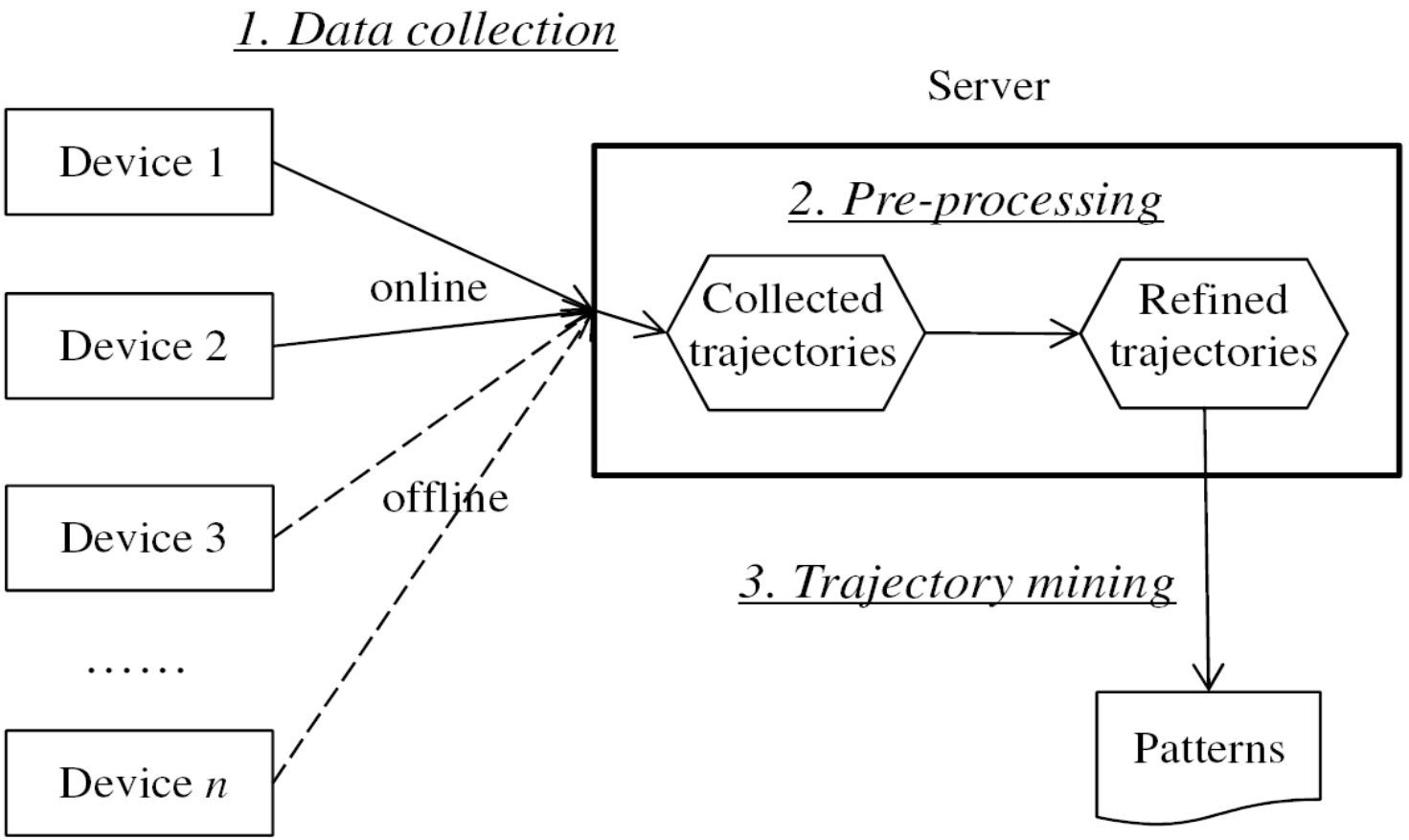
JAMES COOK UNIVERSITY AUSTRALIA



Stop/Move Representation

JAMES COOK UNIVERSITY AUSTRALIA

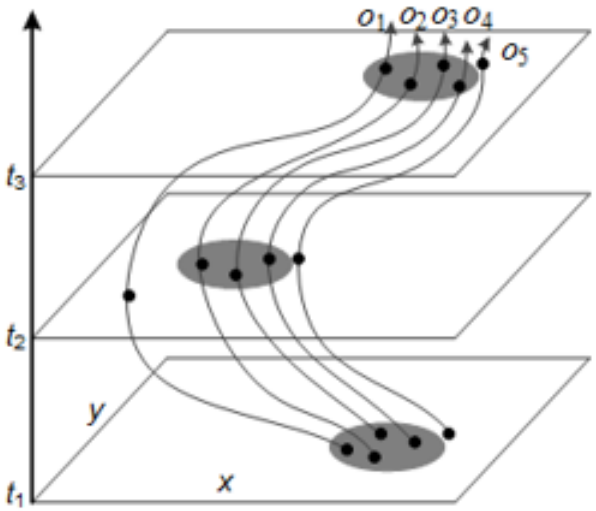
Trajectory Data Mining



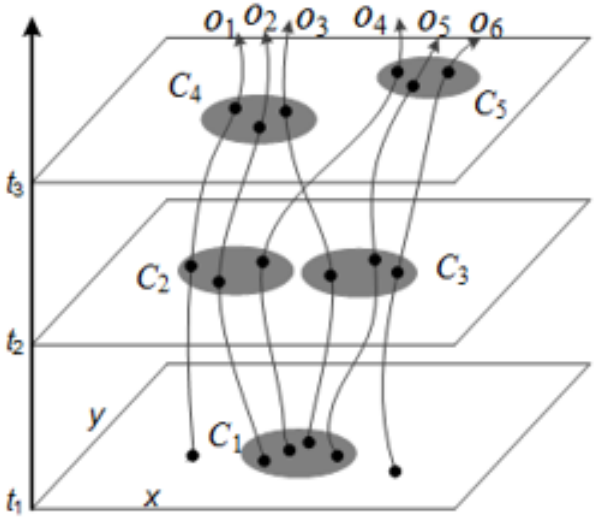
Trajectory Data Mining

JAMES COOK UNIVERSITY AUSTRALIA

- Categories of patterns:
  - moving together patterns,
  - trajectory clustering,
  - periodic patterns and
  - frequent sequential patterns



A) Flock, convoy and swarm

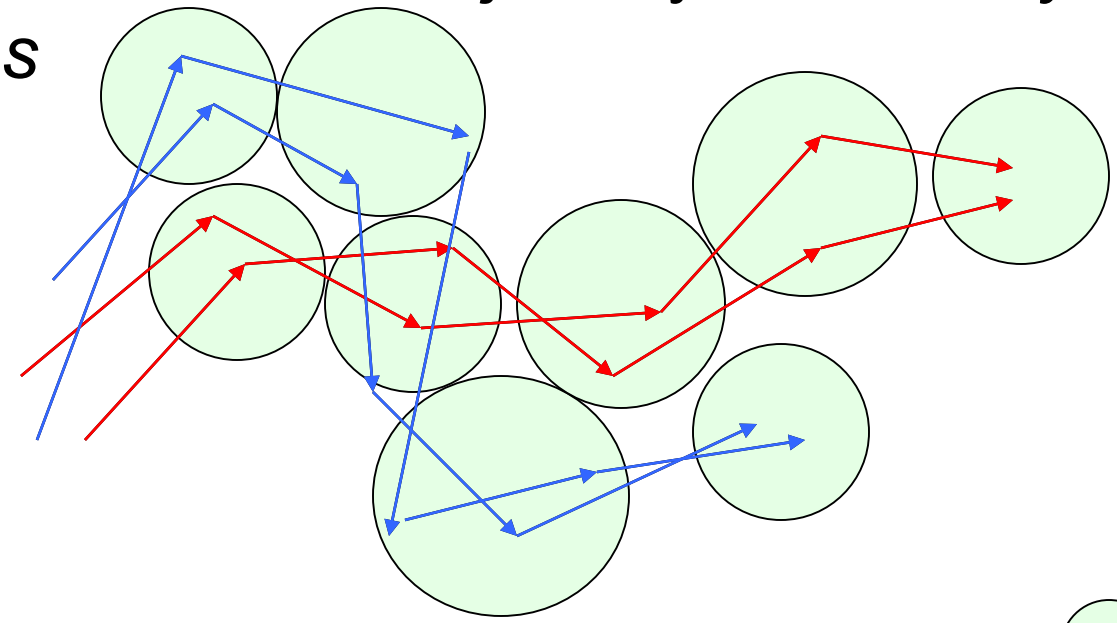


B) Gathering

Trajectory Clustering

JAMES COOK UNIVERSITY AUSTRALIA

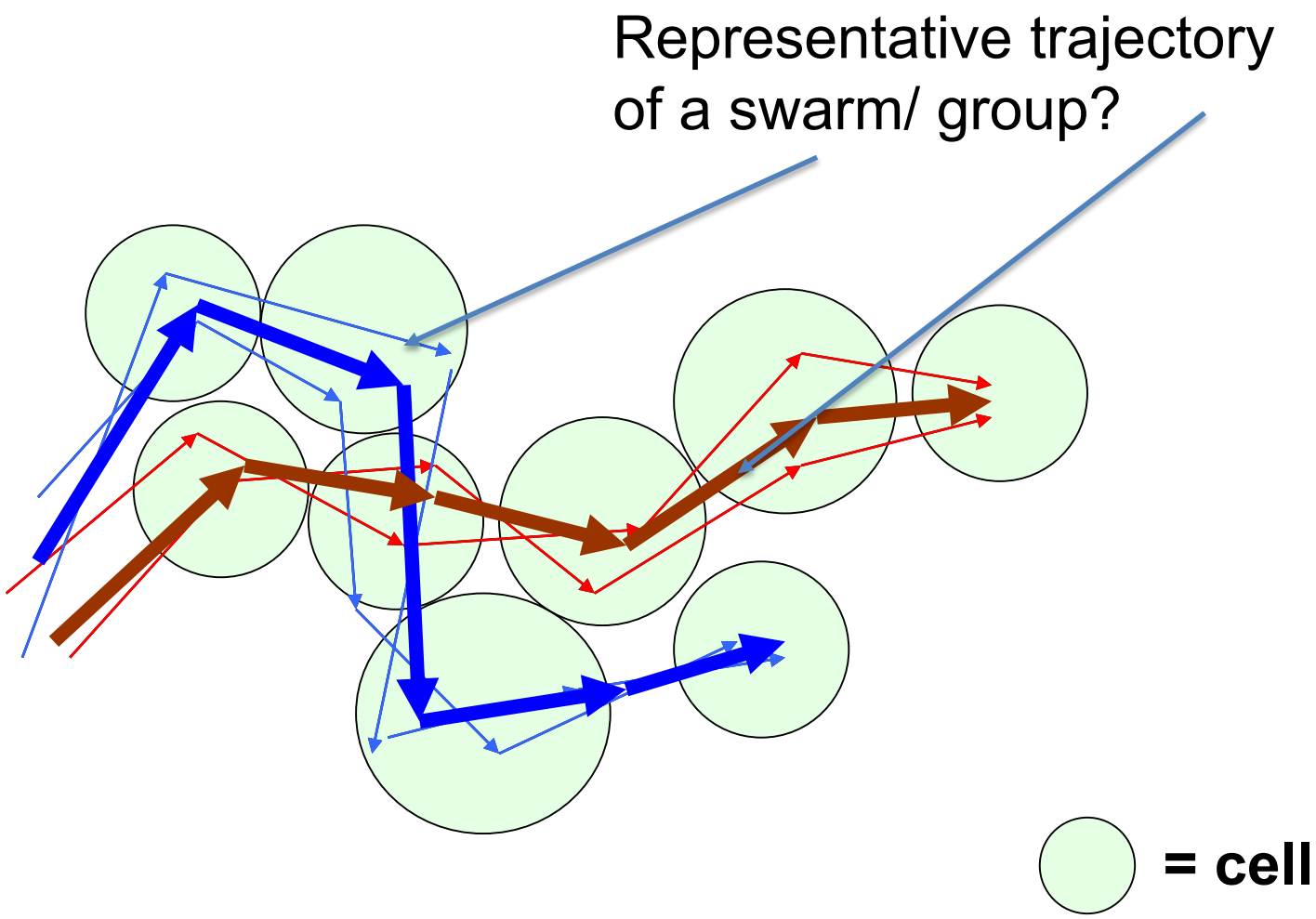
- Group similar trajectories geometric proximity in spatial/spatiotemporal space.
- Find a representative trajectory from *many trajectories*



= cell

Trajectory Clustering

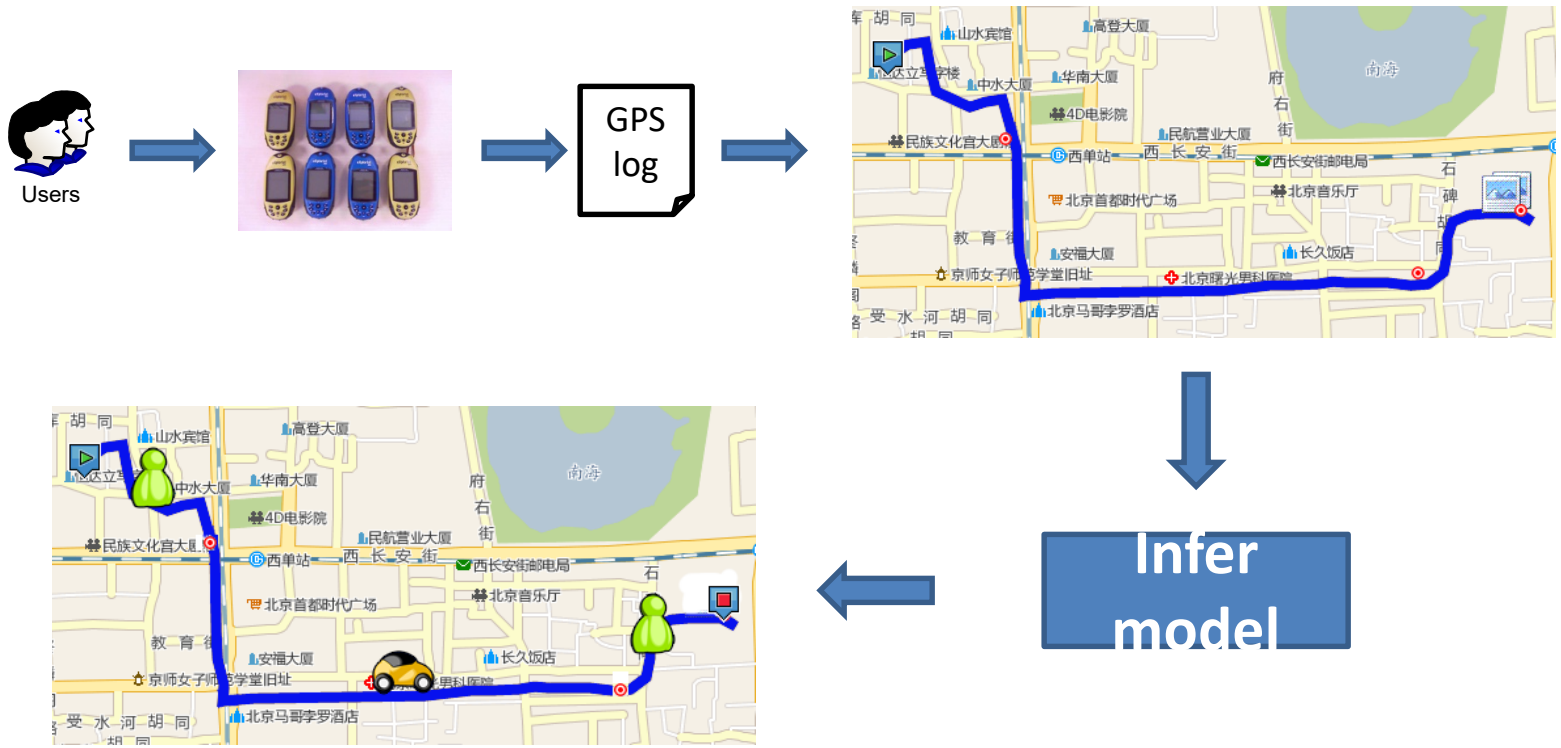
JAMES COOK UNIVERSITY AUSTRALIA



Trajectory Classification

JAMES COOK UNIVERSITY AUSTRALIA

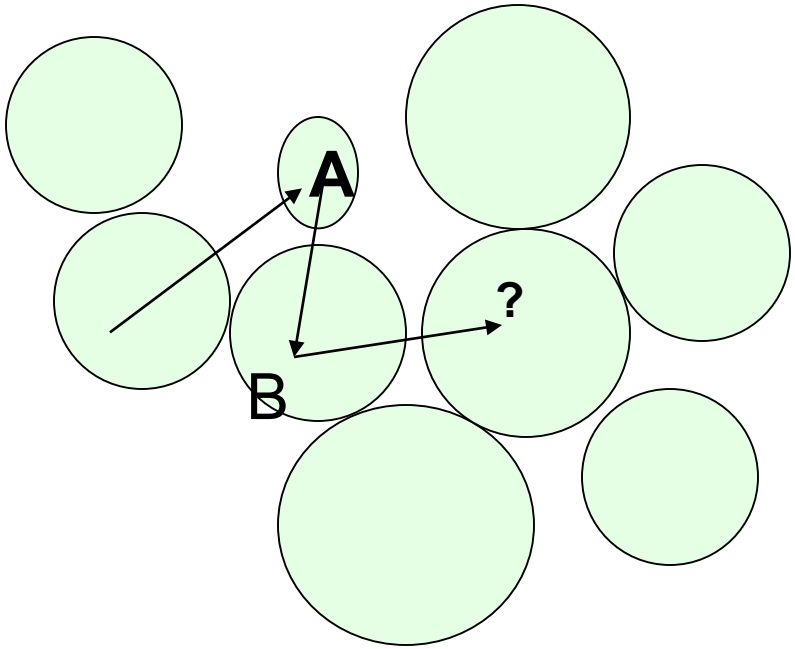
- With supervised learning, classify trajectories into activities like hiking/ dining or different modes (walking/ driving)



Trajectory Classification

JAMES COOK UNIVERSITY AUSTRALIA

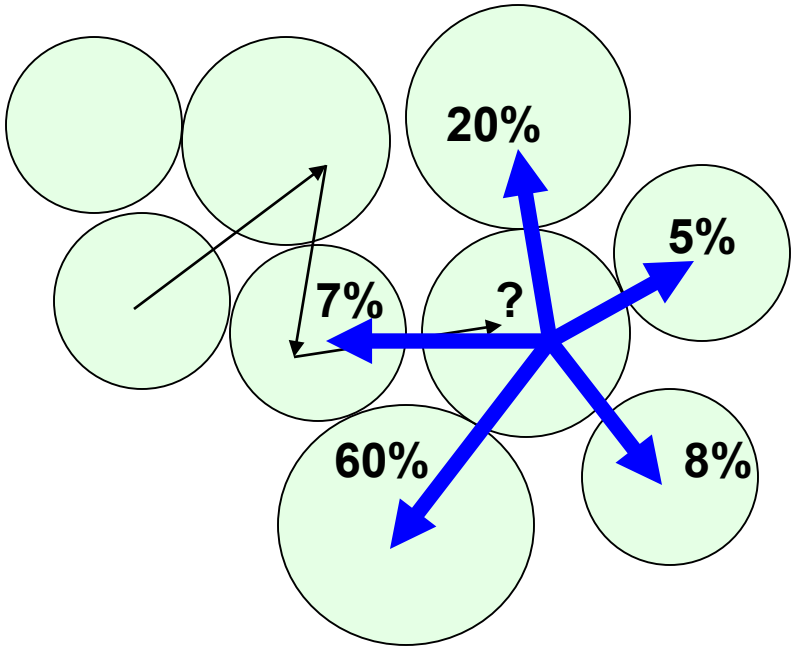
- Predict next move.
  - If it is driving activity, where is next place of interest after A / B?



Trajectory Classification

JAMES COOK UNIVERSITY AUSTRALIA

- Obtain next destination with probability.
  After drinks and eating, next?



Spatio-Temporal Periodic Pattern Mining

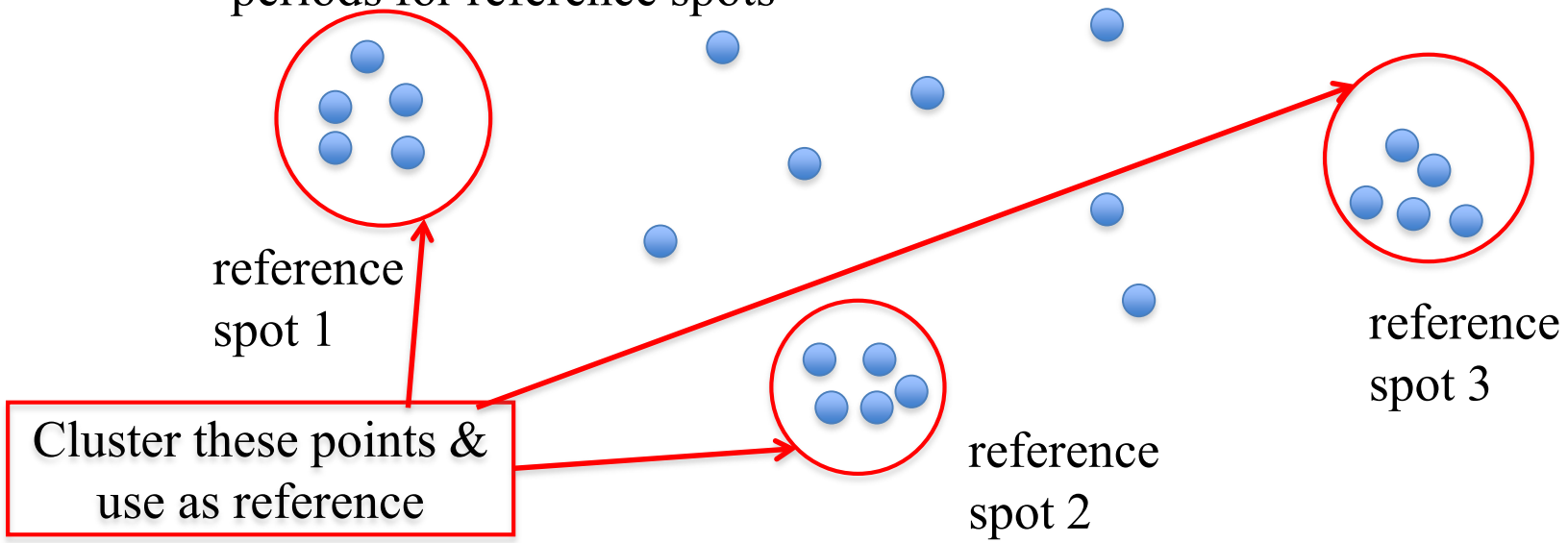
JAMES COOK UNIVERSITY AUSTRALIA

- Periodic patterns are trajectories periodically executed by a moving object. For eg. regular movement patterns from office staff, which are rather similar each working day.
- There are 2 main approaches:
  - Fixed Period Approach
  - Reference Spot Approach

Spatio-Temporal PPM

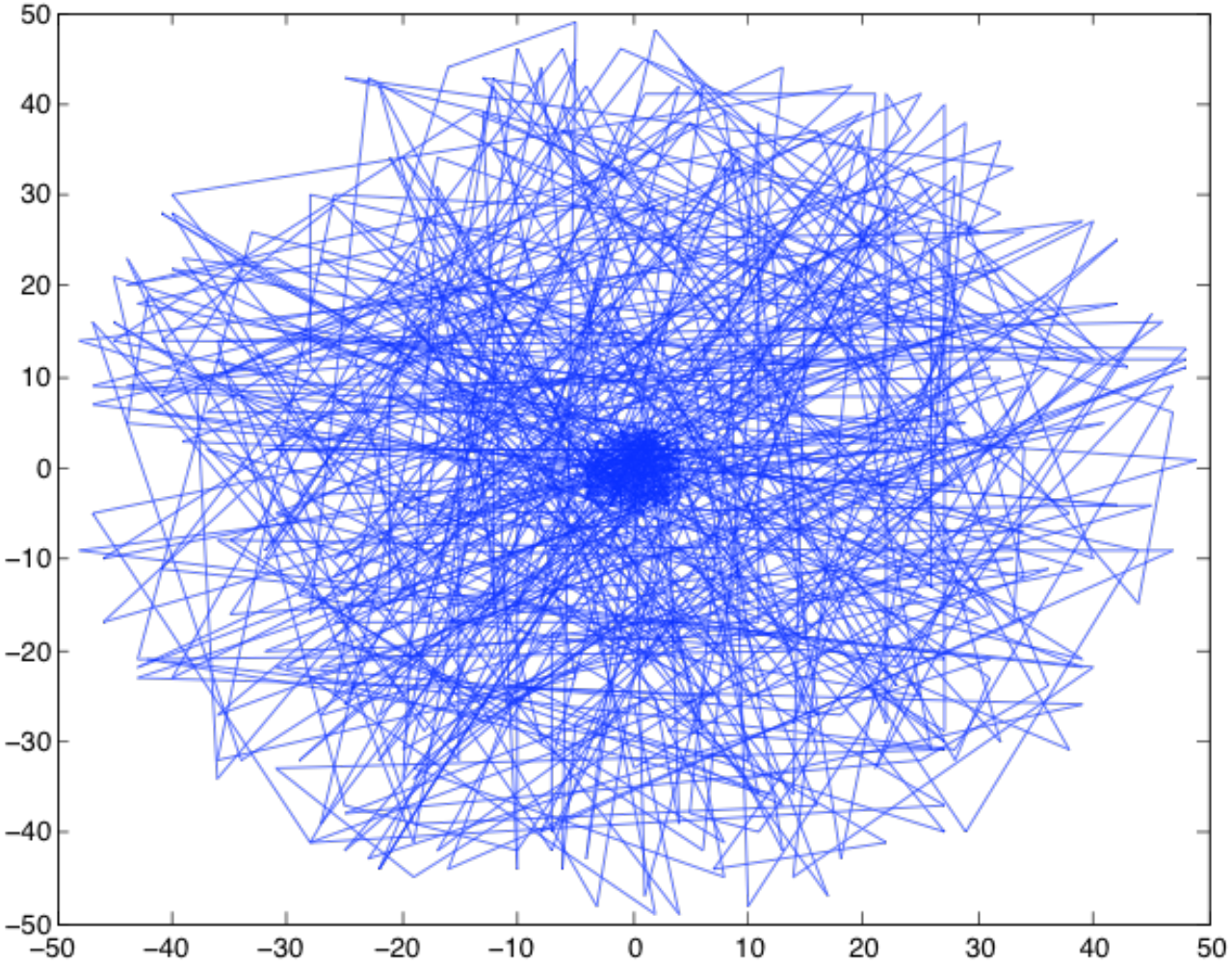
JAMES COOK UNIVERSITY AUSTRALIA

- Fixed (Time) Period Approach
  - To segment the long trajectory into a set of smaller (shorter) sub-trajectories based on a given fixed time period
- Reference Spot Approach
  - Find reference spots using clustering algorithms and then find associated periods for reference spots



Periodic Pattern Mining

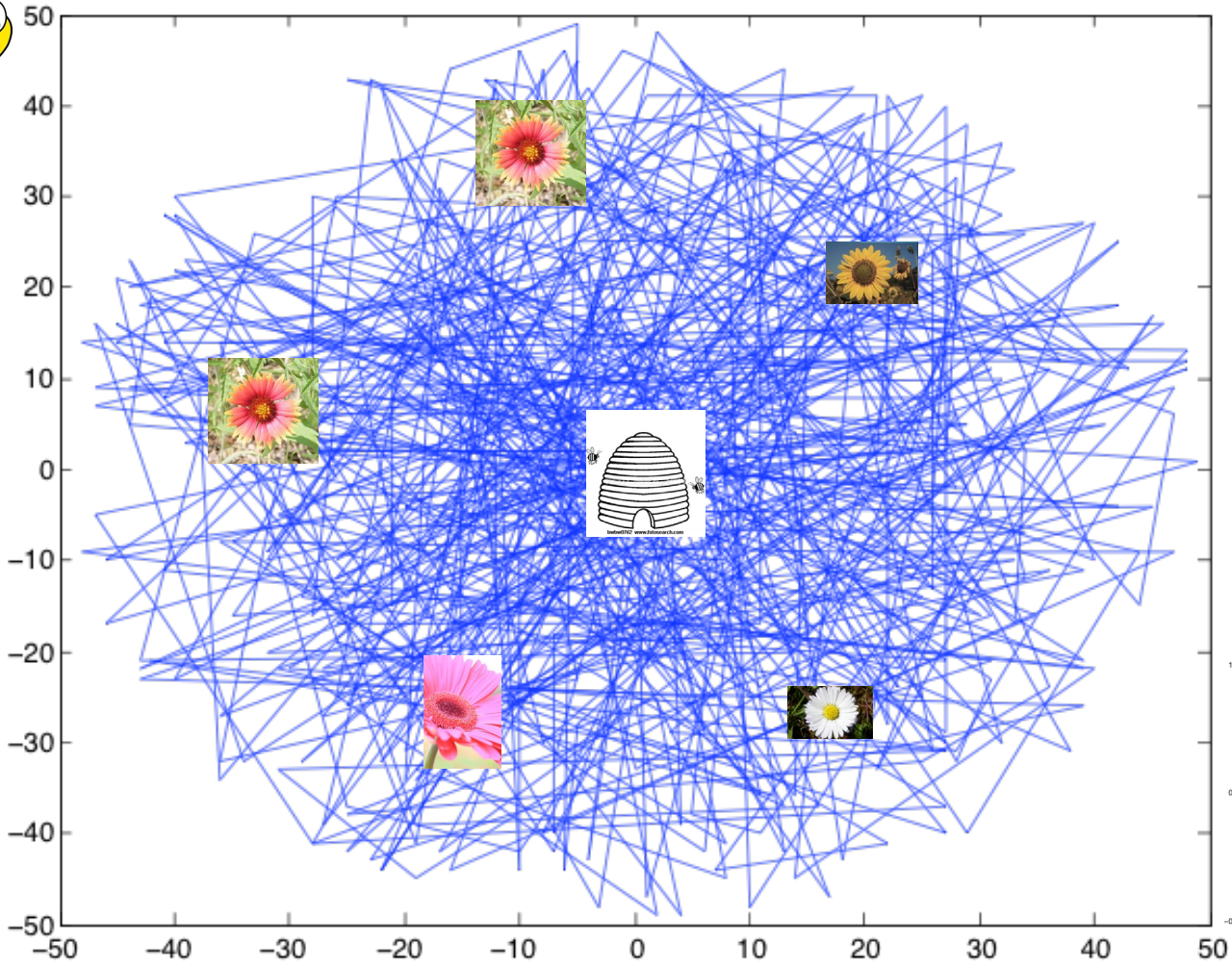
JAMES COOK UNIVERSITY AUSTRALIA

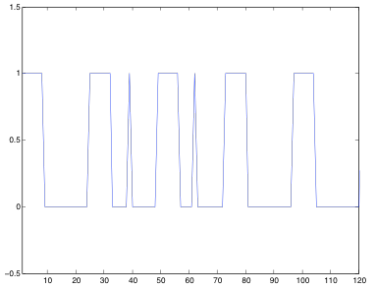


Not easy! Eg. movements of a bee (or bees)

Periodic Pattern Mining

JAMES COOK UNIVERSITY
AUSTRALIA

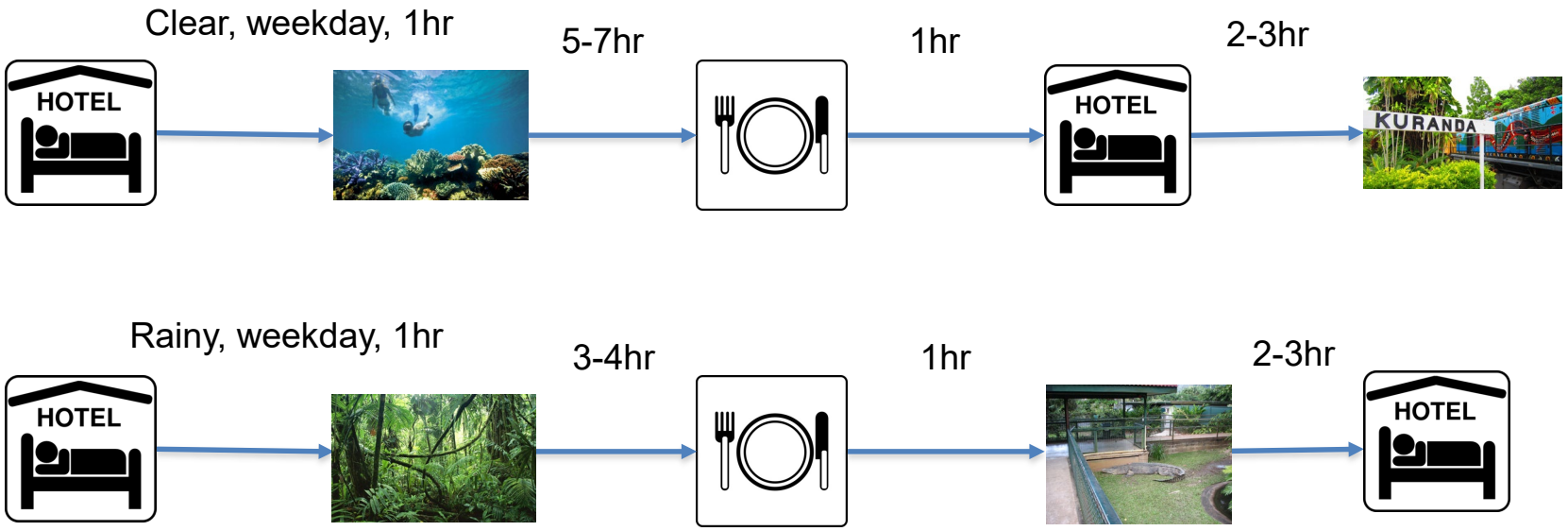




Trajectory Pattern Mining

JAMES COOK UNIVERSITY AUSTRALIA

- TPM considers spatio-temporal information
- In addition, add on aspatial *semantic information* to produce richer patterns



Open Challenges

JAMES COOK UNIVERSITY AUSTRALIA

- Incorporate semantics – semantic trajectory data mining by incorporating aspatial information
- Techniques largely the same:
  - classification is still in its infancy
  - Association mining (more used)
  - Lots of pre-processing with uncertainties and noise handling