Understand individual mobility and accessibility using smartphone-based activity survey: an object-oriented framework

Ying Song
Geography, Environment and Society
University of Minnesota – Twin Cities

Yingling Fan
Humphrey School of Public Affairs
University of Minnesota – Twin Cities

Julian Wolfson
Division of Biostatistics, School of Public Health
University of Minnesota – Twin Cities

Individuals' Movement in Urban Environment

❖ Human Mobility and Accessibility

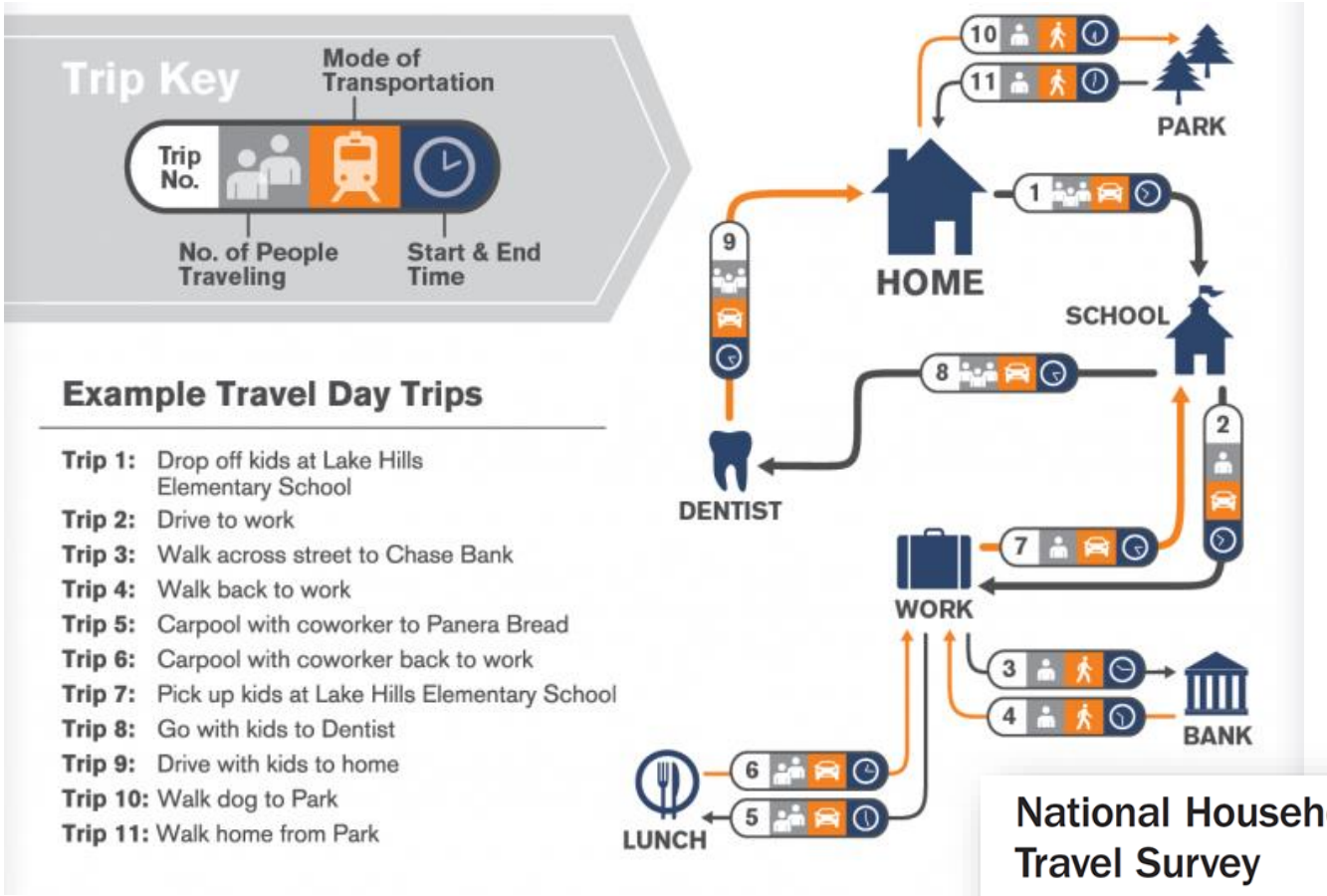
Movement Types		Temporal	
		Non-recurrent	Recurrent
Spatial	Local	**Residential Mobility**	**Commuting**
	Long distance	**Migration**	**Seasonal Work**

Sustainable Mobility (Banister 2008)

- individual schedule
- accessibility to destinations
- multimodal

Individuals' Movement in Urban Environment

❖ Data Collection - Travel Survey



Rich in contents

Retrospective

"Non-spatial"

Graph Source

Individuals' Movement in Urban Environment

❖ Data Collection - Location-Aware Technologies


Mobile


TomTom

Garmin

Safety

Commute

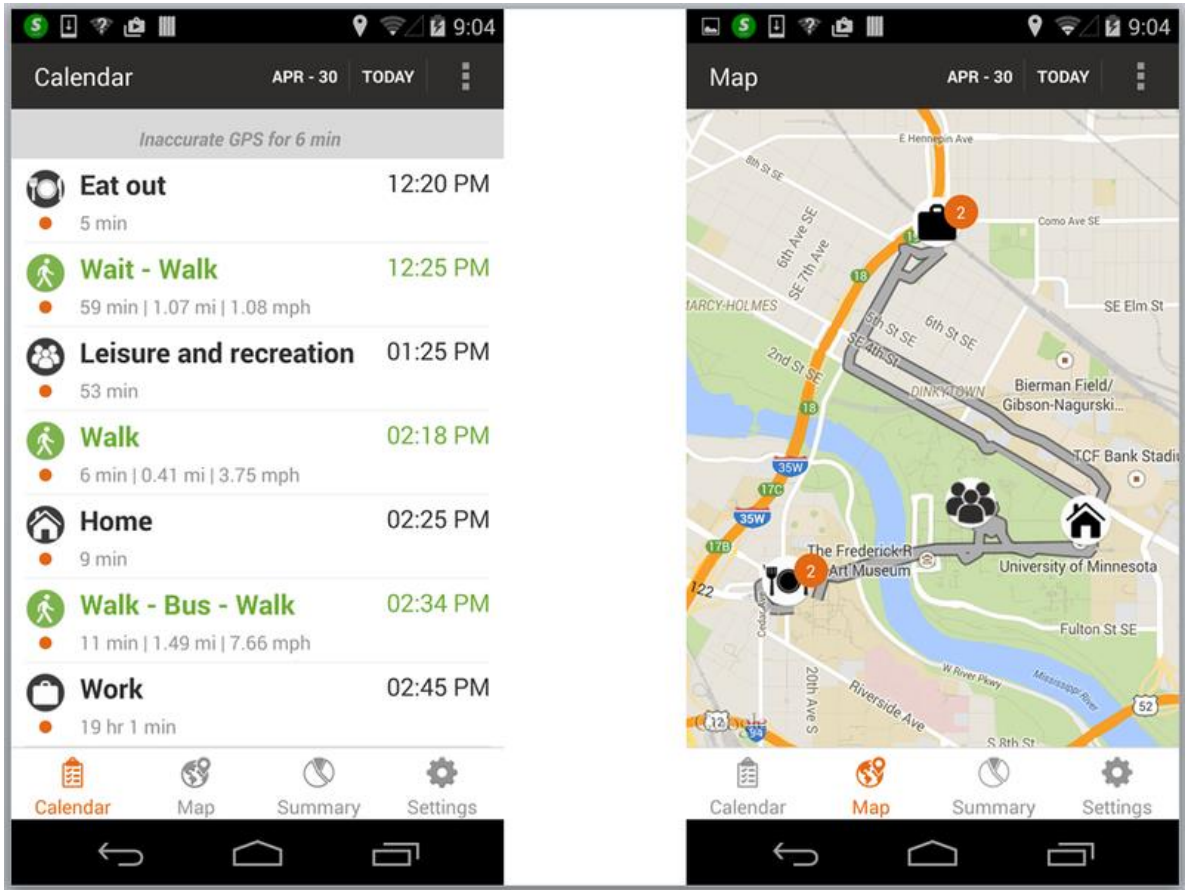
Massive but Thin

Individuals' Movement in Urban Environment

❖ Data Collection - Smartphone-based Survey Apps

Dynamica

- ✓ user profile
- ✓ trip purpose
- ✓ travel mode
- ✓ schematic attributes
  - emotion status
  - trip satisfaction
  - estimated costs
  - …



Objectives

- **An object-oriented framework**

1. **Data Management**

   1.1 Conceptual data representations

   1.2 Logical data structures

   1.3 Physical data storage, index, …

2. **Data Analysis**

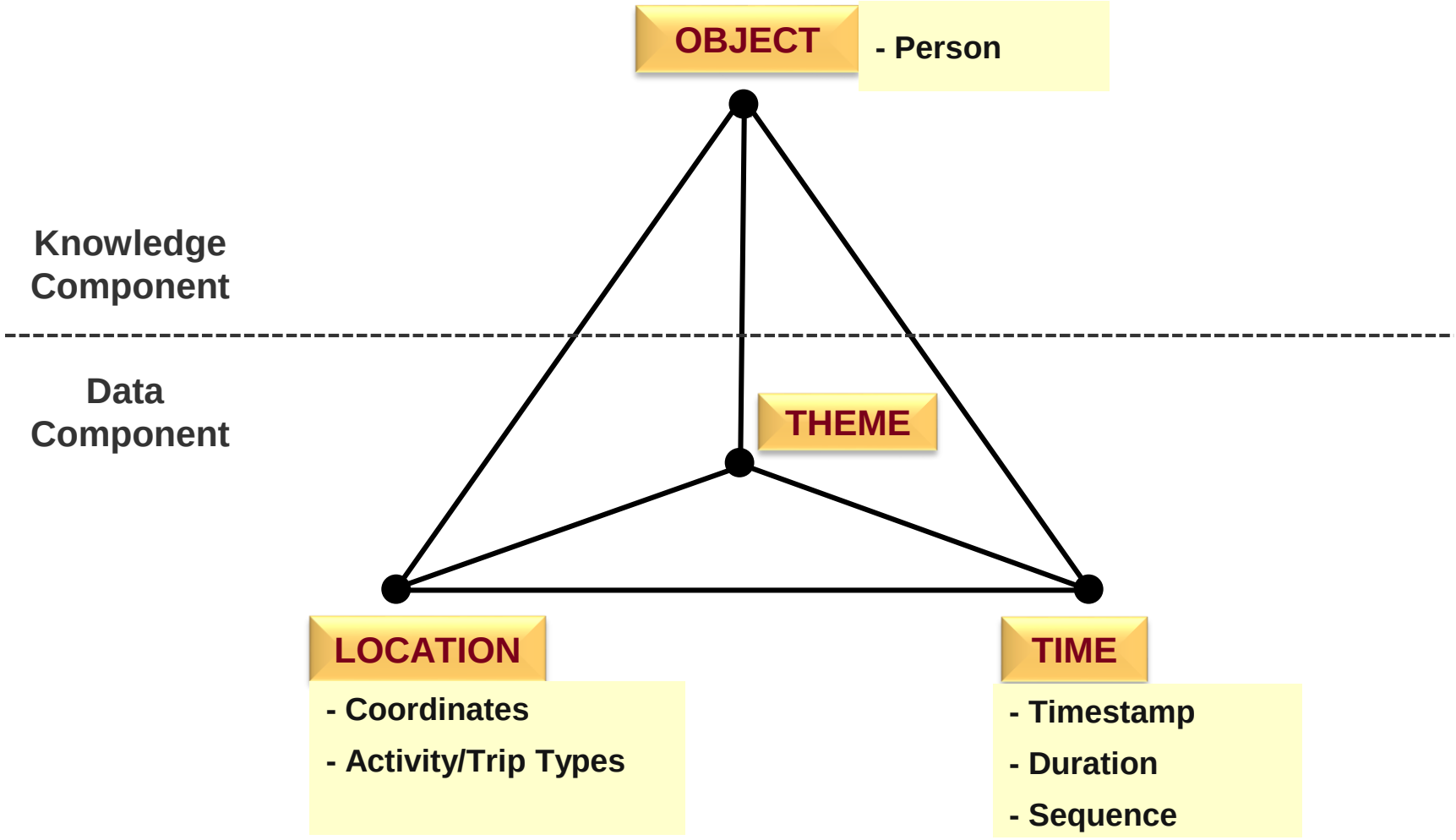
   2.1 Partonomy & Taxonomy

   2.2 Approaches

   - Statistics
   - Data Mining

1.1 Conceptual: Data Representation

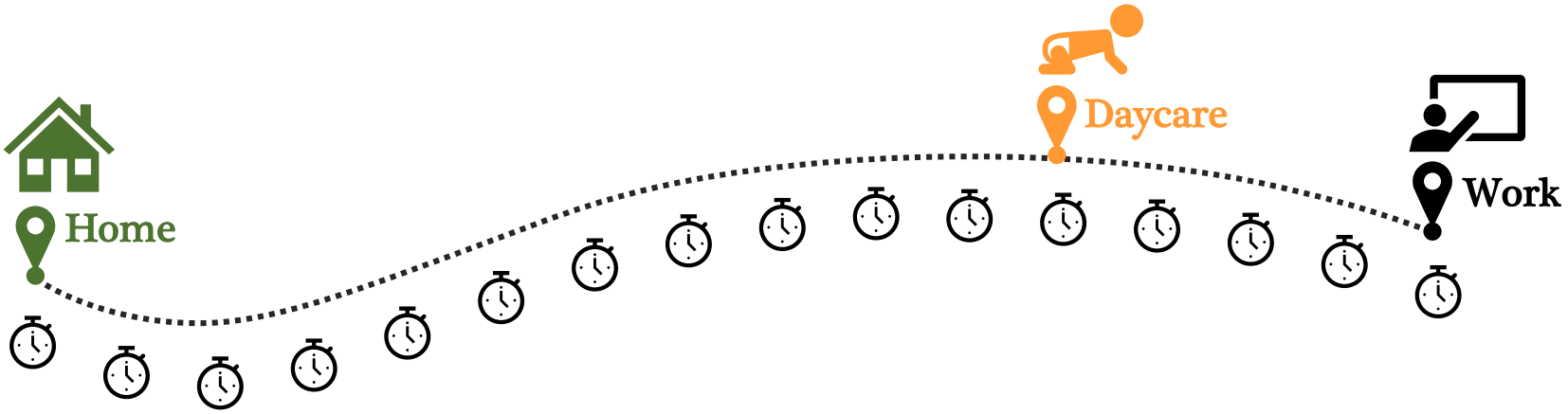
❖ Pyramid Framework



Mennis et al. 2000

1.1 Conceptual: Data Representation

❖ Data Component

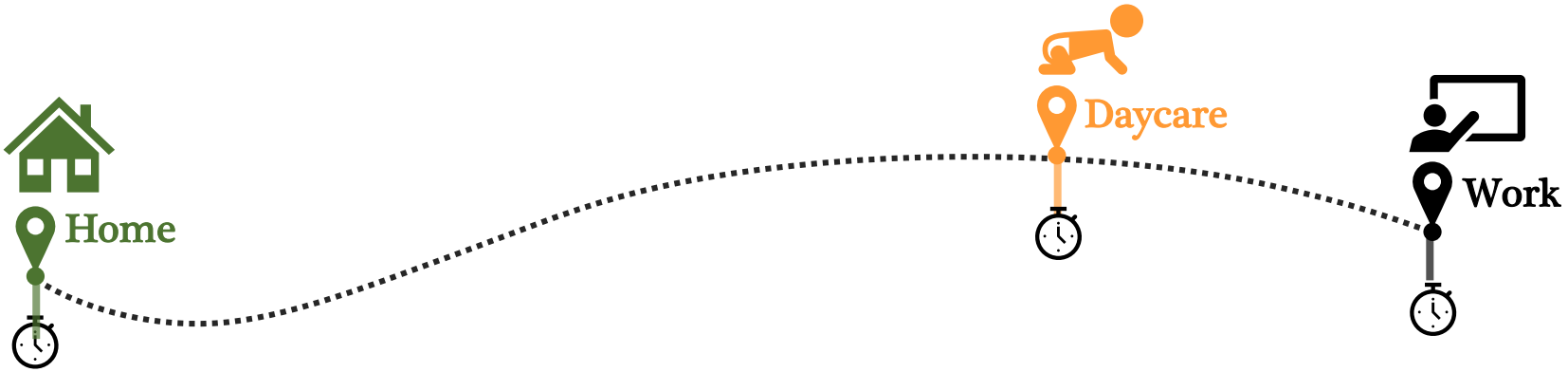


Lagrangian: GPS-based tracking data

Location	a spatial location
Time	a timestamp
Theme	mobility status (e.g. odometer, speed, energy level…)

1.1 Conceptual: Data Representation

❖ Data Component

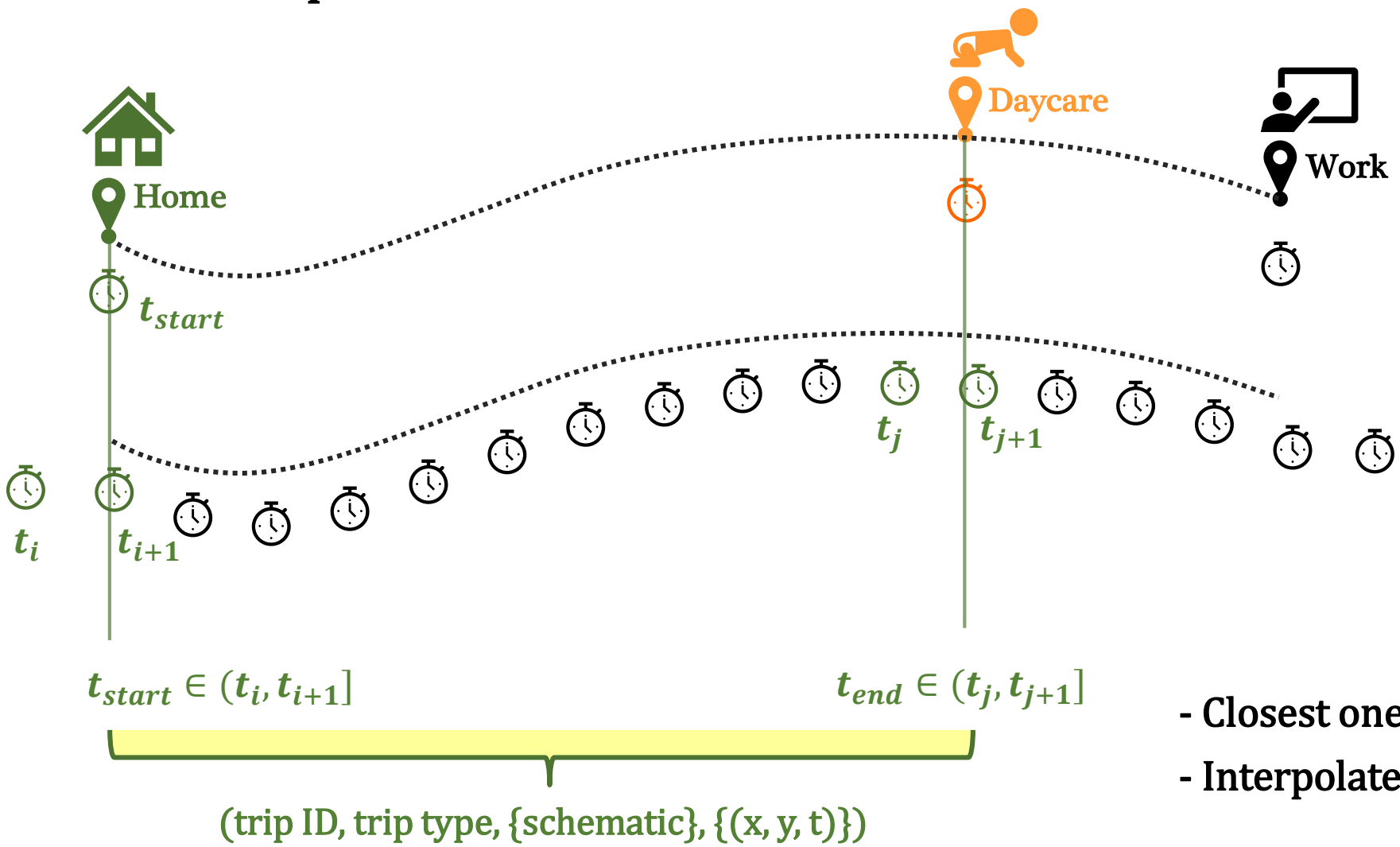


Eulerian: Smartphone-based activity survey

Location	an activity or a trip
Time	the duration of the activity or trip
Theme	experience (e.g. emotion status, trip routes, costs, …)

1.1 Conceptual: Data Representation

❖ Relationship

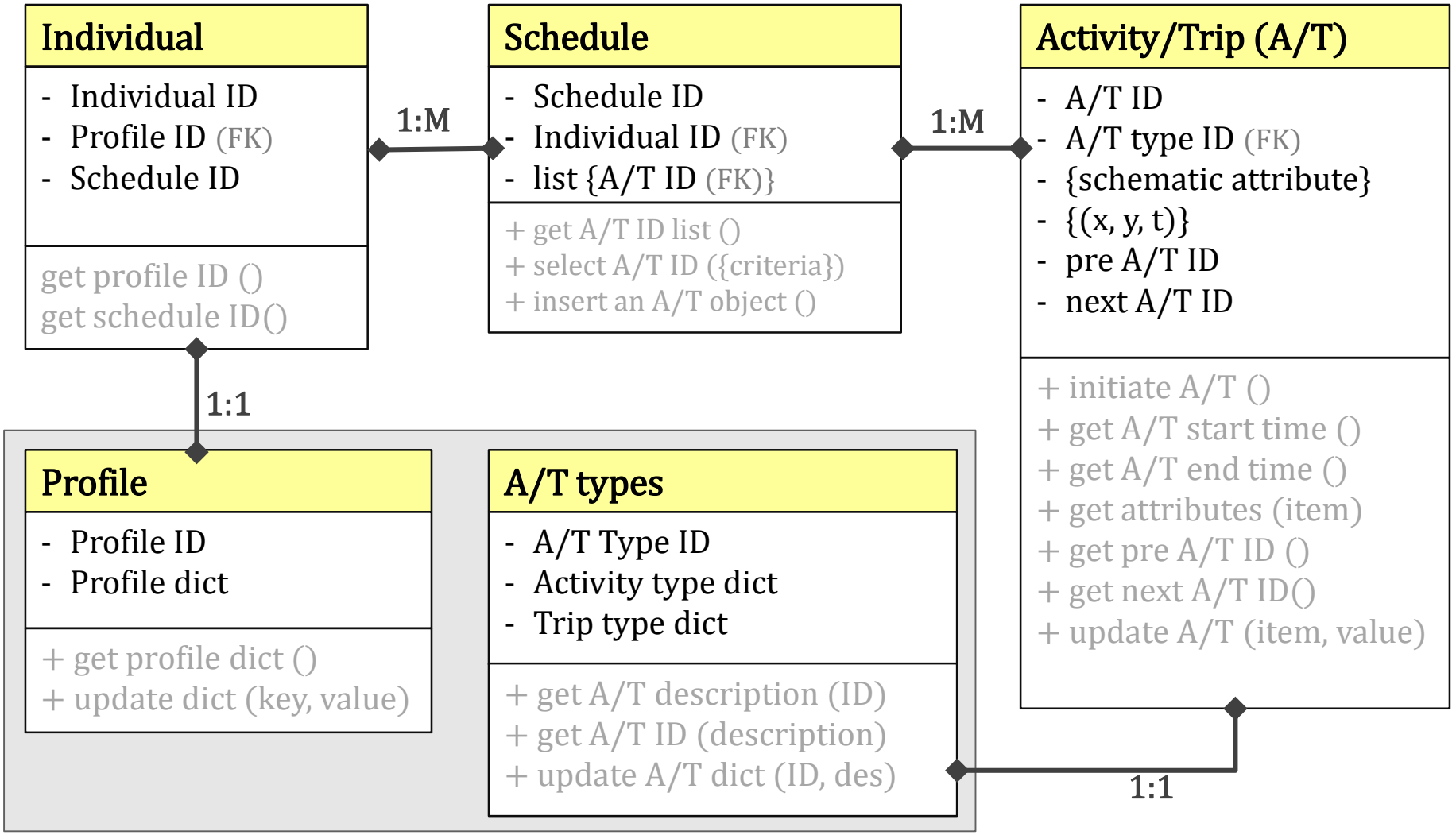


- Closest one
- Interpolated

yingsong@umn.edu

1.1 Logical: Data Structure

❖ **Objects, Attributes and Behaviors** (selected)



2.1 Taxonomy & Partonomy

❖ Knowledge Component

		Lagrangian GPS-based tracking data	Eulerian Smartphone-based activity survey
Knowledge	Partonomy	• space-time locations • stops and moves • movement trajectories	• activities and their locations • trips and their travel modes • individuals with trips and activities
	Taxonomy	**space-time trajectories**, each with a sequence of tuple {(x, y, t, {status})}	**Individual participants,** each with an ordered series of tuple {(activity/trip, start t, end t, {schematic})}

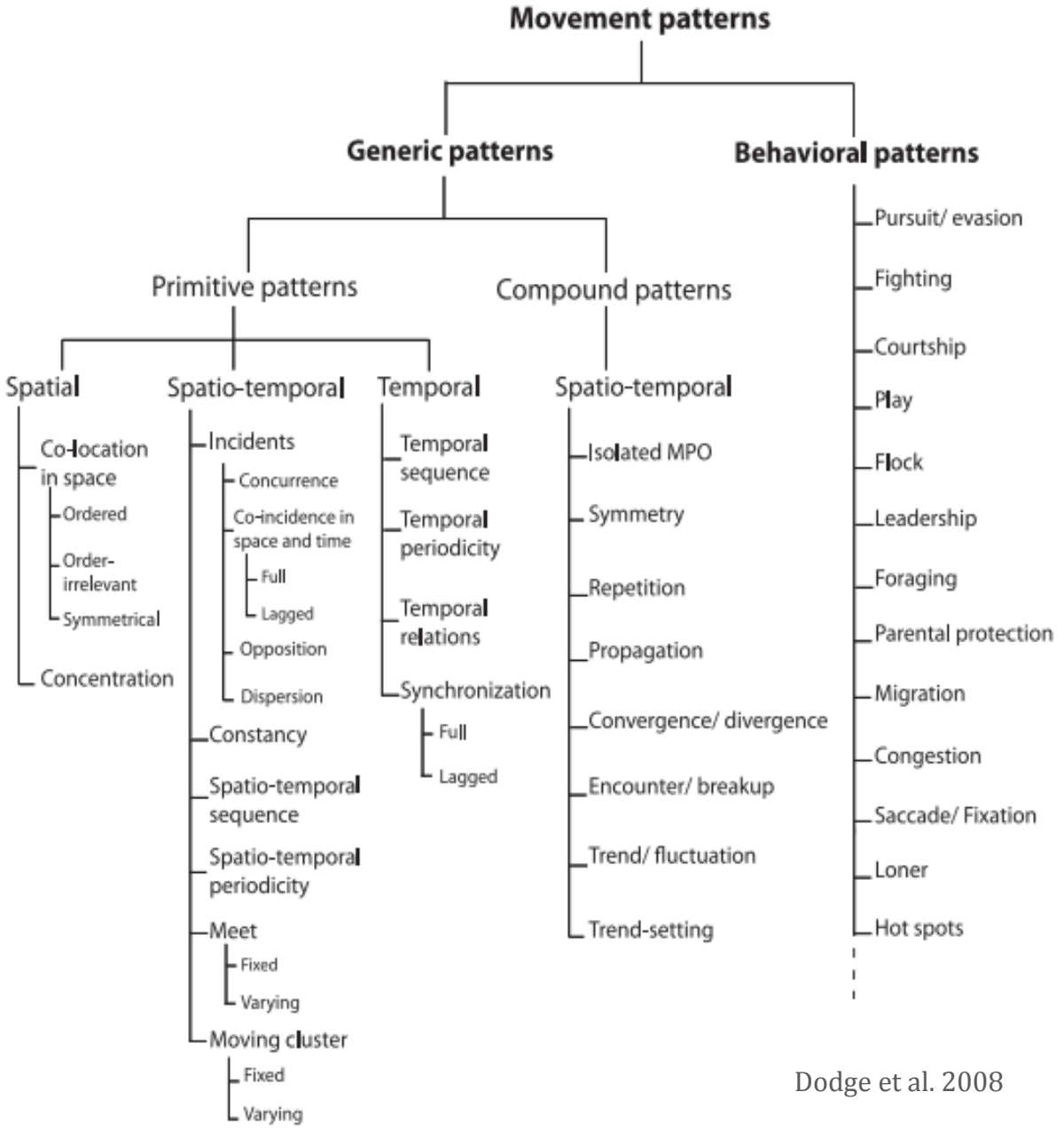
2.1 Taxonomy & Partonomy

❖ Mobility Patterns

1. Movement parameters

Activity/Trip (A/T)
- {schematic attribute} - {(x, y, t)}

- Individual group
- Schedule feature
- Activity/Trip type



Dodge et al. 2008

2.1 Taxonomy & Partonomy

❖ Mobility Patterns

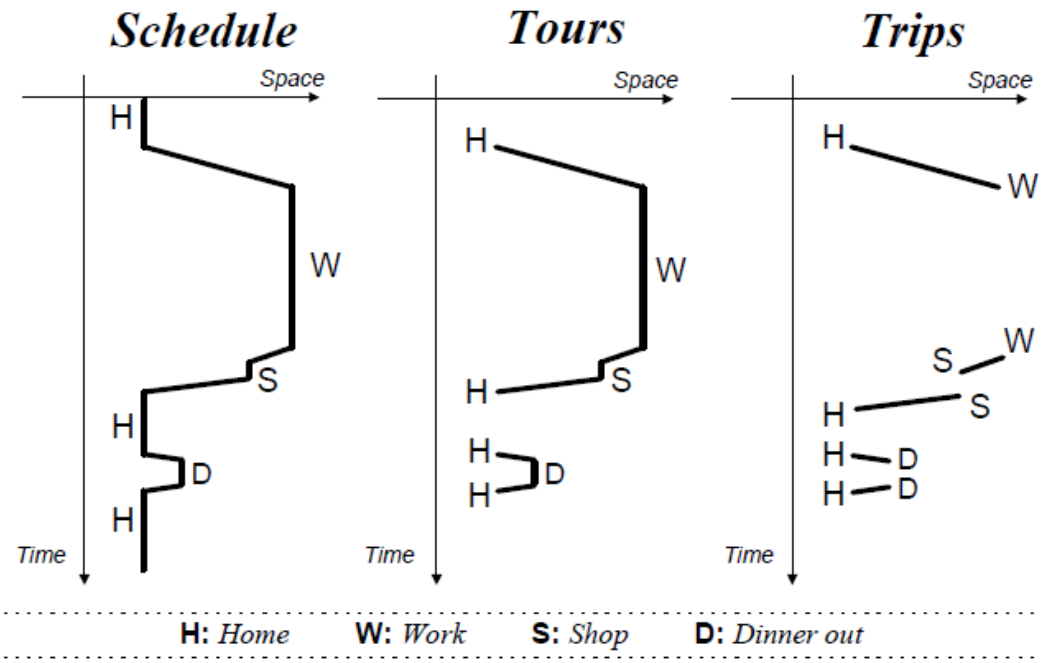
2. Individual Schedule

Schedule

- A/T ID list

Activity/Trip (A/T)

- A/T type ID
- pre A/T ID
- next A/T ID
- {schematic attribute}
- {(x, y, t)}



Ben-Akava (MIT courses)

2.2 Approaches

Data Mining	Statistics
Explorative – Dig out the data first, discover novel patterns and then make theories.	Confirmative – Provide theory first and then test it using various statistical tools.
Involves Data Cleaning	Statistical methods applied on Clean Data
Usually involves working with large datasets.	Usually involves working with small datasets.
Makes generous use of heuristics think	There is no scope for heuristics think.
Inductive process	Deductive (Does not involve making any predictions)
Numeric and Non-Numeric Data	Numeric Data
Less concerned about data collection.	More concerned about data collection.

https://www.dezyre.com/article/data-mining-vs-statistics-vs-machine-learning/349

2.2 Approaches

1. Statistics

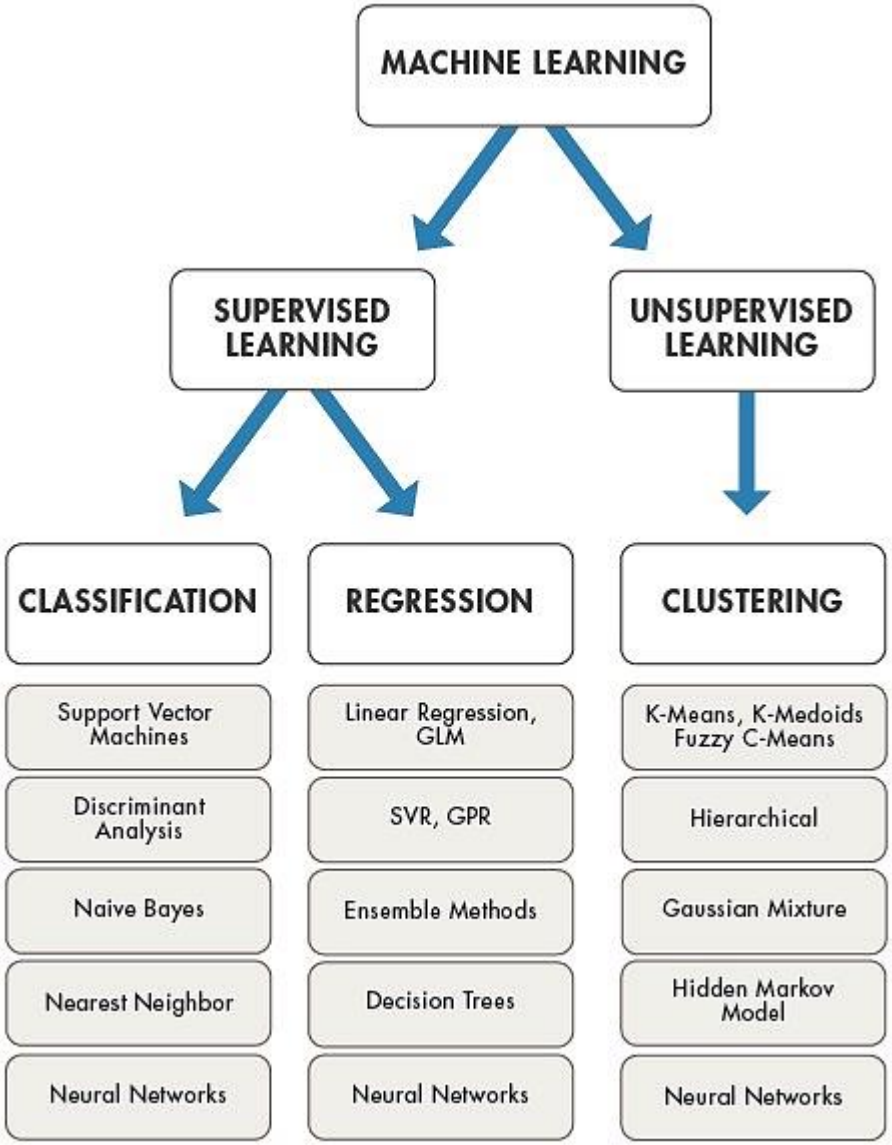
- quantifies data from sample
- estimates population behavior

2. Data mining

- finds out pattern in data

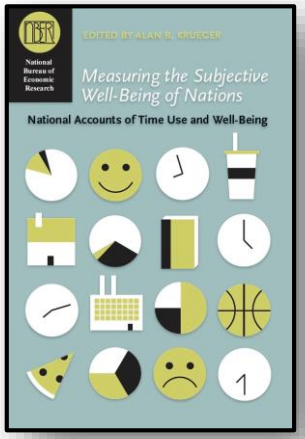
3. Machine learning

- learns from training data
- predicts or estimates future



Bhattacharjee: https://medium.com/technology-nineleaps/popular-machine-learning-algorithms-a574e3835ebb

Example Study: subjective wellbeing



The Currency of Life (Krueger et al. 2009)

Economics time use, goods, and utility

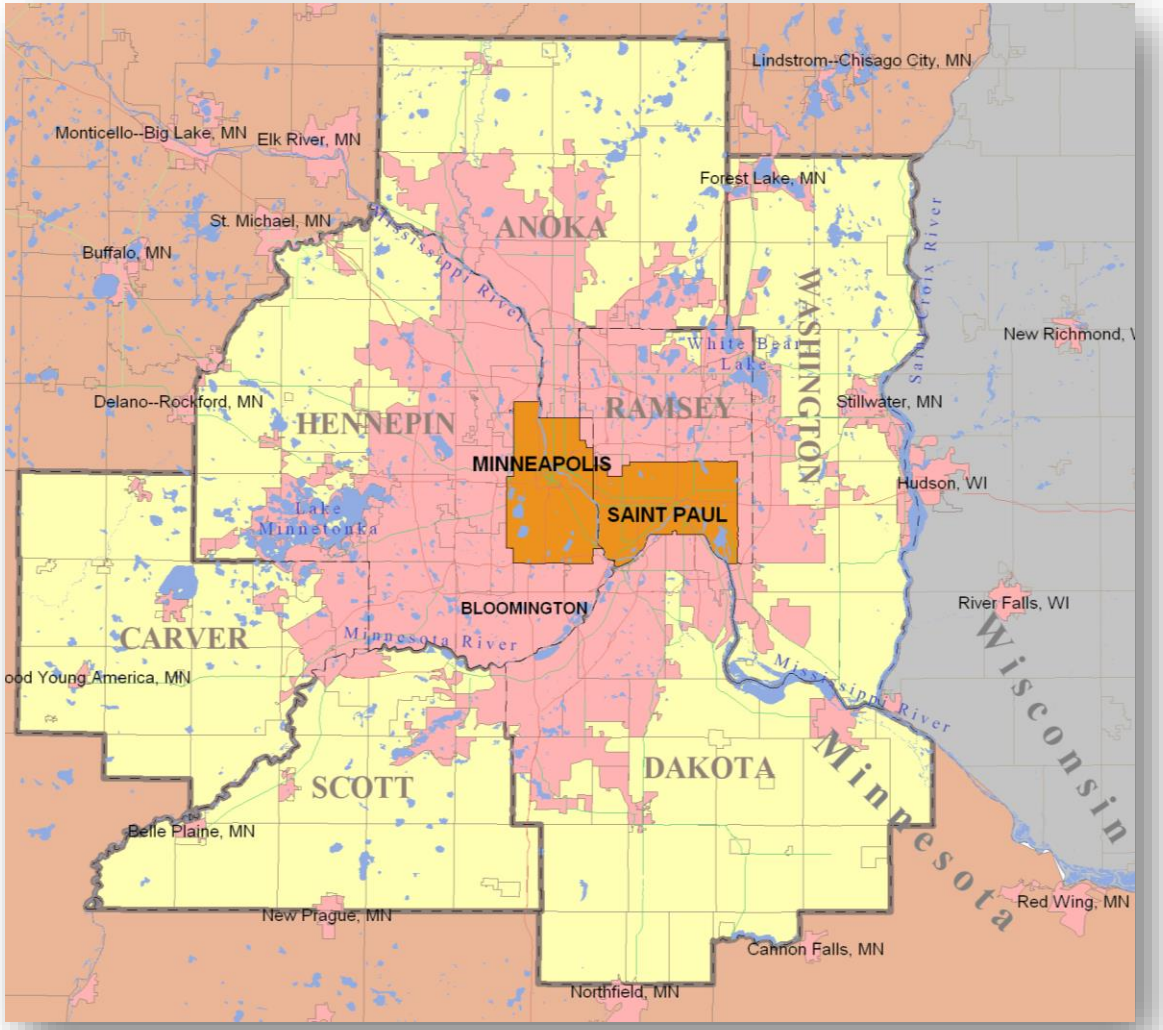
Psychology subjective well- being

Measure U-Index a misery index of sorts

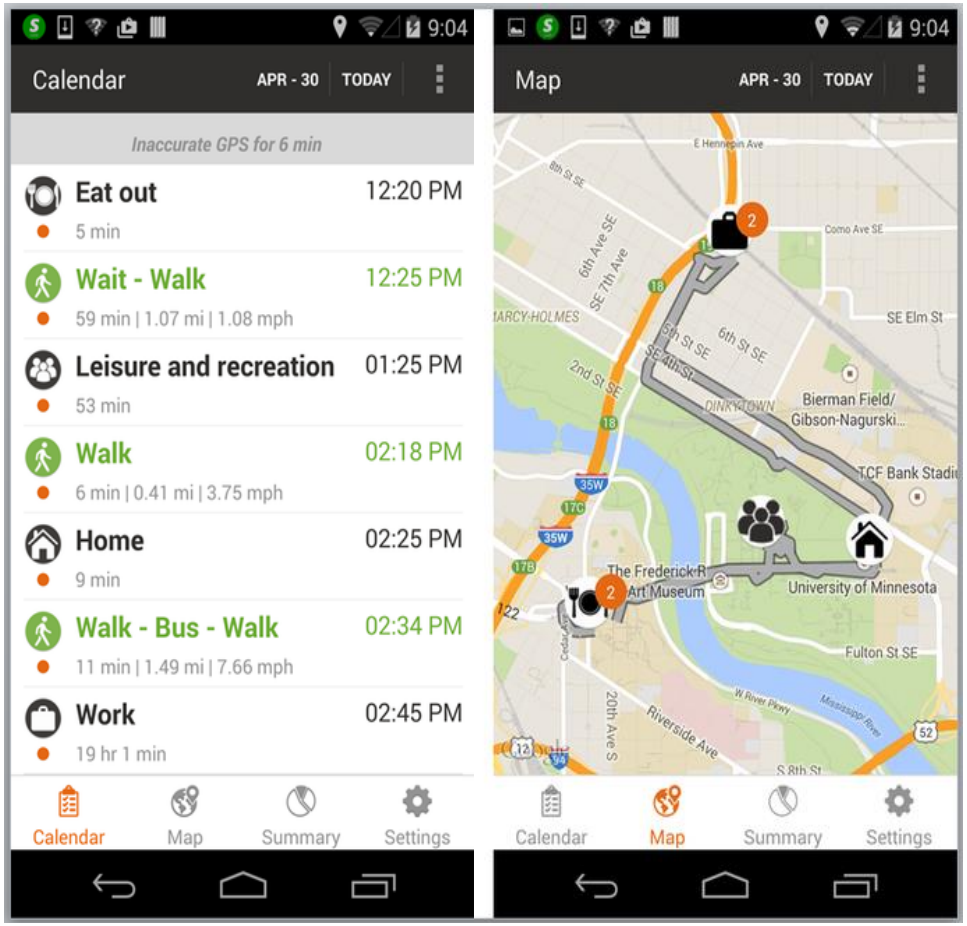
0. Data and Study Area

- Twin Cities Metro Area
- 373 users
- 1 week period
- 25,699 activities/trips
- 6 emotions, level 1-7

  - Happy
  - Meaningful
  - Sad
  - Tired
  - Stressful
  - Painful



1.1 Data Representation



Activity Type	Trip Type
Home	Car
Work	Bus
Education	Rail
Personal business	Shuttle
Shop	In vehicle
Eat out	Walk,
Leisure/Recreation	Bike
Other/Unknown	Wait
	Other/Unknown

Schematic Attributes 6 emotions

Happy	Sad	Stressful
Meaningful	Tired	Painful

(7 levels of strength; average by activity/trip)

GPS tracking data (3 second)

1.2 Data Structure (in process)

Double-linked list

```
class ActNode:
    def __init__(self, node_id=None, node_tuple=None):
        if node_id and node_tuple:
            self.aID = node_id
            self.aTuple = node_tuple
            self.nextNode = None
```

[activity/trip]
- a tuple of values

```
class OrderedActList:
    def __init__(self):
        self.head = None
```

[activity sequence]
- ordered linked list

```
class AppActivitySpace():
    def __init__(self):
        self.tripDict = {}
        self.actDict = {}
```

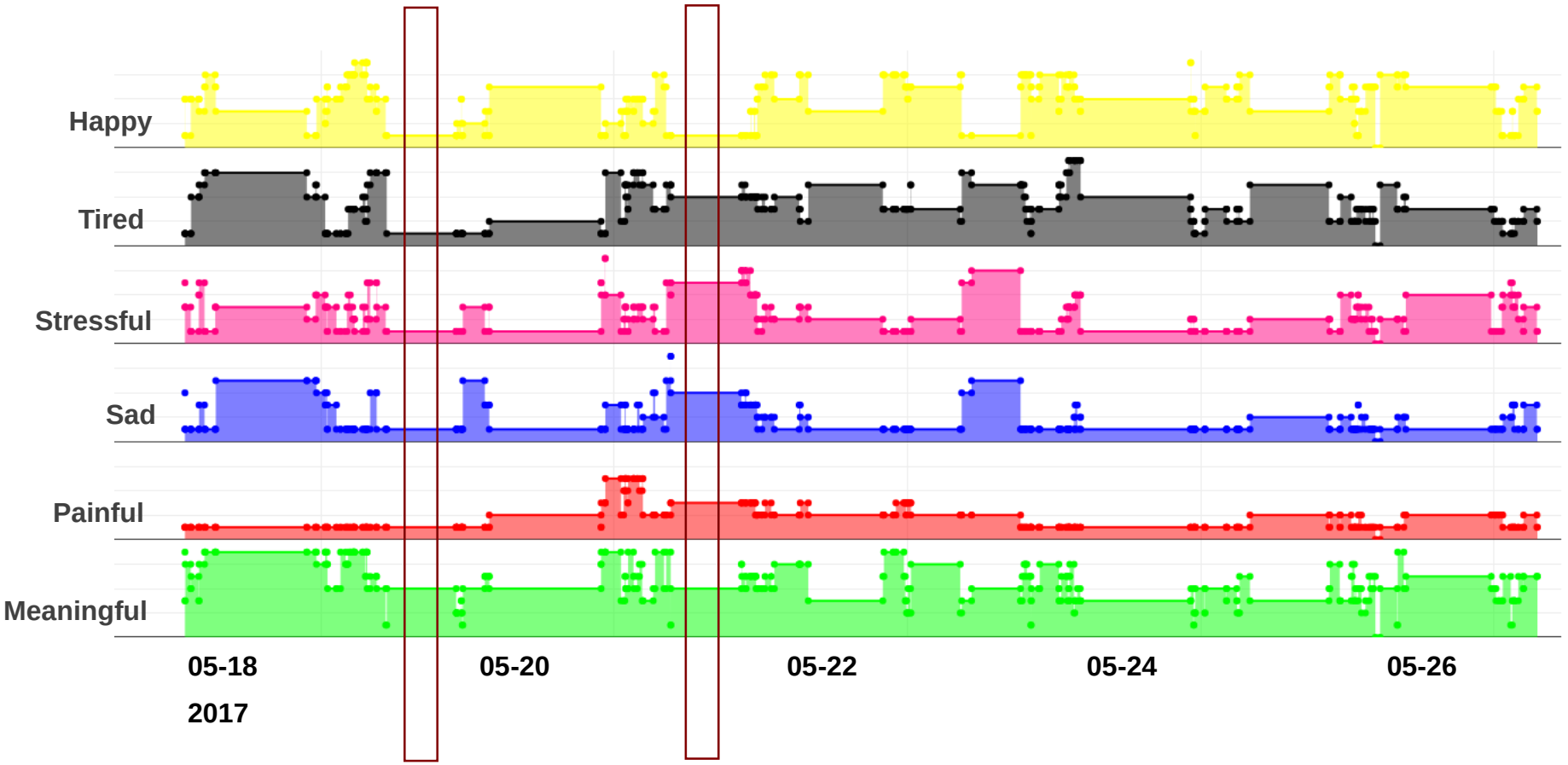
[activity types]

```
class AppUser():
    def __init__(self, uid=None, act_list=None):
        if uid and isinstance(act_list, OrderedActList):
            self.uID = uid
            self.actList = act_list
        elif not uid:
            raise ValueError("input user ID cannot be NONE")
        else:
            raise TypeError("input activity list is not an OrderedActList data type")
```

[user]
- with activity list

2.1 Analysis – Visual Exploration

1. Emotion status across time



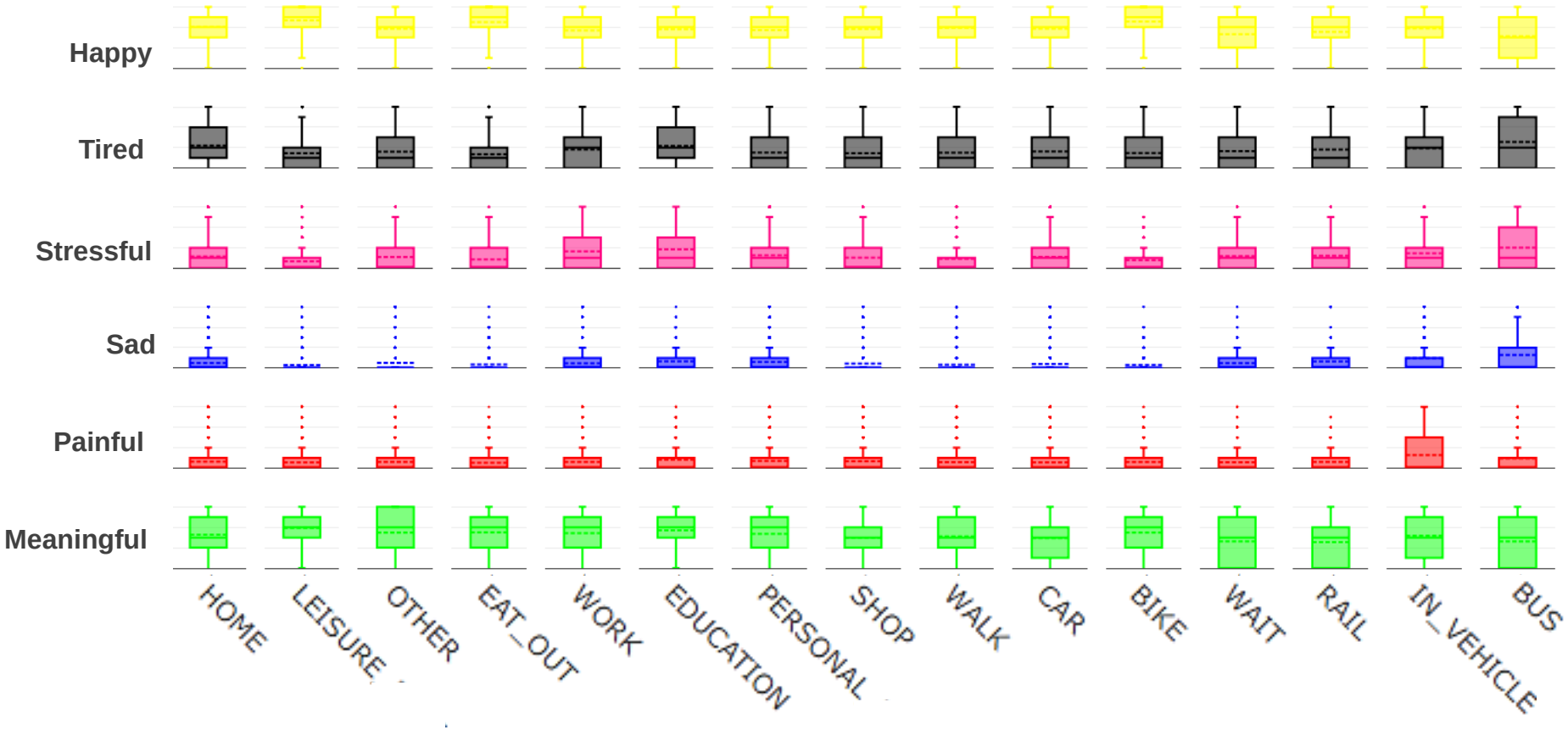
Happiness is independent from other five emotions?

https://plot.ly/~AliSong7/84.embed

2.1 Analysis – Visual Exploration

2. Emotion status by activity and mode

Mean & Variations

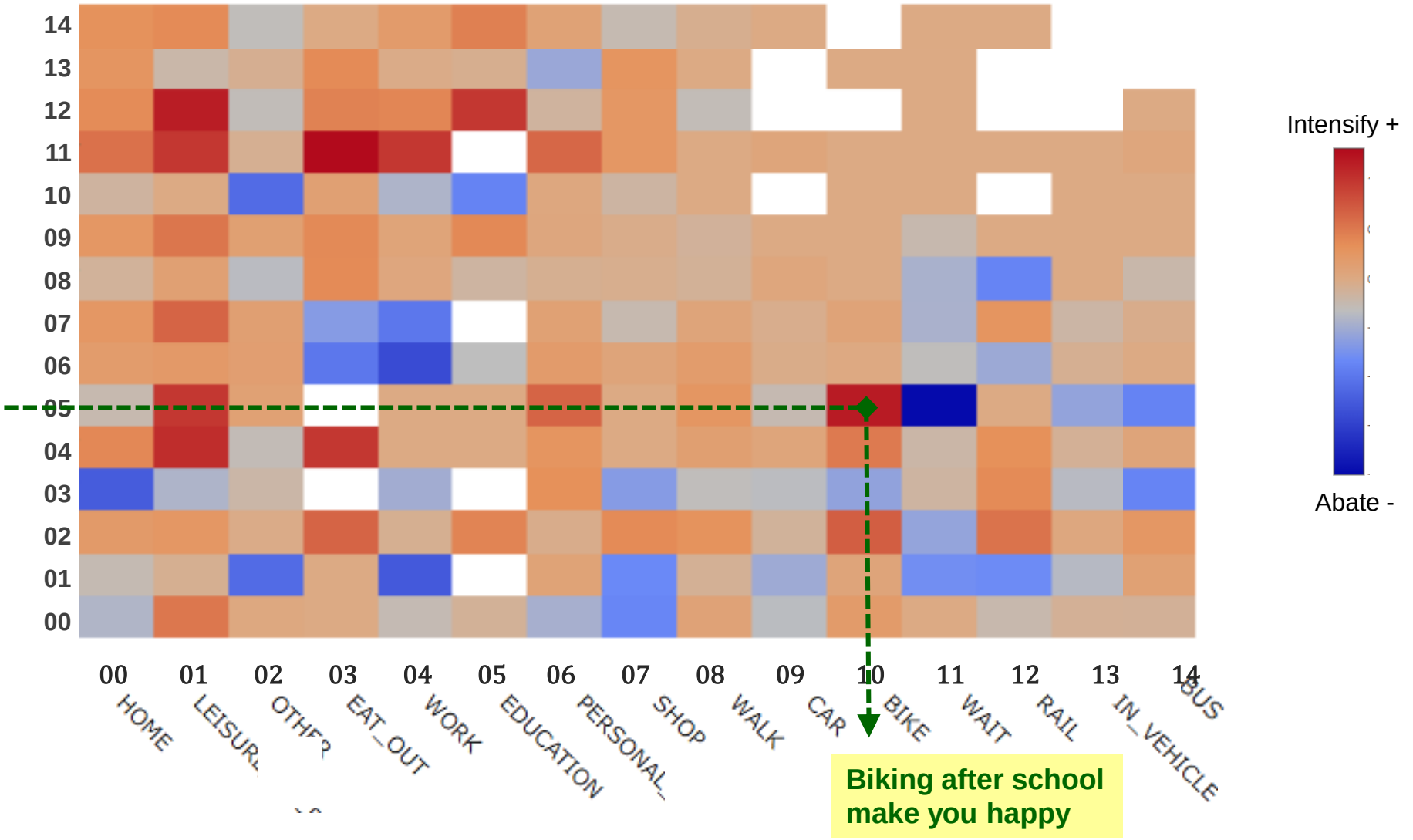


https://plot.ly/~AliSong7/86.embed

2.1 Analysis – Visual Exploration

3. Transition matrix for emotion changes

https://plot.ly/~AliSong7/92.embed

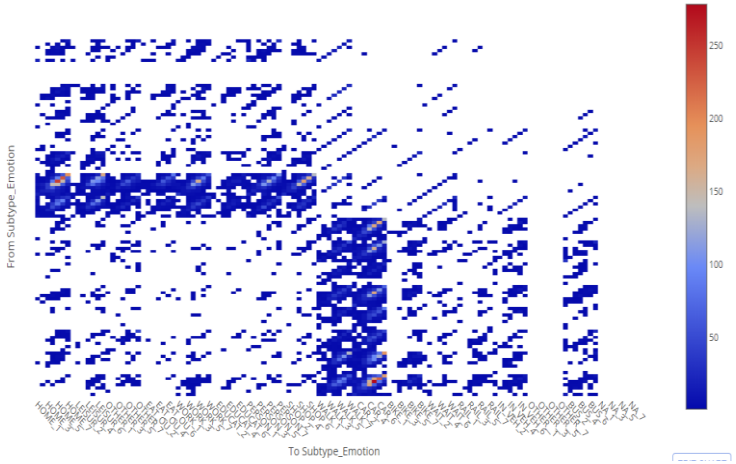


2.2 Analysis – time series analysis

(1) Step Patterns - Markov models

- **Continuous-time semi-Markov**
  jump rate & holding times

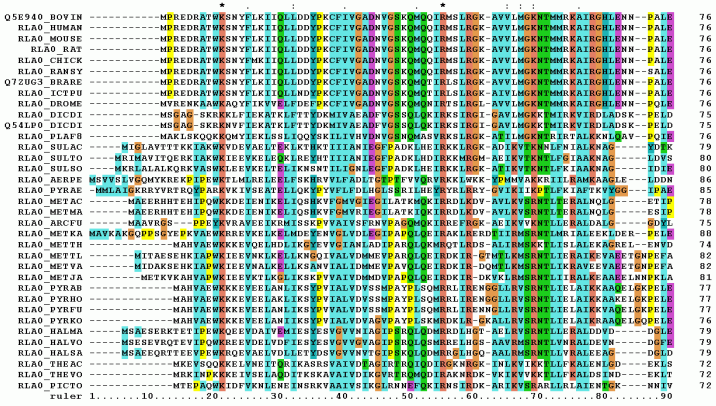
- **Hidden Markov chain**
  emotions ⇔ activities and modes



(2) Sequential Patterns - Path alignment and clustering

- **One-dimension**
  behavior or emotion; feature-based clustering

- **Multidimension**
  sequences of tuples (behavior, emotion)



Ying Song yingsong@umn.edu