Days of Our Lives

Assessing Day Similarity from Location Traces

James Biagioni and John Krumm

Microsoft® Research

Days of our Lives

Days of their lives



They’re still doing this?!?



“Ordinary people”



Going to work...



Kids’ soccer game...



Shopping…



Spending time at home...



Days and similarity

1) Find representation of days
   - Encoding of activity sequences

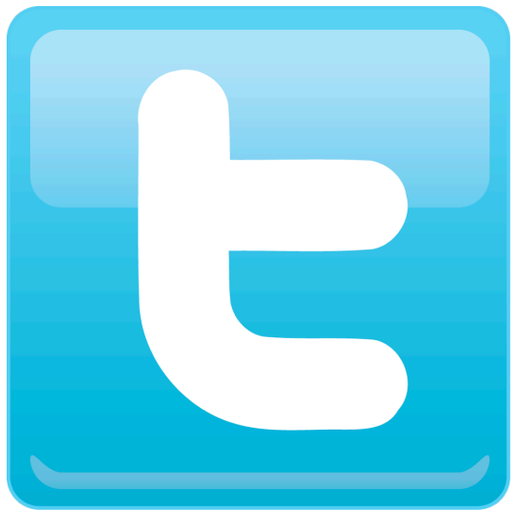
2) Develop “day similarity” metric
   - Corresponding to people’s notion of *similar*

Why?

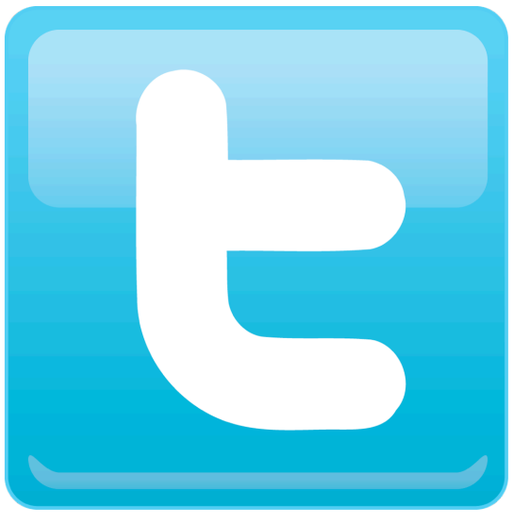
- Identification of anomalous days
  - May indicate confusion (important to detect)
  - Change of habits
- Clustering of similar days
  - Assess variety
  - Prediction of how days will evolve

Encoding a day

Encoding a day

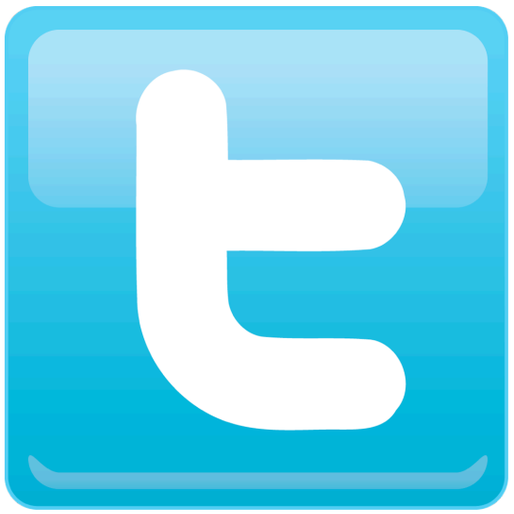


Encoding a day





Encoding a day

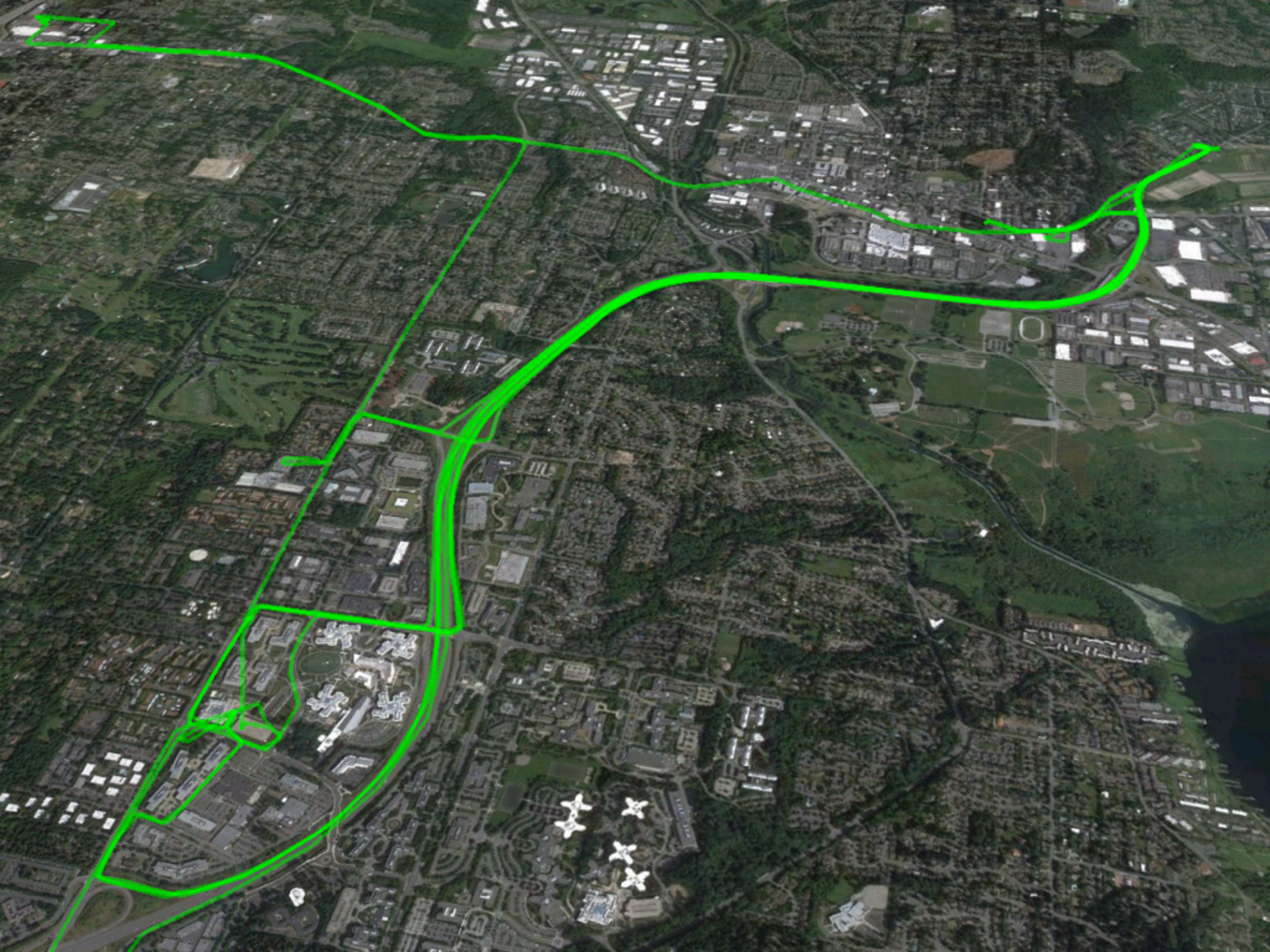






The good and the bad

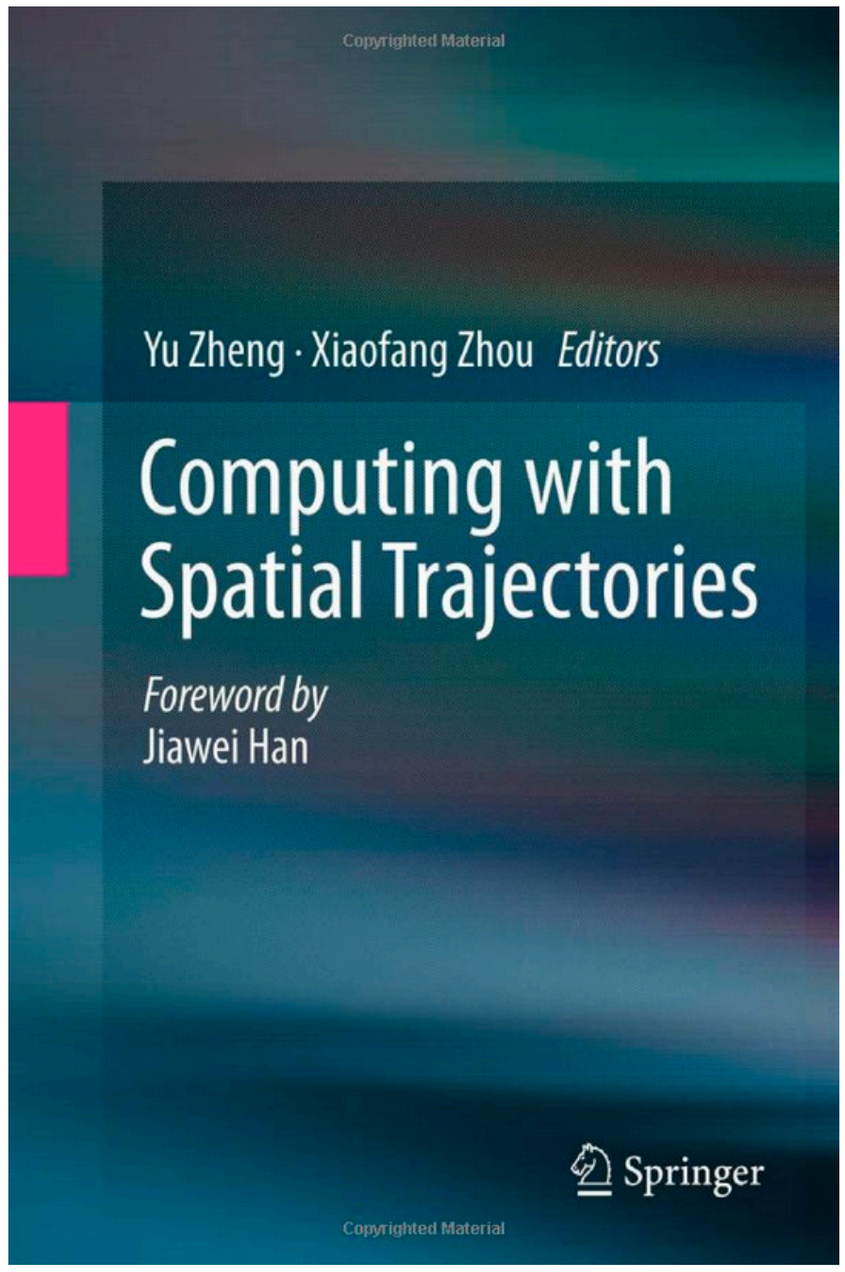
- The good
  - Excellent at isolating activities
- The bad
  - Impose heavy burden on users




Location is...
1) Constantly existent state
2) Almost always measurable

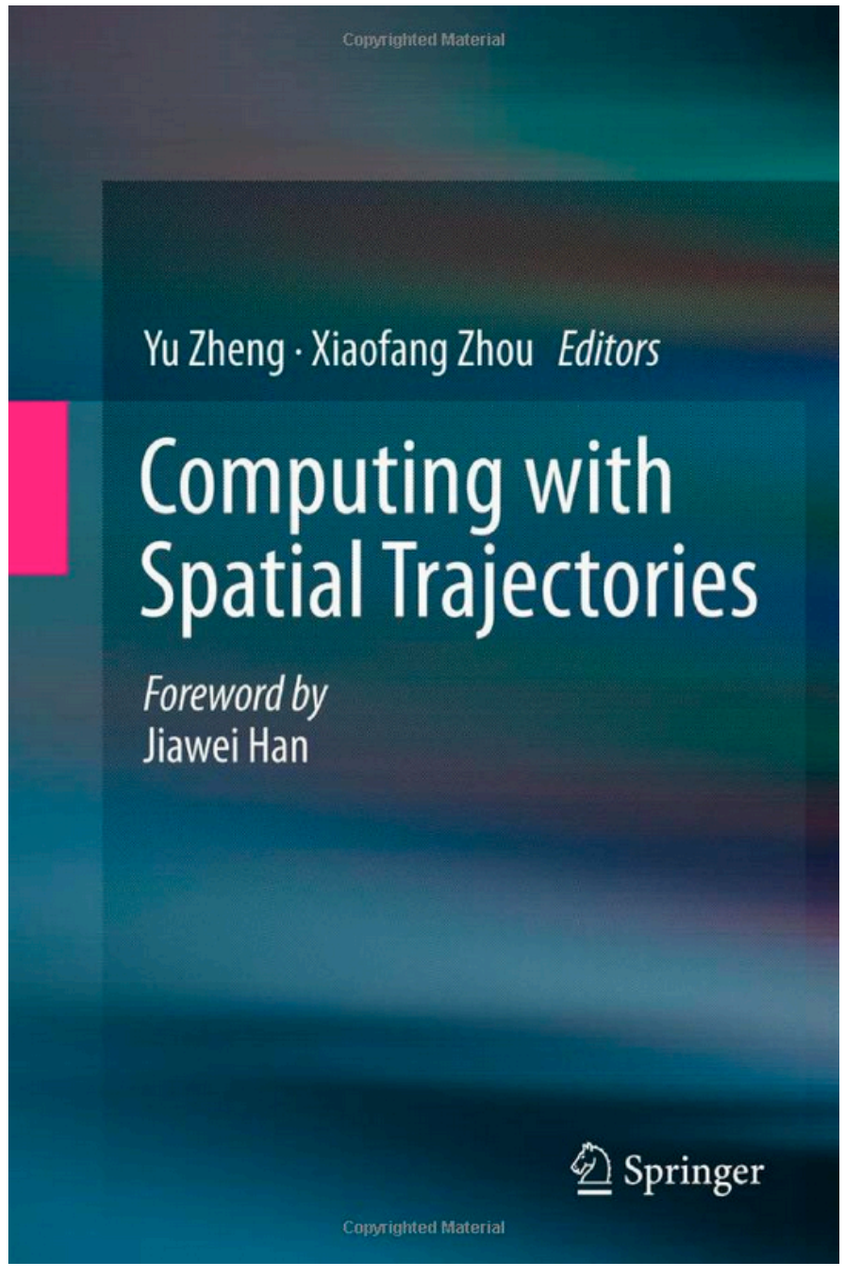
Prior work

- Location trace similarity
  - Machine processing



Prior work

- Location trace similarity
  - Machine processing
- Day similarity
  - Anomaly detection



Anomaly detection literature

- Giroux et al. (ACM PETRA, 2008)
- Xian and Gong (IEEE ToPAMI, 2008)
- Ma (ACM SIGACCESS, 2009)
- Patterson et al. (ACM UbiComp 2012)

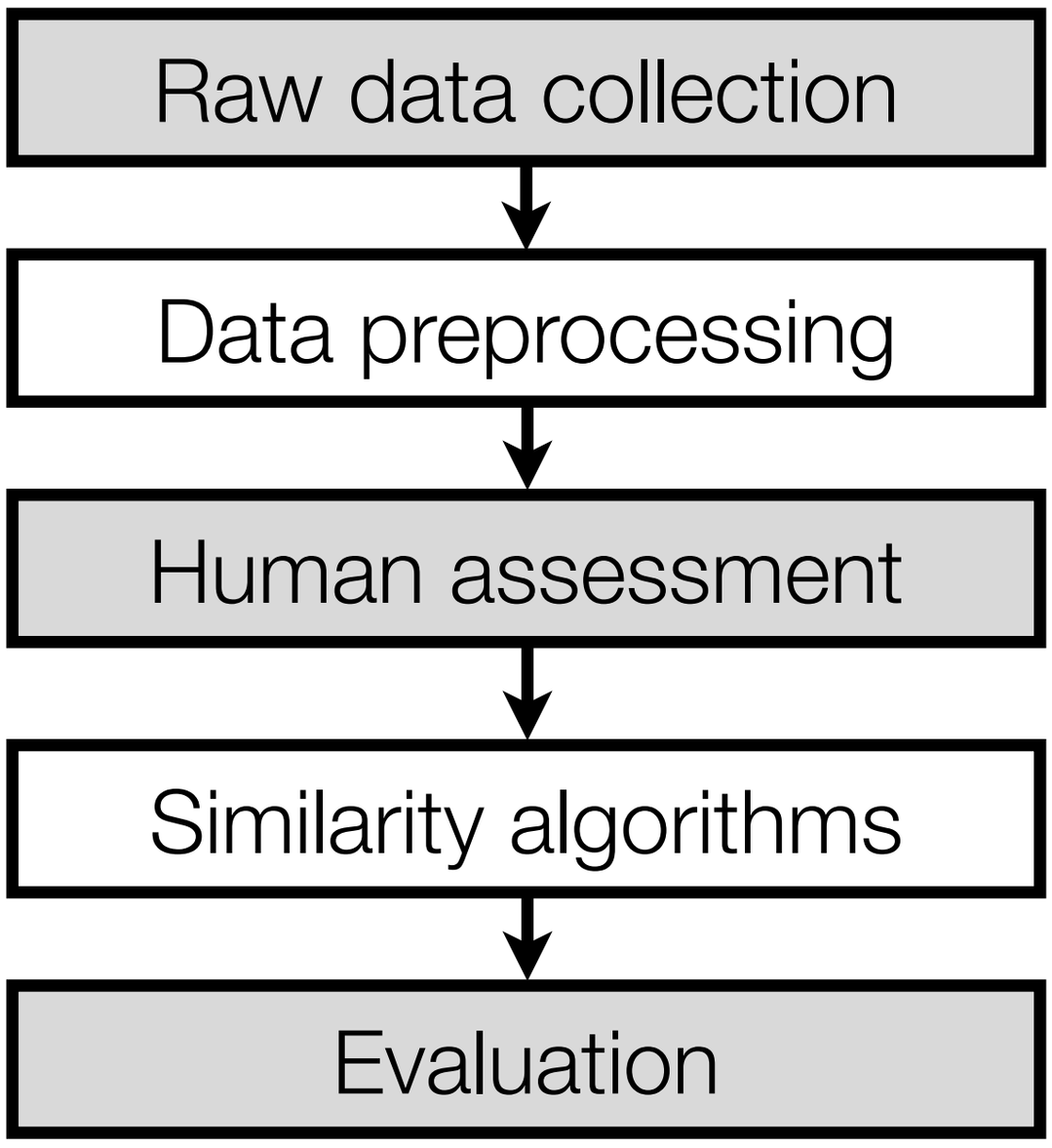
Key features

- Depend on learning “normal” behavior
  - Per-subject training required
- When observations are “similar”
  - Researcher-defined

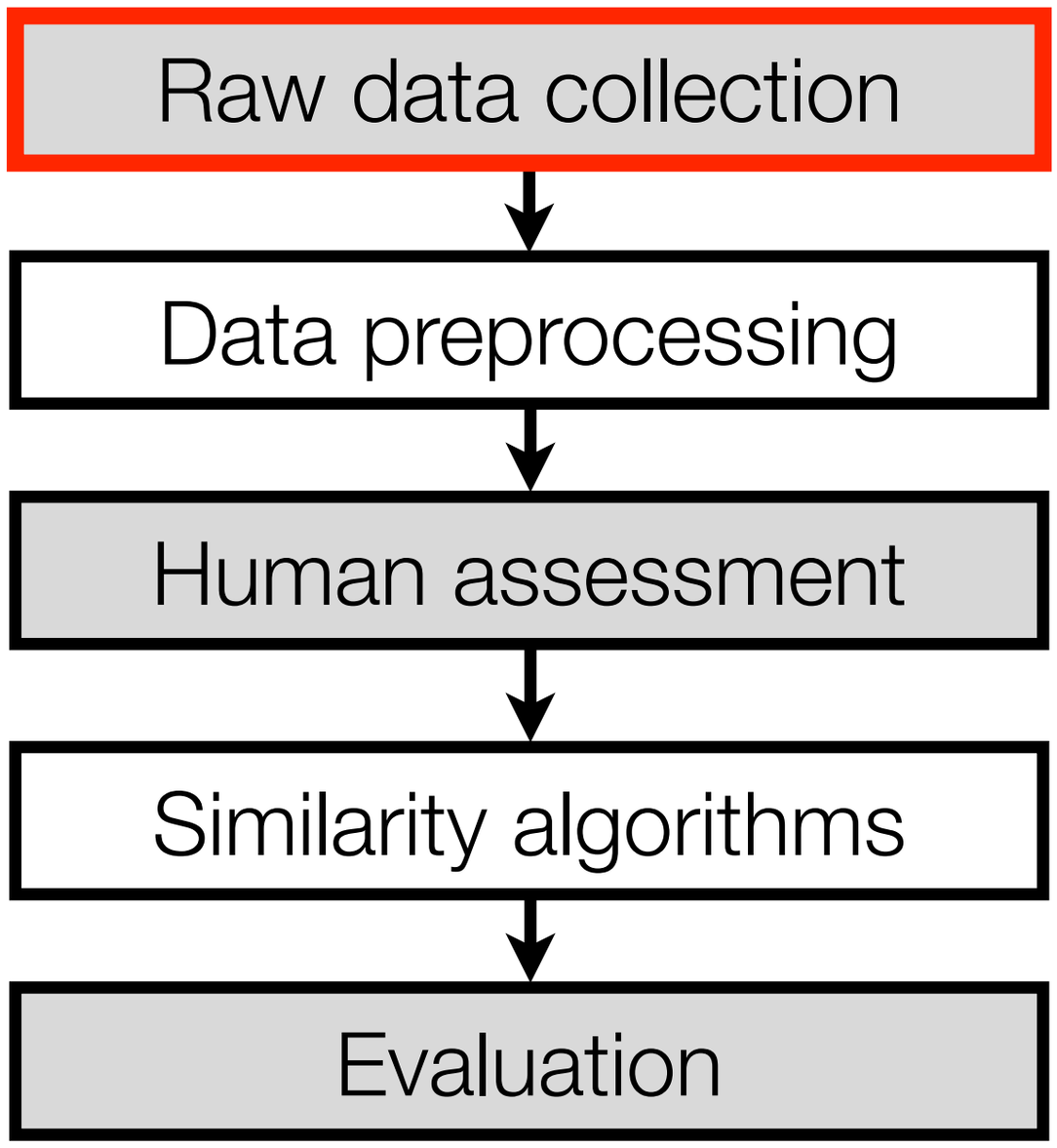
Our goals

- Works “off the shelf”
  - No dependence on training
  - Works well across multiple people
- When behavior is “similar”
  - Human-subject-defined

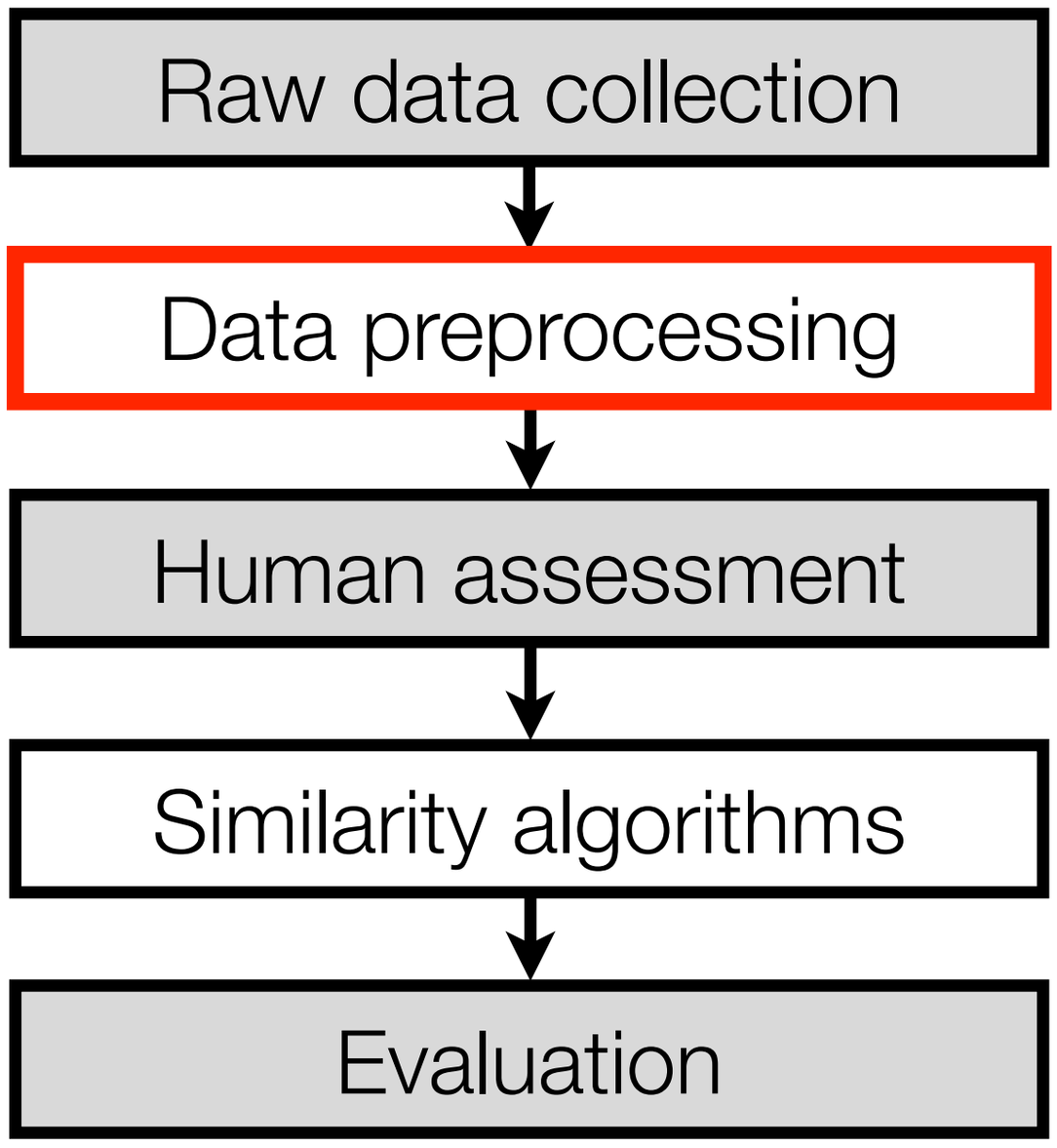
Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

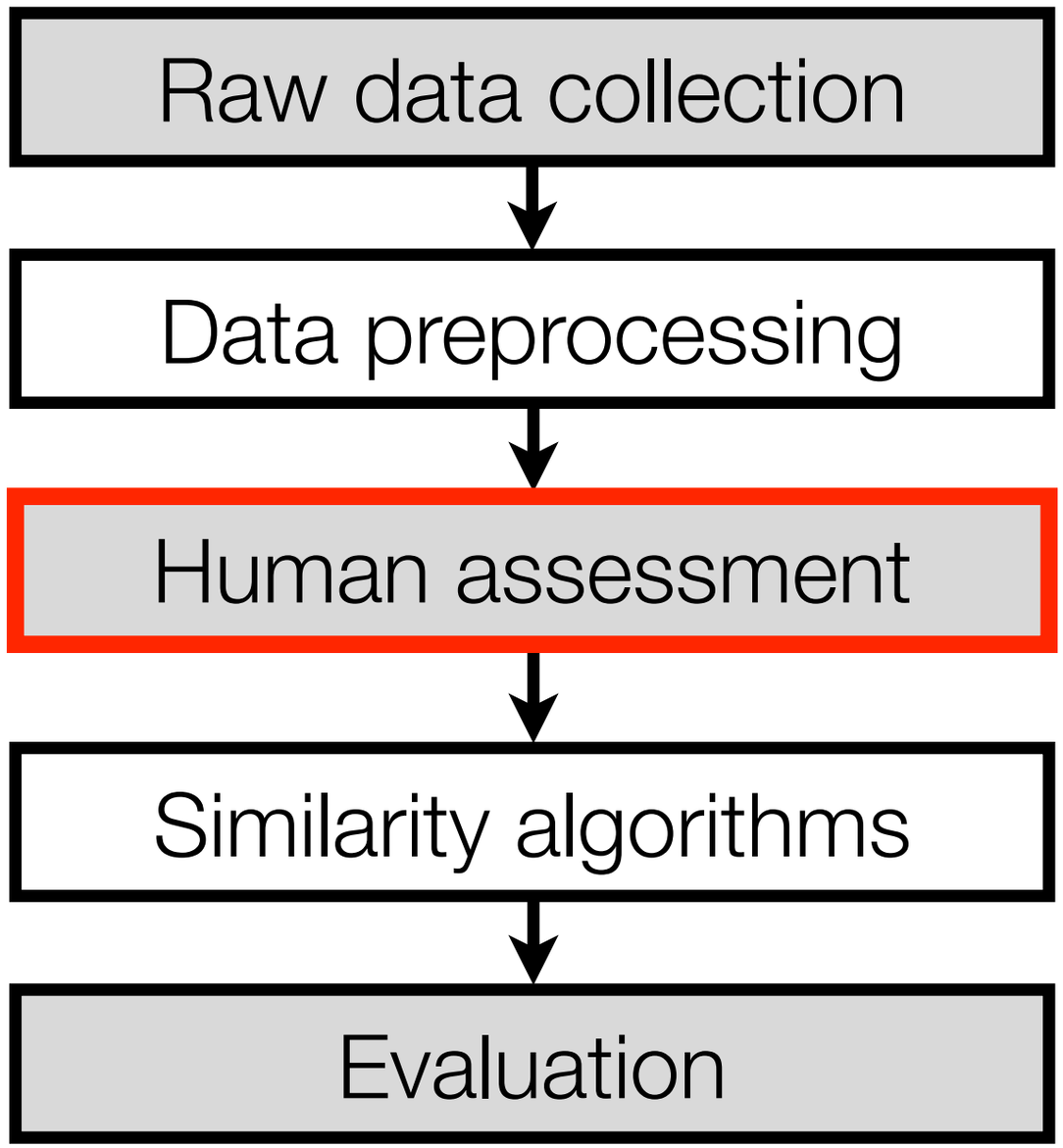
Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

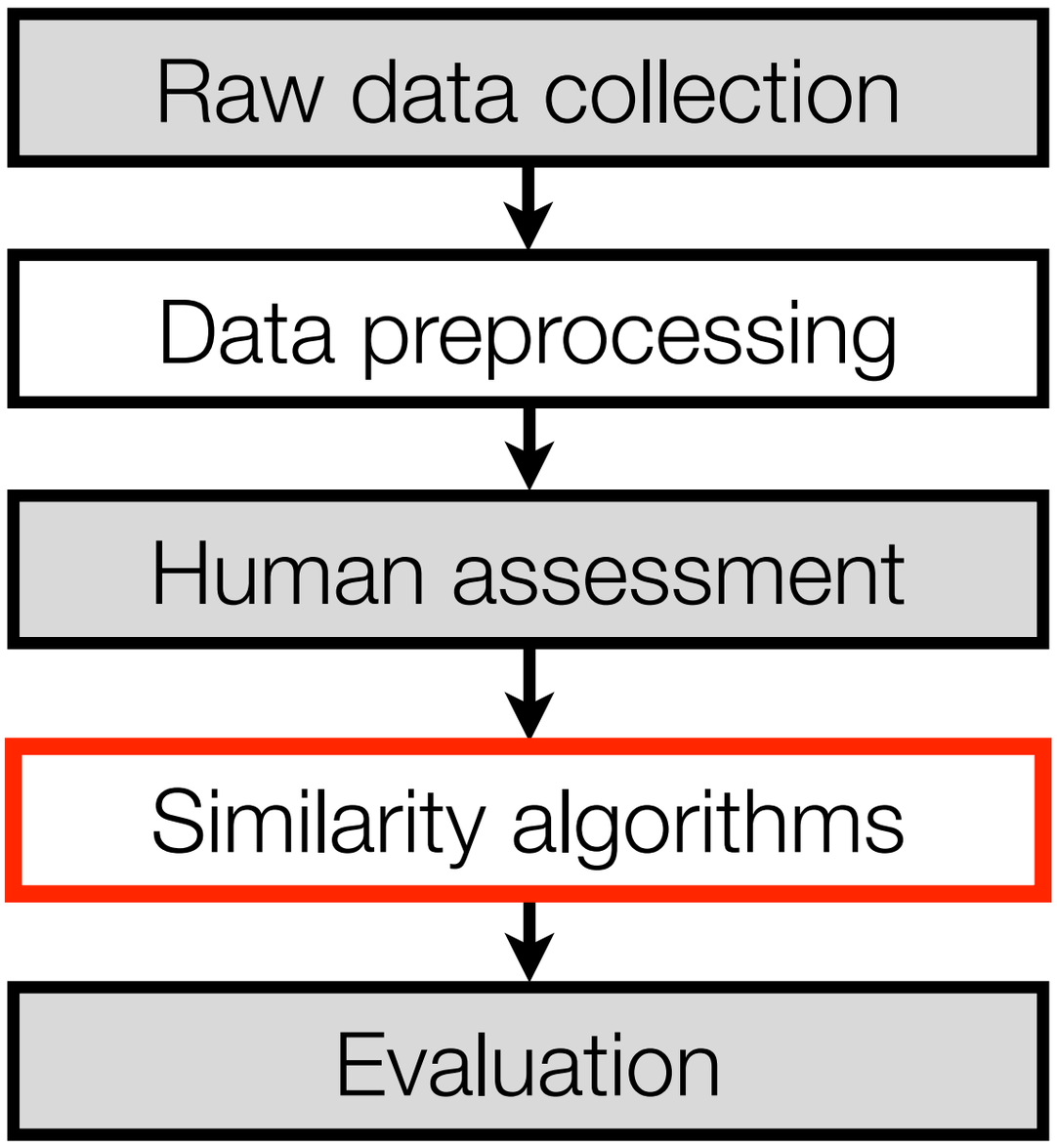
Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

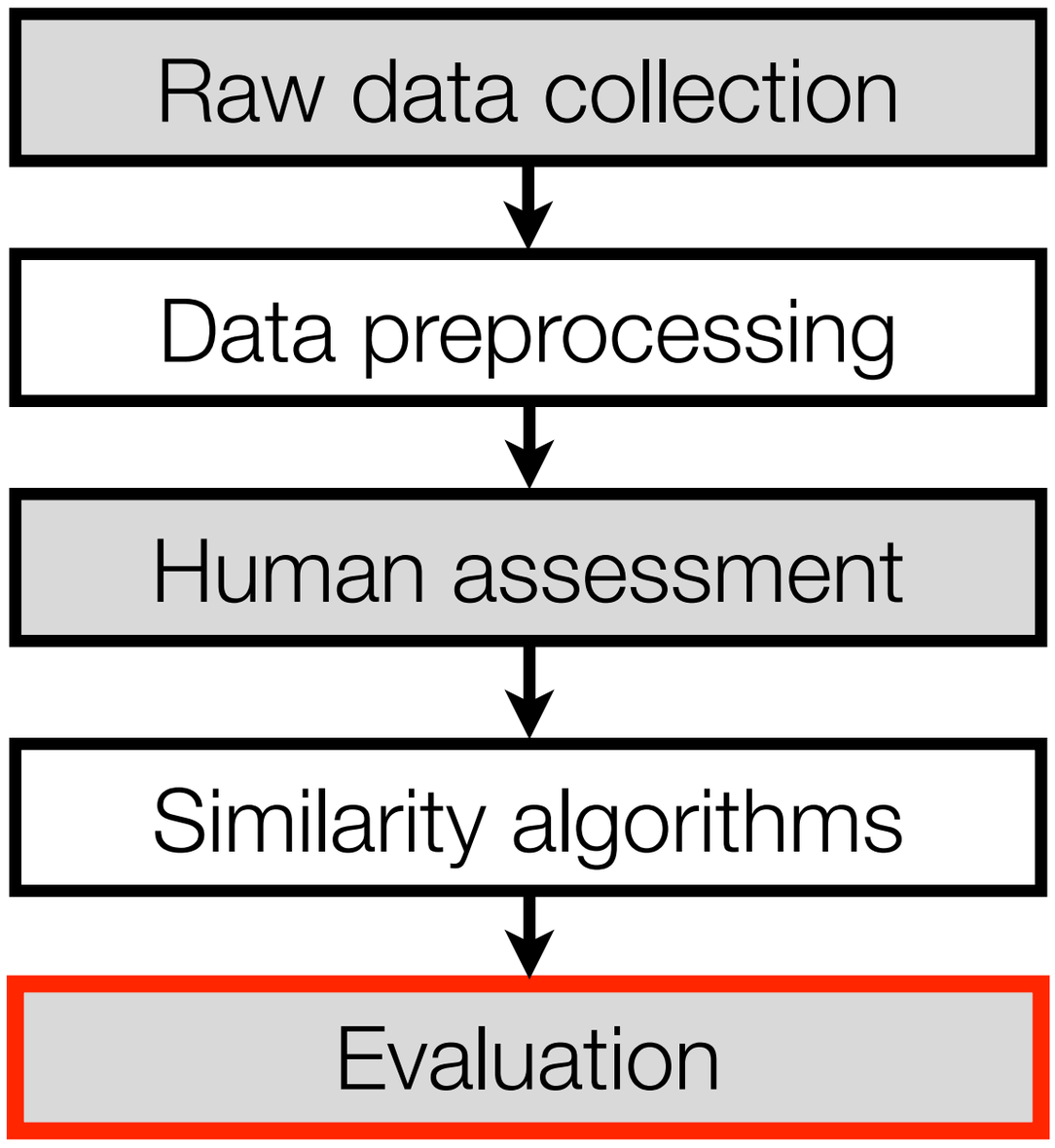
Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

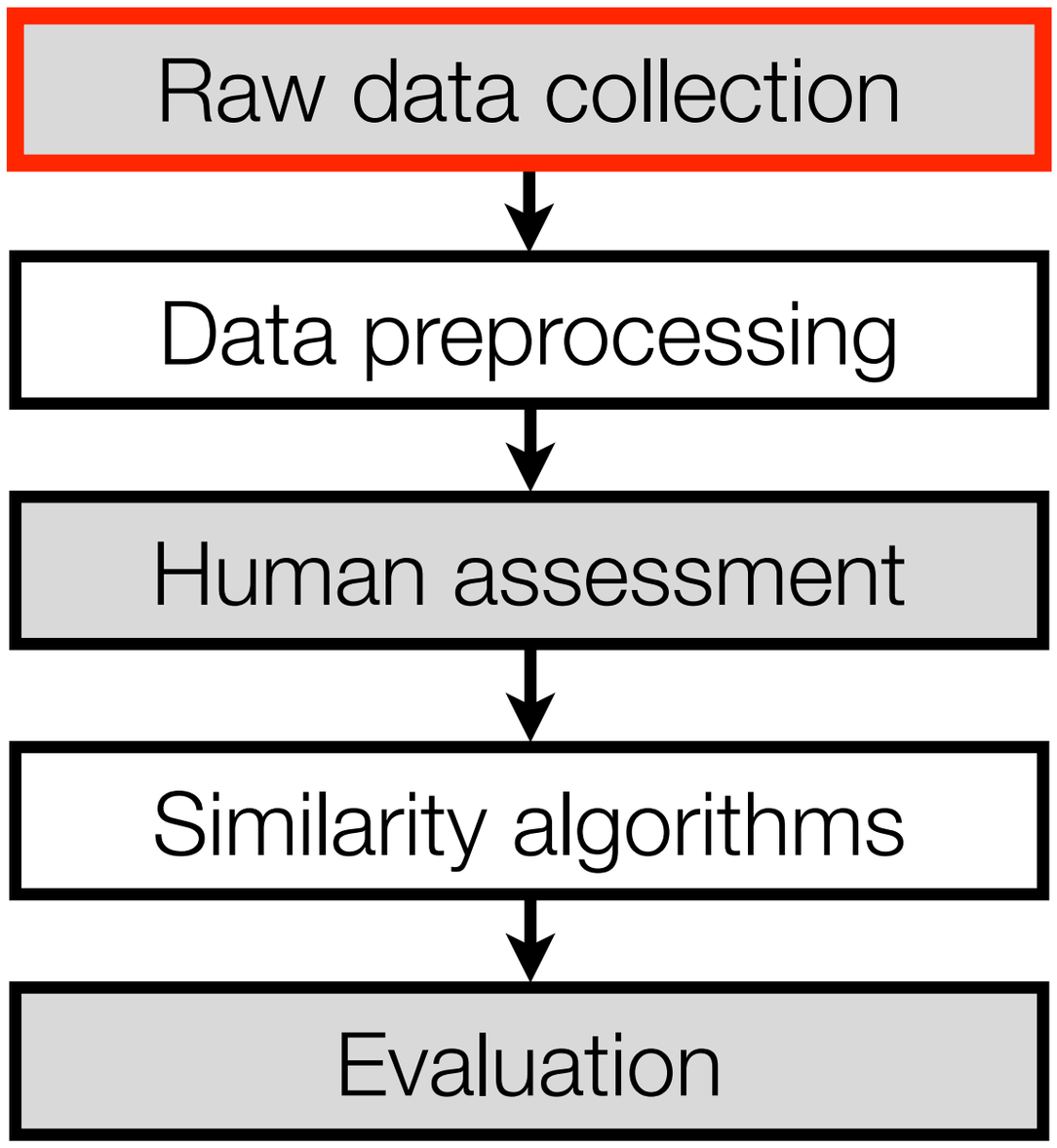
Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Willing volunteers



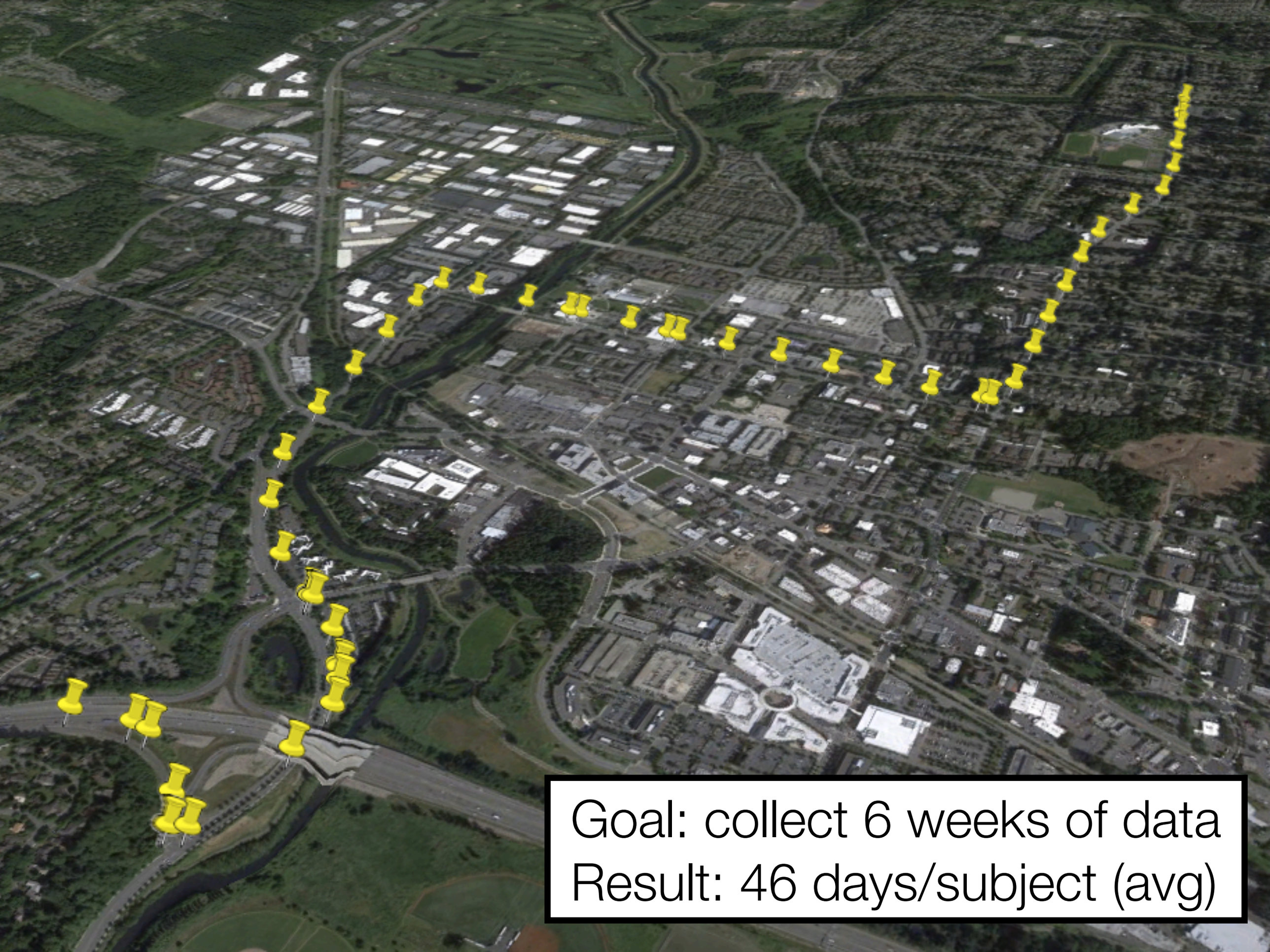
GPS logger

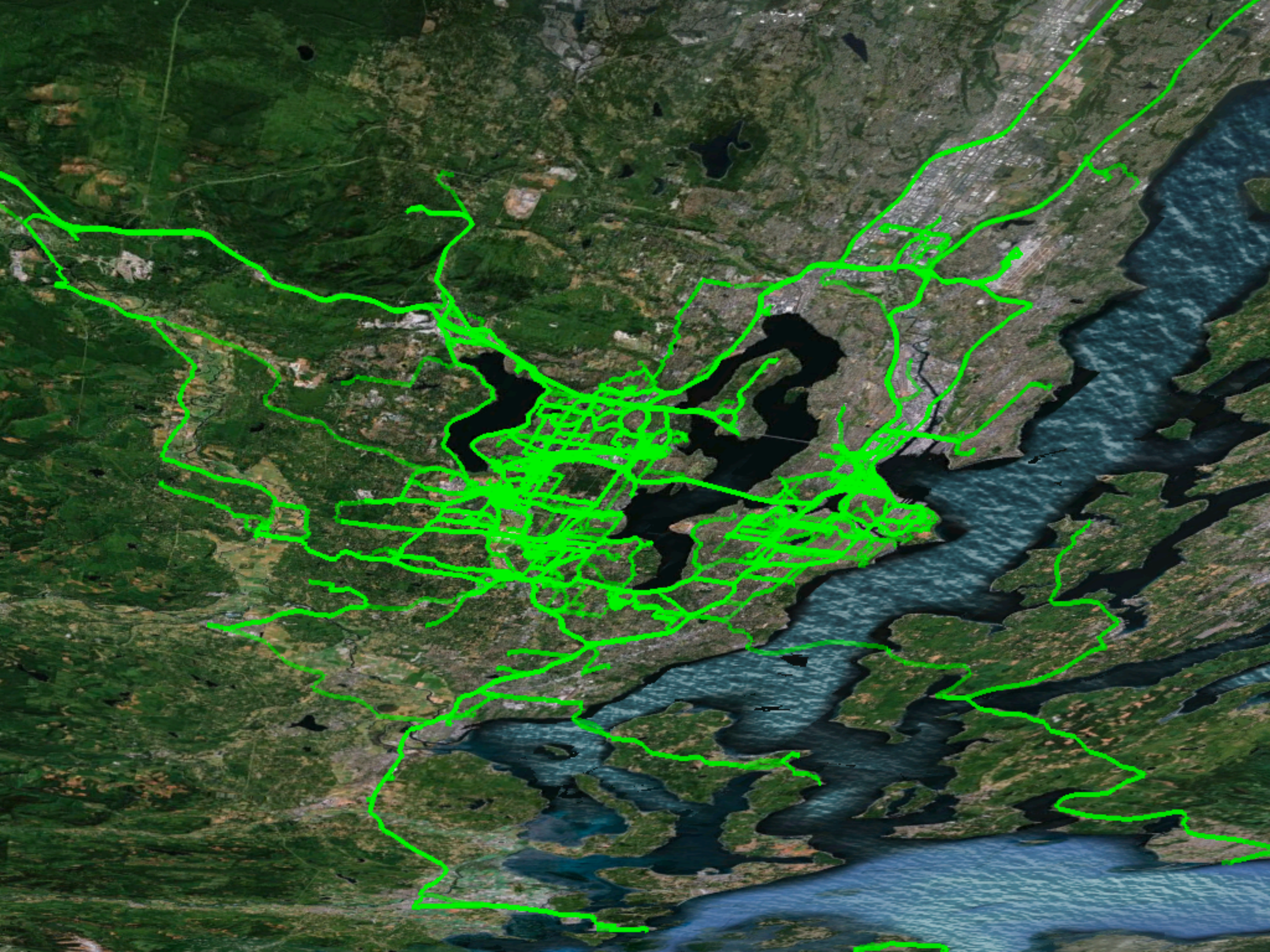


Primary vehicle

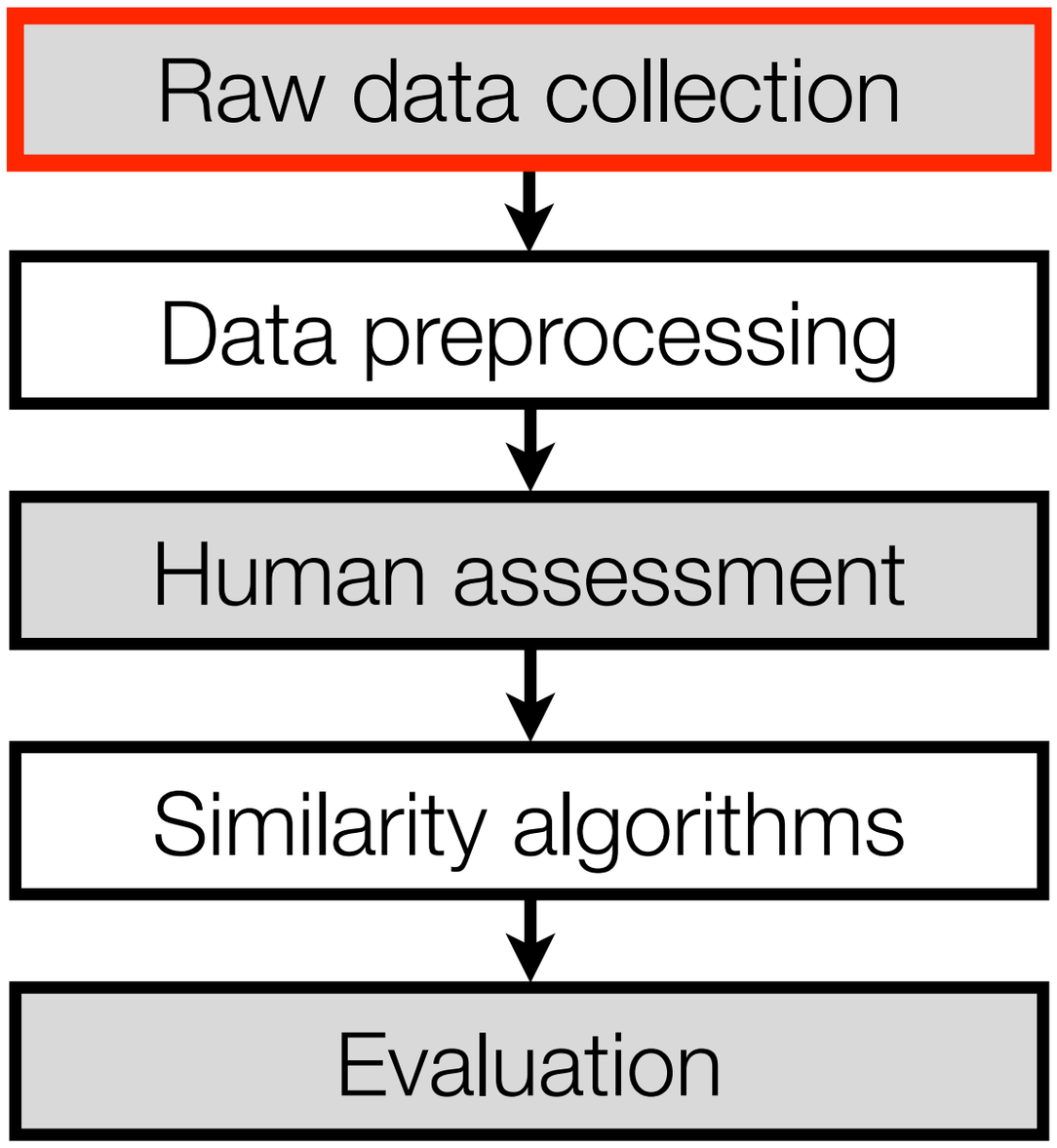




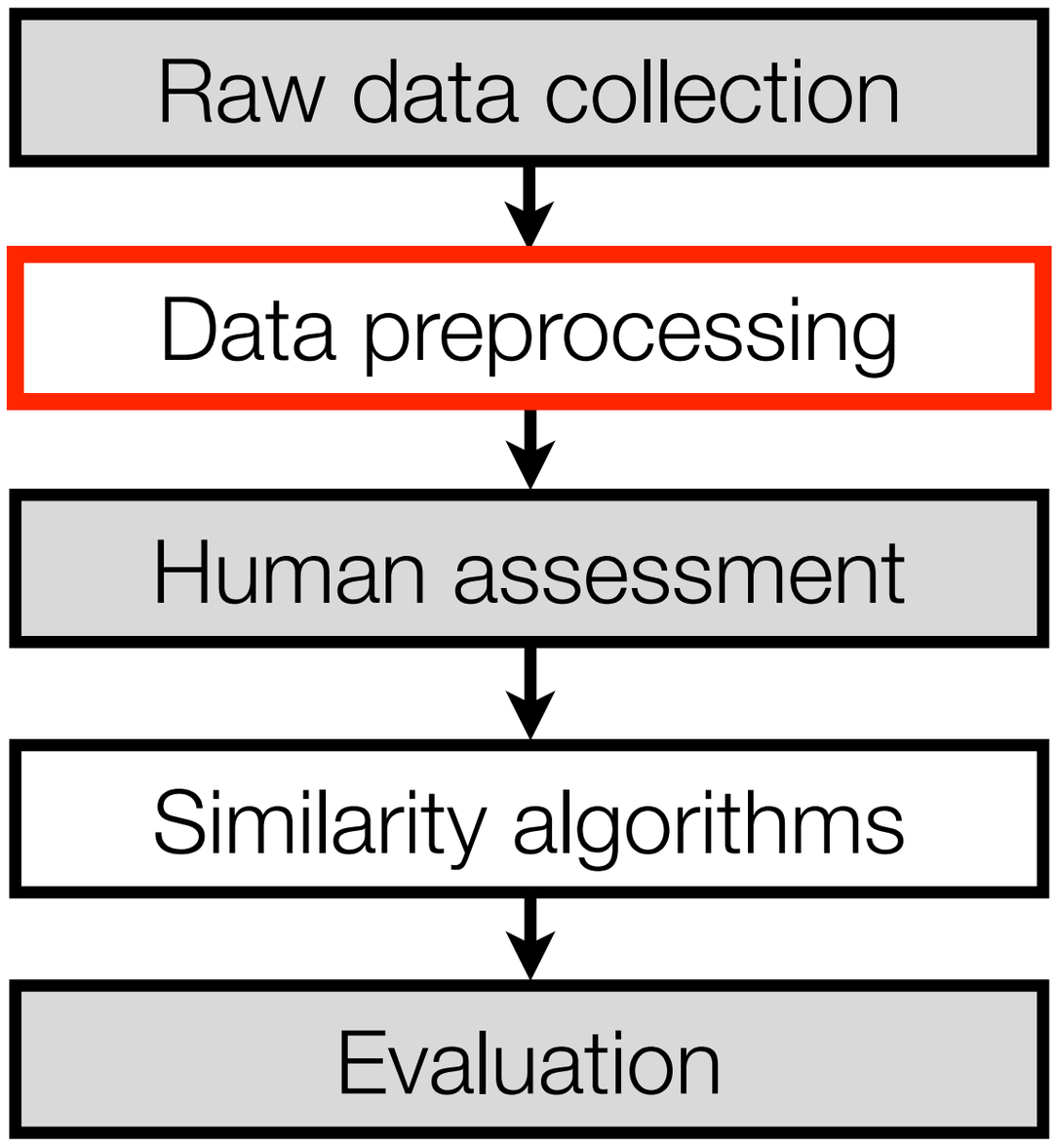

Goal: collect 6 weeks of data
Result: 46 days/subject (avg)



Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

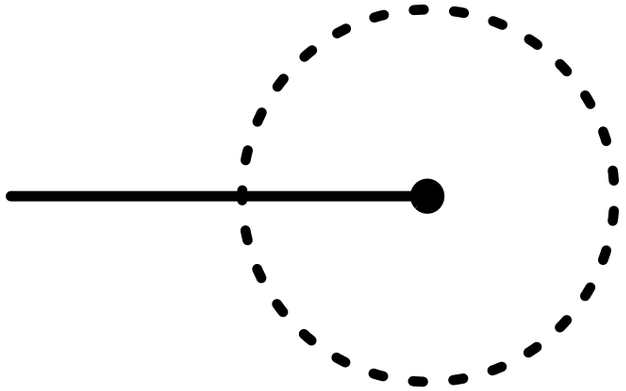
Raw location traces

- Want to represent traces as “days”
  - Sequence of activities
- Where location traces are “stopped”
  - Indicates presence of activity

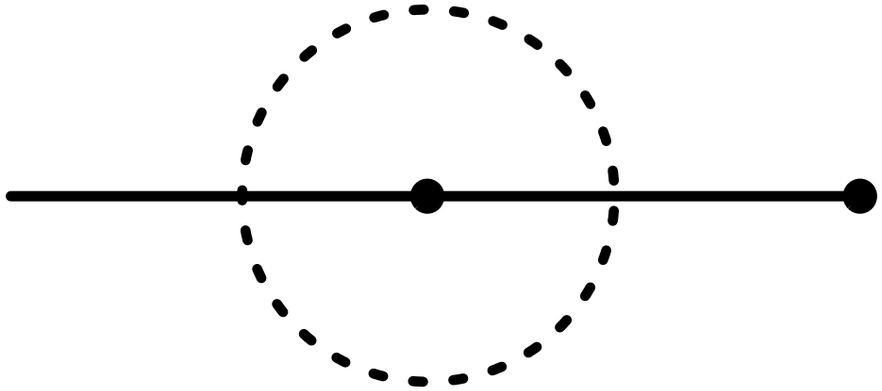
Stop detection



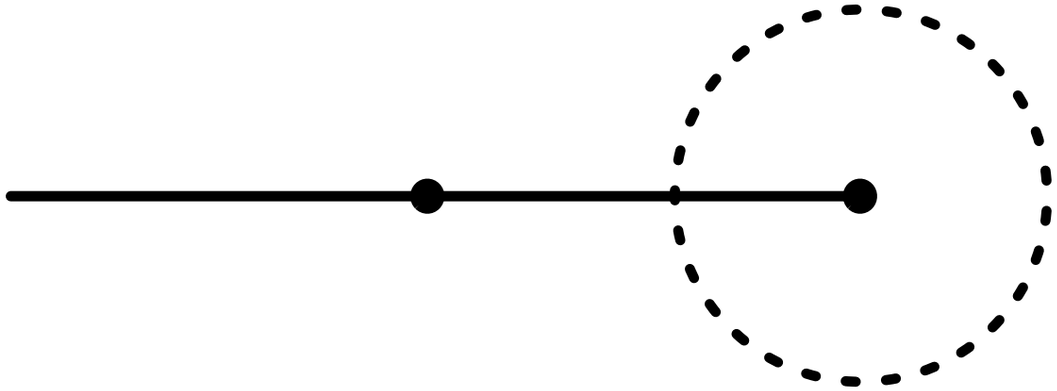
Stop detection



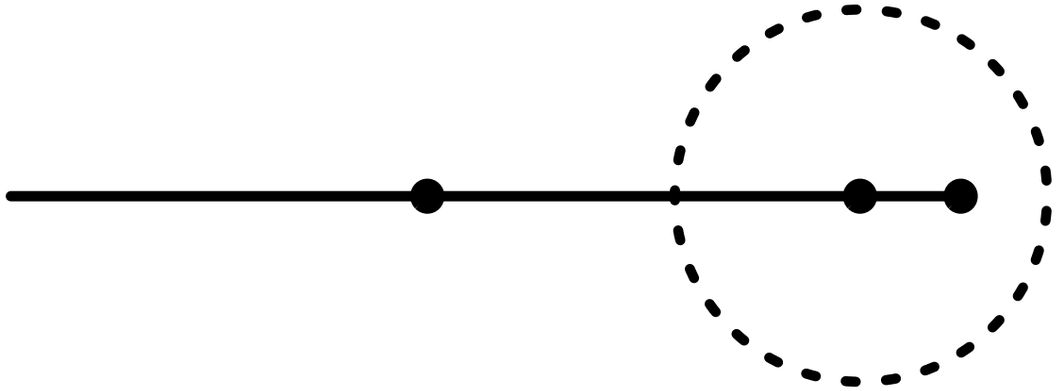
Stop detection



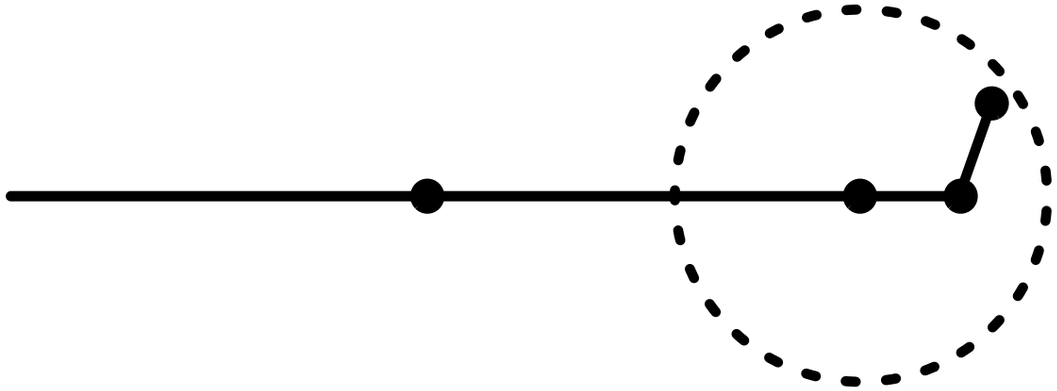
Stop detection



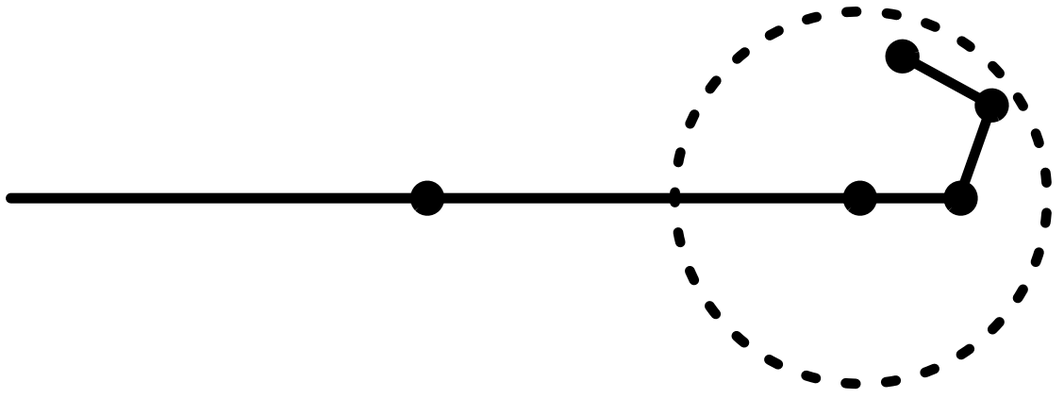
Stop detection



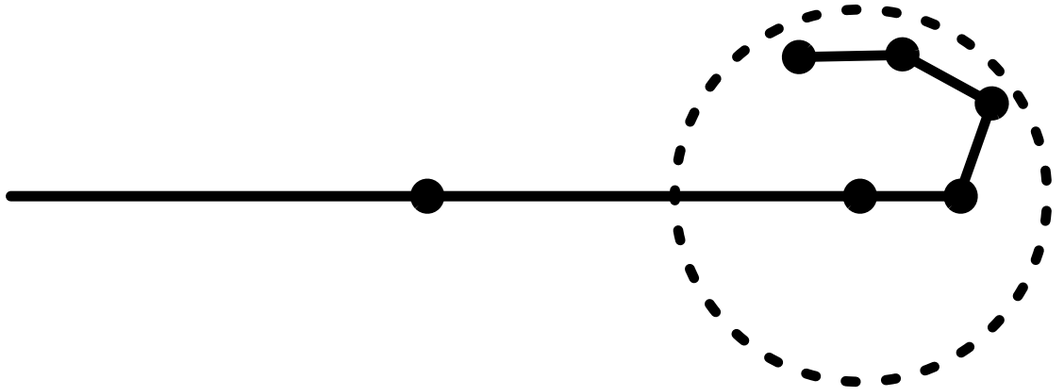
Stop detection



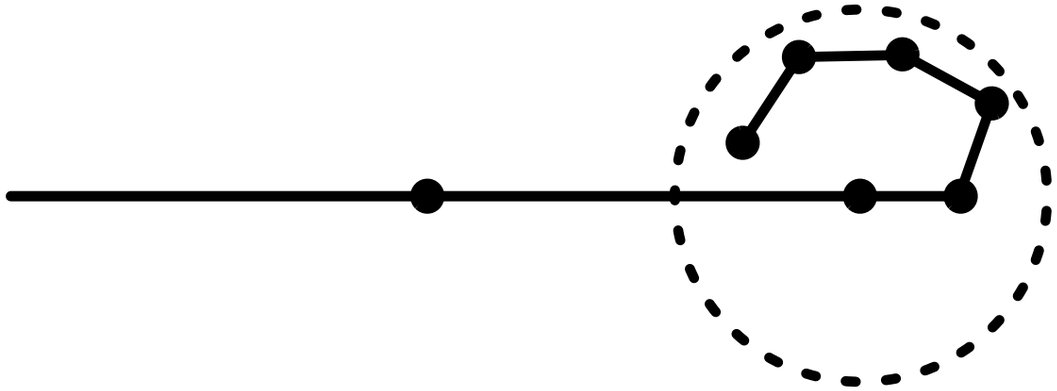
Stop detection



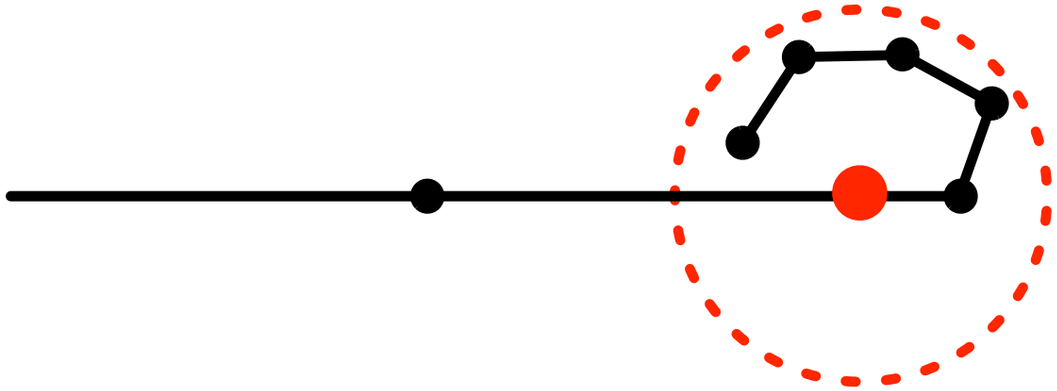
Stop detection



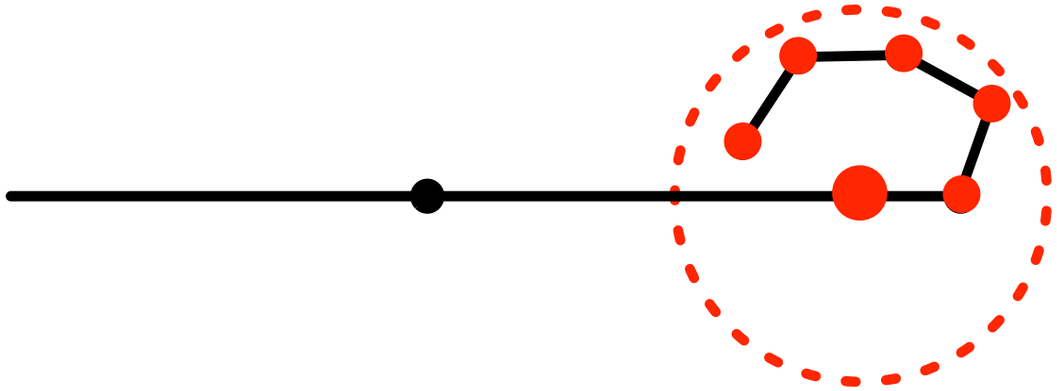
Stop detection



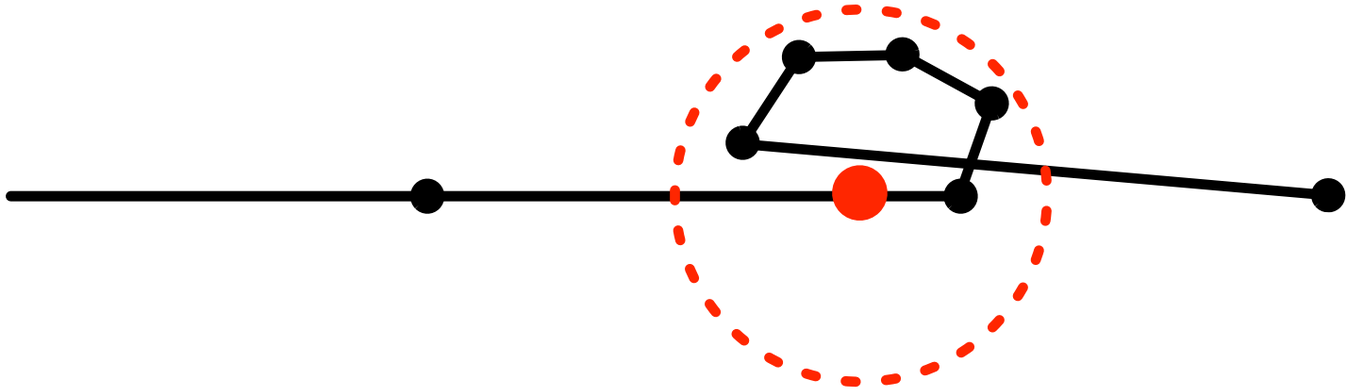
Stop detection



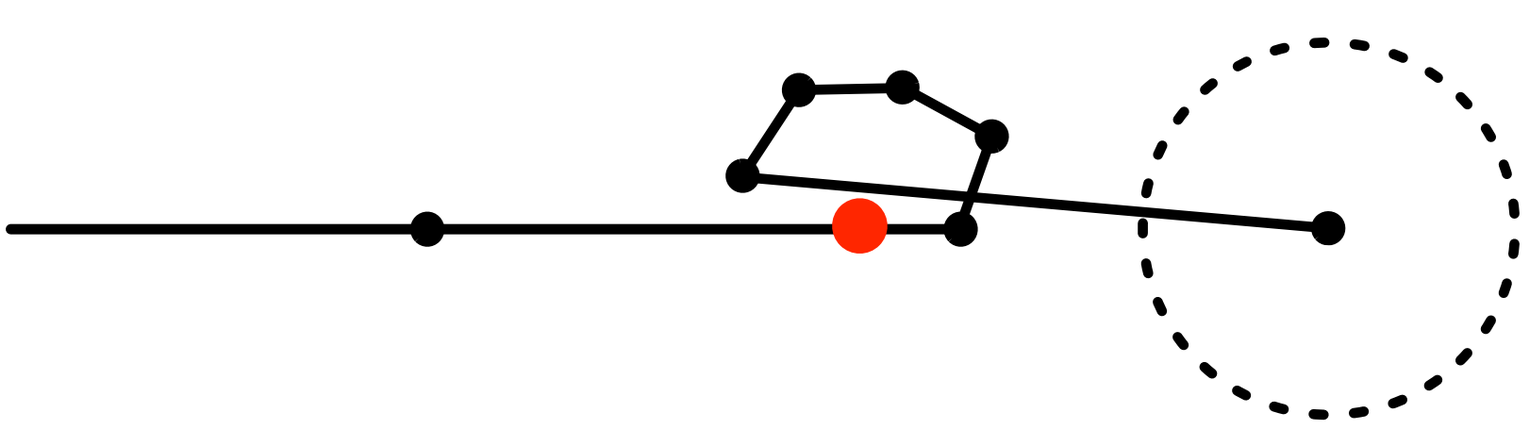
Stop detection



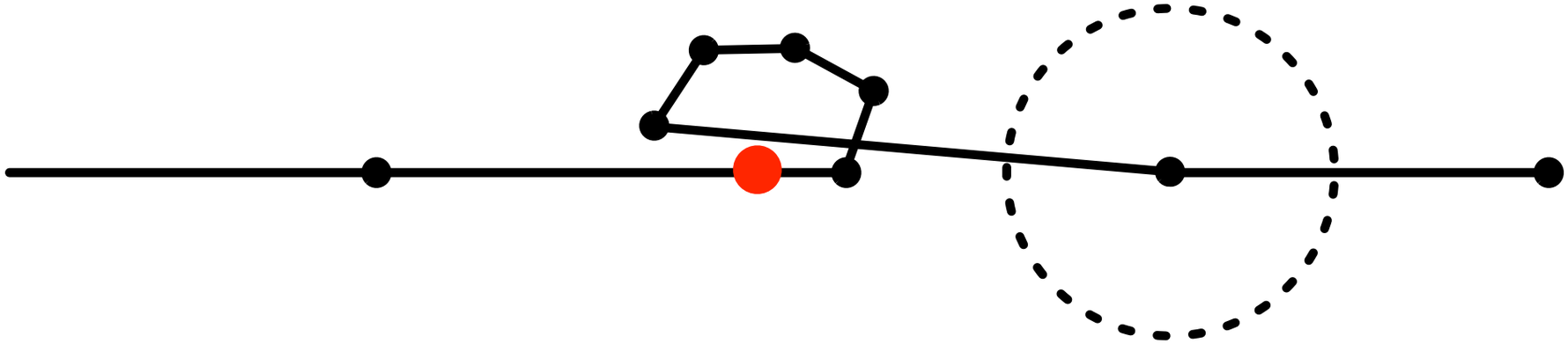
Stop detection



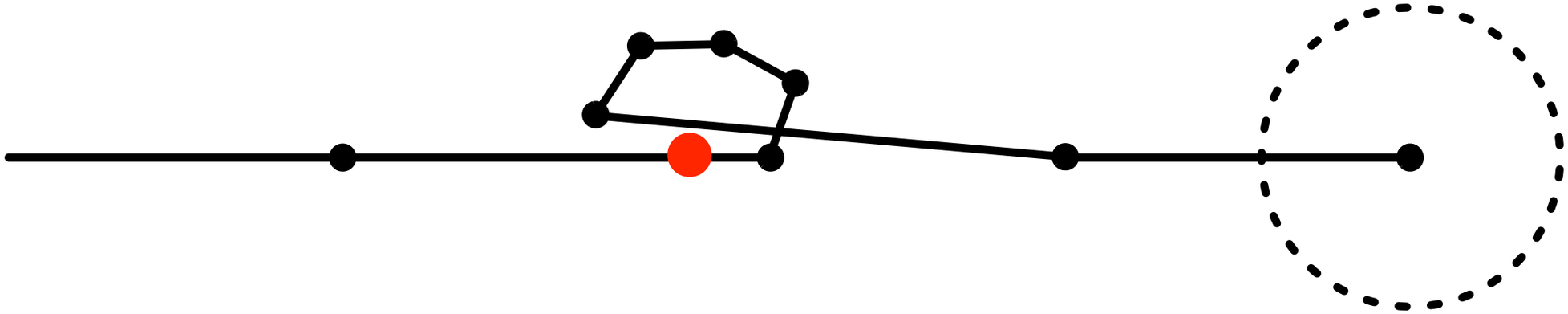
Stop detection



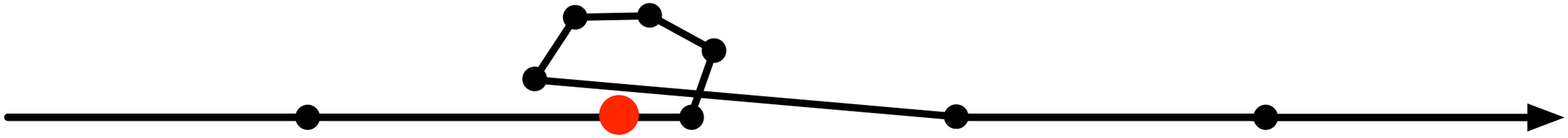
Stop detection

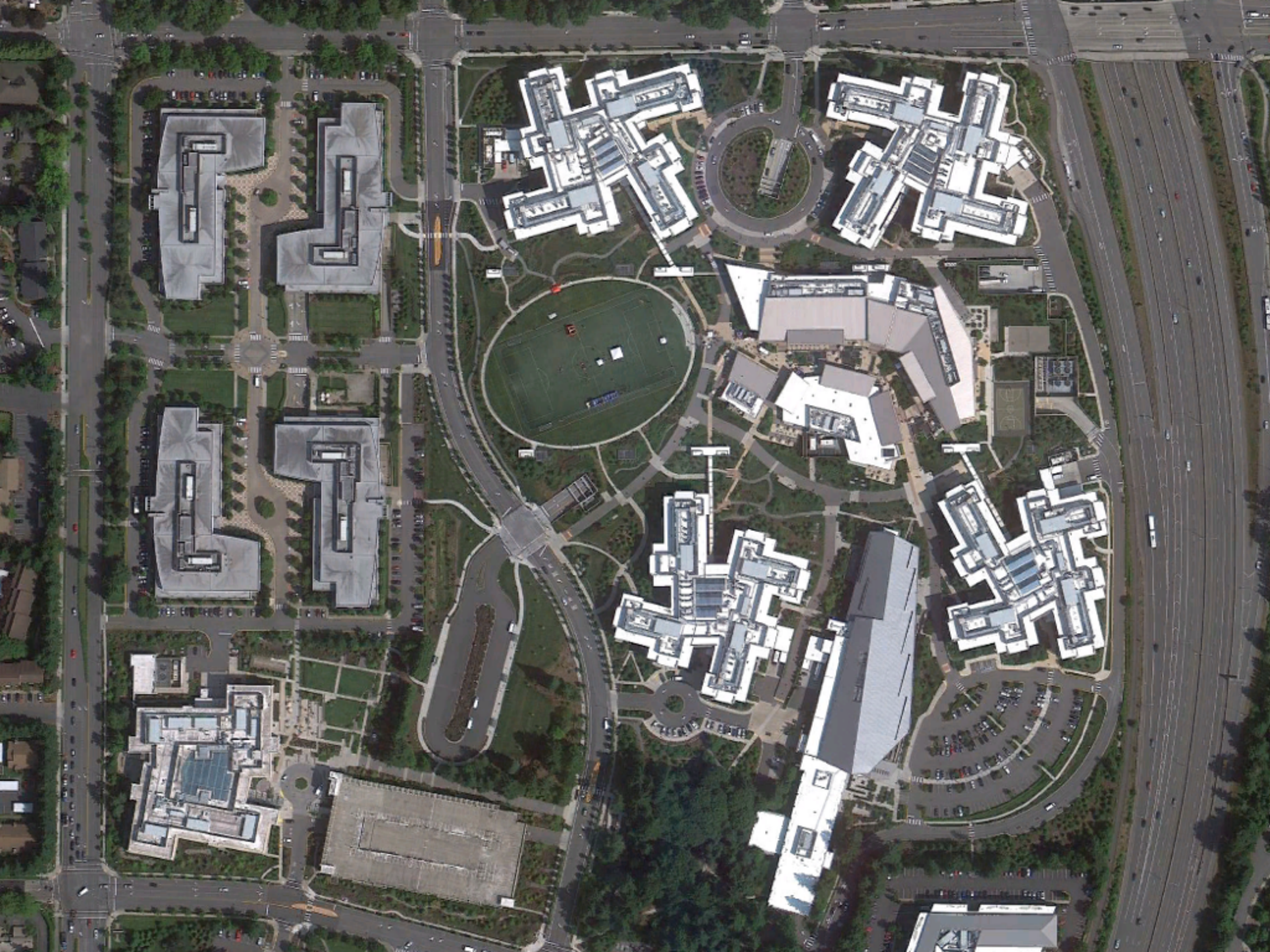


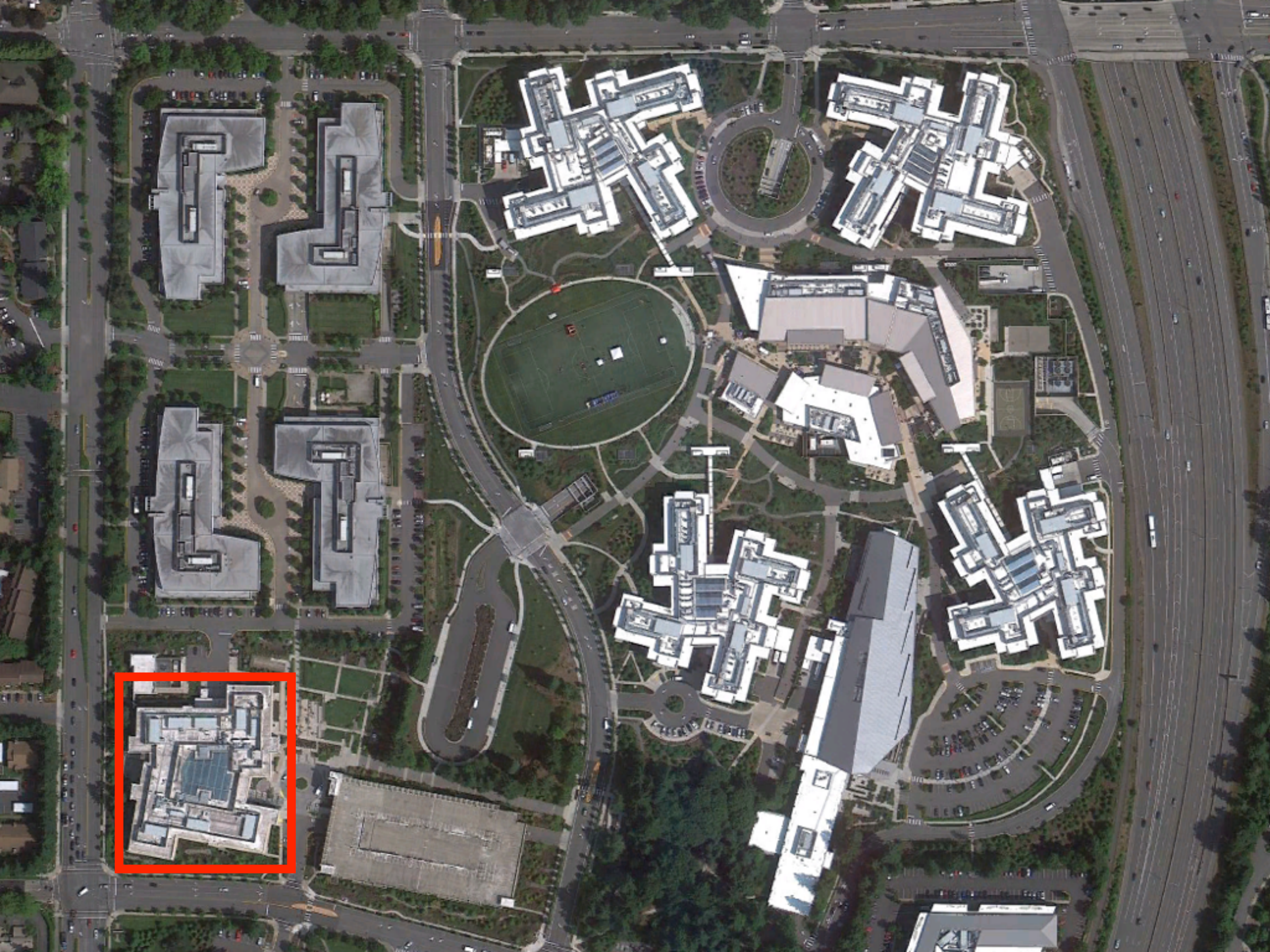
Stop detection

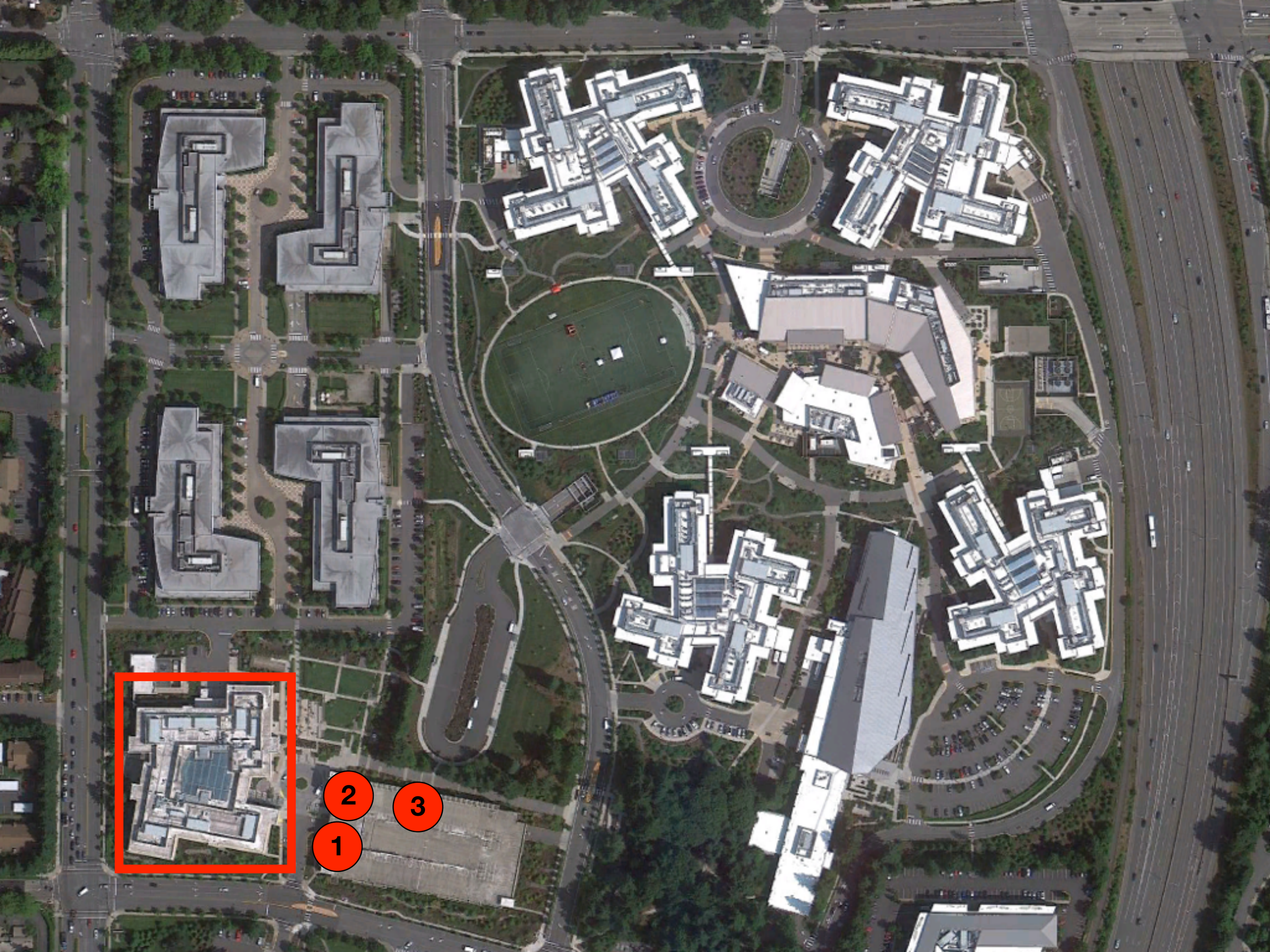


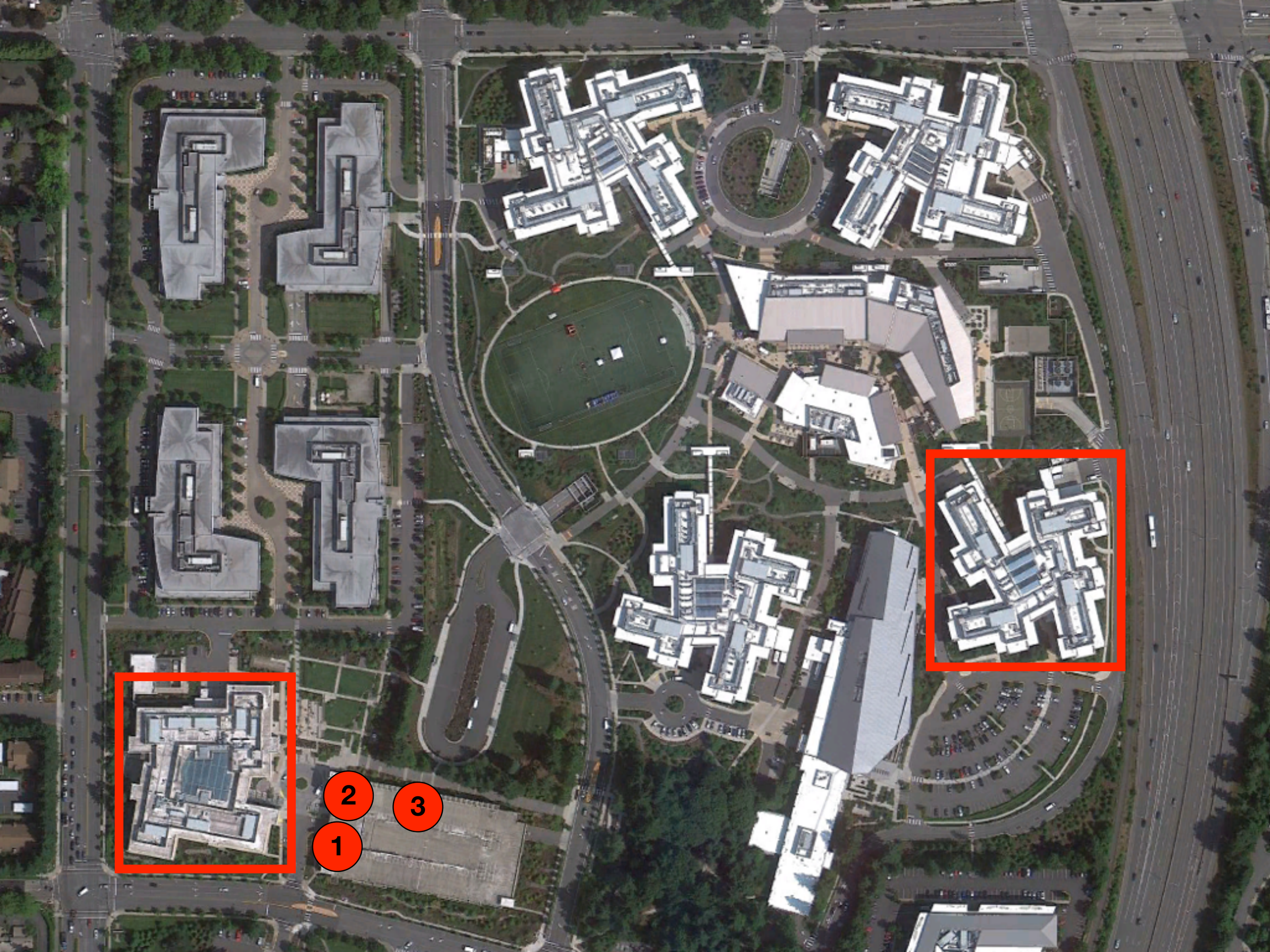
Stop detection

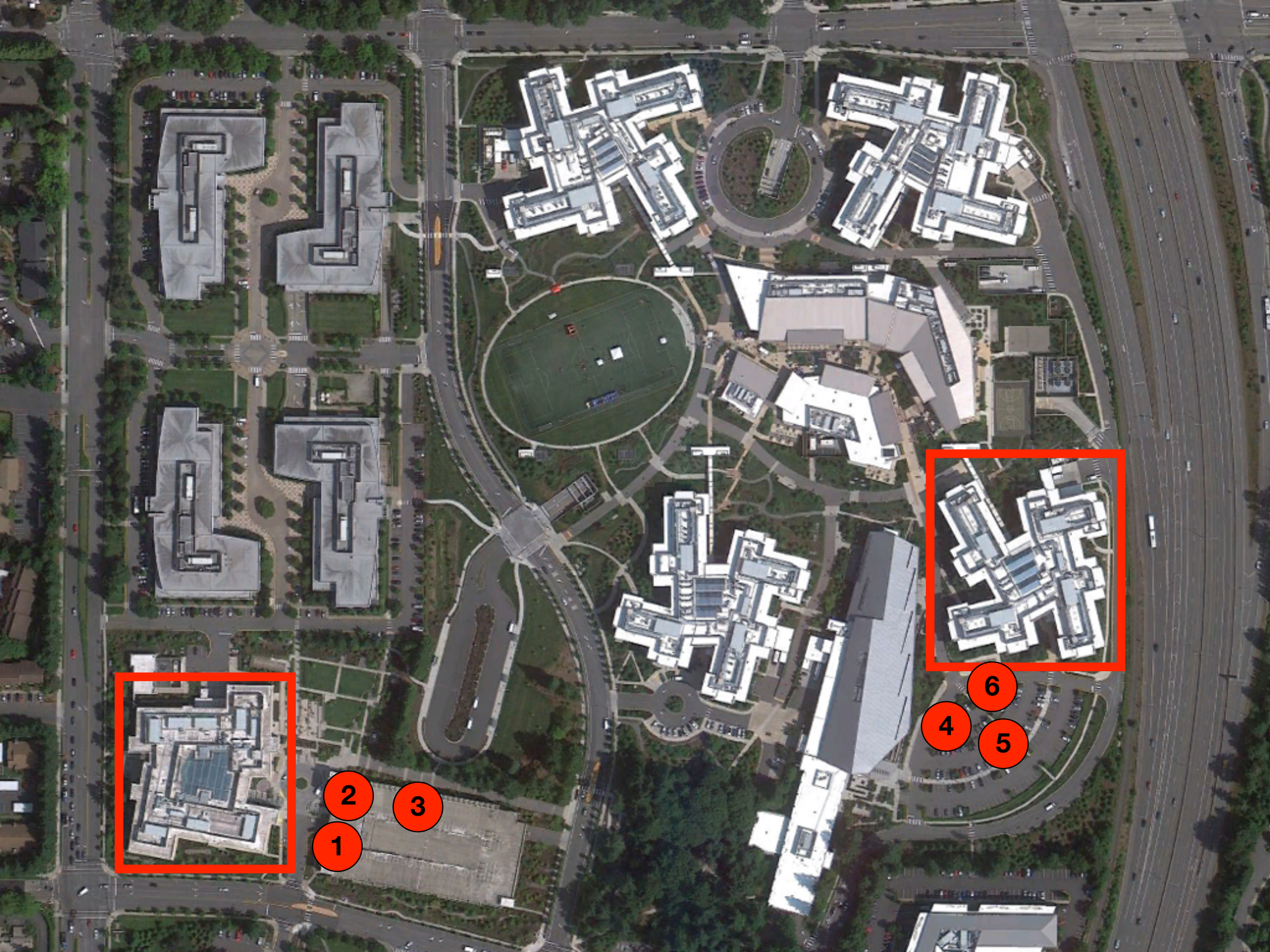


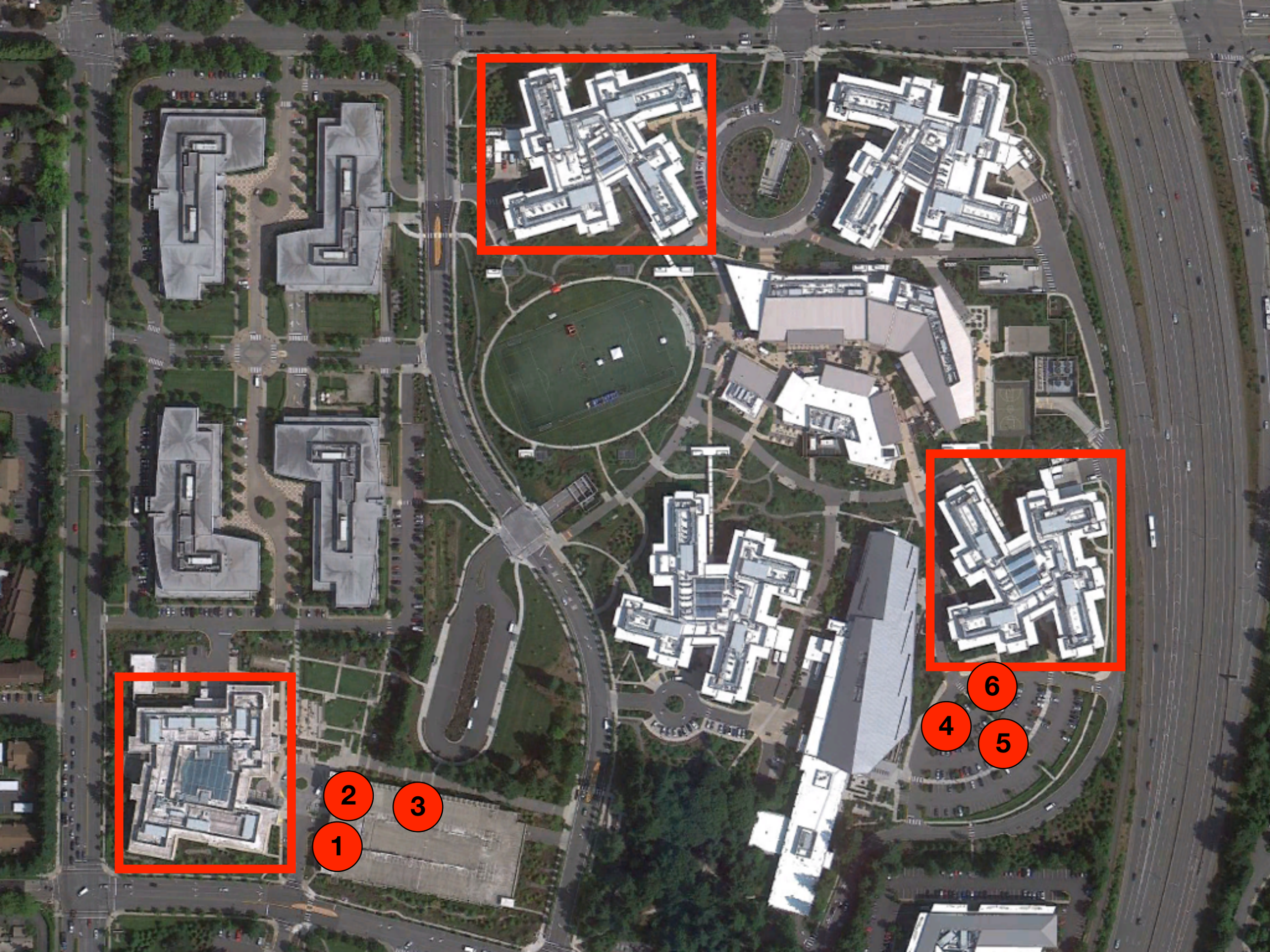


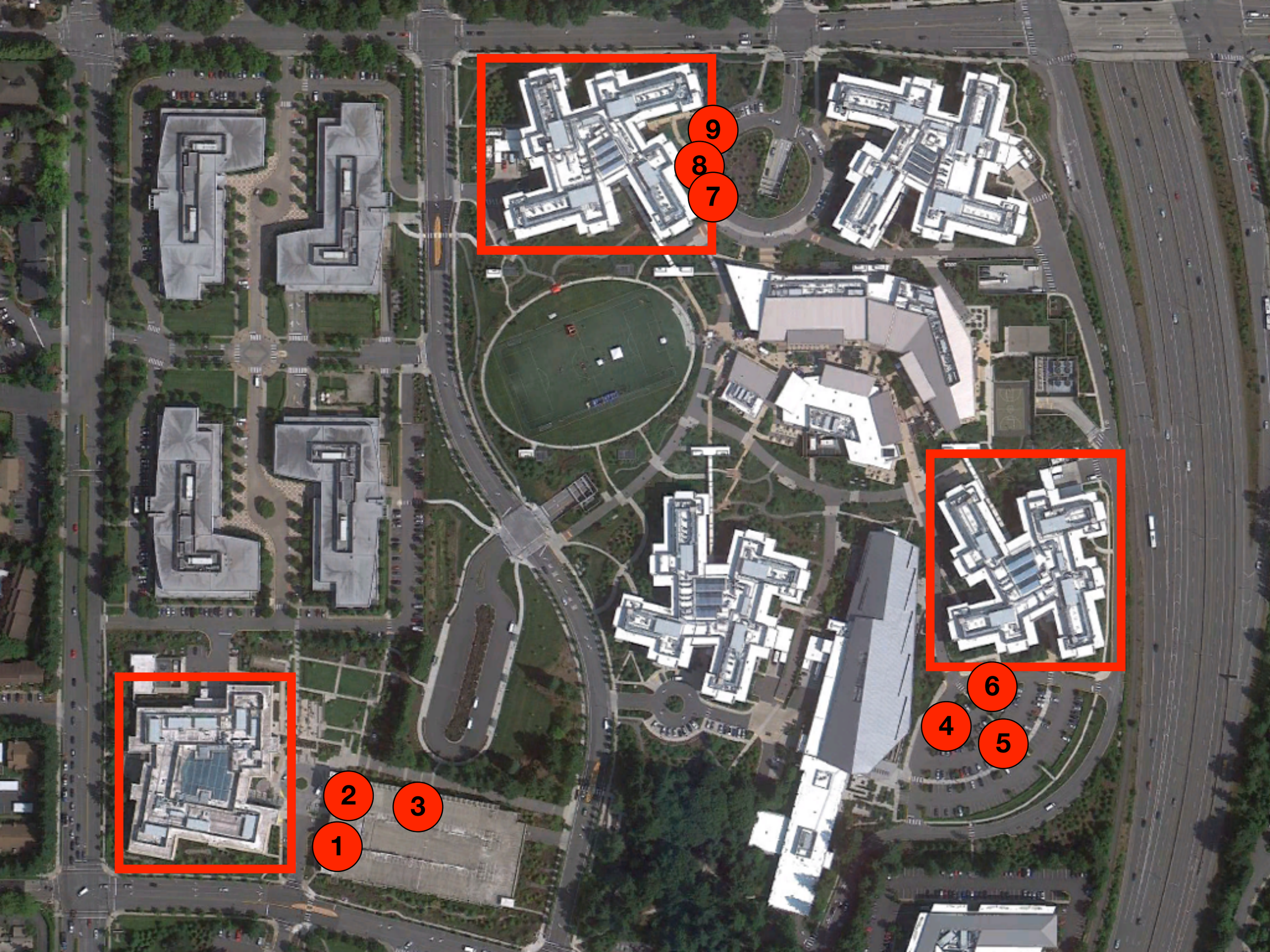


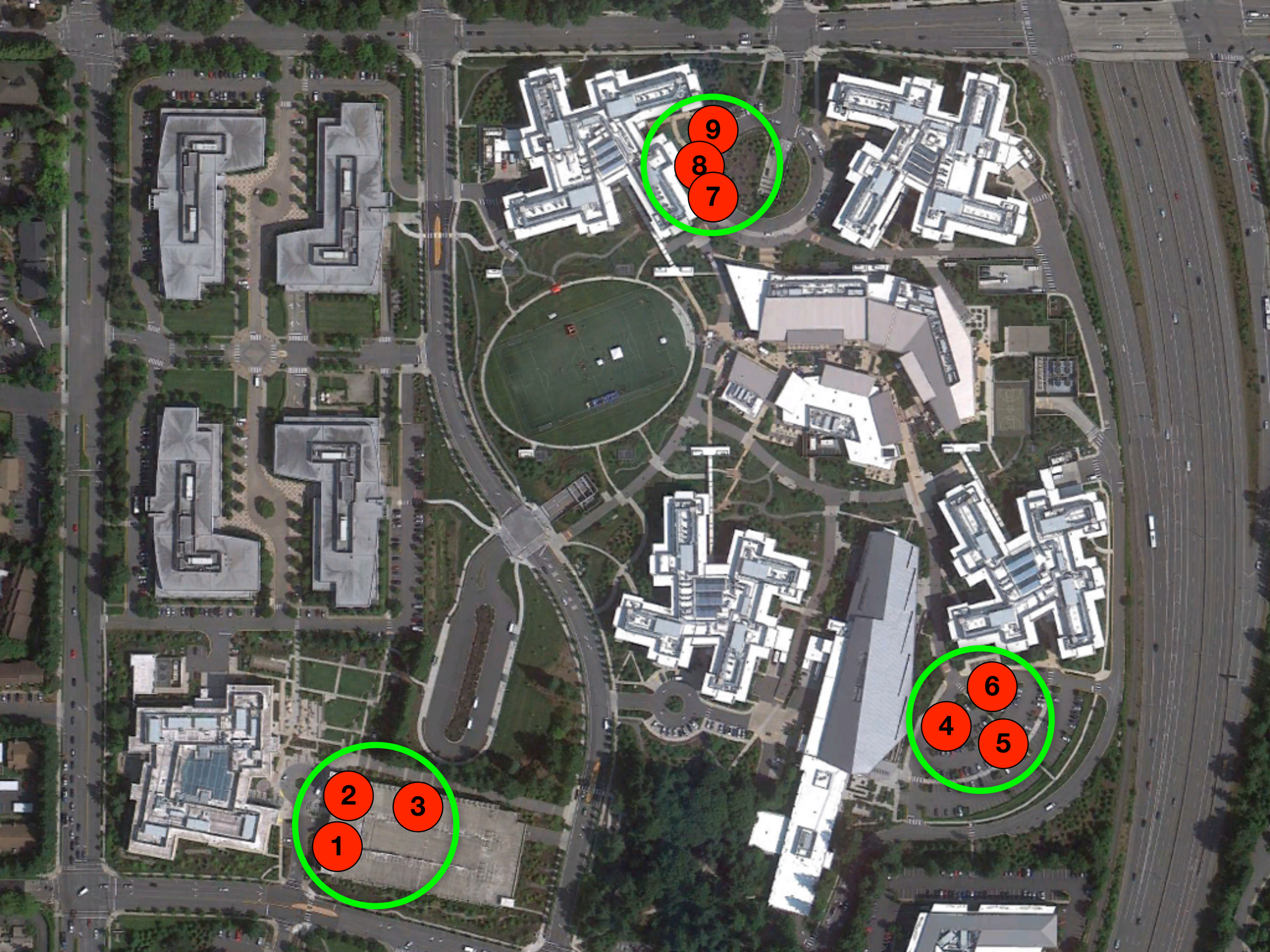

2
3
1

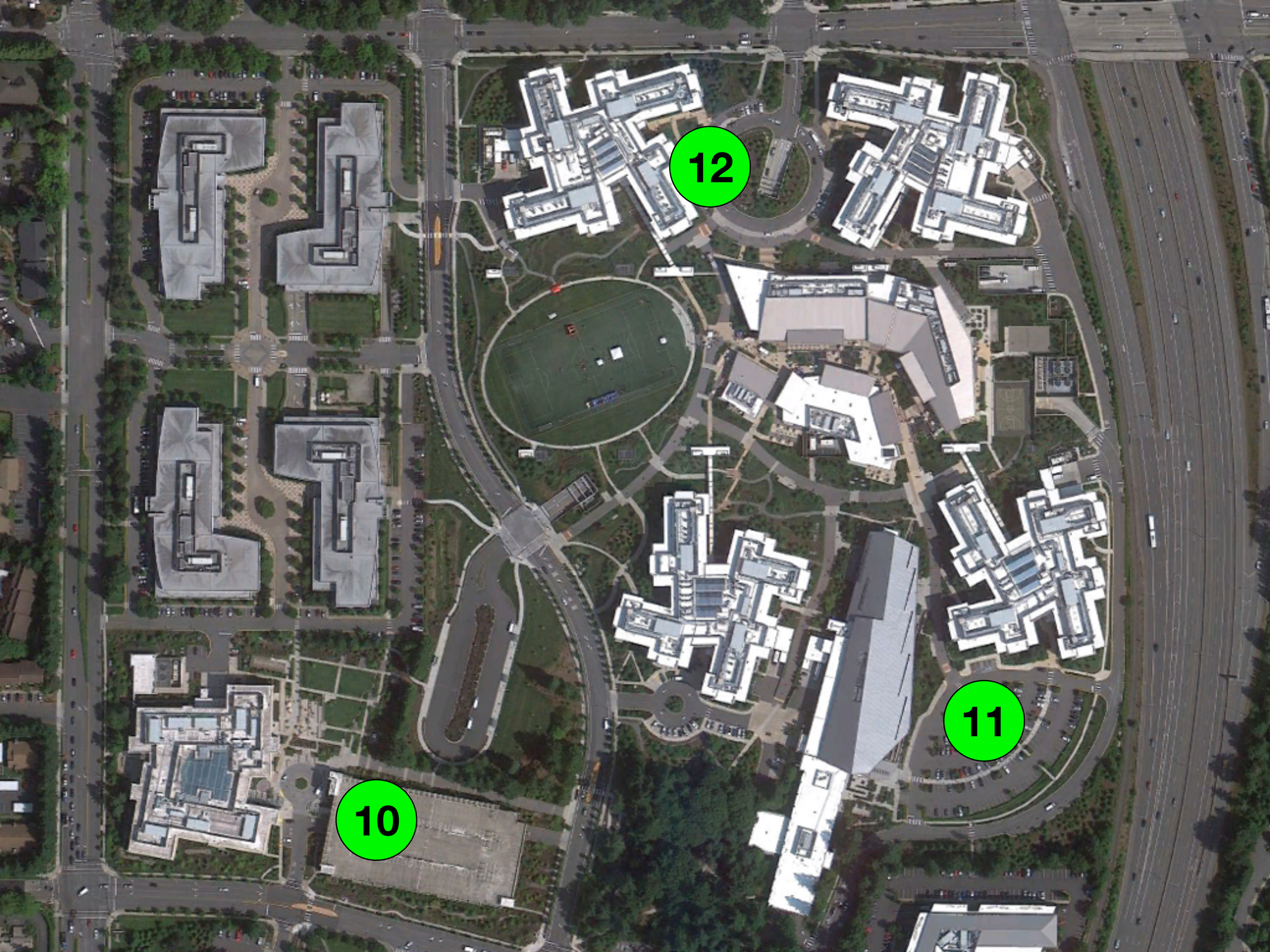

2
3
1


1
2
3
4
5
6


1
2
3
4
5
6


9
8
7
6
4
5
2
3
1


9
8
7
2
3
1
6
4
5

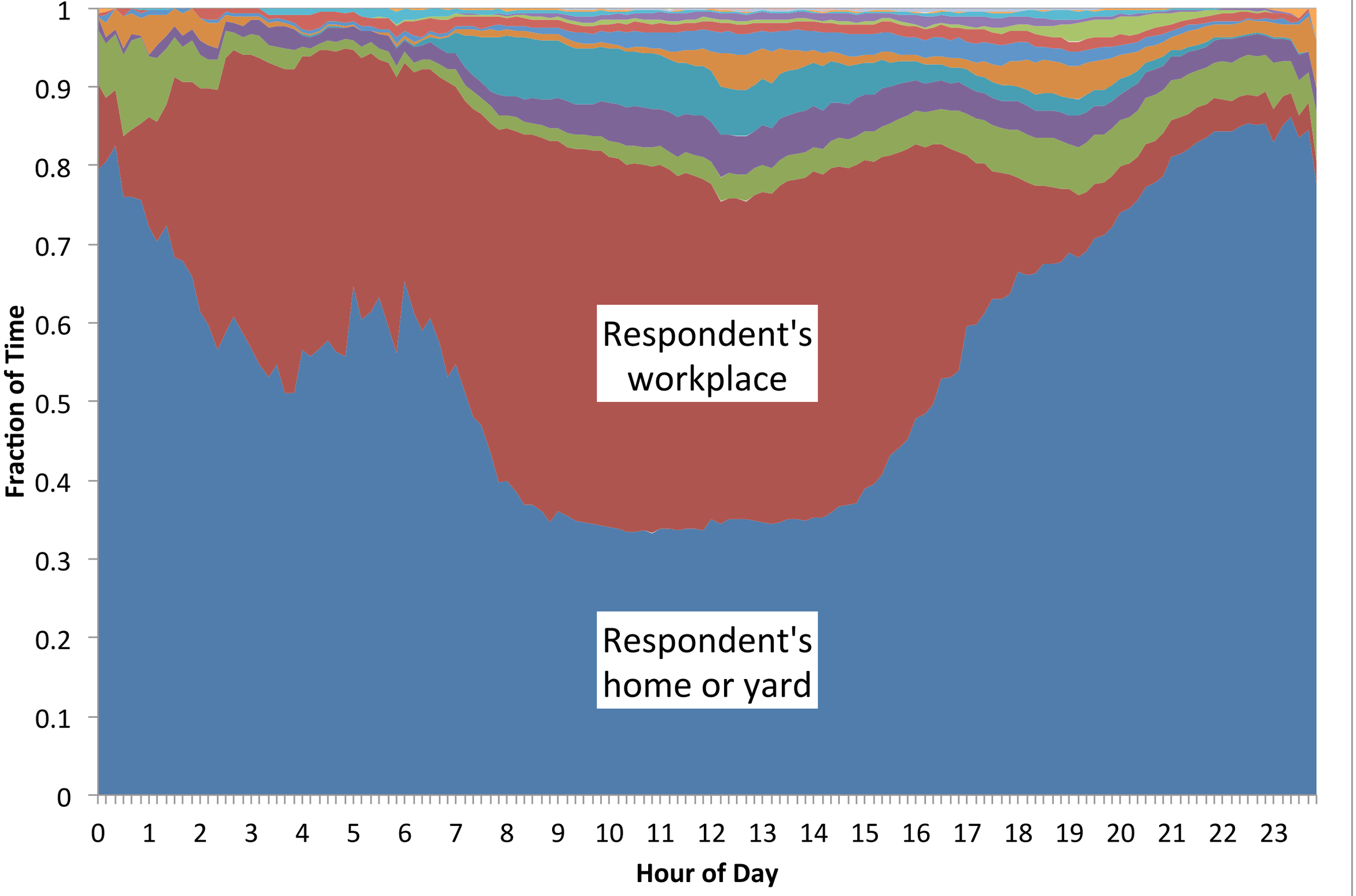

12
11
10

Home/work classification

- Used aggregated stop information
- Helpful for our subjects
  - Home and work occur frequently
- Leveraged ATUS data
  - How and where people spend their time

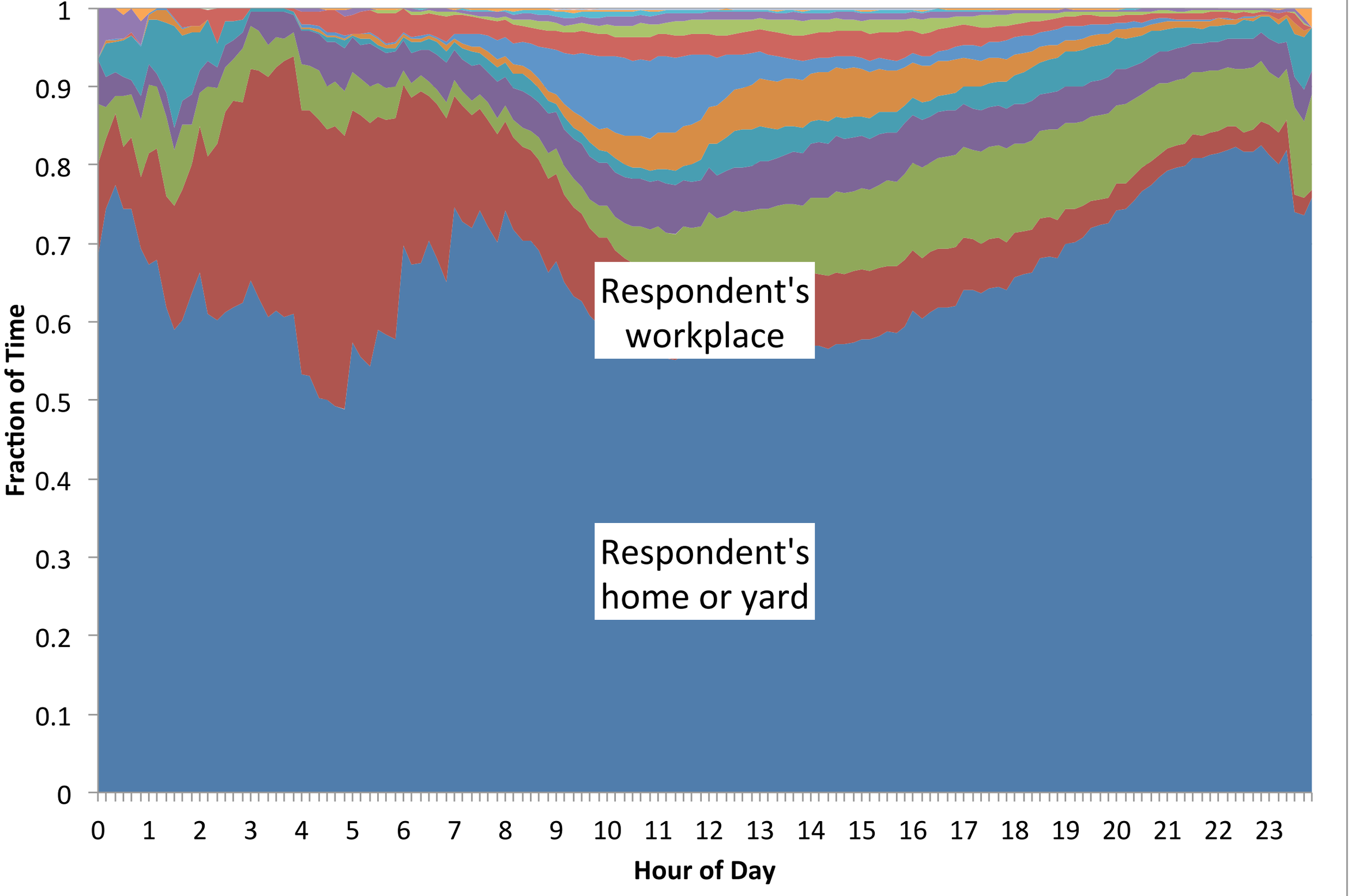
Fraction of Time at Place (weekday)

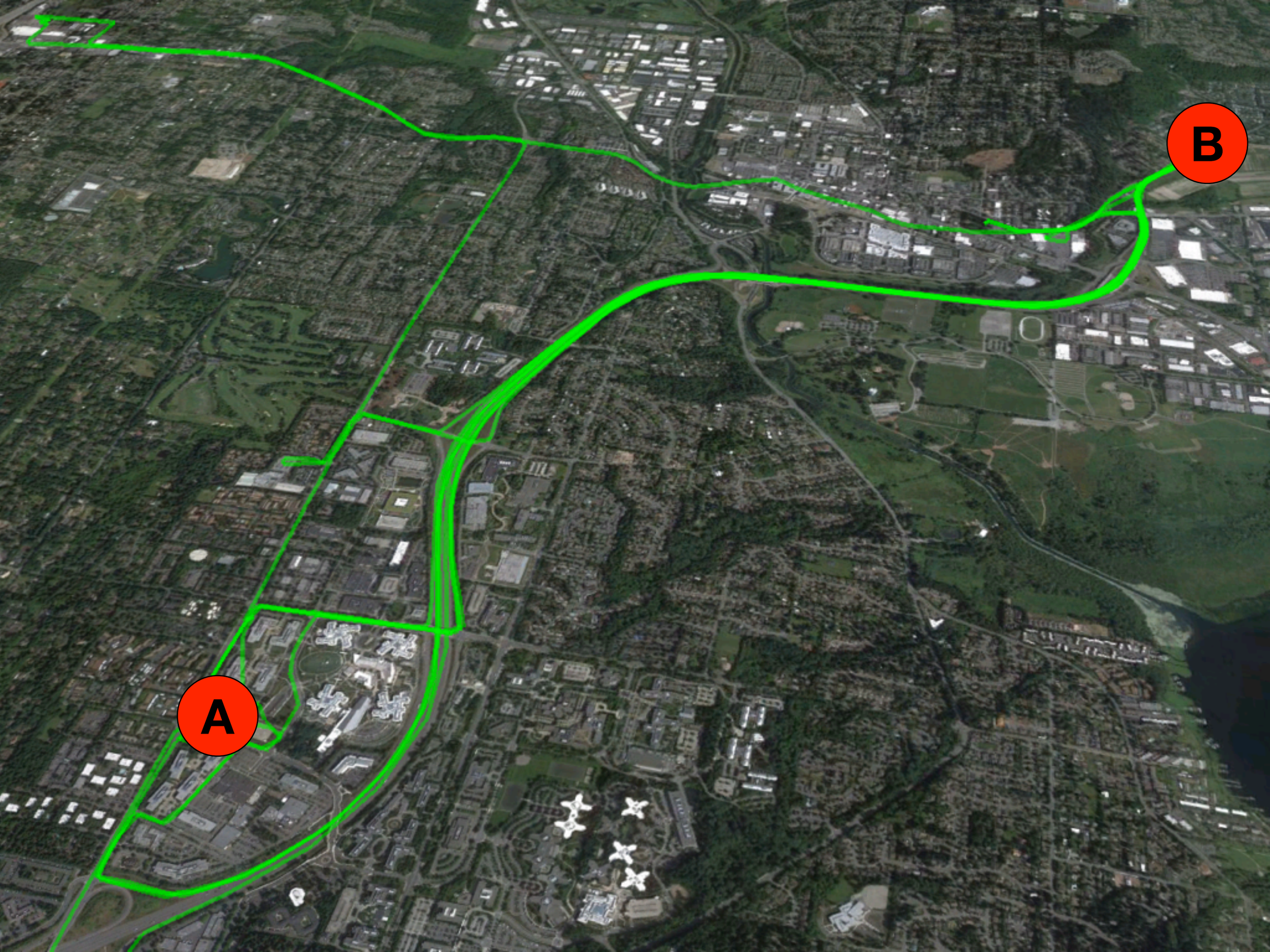
American Time Use Survey 2008

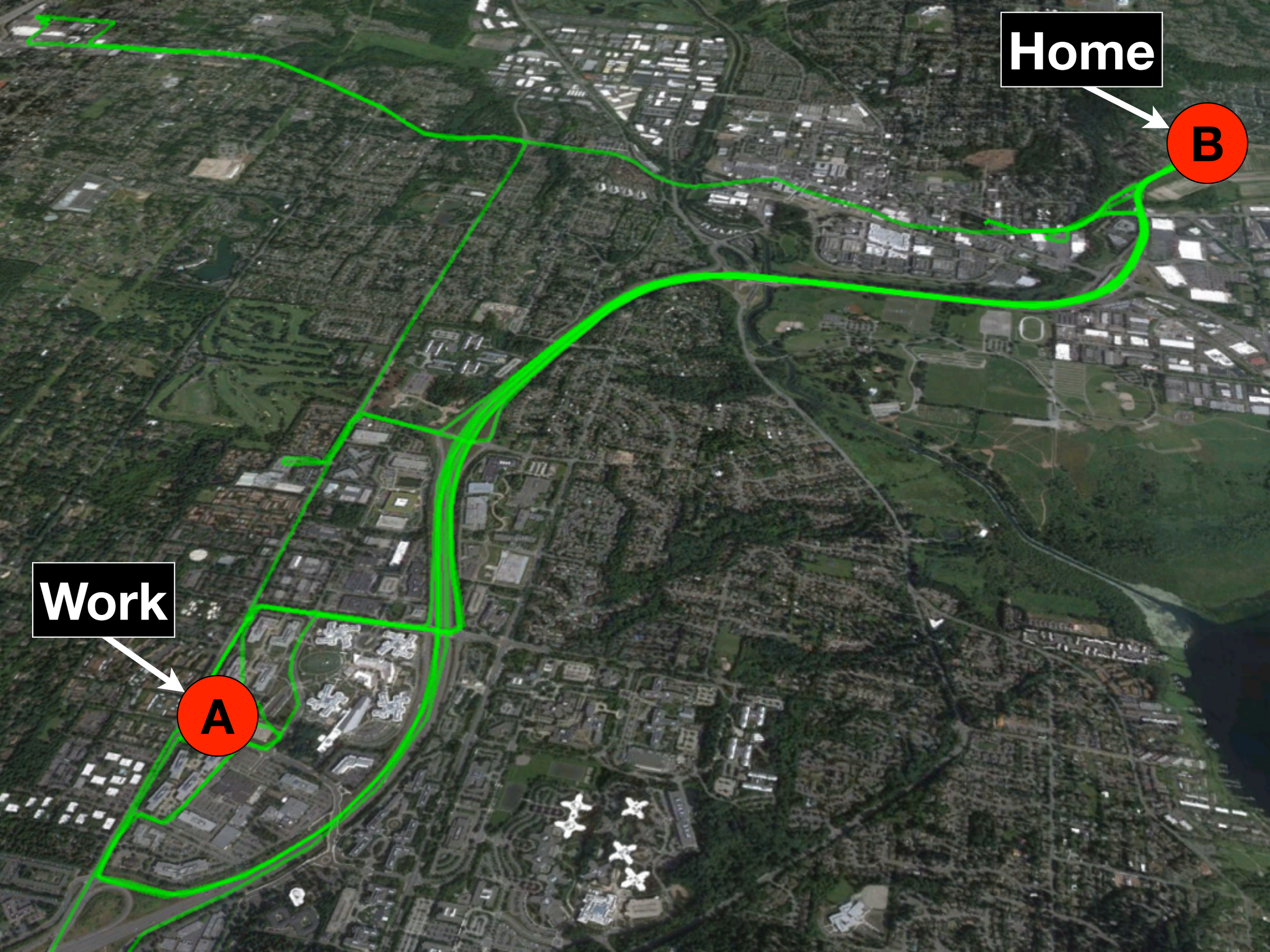


Fraction of Time at Place (weekend)

American Time Use Survey 2008




B
A


Home
B
Work
A

Symbolic location trace

```
47.643191 -122.139533
47.644055 -122.140175
47.645260 -122.140190
47.645953 -122.139900
47.645916 -122.138236
47.645920 -122.137200
47.645965 -122.135611
47.645900 -122.135735
47.645878 -122.135735
47.645870 -122.135735
47.645926 -122.135788
47.645950 -122.135803
47.645961 -122.135803
47.645970 -122.135788
47.645961 -122.135611
47.646061 -122.134353
```

Symbolic location trace

```
47.643191 -122.139533 A
47.644055 -122.140175 A
47.645260 -122.140190 A
47.645953 -122.139900 -
47.645916 -122.138236 -
47.645920 -122.137200 -
47.645965 -122.135611 -
47.645900 -122.135735 B
47.645878 -122.135735 B
47.645870 -122.135735 B
47.645926 -122.135788 -
47.645950 -122.135803 -
47.645961 -122.135803 -
47.645970 -122.135788 C
47.645961 -122.135611 C
47.646061 -122.134353 C
```

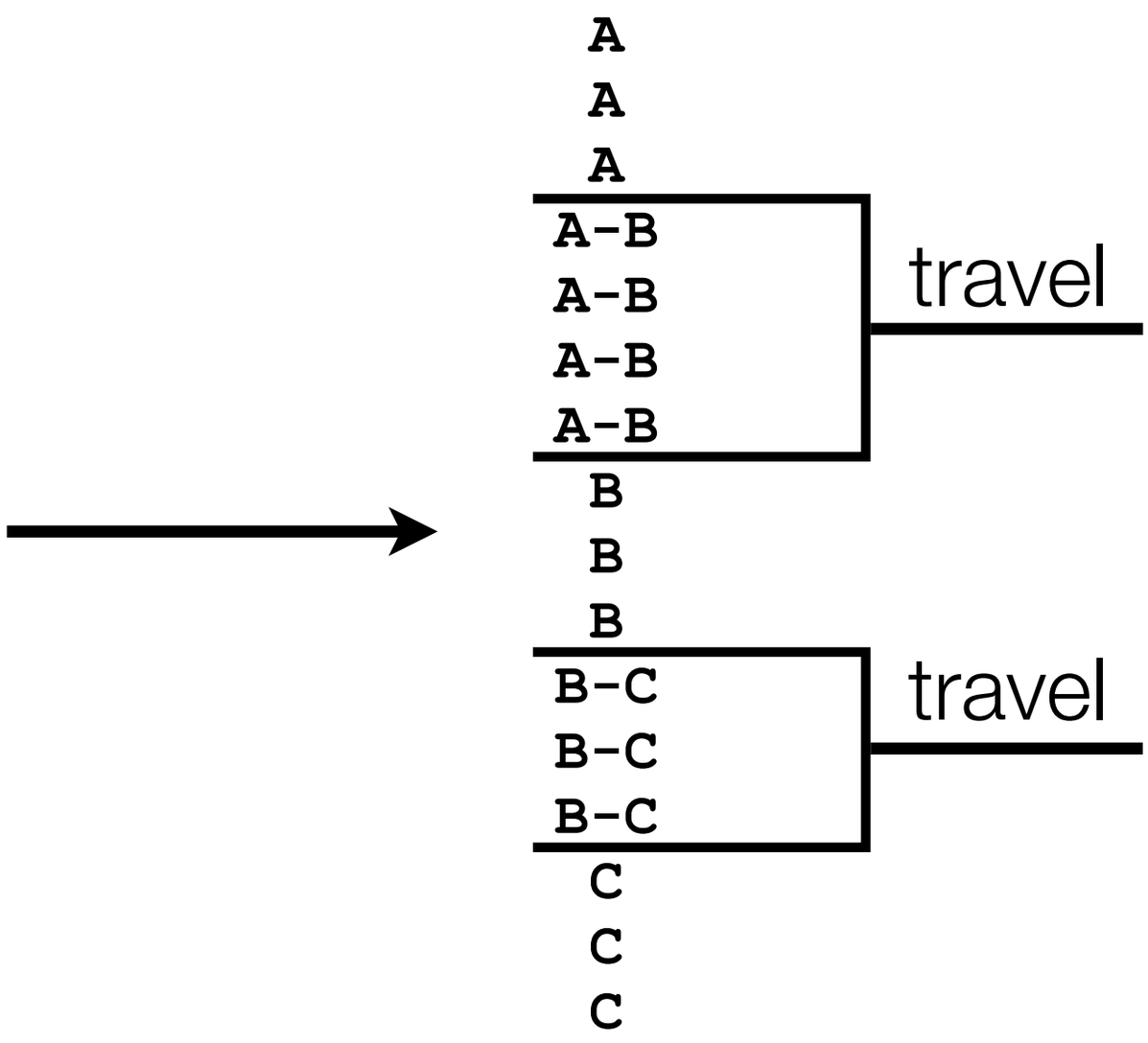
Symbolic location trace

```
47.643191 -122.139533 A                    A
47.644055 -122.140175 A                    A
47.645260 -122.140190 A                    A
47.645953 -122.139900 -                    -
47.645916 -122.138236 -                    -
47.645920 -122.137200 -                    -
47.645965 -122.135611 -                    -
47.645900 -122.135735 B                    B
47.645878 -122.135735 B                    B
47.645870 -122.135735 B                    B
47.645926 -122.135788 -                    -
47.645950 -122.135803 -                    -
47.645961 -122.135803 -                    -
47.645970 -122.135788 C                    C
47.645961 -122.135611 C                    C
47.646061 -122.134353 C                    C
```



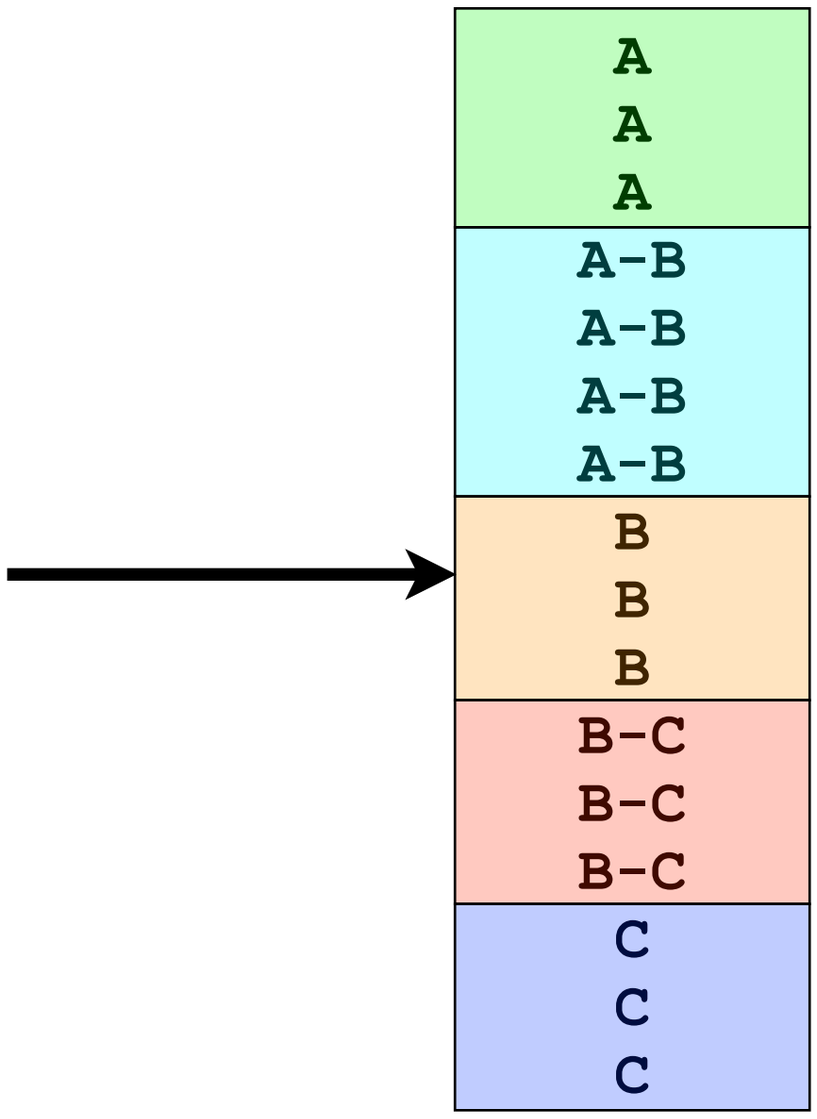
Symbolic location trace

```
47.643191 -122.139533 A
47.644055 -122.140175 A
47.645260 -122.140190 A
47.645953 -122.139900 -
47.645916 -122.138236 -
47.645920 -122.137200 -
47.645965 -122.135611 -
47.645900 -122.135735 B
47.645878 -122.135735 B
47.645870 -122.135735 B
47.645926 -122.135788 -
47.645950 -122.135803 -
47.645961 -122.135803 -
47.645970 -122.135788 C
47.645961 -122.135611 C
47.646061 -122.134353 C
```

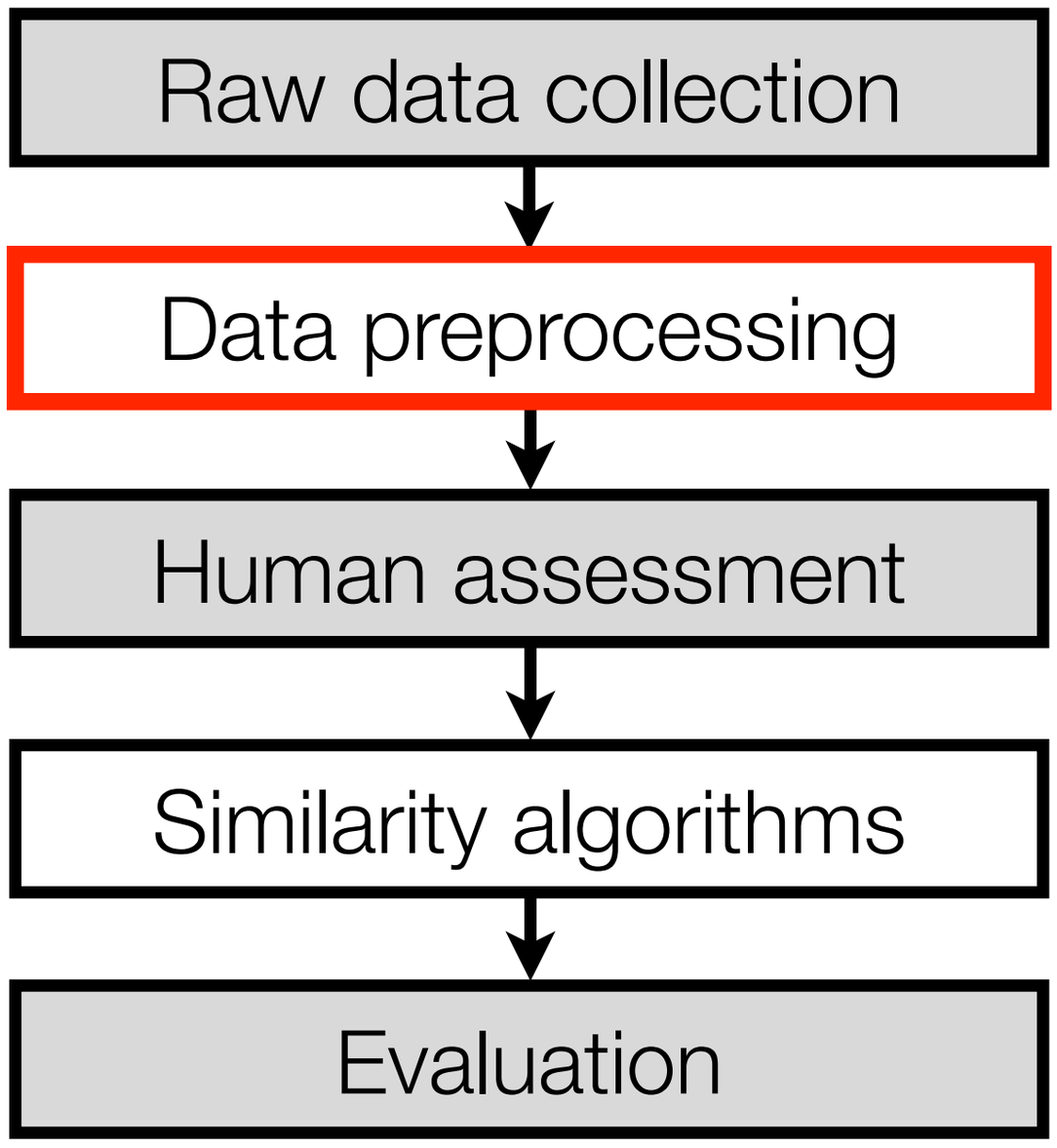


Symbolic location trace

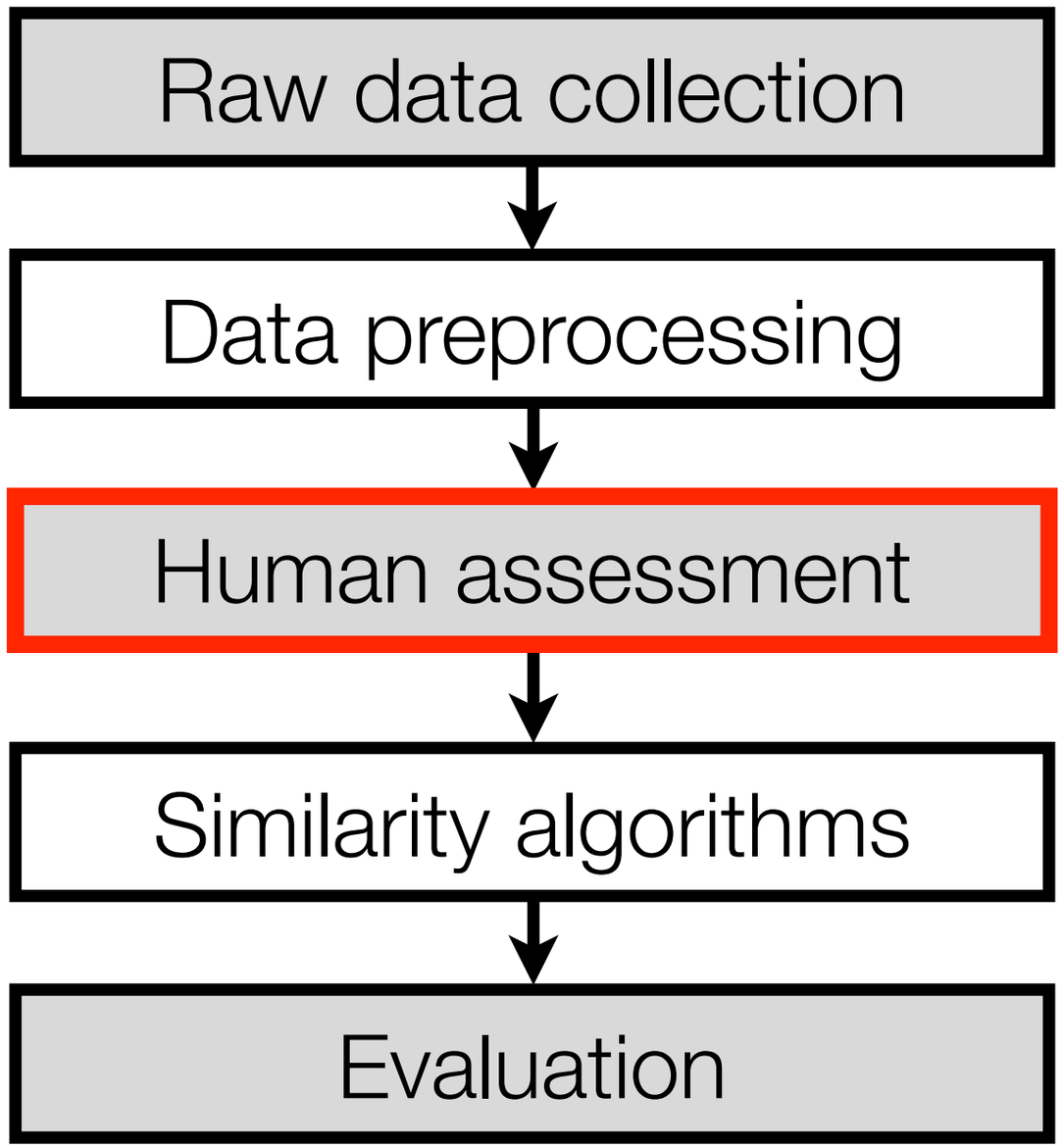
```
47.643191 -122.139533 A
47.644055 -122.140175 A
47.645260 -122.140190 A
47.645953 -122.139900 -
47.645916 -122.138236 -
47.645920 -122.137200 -
47.645965 -122.135611 -
47.645900 -122.135735 B
47.645878 -122.135735 B
47.645870 -122.135735 B
47.645926 -122.135788 -
47.645950 -122.135803 -
47.645961 -122.135803 -
47.645970 -122.135788 C
47.645961 -122.135611 C
47.646061 -122.134353 C
```



Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Human assessment

- Subjective evaluation
  - Captures “day similarity”
- Subjects’ evaluate their own data
  - Ground truth data collection

“Demo”

Days

Browse | Load Trace

Start Survey | - k= 2 + | Show Patterns | Find Patterns

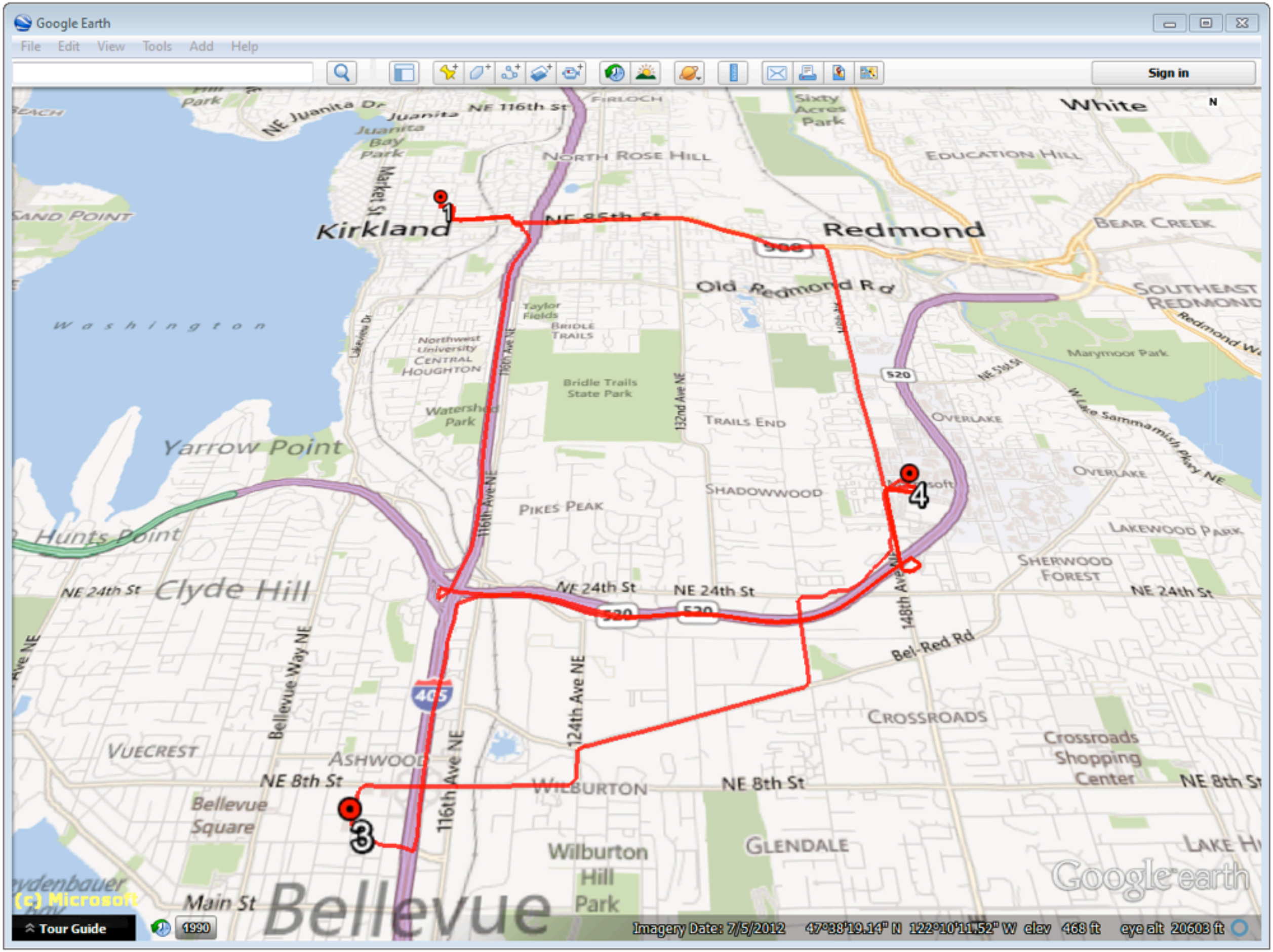
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11 Wed 7/11	12 Thu 7/12	13 Fri 7/13	14 Sat 7/14
15 - 21 Jul	15	16 Mon 7/16	17 Tue 7/17	18 Wed 7/18	19	20	21
22 - 28 Jul	22 Sun 7/22	23 Mon 7/23	24 Tue 7/24	25 Wed 7/25	26 Thu 7/26	27 Fri 7/27	28
29/7 - 4/8	29 Sun 7/29	30 Mon 7/30	31 Tue 7/31	01 Aug Wed 8/1	2 Thu 8/2	3 Fri 8/3	4
5 - 11 Aug	5	6 Mon 8/6	7 Tue 8/7	8 Wed 8/8	9 Thu 8/9	10 Fri 8/10	11
12 - 18 Aug	12	13 Mon 8/13	14	15	16	17	18
19 - 25 Aug	19	20 Mon 8/20	21 Tue 8/21	22 Wed 8/22	23 Thu 8/23	24 Fri 8/24	25
26/8 - 1/9	26	27 Mon 8/27	28 Tue 8/28	29 Wed 8/29	30	31	01 Sep

Days

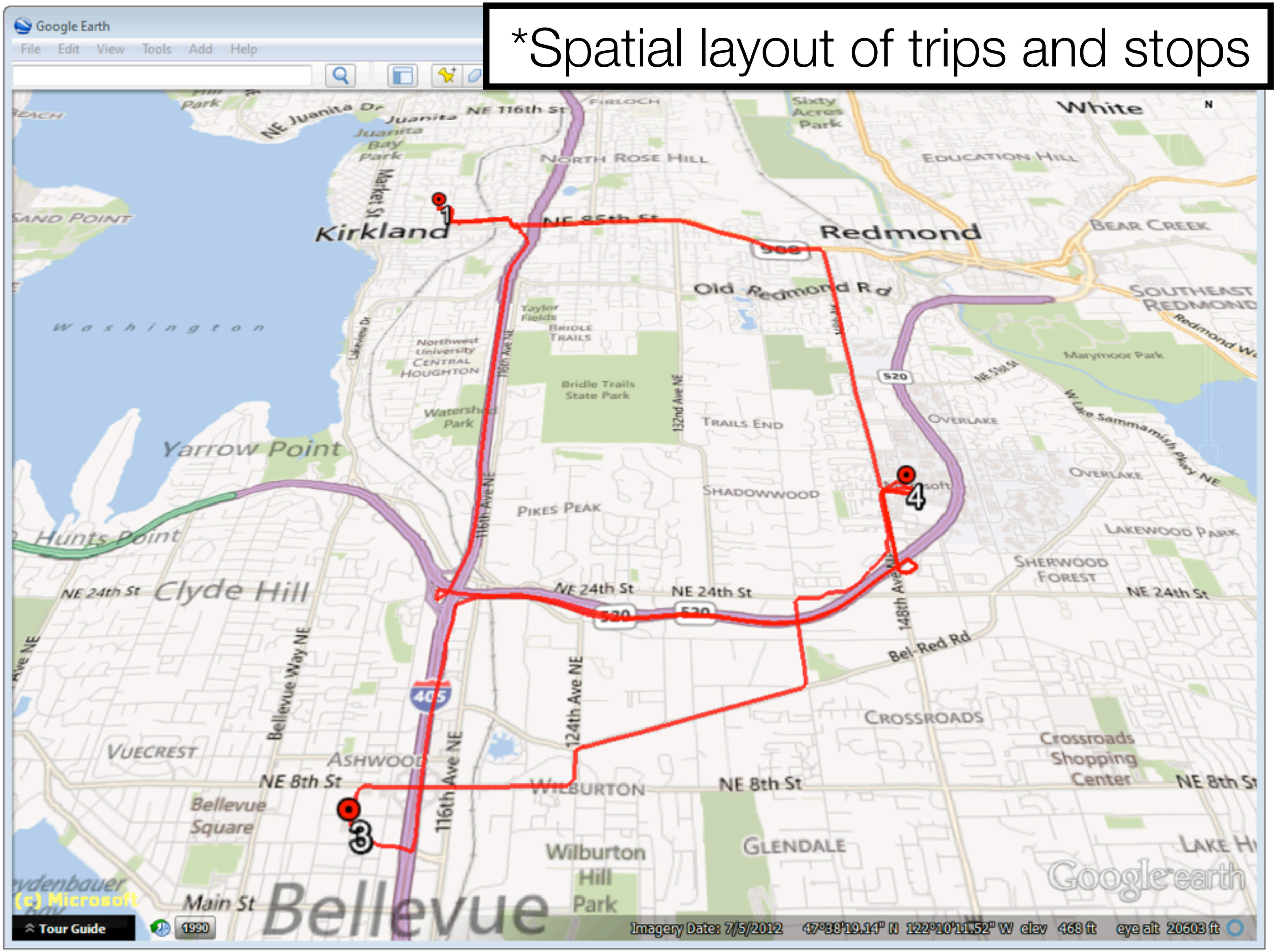
Browse | Load Trace

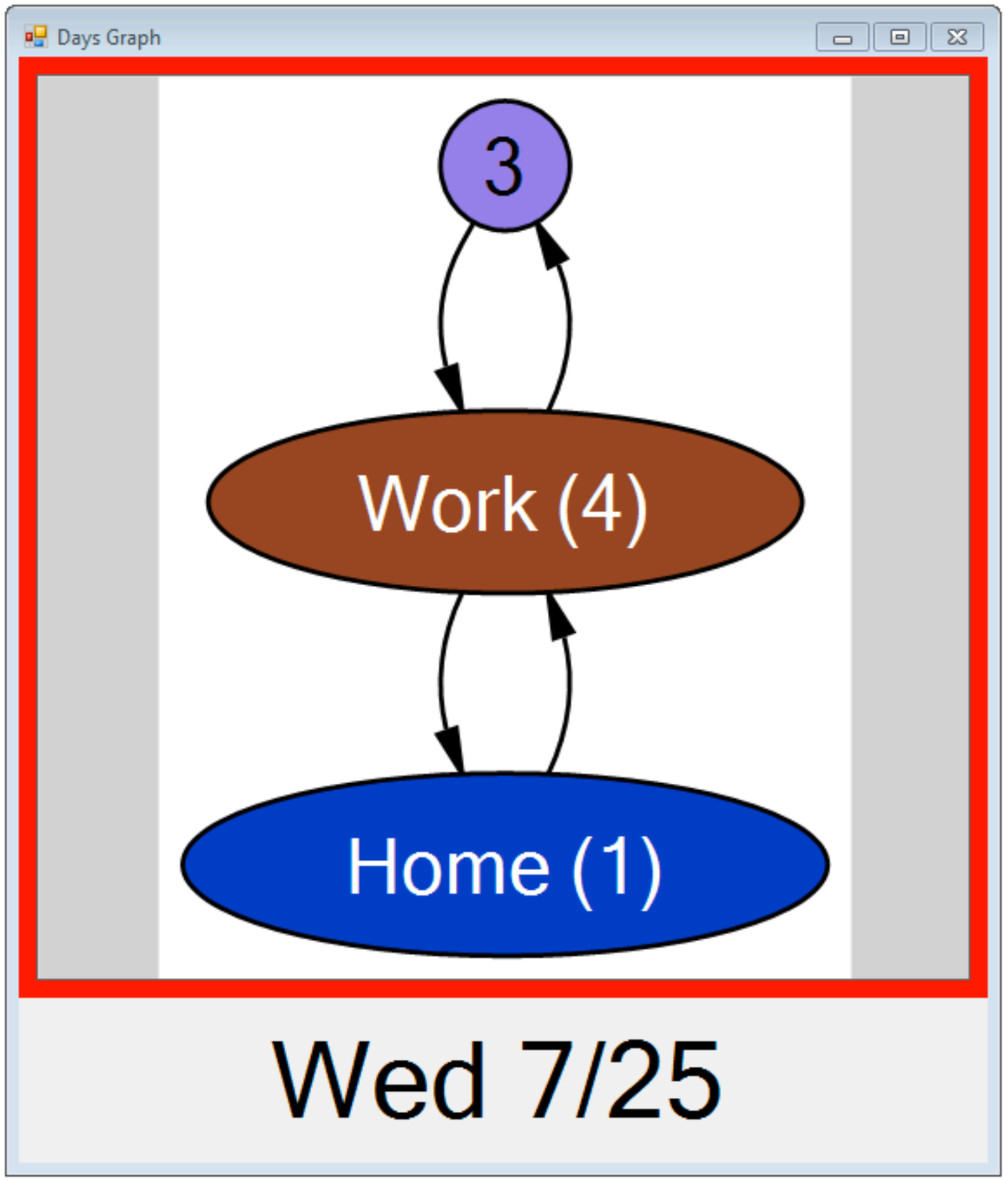
Start Survey | - k=2 + | Show Patterns | Find Patterns

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11 Wed 7/11	12 Thu 7/12	13 Fri 7/13	14 Sat 7/14
15 - 21 Jul	15	16 Mon 7/16	17 Tue 7/17	18 Wed 7/18	19	20	21
22 - 28 Jul	22 Sun 7/22	23 Mon 7/23	24 Tue 7/24	25 Wed 7/25	26 Thu 7/26	27 Fri 7/27	28
29/7 - 4/8	29 Sun 7/29	30 Mon 7/30	31 Tue 7/31	01 Aug Wed 8/1	2 Thu 8/2	3 Fri 8/3	4
5 - 11 Aug	5	6 Mon 8/6	7 Tue 8/7	8 Wed 8/8	9 Thu 8/9	10 Fri 8/10	11
12 - 18 Aug	12	13 Mon 8/13	14	15	16	17	18
19 - 25 Aug	19	20 Mon 8/20	21 Tue 8/21	22 Wed 8/22	23 Thu 8/23	24 Fri 8/24	25
26/8 - 1/9	26	27 Mon 8/27	28 Tue 8/28	29 Wed 8/29	30	31	01 Sep

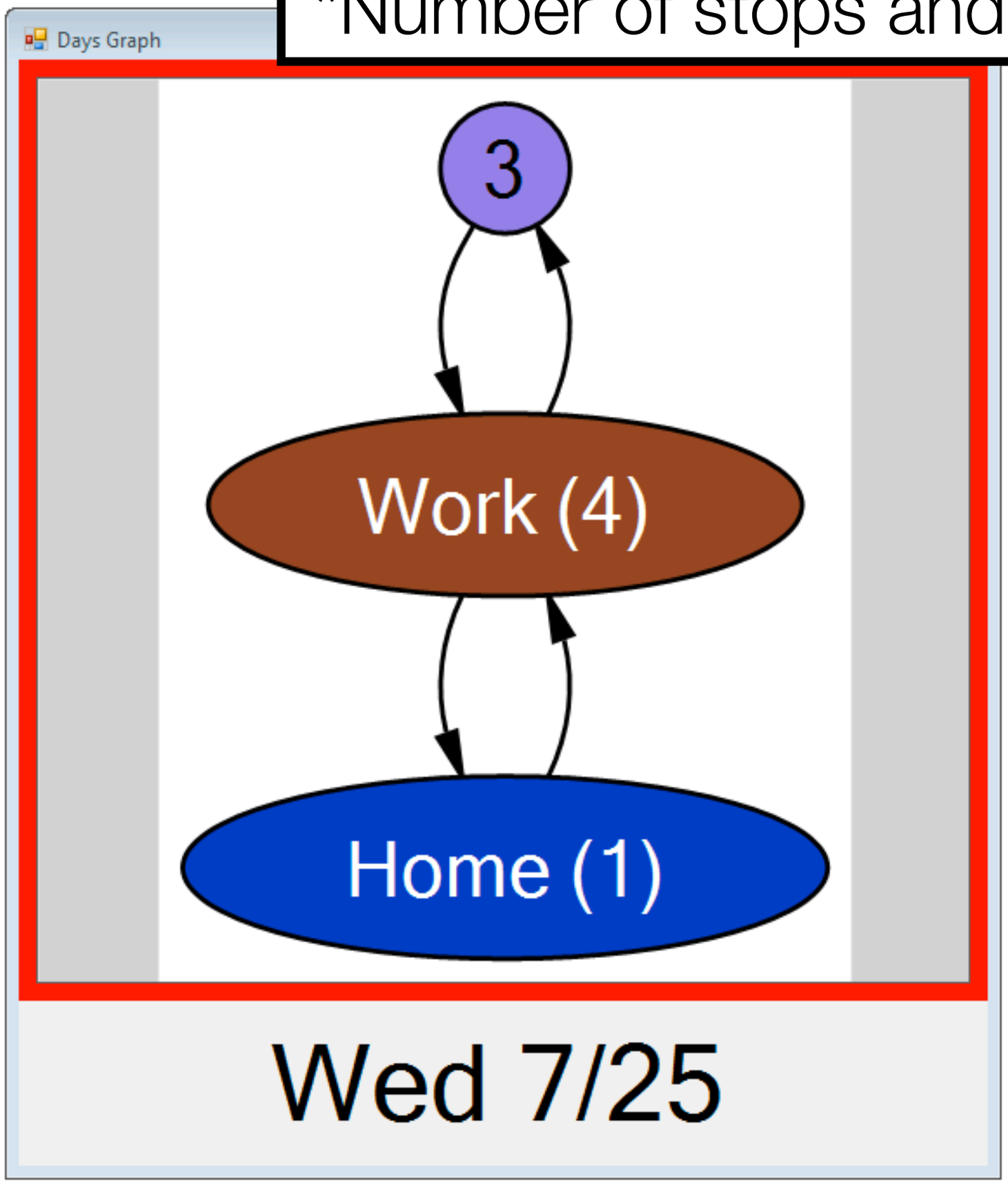

Google Earth
File Edit View Tools Add Help
Sign in
Kirkland
Redmond
Bellevue
Yarrow Point
Hunts Point
Clyde Hill
Medina
Sand Point
Washington
White
Education Hill
North Rose Hill
Firloch
Sixty Acres Park
Juanita Bay Park
Juanita
Market St
NE 116th St
NE 85th St
Old Redmond Rd
Bear Creek
Southeast Redmond
Redmond Way
Marymoor Park
W Lake Sammamish Pkwy NE
Overlake
Lakewood Park
Sherwood Forest
NE 24th St
Bel-Red Rd
Crossroads
Crossroads Shopping Center
NE 8th St
Vuecrest
Ashwood
Wilburton
Glendale
Lake Hills
Bellevue Square
Wilburton Hill Park
Main St
Bellevue Way NE
116th Ave NE
124th Ave NE
132nd Ave NE
148th Ave NE
Taylor Fields
Bridle Trails
Bridle Trails State Park
Trails End
Shadowwood
Pikes Peak
Watershed Park
Northwest University
Central Houghton
Lakeview Dr
520
405
908
1
3
4
(c) Microsoft
Google earth
Tour Guide
1990
Imagery Date: 7/5/2012 47°38'19.14" N 122°10'11.52" W elev 468 ft eye alt 20603 ft

*Spatial layout of trips and stops

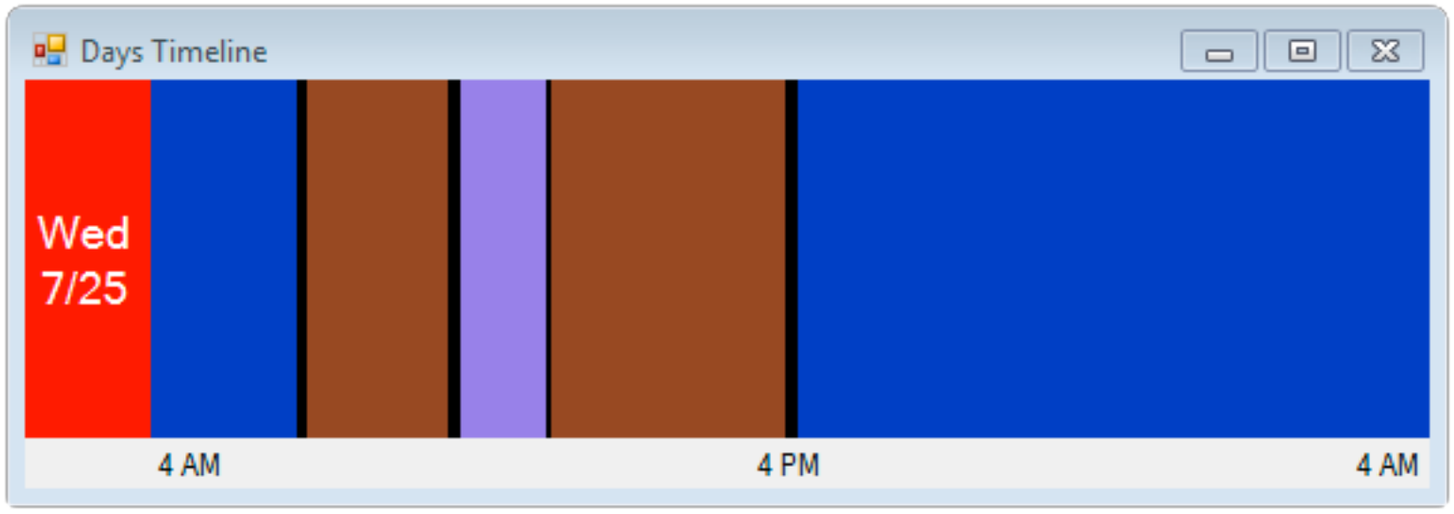



Days Graph
3
Work (4)
Home (1)
Wed 7/25

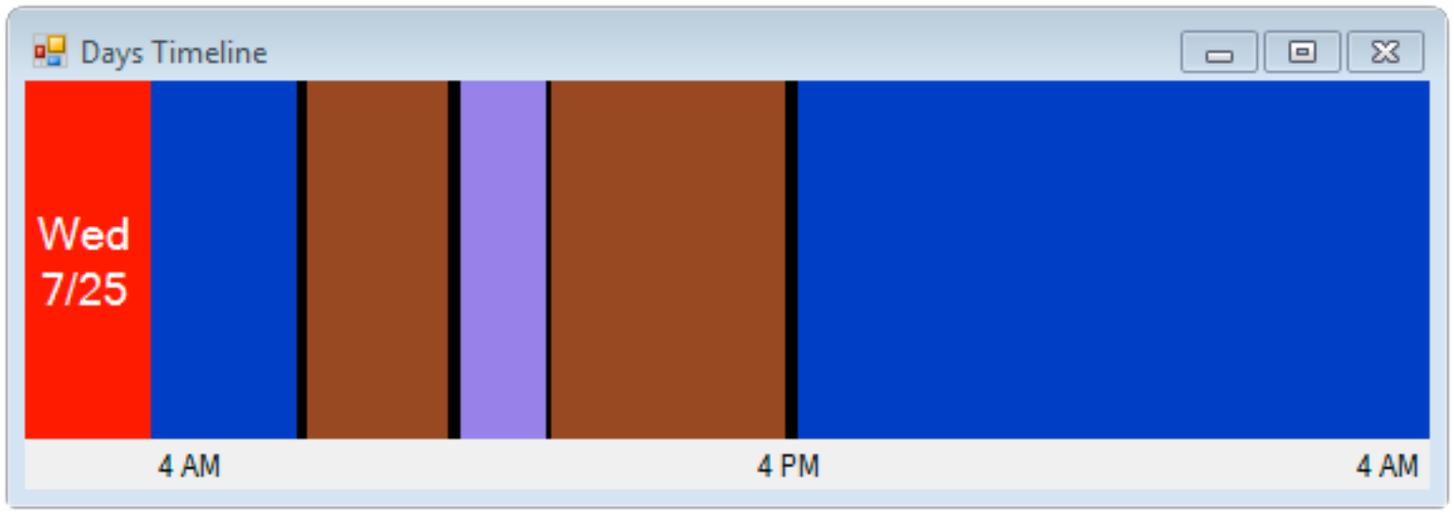
*Number of stops and transitions



Wed 7/25


Days Timeline
Wed
7/25
4 AM
4 PM
4 AM

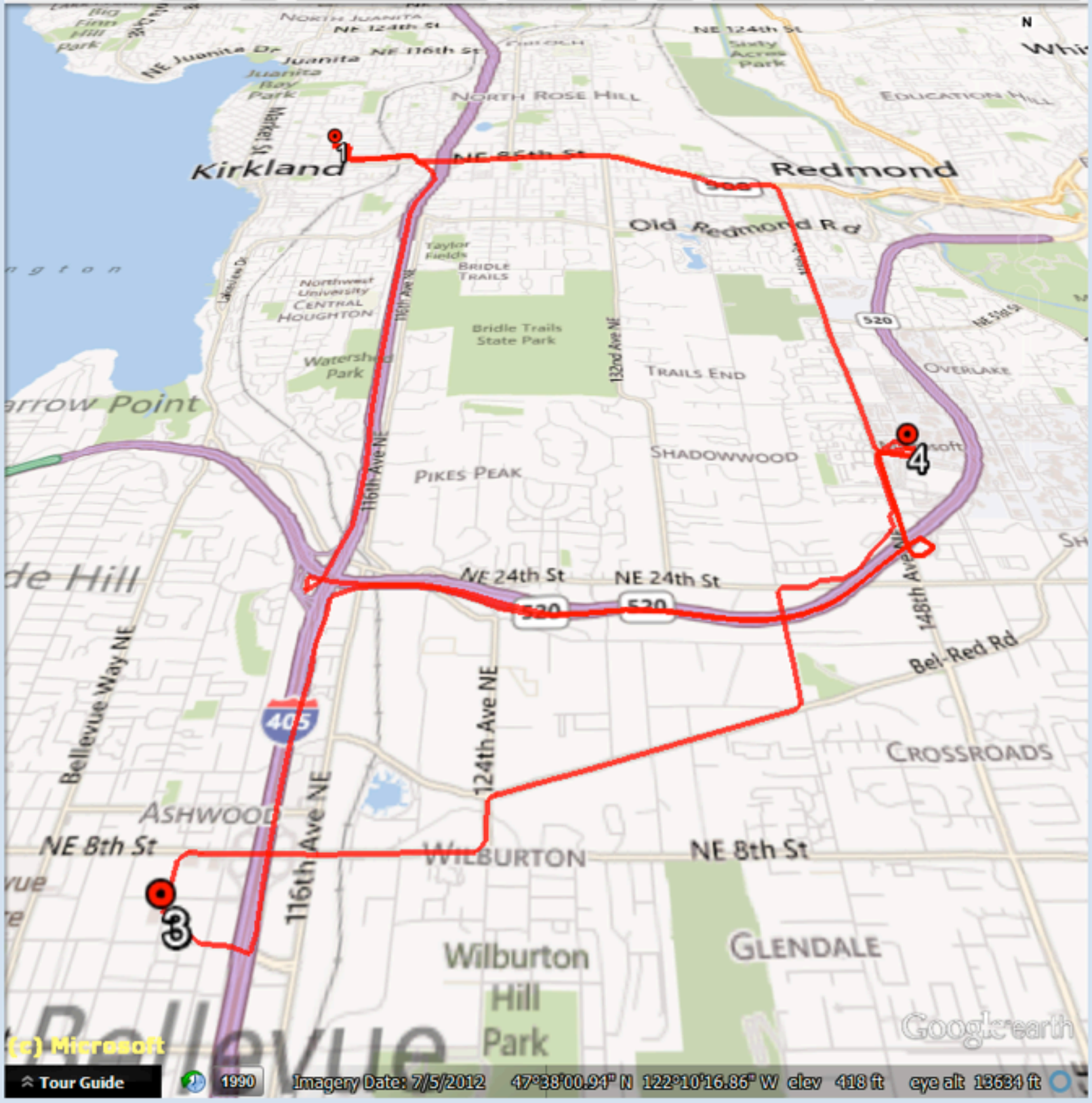
*Temporal relationships



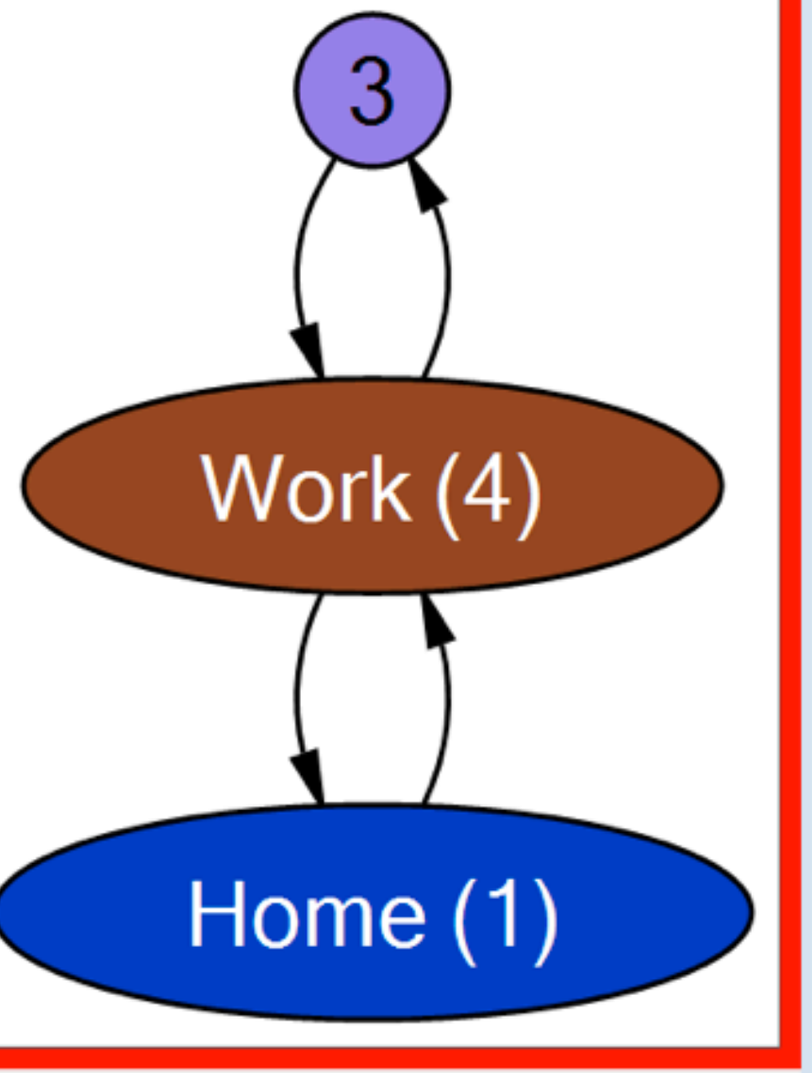
Google Earth

File Edit View Tools Add Help

Sign in

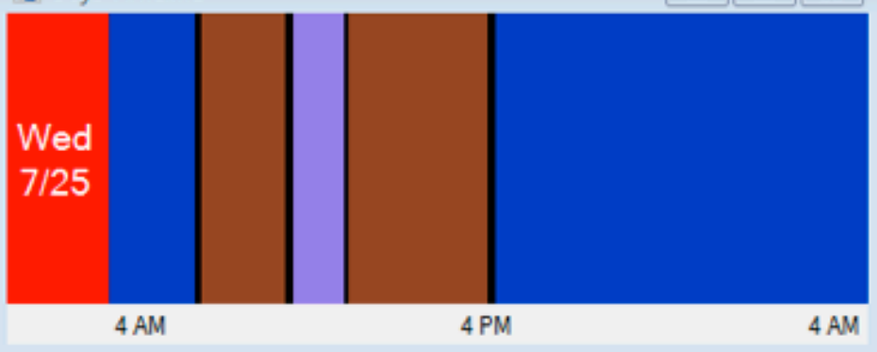


Days Graph



Wed 7/25

Days Timeline



Google Earth

File Edit View Tools Add Help

Sign in

Redmond

Old Redmond Rd

NE 124th St

NE 116th St

NE 24th St

Bel-Red Rd

NE 8th St

148th Ave NE

4th Ave NE

132nd Ave NE

Bridle Trails State Park

Marymoor Park

Sixty Acres Park

Crossroads Shopping Center

Crossroads Park

Lake Sammamish Pkwy NE

Redmond Way

Tour Guide

1990

Imagery Date: 7/5/2012 47°38'30.30" N 122°08'43.43" W elev 399 ft eye alt 13686 ft

(c) Microsoft

Days Graph

3

Work (4)

Home (1)

Wed 7/25

Days Timeline

Wed 7/25

4 AM

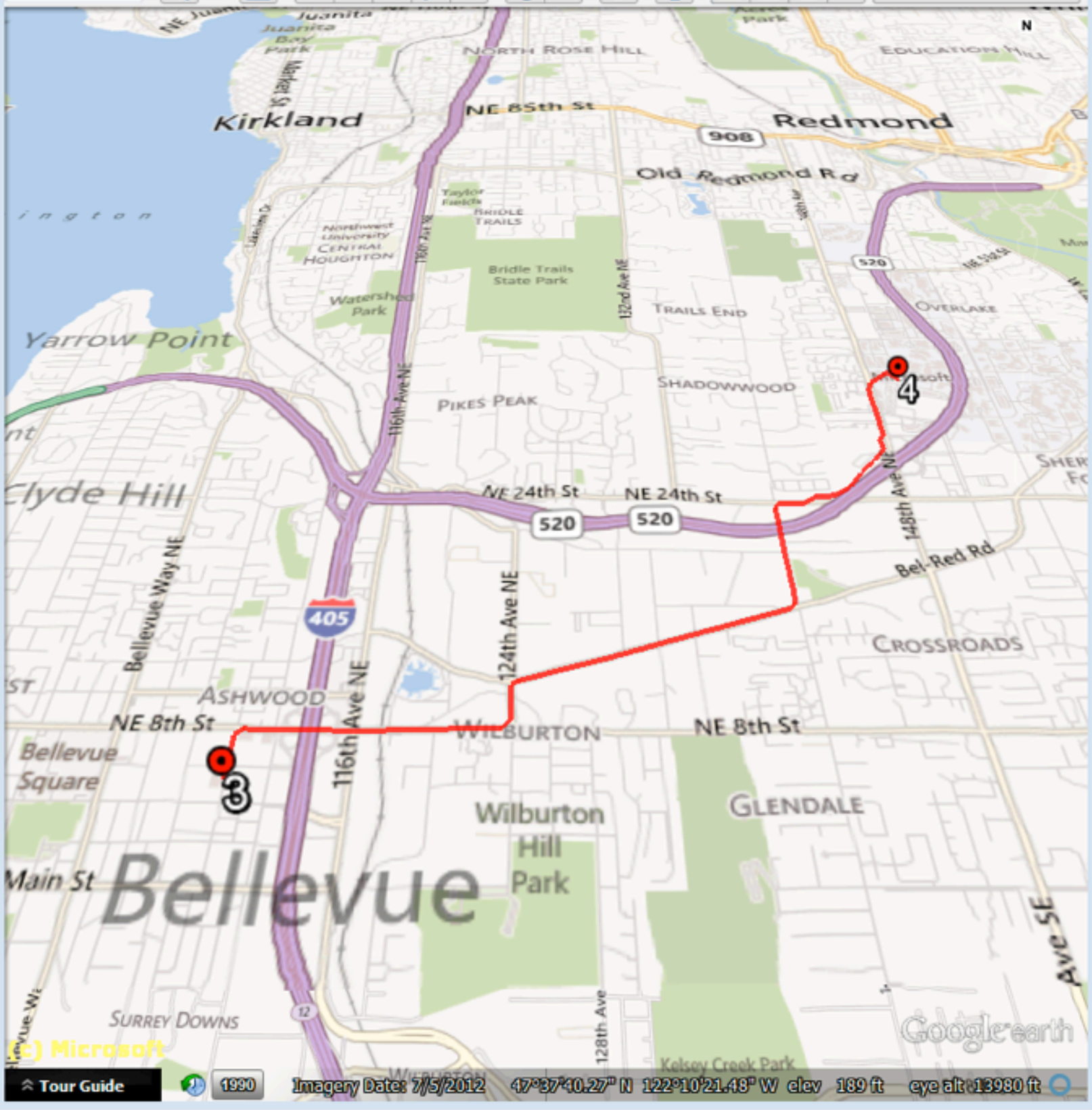
4 PM

4 AM

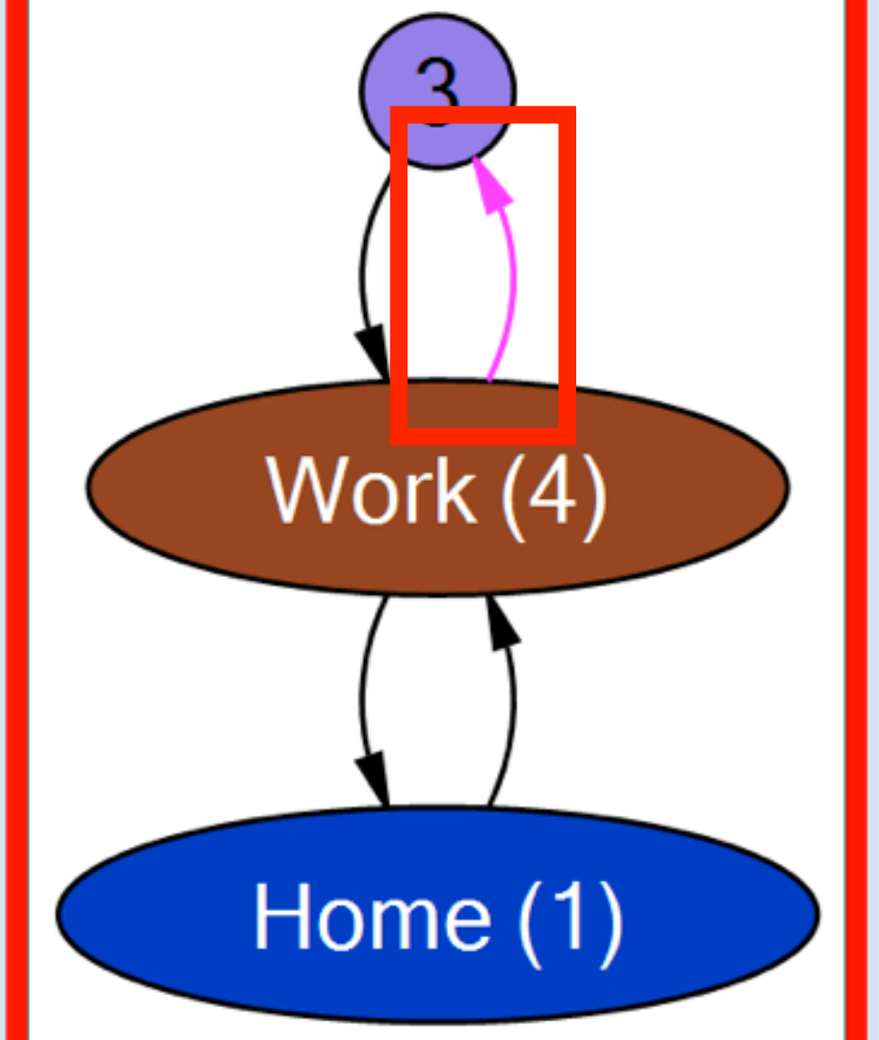
Google Earth

File Edit View Tools Add Help

Sign in

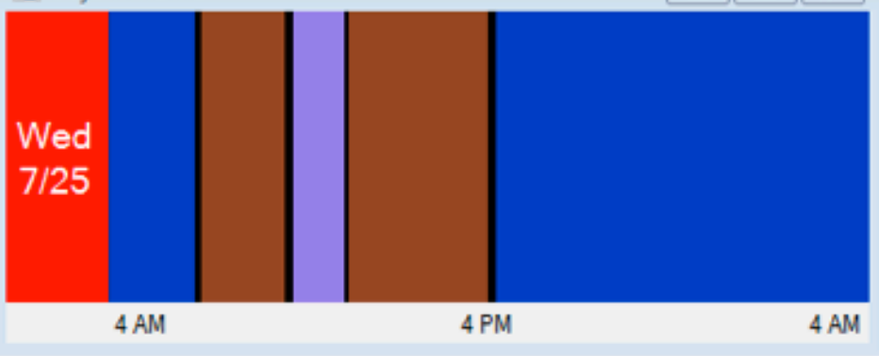


Days Graph



Wed 7/25

Days Timeline



Google Earth

File Edit View Tools Add Help

Sign in

Kirkland

Redmond

Yarrow Point

Clyde Hill

Bellevue

Bellevue Square

Wilburton Hill Park

Bridle Trails State Park

Old Redmond Rd

NE 24th St

NE 8th St

Tour Guide

1990

Imagery Date: 7/5/2012 47°38'00.75" N 122°10'44.84" W elev 246 ft eye alt 16120 ft

(c) Microsoft

Days Graph

3

Work (4)

Home (1)

Work (4)

Home (1)

21

Wed 7/25

Wed 8/29

Days Timeline

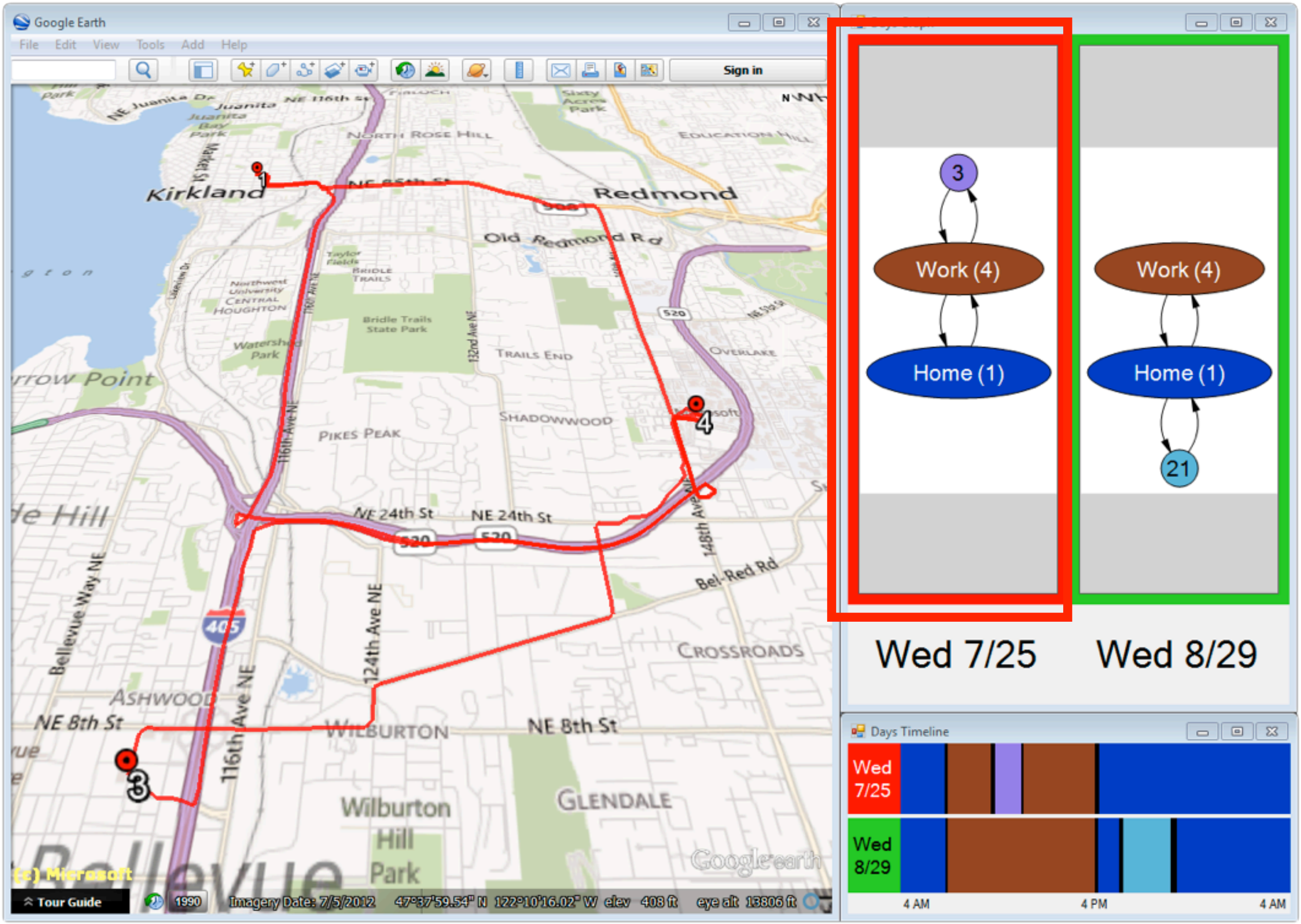
Wed 7/25

Wed 8/29

4 AM

4 PM

4 AM


Google Earth
File Edit View Tools Add Help
Sign in
Kirkland
Redmond
Old Redmond Rd
NE 24th St
NE 8th St
Bellevue
Wilburton
Wilburton Hill Park
Crossroads
Glendale
Ashwood
Pikes Peak
Shadowwood
Trails End
Bridle Trails State Park
North Rose Hill
Education Hill
Overlake
Bel-Red Rd
116th Ave NE
124th Ave NE
132nd Ave NE
148th Ave NE
Bellevue Way NE
(c) Microsoft
Tour Guide
Imagery Date: 7/5/2012 47°37'59.54" N 122°10'16.02" W elev 408 ft eye alt 13806 ft
3
Work (4)
Home (1)
Work (4)
Home (1)
21
Wed 7/25
Wed 8/29
Days Timeline
Wed 7/25
Wed 8/29
4 AM
4 PM
4 AM

Google Earth

File Edit View Tools Add Help

Sign in

Kirkland

Redmond

Yarrow Point

Clyde Hill

Bellevue

Old Redmond Rd

NE 24th St

NE 8th St

520

405

Bellevue Square

Wilburton Hill Park

Tour Guide

1990

Imagery Date: 7/5/2012 47°37'54.16" N 122°10'44.69" W elev 202 ft eye alt 13966 ft

(c) Microsoft

Days Graph

3

Work (4)

Home (1)

Work (4)

Home (1)

21

Wed 7/25

Wed 8/29

Days Timeline

Wed 7/25

Wed 8/29

4 AM

4 PM

4 AM

Days Graph

Home (1)

Work (4)

8

Home (1)

Mon 7/16

Sun 7/29

Days Surv...

Survey trial #1/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

5

Home (1)

Work (4)

Home (1)

Work (4)

Thu 7/26

Tue 7/31

Days Timeline

Mon 7/16

Sun 7/29

4 AM

4 PM

4 AM

Days Timeline

Thu 7/26

Tue 7/31

4 AM

4 PM

4 AM

Days Graph

Home (1)

Work (4)

8

Home (1)

Mon 7/16

Sun 7/29

Days Surv...

Survey trial #1/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

5

Home (1)

Work (4)

Home (1)

Work (4)

Thu 7/26

Tue 7/31

Days Timeline

Mon 7/16

Sun 7/29

4 AM

4 PM

4 AM

Days Timeline

Thu 7/26

Tue 7/31

4 AM

4 PM

4 AM

Days Graph

Home (1)

Work (4)

8

Home (1)

Mon 7/16

Sun 7/29

Days Surv...

Survey trial #1/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

5

Home (1)

Work (4)

Home (1)

Work (4)

Thu 7/26

Tue 7/31

Days Timeline

Mon 7/16

Sun 7/29

4 AM

4 PM

4 AM

Days Timeline

Thu 7/26

Tue 7/31

4 AM

4 PM

4 AM

"Yes, I did the same things today as I did yesterday."

Google Earth

File Edit View Tools Add Help

Sign in

Kirkland
Redmond
Bridle Trails State Park
Watershed Park
Everest Park
Taylor Fields
Northwest University
Central Houghton
Bridle Trails
Compton Trails
Pikes Peak
Bellemead
Trails End
Shadowwood
Grass Lawn
South Rose Hill
Rose Hill
Sammamish Valley
River Trail
Overlake
Redmond Town Center
Education Hill
Jonathan Hartman Park
Norkirk
Highlands
Ferriton
NE 68th St
NE 70th St
NE 37th St
NE 40th St
NE 38th St
NE 32nd St
NE 30th St
NE 51st St
Redmond Way
Lake St S
Lakeview Dr
Market St
Washington Blvd NE
Northup Way
116th Ave NE
124th Ave NE
132nd Ave NE
134th Ave NE
148th Ave NE
405
520
202
1
4
8

(c) Microsoft

Google earth

Tour Guide

1990

Imagery Date: 7/5/2012 47°39'17.27" N 122°09'56.27" W elev 511 ft eye alt 13185 ft

Google Earth

File Edit View Tools Add Help

Sign in

Kirkland
Redmond
Bridle Trails State Park
Northwest University
Central Houghton
Taylor Fields
Bridle Trails
Compton Trails
Pikes Peak
Shadowwood
Trails End
Overlake
Grass Lawn
South Rose Hill
Rose Hill
Sammamish Valley
Education Hill
Jonathan Hartman Park
Nike Park
River Trail
Redmond Town Center
Town Center
Everest Park
Feriton
Norkirk
Highlands
Watershed Park
Bellemead
NE 70th St
NE 68th St
NE 37th St
NE 40th St
NE 38th St
NE 36th St
NE 32nd St
NE 51st St
132nd Ave NE
134th Ave NE
116th Ave NE
148th Ave NE
124th
Lake St S
Lakeview Dr
405
520
202

1 1 4 4 5

(c) Microsoft

Google earth

Tour Guide 1990

Imagery Date: 7/5/2012 47°39'21.33" N 122°09'36.07" W elev 452 ft eye alt 13809 ft

Days Graph

Home (1)

Work (4)

8

Home (1)

Mon 7/16

Sun 7/29

Days Surv...

Survey trial #1/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

5

Home (1)

Work (4)

Home (1)

Work (4)

Thu 7/26

Tue 7/31

Days Timeline

Mon 7/16

Sun 7/29

4 AM

4 PM

4 AM

Days Timeline

Thu 7/26

Tue 7/31

4 AM

4 PM

4 AM

Days Graph

Home (1)

Work (4)

8

Home (1)

Mon 7/16

Sun 7/29

Days Surv...

Survey trial #1/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

5

Home (1)

Work (4)

Home (1)

Work (4)

Thu 7/26

Tue 7/31

Days Timeline

Mon 7/16

Sun 7/29

4 AM

4 PM

4 AM

Days Timeline

Thu 7/26

Tue 7/31

4 AM

4 PM

4 AM

Days Graph

Work (4)

5

Home (1)

3

Work (4)

Home (1)

Thu 7/12

Wed 7/25

Days Surv...

Survey trial #7/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

Home (1)

Work (4)

3

Home (1)

Work (4)

Mon 7/30

Wed 8/8

Days Timeline

Thu 7/12

Wed 7/25

4 AM

4 PM

4 AM

Days Timeline

Mon 7/30

Wed 8/8

4 AM

4 PM

4 AM

Days Graph

Work (4)

5

Home (1)

3

Work (4)

Home (1)

Thu 7/12

Wed 7/25

Days Surv...

Survey trial #7/10

Left

←

Which pair of days are most similar?

→

Right

Days Graph

Home (1)

Work (4)

3

Home (1)

Work (4)

Mon 7/30

Wed 8/8

Days Timeline

Thu 7/12

Wed 7/25

4 AM

4 PM

4 AM

Days Timeline

Mon 7/30

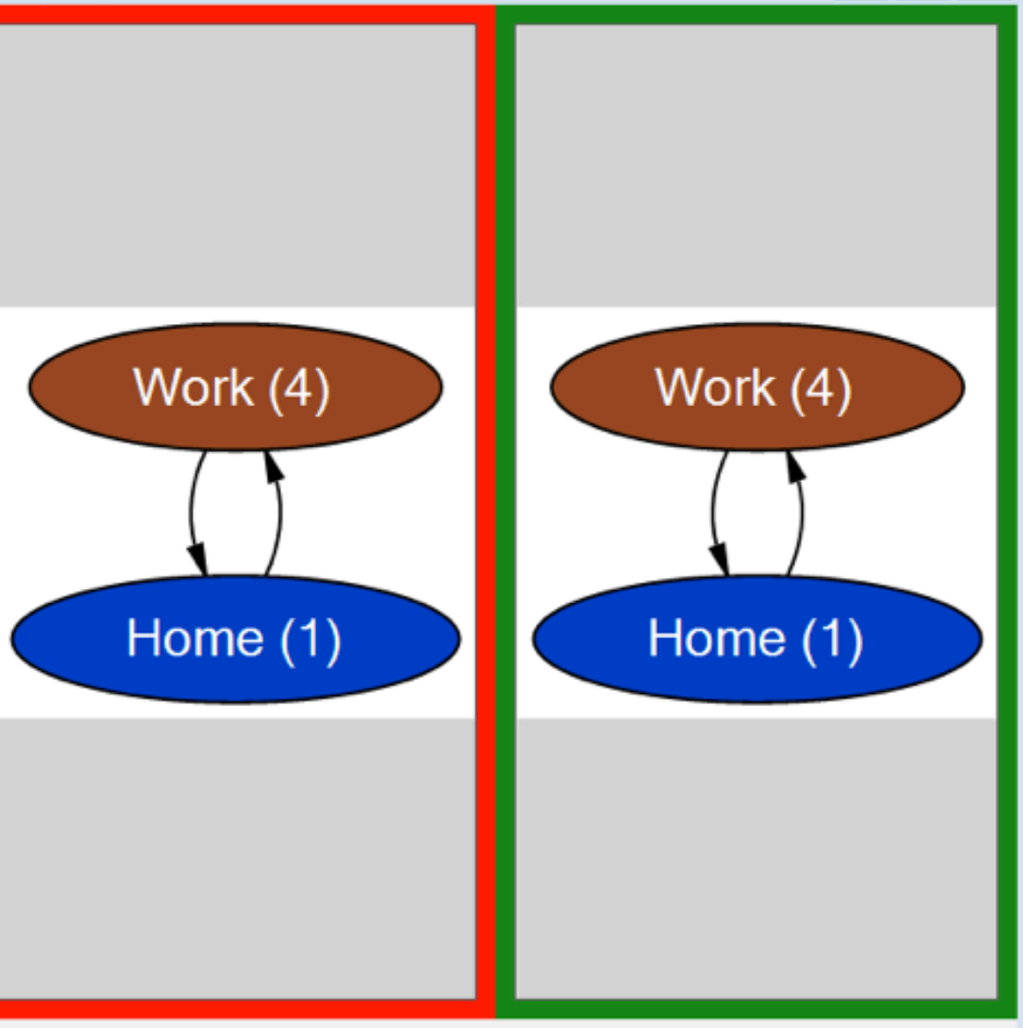
Wed 8/8

4 AM

4 PM

4 AM

Days Graph



Mon 7/16 Tue 8/21

Days Surv...

Survey trial #10/10
Last one!

Left

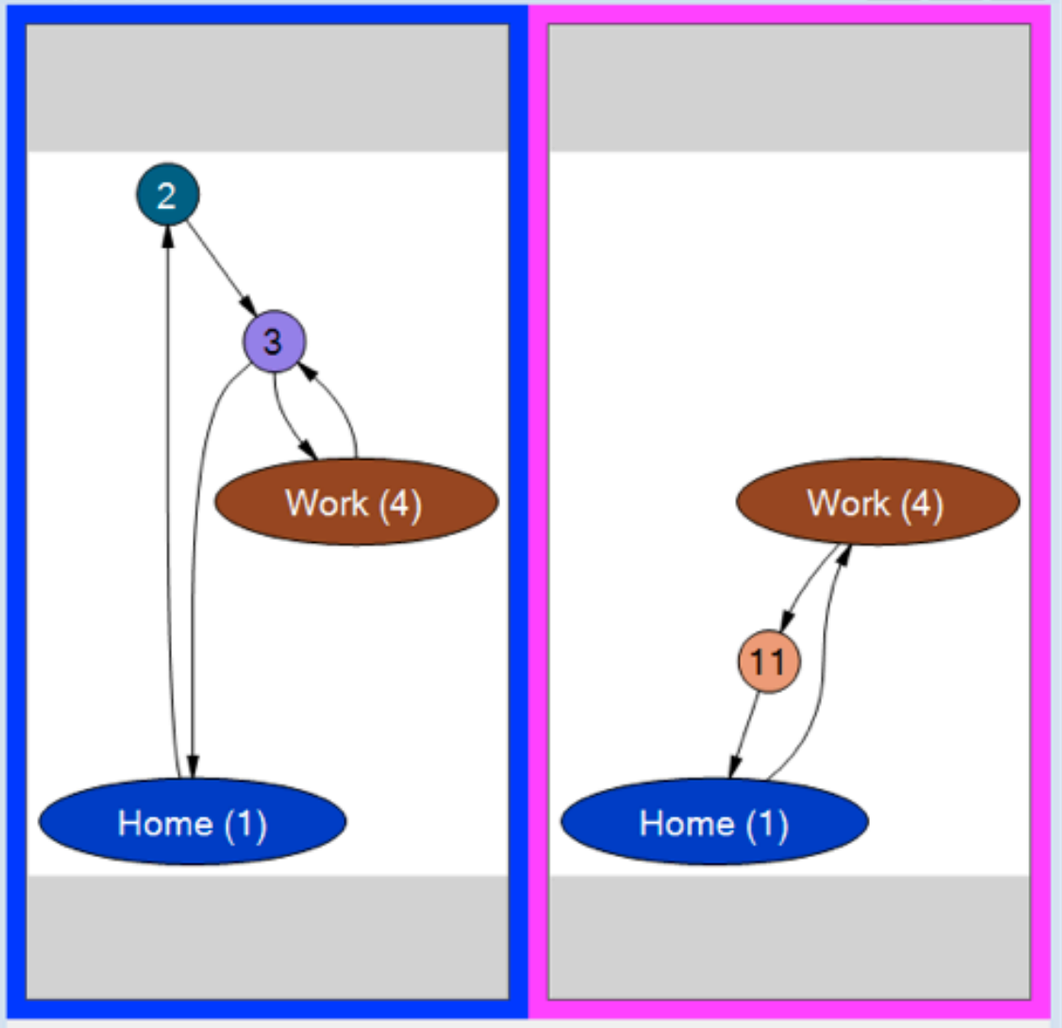
←

Which pair of days are most similar?

→

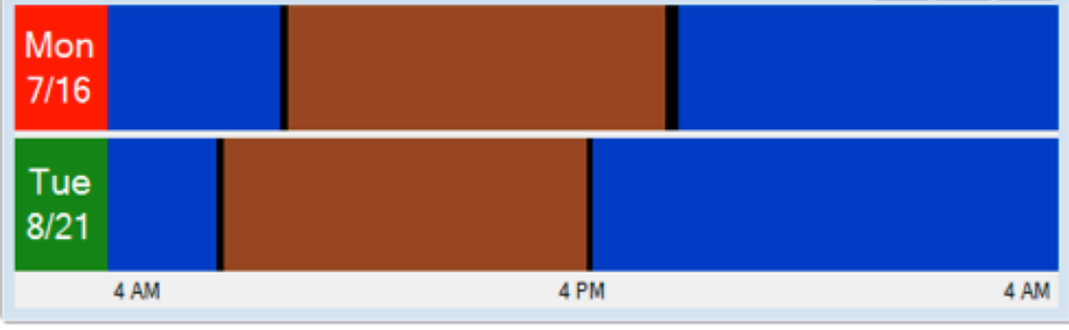
Right

Days Graph

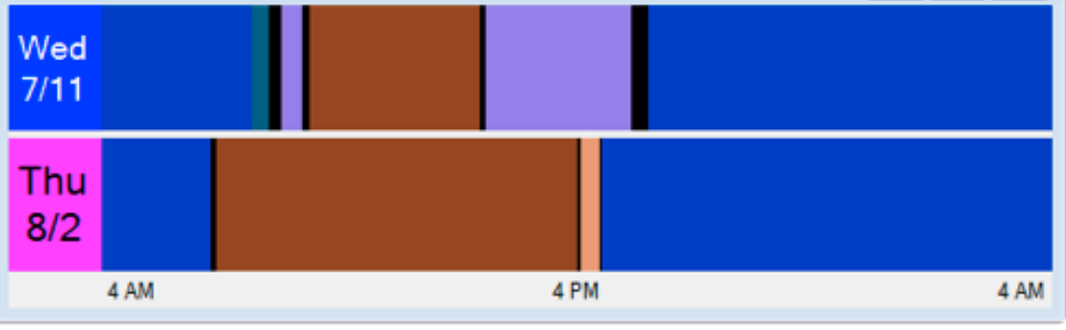


Wed 7/11 Thu 8/2

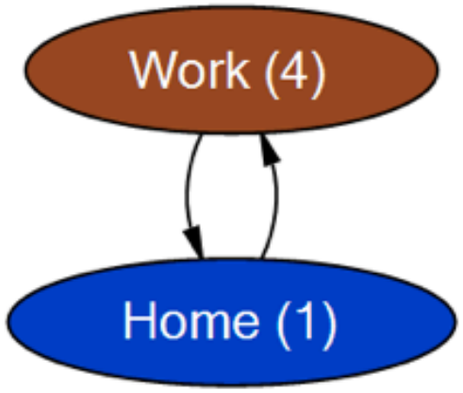
Days Timeline

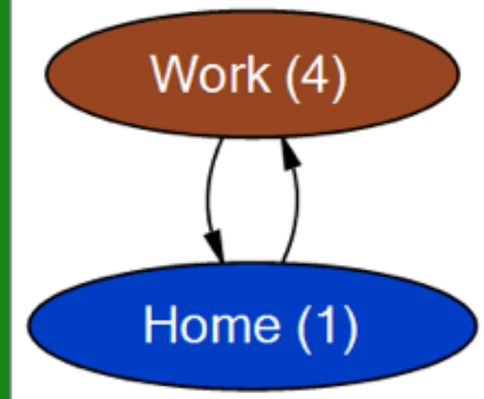


Days Timeline



Days Graph





Mon 7/16 Tue 8/21

Days Surv...

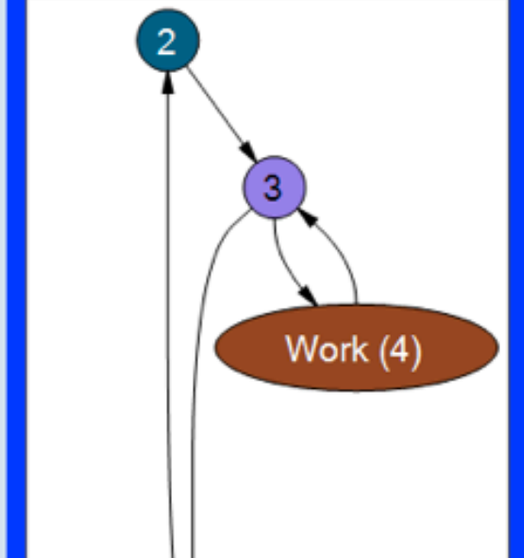
Survey trial #10/10
Last one!

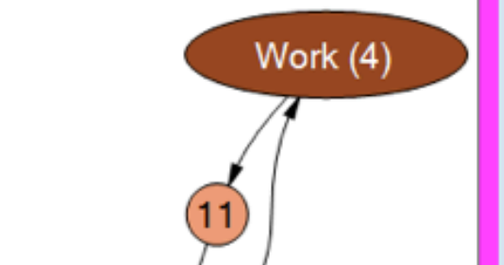
Left

Which pair of days are most similar?

Right

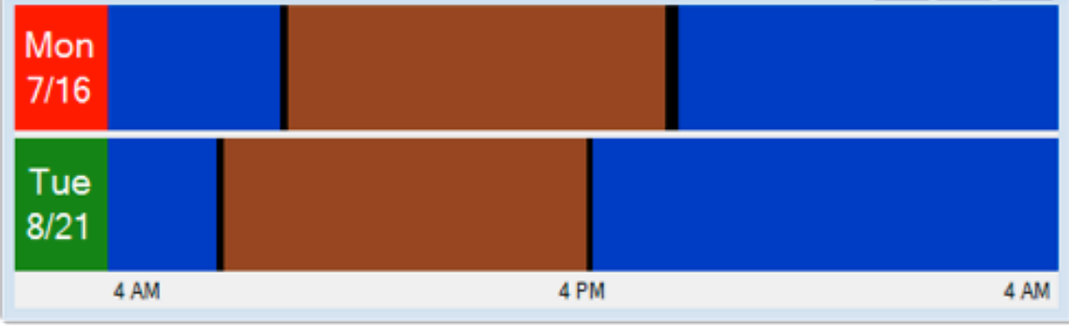
Days Graph



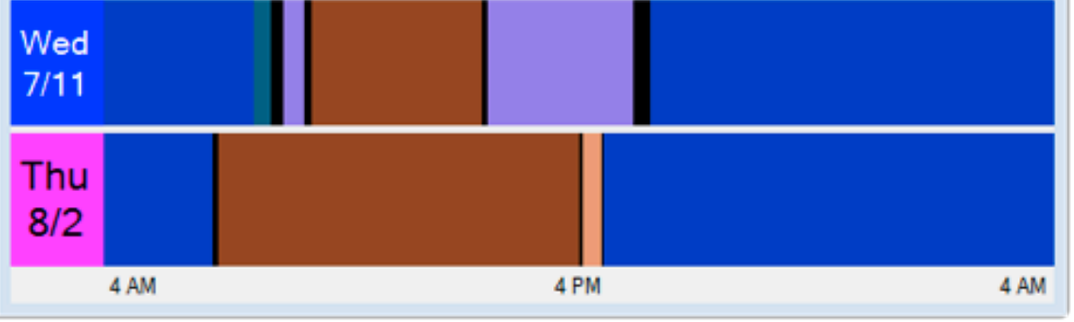


Wed 7/11 Thu 8/2

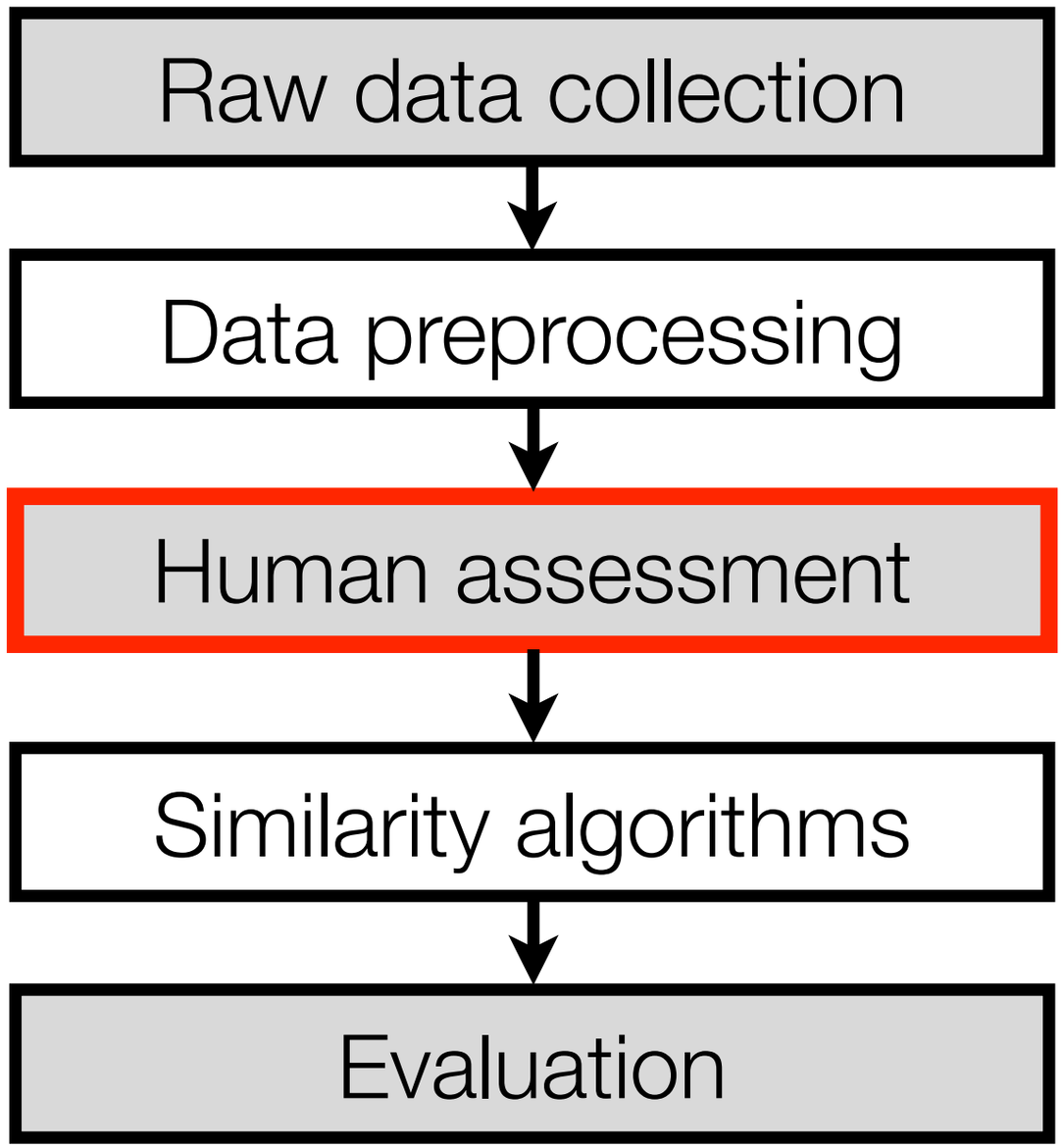
Days Timeline



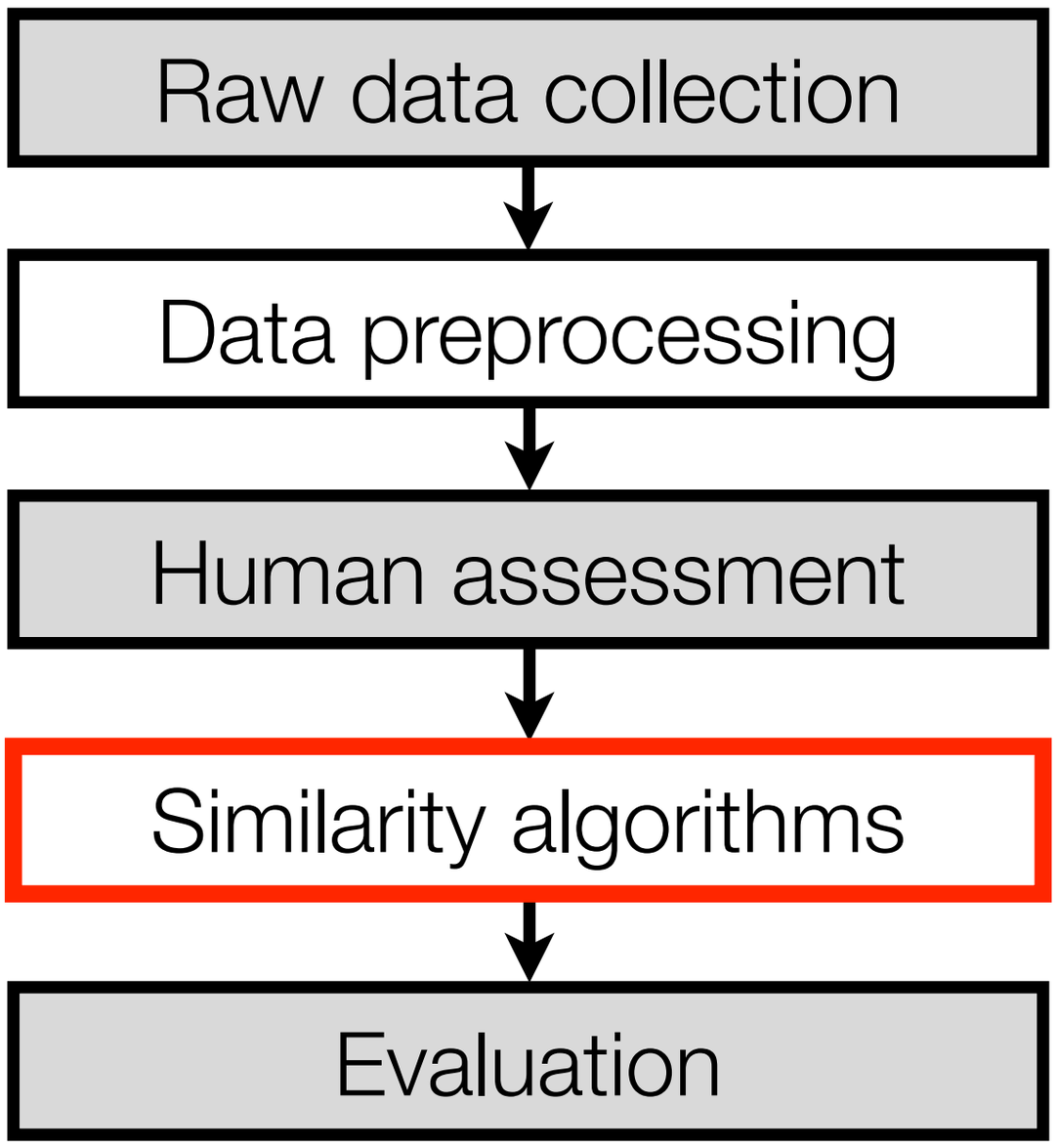
Days Timeline



Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

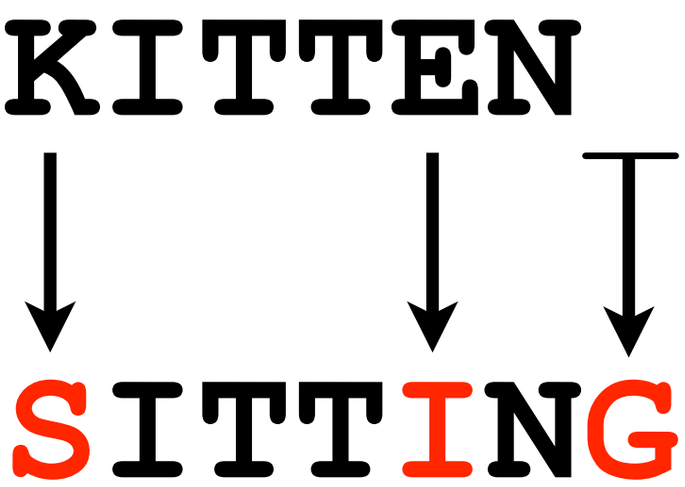
Similarity algorithms

- Edit distance
- Distance sensitive edit distance
- Stop type distance
- Sum of pairs distance

Edit distance

- Levenshtein distance
- Transform one string into another
  - Edit operations: insertion, deletion, substitution
- Example: kitten → sitting
  - Distance = 3

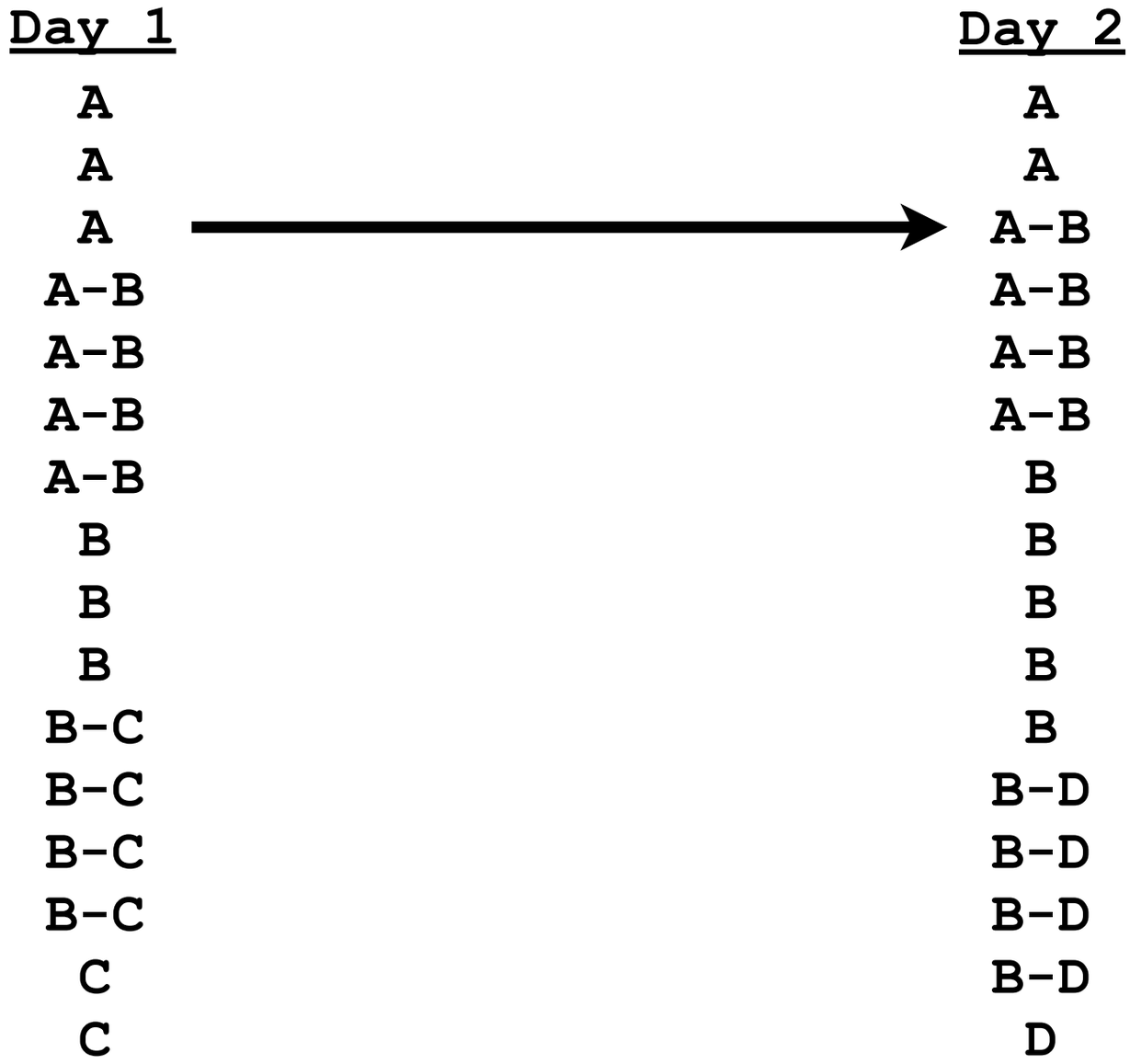
kitten → sitting



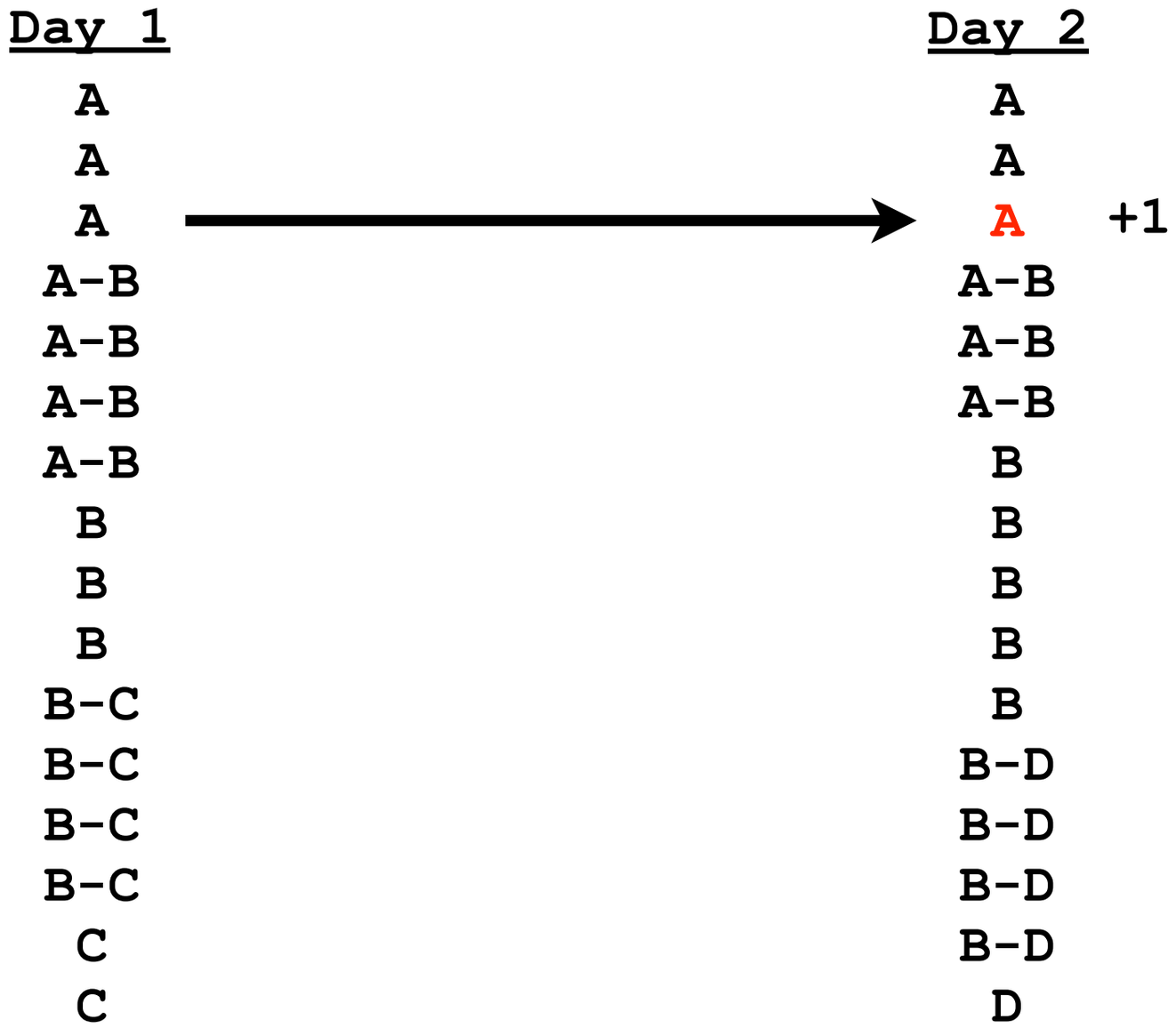
Edit distance between days

Day 1		Day 2
A		A
A		A
A		A-B
A-B		A-B
A-B		A-B
A-B		A-B
B		B
B		B
B		B
B-C		B
B-C		B-D
B-C		B-D
B-C		B-D
B-C		B-D
C		D
C		

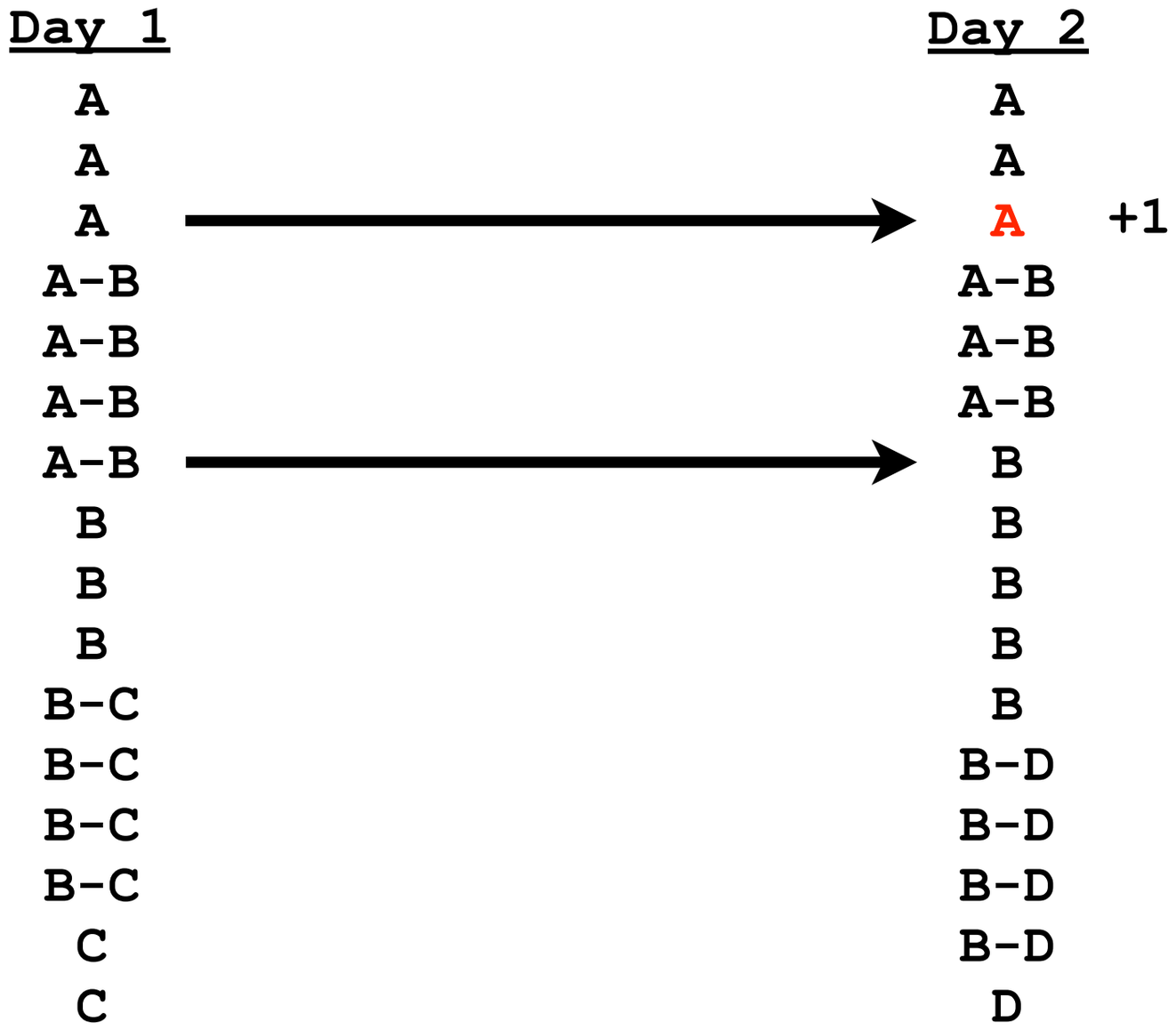
Edit distance between days



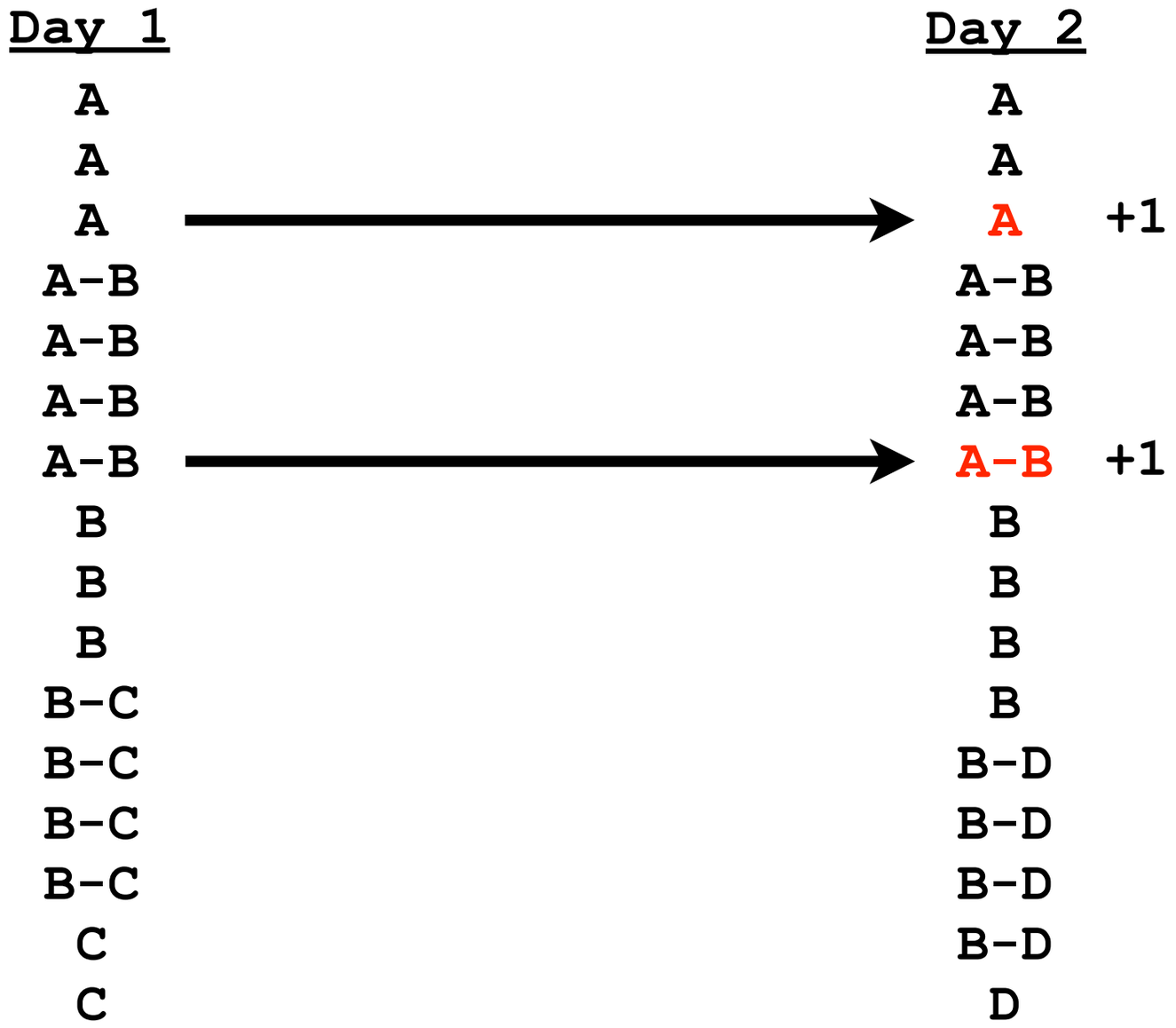
Edit distance between days



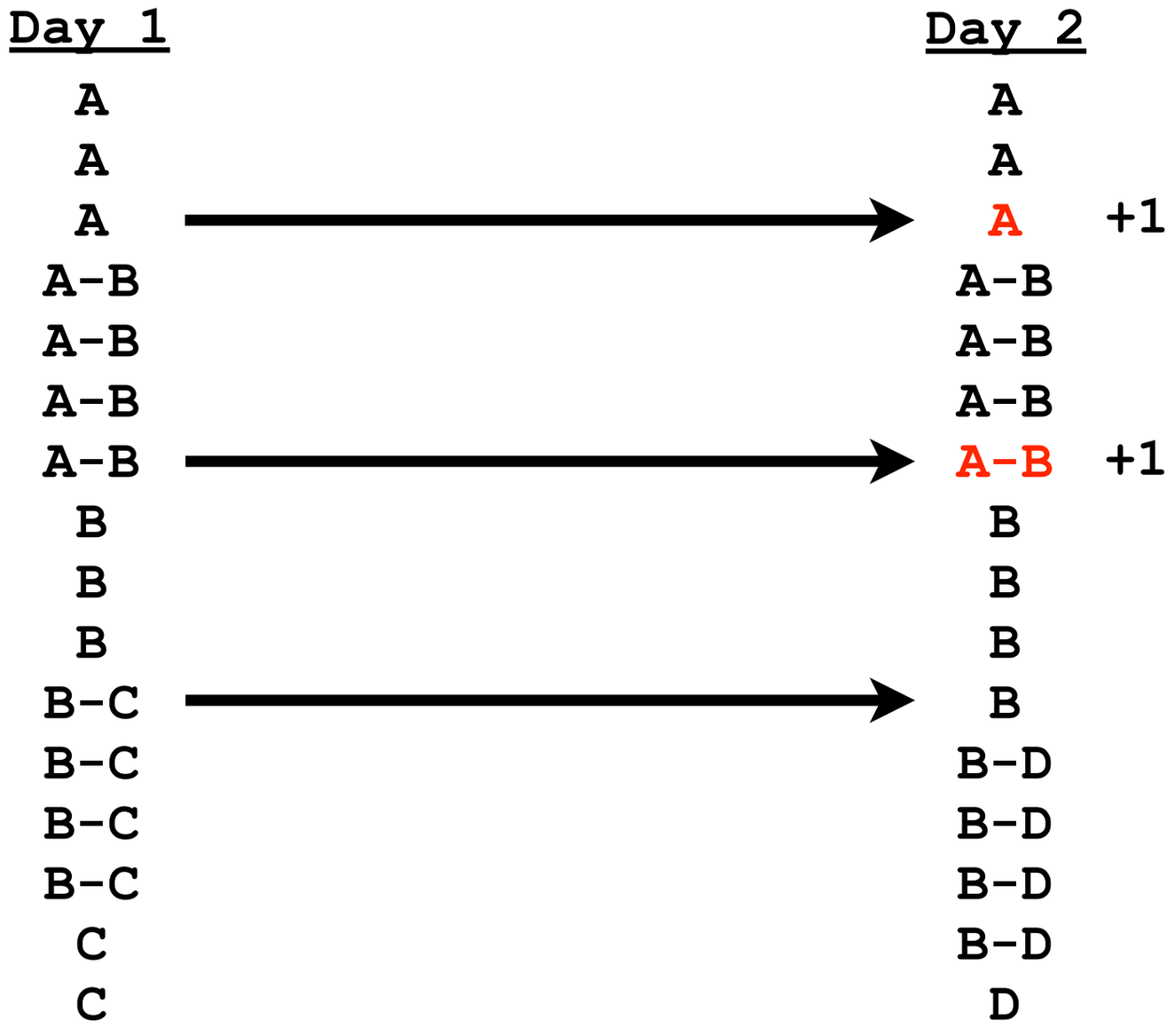
Edit distance between days



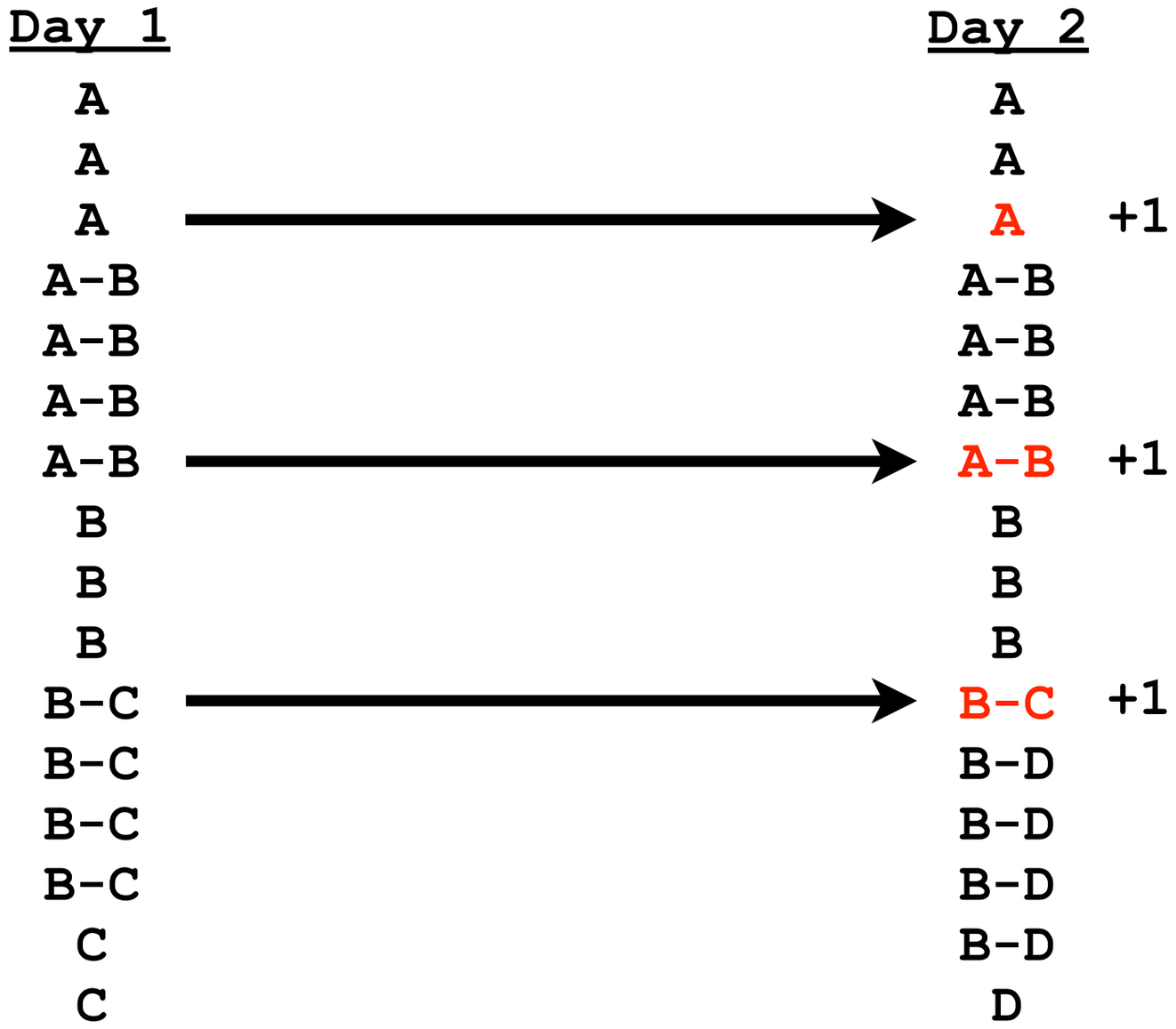
Edit distance between days



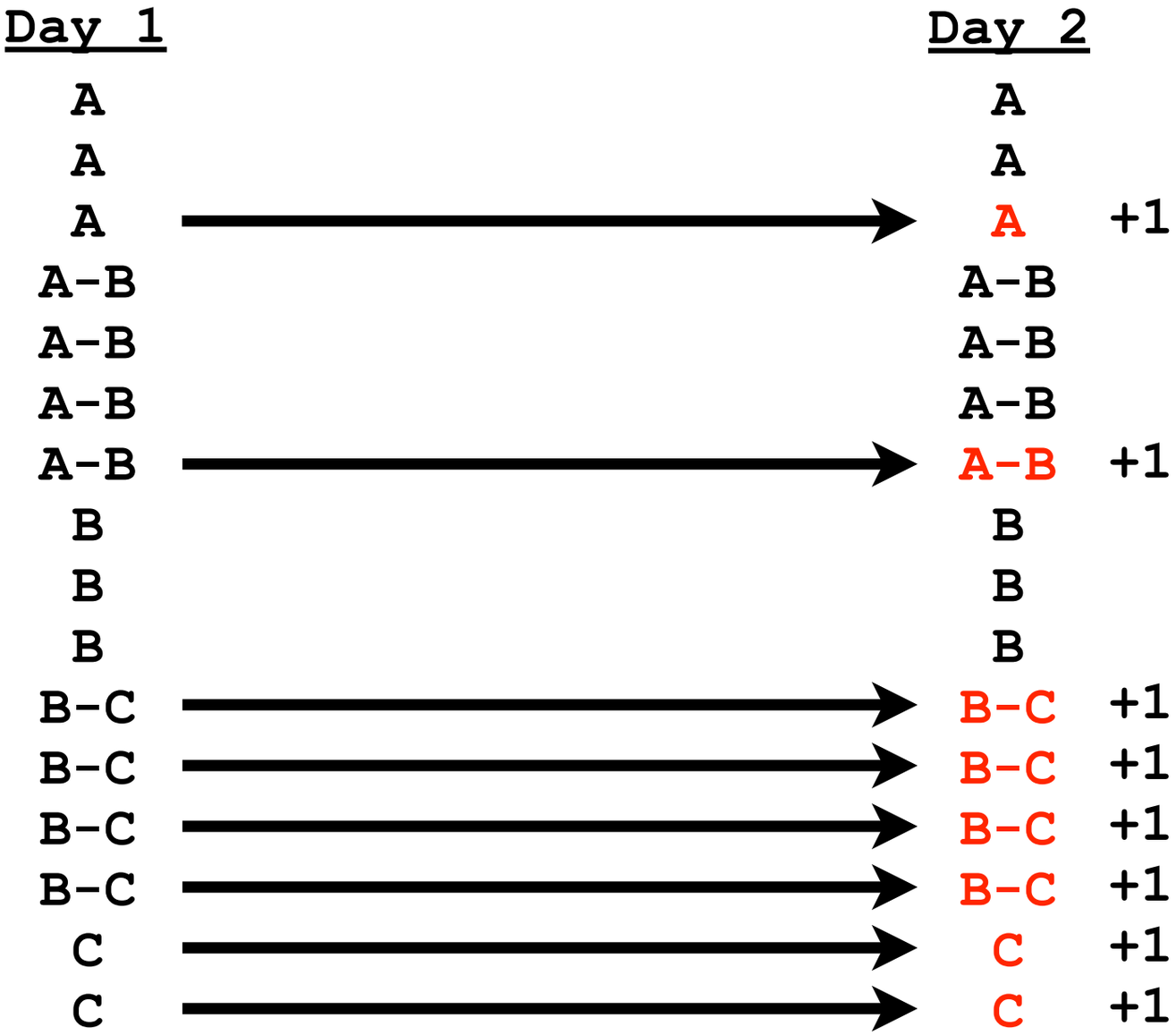
Edit distance between days



Edit distance between days

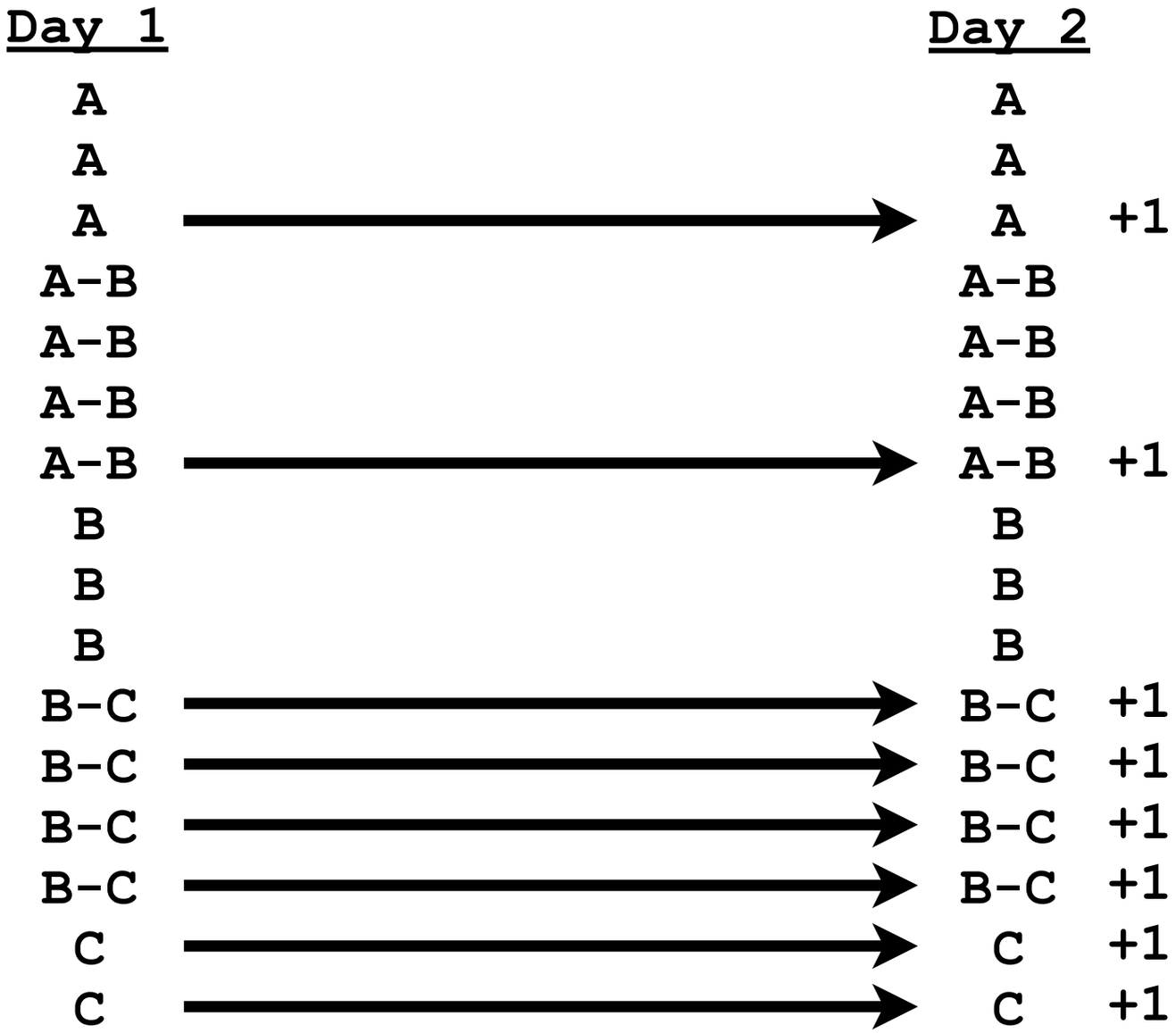


Edit distance between days

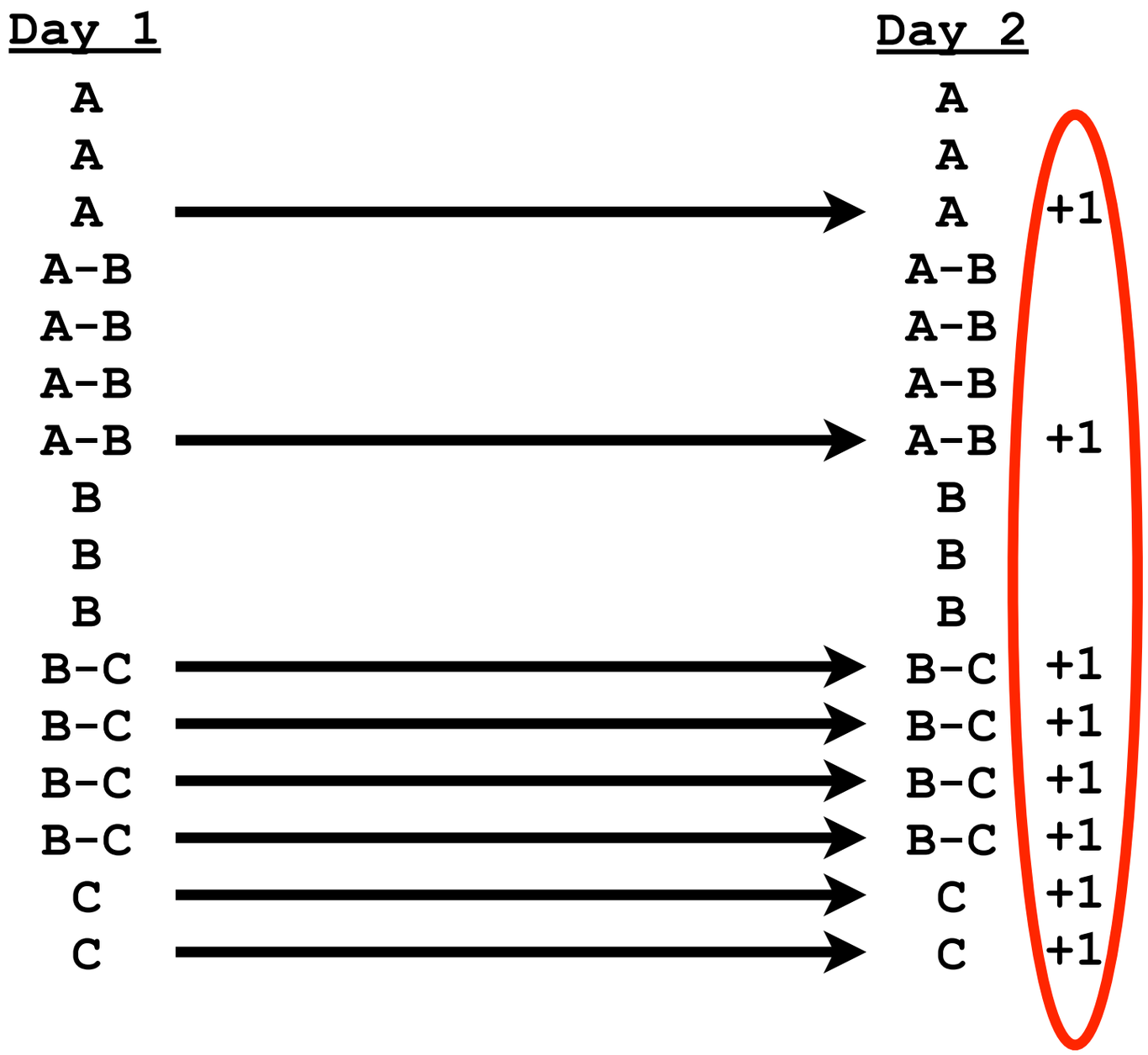


Distance sensitive edit distance

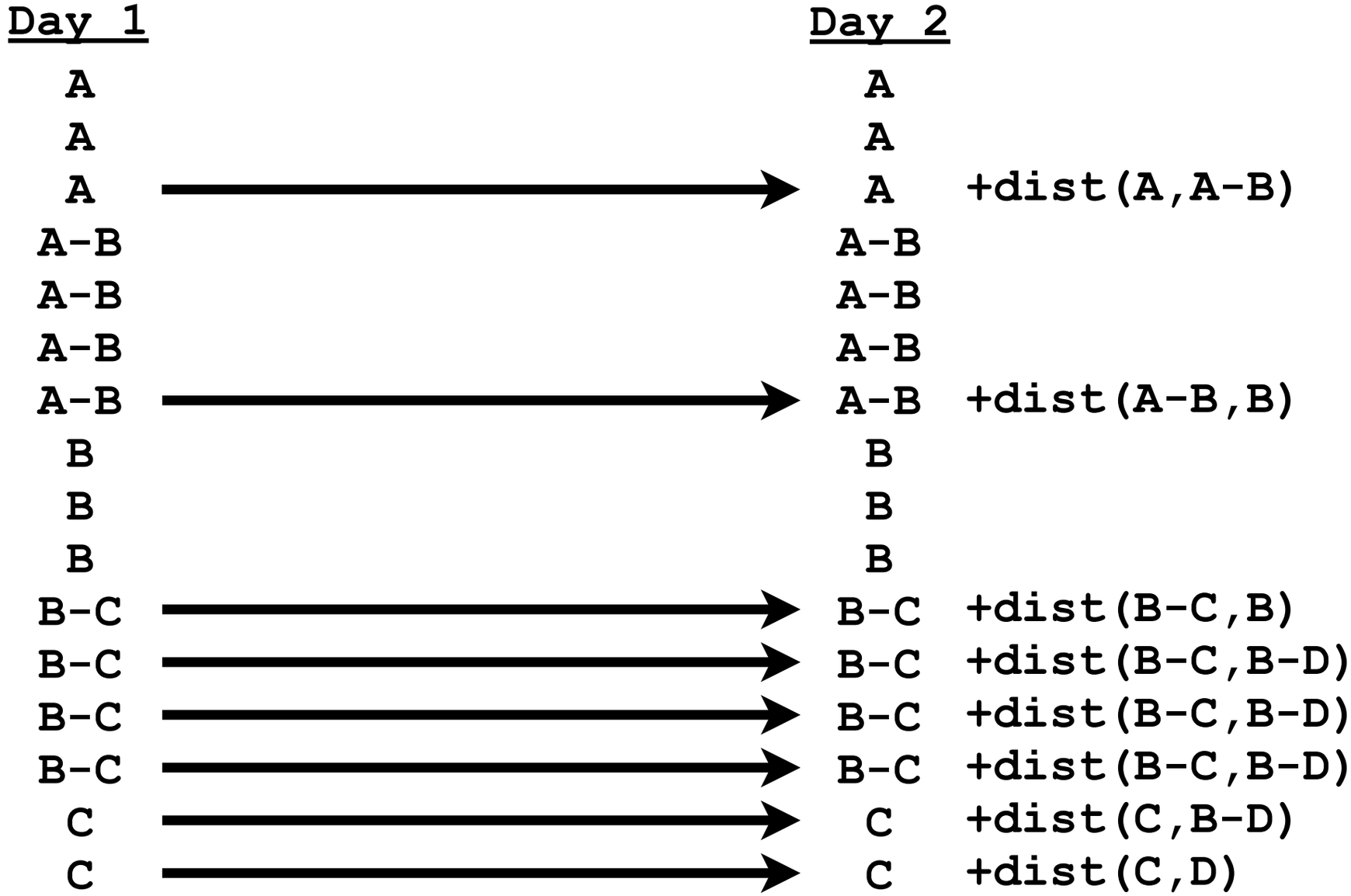
Geographic distance cost



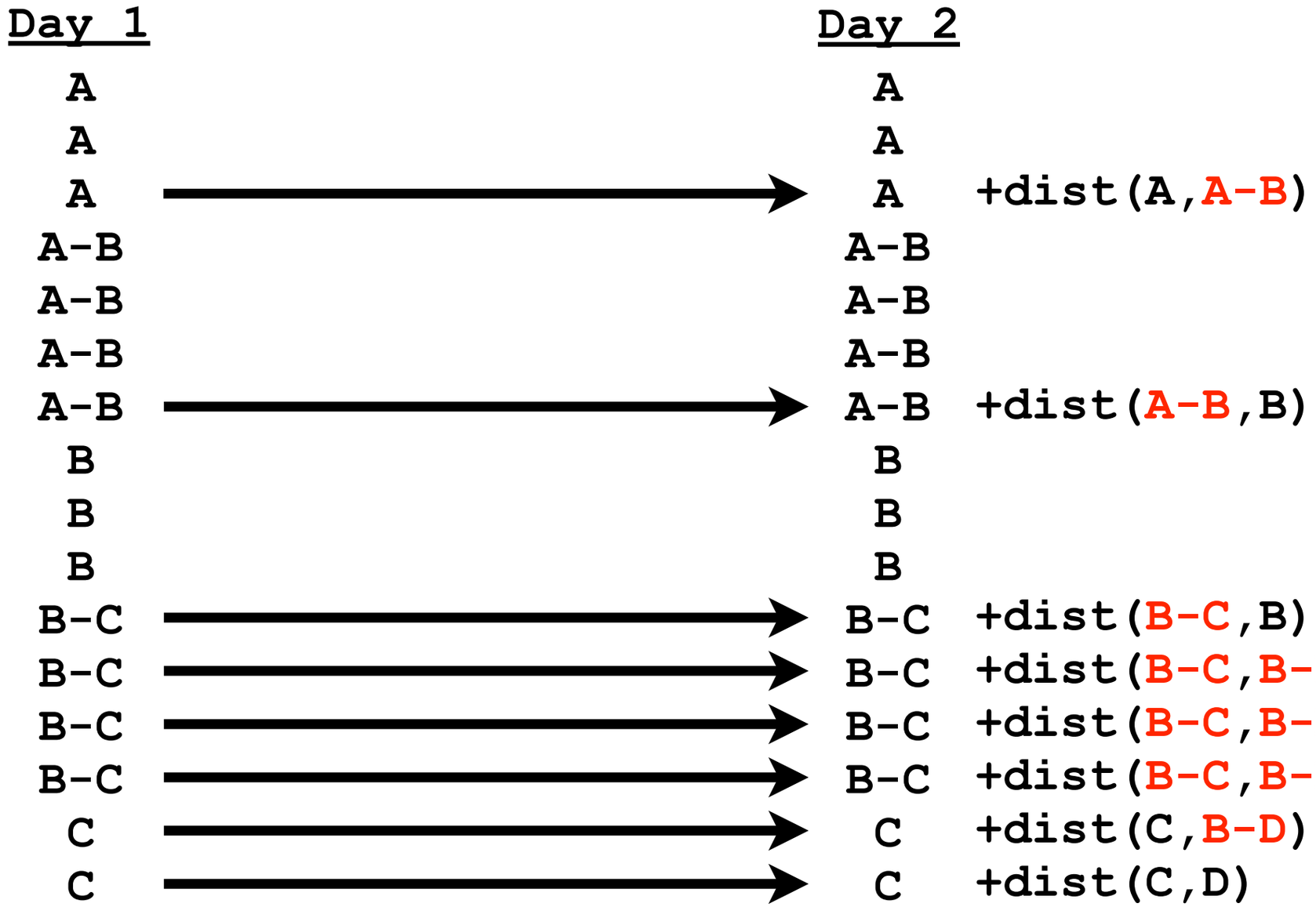
Geographic distance cost



Geographic distance cost



Geographic distance cost



Symbol string transformation

Day 1	Day 2
A	A
A	A
A	A-B
A-B	A-B
A-B	A-B
A-B	A-B
A-B	B
B	B
B	B
B	B
B-C	B
B-C	B-D
B-C	B-D
B-C	B-D
B-C	B-D
C	B-D
C	D

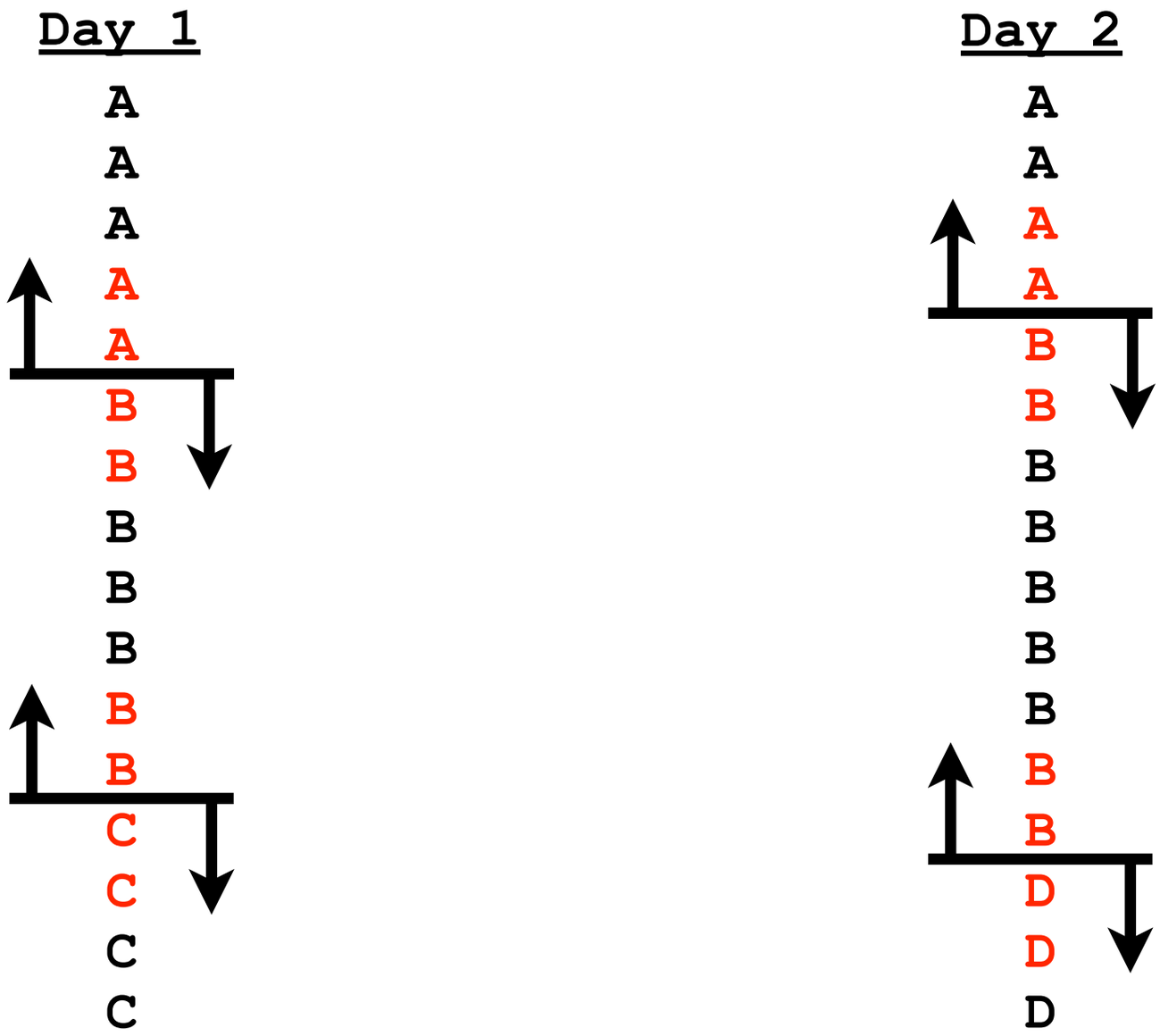
Symbol string transformation

Day 1	Day 2
A	A
A	A
A	A-B
A-B	A-B
A-B	A-B
A-B	A-B
B	B
B	B
B	B
B-C	B
B-C	B-D
B-C	B-D
B-C	B-D
C	B-D
C	D

Symbol string transformation

Day 1	Day 2
A	A
A	A
A	A-B
A-B	A-B
A-B	A-B
A-B	A-B
B	B
B	B
B	B
B	B
B-C	B
B-C	B-D
B-C	B-D
B-C	B-D
B-C	B-D
C	B-D
C	D

Symbol string transformation

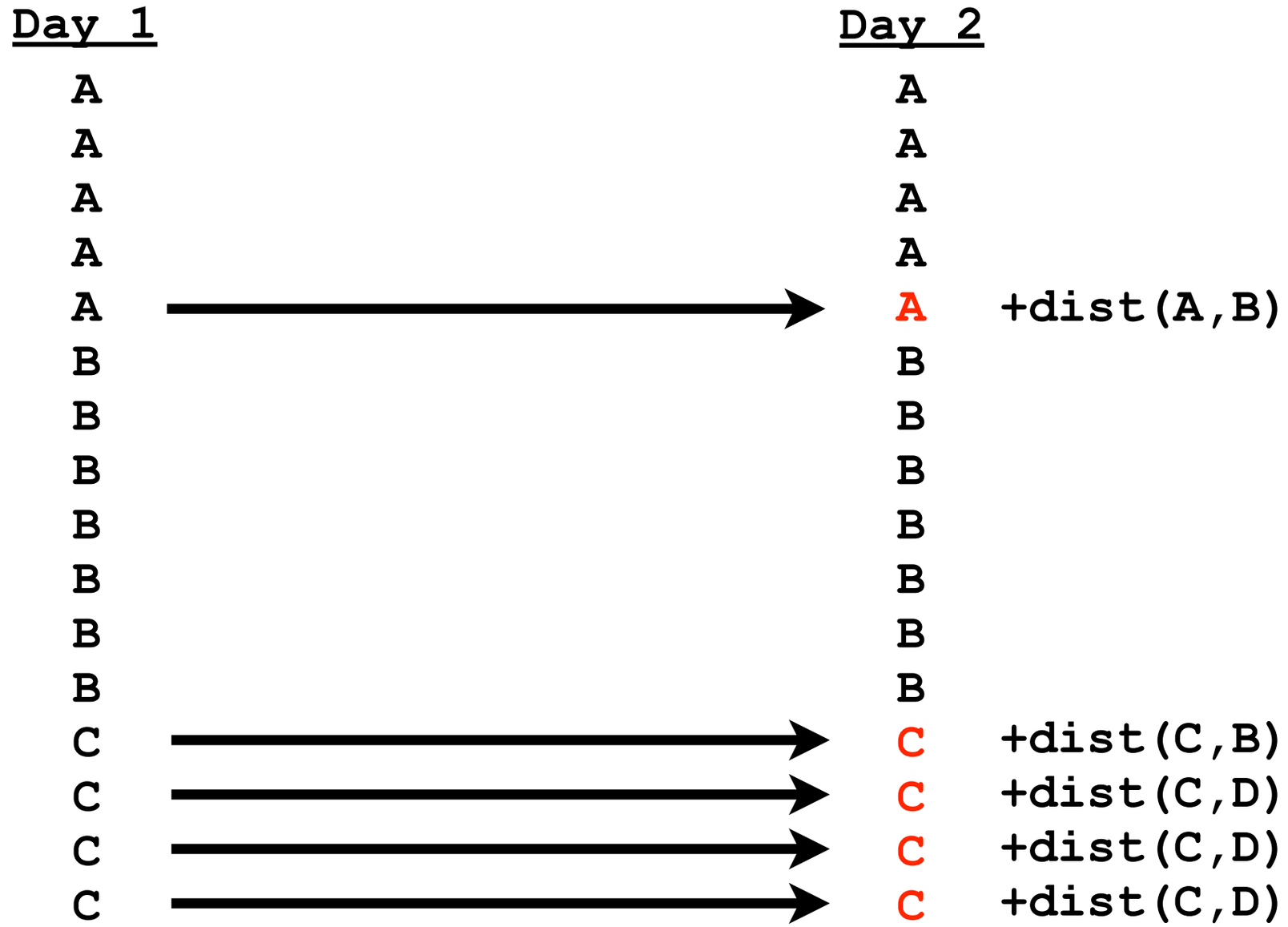


Distance sensitive edit distance

Day 1	Day 2
A	A
A	A
A	A
A	A
A	B
B	B
B	B
B	B
B	B
B	B
B	B
C	B
C	D
C	D
C	D



Distance sensitive edit distance



Stop type distance

bing Local search

- Geographic business database
- Input: coordinate pair, search radius
- Output(s): location, distance, category

bing Local search example

Input: (47.641954, -122.140497), 250m

bing Local search example

Input: (47.641954, -122.140497), 250m
Output:
1: Studio C, 141m, Hair & Beauty

bing Local search example

Input: (47.641954, -122.140497), 250m
Output:
1: Studio C, 141m, Hair & Beauty

Location-type distribution:
(Hair & Beauty, 1.0)

bing Local search example

Input: (47.625788, -122.148022), 250m

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

```
weight = 1 - (distance from stop/search radius)
```

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

1: Employment Services = 1 - (X/250)
2: Legal Services = 1 - (X/250)
3: Dental Care = 1 - (X/250)

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

1: Employment Services = 1 - (61/250) = 0.756
2: Legal Services = 1 - (64/250) = 0.744
3: Dental Care = 1 - (202/250) = 0.192

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

1: Employment Services = 1 - (61/250) = 0.756
2: Legal Services = 1 - (64/250) = 0.744
3: Dental Care = 1 - (202/250) = 0.192

Sum = 0.756 + 0.744 + 0.192 = 1.692

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

1: Employment Services = 0.756/1.692 = 0.447
2: Legal Services = 0.744/1.692 = 0.440
3: Dental Care = 0.192/1.692 = 0.113

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care

Location-type distribution:
(Employment Services, 0.447),
(Legal Services,0.440),
(Dental Care, 0.113)

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care
4: Dickens and Mason DDS, 205m, Dental Care

Location-type distribution:
(Employment Services, 0.404),
(Legal Services,0.397),
(Dental Care, 0.103),
(Dental Care, 0.096)

bing Local search example

Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care
4: Dickens and Mason DDS, 205m, Dental Care

Location-type distribution:
(Employment Services, 0.404),
(Legal Services,0.397),
(Dental Care, 0.103 + 0.096)

bing Local search example

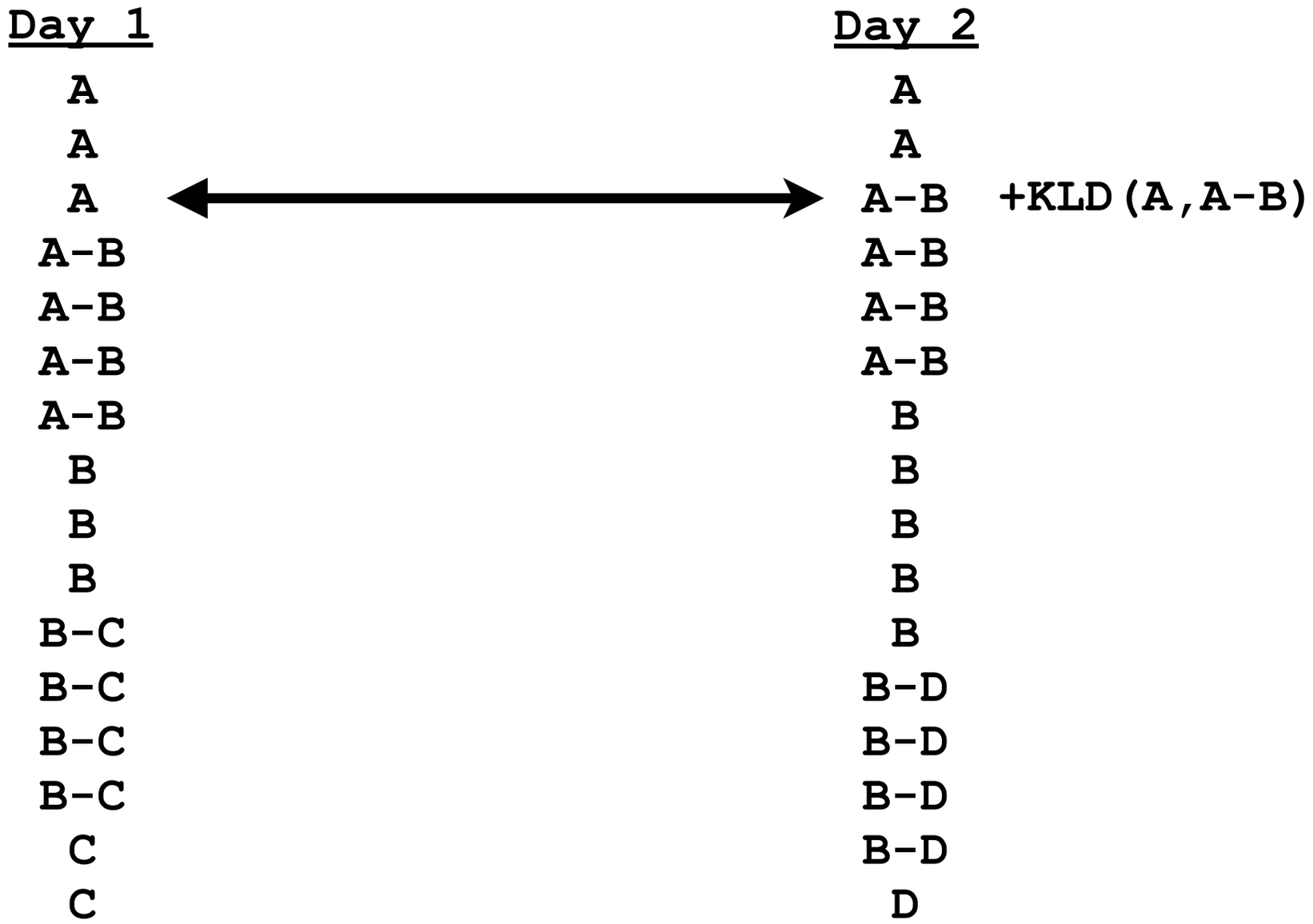
Input: (47.625788, -122.148022), 250m
Output:
1: A and H Partners, 61m, Employment Services
2: Deborah A Bianco PLLC, 64m, Legal Services
3: Rodney B Wentworth DDS, 202m, Dental Care
4: Dickens and Mason DDS, 205m, Dental Care

Location-type distribution:
(Employment Services, 0.404),
(Legal Services,0.397),
(Dental Care, 0.199)

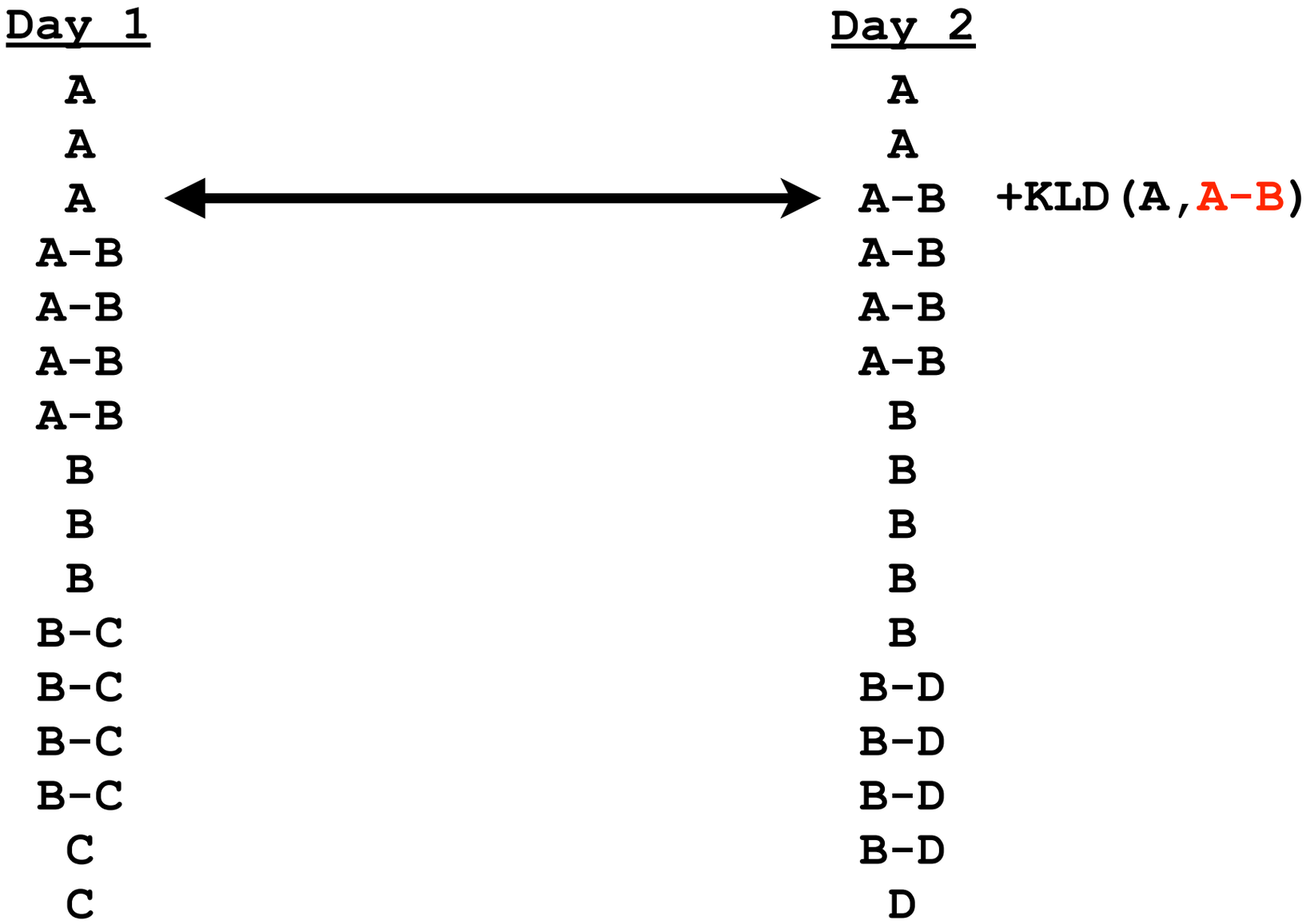
Compute stop type distance

Day 1	Day 2
A	A
A	A
A	A-B
A-B	A-B
A-B	A-B
A-B	A-B
A-B	B
B	B
B	B
B	B
B-C	B
B-C	B-D
B-C	B-D
B-C	B-D
B-C	B-D
C	B-D
C	D

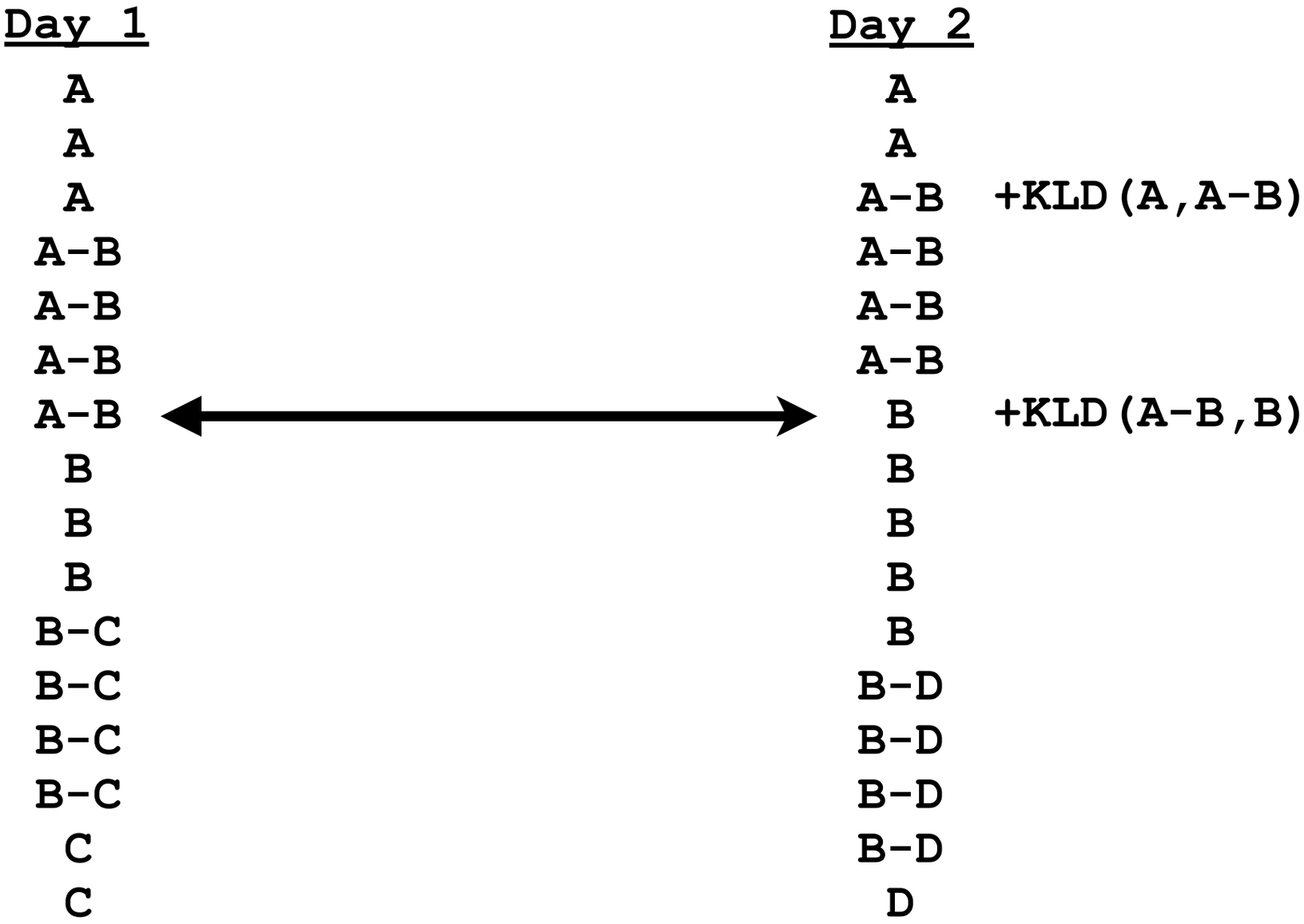
Compute stop type distance



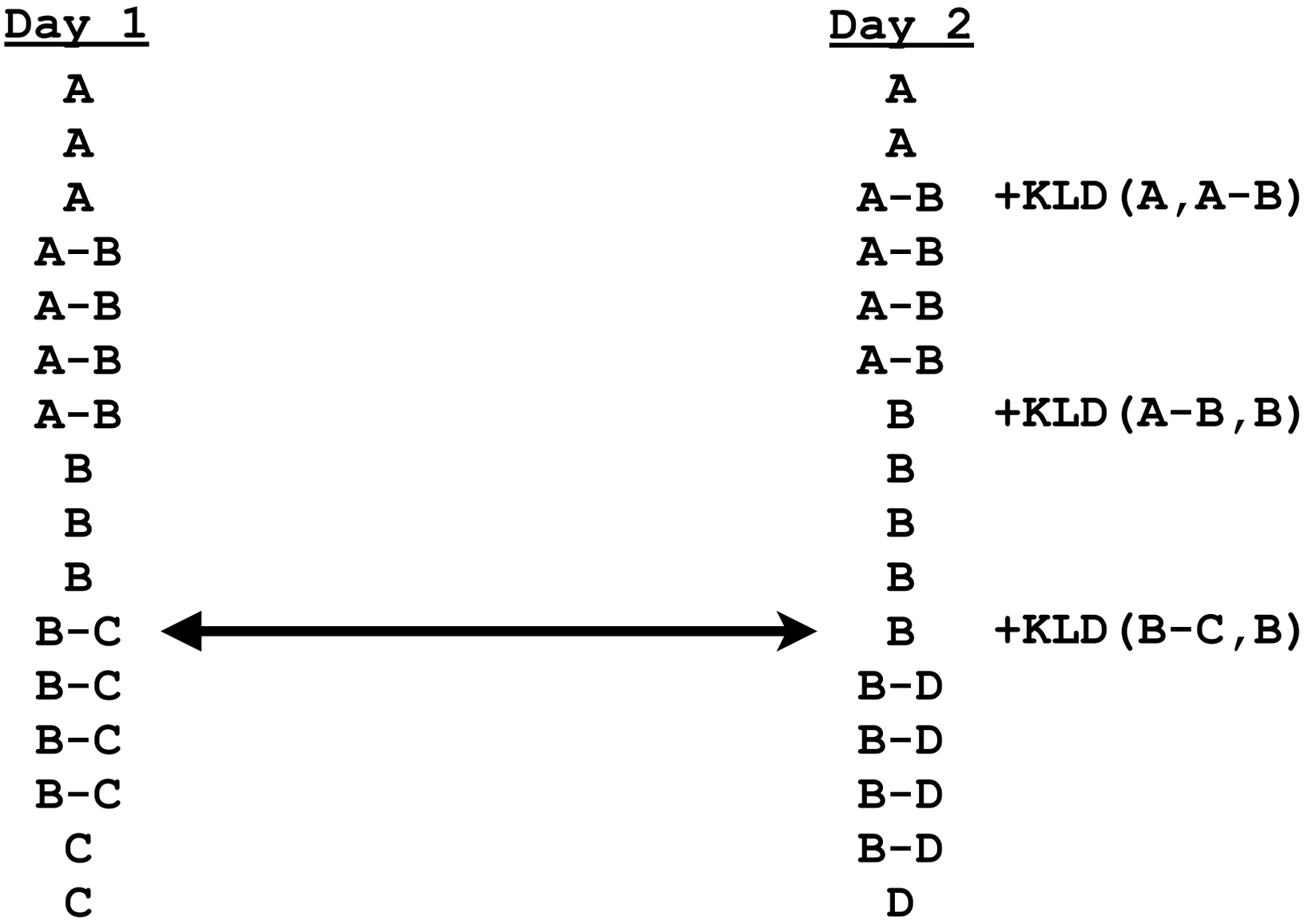
Compute stop type distance



Compute stop type distance



Compute stop type distance



Compute stop type distance

Day 1	Day 2	
A	A	
A	A	
A	A-B	+KLD(A,A-B)
A-B	A-B	
A-B	A-B	
A-B	A-B	
A-B	B	+KLD(A-B,B)
B	B	
B	B	
B	B	
B-C	B	+KLD(B-C,B)
B-C	B-D	+KLD(B-C,B-D)
B-C	B-D	+KLD(B-C,B-D)
B-C	B-D	+KLD(B-C,B-D)
C	B-D	+KLD(C,B-D)
C	D	+KLD(C,D)



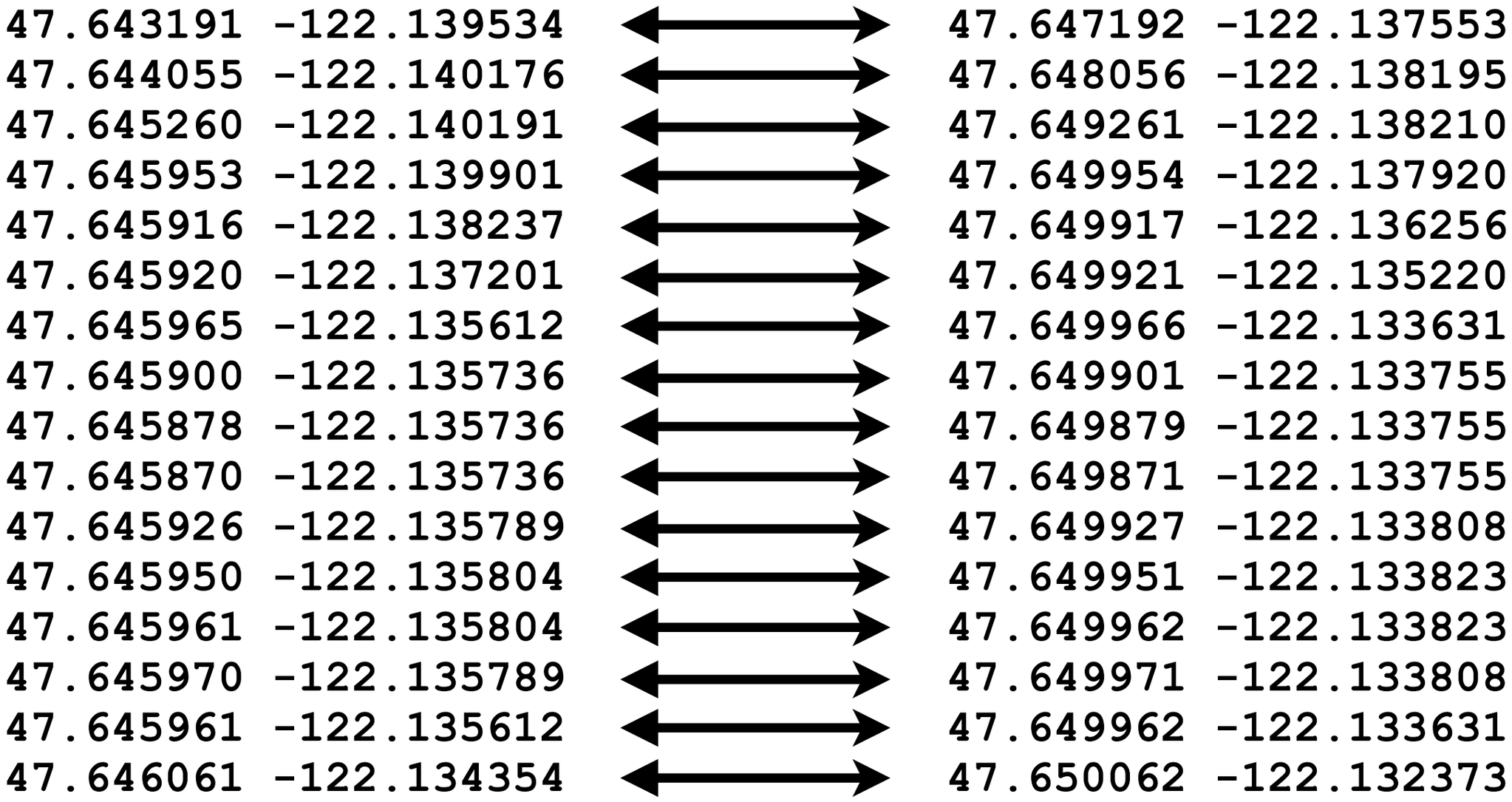
Sum of pairs distance

Sum of pairs distance

```
47.643191 -122.139534
47.644055 -122.140176
47.645260 -122.140191
47.645953 -122.139901
47.645916 -122.138237
47.645920 -122.137201
47.645965 -122.135612
47.645900 -122.135736
47.645878 -122.135736
47.645870 -122.135736
47.645926 -122.135789
47.645950 -122.135804
47.645961 -122.135804
47.645970 -122.135789
47.645961 -122.135612
47.646061 -122.134354
```

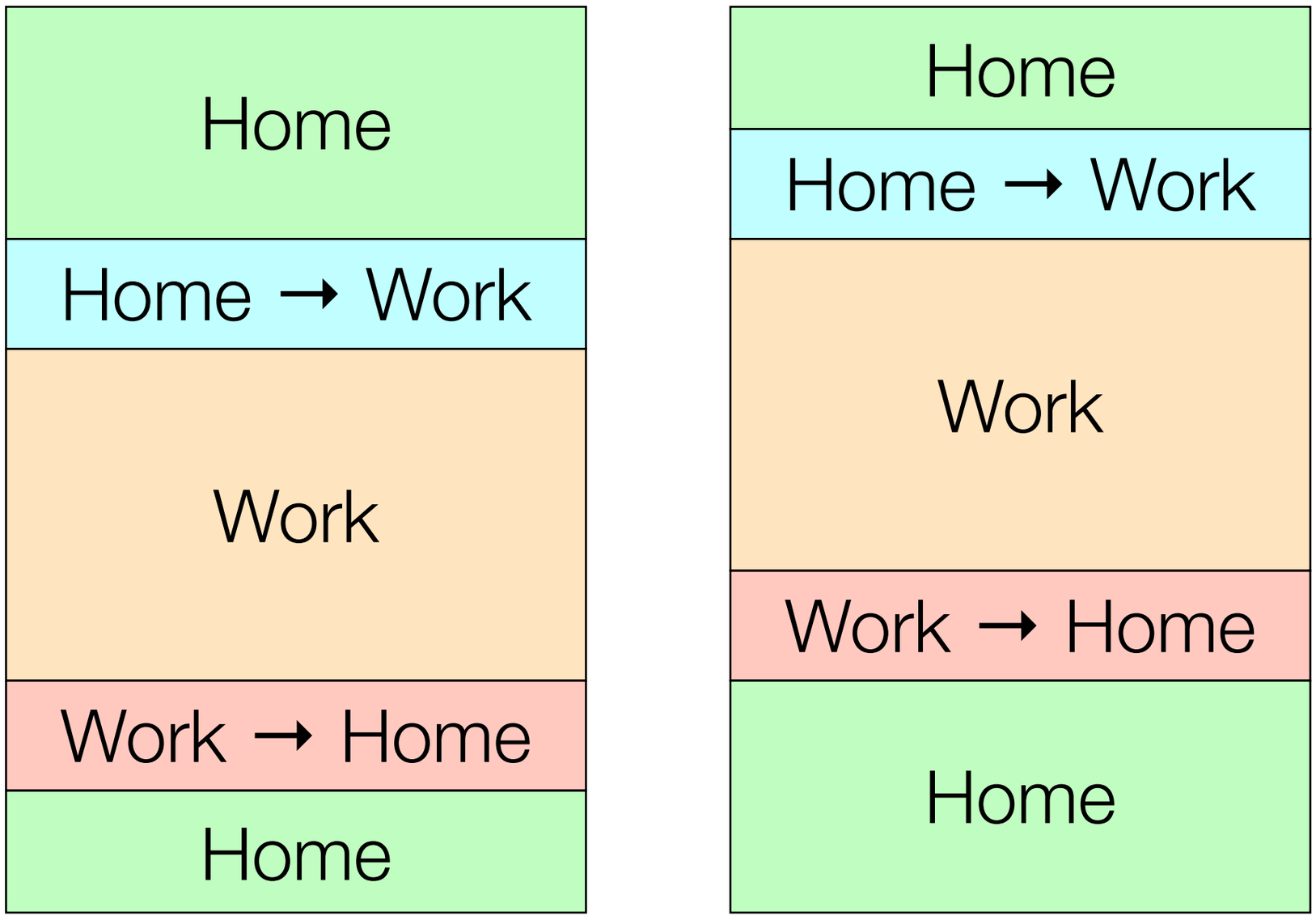
```
47.647192 -122.137553
47.648056 -122.138195
47.649261 -122.138210
47.649954 -122.137920
47.649917 -122.136256
47.649921 -122.135220
47.649966 -122.133631
47.649901 -122.133755
47.649879 -122.133755
47.649871 -122.133755
47.649927 -122.133808
47.649951 -122.133823
47.649962 -122.133823
47.649971 -122.133808
47.649962 -122.133631
47.650062 -122.132373
```

Sum of pairs distance

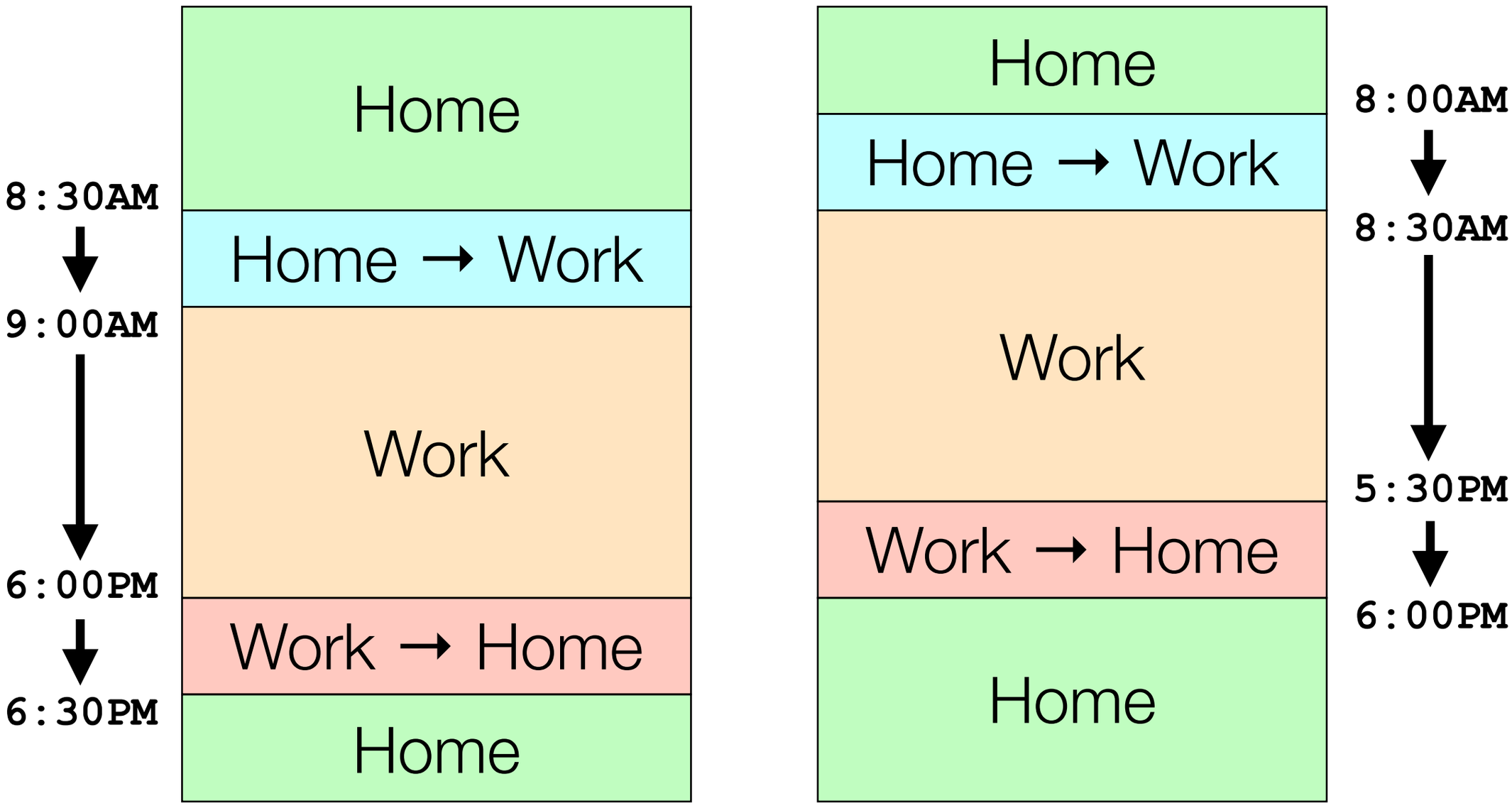


Dynamic time warping

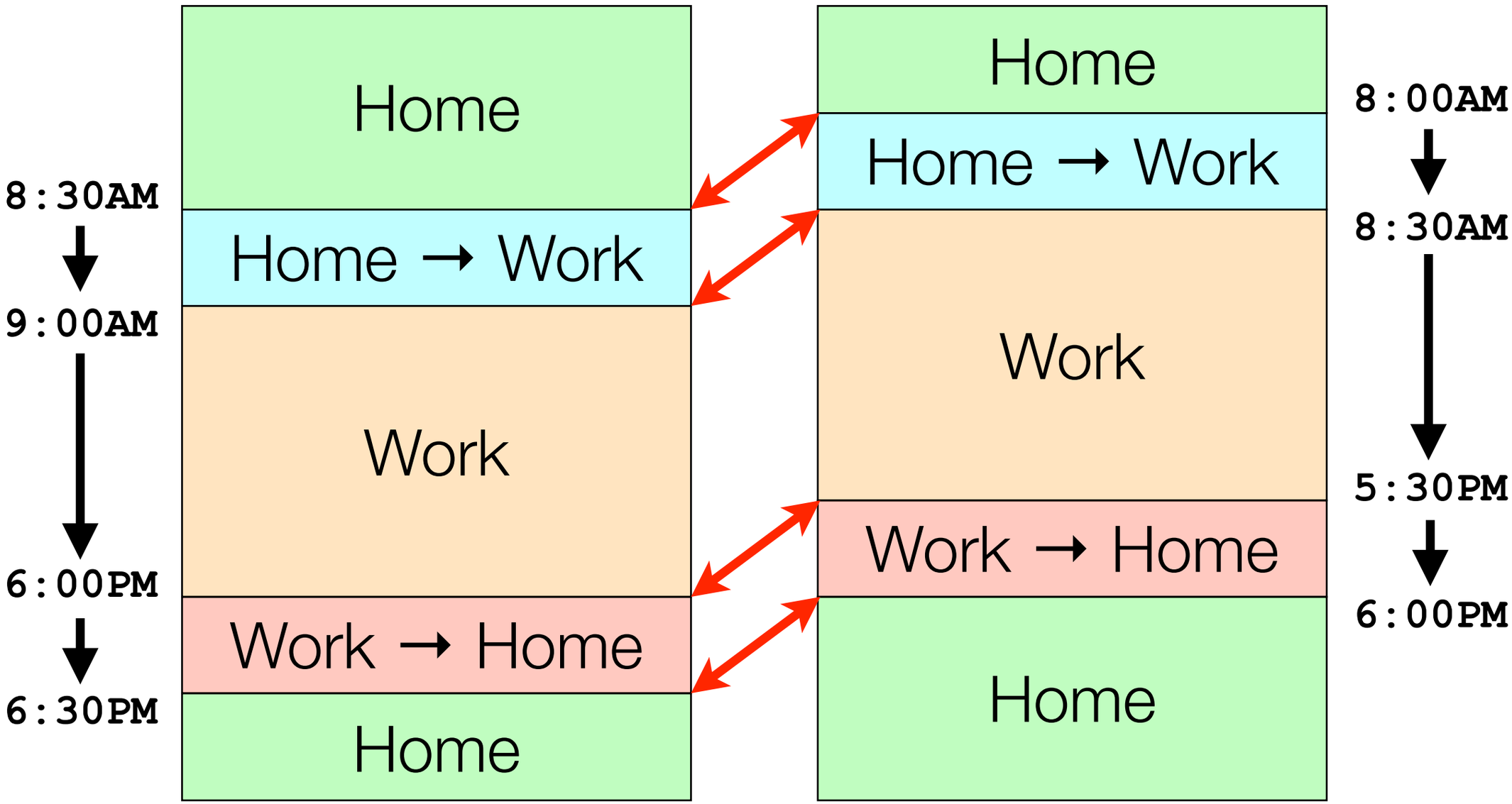
Dynamic time warping



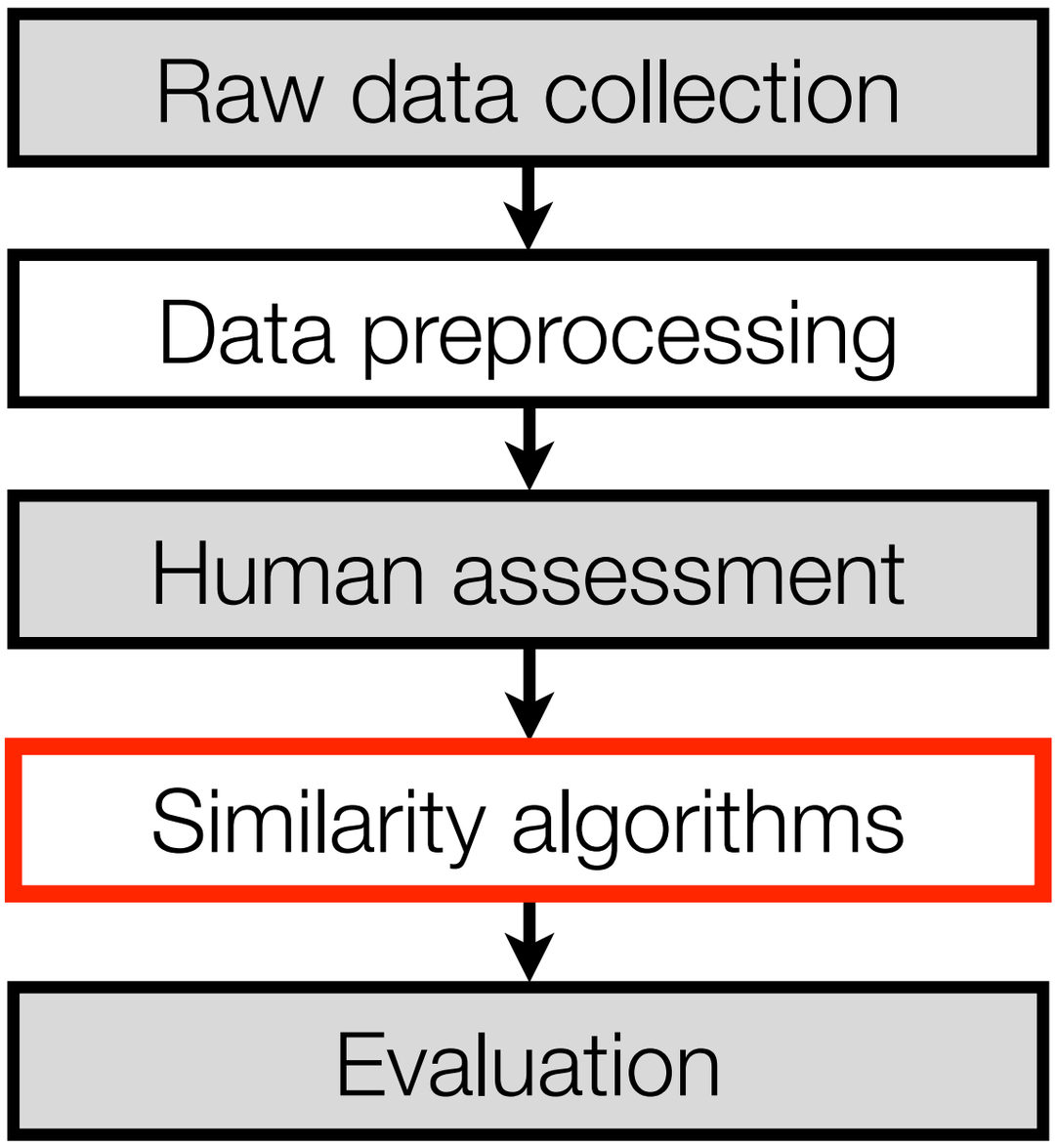
Dynamic time warping



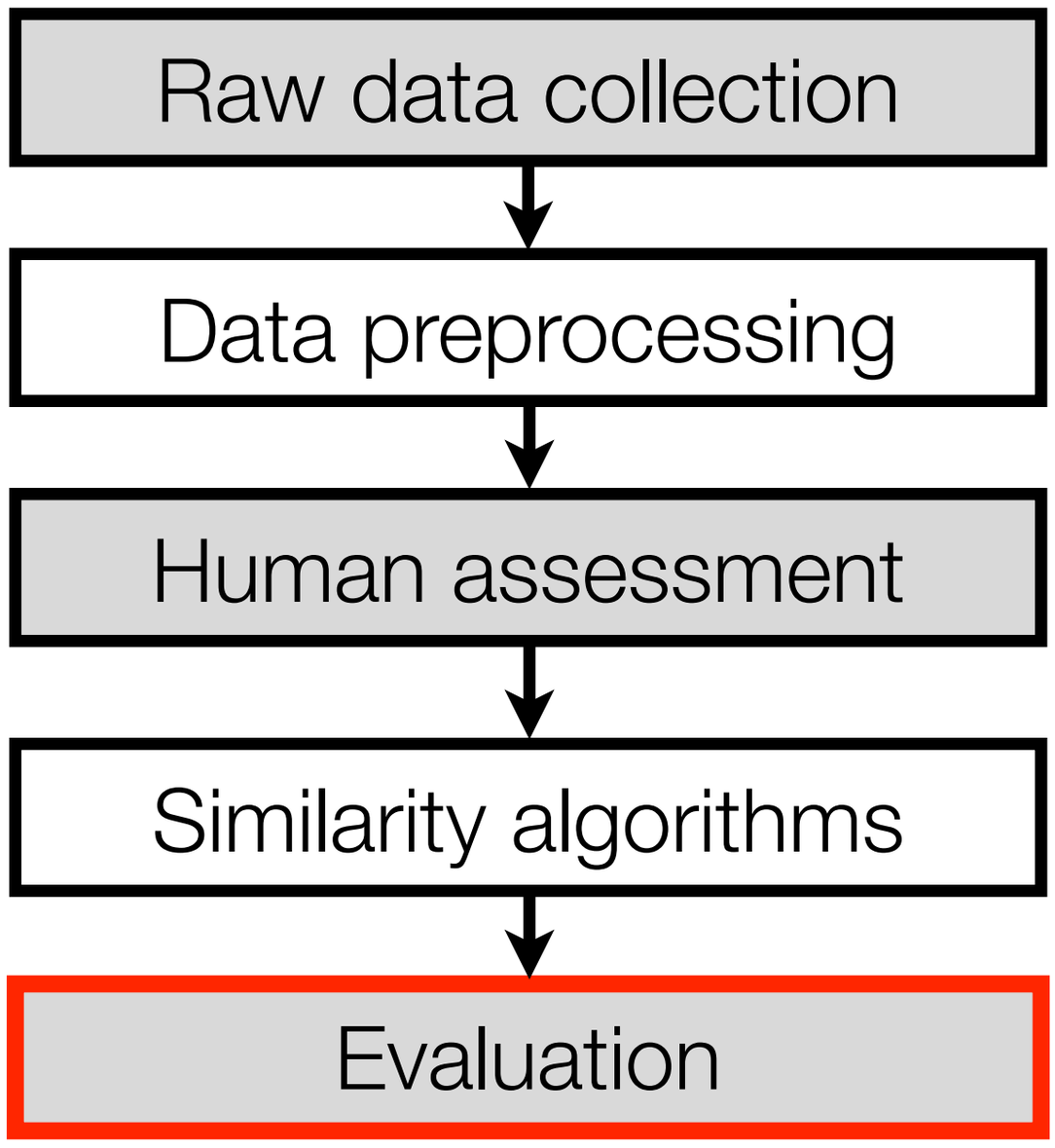
Dynamic time warping



Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map

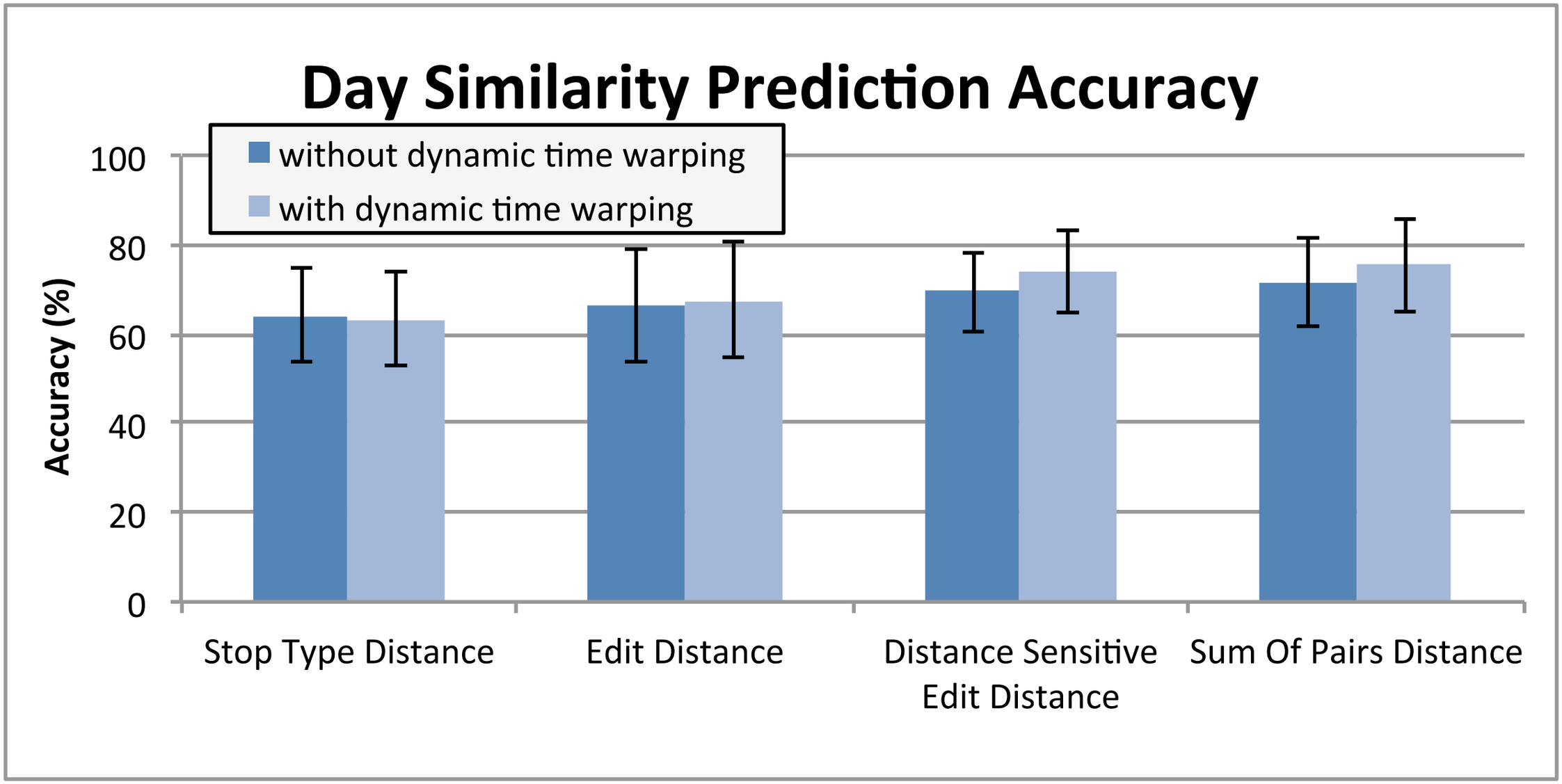

Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Algorithm evaluation

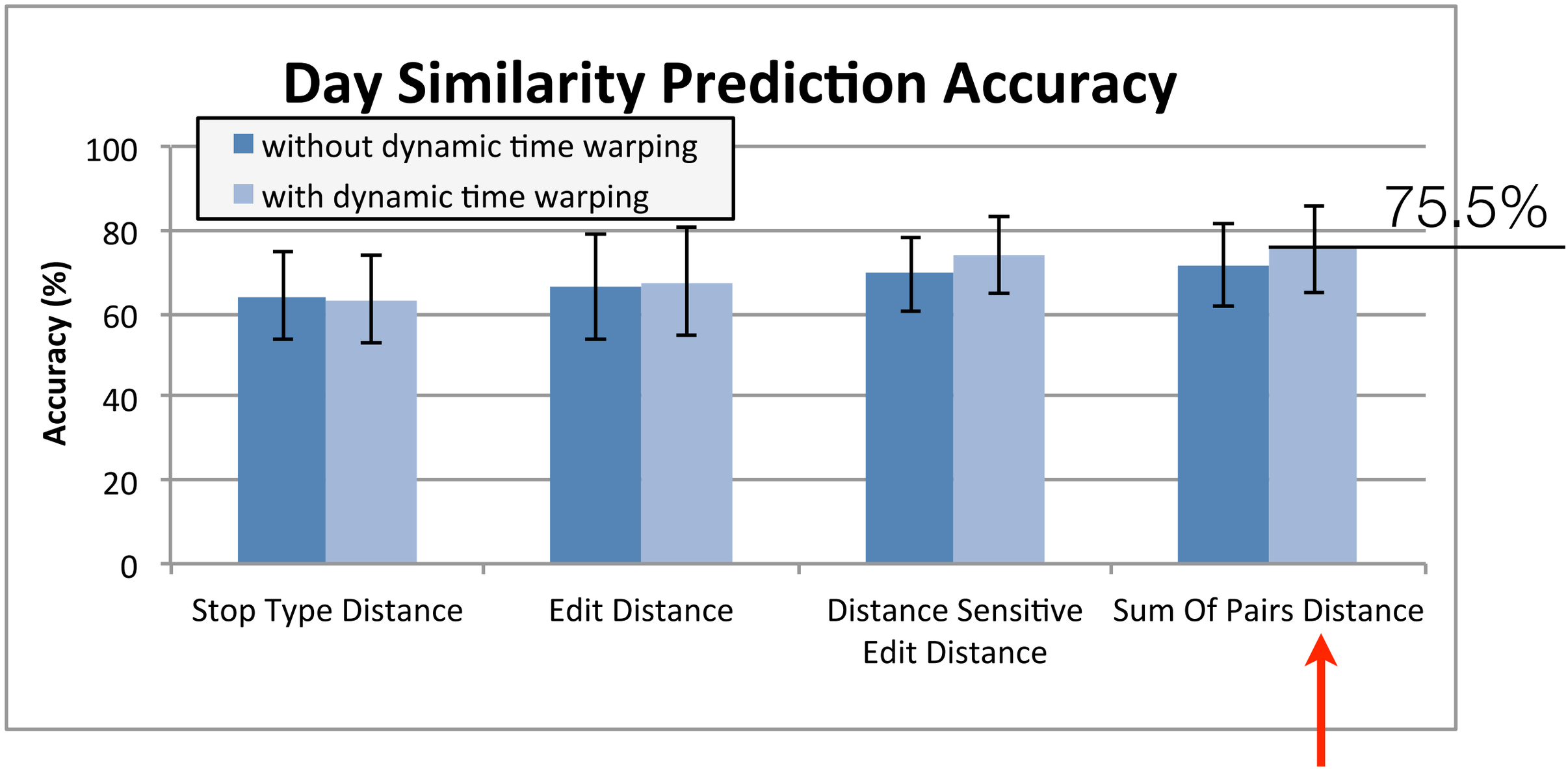
- Provided same dataset

- Assess which pairs were most similar

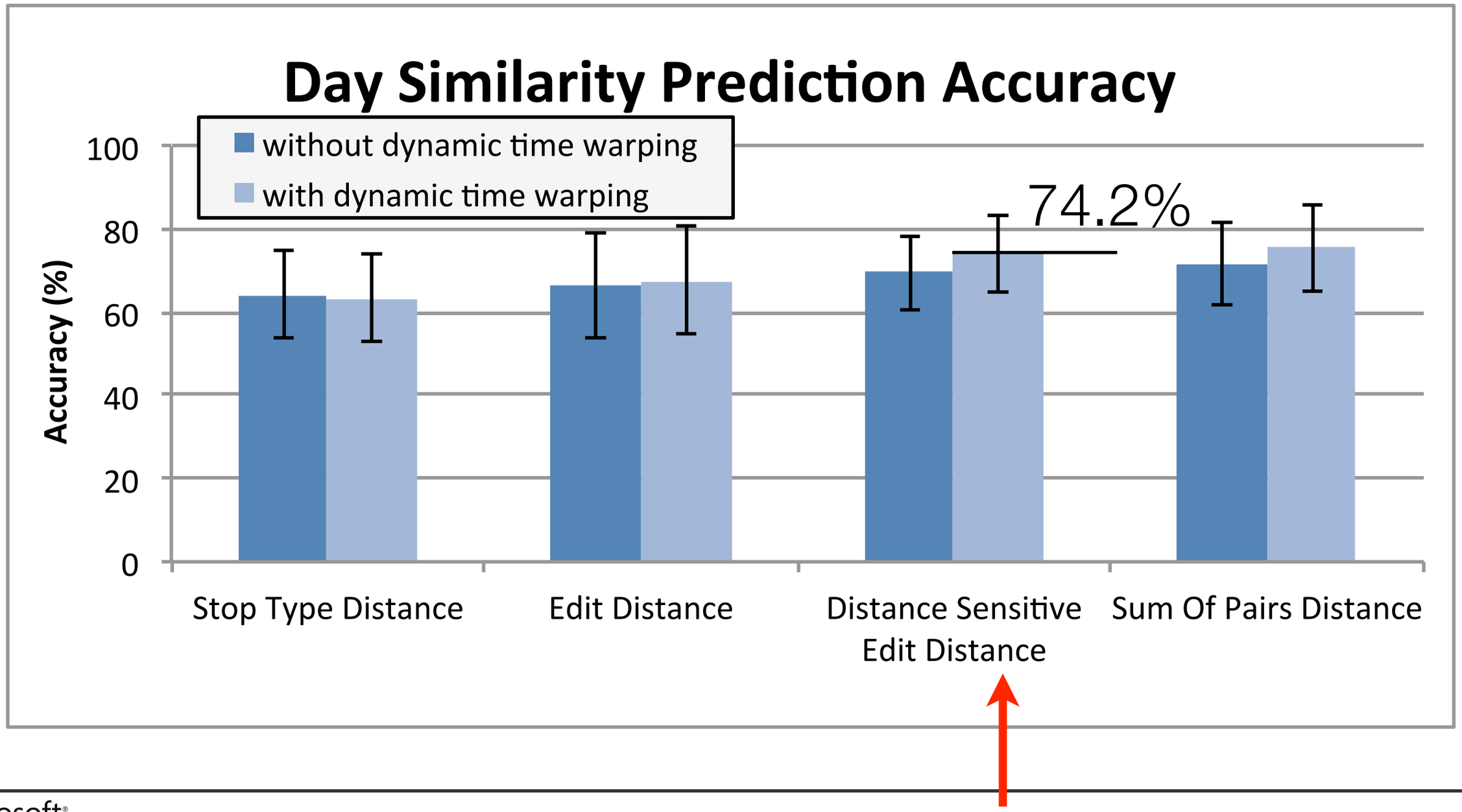
Overall accuracy results



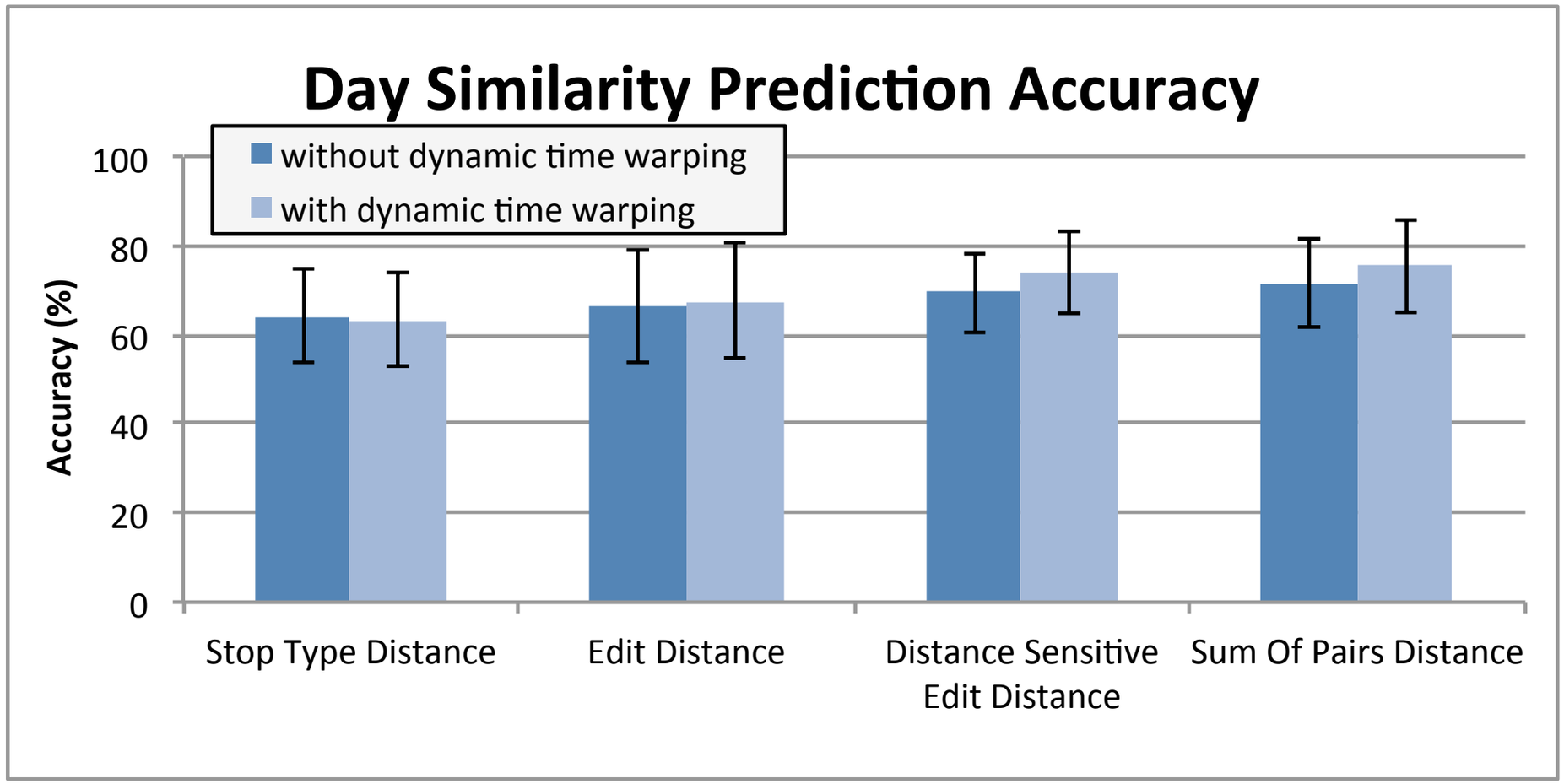
Overall accuracy results



Overall accuracy results



Accuracy results



$$F(7,29) = 11.22,\ p = 5.45 \times 10^{-12}$$

Pairwise *t*-tests

Algorithm	Wins	Losses
Edit Distance w/o DTW	0	3
Edit Distance w/DTW	0	1
Distance Sensitive Edit Distance w/o DTW	2	2
Distance Sensitive Edit Distance w/DTW	5	0
Stops Categories Distance w/o DTW	0	4
Stops Categories Distance w/DTW	0	4
Sum of Pairs Distance w/o DTW	3	0
Sum of Pairs Distance w/DTW	4	0

Pairwise *t*-tests

Algorithm	Wins	Losses
Edit Distance w/o DTW	0	3
Edit Distance w/DTW	0	1
Distance Sensitive Edit Distance w/o DTW	2	2
Distance Sensitive Edit Distance w/DTW	5	0
Stops Categories Distance w/o DTW	0	4
Stops Categories Distance w/DTW	0	4
Sum of Pairs Distance w/o DTW	3	0
Sum of Pairs Distance w/DTW	4	0

Pairwise *t*-tests

Algorithm	Wins	Losses
Edit Distance w/o DTW	0	3
Edit Distance w/DTW	0	1
Distance Sensitive Edit Distance w/o DTW	2	2
Distance Sensitive Edit Distance w/DTW	5	0
Stops Categories Distance w/o DTW	0	4
Stops Categories Distance w/DTW	0	4
Sum of Pairs Distance w/o DTW	3	0
Sum of Pairs Distance w/DTW	4	0

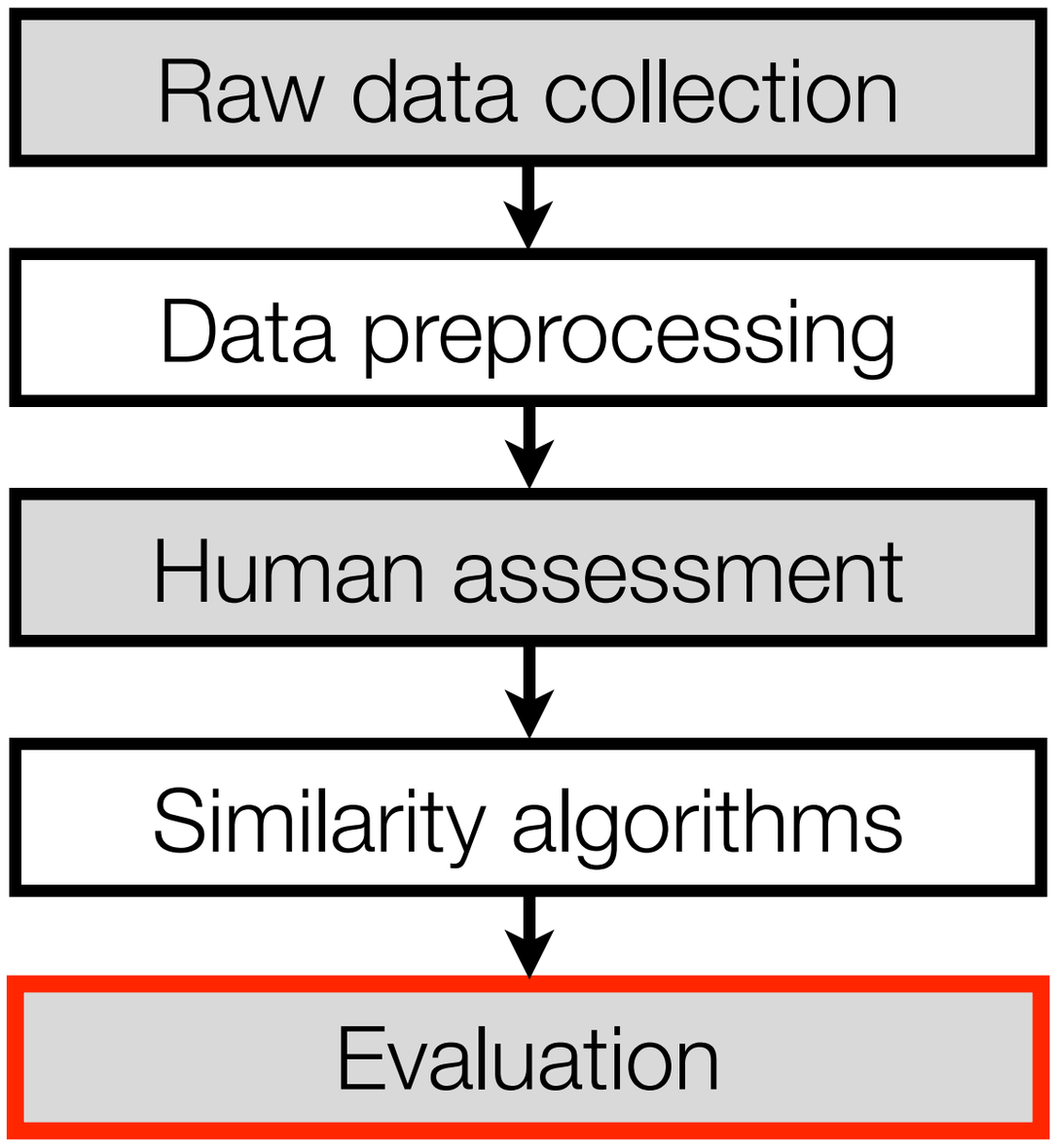
Best performing algorithms

- No significant performance difference
- Sum of pairs distance w/DTW
- Distance sensitive edit distance w/DTW

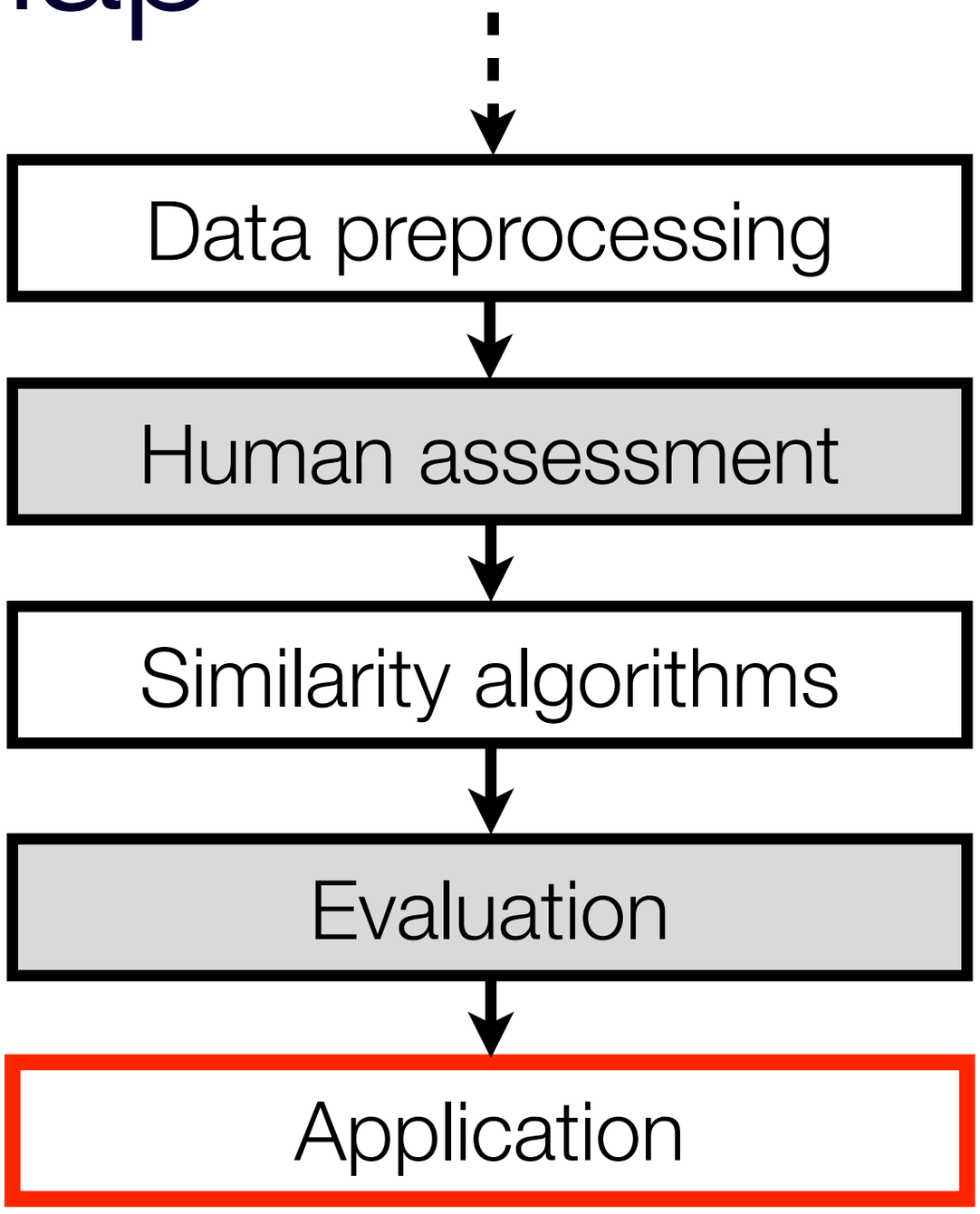
Best performing algorithms

- No significant performance difference
- **Sum of pairs distance w/DTW**
- Distance sensitive edit distance w/DTW

Road map


Raw data collection
Data preprocessing
Human assessment
Similarity algorithms
Evaluation

Road map


Data preprocessing
Human assessment
Similarity algorithms
Evaluation
Application

Application

- Clustering
- Uses similarity measure
  - Find groups of similar days
- Tested as part of human assessment
  - Used standard edit distance

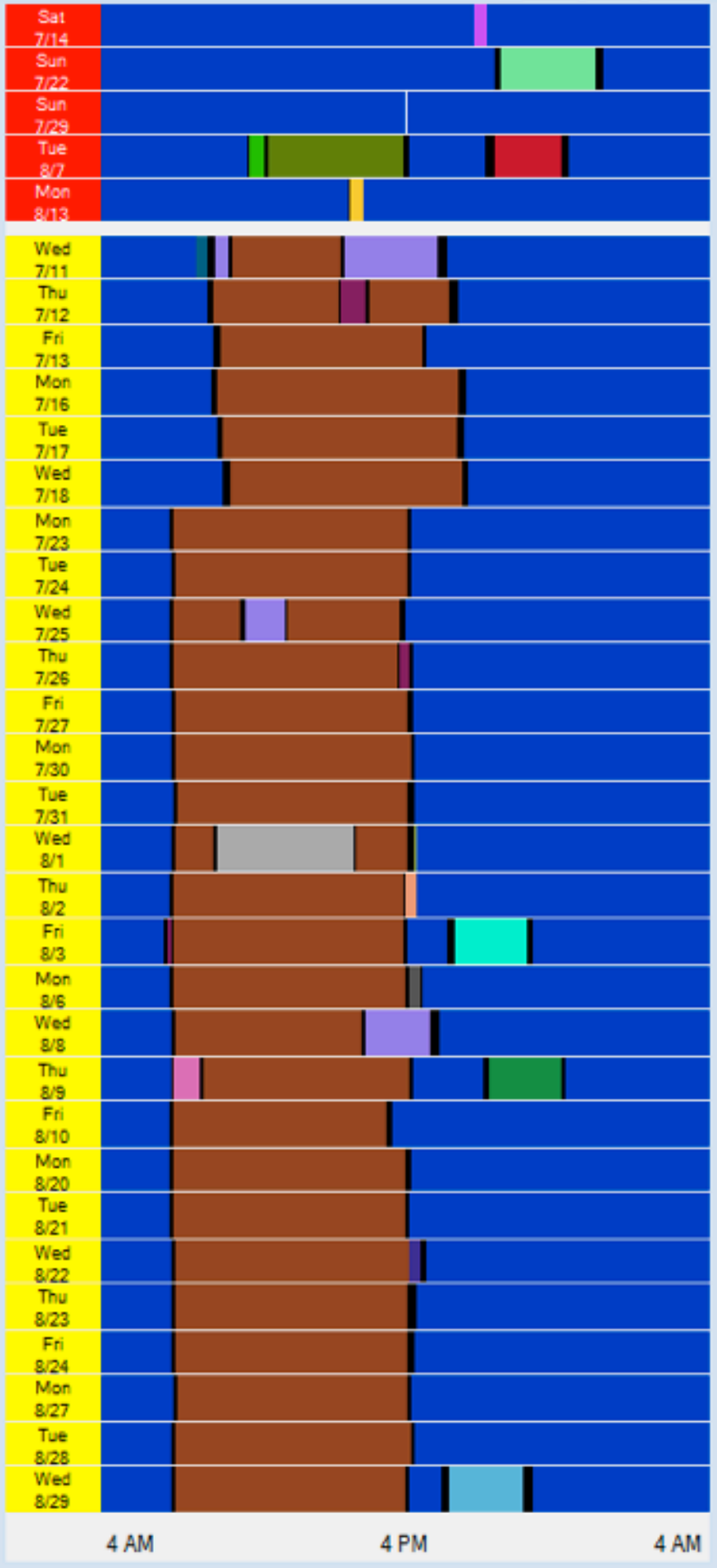
Days

Browse | Load Trace

Start Survey | - | k= 2 | + | Show Patterns | Find Patterns

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11	12	13	14
				Wed 7/11	Thu 7/12	Fri 7/13	Sat 7/14
15 - 21 Jul	15	16	17	18	19	20	21
		Mon 7/16	Tue 7/17	Wed 7/18			
22 - 28 Jul	22	23	24	25	26	27	28
	Sun 7/22	Mon 7/23	Tue 7/24	Wed 7/25	Thu 7/26	Fri 7/27	
29/7 - 4/8	29	30	31	01 Aug	2	3	4
	Sun 7/29	Mon 7/30	Tue 7/31	Wed 8/1	Thu 8/2	Fri 8/3	
5 - 11 Aug	5	6	7	8	9	10	11
		Mon 8/6	Tue 8/7	Wed 8/8	Thu 8/9	Fri 8/10	
12 - 18 Aug	12	13	14	15	16	17	18
		Mon 8/13					
19 - 25 Aug	19	20	21	22	23	24	25
		Mon 8/20	Tue 8/21	Wed 8/22	Thu 8/23	Fri 8/24	
26/8 - 1/9	26	27	28	29	30	31	01 Sep
		Mon 8/27	Tue 8/28	Wed 8/29			

Days Timeline



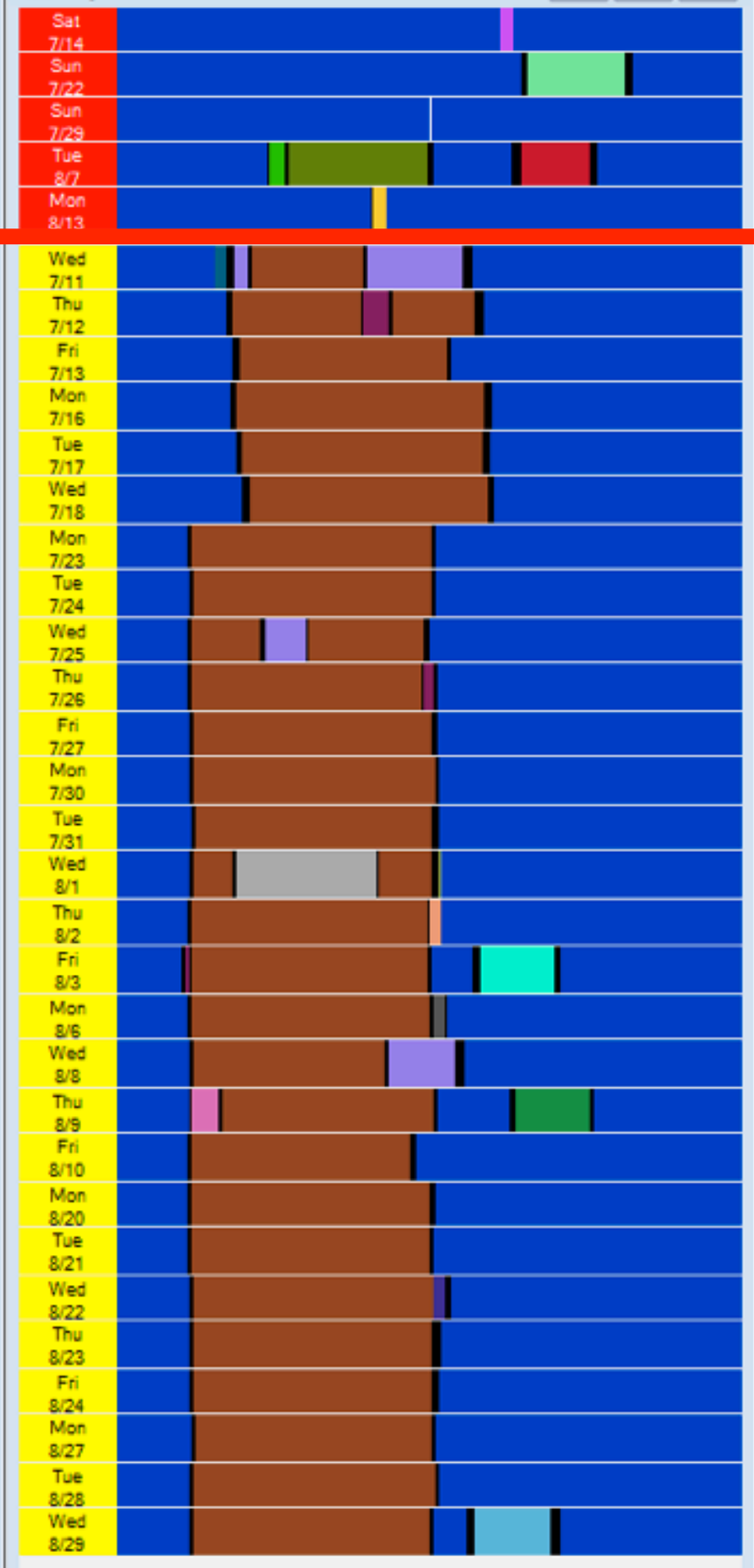
Days

Browse | Load Trace

Start Survey | - | k= 2 | + | Show Patterns | Find Patterns

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11 Wed 7/11	12 Thu 7/12	13 Fri 7/13	14 Sat 7/14
15 - 21 Jul	15	16 Mon 7/16	17 Tue 7/17	18 Wed 7/18	19	20	21
22 - 28 Jul	22 Sun 7/22	23 Mon 7/23	24 Tue 7/24	25 Wed 7/25	26 Thu 7/26	27 Fri 7/27	28
29/7 - 4/8	29 Sun 7/29	30 Mon 7/30	31 Tue 7/31	01 Aug Wed 8/1	2 Thu 8/2	3 Fri 8/3	4
5 - 11 Aug	5	6 Mon 8/6	7 Tue 8/7	8 Wed 8/8	9 Thu 8/9	10 Fri 8/10	11
12 - 18 Aug	12	13 Mon 8/13	14	15	16	17	18
19 - 25 Aug	19	20 Mon 8/20	21 Tue 8/21	22 Wed 8/22	23 Thu 8/23	24 Fri 8/24	25
26/8 - 1/9	26	27 Mon 8/27	28 Tue 8/28	29 Wed 8/29	30	31	01 Sep

Days Timeline



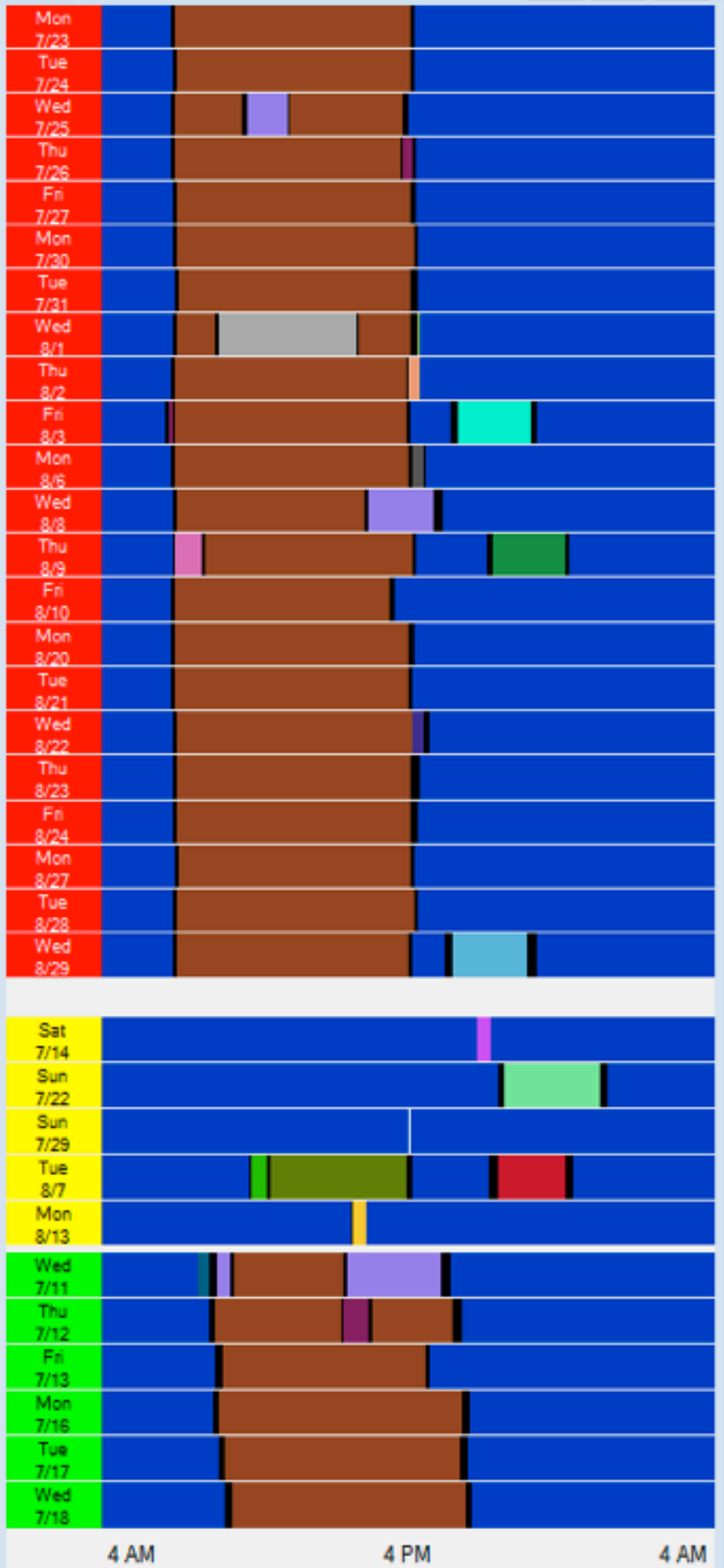
Days

Browse | Load Trace

Start Survey | - k= 3 + | Show Patterns | Find Patterns

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11	12	13	14
				Wed 7/11	Thu 7/12	Fri 7/13	Sat 7/14
15 - 21 Jul	15	16	17	18	19	20	21
		Mon 7/16	Tue 7/17	Wed 7/18			
22 - 28 Jul	22	23	24	25	26	27	28
	Sun 7/22	Mon 7/23	Tue 7/24	Wed 7/25	Thu 7/26	Fri 7/27	
29/7 - 4/8	29	30	31	01 Aug	2	3	4
	Sun 7/29	Mon 7/30	Tue 7/31	Wed 8/1	Thu 8/2	Fri 8/3	
5 - 11 Aug	5	6	7	8	9	10	11
		Mon 8/6	Tue 8/7	Wed 8/8	Thu 8/9	Fri 8/10	
12 - 18 Aug	12	13	14	15	16	17	18
		Mon 8/13					
19 - 25 Aug	19	20	21	22	23	24	25
		Mon 8/20	Tue 8/21	Wed 8/22	Thu 8/23	Fri 8/24	
26/8 - 1/9	26	27	28	29	30	31	01 Sep
		Mon 8/27	Tue 8/28	Wed 8/29			

Days Timeline



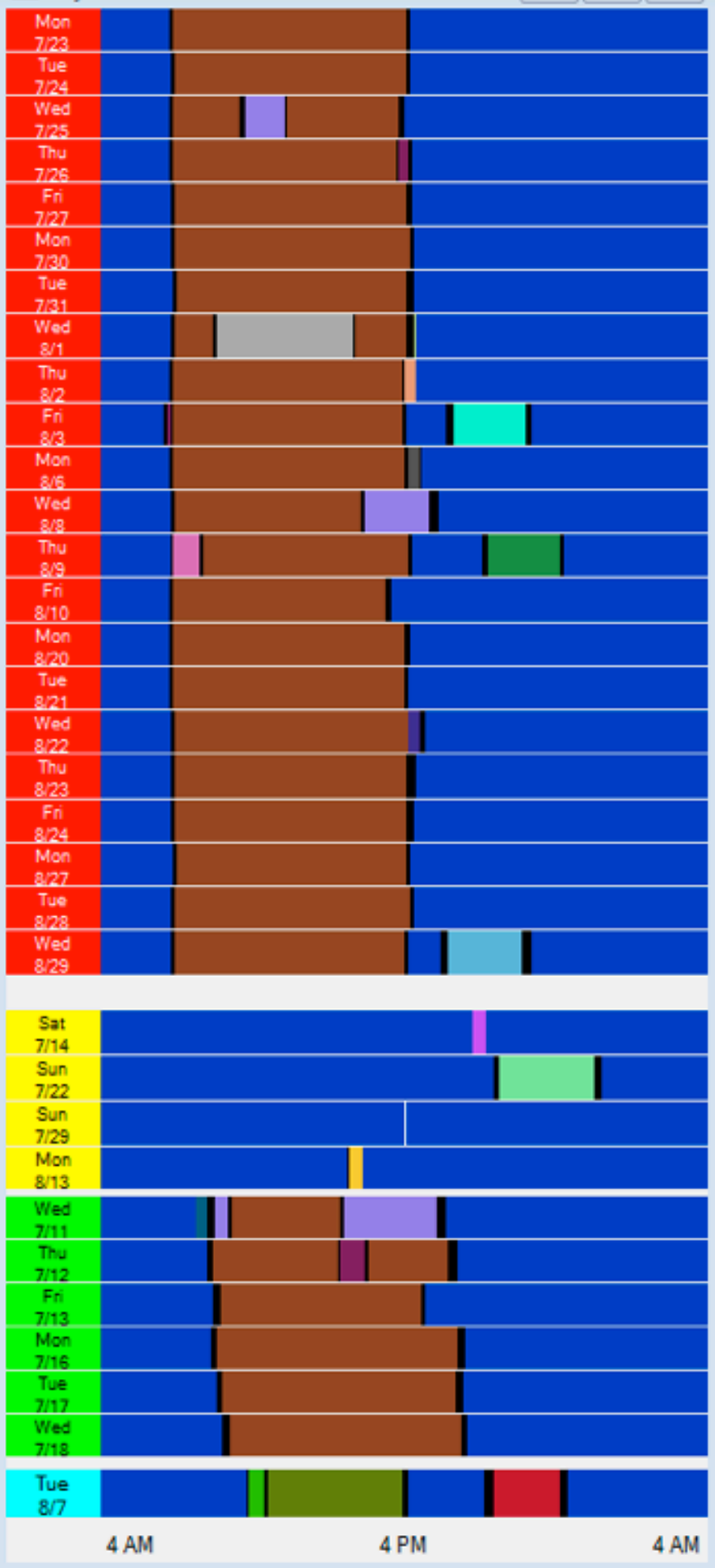
Days

Browse | Load Trace

Start Survey | - k= 4 + | Show Patterns | Find Patterns

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11	12	13	14
				Wed 7/11	Thu 7/12	Fri 7/13	Sat 7/14
15 - 21 Jul	15	16	17	18	19	20	21
		Mon 7/16	Tue 7/17	Wed 7/18			
22 - 28 Jul	22	23	24	25	26	27	28
	Sun 7/22	Mon 7/23	Tue 7/24	Wed 7/25	Thu 7/26	Fri 7/27	
29/7 - 4/8	29	30	31	01 Aug	2	3	4
	Sun 7/29	Mon 7/30	Tue 7/31	Wed 8/1	Thu 8/2	Fri 8/3	
5 - 11 Aug	5	6	7	8	9	10	11
		Mon 8/6	Tue 8/7	Wed 8/8	Thu 8/9	Fri 8/10	
12 - 18 Aug	12	13	14	15	16	17	18
		Mon 8/13					
19 - 25 Aug	19	20	21	22	23	24	25
		Mon 8/20	Tue 8/21	Wed 8/22	Thu 8/23	Fri 8/24	
26/8 - 1/9	26	27	28	29	30	31	01 Sep
		Mon 8/27	Tue 8/28	Wed 8/29			

Days Timeline

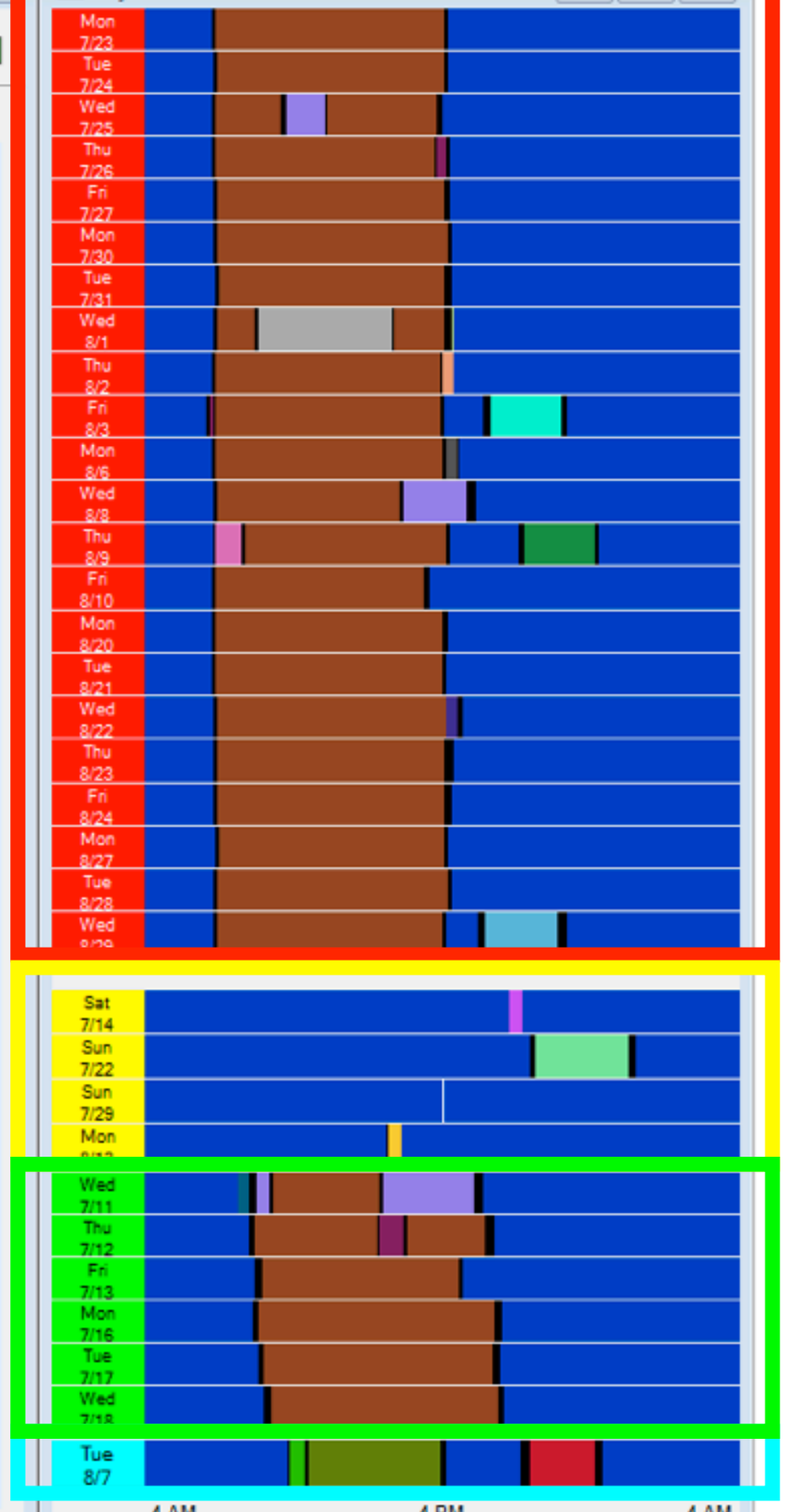


Days

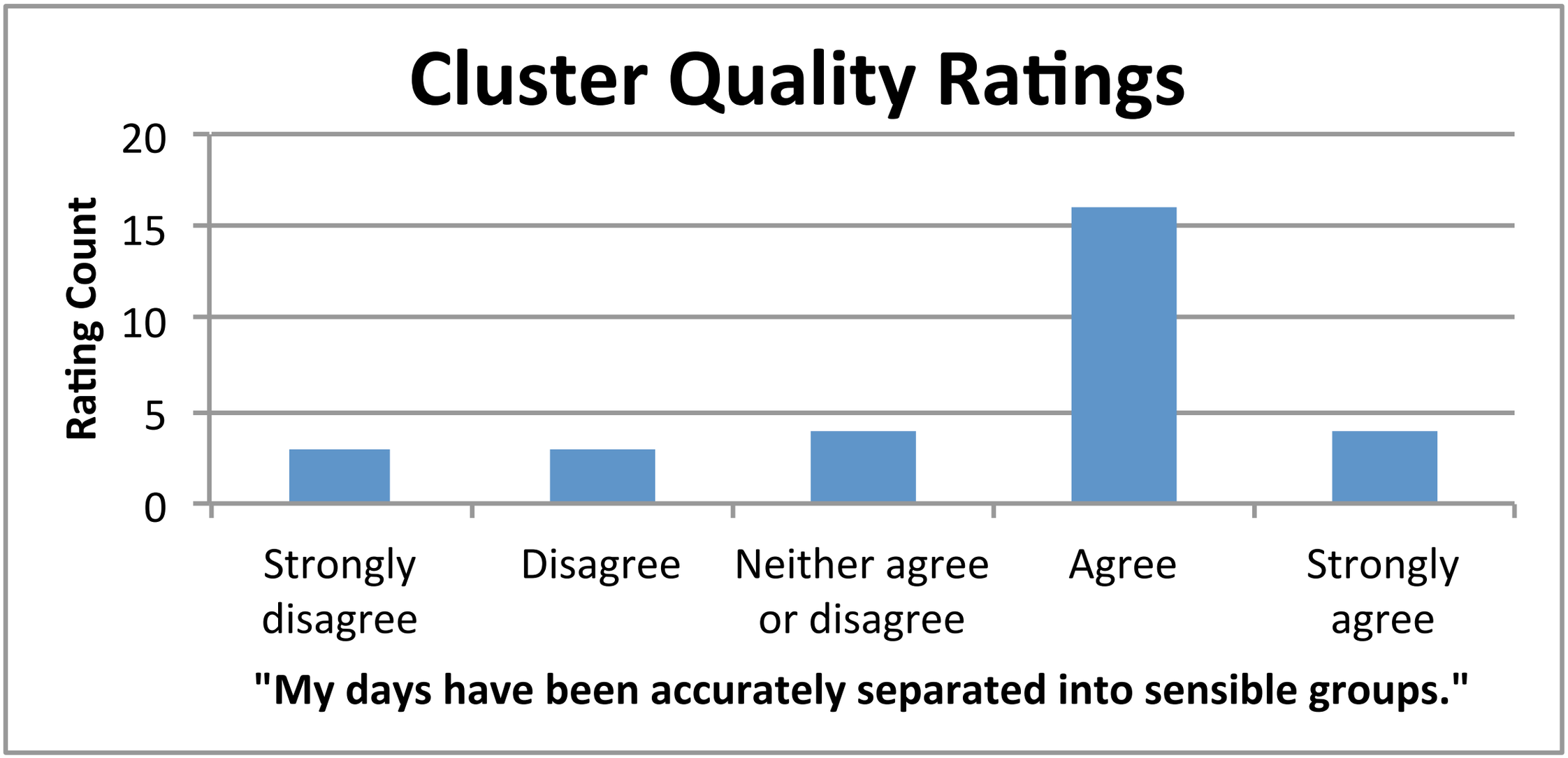
Browse | Load Trace

Start Survey | - k= 4 + | Show Patterns | Find Patterns

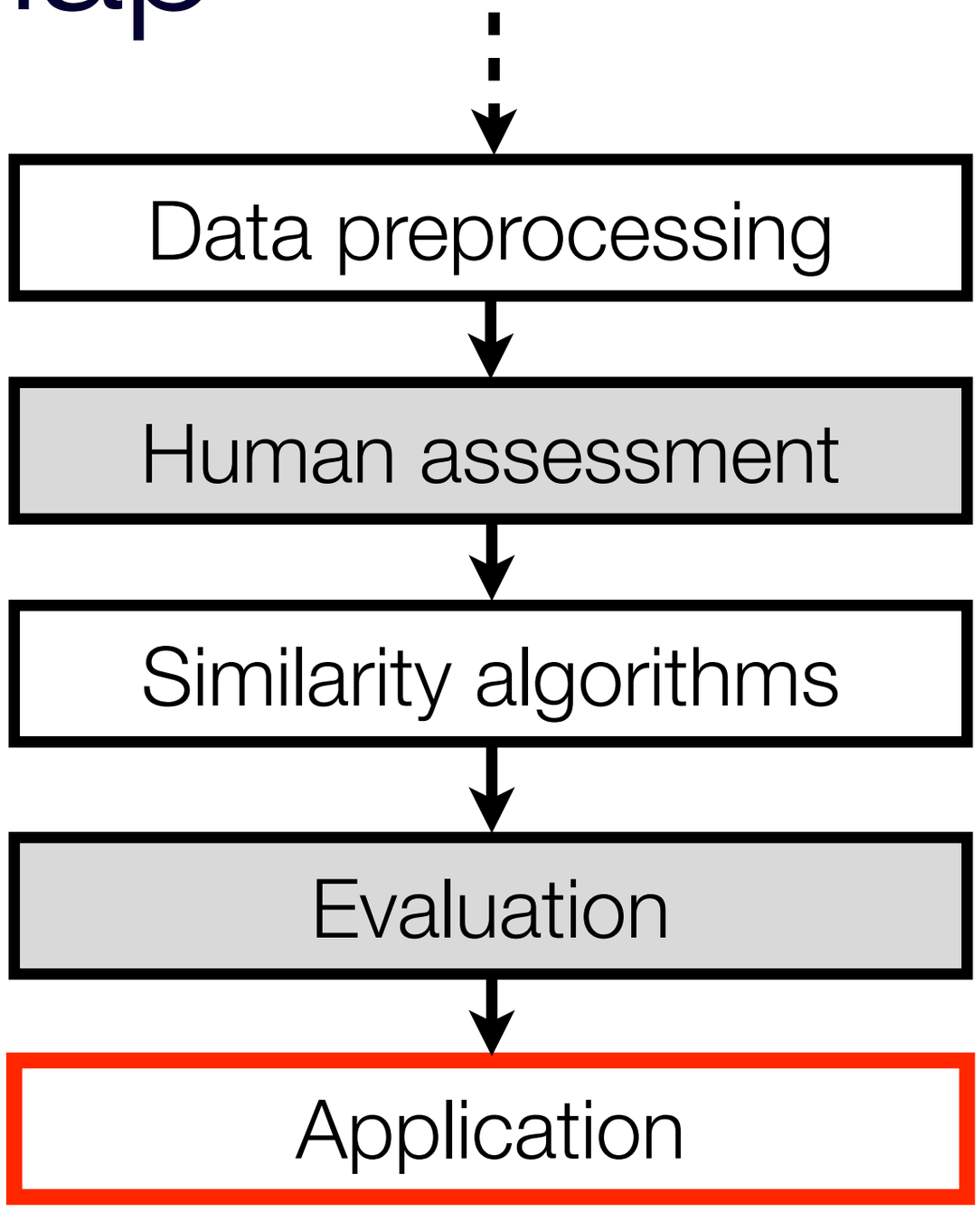
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8 - 14 Jul	08 Jul	9	10	11	12	13	14
				Wed 7/11	Thu 7/12	Fri 7/13	Sat 7/14
15 - 21 Jul	15	16	17	18	19	20	21
		Mon 7/16	Tue 7/17	Wed 7/18			
22 - 28 Jul	22	23	24	25	26	27	28
	Sun 7/22	Mon 7/23	Tue 7/24	Wed 7/25	Thu 7/26	Fri 7/27	
29/7 - 4/8	29	30	31	01 Aug	2	3	4
	Sun 7/29	Mon 7/30	Tue 7/31	Wed 8/1	Thu 8/2	Fri 8/3	
5 - 11 Aug	5	6	7	8	9	10	11
		Mon 8/6	Tue 8/7	Wed 8/8	Thu 8/9	Fri 8/10	
12 - 18 Aug	12	13	14	15	16	17	18
		Mon 8/13					
19 - 25 Aug	19	20	21	22	23	24	25
		Mon 8/20	Tue 8/21	Wed 8/22	Thu 8/23	Fri 8/24	
26/8 - 1/9	26	27	28	29	30	31	01 Sep
		Mon 8/27	Tue 8/28	Wed 8/29			



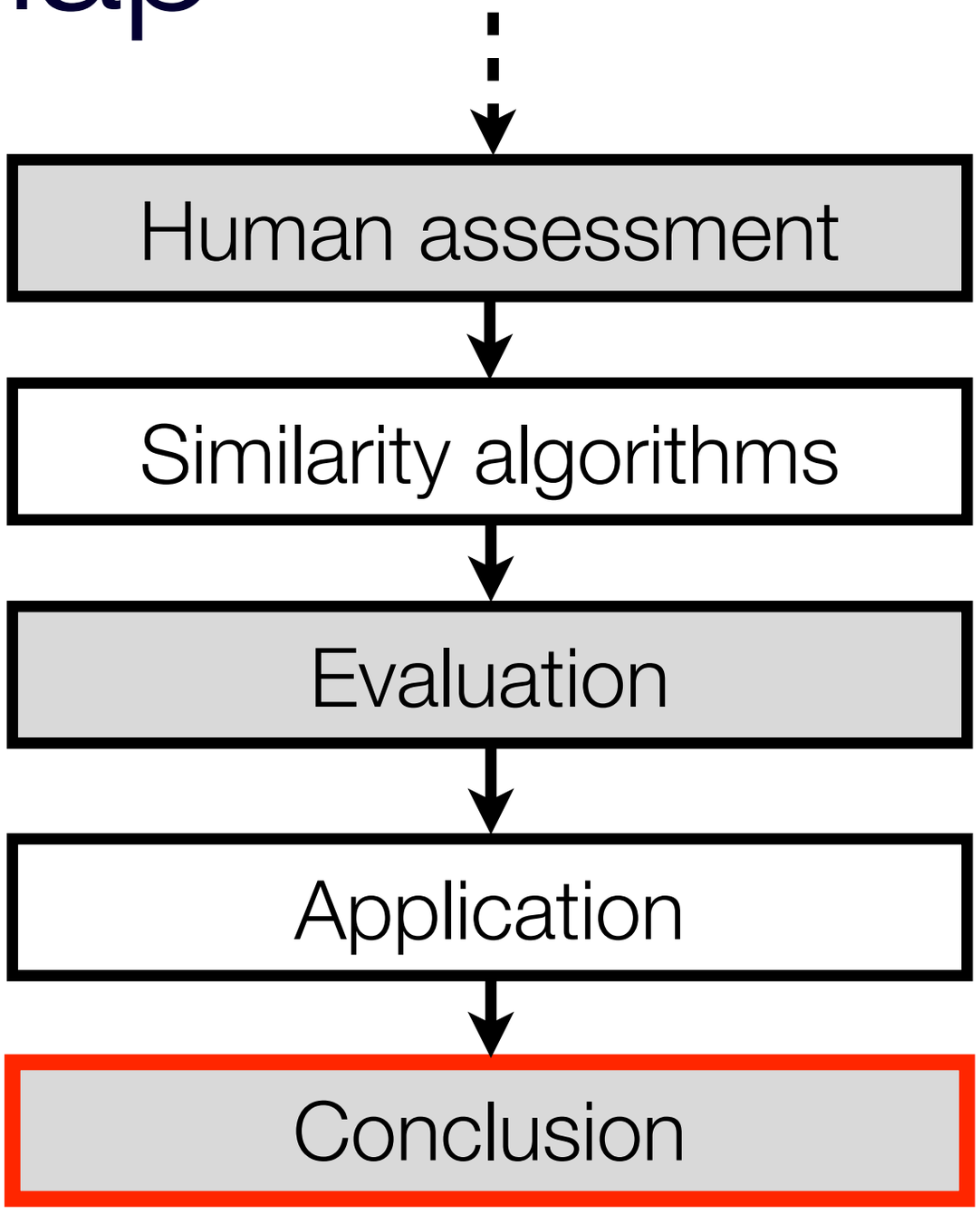
Clustering evaluation



Road map


Data preprocessing
Human assessment
Similarity algorithms
Evaluation
Application

Road map


Human assessment
Similarity algorithms
Evaluation
Application
Conclusion

Conclusions

- Location traces for encoding days
- 30-subject user study
- Assessed 8 similarity algorithms
- Clustering application

Future work

- Better stop-type classification
  - Limited to business-type distributions
- Other similarity algorithms
- Anomaly detection experiments
  - Expect this would work well

Thanks!

Questions?

Microsoft®
Research