Exploring Application Usage Patterns of Smart Phones for Discovering Personal Semantic Regions

Presenter: Chun-Ta Lu

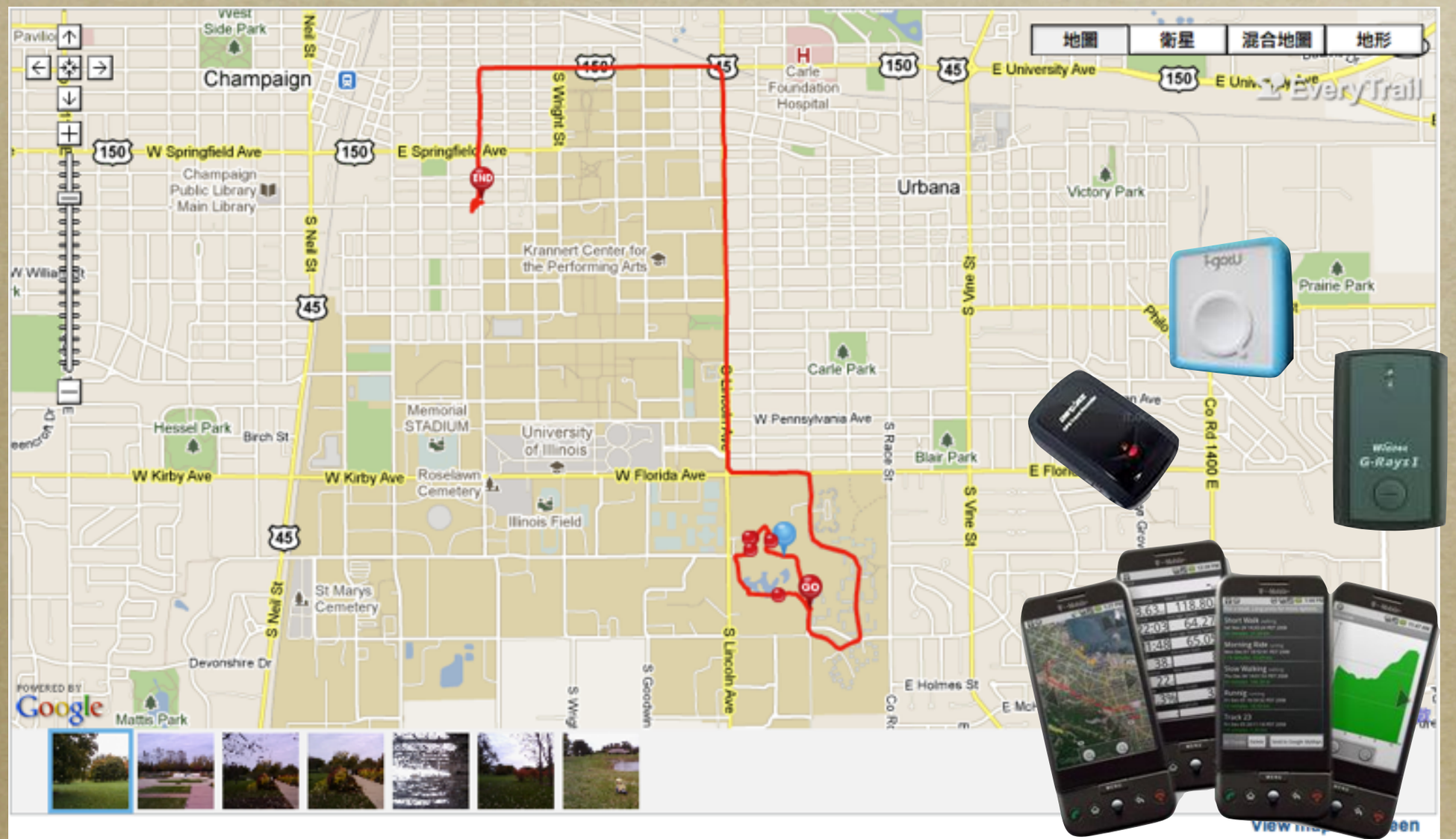
Outline

- *Motivation*
- *Goal*
- *Related Works*
- *Framework*
  - *Region Extraction*
  - *Personal Semantic Extraction*
- *Experiment*
- *Conclusion*

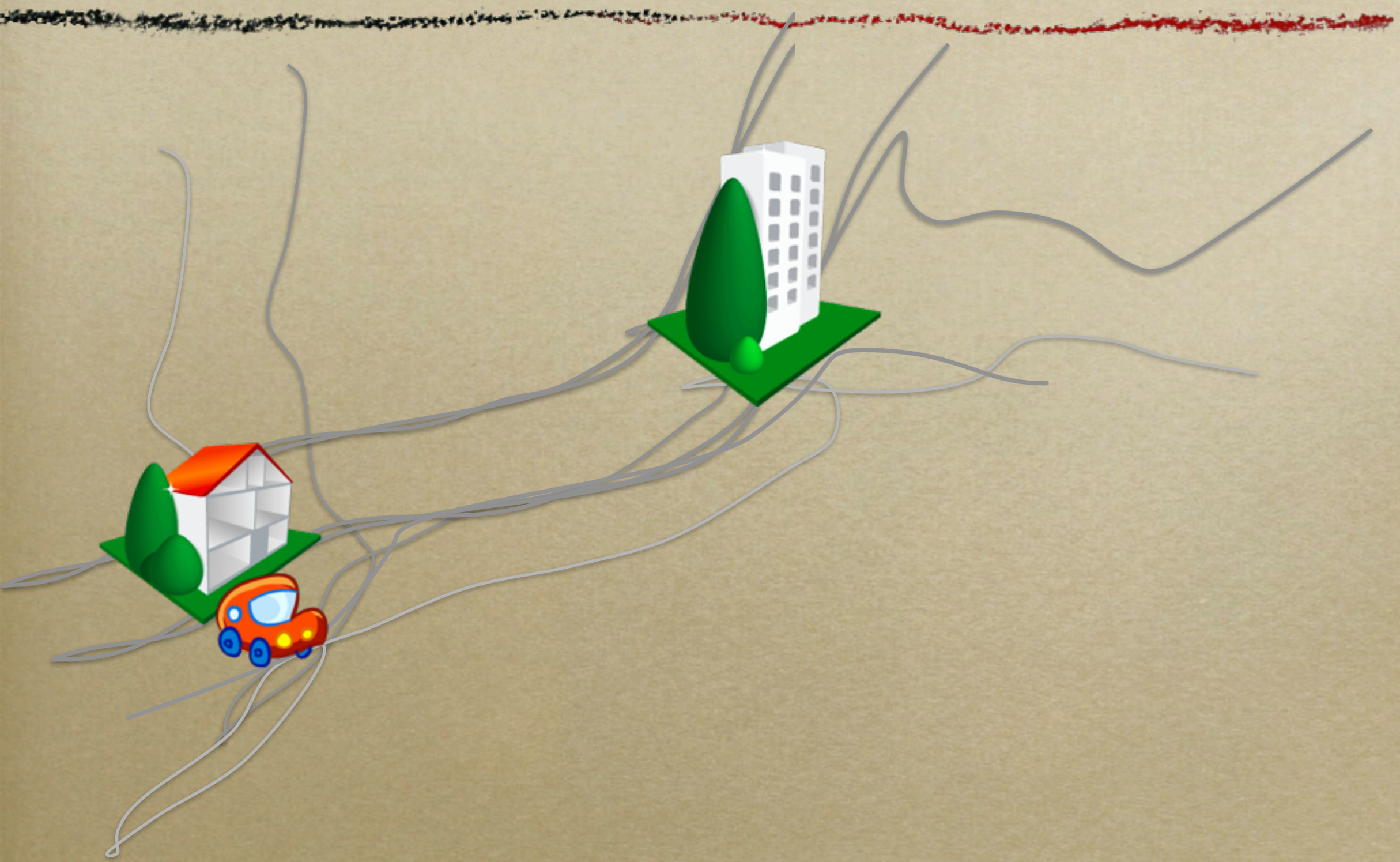
Motivation (1/5)

Rapid development of positioning techniques, users can easily collect their trajectories

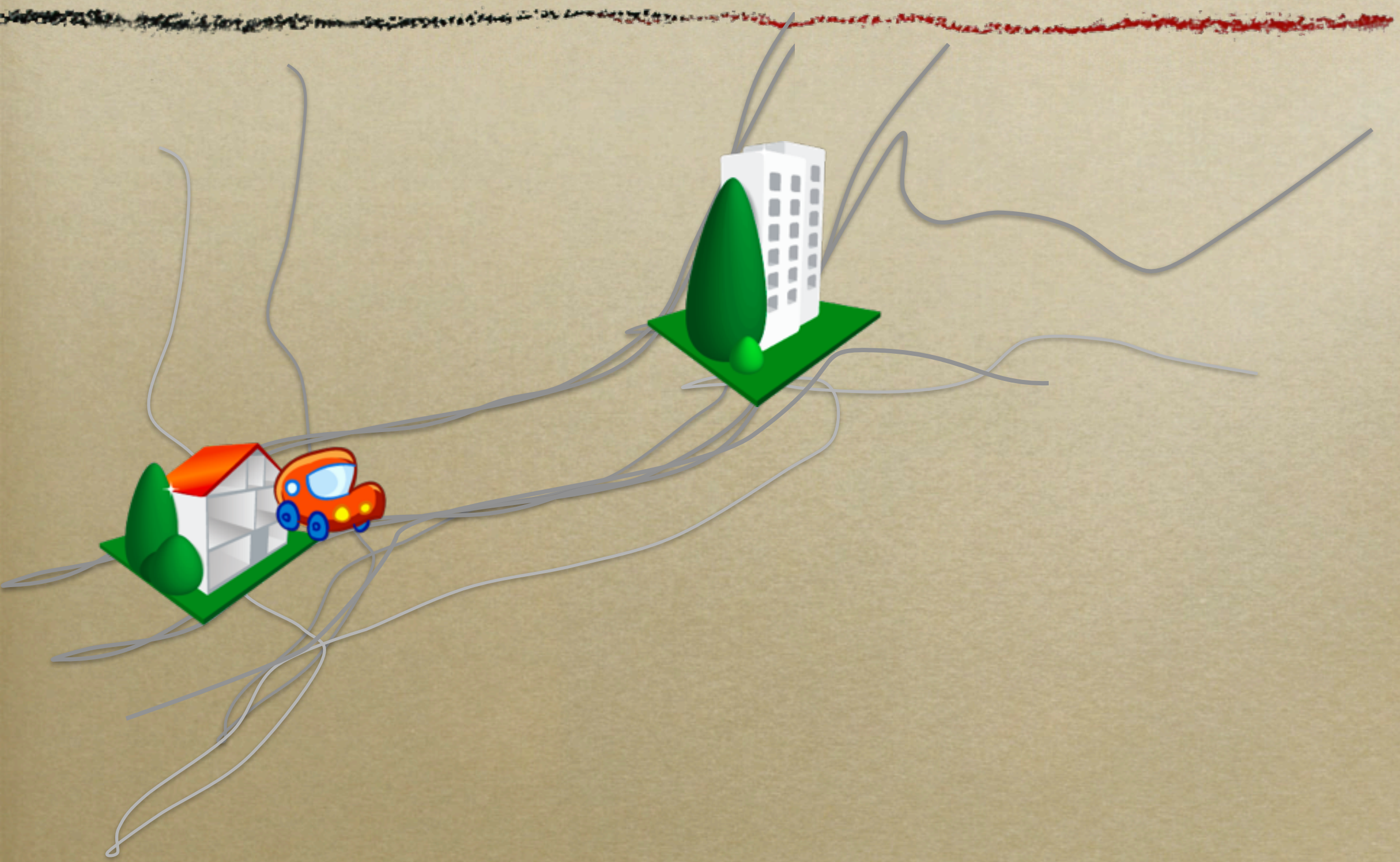
- GPS Logger, smart phones and navigation devices



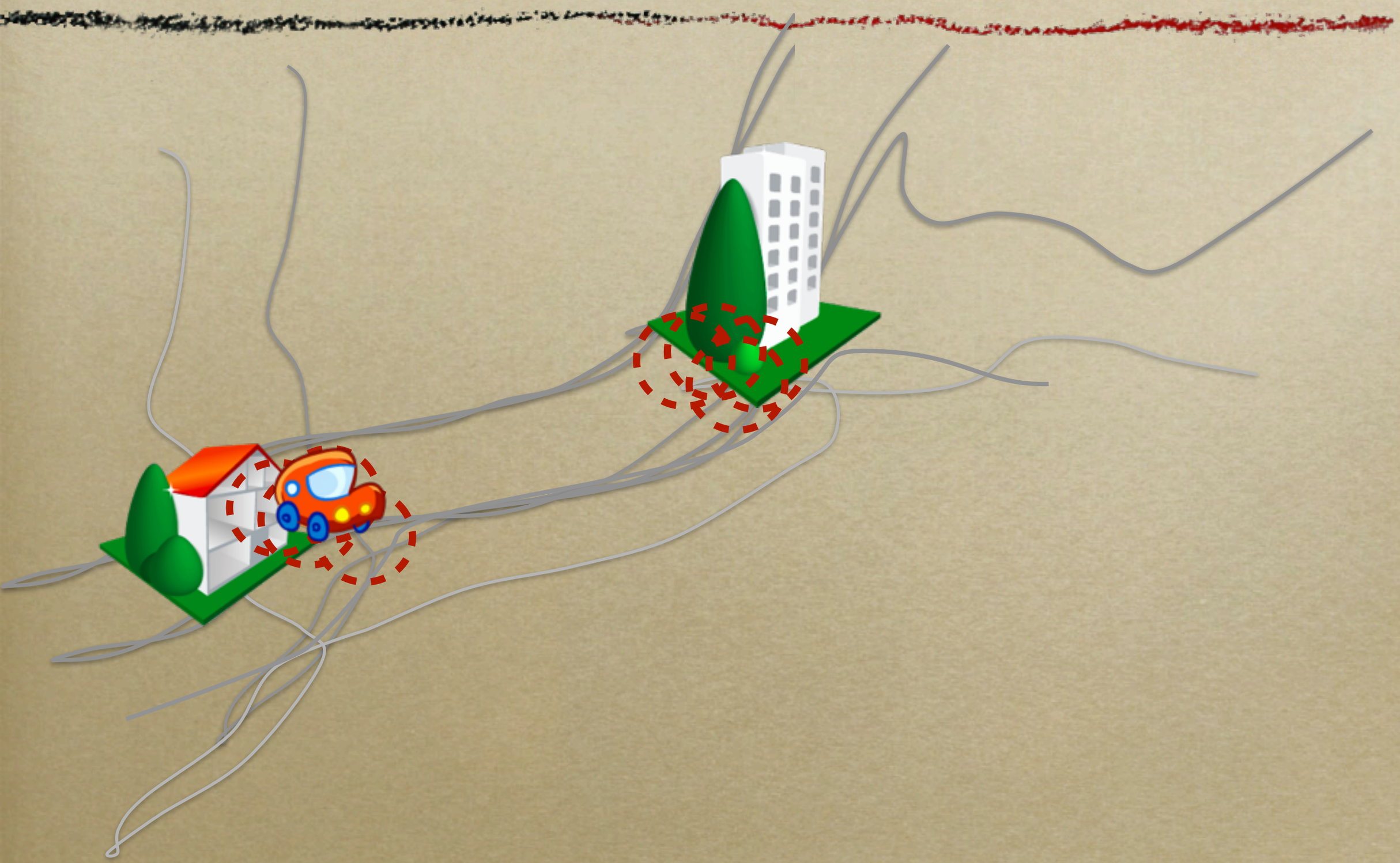
Motivation (2/5)



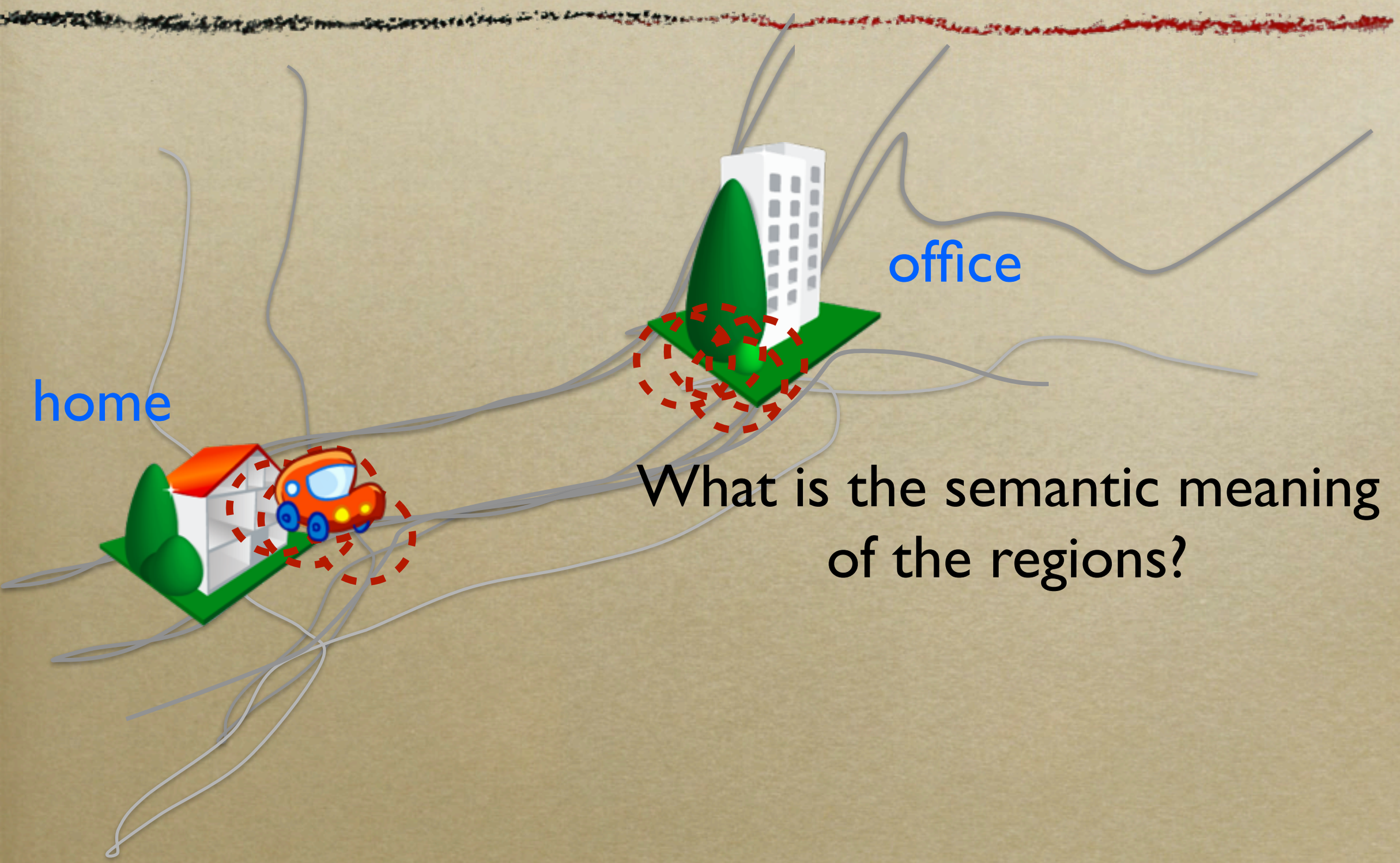
Motivation (2/5)



Motivation (2/5)

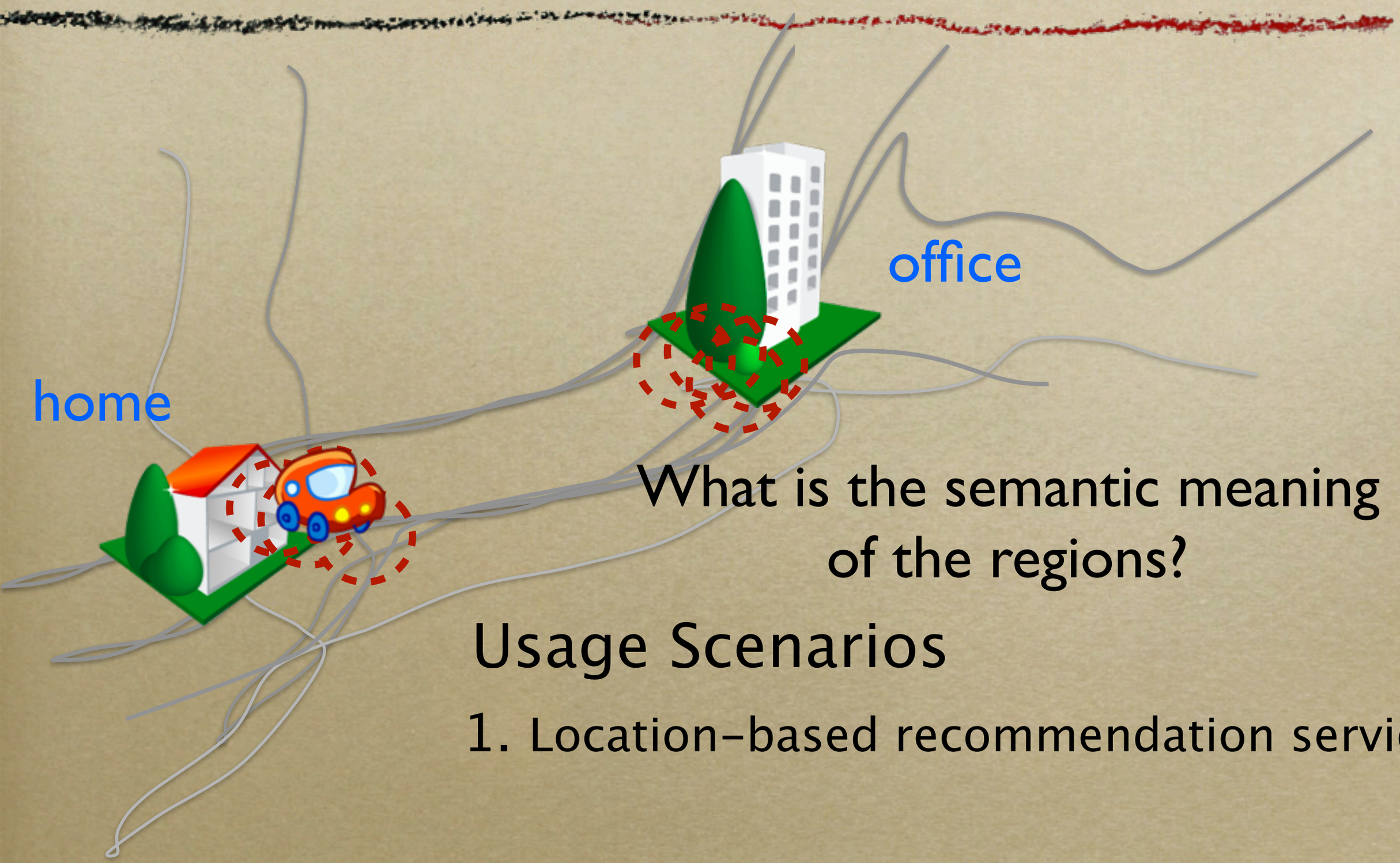


Motivation (2/5)



What is the semantic meaning of the regions?

Motivation (2/5)

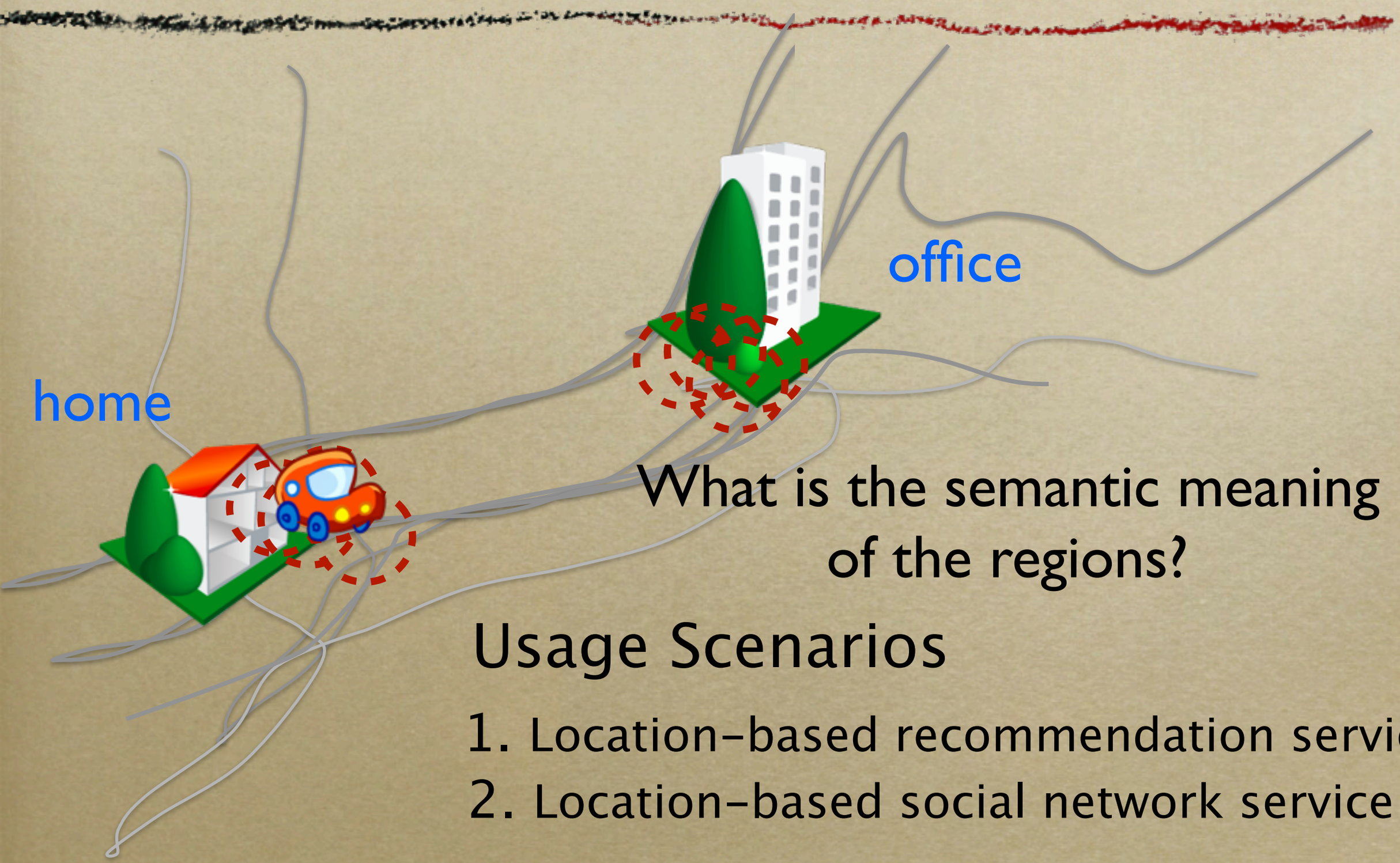


What is the semantic meaning of the regions?

Usage Scenarios

1. Location-based recommendation service

Motivation (2/5)

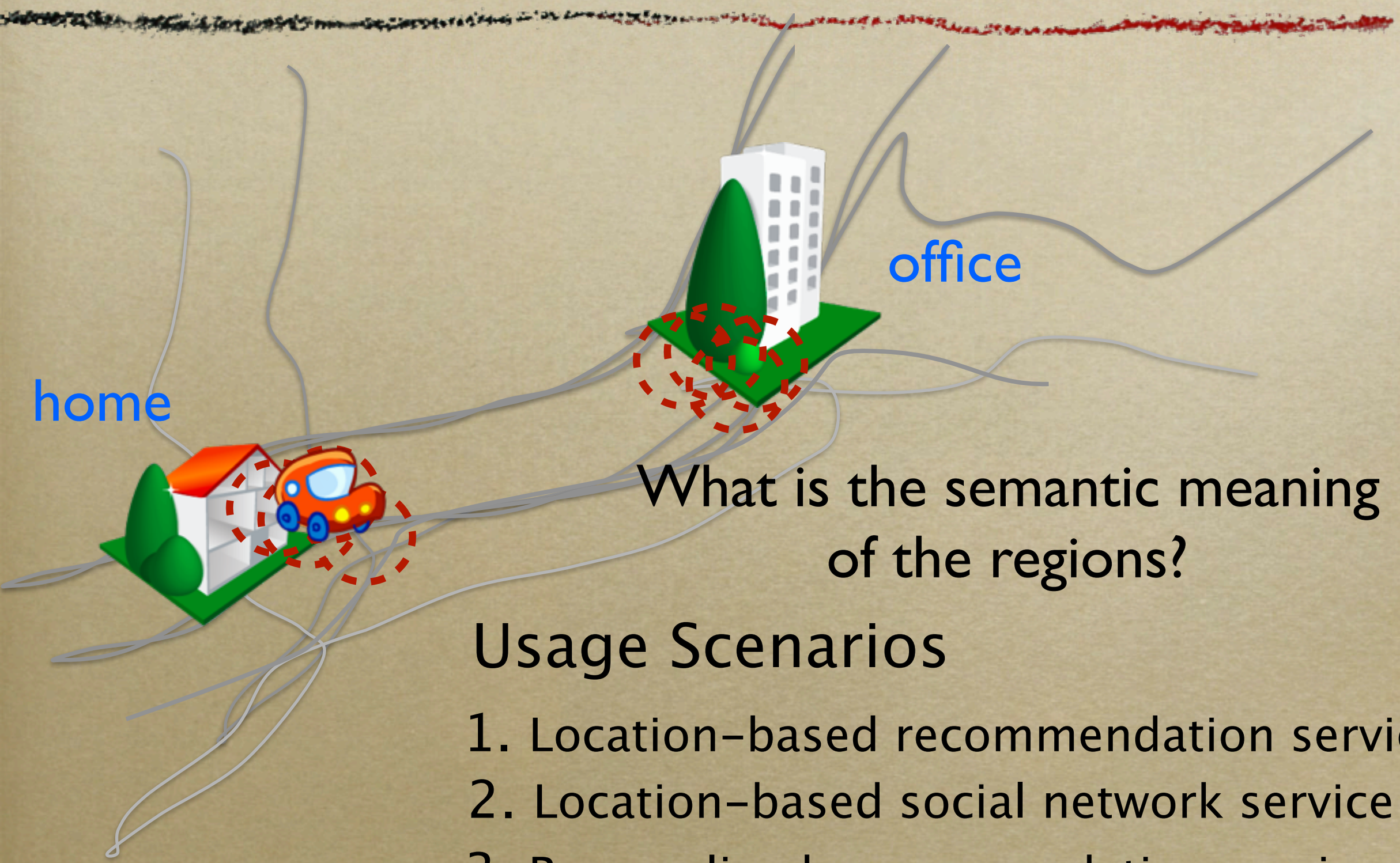


What is the semantic meaning of the regions?

Usage Scenarios

1. Location-based recommendation service
2. Location-based social network service

Motivation (2/5)

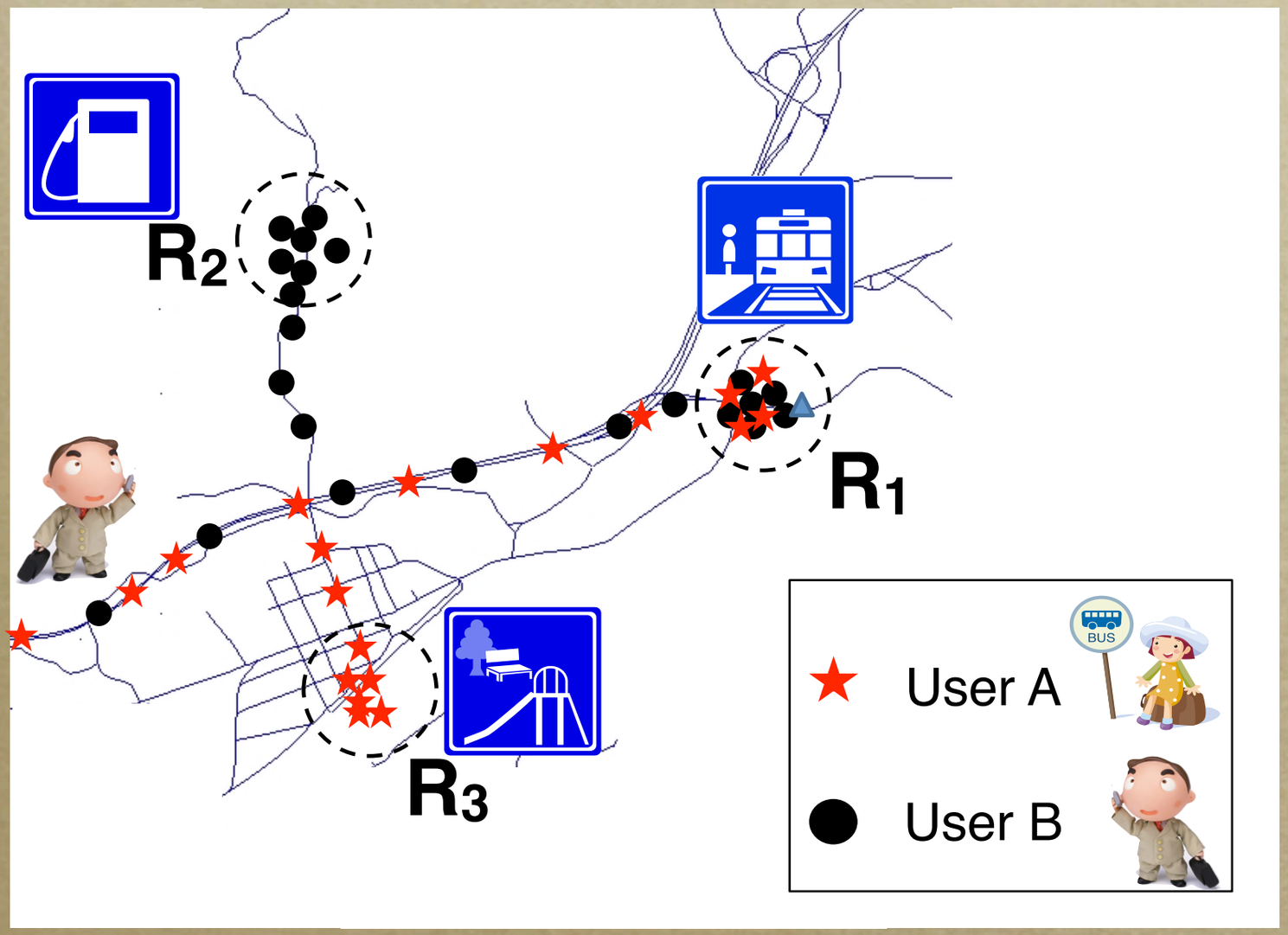


What is the semantic meaning of the regions?

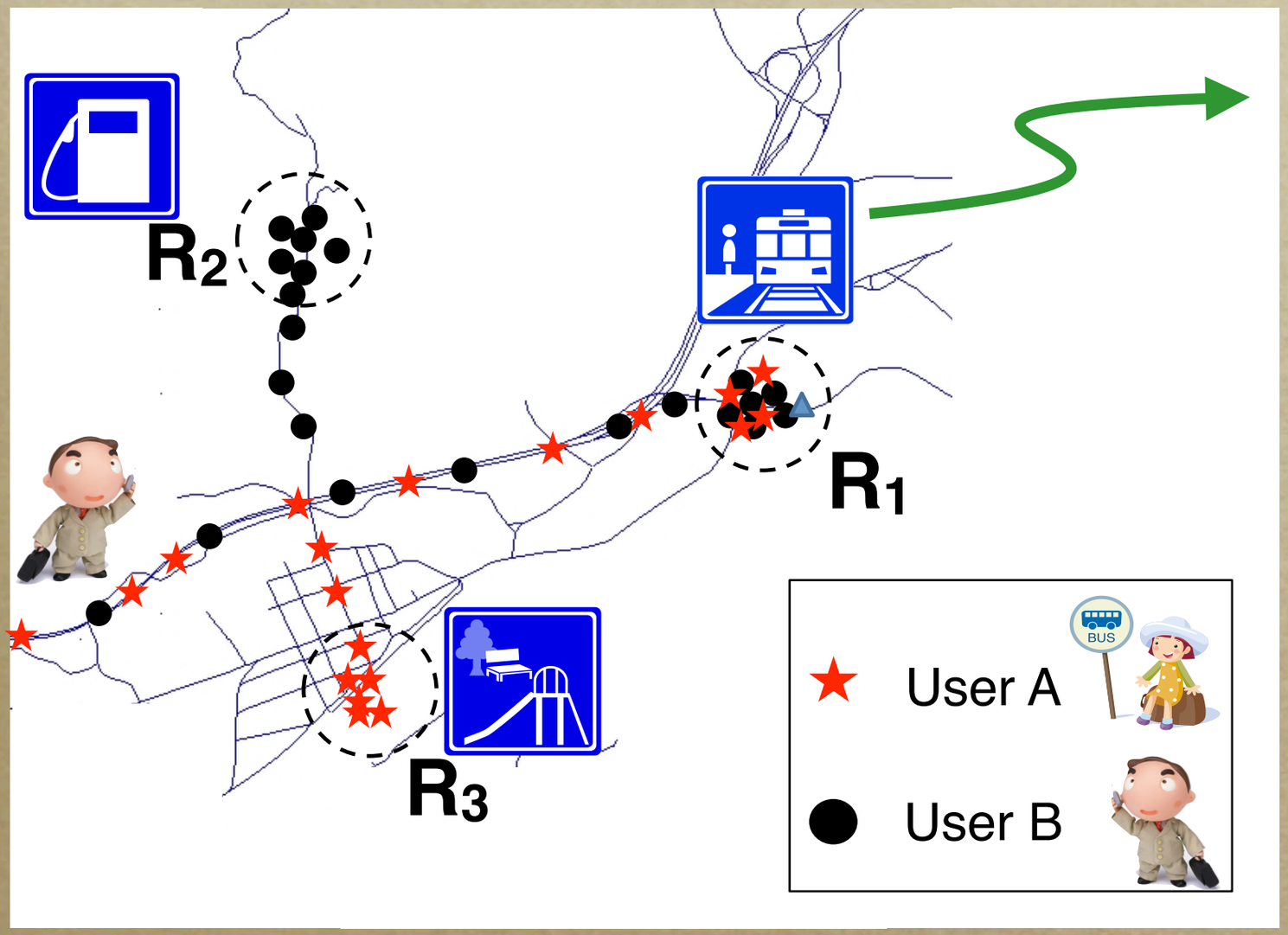
Usage Scenarios

1. Location-based recommendation service
2. Location-based social network service
3. Personalized recommendation service

Motivation (3/5)



Motivation (3/5)

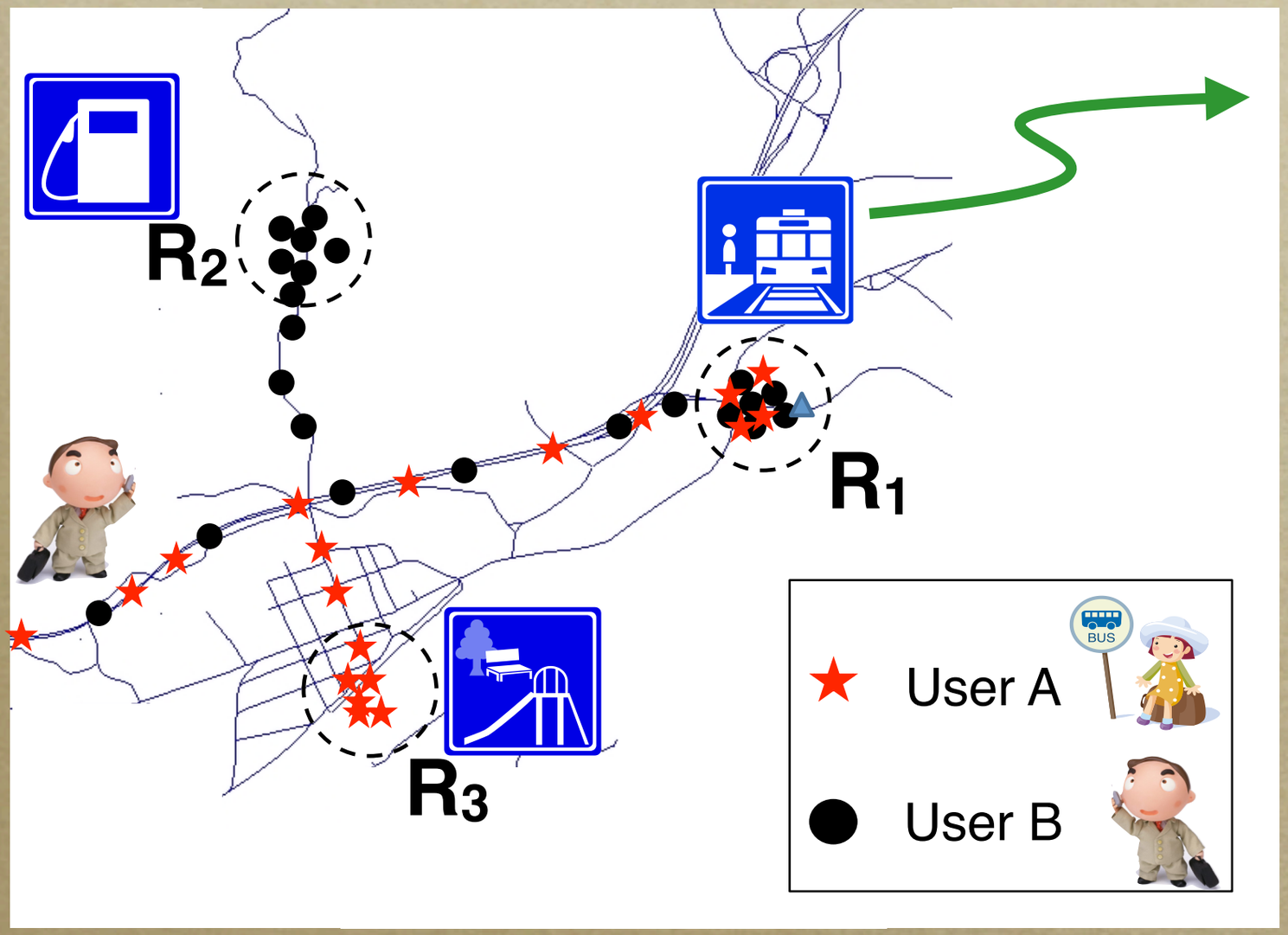


geographic information

⇩

semantic of region

Motivation (3/5)



geographic information

⇩

semantic of region

BUT!!!

Different users have different behavior

Motivation (4/5)

Motivation (4/5)

Motivation (4/5)

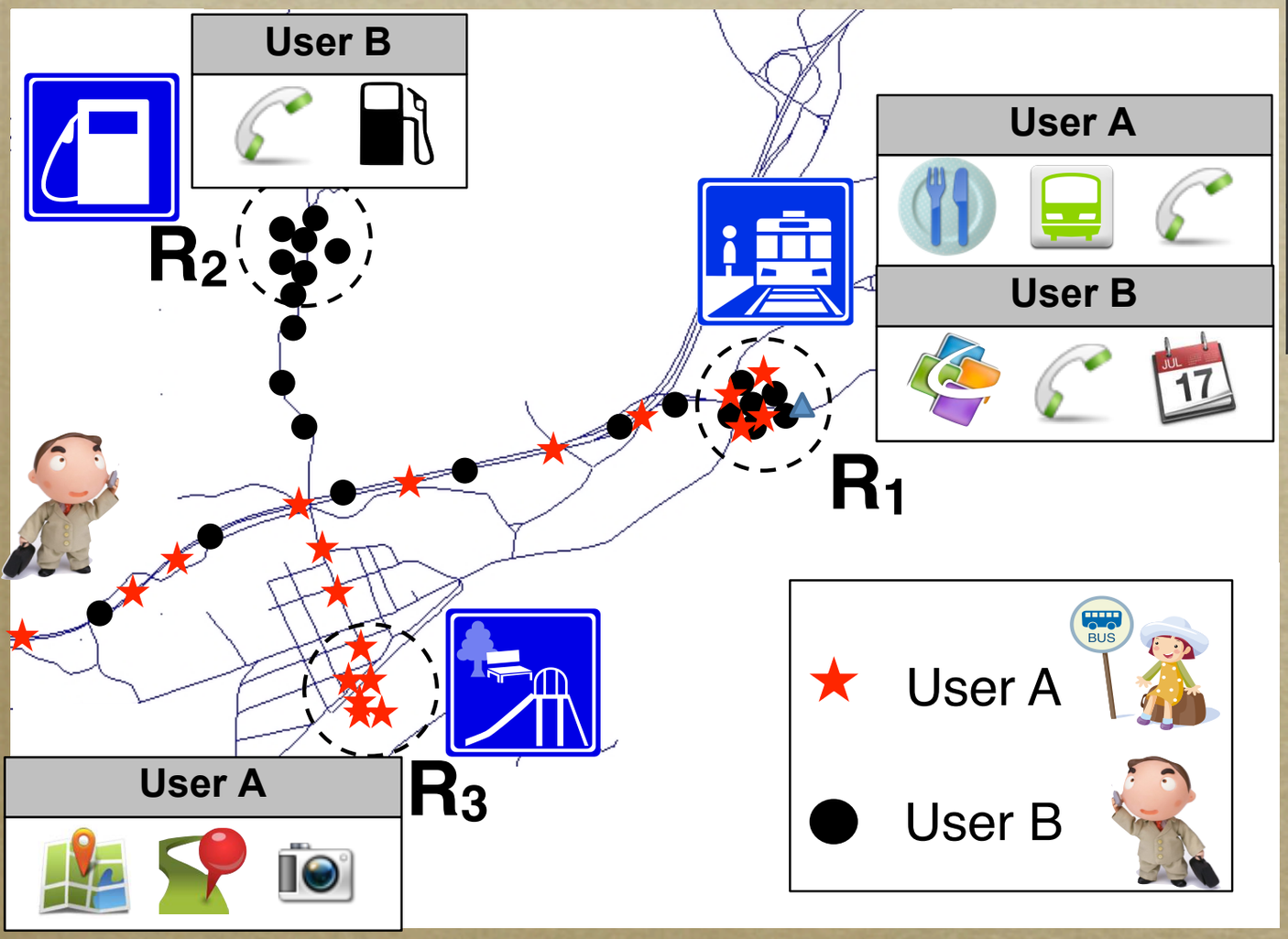
Motivation (4/5)

Motivation (4/5)

Motivation (4/5)

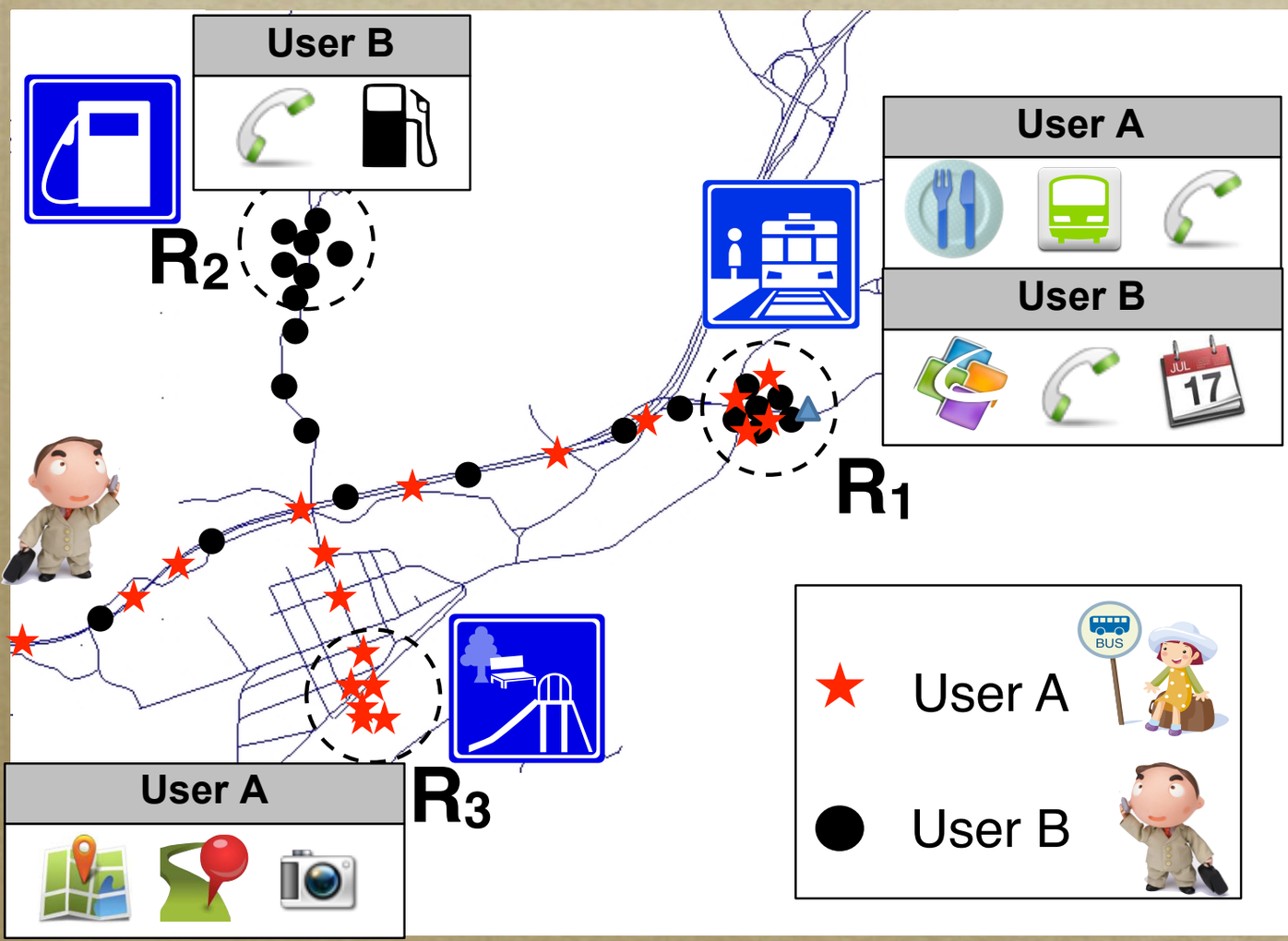
App reflect behavior

Motivation (5/5)



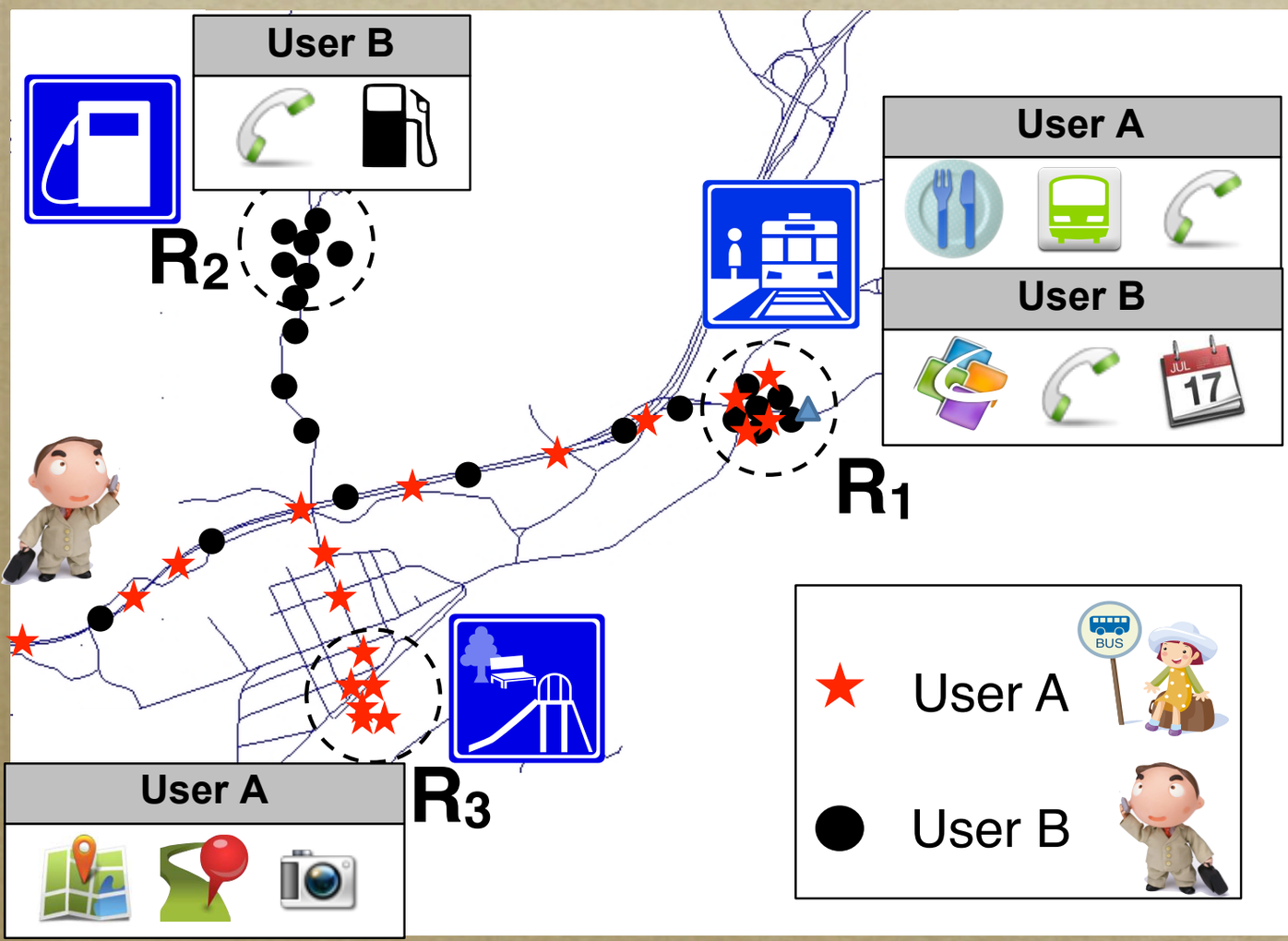
	R_1	R_2	R_3
User A			
User B			

Motivation (5/5)



	R_1	R_2	R_3
User A			
User B			

Motivation (5/5)



	R_1	R_2	R_3
User A			
User B			

App reflect behavior

	R_1	R_2	R_3
User A	*Travel*		*Travel*
User B	*Office*	*Shopping*	

Outline

- *Motivation*
- *Goal*
- *Related Works*
- *Framework*
  - *Region Extraction*
  - *Personal Semantic Extraction*
- *Experiment*
- *Conclusion*

Goal

Discover personal semantic regions

Input:

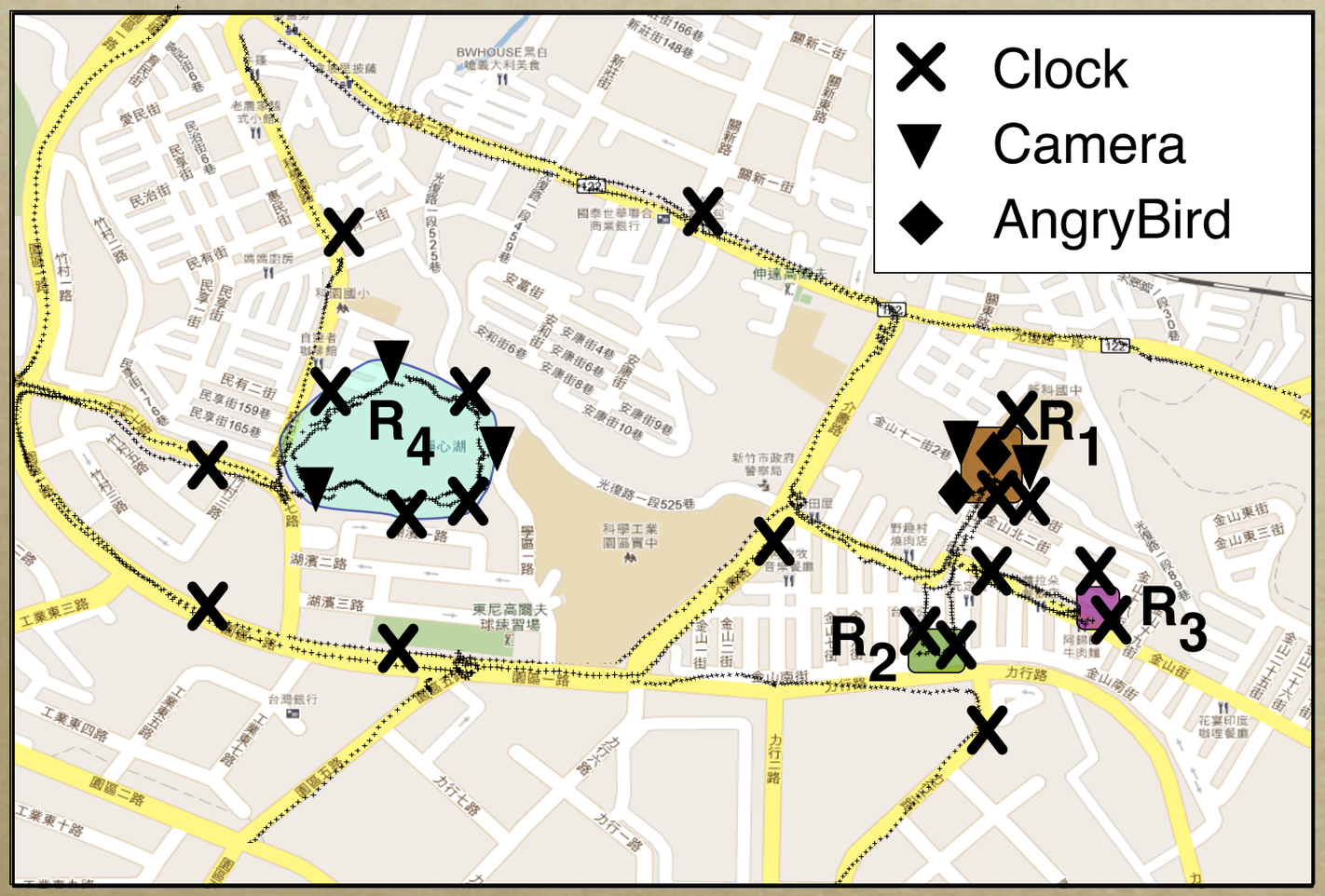
- Geographic trajectories
- App usage logs

Output:

- Personal semantic regions
  - ➡ Region
  - ➡ geographic semantic
  - ➡ behavior semantic

Issue:

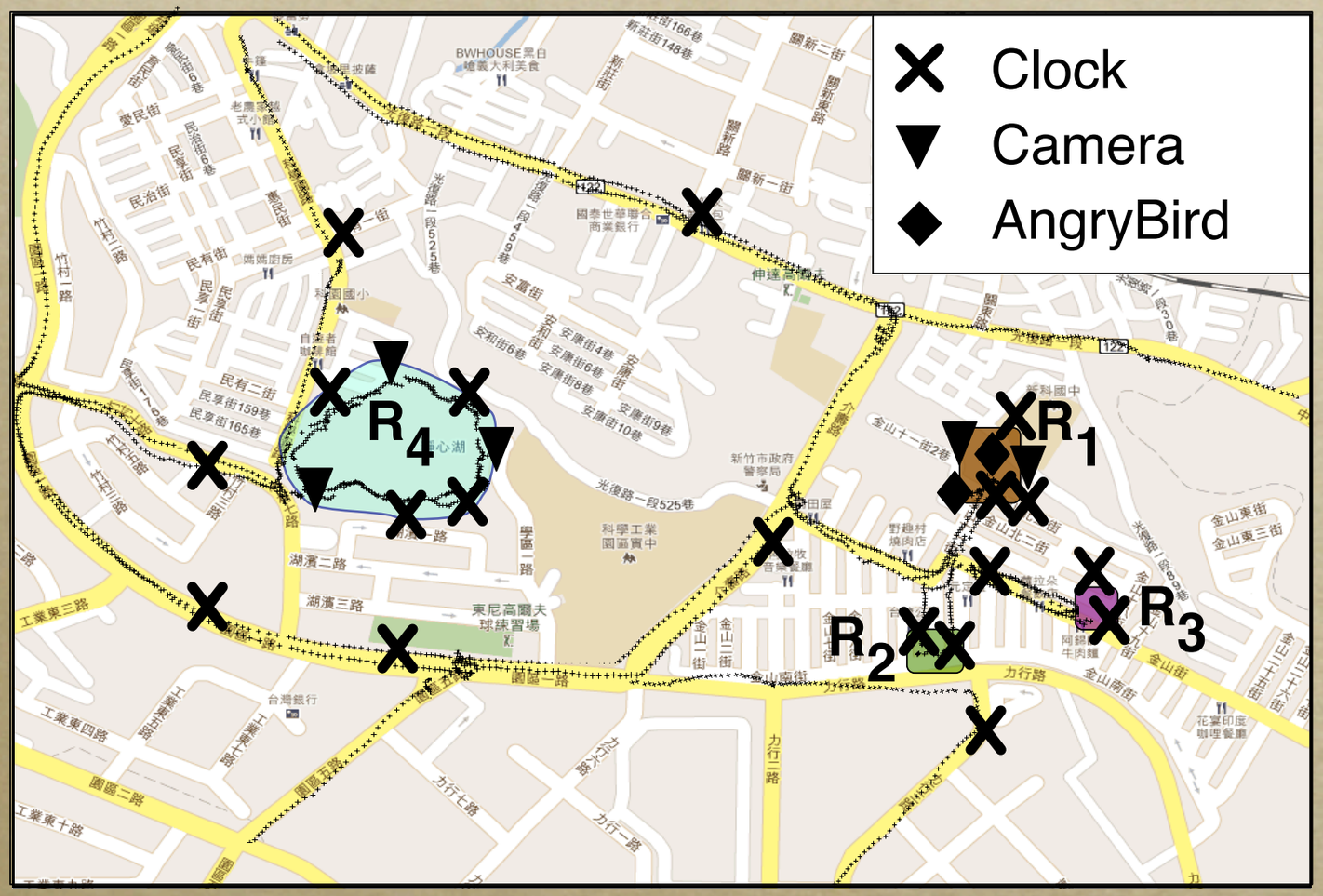
Semantic Extraction from App



	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

Issue:

Semantic Extraction from App

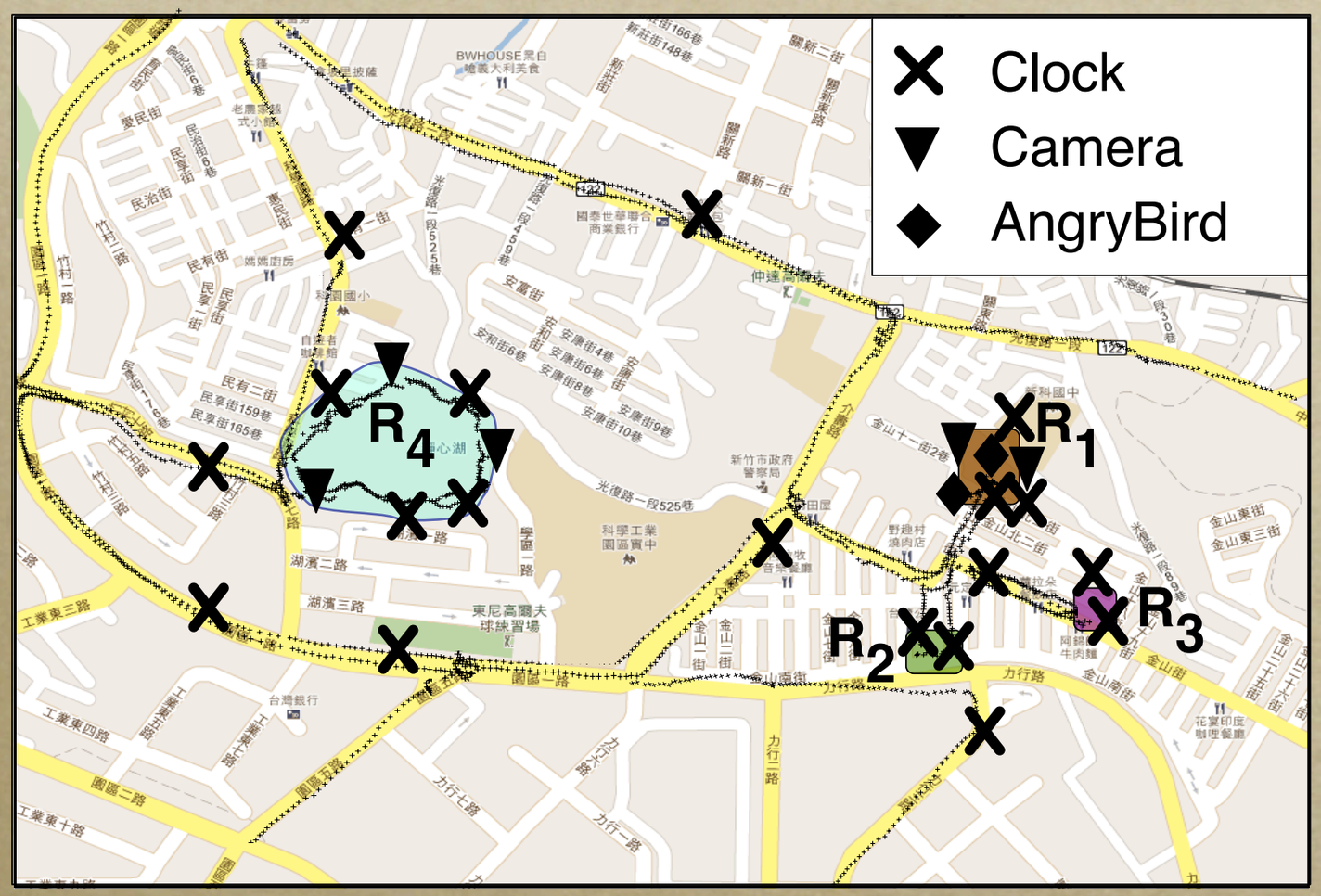


	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

Which App can reflect user behavior on a location?

Issue:

Semantic Extraction from App



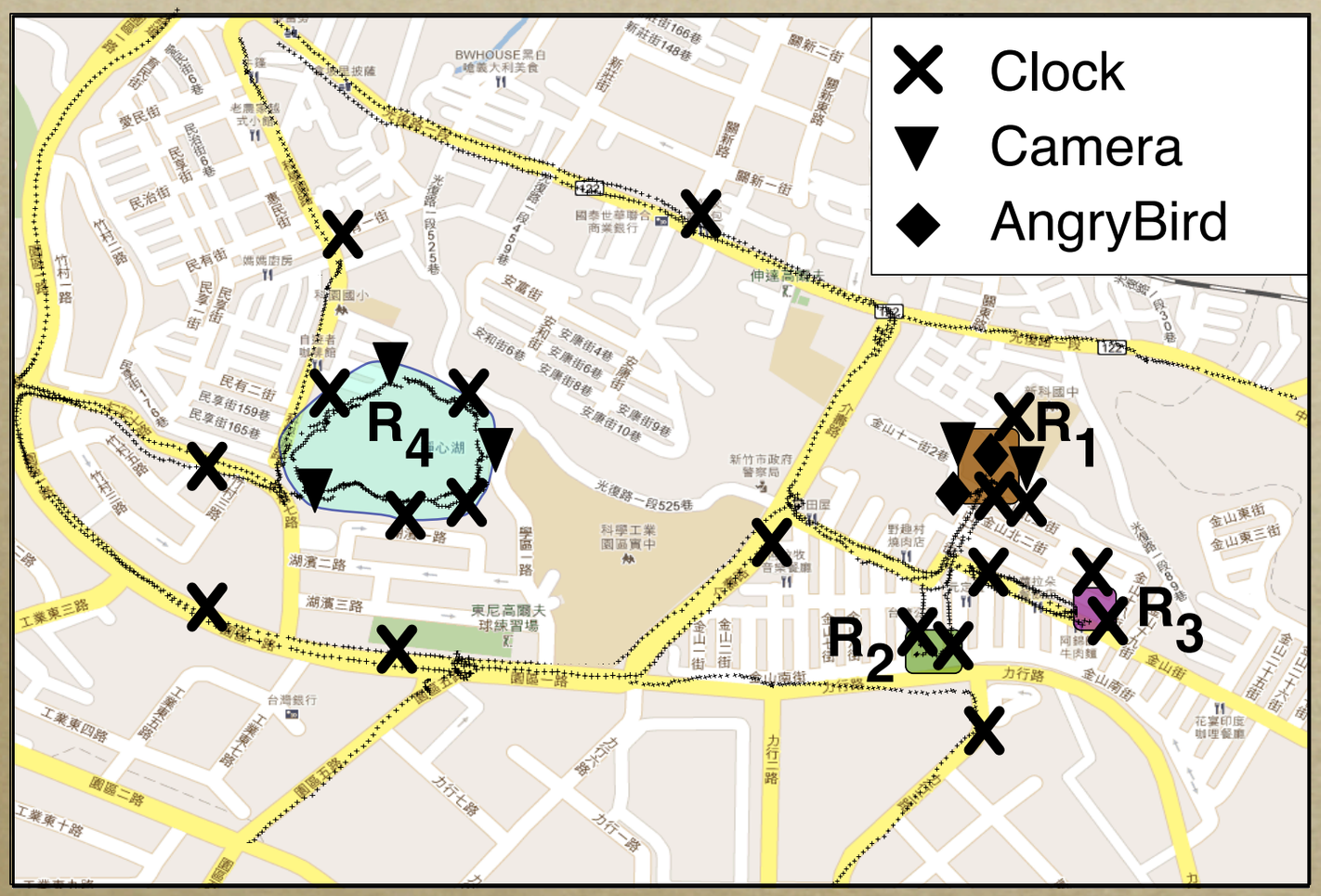
	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

representative Apps

Which App can reflect user behavior on a location?

Issue:

Semantic Extraction from App



	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

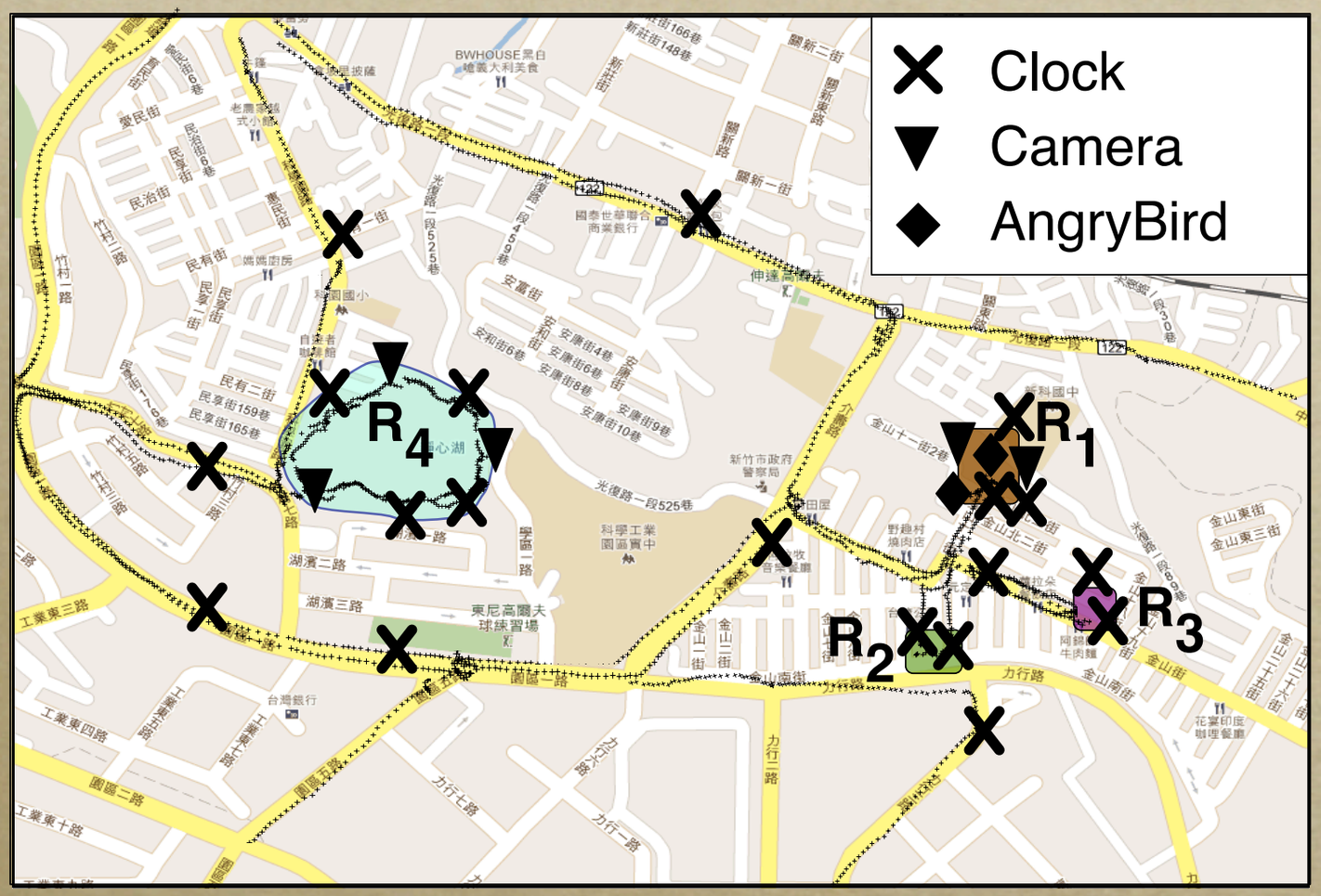
representative Apps

Which App can reflect user behavior on a location?

How many representative Apps?

Issue:

Semantic Extraction from App



	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

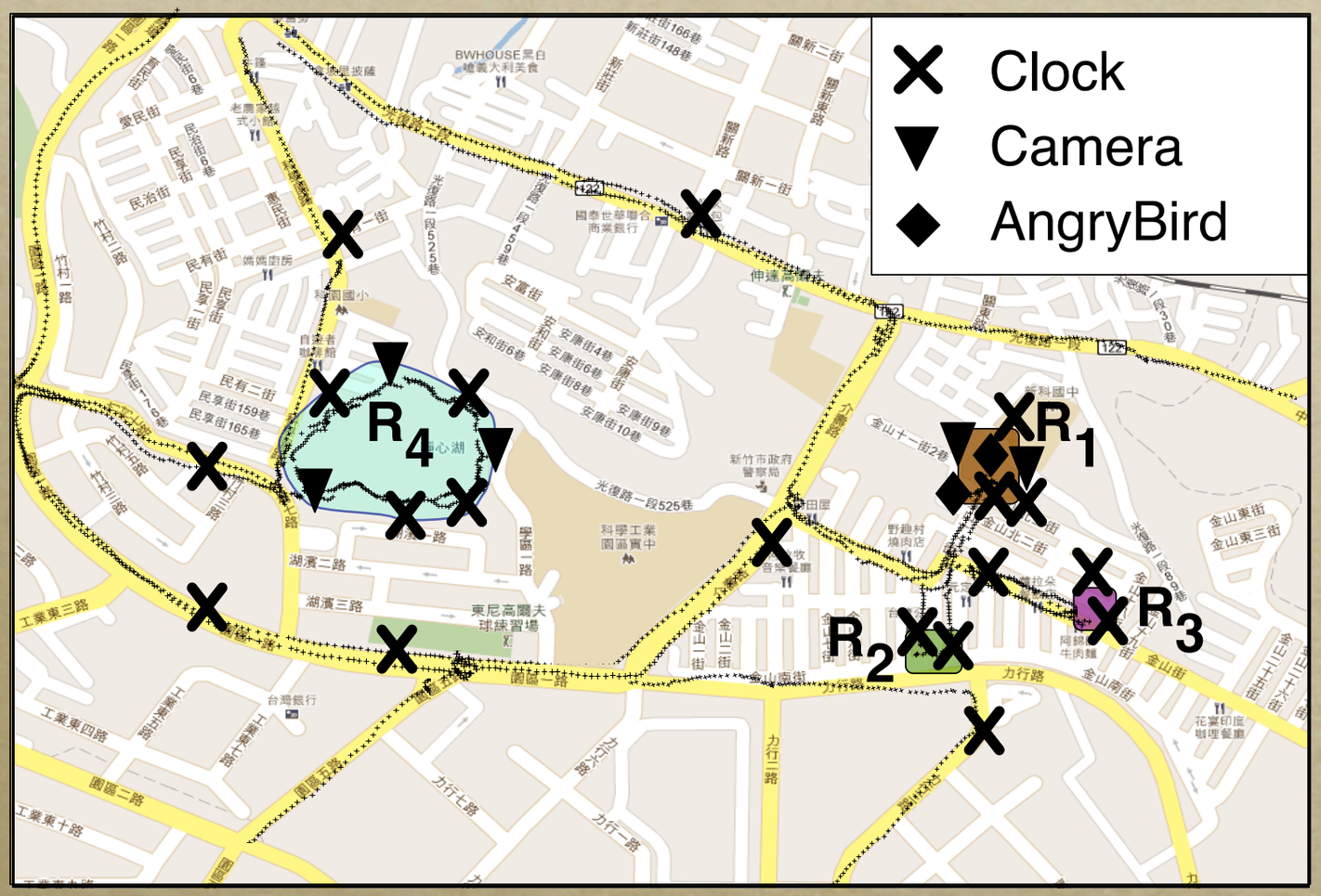
representative Apps

Which App can reflect user behavior on a location?

How many representative Apps? K

Issue:

Semantic Extraction from App



	R_1	R_2	R_3	R_4
Clock	3	2	2	4
Camera	2	0	0	3
Angry Bird	2	0	0	0

representative Apps

Which App can reflect user behavior on a location?

How many representative Apps? K

App usage pattern: K representative Apps

Outline

- *Motivation*
- *Goal*
- *Related Works*
- *Framework*
  - *Region Extraction*
  - *Personal Semantic Extraction*
- *Experiment*
- *Conclusion*

Related Work

Region Extraction:

(1) spatial density based clustering [22]

(2) stay points [11,12]

Semantic Annotation:

(1) geographic information (POIs) [11]

(2) App usage mining [27]

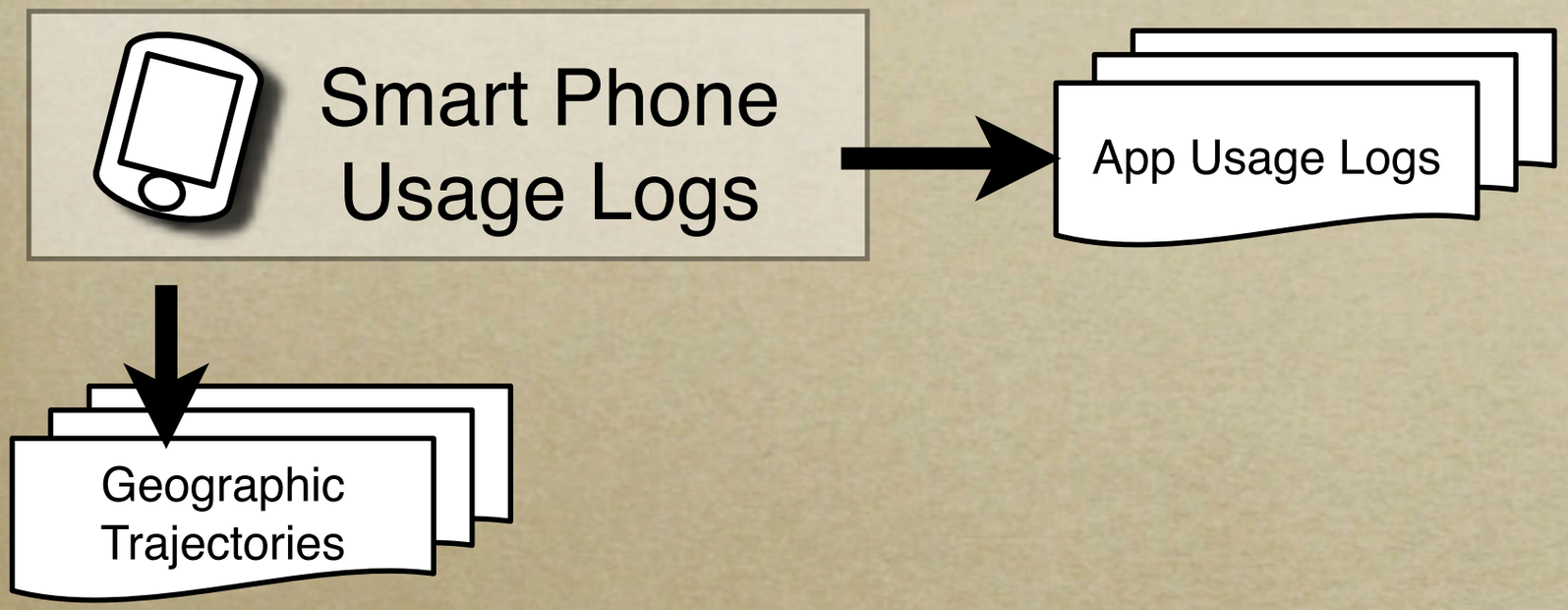
[22] M. Ester, et. al. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,”, KDD (1996)
[12] Zheng, et. at. “Mining Interesting Locations and Travel Sequences From GPS Trajectories.” ,WWW (2009)
[11] X. Cao, et. al. “Mining Significant Semantic Locations From GPS Data,” PVLDB (2010)
[27] N. Eagle et.al., “Reality Mining: Sensing Complex Social Systems,” PerCom (2006)

Outline

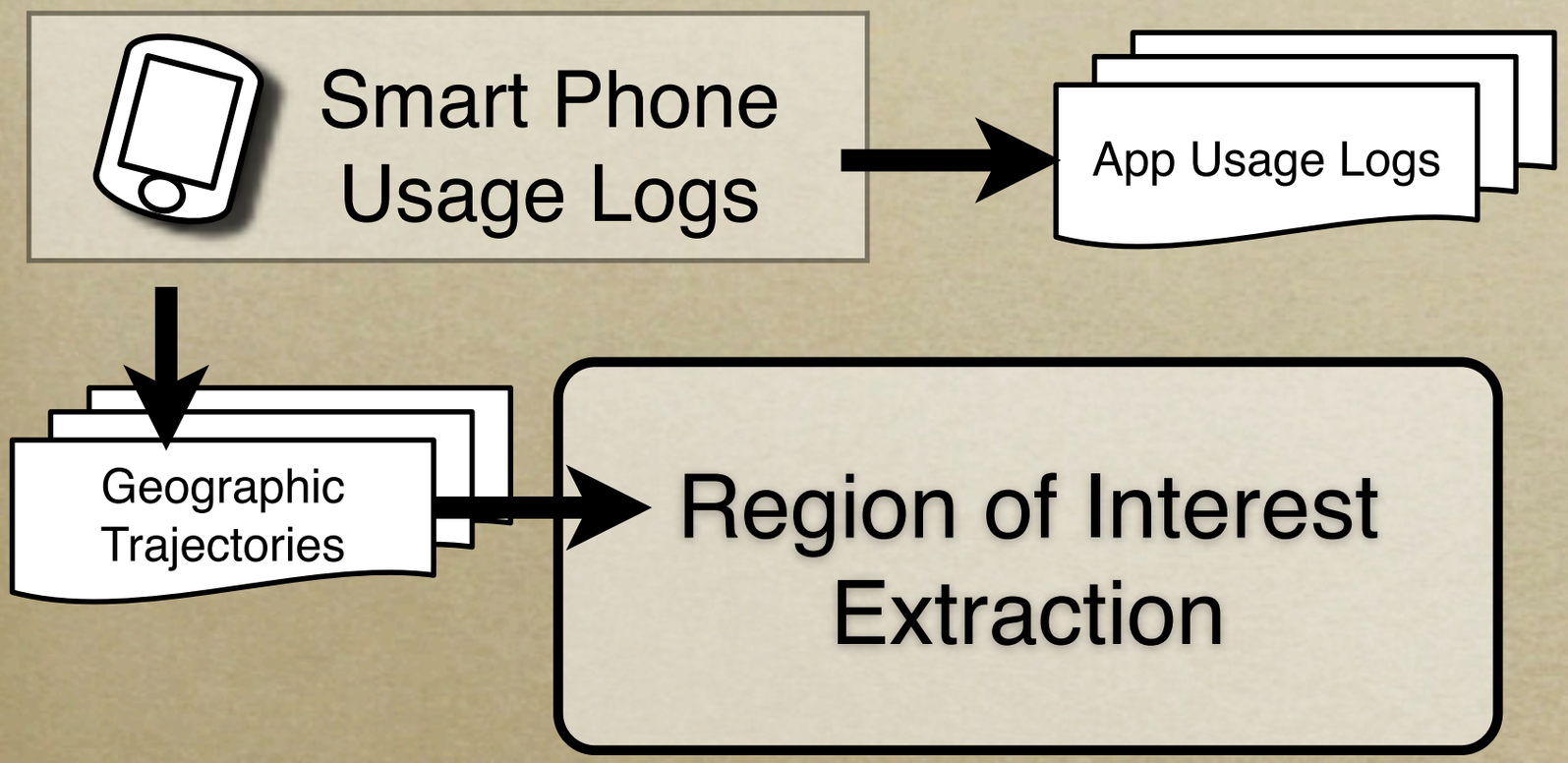
- *Motivation*
- *Goal*
- *Related Works*
- *Framework*
  - *Region Extraction*
  - *Personal Semantic Extraction*
- *Experiment*
- *Conclusion*

Framework

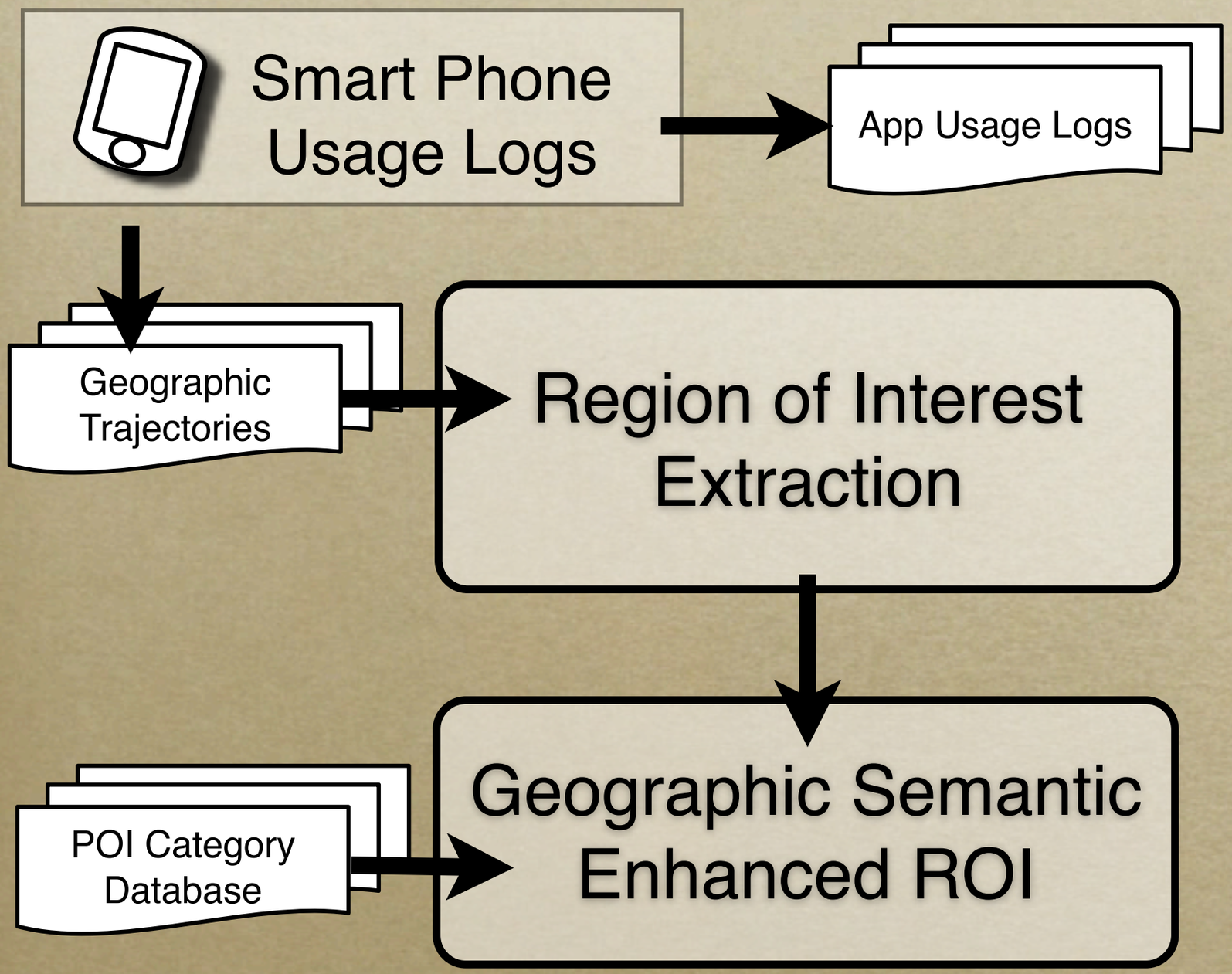
Framework


Smart Phone
Usage Logs
App Usage Logs
Geographic
Trajectories

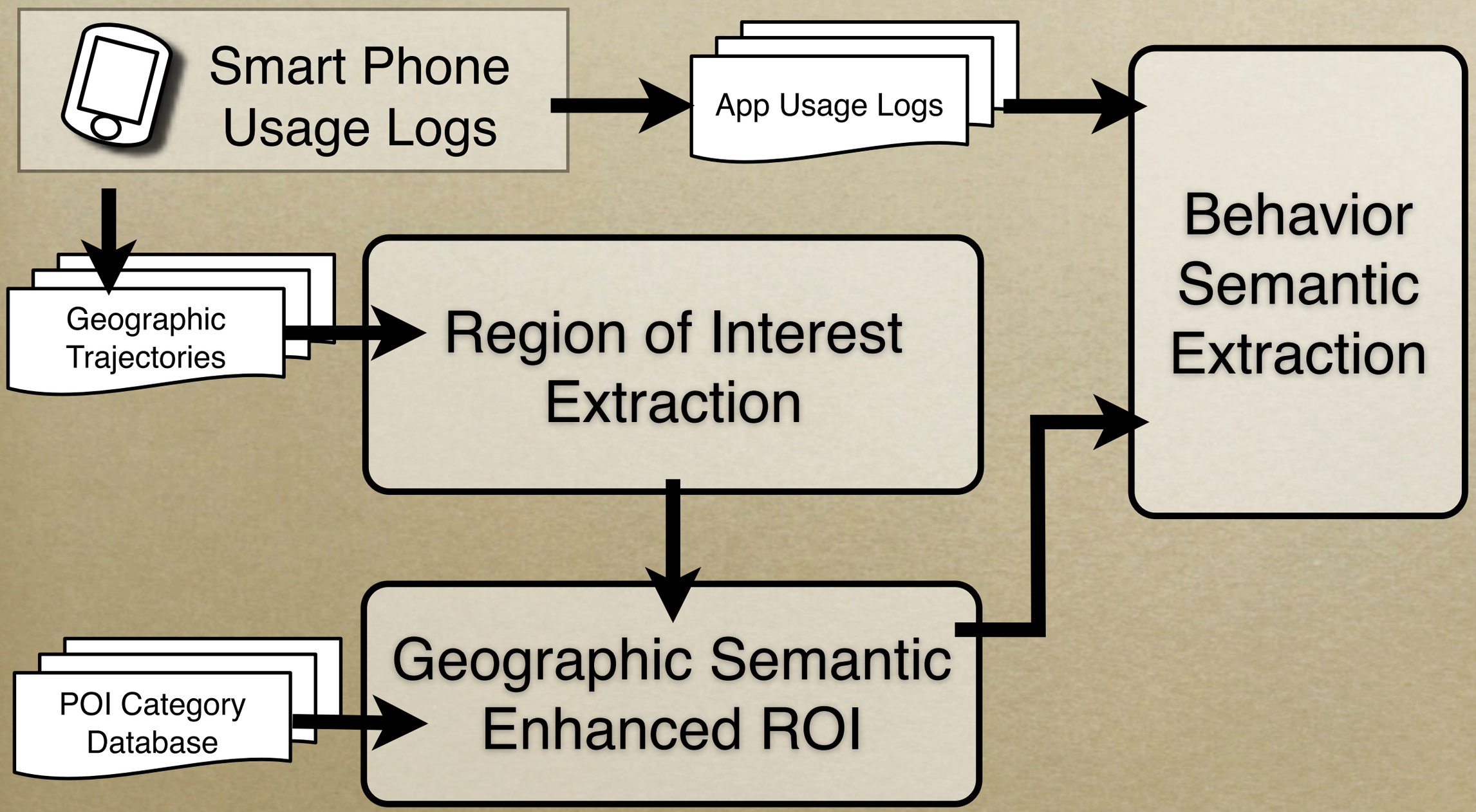
Framework


Smart Phone Usage Logs
App Usage Logs
Geographic Trajectories
Region of Interest Extraction

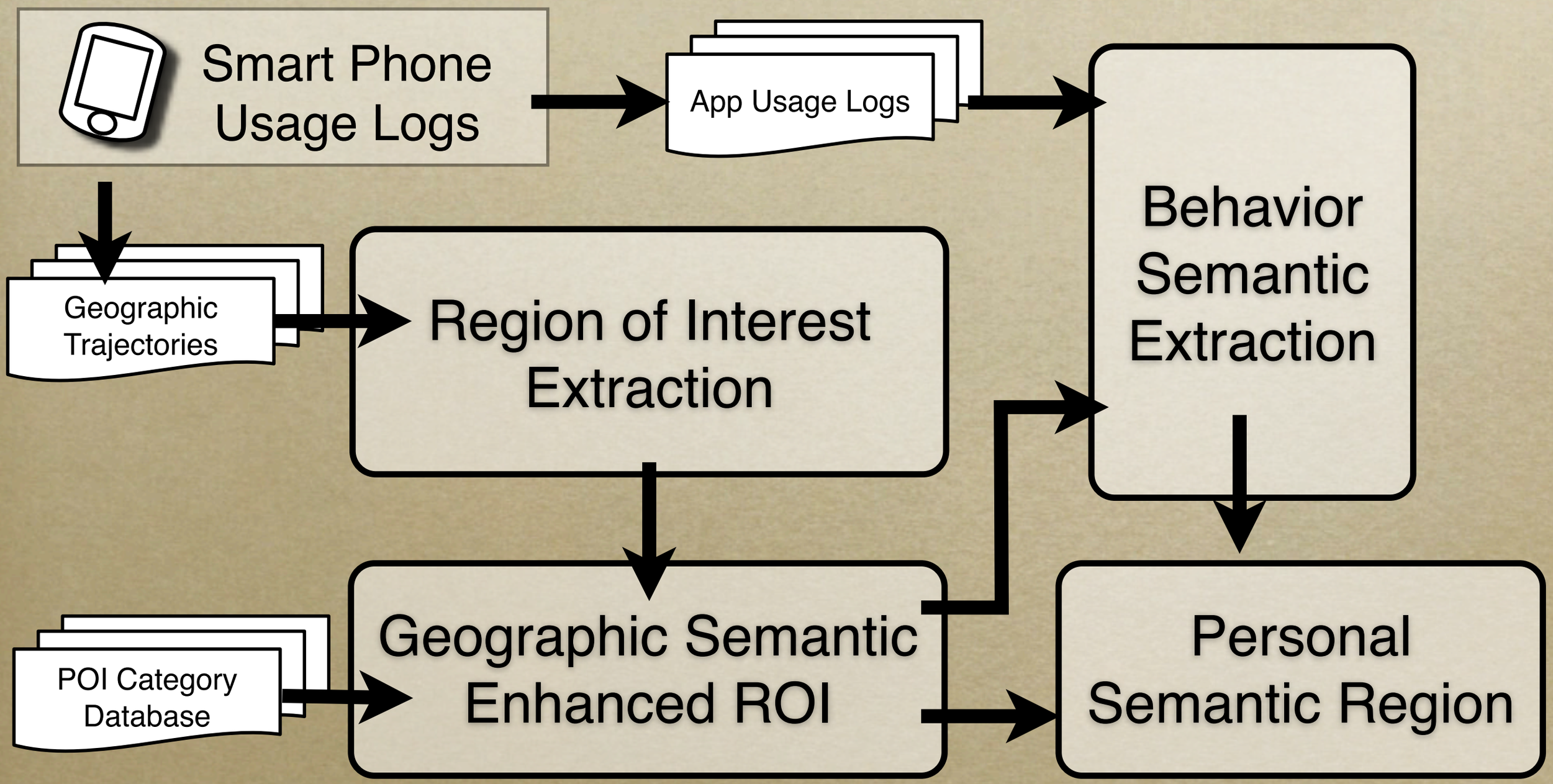
Framework


Smart Phone Usage Logs
App Usage Logs
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI

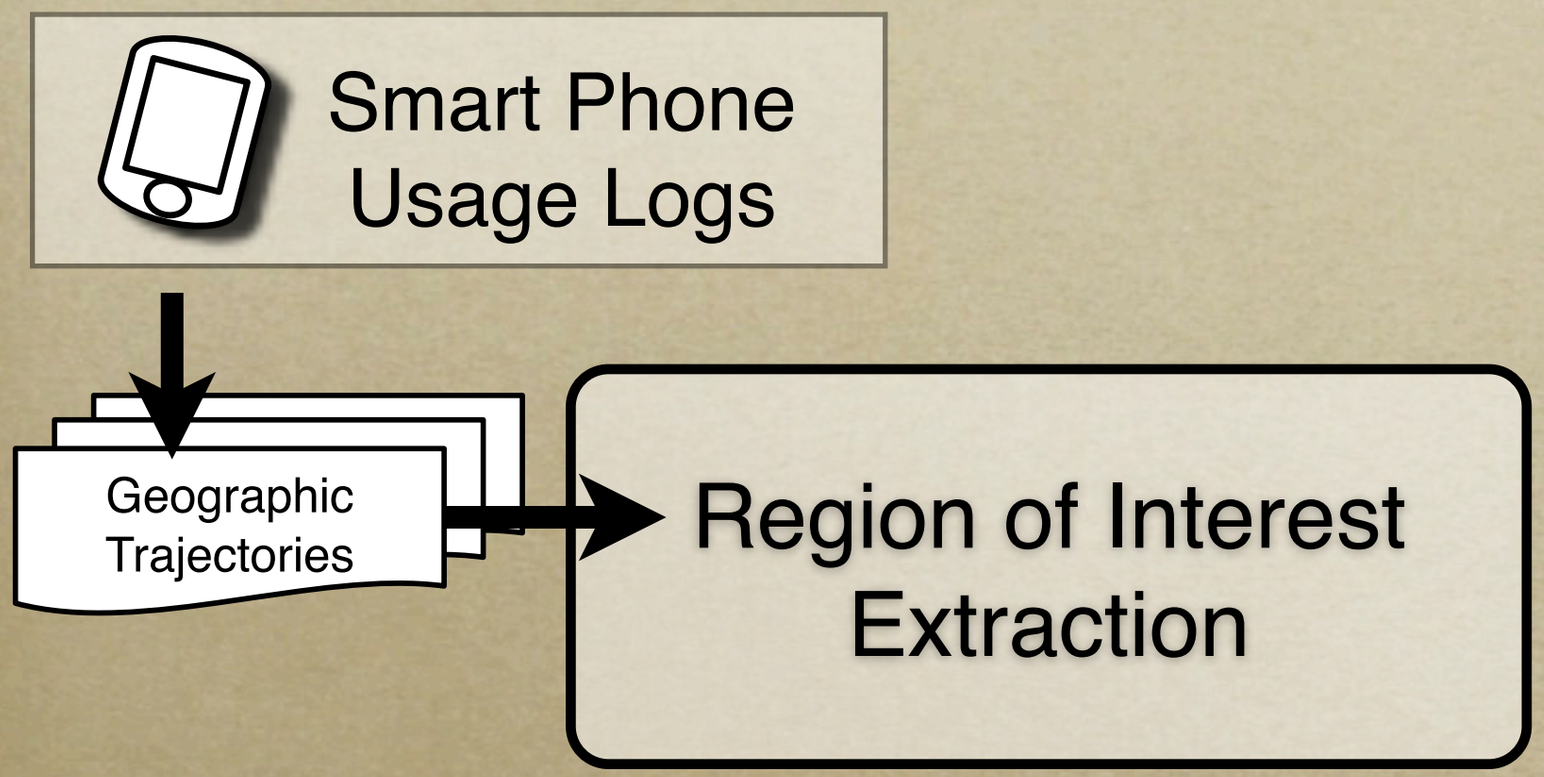
Framework


Smart Phone Usage Logs
App Usage Logs
Behavior Semantic Extraction
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI

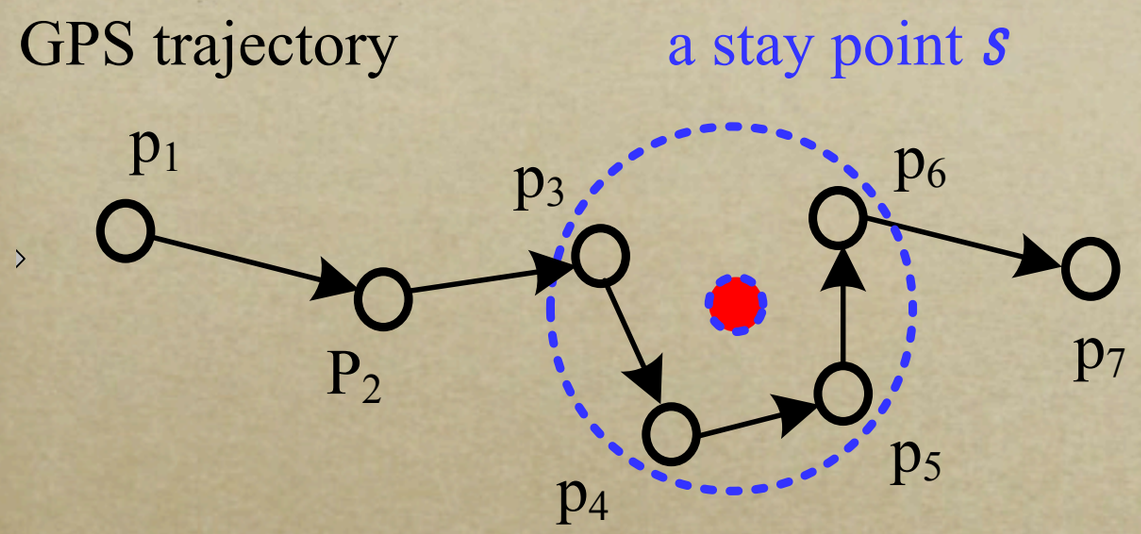
Framework


Smart Phone Usage Logs
App Usage Logs
Behavior Semantic Extraction
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI
Personal Semantic Region

Framework


Smart Phone
Usage Logs
Geographic
Trajectories
Region of Interest
Extraction

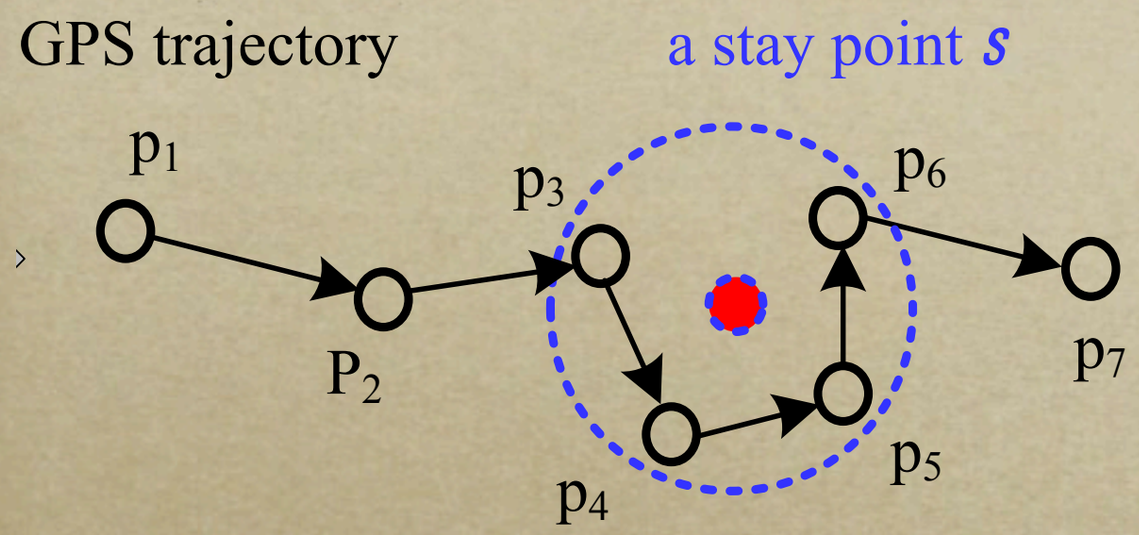
Issue: Region Extraction (1/2)



A stay point s can detected from consecutive GPS points $P=\{pm, pm+1, \ldots, pn\}$, where $\forall m<i\leq n$, $Dist(pm,pi) \leq D_{threh}$ and $pn.T-pm.T \geq T_{threh}$

Zheng, et. at. "Mining Interesting Locations and Travel Sequences From GPS Trajectories." ,WWW (2009)

Issue: Region Extraction (1/2)

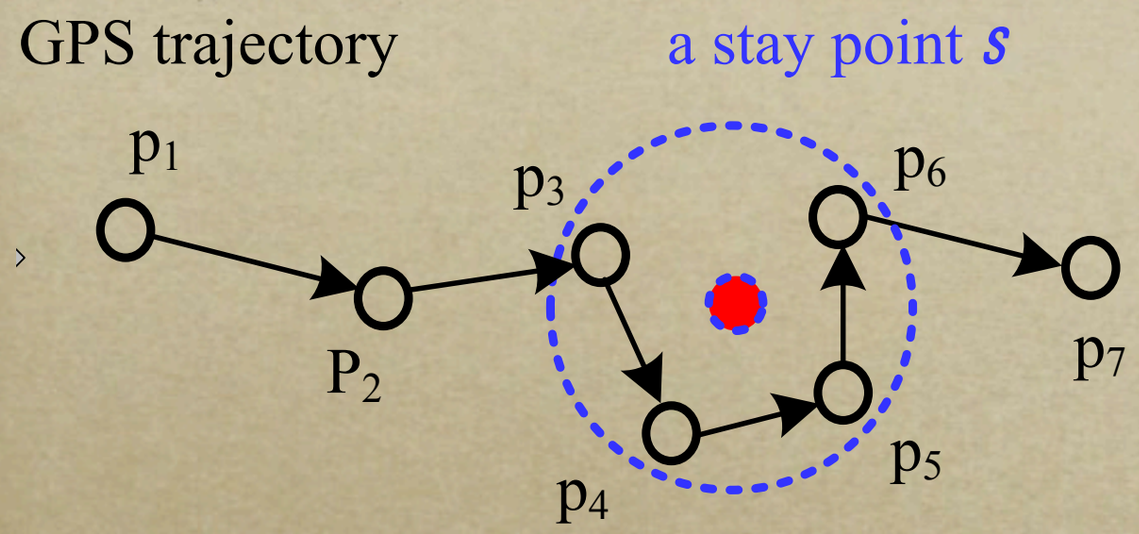


A stay point s can detected from consecutive GPS points $P=\{pm, pm+1, \ldots, pn\}$, where $\forall m<i\leq n$, $Dist(pm,pi) \leq D_{threh}$ and $pn.T-pm.T \geq T_{threh}$

Universal distance limit cannot find all regions

Zheng, et. at. "Mining Interesting Locations and Travel Sequences From GPS Trajectories." ,WWW (2009)

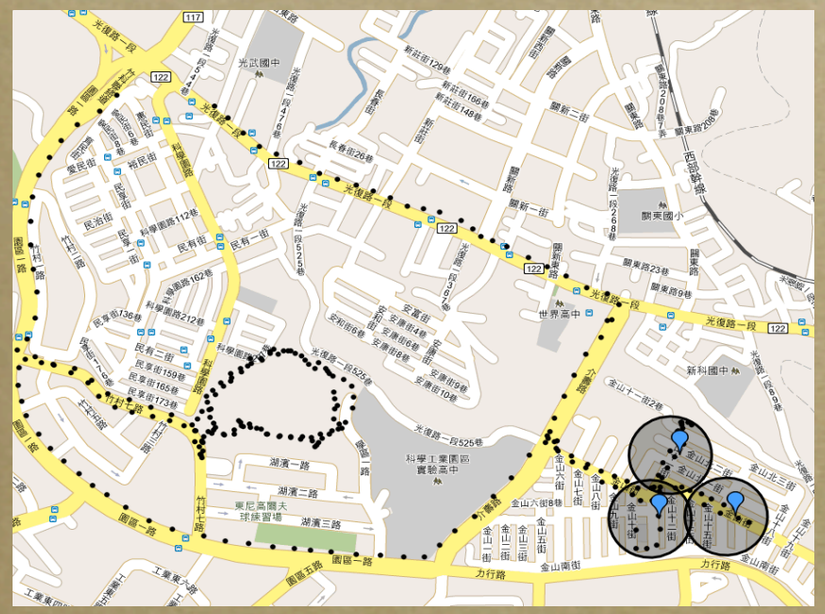
Issue: Region Extraction (1/2)



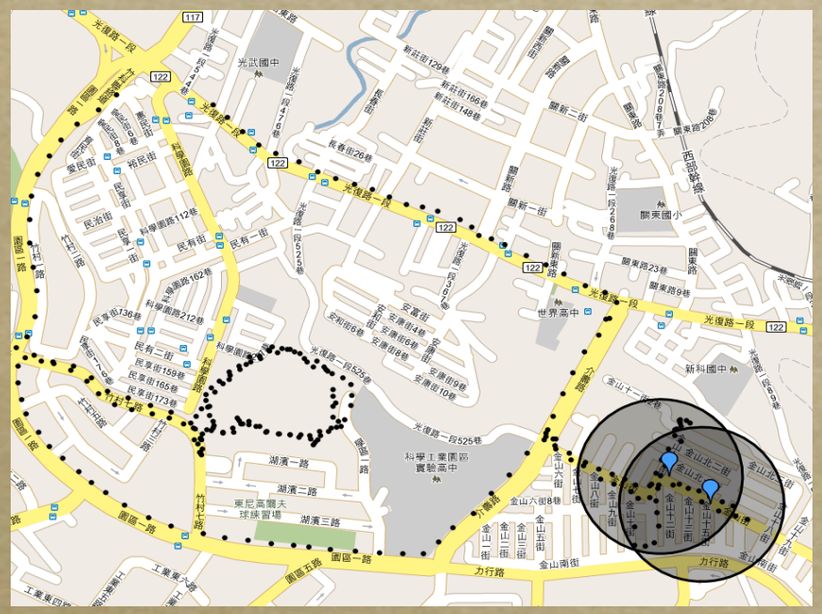
A stay point s can detected from consecutive GPS points $P=\{pm, pm+1, \ldots, pn\}$, where $\forall m<i\leq n$, $Dist(pm,pi) \leq D_{threh}$ and $pn.T-pm.T \geq T_{threh}$

Universal distance limit cannot find all regions

T: 10 mins, D: 100 meters



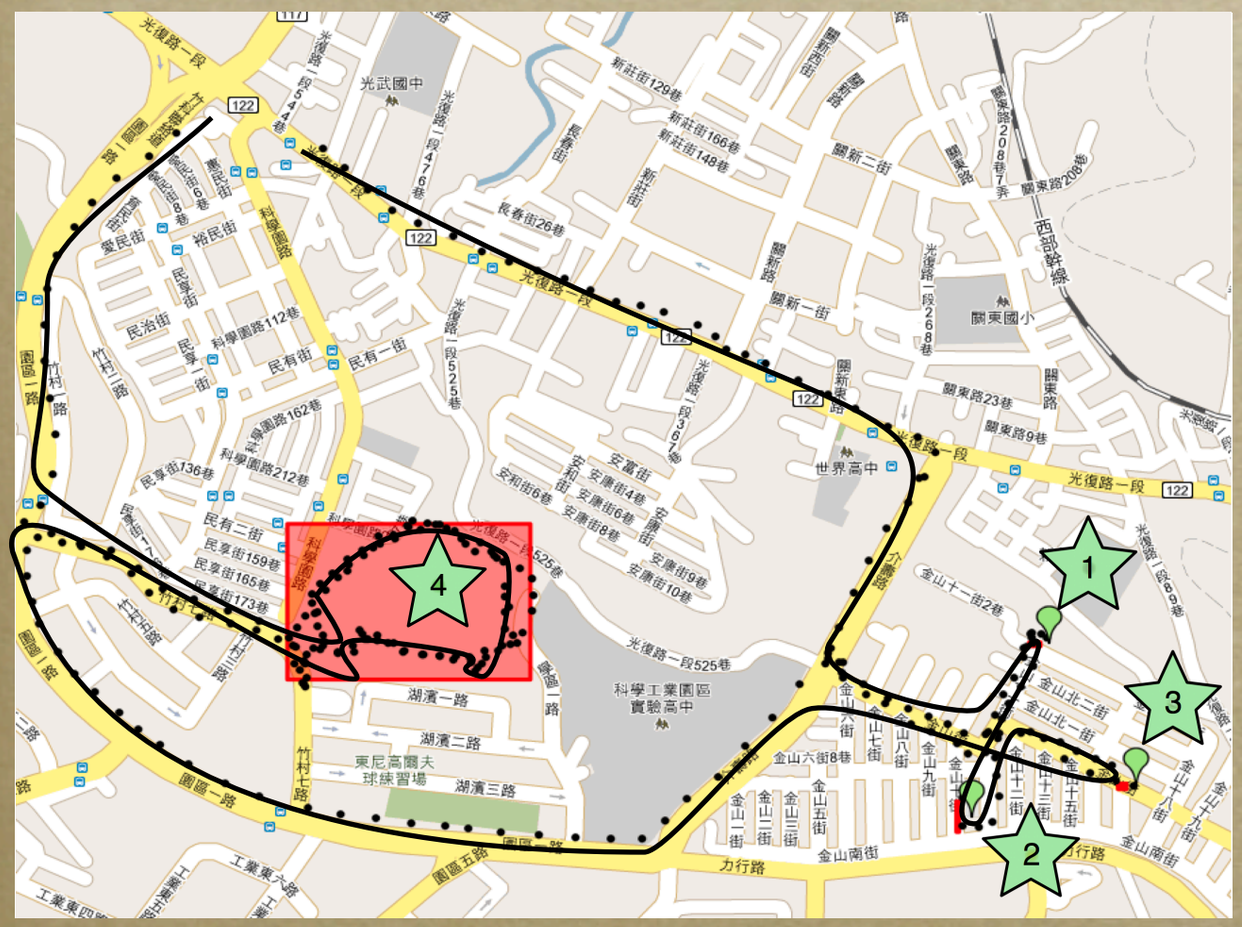
T: 10 mins, D: 200 meters



T: 10 mins, D: 250 meters

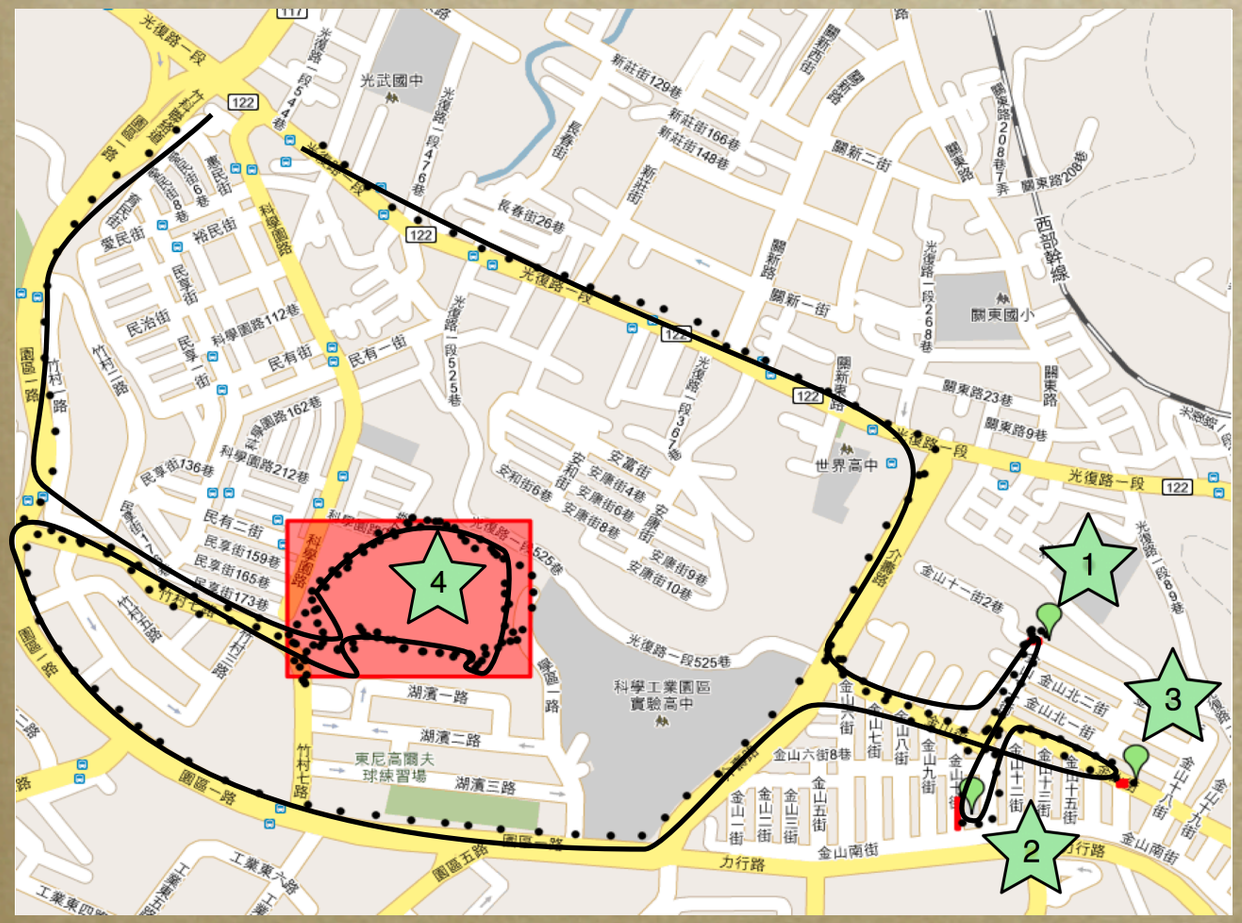
Zheng, et. at. "Mining Interesting Locations and Travel Sequences From GPS Trajectories." ,WWW (2009)

Issue: Region Extraction (2/2)



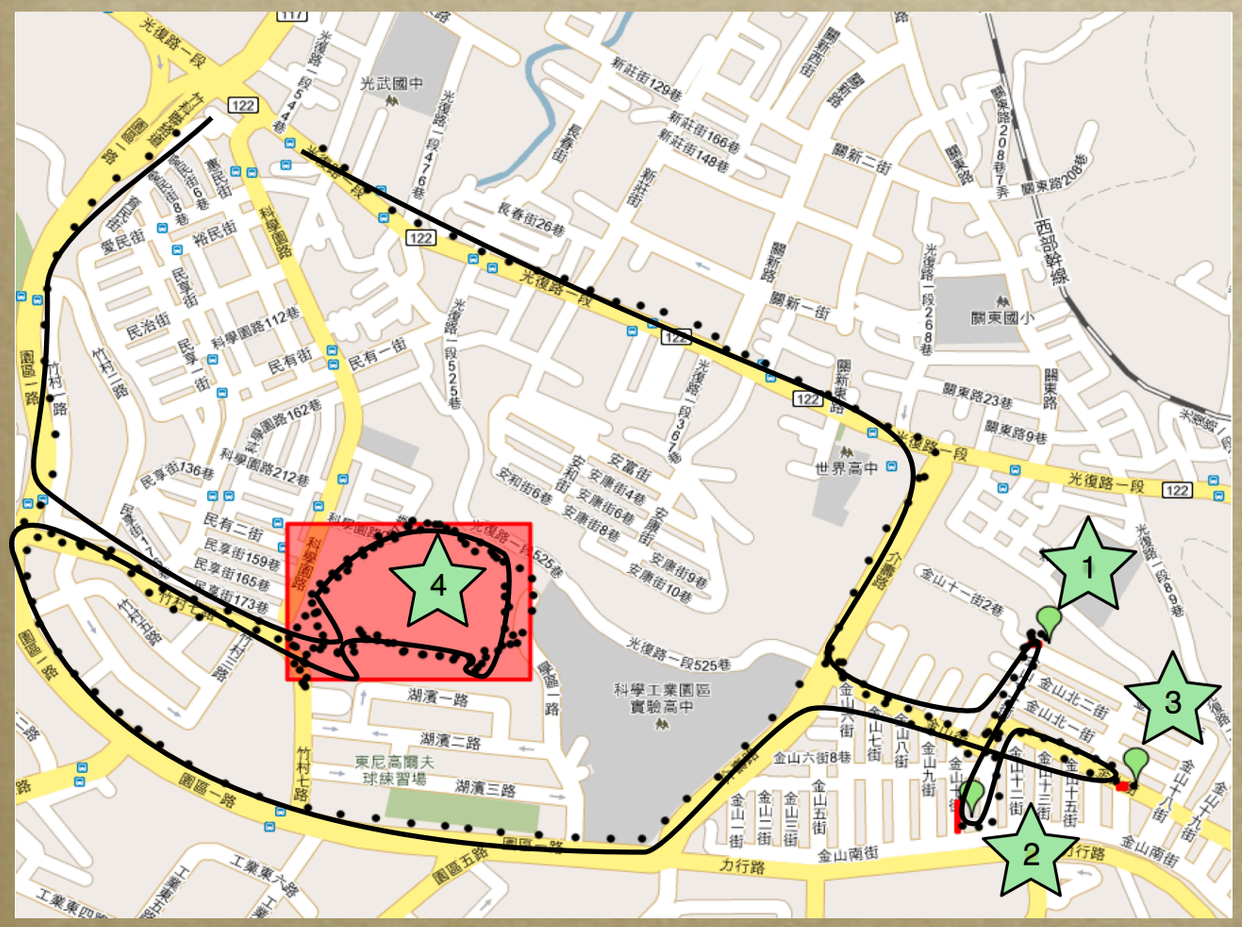
Issue: Region Extraction (2/2)

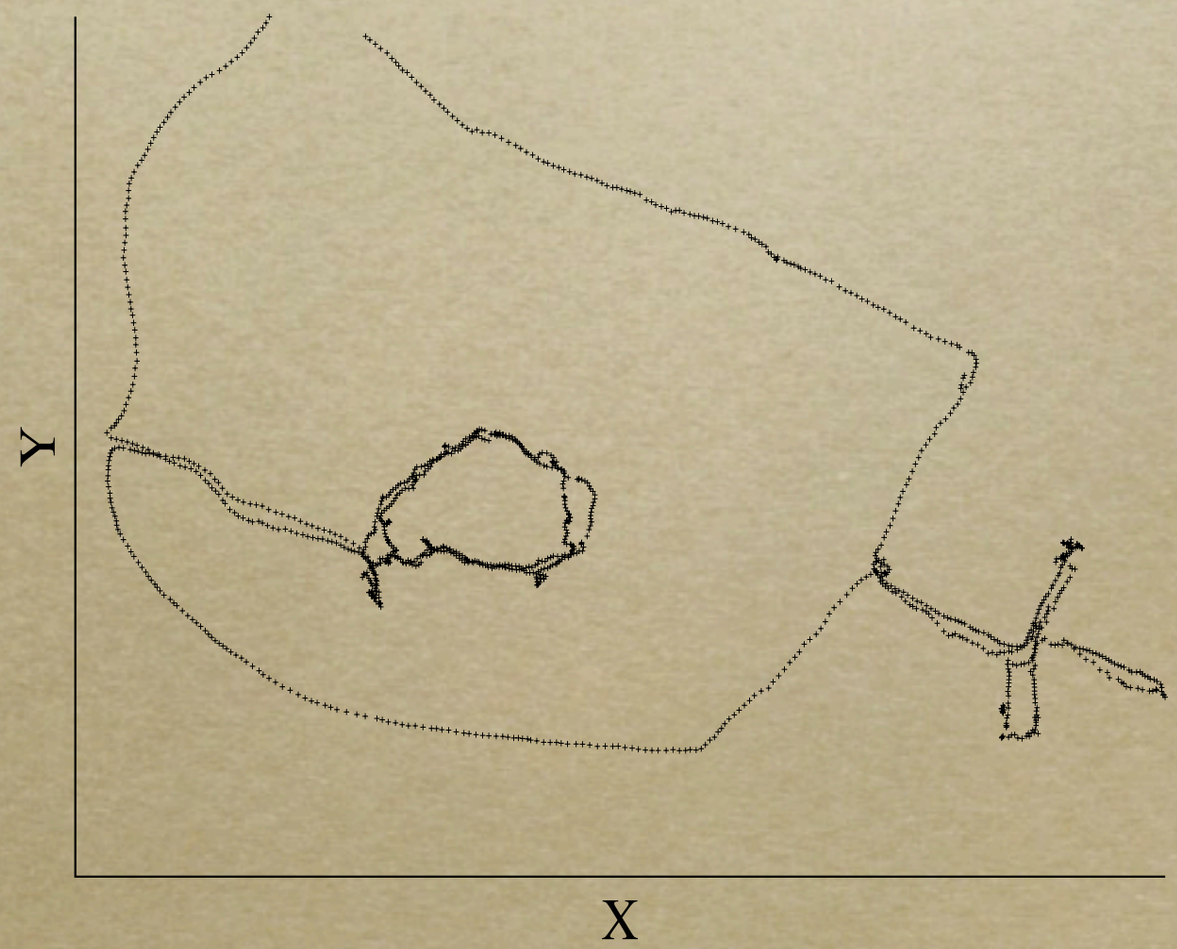
- Various time gap



Issue: Region Extraction (2/2)

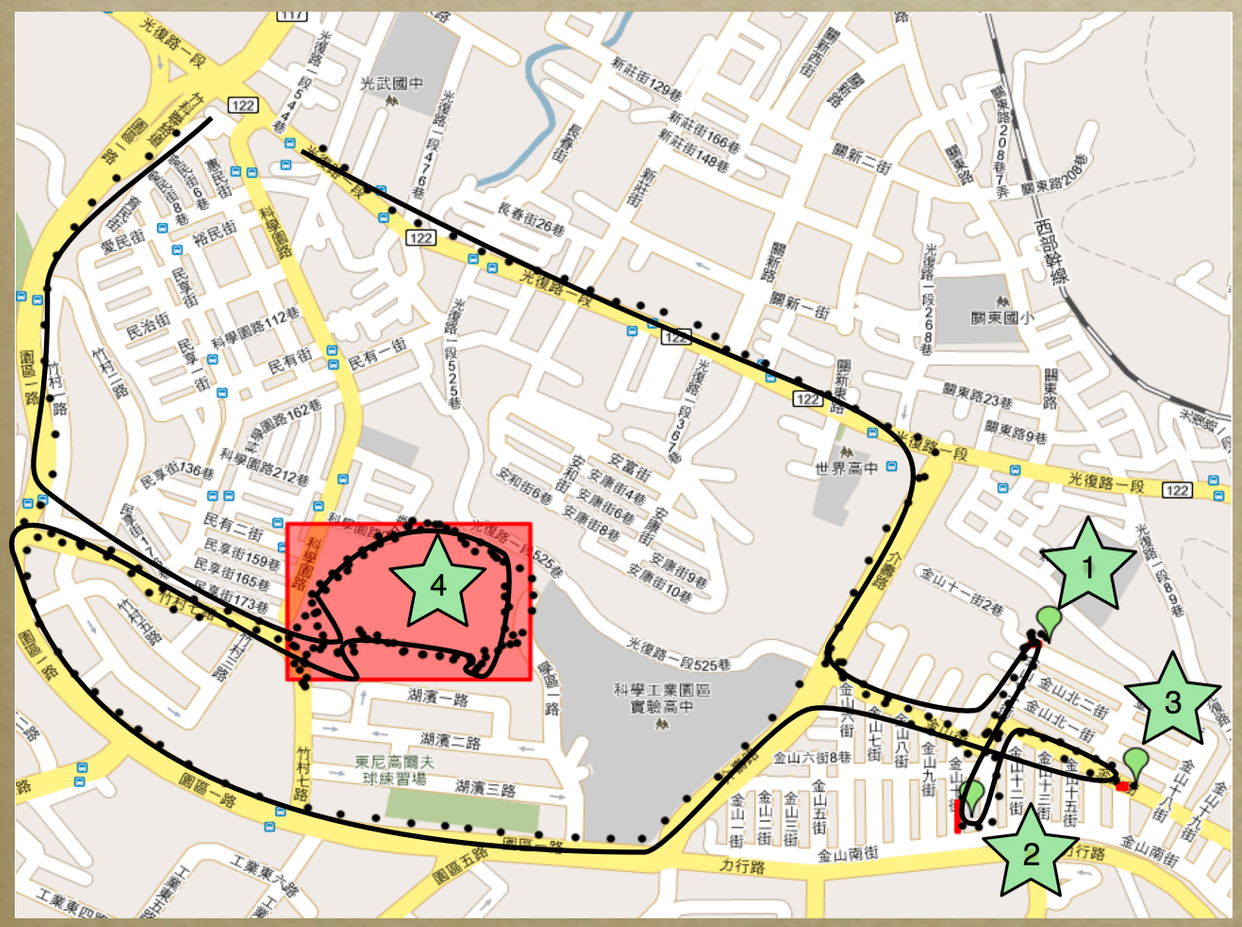
- Various time gap

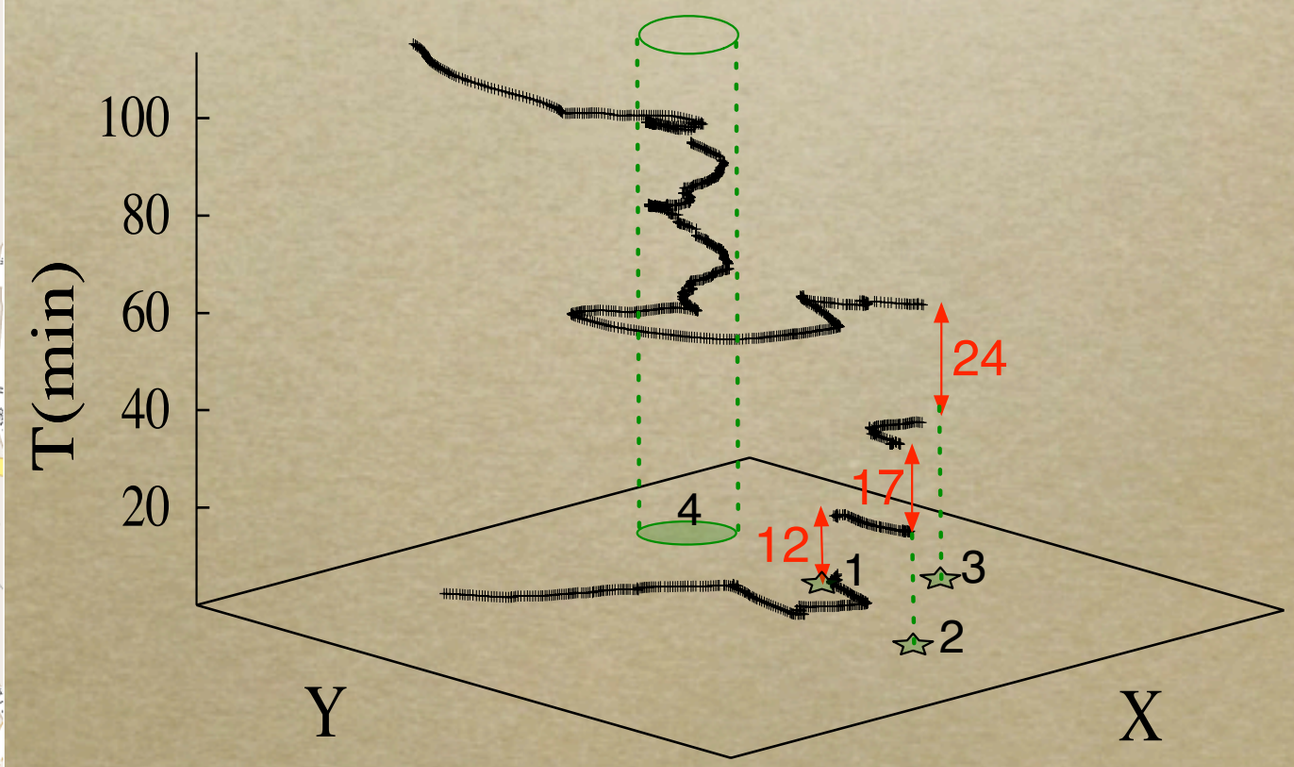




Issue: Region Extraction (2/2)

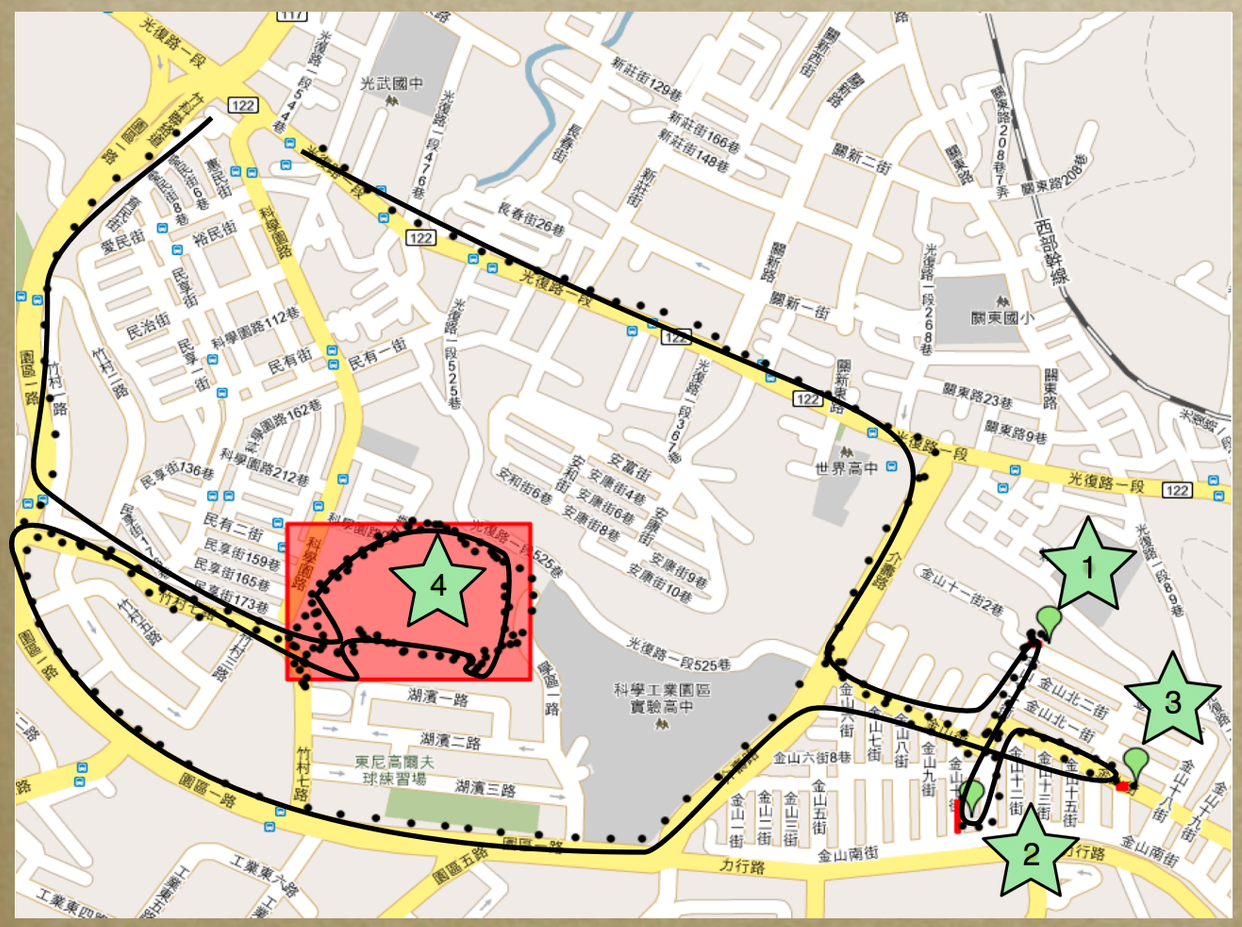
- Various time gap

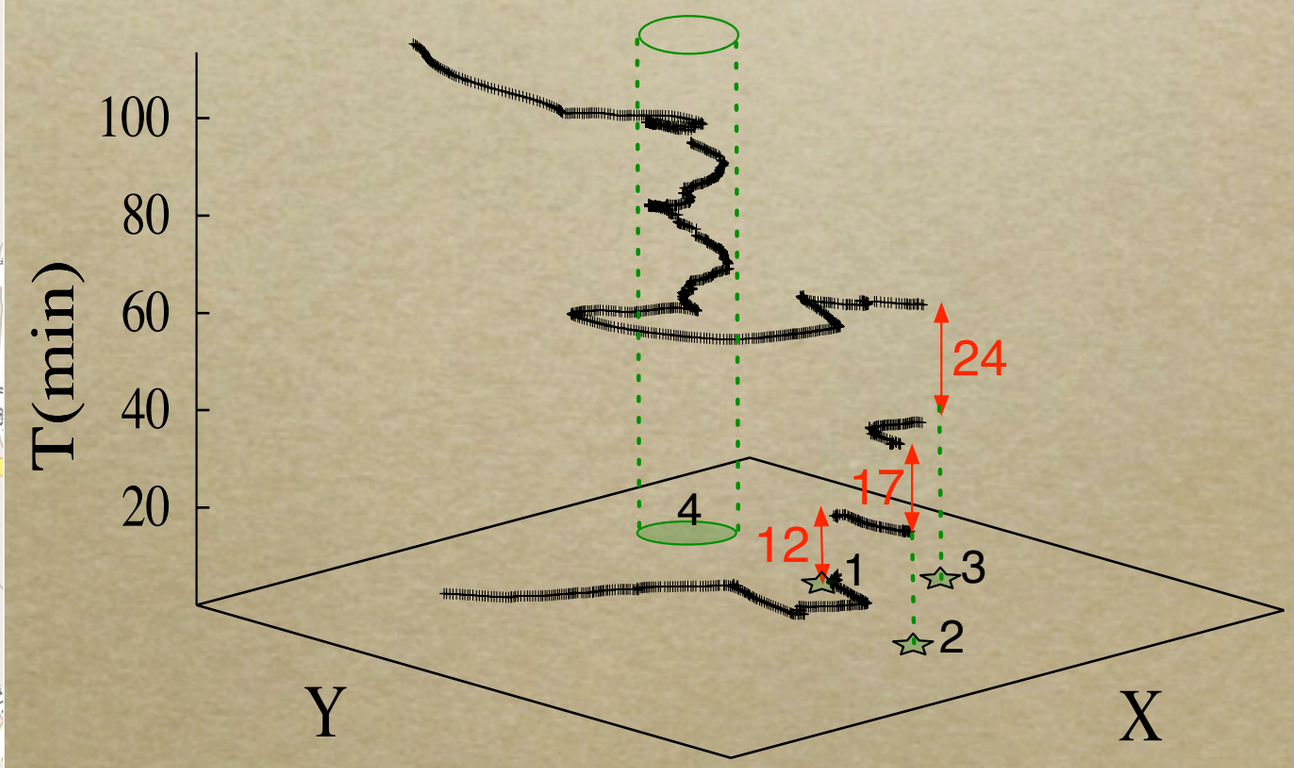




Issue: Region Extraction (2/2)

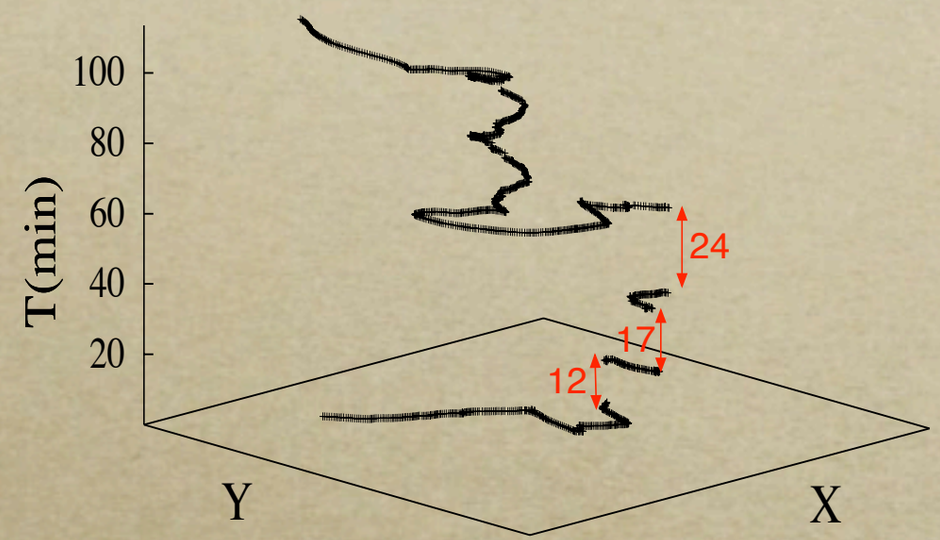
- Various time gap
- Different region size



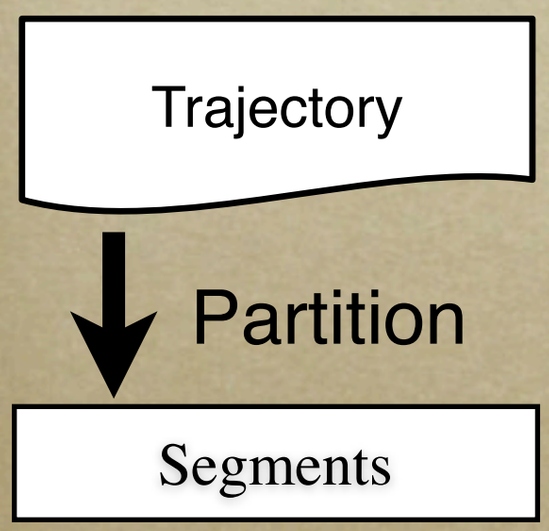


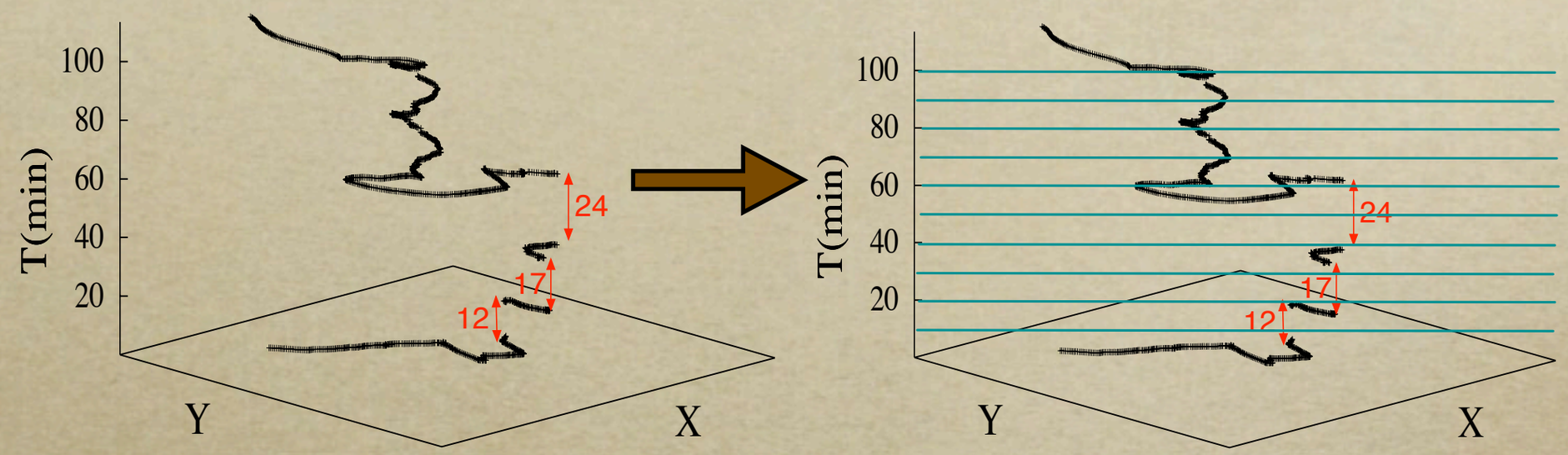
Regions of Interest Extraction - Sequential Density Clustering

Trajectory

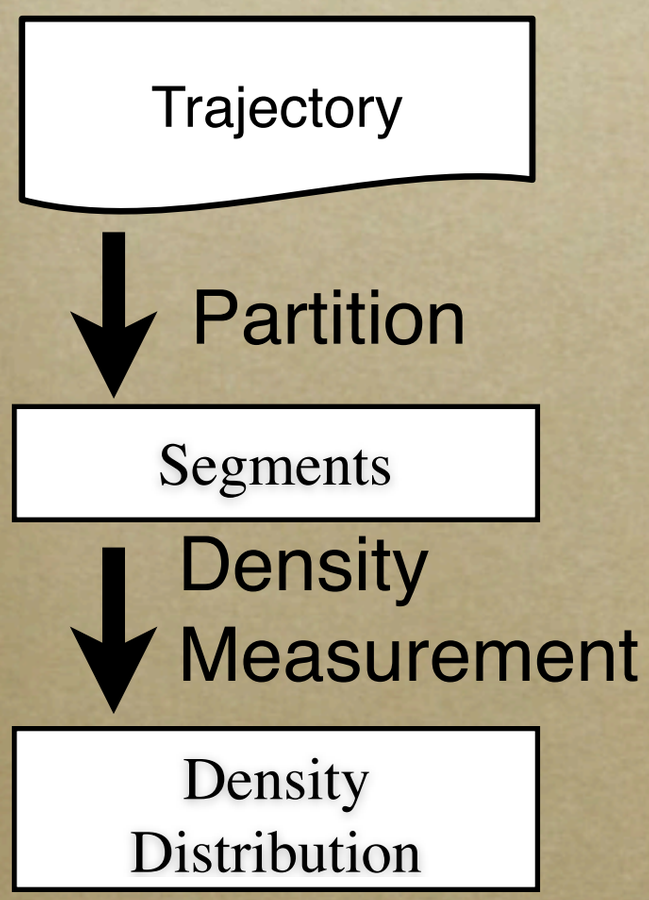


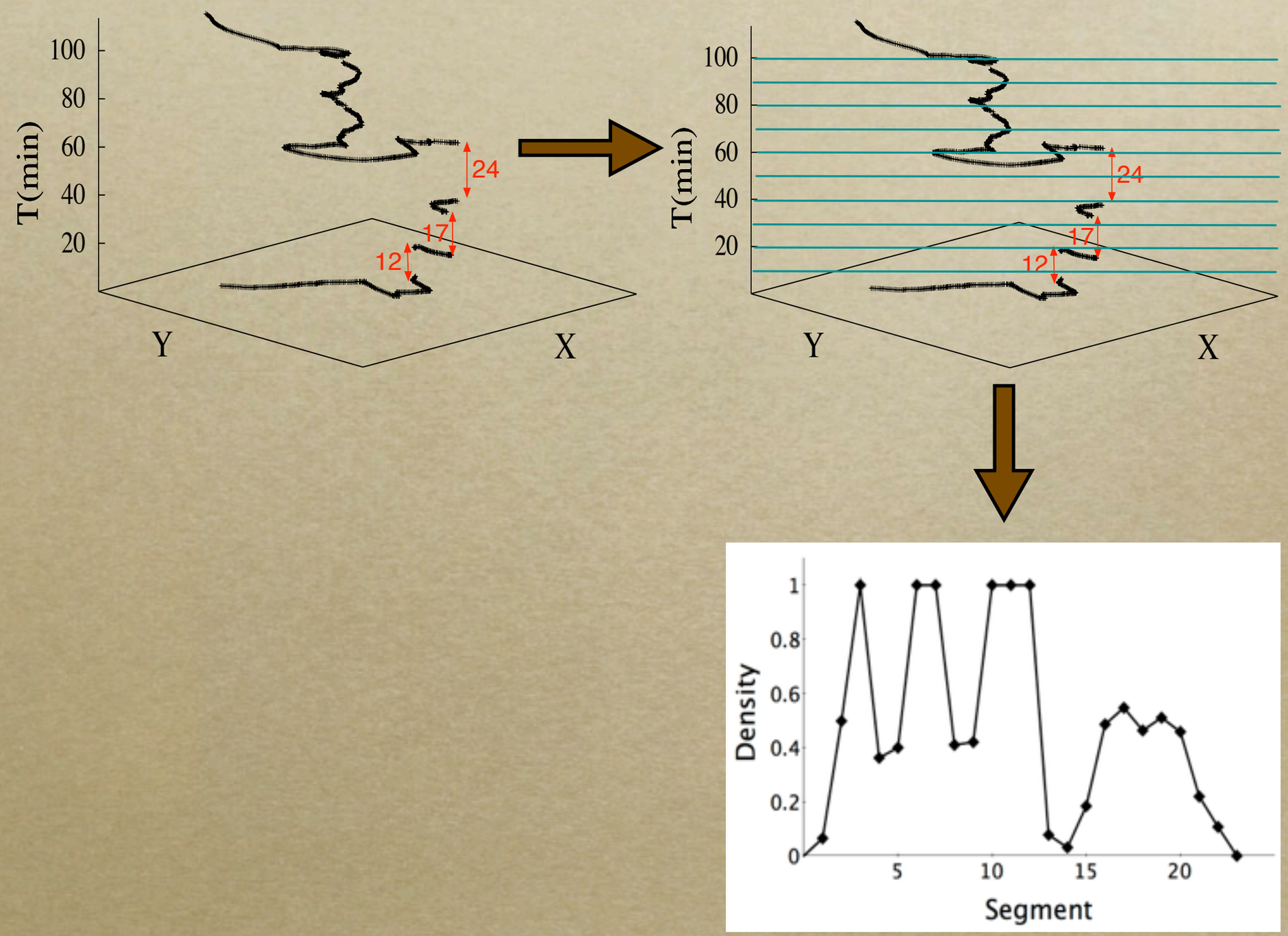
Regions of Interest Extraction - Sequential Density Clustering



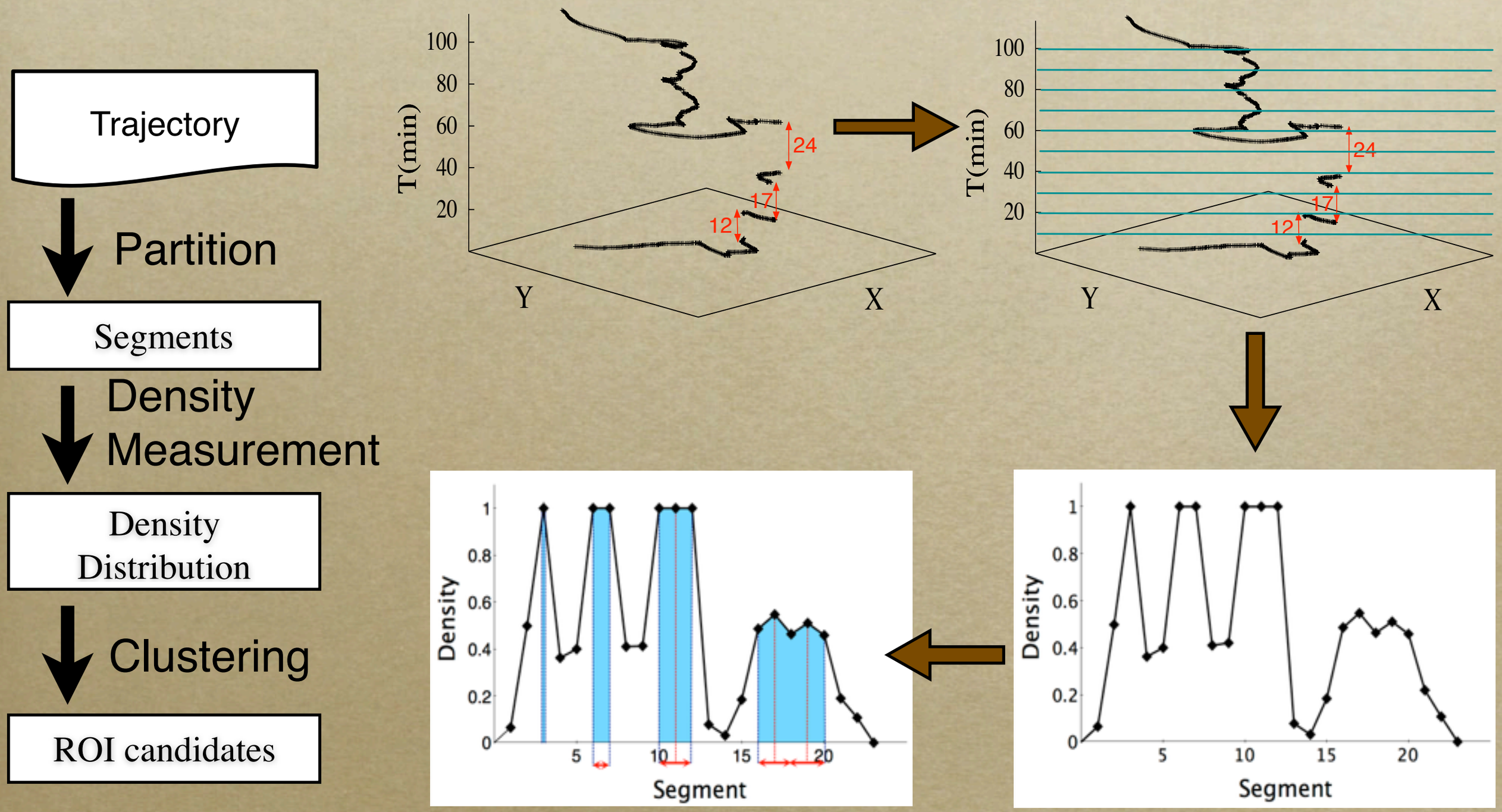


Regions of Interest Extraction - Sequential Density Clustering





Regions of Interest Extraction - Sequential Density Clustering



Regions of Interest Extraction - Sequential Density Clustering

Trajectory

↓ Partition

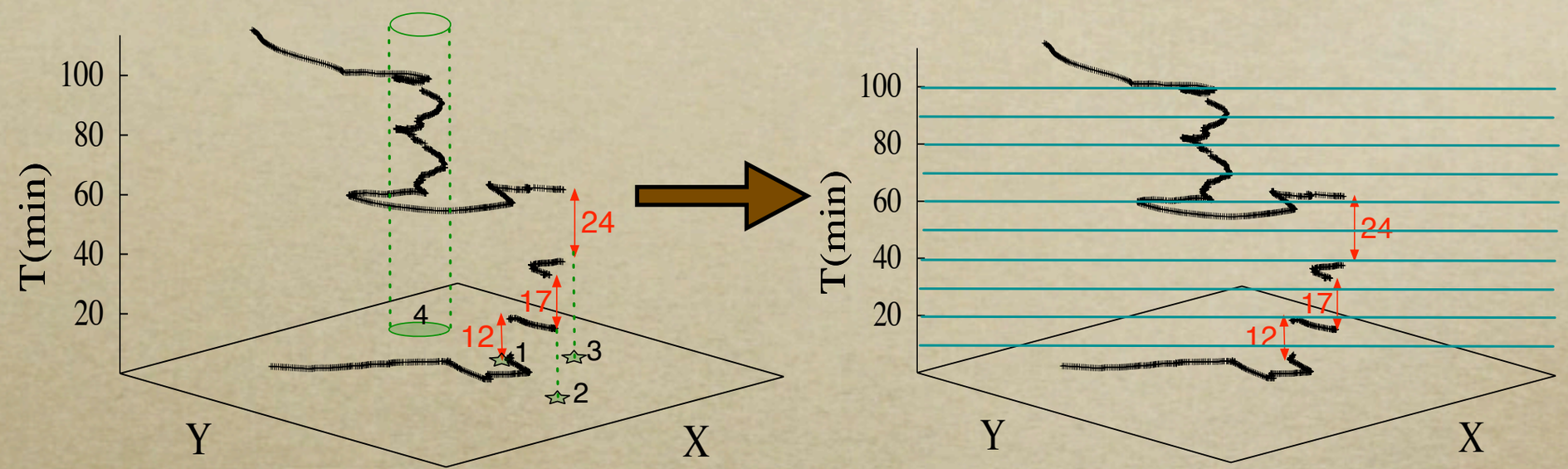
Segments

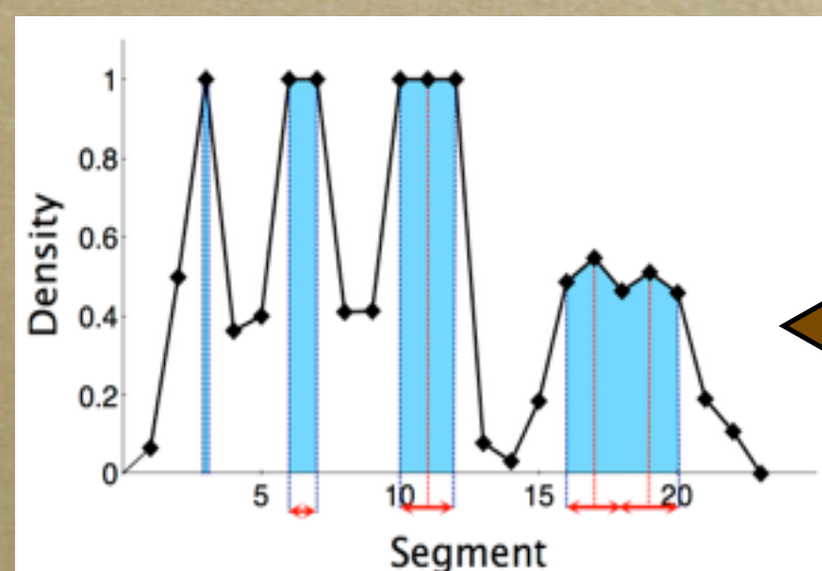
↓ Density Measurement

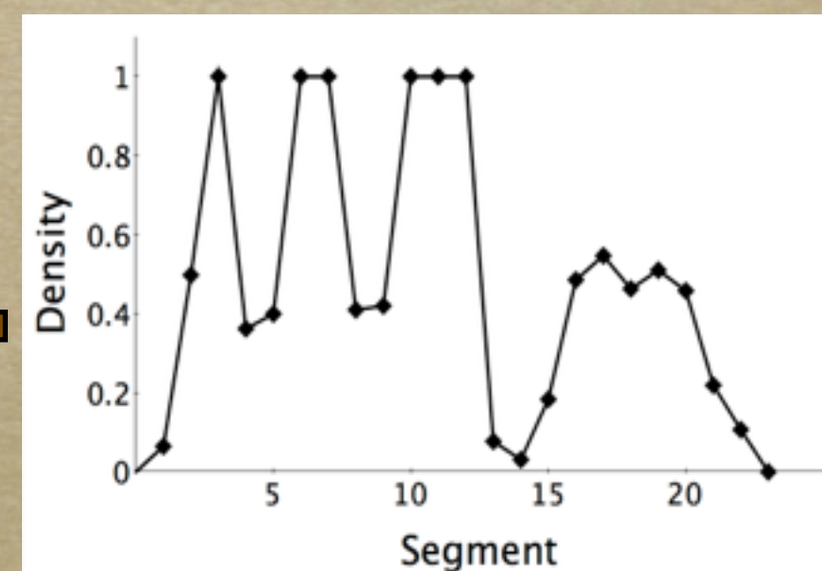
Density Distribution

↓ Clustering

ROI candidates

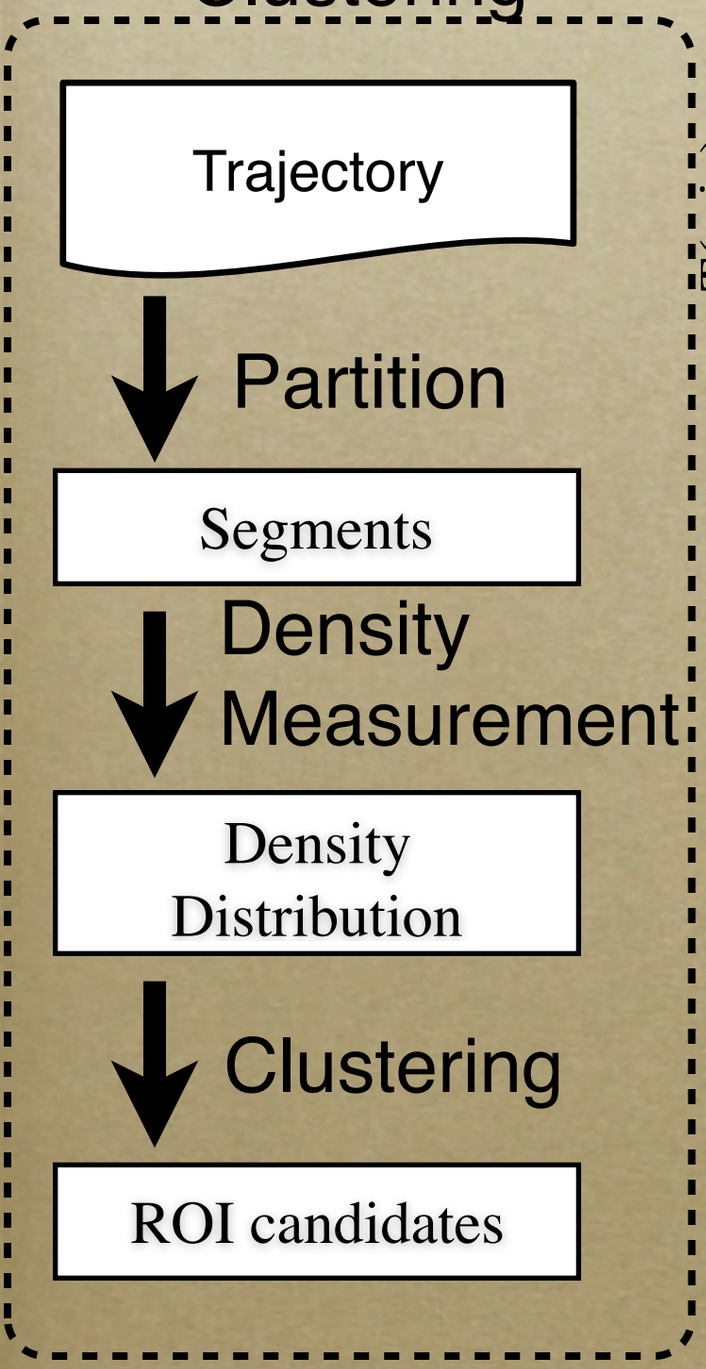


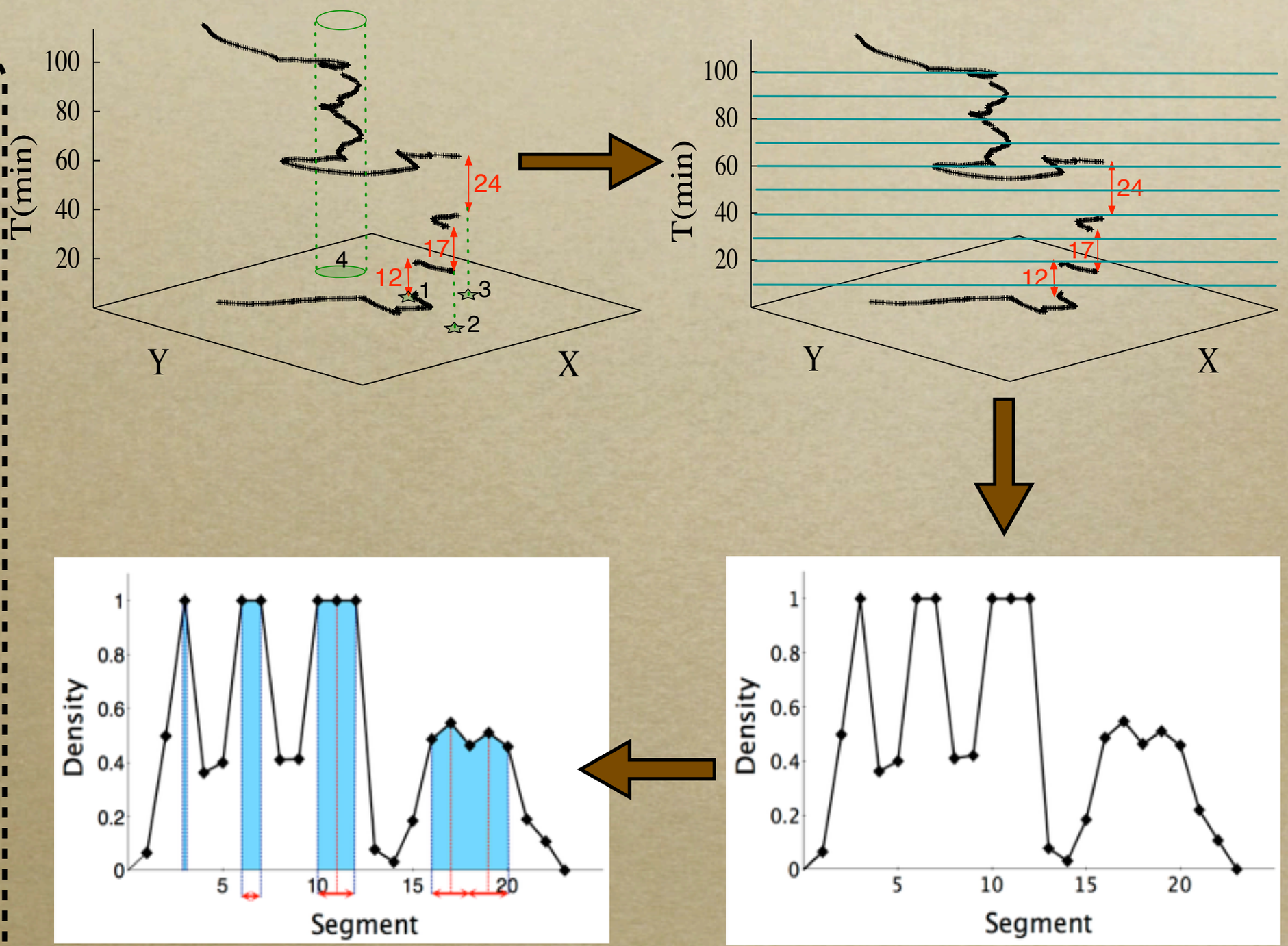




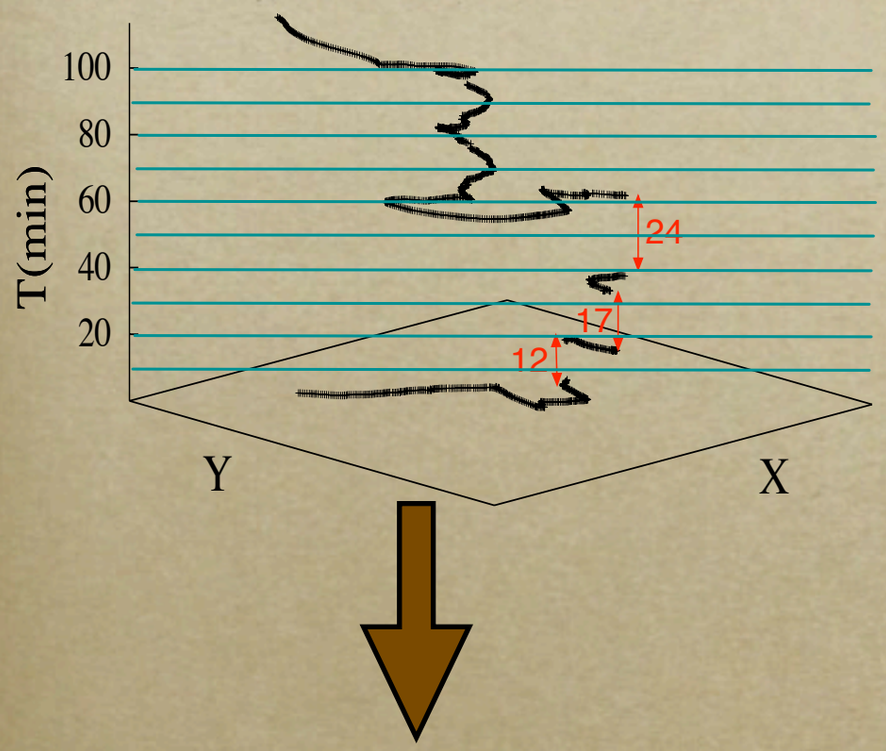
Regions of Interest Extraction - Sequential Density Clustering

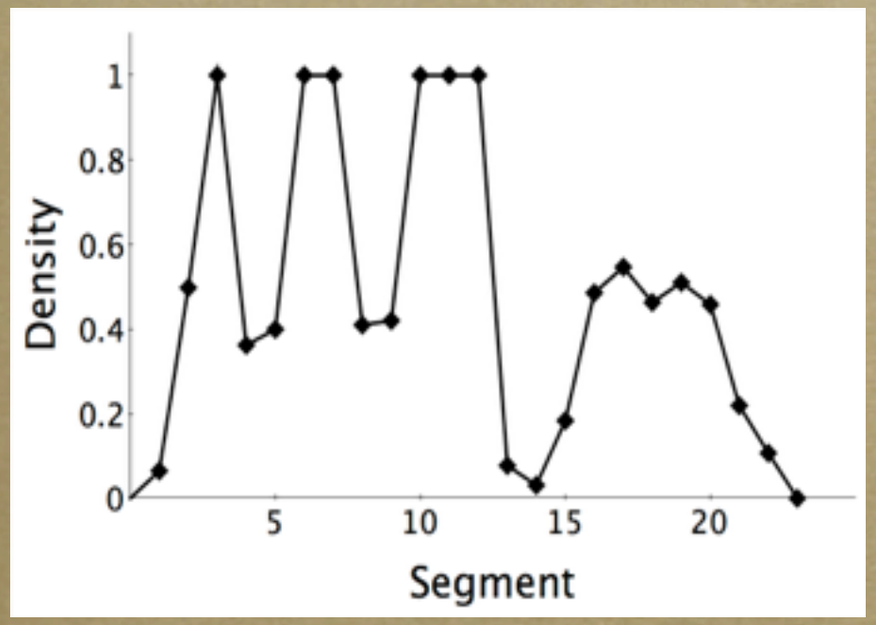
Sequential Density Clustering





Sequential Density Clustering: Spatiotemporal Density

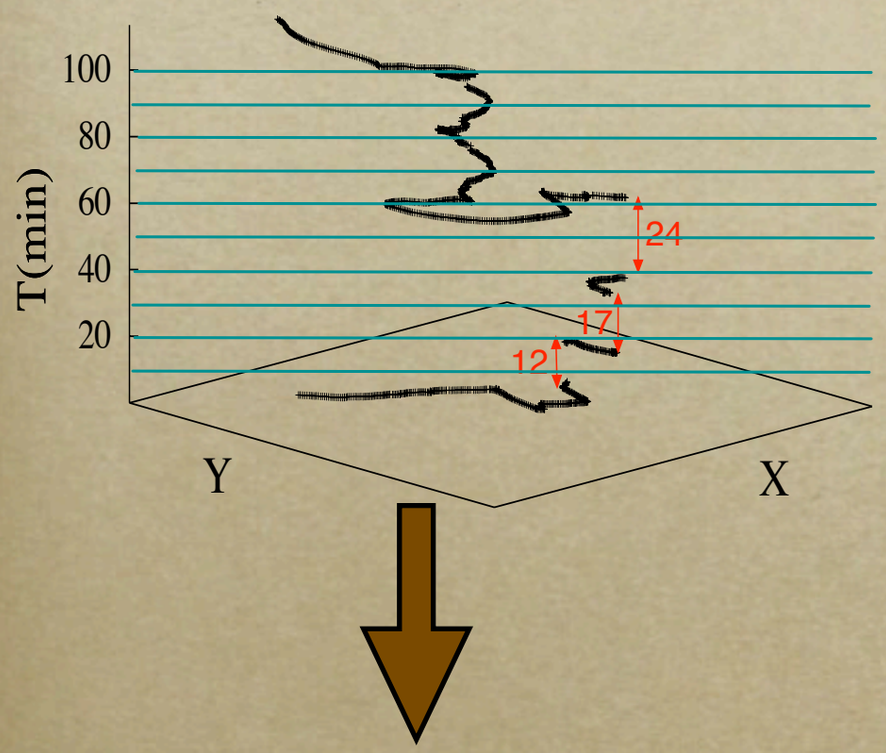


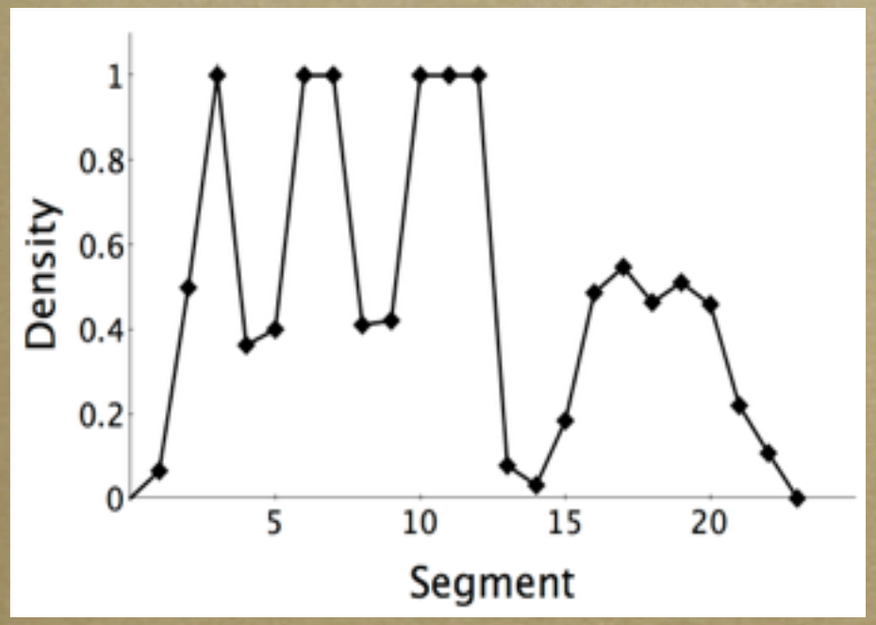


Spatial dissimilarity:

$$Cost(S_l) = \sum_{i=1}^{n} D_{E^2}(o_i, c) = \sum_{i=1}^{n} \| o_i - c \|^2$$

Sequential Density Clustering: Spatiotemporal Density





Take time gap as a weight

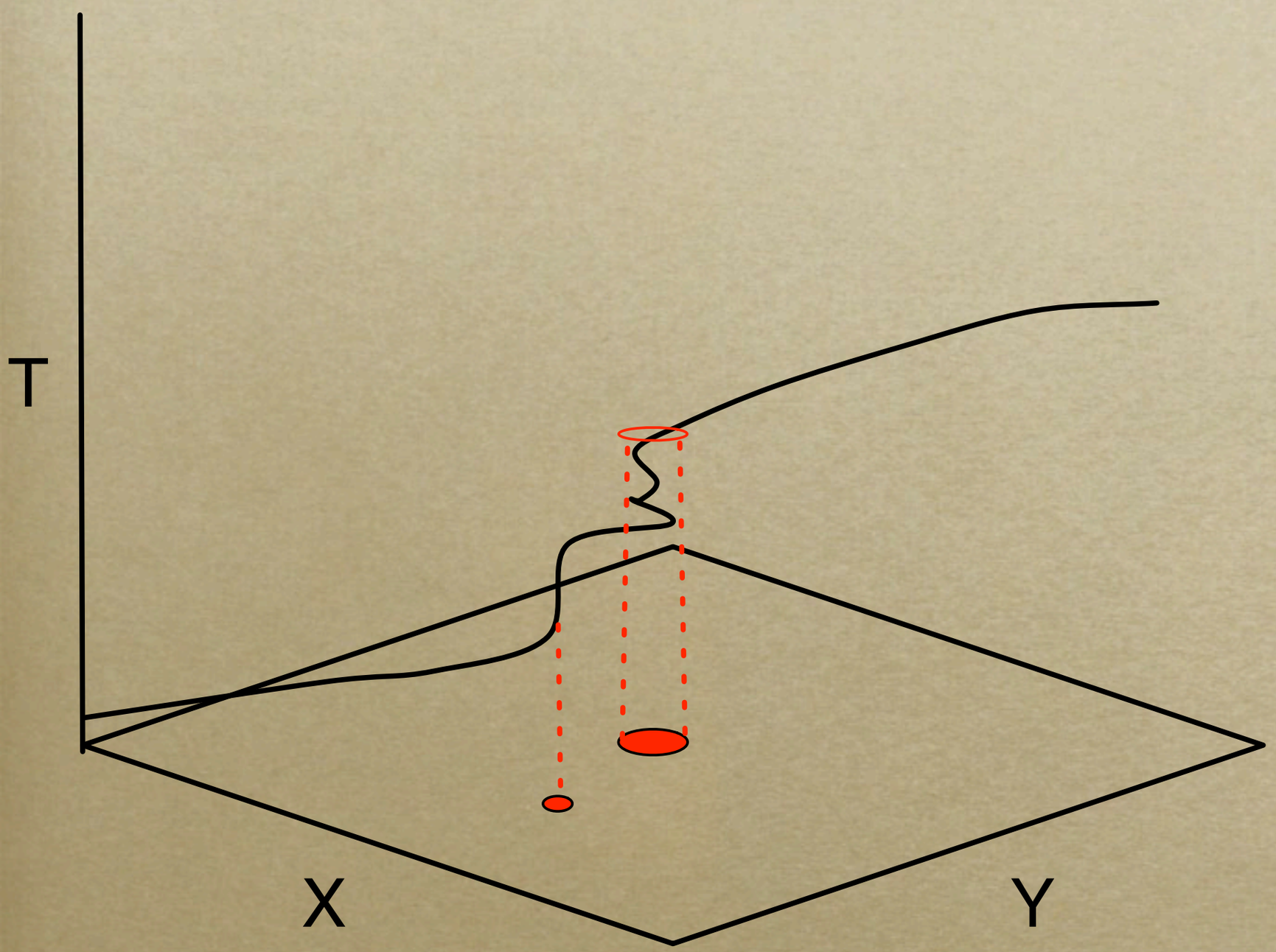
Spatial dissimilarity:

$$Cost(S_l) = \sum_{i=1}^{n} D_{E^2}(o_i, c) = \sum_{i=1}^{n} \| o_i - c \|^2$$

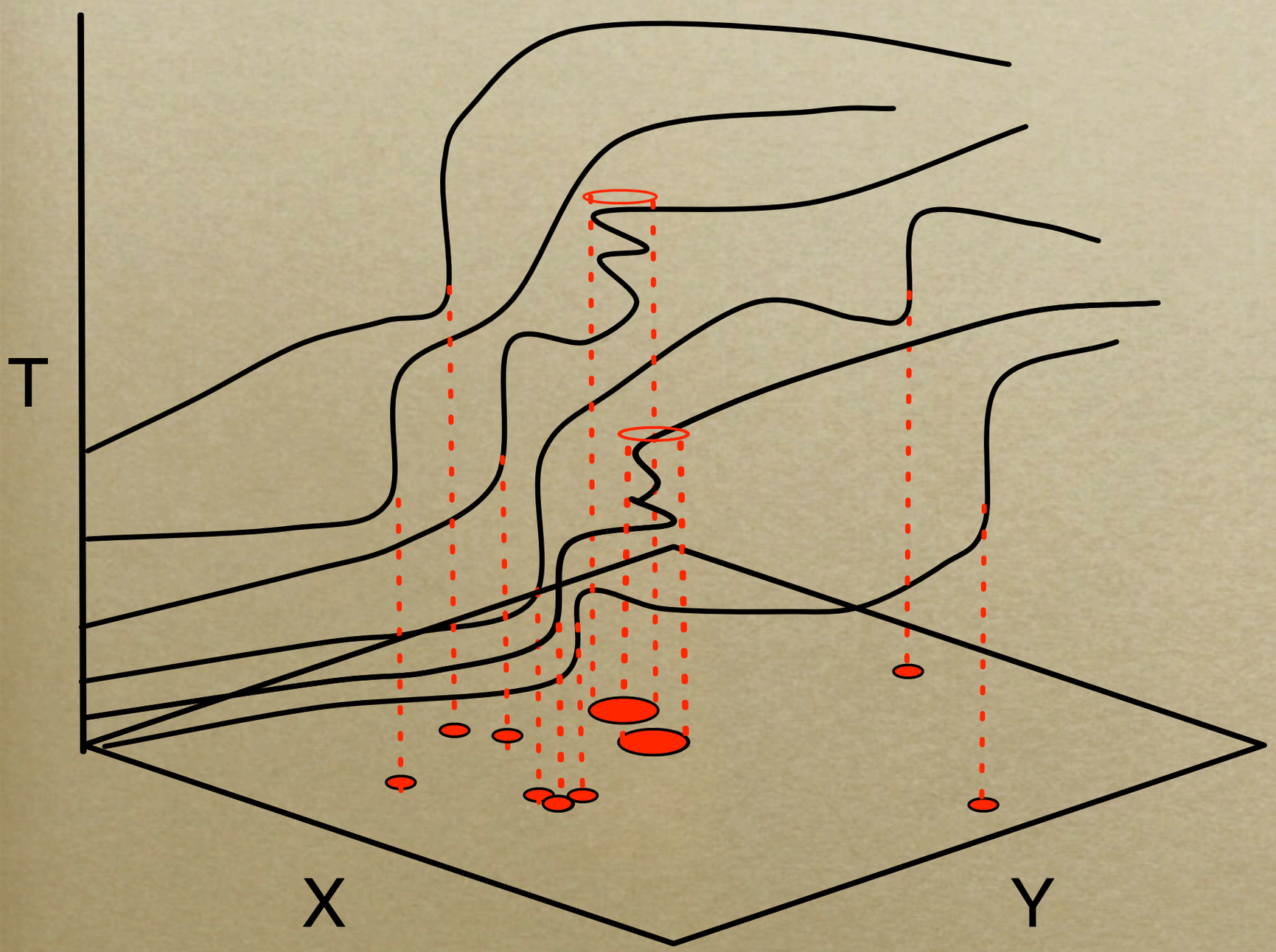
$$w_k = \frac{(t_k - t_{k-1}) + (t_{k+1} - t_k)}{2}$$

$$Density(s_l) = \log(1 + \frac{1}{WCost(S_l) + \gamma})$$

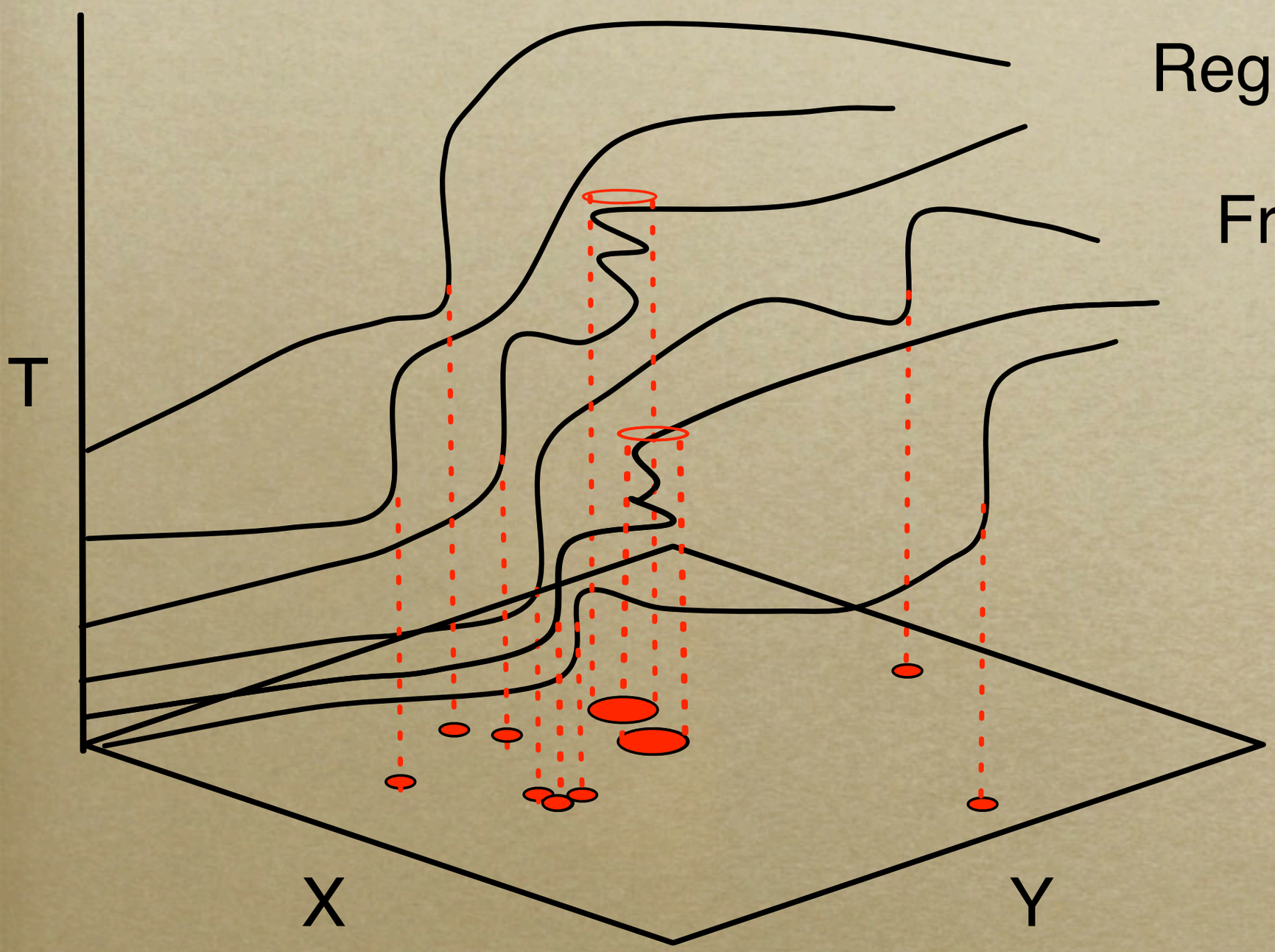
Regions of Interest Extraction - Region Clustering



Regions of Interest Extraction - Region Clustering



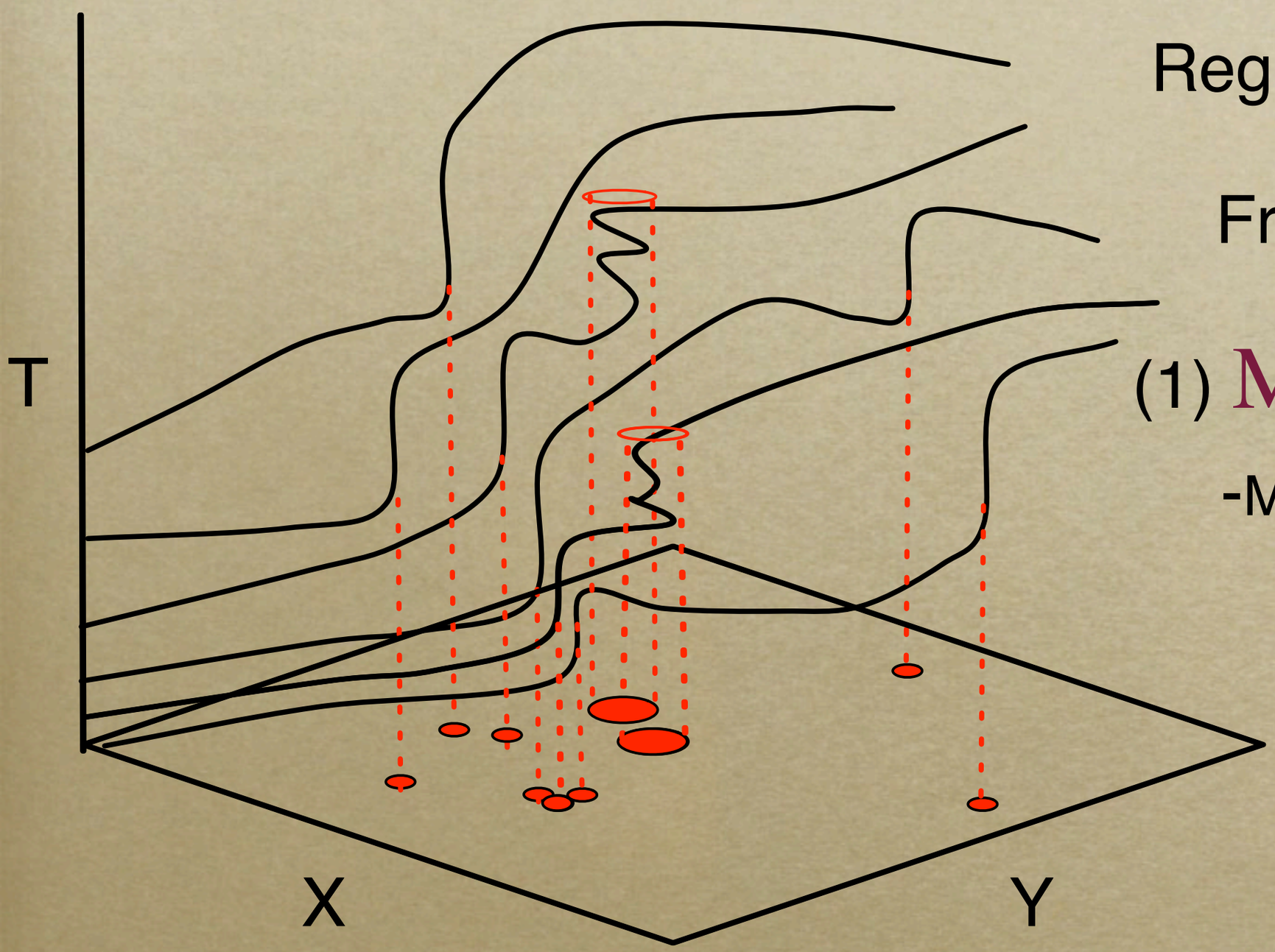
Regions of Interest Extraction - Region Clustering



Region of Interests (ROI)

Frequently stay

Regions of Interest Extraction - Region Clustering



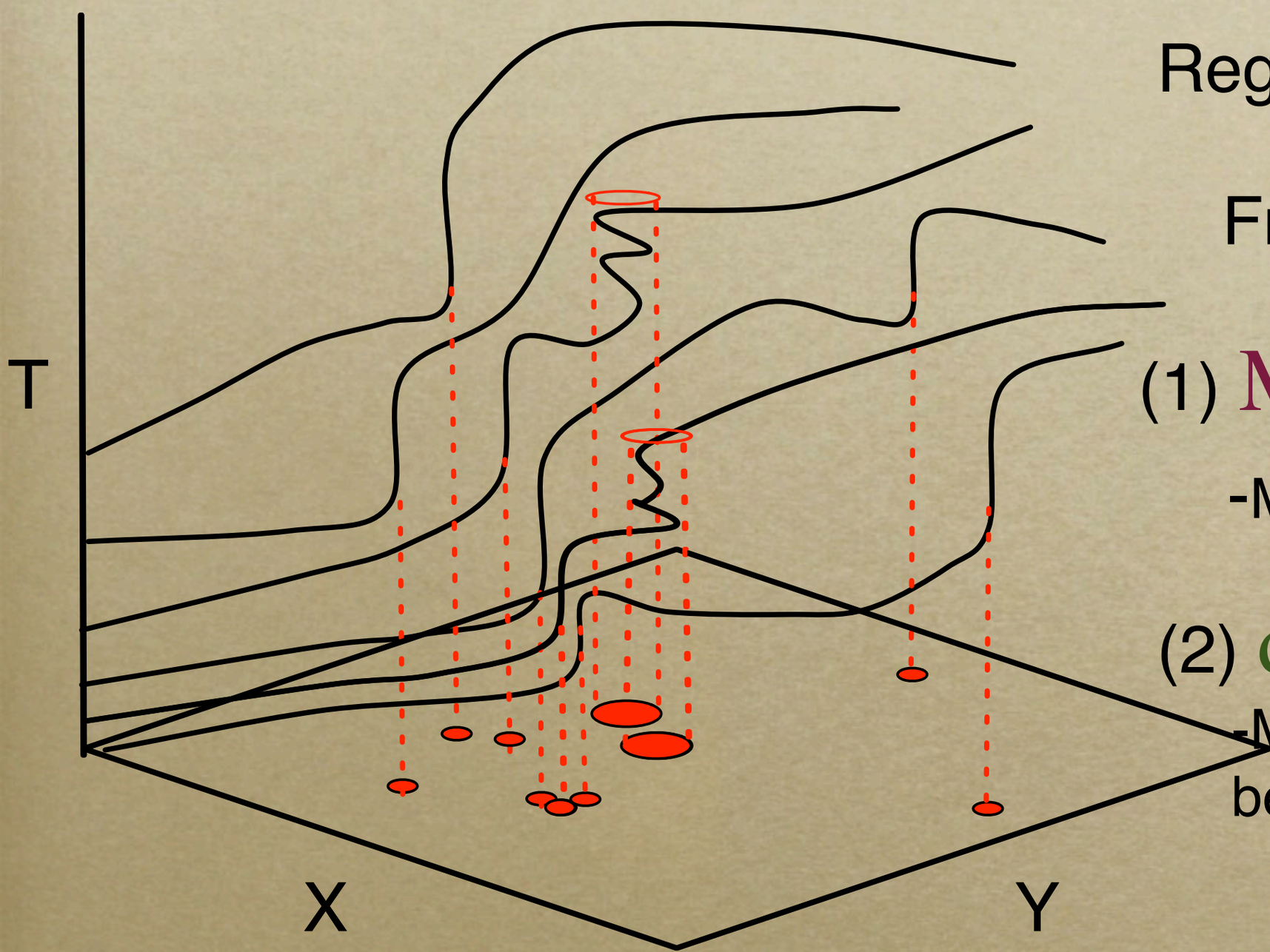
Region of Interests (ROI)

Frequently stay

(1) MinSR

-Minimum number of stay

Regions of Interest Extraction - Region Clustering



Region of Interests (ROI)

Frequently stay

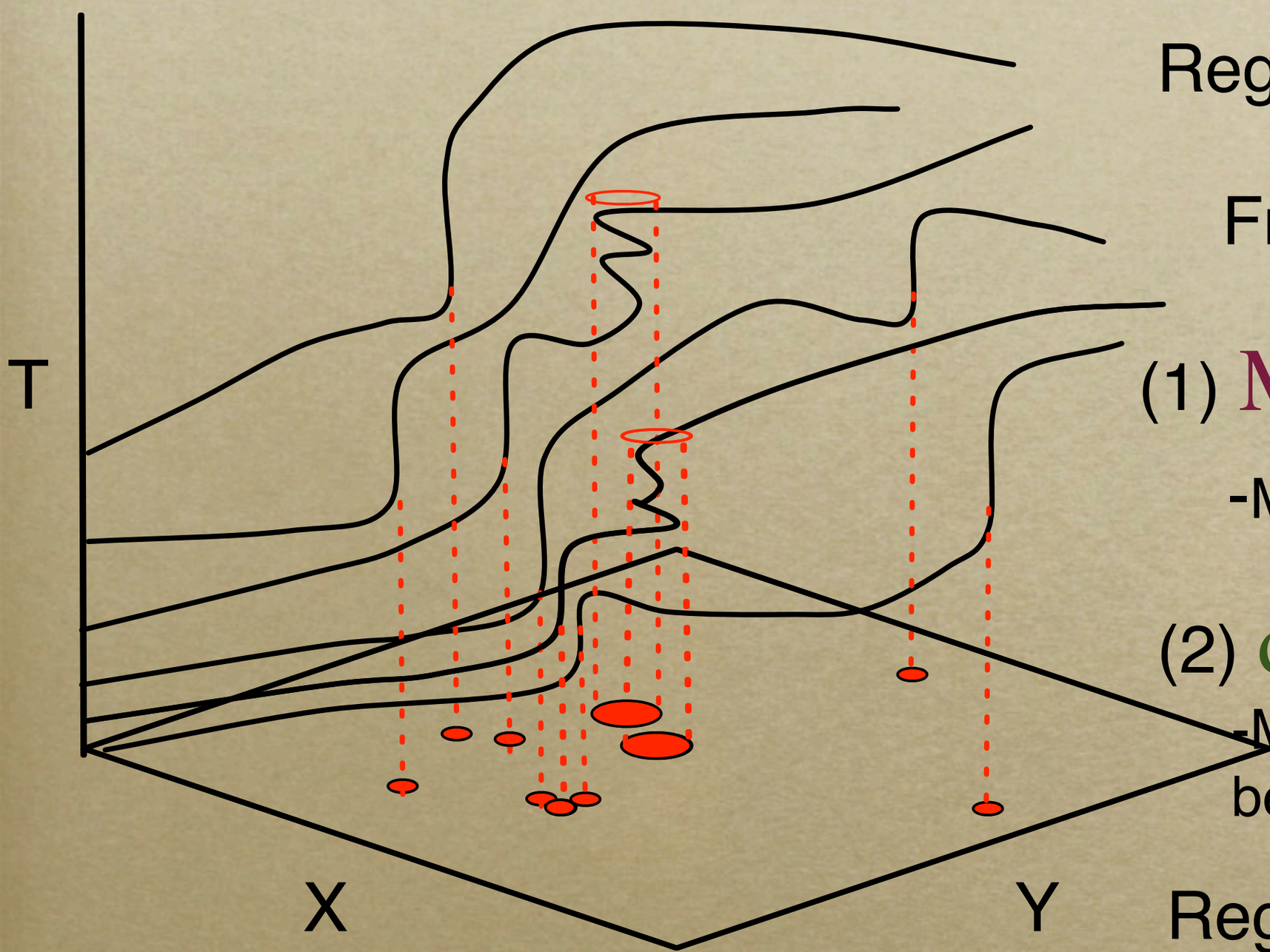
(1) MinSR

-Minimum number of stay

(2) d_h

-Maximum distance between stays

Regions of Interest Extraction - Region Clustering



Region of Interests (ROI)

Frequently stay

(1) MinSR

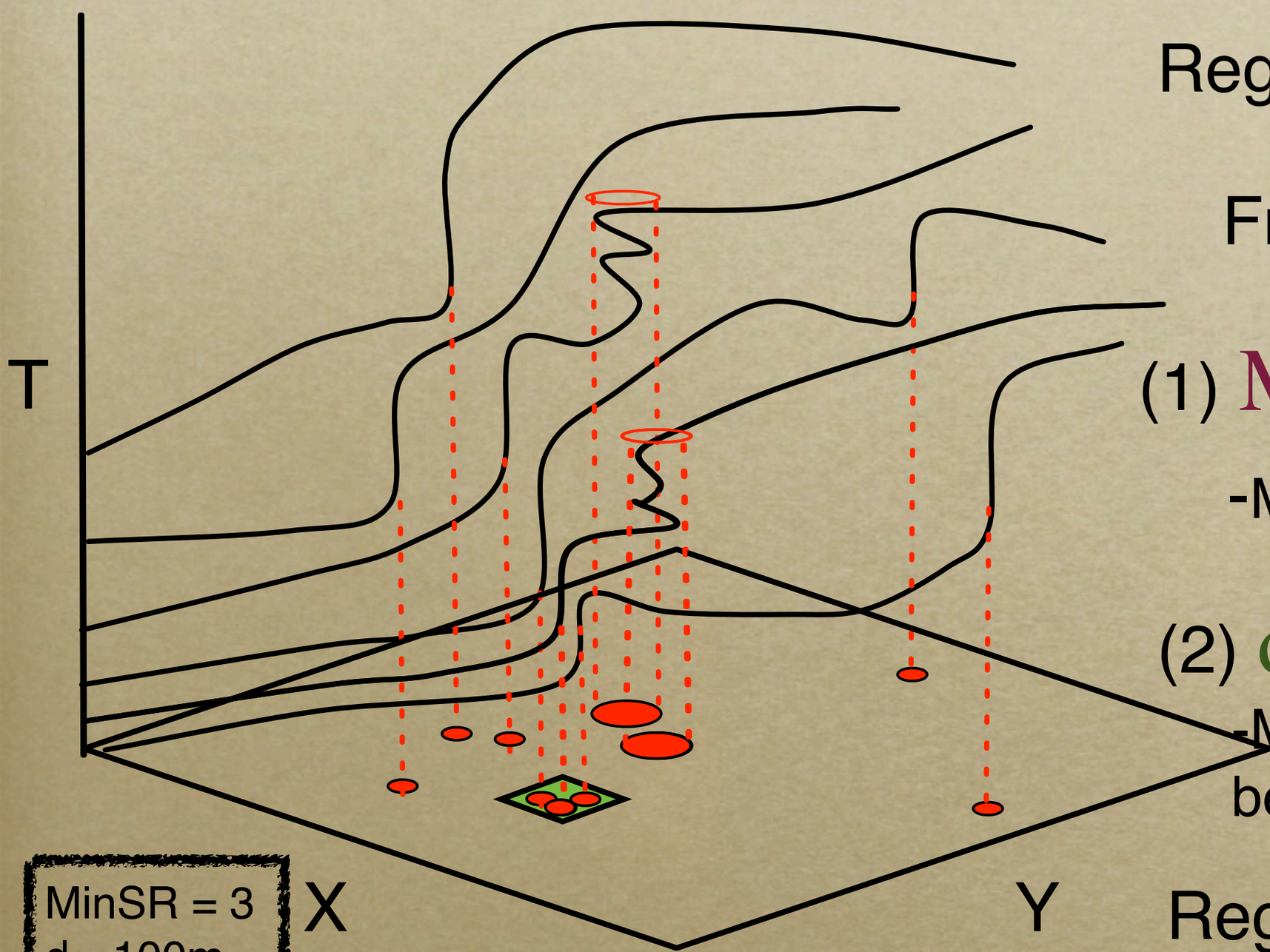
-Minimum number of stay

(2) d_h

-Maximum distance between stays

Region Clustering: SNN

Regions of Interest Extraction - Region Clustering



MinSR = 3
d_h=100m

Region of Interests (ROI)

Frequently stay

(1) MinSR

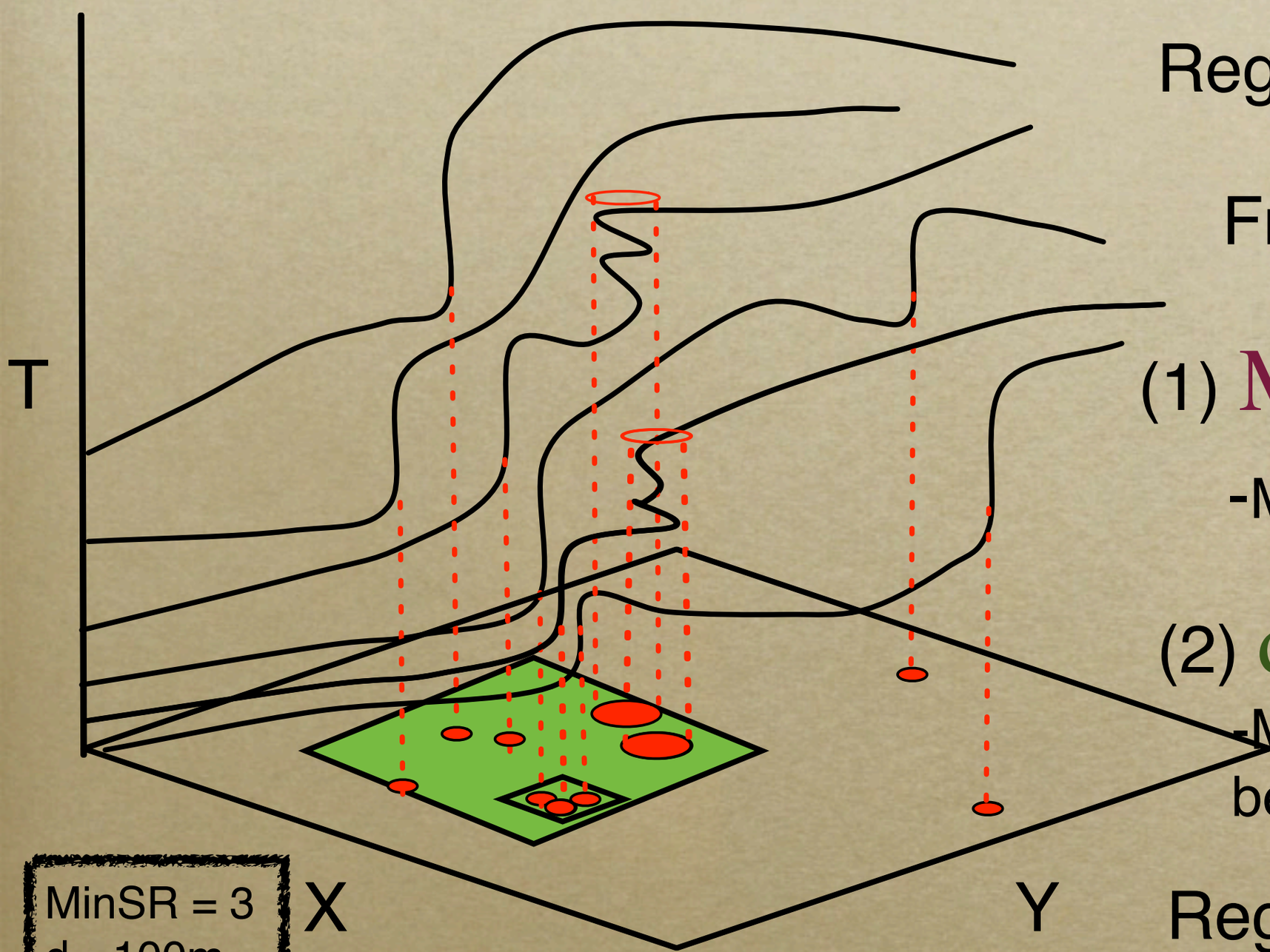
-Minimum number of stay

(2) d_h

-Maximum distance between stays

Region Clustering: SNN

Regions of Interest Extraction - Region Clustering



MinSR = 3
d_h=100m

Region of Interests (ROI)

Frequently stay

(1) MinSR

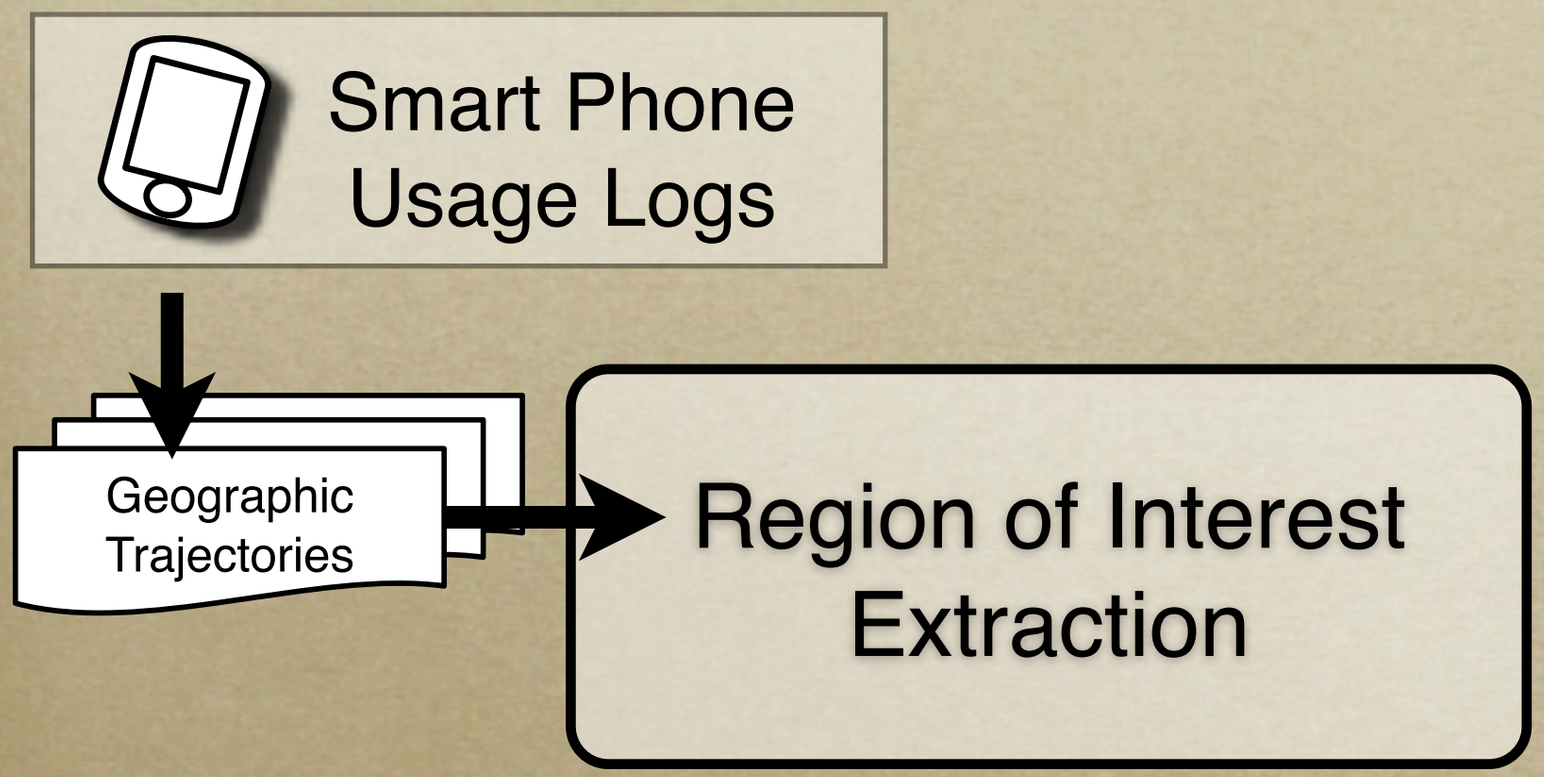
-Minimum number of stay

(2) d_h

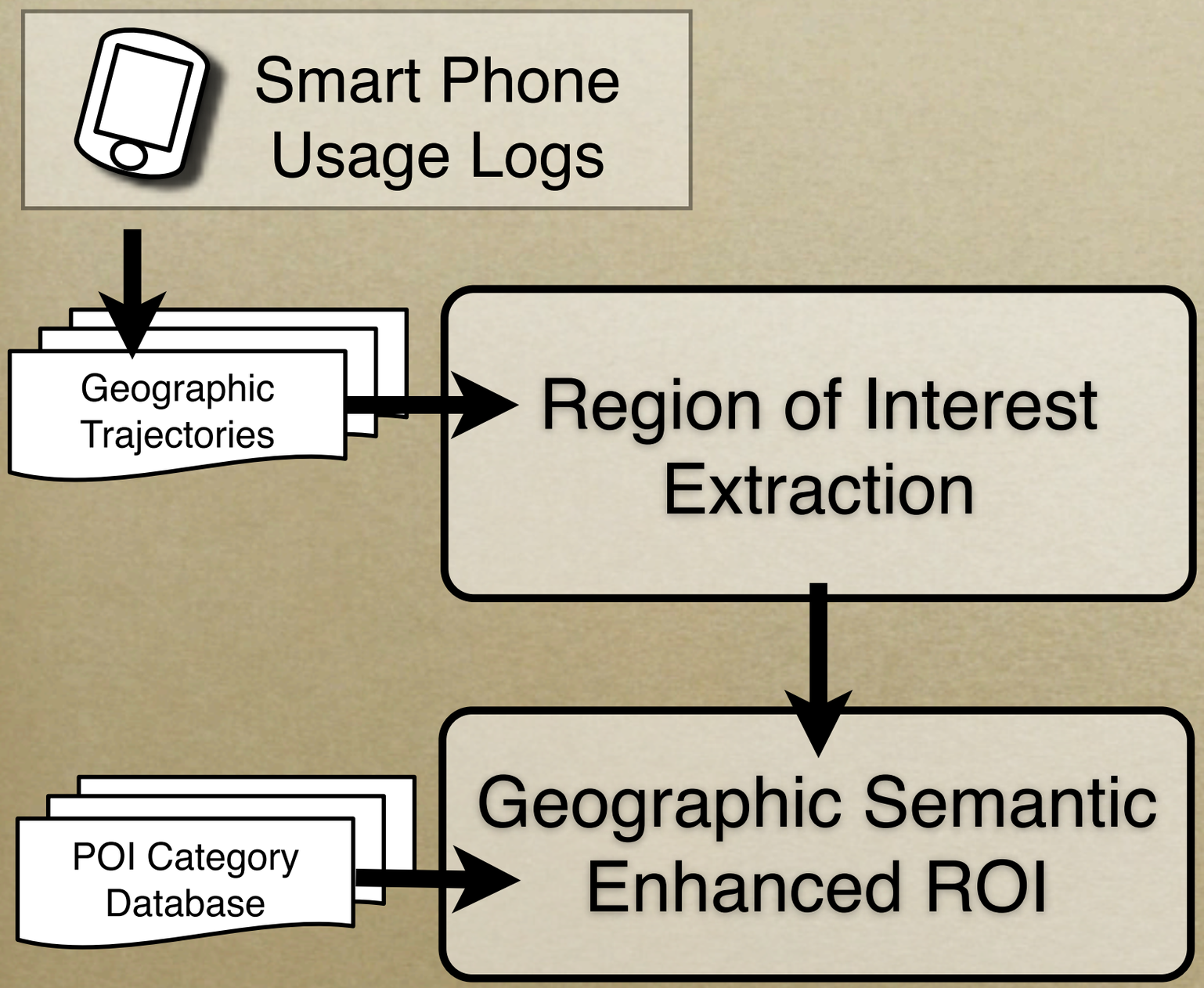
-Maximum distance between stays

Region Clustering: SNN

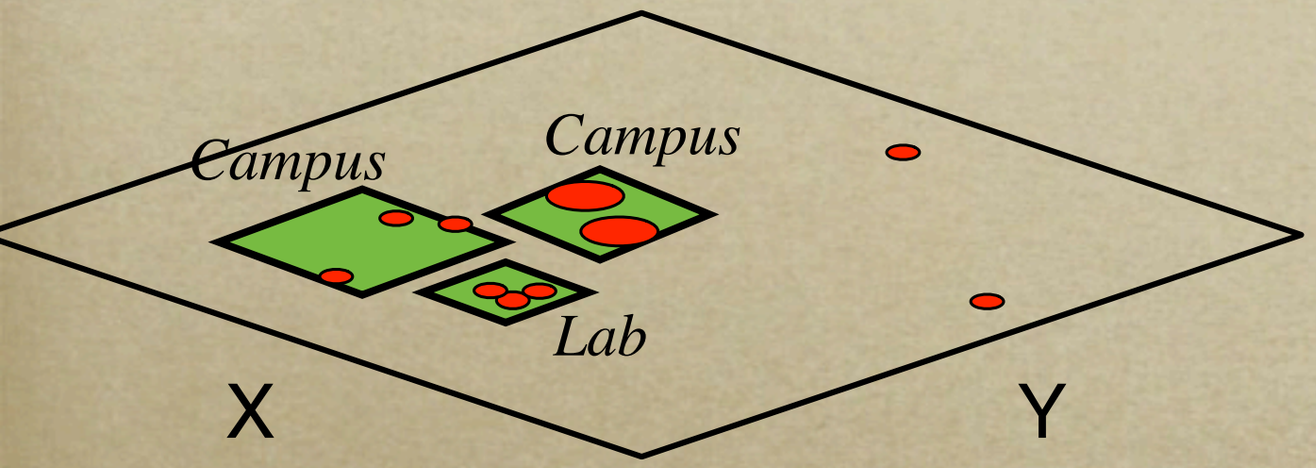
Framework


Smart Phone
Usage Logs
Geographic
Trajectories
Region of Interest
Extraction

Framework

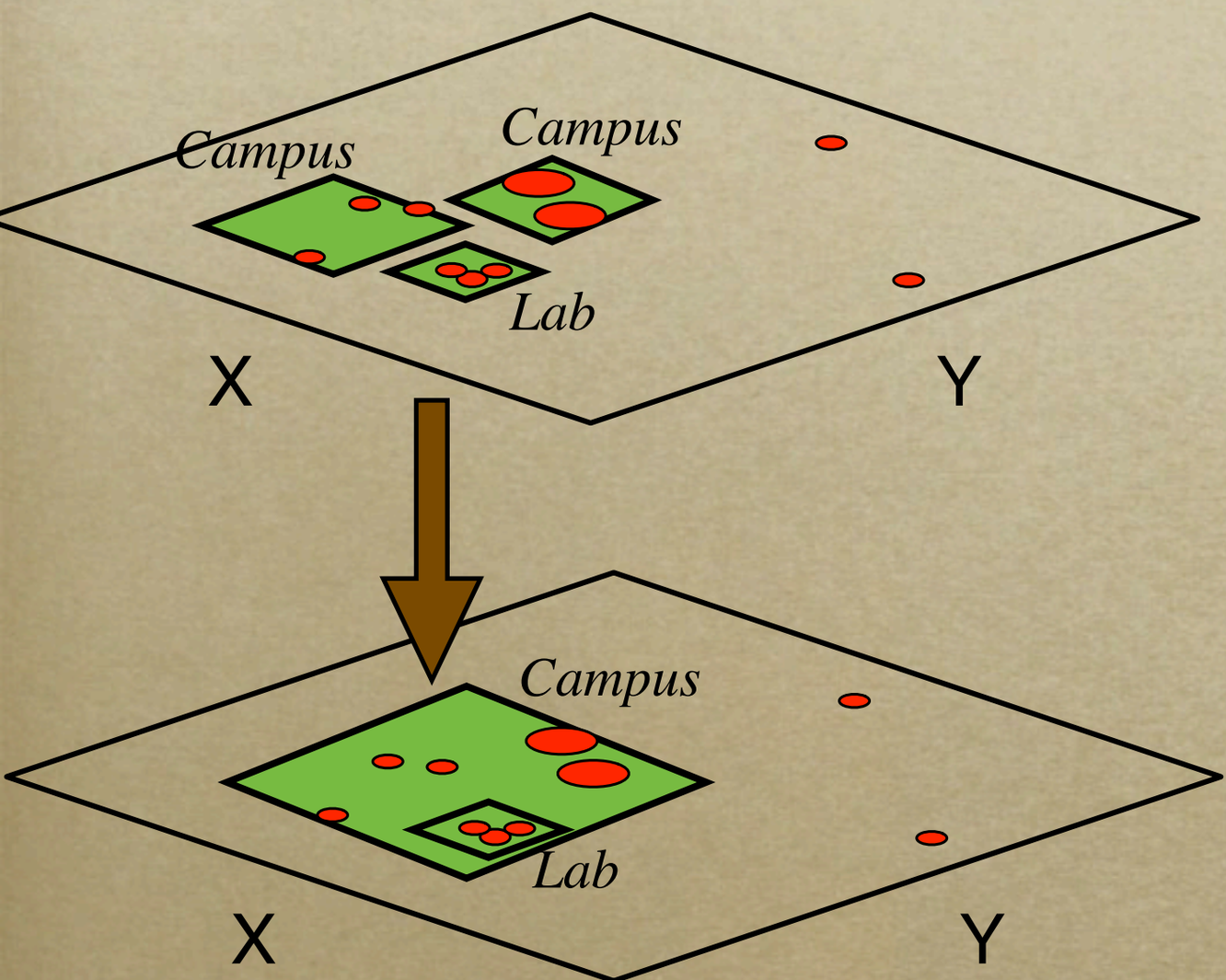

Smart Phone
Usage Logs
Geographic
Trajectories
Region of Interest
Extraction
POI Category
Database
Geographic Semantic
Enhanced ROI

Geographic Semantic Enhanced ROI



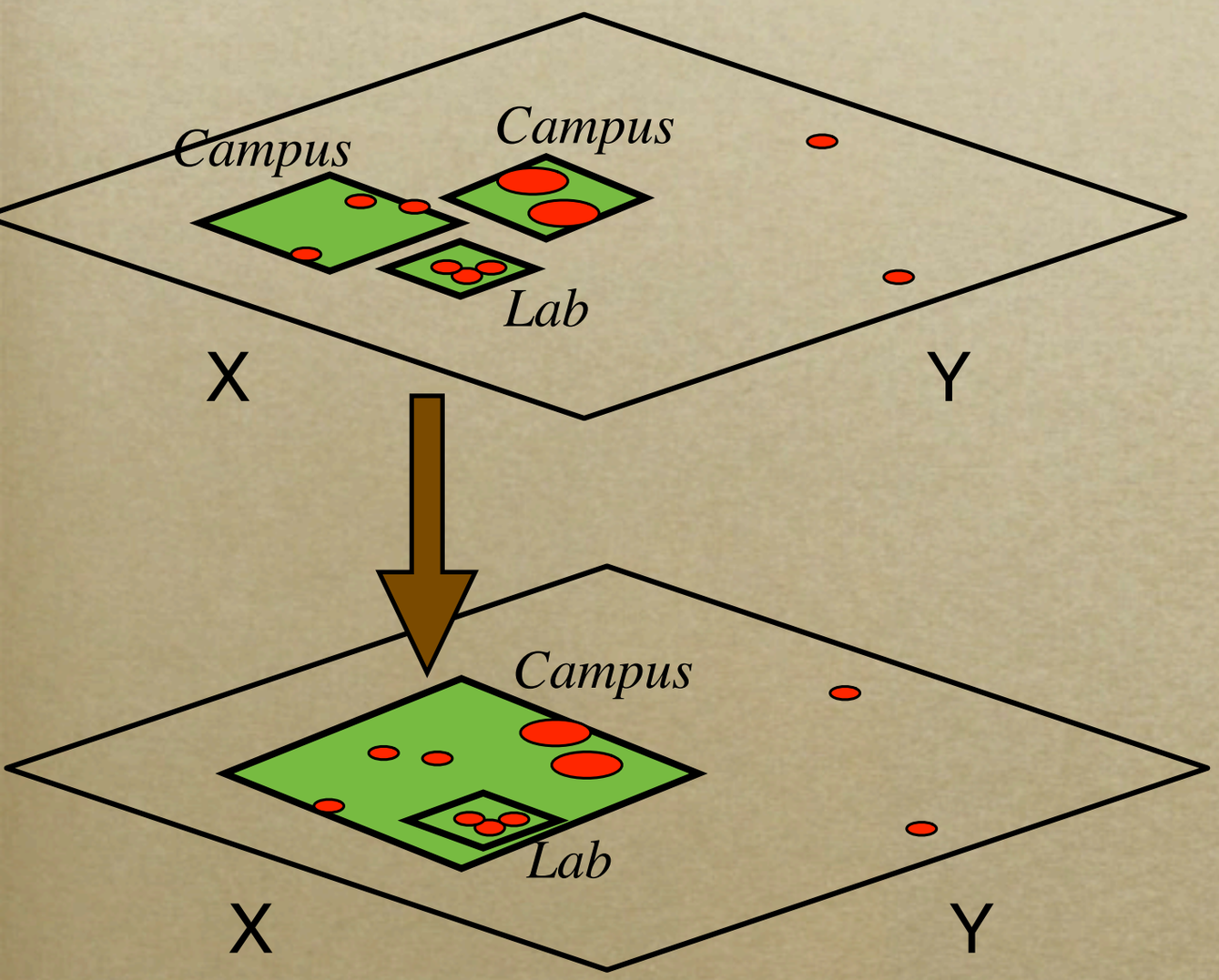
points of interest (POIs)
=> geographic semantic

Geographic Semantic Enhanced ROI



points of interest (POIs)
=> geographic semantic

Geographic Semantic Enhanced ROI



points of interest (POIs)
=> geographic semantic

Adopt SEM-CLS method [11]

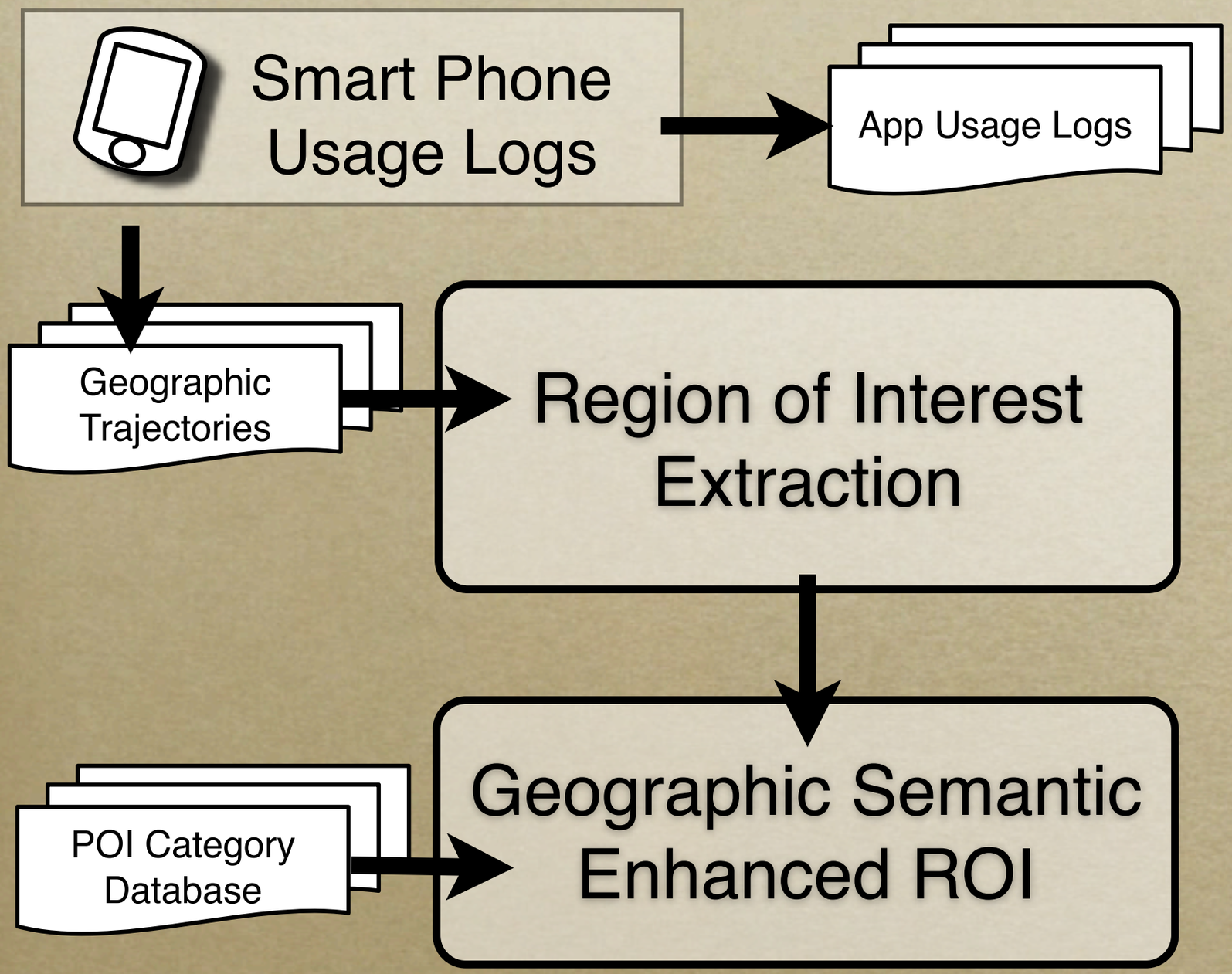
1. Split
   split cluster contains different geographic semantic

2. Merge
   merge similar clusters

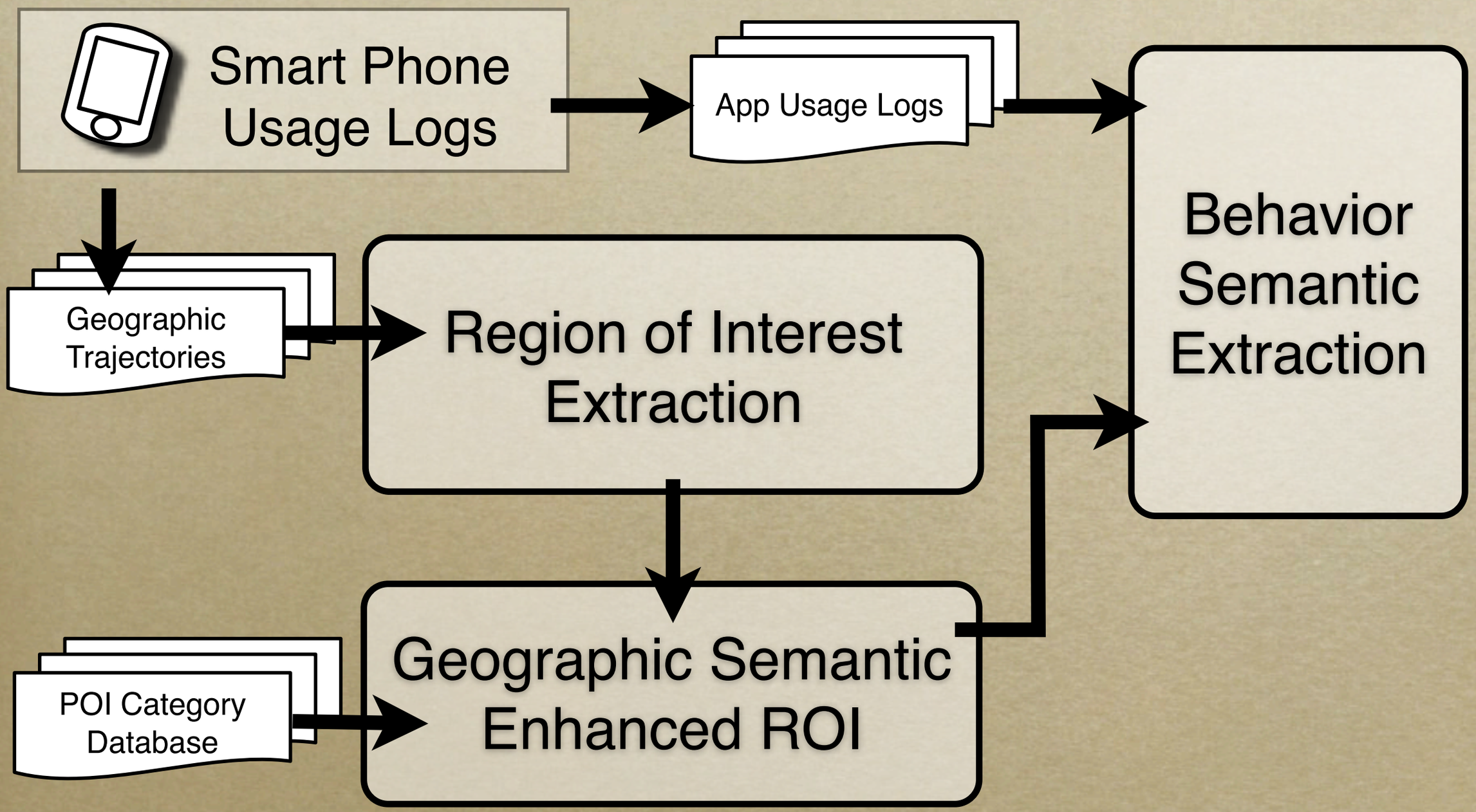
$$Sim(R_1, R_2) = \frac{\overrightarrow{Sem_{l1}} \cdot \overrightarrow{Sem_{l2}}}{\| \overrightarrow{Sem_{l1}} \| \| \overrightarrow{Sem_{l2}} \|} + \frac{min(t_{ae1}, t_{ae2})}{max(t_{ae1}, t_{ae2})} + \frac{min(t_{as1}, t_{as2})}{max(t_{as1}, t_{as2})}$$

[11] X. Cao, et. al. “Mining Significant Semantic Locations From GPS Data,” PVLDB (2010)

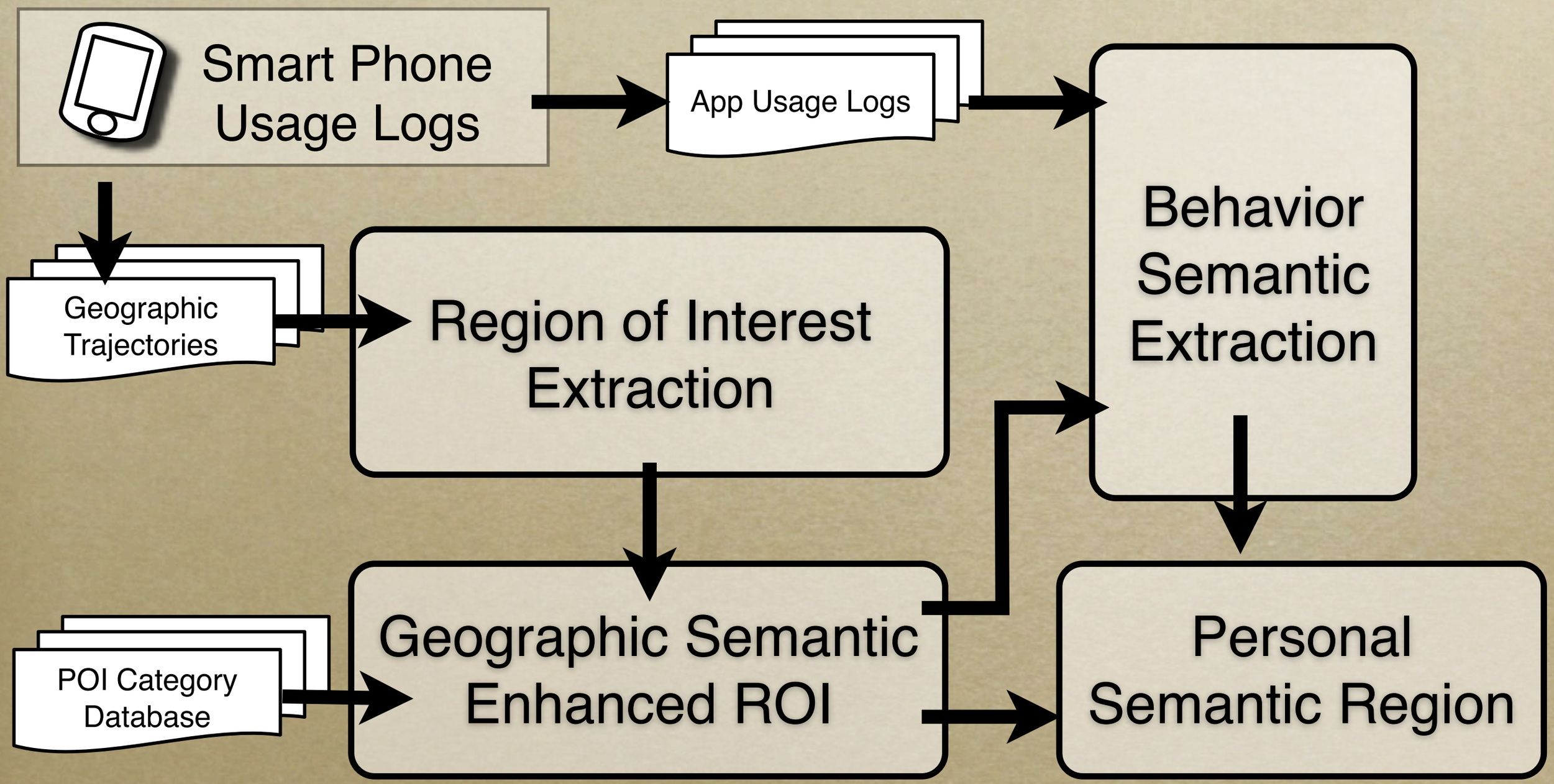
Framework


Smart Phone Usage Logs
App Usage Logs
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI

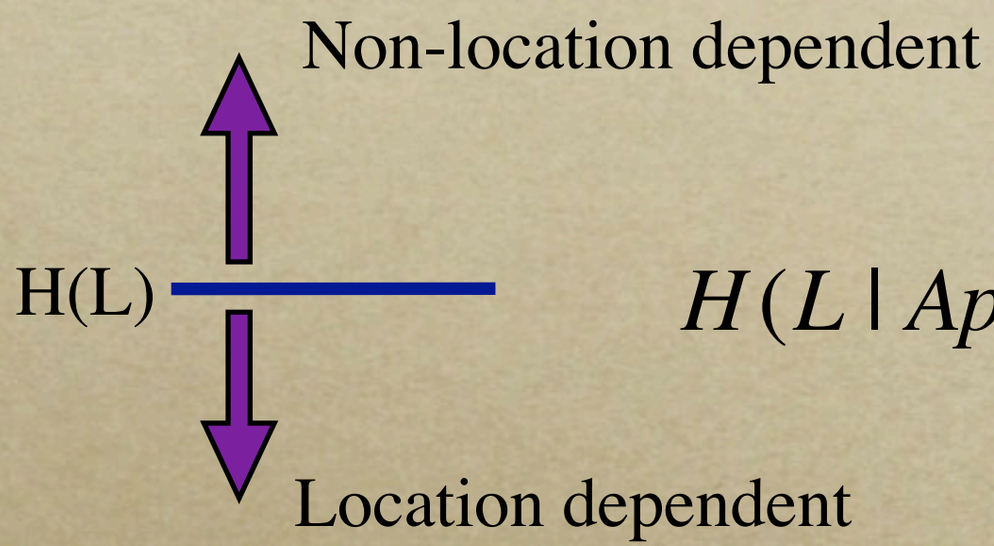
Framework


Smart Phone Usage Logs
App Usage Logs
Behavior Semantic Extraction
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI

Framework


Smart Phone Usage Logs
App Usage Logs
Behavior Semantic Extraction
Geographic Trajectories
Region of Interest Extraction
POI Category Database
Geographic Semantic Enhanced ROI
Personal Semantic Region

Behavior Semantic Extraction(1/2)
Location Dependency Measurement



$$H(L \mid App_i) = -\sum_{j=1}^{|L|} P(L_j) * \log_2 P(L_j)$$

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$H(L \mid APP_1) = -5 * \frac{10}{50} * \log_2(\frac{10}{50}) = 2.32$$

$$H(L \mid APP_2) = -\frac{5}{7} * \log_2(\frac{5}{7}) - 2 * \frac{1}{7} * \log_2(\frac{1}{7}) = 1.15$$

$$H(L \mid APP_3) = -3 * \frac{7}{21} * \log_2(\frac{7}{21}) = 1.58$$

Behavior Semantic Extraction(2/2)
App Usage Pattern

Definition:

An App usage pattern Γ_L is a ranking list of top K representative Apps on location L

Ranking functions:

Application Frequency:

$$AF_{i,j} = \frac{n_{i,j}}{\sum_{k=1}^{|L_j|} n_{k,j}}$$

Application Frequency-Inverse Location Frequency

$$AF-ILF_{i,j} = AF_{i,j} * \log_2\left(\frac{|L|}{|\{l : App_i \in L\}|}\right)$$

Application Frequency-Relative Location Frequency

$$AF-RLF_{i,j} = AF_{i,j} * \frac{n_{i,j}}{\sum_q n_{i,q}}$$

Ranking Function: Application Frequency

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF_{i,j} = \frac{n_{i,j}}{\sum_{k=1}^{|L_j|} n_{k,j}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.556	0.455	0.556	1	1
App_2	0.056	0.227	0.056	0	0
App_3	0.389	0.318	0.389	0	0

$$AF_{1,2} = \frac{10}{10+5+7} = \frac{10}{22}$$

$$AF_{2,2} = \frac{5}{10+5+7} = \frac{5}{22}$$

$$AF_{3,2} = \frac{7}{10+5+7} = \frac{7}{22}$$

Ranking Function: Application Frequency

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF_{i,j} = \frac{n_{i,j}}{\sum_{k=1}^{|L_j|} n_{k,j}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.556	0.455	0.556	1	1
App_2	0.056	0.227	0.056	0	0
App_3	0.389	0.318	0.389	0	0

$$AF_{1,2} = \frac{10}{10+5+7} = \frac{10}{22}$$

$$AF_{2,2} = \frac{5}{10+5+7} = \frac{5}{22}$$

$$AF_{3,2} = \frac{7}{10+5+7} = \frac{7}{22}$$

Ranking Function: Application Frequency

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF_{i,j} = \frac{n_{i,j}}{\sum_{k=1}^{|L_j|} n_{k,j}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.556	0.455	0.556	1	1
App_2	0.056	0.227	0.056	0	0
App_3	0.389	0.318	0.389	0	0

$$AF_{1,2} = \frac{10}{10+5+7} = \frac{10}{22}$$

$$AF_{2,2} = \frac{5}{10+5+7} = \frac{5}{22}$$

$$AF_{3,2} = \frac{7}{10+5+7} = \frac{7}{22}$$

Ranking Function: Application Frequency

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF_{i,j} = \frac{n_{i,j}}{\sum_{k=1}^{|L_j|} n_{k,j}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.556	0.455	0.556	1	1
App_2	0.056	0.227	0.056	0	0
App_3	0.389	0.318	0.389	0	0

$$AF_{1,2} = \frac{10}{10+5+7} = \frac{10}{22}$$

$$AF_{2,2} = \frac{5}{10+5+7} = \frac{5}{22}$$

$$AF_{3,2} = \frac{7}{10+5+7} = \frac{7}{22}$$

Ranking Function: AF-ILF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF-ILF_{i,j} = AF_{i,j} * \log_2\left(\frac{|L|}{|\{l : App_i \in L\}|}\right)$$

	L_1	L_2	L_3	L_4	L_5
App_1	0	0	0	0	0
App_2	0.041	0.167	0.041	0	0
App_3	0.287	0.234	0.287	0	0

$$AF-ILF_{1,2} = \frac{10}{22} * \log_2\left(\frac{5}{5}\right) = 0$$

$$AF-ILF_{2,2} = \frac{5}{22} * \log_2\left(\frac{5}{3}\right) = 0.167$$

$$AF-ILF_{3,2} = \frac{7}{22} * \log_2\left(\frac{5}{3}\right) = 0.234$$

Ranking Function: AF-ILF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF-ILF_{i,j} = AF_{i,j} * \log_2\left(\frac{|L|}{|\{l : App_i \in L\}|}\right)$$

	L_1	L_2	L_3	L_4	L_5
App_1	0	0	0	0	0
App_2	0.041	0.167	0.041	0	0
App_3	0.287	0.234	0.287	0	0

$$AF-ILF_{1,2} = \frac{10}{22} * \log_2\left(\frac{5}{5}\right) = 0$$

$$AF-ILF_{2,2} = \frac{5}{22} * \log_2\left(\frac{5}{3}\right) = 0.167$$

$$AF-ILF_{3,2} = \frac{7}{22} * \log_2\left(\frac{5}{3}\right) = 0.234$$

Ranking Function: AF-ILF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF-ILF_{i,j} = AF_{i,j} * \log_2\left(\frac{|L|}{|\{l : App_i \in L\}|}\right)$$

	L_1	L_2	L_3	L_4	L_5
App_1	0	0	0	0	0
App_2	0.041	0.167	0.041	0	0
App_3	0.287	0.234	0.287	0	0

$$AF-ILF_{1,2} = \frac{10}{22} * \log_2\left(\frac{5}{5}\right) = 0$$

$$AF-ILF_{2,2} = \frac{5}{22} * \log_2\left(\frac{5}{3}\right) = 0.167$$

$$AF-ILF_{3,2} = \frac{7}{22} * \log_2\left(\frac{5}{3}\right) = 0.234$$

Ranking Function: AF-ILF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF-ILF_{i,j} = AF_{i,j} * \log_2\left(\frac{|L|}{|\{l : App_i \in L\}|}\right)$$

	L_1	L_2	L_3	L_4	L_5
App_1	0	0	0	0	0
App_2	0.041	0.167	0.041	0	0
App_3	0.287	0.234	0.287	0	0

$$AF-ILF_{1,2} = \frac{10}{22} * \log_2\left(\frac{5}{5}\right) = 0$$

$$AF-ILF_{2,2} = \frac{5}{22} * \log_2\left(\frac{5}{3}\right) = 0.167$$

$$AF-ILF_{3,2} = \frac{7}{22} * \log_2\left(\frac{5}{3}\right) = 0.234$$

Ranking Function: AF-RLF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF - RLF_{i,j} = AF_{i,j} * \frac{n_{i,j}}{\sum_q n_{i,q}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.111	0.09	0.111	0.2	0.2
App_2	0.008	0.162	0.008	0	0
App_3	0.13	0.106	0.13	0	0

$$AF - RLF_{1,2} = \frac{10}{22} * \frac{10}{10 * 5} = 0.09$$

$$AF - RLF_{2,2} = \frac{5}{22} * \frac{5}{1 + 5 + 1} = 0.162$$

$$AF - RLF_{3,2} = \frac{7}{22} * \frac{7}{7 * 3} = 0.106$$

Ranking Function: AF-RLF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF - RLF_{i,j} = AF_{i,j} * \frac{n_{i,j}}{\sum_q n_{i,q}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.111	0.09	0.111	0.2	0.2
App_2	0.008	0.162	0.008	0	0
App_3	0.13	0.106	0.13	0	0

$$AF - RLF_{1,2} = \frac{10}{22} * \frac{10}{10 * 5} = 0.09$$

$$AF - RLF_{2,2} = \frac{5}{22} * \frac{5}{1 + 5 + 1} = 0.162$$

$$AF - RLF_{3,2} = \frac{7}{22} * \frac{7}{7 * 3} = 0.106$$

Ranking Function: AF-RLF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF - RLF_{i,j} = AF_{i,j} * \frac{n_{i,j}}{\sum_q n_{i,q}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.111	0.09	0.111	0.2	0.2
App_2	0.008	0.162	0.008	0	0
App_3	0.13	0.106	0.13	0	0

$$AF - RLF_{1,2} = \frac{10}{22} * \frac{10}{10 * 5} = 0.09$$

$$AF - RLF_{2,2} = \frac{5}{22} * \frac{5}{1 + 5 + 1} = 0.162$$

$$AF - RLF_{3,2} = \frac{7}{22} * \frac{7}{7 * 3} = 0.106$$

Ranking Function: AF-RLF

	L_1	L_2	L_3	L_4	L_5
App_1	10	10	10	10	10
App_2	1	5	1	0	0
App_3	7	7	7	0	0

$$AF - RLF_{i,j} = AF_{i,j} * \frac{n_{i,j}}{\sum_q n_{i,q}}$$

	L_1	L_2	L_3	L_4	L_5
App_1	0.111	0.09	0.111	0.2	0.2
App_2	0.008	0.162	0.008	0	0
App_3	0.13	0.106	0.13	0	0

$$AF - RLF_{1,2} = \frac{10}{22} * \frac{10}{10 * 5} = 0.09$$

$$AF - RLF_{2,2} = \frac{5}{22} * \frac{5}{1 + 5 + 1} = 0.162$$

$$AF - RLF_{3,2} = \frac{7}{22} * \frac{7}{7 * 3} = 0.106$$

Outline

- *Motivation*
- *Goal*
- *Related Works*
- *Framework*
  - *Region Extraction*
  - *Personal Semantic Extraction*
- *Experiment*
- *Conclusion*

Experiment

Personal Semantic Region Discovery

- Dataset
- Setting
- Case Study
  - comparison between different users
- App usage pattern
  - comparison between ranking function
  - impact of parameter
- Region extraction
  - comparison between SDC and SP

Experiment: Dataset

Dataset:

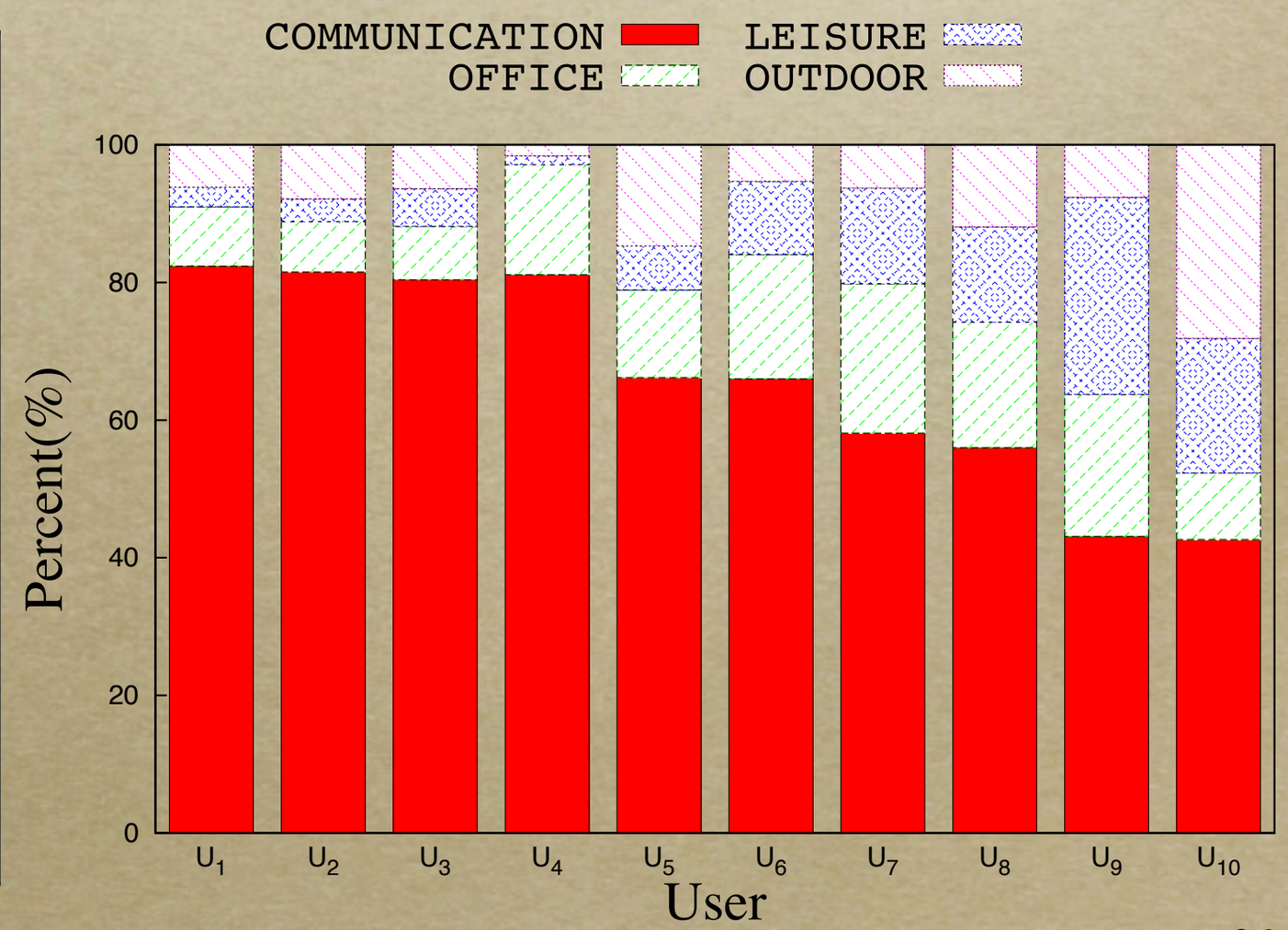
track 14 users over 10 months
use subset of 10 users during 2010-07 to 2010-12

Experiment: Dataset

Dataset:

track 14 users over 10 months
use subset of 10 users during 2010-07 to 2010-12

Behavior Semantic	App Category
COMMUNICATION	Communication
OFFICE	Finance, Business, Productivity, Tools, Personalization
LEISURE	Entertainment, Games, Photograph, Media & Video, Music & Audio, Books & Reference, News & Magazine
OUTDOOR	Social, Transportation, Travel & Local, Weather, Health & Fitness, Shopping, Sports



Experiment: Setting

Region extraction:

Experiment: Setting

Region extraction:

- Parameters:
  - minimum stay duration (Period) > 15 minutes
  - minimum number of stay (MinSR) = 3
  - maximum distance (d_h) = 400 meters

Experiment: Setting

Region extraction:

- Parameters:
  - minimum stay duration (Period) > 15 minutes
  - minimum number of stay (MinSR) = 3
  - maximum distance (d_h) = 400 meters
- Result:

  3446 ROI candidates and 92 ROIs
  from 0.3 million location points

Experiment: Setting

Region extraction:

- Parameters:
  - minimum stay duration (Period) > 15 minutes
  - minimum number of stay (MinSR) = 3
  - maximum distance (d_h) = 400 meters
- Result:

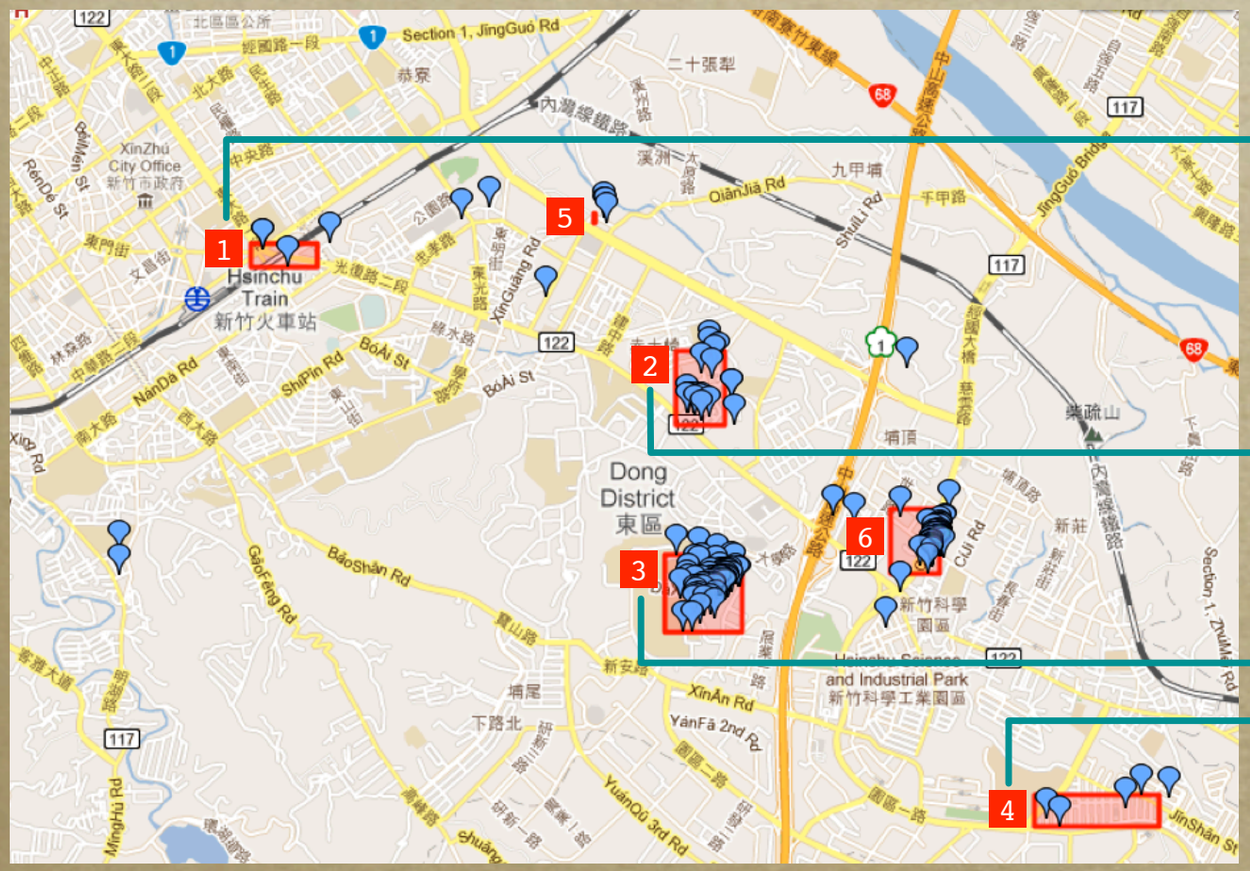
  3446 ROI candidates and 92 ROIs
  from 0.3 million location points

Semantic Annotation:

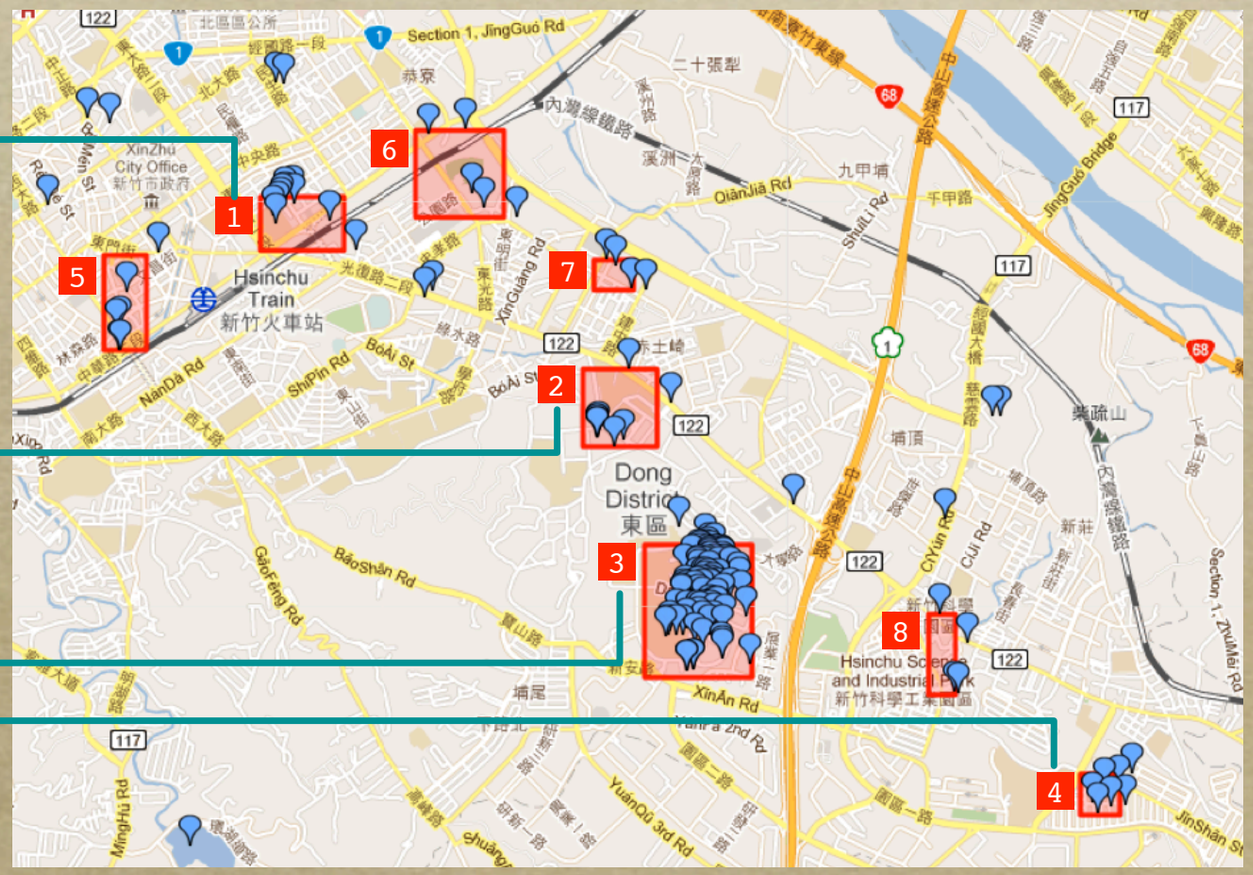
- Parameters:
  - App usage pattern size (K) = 3

Experiment: Case Study

Different users on common places



Personal ROIs of User U_7



Personal ROIs of User U_3

Experiment: Case Study

Region	Geographic Semantic	Behavior Semantic	
		User U_3	User U_7
ℓ_1	Hsinchu Train Station	OUTDOOR	LEISURE
ℓ_2	NTHU night market	LEISURE	OUTDOOR
ℓ_3	NCTU Campus	WORK	COMMUNICATION
ℓ_4	JinShan st.	LEISURE	LEISURE

Region	User U_3		User U_7	
	App Usage Pattern (K=3)	AF-RLF Score	App Usage Pattern (K=3)	AF-RLF Score
ℓ_1	PATS (Travel & Local)	0.023	GoogleMap (Travel & Local)	0.052
	NubiNews (News & Magazine)	0.017	Fxcamera (Photograph)	0.049
	FriendStream (Social)	0.004	AngelPiano (Games)	0.044
ℓ_2	Album (Media & Video)	0.013	Facebook (Social)	0.035
	Skype (Communication)	0.01	Calendar (Productivity)	0.023
			Skype (Communication)	0.017
ℓ_3	Clock (Tools)	0.036	MMS (Communication)	0.027
	NubiNews (News & Magazine)	0.018	Calendar (Productivity)	0.025
	Facebook (Social)	0.017	Bump (Social)	0.023
ℓ_4	NubiNews (News & Magazine)	0.045	SkyMap (Books & Reference)	0.015
	FriendStream (Social)	0.014	Bubbles (Games)	0.009
			Naruto (Games)	0.008

Experiment: App Usage Pattern Measurements

- User verification:

  Ask user whether this App is a representative App?

  ➡Yes, No Idea, No

Experiment: App Usage Pattern Measurements

- User verification:

  Ask user whether this App is a representative App?

  ➡Yes, No Idea, No

  Assume "Yes" Apps are the representative Apps.

Experiment: App Usage Pattern Measurements

- User verification:

  Ask user whether this App is a representative App?

  ➡Yes, No Idea, No

  Assume "Yes" Apps are the representative Apps.

- Measurement:

  Because each representative App usage record can represent for user behavior, we random select "Yes" App records in the region and examine whether we can find it in the App usage pattern.

  -Precision: $$\frac{\|\{\text{"}Yes\text{" App in App usage pattern}\}\|}{\|\{\text{"}Yes\text{" App record}\}\|}$$

  -nDCG: $$DCG_K = \sum_{i=1}^{K} \frac{2^{rel_i} - 1}{\log_2(1+i)} \ , nDCG_K = \frac{DCG_K}{IDCG_K}$$

Experiment: App Usage Pattern

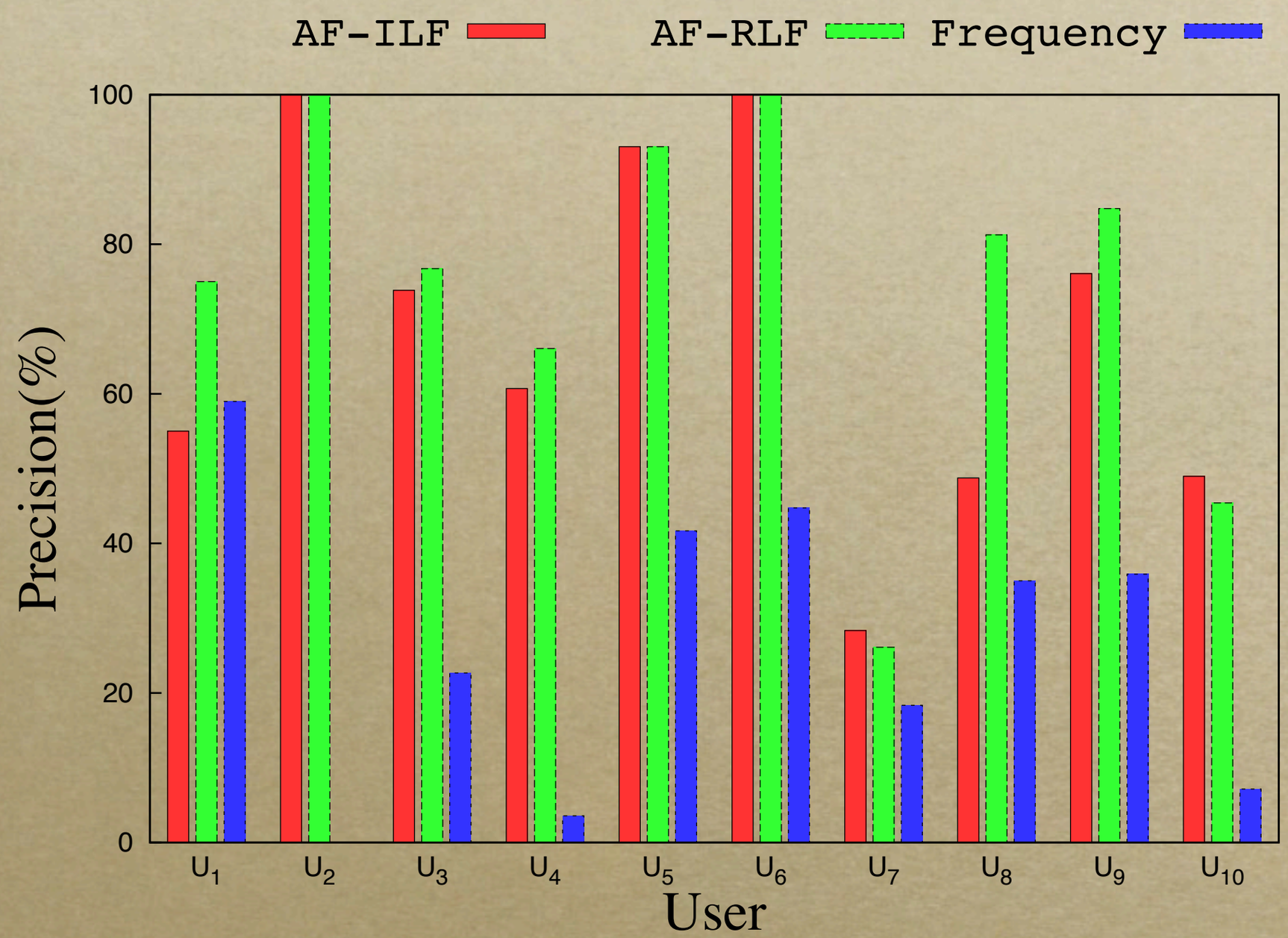
- Compare effectiveness of ranking functions:
  - ranking functions
    - *-Frequency*
    - *-AF-ILF*
    - *-AF-RLF*

Experiment: App Usage Pattern

- Compare effectiveness of ranking functions:
  - ranking functions
    - *-Frequency*
    - *-AF-ILF*
    - *-AF-RLF*

- Investigate impact of App usage pattern size (K)

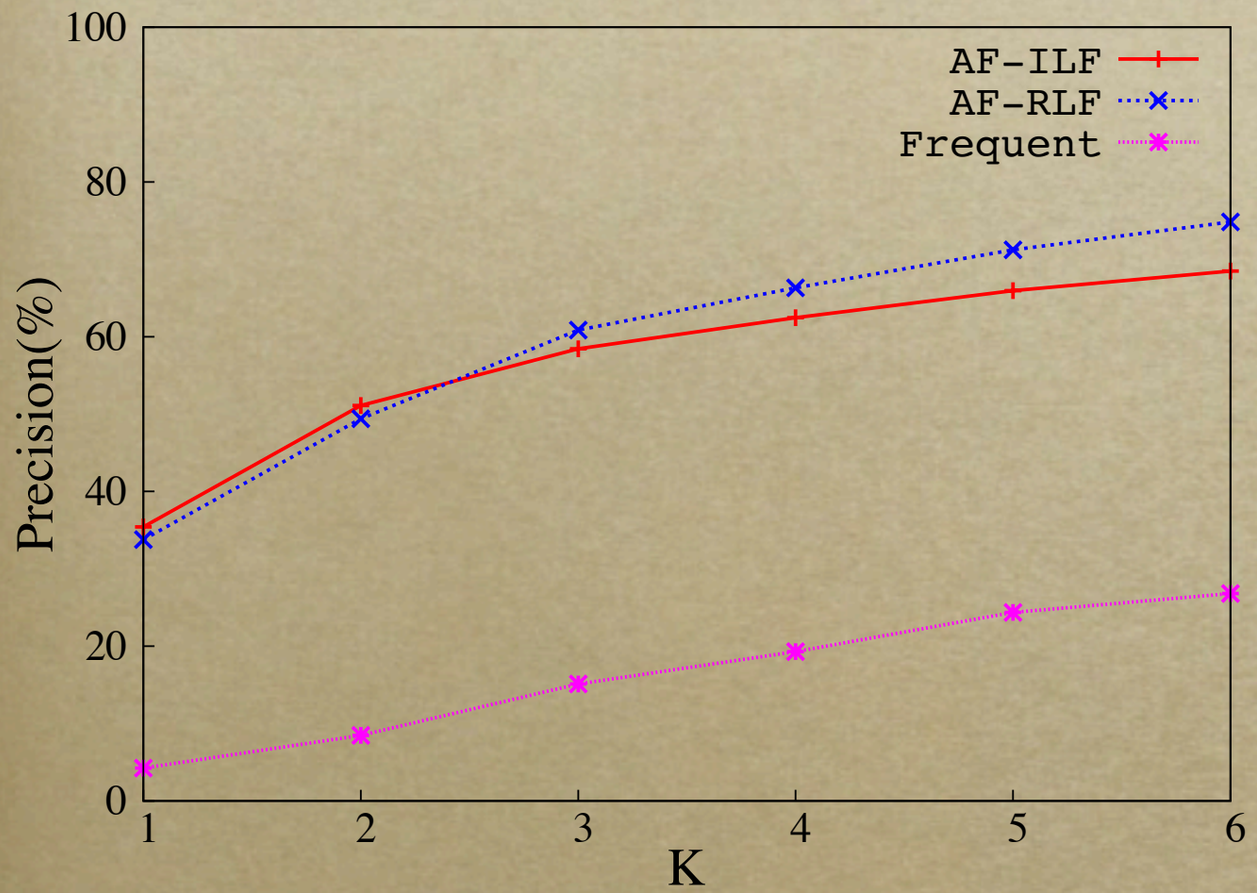
Experiment: App Usage Pattern
Comparison between ranking functions

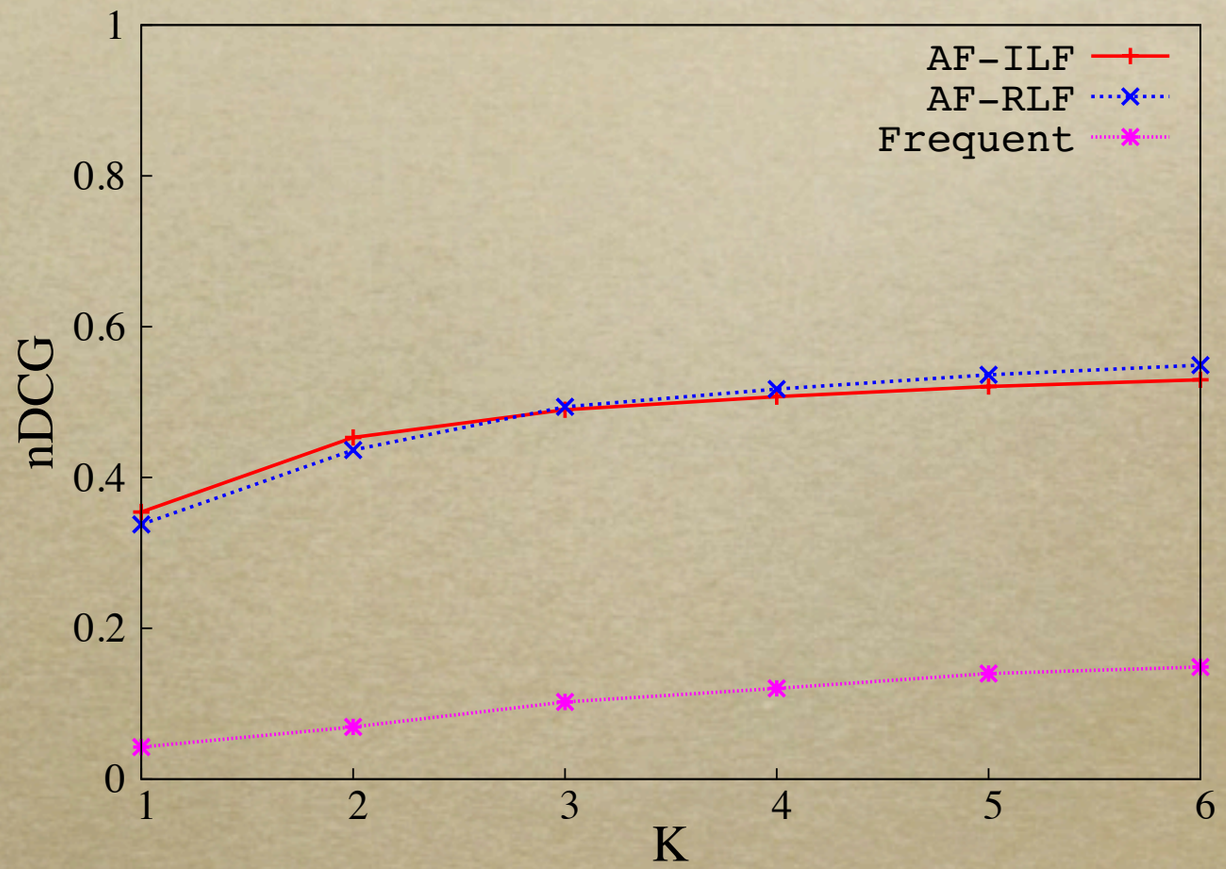


K=6

AF-RLF dominates other two approaches

Experiment: App Usage Pattern
Impact of App usage pattern size (K)





Precision and nDCG increase when the size of list increase

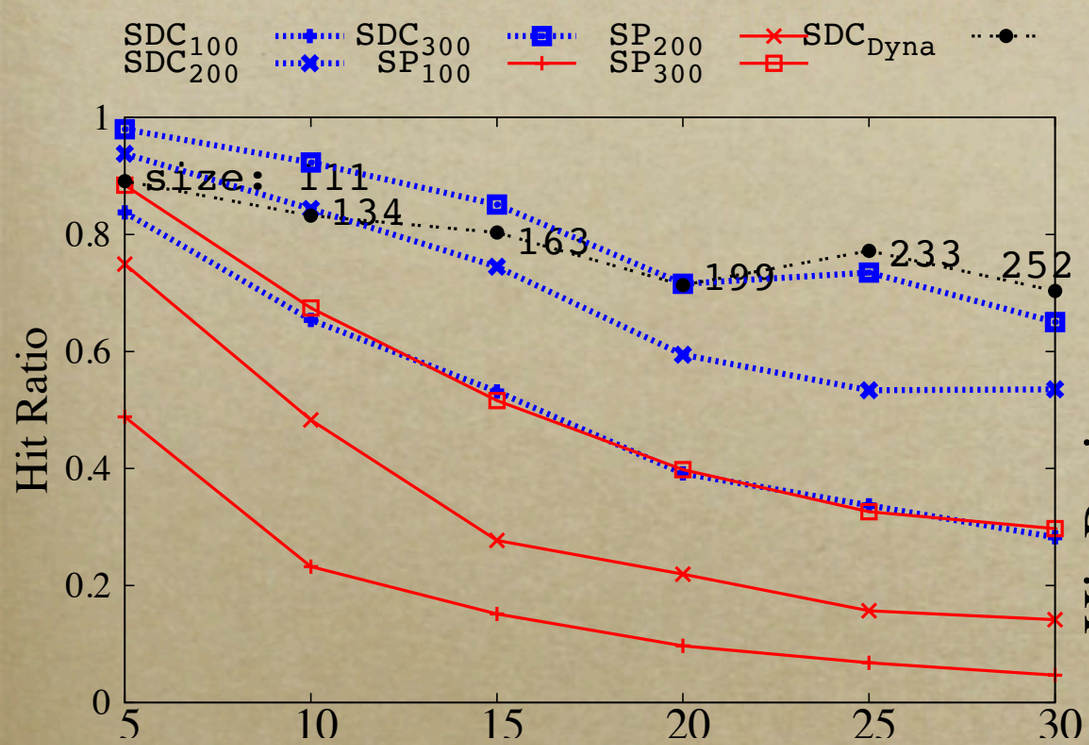
Experiment: Region Extraction

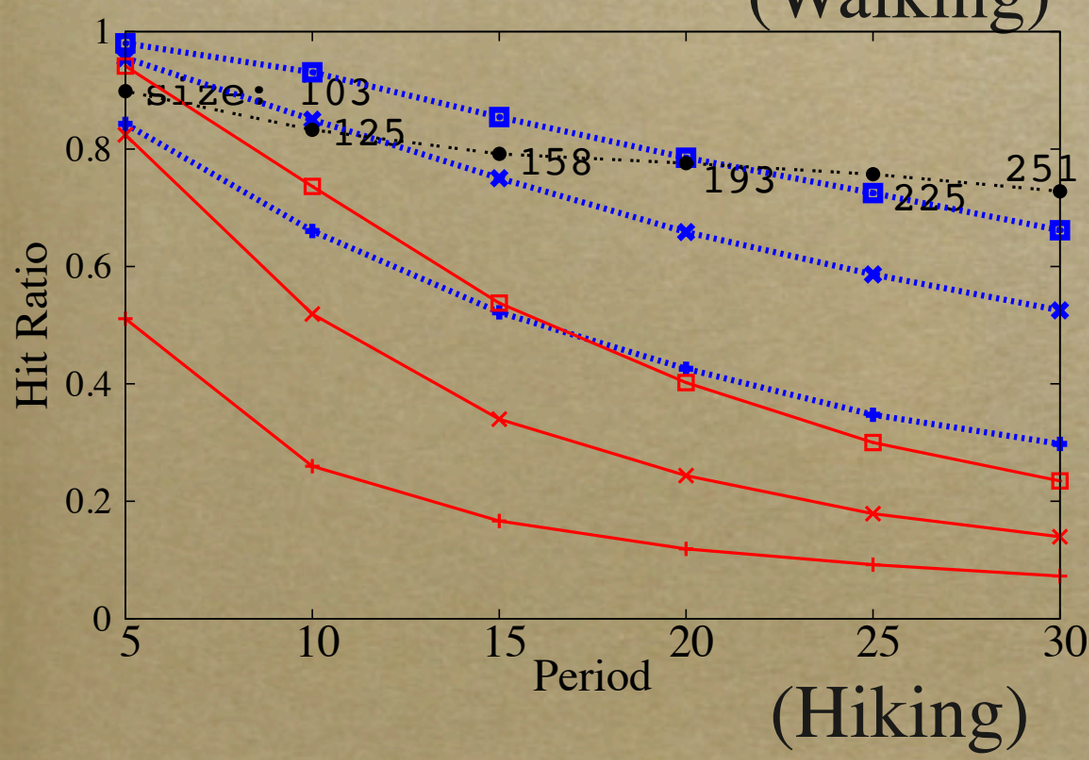
Activity	# Trajectory	# Photo
Hiking	3839	33065
Road Biking	5032	11968
Walking	955	4685

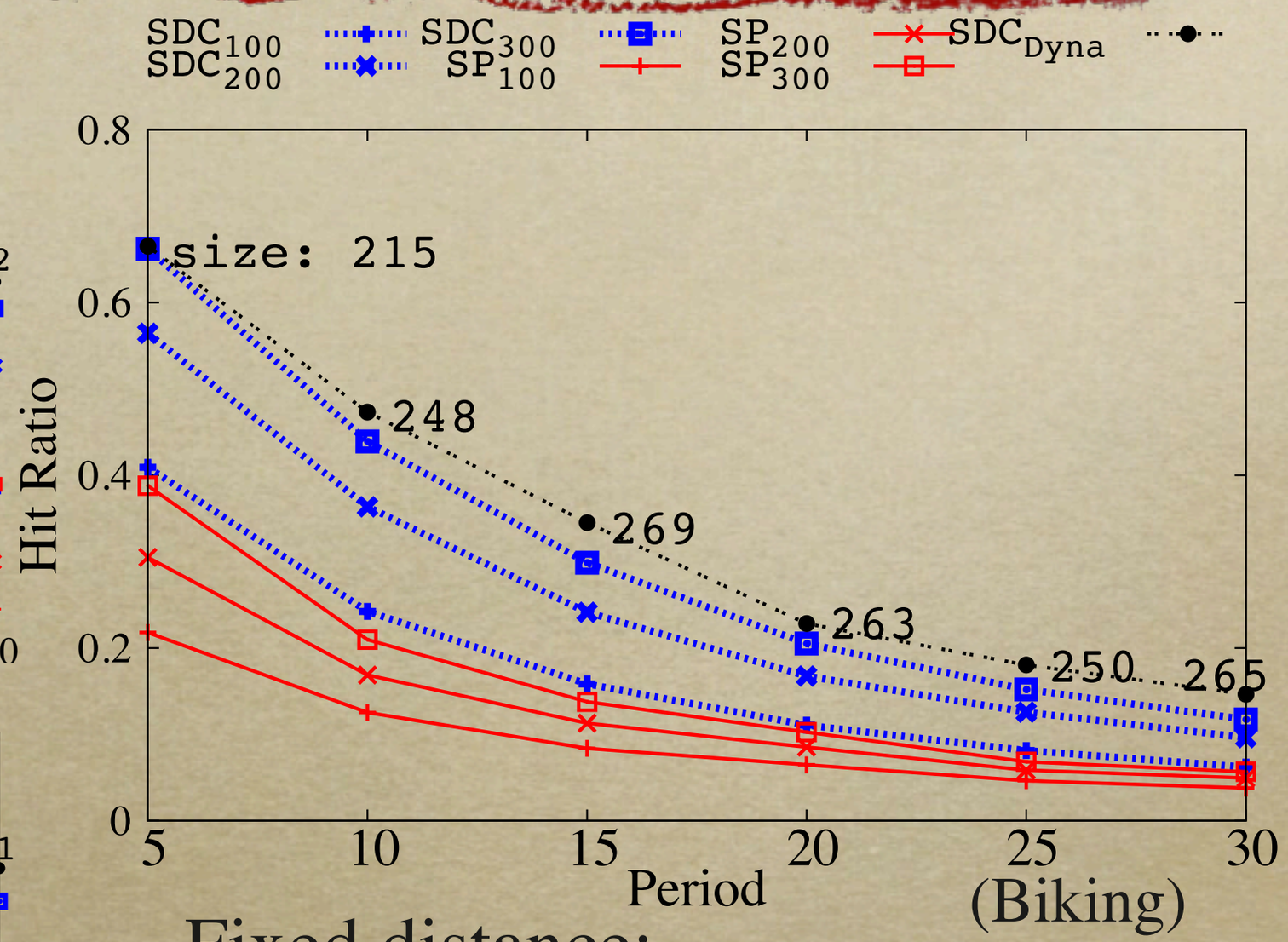
Dataset of three activities in California from Everytrail

- Compare hit ratio of our method (SDC) with Stay Point (SP)
  - Hit ratio: ratio of photos inside discovered region

Experiment: Region Extraction Comparison between SDC and SP







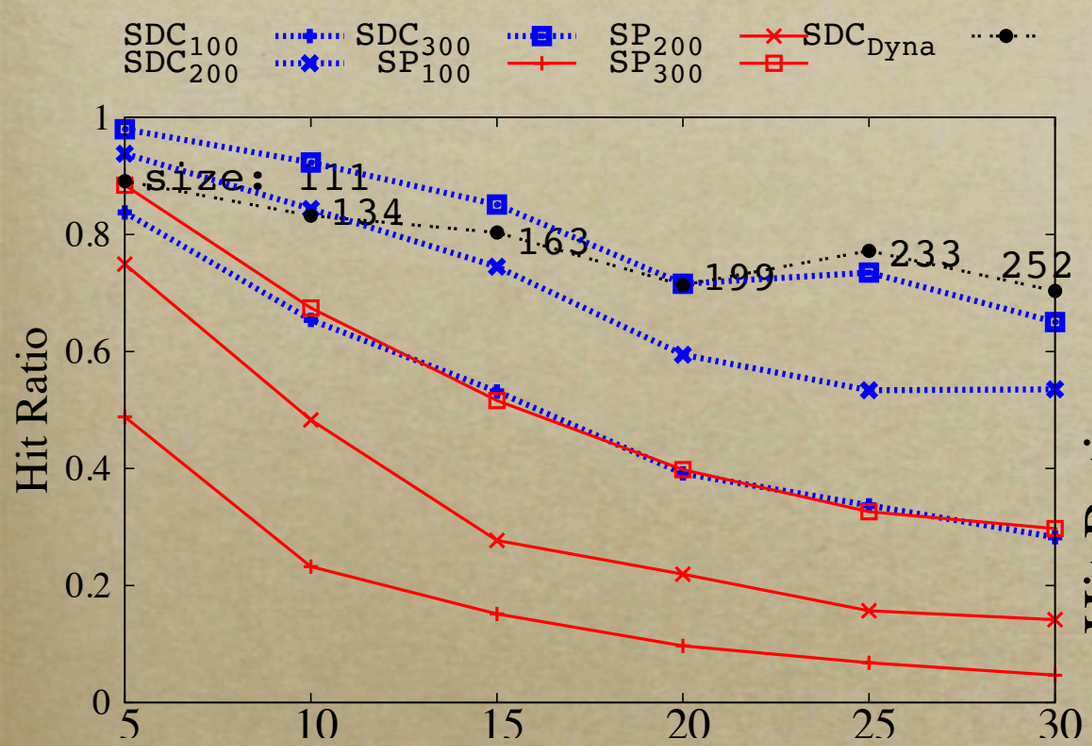
Fixed distance:

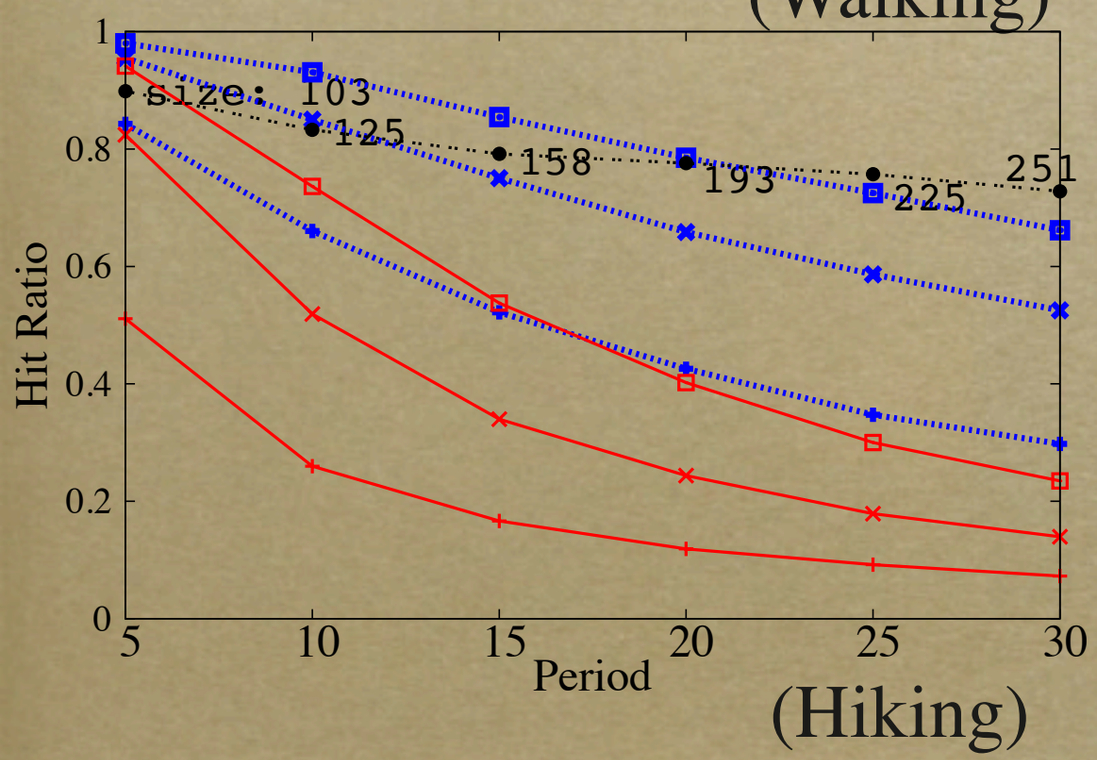
- SP: Stay Point
- SDC: Sequential Density Clustering

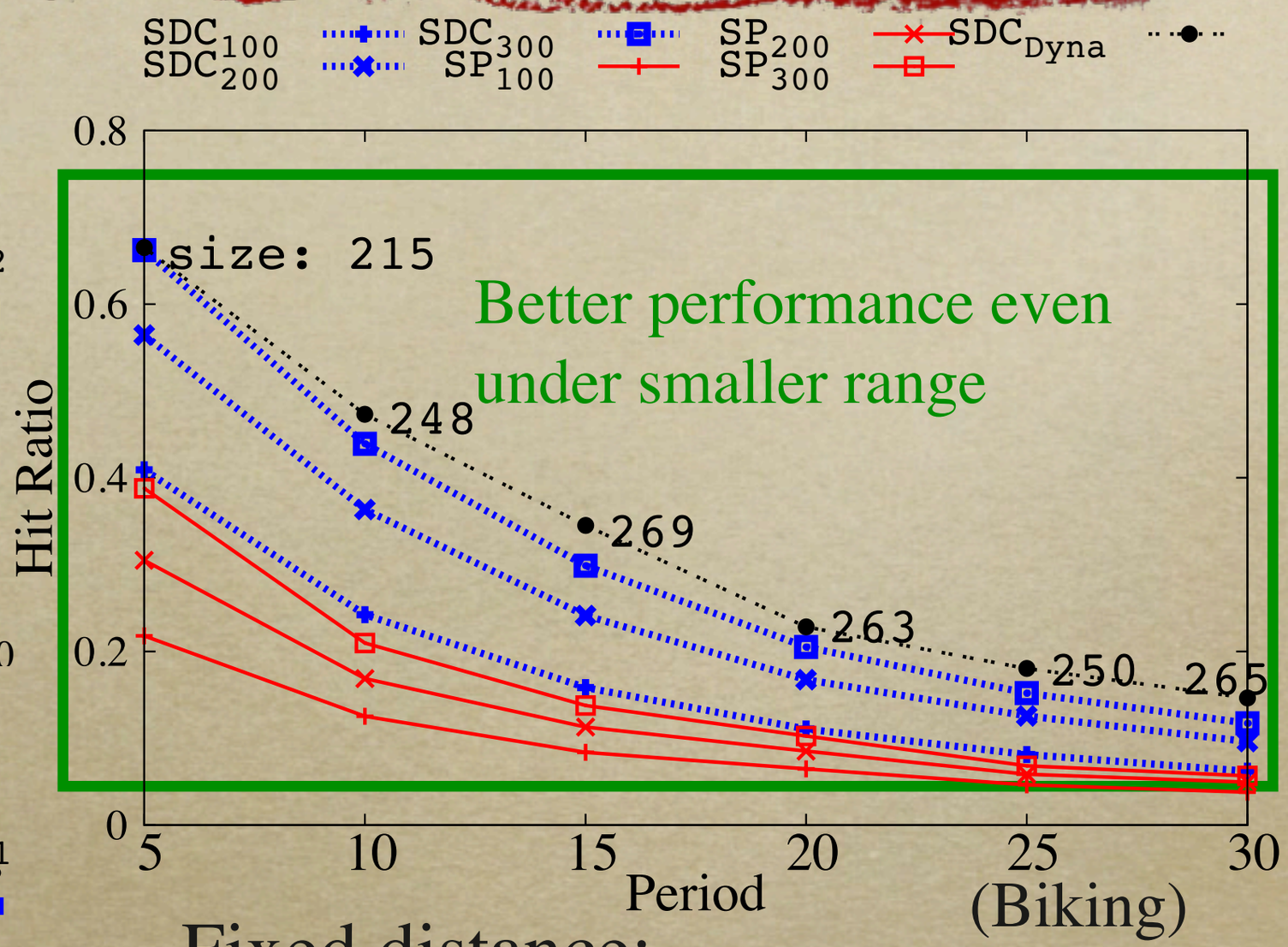
Variable distance:

- SDC_{Dyna}: SDC under dynamic range

Experiment: Region Extraction Comparison between SDC and SP







Fixed distance:

- SP: Stay Point
- SDC: Sequential Density Clustering

Variable distance:

- SDC_{Dyna}: SDC under dynamic range

Conclusion

- Exploring App usage patterns for discovering personal semantic regions
- Develop a framework of mining personal semantic regions
  - Propose sequential density clustering to discover regions of interests
  - Propose ranking list to extract representative Apps
    - *-Frequency (Application Frequency)*
    - *-AF-ILF (Application Frequency-Inverse Location Frequency)*
    - *-AF-RLF (Application Frequency-Relative Location Frequency)*
- Experiment shows our method is more accurate than existing schemes (stay points).

Thank You