

Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

Xuemei Liu, James Biagioni, Jakob Eriksson,
Yin Wang, George Forman, Yanmin Zhu



UIC Bits Networked
UNIVERSITY OF ILLINOIS
AT CHICAGO Systems Laboratory
COLLEGE OF ENGINEERING

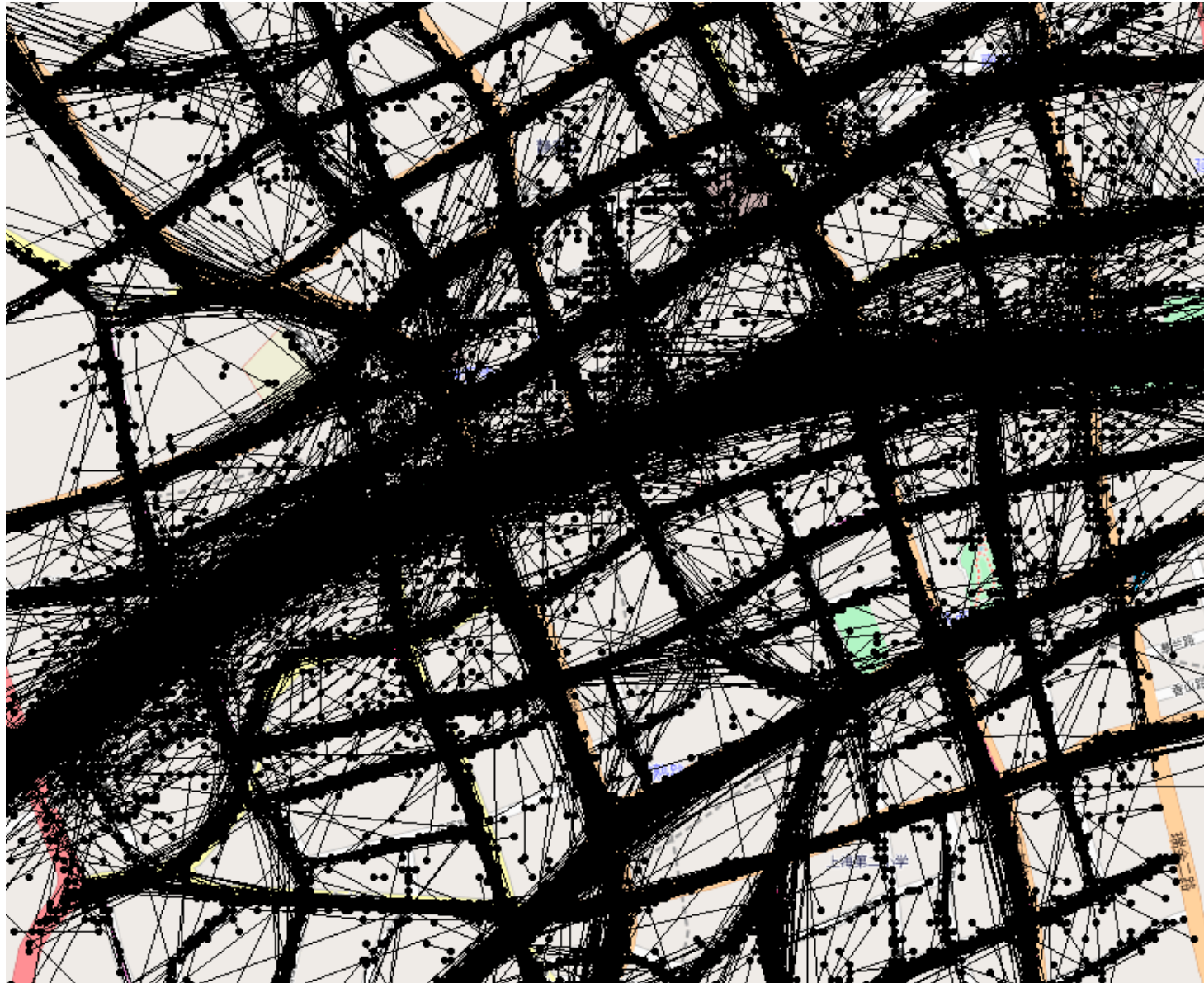
Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

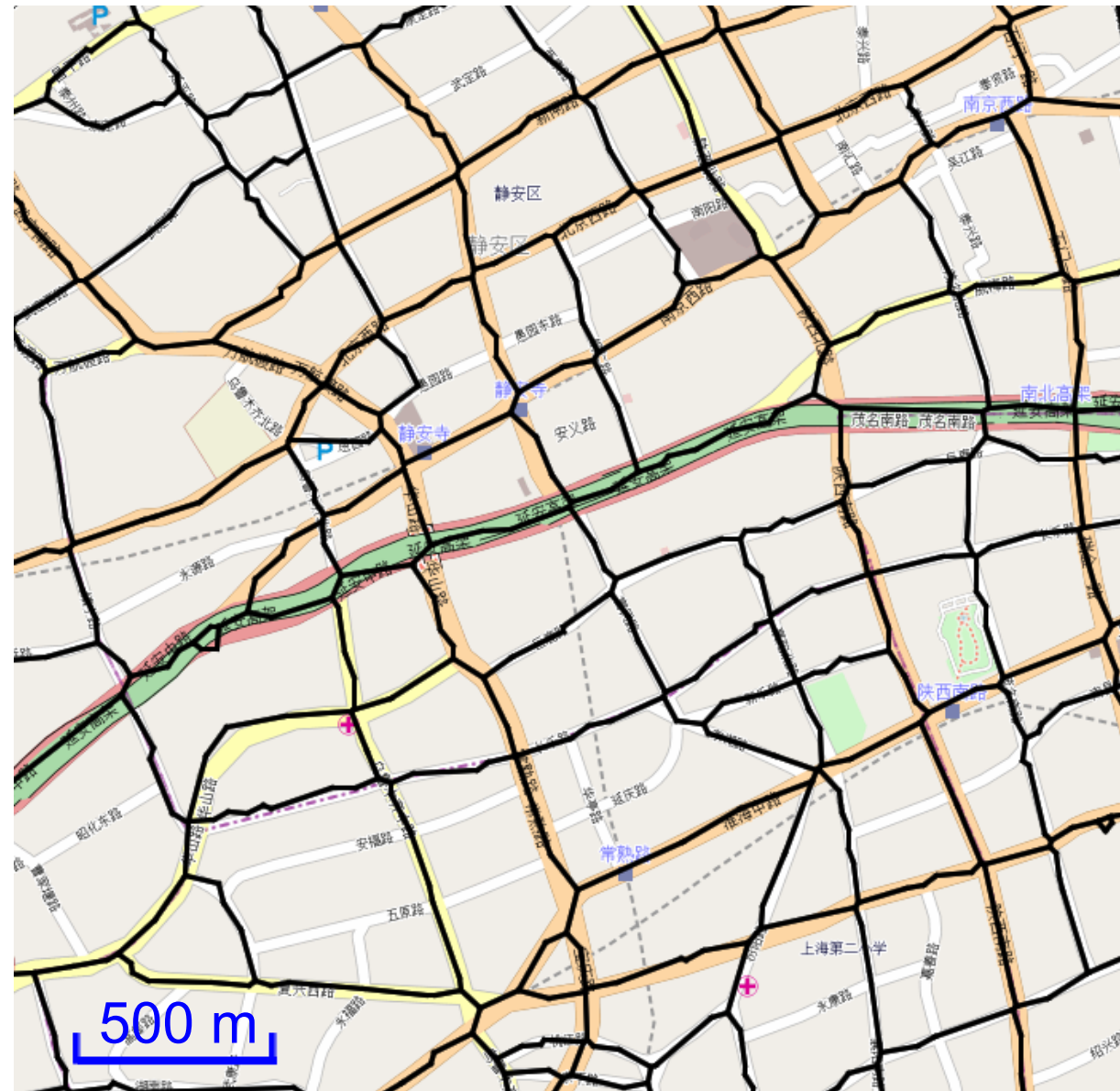
Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

Raw GPS traces

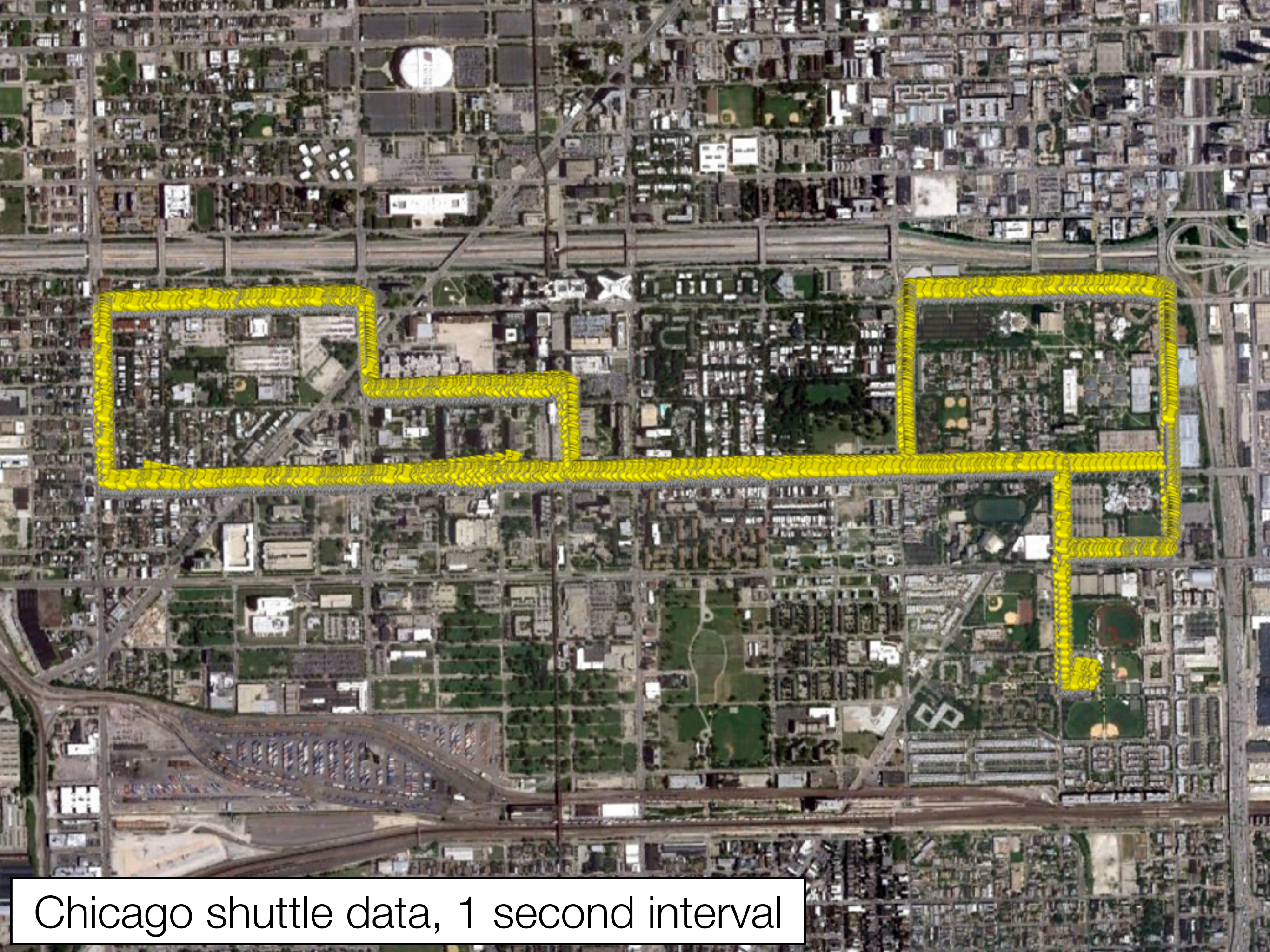


Inferred road map



Mining Large-Scale, **Sparse** **GPS Traces** for Map Inference

Comparison of Approaches



Chicago shuttle data, 1 second interval

Shanghai taxi data, 16/61 second interval



Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

2 hours of Shanghai data



Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

Mining Large-Scale, Sparse GPS Traces for Map Inference

Comparison of Approaches

Existing approaches

- ▶ *k*-Means clustering
 - Edelkamp & Schrödl (2003)
- ▶ Kernel density estimation
 - Davies et al. (2006)
- ▶ Trace merging
 - Liu et al. (2012)

Why infer maps?

Road surveys



Rural/developing areas



New road construction



Road closures



Road closures



Opportunistic data collection



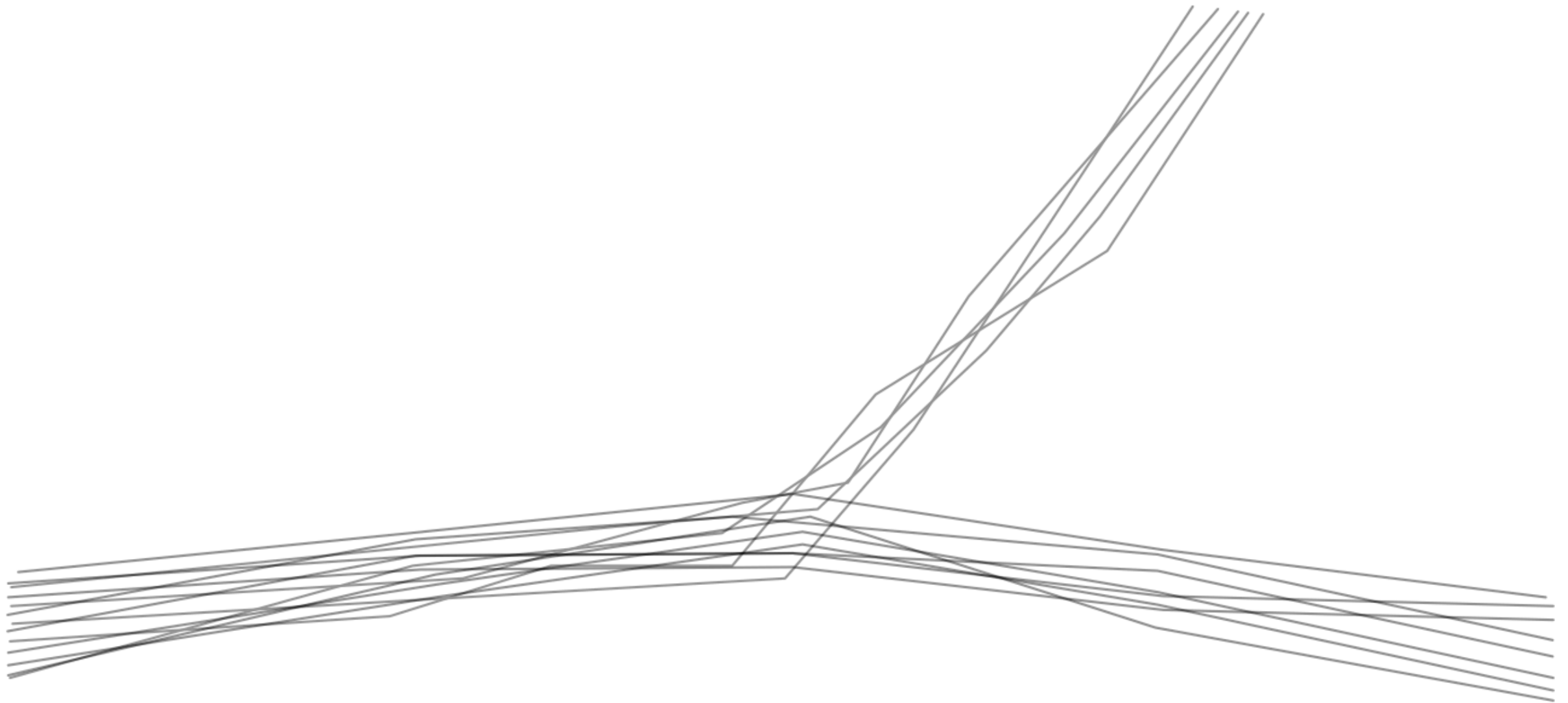
Existing approaches

- ▶ *k*-Means clustering
 - Edelkamp & Schrödl (2003)
- ▶ Kernel density estimation
 - Davies et al. (2006)
- ▶ Trace merging
 - Liu et al. (2012)

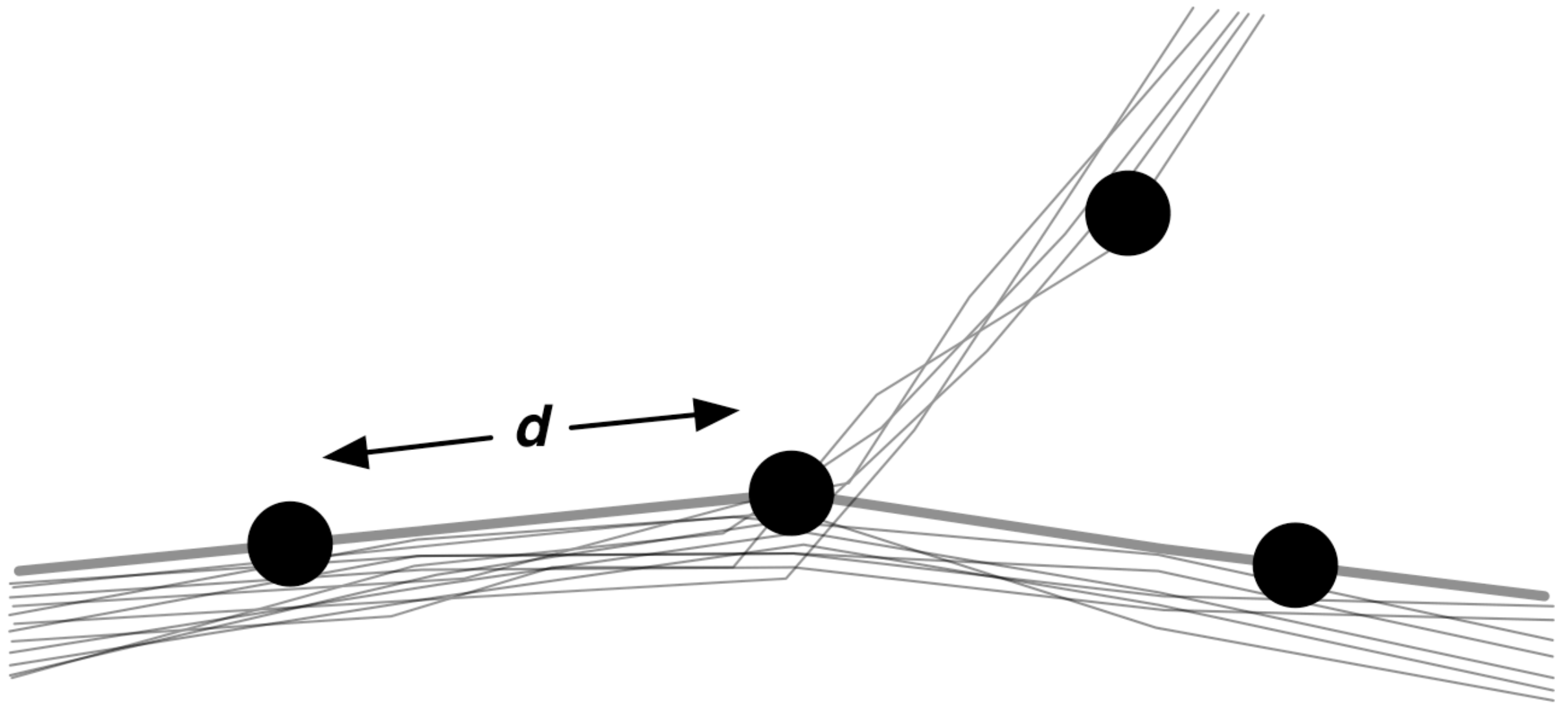
k -Means Clustering

Edelkamp & Schrödl (2003)

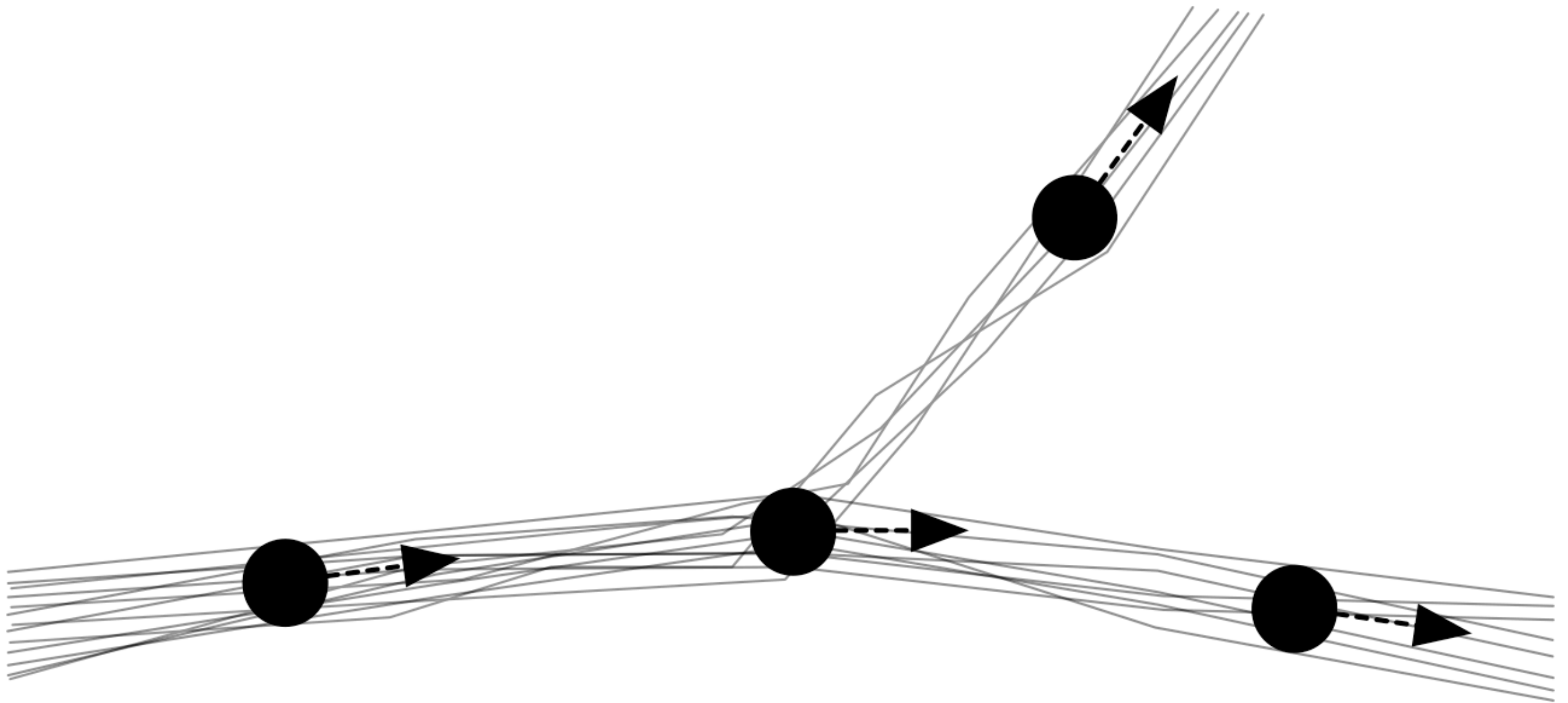
Raw GPS traces



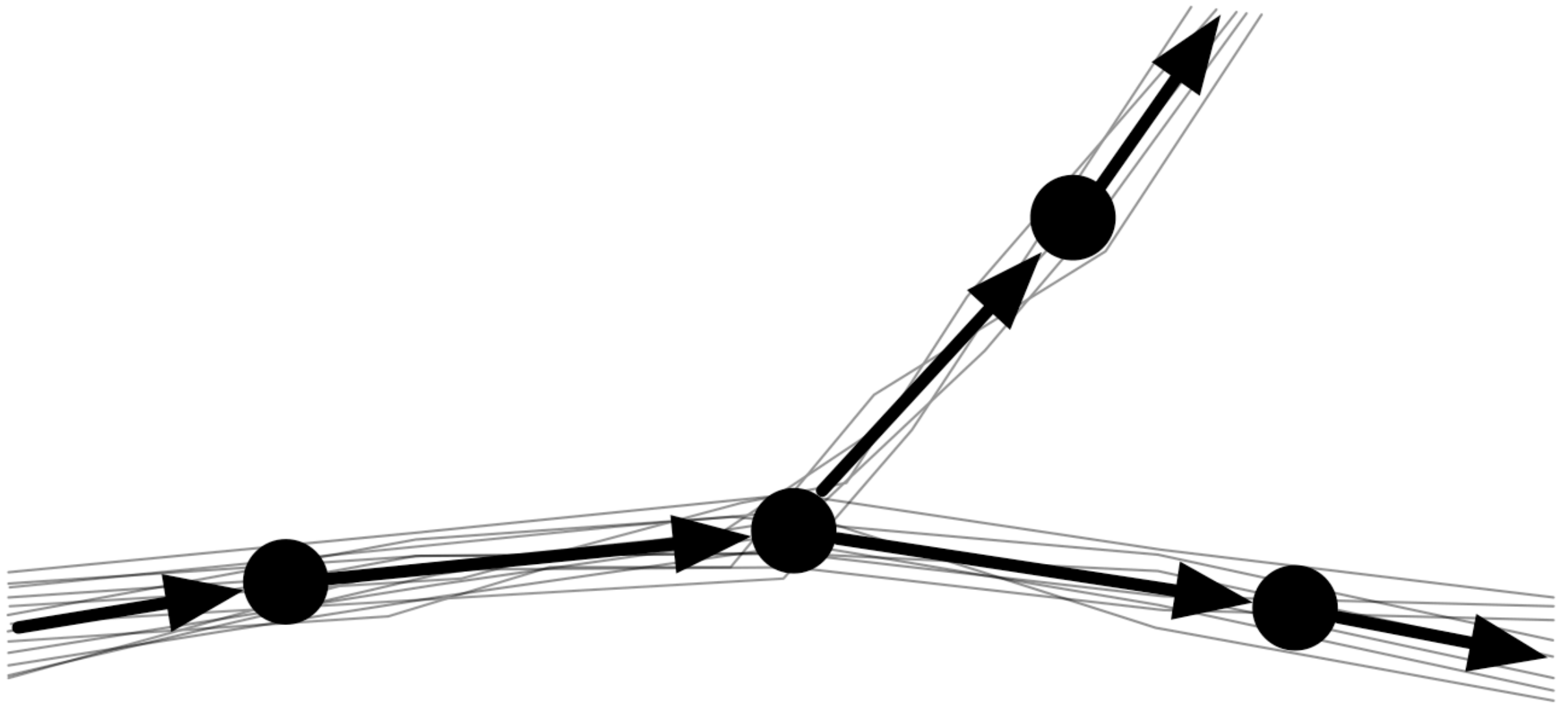
Drop seeds



Adjust seeds



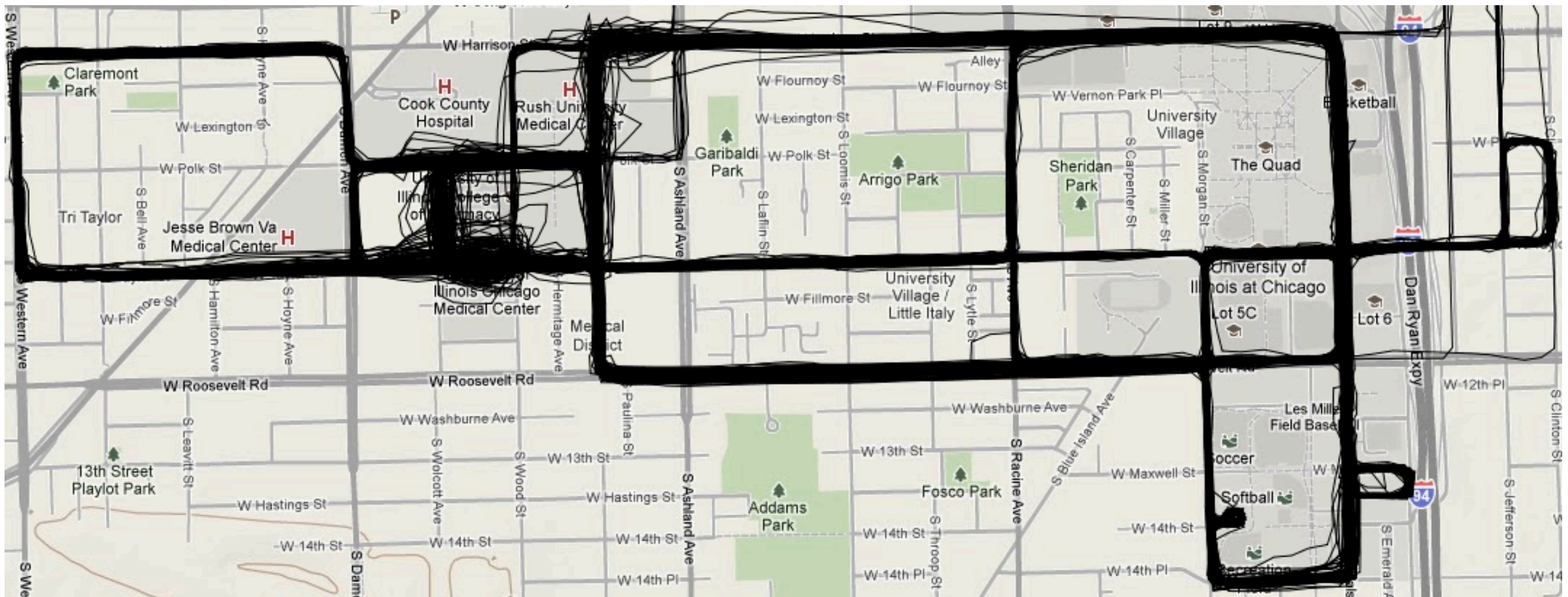
Link seeds



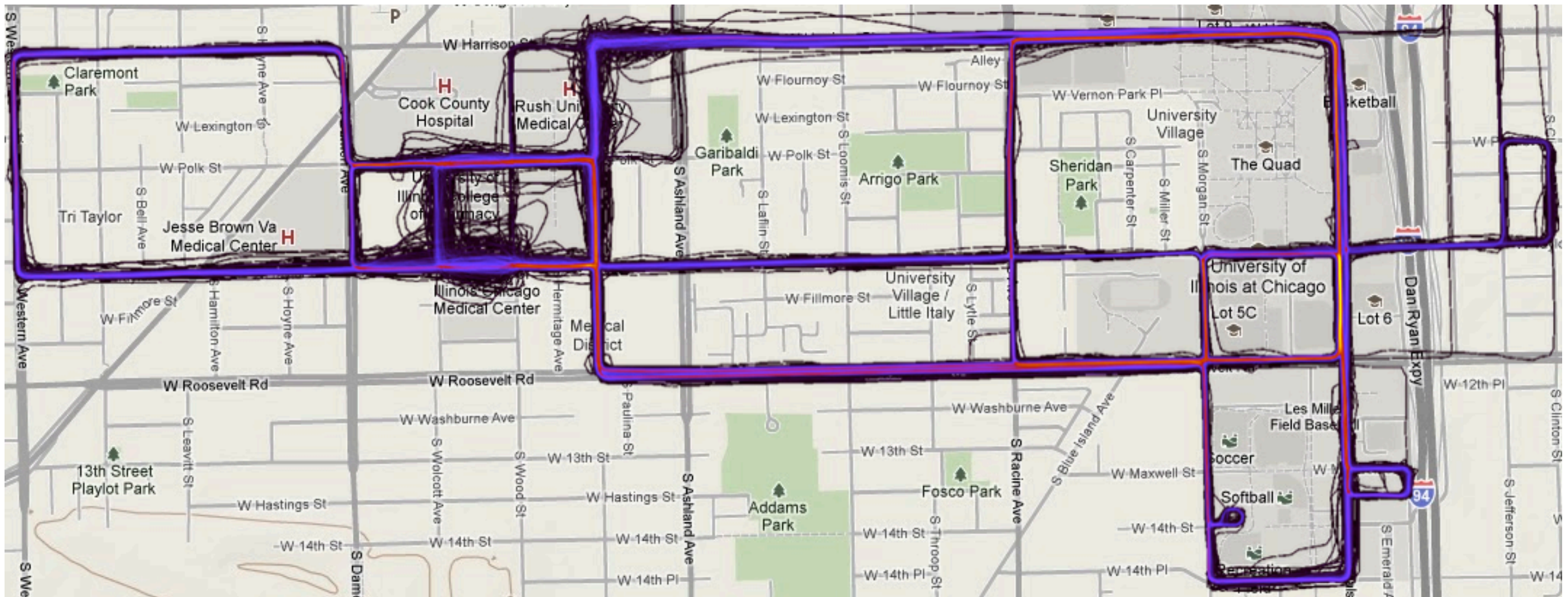
Kernel Density Estimation

Davies et al. (2006)

Raw GPS traces

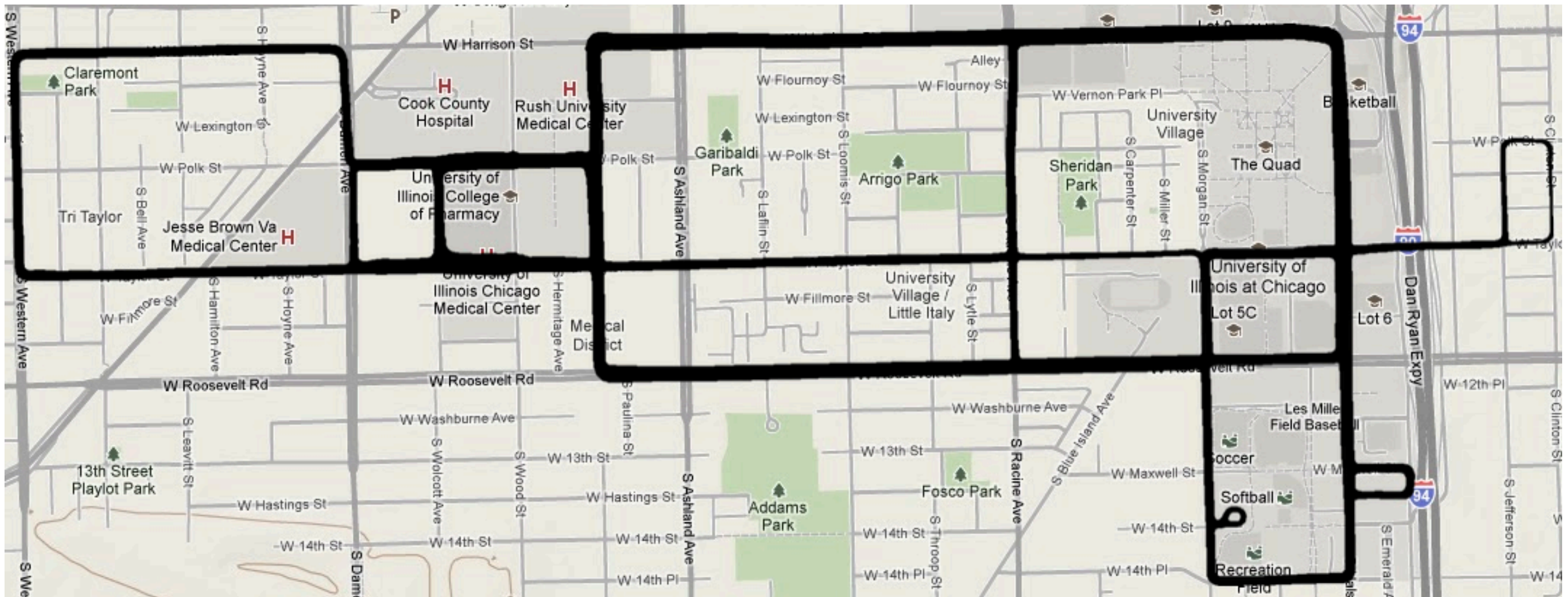


2-D histogram

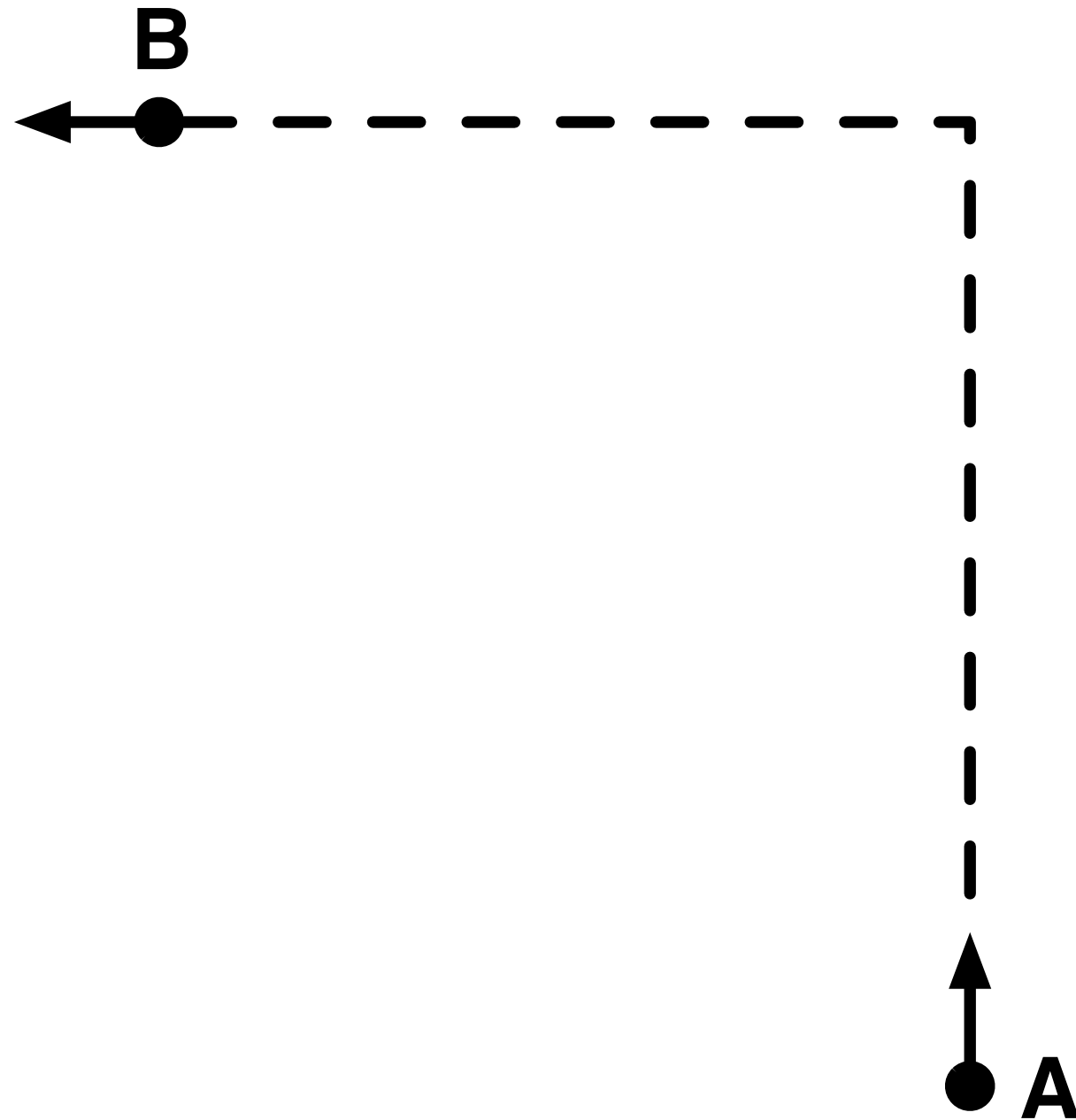


The map displays the University City neighborhood in Chicago, bounded by W Harrison St to the north and W 14th St to the south. Key streets shown include W Lexington St, W Polk St, W Fillmore St, W Roosevelt Rd, W Washburne Ave, W 13th St, W Hastings St, W 14th St, and W 14th Pl. Major landmarks include Cook County Hospital, Rush University Medical Center, University Village, and the University of Illinois at Chicago. Parks such as Claremont Park, Garibaldi Park, Addams Park, and Fosco Park are also marked. The map is overlaid with a thick black outline and a blue outline, indicating specific areas of interest.

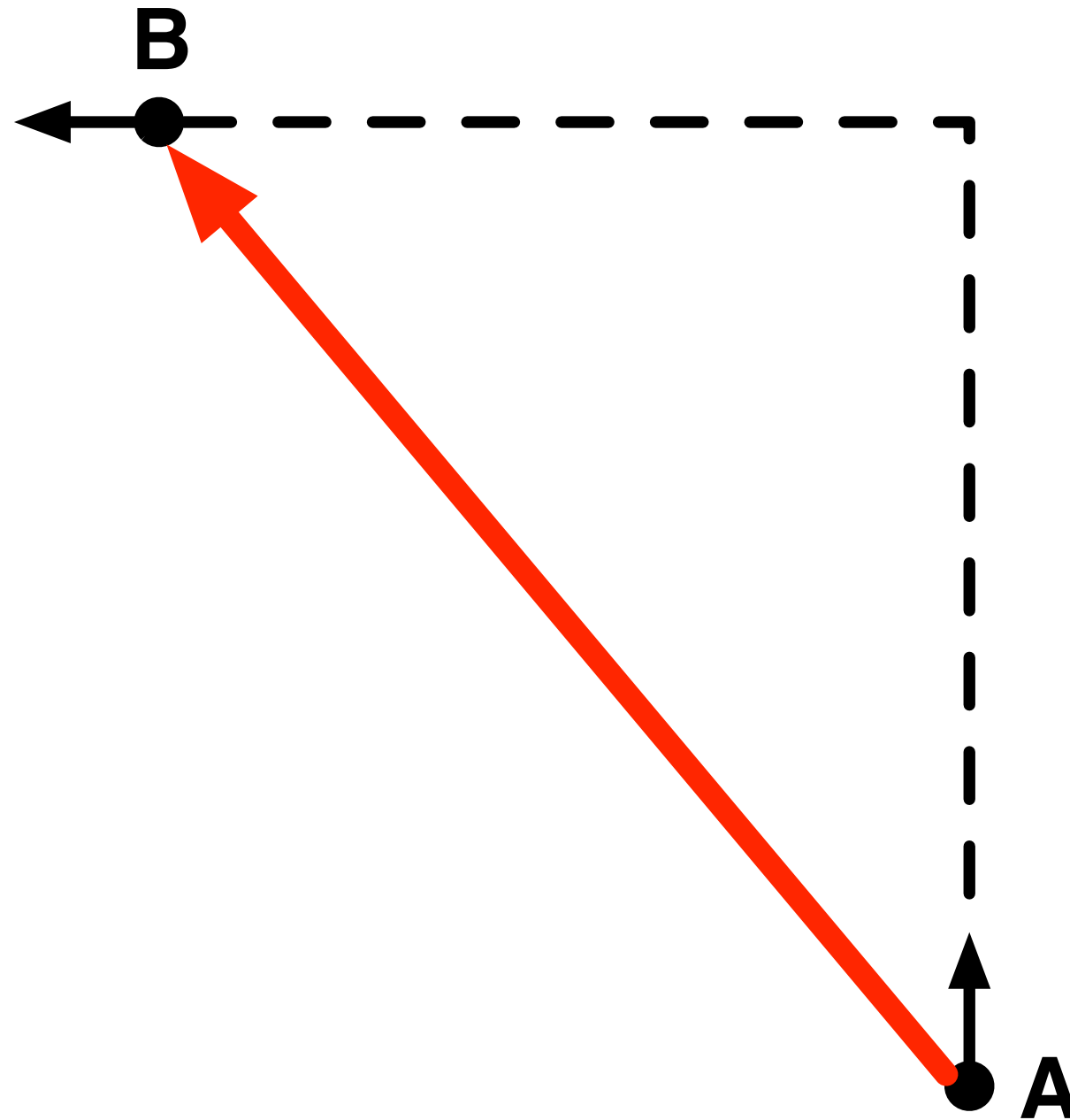
Thresholded image



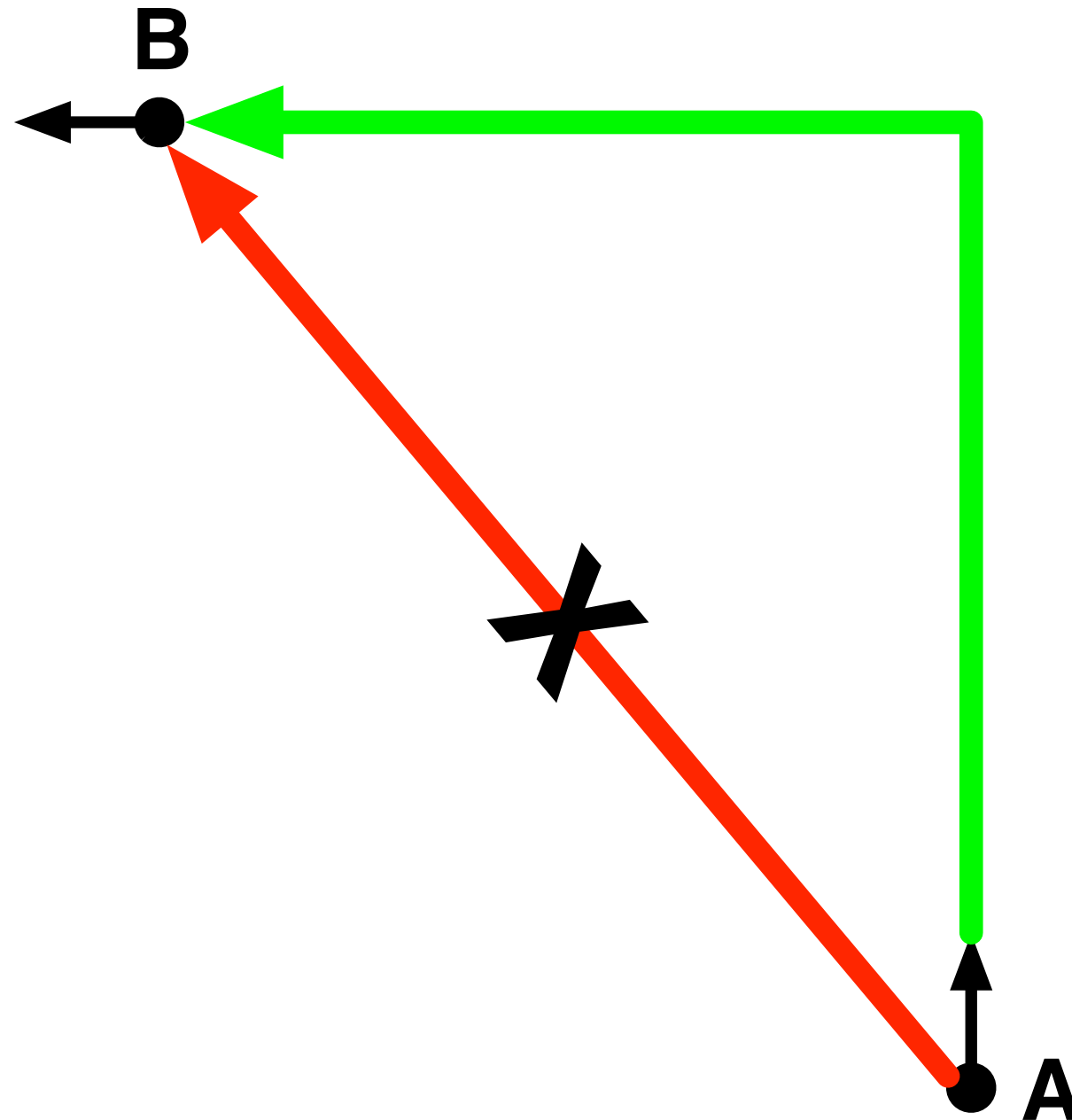
Sparse GPS samples



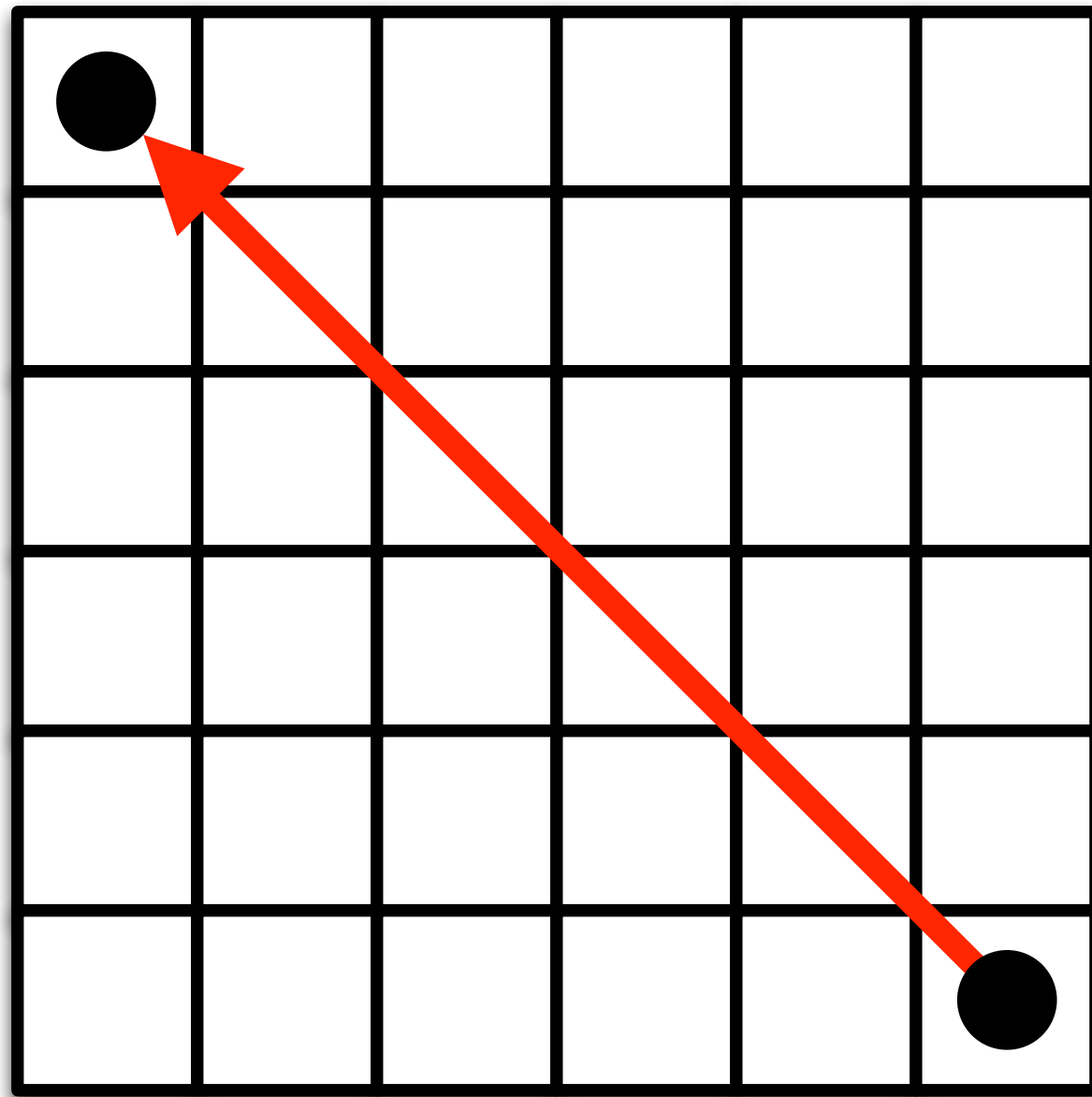
Incorrect trajectory



Actual trajectory



Current method



Current result

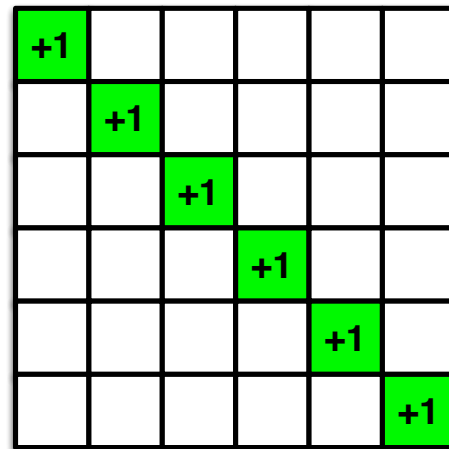
+1					
	+1				
		+1			
			+1		
				+1	
					+1

Proposed result

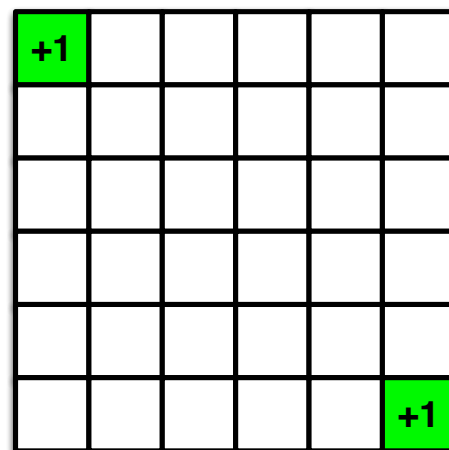
+1					
					+1

KDE variants

- ▶ KDE “lines”



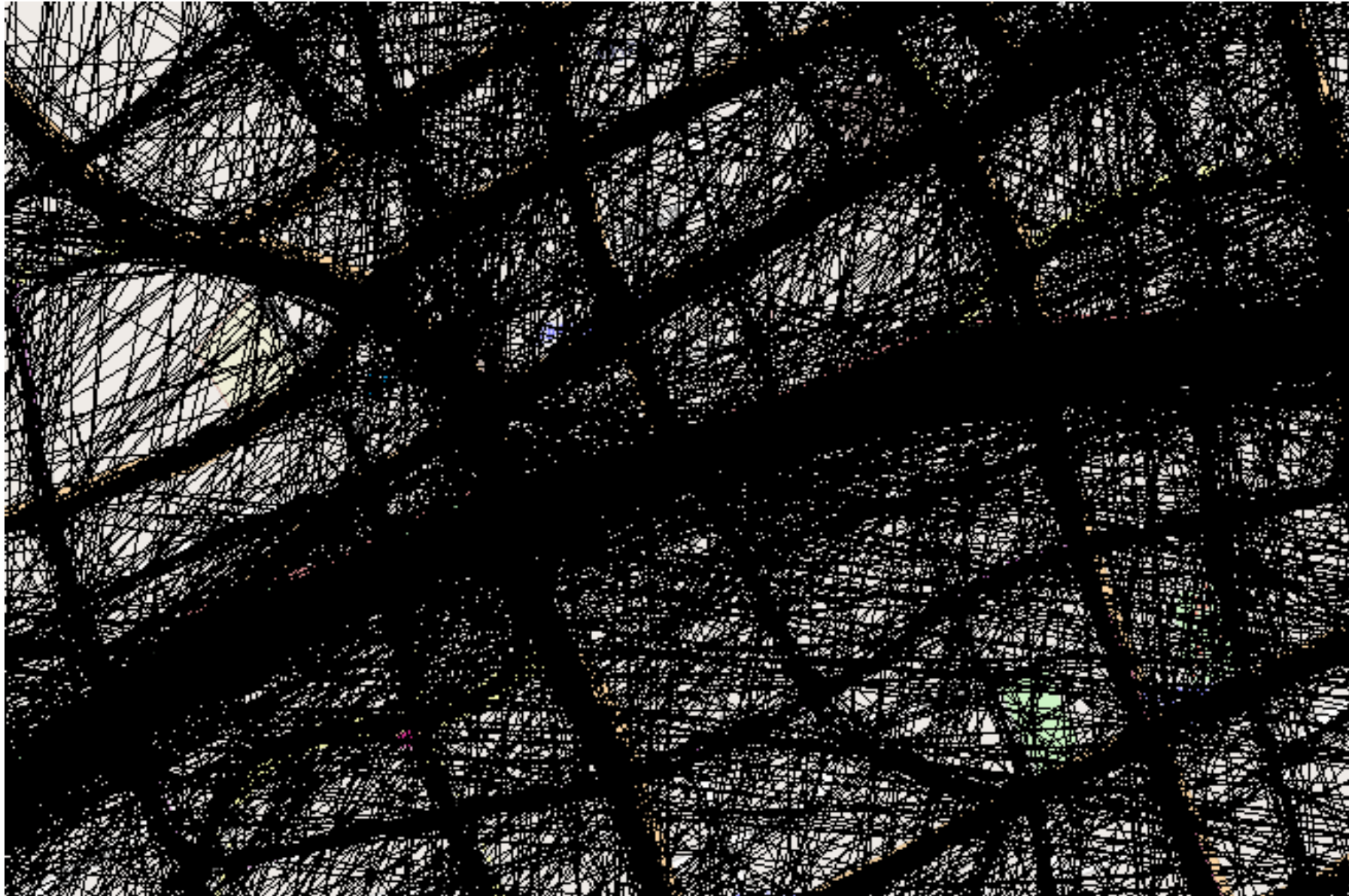
- ▶ KDE “points”



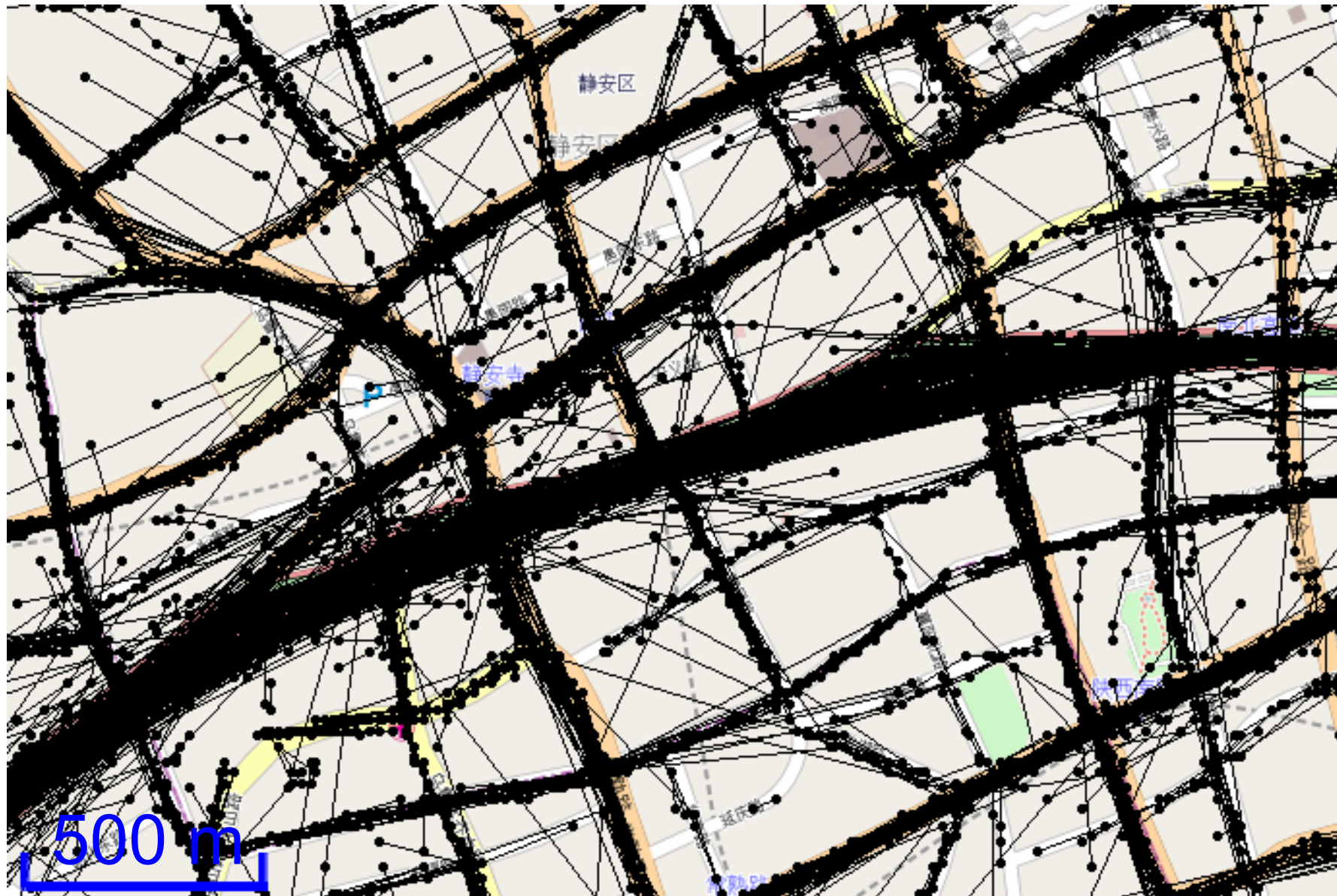
Trace Merging

Liu et al. (2012), “TC1”

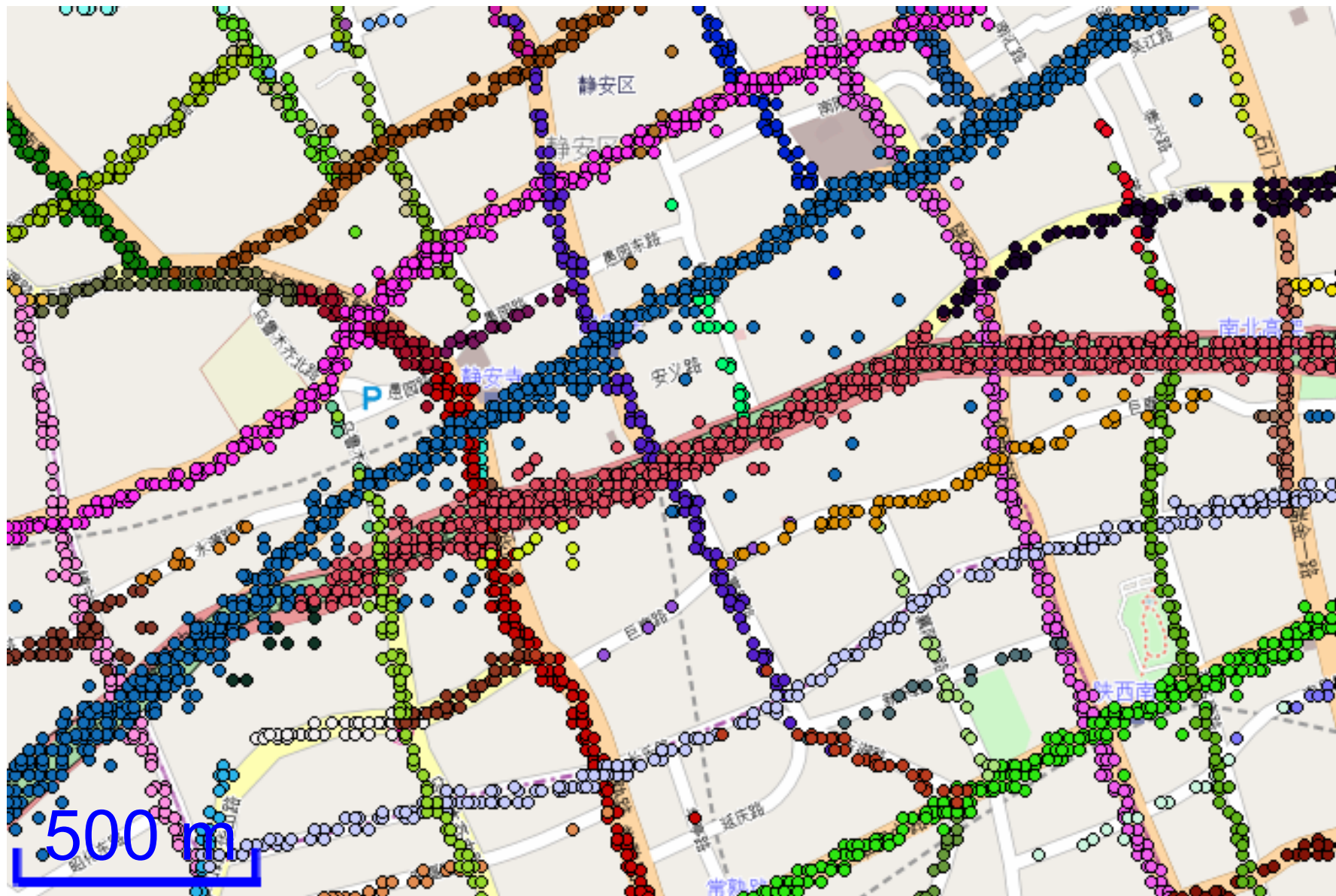
Raw GPS traces



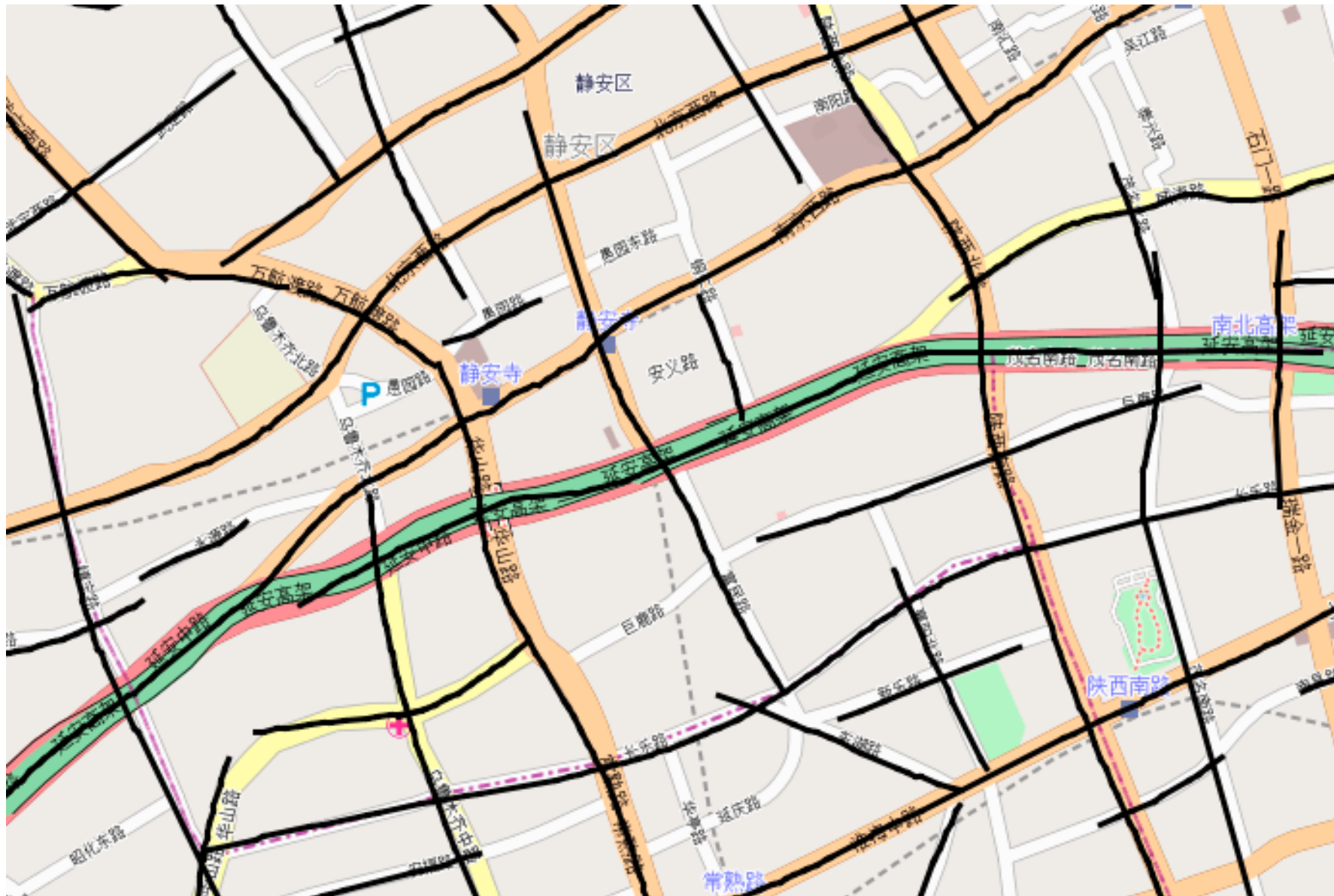
Segment selection



Clustering

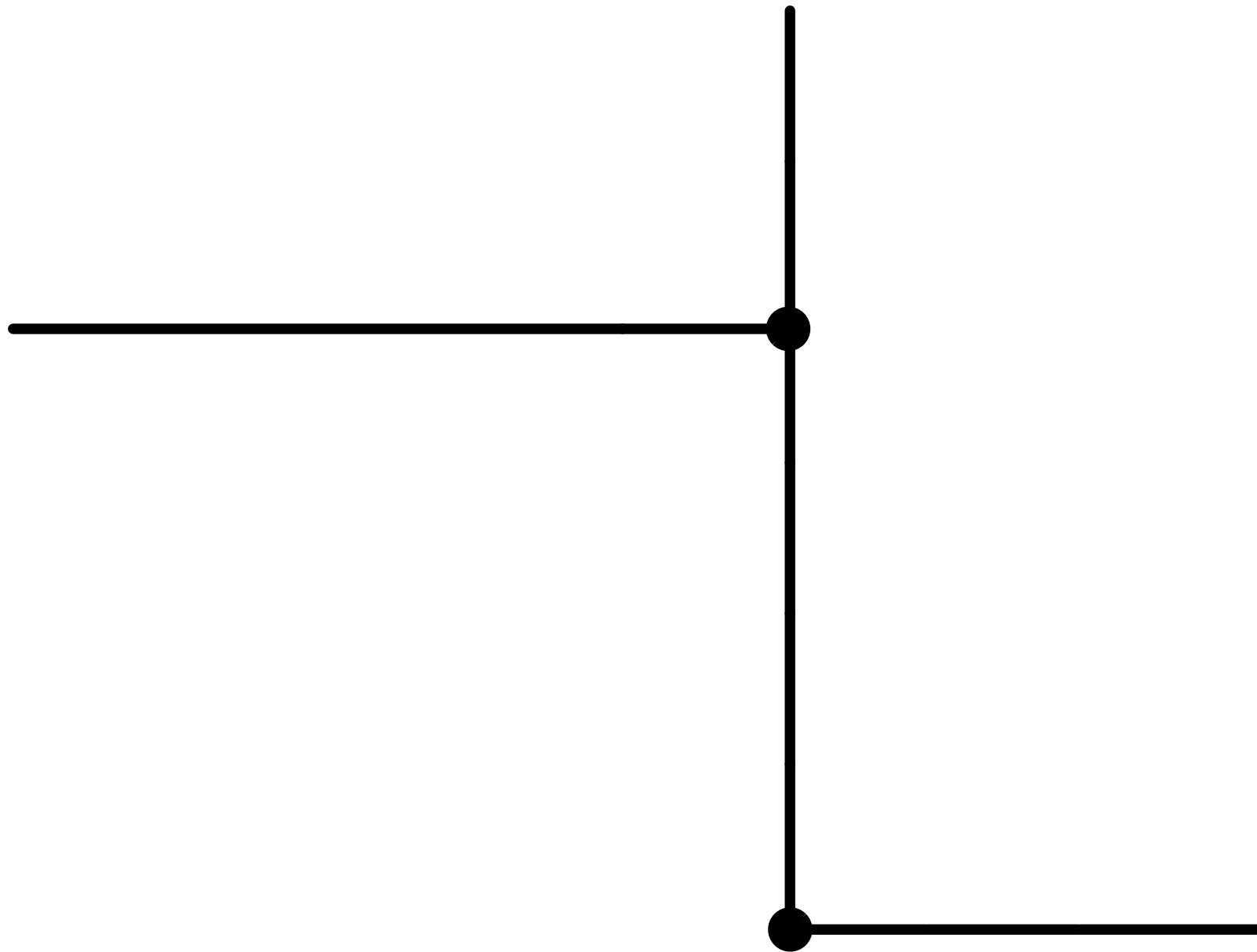


Map extraction

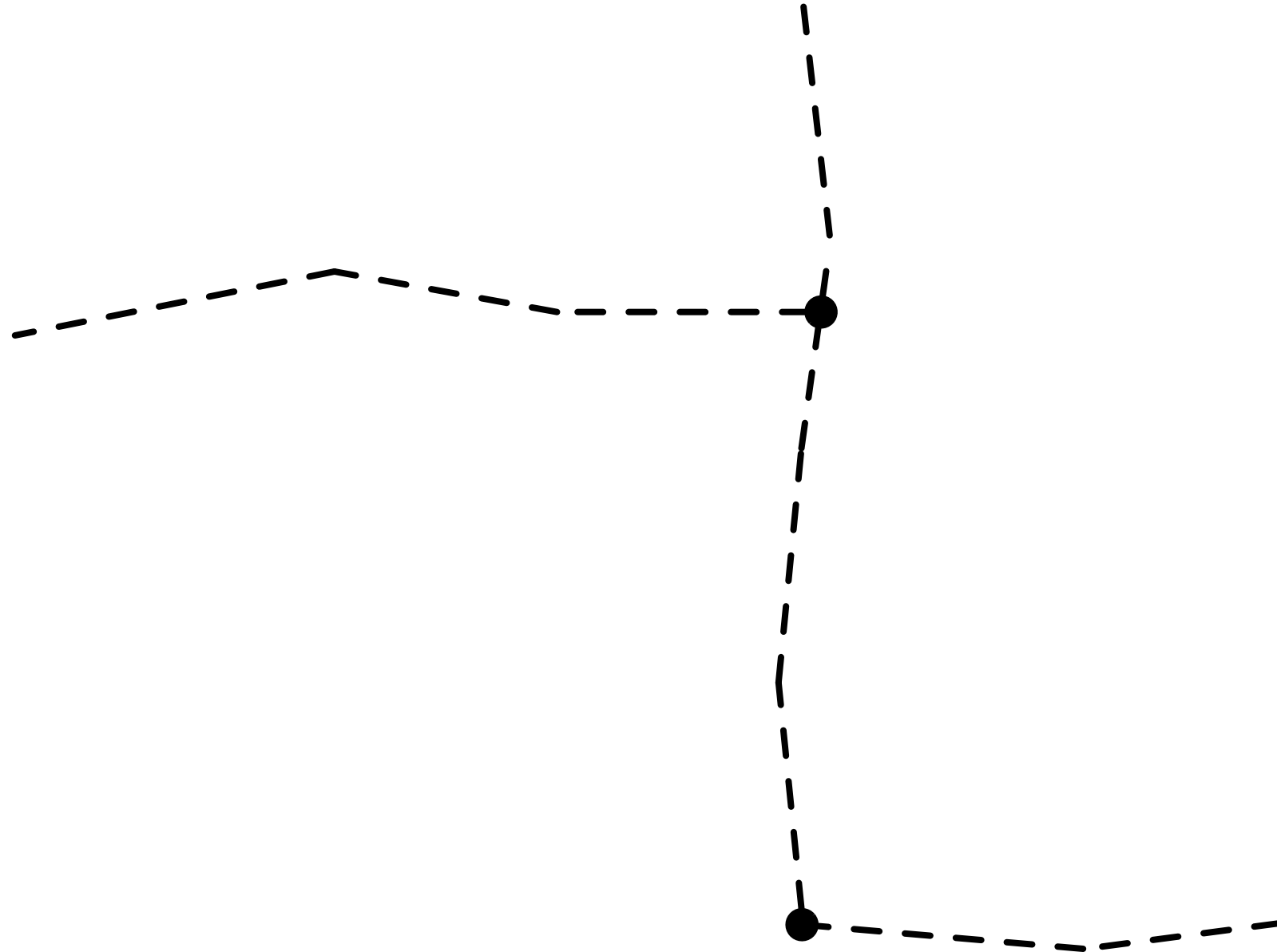


Quantitative Evaluation

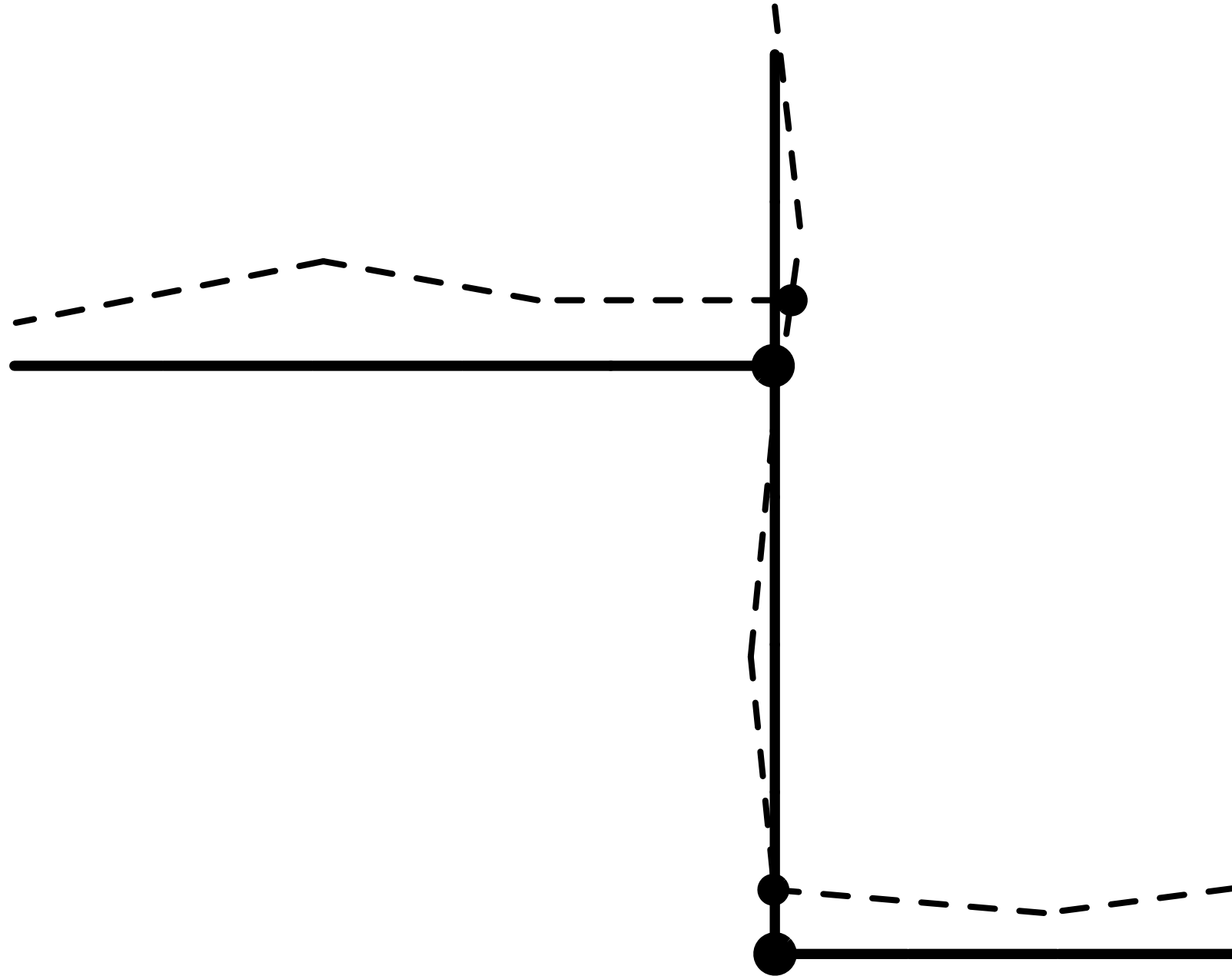
Ground truth map



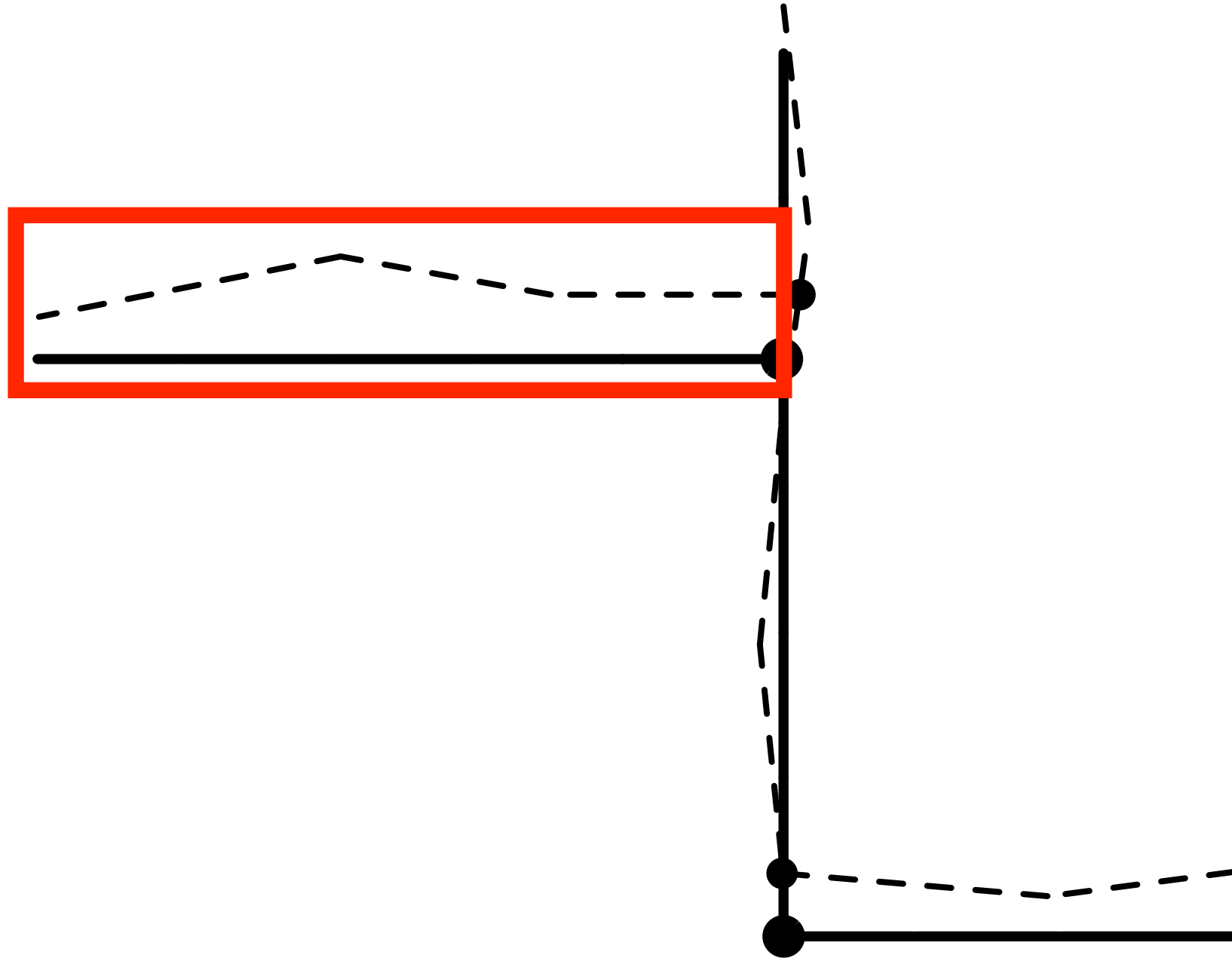
Inferred map



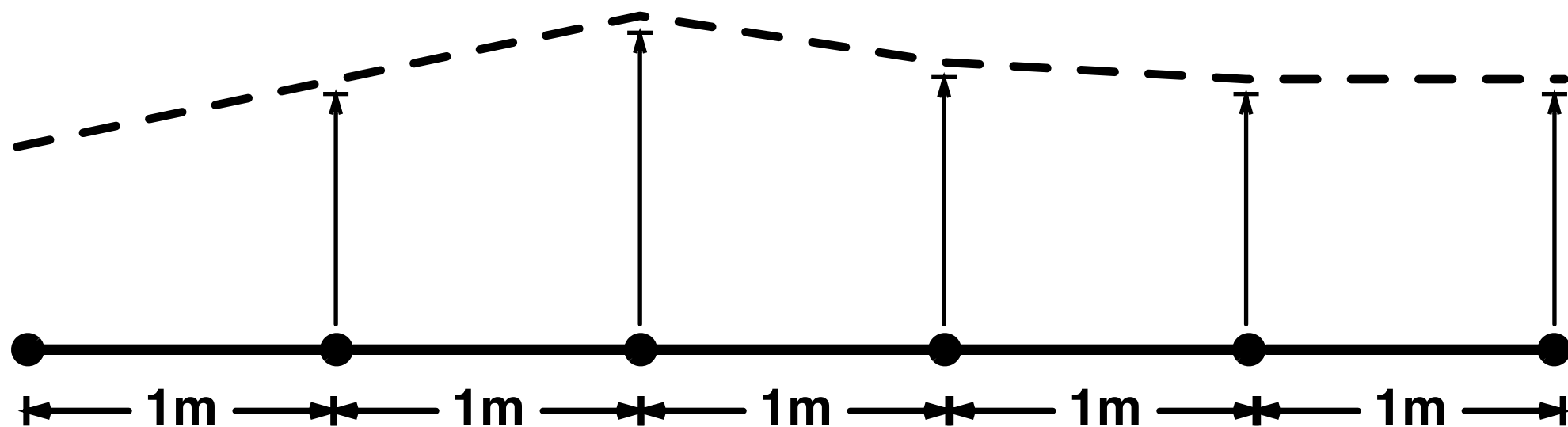
Overlaid maps



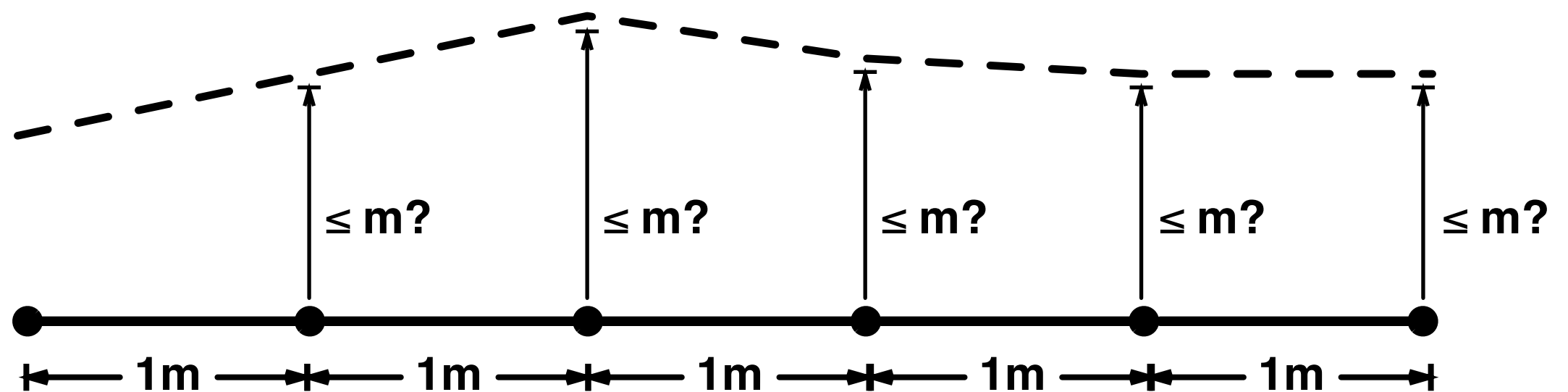
Overlaid maps



True positive length



True positive length



true positive length = $\# \leq m$

Evaluation metrics

$||Inferred|| = \text{total inferred road length (m)}$

$||Ground Truth|| = \text{total ground truth road length (m)}$

Evaluation metrics

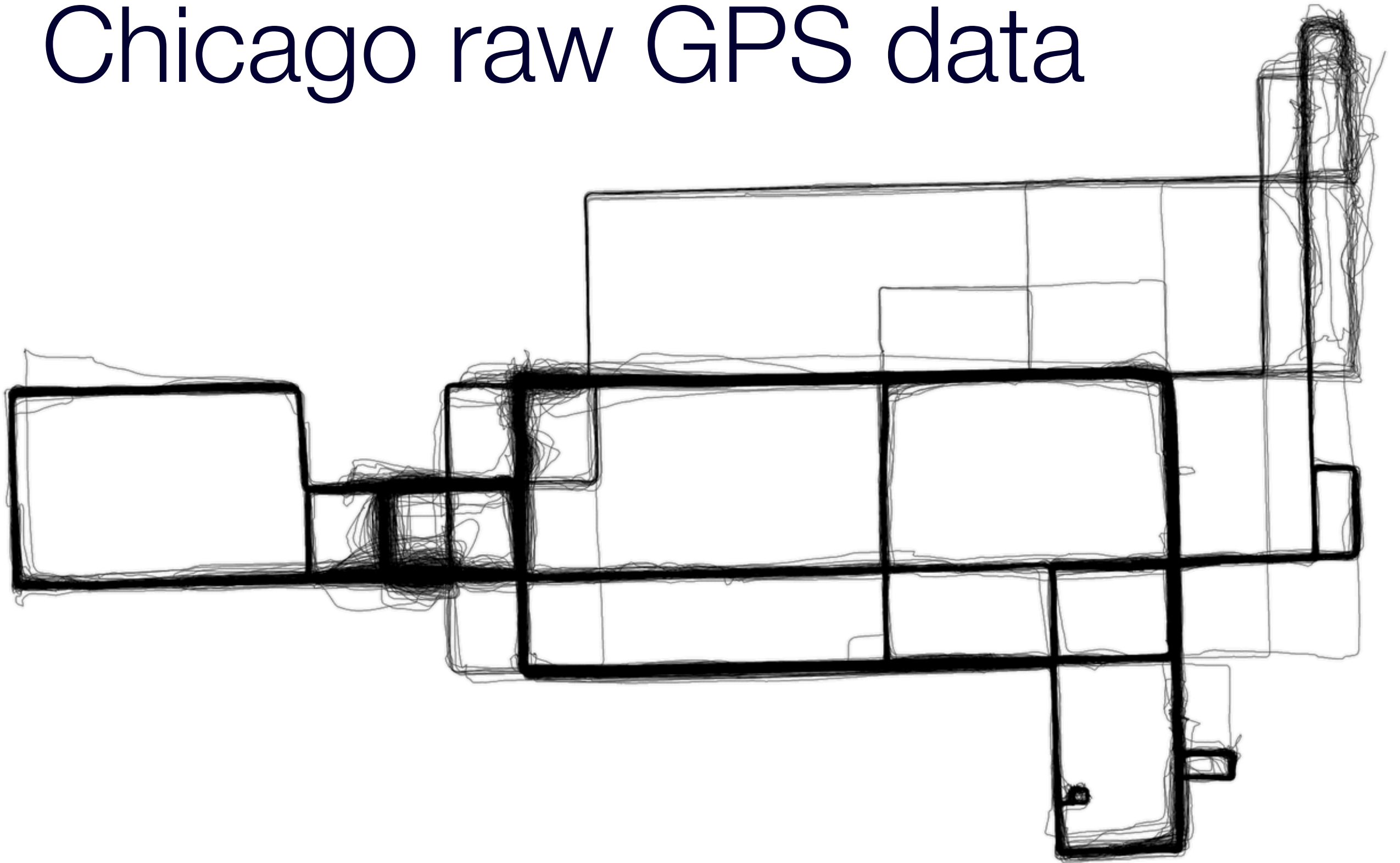
$tp = \text{true positive length}$

$$\text{precision} = \frac{tp}{||Inferred||}, \quad \text{recall} = \frac{tp}{||Ground Truth||}$$

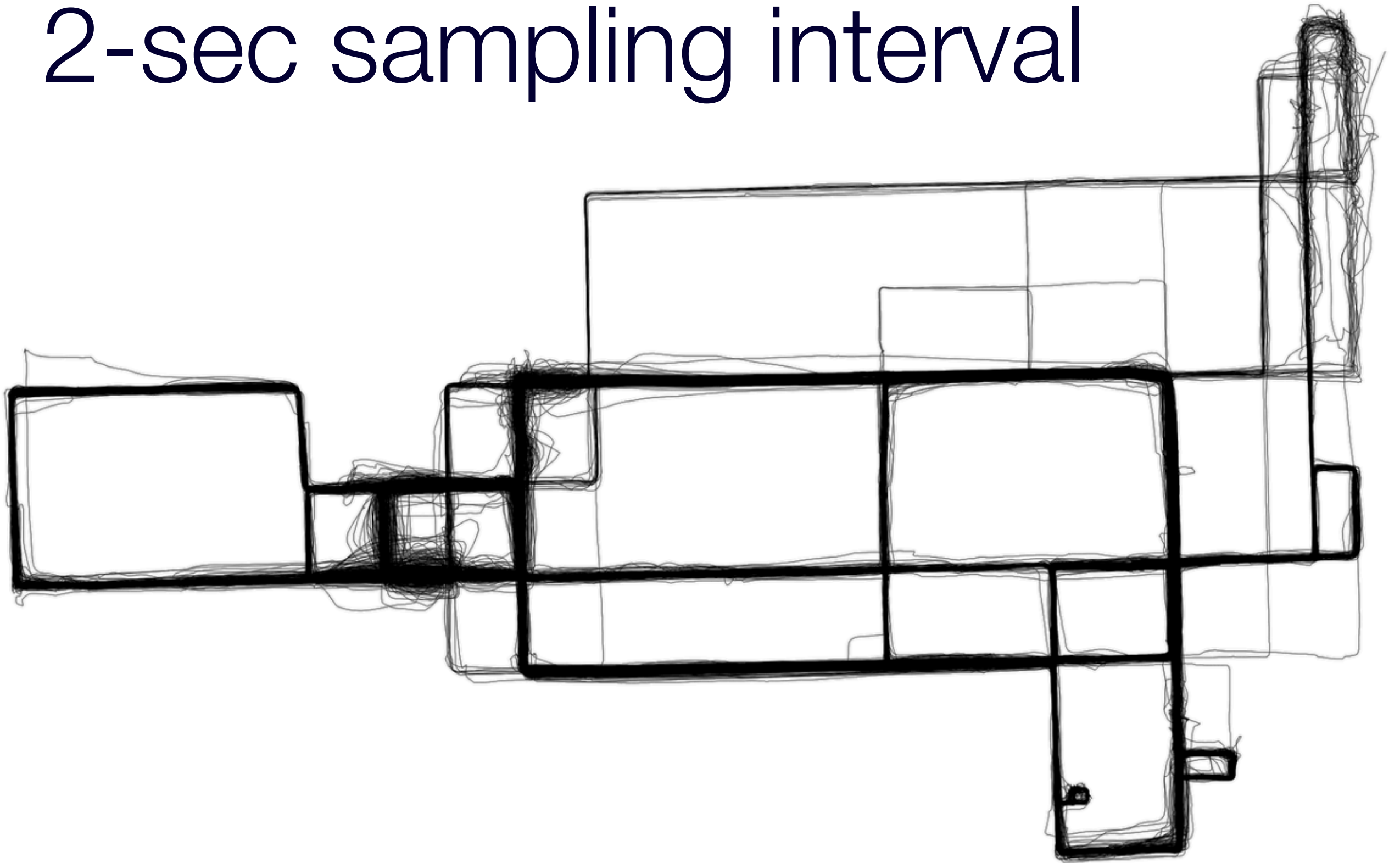
$$F\text{-measure} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Chicago Evaluation

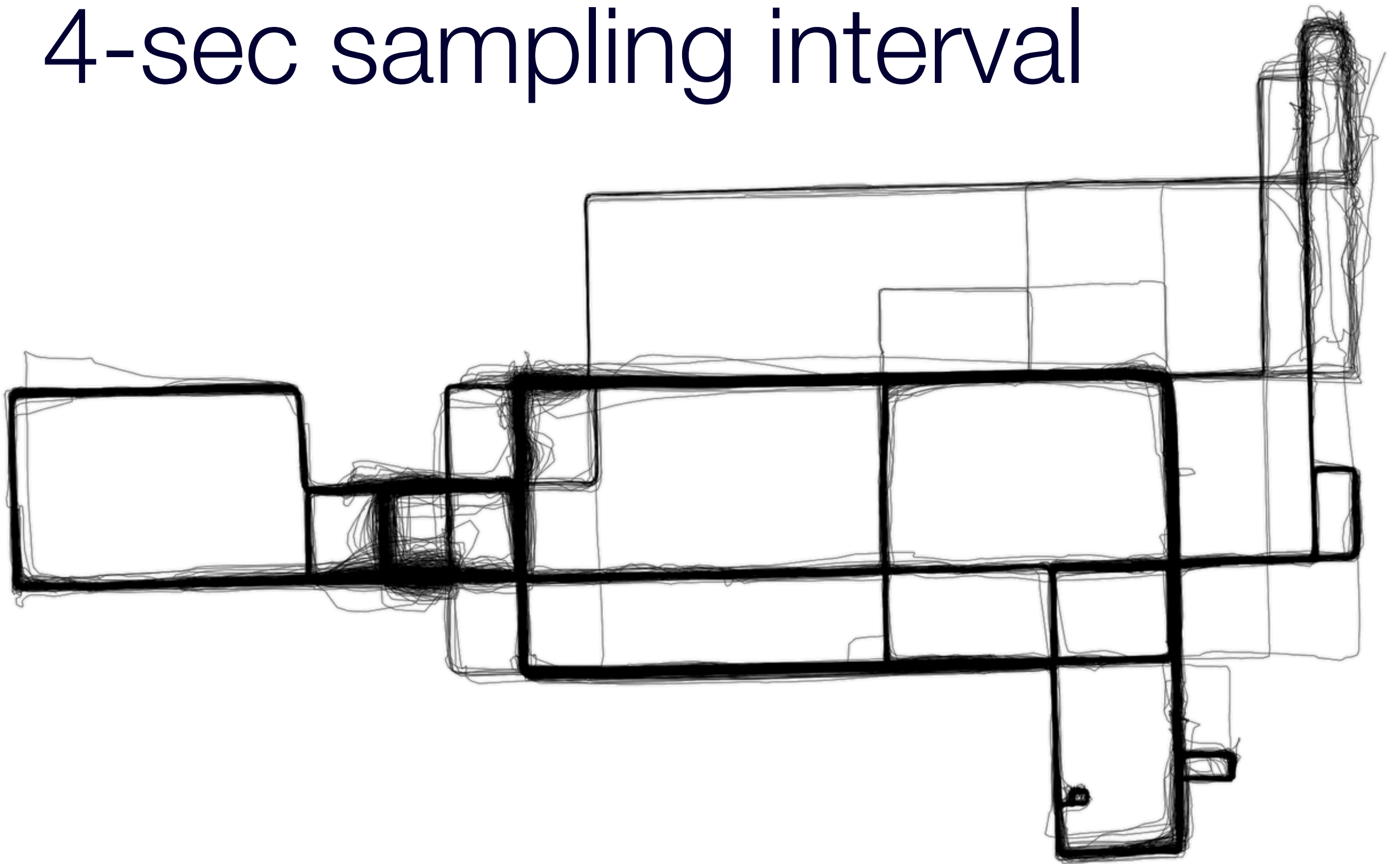
Chicago raw GPS data



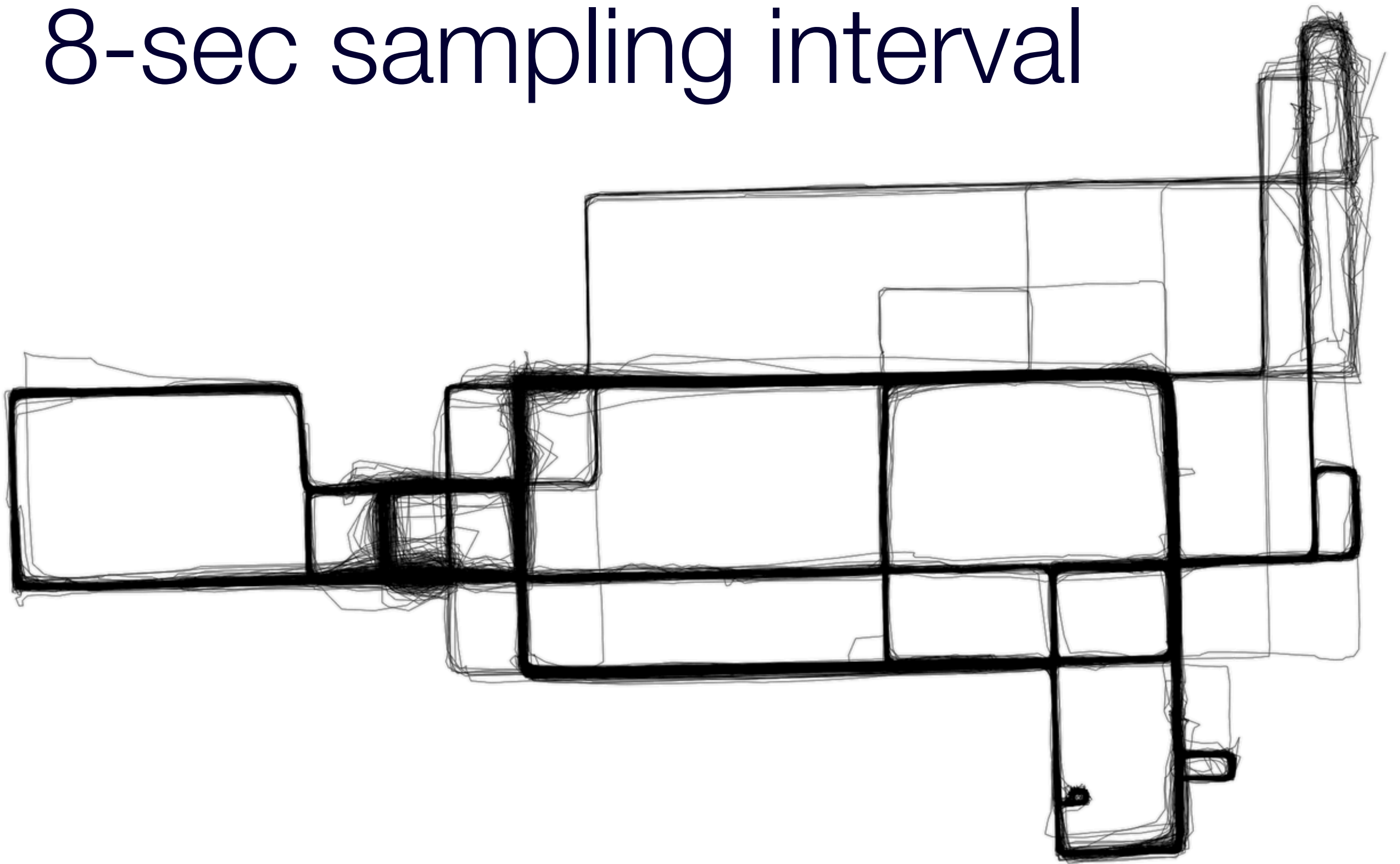
2-sec sampling interval



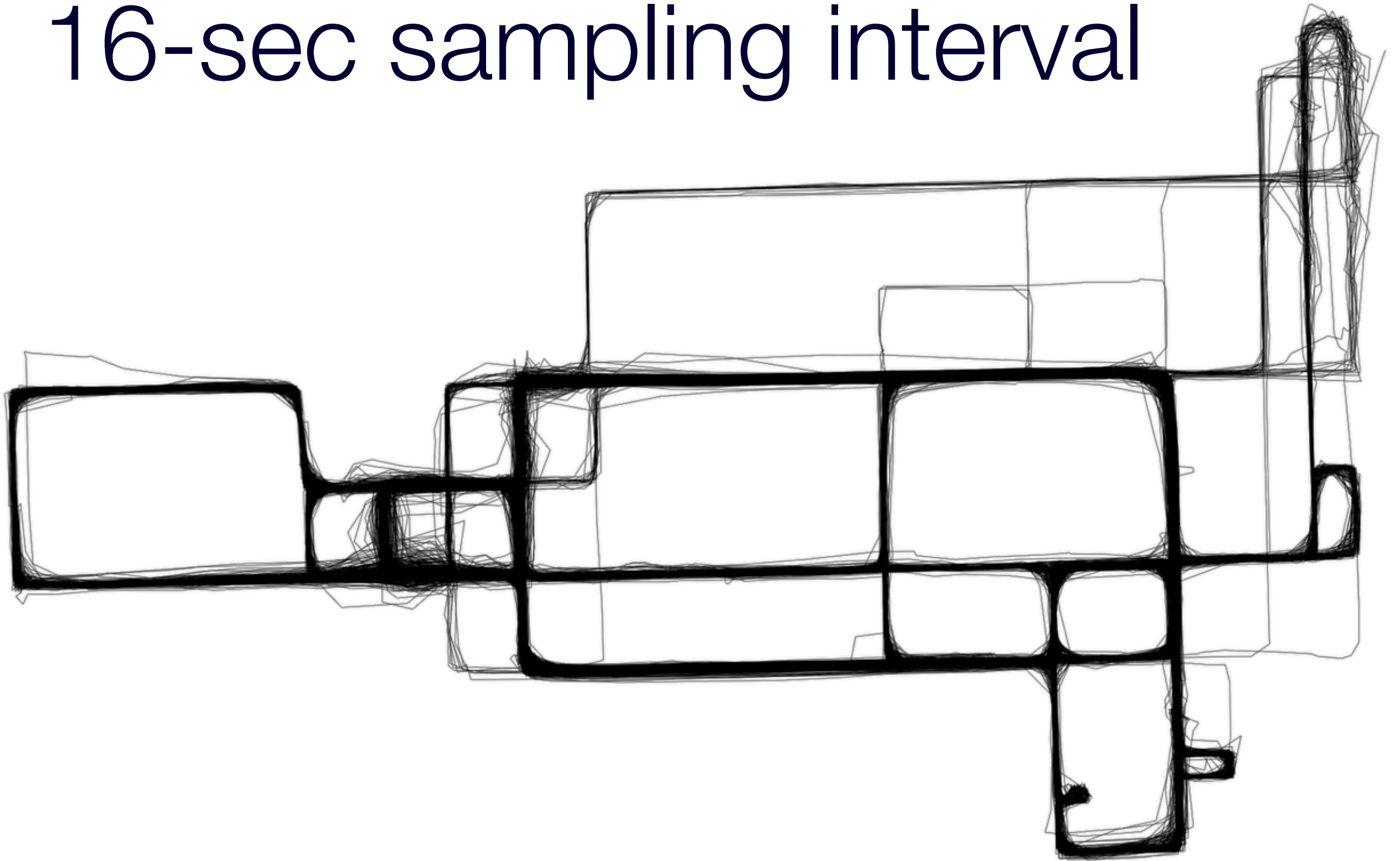
4-sec sampling interval



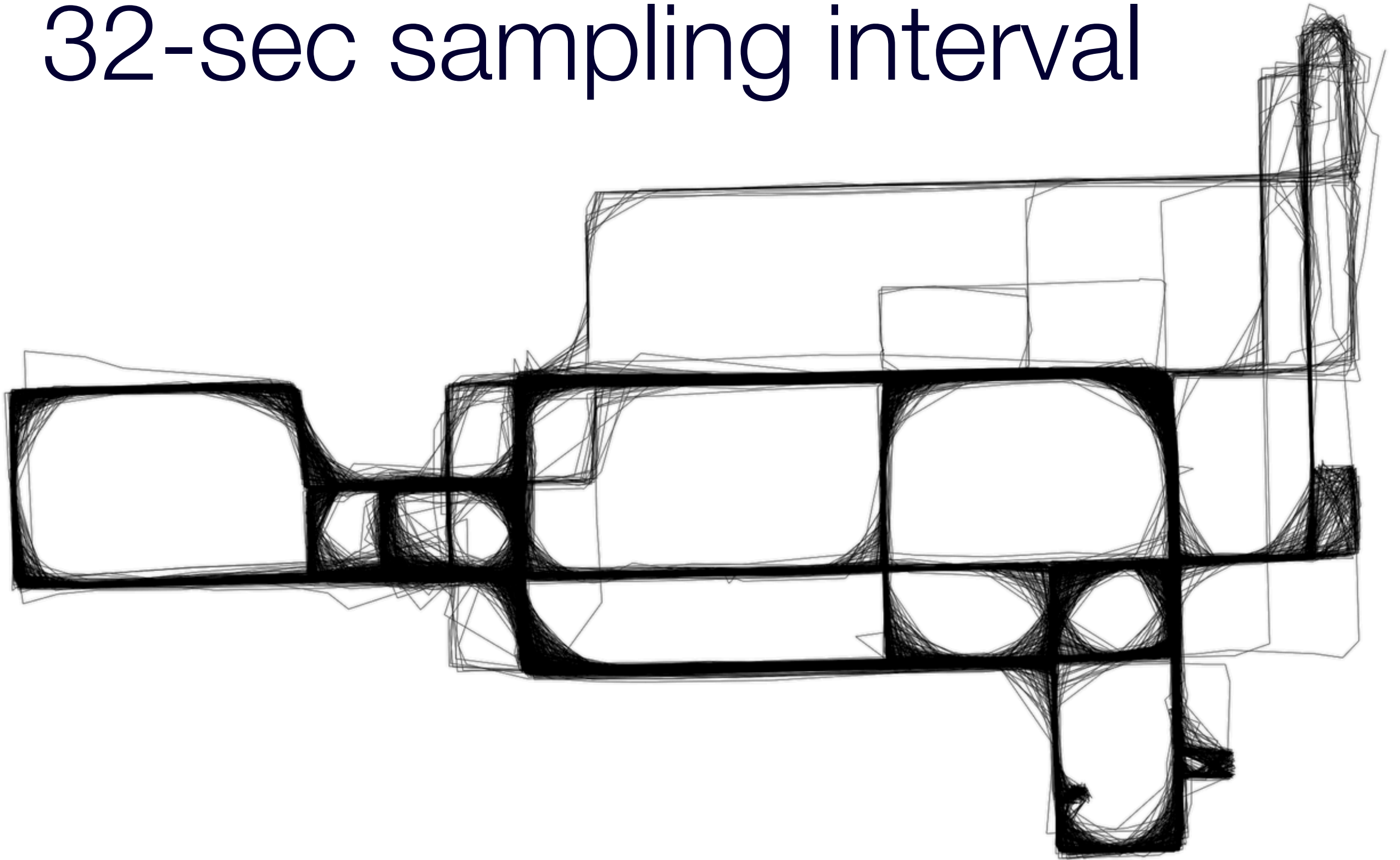
8-sec sampling interval



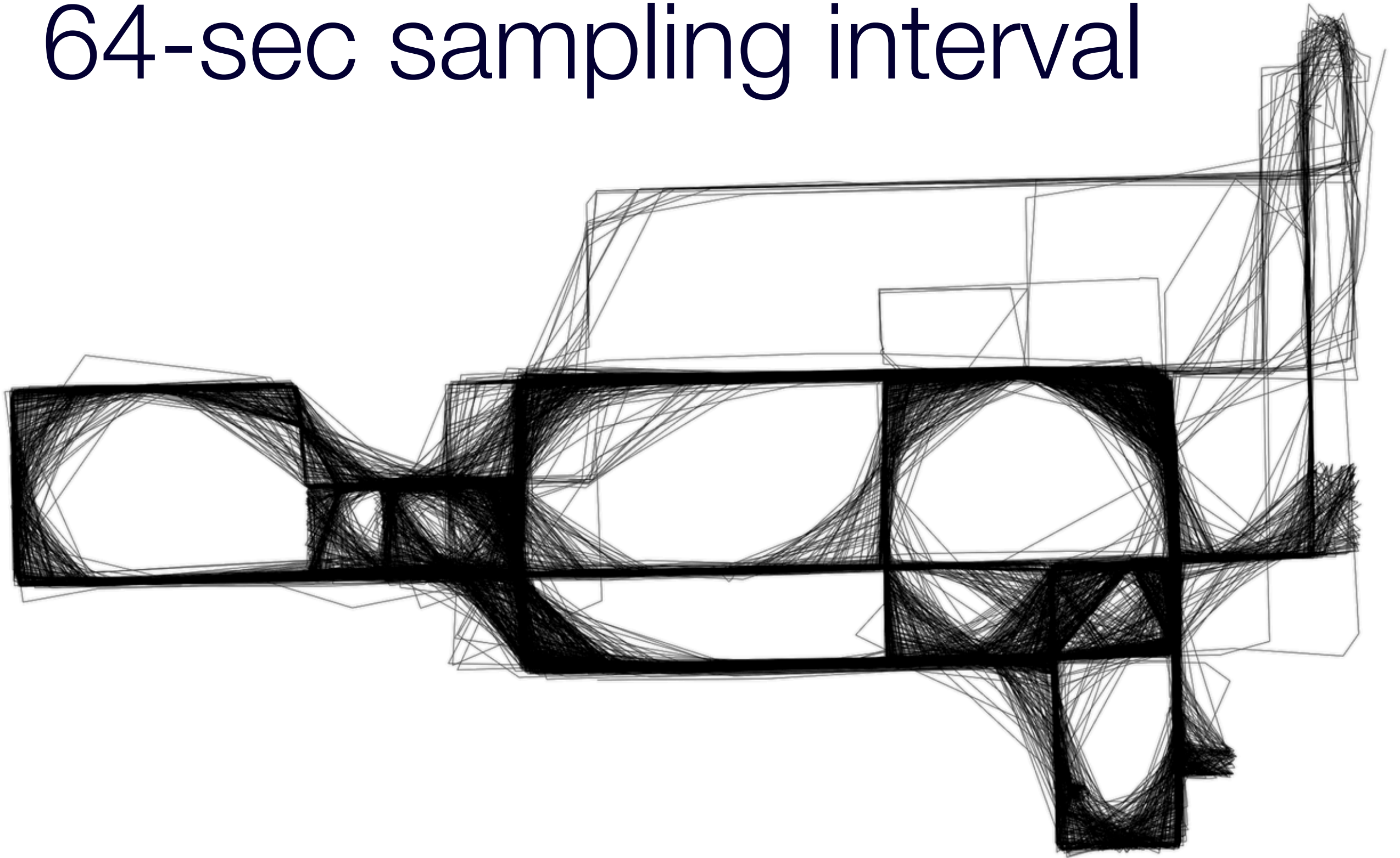
16-sec sampling interval



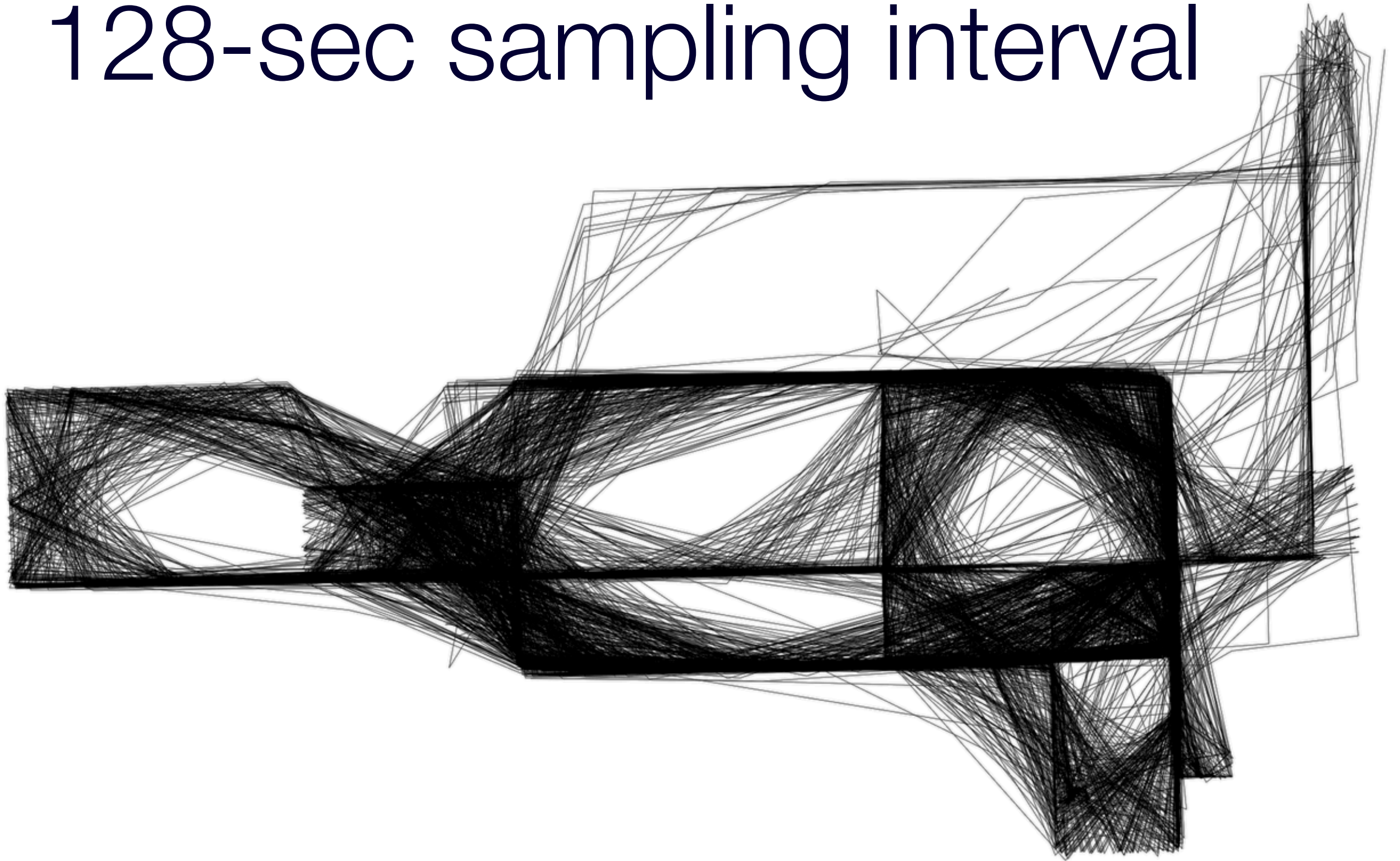
32-sec sampling interval



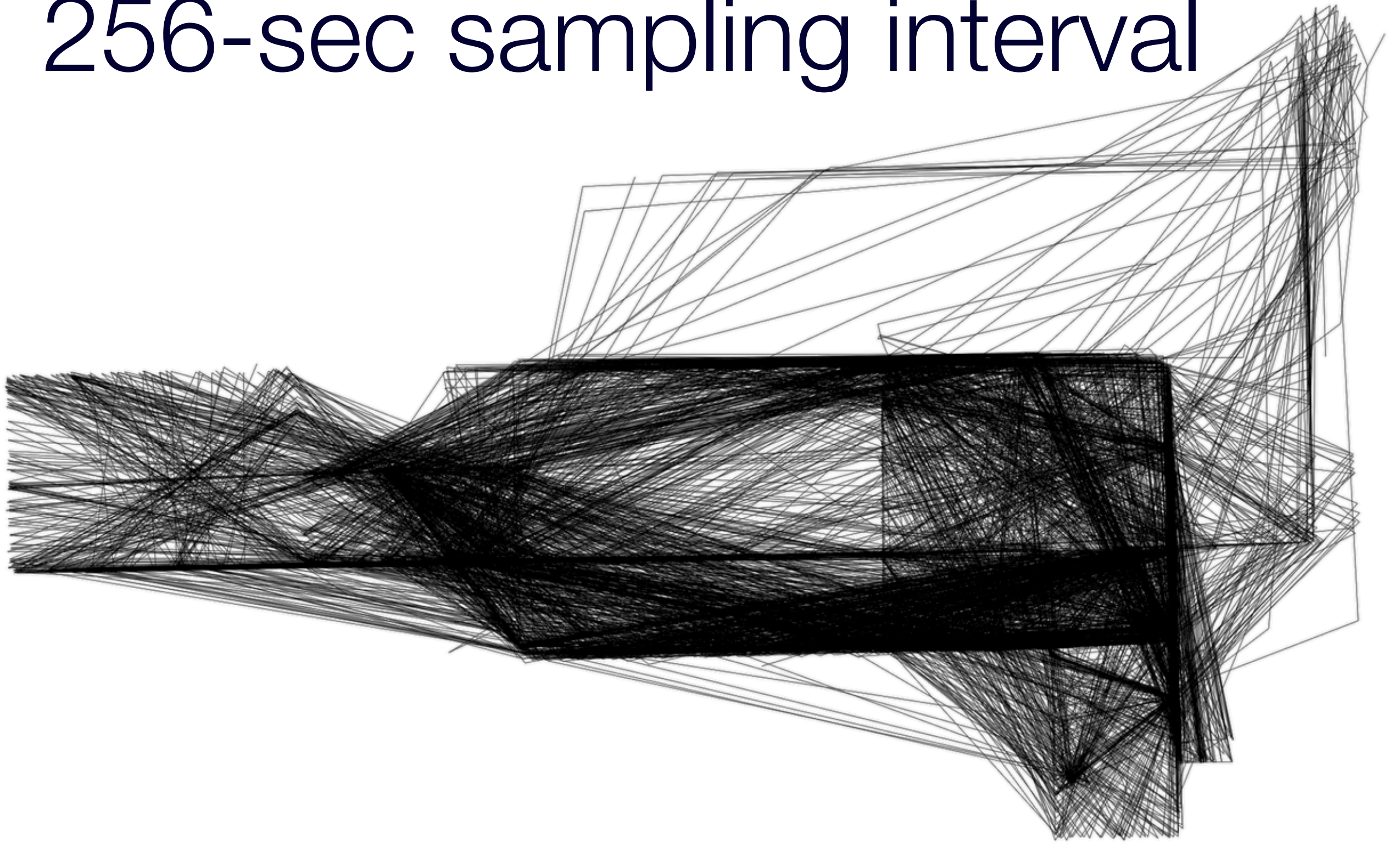
64-sec sampling interval



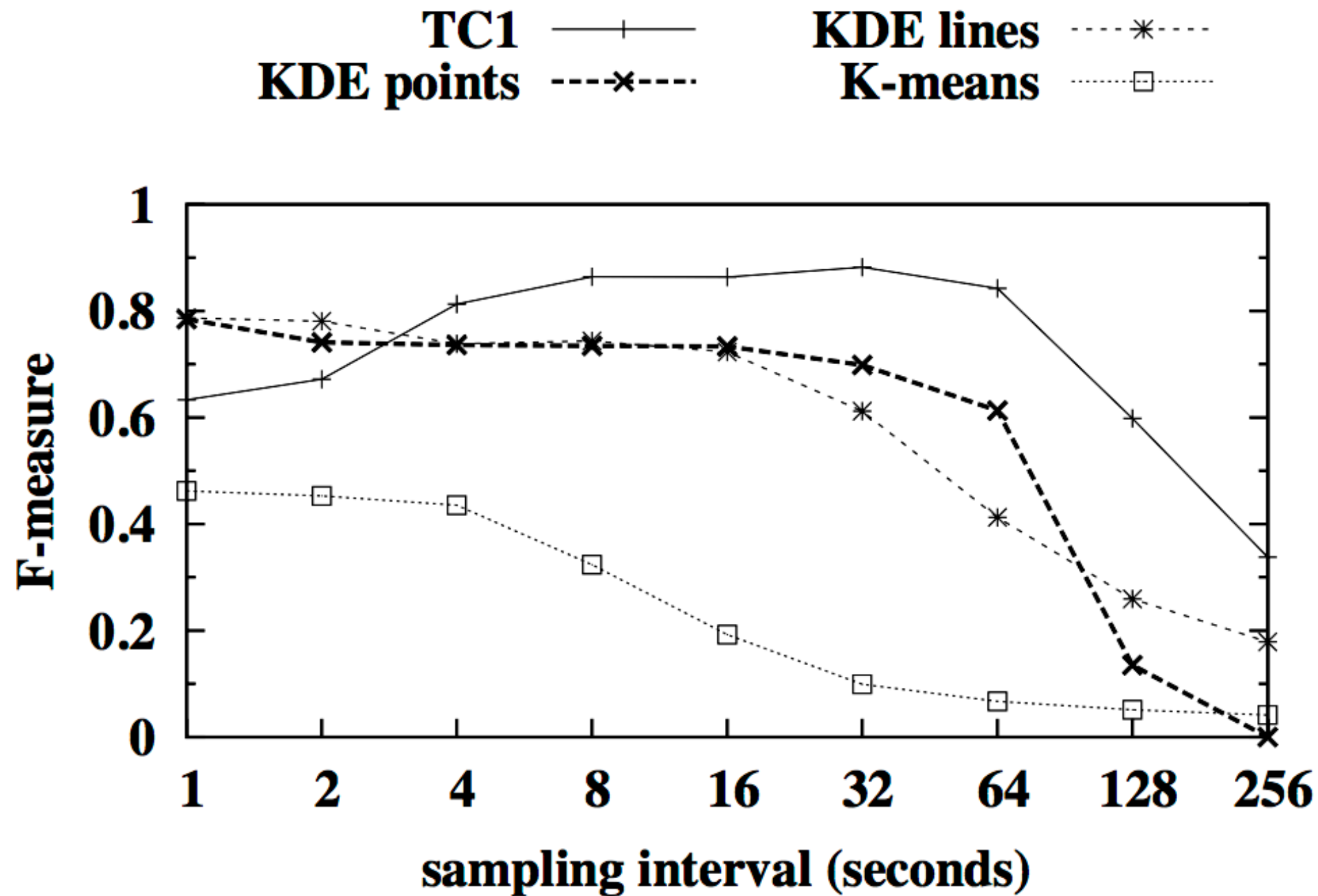
128-sec sampling interval



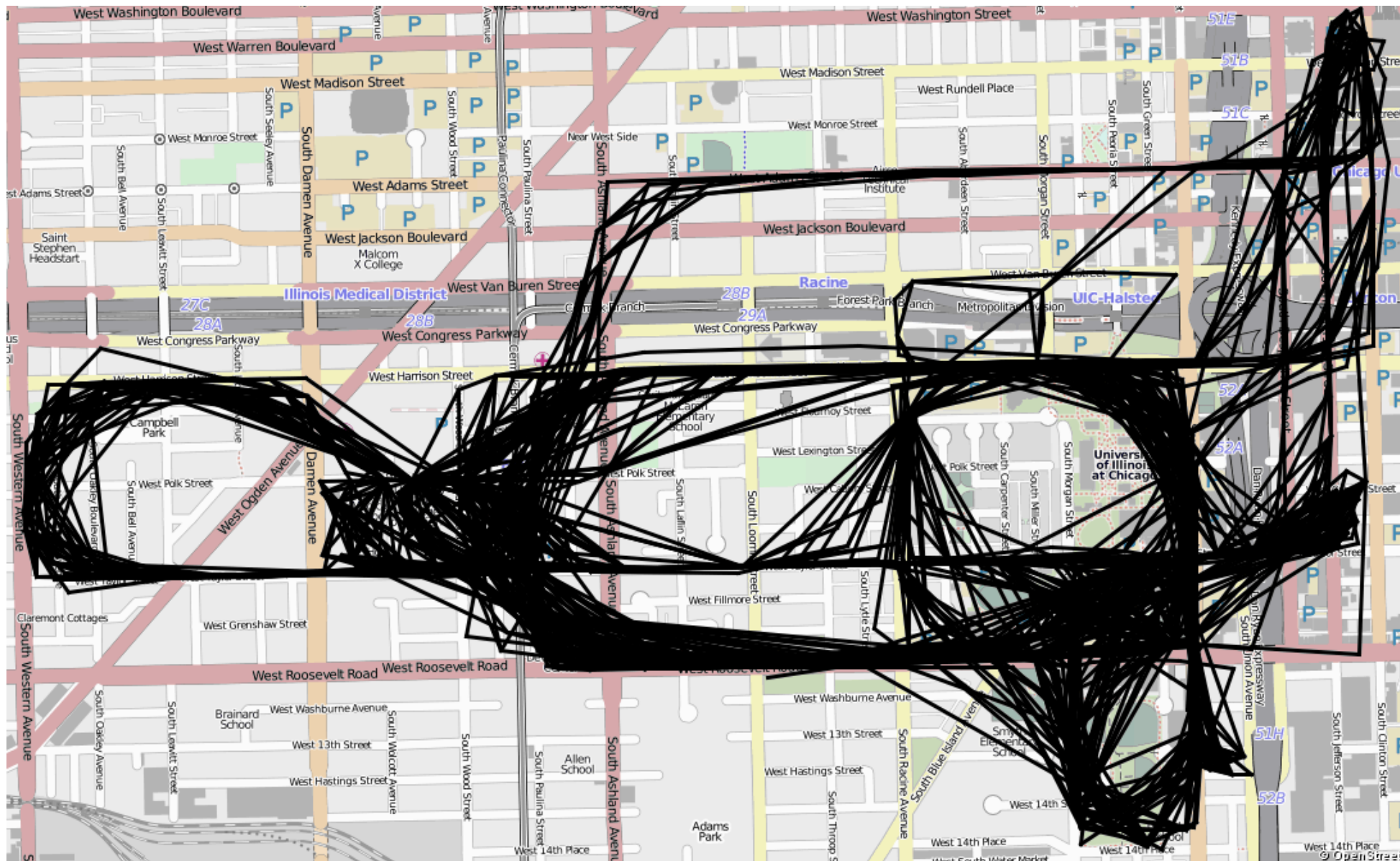
256-sec sampling interval



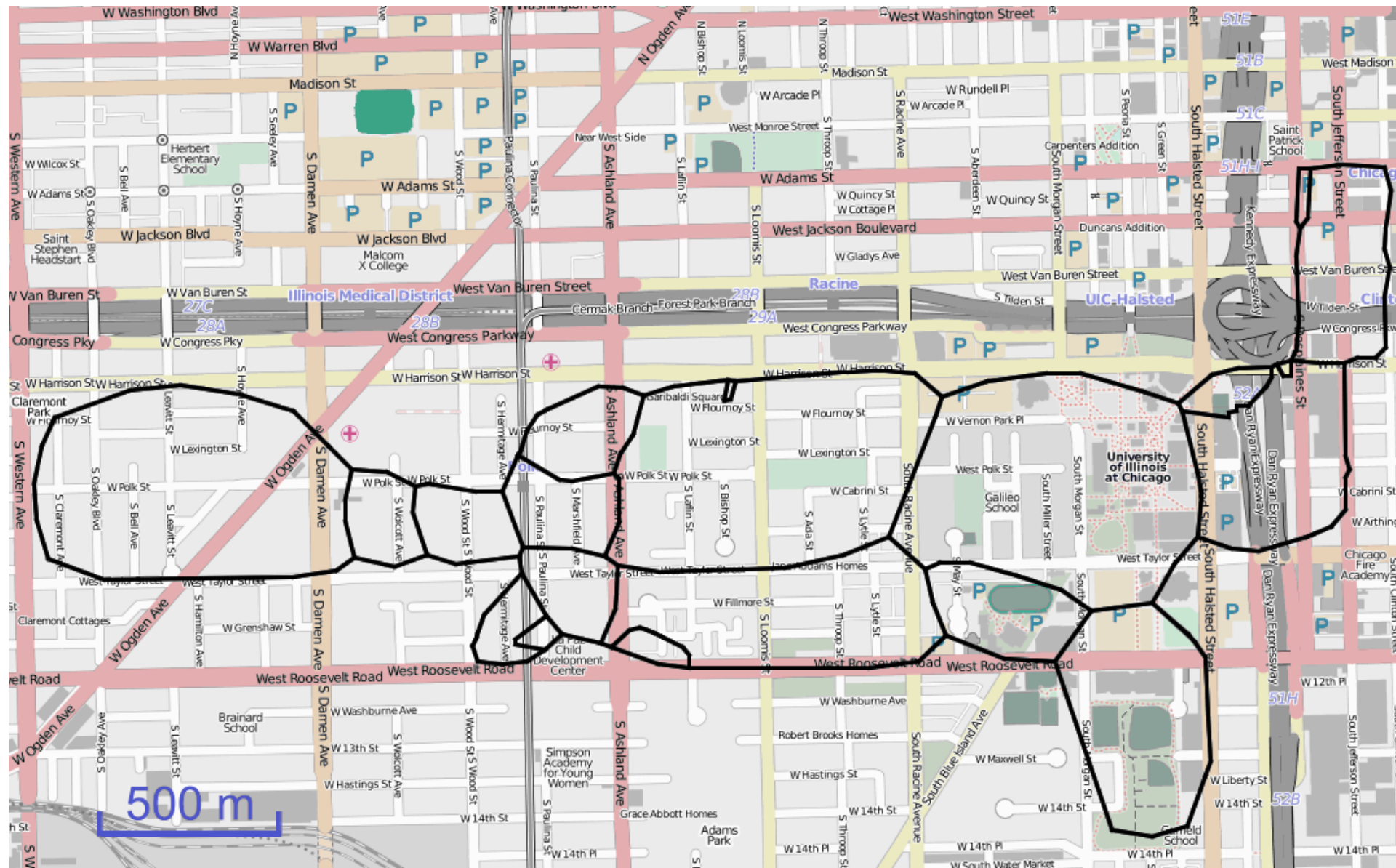
Overall comparison



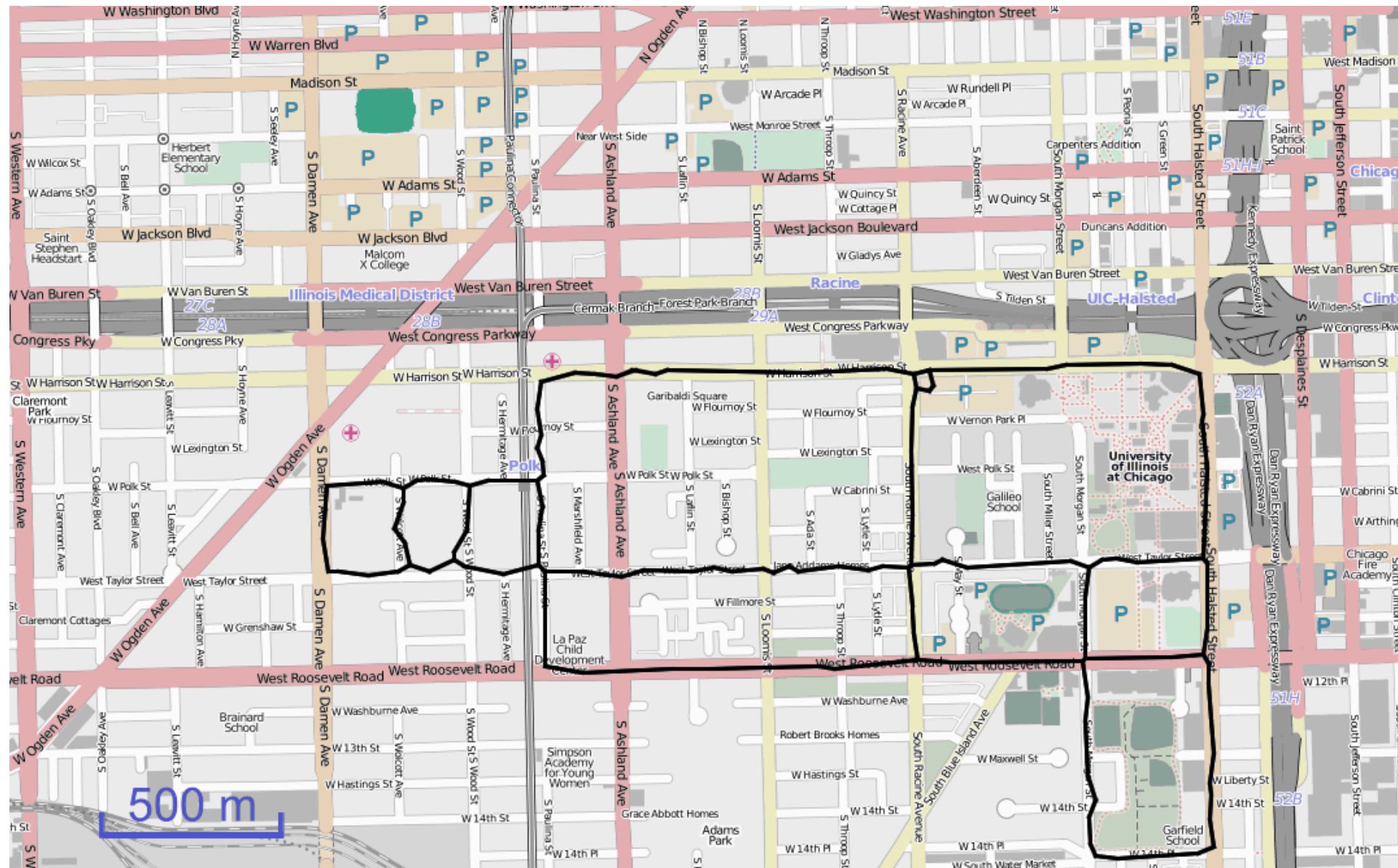
K-means



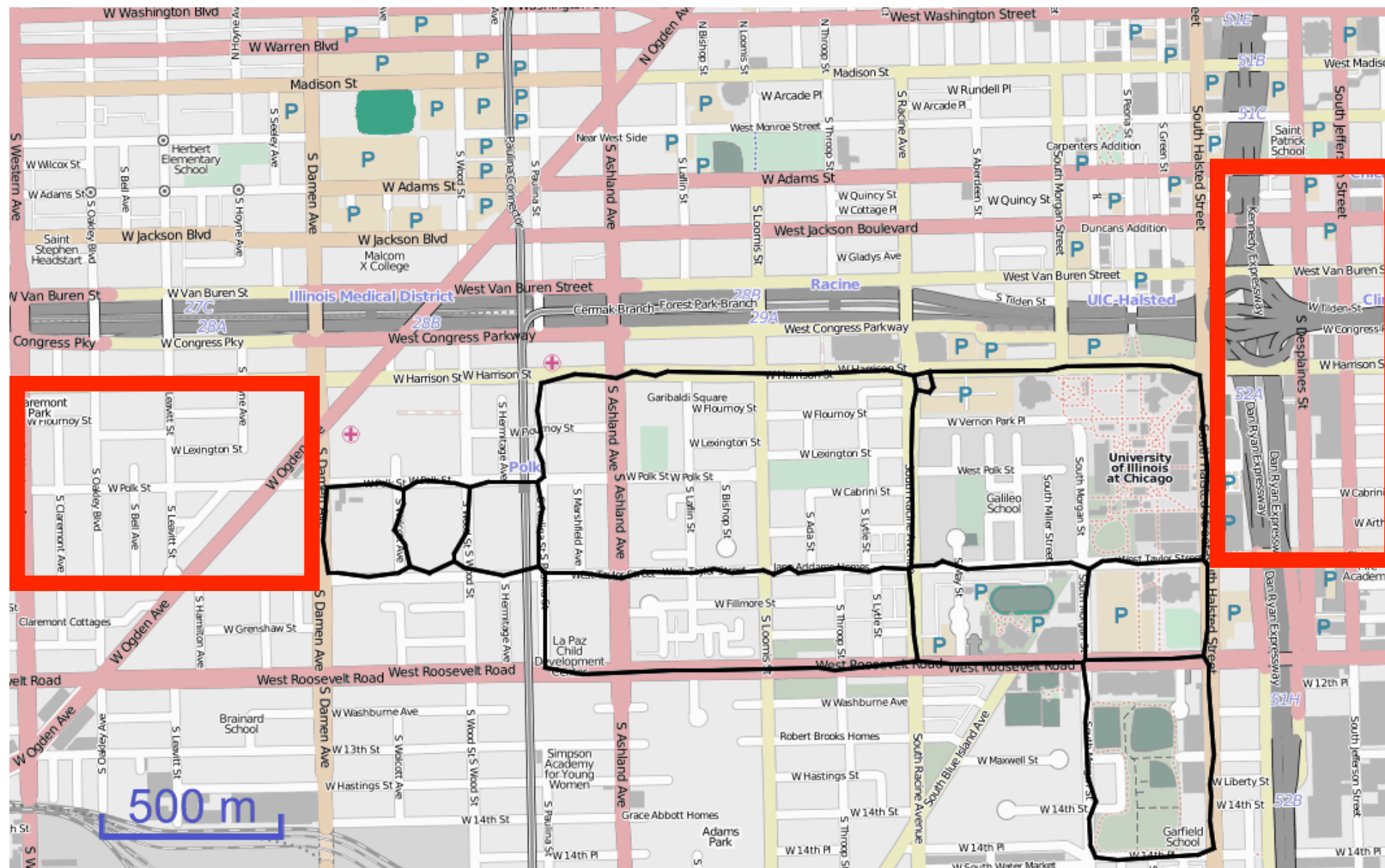
KDE lines



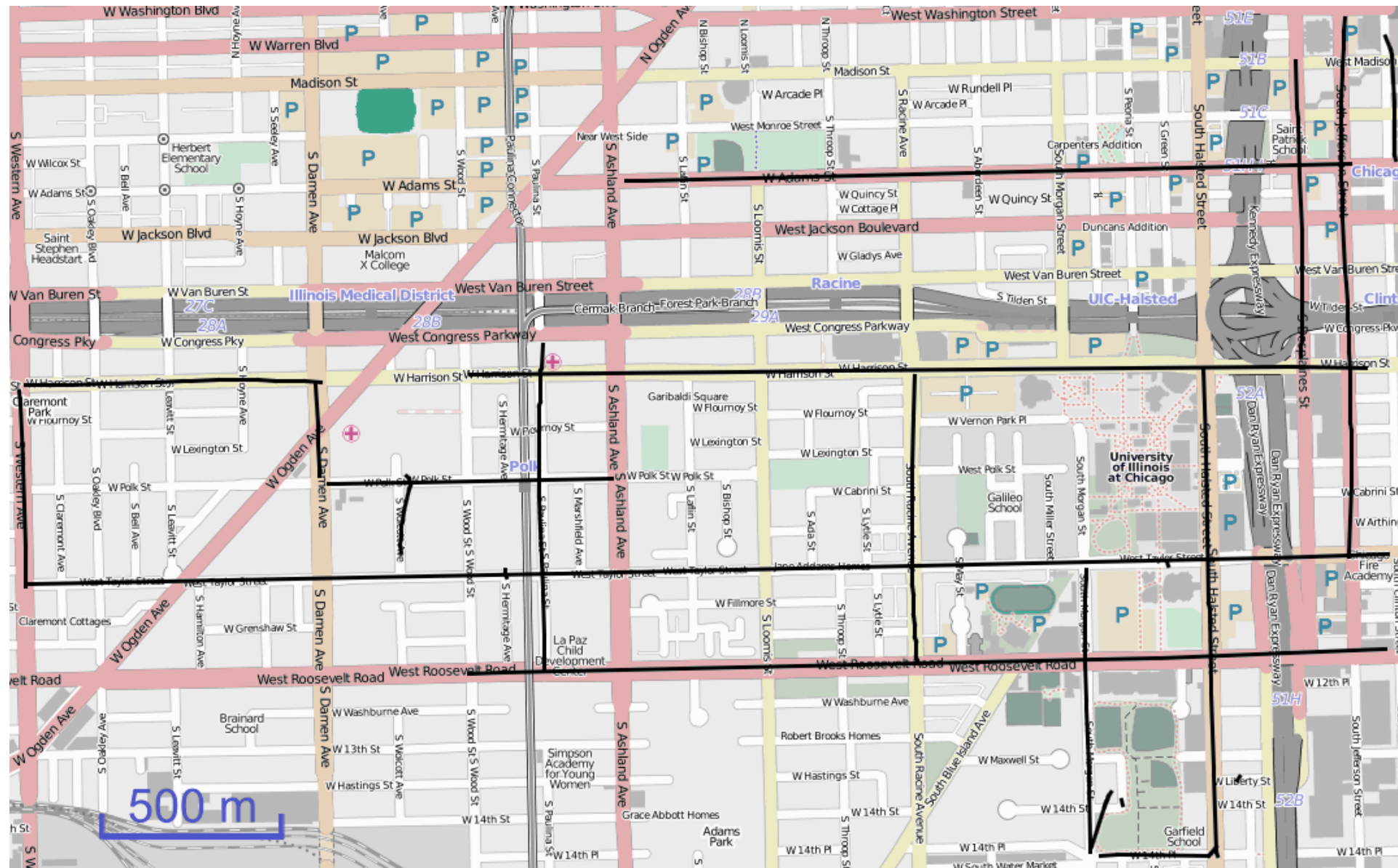
KDE points



KDE points

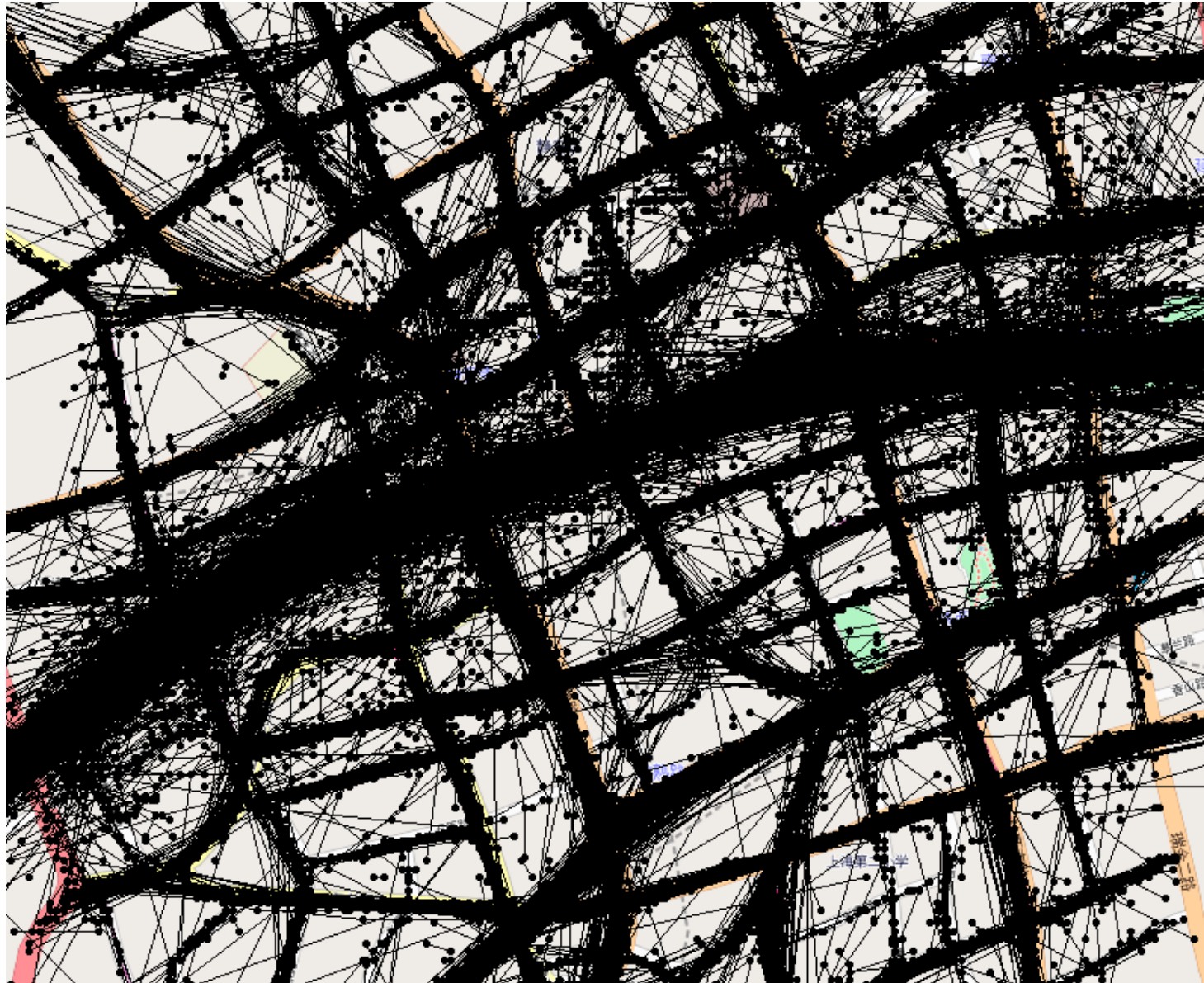


TC1

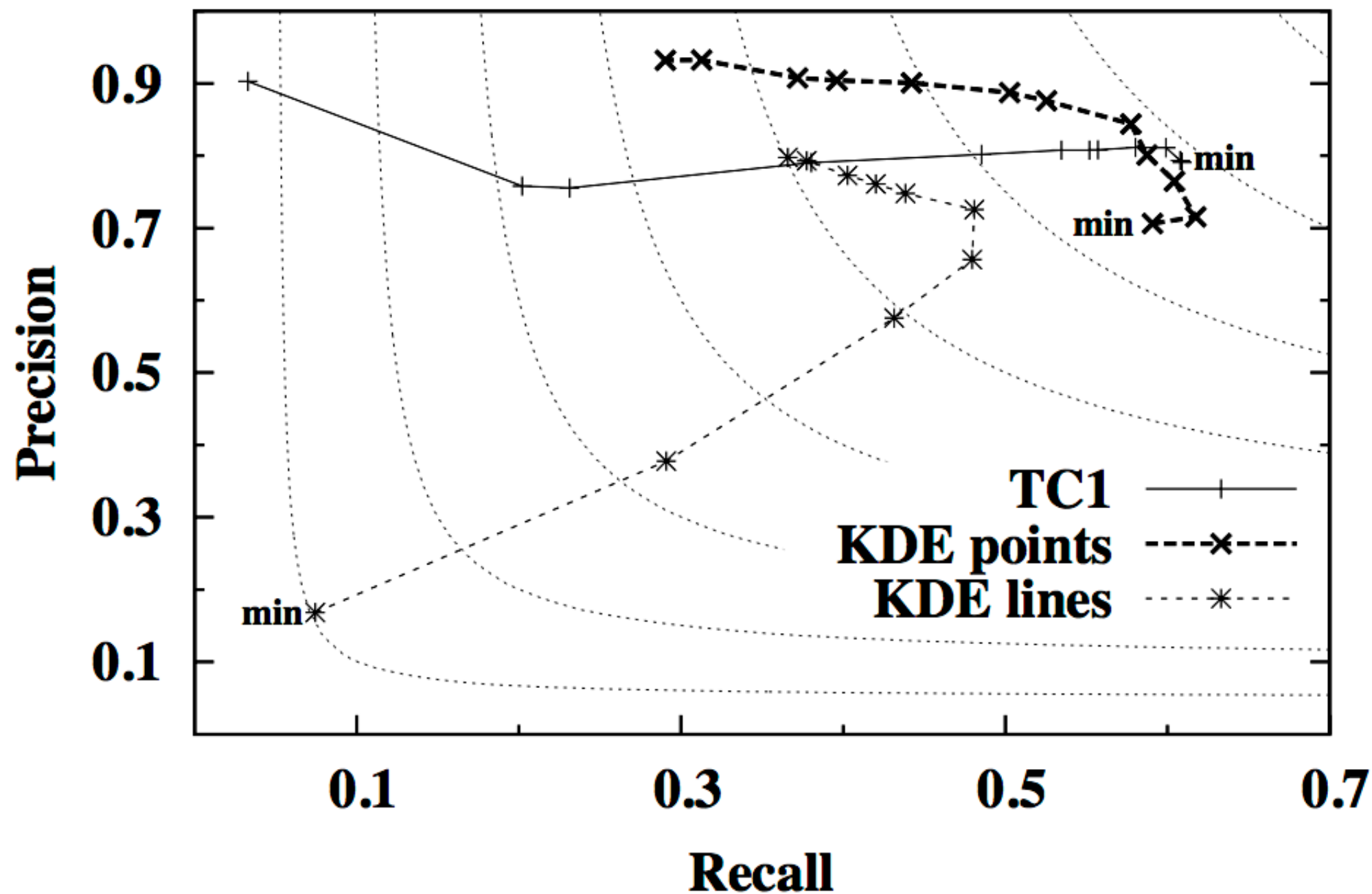


Shanghai Evaluation

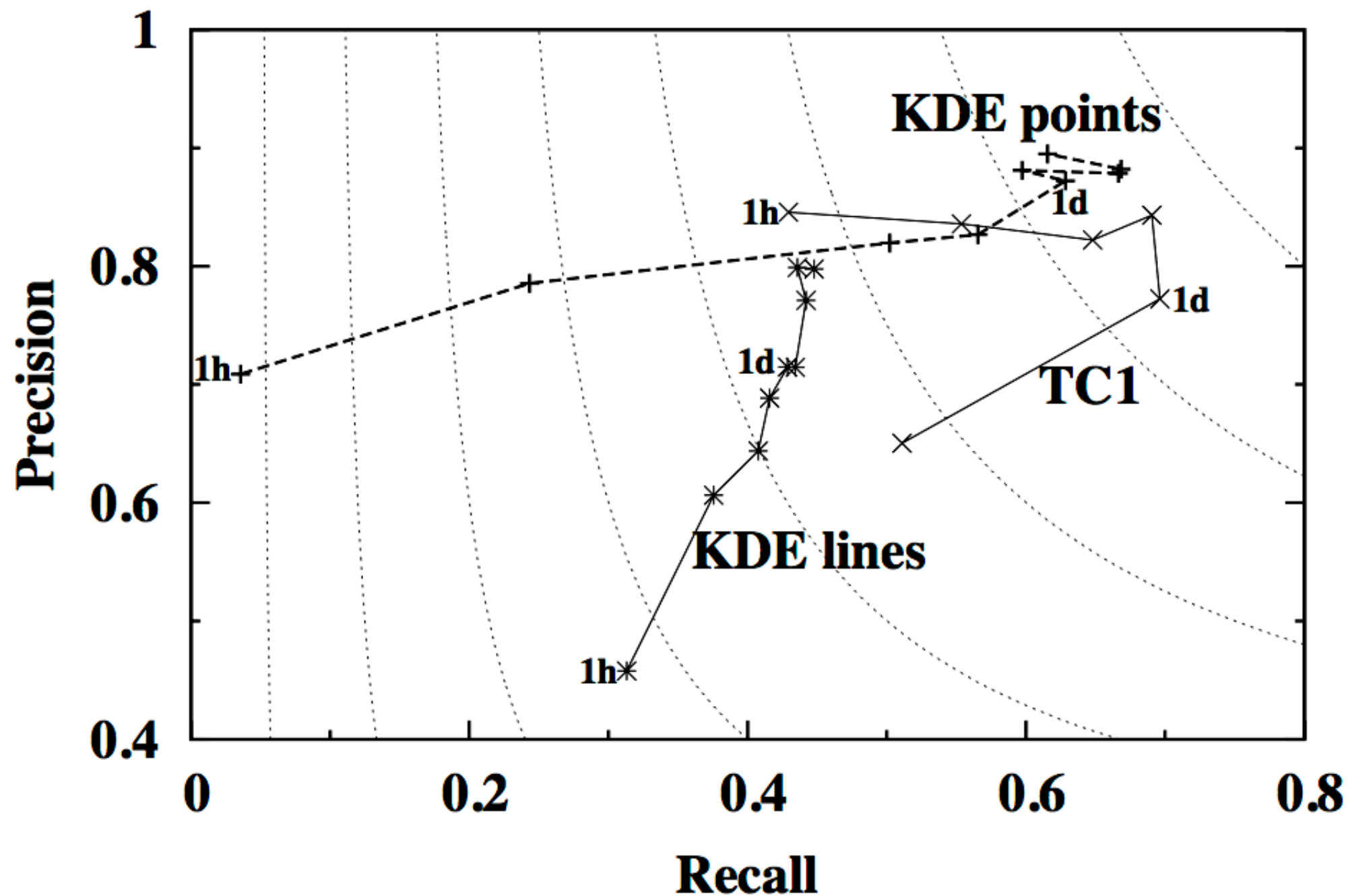
Shanghai raw GPS data



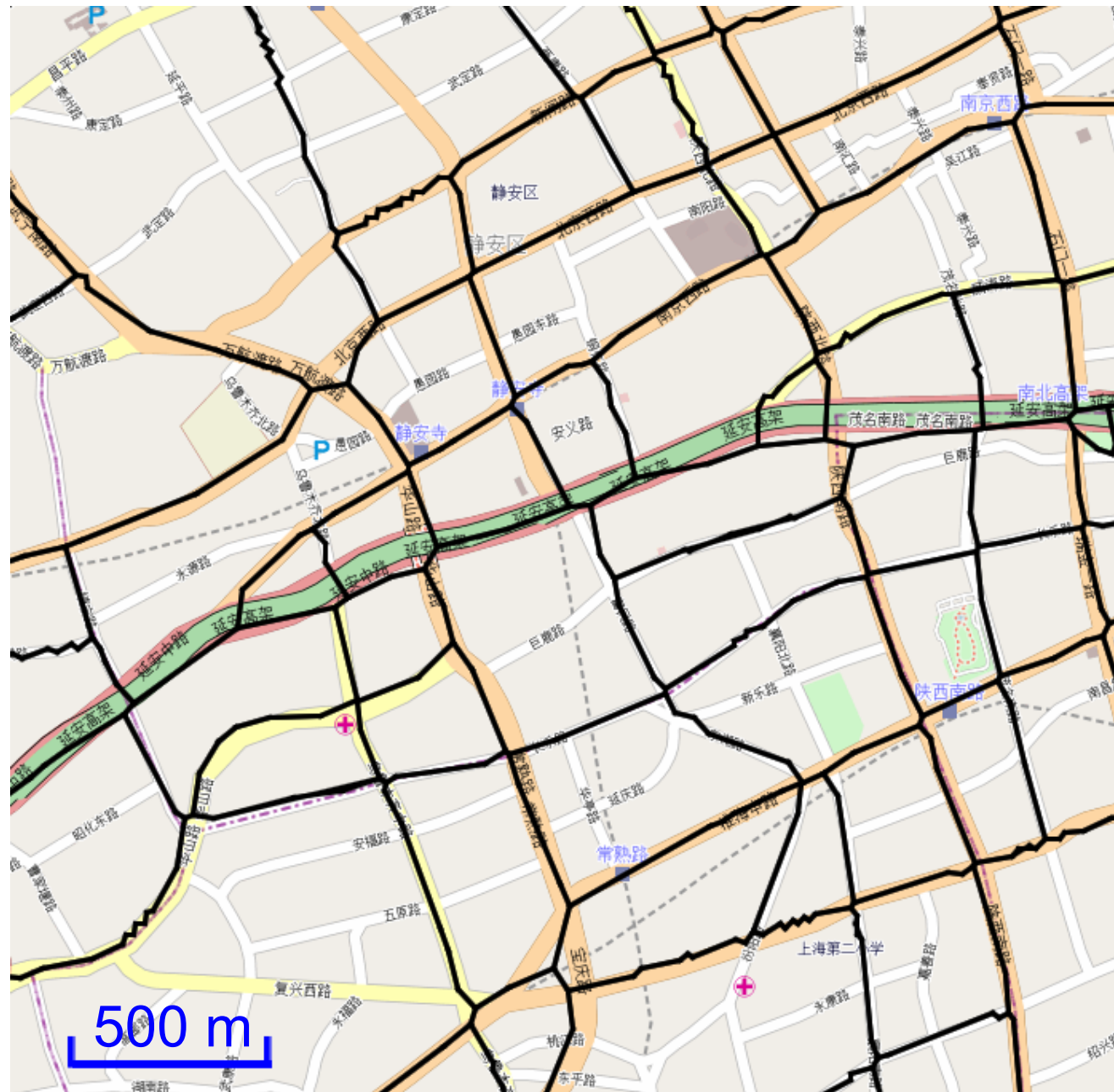
Precision/recall vs. threshold



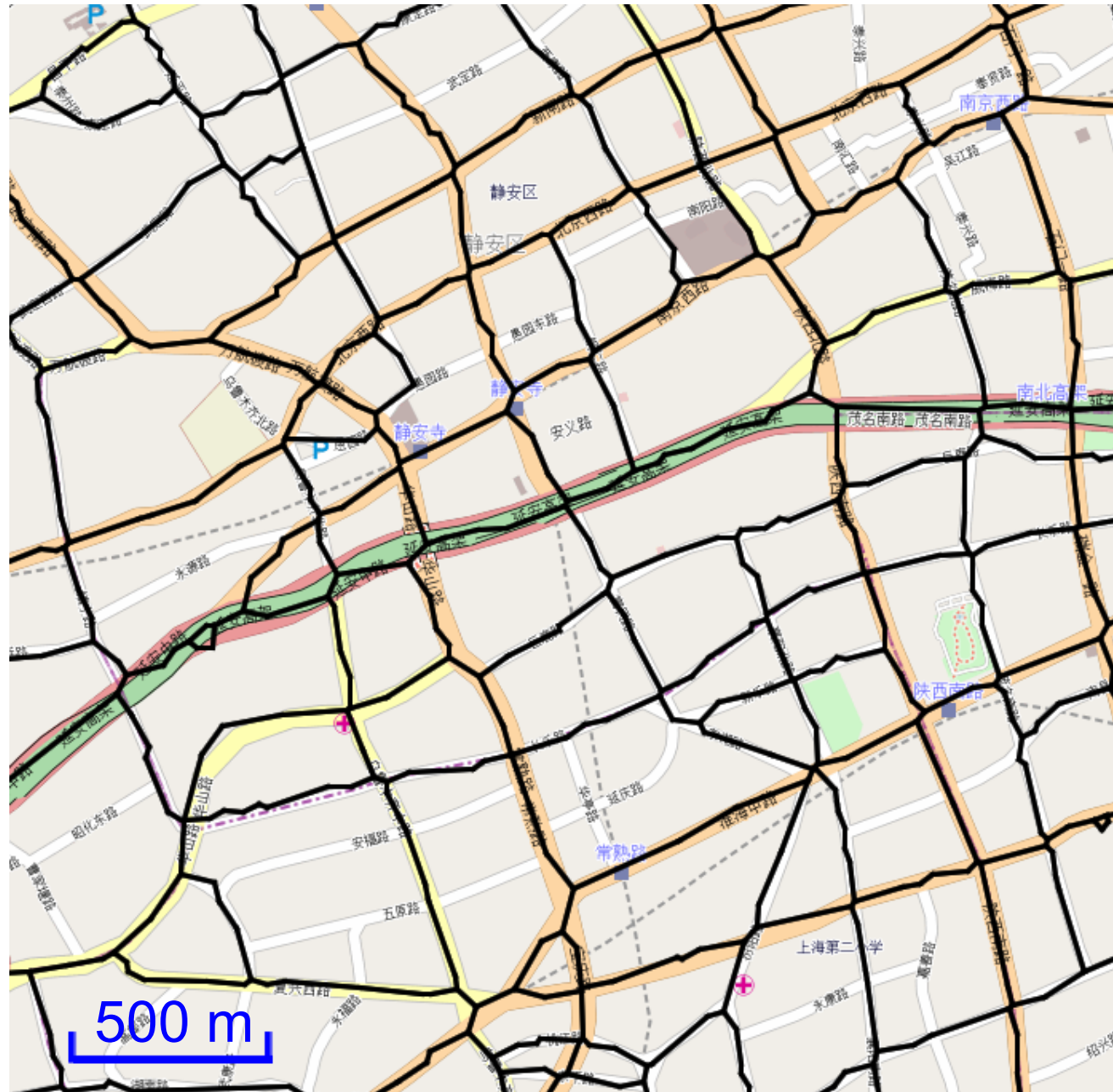
Precision/recall vs. data size



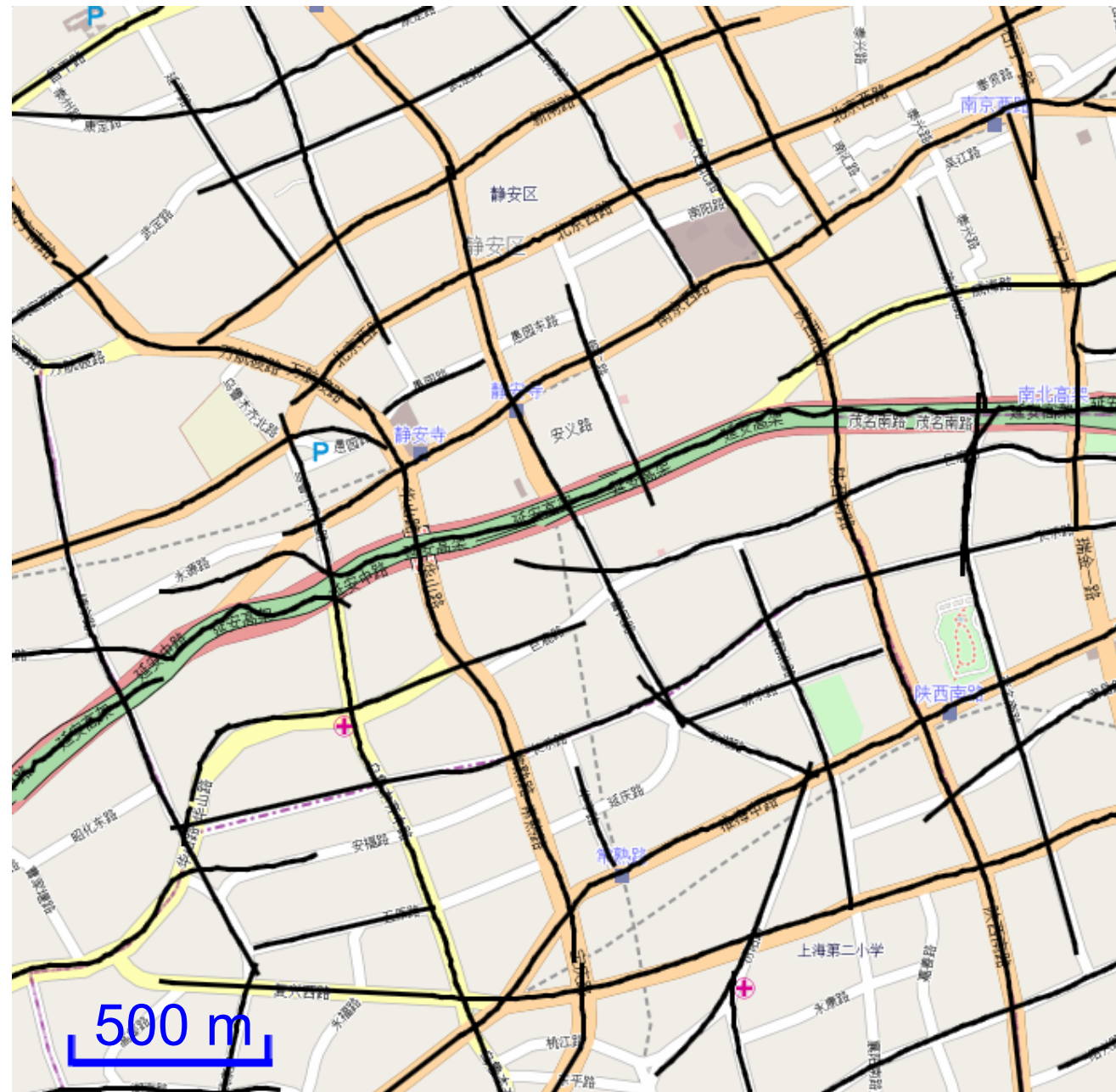
KDE lines



KDE points

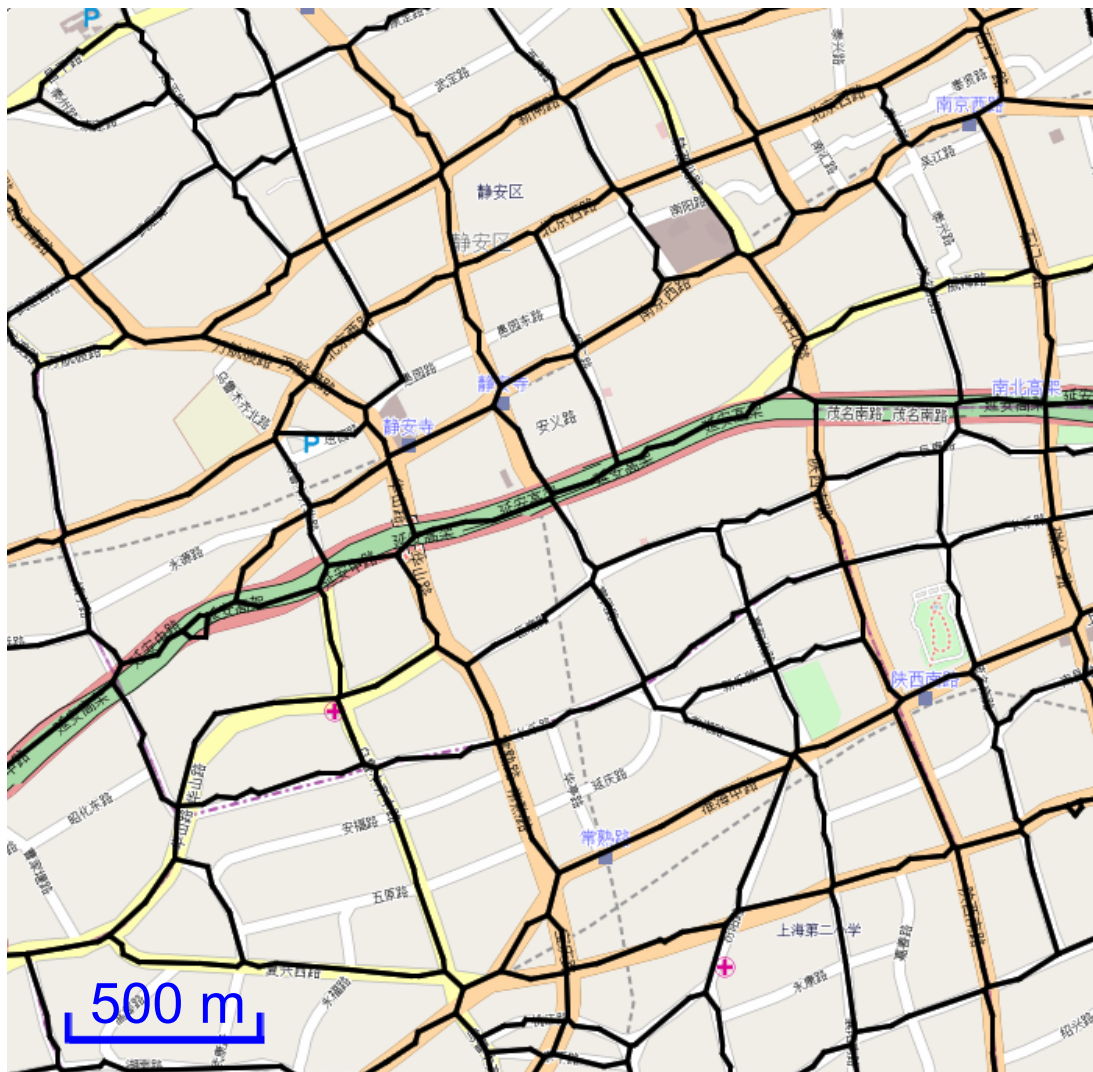


TC1

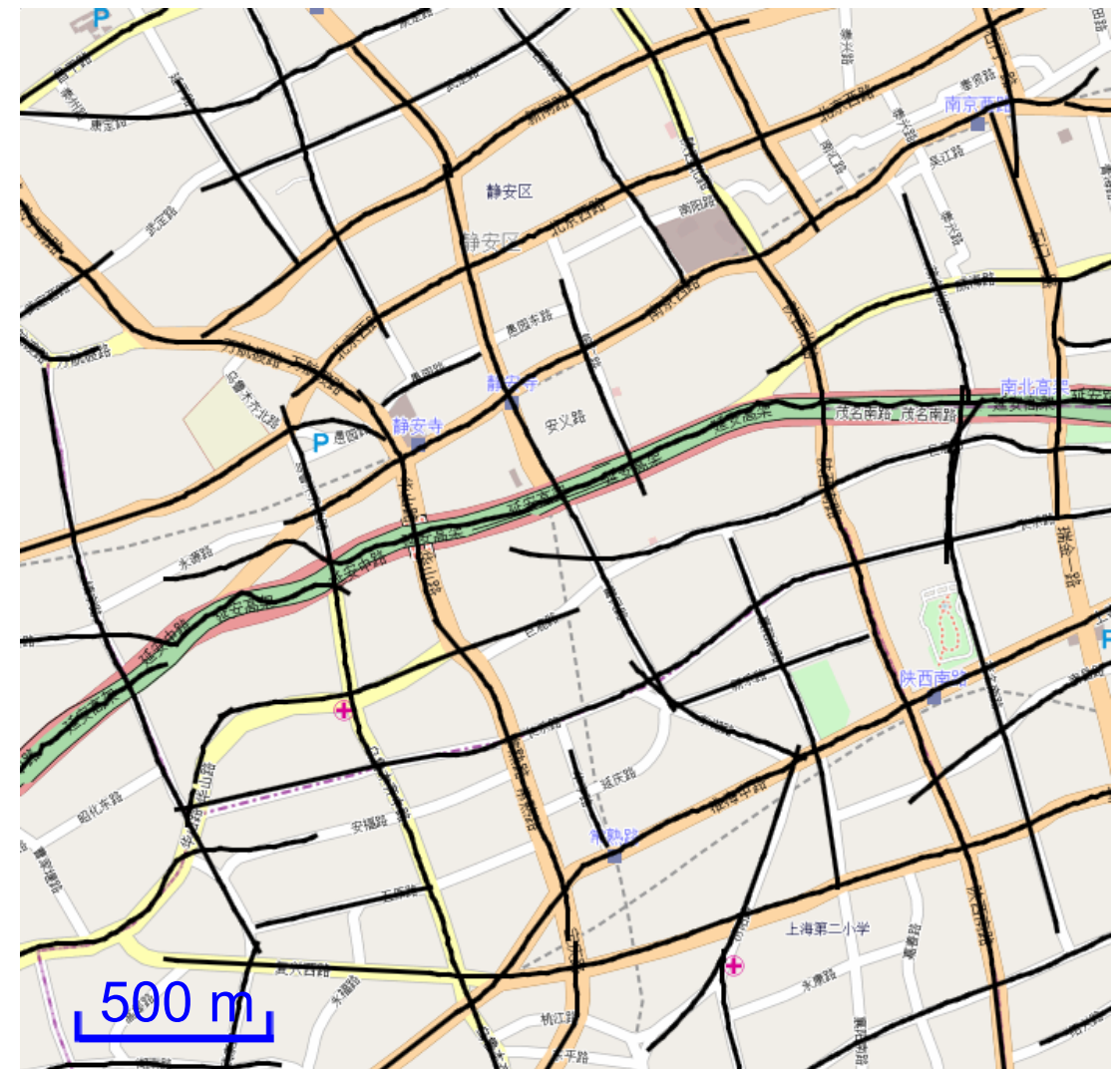


What have we learned?

KDE points



TC1



VS.

Tale of the tape

KDE points	vs.	TC1
X	sparsity	✓
✓	scalability	X
✓	intersections	X
X	centerlines	✓

Future work



Thanks!

Questions?



UIC Bits Networked
UNIVERSITY OF ILLINOIS
AT CHICAGO Systems Laboratory
COLLEGE OF ENGINEERING