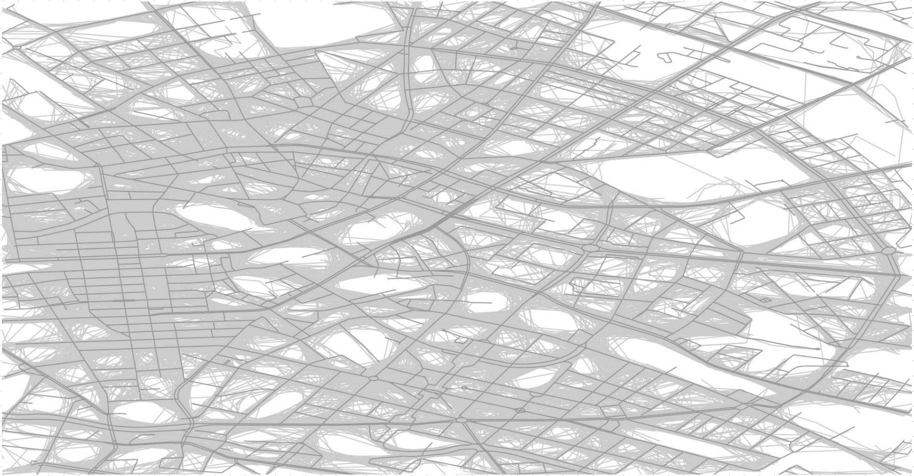
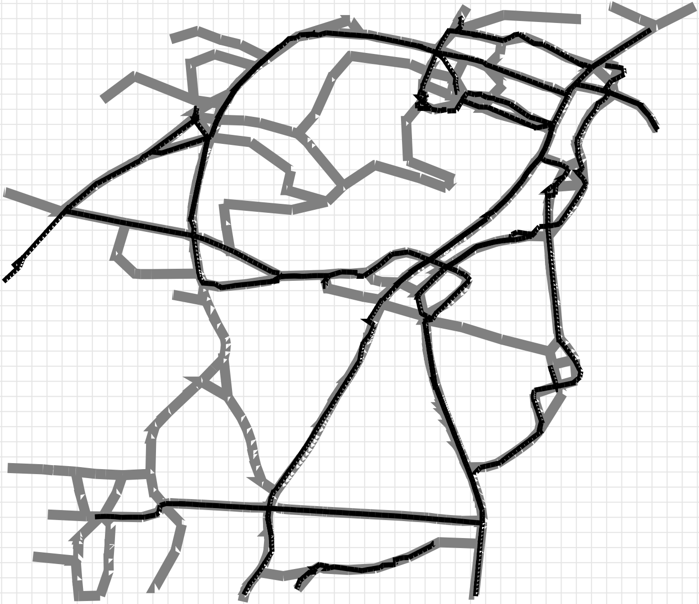
Map Construction and Comparison

Using Local Structure

Brittany Terese Fasy, Tulane University
joint work with M. Ahmed and C. Wenk

6 February 2014
SAMSI Workshop





Examples

Examples

- Road networks:

Examples

- Road networks: GPS trajectories of cars.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths:

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths:

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies:

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems:

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems: Marron's brain vessels.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems: Marron's brain vessels.
- Hurricane paths:

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems: Marron's brain vessels.
- Hurricane paths: historical paths.

Examples

- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems: Marron's brain vessels.
- Hurricane paths: historical paths.
- **Networks:**

Examples

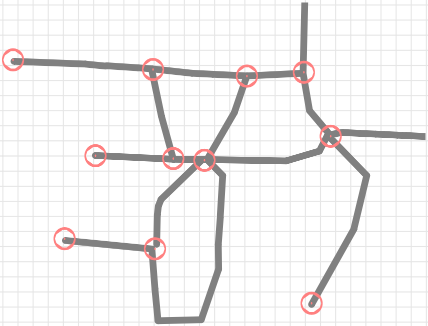
- Road networks: GPS trajectories of cars.
- Hiking paths: GPS paths of people.
- Migration paths: GPS on animals.
- Fillaments of galaxies: point cloud data.
- Biological systems: Marron's brain vessels.
- Hurricane paths: historical paths.
- **Networks: path-constrained trajectories.**

Part I: Map Construction

Road Network Representation

Definition (Road Network)

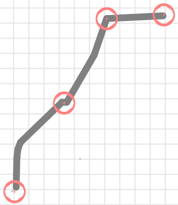
A road network $G = (V, E)$ is an embedded 1-complex. Edges E represent the *roads* or *streets* and the vertices V represent the *intersections*.



Road Network Representation

Definition (Road Network)

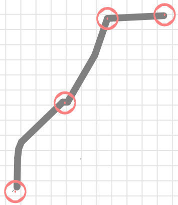
A road network $G = (V, E)$ is an embedded 1-complex. Edges E represent the *roads* or *streets* and the vertices V represent the *intersections*.



The Data

GPS Trajectories

street-path: $p \colon I \rightarrowtail G$
noise: $\varepsilon \colon I \to \mathbb{R}^2$



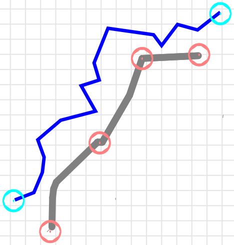
Definition (Link Length)

The *Link Lengh* of a path is the number of edges in the path.

The Data

GPS Trajectories

street-path: $p\colon I \rightarrowtail G$
noise: $\varepsilon\colon I \to \mathbb{R}^2$
trajectory: $\hat{p} = p + \varepsilon$

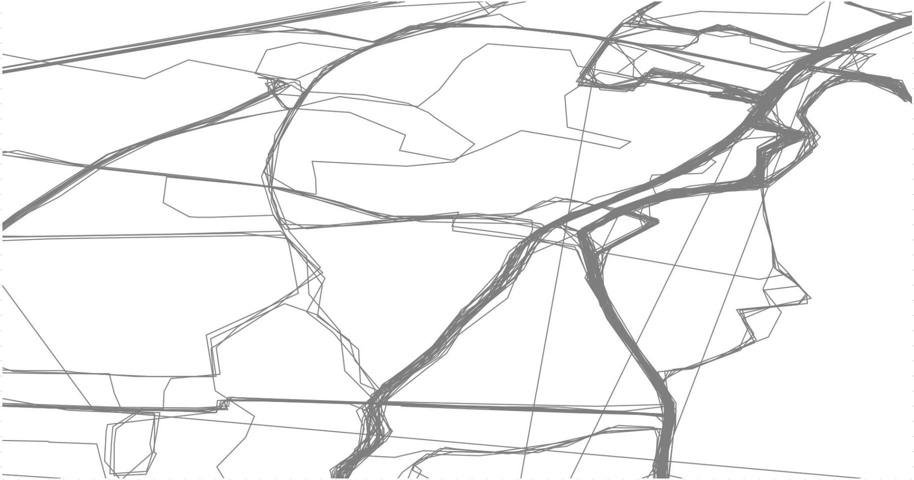


Definition (Link Length)

The *Link Lengh* of a path is the number of edges in the path.

Trajectory Data

Athens Trajectories from School Bus Data



The Reconstruction Problem

Problem Statement

Given a set of (constrained) trajectories, extract the underlying geometric graph structure, in particular:

- Intersections (Vertices of degree > 2)
- Streets (Edges: PL paths between vertices)

Existing Reconstructing Algorithms

http://www.mapconstruction.org

[ACCGGM-12] M. Aanjaneya, F. Chazal, D. Chen, M. Glisse, L. Guibas, and D. Morozov. Metric Graph Reconstruction from Noisy Data, 2012.

[BE-12a] J. Biagioni, J. Eriksson. Map Inference in the Face of Noise and Disparity. SIGSPATIAL, 2012.

[KP-12] S. Karagiorgou and D. Pfoser. On Vehicle Tracking Data-Based Road Network Generation. SIGSPATIAL, 2012.

[AW-12] M. Ahmed, C. Wenk. Constructing Street Networks from GPS Trajectories. European Symp. on Algorithms, 2012.

[OTHER] There were too many algorithms to cite here . . .

Incremental Fréchet Matching [AW]

Assumptions

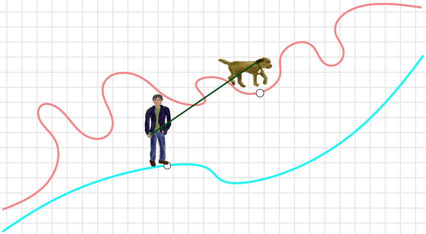
1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.

Assumption A1

Trajectories are Close to Paths

Definition (Fréchet Distance)

$d_F(p, \hat{p}) =$
$\inf_\alpha \max_t d\left(p(t), \hat{p}(\alpha(t))\right).$



Assumption A1

$d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.

Incremental Fréchet Matching [AW]

Assumptions

1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.

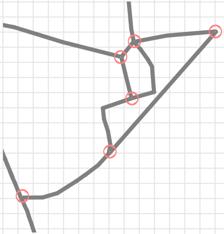
Incremental Fréchet Matching [AW]

Assumptions

1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.
2. A2: If two paths in G come close enough, they intersect.

Assumption A2

Close Paths Intersect

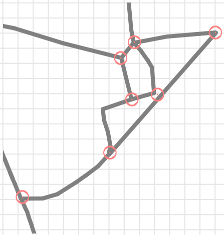


Assumption A2

If two paths in G come close enough, they intersect.

Assumption A2

Close Paths Intersect



Assumption A2

If two paths in G come close enough, they intersect.

Incremental Fréchet Matching [AW]

Assumptions

1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.
2. A2: If two paths in G come close enough, they intersect.

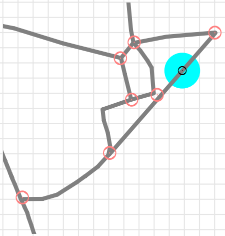
Incremental Fréchet Matching [AW]

Assumptions

1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.
2. A2: If two paths in G come close enough, they intersect.
3. A3: For each street s and a ball B far enough away from the endpoints of s, $s \cap B$ is two points.

Assumption A3

Roads Are Nice



Assumption A3

For each street s and a ball B far enough away from the endpoints of s, $s \cap B$ is two points.

Incremental Fréchet Matching [AW]

Assumptions

1. A1: $d_F(p, \hat{p}) \leq \frac{\varepsilon}{2}$ for all trajectories $\hat{p}$.
2. A2: If two paths in G come close enough, they intersect.
3. A3: For each street s and a ball B far enough away from the endpoints of s, $s \cap B$ is two points.

Incremental Fréchet Algorithm [AW]

For each trajectory t:

1. Find the portion of the map closest to t.
2. Add edges if necessary.
3. Simplify:

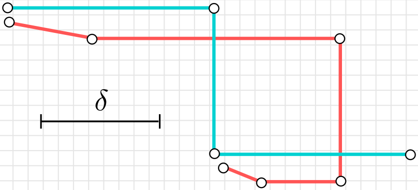
Incremental Fréchet Algorithm [AW]

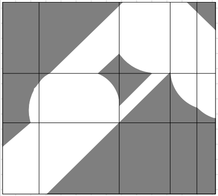
For each trajectory t:

1. Find the portion of the map closest to t.
2. Add edges if necessary.
3. Simplify:

Computing the Fréchet Distance

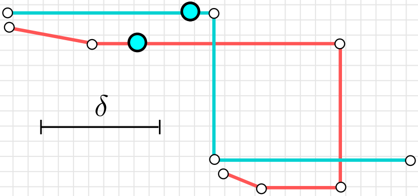
The Free Space Diagram

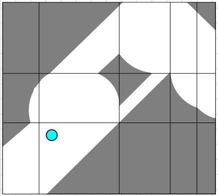




Computing the Fréchet Distance

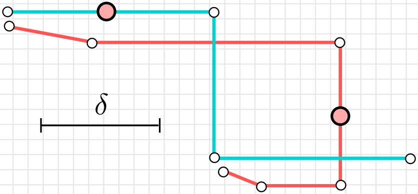
The Free Space Diagram

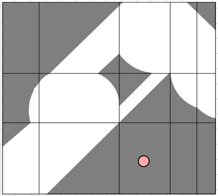




Computing the Fréchet Distance

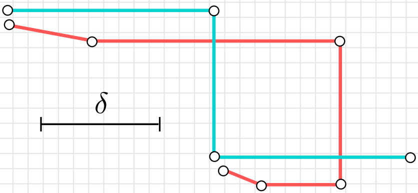
The Free Space Diagram

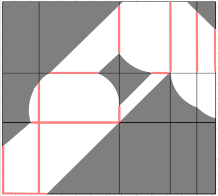




Computing the Fréchet Distance

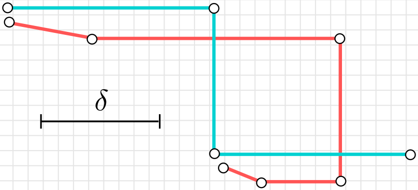
The Free Space Diagram

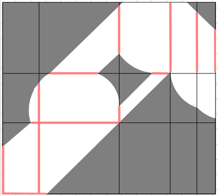




Computing the Fréchet Distance

The Free Space Diagram

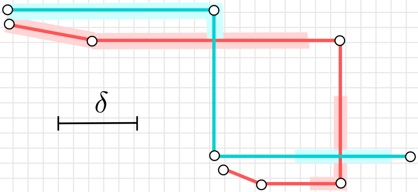




- $O(mn \log(mn))$ using parametric search [AG-95].
- $O(mn \log^2(mn))$ randomized approach [CW-10].

Finding Closest Match ...

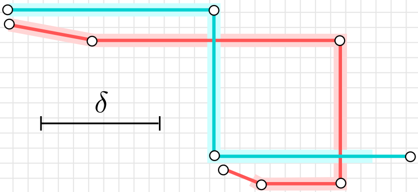
Partial Fréchet Distance

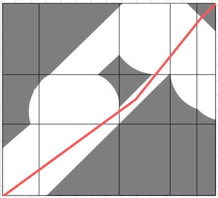


Let grey region have weight 1 and white region weight 0. Find the shortest path from one end to the other under this metric.

Finding Closest Match ...

Partial Fréchet Distance

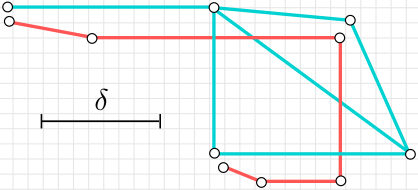


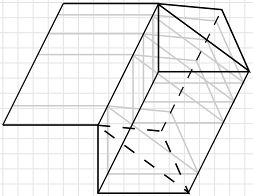


Let grey region have weight 1 and white region weight 0. Find the shortest path from one end to the other under this metric.

Best Match in Graph

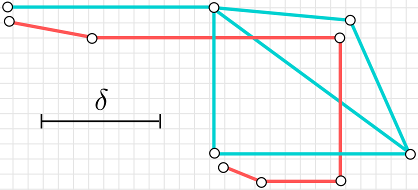
[BBW-09]

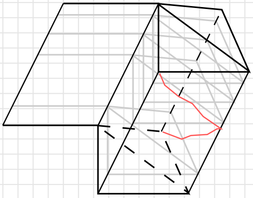




Best Match in Graph

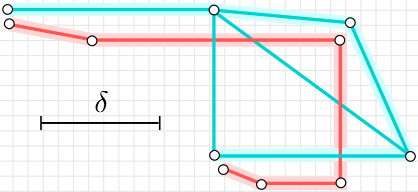
[BBW-09]

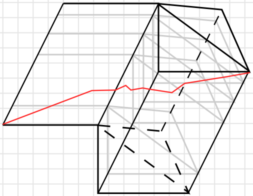




Best Match in Graph

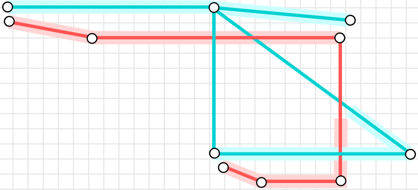
[BBW-09]





Best Match in Graph

[BBW-09]

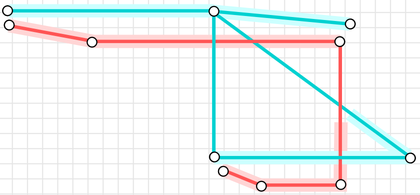


Incremental Fréchet Algorithm [AW]

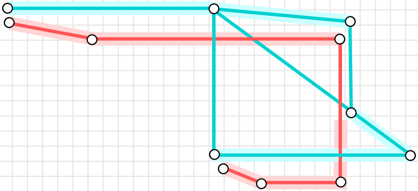
For each trajectory t:

1. Find the portion of the map closest to t.
2. Add edges if necessary.
3. Simplify:

Adding Edges



Adding Edges



Incremental Fréchet Algorithm [AW]

For each trajectory t:

1. Find the portion of the map closest to t.
2. Add edges if necessary.
3. Simplify:

Incremental Fréchet Algorithm [AW]

For each trajectory t:

1. Find the portion of the map closest to t.
2. Add edges if necessary.
3. Simplify: Min-link curve simplification algorithm.

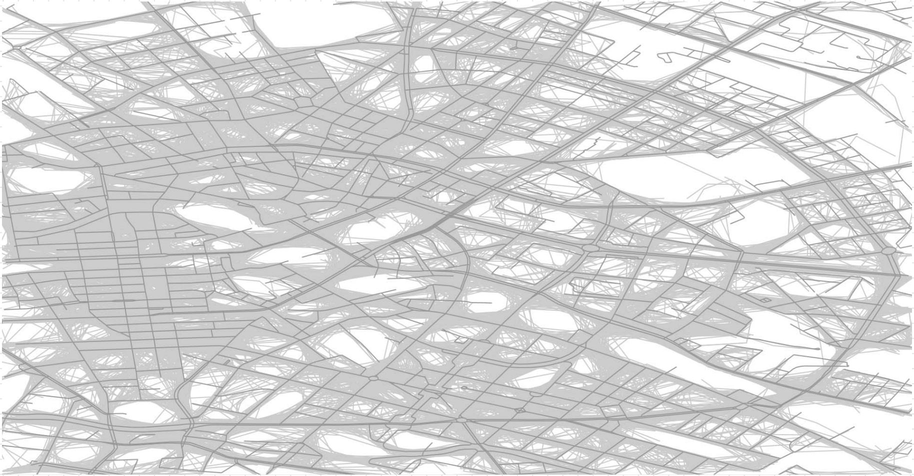
Dataset 1: Berlin

www.openstreetmaps.org

- 27,188 GPS trajectories from school buses.
- D = 5.2km x 6.1km.
- Trajectories have 7 observations on average.

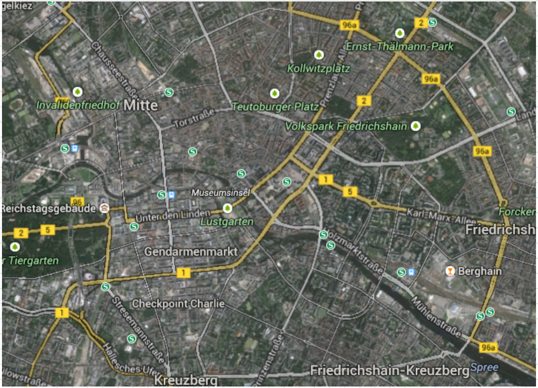
Dataset 1: Berlin

www.openstreetmaps.org



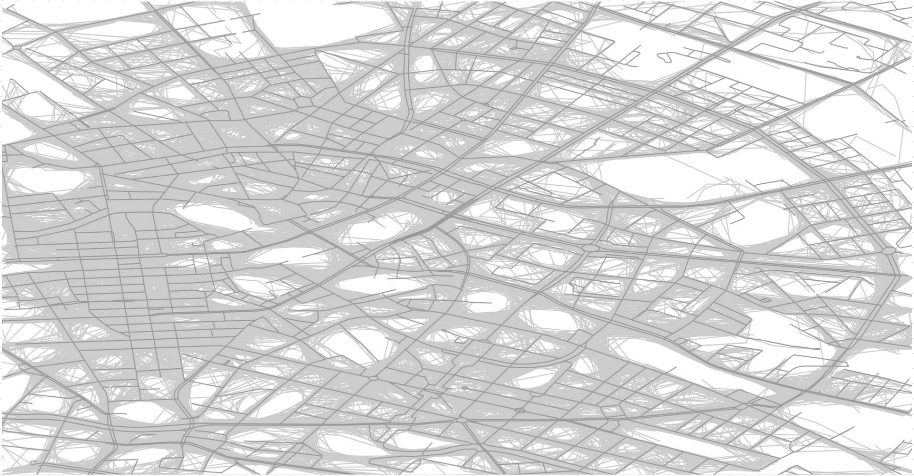
Dataset 1: Berlin

www.openstreetmaps.org



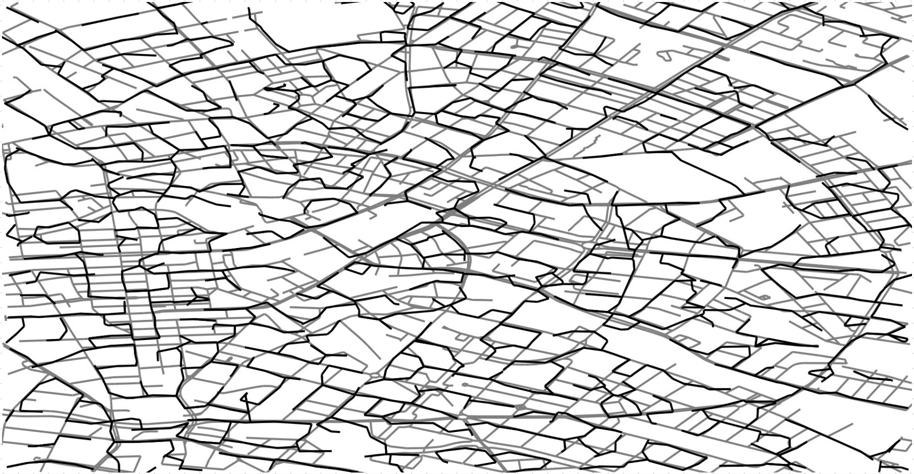
Dataset 1: Berlin

www.openstreetmaps.org



Dataset 1: Berlin

www.openstreetmaps.org



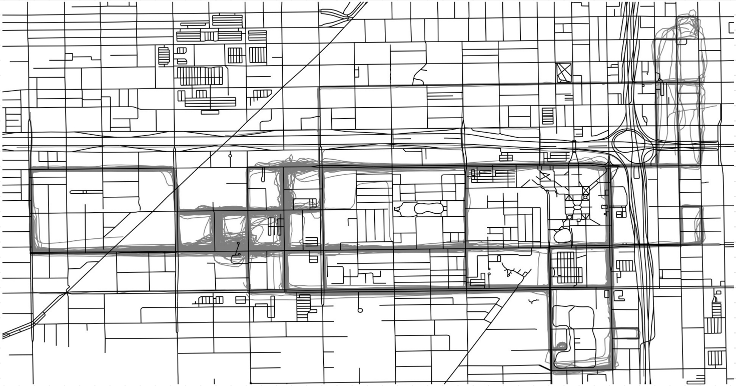
Dataset 2: Chicago

www.openstreetmaps.org

- 889 GPS trajectories from university buses.
- D = 7 km x 4.5 km
- Trajectories: 100-300 samples (avg. 3.22 km).

Dataset 2: Chicago

www.openstreetmaps.org

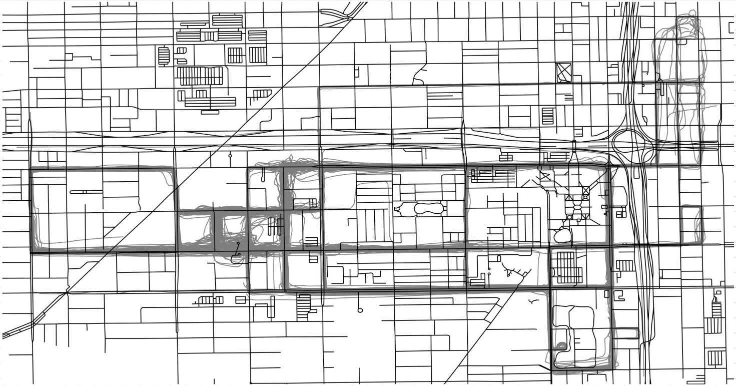


Dataset 2: Chicago

www.openstreetmaps.org

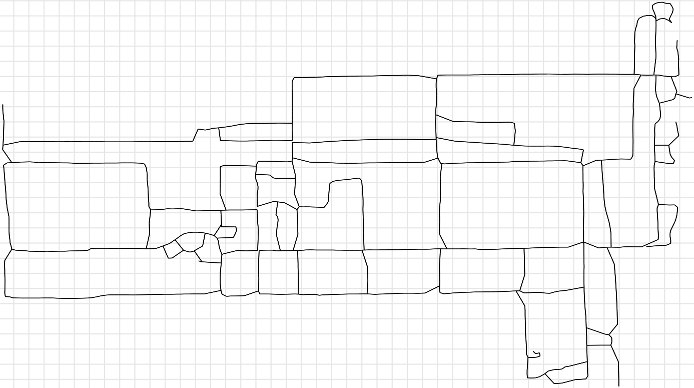
Dataset 2: Chicago

www.openstreetmaps.org



Dataset 2: Chicago

www.openstreetmaps.org



Dataset 2: Chicago

www.openstreetmaps.org

Part II: Map Comparison

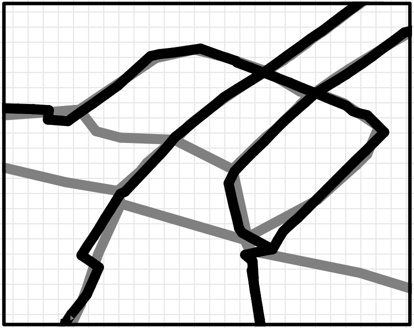
Different Approaches

[BE-12b] J. Biagioni, J. Eriksson. James Biagioni and Jakob Eriksson. Inferring Road Maps from GPS Traces: Survey and Comparative Evaluation. In 91st Annual Meeting of the Transportation Research Board, 2012.

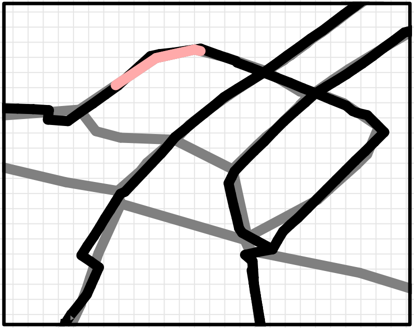
[AFHW-14] M. Ahmed, B. Fasy, K. Hickmann, C. Wenk. Path-Based Distance for Street Map Comparison. Arxiv.

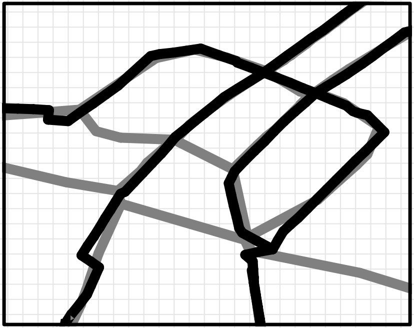
[AFW-14] M. Ahmed, B. Fasy, C. Wenk. Local Homology Based Distance Between Maps. In Submission.

Path-Based Methods

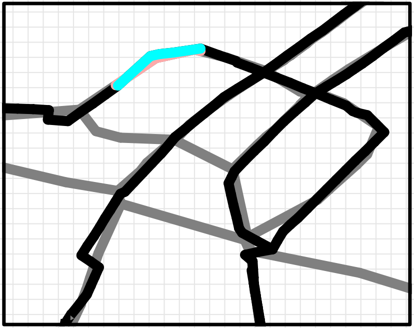


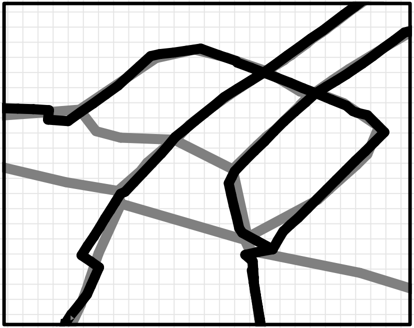
Path-Based Methods



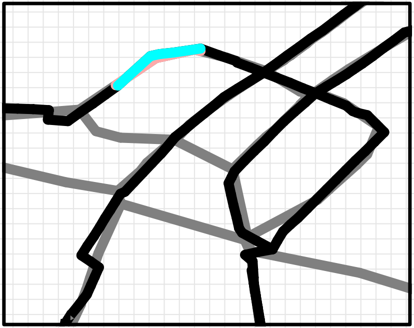


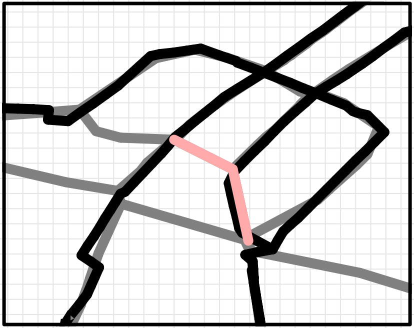
Path-Based Methods



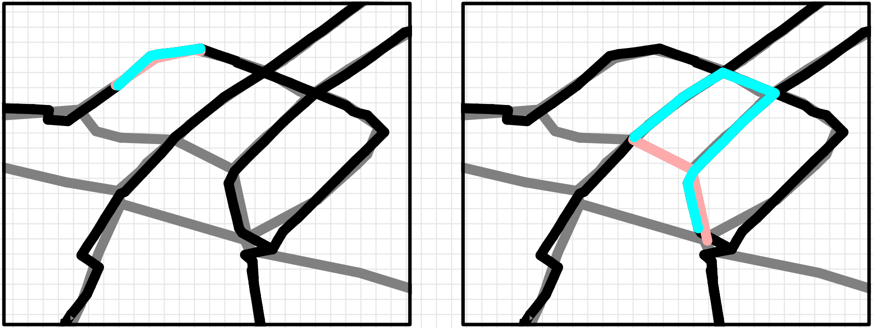


Path-Based Methods

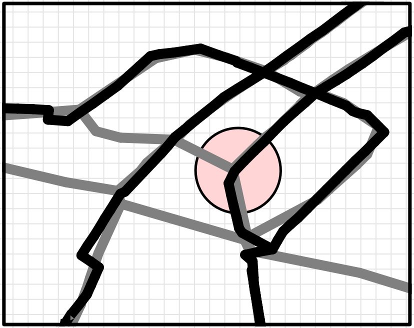




Path-Based Methods



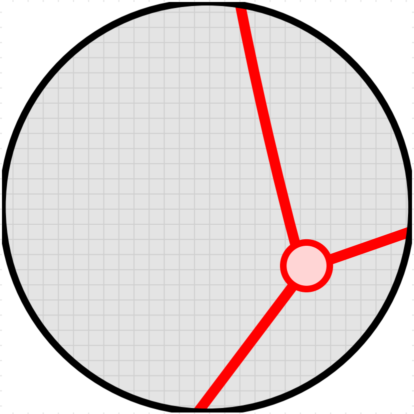
Local Topology Approach [AFW]



We will use the embedding and the local homology to create a *local distance signature*.

Local Distance Signature

Finding the Local Persistence Diagram

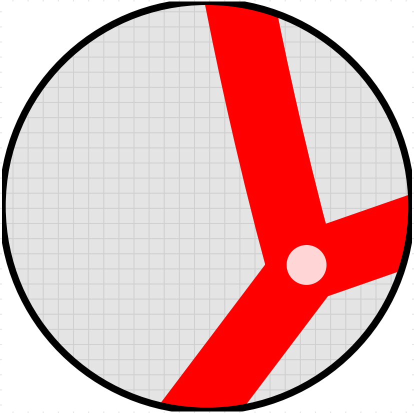


$D \subset \mathbb{R}^2$ is the compact domain.
$G_1 \subset D$ is the road network.
$r > 0$ is the scale.
$x \in D$.

$$LG_1(x, 0) = (G_1 \cap B_r(x))/\partial B_r(x)$$

Local Distance Signature

Finding the Local Persistence Diagram

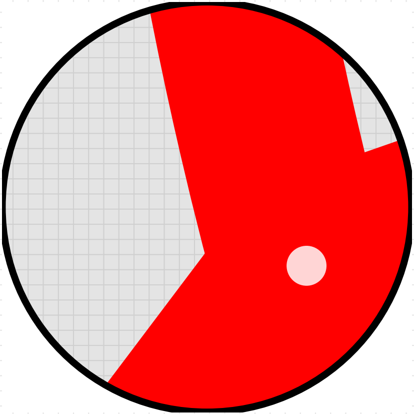


$D \subset \mathbb{R}^2$ is the compact domain.
$G_1 \subset D$ is the road network.
$r > 0$ is the scale.
$x \in D$.

$$LG_1(x, \varepsilon) = ({G_1}^{\varepsilon} \cap B_r(x))/\partial B_r(x)$$

Local Distance Signature

Finding the Local Persistence Diagram

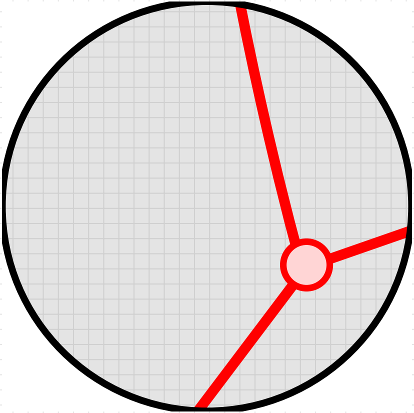


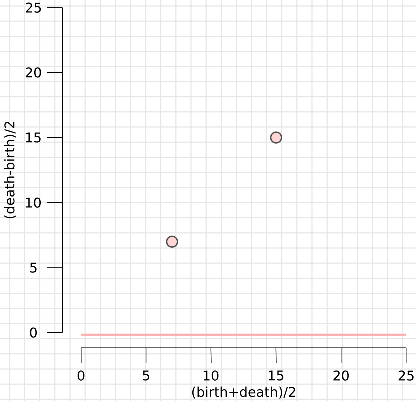
$D \subset \mathbb{R}^2$ is the compact domain.
$G_1 \subset D$ is the road network.
$r > 0$ is the scale.
$x \in D$.

$$LG_1(x, 2\varepsilon) = ({G_1}^{2\varepsilon} \cap B_r(x))/\partial B_r(x)$$

Local Distance Signature

Finding the Local Persistence Diagram

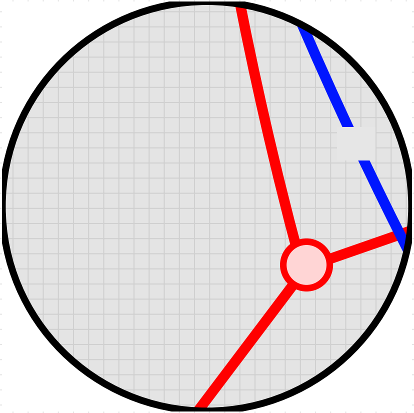


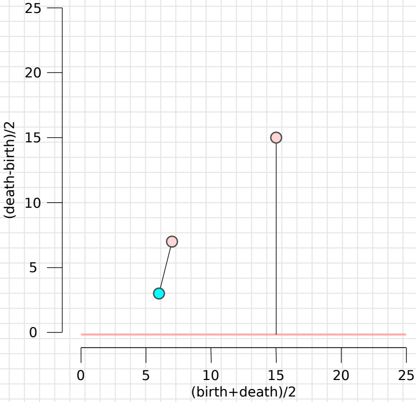


$$LG_1(x, 0) = (G_1 \cap B_r(x))/\partial B_r(x)$$

Local Distance Signature

Computing the Local Distance

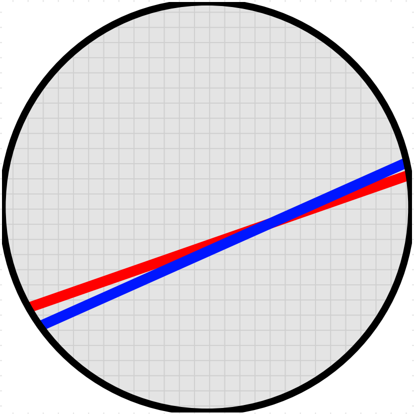


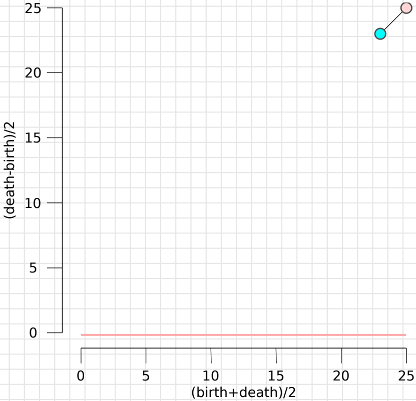


$$lhd(x,r) := W_\infty(\mathrm{Dgm}_1(LG_1(x,r)), \mathrm{Dgm}_1(LG_2(x,r)))$$

Local Distance Signature

Computing the Local Distance





$$lhd(x,r) := W_\infty(\mathrm{Dgm}_1(LG_1(x,r)), \mathrm{Dgm}_1(LG_2(x,r)))$$

Local Topology Based Distance

Definition (Local Homology Distance)

$$d^{LH}(G_1, G_2) = \frac{1}{|D|} \int_D lhd(x, r)\, dx$$

Local Topology Based Distance

Definition (Local Homology Distance)

$$d^{LH}(G_1, G_2) = \frac{1}{|D|} \int_{r_0}^{r_1} \int_D lhd(x, r)\, dx\, dr$$

Local Topology Based Distance

Definition (Local Homology Distance)

$$d^{LH}(G_1, G_2) = \frac{1}{|D|} \int_{r_0}^{r_1} \omega_r \int_D lhd(x, r)\, dx\, dr$$

Local Topology Based Distance

Definition (Local Homology Distance)

$$d^{LH}(G_1, G_2) = \frac{1}{|D|} \int_{r_0}^{r_1} \omega_r \int_D lhd(x, r)\, dx\, dr$$

Variants

1. D is a rectangular domain.
2. $D = (G_1 \cup G_2) + B_\delta$
3. $r_0 = r_1$ (fixed radius)
4. $r_0 = 0 \implies$ metric

Local Topology Based Distance

Definition (Local Homology Distance)

$$d^{LH}(G_1, G_2) = \frac{1}{|D|} \int_{r_0}^{r_1} \omega_r \int_D lhd(x, r)\, dx\, dr$$

Variants

1. D is a rectangular domain.
2. $D = (G_1 \cup G_2) + B_\delta$
3. $r_0 = r_1$ (fixed radius)
4. $r_0 = 0 \implies$ metric

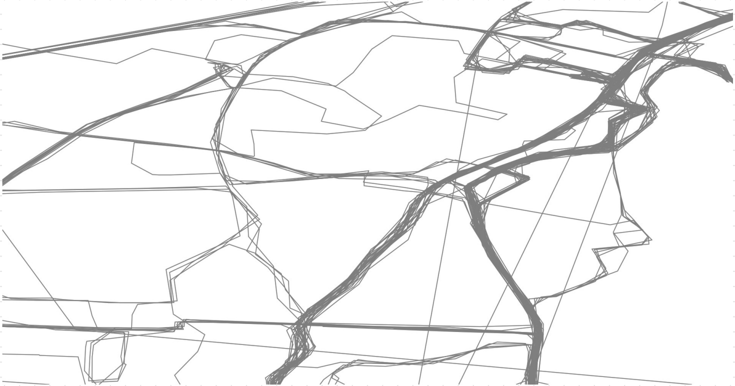
Dataset 3: Athens

www.openstreetmaps.org

- 129 GPS trajectories from school buses.
- D = 2.6 km x 6 km
- Two reconstruction algorithms: [BE-12a] and [KP-12].
- Trajectories: 13-47 samples

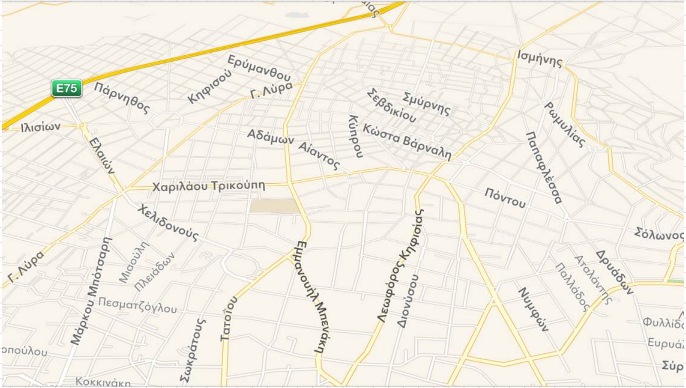
Dataset 3: Athens

www.openstreetmaps.org



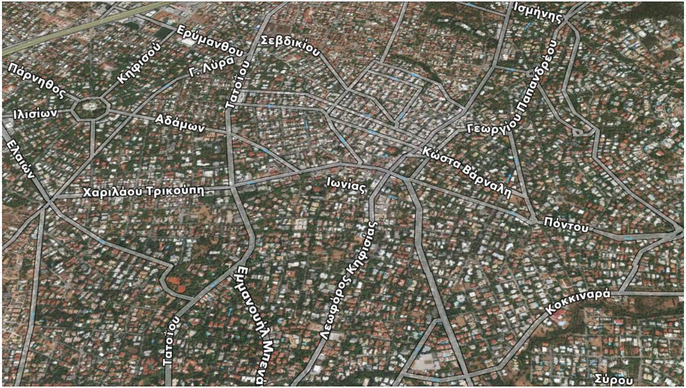
Dataset 3: Athens

www.openstreetmaps.org



Dataset 3: Athens

www.openstreetmaps.org



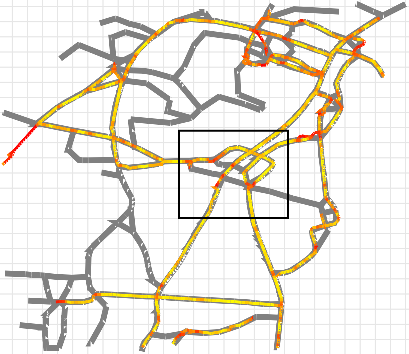
Thanks Fabrizio and Jisu!

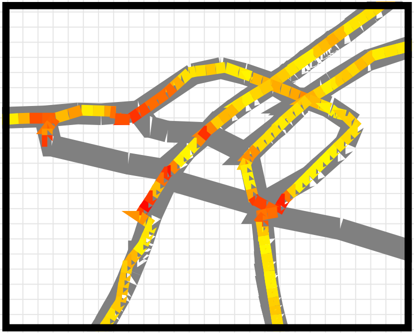
Results

Athens: Comparing Two Different Reconstructions

Results

Athens: Comparing Two Different Reconstructions





The Bootstrap

... When the Ground Truth is Unknown

The Bootstrap

- n input trajectories S_n.
- Compute $\widehat{G} = \widehat{G}(S_n)$.

The Bootstrap

... When the Ground Truth is Unknown

The Bootstrap

- n input trajectories S_n.
- Compute $\widehat{G} = \widehat{G}(S_n)$.
- Find subsample S_n^*.
- Compute $\widehat{G}^* = \widehat{G}(S_n^*)$.

The Bootstrap

... When the Ground Truth is Unknown

The Bootstrap

- n input trajectories S_n.
- Compute $\widehat{G} = \widehat{G}(S_n)$.
- Find subsample S_n^*.
- Compute $\widehat{G}^* = \widehat{G}(S_n^*)$.
- Repeat B times, obtaining $\widehat{G}_1^*, \ldots \widehat{G}_B^*$.
- Let $T_j = d^{LH}(\widehat{G}, \widehat{G}_j^*)$.

The Bootstrap

... When the Ground Truth is Unknown

The Bootstrap

- n input trajectories S_n.
- Compute $\widehat{G} = \widehat{G}(S_n)$.
- Find subsample S_n^*.
- Compute $\widehat{G}^* = \widehat{G}(S_n^*)$.
- Repeat B times, obtaining $\widehat{G}_1^*, \dots \widehat{G}_B^*$.
- Let $T_j = d^{LH}(\widehat{G}, \widehat{G}_j^*)$.

$$Z_\alpha = \inf\{z \ : \ \frac{1}{B}\sum_1^B I(T_j > z) \le \alpha\}$$

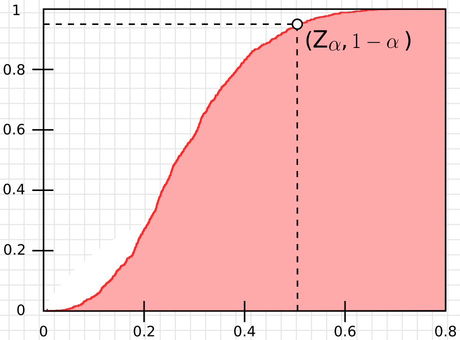
The Bootstrap

... When the Ground Truth is Unknown

The Bootstrap

- n input trajectories S_n.
- Compute $\widehat{G} = \widehat{G}(S_n)$.
- Find subsample S_n^*.
- Compute $\widehat{G}^* = \widehat{G}(S_n^*)$.
- Repeat B times, obtaining $\widehat{G}_1^*, \ldots \widehat{G}_B^*$.
- Let $T_j = d^{LH}(\widehat{G}, \widehat{G}_j^*)$.

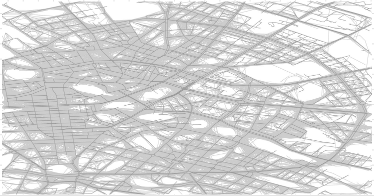
$$Z_\alpha = \inf\{z \ : \ \frac{1}{B}\sum_1^B I(T_j > z) \leq \alpha\}$$

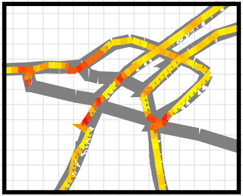


Ongoing

- Apply this to different construction algorithms versus the ground truth to rank algorithms.
- Implementation of distance measure.
- Improve theoretical guarantees.
- Input Model: other noise models?
- Output Model: road category, direction, intersection regions, ...

Summary





Map Construction

[AW-12] Constructing Street Networks from GPS Trajectories.

Map Comparison

[AFW-14] Local Homology Based Distance Between Maps.

More References

[AG-95] H. Alt, M. Godau. Computing the Frchet Distance Between Two Polygonal Curves. IJCGA, 1995.

[BBW-09] K. Buchin, M. Buchin, Y. Wang. Exact Algorithms for Partial Curve Matching via the Fréchet Distance. SODA 2009.

[CW-10] A. F. Cook IV, C. Wenk. Geodesic Frchet Distance Inside a Simple Polygon. ACM TALG 7(1), 2010.

Thank You!

Contact me: brittany.fasy@alumni.duke.edu