Quality of Maps

Challenge 1: Building the right ground truth

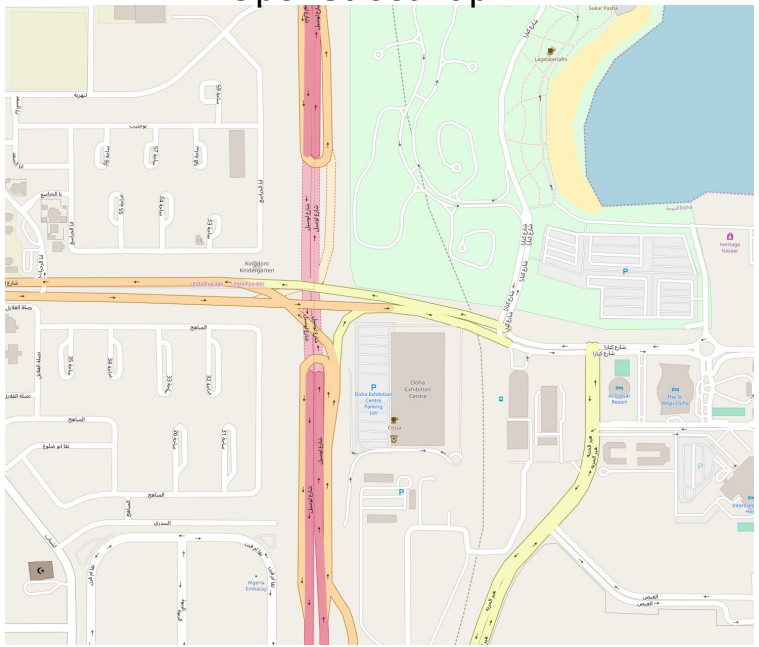
- Examples
- Deepglobe challenge @CVPR by Facebook /SpaceNet
  - Complaints about ground truth
  - Provide example (Favyen/Songtao)
- SpaceNet Challenge

Challenge 1: Building the right ground truth

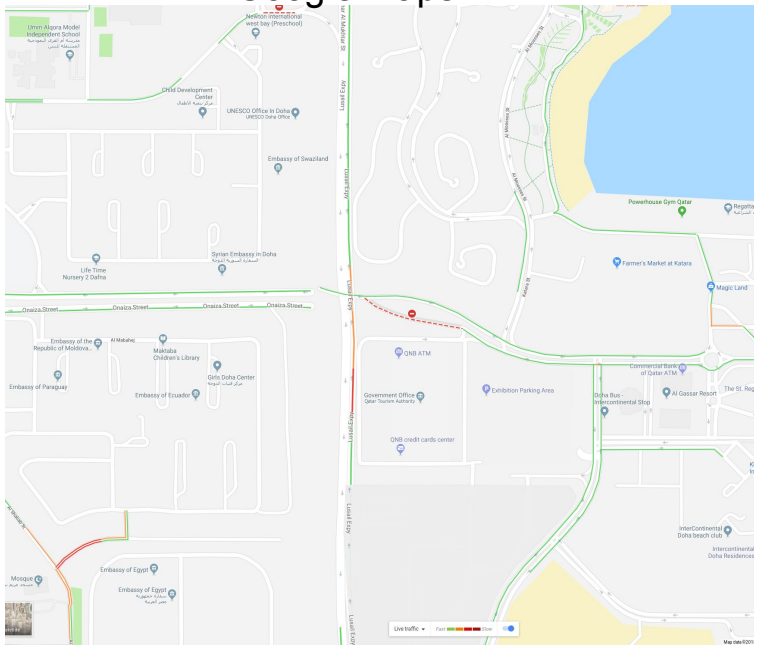
OSM is good for large scale implementations/comparisons

But, noisy, incomplete, and erroneous in some areas.

OpenStreetMap

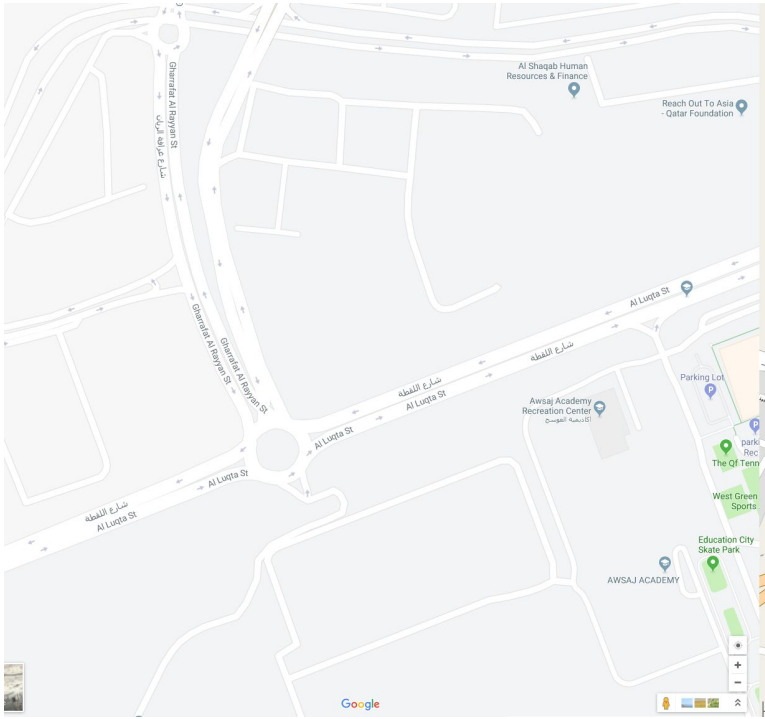


Google Maps

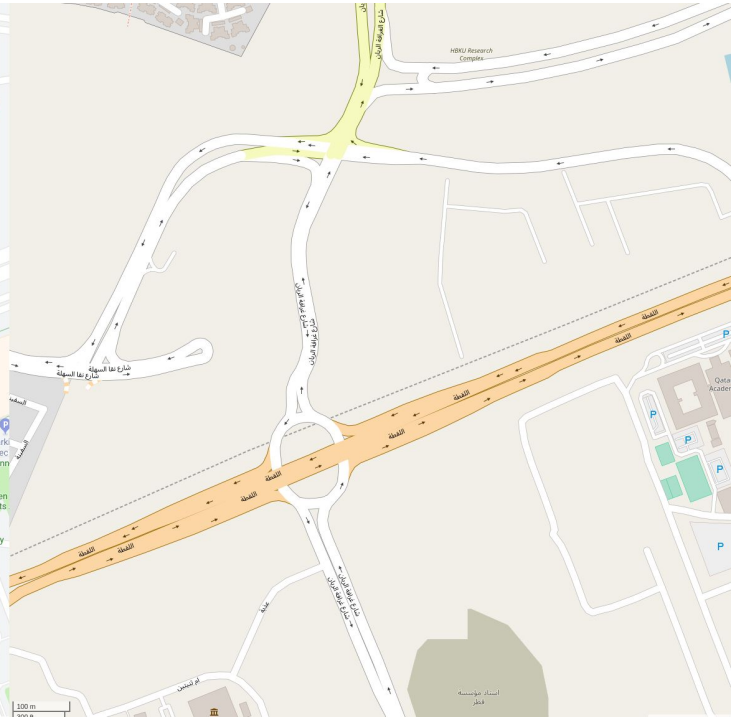


Challenge 1: Building the right ground truth

Google Maps

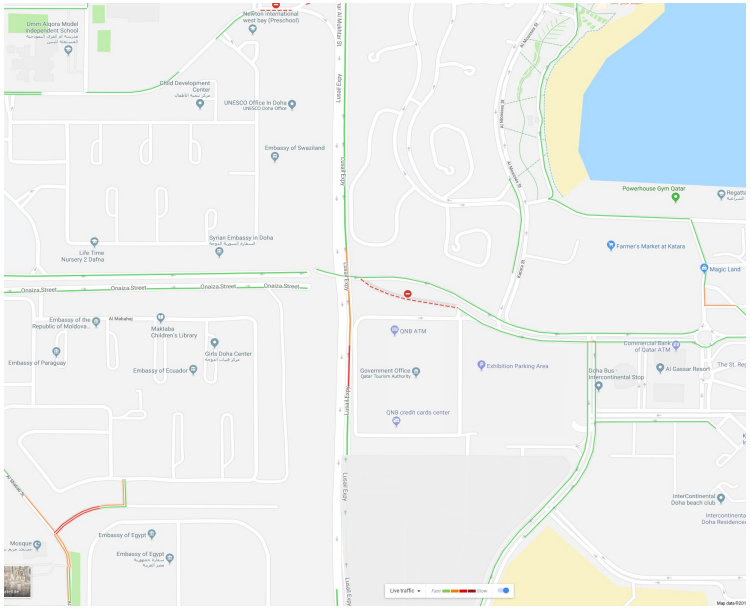


OpenStreetMap



Challenge 1: Building the right ground truth

Even Google is not that accurate!

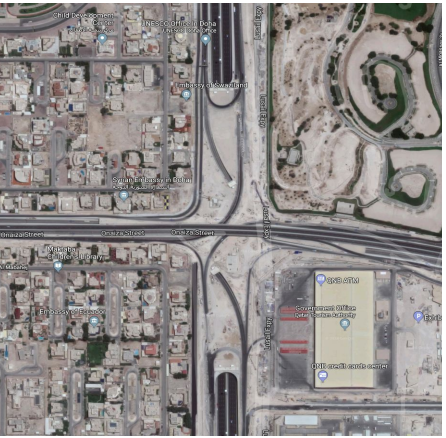


Challenge 1: Building the right ground truth

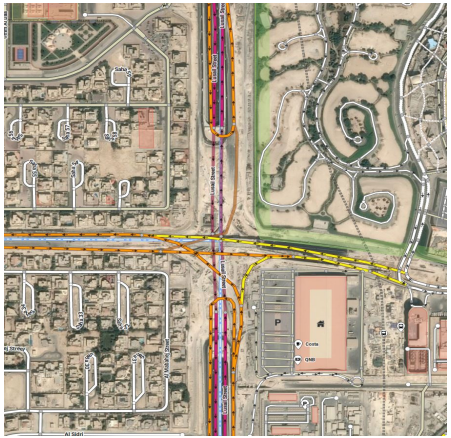
- Few more examples
- Should one use proprietary maps?
  - Cost?
  - Availability?
  - Scalability?

Challenge 2: Getting the right data

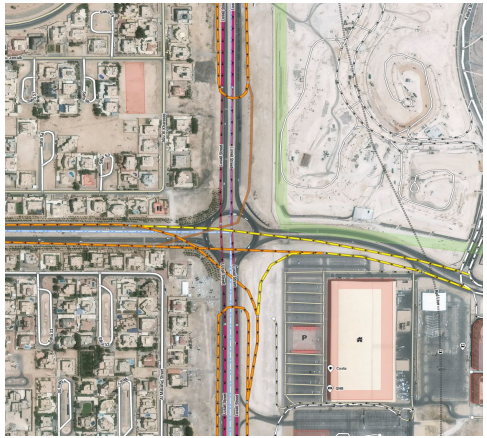
Satellite images?



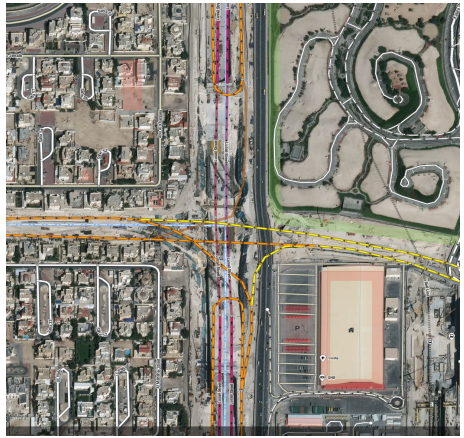
Google



Bing



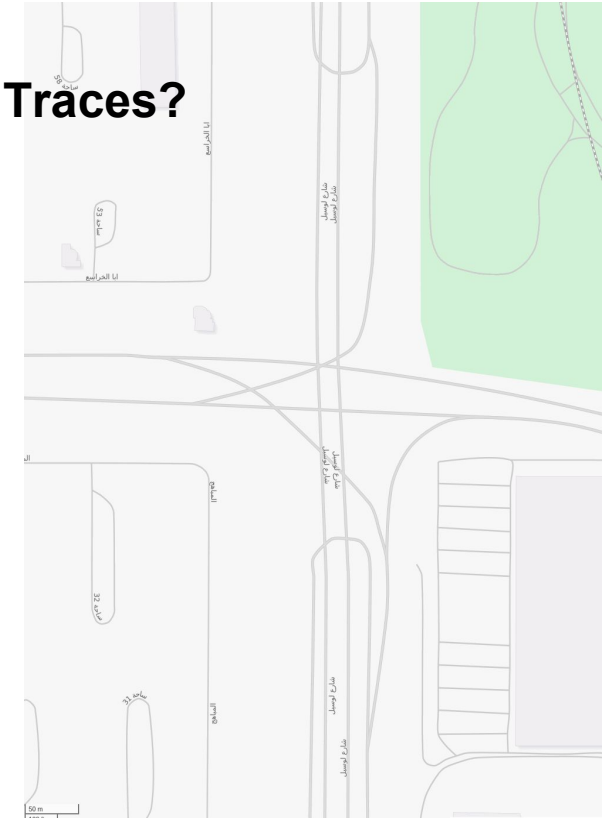
DigitalGlobe Standard

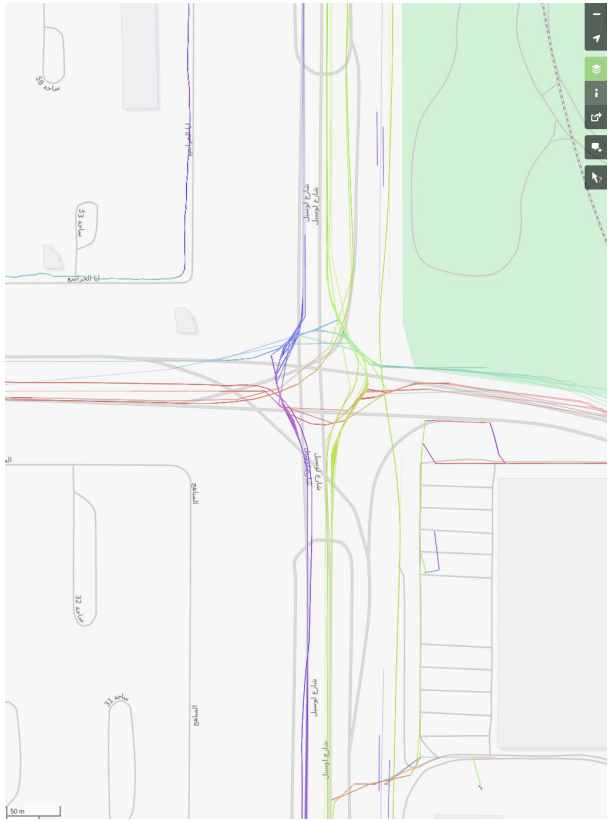


DigitalGlobe Premium

Challenge 2: Getting the right data

GPS Traces?





Challenge 2: Getting the right data

Add some examples, from Doha changes, youtube video, etc.

Challenge 3: Define the right evaluation Metric

Aspects to consider

- Structure and Geometry
- Function and Topology and Routing
- Humans vs. Machines:
  - Beauty, smoothness, etc.

Three Main Categories

- Pixel based distances
  - Precision/recall/F1
- Point set-based distances
  - Hausdorff
  - Holes and Marbles
- Path based distance
  - Path Length - APLS (SpaceNet Challenge)
  - Path Similarity - Frechet.
- Intersection over Union?

Mahmuda Ahmed, Sophia Karagiorgou, Dieter Pfoser, and Carola Wenk. 2015. **A comparison and evaluation of map construction algorithms using vehicle tracking data**. *Geoinformatica* 19, 3 (July 2015), 601-632

Pixel based distance

Pixel based distance

- Prominent in computer vision community
  - Prediction, recall, F1
  - Relaxed precision, recall, F1
- Adopted in DeepGlobe 2018 @CVPR

Pixel based distance: Problem

- Ignores topology
  - Connections, intersections, under/over passes, etc.
  - Not good for routing

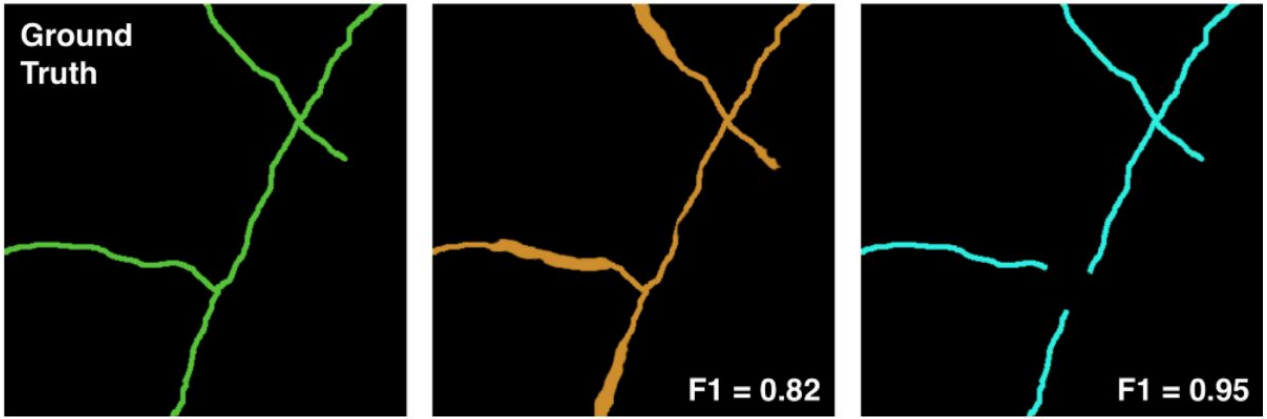


Figure 3. Example of the pitfalls of the pixel-based F1 metric. Left: Ground truth road mask in green. Middle: proposal mask in orange, which achieves a decent F1 score of 0.82 since most pixels are correctly labeled despite some inconsistencies in road width. Right: proposal mask in cyan, which achieves a superior F1 score of 0.95 since fewer pixels are classified incorrectly. For routing purposes, however, the rightmost plot is clearly inferior since it misses an important intersection and severs a road. The fact that pixel-based F1 scores incentivize the rightmost plot over the middle plot is suboptimal.

Credit. SpaceNet Challenge.

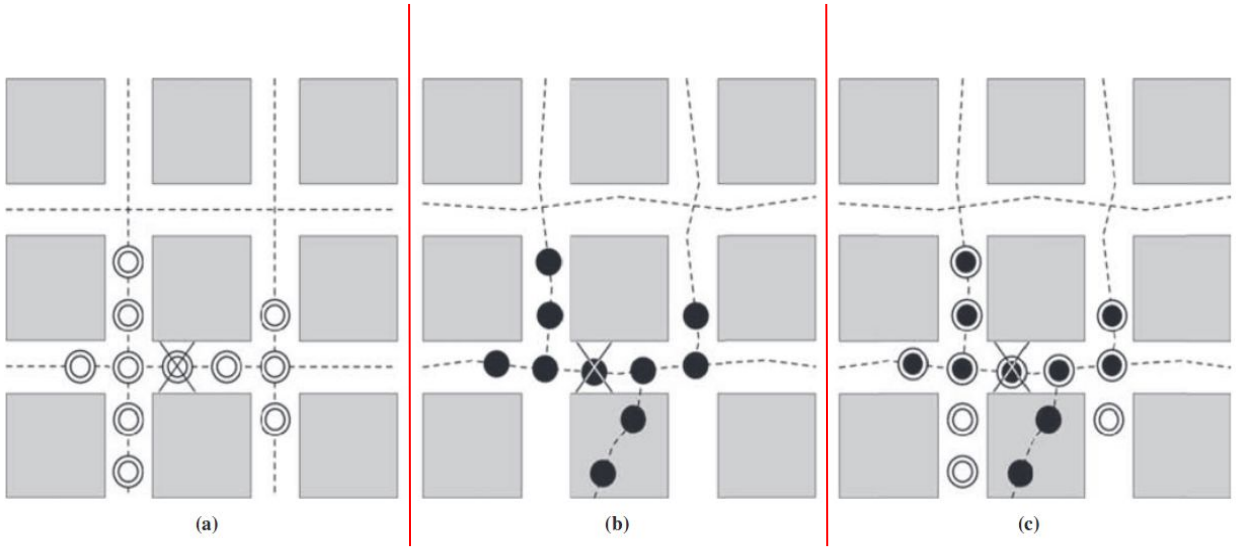
Point set-based distance

Point set-based distances

- Each map is represented as a set of points
- Compute distance between points
- Hausdorff distance
  - No Topology.
  - Given two sets A, B; $d(A,B) = \max_a \min_b (d(a, b))$
  - Very sensitive to outliers!
- Holes and Marbles
  - Sample the two maps
  - TOPO and GEO

Holes and Marbles

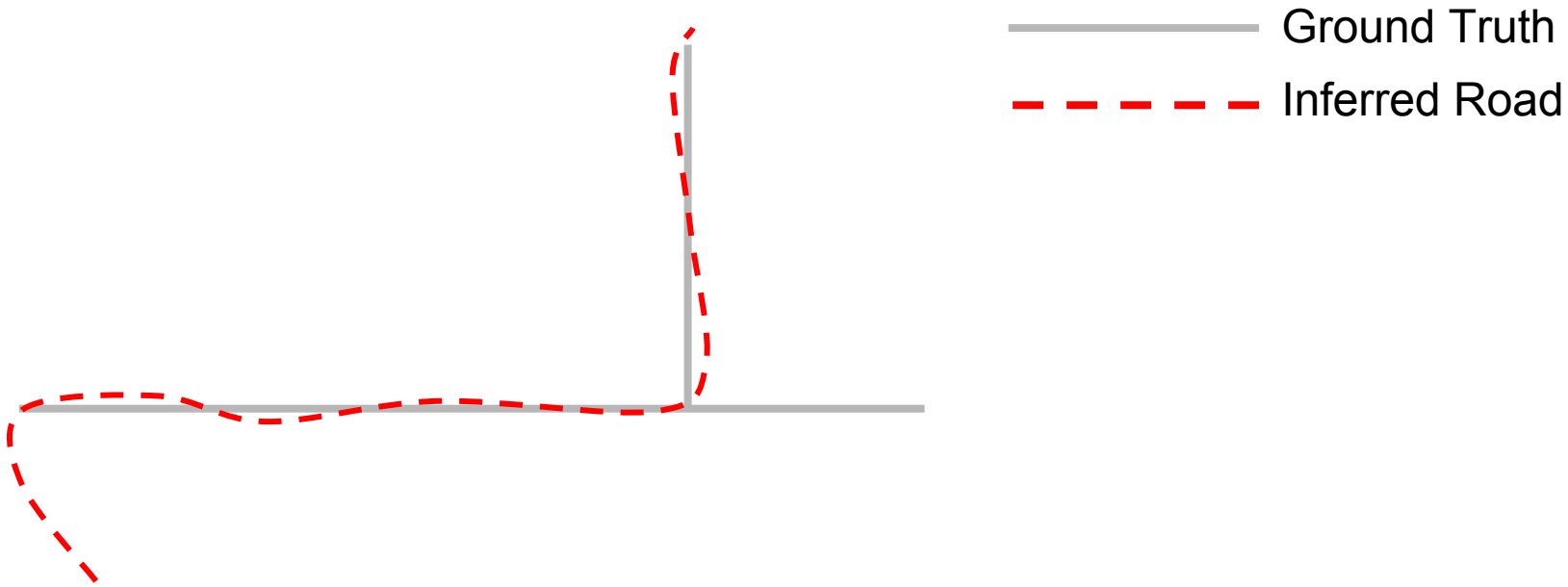
- GEO
- TOPO



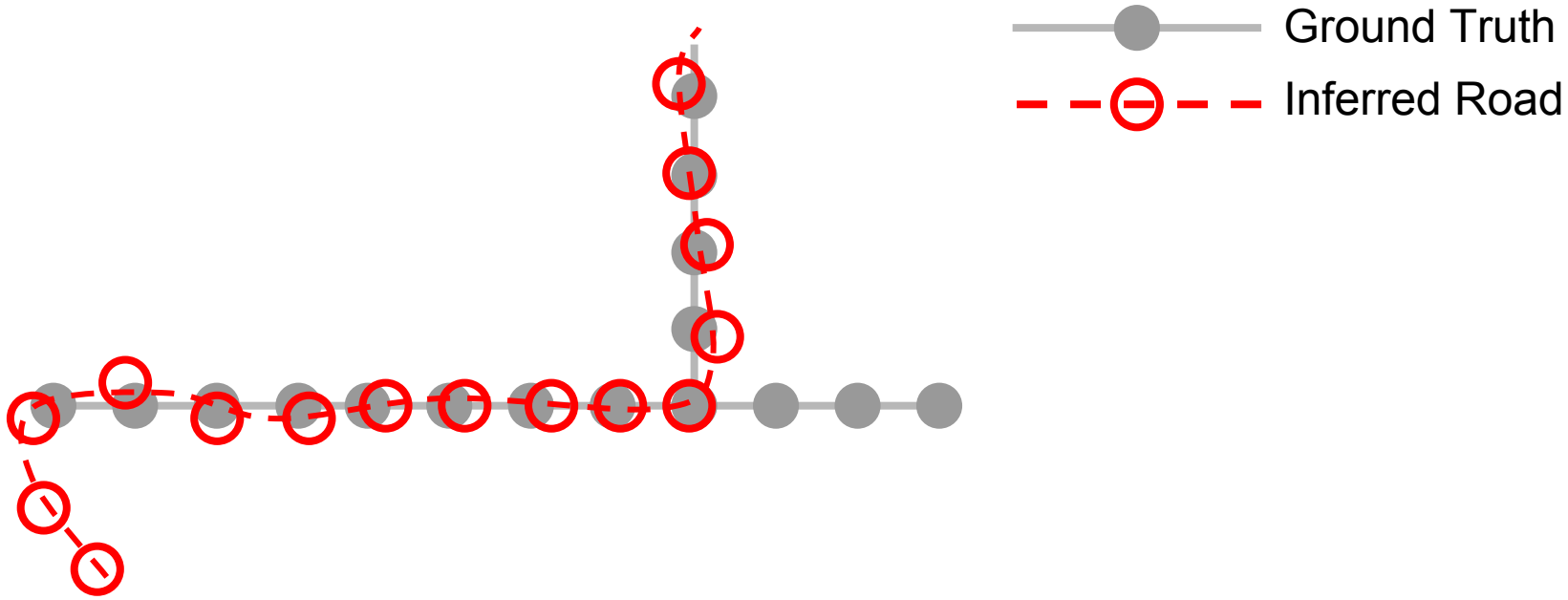
Evaluation Metrics

- Geometric evaluation (GEO)
- Graph-Sampling Based Distance (TOPO)
- Shortest Path Based Distance

Geometric evaluation (GEO)

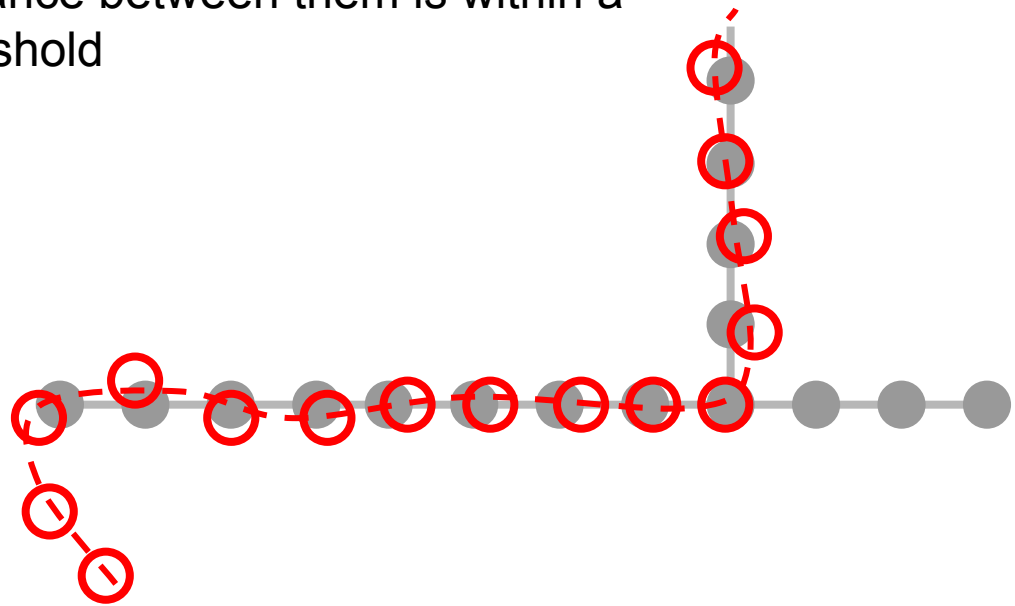


Geometric evaluation (GEO)



Geometric evaluation (GEO)

and could be matched if the distance between them is within a threshold

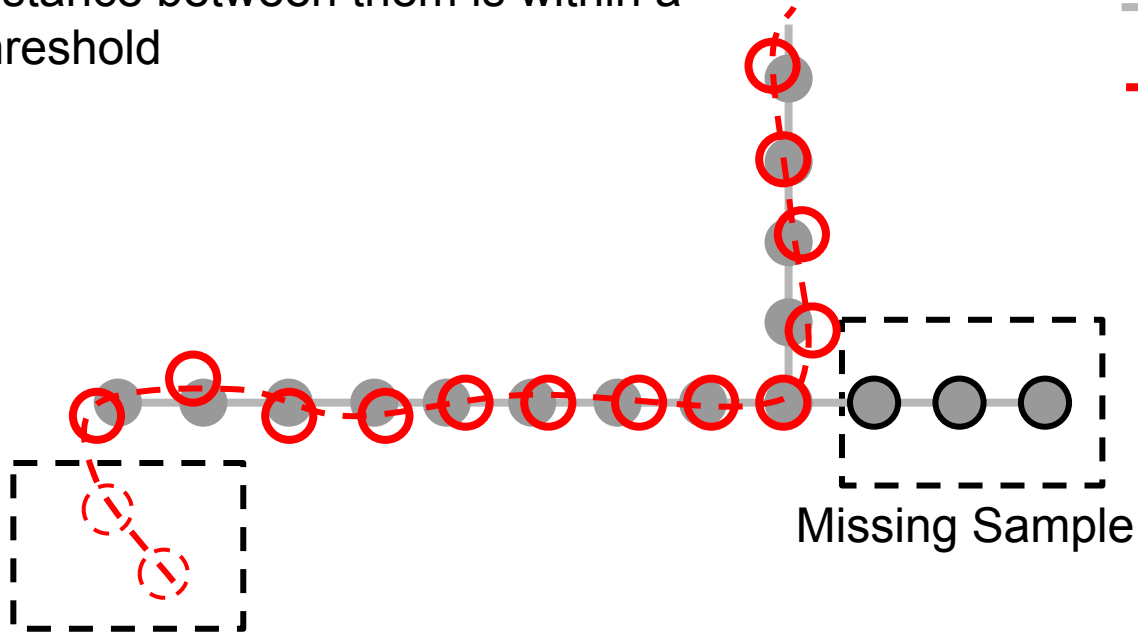


Ground Truth

Inferred Road

Geometric evaluation (GEO)

and could be matched if the distance between them is within a threshold



Ground Truth

Inferred Road

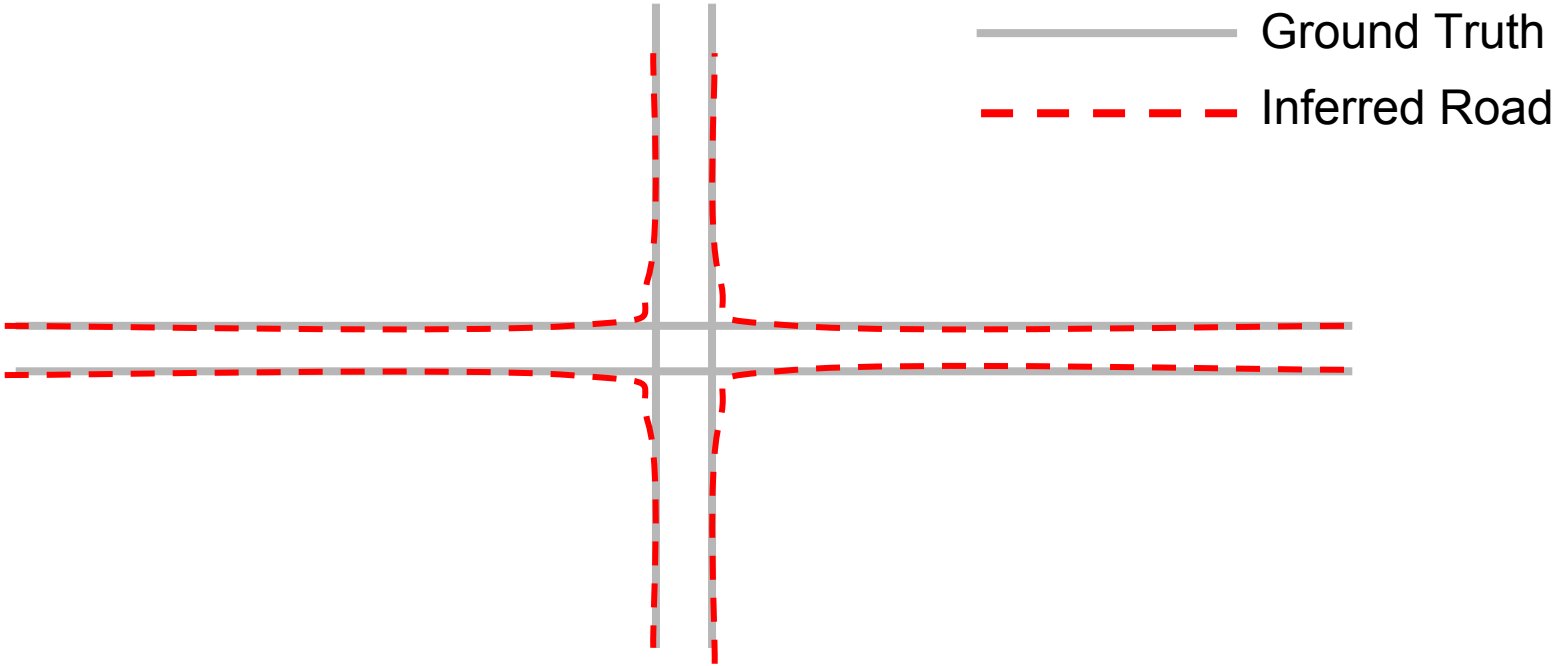
Overall performance

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

$$precision = 1 - \frac{|spurious\ samples|}{|inferred\ samples|}$$

$$recall = 1 - \frac{|missing\ samples|}{|ground\ truth\ samples|}$$

Geometric evaluation (GEO) Limitation



Graph-Sampling Based Distance (TOPO)

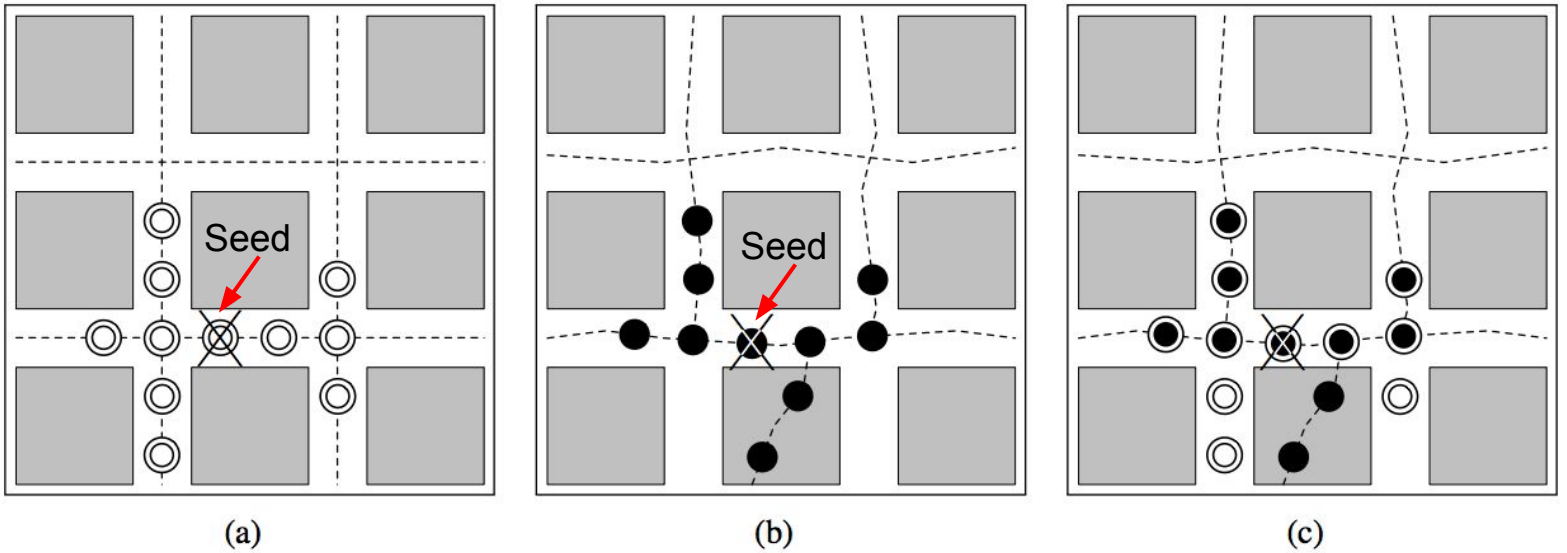
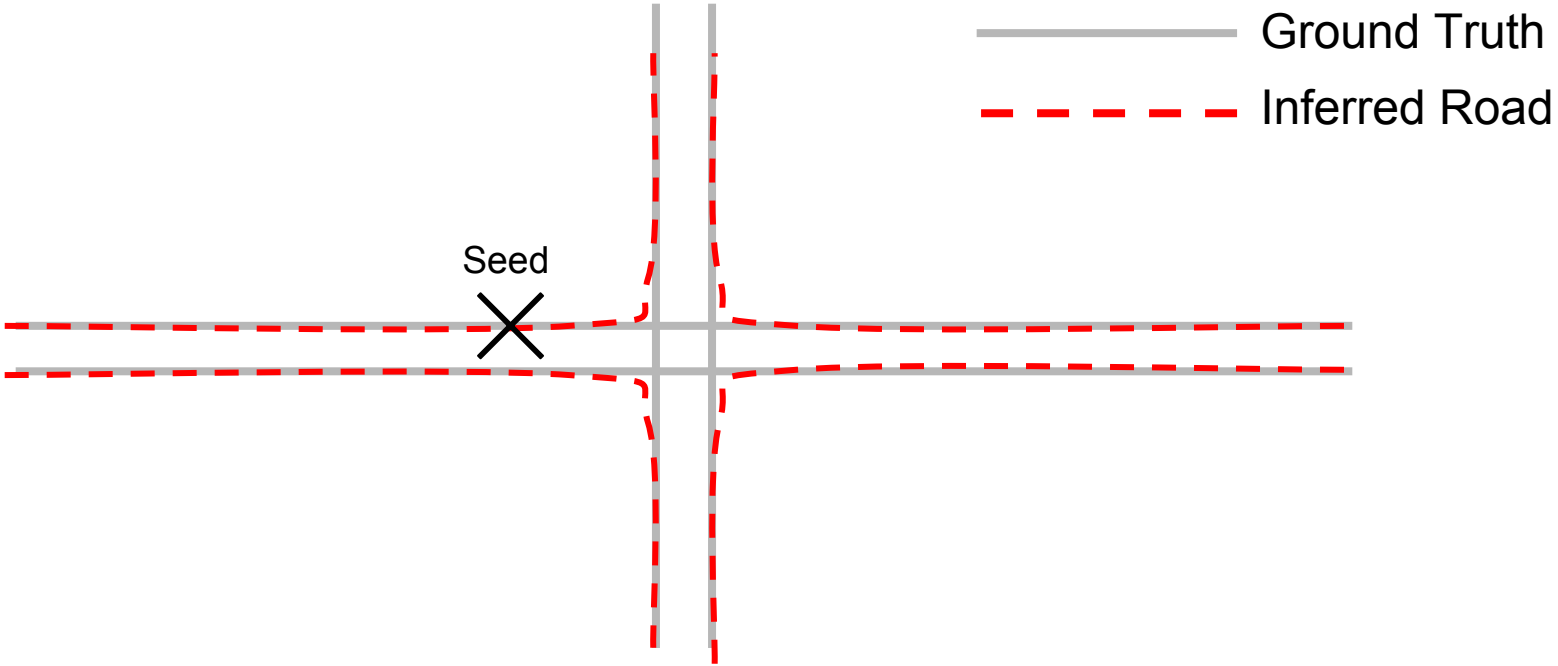


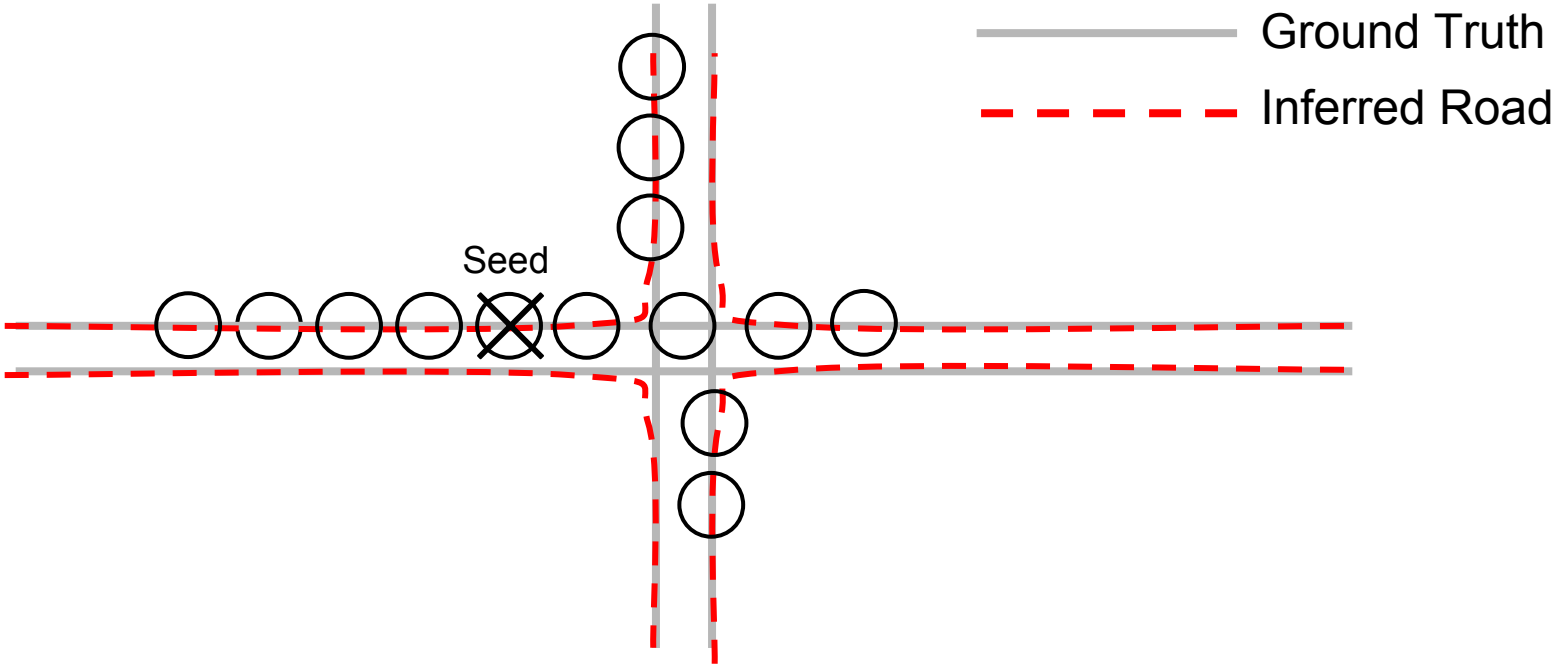
FIGURE 4 Overview of map comparison algorithm. (a) Holes are dropped at even intervals along edges of the ground truth map. (b) Marbles are dropped at even intervals along edges of the generated map. (c) Marbles from generated map fill holes where the maps overlap.

Eriksson, Jakob. "INFERRING ROAD MAPS FROM GPS TRACES: SURVEY AND COMPARATIVE EVALUATION 2 James Biagioni* 3 Ph. D. Student 4 Department of Computer Science 5."

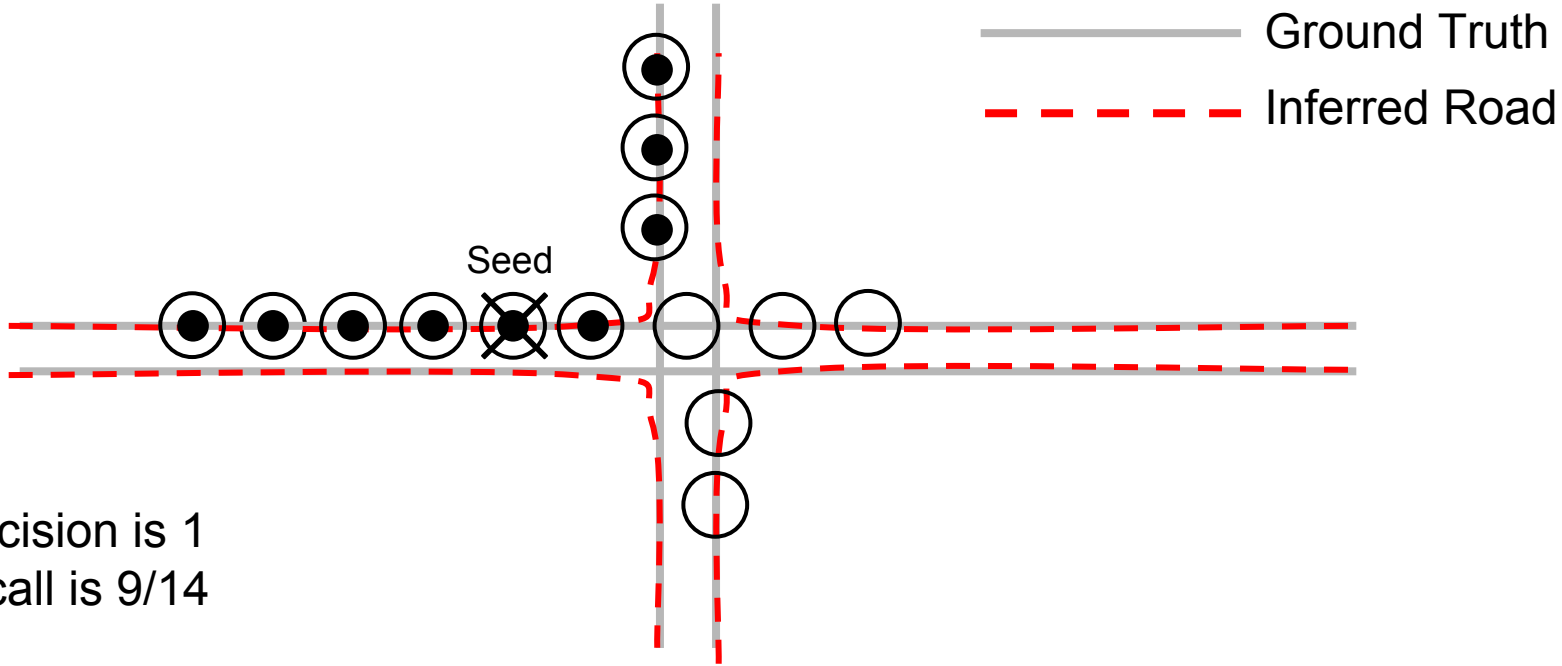
Graph-Sampling Based Distance (TOPO)



Graph-Sampling Based Distance (TOPO)



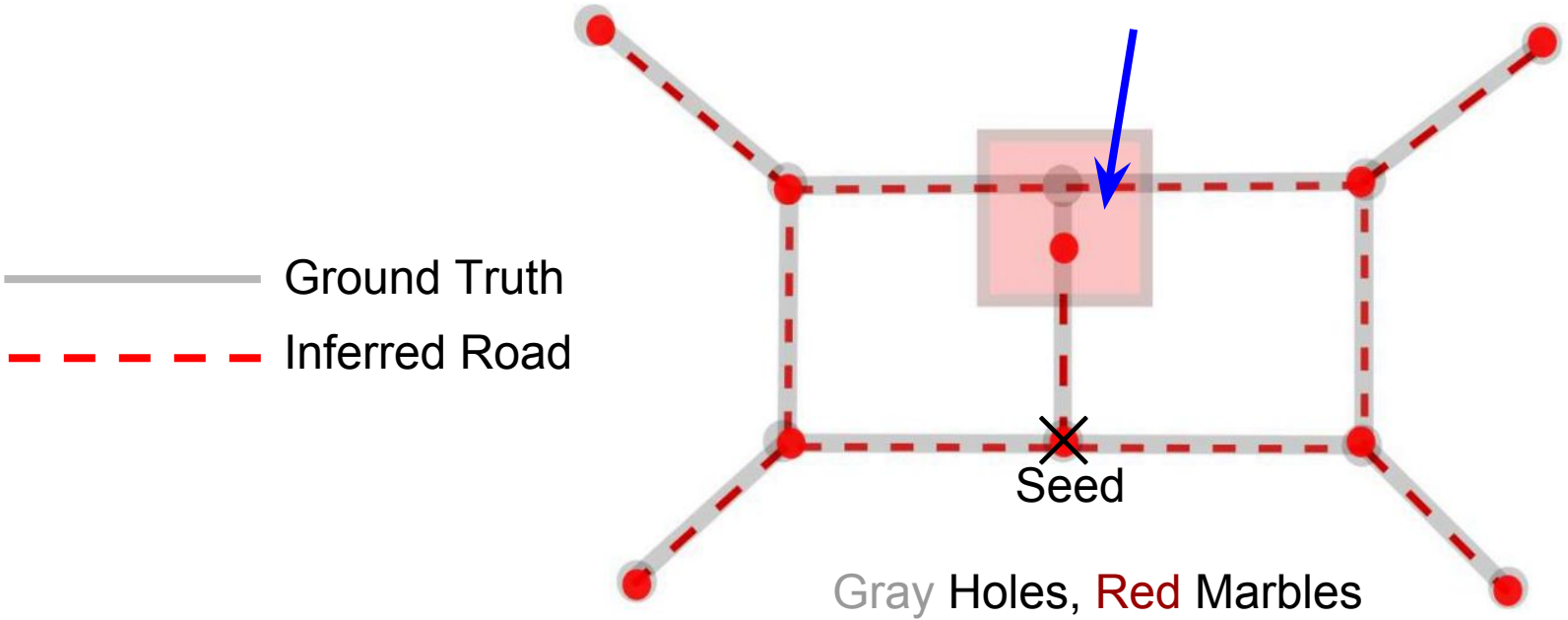
Graph-Sampling Based Distance (TOPO)



Precision is 1
Recall is 9/14

Graph-Sampling Based Distance (TOPO) Limitation

TOPO may fail to capture the broken connection

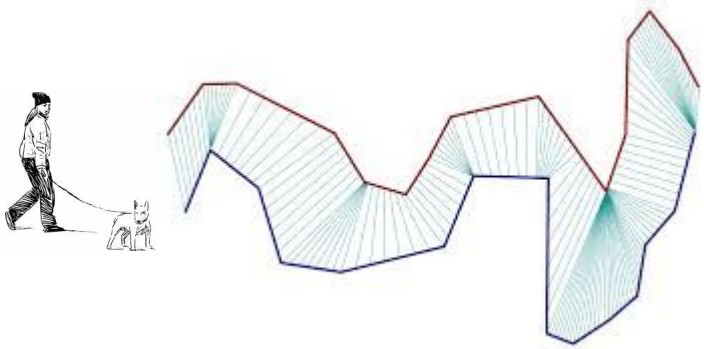


Ahmed, Mahmuda, et al. "A comparison and evaluation of map construction algorithms using vehicle tracking data." *GeoInformatica* 19.3 (2015): 601-632.

Path based distance

Path based distances

- Each map is represented by a set of paths.
- Unique paths, shortest paths, sampling
- Path similarity
  - Discrete frechet distance.
  - Frechet distance idea: the minimum leash length such that a man and a dog can walk on the two curves (paths) generated from G.
  - $d(P_A, P_B) = \min_{a \in PA} \max_{b \in PB} (\text{Frechet}(a, b))$
- Path length
  - APLS



APLS: Average Path Length Similarity

- Graph theory based
- Considers topology
- Good for routing

$$C = 1 - \frac{1}{N} \sum min\left\{1, \frac{|L(a,b) - L(a',b')|}{L(a,b)}\right\}$$

$$N = number\ of\ unique\ paths$$
$$L(a,b) = length\ of\ path(a,b)$$

Equation 1. APLS metric. Node a′ is the node in the proposal graph G′ nearest the location of node a in the ground truth graph G. L(a,b) denotes a path distance in the ground truth graph G, and L(a′, b′) denotes the path length between the corresponding nodes in the proposal graph.

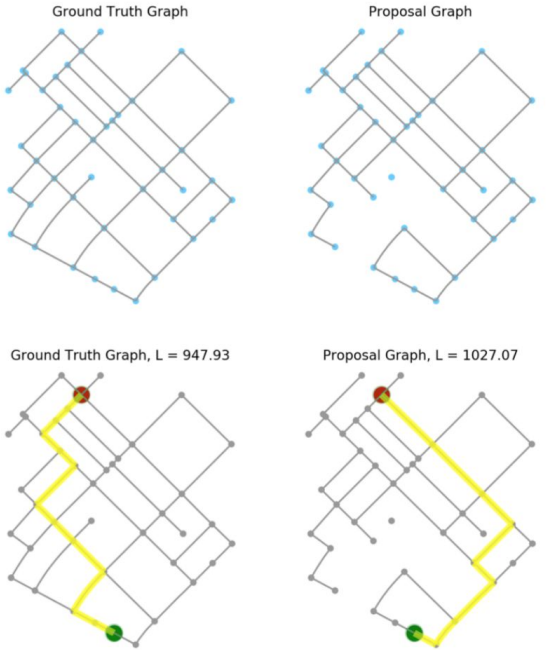


Figure 4. Demonstration of path length difference between sample ground truth and proposal graphs. Upper Left: Ground truth graph. Upper Right: Proposal graph with 30 edges removed. Lower Left: Shortest path between source (green) and target (red) node in the ground truth graph is shown in yellow, with a path length of ~948 meters. Lower Right: Shortest path between source and target node in the proposal graph, with a path length of ~1027 meters; this difference in length forms the basis for our graph similarity metric. Plotting is accomplished via the excellent osmnx package.

APLS vs. F1

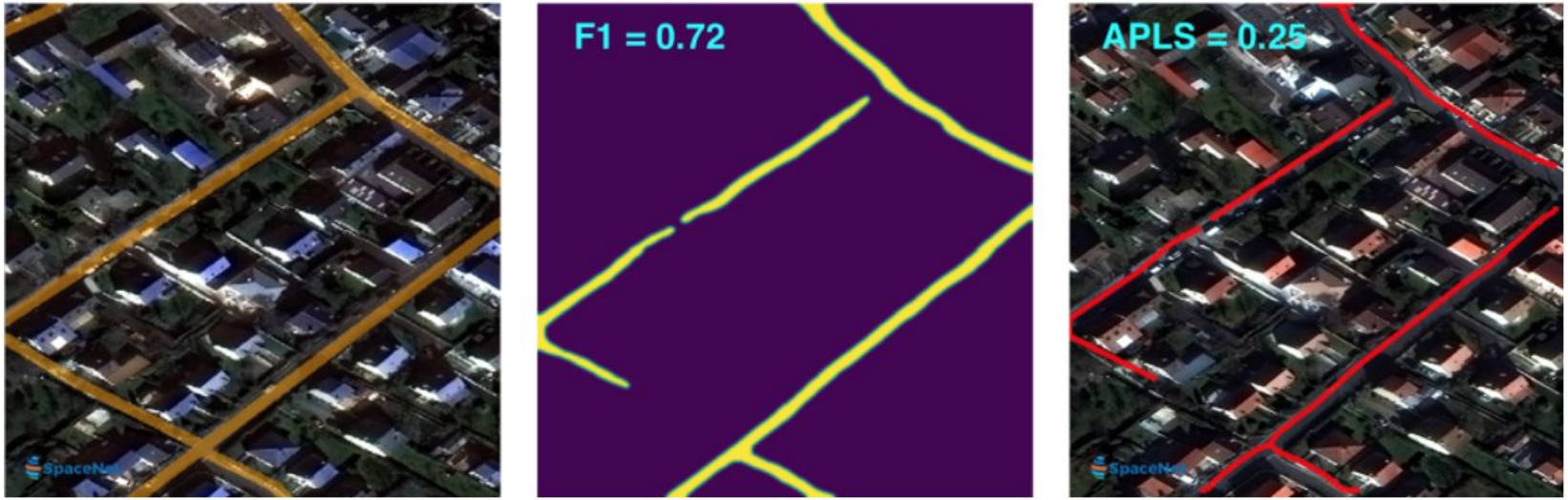


Figure 6. Comparison of F1 and APLS scores for a given proposal. Left: Sample satellite image with ground truth mask overlaid in orange. Middle: Proposal mask in yellow, yielding an F1 score of 0.72. Right: Proposal graph (red) inferred from the proposal mask; missing intersections and road segments are heavily penalized, yielding an APLS score of 0.25.

APLS

How to map nodes in the GT to nodes in the generated map?

Node semantics: we define **control** nodes as road **endpoints**, **intersections**, or **midpoints** along routes of interest.

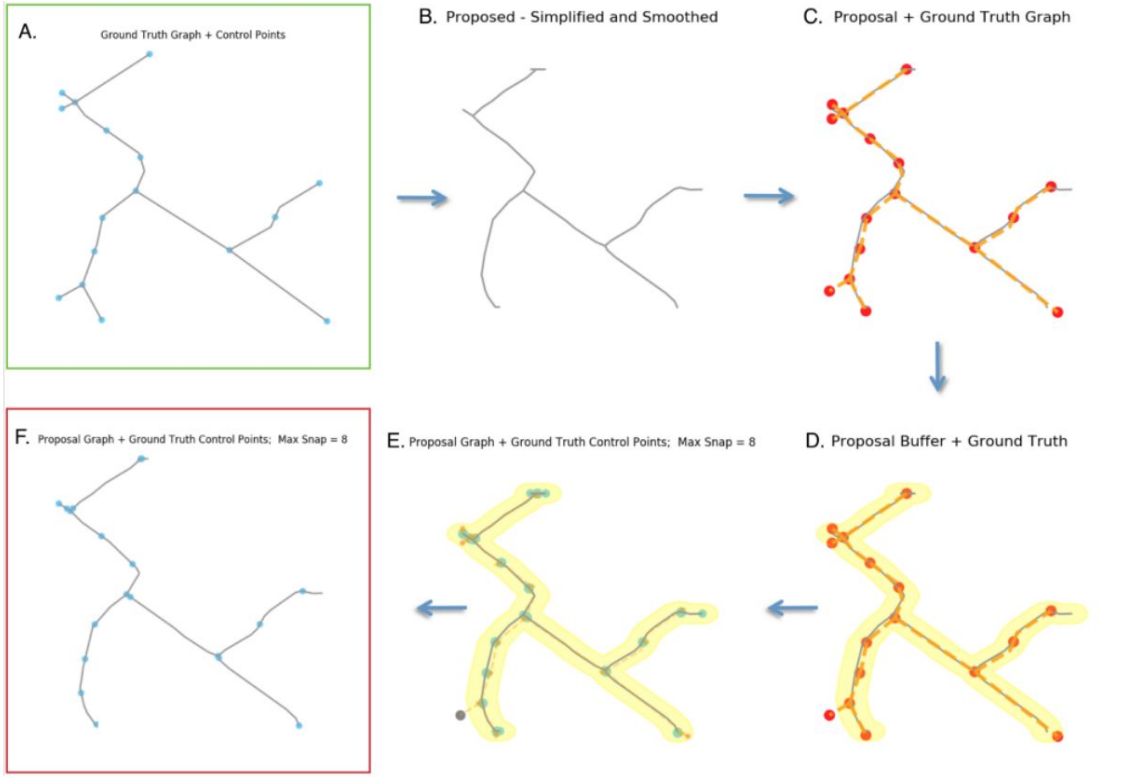


Figure 1. Node snapping procedure. A: Ground truth graph with control nodes. B: Proposal graph. C: Ground truth graph (orange) and ground truth control points (red) overlaid on proposal graph (grey). D: Same as graph C, but showing an 8 meter buffer (yellow) around the proposal graph. E: Ground truth control nodes are injected into the proposal graph at the nearest edge, except for nodes outside the buffer (grey). F: Final proposal graph with nodes injected at the appropriate location to compare to graph A.

APLS: Symmetric comparison

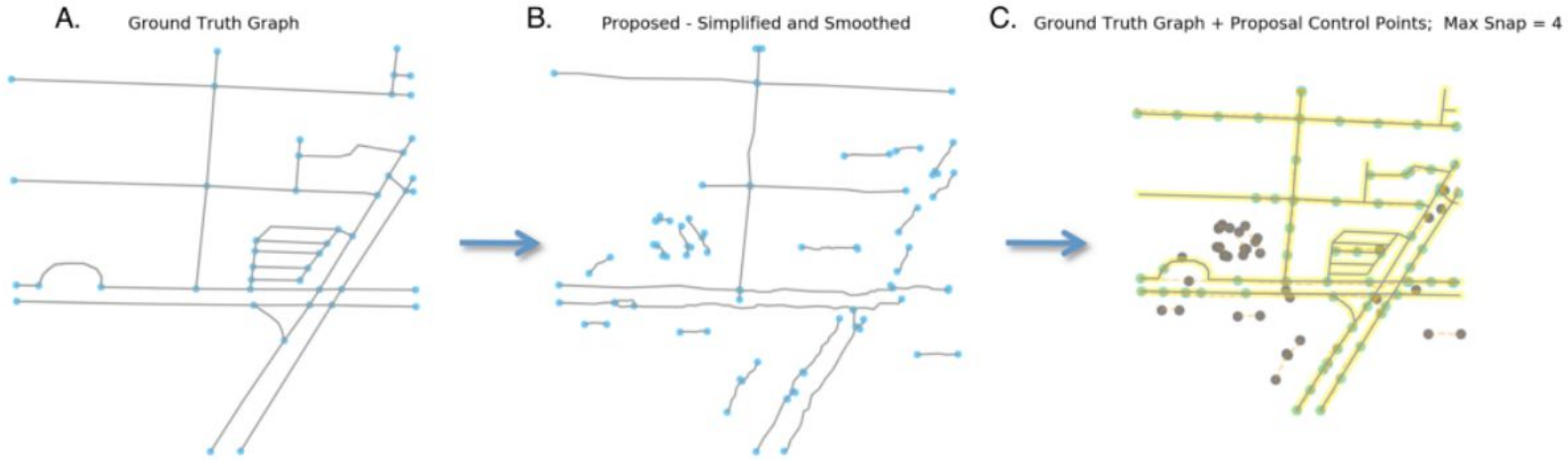
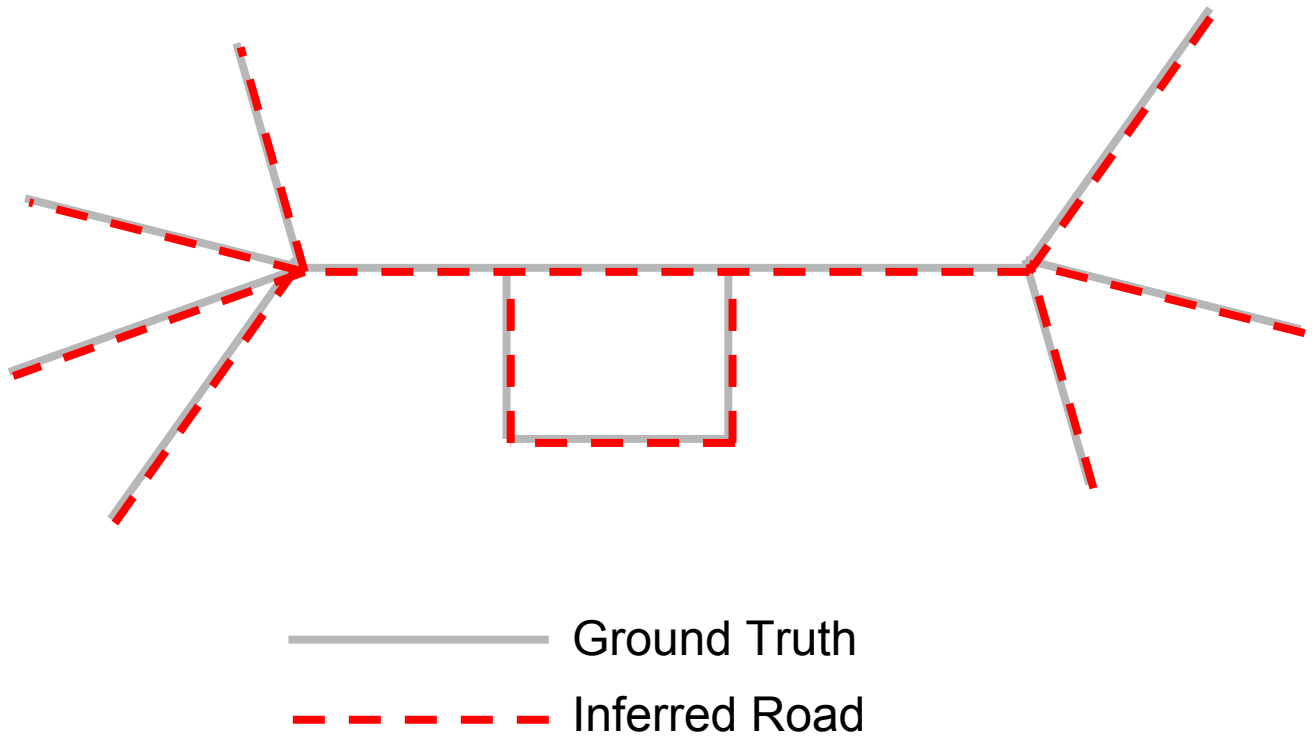
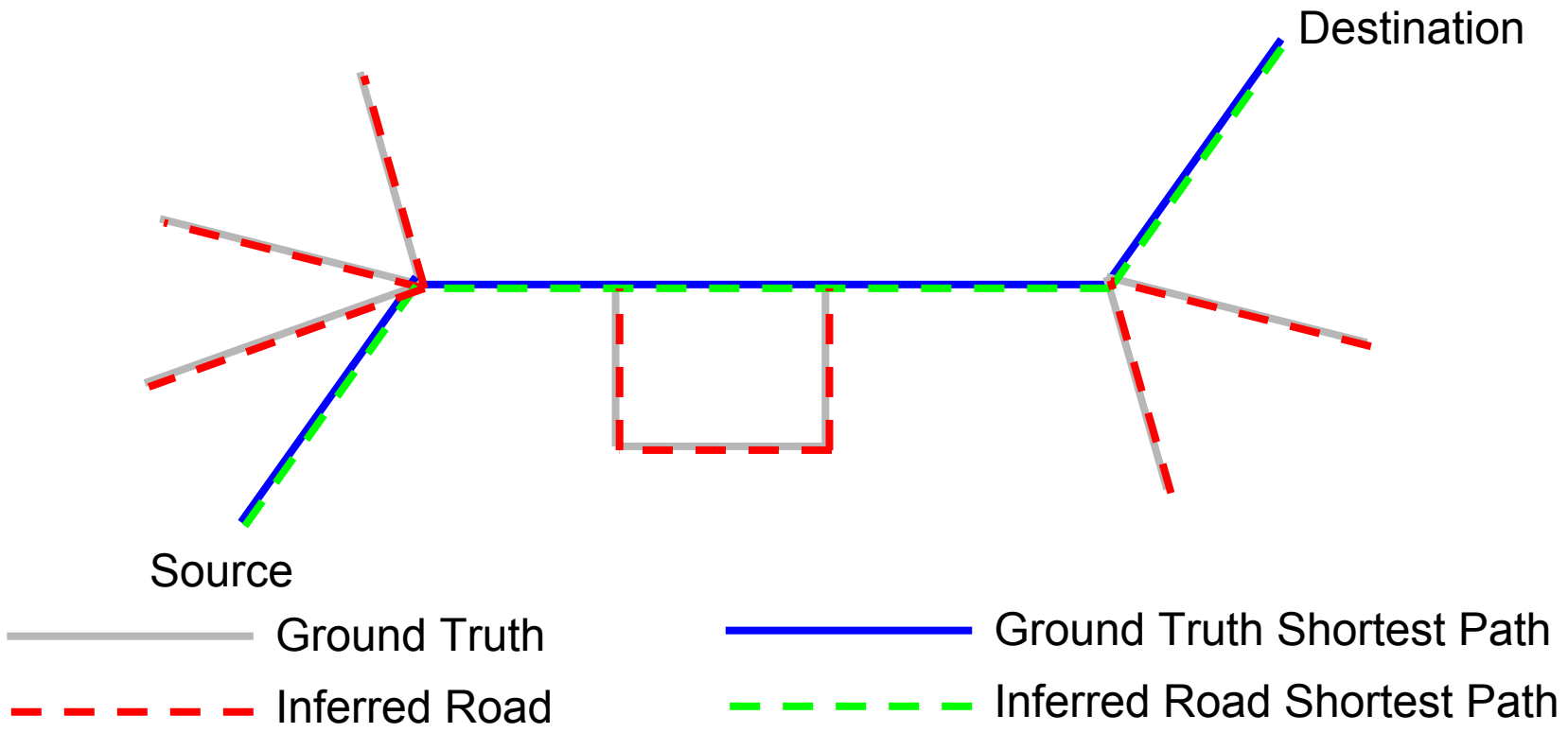


Figure 2. Illustration of the need to apply Equation 1 symmetrically (i.e.: ground truth to proposal, and proposal to ground truth). A: Ground truth graph. B: Proposal graph with many short, spurious connections well outside the buffer. C: Proposal nodes snapped onto the ground truth graph; snapping proposal control points onto the ground truth graph and then computing the metric penalizes the many extraneous predictions (grey nodes).

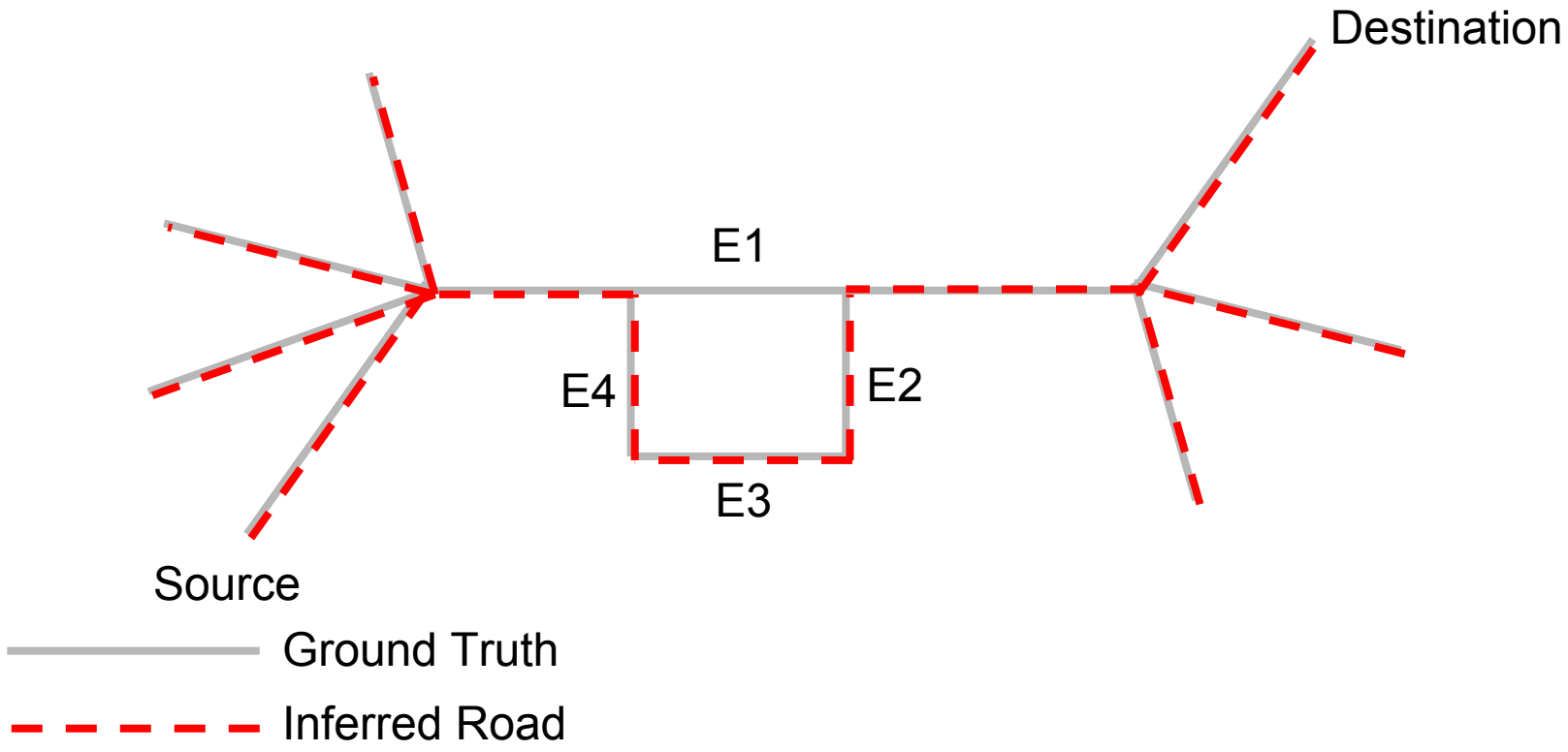
Shortest Path Based Distance



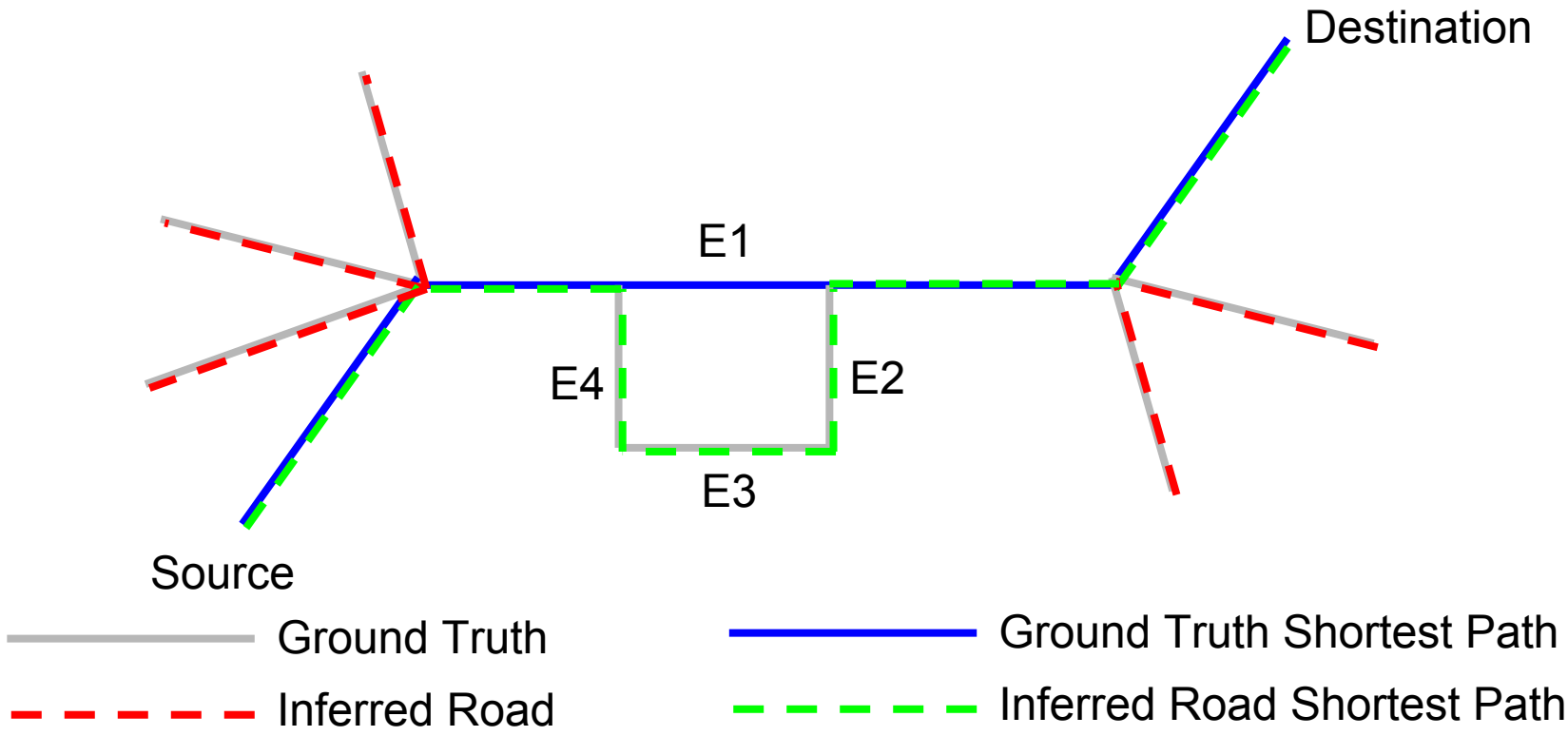
Shortest Path Based Distance



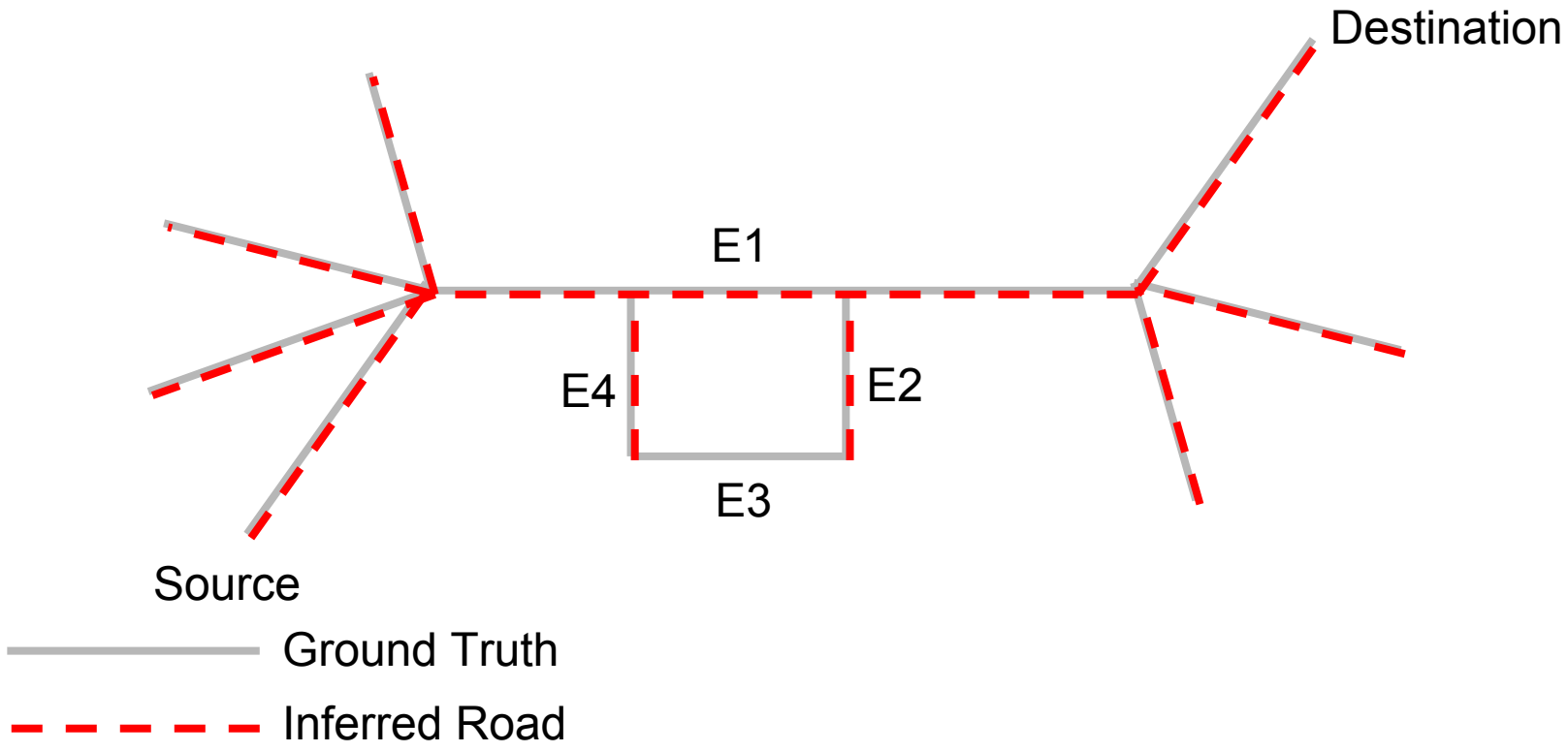
Shortest Path Based Distance



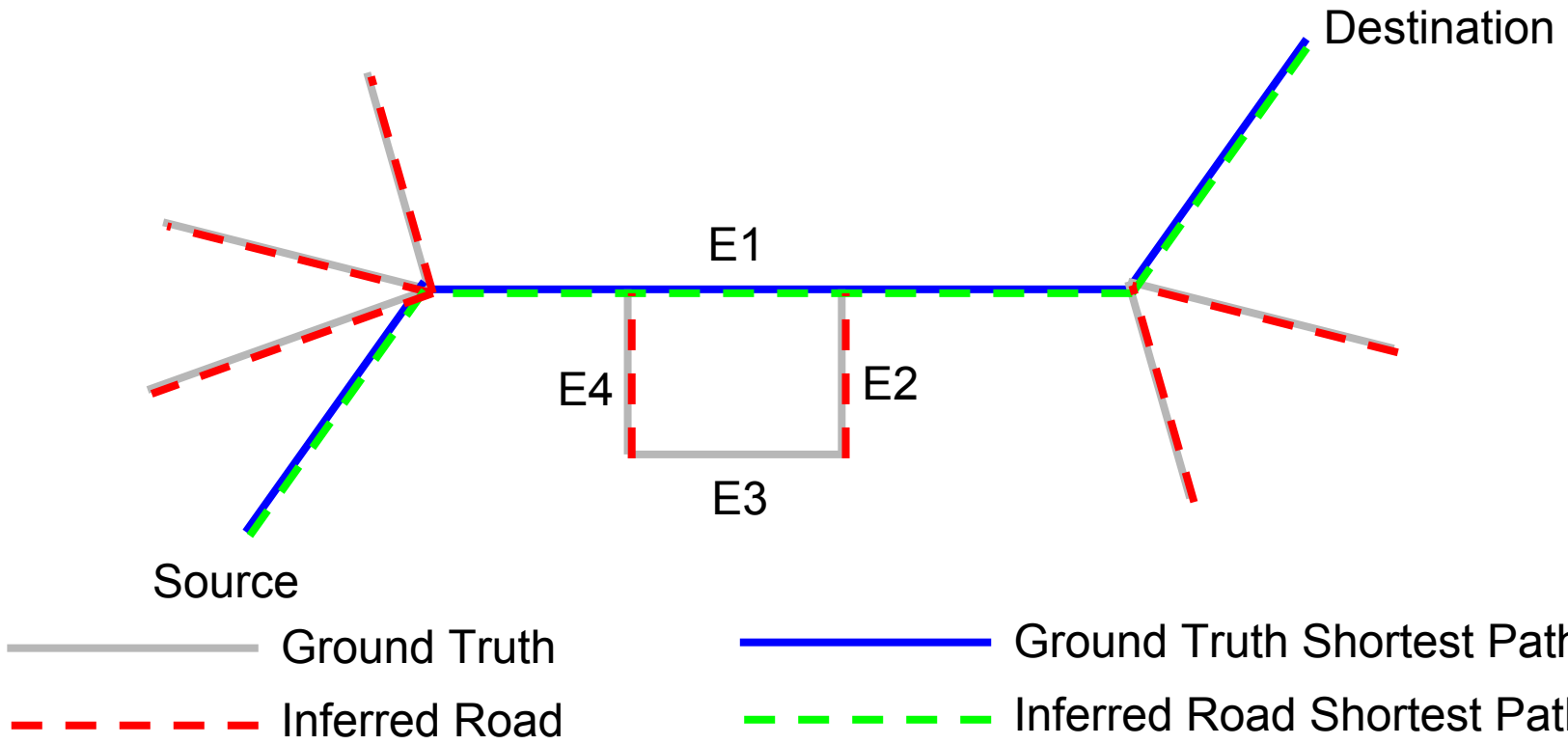
Shortest Path Based Distance



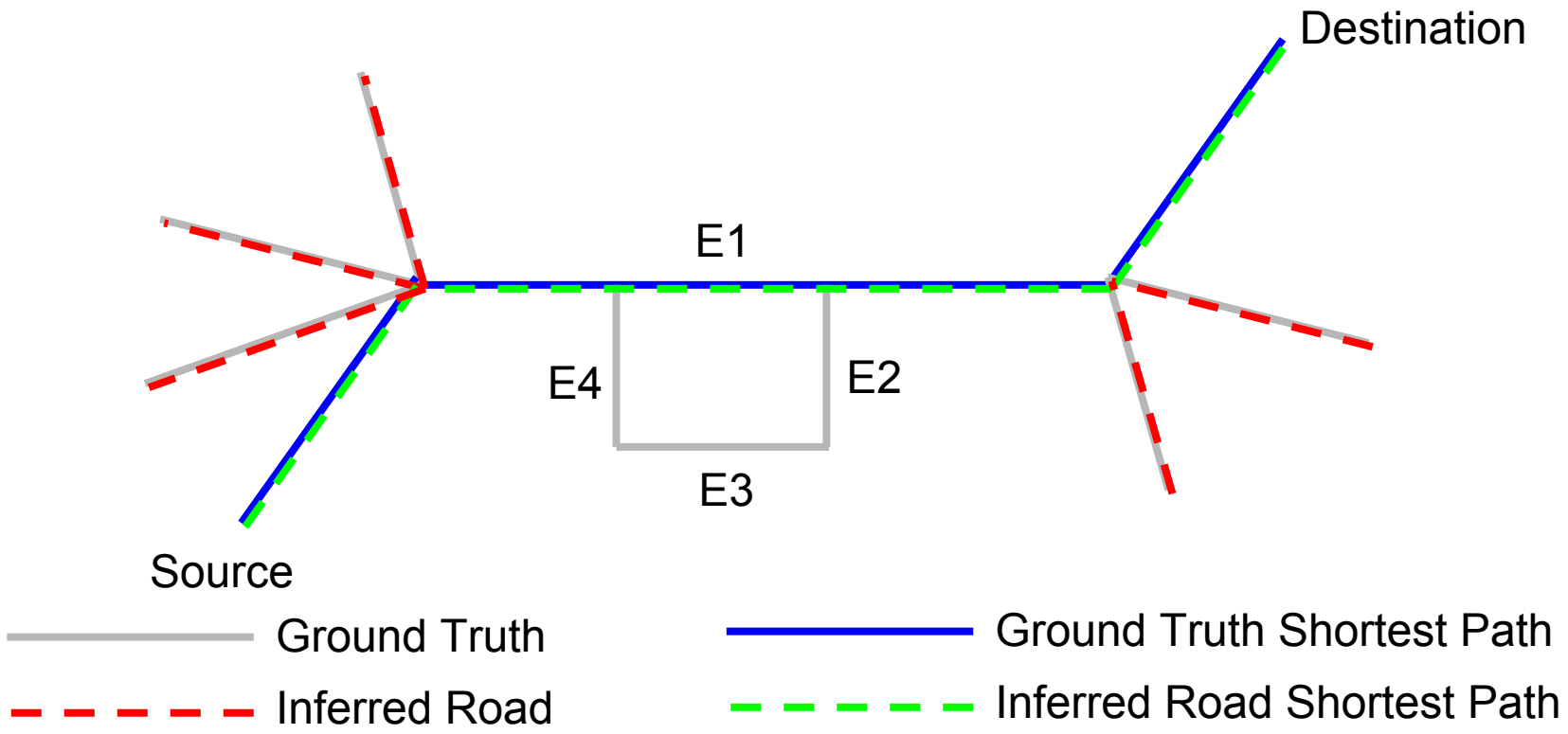
Shortest Path Based Distance



Shortest Path Based Distance



Shortest Path Based Distance



Literature review

Check all Map Making papers, especially recent ones, and summarize metrics they used.

References

- [relaxed Precision/Recall] Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*.
- [relaxed Precision/Recall] Mnih and G. Hinton, “Learning to detect roads in high-resolution aerial images,” ECCV , pp. 210–223, 2010

Methods from Ahmed et al.

-

Challenge 3: Define the right evaluation Metric

- Directed Hausdorff distance
- Path based distance
- Shortest paths based distance
- Holes and marbles
- Intersection over Union
- Precision/recall/F1 , relaxed precision/recall (pixel within x labeled pixel.)

Mahmuda Ahmed, Sophia Karagiorgou, Dieter Pfoser, and Carola Wenk. 2015. **A comparison and evaluation of map construction algorithms using vehicle tracking data**. *Geoinformatica* 19, 3 (July 2015), 601-632