TRAFFIC FLOW ESTIMATION FROM CELLULAR NETWORK DATA

Master's thesis presentation

Nils Breyer

August 27, 2015

LiU/ITN, UC Berkeley (USA)

WHY ARE LINKFLOWS OF INTEREST?

Linkflow

number of vehicles or people using a link during a certain time

WHY ARE LINKFLOWS OF INTEREST?

Linkflow

number of vehicles or people using a link during a certain time

Usage

- Traffic planning
- Traffic control

DATASOURCES

Travel surveys

Link counts

DATASOURCES

Travel surveys

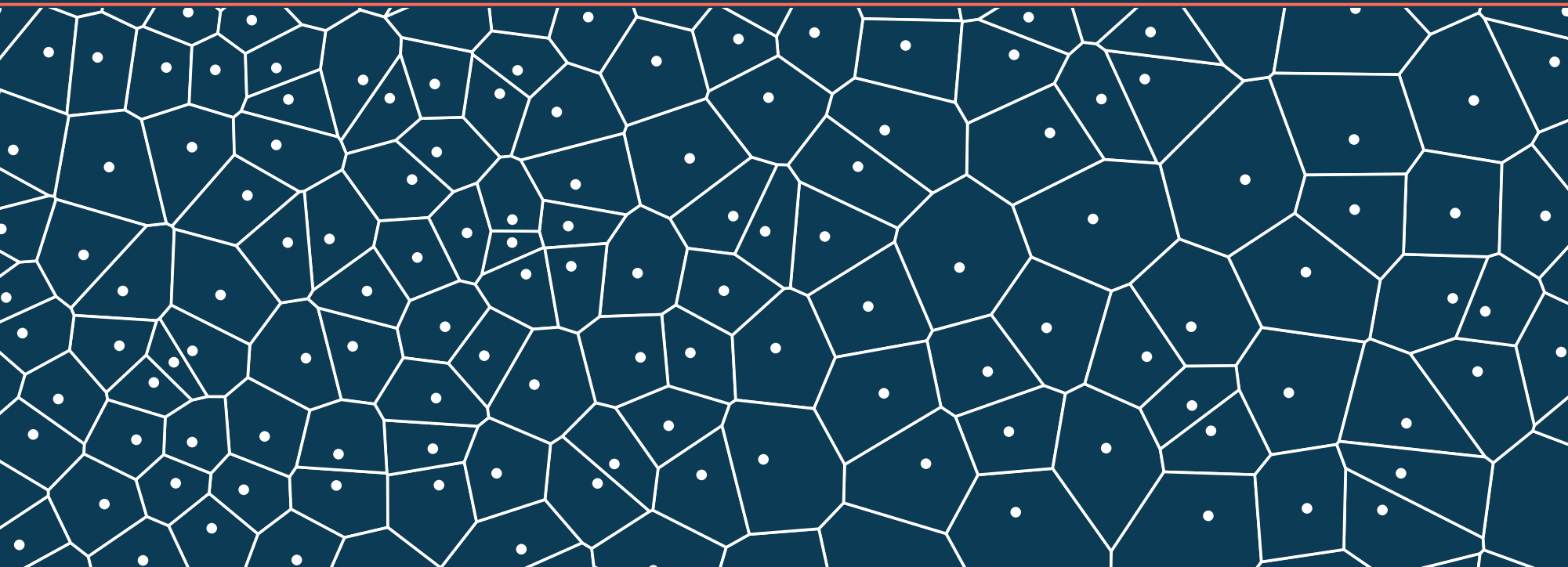
Link counts

Cellular networks

OUTLINE

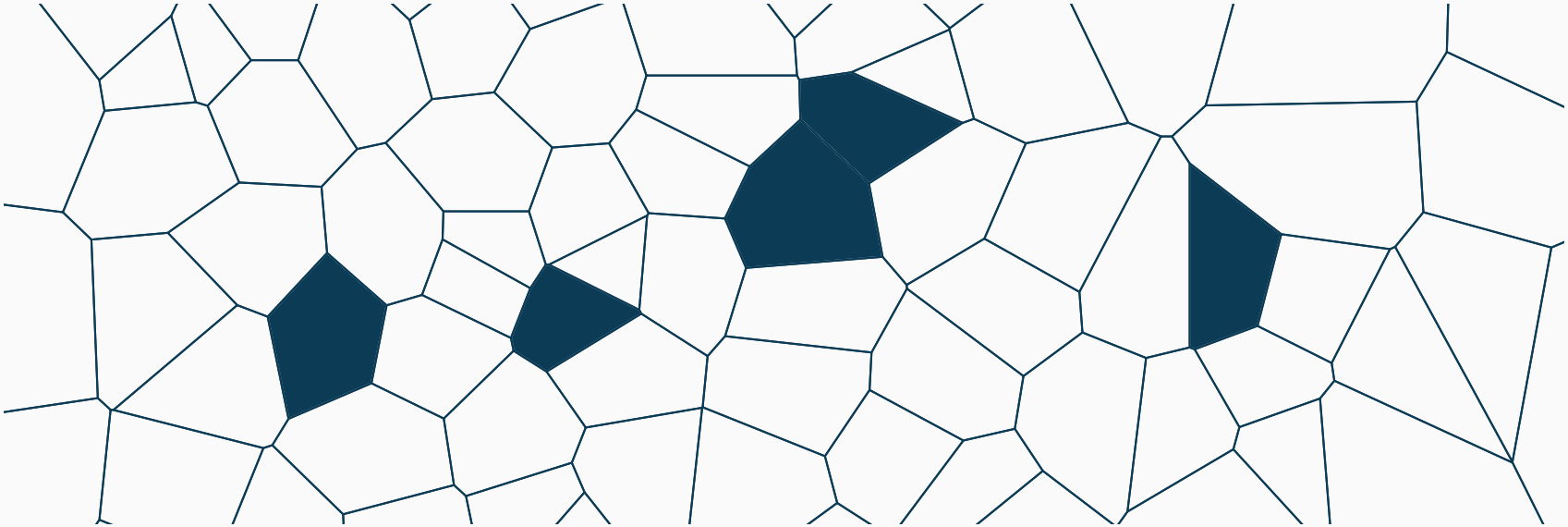
- Overview of the model
- Trip extraction and travel demand estimation
- Cellpath routing
- Network loading
- Results & validation

Overview of the model



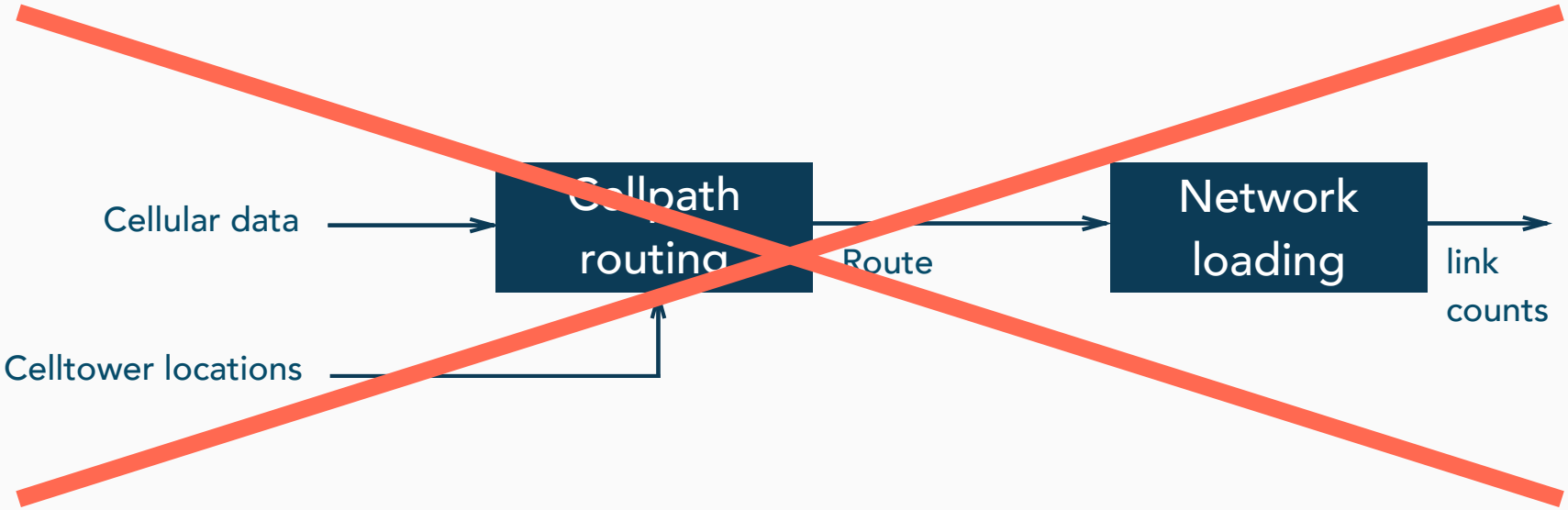
Cellpath

a sequence of cells that a user connected to along a trip

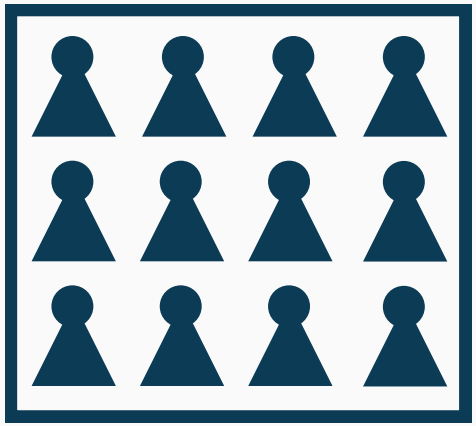


OVERVIEW OF THE MODEL

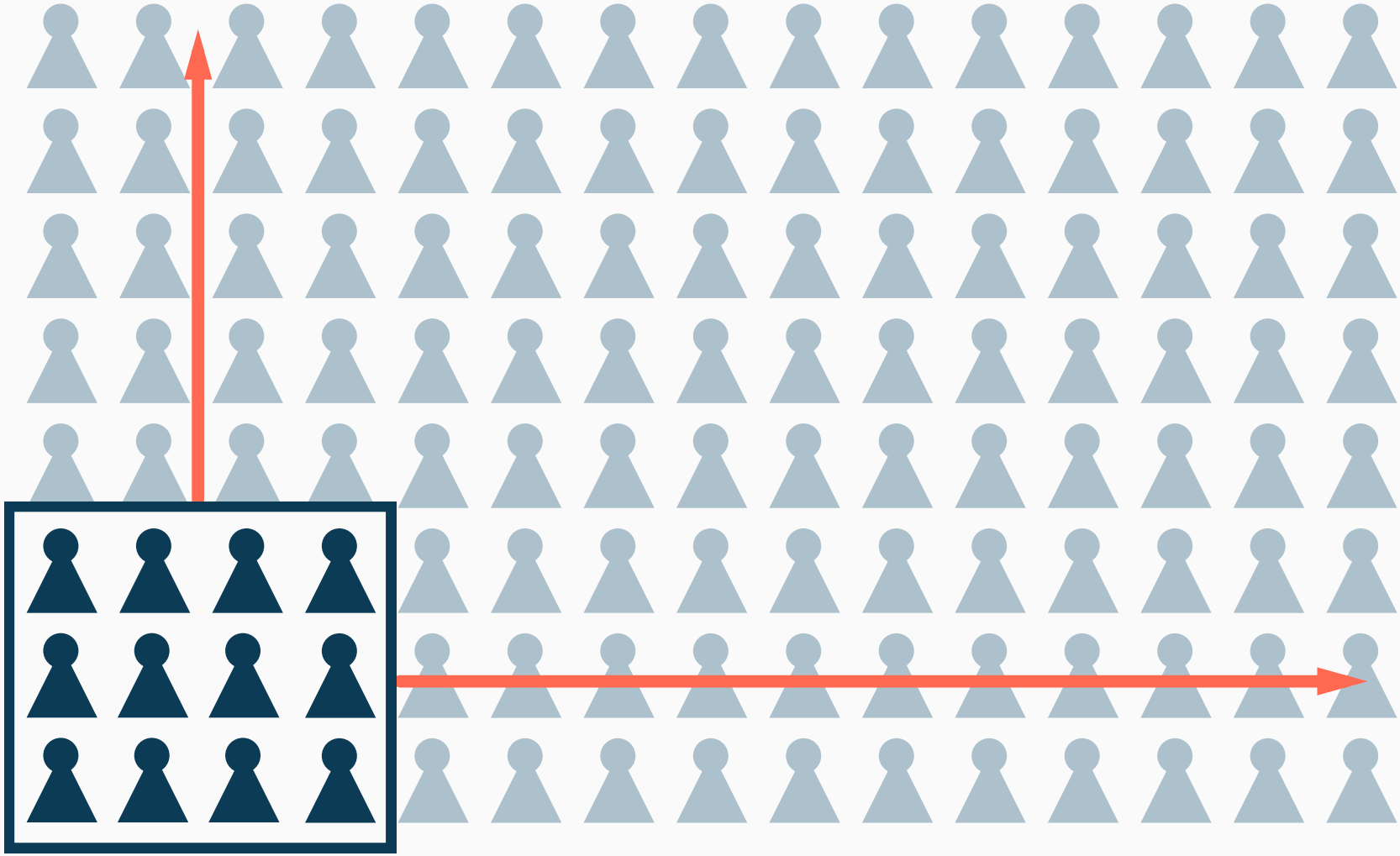
OVERVIEW OF THE MODEL



PROBLEM: DATA ONLY FOR A SAMPLE



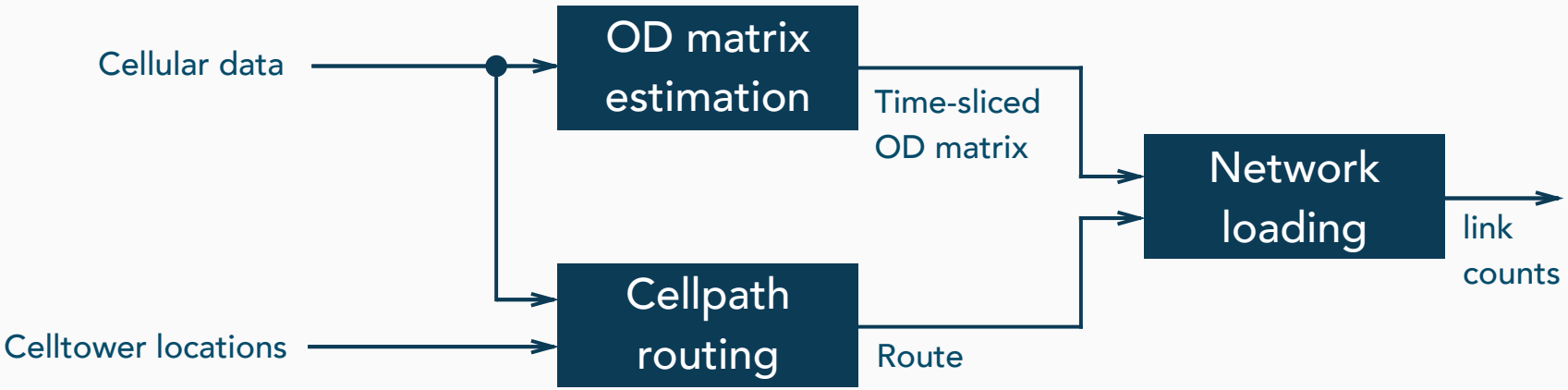
PROBLEM: DATA ONLY FOR A SAMPLE



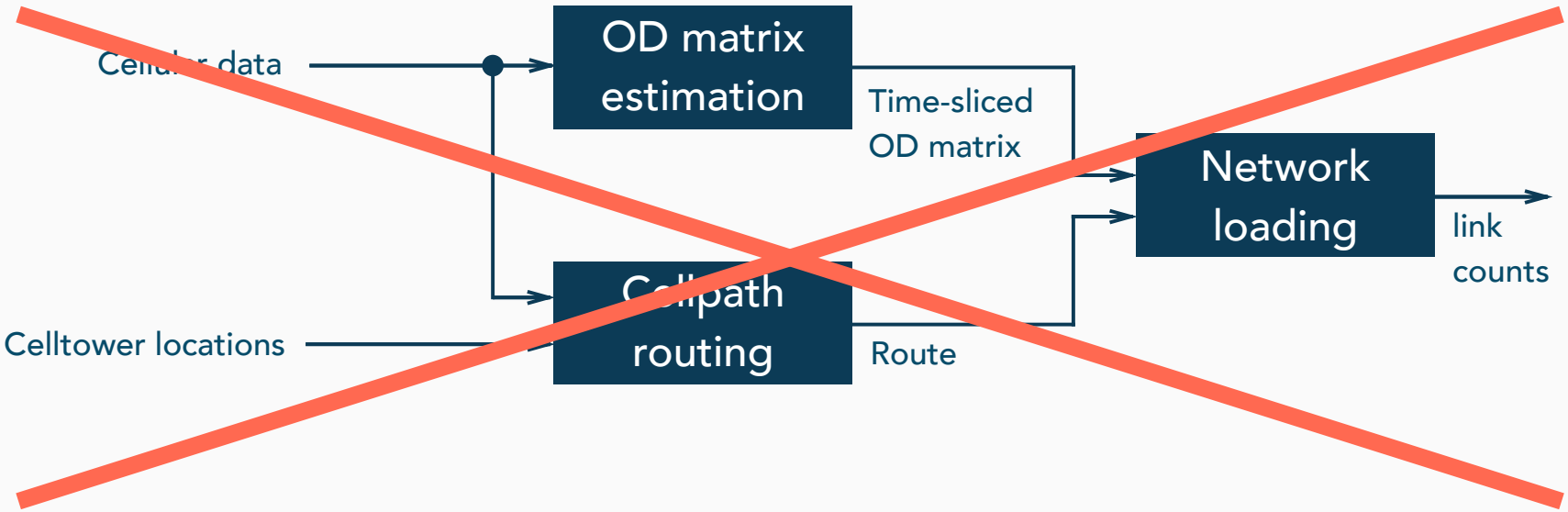
OD-matrix

A matrix containing the traffic demand between pairs of origin and destination zones

OVERVIEW OF THE MODEL



OVERVIEW OF THE MODEL



PROBLEM: DATA CONTAINS NOT ONLY MOVEMENTS

User ID	*Timestamp*	*Cell ID*
1	2013-10-01 06:50:00	1
1	2013-10-01 08:10:00	1
1	2013-10-01 08:10:00	3
2	2013-10-01 08:20:00	2

Table 1: An example of a cellular network dataset

PROBLEM: DATA CONTAINS NOT ONLY MOVEMENTS

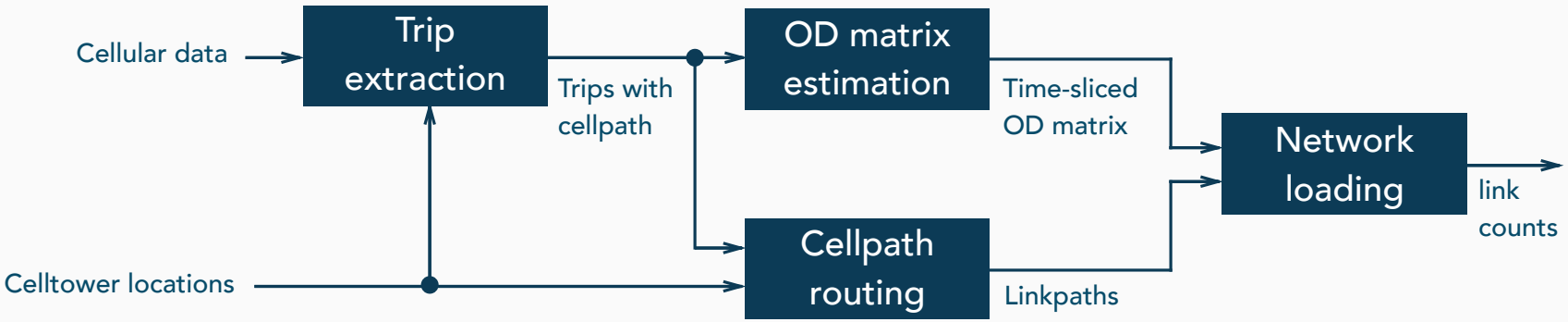
User ID	*Timestamp*	*Cell ID*
1	2013-10-01 06:50:00	1
1	2013-10-01 08:10:00	1
1	2013-10-01 08:10:00	3
2	2013-10-01 08:20:00	2

Table 1: An example of a cellular network dataset

Datatypes

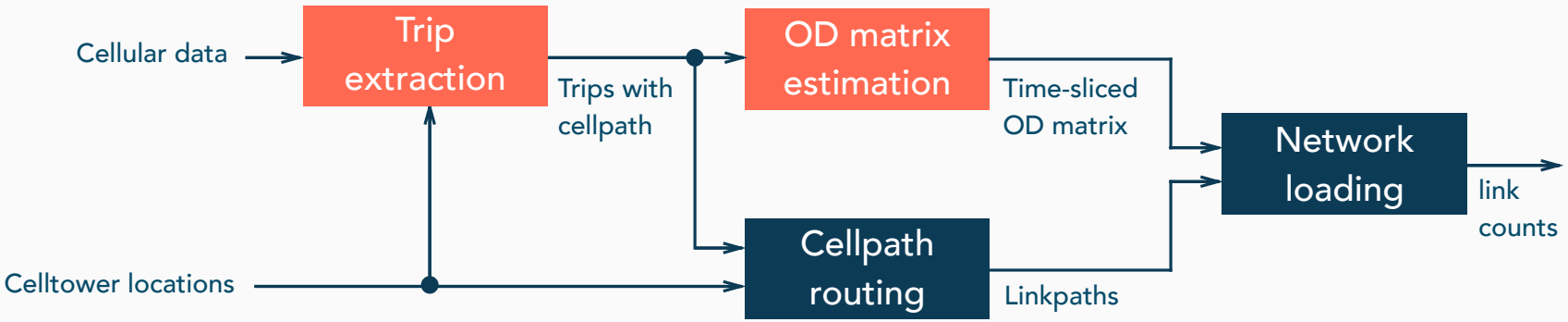
- Call detail records (CDR)
- Handover data

OVERVIEW OF THE MODEL



Trip extraction & travel demand estimation

OVERVIEW OF THE MODEL



Challenges

- Depends on the datasource
- Cell-switching due to network balancing
- Low resolution in time (CDR)

TRIP EXTRACTION

Challenges

- Depends on the datasource
- Cell-switching due to network balancing
- Low resolution in time (CDR)

Solutions

- algorithmic approach (CDR)
- movement efficiency metric (handover data)

Challenges

- Additional data necessary
- Sample might not be representative
- Vehicles $\neq$ people

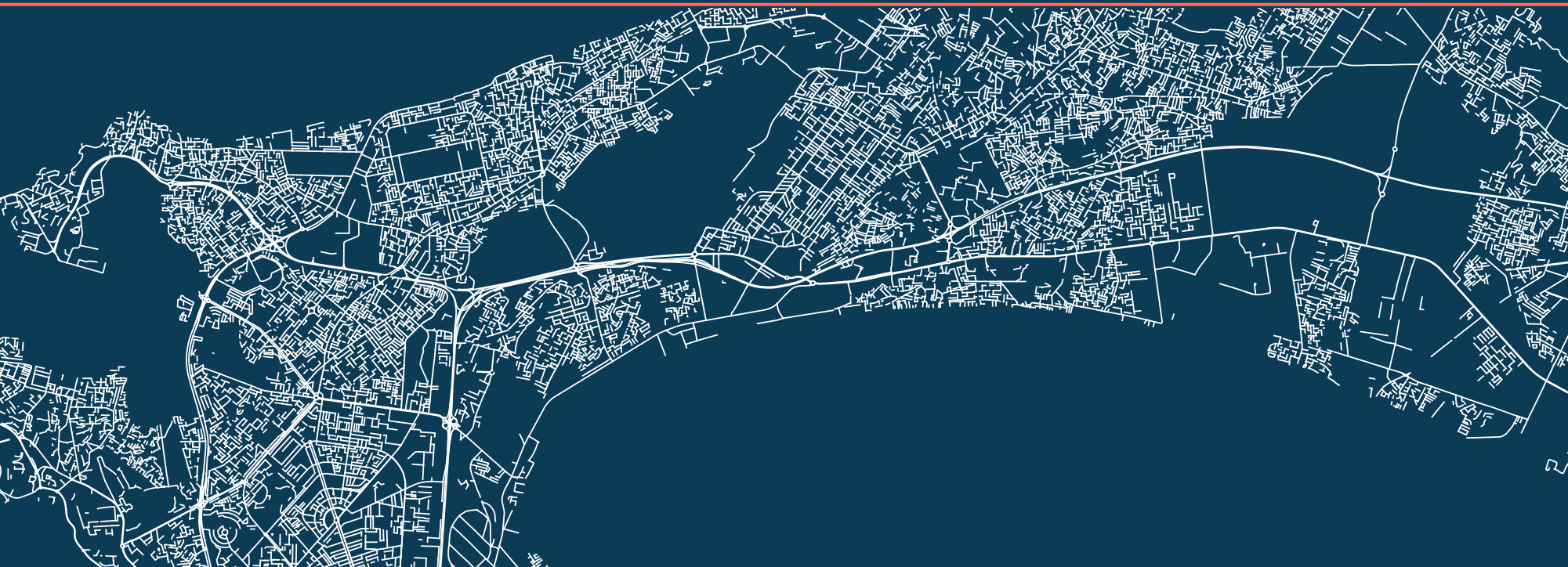
Challenges

- Additional data necessary
- Sample might not be representative
- Vehicles $\neq$ people

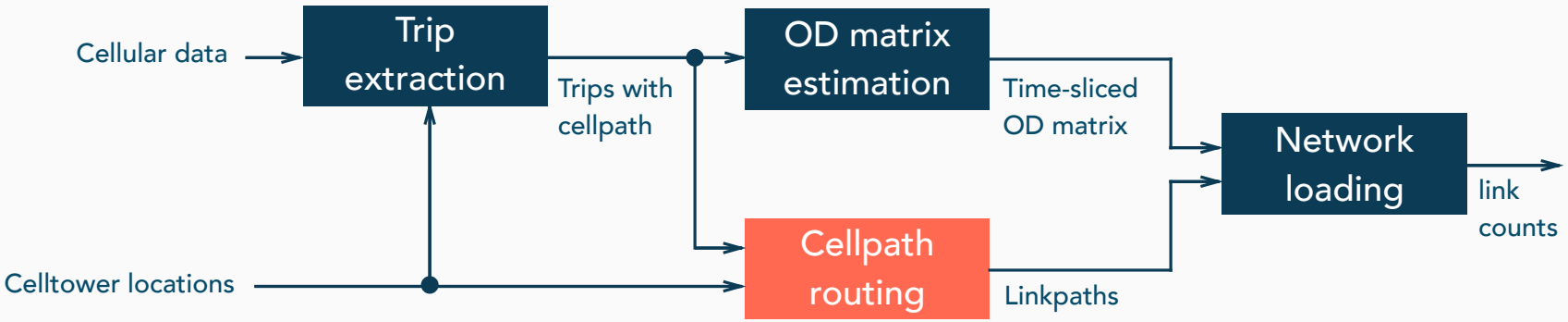
Solutions

- Upscaling using total population
- Data fusion with census data

Cellpath routing



OVERVIEW OF THE MODEL



CELLPATH ROUTING PROBLEM

Input

Cellpath $p = (b_1, b_2, ..., b_n)$

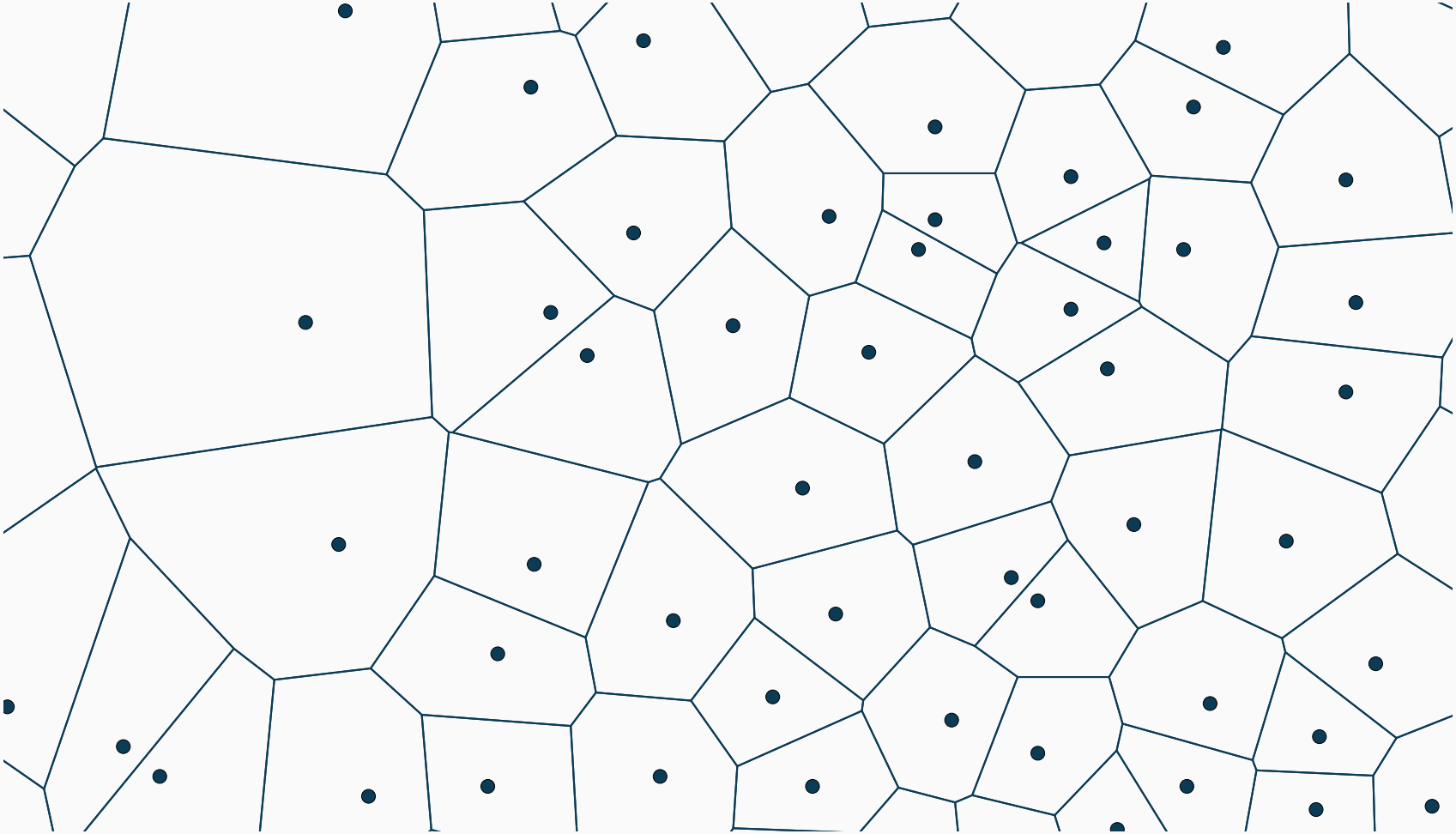
Output

A route consisting of connected links on the road network

Goal

Recover the original route that the user took

VORONOI DIAGRAM



CELLPATH ROUTING ALGORITHMS

Shortest-path Routing

shortest path between first and last cell in cellpath

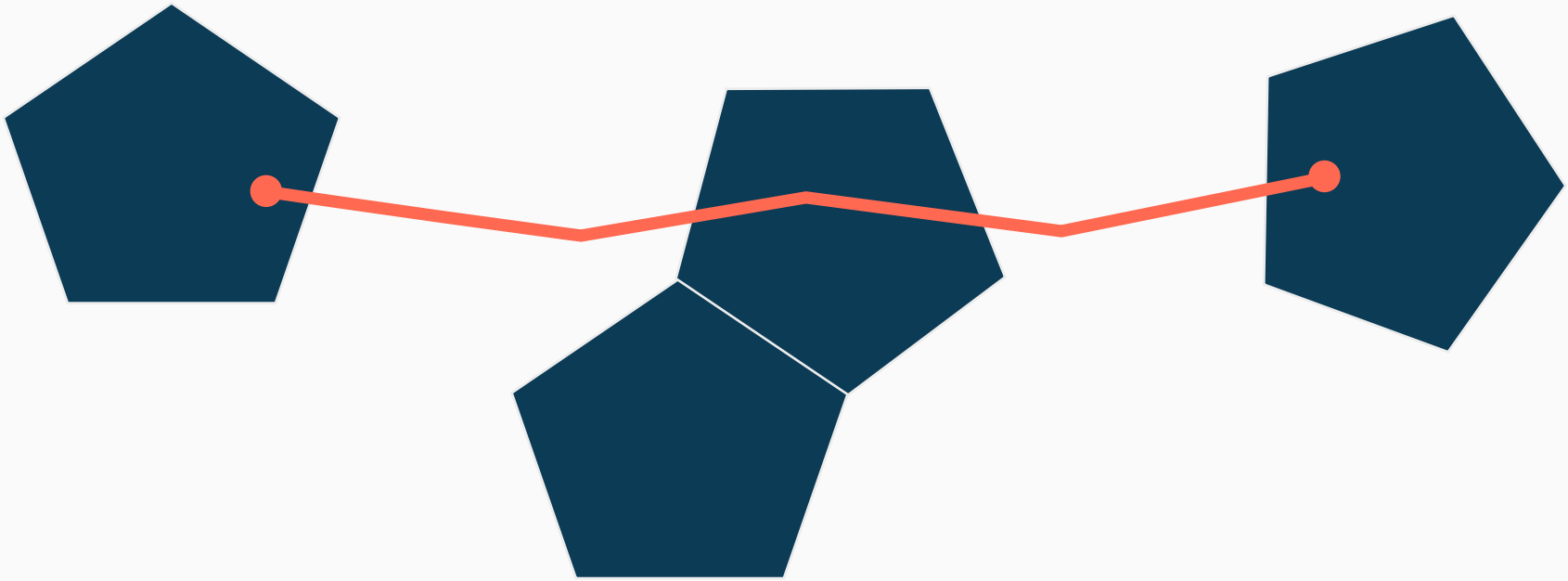
Strict Voronoi Routing

route must go through every cell in the cellpath

Lazy Voronoi Routing

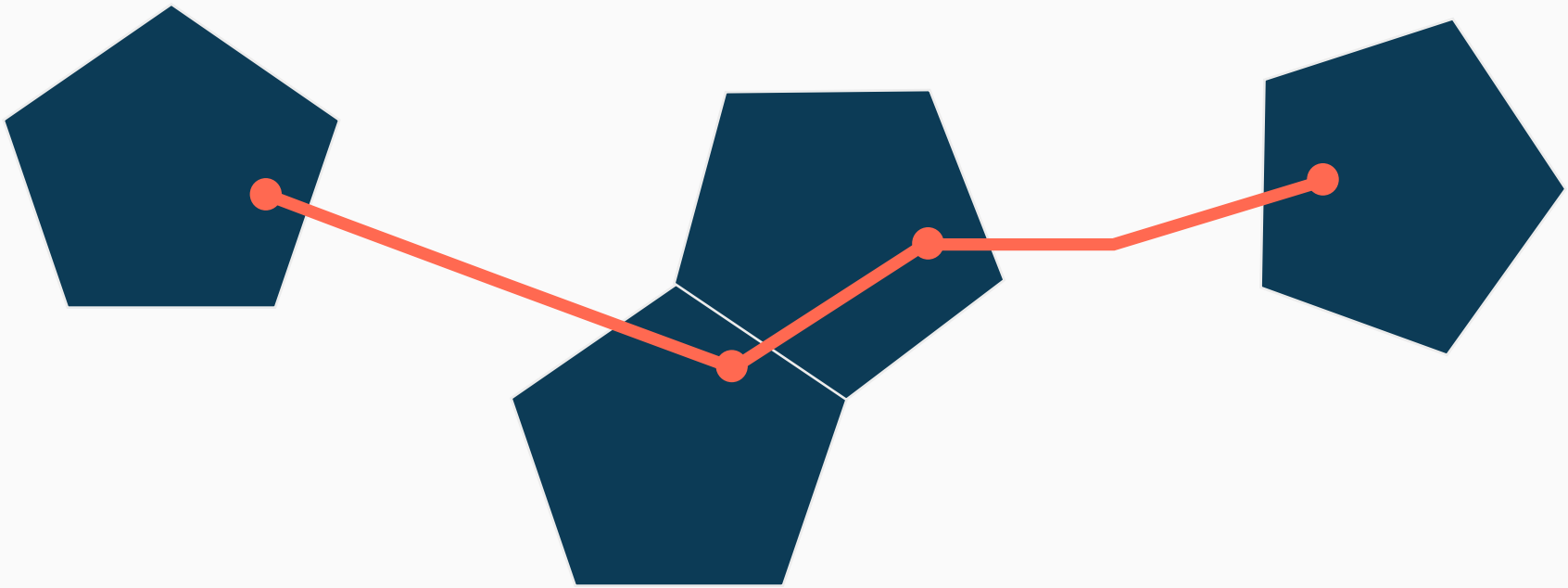
route is encouraged to go through cells in the cellpath

SHORTEST-PATH ROUTING



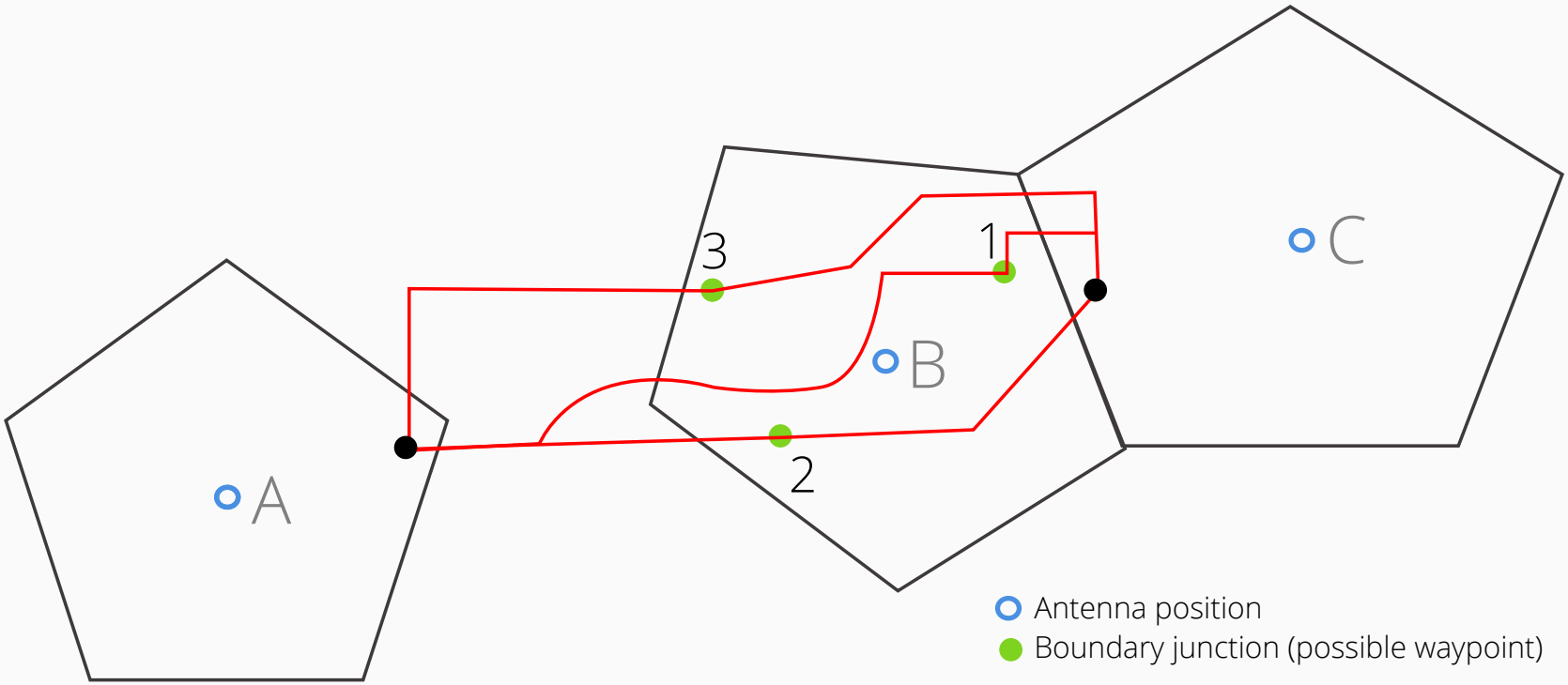
shortest path between first and last cell in cellpath

STRICT VORONOI ROUTING

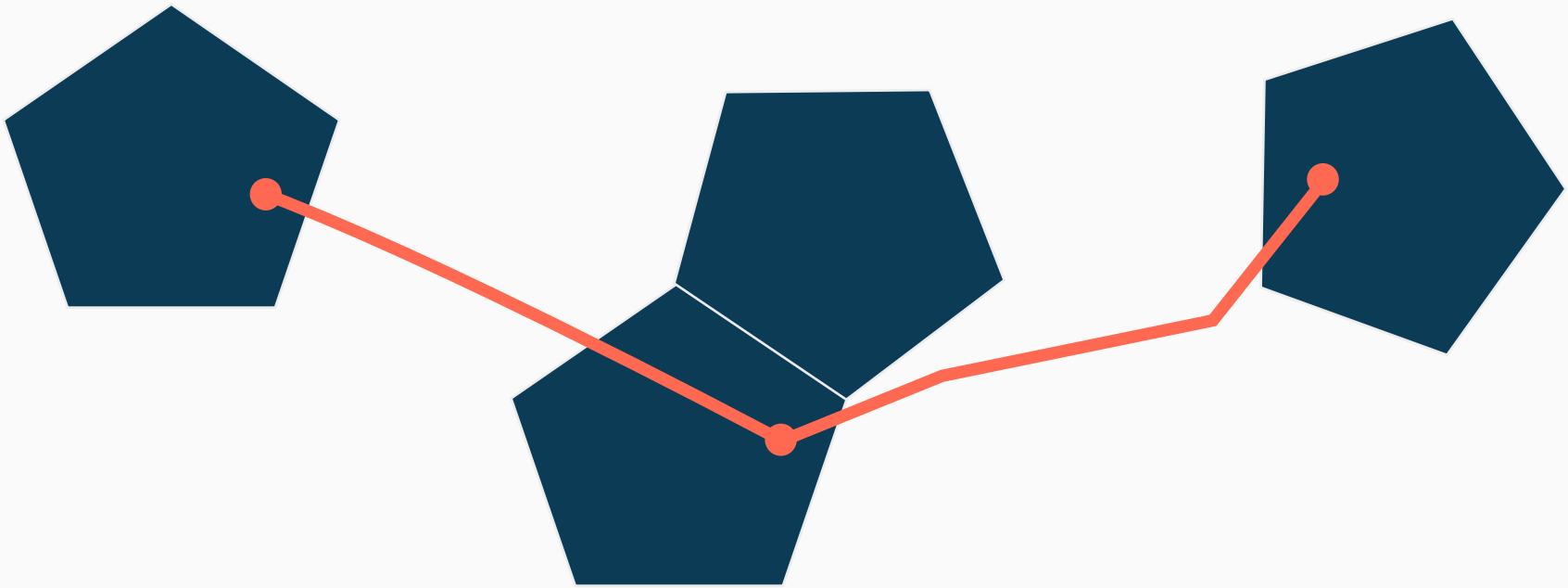


route must go through every cell in the cellpath

WAYPOINT SEARCH

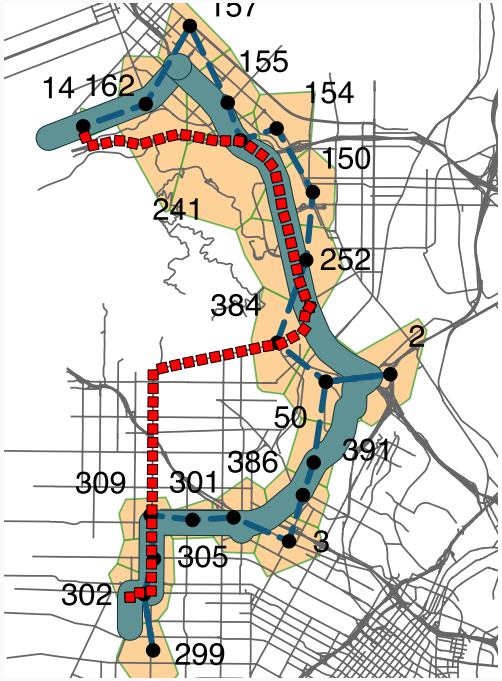


LAZY VORONOI ROUTING

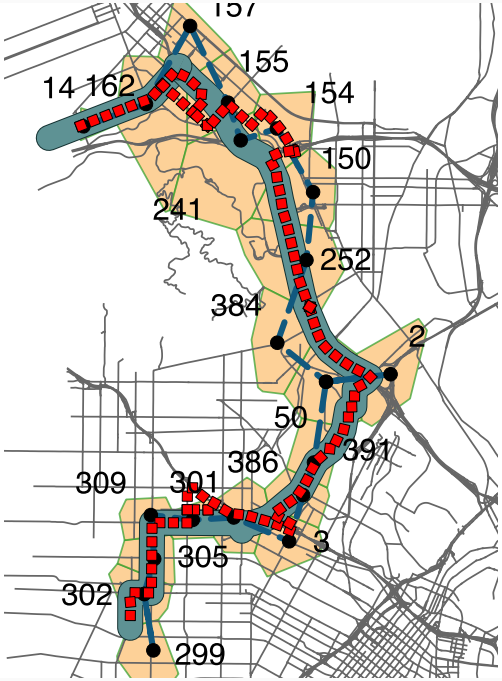


route is encouraged to go through cells in the cellpath

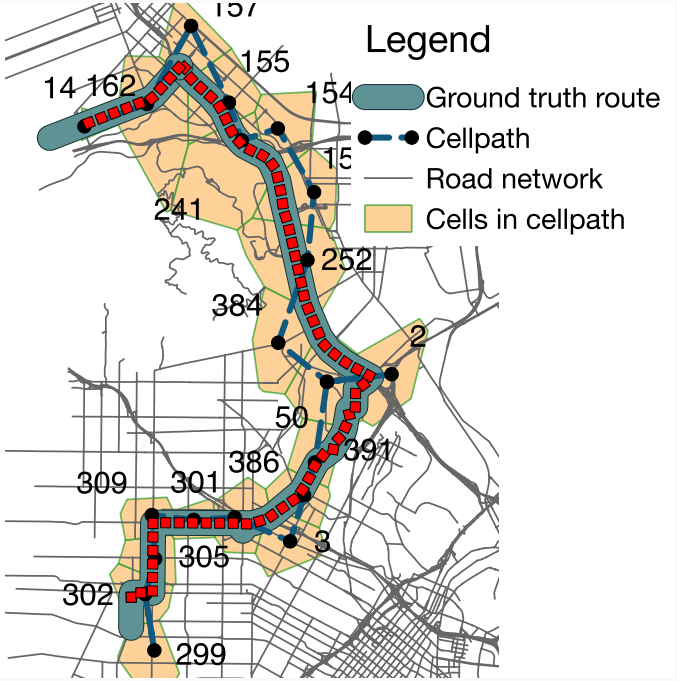
COMPARISION BETWEEN ALGORITHMS



Shortest-path

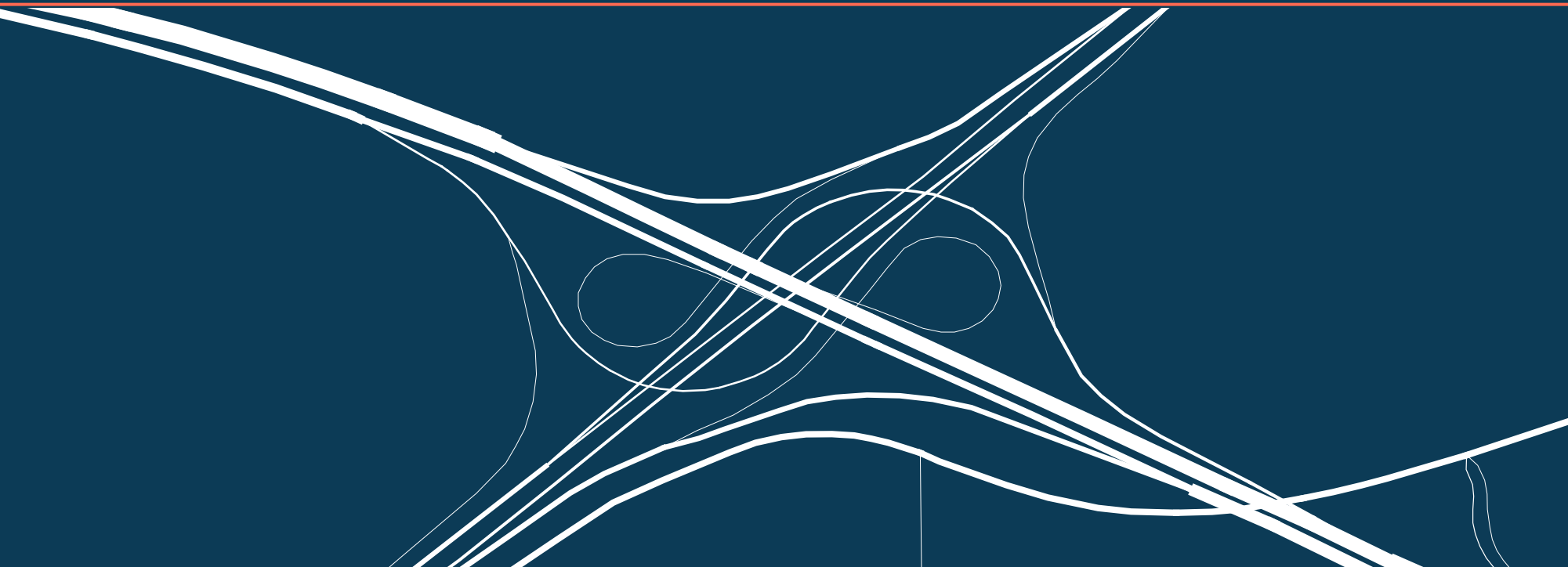


Strict Voronoi

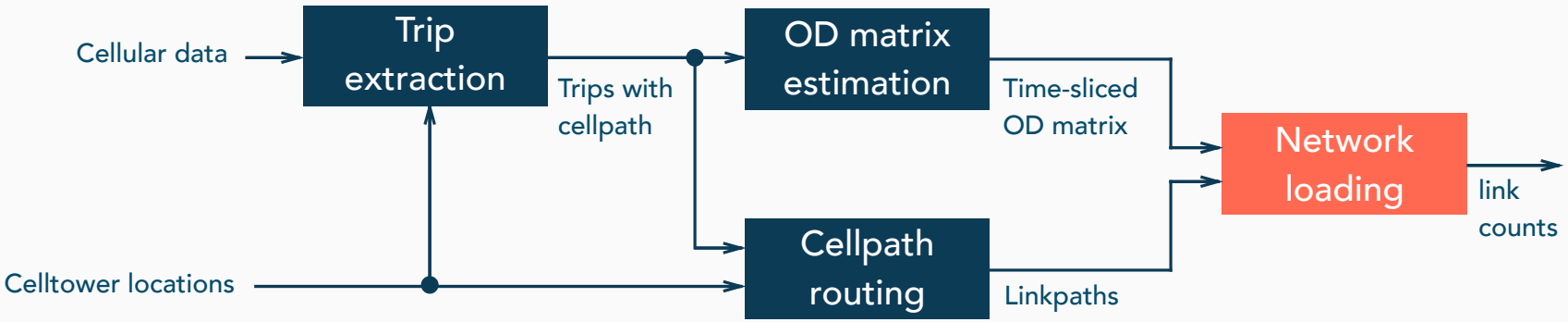


Lazy Voronoi

Network loading



OVERVIEW OF THE MODEL



SINGLE OD-PAIR

Shortest-path

SINGLE OD-PAIR

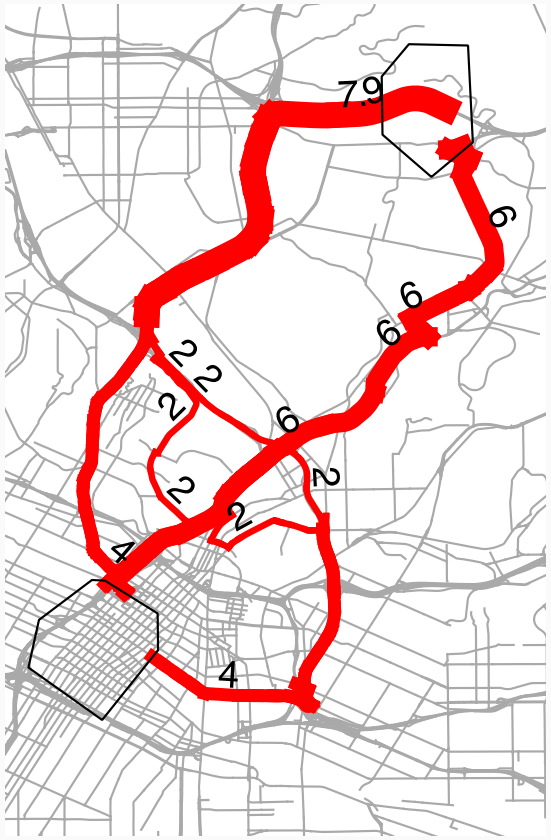
Shortest-path

Strict Voronoi

SINGLE OD-PAIR

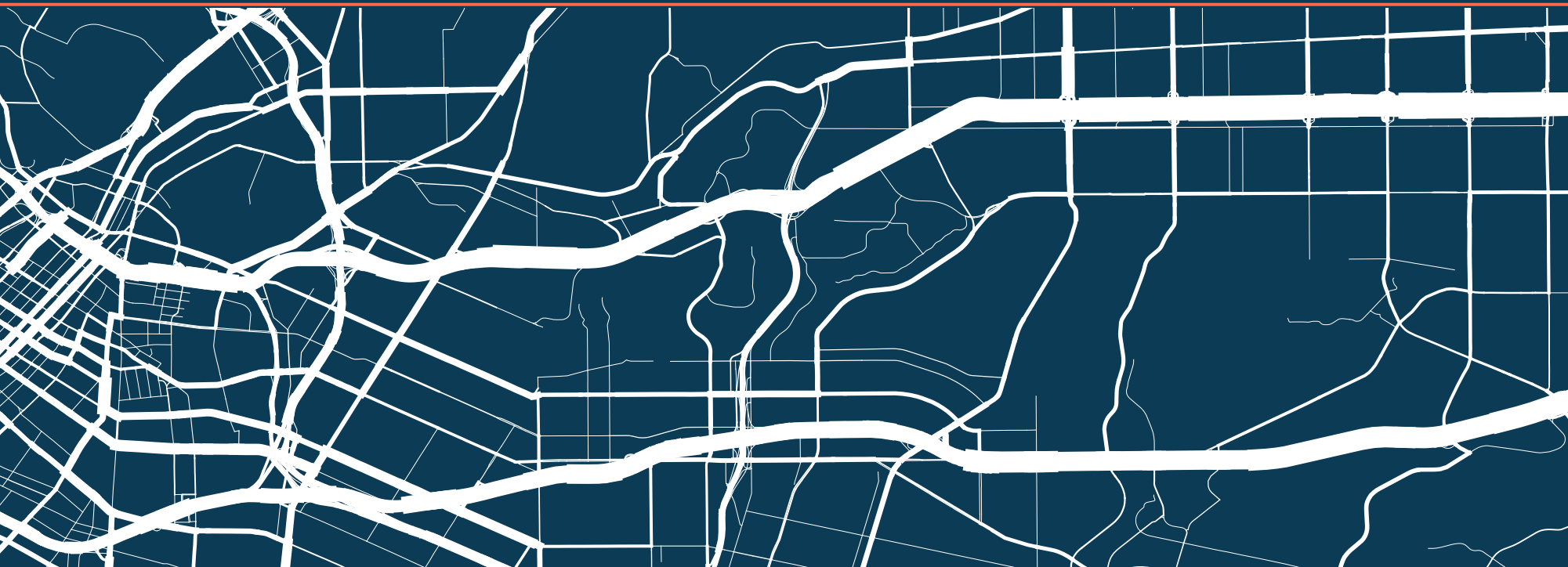
Shortest-path

Strict Voronoi



Lazy Voronoi

Results & validation



LOS ANGELES DATASET

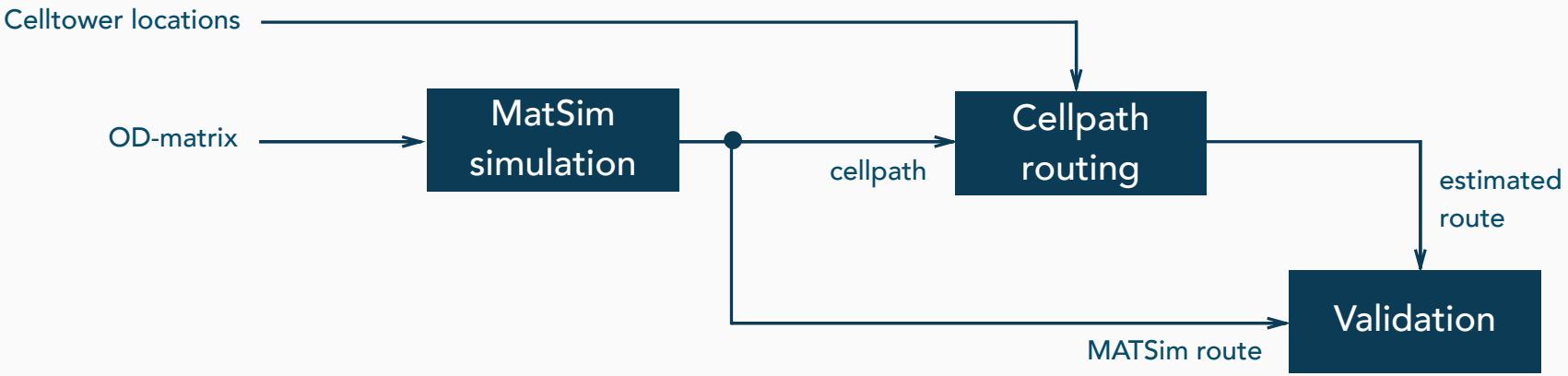
Coverage

I-210 corridor in Los Angeles, USA

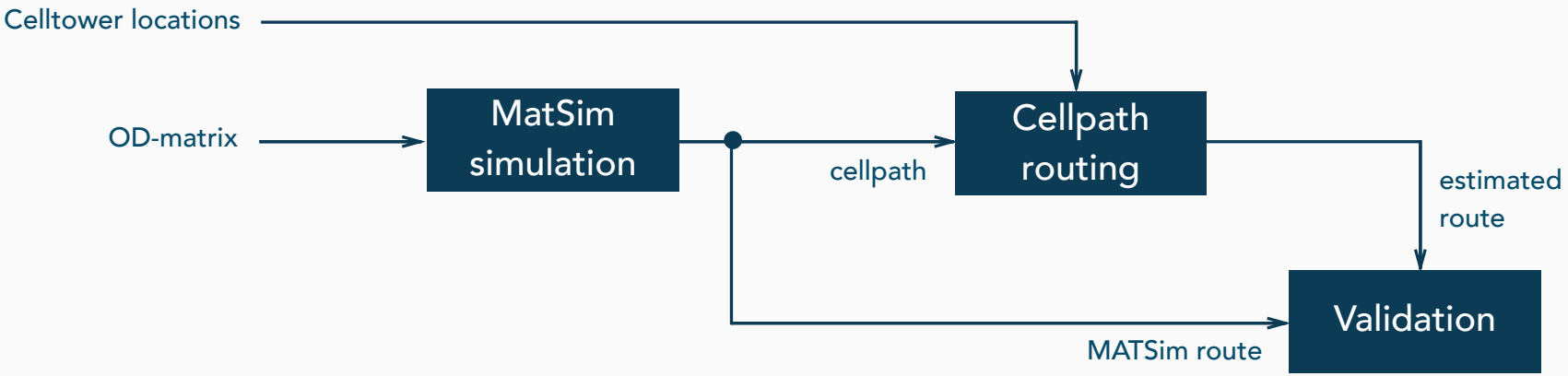
Input data

- Antenna positions
- OD-matrix estimated from real cellular data (UC Berkeley)
- Simulated cellpaths (UC Berkeley)
- OpenStreetMap road network

ROUTE VALIDATION



ROUTE VALIDATION

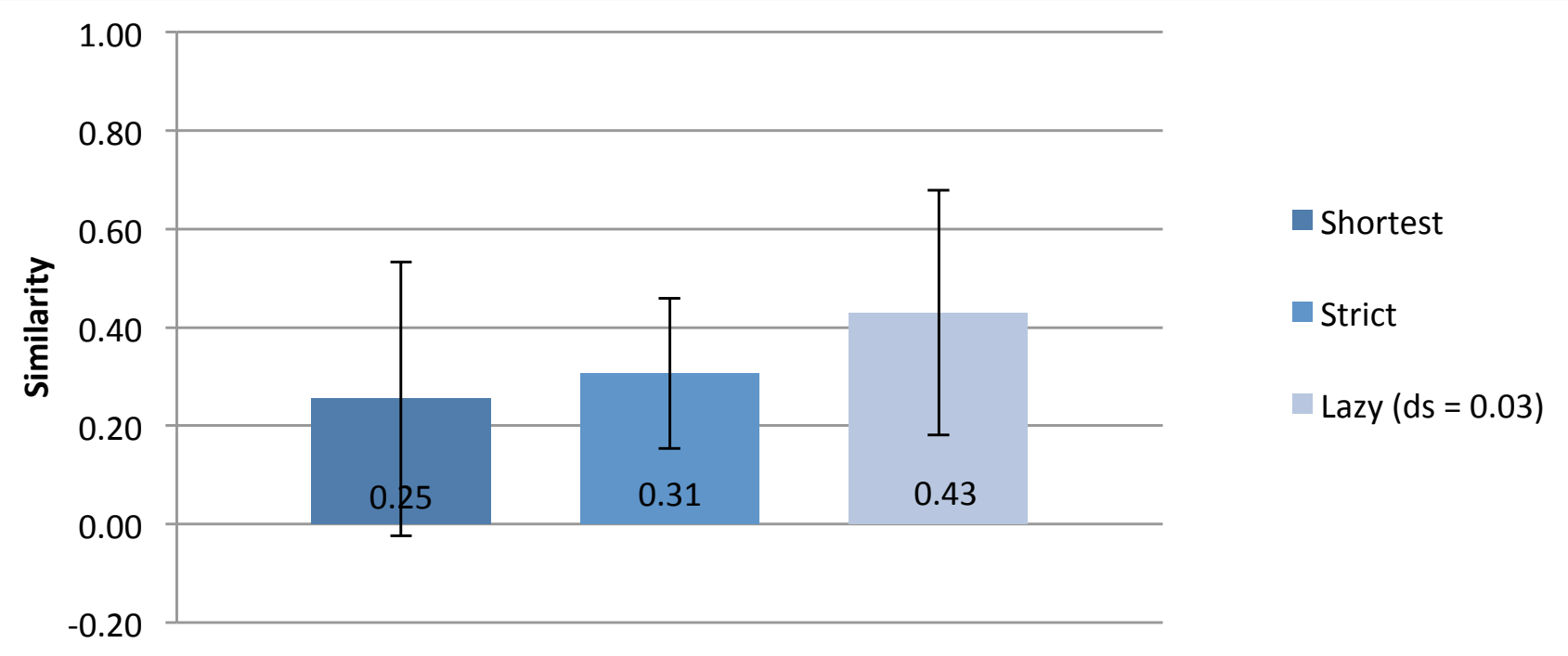


Definition: Route similarity metric

Two routes $r_a = (V_a, E_a)$ and $r_b = (V_b, E_b)$ have a *route similarity* of:

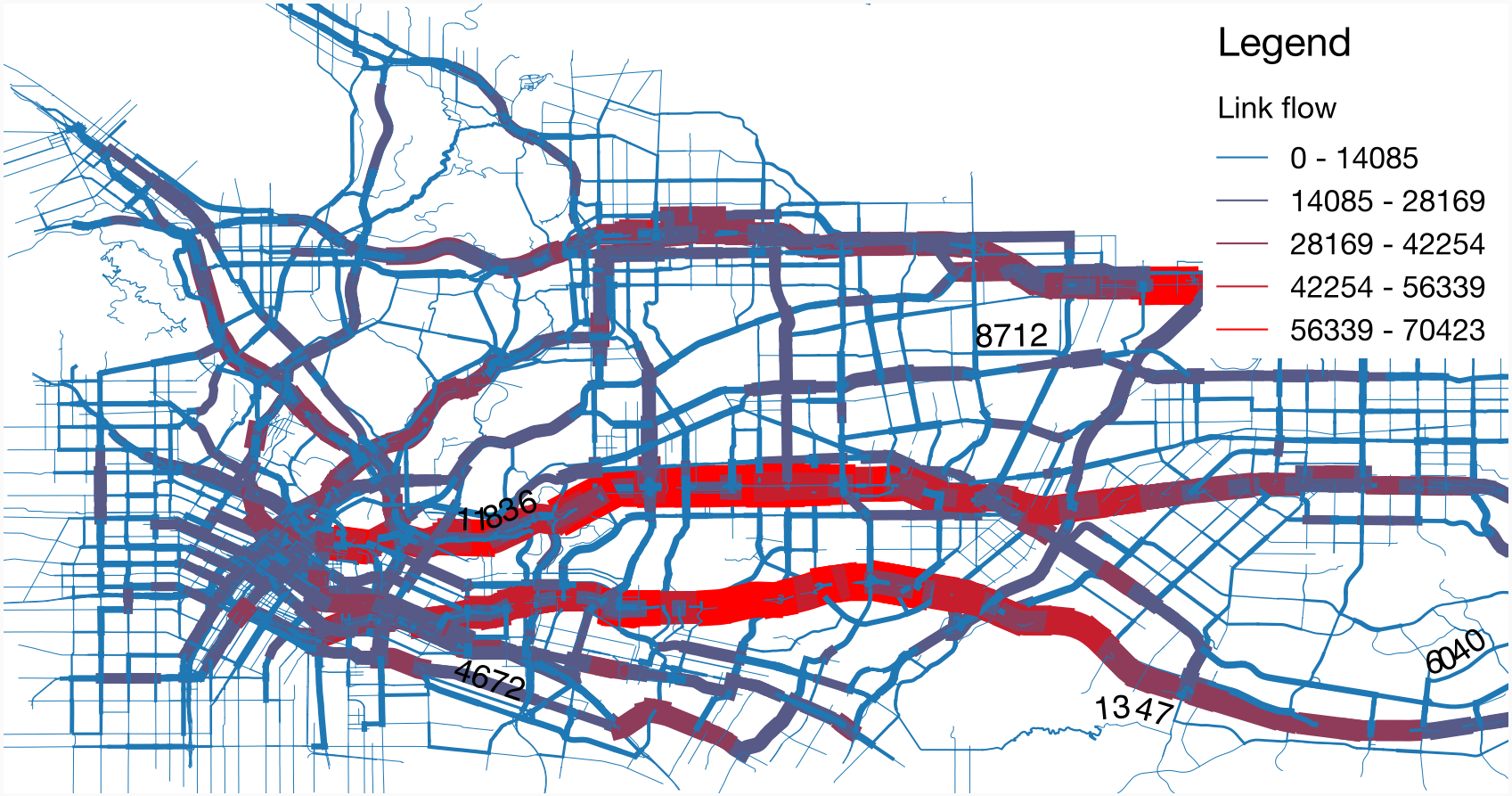
$$s(a, b) := \frac{|V_a \cap V_b|}{|V_a \cup V_b|}$$

ROUTE VALIDATION



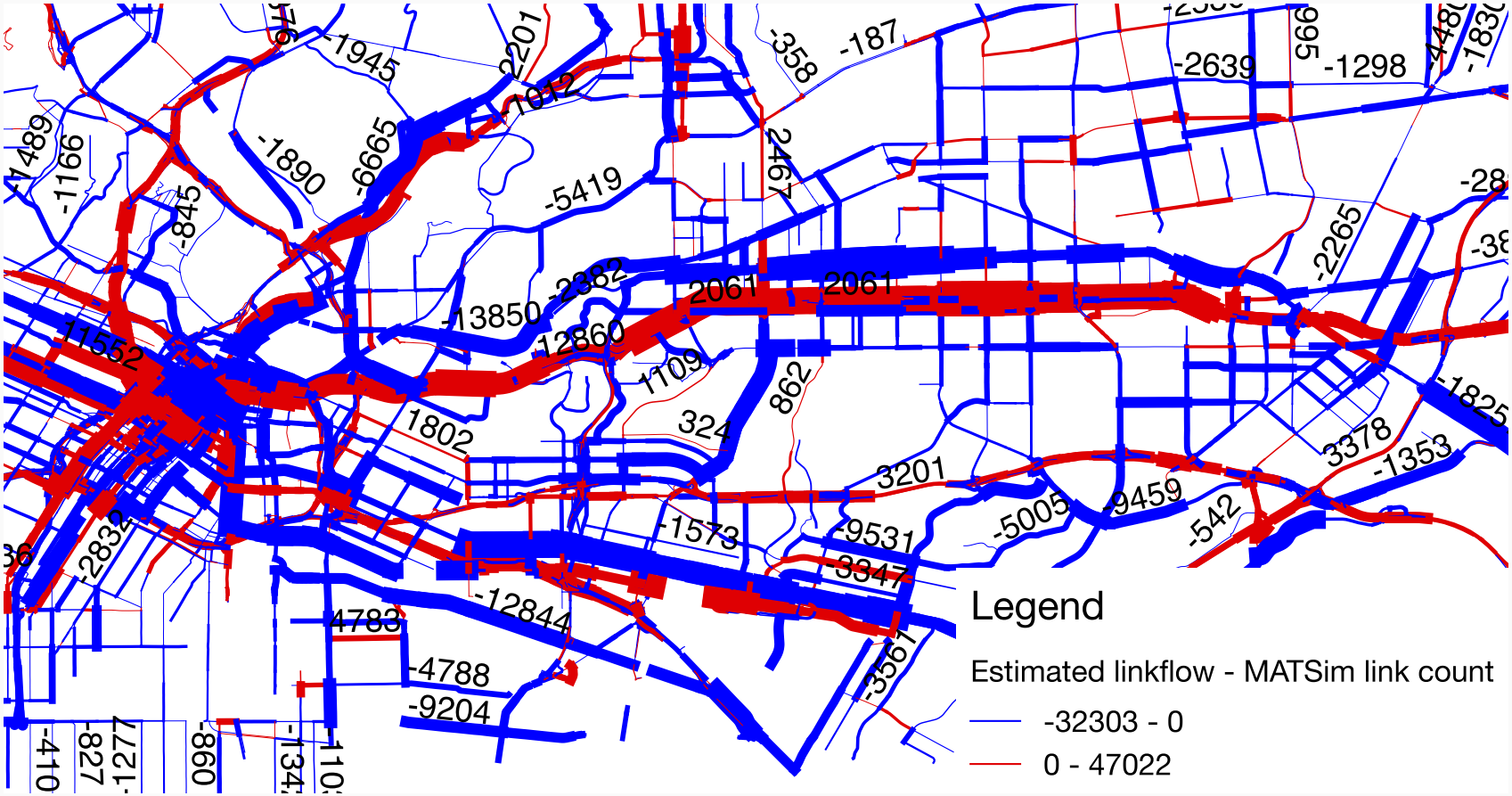
average of 1000 randomly selected routes

NETWORK LOADING RESULT

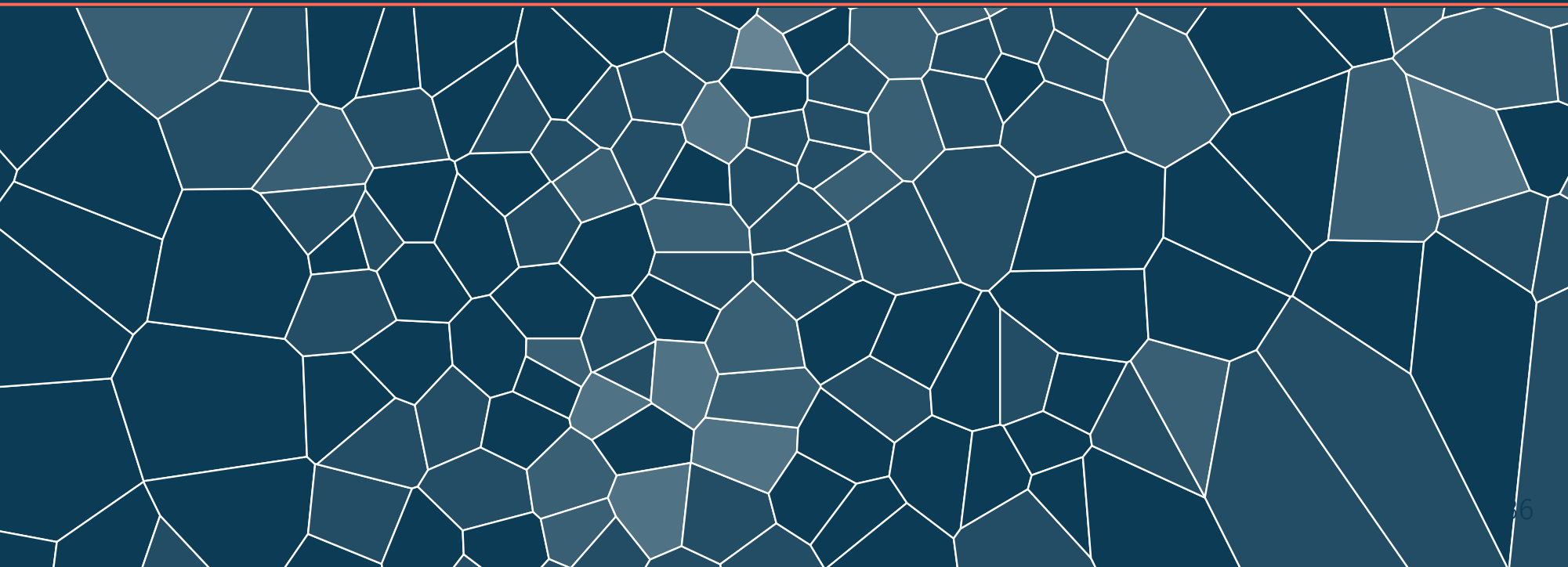


using Lazy Voronoi routing

NETWORK LOADING VALIDATION



Conclusions



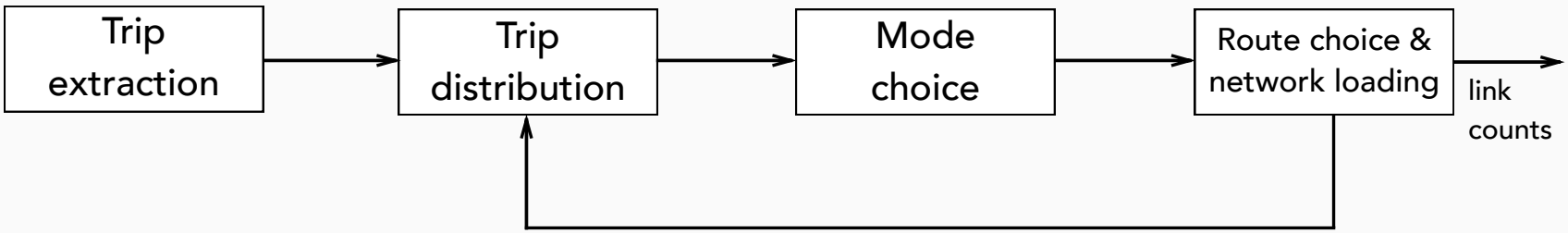
CONCLUSIONS

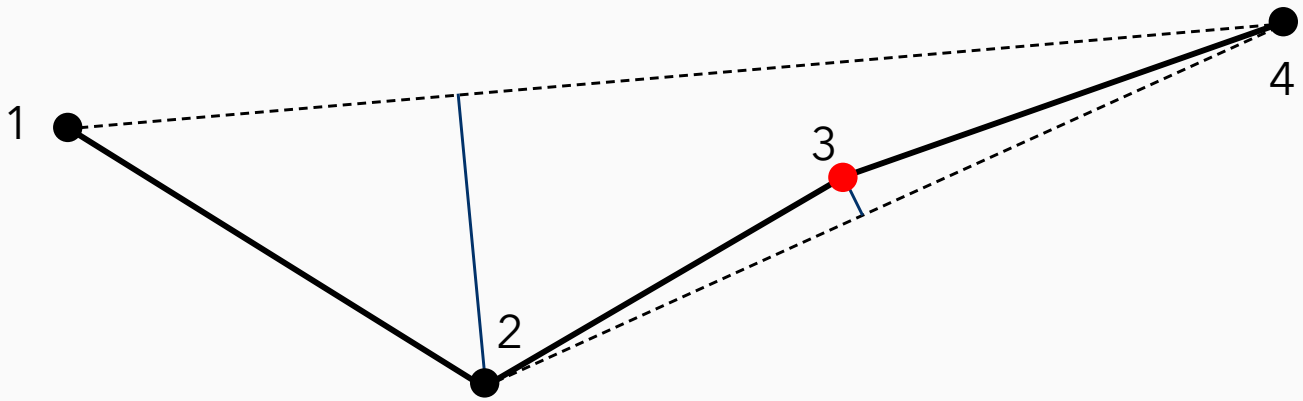
1. Lazy Voronoi improves route recoverage over Shortest-path and Strict Voronoi routing
2. Route similarity not enough for a precise network loading
3. Data fusion with traffic counts could improve the result

Nils Breyer

nilsbreyer.eu

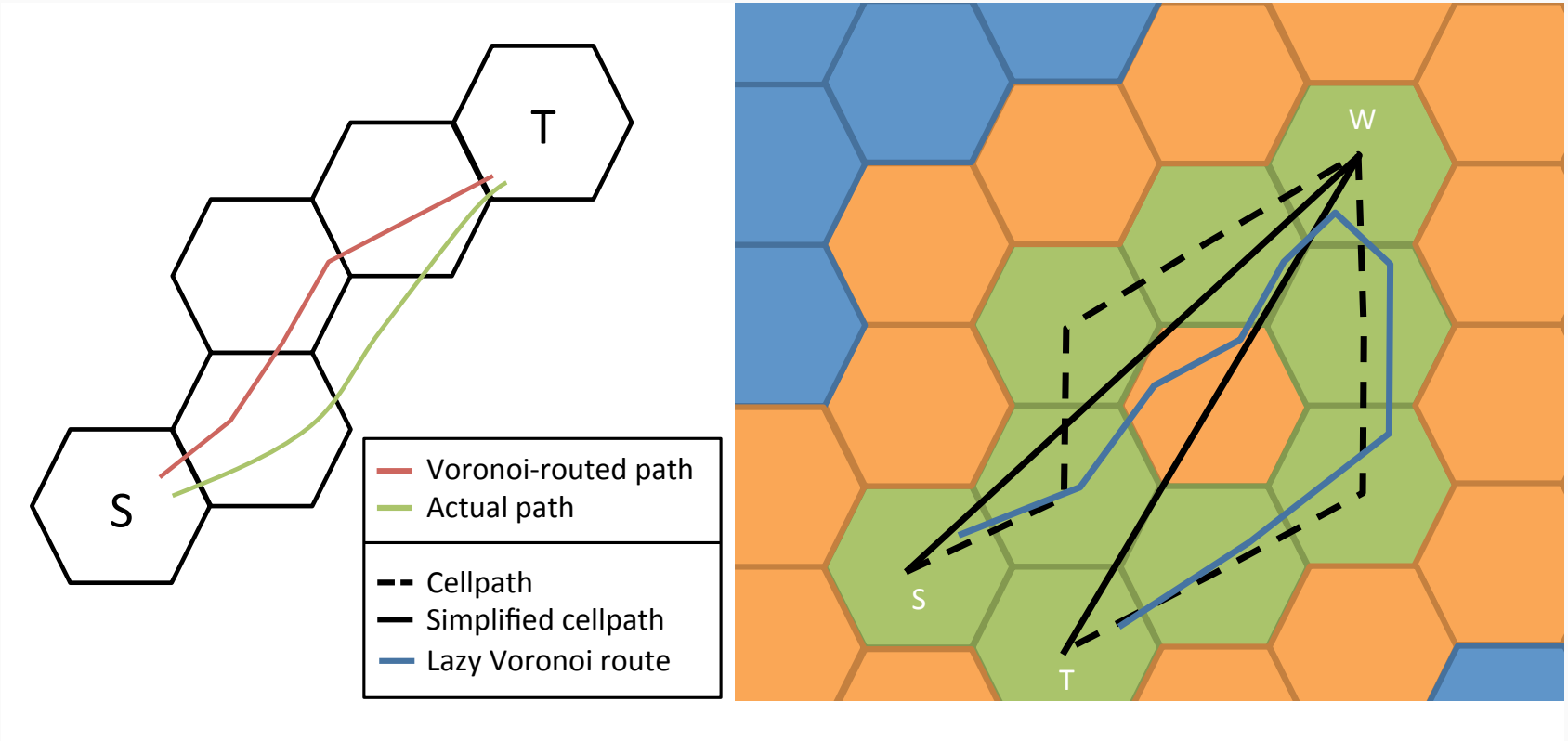
4-stage model





Ramer-Douglas Peuker algorithm

LAZY VORONOI ROUTING DETAILS



LA NETWORK LOADING GEH