Geometric Embedding of Road distance as a metric for Geospatial Analysis

Shashwathy Kannan - CS19B041

May 17, 2023

Outline

- → Motivation
- → Objectives
- → Input Resources
- → Embedding on Road Networks
- → Analysis
- → Applications
- → Conclusion

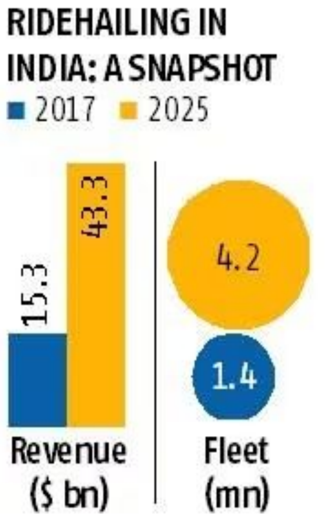
Motivation

Why?

VRP in logistics and ride-sharing is indispensable

MARKET POTENTIAL FROM KEY MOBILITY CLUSTER IN INDIA

	Value ($ billion) 2018	Value ($ billion) 2030
Traditional carsharing	0.069	0.592
P2P carsharing	0.007	0.298
Corporate sharing	0	0.172
Dynamic shuttle	0.132	23.626
MaaS	0	3.356
Taxi services	15.39	61.66
Total	**15.598**	**89.704**



Why?

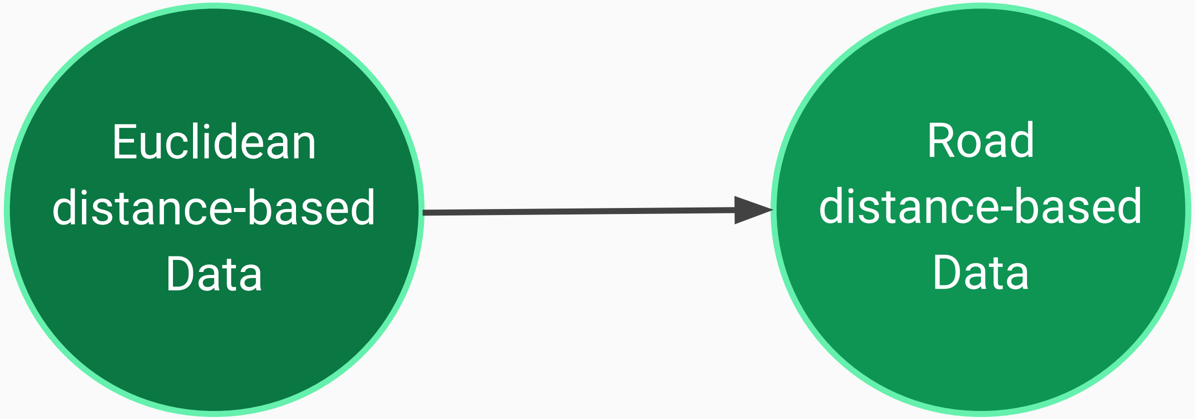
Lack of Real-world distance based dataset

Improve precision of Routing Algorithms

Objective

Aim

Methodical Conversion

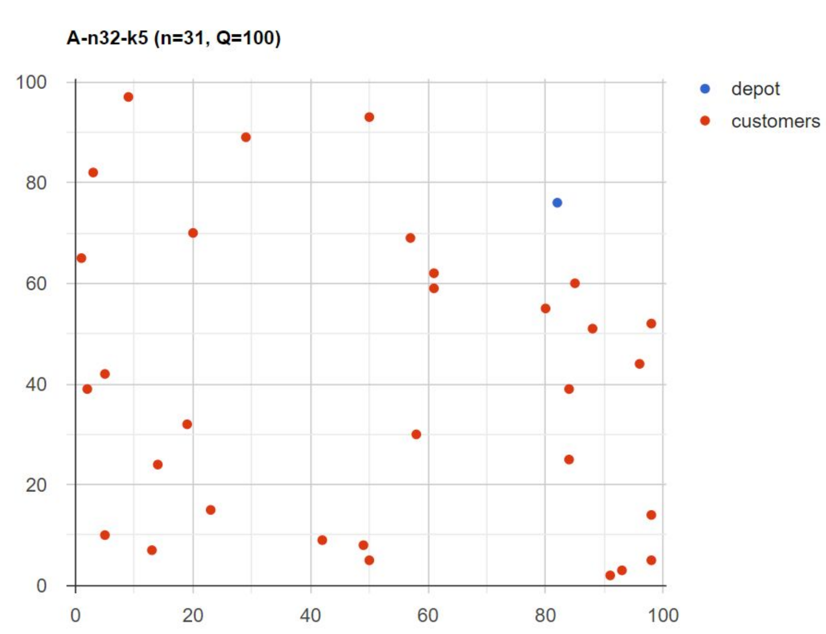


Input Resources

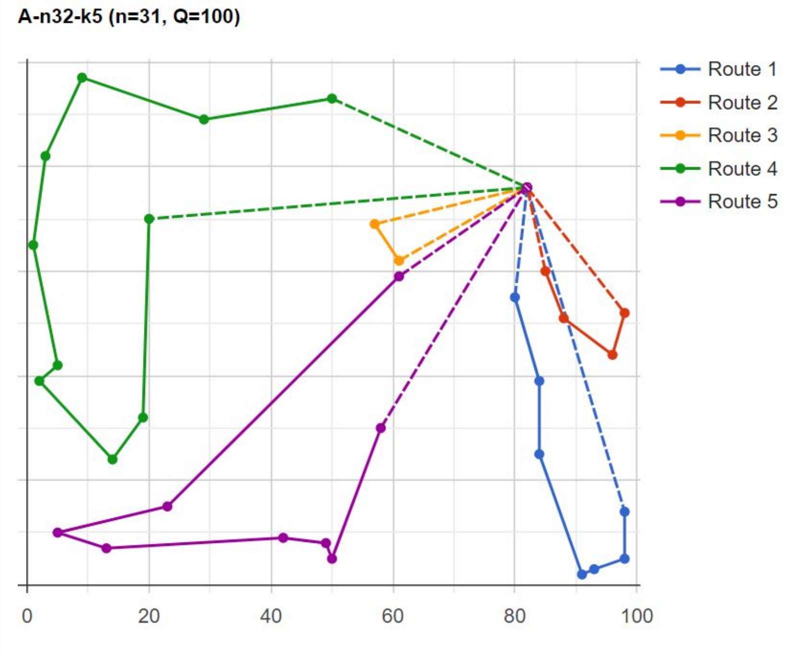
CVRP

- Definition and Constraints:
  - ***n*** customers and a single depot
  - Each customer has a specific **demand**
  - Each container has a maximum demand **capacity** it can accommodate
- Objective:
  - Reduce the total combined distance of routes

Sample Data



A-n32-k5 (n=31, Q=100)
100
80
60
40
20
0
0
20
40
60
80
100
depot
customers



A-n32-k5 (n=31, Q=100)
0
20
40
60
80
100
Route 1
Route 2
Route 3
Route 4
Route 5

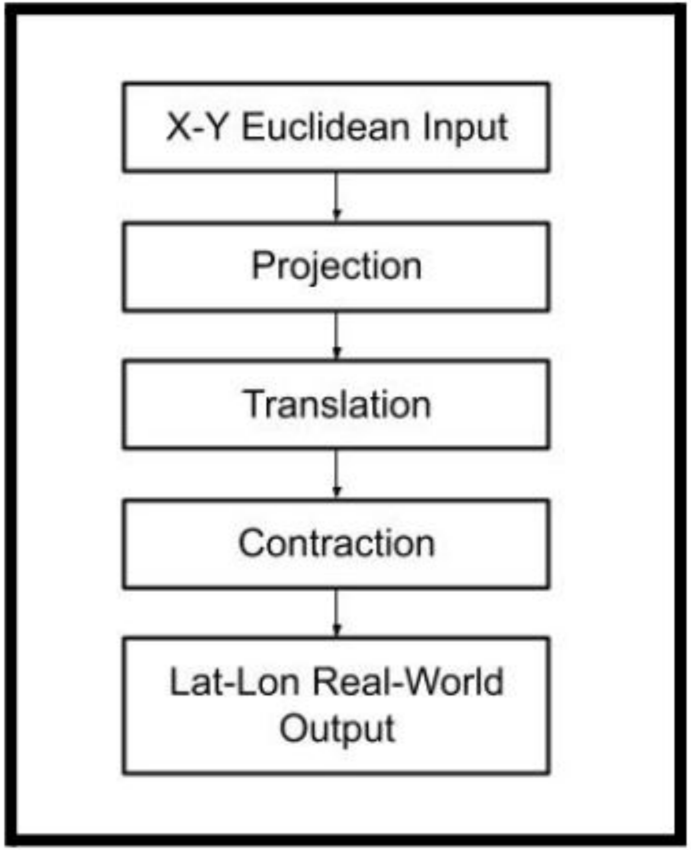
Benchmark Datasets

- Benchmark Dataset - Augerat (1995) was used from CVRPLIB, DIMACS, Rutgers University
- ***N*** ranges from 30 to 80
- 27 datasets in all

NAME : A-n6-k5
COMMENT : (Augerat et al, No of trucks: 5, Optimal value: /*cost*/)
TYPE : CVRP
DIMENSION : 6
EDGE_WEIGHT_TYPE : EUC_2D
CAPACITY : 50
NODE_COORD_SECTION
1 82 76
2 96 44
3 50 5
4 49 8
5 13 7
6 29 89
DEMAND_SECTION
1 0
2 19
3 21
4 6
5 19
6 7
DEPOT_SECTION
1
-1
EOF

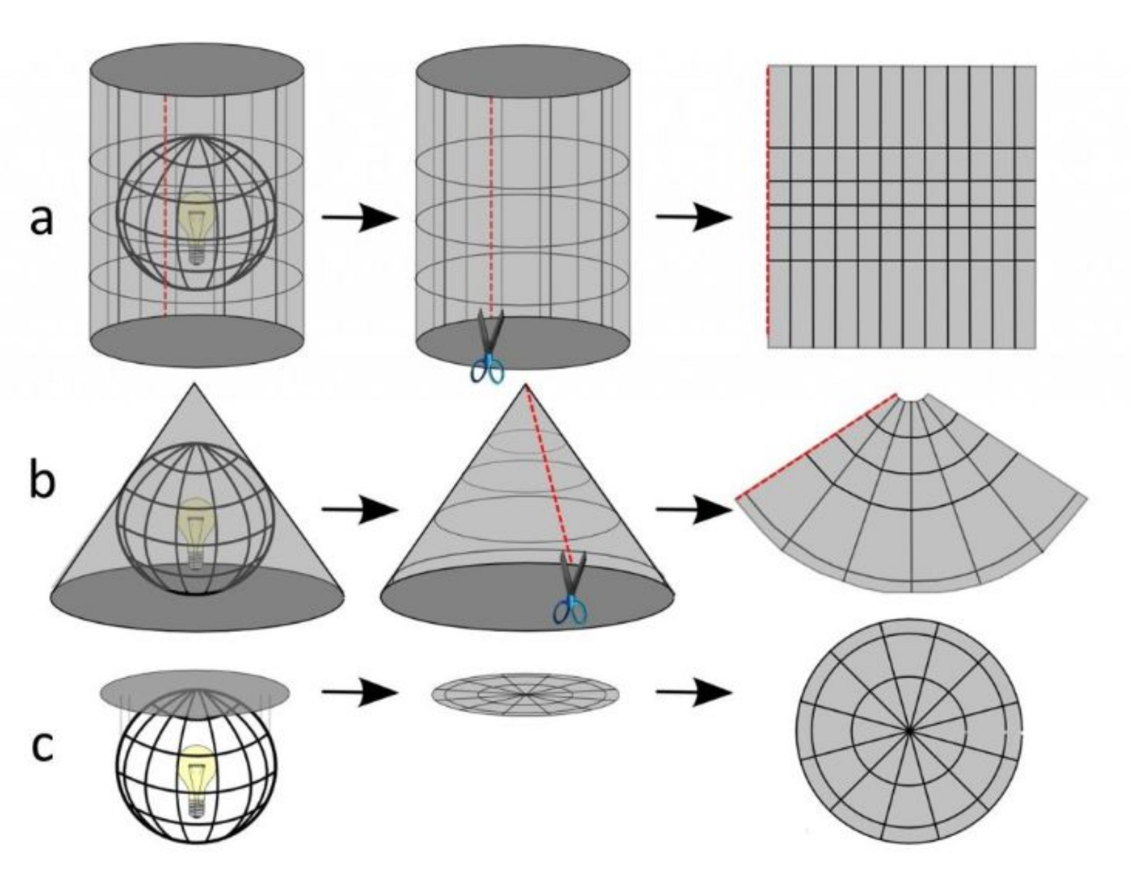
Embedding on Road Networks

Transformation Algorithm



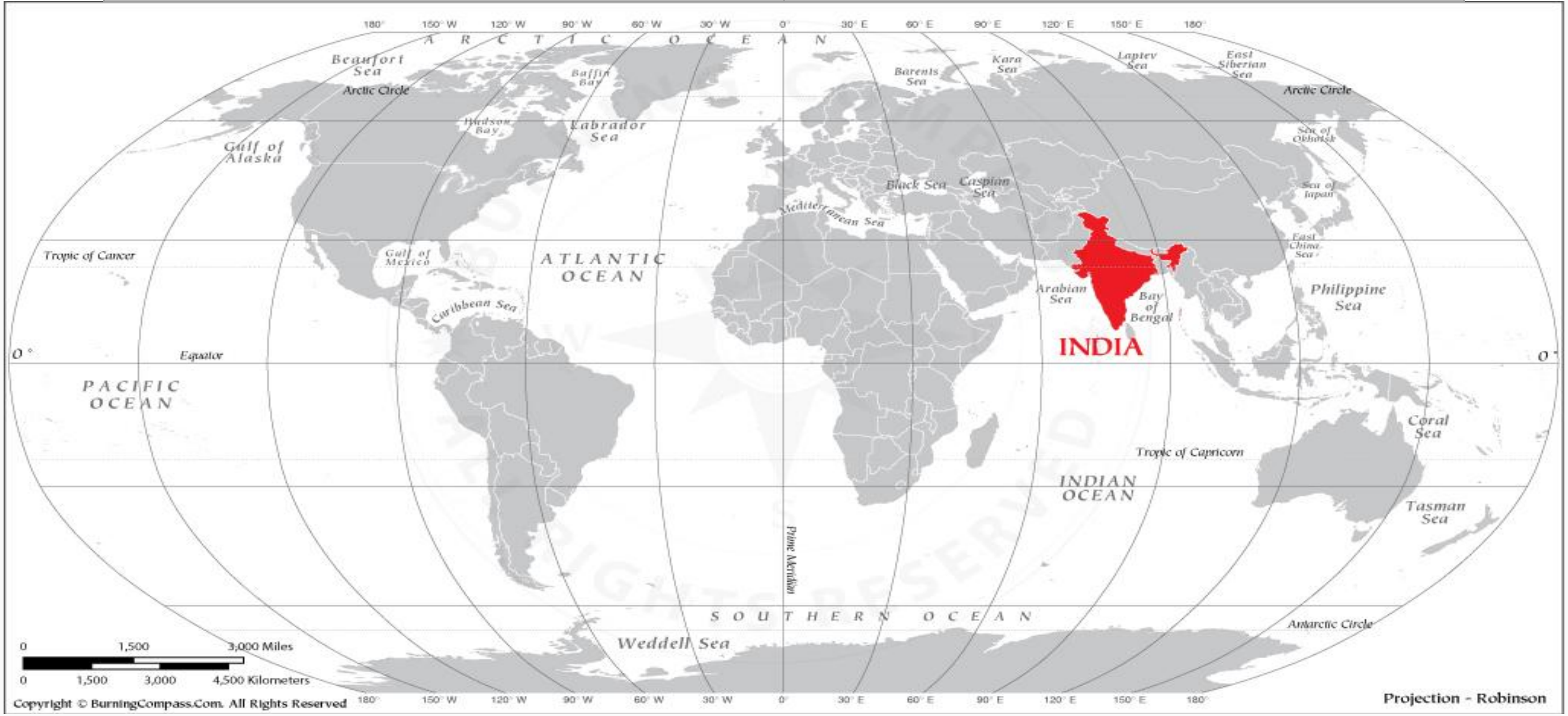
Projection

Developable Surfaces

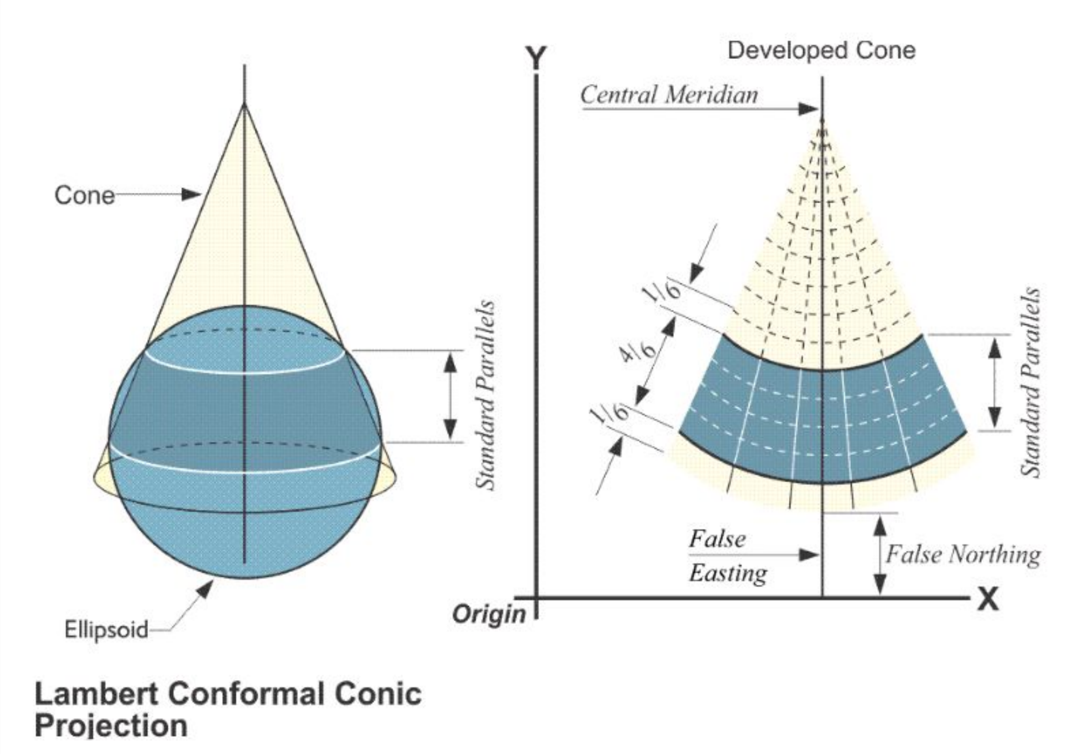


Suitable Map Projections

Lambert **Conical** Conformal Projection	Mercator **Cylindrical** Projection

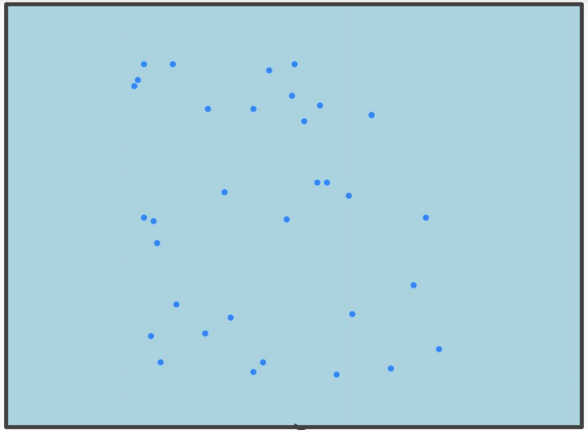


Lambert Conical Conformal Projection



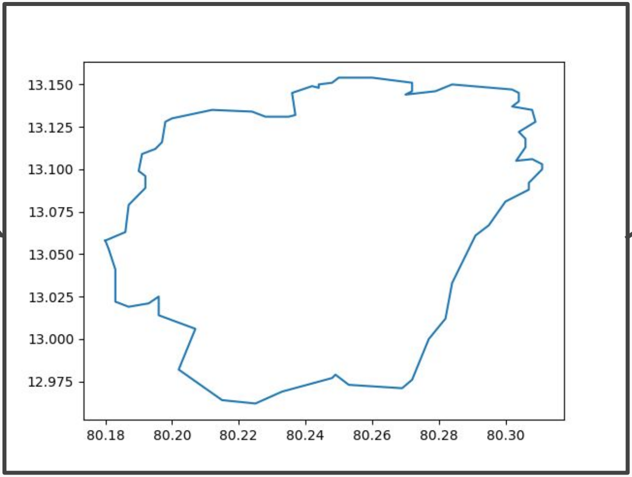
Disproportionate Areas

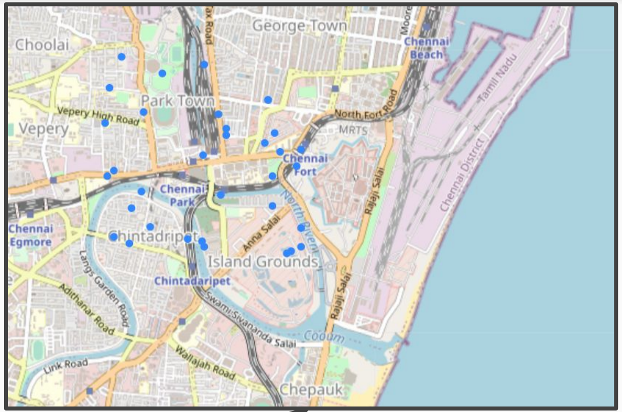
Translation and Contraction



Projected points

Bounding Polygon





Compressed to fit

Time Complexity

- Shapley's contain function is O(V) where V is the number of vertices in the polygon.
  In order to optimize this, one can alter the variables in Polygon creation by OSRM to obtain smoother or a detailed polygon.

- The compression factor also bears an inverse relationship in the time complexity

- Overall, O(N*V/k) where N is the total number of input points, V the number of vertices in the polygon and k is the compression factor

Embedding

Road distance d_Y (f(x), f(y)) can be modeled as a **D-embedding** to Euclidean distance d_X (x, y) with a **distortion D,** where D ≥ 1 is a real number, if there is a number r > 0 such that for all x, y ∈ X

$$r \cdot d_X(x, y) \leq d_Y(f(x), f(y)) \leq D \cdot r \cdot d_X(x, y)$$

This is an approximation embedding as an isometric embedding where,

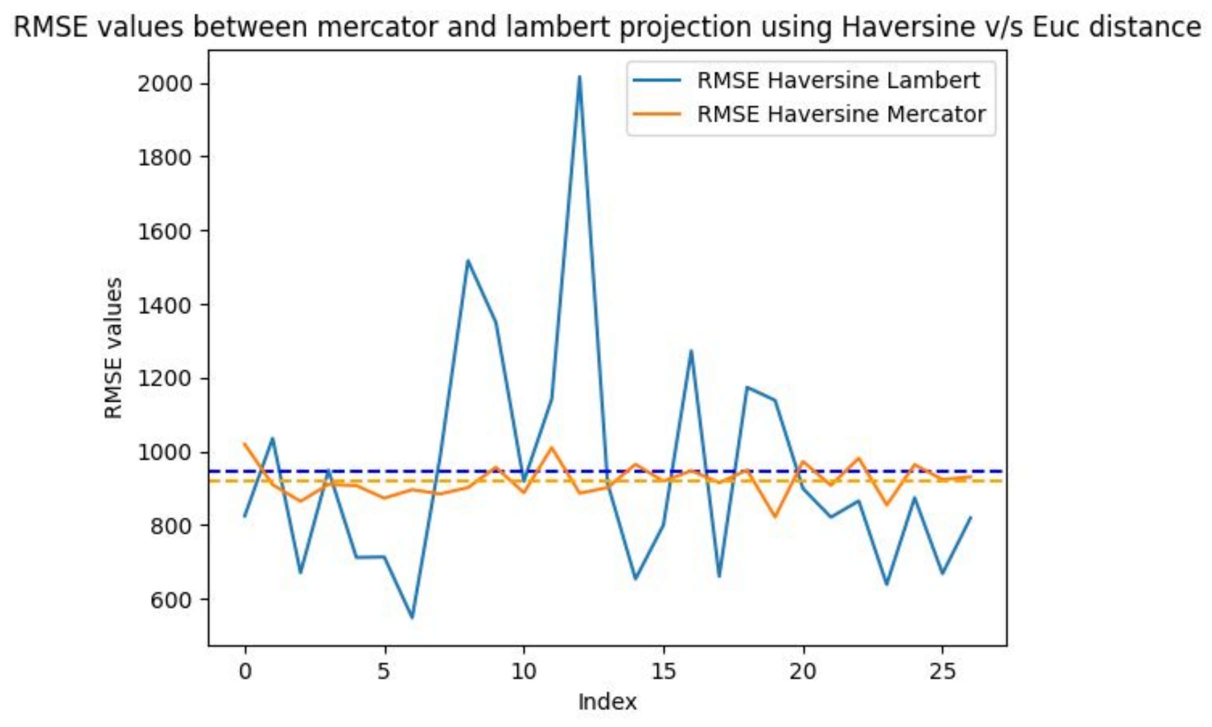
$$d_{road}[f(x), f(y)] = d_{euc}[x, y]$$

This was taken from Matousek, J., Lecture Notes on Metric Embeddings. Institute of Theoretical Computer Science, ETH Zurich, Switzerland, 2013.

Analysis

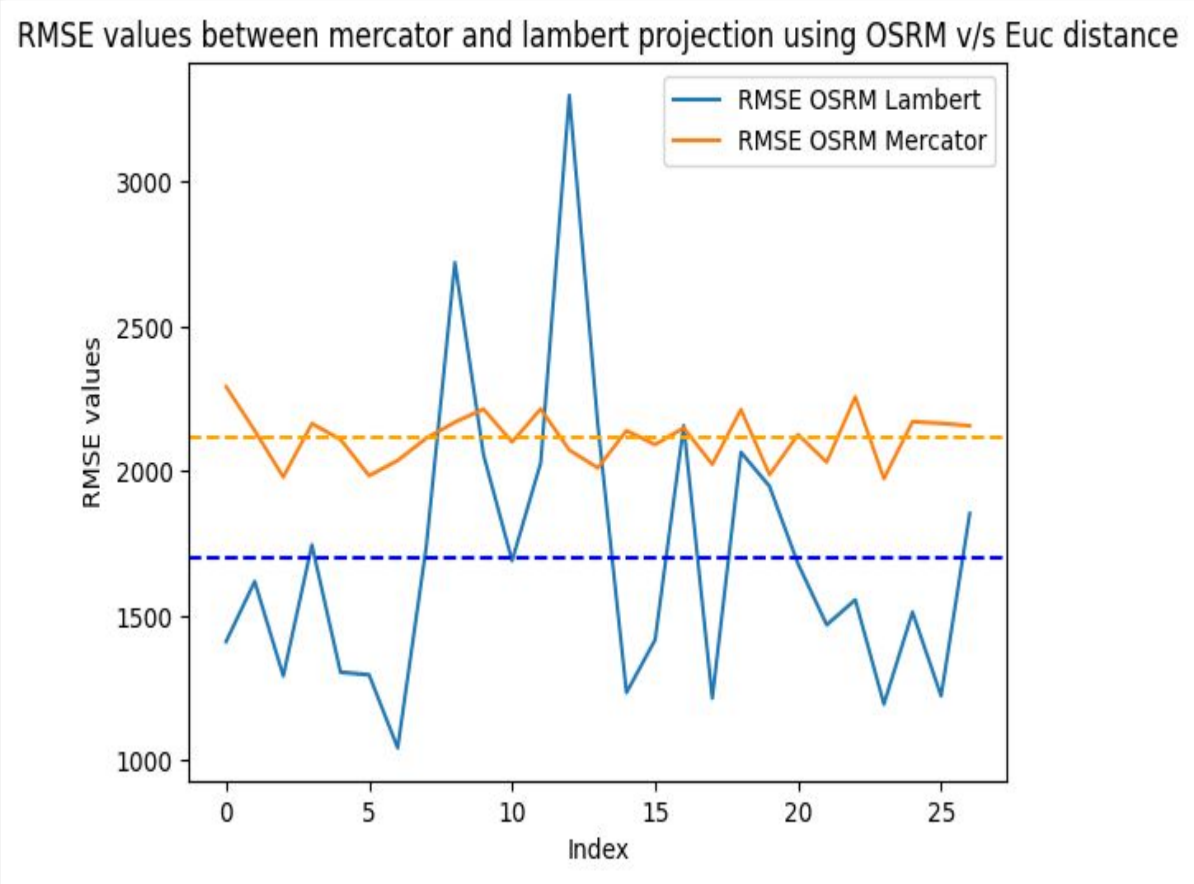
RMSE

- The Mercator projection method is suitable while considering Haversine distance

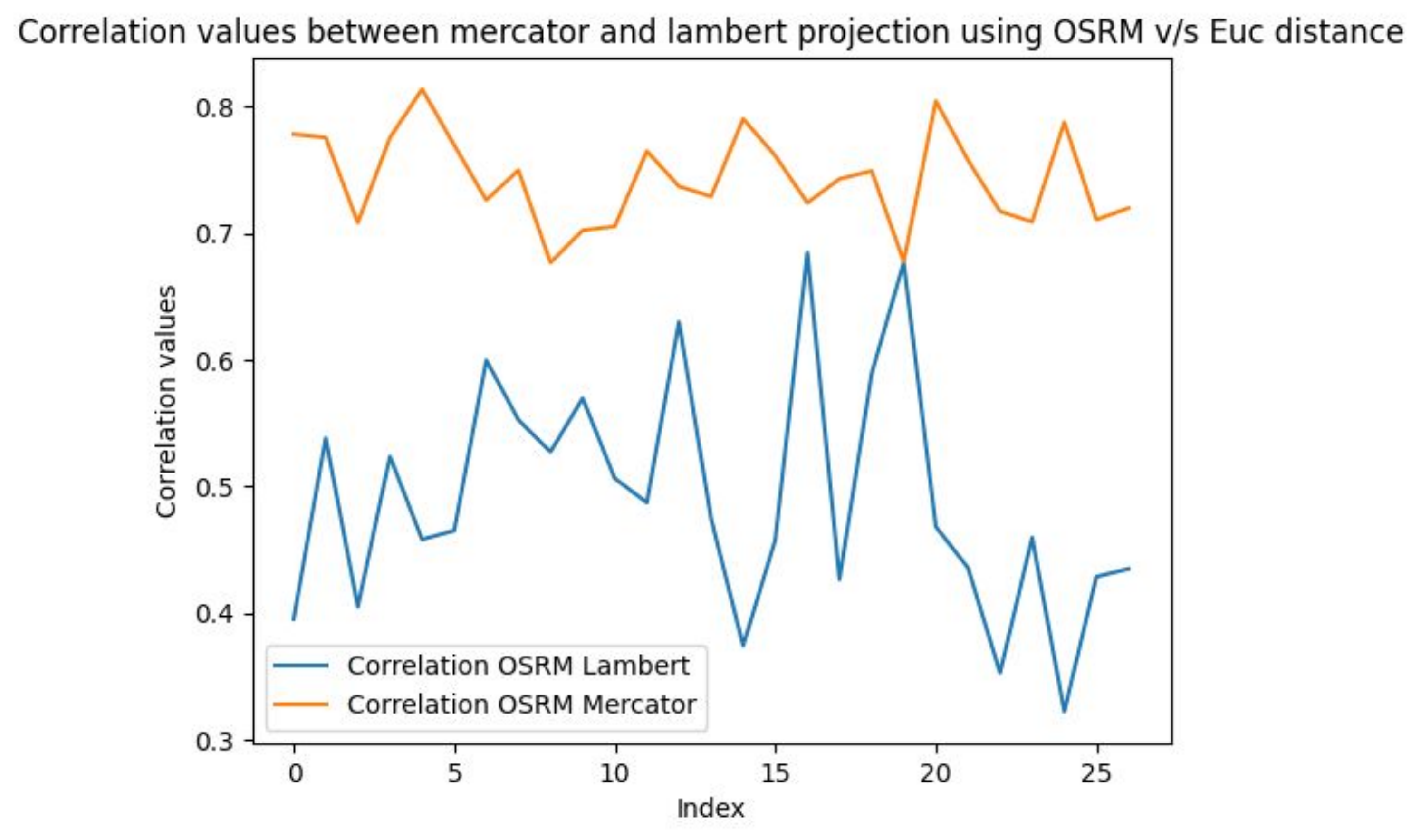


RMSE

❖ The Lambert conical conformal projection is a better fit when considering Road distance in the case of Chennai city



Correlation Coefficient

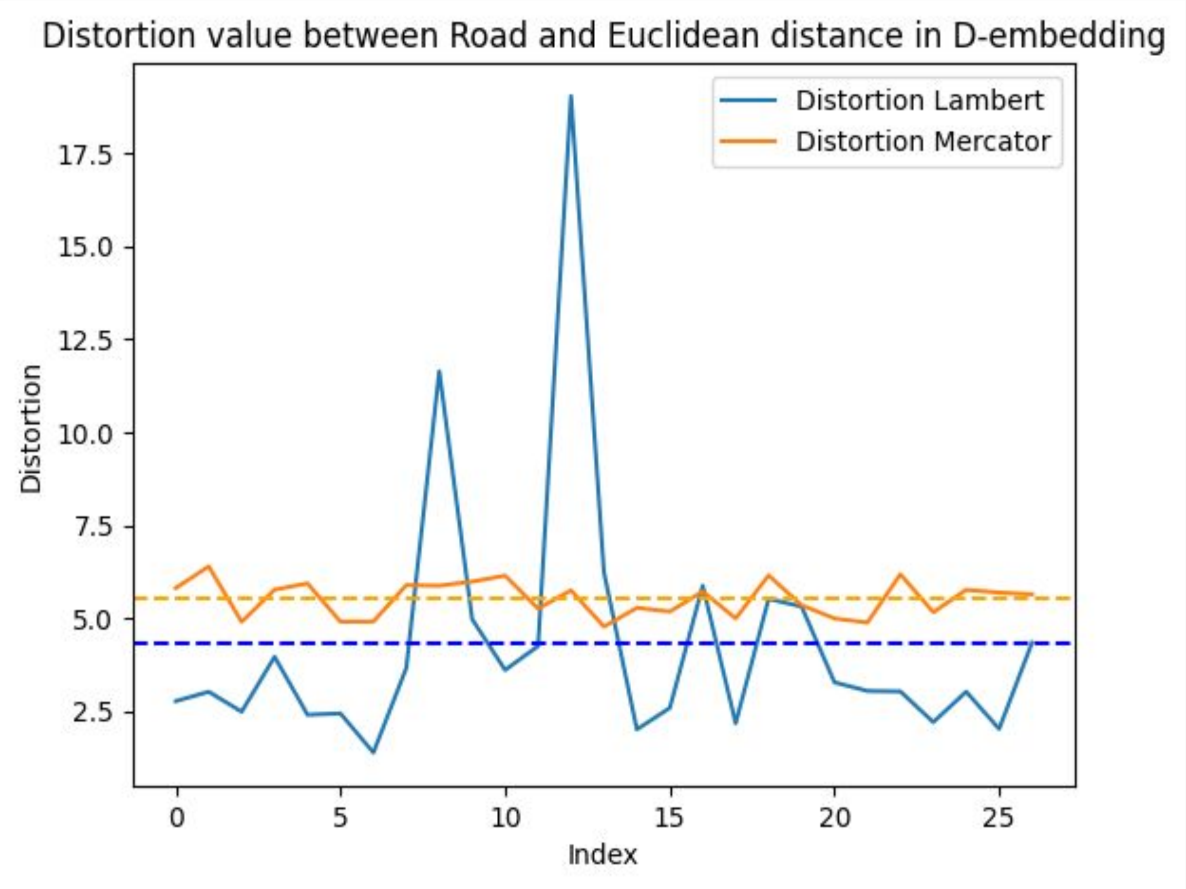


Distortion

- A lower value of D indicates a better embedding, and thus, the Lambert projection was found to be a superior embedding method for the Road distance metric in Chennai city

- D was calculated by dividing the sum of MSE by sum of euclidean distance squares

- R here is the greatest lower bound (was determined using typical Shepard-Kruska' inbuilt algorithm)

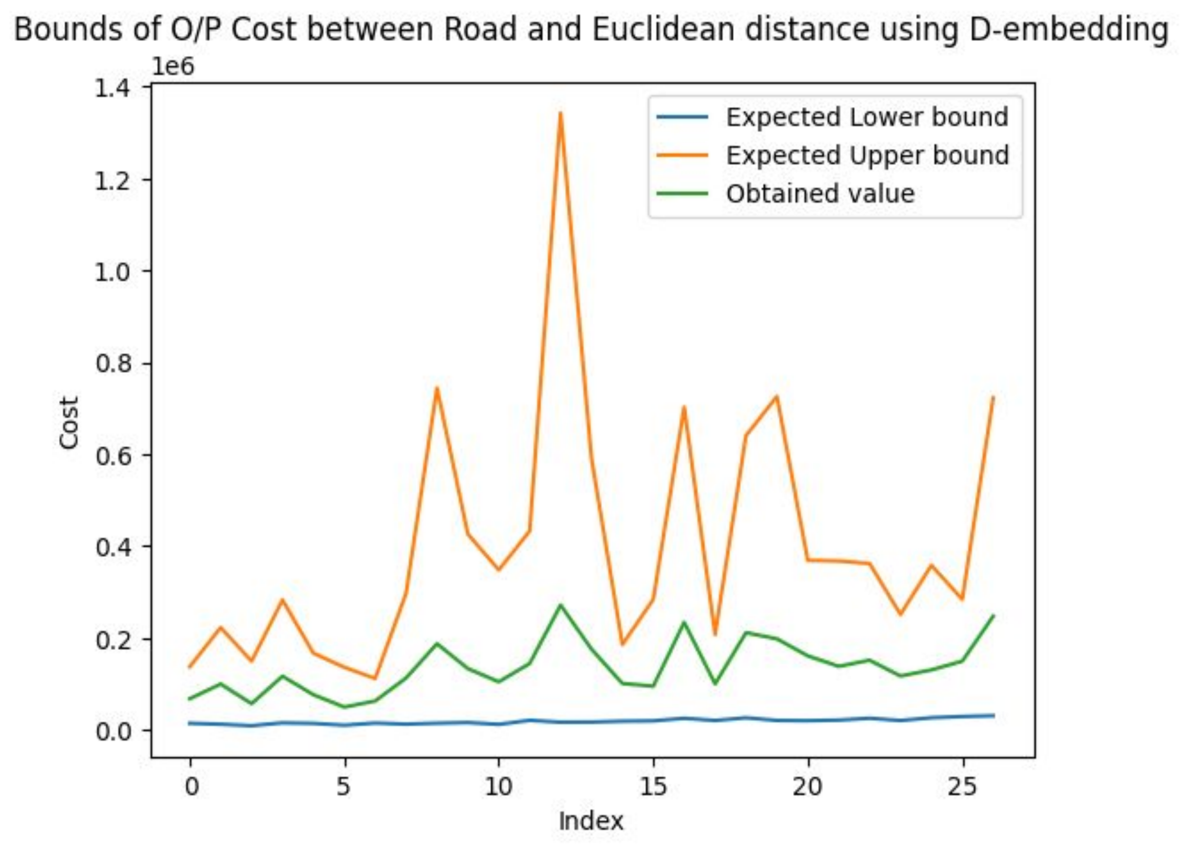


Bounds for Routing Algorithm

- ➔ We try to predict the bounds of output for road distance data by utilizing the cost associated with euclidean distance

- ➔ The obtained output values falll within the predicted upper and lower bounds, indicating that our model has effectively captured the relationship between road and euclidean distance

- ➔ However, this result may be specific to this particular dataset and still requires further testing before making a generalization

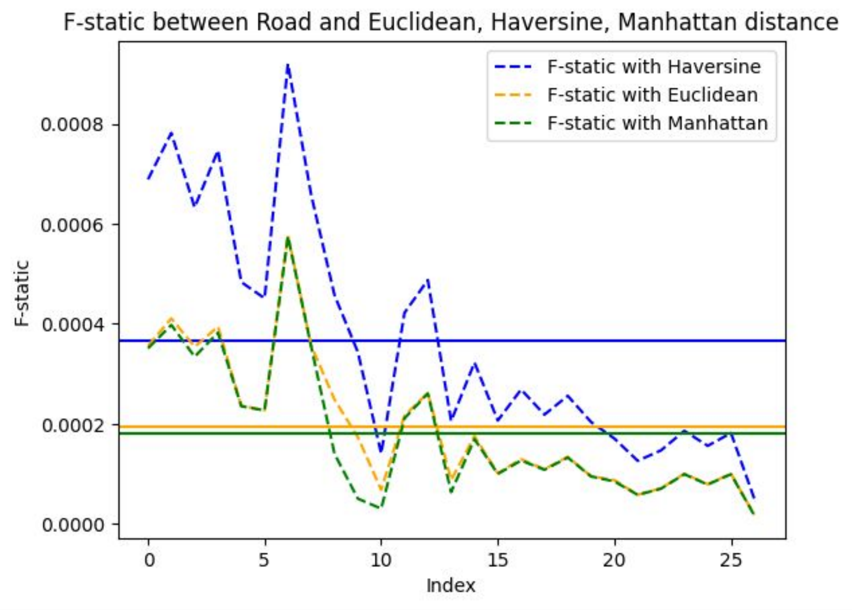


Underlying City Network as a Standard Metric

F-Statistic

The code applies the k-means clustering algorithm to find how similar the clusters are in each metric

A high F-statistic indicates that the clusters are significantly different from each other, meaning the metric used is ineffective in capturing the underlying structure of the road network

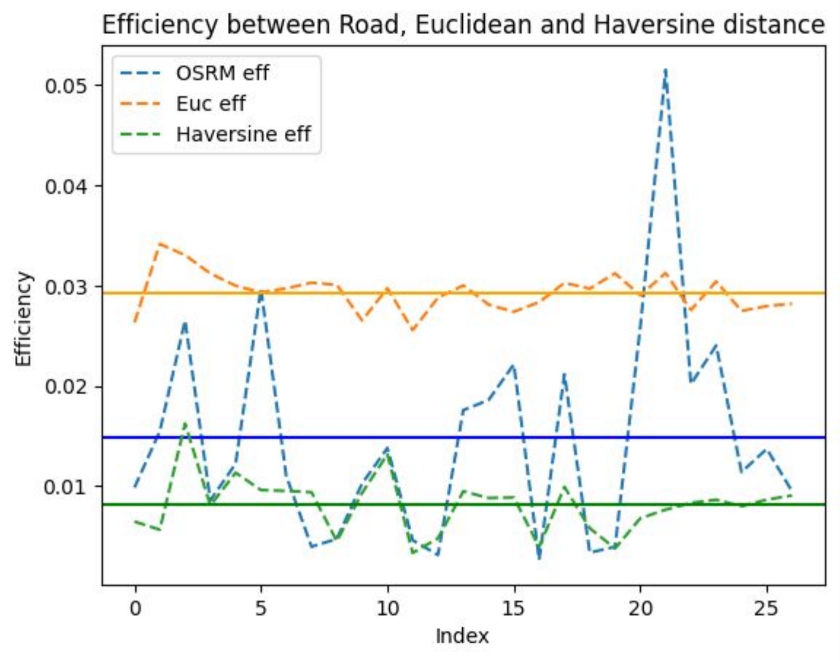


Underlying City Network as a Standard Metric

Simple Efficiency Metric

We consider the harmonic mean efficiency for each value in the distance matrix.

This is suitable for local travel and transportation as it gives more weight to smaller distances.



Applications

Applications

City Planning

- Building new roads
- Optimising existing network
- Allocating resources for public transportation

Logistics and Delivery

- Optimising delivery routes
- Setting up headquarters or warehouses in ideally connected locations

Environmental Impact

- Locating areas contributing to air pollution due to traffic

Conclusions

- → Methodical Transformation - Inverse Map projection
- → City's Road Network analysis
- → Geometric Embedding of road distance