Bachelor Thesis
Assessment of approximation method for TSP path length on a road network: a simulation study

Koen Stevens
S5302137

April 6, 2025

1 Introduction

The Traveling Salesman Problem is an important problem in operations research. It is particularly relevant for last-mile carriers and other logistics companies where efficient routing directly impacts cost, time and service quality. Efficient approximation methods are important for practical applications where exact solutions are too computationally intensive to conduct. Beardwood, Halton, and Hammersley (1959) proved the relation

$$L \to \beta\sqrt{nA}, \text{ as } n \to \infty \tag{1}$$

as an estimation for the length of the shortest TSP path measured by euclidean distance through n random locations inside an area in $\mathbb{R}^2$ with area A, where β is some proportionality constant. This is an asymptotic result, but it also holds for small n in areas of regular shape, such as triangles or squares. 1 underestimates the true value however, if n is small in elongated shape. Daganzo (1984). The main objective of this research is to estimate β and assess how well formula 1, approximates the length of the optimal TSP path in a real-world area.

2 Literature Review

3 Research Question

A few questions arise in order to tackle this problem:

- What is β in formula 1 when applied to a real road network?
- How well does this formula predict the length of the TSP path on a road network?

- Do the value of β and the performance of the formula vary across different types of areas (rural or urban, grid-like or organic)?

4 Hypotheses

I expect to find a higher β than the β that has been estimated previously in simulations where the $\mathbb{R}^2$ was used and no road network was considered, since traveling over a road network is always just as long or longer than traveling in a straight line. I also expect this β to vary when different types of areas are considered.

5 Data

The data used is OpenStreetMap data for the Province of Groningen.

6 Methodology

By comparing the estimated and actual TSP path lengths for simulated sets of locations within the area, this study aims to evaluate the accuracy of this formula in a real-world setting.

1. Make a connected graph of the road network of Groningen, using the igraph module in python.
2. Simulate locations on this graph.
3. Make a distance matrix of these locations.
4. Use an efficient method to find the shortest TSP path (LKH).
5. Compute the length of this TSP path.
6. Conduct 2-5 a number of times and estimate β.
7. Assess how good of a predictor the formula $L = \beta\sqrt{nA}$ is.

References

Beardwood, Jillian, John H Halton, and John Michael Hammersley (1959). The shortest path through many points. In *Mathematical proceedings of the Cambridge philosophical society*, Volume 55, pp. 299–327. Cambridge University Press.

Daganzo, Carlos F (1984). The length of tours in zones of different shapes. *Transportation Research Part B: Methodological 18*(2), 135–145.

7 Appendix

Some code of the current progress.
Some visualizations of the road network and paths generated with shortest path algorithm.