William Rodman

Honors Thesis

Computer Science Department

12 September 2023

Practical Implementations of the Traversal Distance

As digital road network datasets such as Apple Maps and OpenStreetMap have become increasingly important in the last two decades. The petabytes of data these platforms have collected has created a need to research algorithms capable of measuring the distance between two road network datasets.

In computer science, road networks are stored in a data structure called a geometric graph. One algorithm capable of measuring the distance between geometric graphs is the Traversal Distance. Given a distance epsilon, the Traversal Distance algorithm traverses a geometric graph to compute the free space between another geometric graph, then checks the free space projects over the entire first geometric graph. Applying binary search to the Traversal Distance algorithm will find the minimum distance epsilon. **My thesis** will determine how effective the Traversal Distances is at measuring distance between two geometric graphs for different types of datasets.

This problem first requires compiling datasets that includes labeled pairs of geometric graphs; examples of datasets include digital road networks, leaf patters and mandarin characters. Next, recording the Traversal Distances precision by computing the minimum epsilons of the datasets. Finally, determining the Traversal Distances effectiveness relative to other graph distance algorithms by comparing their precision across datasets and algorithms; other algorithms include the Hausdorff Distance and Graph Sampling Distance. The Hausdorff

Distance measures spaces between polynomial curves by measuring how far their subset of space is from each other, and the Graph Sampling Distance measures distance between two geometric graphs by computing a one-to-one matching of points along both graphs, then measuring distance between matched points.

While Traversal Distance is a well-defined method of measuring distance, there is a **current gap in understanding** of how effective the algorithm is after implementation. Therefore, there is a need to investigate how effective the algorithm is when implemented and applied to real world data. **My hypothesis** is the Traversal Distance is a polynomial time algorithm more effective at measuring distance between specific types of datasets than the Hausdorff Distance and Geometric Graph Distance. The rationale being that the Geometric Graph Distance is potentially more effective but hard to compute for large datasets as a non-polynomial time algorithm. In addition, the Traversal Distance is more effective than the Hausdorff Distance for measuring distance between geometric graphs. After completing this project, the **expected outcome** will be a concrete understanding of when to use the Traversal Distance in practical applications.

Bibliography

Graph Sampling for Map Comparison. Erfan Hosseini Sereshgi, Carola Wenk.

Distance Measures for Geometric Graphs. Sushovan Majhi, Carola Wenk

Matching Planar Maps. Carola Wenk.

Measuring the Similarity of Geometric Graphs.