Network Clustering Using Detector Data Considering Road Classes

Master's Thesis of Usman Maqsood

Mentoring:
M.Sc. Yunfei Zhang,
M.Sc. Yamam Alayasreih

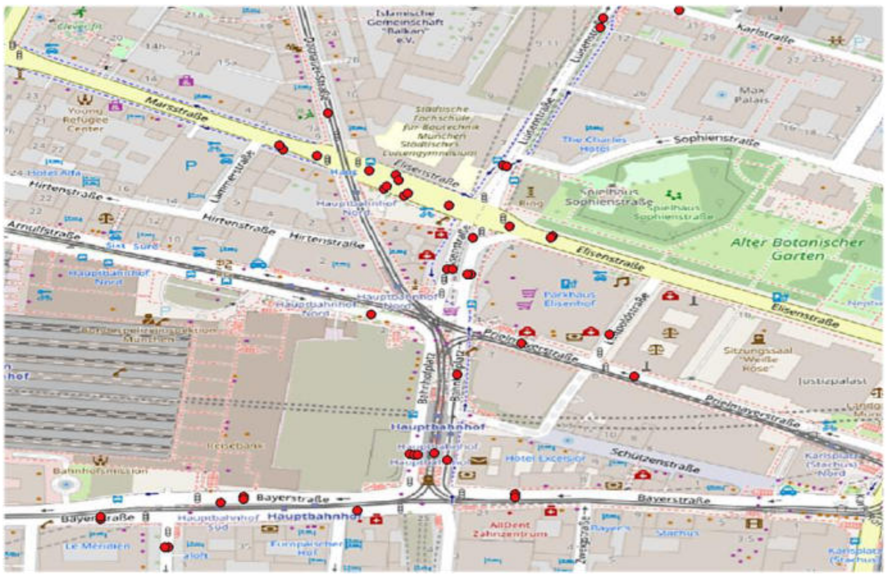

Figure 1: Stationary Loop Detectors on Munich Road Map

Introduction and Objective

This master thesis mainly focuses on developing a novel clustering technique for traffic state estimation that incorporates road class information alongside geographic and traffic pattern data. In many urban traffic studies, the clustering of loop detector data is solely based on geospatial information and traffic behavior. However, these studies often overlook the significance of incorporating road classes, such as motorways, primary roads, secondary roads, and residential roads, which can influence the clustering of loop detectors. This study addresses the gap by proposing a method that incorporates road class information into the clustering process. Specifically, the research addresses two central questions: (1) What indicators are most suitable for evaluating the quality of clustering algorithms in traffic network analysis? and (2) How does the integration of road class information affect the performance of clustering algorithms compared to those that rely solely on spatial or traffic metrics? Through this investigation, the thesis seeks to demonstrate measurable improvements in clustering quality.

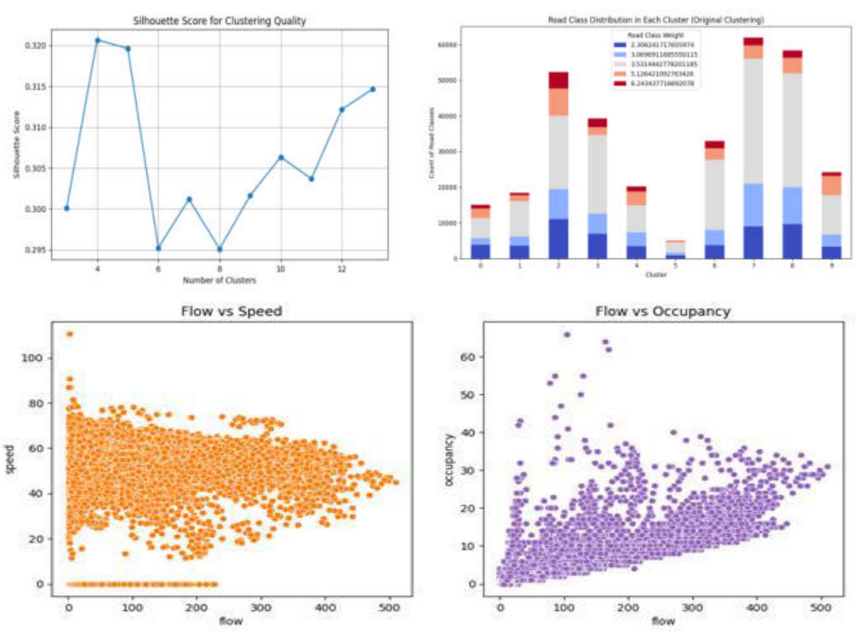

Figure 2 : Sensitivity Analysis and Relation of Speed/Flow/Occupancy

Methodology

The study utilizes stationary loop detector data from the city of Munich, capturing traffic metrics such as flow, occupancy, and speed. Initial preprocessing involved filtering out incomplete or unreliable readings, resulting in a cleaned dataset that retained approximately 62% of the original records. The geographic coordinates of each detector were then enriched with semantic information by matching them to road segments using QGIS and OpenStreetMap data, thereby assigning road class labels to each detector. Clustering was performed incrementally in four stages: **Step 1 -** Geographic Clustering: Detectors were grouped using K-Means based solely on longitude and latitude to establish a spatial baseline. **Step 2 -** Traffic Data Integration: Occupancy data was added to capture temporal traffic behavior. **Step 3 -** Flow Data Addition: Traffic flow was included to reflect demand variability. **Step 4 -** Road Class Weighting: Road class was integrated using two weighting strategies: ratio-based and logarithmic. Each class (e.g., motorway, primary, residential) was assigned a numeric weight, representing its importance in urban mobility. Each stage was evaluated using Key Performance Indicators (KPIs) such as the Silhouette Score, Davies-Bouldin Index (DBI), Calinski-Harabasz Score (CHS), and Occupancy/Flow variances to identify the most meaningful cluster structure.

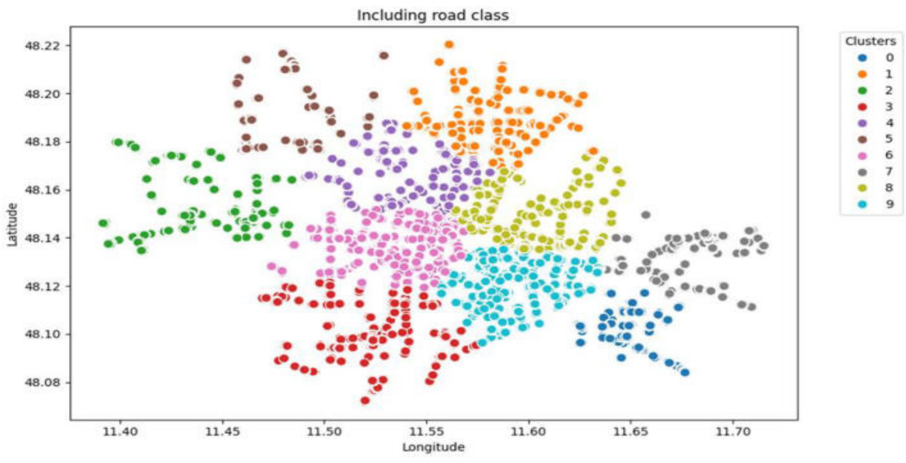

Figure 3 : Clustering based on Geospatial, Traffic & Road Classes

Results and Conclusion

The clustering analysis revealed that a 10-cluster configuration provided the best trade-off between cohesion and separation. This setup resulted in the lowest Davies-Bouldin Index (21.31) and a balanced flow and occupancy variance across clusters, indicating optimal internal compactness and external separability. Clustering based solely on spatial features produced overly broad groups with limited interpretability. However, adding traffic metrics and especially road class weights significantly improved cluster quality and alignment with actual traffic behavior. Both ratio-based and logarithmic weighting methods for road classes demonstrated notable improvements over traditional methods. The most homogeneous and compact clusters were achieved using a combination of all features: geographical coordinates, occupancy, flow, and road class. These refined clusters better captured the diversity of Munich's urban road network. In conclusion, the findings indicated positive outcomes when road classes were incorporated in grouping loop detector data.