Optimizing Map Feature Collection Routing Using Graph Database Technology

Kaede Hasegawa[1], Prof. Antonis Bikakis[1], Akos Magdo[2]
[1]University College London, [2]HERE Technologies

Background and Motivation

Accurate digital maps are crucial for efficiently meeting logistics demands, requiring frequent updates to road features. This creates the challenge of **finding an optimal path that visits all roads needing data collection, known as the Rural Postman Problem (RPP)**. Traditionally, RPP has been addressed using algorithms developed with tabular databases, such as relational databases. However, graph databases, which represent data in a network structure, can better capture road relationships, offering potential advantages in solving RPP [1]. Therefore, **this study explores the novel application of graph databases and their query language for developing optimization algorithms to address the RPP**.

Data and Methods

Using German road network shapefiles provided by HERE Technologies, 3 regions containing 100, 300, and 501 roads were converted into 3 graph databases in Neo4j via CSV import. These databases were tested with 3, 5, and 10 specified edges over 3 trials, averaging performance based on distance and time. In the graph database, **nodes represent junctions, edges represent roads, and properties store their characteristics** (Figure 1).

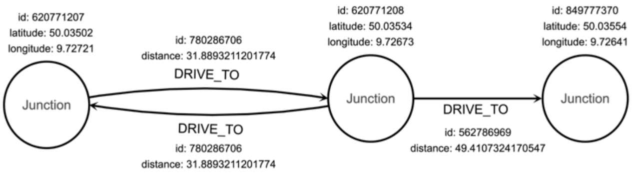

Figure 1: Road Network in Graph Database

4 RPP algorithms were implemented:

(a) **Nearest Neighbour** (NN) [2]: Existing heuristic that selects closest specified edge

(b) **Monte Carlo** (MC) [3]: Existing heuristic that randomly selects edges to form a path

(c) **Genetic Algorithm** (GA) [4]: Existing metaheuristic that evolves a population of solutions using mutation & natural selection

(d) Cypher-Only Algorithm (CA): A novel algorithm that uses Cypher alone

All algorithms implemented **proposed graph transformation**, which leveraged **A* Algorithm** from Neo4j (a heuristic-based shortest path algorithm). It simplified the problem by connecting end nodes of the specified edges with their corresponding shortest paths.

Key Findings

Results in Figure 2 revealed the following:

- **Monte Carlo** produced the **shortest** routes.
- **Nearest Neighbour** computed the **fastest.**
- **Cypher Algorithm** generated shortest routes but **encountered timeouts** when the specified edges exceeded 3.
- **Genetic Algorithm performed the worst** in both measures, contrary to existing studies.

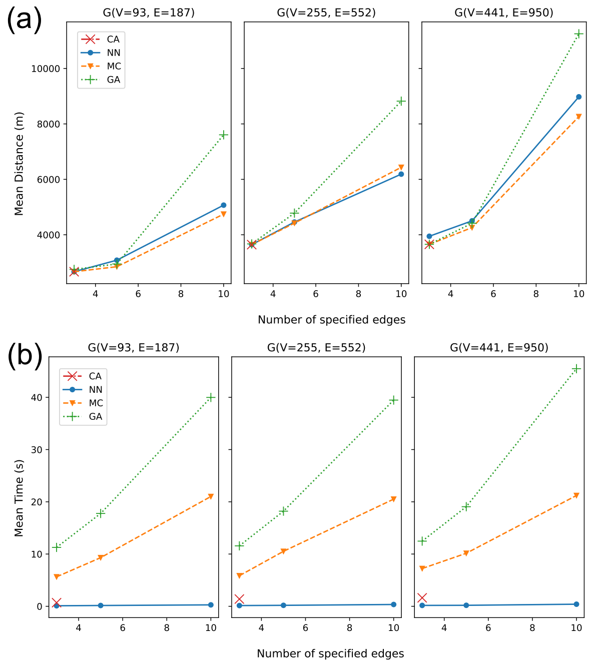

Figure 2: Algorithm Performance per Graph Condition in (a) Distance and (b) Computed Time

Value of the Research

The findings highlight the limitations of using Cypher alone for the RPP, particularly in larger instances. However, its strong path-finding ability in smaller instances and high performance of other Cypher-integrated algorithms suggest **significant potential for the novel use of Neo4j and Cypher in solving the RPP**.

Future research directions include:

- Implementing **exploitative strategies** (e.g., hybrid methods) in the Genetic Algorithm to address its possible over-exploration.
- Incorporating **real-life factors** (e.g., customised weighting) to enhance real-life applicability, given the simplified setup.
- **Comparing** these graph-based algorithms with those using **relational databases** to assess the effectiveness of graph databases.

References:
[1] Miler, M., Medak, D., and Odobašić, D.: 'The shortest path algorithm performance comparison in graph and relational database on a transportation network', Promet-Traffic&Transportation, 2014, 26, (1), pp. 75-82
[2] Adnan, S., Abood, e.W., and Abdulmuhsin, W.: 'The multi-point delivery problem: Shortest Path Algorithm for Real Roads Network using Dijkstra', Journal of Physics: Conference Series, 2020, 1530, pp. 012040
[3] De Córdoba, P.F., Raffi, L.G., and Sanchis, J.M.: 'A heuristic algorithm based on Monte Carlo methods for the rural postman problem', Computers & operations research, 1998, 25, (12), pp. 1097-1106
[4] Kang, M.-J., and Han, C.-G.: 'Solving the rural postman problem using a genetic algorithm with a graph transformation', in Editor (Ed.)^(Eds.): 'Book Solving the rural postman problem using a genetic algorithm with a graph transformation' (1998, edn.), pp. 356-360