Math Extended Essay

Veer Mehta

1.0 Journey

1. Got the data about route of a bus from the school transport department.

2. Looked up at various methods to optimise the route the bus travels. Realised that it is a travelling salesman problem. Travelling salesman problem is one in which a salesman starts to travel from one city and travels to various different cities and returns to the same city from where she/he started. Each city needs to be covered only once. The shortest route has to be identified.

3. Searched about various methods to solve the TSP. There are major 3 types of methods.

   - Exact algorithms
     - Guaranteed to find an optimal solution
     - Would not be practically possible to calculate large problem instances due to their exponential time complexity.
     - It includes Brute force, Dynamic programming and Branch & Bound methods.
   - Heuristics
     - Faster than exact algorithms and can provide good solutions in a reasonable amount of time.
     - Do not guarantee optimal solution
     - It includes Nearest neighbour, genetic algorithms and ant colony optimisation
   - Approximations
   - Provide a solution that is guaranteed to be within a certain factor of the optimal solution.
   - It includes Christofides' algorithms and 2-approximation algorithm

4. The research question required a very accurate answer, and this will be provided by exact algorithms. Choice between branc & bound and dynamic programming.

   Advantages of dynamic programming:

   - Guaranteed optimality: Dynamic programming guarantees that the solution found is optimal, meaning it provides the shortest possible Hamiltonian cycle that visits all nodes exactly once.

   - Faster for small instances: Dynamic programming is generally faster than branch and bound for small instances of the TSP (up to 20-25 nodes), due to its efficient use of memory and simpler computation.

   - Easier to implement: Dynamic programming is generally easier to implement than branch and bound, as it requires fewer steps and less complex code.

   Disadvantages of dynamic programming

   - Memory requirements: Dynamic programming requires a large amount of memory to store intermediate results, which can be a challenge for large instances of the TSP.

   - Inefficient for large instances: Dynamic programming can be inefficient for large instances of the TSP (more than 25 nodes), as the number of sub-problems grows exponentially with the size of the problem. This can lead to very long computation times.

   - Not suitable for asymmetric TSP: Dynamic programming is only suitable for symmetric TSP, where the distance between two nodes is the same in both directions.

   Advantages of Branch bound

   - Good for large instances: Branch and bound is generally better suited for large instances of the TSP, as it uses a systematic search process to explore the solution space and can prune large parts of the search tree, leading to faster computation times.

   - Suitable for asymmetric TSP: Branch and bound can be used to solve both symmetric and asymmetric TSP.

- Can handle constraints: Branch and bound can be extended to handle additional constraints, such as time windows or capacity limits.

Disadvantages for Branch & Bound

- No guarantee of optimality: Branch and bound does not guarantee that the solution found is optimal, as it relies on heuristics to prune the search tree and may miss the global minimum.
- More complex implementation: Branch and bound requires more complex implementation than dynamic programming, due to the need to handle branching and pruning of the search tree.
- Higher computation time for small instances: Branch and bound may take longer to solve small instances of the TSP due to the overhead of branching and pruning the search tree.

5. Dynamic program is well suited for the bus route optimisation and hence it was chosen.

6. Requirement for solving any type of TSP is to create a distance matrix, which involves distance from each stop to other stop. The matrix was calculated using google maps where all the addresses that the bus covered were plotted and distances were provided by google maps.

7. I learned linear programming from youtube videos. Solved an example and tried to apply it to the distance matrix. However soon i realised that using dynamic programming with 15 stops would be very large to calculated by hand.

8. Solution to the problem is that i can cluster the 15 stops into smaller groups and then calculate the shortest path using dynamic programming for each cluster and then connecting all the clusetrs to get the final optimised route.

9. Next problem is how to cluster. Hence the method of clustering i found is K means clustering algorithm which helps to form clusters.

2.0   Procedure to follow

1. Construct a distance matrix

2. Cluster the stops into 3 groups using k means clustering method

3. Apply dynamic programming to the clusters and identify the best route

4. Connect the clusters and the school vertex

5. Get the optimised route