The Generalized Travelling Salesman Problem

Finding the shortest possible USA 49 States Tour

Brandon Yeh

1   Overview

In the Generalized Travelling Salesman Problem (GTSP), the nodes V are partitioned into k clusters $V_1, V_2, ..., V_k$ and the goal is to find a minimum cost circuit that includes exactly one node in each cluster.

In this report, a well known genetic algorithm published by John Silberholz and Bruce L. Golden is used to attempt to find a good solution to the 49 tour problem. http://josilber.scripts.mit.edu/GTSP.pdf.

However, this genetic algorithm only tests instances with up to 1000 nodes and a large number of clusters with approximately 5 nodes each. There is no computational data provided to show the performance of such an algorithm with siginificantly greater nodes and few smaller clusters. More specifically, this report will outline the performance of such an algorithm that contains 49 clusters of, on average, 2300 nodes each and 115475 nodes in total.

Implementation of this algorithm is done in Java 8 Version 74, using only the native library provided in the SDK. The computer used to run the computational tests have the following properties:

1. Acer Aspire 5755G-9417 with an Intel Core i7-2670QM 2.2GHz Processor
2. Operating System: Ubuntu 14.04.4 LTS running Linux Kernel 3.13.0-83-generic (i686)
3. 2x Crucial 4GB Single DDR3 1600 MT/s PC3-12800 CL11 SODIMM 204-Pin 1.35V/1.5V Notebook Memory CT51264BF160BJ
4. Samsung Electronics 840 EVO-Series 250GB 2.5-Inch SATA III Single Unit Version Internal Solid State Drive MZ-7TE250BW

2   Problem Formulation

This section of the report will formally define the problem that the modified genetic algorithm will solve.

The following data is to be used as input to the genetic algorithm:

1. A list of 49 States in the continental USA
2. A list of 115475 cities in the above 49 states provided from the following url: http://www.math.uwaterloo.ca/tsp/data/usa/usa115475_cities.txt

Furthermore, the following assumptions are made to simplify the problem:

1. The travel distance between 2 cities provided in the list of cities above is the Euclidean distance between the pair of cities. This identical to the EUC_2D method in TSPLIB.

Finally, the problem is defined formally:

> Given a list of 49 states in the continental USA and a list of cities within each state, find the shortest tour through each 49 states, visiting exactly one city in each state.

3 The Genetic Algorithm

As noted in the overview, the algorithm used to sove this 49 State USA tour is designed by John Silberholz and Bruce L. Golden albiet, with a few modifications.

3.1 Input Parameters

The following input parameters are defined in the algorithm itself:

```
private int INITIAL_POPULATION_SIZE;
private int NUMBER_OF_ISOLATED_POPULATIONS;
private int TOTAL_GENERATIONS;
private int REPLICATON_SIZE;
private int REPRODUCTION_SIZE;
private int TERMINATION_CONDITION;
```

We first define the size of each isolated population as well as the number of isolated populations

3.2 Isolated Population Generation

3.3 Natural Selection

3.4 Reproduction

3.5 Isolated Population Merging

4 Heuristics and Subroutines

5 Testing

5.1 Summary