INFO 6205 Spring 2022 Project

Traveling Salesman Problem

Created by:

Aishwarya Venkatesan, <u>venkatesan.ai@northeastern.edu</u> ,NUID - 001569213

Hinal Patel, patel.hina@northeastern.edu, NUID - 002766180

Introduction:

Aim: To Solve the Traveling Salesman Problem (TSP) is an NP-hard algorithm for finding the shortest tour of n cities (or, in general, points in a two-dimensional space).

Approach: Solving the Traveling Salesman Problem (TSP) based on Christofides algorithm is a heuristic algorithm that combines minimum spanning tree (MST) and Eulerian circuit concepts to find an approximate solution with a guarantee of being within a factor of 1.5 of the optimal solution.

Steps of Christofides algorithm

- 1. **Input**: The input to Christofides algorithm consists of a complete graph with weighted edges, where each vertex represents a city to be visited by the salesman, and each edge represents the distance (or cost) between two cities.
- 2. **Step 1: Minimum Spanning Tree (MST)**: The first step of the Christofides algorithm is to construct a minimum spanning tree (MST) of the input graph. The MST is a tree that spans all the vertices of the graph and has the minimum possible total edge weight. This is achieved by using Prim's algorithm.
- 3. **Step 2: Finding Odd-degree Vertices**: In the MST obtained from Step 1, some vertices may have odd degrees (i.e., an odd number of incident edges). These vertices represent the cities that are visited an odd number of times in the MST, and hence will have odd degrees. In this step, we need to find all the vertices in the MST that have odd degrees, as they are potential starting and ending points for the Eulerian circuit.
- 4. **Step 3: Perfect Matching on Odd-degree Vertices**: Next, we need to find a perfect matching (i.e., a set of non-overlapping edges that cover all vertices) on the set of odd-degree vertices obtained from Step 2. This can be done efficiently using any algorithm for finding a minimum-weight perfect matching. The perfect matching

- ensures that every vertex in the set of odd-degree vertices has an even degree, which is a requirement for Eulerian circuits.
- 5. **Step 4: Combining MST and Perfect Matching**: In this step, we combine the MST obtained from Step 1 with the perfect matching obtained from Step 3 to obtain a connected graph that has all vertices with even degrees. One way to do this is to add the edges of the perfect matching to the MST, creating a multigraph (i.e., a graph that allows multiple edges between the same pair of vertices). We then convert this multigraph into a simple graph (i.e., a graph with no multiple edges) by removing duplicate edges and obtaining the sum of the weights of the remaining edges.
- 6. **Step 5: Eulerian Circuit**: The next step is to find an Eulerian circuit in the graph obtained from Step 4. An Eulerian circuit is a closed path that visits every edge of a graph exactly once. In the graph obtained from Step 4, all vertices have even degrees, which guarantees the existence of an Eulerian circuit. The Eulerian circuit represents a cycle that passes through all the vertices of the graph and visits each edge exactly once.
- 7. **Step 6: Shortcutting the Eulerian Circuit**: In this step, we shortcut the Eulerian circuit obtained from Step 5 to obtain a Hamiltonian circuit, which is a valid solution to the TSP. A Hamiltonian circuit is a cycle that visits every vertex of a graph exactly once. To do this, we start at an arbitrary vertex of the Eulerian circuit and traverse the circuit, skipping already visited vertices and adding the unvisited ones to the Hamiltonian circuit. This process is repeated until all vertices have been visited, resulting in a Hamiltonian circuit.
- 8. Step 7: Calculating the Total Distance: Finally, we calculate the total distance

After Calculating initial tour distance using Christofides algorithm we optimize the tour by using tactical and strategic approach which includes :

- Tactical: 2-opt and/or 3-opt improvement;
- Strategic: simulated annealing, ant colony optimization, genetic algorithms, etc.

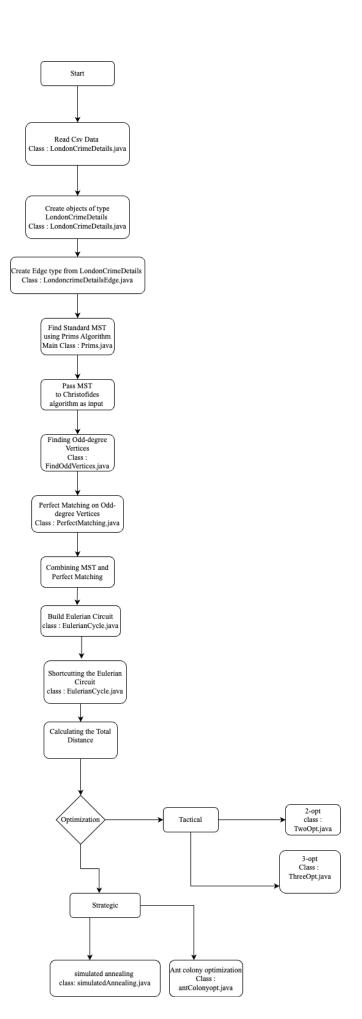
Program

- **Data Structures included**: LinkedList, List, Arrays, Stack, Queue, LinkedHashSet,Set,HashMap.
- Algorithm: Prims Algorithm, Christofides algorithm, 2-opt and/or 3-opt improvement, simulated annealing, ant colony optimization, genetic algorithms, etc.
- **Invariants**: Invariants are mathematical properties or characteristics of a problem that remain constant or unchanged throughout the problem-solving process. In the Traveling Salesman Problem (TSP), there are several invariants that are important to consider

- when developing algorithms or analyzing solutions. Here are some of the most important invariants in TSP.
- Symmetry: TSP is a symmetric problem because the distance between two cities A and B is the same as the distance between B and A. This means that the order in which we visit the cities does not affect the total distance traveled. For example, if we have a TSP with three cities A, B, and C, the distance from A to B is the same as the distance from B to A. Therefore, we can assume without loss of generality that we always start and end at a fixed city, such as city A.
- Triangle Inequality: The triangle inequality states that the shortest distance between two points is a straight line. In TSP, this means that the distance between two cities A and C is always less than or equal to the sum of the distances between A and B, and between B and C. This is important because it allows us to eliminate certain routes that would be longer than other routes that visit the same cities.
- Subtour Elimination: A subtour is a subset of cities that form a closed loop. In TSP, we want to find a solution that visits all cities exactly once, which means that there should be no subtours. Therefore, an invariant in TSP is that any solution must satisfy the condition that there are no subtours.
- Optimality Condition: An optimal solution to TSP must visit each city exactly once and return to the starting city, and its total distance must be the minimum possible among all such tours. This means that any solution that violates this condition is not optimal.

By considering these invariants, we can develop algorithms that guarantee certain properties of the solutions they find. For example, the Christofides algorithm guarantees a solution that is within a factor of 1.5 of the optimal solution by exploiting the invariants of TSP.

Flowchart:



Observations:

The Traveling Salesman Problem (TSP) is a difficult problem to solve, and there is no single algorithm that works best for all instances of the problem. The Christofides algorithm is a good starting point for obtaining an initial solution, which provides an approximation solution within a factor of 3/2 of the optimal solution and has a time complexity of O(n^3), making it fast and easy to implement. However, the Christofides algorithm can produce suboptimal solutions for some instances of the TSP problem and may not work well for large TSP instances, which have a time complexity of O(n^2 log n) due to the need to compute a minimum spanning tree.

To improve the solution obtained from the Christofides algorithm, 2-opt and 3-opt optimization techniques can be used, which have a time complexity of O(n^2) and O(n^3), respectively, and can be effective for moderately-sized TSP instances. However, for large TSP instances, the time complexity can be prohibitively high, and these techniques may not find the optimal solution for all instances of the TSP problem.

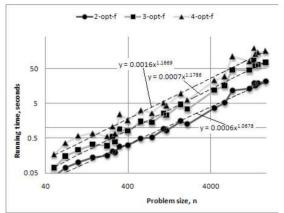
Simulated annealing is a metaheuristic algorithm that can be effective for finding good solutions to TSP problems and has a time complexity of $O(k * n^2)$, where k is the number of iterations. However, it requires careful tuning of the temperature parameter, and the number of iterations needed to achieve a good solution can be significant, making it computationally expensive for large TSP instances.

Ant colony optimization is a metaheuristic algorithm that can be effective for finding good solutions to TSP problems, especially for large TSP instances, and has a time complexity of O(m * n^2), where m is the number of ants. However, it requires careful tuning of the parameters, including the pheromone update rule, the ant movement rule, and the evaporation rate, and can also be computationally expensive for large TSP instances.

Graphical Analysis:

Comparing 2-opt and 3-opt

	4-opt-f		4-opt-f with greedy start and k-swap-kicks							
Problem name	Random	k							Best found	
	start	start	2	3	4	5	6	10	15	(k)
pla7397	4.85	2.94	2.01	1.74	1.62	2.04	1.86	2.28	2.71	1.34 (4)
rl11849	6.30	2.85	2.52	2.58	2.52	2.82	2.88	3.14	3.13	2.23 (4)
usa13509	5.11	3.07	2.61	2.72	2.59	2.98	2.93	3.21	3.19	2.47 (2)
brd14051	5.30	3.03	2.77	2.89	2.94	3.07	3.09	3.06	3.27	2.66 (2)
d15112	5.25	2.92	2.71	2.82	2.87	2.94	3.01	3.10	3.16	2.58 (2)
d18512	5.37	2.84	2.68	2.82	2.84	2.94	2.94	3.04	3.07	2.54(2)



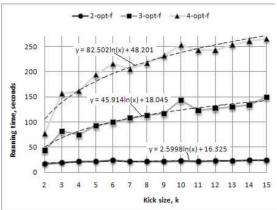
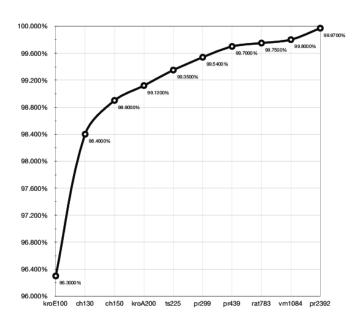


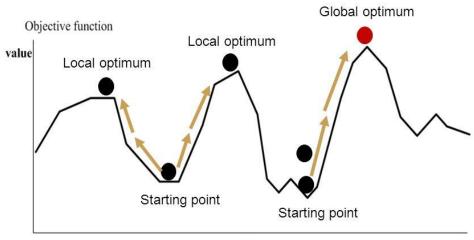
Figure 5. Running time dependence on problem size (k-swap-kick size = 4, logarithmic scale)

Figure 6. Running time dependence on k-swap-kick size (problem d18512)

1b]



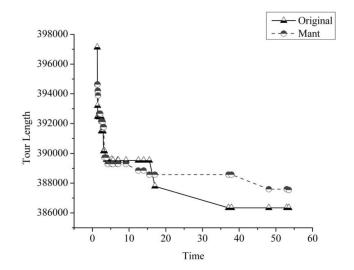
Simulated Annealing (3/3)



search space

Page • 16

1d]



Result:

SS of MST

8.0.714.5 1.62-358 0.35555, 157-162 0.55708, 157-462 0.49622, 298-402 0.3741, 138-298 0.17085, 298-410 0.38723, 150-358 0.46464, 410-573 0.65217, 402-517 0.82232, 413-517 0.44904, 413-579 0.41328, 135-579 0.39471, 135-301 0.10257, 135-422 0.28313, 301-456 0.29708, 38-301 0.588976, 38-317 0.22609, 317-353 0.14554, 33-353 0.11978, 159-579 0.41488, 97-519 0.61489, 34-519 0.51469, 34-519 0.50272, 373-532 0.09728, 337-533 0.19577, 533-458 0.25975, 103-545 0.08008, 419-533 0.57085, 413-42 0.28313, 301-456 0.2723, 373-532 0.09728, 337-533 0.29577, 533-548 0.25975, 103-545 0.08008, 419-533 0.57085, 413-42 0.28686, 289-545 0.26007, 289-546, 289-545 0.6007, 235-550 0.165676, 289-546 0.26007, 412-540 0.00008, 466-350 0.08008, 419-533 0.08887, 191-543 0.09538, 232-346 0.15976, 139-232 0.08008, 191-250 0.17749, 151-250 0.09019, 161-174 0.14657, 259-466 0.15506, 132-466 0.06021, 240-292 0.10794, 108-120 0.14488, 292-523 0.10919, 255-523 0.26183, 46-235 0.28183, 46-131 0.05809, 116-592 0.186614, 47550 0.19887, 349-441 0.30086, 489-447 0.09008, 466-250 0.09008

SS of MST Continue

8. 946.42, 144-319 1. 45815, 94-233 1.57525, 233-275 8.33192, 48-275 1.9536, 48-549 1.32078, 271-540 1.36875, 6-271 8.06855, 6-521 8.07825, 238-263 4.71207, 74-682 1.03416, 74-268 8.46949, 266-565 1.23166, 271-574 1.35978, 488-574 1.30889, 118-468 8.75927, 118-147 8.99282, 147-434 8.69883, 264-434 1.22113, 264-417 1.19711, 58-574 1.55619, 248-484 1.69116, 222-538 1.79344, 228-538 1.09252, 138-228 1.09745, 138-228 1.09745, 138-238 8.25788, 228-463 1.02186, 228-538 8.32743, 388-383 8.47969, 338-383 8.48165, 338-358 8.35817, 535-552 8.19110, 88-552 8.09715, 88-126 8.65359, 126-393 8.57762, 393-557 8.35488, 557-575 8.45555, 375-575 8.17377, 227-375 8.18517, 527-557 8.06679, 212-2201 1.46360, 18-212 8.54243, 18-218 8.62367, 18-26 8.65137, 26-411 8.92919, 189-411 1.32546, 189-219 1.37667, 201-354 1.43379, 87-411 1.64457, 527-557 8.06679, 212-2201 1.7169, 2-425 8.39116, 341-425 8.27861, 32-425 8.46833, 341-566 8.79981, 2-211 8.1854, 48-211 1.24569, 46-512 8.19284, 483-512 8.19343, 42-483 8.89889, 32-534 8.98889, 327-33 8.5782, 393-557 8.35489, 527-559 8.17374, 295-559 8.17314, 595-554 8.17377, 227-375 8.18727, 201-218 8.18547, 495-218 8.18517, 528-548 8.18518, 528-548

number of odd vertices 23

perfect matching pairs [243-576 0.00000, 182-466 0.00000, 51-177 0.00000, 103-545 0.00000, 161-250 0.00019, 135-301 0.10257, 33-353 0.11078, 108-120 0.14188, 235-503 0.15687, 139-360 0.15770, 171-35 0.16062, 77-544 0.1625, 138-298 0.17065, 398-570 0.17741, 241-376 0.18680, 119-521 0.20022, 406-566 0.26875, 292-467 0.24607, 183-290 0.32577, 287-561 0.33304, 39-431 0.36025, 3-279 0.37585, 2-425 0.39110, 216-383 0.41162, 25-468 0.41878, 45-272 0.44505, 481-492 0.450008, 409-577 0.45912, 1-143 0.4665, 283-462 0.47731, 168-460 0.48170, 422-579 0.48311, 184-256 0.48703, 19-84 0.49341, 253-564 0.58907, 34-519 0.65081, 153-167 0.65328, 529-557 0.66670, 188-555 0.57823, 325-547 0.86283, 14-259 0.70977, 156-192 0.72085, 407-533 0.80562, 54-553 0.88206, 162-402 0.89845, 179-405 0.99247, 356-423 0.99656, 421-550 0.99667, 96-319 1.00984, 246-456 1.01107, 187-214 1.02299, 61-528 1.04419, 278-309 1.07211, 349-522 1.10182, 371-484 1.12408, 102-386 1.18726, 150-269 1.19324, 264-417 1.19711, 340-40 1.21320, 357-441 1.24278, 367-489 1.26700, 188-327 1.27562, 131-281 1.28360, 296-559 1.29339, 271-574 1.35998, 348-502 1.37428, 18-411 1.38466, 28-63 1.41615, 397-573 1.47746, 20-154 1.48077, 22-60 1.49806, 134-571 1.51983, 24-499 1.62884, 208-268 1.63329, 88-176 1.67171, 227-345 1.77337, 273-508 1.79424, 313-428 1.87346, 5-540 1.87757, 387-498 1.91652, 433-649 1.99640, 91-494 2.07635, 117-539 2.19171, 64-416 2.19786, 88-362 2.22520, 144-457 2.30198, 220-354 2.39305, 59-85 2.62324, 461-538 2.69966, 92-476 2.76853, 439-448 2.85806, 497-525 2.87879, 87-323 2.93692, 4-572 3.08576, 56-565 3.18374, 42-311 3.20899, 195-218 3.40476, 52-377 3.42055, 59-209 3.43723, 12-389 3.56078, 104-169 3.56161, 263-424 3.62318, 302-546 3.99937, 198-277 4.08123, 110-563 4.08434, 372-438 4.32685, 29-347 4.34645, 196-365 4.52037, 459-478 4.85655, 86-566 5.29052, 58-361 7.83514, 163-331 8.91107, 294-359 9.25799, 129-248 11.48198, 199-280 13.39033, 23-202 38.06727]

value of start point of tour 418

Hamiltonian Tour

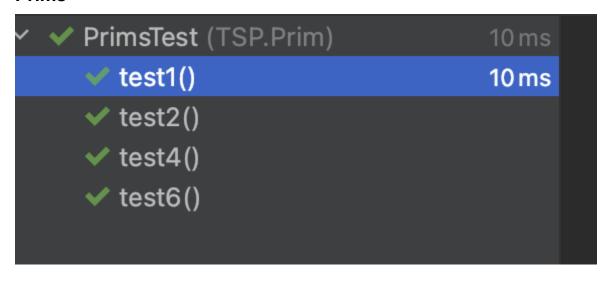
namilton tour path weight in km 865.6042363609656 hamilton tour path weight in metres 865604.2363609656

Two Opt

twoOptCalculation method {tourWeight=641.6928552968773, tour=[418, 586, 341, 425, 32, 2, 287, 42, 483, 512, 48, 211, 581, 95, 494, 567, 91, 81, 536, 459, 44, 478, 223, 429, 348, 582, 558, 544, 77, 452, 588, 52, 179, 36, 288, 495, 346, 377, 318, 155, 525, 548, 497, 8, 43, 487, 53, 244, 65, 199, 236, 351, 285, 83, 136, 71, 486, 566, 578, 142, 176, 88, 385, 79, 478, 426, 334, 241, 376, 288, 541, 60, 78, 311, 569, 22, 382, 414, 427, 528, 61, 343, 515, 158, 187, 214, 531, 465, 227, 269, 31, 55, 472, 445, 455, 555, 544, 577, 217, 408, 446, 324, 186, 302, 318, 171, 293, 329, 546, 488, 571, 89, 392, 24, 561, 287, 490, 409, 11, 257, 315, 263, 556, 424, 437, 75, 208, 295, 456, 454, 267, 362, 68, 386, 102, 146, 395, 347, 29, 194, 249, 266, 96, 562, 319, 144, 193, 270, 23, 340, 404, 94, 570, 398, 245, 268, 208, 457, 166, 415, 276, 564, 253, 312, 471, 233, 275, 48, 549, 284, 69, 316, 333, 412, 198, 16, 277, 50, 14, 259, 320, 370, 328, 384, 209, 12, 203, 274, 251, 389, 17, 224, 231, 132, 145, 256, 184, 416, 64, 350, 304, 163, 488, 197, 581, 303, 101, 331, 505, 279, 3, 127, 367, 242, 489, 332, 443, 123, 495, 576, 243, 520, 444, 461, 283, 462, 538, 128, 551, 381, 336, 338, 338, 339, 351, 359, 296, 398, 334, 339, 371, 484, 222, 530, 220, 139, 105, 463, 288, 383, 336, 535, 552, 88, 126, 393, 529, 557, 575, 375, 229, 195, 355, 499, 218, 258, 213, 359, 294, 486, 374, 440, 436, 391, 252, 438, 327, 518, 428, 477, 78, 265, 344, 352, 397, 509, 107, 410, 138, 298, 573, 508, 210, 226, 97, 159, 579, 135, 301, 456, 485, 98, 131, 46, 235, 503, 1, 169, 262, 288, 124, 237, 511, 458, 230, 198, 431, 116, 492, 481, 516, 474, 41, 559, 82, 234, 296, 291, 441, 357, 297, 460, 99, 168, 247, 104, 407, 419, 289, 545, 103, 533, 532, 373, 519, 344, 421, 518, 373, 373, 248, 469, 433, 93, 205, 468, 253, 308, 568, 56, 364, 185, 181, 114, 84, 19, 514, 49, 282, 45, 272, 565, 260, 74, 482, 534, 67, 271, 574, 58, 408, 118, 147, 434, 417, 264, 593, 519, 474, 41, 559, 366, 354, 353, 333, 377, 388, 529, 587, 597, 216, 516, 376, 229, 486, 353, 333, 317, 38, 422, 4

Three Opt

Unit Test Sample Prims



Perfect Matching

```
PerfectMatchingTest (TSP.Prim) 14 ms

✓ test2() 14 ms

Process finished with exit code 0
```

Mathematical Analysis:

Christofides Algorithm

We can improve on the MST algorithm by slightly modifying the MST. Define an Euler tour of a graph to be a tour that visits every edge in the graph exactly once. As before, find the minimum spanning tree T of G rooted at some node r. Compute the minimum cost perfect matching M of all the odd degree vertices, and add M to T to create T/. Let H be the list of vertices of Euler tour of T/ with duplicate vertices removed. Return the cycle that visits vertices in the order of H.

Approximation Ratio We will show that the Christofides algorithm is a 2 3 -approximation algorithm for the metric TSP problem. We first note that an Euler tour of T/ = T U M exists because all vertices are of even degree. We now bound the cost of the matching M. Lemma 3 $c(M) \le 1 \ 2 \ c(H*)$. Proof. Consider the optimal solution H/ to the TSP of just the odd degree vertices of T. We can break H/ to two perfect matchings M1 and M2 by taking every other edge. Because M is the minimum cost perfect matching, we know that $c(M) \le \min(c(M1), c(M2))$. Furthermore, because H/ only visits a subset of the graph, $c(H/) \le c(H*)$. Therefore, $c(H/) \le c(H*) \Rightarrow c(M) \le 1 \ c(H*)$. 2 The cost of Euler tour of T/ is c(T) + c(M) since it visits all edges exactly once. We know that $c(T) \le c(H*)$ as before (Lemma 1). Using Lemma 3 along with Lemma 1, we get $c(T) + c(M) \le c(H*) + 2 \ 1 \ c(H*) = 2 \ 3 \ c(H*)$. Finally, removing duplicates further reduces the cost by triangular inequality. Therefore, $c(H) \le c(T/) = c(T) + c(M) \le 3 \ 2 \ c(H*)$

Conclusion:

In conclusion, the choice of optimization algorithm depends on the specific instance of the TSP problem and the trade-off between computational complexity and solution quality. For smaller TSP instances, the Christofides algorithm or 2-opt and 3-opt optimization may be sufficient due

to their lower time complexity. For larger TSP instances, simulated annealing or ant colony optimization may be necessary to find good solutions, despite their higher time complexity. However, both simulated annealing and ant colony optimization require careful tuning of their parameters to achieve good results.

References:

To Code Christofides algorithm: https://www.youtube.com/watch?v=GiDsilBOVoA

To Code 2-opt : https://en.wikipedia.org/wiki/2-opt

To Code 3-opt: https://en.wikipedia.org/wiki/3-opt

1a] img src: https://isd.ktu.lt/it2011/material/Proceedings/1_AI_5.pdf

1b] img src:

https://www.researchgate.net/profile/Kamil-Rocki/publication/229476110/figure/download/fig 1/AS:670005742034983@1536753225918/Percentage-of-time-spent-on-2-opt-local-search-only-during-a-single-Iterative-Local.png

1c] img src:

https://slideplayer.com/slide/7698916/25/images/16/Simulated+Annealing+%283%2F3%29.jpg

1d] img src: https://personal.ntu.edu.sg/elpwang/PDF_web/08_ICNC_LiBifan.pdf