Held Karp Algorithm

Group Members
Hitarth Bhatt - 202201024
Tanmay Singh -202201135
Ram Kulkarni - 202201354
Tanay Jain - 202201026
Gireesh Reddy - 202201122
Shouraya Nahar -202201296*

May 2023

[1]

*__SC205 Final Project(Course Instructor: Prof. Manish K. Gupta)__
[1]source:https://ecampus.daiict.ac.in/webapp/intranet/index.jsp

Contents

Abstract

This project is on one of the application of graph theory. In this presentation, Held Karp algorithm is used for solving travelling salesman problem.

1 INTRODUCTION

The **Held–Karp algorithm**, also called **Bellman–Held–Karp algorithm**, is a dynamic programming algorithm proposed in 1962 independently by Bellman and by Held and Karp to solve the traveling salesman problem (TSP), in which the input is a distance matrix between a set of cities, and the goal is to find a minimum-length tour that visits each city exactly once before returning to the starting point. It finds the exact solution to this problem, and to several related problems including the Hamiltonian cycle problem, in exponential time.

Dynamic programming is both a mathematical optimization method and a computer programming method. The method was developed by Richard Bellman in the 1950s.

2 Application

The general applications of the TSP are ranging from mapping genes to scheduling space-based telescopes. Some of the real-life applications of this algorithm are the following:

2.1 SCHOOL BUS ROUTING

Scheduling a single run of a school bus to pick up waiting children is naturally a TSP. Indeed, a number of companies specialize in school bus routing software, offering school districts a quick way to optimize their pickup schedules via the solution of the TSP.

2.2 GENOME SEQUENCING

The mapping of the human genome is one of the great achievements in the history of science.
A focus of this work before finding the genome sequence is the accurate placement of markers that serve as landmarks for the genome maps. The TSP plays an important role here, providing a tool for building sequences from experimental data on the proximity of individual pairs of markers.

2.3 PRINTED CIRCUIT BOARDS

Printed circuit boards can be found in many common electronic devices. They are typically used to mount the integrated circuits, allowing chips to be combined with other hardware to obtain some specific functionality. While scan chains arise in chip design, a classic application of the TSP arises in the production of the basic boards.

2.4 GUIDING LASERS FOR CRYSTAL ART

The focal point of the laser beam is used to create fractures at specified three-dimensional points in the crystal, creating tiny points that are visible in the clear material. The TSP is again to route the laser through the points to minimize the production time.

2.5 AIMING TELESCOPES AND X-RAYS

Sites such as planets, stars, and galaxies and the observations are to be made with some form of telescope. Here the process of rotating equipment into position is called slewing. For large-scale telescopes, slewing is a complicated and time-consuming procedure, handled by computer-driven motors. In this setting, a TSP tour that minimizes the total slewing time for a set of observations can easily be implemented as part of an overall scheduling process.

3 Formulation and Solving of Mathematics

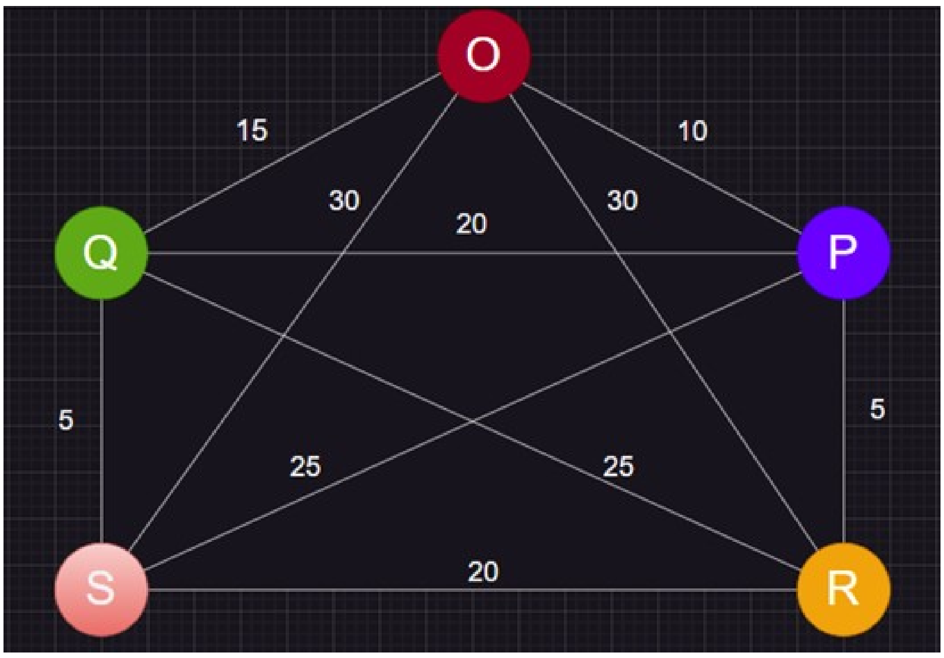
3.1 Problem Statement

A School bus starting from School needs to visit all the specified bus stops, where the distance between all the stops are known and each bus stop needs to be visited once. What is the shortest possible route that the bus visit each stop exactly once and returns back to school?

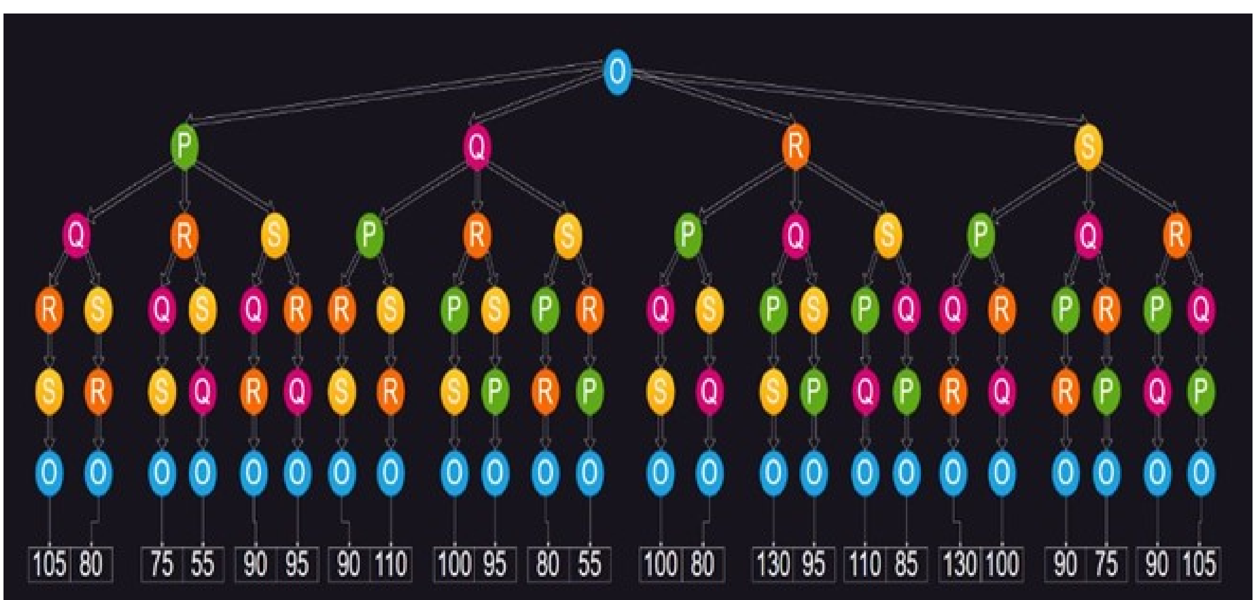
[5]

3.2 Solution

The bus in this case has to visit each and every bus stop (node) exactly so once this Implies the bus is following the Hamiltonian path. But this problem is more complex than the Hamiltonian cycle as shortest path is also to be found. Given below graph depicts the five bus stops which the school bus has to travel. Bus has to start from the bus stop O(school), where many possible routes are possible to travel.



We have to find the shortest route from all those routes. First of all, we will draw a tree which has all the possible routes and try to find the shortest route among all the routes.



This is the recursive tree where ‘O’ is the starting bus stop and we have to visit remaining bus stops in any order we want.
So we can go from ‘O’ to ‘P’ or ’O’ to ‘Q’ or ‘O’ to ‘R’ or ‘O’ to ‘S’. This process will go again and again.

After drawing all the routes and summing all the distances of all such cases (24 routes) we will get the shortest route(s) for school bus.
Now let's solve this problem using Held-Karp algorithm.
Held-Karp algorithm is nothing but a Dynamic programming algorithm. Dynamic Programming can be applied only if main problem can be divided into sub-problems. From the above recursive tree we derived a formula which is stated below
C(I,S)=min{ D(I, k) + C(k, S-k)} ,where k belongs to set S. This is the dynamic programming formula.
The cost function is termed as C, starting bus stop is I, D is the distance between two bus stops, S is set of remaining bus stops, we can go of any of the k bus stops that belong to the set S.
The table drawn below, shows the distances between any two bus stops:

	O	P	Q	R	S
O	0	10	15	30	30
P	10	0	20	5	25
Q	15	20	0	25	5
R	30	5	25	0	20
S	30	25	5	20	0

In our problem the starting bus stop(school) is O and we have to travel remaining bus stops in the shortest possible path. Given below is the general formula considering O as starting bus stop
C(O,{P,Q,R,S})= MIN{ D(O,K)+C(K,P,Q,R,S-K)},where S={P,Q,R,S}

3.3 Calculation

In our problem O is the starting bus stop(school).
STEP 1: We need to find the distance between O and remaining bus stops such that the bus passes through exactly one bus stop.

$C_{OP(R)} = D_{OR} + D_{RP} = 35$

$C_{OP(Q)} = D_{OQ} + D_{QP} = 35$

$C_{OP(S)} = D_{OS} + D_{SP} = 55$

$C_{OR(P)} = D_{OP} + D_{PR} = 15$

$C_{OR(Q)} = D_{OQ} + D_{QR} = 40$

$C_{OR(S)} = D_{OS} + D_{SR} = 50$

$C_{OQ(R)} = D_{OR} + D_{RQ} = 55$

$C_{OQ(P)} = D_{OP} + D_{PQ} = 30$

$C_{OQ(R)} = D_{OS} + D_{SQ} = 35$

$C_{OS(P)} = D_{OP} + D_{QR} = 35$

$C_{OS(Q)} = D_{OQ} + D_{QS} = 20$

$C_{OS(R)} = D_{OR} + D_{RS} = 50$

STEP 2: We need to find the distance between O and remaining bus stops such that the bus passes through exactly two bus stops.(using the info of step 1)

$$\begin{aligned} C_{OP(QR)} &= MIN\{C_{OQ(R)} + D_{QP}, C_{OR(Q)} + D_{RP}\} \\ &= MIN\{55 + 20, 40 + 5\} \\ &= 45 \end{aligned}$$

$$\begin{aligned} C_{OP(QS)} &= MIN\{C_{OQ(S)} + D_{QP}, C_{OS(Q)} + D_{SP}\} \\ &= MIN\{35 + 20, 25 + 20\} \\ &= 45 \end{aligned}$$

$$\begin{aligned} C_{OP(RS)} &= MIN\{C_{OS(R)} + D_{SP}, C_{OR(S)} + D_{RP}\} \\ &= MIN\{50 + 25, 50 + 5\} \\ &= 55 \end{aligned}$$

$$\begin{aligned} C_{OQ(PR)} &= MIN\{C_{OP(R)} + D_{PQ}, C_{OR(P)} + D_{RQ}\} \\ &= MIN\{15 + 20, 15 + 25\} \\ &= 40 \end{aligned}$$

$$\begin{aligned} C_{OQ(SR)} &= MIN\{C_{OS(R)} + D_{SQ}, C_{OR(S)} + D_{RQ}\} \\ &= MIN\{50 + 5, 50 + 25\} \\ &= 55 \end{aligned}$$

$$\begin{aligned} C_{OQ(PS)} &= MIN\{C_{OP(S)} + D_{PQ}, C_{OS(P)} + D_{SQ}\} \\ &= MIN\{55 + 20, 35 + 5\} \\ &= 40 \end{aligned}$$

$$\begin{aligned} C_{OR(PQ)} &= MIN\{C_{OP(Q)} + D_{PR}, C_{OQ(P)} + D_{QR}\} \\ &= MIN\{35 + 5, 30 + 25\} \\ &= 40 \end{aligned}$$

$$\begin{aligned} C_{OR(PS)} &= MIN\{C_{OP(S)} + D_{PR}, C_{OS(P)} + D_{SR}\} \\ &= MIN\{55 + 5, 35 + 20\} \\ &= 55 \end{aligned}$$

$$\begin{aligned} C_{OR(QS)} &= MIN\{C_{OS(Q)} + D_{SR}, C_{OQ(S)} + D_{QR}\} \\ &= MIN\{20 + 20, 35 + 25\} \\ &= 40 \end{aligned}$$

$$\begin{aligned} C_{OS(PQ)} &= MIN\{C_{OP(Q)} + D_{PS}, C_{OQ(P)} + D_{QS}\} \\ &= MIN\{35 + 25, 30 + 5\} \\ &= 35 \end{aligned}$$

$$\begin{aligned} C_{OS(PR)} &= MIN\{C_{OP(R)} + D_{PS}, C_{OR(P)} + D_{RS}\} \\ &= MIN\{35 + 35, 15 + 20\} \\ &= 35 \end{aligned}$$

$$\begin{aligned} C_{OS(QR)} &= MIN\{C_{OR(Q)} + D_{RS}, C_{OQ(R)} + D_{QS}\} \\ &= MIN\{40 + 20, 55 + 5\} \\ &= 60 \end{aligned}$$

Step 3 : We need to find the distance between O and remaining bus stops such that the bus passes through exactly three(all) bus stops(using the info of step 2)

$$\begin{aligned} C_{OP(QRS)} &= MIN\{C_{OQ(RS)} + D_{QP}, C_{OR(QS)} + D_{RP}, C_{OS(RQ)} + D_{SP}\} \\ &= MIN\{60 + 20, 40 + 5, 60 + 25\} \\ &= 45 \end{aligned}$$

$$\begin{aligned} C_{OQ(PRS)} &= MIN\{C_{OP(RS)} + D_{PQ}, C_{OR(PS)} + D_{RQ}, C_{OS(RP)} + D_{SQ}\} \\ &= MIN\{55 + 20, 55 + 25, 35 + 5\} \\ &= 40 \end{aligned}$$

$$\begin{aligned} C_{OR(QRP)} &= MIN\{C_{OQ(PS)} + D_{QR}, C_{OP(QS)} + D_{PR}, C_{OS(PQ)} + D_{SR}\} \\ &= MIN\{40 + 25, 45 + 5, 35 + 20\} \\ &= 50 \end{aligned}$$

$$\begin{aligned} C_{OS(QRP)} &= MIN\{C_{OQ(RP)} + D_{QS}, C_{OR(QP)} + D_{RS}, C_{OP(RQ)} + D_{PS}\} \\ &= MIN\{40 + 5, 40 + 20, 45 + 25\} \\ &= 45 \end{aligned}$$

STEP 4:: Now we have to add back the distances to get back to O in every case of step 3 to find the shortest among all.

In $C_{OP(QRS)}$ shortest distance will MIN of $C_{OP(QRS)} + D_{PO}$
$= 45{+}10 = 55$
In $C_{OQ(PRS)}$ shortest distance will MIN of $C_{OQ(PRS)} + D_{QO}$
$= 40{+}15 = 55$
In $C_{OR(QPS)}$ shortest distance will MIN of $C_{OR(QPS)} + D_{RO}$
$= 50{+}30 = 80$
In $C_{OS(QRP)}$ shortest distance will MIN of $C_{OS(QRP)} + D_{SO}$
$= 45{+}30 = 75$
So from the above calculations we can conclude that shortest path to cover all the bus stops and return to school is

1. $O \rightarrow Q \rightarrow S \rightarrow R \rightarrow P \rightarrow O$

2. $O \rightarrow P \rightarrow R \rightarrow S \rightarrow Q \rightarrow O$

So 55 is the minimum distance which bus has to travel to cover all the bus stops.
Here both the paths are same, only the direction of the paths are different. The school bus can choose any one of the path from these two so that the bus has to travel minimum distance.

3.4 Analysis

TSP belongs to the class of combinatorial optimization known as NP-complete.
TSP is classified as NP hard because no quick solution and it becomes more difficult to calculate best route when the no of nodes increases. Since it takes exponential time to solve NP, the solution cannot be checked in polynomial time.
The naive method to solve the problem is just checking all the possible cases, which has time complexity O(n!) and takes lot of time because there will be many cases which we have to check and then compare.
On the other hand in Held Karp algorithm we have to find minimum distance to go from O to every bus stop passing through all the possible cases of bus stops (1,2,3...n-1).
$^nC_0 +^n C_1 +^n C_2 + ... +^n C_n = 2^n$

This algorithm considers 2n subsets. For each subset, one computes the cost function for all of its elements, which there are no more than n of them. For each execution of the cost function, one again computes no more than n values based on the values obtained from the previous step. In total, the time complexity of this algorithm is: $O(n^2 * 2^n)$ Another way to find the time complexity is:
If we solve the recursive equation we will get $(n-1)2^{(n-2)}$ sub-problems, which is $\mathbf{O}(n2^n)$.
Each sub-problem will take O(n) time (finding path to remaining (n-1) nodes).Therefore total Time Complexity is

$$O(n2^n) * O(n) = O(n^2 2^n).$$

Space complexity is also number of sub-problems which is $O(n2^n)$.
Time complexity is exponential and space complexity is also exponential.

4 Pseudo Code

Data: Graph G with n vertices
Result: Optimal cost of the Bus Stops Problem
for $k := 2$ **to** n **do**
  $g(\{k\}, k) := d(1, k)$;
end
for $s := 2$ **to** $n-1$ **do**
  for *all* $S \subseteq \{2, \ldots, n\}$ *with* $|S| = s$ **do**
    for *all* $k \in S$ **do**
      $g(S, k) := \min_{m \neq k, m \in S} [g(S \setminus \{k\}, m) + d(m, k)]$;
    end
  end
end
$opt := \min_{k \neq 1} [g(\{2, \ldots, n\}, k) + d(k, 1)]$;
return *opt*;

[2]

5 Advantages Of This Algorithm

1. Performs better against other global optimisation techniques such as neutral net, genetic algorithms, simulated annealing.

2. Can be used in dynamic application that is adapts to changes such as new distances.
3. The most efficient routes for drivers to take when out on the road.
4. The better it is solved, the less distance is travelled.
5. More fuel is saved. With the ever-increasing importance of reducing our carbon footprint, reducing the distances travelled directly contributes to the amount of fossil fuels being burned. The ability to effect such important change through problem solving is a bonus for the greater society too.
6. Less hours are required by the drivers.

6 Disadvantages Of This Algorithm

It is quite difficult to solve TSP as it is known as NP-hard. So, with an increasing the amount of nodes, the complexity of solving TSP increases exponentially hence it becomes infeasible to solve the problem.

7 CONCLUSION

- In general, the travelling salesman problem through naïve method is a bit complex to solve. If there is a way to break this problem into smaller component to break this into smaller component problems, the components will be at least as complex as the original one.
- The easiest and the most expensive solution is to try all possibilities but the problem with this method is that for N bus stops you have (N-1) factorial possibilities.
- This becomes very difficult to find. This approach is infeasible even for slightly high number of vertices, so we use the dynamic programming(Held-Karp algorithm) approach to solve this problem.

8 Group Members And Their Contributions

1. **Hitarth Bhatt(202201024):-** He is the group Leader and he has done all the documentation like latex, biblography, ppt. (Contribution 16.67%)
2. **Tanmay Singh(202201135):-** He is the one who is working on the code of the algorithm and the website with Tanay Jain.(Contribution 16.67%)
3. **Ram Kulkarni(202201354):-** He is the one who has done the research, formulatiom, calculation and preparing material for documentation with Gireesh Reddy and Shouraya Nahar. (Contribution 16.67%)
4. **Tanay Jain(202201026):-** He is also the one who is working on the code of the algorithm and the website with Tanmay Singh.(Contribution 16.67%)
5. **Gireesh Reddy(202201122):-**He is the one who has done the research, formulatiom, calculation and preparing material for documentation with Ram Kulkarni. and Shouraya Nahar. (Contribution 16.67%)
6. **Shouraya Nahar(202201296):-**He is the one who has done the research, formulatiom, calculation and preparing material for documentation with Ram Kulkarni and Gireesh Reddy (Contribution 16.67%)

References

[1] D.L. Applegate, R.E. Bixby, V. Chvátal, and W.J. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton Series in Applied Mathematics. Princeton University Press, 2011.

[2] Subham Datta. Traveling salesman problem – dynamic programming approach. https://www.baeldung.com/cs/tsp-dynamic-programming, 2023. For Pseudo Code.

[3] Brendan Tobin Caroline Hassett. The travelling salesman problem [and how to solve it]. https://smartroutes.io/blogs/the-travelling-salesman-problem/, 2022.

[4] Kodeeswaran. Travelling salesman problem (tsp). https://www.interviewbit.com/blog/travelling-salesman-problem/?amp=1, 2023.

[5] Petrovv. Concept school bus route on the map. https://shorturl.at/pDNUV. For the Bus Pic.

[6] Stark Rayn. Held karp algorithm. https://youtu.be/jUYAJ72m8P0, 2020.

[7] Tim Roughgarden. Algorithms for np-hard problems. https://youtu.be/9DwFwkAwnmY, 2020.